The Martingale Approach for Concentration and Applications in Information Theory, Communications and Coding

Igal Sason

Department of Electrical Engineering
Technion - Israel Institute of Technology
Haifa 32000, Israel
E-mail: sason@ee.technion.ac.il

Abstract

This chapter introduces some concentration inequalities for discrete-time martingales with bounded increments, and it exemplifies some of their potential applications in information theory and related topics. The first part of this chapter introduces briefly discrete-time martingales and the Azuma-Hoeffding & McDiardmid's inequalities which are widely used in this context. It then derives these refined inequalities, followed by a discussion on their relations to some classical results in probability theory. It also considers a geometric interpretation of some of these inequalities, providing an insight on the inter-connections between them. The second part exemplifies the use of these refined inequalities in the context of hypothesis testing, information theory, communications, and coding. The chapter is concluded with a discussion on some directions for further research.

Index Terms

Concentration of measures, error exponents, Fisher information, hypothesis testing, information divergence, large deviations, martingales, moderate deviations principle.

I. INTRODUCTION

Inequalities providing upper bounds on probabilities of the type $\mathbb{P}(|X - \overline{x}| \geq t)$ (or $\mathbb{P}(X - \overline{x} \geq t)$ for a random variable (RV) X, where $\overline{x}$ denotes the expectation or median of X, have been among the main tools of probability theory. These inequalities are known as concentration inequalities, and they have been subject to interesting developments in probability theory. Very roughly speaking, the concentration of measure phenomenon can be stated in the following simple way: "A random variable that depends in a smooth way on many independent random variables (but not too much on any of them) is essentially constant" [75]. The exact meaning of such a statement clearly needs to be clarified rigorously, but it will often mean that such a random variable X concentrates around $\overline{x}$ in a way that the probability of the event $\{|X - \overline{x}| > t\}$ decays exponentially in t (for $t \geq 0$). The foundations in concentration of measures have been introduced, e.g., in [3, Chapter 7], [15, Chapter 2], [16], [42], [47], [48, Chapter 5], [50], [74] and [75]. Concentration inequalities are also at the core of probabilistic analysis of randomized algorithms (see, e.g., [3], [23], [50] and [61]).

The Chernoff bounds provide sharp concentration inequalities when the considered RV X can be expressed as a sum of n independent and bounded RVs. However, the situation is clearly more complex for non-product measures where the concentration property may not exist. Several techniques have been developed to prove concentration of measures. Among several methodologies, these include Talagrand's concentration inequalities for product measures (e.g., [74] and [75] with some information-theoretic applications in [40] and [41]), logarithmic-Sobolev inequalities (e.g., [23, Chapter 14], [42, Chapter 5] and [47] with information-theoretic aspects in [37], [38]), transportation-cost inequalities which originated from information theory (e.g., [23, Chapters 12, 13] and [42, Chapter 6]), and the martingale approach (e.g., [3, Chapter 7], [50] with information-theoretic aspects in, e.g., [45], [60], [61], [81]). This chapter mainly considers the last methodology, focusing on discrete-time martingales with bounded jumps.

The Azuma-Hoeffding inequality is by now a well-known methodology that has been often used to prove concentration phenomena for discrete-time martingales whose jumps are bounded almost surely. It is due to Hoeffding [34] who proved this inequality for $X = \sum_{i=1}^{n} X_i$ where $\{X_i\}$ are independent and bounded RVs, and Azuma [7] later extended it to bounded-difference martingales. It is noted that the Azuma-Hoeffding inequality for a bounded martingale-difference sequence was extended to centering sequences with bounded differences [51]; this