Upgrading Model Selection Criteria with Goodness of Fit Tests for Practical Applications

Riccardo Rossi , Andrea Murari, Pasquale Gaudio and Michela Gelfusa

Abstract:

The Bayesian information criterion (BIC), the Akaike information criterion (AIC), and some other indicators derived from them are widely used for model selection. In their original form, they contain the likelihood of the data given the models. Unfortunately, in many applications, it is practically impossible to calculate the likelihood, and, therefore, the criteria have been reformulated in terms of descriptive statistics of the residual distribution: the variance and the mean-squared error of the residuals. These alternative versions are strictly valid only in the presence of additive noise of Gaussian distribution, not a completely satisfactory assumption in many applications in science and engineering. Moreover, the variance and the mean-squared error are quite crude statistics of the residual distributions. More sophisticated statistical indicators, capable of better quantifying how close the residual distribution is to the noise, can be profitably used. In particular, specific goodness of fit tests have been included in the expressions of the traditional criteria and have proved to be very effective in improving their discriminating capability. These improved performances have been demonstrated with a systematic series of simulations using synthetic data for various classes of functions and different noise statistics.