Lecture 02: Mathematical Basics (Inequalities)

Arithmetic Mean - Geometric Mean Inequality

Theorem (Basic AM-GM)

For $a, b > 0$, we have

$$\frac{a+b}{2} \geqslant \sqrt{ab}$$

And equality holds if and only if $a = b$.

Proof.

$$\frac{a+b}{2} \geqslant \sqrt{ab} \iff \left(\sqrt{a} - \sqrt{b}\right)^2 \geqslant 0$$

The second statement is true for all reals. And, equality holds if and only if $a = b$. □

Geometric Mean - Harmonic Mean Inequality

Theorem (Basic GM-HM)

For $a, b > 0$, we have

$$\sqrt{ab} \geqslant \left(\frac{\frac{1}{a} + \frac{1}{b}}{2} \right)^{-1}$$

And equality holds if and only if $a = b$.

Think: Proof? (it is a consequence of the Basic AM-GM Inequality)

- First step is to note that

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

This can be proven by induction on n and using the Basic AM-GM

Generalizing AM-GM Inequality

Theorem ((Slight) Generalization of AM-GM)

For $\alpha_1, \dots, \alpha_n \in \mathbb{Q}^+$ such that $\sum_{i=1}^{n} \alpha_i = 1$ and $a_1, \dots, a_n \geqslant 0$, we have

$$\sum_{i=1}^{n} \alpha_i a_i \geqslant \prod_{i=1}^{n} a_i^{\alpha_i}$$

And, equality holds if and only if $a_1 = \cdots = a_n$.

Think: Prove using Basic AM-GM. Let $\alpha_i = p_i/q_i$ where p_i and q_i are relatively prime integers. Let N be the L.C.M. of $\{q_1, \dots, q_n\}$ Consider $(p_i/q_i)N$ copies of a_i, for $i \in [n]$ and apply AM-GM on the N numbers

Think: Generalize GM-HM analogously

Further Generalization: Generalization to $\alpha_i \in \mathbb{R}$ will be done later

Theorem (Jensen's Inequality)

Let f be a convex downward function in the range R. Let $\mathbb{X}$ be a probability distribution over $x_1, \dots, x_n \in R$. We have

$$\mathbb{E}\left[f(\mathbb{X})\right] \geqslant f\left(\mathbb{E}\left[\mathbb{X}\right]\right)$$

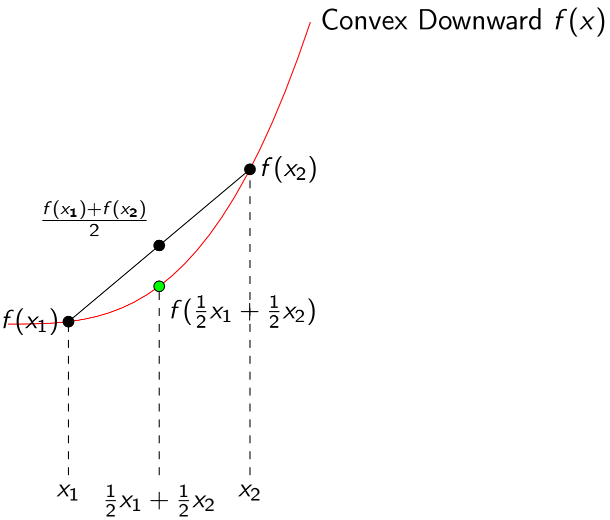
Equality holds if and only if all x_i are identical for i such that $\mathbb{P}\left[\mathbf{X} = i\right] > 0$.

Clarification: "Convex Downward" function is a function that looks like the function $f(x) = x^2$ and *does not* look like the function $f(x) = \sqrt{x}$

Proof Intuition: Use induction on n. Base case of $n = 2$ is proven using: "the chord between two points lies above the function between the two points."

Think: Analogous statement for convex upwards function

Intuition



Jensen's Inequality says that $\frac{f(x_1)+f(x_2)}{2}$ is higher than $f(\frac{1}{2}x_1 + \frac{1}{2}x_2)$

Application: AM-GM from Jensen's Inequality

- Let $f(x) = \log x$, for $x > 0$
- Note that $f(x)$ is "convex upwards" (i.e., it looks like $\sqrt{x}$)
- Let $\mathbb{X}$ be the random variable over $[n]$ that outputs i with probability α_i
- Let $x_i = a_i$ for $i \in [n]$
- By Jensen's Inequality we have $\mathbb{E}\left[f(\mathbb{X})\right] \leqslant f\left(\mathbb{E}\left[\mathbb{X}\right]\right)$
- This equivalent to

$$\sum_{i\in[n]} \alpha_i \log x_i \leqslant \log\left(\sum_{i\in[n]} \alpha_i x_i\right)$$

- Exponentiating both sides, we get the AM-GM inequality: $\prod_{i\in[n]} x_i^{\alpha_i} \leqslant \sum_{i\in[n]} \alpha_i x_i$

Generalized AM-GM

Theorem (Generalized AM-GM)

For $\alpha_1, \ldots, \alpha_n \in \mathbb{R}^+$ *such that* $\sum_{i=1}^{n} \alpha_i = 1$ *and* $a_1, \ldots, a_n \geqslant 0$, *we have*

$$\sum_{i=1}^{n} \alpha_i a_i \geqslant \prod_{i=1}^{n} a_i^{\alpha_i}$$

And, equality holds if and only if $a_1 = \cdots = a_n$.

Cauchy–Schwarz Inequality

Theorem (Cauchy–Schwarz Inequality)

Let $a_1, \ldots, a_n, b_1, \ldots, b_n \geqslant 0$. Then the following holds

$$\sum_{i=1}^{n} a_i b_i \leqslant \left(\sum_{i=1}^{n} a_i^2 \right)^{1/2} \left(\sum_{i=1}^{n} b_i^2 \right)^{1/2}$$

Equality holds if and only if a_i / b_i is a constant for all $i \in [n]$.

Proof Outline:

- Prove the theorem for $n = 2$ using AM-GM inequality
- Prove for $n > 2$ using induction

Hölder's inequality

Theorem (Hölder's inequality)

Let $a_1, \ldots, a_n, b_1, \ldots, b_n \geqslant 0$. Let $p, q > 0$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then the following holds

$$\sum_{i=1}^{n} a_i b_i \leqslant \left(\sum_{i=1}^{n} a_i^p \right)^{1/p} \left(\sum_{i=1}^{n} b_i^q \right)^{1/q}$$

Equality holds if and only if a_i^p / b_i^q is a constant for all $i \in [n]$.

Proof Outline:

- Assume the inequality holds for $n = 2$
- Use induction to extend the inequality to extend to $n > 2$

In this section we prove the full Hölder's inequality in one-shot. The case of $n = 2$ is just a restriction of the analysis below to $n = 2$.

- Note that $p, q > 0$ such that $\frac{1}{p} + \frac{1}{q} = 1$ implies that $p, q > 1$
- Consider the function $f(x) = x^p$
- For $p > 1$, this function is convex downwards
- Let $x_i = a_i / b_i^{q/p}$
- Let $\alpha_i = \Lambda \cdot b_i^{1+\frac{q}{p}}$, where Λ is the normalizing constant such that $\sum_{i \in [n]} \alpha_i = 1$
- By Jensen's Inequality on $f(x) = x^p$ we have:

$$\sum_{i \in [n]} \alpha_i x_i^p \geqslant \left(\sum_{i \in [n]} \alpha_i x_i \right)^p$$

Let be first find what is the value of Λ.

- $\sum_{k\in[n]} \alpha_k = \sum_{k\in[n]} \Lambda \cdot b_k^{1+\frac{q}{p}} = 1$
- Note that $\frac{1}{p} + \frac{1}{q} = 1$. Multiplying both sides by q, we get $\frac{q}{p} + 1 = q$
- Now, we can substitute $\frac{q}{p} + 1 = q$ to get

$$\sum_{k\in[n]} \alpha_k = \Lambda \sum_{k\in[n]} b_k^q = 1$$

- This implies that

$$\Lambda = 1/ \sum_{k\in[n]} b_k^q$$

- Now, we can conclude that

$$\alpha_i x_i = \frac{a_i b_i}{\sum_{k\in[n]} b_k^q}, \text{ and}$$

$$\alpha_i x_i^p = \Lambda \cdot b_i^{1+\frac{q}{p}} \cdot \frac{a_i^p}{b_i^q} = \frac{a_i^p}{\sum_{k\in[n]} b_k^q} \quad \text{Using the fact } \frac{q}{p} + 1 = q$$

Now, let us substitute these values to simplify the equation we had obtained by applying the Jensen's Inequality.

$$\sum_{i\in[n]} \alpha_i x_i^p \geqslant \left(\sum_{i\in[n]} \alpha_i x_i \right)^p$$

$$\iff \quad \sum_{i\in[n]} \frac{a_i^p}{\sum_{k\in[n]} b_k^q} \geqslant \left(\sum_{i\in[n]} \frac{a_i b_i}{\sum_{k\in[n]} b_k^q} \right)^p$$

We continue this simplification in the next page

Base Case of $n = 2$

$$\sum_{i\in[n]} \frac{a_i^p}{\sum_{k\in[n]} b_k^q} \geqslant \left(\sum_{i\in[n]} \frac{a_i b_i}{\sum_{k\in[n]} b_k^q}\right)^p$$

$$\Longleftrightarrow \quad \left(\frac{a_i^p}{\sum_{k\in[n]} b_k^q}\right)^{1/p} \geqslant \sum_{i\in[n]} \frac{a_i b_i}{\sum_{k\in[n]} b_k^q}$$

$$\Longleftrightarrow \quad \left(\sum_{k\in[n]} b_i^q\right)^{1-\frac{1}{p}} \left(\sum_{i\in[n]} a_i^p\right)^{1/p} \geqslant \sum_{i\in[n]} a_i b_i$$

$$\Longleftrightarrow \quad \left(\sum_{k\in[n]} b_i^q\right)^{1/q} \left(\sum_{i\in[n]} a_i^p\right)^{1/p} \geqslant \sum_{i\in[n]} a_i b_i \qquad \because 1 - \frac{1}{p} = \frac{1}{q}$$

This completes the proof of the Höder's Inequality

Taylor and Maclaurin Series

- Let f be an infinitely differentiable function
- $f^{(n)}(x)$ represents $\frac{d^n f}{dx^n}(x)$. For $n = 0$, $f^{(n)}(x)$ represents $f(x)$
- Taylor Series of f around x_0 is given by

$$f(x) = \sum_{n \geqslant 0} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

- Maclaurin Series of f is the Taylor series with $x_0 = 0$
- Define the truncation of the Taylor series of f up to N terms as follows

$$T_{f,N,x_0}(x) := \sum_{n=0}^{N} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n$$

- The remainder function is defined as follows

$$R_{f,N,x_0}(x) := f(x) - T_{f,N,x_0}(x)$$

Lagrange Form

Theorem (The Remainder Theorem)

Suppose f is $N+1$ differentiable function. There exists c between x_0 and x such that

$$R_{f,N,x_0} = \frac{f^{(N+1)}(c)}{(N+1)!}(x - x_0)^{N+1}$$

This theorem bounds the error between $f(x)$ and the truncation $T_{f,N,x_0}(x)$

- Let $f(x) = \exp(-x)$
- Note that $f^{(n)}(x) = (-1)^n \exp(-x)$
- Note that the Taylor series of $f(x)$ is

$$f(x) = \exp(-x) = 1 - x + \frac{x^2}{2} - \frac{x^3}{6} + \cdots$$

- Note that $T_{f,1,0}(x) = 1 - x$ and $T_{f,2,0} = 1 - x + \frac{x^2}{2}$
- By applying the remainder theorem, we get

$$\exp(-x) - (1 - x) = R_{f,1,0} = \frac{\exp(-c)}{2!} x^2 \geqslant 0$$

$$\exp(-x) - (1 - x + \frac{x^2}{2}) = R_{f,2,0} = \frac{-\exp(-c')}{3!} x^3 \leqslant 0$$

- This implies that $1 - x \leqslant \exp(-x) \leqslant 1 - x + x^2/2$

Plots