The University of Memphis
Computer Science Division
Computational Neurodynamics Seminar

Monday, April 21, 2003, 2 p.m.
Room 245, Dunn Hall

Model Structures Selection for Time Series Prediction

Amaury Lendasse
Université Catholique de Louvain, Belgium

Abstract

Time series forecasting is a great challenge in many fields. In finance, one forecasts stock exchange courses or indices of stock markets; data processing specialists forecast the flow of information on their networks; producers of electricity forecast the load of the following day. The common point to their problems is: how can one analyze and use the past to forecast the future? Many techniques exist, including linear methods (ARX, ARMA…) and nonlinear methods (artificial neural networks). It is thus necessary to compare the various models (for example with regards to their performances and complexity) and choose the best one. The ranking of the models is made according to some criterion like the generalization error, usually defined as the average error that a model would make on an infinite-size and unknown test set independent from the learning one. In practice the generalization error can only be estimated, but there exists some methods to provide such an estimation: the AIC or BIC criteria as well as other well-known statistical techniques: the cross-validation and k-fold, the leave-one-out, the bootstrap and its unbiased extension the .632 bootstrap. The seminar is in two parts:

1) We will compare model selection methods, based on experimental estimates of their generalization errors. Experiments in the context of nonlinear time series prediction by Radial-Basis Function Networks show the superiority of the bootstrap methodology over classical cross-validations and leave-one-out.
2) The bootstrap main limitation in practice is the computation time required for assessing an approximation of sufficient reliability (or *accuracy*). We propose a fast approximation of the generalization error using the bootstrap, based on linear approximations of the optimism. We prove experimentally the validity of the linear approximation of the optimism and the links between this approximation and BIC (or AIC) criteria. Finally, we show how to use this approximation to perform efficient bootstrap simulations with reasonable computational complexity.

Amaury Lendasse is a teaching assistant in the Applied Mathematics Department in the Universite Catholique de Louvain. His research concerns time series prediction, Kohonen maps, nonlinear projections, nonlinear approximators, model selection and electricity load prediction. He is author or co-author of about 30 scientific papers in international journals or communications to conferences with reviewing committee. He is currently working towards a Ph.D. degree at the Centre for Systems Engineering and Applied Mechanics (CESAME) in the Universite Catholique de Louvain.