The Catch-Up Phenomenon in Bayesian Inference

Peter Grünwald
CWI, Amsterdam, The Netherlands
Peter.Grunwald@cwi.nl

Abstract

Standard Bayesian model selection/averaging sometimes learn too slowly: there exist other learning methods that lead to better predictions based on less data. We give a novel analysis of this "catch-up" phenomenon. Based on this analysis, we propose the switching method, a modification of Bayesian model averaging that never learns slower, but sometimes learns much faster than Bayes. The method is related to expert-tracking algorithms developed in the COLT literature, and has time complexity comparable to Bayes.

The switching method resolves a long-standing debate in statistics, known as the AIC-BIC dilemma: model selection/averaging methods like BIC, Bayes, and MDL are consistent (they eventually infer the correct model) but, when used for prediction, the rate at which predictions improve can be suboptimal. Methods like AIC and leave-one-out cross-validation are inconsistent but typically converge at the optimal rate. Our method is the first that provably achieves both. Experiments with nonparametric density estimation confirm that these large-sample theoretical results also hold in practice in small samples.