Applied Mathematics and Computation 218 (2011) 4353–4364

Contents lists available at SciVerse ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Discovery of time-inconsecutive co-movement patterns of foreign currencies using an evolutionary biclustering method

Qing-Hua Huang *

School of Electronic and Information Engineering, South China University of Technology, Wushan, Guangzhou, PR China
Guangdong Provincial Key Laboratory of Short-Range Detection and Communication, Guangzhou, PR China

ARTICLE INFO

Keywords:
Foreign exchange rate
Inconsecutive co-movement pattern
Biclustering
Genetic algorithm
Geometric interpretation

ABSTRACT

This paper proposes an evolutionary biclustering algorithm to discover inconsecutive co-movement patterns of different foreign exchange rates. The rows/columns of a bicluster (i.e. a submatrix with a subset of rows and a subset of columns) are not necessarily consecutive. A typical bicluster with constant values on rows and/or columns is represented as a hyperplane in a high-dimensional space and the coefficients of the hyperplane are determined using a genetic algorithm. A detected bicluster demonstrates the co-moving behaviors of a subset of currencies in inconsecutive time periods, indicating that the currencies moved in different manners in some specific time periods. In our experiments, we relate these patterns to the geographically close economic connections and find out the correspondence between the nominal exchange rates and the economic conditions. The findings are useful as a guide for investing foreign currencies.

1. Introduction

Over the past decades, foreign currency exchange rate determination based on international economics and finance has become increasing important since the collapse of the Bretton Woods regime in 1971 [1–4]. Many researchers have worked on modeling and predicting the movement and future trend of floating exchange rates. It is known conventionally that macroeconomic fundamentals significantly influence the exchange rate behavior over the long-run [5–7]. Therefore, earlier models for analyzing the relationship between exchange rate and international macroeconomics consider the power purchasing parity (PPP) and some fundamental variables such as the interest rate and inflation as the core factors which force the exchange rates to reach a final equilibrium [3,7]. However, those models are difficult to validate strictly because of the large variability of economic variables and financial conditions over a long period. Furthermore, an exchange rate often appears to be disconnected from its fundamentals [5,6] due to deviations from PPP and the non-linear relationship between them [8]. The increased volatility in the short-run has thus made the modeling of floating exchange rates a difficult problem [6]. On the other hand, the nominal exchange rate regimes, as well as economic policies, can also affect the variability of macroeconomic quantities [9], resulting in huge macroeconomic changes. Though it is not easy to correctly model exchange rate economics, much attention has been paid to characterization of the natural relationship between exchange rate changes and domestic as well as international fundamentals in recent years [4,5,10].

Nowadays, more and more regional and international economic cooperation organizations are established to strengthen multi-literal trade and coordinate economic policies. A typical and successful example is the European Union (EU) in which 25 countries share a common market and 13 of the 25 countries share the common currency, the EURO. Improved financial

* Address: School of Electronic and Information Engineering, South China University of Technology, Wushan, Guangzhou, PR China.
E-mail address: qhhuang@scut.edu.cn

0096-3003/$ - see front matter
doi:10.1016/j.amc.2011.10.011

and economic connections among the countries within an economic cooperation organization lead to more and more trades and capital flows, hence the closer links of nominal currency exchange rates among the member countries. Discovering the correlations of different currencies would be beneficial to studying the economic and financial connections between different member countries and provide economists and governments with verification evidences that are important for making financial policies. In addition, the correlations are potentially important for guiding investment in the market of foreign exchange rates.

Sopraseuth [9] demonstrated that the floating exchange rates of industrialized countries in a mutual business cycle were associated with high interdependence using a bootstrap test. The similar conclusions can also be drawn by using a conventional clustering technique. However, the economic or political events may often influence the nominal exchange rates in a specific short time period instead of a long time period. Because the specific period associated with a specific event is unknown, statistics based clustering analysis can only be performed with a long time period, hence ignoring important clues for relating short-term fluctuations to the specific economical or political events.

Some attentions have been paid to making use of biclustering methods to analyze the co-movement patterns of exchange rates [11,12]. However, there are drawbacks in those previous studies. In [11], the biclustering method called plaid models detected a predefined small number of biclusters, hence often failing in detection of enough significant patterns. In [12], the authors used Hough transform (HT) to perform the search of biclusters. With a high dimensional data, HT is extremely computational expensive, making this technique impractical for real applications. Furthermore, in those analyses, the biclustering methods were actually regarded as alternatives of conventional clustering methods and took no advantage of their capability of clustering the instances and features simultaneously. As a result, the analysis of co-movement patterns that were not time-consecutive was ignored.

Accordingly, we propose a new biclustering algorithm using a genetic algorithm [13] to discover time-inconsecutive fluctuation patterns among different foreign exchange rates. The algorithm is developed with respect to a geometrical perspective and can find a typical bicluster by detecting a single hyperplane. Because enumerating the currencies is very time-consuming, we use a genetic algorithm to search for a subset of currencies for defining the dimensions of the hyperplane. In the preliminary experiments, the results demonstrate good performance of our algorithm in discovering biclusters from both of synthetic and real data sets with an acceptable computation time.

This paper is organized as follows. The next section provides a brief state-of-the-art discussion of biclustering algorithms. Our biclustering algorithm is detailed in Section 3. We illustrate the preliminary results and corresponding analysis in Section 4, and finally present conclusions in Section 5.

2. A brief state-of-the-art discussion on biclustering algorithms

In conventional clustering methods, a data matrix is partitioned in either the row or the column direction to classify the data into different groups. In biclustering, however, the data are portioned in both row and column directions to discover local subpatterns in the data matrix. The concept of biclustering goes back to the work of Hartigan [14], who called it direct clustering. The term biclustering was first proposed by Cheng and Church [15] in gene microarray data analysis. Because it can overcome the limitation of traditional clustering algorithms in which the rows or columns are clustered globally, biclustering has attracted a lot of attention from researchers in various fields.

In the following context, we present the definitions of several types of biclusters. In a data matrix X with N_r rows and N_c columns, a bicluster is defined as a coherent pattern consisting of a subset of rows and a subset of columns. It can be expressed as a pair (R,C) where $R \subseteq \{1\ldots N_r\}$ is a subset of rows and $C \subseteq \{1\ldots N_c\}$ is a subset of columns. The goal of a biclustering algorithm is to extract all biclusters meeting some evaluation standards. In order to evaluate the quality of a bicluster, Cheng and Church [15] proposed a homogeneity constraint, i.e. the *mean squared residue score* (MSRS), formulated as:

$$h(R,C) = \frac{1}{|R|\cdot|C|}\sum_{i\in R, j\in C}(e_{ij} - e_{iC} - e_{Rj} + e_{RC})^2,$$
$$e_{iC} = \frac{1}{|C|}\sum_{j\in C} e_{ij}, \quad e_{Gj} = \frac{1}{|R|}\sum_{i\in R} e_{ij}, \quad e_{RC} = \frac{1}{|R|\cdot|C|}\sum_{i\in R, j\in C} e_{ij}, \tag{1}$$

where e_{ij} denotes the element value at the ith row and jth column in the bicluster, and $h(R,C)$ the value of MSRS for the bicluster. With a homogeneity threshold δ defining the maximum allowable dissimilarity within the elements of the bicluster, a valid bicluster can be determined if $h(R,C) \leqslant \delta$. The homogeneity threshold is set by users according to their respective applications.

According to Madeira and Oliveira [16], biclusters can be grouped into four major types:

1. Biclusters with constant values,
2. Biclusters with constant values in rows or columns,
3. Biclusters with coherent values,
4. Biclusters with coherent evolutions.