Geometric losses for distributional learning

Arthur Mensch, École Normale Supérieure Paris

Abstract

Building upon recent advances in entropy-regularized optimal transport, and upon Fenchel duality between measures and continuous functions, we propose a generalization of the logistic loss that incorporates a metric or cost between classes. Unlike previous attempts to use optimal transport distances for learning, our loss results in unconstrained convex objective functions, supports infinite (or very large) class spaces, and naturally defines a geometric generalization of the softmax operator. The geometric properties of this loss make it suitable for predicting sparse and singular distributions, for instance supported on curves or hyper-surfaces. We study the theoretical properties of our loss and show-case its effectiveness on two applications: ordinal regression and drawing generation.