

On the Rényi Divergence and the Joint Range of Relative Entropies

Igal Sason

Abstract

This paper starts with a study of the minimum of the Rényi divergence subject to a fixed (or minimal) value of the total variation distance. Relying on the solution of this minimization problem, we determine the exact region of the points $(D(Q||P_1), D(Q||P_2))$ where P_1 and P_2 are any probability distributions whose total variation distance is not below a fixed value, and the probability distribution Q is arbitrary (none of these three distributions is assumed to be fixed). It is further shown that all the points of this convex region are attained by a triple of 2-element probability distributions. As a byproduct of this characterization, we provide a geometric interpretation of the minimal Chernoff information subject to a minimal total variation distance.

Keywords: Chernoff information, Lagrange duality, relative entropy, Rényi divergence, total variation distance.

I. INTRODUCTION

The Rényi divergence, introduced in [22], has been studied so far in various information-theoretic contexts (and it has been actually used before it had a name [24]). These include generalized cutoff rates for hypothesis testing ([1], [7]), generalized guessing moments [9], strong converse theorems for classes of networks [11], channel coding error exponents ([14], [19], [24]), strong data processing theorems for discrete memoryless channels [20], two-sensor composite hypothesis testing [25], and one-shot bounds for various information-theoretic problems [29].

This work starts with a study of the minimum of the Rényi divergence subject to a fixed (or minimal) value of the total variation distance. The derivation of an exact expression for this minimum is initialized by adapting some arguments that have been used by Fedotov *et al.* [10] for the minimization of the relative entropy (a.k.a. Kullback-Leibler divergence), subject to a fixed value of the total variation distance. Our analysis further relies on the Lagrange duality and a solution of the Karush-Kuhn-Tucker (KKT) equations, while asserting strong duality for the studied problem. The use of Lagrange duality significantly simplifies the computational task of the studied minimization problem. The exact expression for the Rényi divergence generalizes, in a non-trivial way, previous studies of the minimization of the relative entropy under the same constraint on the total variation distance (see [10], [15], [21]). The exact expression for this minimum is also compared with known Pinsker-type lower bounds on the Rényi divergence [16] when the total variation distance is fixed. It should be noted that the studied problem minimizes the Rényi divergence w.r.t. all pairs of probability distributions with a total variation distance which is not below a given value; this differs from the type of problems studied in [3] and [18], in connection to the minimization of the relative entropy $D(P||Q)$ with a minimal allowed value of the total variation distance, where the probability distribution (PD) Q was fixed.