JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 75, No. 1, OCTOBER 1992

TECHNICAL NOTE

Some Comments on a Theorem of Hardy and Littlewood

R. SZNAJDER[1] AND J. A. FILAR[2]

Communicated by G. Leitmann

Abstract. In this note, we reconstruct a proof of a classical result due to Hardy and Littlewood. While this result has played an important role in the modern theories of Markov decision processes and stochastic games, it is not that easy to find its proof in the literature in the format in which it has been applied. Furthermore, we supply either examples or complete citations for the other related cases which are not covered by the Hardy-Littlewood theorem.

Key Words. Markov decision processes, stochastic games, method of discounting, method of Cesaro-averaging, Cesaro-summability, Abel-summability.

1. Introduction

Consider a sequence of real numbers $\{a_n\}_{n=0}^{\infty}$ and a scalar $\beta \in [0, 1)$, which we shall frequently refer to, as the discount factor. In the modern literature on discrete dynamic programming, Markov decision processes, and stochastic games, many authors have studied the relationship between two methods of aggregating a sequence of costs/rewards, namely, the method of discounting where

$$f_N(\beta) := (1-\beta) \sum_{n=0}^{N} \beta^n a_n$$

represents the discounted cost over $(N+1)$-stage horizon, and the method

---

[1] Graduate Student, Department of Mathematics and Statistics, University of Maryland at Baltimore County, Baltimore, Maryland.

[2] Professor, Department of Mathematics and Statistics, University of Maryland at Baltimore County, Baltimore, Maryland.

of Cesaro-averaging where

$$\sigma_N := s_N/(N+1) = \left(\sum_{n=0}^{N} a_n\right) \Big/ (N+1)$$

represents the average cost over $(N+1)$-stage horizon. In particular, the infinite-horizon case ($N=\infty$) has been studied extensively, where the limiting behavior of $\lim_{N\to\infty} f_N(\beta)$ as β approaches 1 has been an important tool in analyzing the limiting behavior of σ_N. One of the underlying mathematical results which makes this analysis possible is a classical result often referred to as the Hardy–Littlewood theorem. The modern literature on stochastic games and Markov decision processes contains many references to the above-mentioned theorem, including some discussion as to its correct form; see Gillette (Ref. 1), Liggett and Lippman (Ref. 2), Flynn (Ref. 3), Stern (Ref. 4), Derman (Ref. 5), Heyman and Sobel (Ref. 6), and Thuijsman (Ref. 7).

In view of the above, the present authors have found it surprising that it was not easy to recover (either from Refs. 8, 9, 10, or from anywhere else) the proofs of the precise statements of the Hardy–Littlewood theorem as they were being applied to Markov decision processes and stochastic games. Of course, Hardy and Littlewood were developing a general theory of summability, and could not have anticipated the form in which their theory would find applications some forty or fifty years later!

The contribution of this note is primarily pedagogical. We provide the precise statements of the Hardy–Littlewood theorem, in the form in which they have been applied in the modern literature, and we provide either complete proofs or exact references for the parts whose proofs we omitted. We have taken care to ensure that the latter citations contain the omitted proofs or counterexamples in the same format as they are referred to in this note. We hope that this note will provide a shortcut for other researchers who have an interest in verifying the modern version of Hardy–Littlewood theorem.

2. Preliminaries, Results, and Examples

A sequence $\{a_n\}_{n=0}^{\infty}$ is said to be C-summable to the limit s if

$$\lim_{N\to\infty} \sigma_N = s.$$

A sequence $\{a_n\}_{n=0}^{\infty}$ is said to be A-summable[3] to a if

$$\lim_{\beta\to 1^-} (1-\beta) \sum_{n=0}^{\infty} \beta^n a_n = a.$$

---

[3] Here, $C(A)$-summability stands for Cesaro (Abel)-summability, respectively.

The following classical results are proved in many sources. We refer the reader to Zygmund (Ref. 11, Vol. 1, pp. 74–80).

Theorem 2.1. If the sequence of real numbers $\{a_n\}_{n=0}^{\infty}$ is C-summable to s, then it is also A-summable to s.

Consider an arbitrary sequence $\{a_n\}_{n=0}^{\infty}$ for which the function

$$f(\beta) := (1-\beta) \sum_{n=0}^{\infty} \beta^n a_n$$

is well defined for all β such that $|\beta| < 1$; then the following theorem holds.

Theorem 2.2.

$$\liminf_{N\to\infty} \sigma_N \le \liminf_{\beta\to 1^-} f(\beta) \le \limsup_{\beta\to 1^-} f(\beta) \le \limsup_{N\to\infty} \sigma_N.$$

Of course, Theorem 2.1 also follows trivially from Theorem 2.2. For historical reasons, and for the sake of consistency with Zygmund's elegant presentation (see Ref. 11), we stated both of the above results.

The next theorem contains the result of Hardy and Littlewood which is often used in the theories of Markov decision processes and stochastic games.

Theorem 2.3. Hardy and Littlewood Theorem. Let $\{a_n\}_{n=0}^{\infty}$ be a bounded sequence of real numbers, and let $\lim_{\beta\to 1^-} f(\beta) = a$. Then, $\lim_{N\to\infty} \sigma_N = a$.

In Section 3, we present a nearly self-contained proof of this result. The boundedness condition in the above can be relaxed if we are willing to assume that the sequence is nonnegative. For the proof of the latter result, we refer the reader to Hobson (Ref. 8, Vol. 2, p. 185) and Titchmarsch (Ref. 9, p. 227). Since each bounded sequence can be represented as a difference of two nonnegative sequences, one can see that Theorem 2.3 can also be deduced in this manner. To see that Theorem 2.3 cannot be extended to arbitrary sequences, we give the following example which is related to the example on p. 226 of Ref. 9.

Example 2.1. Consider the following sequence:

$$a_n = \begin{cases} -k, & \text{if } n = 2k-1, \quad k \in N, \\ k+1, & \text{if } n = 2k, \quad k \in N \cup \{0\}. \end{cases}$$

It now follows that

$$f(\beta)=(1-\beta)\sum_{k=0}^{\infty}(k+1)(\beta^{2k}-\beta^{2k+1})$$

$$=\sum_{k=0}^{\infty}(-1)^k(k+1)\beta^k=1/(1+\beta)^2,$$

where the second equality follows from Cauchy's multiplication formula for series,

$$\sum_{k=0}^{\infty}b_k\beta^k=(1-\beta)\sum_{k=0}^{\infty}(b_0+b_1+\cdots+b_k)\beta^k,$$

with

$$b_k=(-1)^k(k+1).$$

Hence,

$$\lim_{\beta\to 1^-}f(\beta)=1/4,$$

and it is easy to check that

$$\liminf_{N\to\infty}\sigma_N=0,$$

while

$$\limsup_{N\to\infty}\sigma_N=1/2.$$

The next example invokes Theorem 2.3 to show that the strict inequality

$$\liminf_{\beta\to 1^-}f(\beta)<\limsup_{\beta\to 1^-}f(\beta)$$

can occur even for a bounded sequence.

Example 2.2. Let us consider the sequence $\{a_n\}_{n=0}^{\infty}$ as follows:

$$0,0,1,1,0,0,0,0,1,1,1,1,1,1,1,1,\ldots,$$

where each block of zeros or ones is of the same length as the sum of the preceding blocks. It is simple to check that

$$f(\beta)=\sum_{k=1}^{\infty}(\beta^{2^{2k-1}}-\beta^{2^{2k}}),$$

$$\lim_{k\to\infty}\sigma_{2^{2k+1}-1}=1/3,\qquad \lim_{k\to\infty}\sigma_{2^{2k}}=2/3.$$

Hence,

$$\liminf_{N\to\infty} \sigma_N < \limsup_{N\to\infty} \sigma_N,$$

which by Theorem 2.3 implies that

$$\liminf_{\beta\to 1^-} f(\beta) < \limsup_{\beta\to 1^-} f(\beta),$$

as well.

Finally, we mention that Liggett and Lippman (Ref. 2) give an elegant example of a zero-one sequence for which

$$\liminf_{N\to\infty} \sigma_N < \liminf_{\beta\to 1^-} f(\beta).$$

Of course, for the negative of this sequence, we have that

$$\limsup_{\beta\to 1^-} f(\beta) < \limsup_{N\to\infty} \sigma_N.$$

Thus, we see that, even for bounded sequences, strict inequalities are possible in every part of the statement of Theorem 2.2.

3. Proof of Theorem 2.3

Here, we present a proof of the Hardy-Littlewood theorem which uses the following fundamental result due to Littlewood; see Zygmund, Ref. 11, pp. 81-83. Our proof is based on ideas which can be found in Hardy and Littlewood (Ref. 12).

Theorem 3.1. Let $\{u_n\}_{n=0}^{\infty}$ be a sequence such that $u_n = O(1/n)$ and $\lim_{\beta\to 1^-} \sum_{n=0}^{\infty} \beta^n u_n$ exists. Then, the series $\sum_{n=0}^{\infty} u_n$ is convergent, and

$$\sum_{n=0}^{\infty} u_n = \lim_{\beta\to 1^-} \sum_{n=0}^{\infty} \beta^n u_n.$$

Proof of Theorem 2.3. Since

$$(1-\beta) \sum_{n=0}^{\infty} \beta^n = 1, \qquad \text{for all } |\beta| < 1,$$

we have that

$$\lim_{\beta\to 1^-} f(\beta) = A$$

if and only if

$$\lim_{\beta \to 1^-} (1-\beta) \sum_{n=0}^{\infty} \beta^n (a_n - A) = 0.$$

Also,

$$\lim_{n \to \infty} \sigma_n = \sigma$$

if and only if

$$\lim_{n \to \infty} [(a_0 - \sigma) + \cdots + (a_n - \sigma)]/(n+1) = 0.$$

Thus, it is enough to show that

$$\lim_{\beta \to 1^-} f(\beta) = 0 \quad \text{implies that } \lim_{n \to \infty} \sigma_n = 0.$$

Consider

$$w_n = \begin{cases} \sigma_n - \sigma_{n-1} = (na_n - s_{n-1})/n(n+1), & \text{if } n \geq 1, \\ \sigma_0 = a_0, & \text{if } n = 0. \end{cases}$$

Since $\{a_n\}_{n=0}^{\infty}$ is bounded, $na_n = O(n)$ and $s_{n-1} = O(n)$, then

$$w_n = O(1/n^2)O(n) = O(1/n).$$

Define an auxiliary function,

$$g(\beta) := \sum_{n=0}^{\infty} w_n \beta^n = \sum_{n=0}^{\infty} a_n \beta^n/(n+1) - \sum_{n=1}^{\infty} s_{n-1} \beta^n / n(n+1).$$

Thus,

$$\begin{aligned} \beta g(\beta) &= \sum_{n=0}^{\infty} a_n \beta^{n+1}/(n+1) - \sum_{n=0}^{\infty} s_n \beta^{n+2}/[(n+1)(n+2)] \\ &= \int_0^{\beta} \left(\sum_{n=0}^{\infty} a_n y^n \right) dy - \left(\int_0^{\beta} dy \right)^2 \left(\sum_{n=0}^{\infty} s_n y^n \right), \end{aligned} \tag{1}$$

where

$$\left(\int_0^{\beta} dy \right)^2 h(y) = \int_0^{\beta} \left[\int_0^{t} h(y)\, dy \right] dt$$

denotes the second antiderivative of $h(y)$. Now, using the well-known identity

$$\sum_{n=0}^{\infty} a_n\beta^n = (1-\beta) \sum_{n=0}^{\infty} s_n\beta^n$$

and recalling that

$$f(\beta)/(1-\beta) = \sum_{n=0}^{\infty} a_n\beta^n,$$

we obtain

$$\sum_{n=0}^{\infty} s_n\beta^n = f(\beta)/(1-\beta)^2. \tag{2}$$

Using the hypothesis that

$$\lim_{\beta\to 1^-} f(\beta) = 0,$$

we obtain from (1) and (2) that

$$\beta g(\beta) = \int_0^{\beta} f(y)/(1-y)\, dy - \int_0^{\beta} \left[\int_0^{t} f(y)/(1-y)^2\, dy \right] dt.$$

Integrating the second term by parts, we obtain

$$\begin{aligned} \beta g(\beta) &= \int_0^{\beta} f(y)/(1-y)\, dy - \int_0^{\beta} (\beta - y) f(y)/(1-y)^2\, dy \\ &= (1-\beta) \int_0^{\beta} f(y)/(1-y)^2\, dy \\ &= (1-\beta) \cdot [(1/(1-\beta)) \cdot o(1)] \to 0, \qquad \text{as } \beta \to 1^-. \end{aligned}$$

We have now proved that

$$g(\beta) = \sum_{n=0}^{\infty} w_n\beta^n \to 0, \qquad \text{as } \beta \to 1^-.$$

Recalling that $w_n = O(1/n)$, we now obtain from Theorem 3.1 that

$$\sum_{n=0}^{\infty} w_n = 0.$$

Since

$$\sigma_n = \sum_{k=0}^{n} w_k,$$

we have proved that

$$\lim_{n\to\infty} \sigma_n = 0,$$

as required. □

The numerical experiments show that the function in Example 2.2 approximately satisfies

$$\liminf_{\beta\to 1} f(\beta)\approx 0.4984, \qquad \limsup_{\beta\to 1} f(\beta)\approx 0.5027,$$

which, in comparison with

$$\liminf_{N\to\infty} \sigma_N = 1/3, \qquad \limsup_{N\to\infty} \sigma_N = 2/3,$$

shows that this single function strictly distinguishes all the terms in Theorem 2.2.

References

1. GILLETTE, D., *Stochastic Games with Zero Stop Probabilities*, Contributions to the Theory of Games, Edited by A. W. T. M. Dresher and P. Wolfe, Annals of Mathematics Studies, Vol. 39, pp. 179–188, 1957.
2. LIGGETT, T., and LIPPMAN, S., *Stochastic Games with Perfect Information and Time Average Payoff*, SIAM Review, Vol. 11, pp. 604–607, 1969.
3. FLYNN, J., *Averaging vs. Discounting in Dynamic Programming: A Counterexample*, Annals of Statistics, Vol. 2, pp. 411–413, 1974.
4. STERN, M., *On Stochastic Games with Limiting Average Payoff*, PhD Thesis, University of Illinois at Chicago, 1975.
5. DERMAN, C., *Finite State Markovian Decision Processes*, Academic Press, New York, New York, 1970.
6. HEYMAN, D. P., and SOBEL, M. J., *Stochastic Models in Operations Research*, Vol. 2, McGraw-Hill, New York, New York, 1984.
7. THUIJSMAN, F., *Optimality and Equilibria in Stochastic Games*, PhD Thesis, Rijksuniversiteit Limburg, Maastricht, The Netherlands, 1989.
8. HOBSON, E. W., *The Theory of Functions of a Real Variable and the Theory of Fourier Series*, 2nd Edition, Cambridge University Press, Cambridge, England, 1926.
9. TITCHMARSCH, E. C., *The Theory of Functions*, 2nd Edition, Oxford University Press, London, England, 1939.
10. WIDDER, D. V., *The Laplace Transform*, Princeton University Press, Princeton, New Jersey, 1941.
11. ZYGMUND, A., *Trigonometric Series*, Cambridge University Press, Cambridge, England, 1968.
12. HARDY, G. H., and LITTLEWOOD, J. E., *Notes on the Theory of Series (XVI): Two Tauberian Theorems*, Journal of London Mathematical Society, Vol. 6, pp. 281–286, 1931.