Parameter clustering: encouraging similarities between estimates via euclidean distance regularization

Tso-Jung Yen

Institute of Statistical Science, Academia Sinica

Abstract

In statistical estimation, one important goal is to obtain a model that has better ability in prediction but fewer parameters for interpretation. Such parsimony requirement leads statisticians to develop various techniques for reducing the effective number of parameters in the model. In this paper we propose a penalized estimation method to fulfill this requirement. The method aims to reduce the effective number of parameters by estimating parameters with identical values. It imposes l_2 -norm penalty functions on differences between pairs of the parameters. Under this setting, the method is able to shrink the differences to zero, yielding identical estimates for the parameters. To numerically carry out the method, we first formulate the problem as a constrained optimization problem, and then solve the constrained optimization problem by developing an iterative algorithm based on the alternating direction method of multipliers. Simulation studies show that the method can simultaneously identify the number of effective parameters and deliver collaborative estimates for these parameters. We discuss several applications and a proposal for carrying out this method via distributed optimization.

Keywords: Euclidean distance; Fused lasso; l_2 -norm regularization; Alternating direction method of multipliers; Block splitting algorithms.