Comparison of Second-order Polynomial Model Selection Methods: an Experimental Survey

Grace W Rumantir
School of Computer Science and Software Engineering
Monash University – Clayton Vic 3168 Australia
gwr@csse.monash.edu.au

This abstract gives an overview of the work described in (Rumantir 1999). The paper compares some of the most commonly cited model selection criteria that claim to have the mechanism to balance between model complexity and goodness of fit. The model chosen by any of the methods is claimed to be a parsimonious description of the data at hand, therefore has predictive power for future data.

This work tests the robustness of each method in support of the above claim in performing two tasks. First, to recover models from artificial data generated from known true models. Second, to parsimoniously select models from real data for the purpose of forecasting hurricane intensity change. The models selected are then compared with the benchmark models being used in operation.

In addition to polynomials with single independent variables, the products of variables up to the second order are also considered. A method based on the Minimum Message Length (MML) principle has been formulated and is compared with the methods found in the literature. A common gradient descent search strategy has been developed and is used with all of the model selection criteria. The summary of the methods examined is given in the table below.

Method		Reference
Minimum Message Length	MML	Wallace 1987
Minimum Description Length	MDL	Rissanen 1978
Stochastic Complexity	SC	Rissanen 1987
Akaike's Information Criterion	AIC	Akaike 1973
Corrected AIC	CAICF	Bozdogan 1987
Bayesian Information Criterion	BIC	Schwarz 1978
Structured Risk Minimisation	SRM	Vapnik 1995
Mallows' C_p	C_p	Mallows 1973
F-to-enter & F-to-exit	F-test	in Miller 1990
Adjusted Coeff. of Determination	adjR^2	Ezekiel 1930
Generalized Cross Validation	GCV	Wahba 1979
Predictive sum of squares	PRESS	Allen 1974

Based on the experiments with artificial data and real atmospheric data for hurricane intensity change forecasting, it is shown that MML, MDL, CAICF, SRM and SC methods are good candidates for fully automated model selection tasks. Given a noisy data set, the methods can reliably converge to a true model (if one exists) or to a reasonably parsimonious model.

AIC, BIC, C_p, F-test and adjR^2 have overfitted the training data which suggests that when comparing two models with different complexity, the increase in the penalty terms for model complexity is not sufficient compared to the decrease in the terms for goodness of fit. This prompted the doubt that the balancing mechanism of the methods might not be robust enough for automated model selection task.

PRESS and GCV have overfitted the test data which not only suggests that their performance on test data cannot guarantee convergence into parsimonious models, but also necessitates the availability of a third independent data set for validation of model selection.

For the task of building models for hurricane intensity change forecasting, the methods MML, MDL, CAICF, SRM and SC come up with models with superior predictive performance than the standard benchmark models used in operation. The results emphasize two things. First, the importance of having homogenous data for training and test data sets for a model selection method to pick up regularities in the training data that can be extrapolated into the test data set and beyond. Second, that the unavoidable practice of using non-exhaustive search strategy on large search space pronounces the influence of selection bias in determining to which local minimum a model selection method would converge.

Acknowledgments

The atmospheric data discussed in the paper was generously supplied by Chris Landsea of the Hurricane Research Centre NOAA, Miami Florida. The author is grateful to Chris Wallace for guidance in the development of the MML method. The author is a recipient of the Australian Postgraduate Award (Industry).

References

Rumantir, G. W. 1999. Comparison of second-order polynomial model selection methods: an experimental survey. Technical report, School of Computer Science and Software Engineering, Monash University.