RESEARCH ARTICLE

AESTIMATIO, THE IEB INTERNATIONAL JOURNAL OF FINANCE, 2014. **9**: 68-89 DOI:10.5605/IEB.9.4

The sheaf of straight lines methodology. An application to the Spanish stock market

Ferrán Aranaz, Magdalena
Márquez de la Cruz, Elena

RECEIVED: 2 OCTOBER 2013
ACCEPTED: 9 FEBRUARY 2014

Abstract

Financial markets, and particularly stock markets, have become an essential factor in understanding the behavior of the real economies. The analysis of how the stock markets have responded to the recent financial and economic crisis is crucial to understanding the economy's behavior. However, its analysis is not an easy task due, among many other reasons, to the high number of implicated variables. Usually, different graphical techniques are applied to analyze the evolution of the stock markets and to forecast the expected trend of the quoted stock prices. This paper proposes a new graphical methodology to analyze stock market behavior: the Sheaf of Straight Lines Methodology (SoSLM). The main advantage of this methodology is that it helps to compare time paths of a set of time series for a given time period. This makes it possible to better understand the evolution of one of them with respect to the others, in relation to the formed sheaf of straight lines. Simultaneously, it eliminates the difficulties arising when a high number of time series are compared through the joined graphical representation, either in a unique chart or in separated ones. So, it is particularly suitable for application to financial markets. This paper applies the SoSLM to the 35 constituents of the Spanish IBEX 35 index from June to September of 2013. The Spanish stock market has been hit hard by the recent financial crisis, although it has shown signs of recovery since June 2013; so, it presents an interesting case for analysis.

Keywords:
Sheaf of straight lines methodology; Time series analysis; Stock Market; IBEX 35
JEL classification:
C40, G10, Y10.

Ferrán Aranaz, M. ✉ Dpto. de Estadística e I.O. III Facultad de Estudios Estadísticos. Universidad Complutense de Madrid. Avda. Puerta de Hierro s/n. 28040 Madrid. ✆+34 91 394 4022. Email: mferrana@ucm.es

Márquez de la Cruz, E. Dpto. de Economía Aplicada III. Facultad de CC. Económicas y Empresariales. Universidad Complutense de Madrid. Campus de Somosaguas. Pozuelo de Alarcón 28223 (Madrid), Spain. ✆ +34 91 394 2435. Fax: +34 91 394 25 82. Email: emarquez@ccee.ucm.es

AESTIMATIO, THE IEB INTERNATIONAL JOURNAL OF FINANCE, 2014. **9**: 68-89

La metodología del haz de rectas. **Una aplicación al mercado bursátil español**

Ferrán Aranaz, Magdalena
Márquez de la Cruz, Elena

Resumen

Los mercados financieros, y en particular los mercados bursátiles, se han convertido en un factor esencial para comprender el comportamiento de las economías reales. El análisis de cómo han respondido estos mercados a la reciente crisis económica y financiera es crucial para comprender el comportamiento de la economía. Sin embargo, su estudio no es una tarea sencilla, entre otras muchas razones, por el elevado número de variables implicado en su análisis. Es habitual el empleo de diferentes métodos gráficos para analizar la evolución de los mercados y para predecir la tendencia esperada de los precios de los activos cotizados. Este trabajo propone una nueva metodología gráfica para el análisis del comportamiento del mercado bursátil: la Metodología del Haz de Rectas. La ventaja esencial de esta metodología es que permite comparar las sendas de un conjunto de series temporales para un período de tiempo dado de un modo más sencillo que otros métodos alternativos, haciendo posible una mejor comprensión de la evolución de una de las variables del haz de rectas con respecto a las demás. Al mismo tiempo, elimina las dificultades que surgen cuando se compara un elevado número de series temporales a través de su representación gráfica, ya sea en un único gráfico o en gráficos separados. Por lo tanto, su aplicación al mercado bursátil puede resultar especialmente útil. En este trabajo se aplica la metodología SoSLM a los 35 componentes del índice bursátil español IBEX 35 para el período comprendido entre junio y septiembre de 2013. El mercado bursátil español ha sido duramente golpeado por la reciente crisis financiera, si bien parece mostrar signos de recuperación a partir de junio de 2013. Es, por tanto, un caso de análisis de gran interés.

Palabras clave:

Metodología del haz de rectas, análisis de series temporales, mercado bursátil, IBEX35.

1. Introduction

Line Charts are the standard way of representing historical data; a simple look at the chart can help to detect patterns in the temporal path of the analyzed variable. However, when the objective is to compare the time paths of a set of variables, different problems arise, not only when they are represented in a unique chart (scale differences among others) but also if multiple charts are drawn; even if the number of variables is not significantly large.

Although nowadays computer technology offers excellent tools to facilitate the representation of space-time data (see Andrienko *et al.*, 2003, for a detailed analysis), there are few studies related to the representation of massive time series (see Lin *et al.,* 2005). The graphical methodology of the Sheaf of Straight Lines (SoSLM from now on) proposed in this paper can be understood as a way to simplify the comparison of the time series by providing a set of 'summary' time series. These make it easier to perform comparisons among them. The main advantage of this methodology is that it helps to compare time paths of a set of time series for a given time period, making it possible to better understand the evolution of the variables included in such a set. At the same time, the proposed methodology eliminates the above mentioned difficulties when comparing a high number of time series through the joined graphical representation, either in a unique chart or in separate ones.

The mentioned difficulties are also present in the case of the analysis of stock market data. Since the SoSLM methodology facilitates the comparison among a high number of variables, it becomes especially useful in the case of the study of stock markets, where a high number of stocks prices are usually implicated when the market performance needs to be analyzed.

So, centered on the current importance of stock markets and with the objective of making understanding the evolution of stock prices easier, this paper will illustrate the SoSLM. We will apply it to the observed evolution of the IBEX 35 Spanish stock market index during the summer of 2013; specifically, we will apply the SoSLM to the 35 firms including in the IBEX 35 index from June to September of 2013. The Spanish economy has been one of the more deeply affected European economies during the recent financial and economic crisis. The Spanish bank sector, essential in the IBEX35 index as we will see later, has been especially affected. However, the Spanish stock market seems to have reversed the trend from June of 2013 onward. So, our sample period starts at June of 2013 and ends at September, when this paper was written.

As is well known, the IBEX 35 is the reference index for the Spanish stock market. It is composed of the 35 securities listed on the Stock Exchange Interconnection System of

the four Spanish Stock Exchanges, which were most liquid during the control period. The weight of each stock inside the index is in relation to its market capitalization.[1] Table 1 shows the constituents of the index for our sample period.

Table 1. IBEX 35 constituents at 1st of July of 2013

Code	Stock
ABE	ABERTIS INFRAESTRUCTURAS, SERIE A
ANA	ACCIONA
ACX	ACERINOX
ACS	ACS ACTIVIDADES CONST.Y SERVICIOS
AMS	AMADEUS IT HOLDING
MTS	ARCELORMITTAL
POP	BANCO POPULAR ESPAÑOL
SAB	BANCO DE SABADELL
SAN	BANCO SANTANDER
BKT	BANKINTER
BBVA	BANCO BILBAO VIZCAYA ARGENTARIA
BME	BOLSAS Y MERCADOS ESPAÑOLES
CABK	CAIXABANK
DIA	DIA-DISTRIBUIDORA INT. DE ALIMENTACION
EBRO	EBRO FOODS
ENG	ENAGAS
ELE	ENDESA
FCC	FOMENTO CONSTR.Y CONTRATAS(FCC)
FER	FERROVIAL
GAS	GAS NATURAL SDG
GRF	GRIFOLS
IAG	INTERNATIONAL CONSOLIDAT. AIRLINES GROUP
IBE	IBERDROLA
ITX	INDUSTRIAS DE DISEÑO TEXTIL (INDITEX)
IDR	INDRA, SERIE A
JAZ	JAZZTEL
MAP	MAPFRE, S.A.
TL5	MEDIASET ESPAÑA COMUNICACION
OHL	OBRASCON HUARTE LAIN
REE	RED ELECTRICA CORPORACION
REP	REPSOL
SYV	SACYR VALLEHERMOSO
TRE	TECNICAS REUNIDAS
TEF	TELEFONICA
VIS	VISCOFAN

On 2nd January, Bankia and Gamesa were excluded and Viscofan was included; on 23rd April, Jazztel was included in the index and 1st of July, Ebro came in and Abengoa came out.

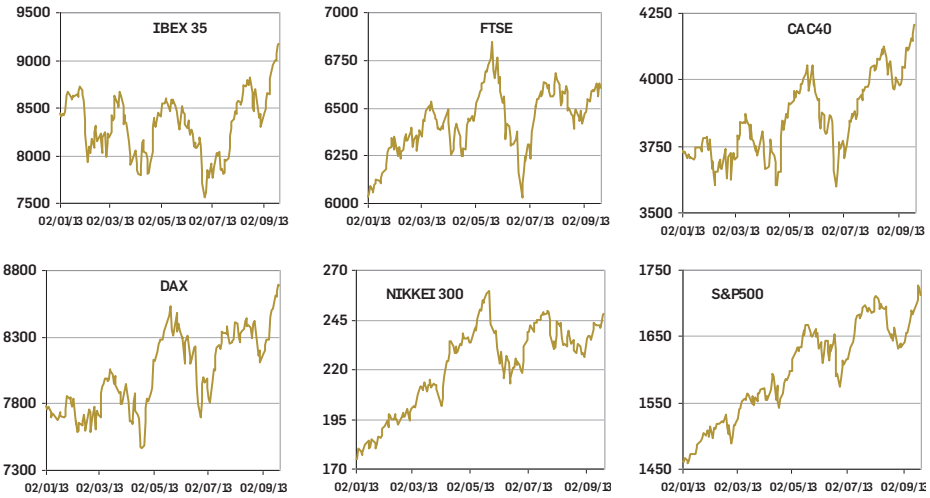
The analysis of the weights of the different stocks on the index is quite interesting; the most important stock is Santander, weighing above 17%; the second is Telefonica with a weight of 13.57%. Adding the weights of these two firms to the third and fourth highest weighted stocks, BBVA and Inditex, respectively, we observe that these four values represent in excess of 53% of the IBEX 35 (see BME, 2013).

As previously mentioned, the Spanish stock market, as many others, has been hit by the recent financial and economic crisis. In fact, on the 1st of October of 2008, the IBEX 35 value was above 11100 points, decreasing dramatically below 6000 on the 24th of July of 2012. This coincided with worst moments of the Spanish sovereign debt crisis, with the risk premium of the Spanish government ten year

[1] *For a detailed description see BME Technical Regulations for the Composition and Calculation of the Sociedad de Bolsas, S.A. Indexes, available at http://www.bolsamadrid.es/docs/Indices/Ibex/normdef-in.pdf*

bonds reaching historical records. However, 2013 seems to be the year in which the Spanish stock market started to recover the path of growth, mainly from mid-year on. Figure 1 shows the daily evolution of the IBEX 35 from the beginning of 2013 until the end of September. It can be observed that the index trend was unclear until the end of June; specifically, the index value reached the 2013 minimum value at 7553.20 on the 21st of June, before beginning to rise through until the end of the sample period. Figure 1 also shows the evolution of some of the most representative international financial market s indexes: CAC40, FTSE, DAX, NIKKEI300 and S&P500. Although, as mentioned above, it is difficult to make comparisons using a line chart because of scale differences, an increasing trend can be observed in all of them through 2013; however, the Spanish stock market seems to have had a more challenging path in attaining the positive trend shown by the other stock markets considered, perhaps due to the above mentioned government risk premium crisis.

Figure 1. IBEX 35, FTSE, CAC40, DAX, NIKKEI300 and S&P500, January-September 2013



Differently from the usually financial newspapers graphs showing the paths of the different stocks included in a market index, and focusing on the Spanish stock market, the SoSLM we propose will allow us to study the representative evolution identified through the paths followed by the 35 stocks included in the index. Within this methodology, the position of each of the stocks in relation to the sheaf of straight lines makes it possible to better understand the situation of one of them with respect to the others, and with respect to the whole market. This is an essential advantage of the SoSML methodology when compared with other graphical representations of the evolution of financial markets.

To accomplish the described objective, the paper is organized as follow: Section 2 analyzes the main advantages of the sheaf of straight lines and describes other graphical methodologies usually applied to the understanding of stock market behavior, remarking on the differences among them. Section 3 describes the essentials of SoSLM and Section 4 applies the methodology to the analysis of the Spanish stock market. Section 5 comments on our main findings.

2. Stock markets graphical analysis

To explain the value of a firm on the stock market and to forecast its possible future value, it is usual to distinguish two different approaches: the technical and the fundamental analysis[2]. Whereas the fundamental analysis focuses on the so called intrinsic value of the stock, the technical approach takes advantage of statistics calculated from the stock markets in order to identify trends and models that can help to anticipate the stock s future evolution[3].

To carry out a fundamental analysis, a study from the general economy and sector situation to the financial and management firm conditions is required. However, the technical analysis is fundamentally based on the graphical analysis of the firm's past prices, together with the study of different statistics on prices and trading volumes calculated from the market data date; with the objective of forecasting the most likely evolution of the stock in the future[4].

Given that the SoSLM is essentially based on the graphical analysis, it is clear that it could be related to the technical approach. However, as we shall see later, both the main purpose and the applied procedures are completely different.

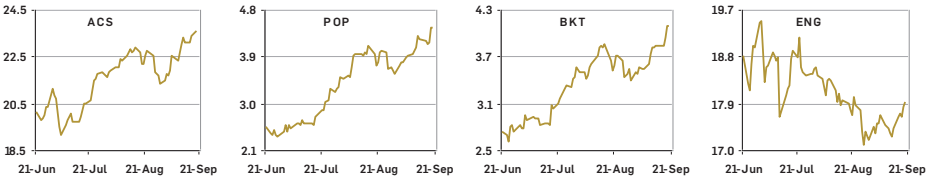
Let us see an example of how the SoSLM can help to better understand the evolution of the stock markets, its main purpose. Consider the case of the following stocks: Actividades de Construcción y Servicios (ACS), Banco Popular Español (POP), Bankinter (BKT) and Enagás (ENG). Figure 2 shows the evolution of the prices of the four companies during the sample period.

[2] *The Dow Theory is usually mentioned as the origin of the technical analysis. For a detailed analysis of this approach, see, among many others, Murphy (1999).*

[3] *Although it is usual to argue that investors do not mix fundamental and technical analysis to make their investment decisions, Cohen et al. (2011) have shown that this is not the case. In fact, they find that investors use some technical along with fundamental tools when deciding their investment strategies.*

[4] *The traditional technical analysis is known as Chartism and it is essentially based on the graphical analysis; more recently, however, the technical analysis has added other statistical techniques. For some recent examples of graphical techniques applied to the stock markets, see Ait Hellal and Meyer (2014) or Lobão and Da Mota Lopes (2014).*

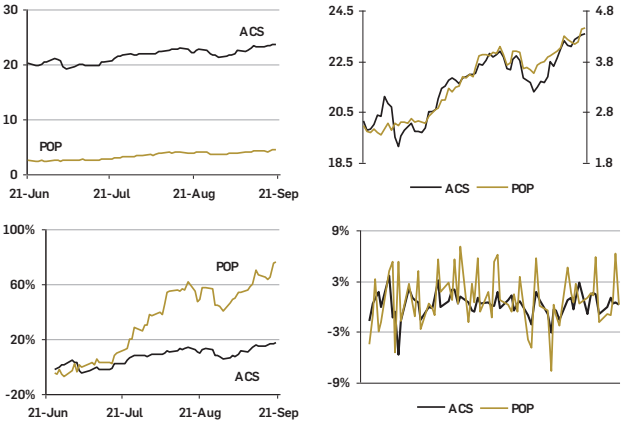
Figure 2. **Acerinox, Banco Popular Español, Bankinter and Enagas, summer 2013**



It seems to be that ACS, POP and BKT have followed a similar increasing path, the evolution of POP and BKT being quite similar. However, ENG seems to follow just the opposite path, clearly decreasing.

Since the data scales differ, a unique figure for the four stocks could help to facilitate the comparative analysis of the price s evolution. What is more, to avoid the difficulties related to the scale, in the case of stock markets data date, it is usual to considerer price rates of growth instead of price levels[5]. Figure 3 shows these ideas for the cases of ACS and POP. The Figure on the top left shows the evolution of both companies although now they are represented in the same chart. This makes the differences in the prices levels clear. The Figure of the top right again shows prices levels but now considering different scales in the vertical axes. Now, it seems that both companies have moved together during the sample period. The Figure on the bottom left shows the prices growth rates taking the first date in the sample as the reference price. So, it can be seen that POP has experienced a higher increase than ACS, close to 80 per cent. A result of the higher daily price increases that are shown by the bottom right Figure: it can be seen that, exceptions apart, the daily rates of return of ACS and POP show the same positive sign; since the values of POP are more extreme, it has experienced a higher increase.

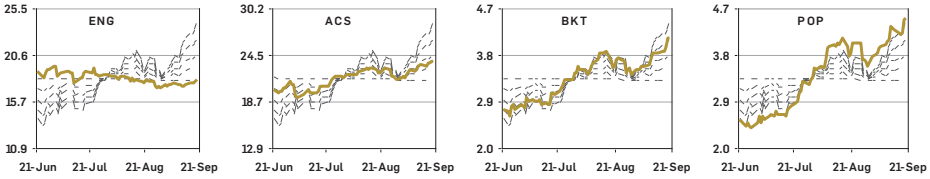
Figure 3. **ACS and POP, summer 2013**



[5] *That is to say, the stock rate of return under the assumption of no dividend payments.*

To end the illustration, Figure 4 shows the SoSLM results for the selected sample (it will be explained later; the only objective now is to illustrate the usefulness of the proposed methodology). The bold line is the evolution of the considered company in each case, whereas the thinner lines show the sheaf of straight lines. Note that it is easy to see that the evolution of each of the considered companies is consistent with the evolution of the four as a whole; additionally, the four considered firms can be ordered as a function of its similarity to the others; so, ENG will be the first one followed by ACS, BKT and, finally, POP. So, the proposed methodology can be useful to avoid the difficulties related to differences in levels or scales when comparing a high number of time series data.

Figure 4. **Enagas, Acerinox, Bankinter and Banco Popular Español, summer 2013, over the SoSLM solution**



To summarize, the SoSLM is a graphical method for studying the relationships among time series of different variables; the methodology is particularly useful when the number of variables to be compared is high. The main purpose of the methodology, different to the technical analysis, is not to forecast the evolution of certain stock prices, but to better understand the behavior of one of the variables when compared with others because they are joined in the configured sheaf of straight lines. To be able to consider information from a properly carried out graphical analysis could be very useful for researchers working out theoretical models to explain economic facts.

3. The sheaf methodology

The SoSLM (Ferrán, 2011a, 2011b and 2013) is a statistical tool that calculates, under the assumption that the underlying structure of the analyzed time series set is a sheaf of straight lines, a set of 'summary' time series that makes it possible to simplify the comparison of their temporal paths. Thus, the basic result of the methodology is a line chart for each time series in which, in order to better understand its evolution in terms of the evolution shown by others, the corresponding temporal path will be represented over a 'summary' time series set.

Before we discuss the methodology and its associated theoretical results, we will introduce the next concept:

Definition 1: A set of J different time series $\{y_{t,j}\}$, $t=1,\ldots,T$, has a sheaf structure of J straight lines if each $y_{t,j}$ is related to another time series, say x_t, by the linear expression:

$$y_{t,j}=b_j+m_jx_t\ ,\quad \forall t,j,\qquad \text{where}\qquad b_j=B_0+B_1m_j\qquad \forall j\ .$$

Next, we show the key results in the SoSLM:

Let $\{y_{t,j}\}$, $j=1,\ \ldots,J$; $t=1,\ldots,T$ be a set of J different time series verifying the following three assumptions:

- All of them show the same mean,

$$\overline{y}_t=\frac{1}{T}\sum_{t=1}^{T}y_{t,j}=\alpha\qquad \forall j\ . \tag{1}$$

- Each of the J time series, $y_{t,j}$, is related to another time series, say x_t, by the linear expression:

$$y_{t,j}=b_jt+m_jx_t+\mu_j\qquad \forall t,j\ . \tag{2}$$

- There exists two coefficients, B_0 and B_1, such that

$$b_j=B_0+B_1m_j\qquad \forall j\ . \tag{3}$$

Then, it can be shown that (see Appendix for proofs):

Remark 1: The Euclidean distance between two any time series $y_{t,j}$ and $y_{t,j'}$ depends on the distance between their respective coefficients m_j and $m_{j'}$; so, it can be used to set a ranking of the J time series. In particular, if $y_{t,q}$, $y_{t,r}$ and $y_{t,p}$ are included in the $\{y_{t,j}\}$ set and they are such that $m_q<m_r<m_p$, then $d(y_{t,q},y_{t,r})<d(y_{t,q},y_{t,p})$, where $d(\cdot)$ is the Euclidean distance. So, it follows that, if a Principal Component Analysis on the matrix of distances between them is carried out, the graphical solution in the space of the first two components, will show the so called Guttman effect[6].

Remark 2: For any pair of time series of the $\{y_{t,j}\}$ set, there exists at least one intersection point among their trajectories. What is more, the intersection points of any two trajectories are the intersection points of all of them.

Remark 3: If more than one intersection point exits, then the difference between any two intersection points is independent from the time series scale.

[6] *The Guttman effect takes place when the graphical representation of the rows or columns of a matrix in the space of the first two dimensions coming from the solution of the Principal Components Analysis, shapes a parabolic arc (Guttman, 1953).*

Remark 4: The time series set $\{\Delta y_{t,j}\}$ has a sheaf of straight lines structure with respect to Δx_t with vertex on $(\Delta x_t, \Delta y_{t,j})=(-B_1, B_0)$ (see Definition 1).

Remark 5: The intersection points of any pair of time series of the $\{\Delta y_{t,j}\}$ set, up to the B_0 value, are the intersection points for all of them.

Remark 6: If J takes a large value, then the structure of the $\{y_{t,j}\}$ set can be summarized in a new set $\{C_{t,k}\}$ of K time series (being K a small value) by choosing K different and representative values over the m_j range.

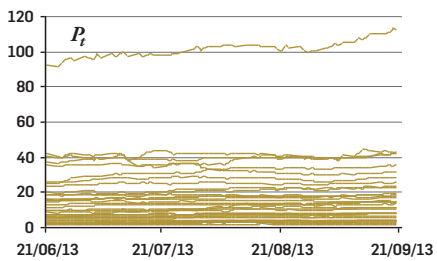
4. Applying the SoSLM to the Spanish stock market

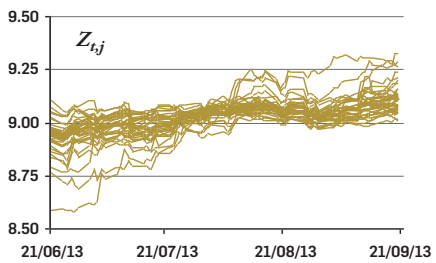
4.1. Extracting the 'summary' time series set

In this section, we use the SoSLM to build a small set of time series that will allows us to describe and to compare daily prices of the thirty-five IBEX35 constituents during the summer of 2013 (we consider the 66 trading days from June 21 to September 20):

$$\{P_{t,j},\ t=1,\ldots,66\},\quad j=1,\ldots,35$$

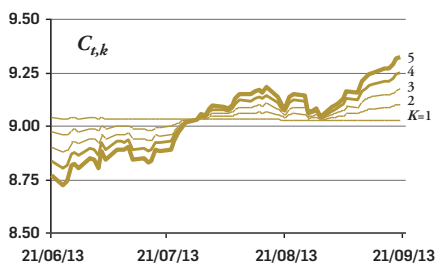
Figure 5. **(From left to right and from top to bottom): Time series set $\{P_{t,j}\}$; Time series set $\{Z_{t,j}\}$; B_0 and B_1 coefficients and m_k, b_k and μ_k values; 'Summary' time series set $\{C_{t,k}\}$**





	B
B_0	-.0004
B_1	.0008

K	m_k	b_k	μ_k
1	0.0566	-0.0004	8.53
2	0.6746	0.0002	2.93
3	1.2926	0.0007	-2.66
4	1.9105	0.0012	-8.26
5	2.5285	0.0017	-13.86



The data has been uploaded from the BME (Bolsas y Mercados Españoles) website. We apply the proposed methodology to the log of the stocks prices, $\{Y_{t,j} = \ln P_{t,j}\}$. We will consider the IBEX35 index as the reference time series. To be specific, if $P_{t,IBEX\,35}$, $t=1,\ldots,66$, is the index value for each of the 66 trading days, the reference time series will be $x_t = \ln P_{t,IBEX\,35}$, $t=1,\ldots,66$.

In the extraction process of the 'summary' time series set we will proceed as follows (see Figure 5 and Table 2):

***Step* 1**: We homogenize the scale of the $\{Y_{t,j}\}$ set. To do so, we consider:

$$Z_{t,j} = Y_{t,j} - \alpha_j + \alpha \quad j = 1,\ldots,35 ,$$

being $\alpha_j = \overline{Y}_j$ and α the chosen common scale (Table 2). If, for example, we take α equal to the average of the referenced time series, $\alpha = \frac{1}{66}\sum_{t=1}^{66} x_t$, then the average of the J time series in $\{Z_{t,j}\}$ set is equal to that of x_t ($\overline{Z}_j = \alpha$, $\forall j$).

***Step* 2**: We calculate the thirty five coefficient vectors $(A_{0,j}, A_{1,j}, A_{2,j})$, $j = 1,\ldots,35$, by estimating the regression equation:

$$\hat{Z}_{t,j} = A_{0,j}t + A_{1,j}x_t + A_{2,j}, \; j = 1,\ldots,35 .$$

Table 2. **R_j^2, $A_{0,j}$, $A_{1,j}$ and $A_{2,j}$ coefficients, and α , α_j , $F_{1,j}$ and $F_{2,j}$ values**

	R_j^2	$A_{0,j}$	$A_{1,j}$	$A_{2,j}$	α=9.03; α_j	$F_{1,j}$	$F_{2,j}$
ABE	0.497	-0.0010	0.694	2.795	2.62	.951	-.277
ANA	0.162	0.0008	0.242	6.823	3.67	.895	-.089
ACX	0.935	0.0010	0.728	2.427	2.04	.940	.299
ACS	0.897	0.0001	1.027	-0.251	3.07	.942	.290
AMS	0.217	-0.0006	0.396	5.476	3.22	.932	-.331
MTS	0.828	0.0019	0.557	3.934	2.27	.867	.427
POP	0.961	0.0039	2.529	-13.936	1.20	-.705	.665
SAB	0.793	0.0020	1.574	-5.251	0.51	-.048	.891
SAN	0.984	-0.0004	1.289	-2.599	1.68	.936	.318
BKT	0.970	0.0010	2.043	-9.453	1.20	-.034	.969
BBVA	0.965	0.0013	1.090	-0.857	1.96	.685	.703
BME	0.920	0.0034	0.122	7.818	3.02	.785	.551
CABK	0.936	0.0000	1.638	-5.763	1.03	.651	.729
DIA	0.914	-0.0005	1.009	-0.067	1.80	.989	.045
EBRO	0.513	-0.0002	0.505	4.477	2.80	.966	-.217
ENG	0.752	-0.0024	0.447	5.075	2.90	.859	-.471
ELE	0.937	0.0006	0.766	2.091	2.86	.960	.223
FCC	0.937	0.0098	1.109	-1.311	2.40	-.830	.484

FER	0.679	-0.0009	0.819	1.663	2.55	.968	-.204
GAS	0.439	-0.0018	0.783	2.020	2.71	.933	-.314
GRF	0.488	-0.0015	1.183	-1.600	3.42	.950	-.125
IAG	0.914	0.0001	1.433	-3.915	1.24	.783	.591
IBE	0.827	-0.0010	0.911	0.839	1.41	.984	-.100
ITX	0.896	0.0006	0.641	3.219	4.62	.988	.061
IDR	0.697	0.0011	0.618	3.411	2.39	.936	.248
JAZ	0.969	0.0027	0.883	0.961	1.90	.429	.874
MAP	0.734	-0.0021	1.262	-2.295	1.00	.967	-.191
TL5	0.515	-0.0002	0.993	0.070	2.02	.927	.052
OHL	0.481	-0.0023	1.079	-0.638	3.32	.930	-.315
REE	0.384	-0.0019	0.654	3.191	3.72	.899	-.405
REP	0.897	0.0003	0.770	2.065	2.86	.984	.099
SYV	0.946	0.0011	2.001	-9.072	1.00	.047	.946
TRE	0.304	-0.0010	0.057	8.555	3.55	.868	-.451
TEF	0.983	-0.0005	1.081	-0.717	2.35	.981	.132
VIS	0.693	0.0004	0.343	5.921	3.68	.976	-.181

***Step* 3**: We calculate the coefficients B_0 and B_1 by estimating the regression equation:

$$\hat{A}_{0,j} = B_0 + B_1 A_{1,j}$$

at the thirty five points $(A_{1,j}, A_{0,j})$, $j=1,\ldots,35$.

***Step* 4**: Within the range of variation of the $A_{1,j}$ values, $j=1,\ldots,J$, we choose m_k representative values, $k=1,\ldots,K$. For example:

$$m_k = m_{k-1} + \theta ,$$

with $\theta = (\max_j A_{1,j} - \min_j A_{1,j})/(K-1)$ and $m_1 = \min_j A_{1,j}$.

***Step* 5**: We calculate the coefficient

$$b_k = B_0 + B_1 m_k,\ k=1,\ldots,K \ .$$

***Step* 6**: We calculate the time series

$$C_{t,k} = b_k t + m_k Y_t + \mu_k,\ k=1,\ldots,K , \tag{4}$$

where $\mu_k = \alpha - \overline{g}_k$, being $g_{t,k} = b_k t + m_k Y_t$.

Therefore, the 'summary' time series set, $\{C_{t,k}\}$, verifies conditions (1), (2) and (3) with respect to x_t; what is more, the time series in $\{C_{t,k}\}$ and $\{Z_{t,j}\}$ present the same scale.

4.2. Interpreting the solution

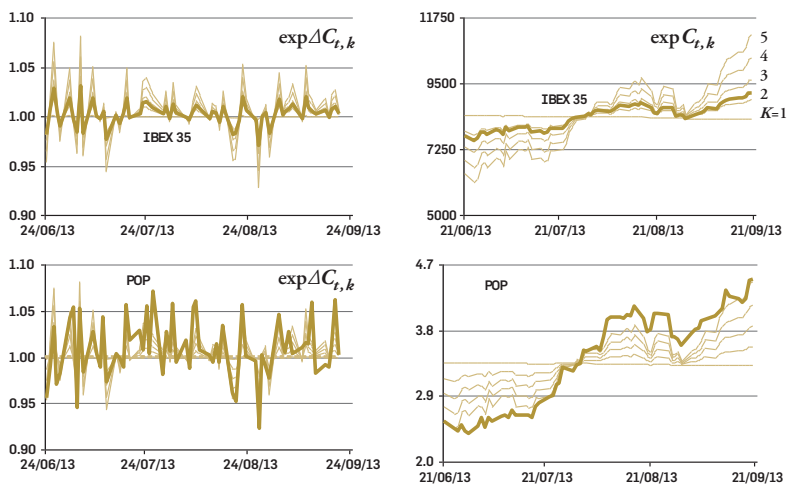
According to the expression of the time series in $\{C_{t,k}\}$ set (see (4) and Figure 4), their fluctuations become more intensive when we move from the first to the fifth 'summary' straight lines. This ordering responds to the relationship with respect to the time series x_t, namely, the time series set $\{\Delta C_{t,k}\}$ has a sheaf of straight lines structure with regard to Δx_t (see Remark 4):

$$\Delta C_{t,k} = b_k + m_k \Delta x_t \quad k = 1, \ldots, 5 \ ,$$

with vertex on $(\Delta x_t \ , \Delta C_{t,k}) = (-B_1 \ , B_0)$.

Thus, the ordering of the times series in $\{C_{t,k}\}$ set is given by the ordering of the time series in $\{\Delta C_{t,k}\}$ set which, in turn, is given by the corresponding coefficients m_k (see Remark 1). Additionally, and as a result of Remark 5, the different cutoff points of the time series set $\{\exp \Delta C_{t,k}\}$ take the value e^{B_0} (Figure 6, *top left*). On the other hand, as a result of Remark 2, the only intersection point (by the end of July) between any two time series of the $\{\exp C_{t,k}\}$ set is also an intersection point for all of them (Figure 6, *top right*). This cutoff point describes two clearly different sub-periods in the sample: in the first one, the fifth curve takes the lowest values, followed by the fourth one, the third one, and so on until the first one, which takes the highest value; the situation is just the opposite in the second sub-period which starts at the cutoff point: in this case is the first curve which takes the lowest values, followed by the second one, the third one, and so on until the fifth curve which takes the highest value.

Figure 6. **(From left to right and from top to bottom): IBEX 35 over the time series set $\{\exp \Delta C_{t,k}\}$; IBEX 35 over the time series set $\{\exp C_{t,k}\}$; POP over the time series set $\{\exp \Delta C_{t,k}\}$; POP over the rescaled time series set $\{\exp C_{t,k}\}$**



4.3. Understanding one of the trajectories

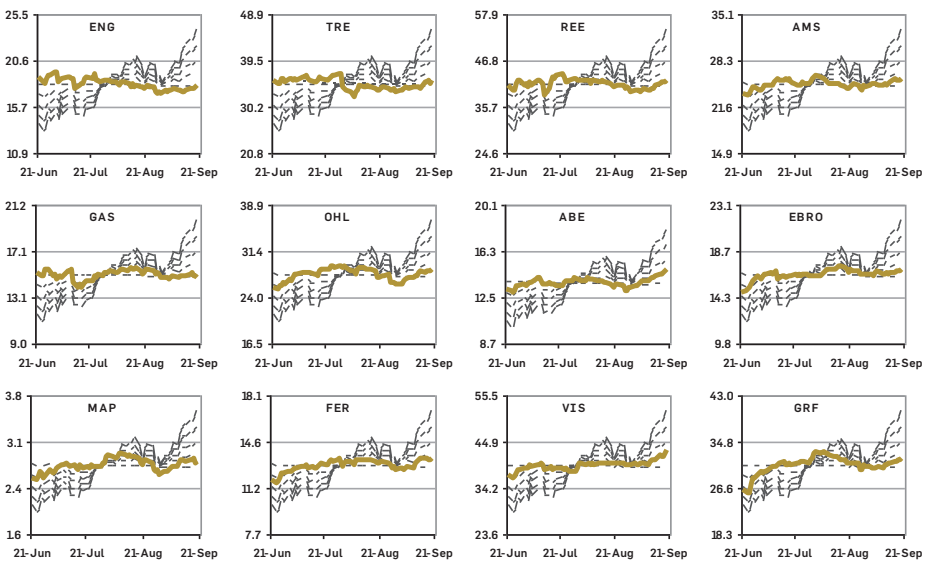
Taking into account that, firstly, the time series in the $\{\exp C_{t,k}\}$ set summarize the thirty-five trajectories of the IBEX 35 constituents and, secondly, the trajectory of IBEX 35 is very close to the first one (Figure 6, *top right*), it is possible to conclude that, in relative terms (compared with its constituents), the values of the IBEX 35 in the second sub-period were not much higher than its values during the first one. But, why is that? The reason behind this result is that, in relative terms, the IBEX35 daily returns were not as high as the returns of some of its constituents; specifically, compared to POP, they were clearly lower (Figure 6, *bottom left*): it can be seen that, exceptions apart, the daily rates of return of IBEX 35 and POP show the same positive sign; since the values of POP are more extreme, it has experienced a higher increase.

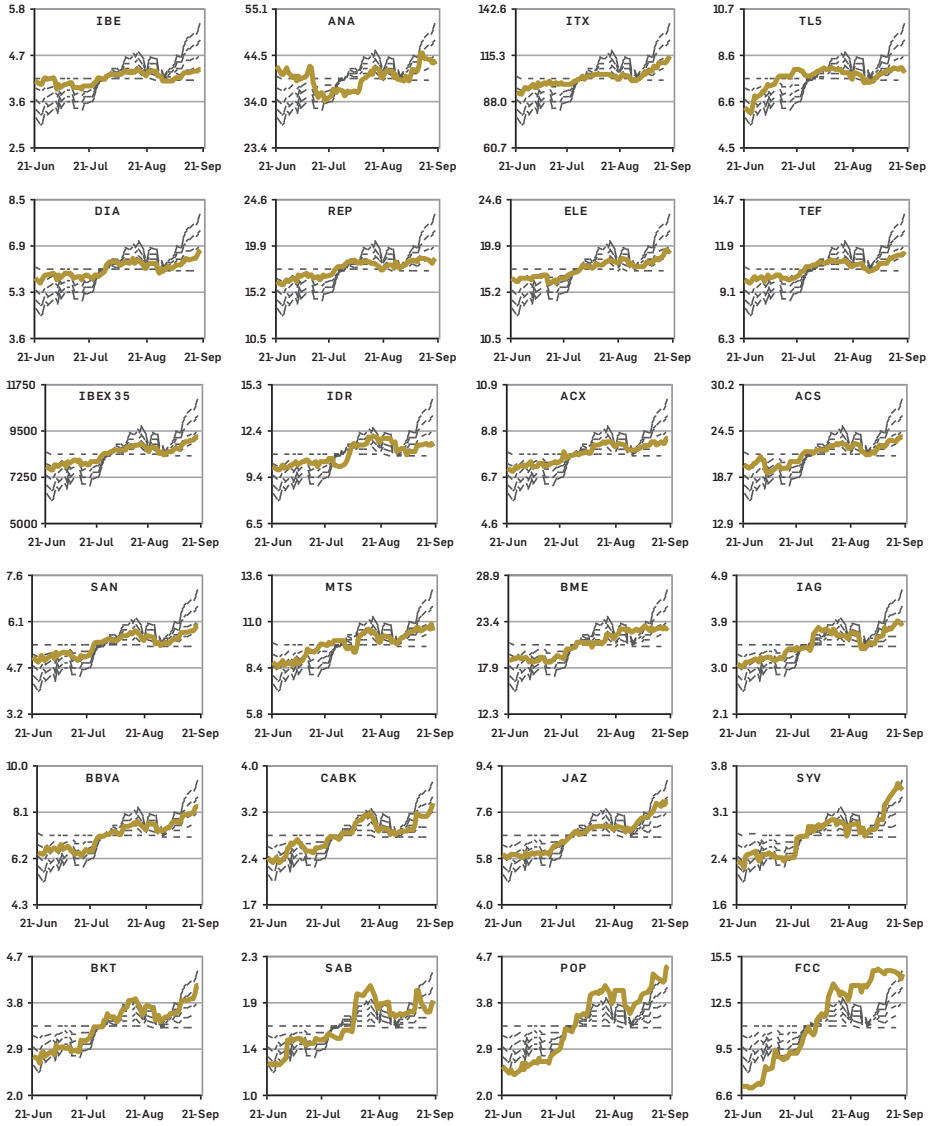
In order to better understand the meaning of the price trajectory of one particular IBEX 35 constituent, $P_{t,j} = \exp Y_{t,j}$, compared with the other ones, we use a graphical representation of the set $\{\exp C^j_{t,k}\}$, where

$$C^j_{t,k} = C_{t,k} - \alpha + \alpha_j \ , \ k = 1, \ldots, 5 \ ,$$

being the scale of the time series set $\{\exp C_{t,k}\}$ the basic reference for the graphical representation of each of the considered variables (Figure 7).

Figure 7. **IBEX 35 constituents over the rescaled time series set $\{\exp C_{t,k}\}$**





Let us see an example. Figure 6 (*bottom right)* shows the POP price trajectory, plotted on the corresponding set of time series $\{\exp C^{j}_{t,k}\}$. It can be seen that, in this case, the variable shows a higher similarity with respect to the fifth curve of the set; in other words, in relative terms (compared to the IBEX 35 and its constituents) we can say that, during the summer of 2013, assuming that the stock market fluctuations respond to a 'sequence of events', the POP responses have been characterized by a high degree of sensitivity to those, which has resulted in a sharp share's price increase.

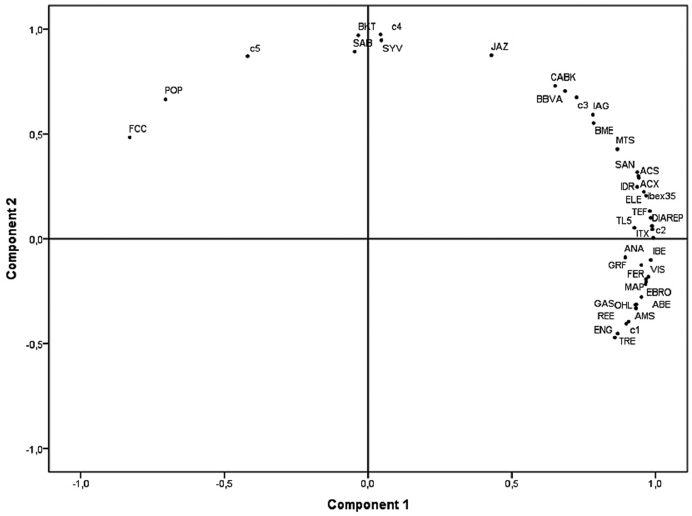
Although the graphical representation of each stock over the 'summary' time series set simplifies the description of the price evolution compared with the other ones, in order to identify similarities and differences among them, comparing the thirty-five

graphical representations would be necessary; to do so, it would be highly desirable that the charts ordering could also contribute to a better understanding. Let us now see how to apply the theoretical basis underlying the process of extraction of the 'summary' time series set to simplify the comparison.

4.4. Comparing historical time series in the Spanish stock market

To establish an ordering of the companies for the sequence of line charts in Figure 7, we will apply a Principal Component Analysis over the matrix of the squared Euclidean distances between each pair of time series of the $\{Z_{t,j}\}$ set[7]. In the corresponding solution for the first two components[8] (Figure 8) it can be seen that the quality of the graphical representation for the different firms is, globally considered, very high[9] and so is the interpretation of their relative positions. Consequently, the ordering of them is established by the parabolic imaginary line that goes through from ENG to FCC.

Figure 8. **Principal Components Analysis applied on the matrix of distances between the IBEX35 constituents: Graphical solution in the space of the first two components**



Besides the thirty five constituents, Figure 8 also includes the IBEX 35 and the 'summary' time series, identified as $c1$, $c2$, ..., $c5$. Note that most of the firms (from ENG to SAN) are concentrated in the first section of the parabolic arc (showing higher values in the first component and lower values in the second one); this is also the case for the IBEX 35 index and for curves $c1$ and $c2$. The reason behind this result is that the firms ENG,

[7] *The justification for this procedure is provided in Ferrán (2011a, 2011b and 2013).*

[8] *Values are given in Table 2.*

[9] *The quality of a point representation is determined by the distance with respect to the origin, being one the highest possible value.*

TRE, REE, AMS, GAS, OHL, ABE and EBRO are very well explained by the first 'summary' time series (Figure 7), while the companies MAP, FER, VIS, GRF, IBE, ITX, TL5, DIA, REP, TEF, ELE, IDR, ACX, ACS, SAN and the IBEX 35 index are properly explained by the second one. In other words, the second group of companies has responded in a more intensive way to the 'sequence of events' that took place during the sample period, resulting in a higher increase in their stock prices. In any case, the response to the events of this second group is lower than the corresponding response of the third group in which MTS, BME, IAG, BBVA and CABK are included (represented close to $c3$ in the parabolic arc). In turns, it will be lower than that of the group made up of JAZ, SYV, BKT and SAB, which will be lower than the response of the fifth and last group of firms, made up of POP and FCC. So, both FCC (Basic Material, Industry and Construction sector) and POP (Financial Services and Real Estate sector) have shown the most intensive response to the events affecting the stock market. This result is not surprising since, as it is well known, the economic sectors of Financial Services and Real Estate and Construction have been the most affected by the recent economic and financial crisis in the Spanish case. However, it is important to note that companies of different sector are included in each of the mentioned groups, showing that both the crisis and the recent incipient recovery process have affected the 35 values included in the IBEX 35 index in a different way to, regardless of sector. So, not all the firms in the Construction sector have responded in the same way to the economic events affecting the stock market, but in a different manner depending on each firm's idiosyncratic factors.

However, if we analyze the IBEX35 firms included in each of the five groups above mentioned in greater detail, we can conclude that the observed IBEX 35 change during the 2013 summer can be explained mainly by the improved performance by the "Financial Services and Real Estate" sectors, and the "Banks and Savings" subsector. In fact, the most important banks are included in groups 3 to 5. This result makes sense since it could reflect that markets have considered that the Spanish financial sector reform program is already giving the expected results.

5. Conclusions

On the 21st of June of 2013 the IBEX 35 reached its minimum value of the year: 7553.2 points; during the summer of 2103, however, the index has showed an increasing trend – reaching the value of 9171.8 on 20th of September. However, the 35 stocks included in the index have been affected in a different manner by this trend change. In this paper, we have applied the SoSLM to compare how the 35 values included in the index have responded to the Spanish stock market recovery process started during the summer of 2013. The results of our analysis allow us to compare the most recent prices of each company relative to both its own past values and those of the other companies.

Specifically, the applied methodology makes it possible to establish an ordering from companies whose responses to the Spanish stock market have been weak to those whose responses have been more intense. We conclude that the observed IBEX 35 recovery process during the past summer of 2013 can be explained mainly by the "Financial Services and Real Estate" sectors, and the "Banks and Savings" subsector. This result is not unexpected since, as is well-known, the Spanish financial sector has been subject to an intensive reform program which seems to have been seen as appropriate by the market.

Acknowledgements

Elena Márquez de la Cruz acknowledges financial support from the Ministry of Economy and Competitiveness, research grant ECO2012-31941.

References

- Ait Hellal, O. and Meyer, G.H. (2014). A wave smoothing algorithm and applications to the financial markets, *AESTIMATIO, the IEB International Journal of Finance*, **8**, pp. 114-131.
- Andrienko, N., Andrienko, G. and Gatalsky, P. (2003). Exploratory spatio-temporal visualization: an analytical review, *Journal of Visual Languages & Computing*, **14**, pp. 503-541.
- BME (2013), Informe Mensual del Mercado Continuo, agosto.
- BME, Technical Regulations for the Composition and Calculation of the Sociedad de Bolsas, S.A.Indexes. Available at http://www.bolsamadrid.es/docs/Indices/Ibex/normdef-in.pdf.
- Cohen, G., Kudryavtsev, A., and Hon-Snir, S. (2011). Stock Market Analysis in Practice: Is It Technical or Fundamental?, *Journal of Applied Finance & Banking,* **1**, pp. 125-38.
- Ferrán Aranaz, M. (2011a), *Una metodología de minería de datos para la agrupación de series temporales: aplicación al sector de la construcción residencial*, PhD Dissertation, Universidad Complutense de Madrid.
- Ferrán Aranaz, M. (2011b), La metodología del haz de rectas para la comparación de series temporales, *Cuadernos de trabajo de la Escuela Universitaria de Estadística*, CT04/2011, Universidad Complutense de Madrid.
- Ferrán Aranaz, M. (2013). El haz de rectas para la comparación de series temporales geográficas, *Estadística Española*, **181**, pp. 123-148.
- Guttman, L. (1953). A Note on Sir Cyril Burt's Factorial Analysis of Qualitative Data, *British Journal of Statistical Psychology*, **6**, pp. 1-4.
- Lin, J., Keogh, E. and Lonardi, S. (2005). Visualizing and discovering non-trivial patterns in large time series databases, *Journal Information Visualization*, **4**(2), pp. 61-82.
- Lobão, J. and Da Mota Lopes, C. (2014). Momentum strategies in the Portuguese Stock Market, *AESTIMATIO, the IEB International Journal of Finance*, **8**, pp. 68-89.
- Murphy, J.J. (1999). *Technical Analysis of the Financial Markets*, New York Institute of Finance, New York.

Appendix: Proofs

Remark 1:

Let's be:

$$\bar{t} = \frac{1}{T}\sum_{t=1}^{T} t = \frac{T+1}{2} \quad \text{and} \quad \bar{x} = \frac{1}{T}\sum_{t=1}^{T} x_t \ .$$

Expressions (2) and (3) imply that:

$$\frac{1}{T}\sum_{t=1}^{T} y_{t,j} = \frac{1}{T}\sum_{t=1}^{T}(b_j t + m_j x_t + \mu_j) = b_j\bar{t} + m_j\bar{x} + \mu_j =$$

$$= (B_0 + B_1 m_j)\bar{t} + m_j\bar{x} + \mu_j = B_0\bar{t} + (B_1\bar{t} + \bar{x})m_j + \mu_j \ .$$

Then, due to (1):

$$\alpha = B_0\bar{t} + (B_1\bar{t} + \bar{x})m_j + \mu_j \ . \quad \forall j$$

Therefore,

$$\mu_j = D_0 + D_1 m_j \quad \forall j,$$

where

$$D_0 = \alpha - B_0\bar{t} \quad \text{and} \quad D_1 = -B_1\bar{t} - \bar{x} \ .$$

Hence, (2) can be also expressed as:

$$y_{t,j} = D_0 + B_0 t + (B_1 t + x_t + D_1)m_j \ .$$

In consequence, if $m_j < m_{j'}$,

$$d(y_{t,j}, y_{t,j'}) = \left(\sum_{t=1}^{T}(D_0 + B_0 t + (B_1 t + x_t + D_1)m_{j'} - (D_0 + B_0 t + (B_1 t + x_t + D_1)m_j))^2\right)^{1/2} =$$

$$= \left(\sum_{t=1}^{T}(B_1 t + y_t + D_1)^2 \cdot (m_{j'} - m_j)^2\right)^{1/2} = (m_{j'} - m_j)\cdot\left(\sum_{t=1}^{T}(B_1 t + x_t + D_1)^2\right)^{1/2} = (m_{j'} - m_j)\cdot \varepsilon \ ,$$

where

$$\varepsilon = \left(\sum_{t=1}^{T}(B_1 t + x_t + D_1)^2\right)^{1/2}$$

is a non-negative constant value independent of both t and k. Therefore, if $m_q < m_r < m_s$ then

$$d(y_{t,q}, y_{t,r}) = (m_r - m_q)\varepsilon < (m_s - m_q)\varepsilon = d(y_{t,q}, y_{t,s}) \ .$$

In other words, under assumptions (1), (2) and (3), the time series from the $\{y_{t,j}\}$ set can be ordered in terms of their similarity: if we assume that the coefficient sequence

m_k is ordered from the smaller to the greater value (given two different series $y_{t,y}$ and $y_{t,j'}$ they must verify that $m_j \neq m_{j'}$):

$$m_1 < m_2 < \ldots < m_J$$

Then, the sequence $y_{t,1}, y_{t,2}, \ldots, y_{t,J}$ is such that

$$\begin{array}{lll} d(y_{t,j}, y_{t,j+1}) < d(y_{t,j}, y_{t,j'}) & j = 1, \ldots, J-2 & j' = j+2, \ldots, J \\ d(y_{t,j}, y_{t,j-1}) < d(y_{t,j}, y_{t,j'}) & j = 3, \ldots, J & j' = 1, \ldots, j-2 \,. \end{array}$$

Remark 2:

If the trajectories of two different time series never intersect, the values of one of them are always greater than the values of the other one:

$$y_{t,j} > y_{t,j'} \qquad \forall t \,.$$

Therefore,

$$\bar{y}_j = \frac{1}{T}\sum_{t=1}^{T} y_{t,j} > \frac{1}{T}\sum_{t=1}^{T} y_{t,j'} = \bar{y}_{j'} \quad ,$$

in contradiction with assumption (1). Hence, there is at least one point where the trajectory of the two time series intersects. Furthermore, if t is the intersection period[10], then,

$$y_{t,j} - y_{t,j'} = 0$$

and, due (5),

$$(B_1 t + x_t + D_1)(m_j - m_{j'}) = 0 \,.$$

Since $m_j \neq m_{j'}$, $\forall j' \neq j$, then, at period t:

$$B_1 t + x_t + D_1 = 0 \,,$$

and hence,

$$(B_1 t + x_t + D_1)(m_j - m_{j'}) = 0 \qquad \forall j' \neq j.$$

Therefore, due to (5), at period t:

$$y_{t,j} = y_{t,j'} \qquad \forall j' \neq j.$$

That is, the intersection points for two trajectories are also the intersection points for the rest.

[10] *The intersection point may appear between two consecutive observations; in such a case the value for t would lie in the time segment defined by the two corresponding time periods. However, for simplicity reasons, the same notation will be used.*

Remark 3:

As we have already seen in Remark 2, if all trajectories intersect at period t then:

$$B_1 t + x_t + D_1 = 0 \ ,$$

and, due to (5),

$$y_{t,j} = D_0 + B_0 t \qquad \forall j \ ,$$

Hence, if all trajectories intersect also at period t', the difference

$$y_{t,j} - y_{t',j} = B_0 \, (t - t')$$

is independent from the time series mean.

Remark 4:

By definitions of $\Delta y_{t,j}$ and Δx_t , the J straight lines:

$$\Delta y_{t,j} = b_j + m_j \Delta x_t$$

derived from (2) are in the sheaf of straight lines with vertex on

$$(\Delta x_t, \Delta y_{t,j}) = (-B_1, B_0) \ ,$$

where B_1 and B_0 are the coefficients on (3).

Remark 5:

At period t it is verified that:

$$\Delta y_{t,j} = b_j + m_j \Delta x_t = B_0 + m_j (\Delta x_t + B_1) \ .$$

Then, at period t ,

$$\Delta y_{t,j} - \Delta y_{t,j'} = (m_j - m_{j'})(\Delta x_t + B_1).$$

Since $m_j \neq m_{j'}$, $\forall j' \neq j$,

$$\Delta y_{t,j} - \Delta y_{t,j'} = 0 \Leftrightarrow \Delta x_t + B_1 = 0 \Leftrightarrow \Delta x_t = -B_1 \Leftrightarrow \Delta y_{t,j} = B_0 \qquad \forall j' \neq j$$

$$y_{t,j} = y_{t,j'} \quad \forall j' \neq j \ .$$