

Measuring Linear and Nonlinear Associations

Masoud Yarmohammadi

Department of Statistics, Payame Noor University, 19395-4697 Tehran, I.R. of Iran

Abstract: In this study, we propose a new approach based on two nonparametric techniques to capture the linear and nonlinear associations. The singular spectrum analysis technique, which is a powerful method for filtering noisy series, is used as a noise reduction method and mutual information is considered for measuring the level of association. The performance of the proposed approach is assessed using the simulated and real time series.

Keywords: Association, mutual information, noise reduction, singular spectrum analysis

INTRODUCTION

Various measures have been considered for measuring the degree of association. The most known measure is the coefficient of linear correlation, but its application, for example, requires a pure linear relationship. This statistics may not be helpful in determining serial dependence if there is some kind of nonlinearity in the data (Granger and Lin, 1994).

It has been shown that a measure based on the mutual information, which captures linear and nonlinear dependencies, without requiring the specification of any kind of assumptions, is better than the linear correlation coefficient to measure association and serial correlation of noisy time series, particularly financial series (Hassani *et al.*, 2009a).

It is clear that the existence of a significant noise level reduces the level of accuracy for measuring association between two series. For example, consider a noisy time series $y_t = s_t + \varepsilon_t$ ($t = 1, \dots, T$) where s_t is a deterministic part and ε_t its stochastic part. Usually the second part is considered as noise.

In this study, we mainly consider two different approaches to obtain the level of association. In the first approach, we capture the level of association directly from noisy series. Thus, we do not consider the existence of the noise level in the first approach. However, in the second approach we first filter the noisy time series, in order to reduce the noise level, and then calculate the measures. It is obvious that the obtained results by the second approach are more effective than the first one if we select a proper method for filtering the series. Here, we provide necessary theoretical background and concisely describe the singular spectrum analysis technique and mutual information, respectively. Using these techniques, we propose an association test for measuring the linear and nonlinear association between two series. The test is based on the filtering approach.

Singular spectrum analysis (SSA): In recent years SSA as a powerful technique of time series analysis has been

developed and applied to many practical problems (Hassani, 2007; Hassani, 2009; Hassani *et al.*, 2009a-c; Hassani and Zhigljavsky, 2009e; Ghodsi *et al.*, 2009; Hassani and Thomako, 2010; Mahmoudvand and Zokaei, 2011). A thorough description of the theoretical and practical foundations of the SSA technique (with several examples) can be found in (Golyandina *et al.*, 2001).

It should be noted that despite the fact that a lot of probabilistic and statistical elements are employed in the SSA-based technique but the technique does not make any statistical assumption concerning either signal or noise while performing the analysis and investigating the properties of the algorithms (Hassani, 2007). This matter can be considered as one of the advantages of the technique against other classical methods which usually rely on some restricted assumptions.

The SSA technique consists of two complementary stages: decomposition and reconstruction and both of which include two separate steps. The original time series is decomposed into a number of additive time series, each of which can be easily identified as being part of the modulated signal, or as being part of the random noise. This is followed by a reconstruction of the original series. Here, we mainly follow Hassani (2007).

Consider the real-valued non-zero time series $Y_T = (y_1, \dots, y_T)$ of sufficient length T . Let $K = T - L + 1$, where L ($L \leq T/2$) is some integer called the window length. Define the matrix

$$X = (x_{ij})_{i,j=1}^{L,K} = [X_1, \dots, X_K] \quad (1)$$

Where $X_j = (y_j, \dots, y_{L+j-1})^T$. We then consider X as multivariate data with L characteristics and $K = T - L + 1$ observations. The columns $X_j = (y_j, \dots, y_{L+j-1})^T$ of X , considered as vectors, lie in an L -dimensional space R_L . Define the matrix XX^T . Singular value decomposition (SVD) of XX^T provides us with the collections of L eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$ and the corresponding eigenvectors U_1, \dots, U_L where U_i is the normalized eigenvector corresponding to the eigenvalue λ ($i = 1, \dots, L$).

A group of r (with $1 \leq r < L$) eigenvectors determine an r -dimensional hyperplane in the L -dimensional space R^L of vectors X_j . If we choose the first r eigenvectors U_1, \dots, U_r , then the squared L_2 -distance between this projection and X is equal to $\sum_{j=r+1}^L \lambda_j$. According to the Basic SSA algorithm, the L -dimensional data is projected onto this r -dimensional subspace and the subsequent averaging over the diagonals allows us to obtain an approximation to the original series.

Mutual information: The mutual information, which is based on the concept of entropy, is a useful technique to capture the relation, either linear or nonlinear, between two series. Below, we provide a brief introduction and mainly follow Hassani *et al.* (2009a). The mutual information of two continuous random variables X and Y can be defined as:

$$I(X;Y) = \iint_{X,Y} P(x,y) \log\left(\frac{P(x,y)}{P(x)P(y)}\right) d_y d_x \quad (2)$$

Where $p(x, y)$ is the joint probability distribution function of X and Y , and $p(x)$ and $p(y)$ are the marginal probability distribution functions of X and Y , respectively. In the discrete case, we replace the integral by a definite double summation. Intuitively, mutual information measures the information that X and Y share: it measures how much knowing one of these variables reduces our uncertainty about the other. Mutual information can be expressed as:

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = H(X) + H(Y) - H(X, Y) \quad (3)$$

Where $H(X)$ and $H(Y)$ are the marginal entropies, $H(X|Y)$ and $H(Y|X)$ are the conditional entropies, and $H(X, Y)$ is the joint entropy of X and Y .

Since $H(X) \geq H(X|Y)$, we have $I(X;Y) \geq 0$; assuming equality iff X and Y are statistically independent. Therefore, the mutual information between the vectors of random variables X and Y can be considered as a measure of dependence between these variables, or better yet, the statistical correlation of X and Y . The statistics defined in Eq. (3) satisfies some of the desirable properties of a good measure of dependence (Hassani *et al.*, 2009a).

The mutual information defined in Eq. (3) takes a value between 0 and infinity, which makes the comparisons difficult between different samples. In this context, the following equality has been used by Granger and Lin (1994) as a standard measure for the mutual information:

$$\lambda = (1 - \exp[-2I(X, Y)])^{\frac{1}{2}} \quad (4)$$

Table 1: The values of λ for different noise levels

	σ_1	σ_2	σ_3	σ_4
Noisy He'non	0.843	0.872	0.891	0.925
ARMA	0.682	0.701	0.729	0.768
GARCH	0.661	0.686	0.711	0.727
Proposed approach	0.618	0.624	0.635	0.648

Note that λ captures the overall dependence, both linear and nonlinear, between X and Y . This measure varies between 0 and 1 being thus directly comparable to the linear correlation coefficient, ρ , based on the relationship between the measures of information theory and variance analysis. Note that here we are not interested in the direction of dependency. Thus, both linear and nonlinear measures are comparable. Our proposed approach consists of two separate stages, in the first stage we filter the noisy series using the SSA technique and in the second stage we apply Eq. (4) to measure the level of association.

EMPIRICAL RESULTS

Below, we shall consider two types of time series; real and artificially generated time series, financial data and a He'non map, respectively. We consider the values of λ before and after noise reduction. Let us first consider the performance of the proposed approach on the chaotic time was such as He'non map (He'non, 1976):

$$\begin{cases} x_{n+1} = 1 + y_n - Ax_n^2 \\ y_{n+1} = Bx_n \end{cases} \quad (5)$$

with usual parameter values: $A = 1.4$ and $B = 0.3$. Here, we follow the same approach considered in Hassani *et al.* (2009c). Therefore, we have generated 3000 observations. Furthermore, we have considered different noise levels to create noise series. Table 1 shows the values of λ before and after filtering for different noise levels (σ_1 : the smallest noise level and σ_4 the largest). Note that the values of λ before adding noise is 0.601, respectively. Therefore, the performance of proposed approach is promising if the captured values are close to 0.601. We also consider Autoregressive Moving Average (ARMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) as linear and nonlinear filtering methods. These models are considered as our benchmark models. We have considered GARCH model as it works well for nonlinear situation and also ARIMA as it is a well known linear model. Therefore, these benchmark models cover both linear and nonlinear situations.

As appears from Table 1, different noise levels give different values of λ . Table 1 confirm that the values of λ obtained from our proposed approach are more robust than the other methods that are considered here.

Table 2: The values of λ for the G7 countries in the first differences

	Canada	Japan	UK	Germany	France	Italy
US	0.82	0.11	0.54	0.35	0.46	0.40
Canada		0.18	0.46	0.39	0.47	0.41
Japan			0.34	0.38	0.33	0.34
UK				0.61	0.59	0.53
Germany					0.62	0.64
France						0.59

Furthermore, the results also show that the values of λ after filtering by ARMA, GARCH and proposed method are more accurate than the values of the noisy series. Similar results has been obtained by Hassani *et al.* (2009a).

Let us now consider the performance of the proposed approach for capturing association between two real series. Here, we mainly follow the approach considered by Menezes *et al.* (2011). The real data set used in our empirical analysis consists of 7 daily stock price series representing the G7 countries: US, Canada, Japan, UK, Germany, France and Italy. The data are the relative price indexes for these markets and cover the period from January, 1st 2000 to January, 1st 2010. All series are flatter than the Gaussian distribution and slightly skewed. The series also show some levels of non-stationarity.

The results in Table 2 indicate that there are different levels of relationships among these countries. The results also show that there are reasonable associations within the EU and North-American countries (the similar results has been concluded by Menezes *et al.* (2011)). However, if we consider these results for original series (not for the first differences) we observe very strong long-run relationships between all markets.

CONCLUSION

In this study, we proposed a new approach based on two nonparametric techniques to capture the linear and nonlinear associations. The proposed test was based on the filtering approach. The performance of the proposed approach using simulation and the real data sets indicated the capability of the proposed test in measuring association between two time series with different structures and features. Expanding this idea to the multivariate cases would be a new development in this area.

ACKNOWLEDGMENT

The author would like to thank the referees for constructive comments which have led to substantial improvements in this paper. This research was supported by a grant from Payame Noor University, Tehran-Iran.

REFERENCES

- Ghodsi, M., H. Hassani, S. Sanei and Y. Hicks, 2009. The use of noise information for detection of temporomandibular disorder. *J. Biomed. Signal Proces. Contr.*, 4(2): 79-85.
- Golyandina, N., V. Nekrutkin and A. Zhigljavsky, 2001. *Analysis of Time Series Structure: SSA and Related Techniques*. Chapman & Hall/CRC, New York-London.
- Granger, G. and J. Lin, 1994. Using the mutual information coefficient to identify lags in nonlinear models. *J. Time Ser. Anal.*, 15: 371-384.
- Hassani, H., 2007. Singular spectrum analysis: Methodology and comparison. *J. Data Sci.*, 5(2): 239-257.
- Hassani, H., 2009. Singular spectrum analysis based on the minimum variance estimator. *Nonlinear Anal-Real*, 11(3): 2065-2077.
- Hassani, H., A. Dionisio and M. Ghodsi, 2009a. The effect of noise reduction in measuring the linear and nonlinear dependency of financial markets. *Nonlinear Anal-Real*, 11(1): 492-502.
- Hassani, H., S. Heravi and A. Zhigljavsky, 2009b. Forecasting european industrial production with singular spectrum analysis. *Int. J. Forecasting*, 25(1): 103-118.
- Hassani, H., M. Zokaei, D. Von Rosen, S. Amiri and M. Ghodsi, 2009c. Does noise reduction matter for curve fitting in growth curve models?, *Comput. Meth. Prog. Bio.*, 96(3): 173-181.
- Hassani, H. and A. Zhigljavsky, 2009e. Singular spectrum analysis: Methodology and application to economics data. *J. Syst. Sci. Complex.*, 22(3): 372-394.
- Hassani, H. and D. Thomakos, 2010. A review on singular spectrum analysis for economic and financial time series. *Statistics Interface*, 3(3): 377-397.
- He'non, M., 1976. A two-dimensional mapping with a strange attractor. *Commun. Math. Phys.*, 50: 69-77.
- Mahmoudvand, R. and M. Zokaei, 2011. A Filter Based Correlation Coefficient by Using Singular Spectrum Analysis. *Proceeding of the 31th Interantional Syposium on Forecasting*. Prague, Czeck, June. 26-29.
- Menezes, R., A. Dionisio and H. Hassani, 2011. On the Globalization of Stock Market: An Application of VECM, SSA Technique and Mutual information to the G7. *Quarterly Review of Economics and Finance*, Forthcoming.