Maximal Functions and their Applications in Ergodic Theory

@sean#8765

April 2023

1 Maximal Functions in Ergodic Theory

1.1 Preliminaries

Maximal functions have a wide variety of applications due to their use of constructing bounds. One such use is the **Maximal Ergodic Theorem**, which can used to prove the **Pointwise Ergodic Theorem**.

First, we introduce some ergodic theoretic terminology, referring to the book by Yves Coudène. Let $(X, \mathcal{M}, \mu)$ be a measure space.

Definition 1.1. A transformation $T : X \to X$ is called *measure-preserving* (with respect to μ) if for every $A \in \mathcal{M}$, we have

$$\mu(T^{-1}(A)) = \mu(A).$$

A quadruple $(X, \mathcal{M}, \mu, T)$ is called a *measure-preserving dynamical system*. An *invariant* (under T) is a function λ on X such that

$$\lambda \circ T = \lambda$$

μ almost everywhere.

The idea of an ergodic transformation is of importance as we will see later.

Definition 1.2. Let $T : X \to X$ be a measure-preserving transformation on a probability space $(X, \mathcal{M}, \mu)$. We call T an *ergodic transformation* if for every $E \in \mathcal{M}$ that satisfies

$$\mu(T^{-1}(E) \triangle E) = 0$$

we have $\mu(E) = 0$ or $\mu(E) = 1$.

We now define a new class of maximal functions Af, contrasting that of the Hardy-Littlewood maximal functions Hf.

Definition 1.3. Let us denote by A_k the operator

$$A_k f = \frac{1}{k}\sum_{j=0}^{k-1} f \circ T^j,$$

and let

$$f_N^* = \sup_{1 \le k \le N} \frac{1}{k}\sum_{j=0}^{k-1} f \circ T^j = \sup_{1 \le k \le N} A_k f.$$

Then we put

$$Af = \sup_{N \in \mathbb{N}} f_N^* = \sup_{n \in \mathbb{N}} \frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j.$$

1.2 The Maximal Ergodic Theorem

We are now ready to state and prove the **Maximal Ergodic Theorem**.

Maximal Ergodic Theorem. Let $(X, \mathcal{M}, \mu, T)$ be a finite measure-preserving dynamical system, let $f \in L^1(X, \mathcal{M}, \mu)$, and let $\lambda \in L^1(X)$ be an invariant under T on X. Furthermore, define

$$E = \{x \in X : Af > \lambda\}.$$

Then

$$\int_E (f - \lambda)\,\mathrm{d}\mu \ge 0.$$

Proof.
First suppose that $f \in L^\infty(X)$. We shall prove the theorem for this case, then extend it to the case for $f \in L^1(X)$. Define

$$E_N := \{x \in X : f_N^* > \lambda\},$$

and hold $N \in \mathbb{N}$ fixed. Notice that

$$(f - \lambda)\chi_{E_N} \ge f - \lambda$$

as $x \in E_N^C$ means that $(f - \lambda)(x) \le 0$. Hence, there exists an integer $m > N$ such that the sum

$$\sum_{j=0}^{m-1} ((f - \lambda)\chi_{E_N})(T^k x)$$

can be broken up as follows:

- There is a string of terms in the sum for which $T^k x \notin E_N$ and hence the characteristic function of E_N makes such terms vanish.
- There is a first index k such that $T^k x \in E_N$, and a string of no more than N terms follows such that the sum is positive (using the inequality established above).
- After the last term in this string, we return to the first point and continue our analysis, finding some combination of strings of 0s and strings of no more than N terms that have positive sums. The full sum can end in the middle of the two cases mentioned.

Hence, we find $k \in \{m-N+1, \ldots, m\}$ such that

$$\sum_{j=0}^{m-1} (f-\lambda)\chi_{E_N}(T^k x) \geq \sum_{j=k}^{m-1} (f-\lambda)\chi_{E_N}(T^k x) \geq -N(\|f\|_\infty + \lambda^+(x)).$$

We integrate the left-hand side and the right-hand side to see that

$$m\int_{E_N} f - \lambda \,\mathrm{d}\mu \geq -N(\|f\|_\infty \mu(X) + \left\|\lambda^+\right\|_1),$$

and dividing by m then taking the limit as $m \to \infty$ yields

$$\int_{E_N} f - \lambda \,\mathrm{d}\mu \geq 0.$$

Taking $N \to \infty$ and applying the **DCT** then provides

$$\int_E f - \lambda \,\mathrm{d}\mu \geq 0,$$

showing the result for $f \in L^\infty(X)$.

We now suppose $f \in L^1(X)$. For each j, define

$$A_j := \{x \in X : |f(x)| \leq j\},$$

and

$$\varphi_j := f\chi_{A_j}.$$

Clearly, $\varphi_j \in L^\infty$ and $\varphi_j \to f$ a.e. and in L^1. For fixed N,

$$(\varphi_j)^*_N \to f^*_N$$

a.e. and in L^1 and

$$\mu(\{(\varphi_j)^*_N > \lambda\} \triangle \{f^*_N > \lambda\}) \to 0.$$

Therefore, if $\Phi = \{x \in X : (\varphi_j)^*_N > \lambda\}$, then

$$0 \leq \int_\Phi \varphi_j - \lambda \,\mathrm{d}\mu \to \int_{E_N} f - \lambda \,\mathrm{d}\mu$$

by the **DCT**. We then let $N \to \infty$ to deduce that

$$\int_E f - \lambda \,\mathrm{d}\mu \geq 0$$

for $f \in L^1(X)$. $\square$

Notice that when λ is constant, the theorem states

$$\int_E f \,\mathrm{d}\mu \geq \lambda\mu(E).$$

This is often taken to be the statement of the **Maximal Ergodic theorem**, and rewriting the statement in the form

$$\mu(\{x \in X : Af > \alpha\}) \leq \frac{1}{\alpha}\int |f| \,\mathrm{d}\mu$$

for some constant α demonstrates the similarity between this class of maximal functions and the Hardy-Littlewood maximal functions.

1.3 The Pointwise Ergodic Theorem

With this tool at hand, we are now able to prove the **Pointwise Ergodic theorem**.

> **Pointwise Ergodic Theorem.** The sequence $\{A_k f\}$ converges almost everywhere for $f \in L^1(X)$, and its limit is integrable.

Proof.
It suffices to show that

$$\int \limsup_{k\to\infty} A_k f \,\mathrm{d}\mu \leq \int f \,\mathrm{d}\mu\,.$$

To see this, notice that this is equivalent to

$$-\int \liminf_{k\to\infty} A_k f \leq -\int f,$$

so that

$$\int \limsup_{k\to\infty} A_k f \leq \int f \leq \int \liminf_{k\to\infty} A_k f \leq \int \limsup_{k\to\infty} A_k f,$$

and hence

$$\int (\limsup_{k\to\infty}(A_k f) - \liminf_{k\to\infty}(A_k f)) = 0 \implies \limsup_{k\to\infty} A_k f = \liminf_{k\to\infty} A_k f \quad \text{a.e.}$$

First consider the positive part of f, namely f^+. For any invariant function $\lambda < \limsup A_k f^+$ with $\lambda \in L^1(X)$, we have $\{Af^+ > \lambda\} = X$, so the **Maximal Ergodic Theorem** gives

$$\int f^+ \geq \int \lambda,$$

and since this holds for any such function $\lambda < \limsup A_k f^+$, we take the supremum to get that

$$\int f^+ \geq \int \limsup_{k\to\infty} A_k f^+.$$

Hence, the positive part $(\limsup A_k f)^+ \leq \limsup A_k f^+$ is integrable, and a similar argument provides the integrability of $(\limsup A_k f)^-$, which shows that $\limsup A_k f \in L^1(X)$.

Let $\varepsilon > 0$ be arbitrary and apply the **Maximal Ergodic Theorem** to $\lambda = \limsup A_k f - \varepsilon$ to deduce that

$$\int f \geq \int \lambda,$$

and since this applies to all $\varepsilon > 0$, we get that

$$\int \limsup_{k\to\infty} A_k f \leq \int f,$$

and we are done since this clearly implies $\lim A_k f \in L^1(X)$. $\square$

A celebrated fact in ergodic theory is an extension of this result, which assigns a value to this limit and states that

$$\lim_{k\to\infty} A_k f = \lim_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} f(T^j x) = \frac{1}{\mu(X)} \int f \, \mathrm{d}\mu$$

μ-a.e. when T is ergodic (of course, $\mu(X) = 1$ in these cases) and μ is invariant (with respect to T). The left side of the inequality is called the *time average* and the right side is called the *space average*.

The ergodic theorems have many wide-ranging applications, including those but not limited to those in probability theory, the theory of Markov chains and stochastic processes, dynamical systems and even analytic number theory. For example, the (strong) **Law of Large Numbers** may be proved using ergodic theory. One surprising result that may be proved using these theorems is that for almost every $x \in [0, 1]$, the elements of the continued fraction expansion of x are unbounded.