An Interval Technical Indicator for Financial Time Series Forecasting

Francisco J. Ruiz[2], Albert Samà[2], German Sánchez[1], José A. Sanabria[1], Núria Agell[1]
[1]ESADE- Universitat Ramon Llull, Campus Sant Cugat
Avda. de la Torre Blanca, 59. 08172 - Sant Cugat del Vallès, Spain;
[2]UPC-BarcelonaTech – Automatic Control Department
Neàpolis Building. Rambla de l'Exposició, 59-69. 08800 Vilanova i la Geltrú. Spain

Abstract

In this work we compare the performance of some standard technical indicators with an interval technical indicator, the *moving interval* (MI), for time series forecasting. MI has the advantage of taking into account the variability of data in the range considered and not only the average, like standard indicators do. However, the use of intervals as input variables require the use of regression methods able to handle with non Euclidean structures. The kernel approach is employed to this end. A recently introduced interval kernel is applied together with the moving interval indicator. The conclusion is that this indicator outperforms the forecasting performance of standard indicators.

1 Introduction

One of the main tasks associated to time series analysis, and financial time series in particular, is forecasting. It can be easily formulated as a multivariable regression problem, that is to find a function f: $R^p \rightarrow R$ such that $x_{i+p+h}=f(x_i,\ldots,x_{i+p-1})$ for every $i\in\{1,\ldots,N-p\}$, where N is the length of the time series, p is the number of input variables in the regression model, i.e. the number of values that may influence the future value, and h is the prediction horizon. If the function f is linear, this model represents an autoregressive model of order p that has been widely studied in the Statistics literature. Others nonlinear methods such as Multilayer Perceptron, Bayesian neural networks, radial basis functions, generalized regression neural networks, K-nearest neighbor regression, CART regression trees, support vector regression and Gaussian processes have also been used in the literature for time series forecasting (see [Ahmed et al 2010] for an exhaustive comparison among these methods).

In all such methods, it is necessary to perform some preprocessing steps to the raw data in order to reduce the complexity of the task and facilitate the interpretation of the results. The main preprocessing step consists on reducing the number of input variables i.e. considering only a reduced set of variables from the set $\{x_i,\ldots,x_{i+p-1}\}$. In this task, it seems appropriate to use of some technical indicators from Technical Analysis theory. Most of the best known technical indicators try to summarize the behavior of the time series by means of averages of past values. In this paper we analyze the suitability of some of these well known technical indicators to be used as input variables in the nonlinear regression problem. In addition, we also propose the use of an interval technical indicator, the moving interval. The advantage of this new technical indicator is that it does not only take into account the average of the past value, but also considers the variability of the time series. Despite this advantage, the use of interval variables does not allow standard regression methods and it is necessary to employ regression algorithms able to work with interval variables. We propose the kernel approach due to its capacity to handle with these kinds of data.

The rest of the paper is organized as follow. The next section is devoted to present some standard technical indicators and also the new interval indicator. In section 3, the kernel approach is briefly described focusing in the use of kernel methods with interval variables. The experiment by using foreign currency exchange rate data in exposed in section 4. Finally, the last section contains conclusions and implications of our results.

2 Technical indicators

The study of the internal information of financial time series such as stock prices or foreign exchange rates evolution is commonly known as Technical Analysis [Kirkpatrick et al.2006]. The word "technical" emphasizes studying the time series itself, instead of external factors expressed in market dynamics. Many authors claim that forecasting methods associated only to time series itself are valueless, since the efficient-market hypothesis (EMS) contradicts the basic tenets of technical analysis by stating that past prices cannot be used to profitably predict future prices, i.e. prices follow an essentially unpredictable random walk [Malkiel 1973]. However, many investors base their expectations on this methodology.

Technical analysis usually employs technical indicators. A technical indicator is a series of data values that are derived

from the time series data. The most used technical indicators are the moving averages that belong to the class of lagging indicators also known as trend-following indicators. There are some variants of moving averages: simple moving averages, exponential moving averages, linear weighted moving averages, moving average convergence divergence, etc.
The objective of the moving average is to smooth out the movement of a time series in order to give a cleaner view or the overall movement of the variable, capturing important patterns in the data, while leaving out noise or other fine-scale phenomena.
The most commonly used type of moving average is the simple moving average (SMA), sometimes referred to as an arithmetic moving average. SMA tries to summarize the past data by computing the mean of successive sets of size k from the time series:

$$SMA_k(t) = \frac{1}{k}\sum_{i=t-k}^{t-1} x_i \tag{1}$$

By using $SMA_k(t)$, we take into account the trend effect of the past k values. For instance, if t refers to days, $SMA_7(t)$ considers the mean of the last week and $SMA_{30}(t)$ the mean of the last month.
Although the SMA helps us discern a trend, it does so after the trend has begun. Thus, this moving average is a lagging indicator. The shorter the period covered by the SMA, the less of a lag there will be. However, using a shorter period also leads to more false signals.
Another moving average that tries to respond to this trade-off associated to the SMA indicator is the exponentially moving average (EMA). Whereas in SMA, the past observations are equally weighted, EMA assigns exponentially decreasing weights as the observation get older. In other words, recent observations are given relatively more important than older observations. The degree of weighting decrease is usually expressed as a constant smoothing factor:

$$\alpha = \frac{2}{k+1} \tag{2}$$

and EMA indicator is calculated by means of the recursive expression:

$$EMA_k(t)=\alpha\cdot x_t+(1-\alpha)\cdot EMA_k(t-1) \tag{3}$$

The initial $EMA_k(1)$ does not exist and $EMA_k(2)$ can be taken as x_1. Still another possibility would be to average the first k observations and start applying (3) from $t \geq k+1$.
EMA indicator is, in fact, a particular case of weighted moving average where the weight associated to each value decreases exponentially. Another popular weighted moving average is the linear weighted moving average (LWMA) where the weight associated to each value decreases linearly. Like EMA, LWMA was also created to overcome the lagging associated with SMA.
Moving Average Convergence-Divergence (MACD) is a technical analysis indicator introduced by Appel in the late 1970s [Appel 1999] and consists on subtracting a long term and a short term moving averages. The most popular formula for the "standard" MACD is the difference between a 26-day and 12-day EMAs. This is the formula that is used in many popular technical analysis programs, and quoted in most technical analysis books on the subject [Kirkpatrick 2006]. The resulting plot forms a line that oscillates above and below zero, i.e. MACD is a centered oscillator.
In this paper, another smooth indicator called moving interval (MI) and recently introduced in [Sánchez et al 2010] is proposed. Unlike SMA, EMA and LWMA, $MI_k(t)$ is not a number but an interval. MI not only takes into account the mean of the past *k* values, it also considers the variability of these values that it is not directly considered in other moving average indicators.

Definition (Moving Interval of order *k*)
Given a time series $\{x_1,\ldots,x_N\}$, it is defined

$$MI_k(t)=\{perc_\pi(x_{t-k},\ldots,x_{t-1}\}) , perc_{100-\pi}(x_{t-k},\ldots,x_{t-1}\}) \tag{4}$$

as a moving interval indicator associated to x_t, where π is the percentage of excluded values for each side of the interval, that is, $MI_k(t)$ is the closed interval defined from the πth and $(100-\pi)$th percentiles of the k previous values $\{x_{t-k},\ldots,x_{t-1}\}$.
Another similar definition of MI is possible by using the mean and the standard deviation, but it is well known that standard deviation is more sensitive to outliers than the difference between percentiles (for instance the Interquartile range).
In figure 1 it is shown the difference between a punctual indicator (SMA) and the interval indicator (MI). By setting k constant. SMA_k results in another scalar time series, while MI_k results in an interval time series.

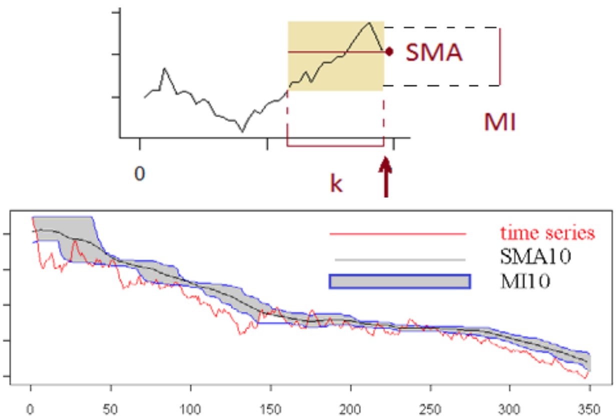

Figure 1. Comparison between SMA and MI indicator.

3 Nonlinear Regression Models: the Kernel approach

SVM are a class of learning algorithm that combines a strong theoretical basis (based on Statistical learning theory), optimization techniques, and the kernel mapping idea. SVM have been used in a variety of areas, for instance, to automatically categorize web pages, to recognize handwritten digits and faces, to identify speakers, to interpret images and also SVM has been employed to predict time series [Kyoung-jae 2003]. Their accuracy is excellent, and in many cases they outperform competing machine learning methods such as Artificial Neural Networks and radial basis function networks due to an excellent generalitzation performance in practice.

SVM belong to the family of kernel methods. The formulation of the linear version of SVM like other 'kernelizable' linear methods solely depends on the dot product between input patterns. This allows replacing this dot product with a suitable function named 'kernel'. The substitution of the dot product by the kernel can be interpreted as mapping the original input patterns into a different higher dimensional feature space where the linear method can perform the learning process (classification, regression,...).

The main advantage of kernel functions is that it is not necessary to consider explicitly this feature space and, hence, the dimension of this space has no effect on the algorithm.

The kernel formalism allows using this kind of methods with data not belonging to a Euclidean space, for instance texts, images, biosequences, trees, and also intervals.

3.1 The concept of kernel

Kernels are first introduced on the SVM approach with the aim of using this machine learning method to establish a nonlinear relationship between input and output variables. However, the advantage of not requiring any Euclidean structure in the input space, only the Hilbert space structure in the named feature space, allows to use SVM with patterns described by variables not belonging to any Euclidean space, for instance patterns described by intervals.

A kernel defined in the set A is a function k from A×A to R that satisfies that for all finite set $\{a_1,\dots,a_n\}$ with arbitrary $n\in N$, the matrix $K=[k(a_i , a_j)]$ (Gramm matrix) is symmetric and positive semidefinite. It is known that for all function k which satisfies this property, there exists a Hilbert space $\{F,\langle\rangle\}$ and a map ϕ from A to F verifying $k(a_i,a_j)=\langle\phi(a_i), \phi(a_j)\rangle$, i.e. the result of applying the function k to the pair of elements (a_i,a_j) from A is equivalent to calculate the dot product of the images by the map ϕ. In addition, if a map ϕ from A to a Hilbert space F can be found, the function $k(a_i,a_j)= \langle\phi(a_i), \phi(a_j)\rangle$ is a kernel [Cristianini et al 2000].

3.2 Intersection kernel

The set of the intervals do not directly have a Euclidean structure. However, it is easy to define kernels in the set of the intervals. One of the easier kernels that can be defined in the set of the intervals is the length of the intersection.

Theorem Let $I_1=[a,b]$, $I_2=]c,d]$ be two real intervals. The function:

$$K_\cap(I_1, I_2)=\text{length}(I_1\cap I_2) \tag{5}$$

is a kernel.

Demonstration Let ϕ a map that associates to each interval $I=[a,b]$ a function g_I from the Hilbert space L_2 of square integrable real function defined on $(-\infty, +\infty)$ of the following form:

$$\phi(I) \equiv g_I(x) = \begin{cases} 1, & \text{if} \quad a \le x \le b; \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

The composition of this function with the usual dot product in L_2 leads to the function $K_\cap$:

$$\begin{aligned} K_\cap([a,b],[c,d]) &= \int_{-\infty}^{+\infty} g_{[a,b]}(x)\cdot g_{[c,d]}(x)\cdot dx \\ &= \text{length}\ ([a,b]\cap[c,d]) \end{aligned} \tag{7}$$

3.3 Extended Intersection kernel

One of the disadvantages of intersection kernel defined above is that it is unable to discriminate between disjoint intervals more or less separated. This fact is particularly serious if the length of the intervals is very small or punctual intervals are involved.

A modification of the intersection kernel that takes into account not only the common part but also the relative distance between them is the intersection kernel with exponential influence introduced in [Ruiz 2006]. This kernel is based on an influence function that extends the range of influence beyond the range of the interval.

Definition (Exponential influence function)
A function f is said to be an exponential influence function from the interval [a,b] of parameter γ if f has the form:

$$f_{[a,b],\gamma}(x) = \begin{cases} e^{-\frac{a-x}{\gamma}}, & \text{if} \quad x < a \\ 1, & \text{if} \quad a \le x \le b \\ e^{-\frac{x-b}{\gamma}} & \text{if} \quad x > b \end{cases} \tag{8}$$

From this exponential influence function the kernel is defined as:

$$K([a,b],[c,d]) = \int_{-\infty}^{+\infty} f_{[a,b]}(x) \cdot f_{[c,d]}(x) \cdot dx \qquad (9)$$

Figure 2 shows the usefulness of the exponential influence function in case of disjoint but near intervals.

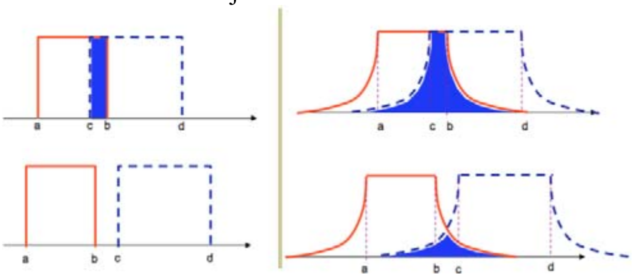

Figure 2. Comparison of intersection kernel with exponential influence function (right) and without it (left) in case of joint (top) and disjoint (bottom) intervals.

3.4 Generalization to Multidimensional Intervals

A *q*-dimensional generalized interval is the Cartesian product of *q* real intervals. Its representation is a hyperrectangle or *q*-orthotope. The most direct way of defining a kernel on the set of multidimensional intervals is using the product of the kernel of each of its components, that is:

If $I_i=(I_{i1}, I_{i2},\ldots,I_{iq})$ and $I_j=(I_{j1}, I_{j2},\ldots,I_{jq})$ are tho q-dimensional intervals:

$$K(I_i,I_j) = K(I_{i1},I_{j1})\cdot K(I_{i2},I_{j2})\cdot\ldots\cdot K(I_{iq},I_{jq}) \qquad (10)$$

where *K* is any kernel defined on the set of real one-dimensional intervals. The function obtained is a kernel due to the fact that the product of kernels is a kernel [Cristianini et al. 2000]. If the one-dimensional kernel used is the intersection kernel, the multidimensional kernel obtained from (10) can be interpreted as the hypervolume of the intersection of the multidimensional hyperectangles.

4 An Application to Forecast Financial Exchange Rate

The value of the exchange rate between two countries means how much the currency of one country is worth in terms of the other country's currency. The exchange rate mechanism is seen as a useful tool in the establishment and creation of a common market between countries. The high volatility and nonlinear behavior of exchange rate encourages the use of a nonlinear regression technique to forecast them.

4.1 Data Set Description

In this research, foreign currency exchange rate data is used for analysis. The dataset comes from the Pacific Exchange Rate Service available online [Pacific 2009]. The period under consideration is from December 15, 1999 until November 14, 2003 (400 observations for each currency). The data set of six world's major currencies is used, specifically, US Dollar (USD), Great Britain Pound Sterling (GBP), Chinese Yuan (CNY), Canadian Dollar (CAD), Japanese Yen (JPY) and Indian Rupee (INR). European Euros (EUR) is selected as the base currency.

4.2 Experimental Setup

In order to compare the use of different technical indicators as a preprocessing step, only two variables were used to forecast the value of x_{i+h}, the last known value, x_i, and one moving average computed from the last 10 values: $\{x_{i-9},\ldots,x_i\}$. The parameter h, the prediction horizon, has been taken from 1 to 5. The moving averages used were SMA, EMA, MACD 10/5 and MI (with a value of 25 for π parameter in equation 4, i.e. by using the interquartile range).

Experiments were run using the statistical software package R [R Project 2011] with packages *kernlab* [Karatzoglou et al. 2004] and *TTR* [Ulrich 2010]. In all cases, ε-SVM for regression [Smola et al. 2004] was used. A Gaussian kernel was adopted when standard point indicators were involved and an intersection kernel with exponential influence when MI is used. In all cases, the parameter of the kernel, either Gaussian or exponential influence width, and the regularization constant C of the SVM algorithm were selected with a grid search tuning. The best parameters found in all training sets were C=10000 and σ=0.001 for SMA, EMA and MACD and C=100 and γ=100 for MI.

For all currencies, a training set of 100 values is used and the rest were used to evaluate the error of the forecast.

The performance metrics used to evaluate the results was the mean absolute percentage error (MAPE). It is defined as:

$$\mathrm{MAPE}(\hat{Y}) = \frac{\sum_{i=1}^{n} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right|}{n} \qquad (11)$$

where Y_i and $\hat{Y}_i$ represent the actual and predicted values respectively, and n is the size of the test set. Unlike other performance metric, MAPE allows to compare values from different scales.

In table 1 and figure 3, the MAPE of the 6 currency exchange rates are shown for forecasting horizons 1 to 5. In all cases, MAPE using MI indicator is lower than MAPE using other indicators.

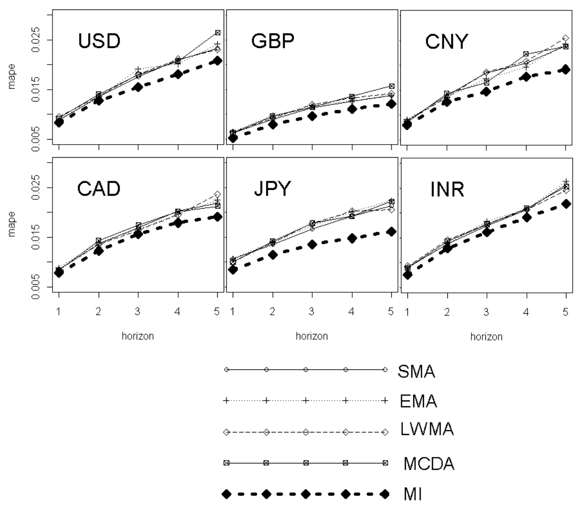

Figure 3. Graphical representation of results from table 1.

Table 1. Results of MAPE as a function of horizon h in all six currency exchange rate data used in this analysis..

	h	1	2	3	4	5
USD	SMA	0.0088	0.0135	0.0175	0.0209	0.0233
	EMA	0.0087	0.0135	0.0191	0.0201	0.0241
	LWMA	0.0095	0.0136	0.0180	0.0211	0.0230
	MCDA	0.0092	0.0141	0.0180	0.0207	0.0264
	MI	**0.0083**	**0.0127**	**0.0155**	**0.0180**	**0.0208**
GBP	SMA	0.0062	0.0090	0.0114	0.0126	0.0138
	EMA	0.0066	0.0097	0.0117	0.0127	0.0138
	LWMA	0.0064	0.0093	0.0119	0.0132	0.0141
	MCDA	0.0063	0.0097	0.0114	0.0136	0.0157
	MI	**0.0053**	**0.0079**	**0.0097**	**0.0111**	**0.0121**
CNY	SMA	0.0088	0.0139	0.0185	0.0204	0.0240
	EMA	0.0091	0.0135	0.0173	0.0196	0.0243
	LWMA	0.0087	0.0136	0.0186	0.0208	0.0254
	MCDA	0.0084	0.0144	0.0165	0.0223	0.0237
	MI	**0.0079**	**0.0126**	**0.0147**	**0.0177**	**0.0192**
CAD	SMA	0.0085	0.0137	0.0170	0.0203	0.0219
	EMA	0.0089	0.0139	0.0167	0.0196	0.0224
	LWMA	0.0085	0.0136	0.0164	0.0194	0.0236
	MCDA	0.0084	0.0143	0.0174	0.0201	0.0212
	MI	**0.0078**	**0.0123**	**0.0156**	**0.0179**	**0.0191**
JPY	SMA	0.0106	0.0135	0.0167	0.0193	0.0213
	EMA	0.0106	0.0141	0.0173	0.0201	0.0223
	LWMA	0.0101	0.0139	0.0178	0.0203	0.0205
	MCDA	0.0100	0.0142	0.0178	0.0192	0.0222
	MI	**0.0085**	**0.0115**	**0.0136**	**0.0148**	**0.0161**
INR	SMA	0.0090	0.0138	0.0175	0.0207	0.0257
	EMA	0.0091	0.0139	0.0183	0.0208	0.0263
	LWMA	0.0093	0.0146	0.0179	0.0207	0.0246
	MCDA	0.0089	0.0144	0.0177	0.0210	0.0253
	MI	**0.0076**	**0.0129**	**0.0161**	**0.0191**	**0.0219**

Conclusion and future works

In this paper, the performance of using interval technical indicators, *moving interval* (MI) as input variables in time series forecasting is evaluated. Unlike standard technical indicators, MI also takes into account the variability of data in the range considered. In order to carry out the study of the MI performance, a new interval kernel, based on the intersection, has been defined. The intersection kernel is extended by an exponential influence function that permits to discriminate between disjoint intervals and also to handle simultaneously with intervals and point values. Results shown in Table 1 and figure 3 prove, first, the advantages of considering this MI indicator in forecasting and, second, the suitability of the intersection kernel employed that can compete with other well established kernels like Gaussian kernels.

Future work will be twofold. On the one hand, we will continue the research on other interval kernels and we will use it with other machine learning methods based on kernel methods and, on the other hand, other most sophisticated interval indicators will be considered.

Acknowledgments

This work is supported by the Spanish project SENSORIAL (TIN2010-20966-C02-02) Spanish Ministry of Education and Science.

References

[Ahmed et al. 2010] N.K Ahmed, A.F. Atiya, N. Gayar, H. El-Shishiny An empirical comparison of machine learning models for time series forecasting. Accepted for publication, to appear in Econometric Reviews, vol. 29, No. 5-6, 2010.

[Appel 1999] Gerard Appel. *Technical Analysis Power Tools for Active Investors*. Financial Times Prentice Hall, 1999.

[Cristianini et al.2000] N. Cristianini and J. Shawe-Taylor. *Support Vector Machines*. Cambridge, United Kingdom: Cambridge University Press, 2000

[Karatzoglou, et al. 2004] Karatzoglou, A., Smola, A., Hornik, K. and Zeileis, A., kernlab - An S4 Package for Kernel Methods in R. *Journal of Statistical Software*, vol 11(9), pp.1-20, 2004.

[Kirkpatrick et al.2006] Kirkpatrick and Dahlquist. *Technical Analysis: The Complete Resource for Financial Market Techmicians.* Financial Times Press, 2006.

[Kyoung-jae 2003]. Kim Kyoung-jae. Financial time series forecasting using support vector machines. *Neurocomputing,* 55 pp307-319. 2003

[Malkiel 1973] Burton Malkiel. *A Random Walk Down Wall Street*. WW Norton & Company. Inc. 1973.

[Pacific 2009] Pacific Exchange Rate Service <URL:http://fx.sauder.ubc.ca/data.html> (Accessed 2009 May).

[R Project 2011] The R Project for Statistical Computing <URL:http://www.r-project.org> (Accessed 2011).

[Sánchez et al 2010] Germán Sánchez, Albert Samà, Francisco J Ruiz and Núria Agell. Moving Intervals for Nonlinear Time Series Forecasting. *Proceeding of the 2010 conference on Artificial Intelligence Research and Development: Proceedings of the 13th International Conference of the Catalan Association for Artificial Intelligence* IOS Press Amsterdam, The Netherlands, The Netherlands. 2010.

[Smola et al. 2004] Smola A. J., Schölkopf B., A tutorial on support vector regression. Statistics and Computing, vol 14(3), pp.199-222, 2004.

[Ruiz 2006] Funciones núcleo sobre Estructuras Cualitativas. Angulo, C., Agell, N. (directors). Ph.D. dissertation, *Universitat Politècnica de Catalunya, Departament d'Enginyeria de Sistemes, Automàtica i Informàtica Industrial*, 2006.

[Ulrich 2010] Ulrich, J., TTR - Technical Trading Rules (R Project)

<URL:http://cran.r-project.org/web/packages/TTR> (Accessed 2010 May).