Neural Network Model for Nonlinearity Detection

Lehel Csató and Péter András

A neural network based algorithm to detect nonlinearities in a specified timeseries is presented. This method tests whether a multilayered neural network can detect nonlinearities in the application domain where it is used (sample data from the future application are given). The use of this test is useful in domains that require huge amounts of computation (in the applications the neural networks are equipped with many hidden neurons). If the neural network gives no significant contribution to the system output, then by eliminating the nonlinear part, a comparative and less "expensive" architecture is obtained.

The aim of this model is twofold: first to give a method to decide which model to use when implementing a forecasting algorithm, and second to combine the linear and nonlinear techniques in order to achieve a better prediction performance.

The network is built up using linear approximators (direct linear input-output connections in neural network terminology) and a nonlinear part. The neural network has the following form:

$$y = A \cdot x + R \cdot f(Q \cdot x) \tag{1}$$

where A is the autoregressive term; Q and R specify the nonlinear part; f is the type of nonlinearity that is used (sigmoid); x and y the system input and output respectively.

The derivation of the learning rules follows the anti-gradient method based on the regular energy (cost) function. The direct connections (A) and the nonlinear components (Q, R) are trained simultaneously.

The learning rules show that in the training phase the learning components (linear and nonlinear) are clearly separated, thus they may run with different training speeds. Applying second order methods to derive learning rules led us to some kind of cooperation between the linear and the nonlinear part.

Since one of the interested and still not formalized area of possible applications is the financial forecasting, the defined architecture has been used to forecast financial timeseries. Numerical applications validate the assumption: in stable markets (NYSE indexes) no significant contribution of the nonlinear term was found, indicating that the nonlinearities that are eventually present cannot be detected by this method. In the second case, with different data taken from an evolving market (BUX indexes), the nonlinear part gave substantial contribution to the network output.