MODEL SELECTION IN LINEAR REGRESSION USING THE MML CRITERION

R A Baxter, D L Dowe

ABSTRACT

We use the minimum message length (MML) criterion to develop a data-based method for model selection in regression problems. We illustrate its application to choosing the degree of a polynomial in least-squares regression. Estimates of the minimum message lengths of model and data given the model are made. The MML criterion proposed here is contrasted with similarly motivated criteria developed using the minimum description length (MDL) principle. The two dominant terms of the MML criterion are the same as Bozdogan's CAICF extension of Akaike's information criterion (AIC). The empirical performance of the criteria are studied in Monte Carlo experiments under various conditions.

The abstract of this paper appeared in [BD94].