Nonlinear Forecasting with Many Predictors by Neural Network Factor Models

Ali Habibnia
Department of Statistics, London School of Economics

Abstract

- This study proposes a nonlinear generalisation of factor models based on artificial neural networks for forecasting financial time series with many predictors. This model would be able to capture both non-linearity and non-normality of a high-dimensional dataset.
- Specification (architecture) of the neural network factor model is determined on the basis of statistical inference and special emphasis is given to data-driven specification.
- Linear factor models can be represented as a special case of this neural network factor model.It means that, if there is no non-linearity between variables, it will work like a linear model.

Construction of FNN

Based on the universal approximation theorem, a single hidden layer feedforward network can approximate arbitrary well any continuous function of n real variables. To show that the neural network models can be seen as a generalization of linear models, we allowed for direct connections from the input variables to the output layer and we assumed that the output transfer function $\{\phi_o(.)\}$ is linear, then the model becomes

$$y_t = \beta_o + \sum_{i\to o} w_{io} x_{it} + \sum_{j\to o} w_{jo}\phi_j\left(\beta_j + \sum_{i\to j} w_{ij}x_{it}\right) + \varepsilon_t,$$

Studies in time series widely used the conventional feedforward neural network trained with the back-propagation algorithm, however, the backpropagation is not an efficient algorithm and it converges slowly. Therfore, Levenberg-Marquardt algorithm and its optimized version with LASSO are implemented in this study which are fast and have stable convergence.

Results

Out-of-sample forecast evaluation results based on different criteria (RMSE, Hit-Rate and Theil) showed that the proposed neural network factor model (NNFM) significantly outperformed linear factor model and Random-Walk approach.

Introduction

Forecasting with factor models are a two-step process:

- **Factor Estimation**, which summarizes the information contained in a large data set in a small number of factors.

$$\underbrace{X_{it}}_{observations} = \Lambda_i \underbrace{f_t}_{factors} + \underbrace{\xi_{it}}_{idiosyncratic} \quad (1)$$

- **Forecasting Equation**, which is the prediction of the variable of interest by using common factors.

$$\underbrace{y_{t+1|t}}_{one\ of\ the\ X} = \lambda\, f_{t+1|t} + \varepsilon_{t+1} \quad (2)$$

Common factors and the idiosyncratic component can be forecast simultaneously or separately.

Financial Forecasting

Financial returns present special features and share the following stylised facts: comovements, non-linearity, non-gausianity (skewness and heavy tails) and leverage effect, which makes the modelling of this variable hard.

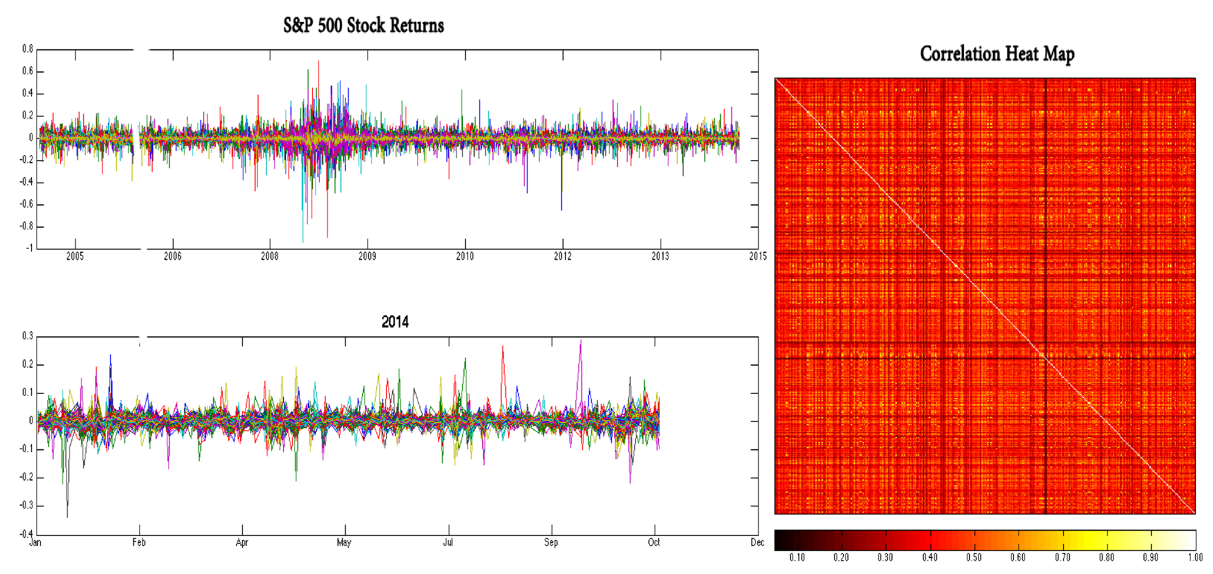


Figure 4: Daily return observations of the 419 companies in S&P500 index

Contribution & Formulation

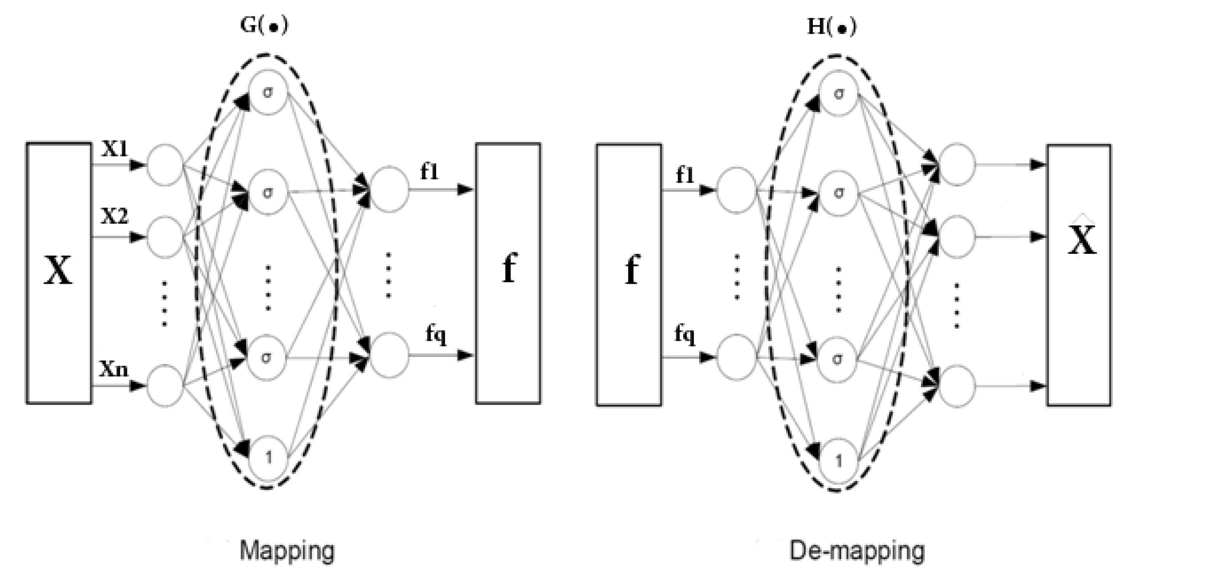


Figure 1: The standard auto-associative neural network architecture for nonlinear PCA (combination of two feed-forward NNs)

- Second extension is a nonlinear factor augmented forecasting equation based on a feed-forward neural network model which can be built in a similar fashion as a statistical model.

$$y_t = \phi_o\left[\beta_o + \sum_{h\to o} w_{ho}\phi_h\left(...\phi_j\left(\beta_j + \sum_{i\to j} w_{ij}x_{it}\right)\right)\right] + \varepsilon_t,$$

- where x_{it} is the value of the ith input node. $\phi_h(.)$, $\phi_j(.)$ and j, h are activation functions (commonly used function is a sigmoidal function)and number of nodes (neurons) used at the hidden layers.

The first extension proposes a **NLPCA (neural network principal component analysis)** as an alternative for factor estimation, which allows the factors to have a nonlinear relationship to the input variables. NLPCA nonlinearly generalizes the classical PCA method by a nonlinear mapping from data to factors. Both neural network parameters and unobservable factors (f) can be optimised simultaneously to minimise the reconstruction error e:

$$e = \hat{X} - X, \quad MSE = E(||\hat{X}(f) - X||^2)$$

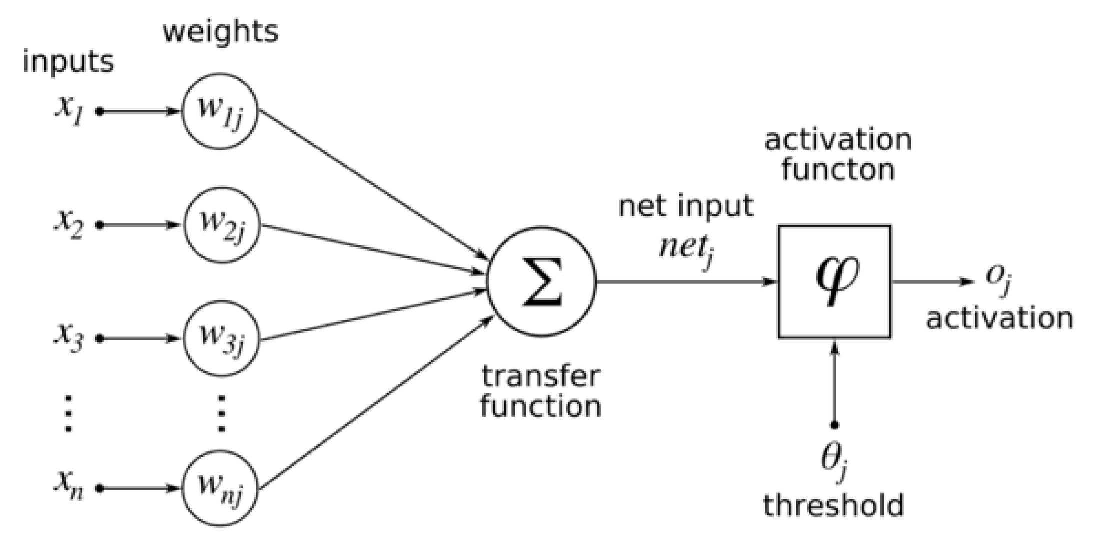
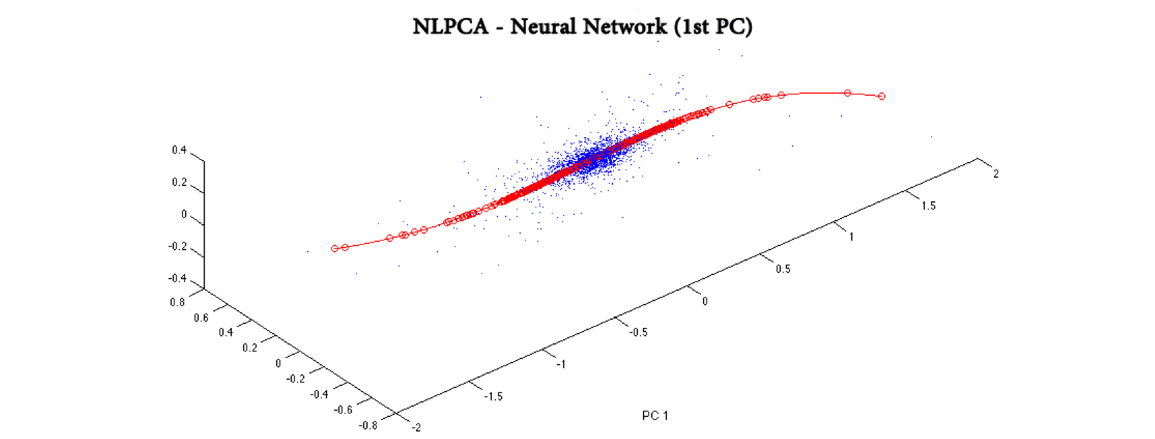


Figure 2: Artificial Neuron configuration

NLPCA on Financial Returns



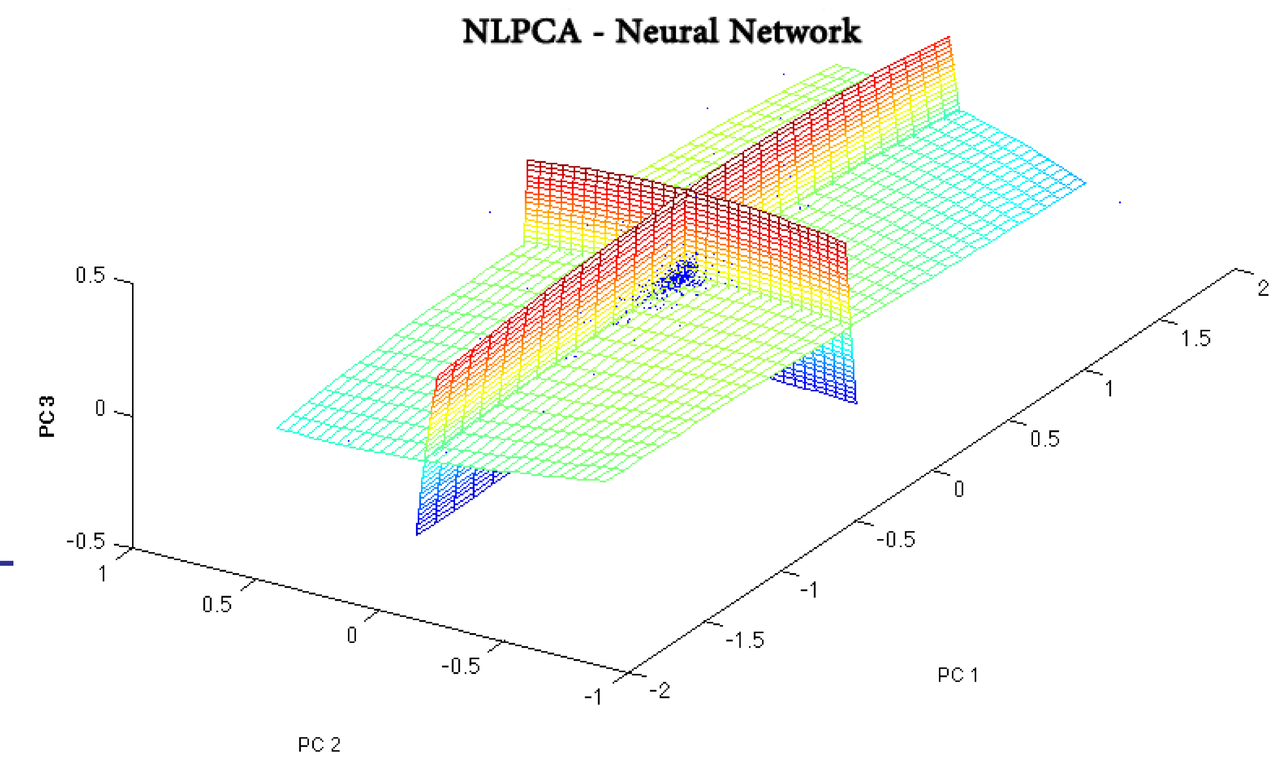


Figure 3: Nonlinear PCA can describe the inherent structure of the data by a curved subspace.

References

[1] J. H. Stock and M. W. Watson. Forecasting using principal components from a large number of predictors. *American Statistical Association*, 97:1167–1179, 2002.

[2] M. Forni et al. The generalized dynamic factor model: Estimation and forecasting. *American Statistical Association*, 100:830–840, 2005.

[3] M. C. Medeiros and T. Terasvirta. Building neural network models for time series: A statistical approach. *J of Forecasting*, 25:49–75, 2006.

[4] C. M. Kuan and H. White. Artificial neural networks: An econometric perspectiv. *Econometric Reviews*, 13:1–91, 1994.

[5] M. Deistler and E. Hamann. Identification of factor models for forecasting returns. *Financial Econometrics*, 3(2):256–281, 2005.

[6] A. N. Gorban and B. M. Kegl(Eds.). *Principal Manifolds for Data Visualization and Dimension Reduction*. Springer, 2008.

[7] M. A. Kramer. Nonlinear principal component analysis using autoassociative neural networks. *AIChE*, 37:233–243, 1991.

Contact Information

Web http://personal.lse.ac.uk/habibnia/
Email a.habibnia@lse.ac.uk
Phone +44 (0)7737842985