Lecture 01: Mathematical Inequalities

Overview I

In today's lecture, we shall cover some techniques to prove fundamental mathematical inequalities. We shall rely on the Lagrange form of the Taylor's Remainder Theorem to prove these results. We emphasize that we shall not prove the theorem itself. The course website provides an additional resource that presents the proof of this result. Interested students are encouraged to go over that proof.

Overview II

We shall, use this theorem to prove the following mathematical inequalities.

1. Jensen's Inequality,
    1. AM-GM-HM Inequality
    2. Cauchy-Schwarz Inequality
    3. Young's Inequality
    4. Hölder's Inequality
2. Approximating $\exp(-x)$ and $\ln(1-x)$ using polynomials, and
3. (In the future, we shall cover) Bonami-Beckner-Gross Hypercontractivity Inequality

Lagrange Form of the Taylor's Remainder Theorem I

Let us begin by recalling the Taylor's Theorem

Theorem (Taylor's Theorem)

$$f(a+\varepsilon) = f(a) + f^{(1)}(a)\frac{\varepsilon}{1!} + f^{(2)}(a)\frac{\varepsilon^2}{2!} + \cdots$$

For example

1. Using $f(x) = \exp(-x)$ and $a = 0$, we get

$$\exp(-\varepsilon) = 1 - \frac{\varepsilon}{1!} + \frac{\varepsilon^2}{2!} - \frac{\varepsilon^3}{3!} + \cdots$$

2. Using $f(x) = \ln(1-x)$ and $a = 0$, we get

$$\ln(1-\varepsilon) = -\frac{\varepsilon}{1} - \frac{\varepsilon^2}{2} - \frac{\varepsilon^3}{3} - \cdots$$

Motovation. Suppose we truncate the infinite Taylor series at the $f^{(k)}\frac{\varepsilon^k}{k!}$ term.

1. Is the truncated series an "overestimation" or an "underestimation"?
2. How good is the quality of approximation?

The Lagrange form of the Taylor Remainder Theorem will help answer this question.

Lagrange Form of the Taylor's Remainder Theorem III

Theorem (Lagrange Form of the Taylor Remainder Theorem)

For every a and ε, *there exists* $\theta \in (0,1)$ *such that*

$$f(a+\varepsilon) = \left(f(a) + f^{(1)}(a)\frac{\varepsilon}{1!} + f^{(2)}(a)\frac{\varepsilon^2}{2!} + \cdots + f^{(k)}(a)\frac{\varepsilon^k}{k!} \right) + f^{(k+1)}(a+\theta\varepsilon)\frac{\varepsilon^{k+1}}{(k+1)!}$$

We refer to the term $R = f^{(k+1)}(a+\theta\varepsilon)\frac{\varepsilon^{k+1}}{(k+1)!}$ as the remainder.

- If the remainder is positive, then the truncation is an underestimation. If the remainder is negative, then the truncation is an overestimation.
- The absolute value of the remainder determines the quality of approximation.

Problem 1. Let $f(x) = \exp(-x)$ and $a = 0$. For $k \geqslant 0$, define $p_k(\varepsilon) = \sum_{i=0}^{k} \frac{(-\varepsilon)^i}{i!}$.
For example, we have $p_0(\varepsilon) = 1$, $p_1(\varepsilon) = 1 - \varepsilon$, $p_2(\varepsilon) = 1 - \varepsilon + \varepsilon^2/2$, and $p_3(\varepsilon) = 1 - \varepsilon + \varepsilon^2/2 - \varepsilon^3/6$, and so on.
For $0 \leqslant \varepsilon \leqslant 1$, apply the Taylor's Remainder Theorem to deduce the following.

1. If k is odd then we have $\exp(-\varepsilon) \geqslant p_k(\varepsilon)$.
2. If k is even then we have $\exp(-\varepsilon) \leqslant p_k(\varepsilon)$.
3. Prove that the absolute value of the remainder when we estimate $\exp(-\varepsilon)$ by $p_k(\varepsilon)$ is at most $\varepsilon^{k+1}/(k+1)!$.

Use the code at Desmos to experiment and develop intuition.

Problem 2. Let $f(x) = \ln(1-x)$ and $a = 0$. For $k \geqslant 0$, define $p_k(\varepsilon) = \sum_{i=1}^{k} \frac{-\varepsilon^i}{i}$.
For example $p_0(\varepsilon) = 0$, $p_1(\varepsilon) = -\varepsilon$, $p_2(\varepsilon) = -\varepsilon - \varepsilon^2/2$, $p_3(\varepsilon) = -\varepsilon - \varepsilon^2/2 - \varepsilon^3/3$, and so on.
For $0 \leqslant \varepsilon \leqslant 1$, apply the Taylor's Remainder Theorem to deduce the following.

1. We have $\ln(1-\varepsilon) \leqslant p_k(\varepsilon)$, for all $k \geqslant 0$.
2. What is the magnitude of the remainder?
3. How will you get a lower bound of $\ln(1-\varepsilon)$?

Use the code at Desmos to experiment and develop intuition.

A High-level Intuitive Summary

- We are using polynomials to estimate any function f
- The "behavior of f" at $(a + \varepsilon)$ is guided by the "properties of f" at the point a!

Definition (Convex Function)

A function f is convex in the range $[a, b]$ if $f^{(2)}$ is positive in $[a, b]$.

For example, the following functions are convex

1. $f(x) = x^2$
2. $f(x) = \exp(x)$
3. $f(x) = \exp(-x)$
4. $f(x) = 1/x$, in $(0, \infty)$

Think: How to define convexity of functions of multiple variables?

Jensen's Inequality I

Jensen's Inequality, intuitively, states the following. Suppose f is a convex function. The secant joining any two points on the curve of f lies above the curve of f.

Theorem (Jensen's Inequality)

For a convex f, we have

$$\frac{f(a)+f(b)}{2} \geqslant f\left(\frac{a+b}{2}\right)$$

Equality holds if and only if $a = b$.

In general, if $\mathbb{X}$ is a probability distribution over a sample space Ω then

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

Jensen's Inequality II

- We can use the Lagrange Form of the Taylor's remainder theorem to prove the Jensen's inequality
- A function f is concave if the function $-f$ is convex. For example, the function $\ln x$, $\ln(1-x)$ in the range $[0, 1)$, $\sqrt{x}$ in the range $[0, \infty)$, and $1/x$, in the range $(-\infty, 0)$ are concave function.
- **Think:** What is Jensen's inequality for concave functions?

Example

- Suppose $f(x) = x^2$. Note that f is convex.
- So, we get the following inequality. For all a, b, we have

$$\frac{a^2 + b^2}{2} \geqslant \left(\frac{a+b}{2}\right)^2$$

Equality holds if and only if $a = b$.

Example Problems I

Use Jensen's Inequality to prove the following mathematical inequalities.

1. **AM-GM Inequality.** For positive a, b, we have

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

   Equality holds if and only if $a = b$.
   Consider the function $f(x) = \ln x$ to prove this inequality.

2. **Cauchy-Schwarz Inequality.** For positive a_1, a_2, b_1, b_2, we have

$$(a_1 b_1 + a_2 b_2) \leqslant \left(a_1^2 + a_2^2\right)^{1/2} \left(b_1^2 + b_2^2\right)^{1/2}$$

   Equality holds if and only if $\frac{a_1}{b_1} = \frac{a_2}{b_2}$.
   Consider the function $f(x) = \ln\left(1 + \exp(x)\right)$.

3. **Young's Inequality.** Let $p, q \geqslant 1$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Such a pair of p and q is referred to as Hölder conjugates. For positive a, b, we have

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q}$$

Equality holds if and only if $a^p = b^q$.
Consider the function $f(x) = \ln x$.

4. **Hölder's Inequality.** For Hölder conjugates p and q, the following holds for positive a_1, a_2, b_1, b_2.

$$(a_1 b_1 + a_2 b_2) \leqslant \left(a_1^p + a_2^p\right)^{1/p} \left(b_1^q + b_2^q\right)^{1/q}$$

What is the equality characterization? What function $f(x)$ will you consider?