Artificial Neural Network for Returns

Application of Non-Linear TSA in Empirical Finance

Alexandra Kim & Markus Mayer

Department of Economics
University of Vienna

January 19, 2010

Preliminaries 1/3

- Artificial Neural Networks (ANNs) or just Neural Networks (NNs) simulate the structure of biological neural networks.
- Network of simple processing elements (neurons), which can exhibit complex global behavior.
- In the artificial case, a function is defined as a composition of other functions, which can further be defined as a composition of functions.

Preliminaries 2/3

- Pro: Able to approximate almost any nonlinear function arbitrarily close. If a time series is characterized by a truly nonlinear dynamic relationship, the ANN will detect these and provide a superior fit, compared to linear time series models.
- Pro: No need to construct a specific parametric nonlinear time series model.

Preliminaries 3/3

- Con: Parameters are difficult, if not possible to interpret, therefore ANNs are often considered as 'black box' models, and constructed mainly for the purpose of pattern recognition and forecasting.
- Con: Danger of overfitting. By incresing the flexibility of the model, it is possible to obtain a almost perfect in-sample fit, but this can only be achieved by fitting the irregular noise of the time series. The result can be a inferior out-of-sample forecast.

universität wien

ANN Terminology 1/2

- We consider the 'single hidden layer feedforward' model, which is the most popular among time series practitioners.
- The network is seen to consist of three different layers:
  - Input layer, consisting of explanatory variables in x_t.
  - The inputs are multipied by so-called connection strengths $\gamma_{i,j}$ as they enter the hidden layer, which consists of q logistic functions $G(\cdot)$.
  - In the hidden layer the linear combinations $x_t'\gamma_j$ are formed and transformed into a value between 0 and 1 by the activation functions $G(\cdot)$. These are multiplied by weights β_j to produce the y_t of the output layer.

ANN Terminology 2/2

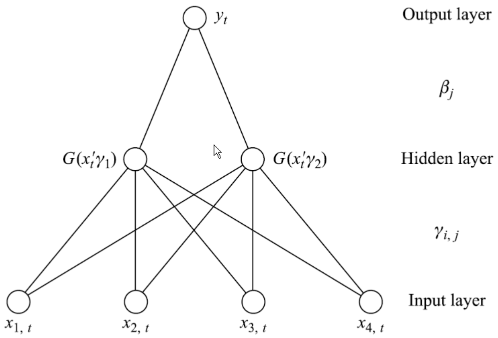

Figure: Single hidden layer feedforward neural network ANN (k,q) with k=4 and q=2.

Consider the following STAR model for a univariate time series y_t,

$$y_t = \phi_0 + \beta_1 G(\gamma[y_{t-1} - c]) + \epsilon_t,$$

where $G(\cdot)$ is the logistic function

$$G(z) = \frac{1}{1 + exp(-z)}$$

An ANN can now be obtained by assuming that the conditional mean of y_t depends on the value of a linear combination of p lagged values $y_{t-1}, ..., y_{t-p}$ relative to the threshold c.

Represenation 2/2

- The model then becomes

$$y_t = \phi_0 + \beta_1 G(x_t' \gamma_1) + \epsilon_t,$$

- The possible nonlinear relationship between y_t and x_t can be modelled by including additional logistic components, which yields

$$y_t = \phi_0 + \sum_{j=1}^{q} \beta_1 G(x_t' \gamma_1) + \epsilon_t.$$

Illustration of Approximation 1/2

- It can be shown, that an ANN of this form can approximate any function arbitrarily close, provided that the number of nonlinear components q is sufficiently large (without proof).
- We consider the last network and assume that only y_{t-1} acts as an input, therefore $x_t = (1, y_{t-1})'$.
- The next slide shows the skeleton of such a network with
  - $\phi_0 = 2,\ q = 3,\ \beta_1 = 8,\ \beta_2 = -12,\ \beta_3 = 6$
  - $G(x_t'\gamma_1) = 1/(1 + exp[-40 - 10y_{t-1}])$
  - $G(x_t'\gamma_2) = 1/(1 + exp[-y_{t-1}])$
  - $G(x_t'\gamma_3) = 1/(1 + exp[20 - 20y_{t-1}])$

Illustration of Approximation 2/2

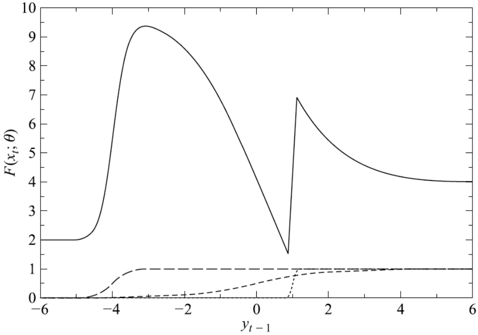

Figure: Skeleton $F(x_t;\theta)$ of an ANN with a single input and $q = 3$ (solid line); values of the activation functions $G(x_t'\gamma_j), j = 1, 2, 3$ are shown on the horizontal axis.

Estimation 1/2

- The parameters in the ANN (k,q) model

$$y_t = x_t'\phi + \sum_{j=1}^{q} \beta_j G(x_t'\gamma_j) + \epsilon_t.$$

can be estimated by minimizing the residual sum of squares function

$$Q_n(\theta) = \sum_{t=1}^{n} [y_t - F(x_t;\theta)]^2,$$

where

$$F(x_t;\theta) = x_t'\phi + \sum_{j=1}^{q} \beta_j G(x_t'\gamma_j).$$

Estimation 2/2

- ANN are not considered as the result of an underlying data generating process. Because we see them as approximating models, they are inherent misspecified.
- Properties of the nonlinear least squares estimator $\hat{\theta}_n$:
  - $\hat{\theta}_n$ converges to θ^* as the sample size n increases without bound.
  - The normalized estimator $\sqrt{n}(\hat{\theta}_n - \theta^*)$ converges to a multivariate normal distribution with mean zero and a covariance matrix, that can be estimated.

universität wien

Model Evaluation and Model Selection

- Implementing an $ANN(p, q)$ requires several decisions to be made:
  - choosing the activation function $G(\cdot)$
  - choosing the number of hidden units q
  - choosing the number of lags p to use as input variables
- In most cases the logistic function is choosen as activation function.
- There are various strategies for choosing p and q. One is to estimate all possible models and select the most appropriate one with the help of selection criteria like AIC.

Forecasting

- A 1-step-ahead forecast y_{n+1} can be computed directly as $\hat{y_{t+1}}|t = x_t'\phi + \sum_{j=1}^{q} \beta_j G(x_t'\gamma_j)$ where $x_t = (1, y_t, ..., y_{t-p+1})$.
- There exists no closed-form expression for multiple-step-ahead forecasts $\hat{y_{t+h}}|t$ where $h > 1$. In this case, we have to rely on simulation techniques.
- Again, the main danger of forecasting with ANNs is overfitting. A method to limit this danger is crossvaliadation.
- Methods for forecast evaluation of ANNs are not different from other models, therefore the main criteria are MSPE & MAPE.

Software for ANN

- All programs, which are used for statistical analysis like SAS, SPSS, Excel and Stata can be used to implement ANNs.
- Besides programs, which are developed especially for ANN, the most commonly used programs for ANNs are MATLAB and recently R.
- We employ the MATLAB code of Shapour and Hossein, which is accessible at IDEAS: A Matlab Code for Univariate Time Series Forecasting (2005)[1]

[1] http://ideas.repec.org/c/wpa/wuwppr/0505001.html

Univariate Time Series Forecasting

- The data we employ is exchange rate data from the ECB database for the Australian Dollar against the Euro on a daily basis from 16 January 2009 to 15 January 2010. [2]
- We want to conduct a 5-step ahead forecast.

[2] http://sdw.ecb.europa.eu/browseSelection.do?DATASET=0FREQ=DCURRENCY=AUDnode=2018794

Appendix: References

- Fan, J., and Yao, Q., (2003): *Nonlinear Time Series: Nonparametric and Parametric Methods.* Springer, New York.
- Frances, P. H., and van Dijk, D., (2000): *Non-Linear Time Series Models in Empirical Finance.* Cambridge University Press, Cambridge.