European Journal of Operational Research 000 (2016) 1–11

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Innovative Applications of O.R.

Nonlinear manifold learning for early warnings in financial markets

Yan Huang[a,1], Gang Kou[b,1], Yi Peng[c,1,*]

[a] School of Economics, Xihua University, Chengdu 610039, China
[b] School of Business Administration, Southwestern University of Finance and Economics, Chengdu, China
[c] School of Management and Economics, University of Electronic Science and Technology of China, Chengdu 610054, China

ARTICLE INFO

Article history:
Received 16 June 2015
Accepted 20 August 2016
Available online xxx

Keywords:
Data mining
Manifold learning
Financial markets
Early warning
Dynamic system

ABSTRACT

A financial market is a complex, dynamic system with an underlying governing manifold. This study introduces an early warning method for financial markets based on manifold learning. First, we restructure the phase space of a financial system using financial time series data. Then, we propose an information metric-based manifold learning (IMML) algorithm to extract the intrinsic manifold of a dynamic financial system. Early warning ranges for critical transitions of financial markets can be detected from the underlying manifold. We deduce the intrinsic geometric properties of the manifold to detect impending crises. Experimental results show that our IMML algorithm accurately describes the attractor manifold of the financial dynamic system, and contributes to inform investors about the state of financial markets.

1. Introduction

Scholars and practitioners have developed increasingly elaborate techniques intended to forecast the approach of financial crises. As a complex, dynamic system, financial markets can exhibit tipping points at which abrupt transitions to a contrasting dynamic regime may occur (Scheffer et al., 2009). This shift is called a critical transition in financial markets, and it is exemplified by systemic market crashes or global crises (Ang & Timmermann, 2012). It is difficult to predict reliably when critical thresholds approach because markets might show little change before reaching the critical point (Scheffer et al., 2009). In addition, shifts in financial markets are usually triggered by stochastic and unpredictable externalities (Sugihara et al., 2012; Battiston et al., 2016). However, investors need adequate warning that an impending crisis is highly probable and reduce potential losses.

The intrinsic complexity and nonlinearity of financial markets make it hard to construct an integral mathematical model to characterize the financial system, and thus the corresponding early warning model is unable to be constructed (Christofides, Eicher, & Papageorgiou, 2016; Kou et al., 2014). However, financial time series are comprehensive reflections of market operations and provide a database for market analysis (Ausín, Galeano, & Ghosh, 2014). In practice, observations about the state of dynamic systems are often one-dimensional time series data. Through Phase Space Reconstruction (PSR), time series data can be reconstructed in a space in which the topology is equivalent to the original dynamic system (Richard, Michael, Andrew, & Ye, 2004). PSR describes the trajectory of the dynamic system in the reconstructed high-dimensional space (He, Liu, Long, & Wang, 2012). Takens' embedding theorem shows that the N-dimensional dynamic system has a low-dimensional structure because the system state is confined to an attractor (M) of dimension $d (d < N)$ in the state space (Han & Christopher, 2011).

An underlying manifold governs dynamic systems and reveals their dynamic nature (Sugihara et al., 2012). Therefore, extracting the intrinsic manifold structure is a primary objective of market research. Manifold learning is a hot topic in the fields of data mining and machine learning, which seek to find the intrinsic low-dimensional embedding structures within high-dimensional data. Our study proposes a manifold learning approach to extract the structure of the manifold underlying high-dimensional phase spaces, explore early warning ranges for critical transitions in markets, and discover further intrinsic structural properties.

Numerous manifold learning methods have been developed, including Isometric Feature Mapping (ISOMAP) (Tenenbaum, Sivlar, & Langford, 2000), Locally Linear Embedding (LLE) (Roweis & Saul, 2000), and Local Tangent Space Alignment (LTSA) (Zhang & Zha, 2004). These methods have successfully discovered the embedded low-dimensional manifold. However, classical manifold learning algorithms are concerned with space geometric characteristics. In financial analysis, data information characteristics—i.e., probability distributions—are important (Huang & Kou, 2014). When

* Corresponding author.
E-mail addresses: kougang@yahoo.com, kou.gang@qq.com, kougang@swufe.edu.cn (G. Kou), pengyi@uestc.edu.cn (Y. Peng).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.ejor.2016.08.058

probability density functions (PDFs) are restricted to form an intrinsic manifold of high-dimensional data, geodesic distance no longer accurately describes the manifold distance (Carter, Raich, Finn, & Hero, 2011).

We selected stock market composite indices as observed in financial time series data. Each data point represents a financial system state, and the distance between them indicates the degree of difference between system states. If the difference is characterized only by the geometric space between data points, the result may not fit the practical significance of the financial analysis but rather cause errors in subsequent analyses. Therefore, this study proposes an IMML algorithm to discover the structure embedded in the high-dimensional phase space, which is reconstructed by the observed financial time series range.

Our study is conducted in three steps. First, we reconstruct an observed financial time series as a high-dimensional phase space. Second, we propose the IMML algorithm and employ it to extract the manifold embedded in the high-dimensional phase space. Third, we use the underlying manifold to detect early warning ranges for critical transitions in markets. In addition to the crisis diagnosis, we implement market prognosis from the perspective of the inherent geometric properties in the manifold.

This study is organized as follows. Section 2 reviews related theory and methods. Section 3 describes our manifold learning method. Section 4 reports the experimental study. Section 5 concludes the paper.

2. Preliminaries

2.1. Phase Space Reconstruction

The theoretical basis of PSR originates in Takens' embedding theorem (Takens, 1981, chap. 21). The theorem shows that complete information about the hidden state of dynamic systems can be preserved in observed time series data. The phase space is a time-delay reconstruction using time-delayed versions of a time series as coordinates for the space. Specifically, given a time series $x = x_n,\ n = 1,\ \dots,N$, a reconstructed phase space matrix X of dimension m and time lag τ is defined by its row vectors:

$$x = [x_{n-(m-1)\tau},\ \dots,x_{n-\tau},\ x_n], \tag{1}$$

where $n = (1 + (m-1)\tau)\ \cdots\ N$ and a row vector x_n is a point in the reconstructed phase space.

A proper time lag can reduce the required RPS dimension. A common heuristic for selecting time lag is to use the first minimum of the automutual information function (Richard et al., 2004). The automutual information function is defined as

$$I_n(X_0, X_1,\ \dots,X_n) = \sum_j (H(x_j) - H(X_0, X_1,\ \dots,X_n)), \tag{2}$$

where $H(x_j)$ is the entropy and $H(X_0, X_1,\ \dots,X_n)$ is the joint entropy of the time series data points. τ is at the first local minimum of mutual information.

Embedding dimension m is another vital parameter for PSR, which is not previously known. Many methods seek to determine the dimension parameter, including the global false nearest-neighbor technique and the Cao method (Cao, 1997). We adopt the Cao method for its robust handling of noise and because it presents no need to set threshold values manually. The related process of calculating m is as follows (Cao, 1997):

$$Y_i(m) = (x_i, x_{i+\tau},\ \dots,x_{i+(m-1)\tau}),\ i = 1, 2,\ \dots,N-(m-1)\tau, \tag{3}$$

where m is the embedding dimension, τ is the time lag, and $Y_i(m)$ denotes the ith reconstruction vector with embedding dimension m. Moreover, let

$$a(i,m) = \frac{\left\|Y_i(m+1) - Y_{n(i,m)}(m+1)\right\|}{\left\|Y_i(m) - Y_{n(i,m)}(m)\right\|} \quad i = 1, 2, \dots, N - m\tau, \tag{4}$$

where $\|Y_k(m) - Y_l(m)\| = \max_{0\le j\le m-1} |x_{k+j\tau} - x_{l+j\tau}|$ and $a(i,m)$ $(1 \le a(i,m) \le N - m\tau)$ is an integer such that $Y_{n(i,m)}(m)$ is the nearest neighbor of $Y_i(m)$ in the m-dimensional reconstructed phase space.

2.2. Manifold Learning

A manifold can be viewed as a nonlinear object that is locally linear (Jamshidi, Kirby, & Broomhead, 2011). For high-dimensional real world data, a perceptually meaningful structure has few degrees of freedom. In other words, high-dimensional data points can be mapped into a surrogate low-dimensional space (Seung & Lee, 2000). Hence, it is possible to construct a mapping that obeys specific properties of the manifold and obtains low-dimensional representation of high-dimensional data while preserving the intrinsic structure underlying the data (Lin & Zha, 2008).

Of the many manifold learning methods, ISOMAP and LLE are the earliest. The key idea of the ISOMAP algorithm is to maintain geodesic distance among points on the manifold and embedded data into low-dimensional space through multidimensional scaling. LLE calculates the reconstruction weights of each point and minimizes embedding cost by solving an eigenvalue problem to preserve the proximity relationship among data. LTSA constructs local linear approximations of the manifold by using a collection of overlapping approximate tangent spaces at each data point and aligns these tangent spaces to obtain a global parameterization of the manifold (Zhang & Zha, 2004). LTSA maps the high-dimensional data points on a manifold to points in a lower-dimension Euclidean space. This mapping is isometric if the manifold is isometric to its parameter space. Local Multidimensional Scaling (LMDS) is a data embedding method based on the alignment of overlapping locally scaled patches, and its inputs are local distances (Yang, 2008). A subset of overlapping patches is chosen by a greedy approximation algorithm of minimum set cover. The patches are aligned to derive global coordinates and minimize a residual measure. LMDS is locally isometric and scales with the number of patches rather than the number of data points. LMDS produces less deformed embedding results than LLE. Also a common nonlinear method for dimension reduction, Kernel Principal Component Analysis (KPCA) is a kernel extension of PCA and a special manifold learning algorithm. KPCA conducts traditional PCA in a kernel feature space, which is nonlinearly related to the input space (Jenssen, 2010).

These manifold learning algorithms use a geodesic distance metric or weight measurement to calculate similarities among data points. In financial practice, considering only the geometric structure of a data space disguises essential characteristics of the data and destroys the proximity relations (topology) of the original data space.

2.3. Information distance metric

The theoretical basis of information distance originates in Shannon information theory and Kolmogorov complexity theory. It is framed as the universal cognitive similarity distance that measures the essential relationship between things (Kolmogorov, 1965). Owing to its parameter-free, feature-free, and alignment-free characteristics, it can be used to manage unstructured and incomprehensible data. The information distance metric (Bennett, Gács, Li, Vitányi, & Zurek, 1998) is the Riemannian metric between PDFs p_1

Please cite this article as: Y. Huang et al., Nonlinear manifold learning for early warnings in financial markets, European Journal of Operational Research (2016), http://dx.doi.org/10.1016/j.ejor.2016.08.058

and p_2. Various pseudo-metrics approximate the information distance, including Kullback–Leibler (KL) divergence, mutual information, and Fisher information distance. In fact, the mutual information between two random variables is equal to the KL divergence (Holger & Sławomir, 2006). Moreover, because the pair of densities approaches each other, the KL divergence is an optimum approximation of the Fisher information distance (Carter et al., 2011). Hence, we adopted the KL divergence to estimate the information distance between financial system state points.

The KL divergence between two probability mass functions occurs naturally in information theory, statistics, and physics (Holger & Sławomir, 2006). The definition of the KL divergence, or relative entropy, can be generalized to non-negative vectors (discrete measures) as follows: for any two vectors $u \geq 0$ with $\|u\|_1 = 1$ and $v \geq 0,\ v \neq 0$, the KL divergence between u and v is given by

$$D(u|v) = \sum_{l=1}^{K} u_l \log \frac{u_l}{v_l}. \tag{5}$$

In this correspondence, $u \geq 0$ and $v \geq 0$ are either both positive vectors or $u_l = 0$ if $v_l = 0$. In that case the lth addend is 0. KL divergence is important in information theory, where it is commonly called the relative entropy of one PDF to another.

3. Manifold learning for financial dynamic system

This section extracts the underlying manifold in dynamic financial systems. First, via PSR, we reconstruct the observed financial time series as a high-dimensional phase space. Second, we propose our IMML algorithm, which we use to extract the manifold embedded in the reconstructed phase space.

3.1. Phase Space Reconstruction from financial time series

As discussed in Section 2.1, time lag τ is calculated using the first minimum of the automutual information function. The reconstructed phase space dimension m is calculated by the Cao method.

Through PSR, financial time series can be reconstructed as phase point vectors $\{X_i^m\}$ after determining the embedding dimension (m) and time lag (τ). These vectors form an $n \times m$ matrix in the phase space as follows:

$$\begin{bmatrix} X_1^m \\ X_2^m \\ \cdots \\ X_n^m \end{bmatrix}^T = \begin{bmatrix} x_1 & x_2 & \cdots & x_m \\ x_2 & x_3 & \cdots & x_{m+1} \\ \cdots & \cdots & \cdots & \cdots \\ x_n & x_{n+1} & \cdots & x_N \end{bmatrix}^T = \begin{bmatrix} PS_x^1 \\ PS_x^2 \\ \cdots \\ PS_x^m \end{bmatrix}. \tag{6}$$

In (7), aligning the vectors $\{X_i^m | i = 1, 2, \ldots, n\}$ in order of time produces m vectors $PS_x^j \in R^n$, $j = 1, 2, \ldots, m$, whose element indices correspond to time. The time series vectors $\{PS_x^j |\ j = 1, 2, \ldots, m\}$ can be considered m-dimensional signals and are denoted by $PS_x^m(t)$ for convenience.

According to Takens' theorem, the reconstructed phase space of a dynamic system can be mapped one-to-one into a low-dimensional attractor from the system state space (Richard et al., 2004), which preserves the dynamic properties of the system. Thus, the above reconstructed phase space R^m would converge to a d-dimensional attractor S by the linear map $F : R^m \to R^d$, in which F is an isometric map preserving the distances between points in the dynamic system. Whitney (1936) shows that an attractor of a dynamic system will be embedded in a manifold. We discover the attractor manifold of the financial dynamic system through manifold learning. Classical manifold learning methods obtain the attractor manifolds by preserving the geodesic distances between points in the state space. In financial practice, data points lie in a probabilistic space; the Euclidean distance metric is not appropriated (De Angelis & Dias, 2014). Thus, we propose an information-metric based manifold learning method to extract the attractor manifold embedded in the reconstructed phase space.

3.2. Information-metric based manifold learning algorithm (IMML)

3.2.1. State space modeling

Through PSR, we obtain n reconstructed vectors $(X_1, X_2, \ldots, X_n)$. Each X_i has m embedding dimensions $(X_i^1, X_i^2, ..., X_i^m)$. Each vector X_i is a dataset consisting of m-dimensional vectors, which is $X_i = (X_i^1, X_i^2, ..., X_i^m)$. Assume each data set X_i has an underlying probability distribution function p_i determined by m-dimensional vector components and the parameters are unknown. Then we can obtain a collection of Probability Density Functions $P = \{p_1, ..., p_n\}$, which lie on manifold S. Each element in manifold S is a probability distribution (p_i). We try to extract S in the space of probability densities using the available information in P; i.e., to find an embedding map $F : p(x) \to y$, where $y \in R^d,\ \ d < m$. Unlike the traditional manifold learning algorithm in Euclidean space, our proposed algorithm discovers low-dimensional embedding in the density space (i.e., a manifold of probability distributions).

3.2.2. Information distance metric for financial data points

To obtain the low-dimensional embedding from the high-dimensional datasets, pair wise sample distance, which measures the informational change between data points, should be maintained. There is a corresponding KL divergence $KL(P, Q)$ between any two probability distributions P and Q, where $KL(P, Q) = E[\log \frac{f(x)}{g(x)}] = \int f(x) \log \frac{f(x)}{g(x)} dx$ and P and Q are described by density functions $f(x)$ and $g(x)$, respectively. Divergence is an approximate distance function that meets the non-negative distance definition but does not satisfy symmetry. Since KL divergence is relative Rényi entropy, Rényi quadratic entropy (Jenssen, 2010) is given by $h(p) = -\log \int p^2(x)dx$, where $p(x)$ is the Probability Density Function generating the dataset or sample $X = x_1, x_2, \ldots, x_n$. For n vectors, each vector has m vector components, and the following information metric model can be formulated.

Suppose $P_i = (p_i^1, p_i^2, ..., p_i^m)$, $(i = 1, 2, .., n)$ is the probability distribution vector of the ith vector.

$$p_i^j \geq 0, \sum_{j=1}^{m} p_i^j = 1, (i = 1, 2, ..., n) \tag{7}$$

The information distance between any two vectors is

$$h(P_i, P_j) = \sum_{t=1}^{m} p_i^t \log \frac{p_i^t}{p_j^t}. \tag{8}$$

Since $h(P_i, P_j) \neq h(P_j, P_i)$, the cross entropy does not satisfy symmetry and can be transformed using the following formula:

$$\text{Let}\quad h(P_i, P_j) = h(P_i, P_j) + h(P_j, P_i) = \sum_{t=1}^{m} p_i^t \log \frac{p_i^t}{p_j^t} + \sum_{t=1}^{m} p_j^t \log \frac{p_j^t}{p_i^t}$$
$$= \sum_{t=1}^{m} p_i^t \log p_i^t + \sum_{t=1}^{m} p_j^t \log p_j^t - \sum_{t=1}^{m} p_i^t \log p_j^t - \sum_{t=1}^{m} p_j^t \log p_i^t \tag{9}$$

p_i and p_j can be obtained by kernel density estimators. A kernel density estimator is a nonparametric PDF model that consists in a linear combination of kernel functions centered upon the data (Leiva-Murillo & Artés-Rodríguez, 2012). Given by $\hat{p(x)} \propto \frac{1}{N} \sum_{i=1}^{N} k(x - x_i, \sigma)$, where $x \in R^D$ and $k(x - x_i, \sigma) = \exp(-\|x - x_i\|^2/\sigma^2)$, $k(x - x_i, \sigma)$ is the kernel function with a given bandwidth σ.

Please cite this article as: Y. Huang et al., Nonlinear manifold learning for early warnings in financial markets, European Journal of Operational Research (2016), http://dx.doi.org/10.1016/j.ejor.2016.08.058

Table 1
Manifold learning for a financial dynamic system.

Input: Financial Time Series Data $X_{n\times 1}$
Output: low-dimensional coordinates Y.

Procedure:
1. Financial time series data $X_{n\times 1}$ are reconstructed into an m-dimensional phase space $X'_{n\times m}$.
(Time lag τ is obtained by automutual information, and initial embedding dimension m results from Cao method.)
2. Extract the Attractor Manifold via IMML
(1) Compute information distance metric h_{ij} by (10) and obtain the relationship matrix H.
(2) Compute the smallest d ($d << m$) non-zero Eigenvalue of matrix $(I-H)(I-H)^T$ and the corresponding Eigenvectors Y by (12).

3.2.3. Low-embedding manifold of financial dynamic system

In the classical LLE method, the local linear structure between neighbors remains unchanged even after dimensionality reduction. For financial time series data points, the concept and scope of neighbors would be extended. Unlike image data sets, the adjacency relationships of financial data points do not entirely depend on the geometric relationships of data points. We assume every data point and its neighbors are located on the same linear manifold. When reproducing a low-dimensional manifold, the corresponding data points in the intrinsic low-dimensional space maintain the same global neighbor relationship. To obtain the low-dimensional representation of datasets, our algorithm constructs an extended local linear structure while retaining the global topological characteristics in the inherent low-dimensional manifold. As mentioned, the relationship metric $h_{ij} = h(P_i, P_j)$ reflects the essential relationships between vectors in the financial time series. Therefore, we can obtain the global relationship metric matrix $H = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1n} \\ \dots & \dots & \dots & \dots \\ h_{i1} & h_{i2} & \dots & h_{in} \\ h_{n1} & h_{n2} & \dots & h_{nn} \end{pmatrix}$, which is the reconstruction weight matrix and may be mapped to the low-dimensional embedding manifold. Similar to the LLE method, we obtain the low-dimensional embedding by solving the following optimization problem:

$$Min \quad \Phi(Y) = \sum_{i=1}^{n} \left\| Y_i - \sum_{j=1}^{n} h_{ij} Y_j \right\|^2. \tag{10}$$

The low-dimensional embedding Y reflects the corresponding reconstruction weight relationship of the sample points in the high-dimensional input space. To eliminate the coordinate translation, rotation, and scaling factors of the low-dimensional embedding, we add two constraints: $\sum_{i=1}^{n} Y_i = 0$ and $\frac{1}{n-1}\sum_{i=1}^{n} Y_i Y_i^T = I$, where I is the identity matrix. Furthermore, (11) can be written as

$$\Phi(Y) = \sum_{i=1}^{n} \left\| Y_i - \sum_{j=1}^{n} h_{ij} Y_j \right\|^2 = \left\| (I-H)^T Y^T \right\|^2 = tr(YMY^T), \tag{11}$$

where $M = (I-H)^T(I-H)$ is the $n \times n$ matrix. To minimize the cost function, low-dimensional embedding Y should be taken as the corresponding eigenvectors $v_1, \dots, v_d$ to the bottom d nonzero eigenvalue of the matrix M, i.e., $Y = [v_1, \dots, v_d]$. We adopt the Renyi information dimension (Renyi, 1959) to determine the intrinsic dimension d. Information dimension D was defined as follows:

$$D = -\lim_{\varepsilon \to 0} \frac{\sum_{i=1}^{N} P_i \log P_i}{\log \varepsilon} \tag{12}$$

where P_i denotes the probability of a point falling into the ith unit; here it denotes the probability distribution of the ith vector obtained above. ε is the standard body, and N is the number of points. Thus, we can obtain the intrinsic dimension by information dimension d.

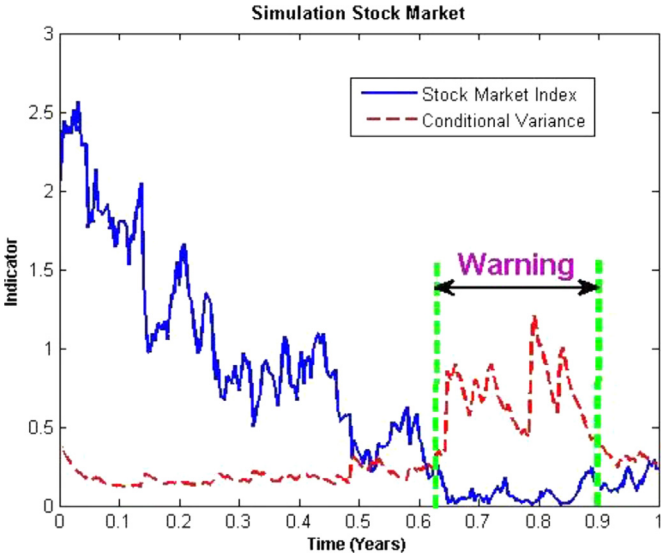


Fig. 1. Simulation stock market.

The proposed manifold learning-based method for a financial dynamic system is summarized as Table 1.

4. Experiments

This section applies our proposed methodology to the early warning for critical transitions in financial markets. Our empirical studies would be conducted on artificial and real datasets, respectively.

4.1. Data descriptions

4.1.1. Artificial time series generated by Monte Carlo simulation

Considering the features of market motion, we employ a Monte Carlo simulation of Stochastic Differential Equations to model stock markets and generate the time series. The simulation market model is:

$$dX_t = \mu X_t dt + D(X_t)\sigma dW_t, \tag{13}$$

where X_t is a state vector of market motion process variables to simulate, μ is a diagonal matrix of expected index returns, D is a diagonal matrix with X_t along the diagonal, σ is a diagonal matrix of standard deviations of index returns, and W_t is a Wiener process vector. The initial values of parameters are $\{\mu = 0.25,\ \sigma = 0.2\}$, simulating 250 trading days. Results are shown in Fig. 1. Traditionally, conditional variance is often used as an early warning indicator in financial time series. Previous studies show that the conditional variance generally exhibits violent oscillations prior to critical transitions in a system (Brock & Carpenter, 2012). Using nonparametric regression, the values of conditional variance would be obtained as also shown in Fig. 1.

Please cite this article as: Y. Huang et al., Nonlinear manifold learning for early warnings in financial markets, European Journal of Operational Research (2016), http://dx.doi.org/10.1016/j.ejor.2016.08.058

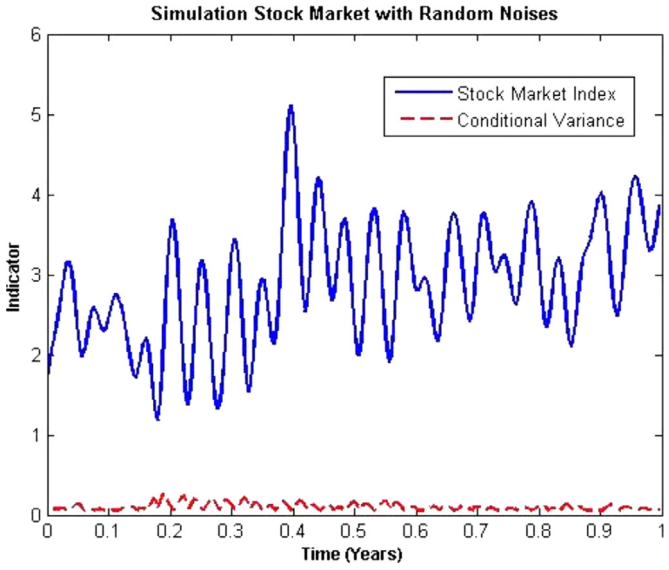


Fig. 2. Simulation stock market with random noises.

In the period spanning 0.6–0.9, the oscillation in conditional variance intensifies, indicating the market is approaching a transition point (Brock & Carpenter, 2012). A crisis warning is issued during this period.

In practice, however, financial markets are disturbed by indiscriminate noise, so that the market index deviates from its real state. Therefore, the simulation should consider noise disturbance. The multiple random noises would be added to the original time series. Results appear in Fig. 2.

As Fig. 2 reveals, noise disturbance renders the original time series unable to reflect the market's actual situation. Moreover, the corresponding conditional variance has lost its early warning effect (Faranda, Emanuele Pons, Giachino, Vaienti, & Dubrulle, 2015). The artificial time series with random noise would be selected in our experiments.

4.1.2. Real datasets

We selected the daily closing prices of CSI 800 and the S&P 500 Index during the period 2005–2015 selected as experimental time series data. The CSI 800 index describes the overall trend of the China A-share market. It offers an overview and running status of price changes in China's securities market. The S&P 500 is a commonly followed US equity index based on the market capitalizations of 500 large companies. It is recognized as one of the best representations of the U.S. stock market and an indicator for the U.S. economy.

4.2. Experimental design

The experiments were conducted as follows:

(1) Phase Space Reconstruction. Using PSR, we constructed the high-dimensional dynamic systems from one-dimensional time series.
(2) Manifold Learning of Financial Dynamic System. Using our proposed IMML method, we discovered the intrinsic attractor manifolds from the reconstructed high-dimensional dynamic flows obtained in step 1.
(3) Early Warning for Critical Transitions in Financial Markets. We adopted the Hidden Markov Model (HMM) on the attractor manifolds to detect early warning signals for critical market transitions.

Table 2
PM values in Fig. 4.

Time series	IMML	LTSA	LMDS	ISOMAP	LLE	KPCA
Simulation data	0.0053	0.0079	0.0086	0.0085	0.0091	0.0083
CSI800	0.0015	0.0058	0.0057	0.0137	0.0396	0.0072
S&P500	0.0037	0.0066	0.0061	0.0118	0.0104	0.0085

In addition, in the empirical study of real datasets, we further derived differential curvatures for the underlying manifold to diagnose the prognosis for a market crisis. The perspective on intrinsic geometric properties in the dataset provides a new tool for diagnosing the robustness of financial markets.

4.3. PSR from financial time series

Let $\{x_t\}_{t=1,2,\ldots,n}$ be the observed time series. Per Takens' theorem, we can reconstruct the vector time series $\{x_t\}_{t=1,2,\ldots,n}$ of n points with $X_t = (x_t, x_{t-1}, x_{t-2}, \ldots, x_{t-(m-1)\tau})$. We adopt the mutual information method and the Cao method to estimate the time lag τ and initial dimension m, respectively. When m and τ are appropriate, the evolution of X_t is topologically equivalent to the underlying dynamic flow. In this experiment, the initial parameters of artificial datasets are $\{m = 8,\ \tau = 3\}$; the two real datasets are $\{m_{CSI800} = 10,\ \tau = 15\}$ and $\{m_{S\&P500} = 8,\ \tau = 23\}$, respectively. Calculation results appear in Fig. 3:

$I(\tau)$ describe how the information content of a signal decreases over prediction range τ. Escalating information loss is related to declining predictability and rising complexity of the signal. The minimum value of $I(\tau)$ indicates the maximum likelihood of no correlation between x_i and $x_{i+\tau}$. Thus, we use the first minimum of $I(\tau)$ as the optimal time lag. Cao defined the evaluation function $E1(d)$ to determine the embedding dimension d, in which the best estimate corresponds to the first minimum of $E1(d)$.

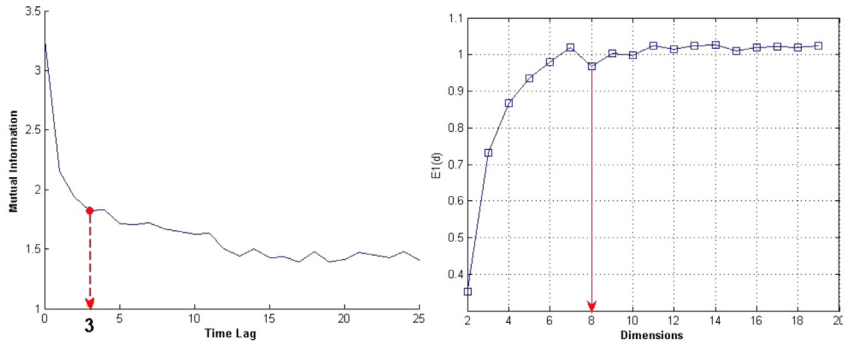
Through the auto-mutual information function and Cao's method, we obtain the original dynamic system. However, the dynamic nature of the original system is usually corrupted by irrelevant components that can disturb useful intrinsic features. Therefore, we propose our IMML to extract the inherent structure embedded in the original dynamic system, which provides data supporting the subsequent early warning analysis.

4.4. Manifold learning for financial dynamic system

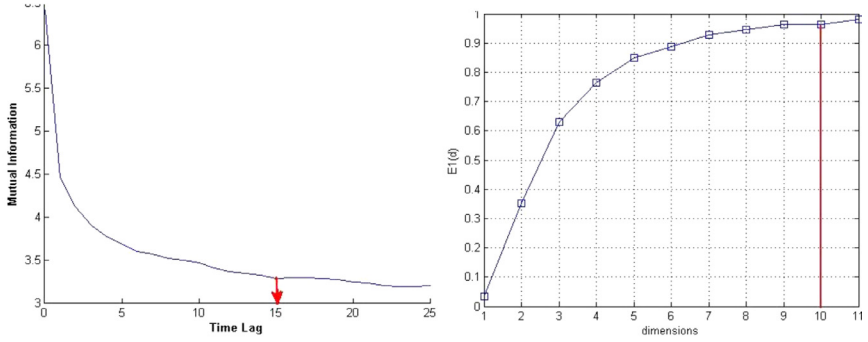
The high-dimensional phase spaces have been obtained in Section 4.3. Therefore, we use our IMML algorithm to extract the underlying manifold S from the high-dimensional dynamic data and to discover the intrinsic structure of the dynamic system. To evaluate the performance of our proposed algorithm, we compare the IMML algorithm with classic and contemporary manifold learning algorithms such as ISOMAP, LLE, KPCA, LMDS, and LTSA. Experimental results appear in Fig. 4. We adopt the Procrustes Measure (PM) to quantify the resulting low-dimensional embedding. PM is a quantitative indicator for nonlinear measurement of goodness that analyses the shape distribution of a data set by statistical analysis. Smaller PM values indicate more accurate embedding (Seber, 2004). MATLAB provides a PM function to compute the corresponding PM values. Evaluation results by PM are in Table 2.

As Table 2 reveals, the IMML method achieves the best results on the three time series datasets. Since the financial practitioners are constrained to probability distributions, IMML exhibits much stronger performance in deciphering practical financial data than the simulation data. The five comparative analysis algorithms use Euclidean distances between sample points to conduct the dimensionality reduction algorithm; however, their performance is unsatisfactory for financial practice, despite their soundness in

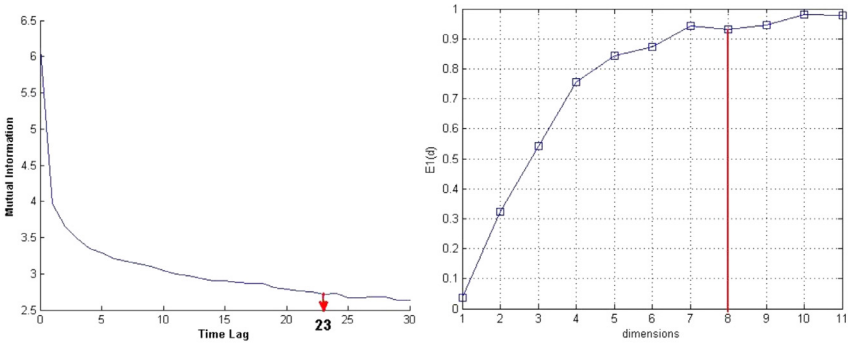
Please cite this article as: Y. Huang et al., Nonlinear manifold learning for early warnings in financial markets, European Journal of Operational Research (2016), http://dx.doi.org/10.1016/j.ejor.2016.08.058



Time Lag and Embedding Dimension for Artificial Time Series



Time Lag and Embedding Dimension for CSI800



Time Lag and Embedding Dimension for S&P500

Fig. 3. Respective results from mutual information and the Cao method.

other fields. Thus, we obtain the accurate low-dimensional representations S embedded in the high-dimensional financial dynamic flows. The three low-dimensional representations are the attractor manifolds that determine the trajectories of the corresponding financial system. These provide data support for the following study.

4.5. Early warning for critical transitions in financial markets

As noted, financial markets have critical points at which a sudden shift to a contrasting regime may occur. It is difficult to predict critical transitions accurately before their tipping points appear (Scheffer et al., 2009). The likelihood of a transition may gradually become evident as a system approaches the point at which a minor trigger can invoke a self-propagating transition to a contrasting state (Scheffer et al., 2012), such as a "flash crash" (Andersen & Bondarenko, 2015). Since it is hard to predict the tipping point and trigger, the gradually increasing likelihoods of such transitions become the practical focus. Therefore, we need to detect points that show a high likelihood of indicating the approach of a transition. Especially, the ranges composed of dense these points should receive close attention. Our detection is implemented on the obtained attractor manifolds.

4.5.1. HMM classifier for critical transitions diagnosis

We adopt the HMM to detect early warning signals for critical transitions in markets. HMMs are effective statistical models for describing and analyzing financial time series data (Maheu & McCurdy, 2009; Rydén, Terasvirta, & Asbrink, 1998). Given an observed time series $\mathcal{X} = \{x_1, \ldots, x_T\}$, the hidden states of the sequence are denoted as $S = \{s_1, s_2, \ldots, s_M\}$ where M is the number of hidden states and N in the observation signals for each state denoted as $V = \{\mathcal{V}_1, \mathcal{V}_2, \ldots, \mathcal{V}_N, \}$. Therefore, an HMM can be briefly defined as $\lambda = (\pi, A, B)$ and has the following components (Rabiner, 1989):

Please cite this article as: Y. Huang et al., Nonlinear manifold learning for early warnings in financial markets, European Journal of Operational Research (2016), http://dx.doi.org/10.1016/j.ejor.2016.08.058

(1) Initial state probability distribution π defined as $\pi = \{\pi_i\}$ where

$$\pi_i = P[s_l = i],\ 1 \le i \le M \tag{14}$$

(2) State transition matrix A defined as $A = \{a_{ij}\},\ 1 \le i,\ j \le M$, where

$$a_{ij} = P[x_{t+1} = s_j | x_t = s_i] \tag{15}$$

a_{ij} represents the probability of the transition for state s_i at time t to state s_j at time $t+1$.

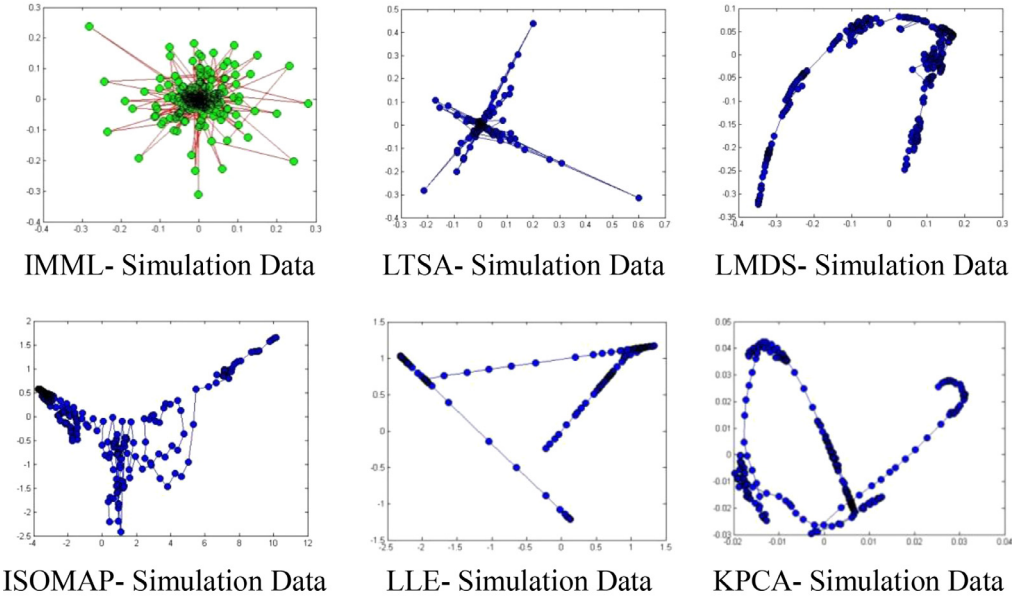
(3) Observation probability distribution B, (i.e., emission probability) defined as $B = \{b_i(k)\},\ 1 \le i \le M,\ 1 \le k \le N$ where $b_i(k) = P[v_k | s_t = i]$ is the probability of observing v_k at state s_i at time t.

The specific steps of detection are as follows.

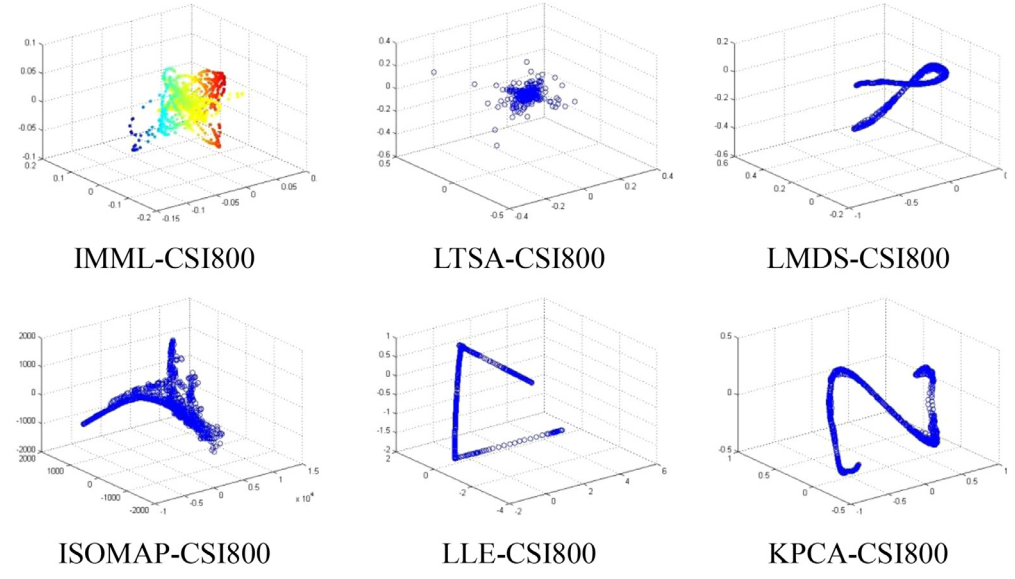
(1) The first step is to estimate the initial state probability distribution π. Here, Gaussian Mixture Models (GMMs) (Figueiredo & Jain, 2002) expose the probability distribution for each point on the manifold. A GMM is defined as

$$p(x) = \sum_{i=1}^{L} w_i p_i(x) = \sum_{i=1}^{L} w_i N\left(x; \mu_i \sum_i\right), \tag{16}$$

where L is the number of mixtures, $N(x; \mu_i, \Sigma_i)$ is a normal distribution with mean μ_i and covariance matrix Σ_i, and w_i is the mixture weight with the constraint $\sum w_i = 1$. Parameters for the GMM are estimated by the well-known expectation–maximization algorithm (Moon, 1996). This iterative method yields a maximum likelihood estimate via the



(a) Manifold about Simulation Data



(b) Manifold about CSI800

Fig. 4. Manifold learning of financial dynamic system.

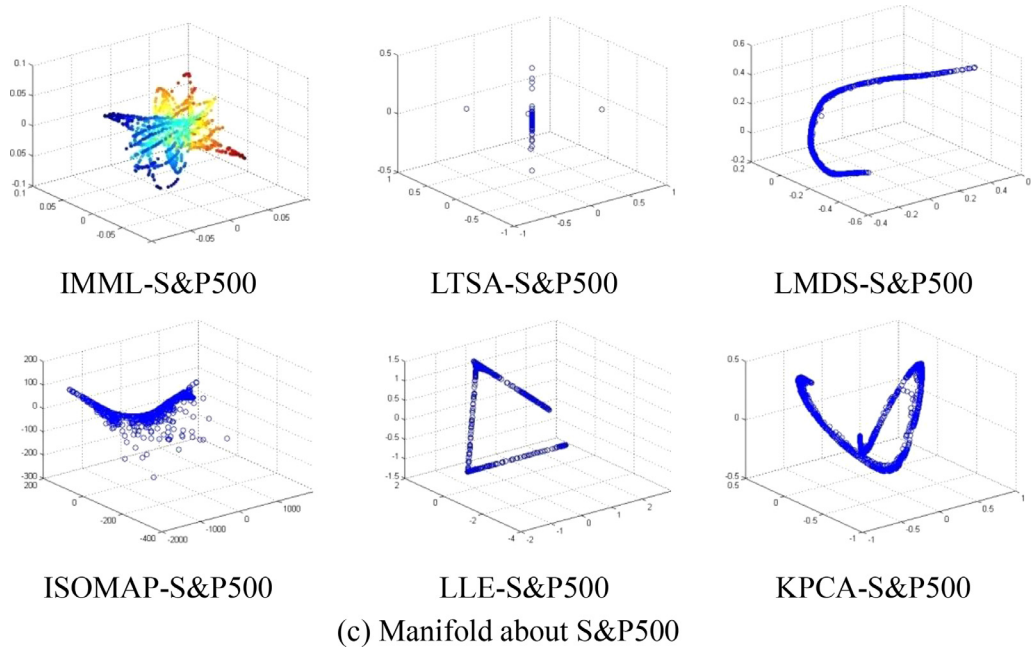


Fig. 4. Continued

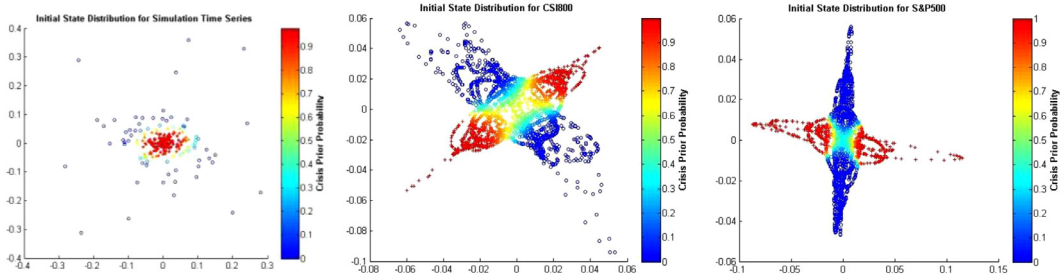


Fig. 5. Initial probability distributions by GMMs. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

estimation formulae:

$$\mu_i' = \frac{\sum_{t=1}^{n} p_i(x_t)x_t}{\sum_{t=1}^{n} p_i(x_t)}, \quad \sum_i^{'} = \frac{\sum_{t=1}^{n} p_i(x_t)(x_t-\mu_i)^T(x_t-\mu_i)}{\sum_{t=1}^{n} p_i(x_t)},$$

$$w_i' = \frac{\sum_{t=1}^{n} p_i(x_t)}{\sum_{t=1}^{n}\sum_{i=1}^{L} p_i(x_t)}. \tag{17}$$

(2) The second step is to construct the state transition matrix A. We obtain the state transition probability distribution by the relationship matrix H based on information divergence in Section 3. With the relationship matrix H, we have

$$a_{ij} = P\left[x_{t+1} = s_j | x_t = s_i\right] = \frac{P(x_{t+1}, x_t)}{P(x_t)} = \frac{h(P_j, P_i)}{h(P_i)} = \frac{h_{ij}}{h_{ii}}. \tag{18}$$

(3) The third step is posterior inference by emission probability. The posterior probability $P(S|x(t))$ for each class S: is

$$P(S|x(t)) = \sum_{c=1}^{S_c} \frac{P_c^S(t)}{\sum_{S'=1}^{S}\sum_{c'=1}^{S_{c'}} P_{c'}^{S'}(t)}, \tag{19}$$

where $P_c^S(1) = \pi_c^S b_c^S(x(1))$, $P_c^S(t) = \sum_{c=1}^{S_c} P_{c'}^S(t-1)a_{cc'}$, $b_c^S(x(t))$ $(1 \le t \le T)$, and $b_c^S(x(t))$ can be derived by expectation–maximization algorithm. The classification of each point would be identified by the highest posterior probability.

4.5.2. Detection results

We assume that observation sequences are independent and identically distributed random variables. Thus, according to the 3σ criterion of the central limit theorem, the effect-thresholds for posterior probabilities, namely three early warning states, would be defined as follows: the blue warning state (S1) indicates a $50\% \sim 70\%$ probability of crisis, a yellow warning state (S2) a $70\% \sim 90\%$ probability, and a red warning state (S3) a $90\% \sim 100\%$ probability. Red warning ranges warrant special attention because they indicate "in crisis or on the verge of crisis."

Through GMM, the initial state distributions appear in Fig. 5. Trajectory densities of the data points would be built in the reconstructed phase spaces.

Fig. 5 shows two initial states, respectively defined as "normal" and "crisis." For each data point, the greater the risk, the closer the state approaches red. For the market simulation, the number of red warning data points exceeds the number of blue points, indicating that the market is at high risk during this period; for the two real datasets, we can roughly see that the number of red warning data points for CSI800 exceeds that for the S&P500. Accordingly, we infer that there was much more risk in the Chinese than in the U.S. market during the past few years.

As mentioned in Section 4.5.1, the state transition probability would be obtained by the relationship matrix H in Section 3. The posterior probabilities for each class would be estimated by EM algorithm. According to the highest posterior probability, the warning states for the data points are determined as shown in Fig. 6.

Please cite this article as: Y. Huang et al., Nonlinear manifold learning for early warnings in financial markets, European Journal of Operational Research (2016), http://dx.doi.org/10.1016/j.ejor.2016.08.058

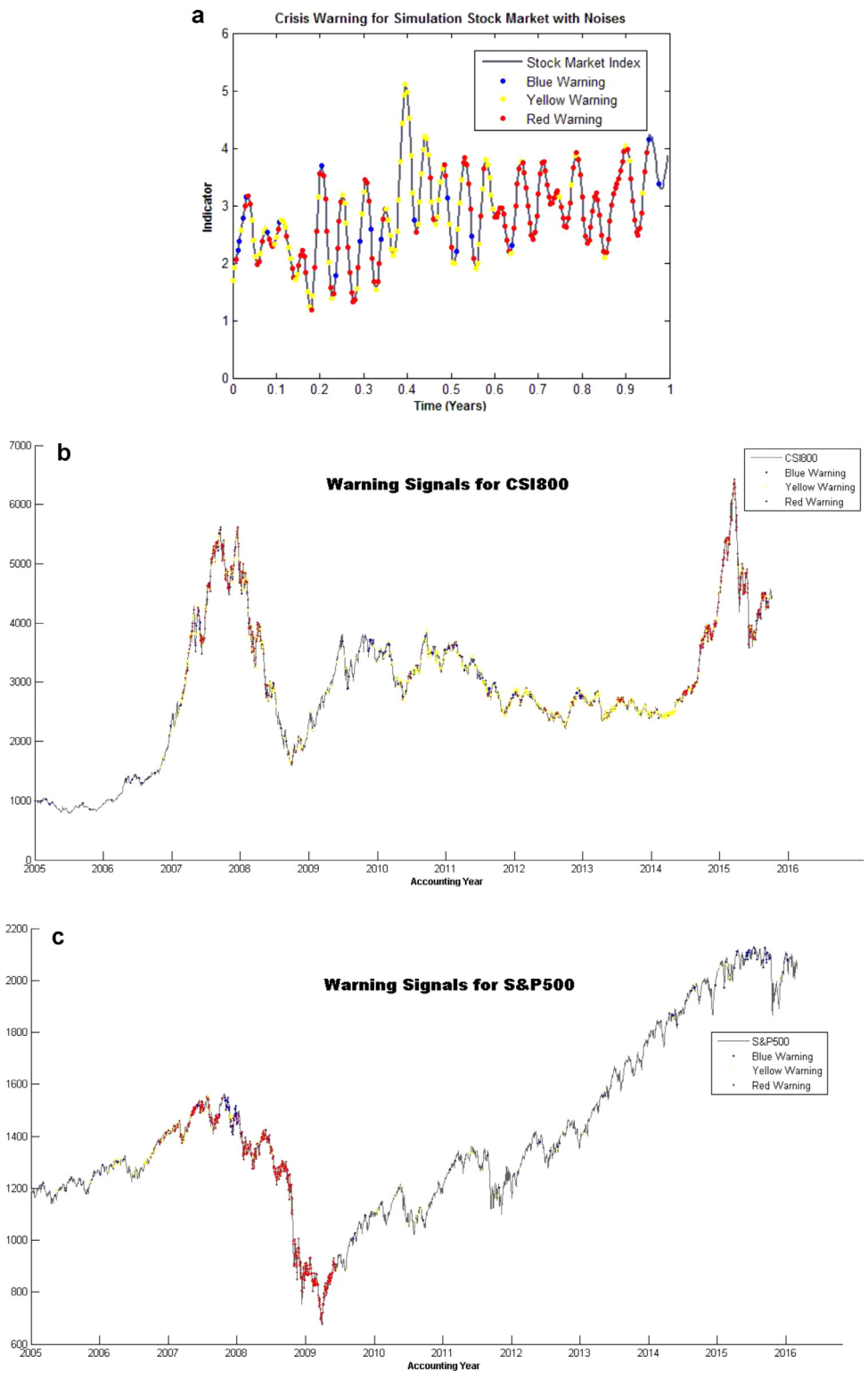


Fig. 6. Early warning signals for critical transitions in financial markets. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

As Fig. 6(a) shows, during the simulated trading days the stock market is threatened by growing crises. Red warning points are concentrated at [0.1, 0.3] and [0.6, 0.9]. Especially in [0.6, 0.9], nearly all warning points are red, indicating the stock market has suffered a crisis. Compared with Fig. 1, our method generates early warning signals consistent with the market's actual situation as reflected by the original data. In addition, our method provides more detailed warning information. Therefore, our method provides reliable early warnings when serious noise disturbances prevent traditional indicators from doing so.

As shown in Fig. 6(b), the warning states indicated by the data points gradually changed from blue to red in 2006–2008, indicating an approaching crisis. In 2006, China's stock market index rose rapidly, reaching an all-time high in 2007, while the yellow

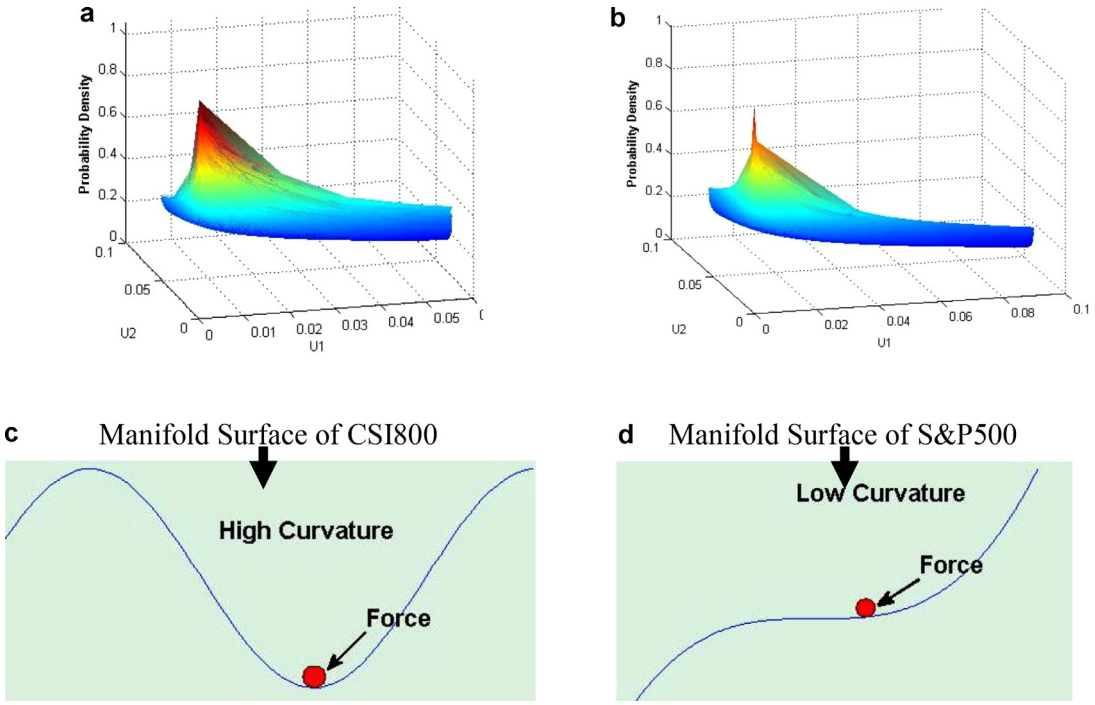


Fig. 7. Manifold surface and curvature.

warning signals reveal the crisis contained within that prosperity. Affected by the US financial crisis, many Chinese enterprises closed or collapsed in 2008, and investors suffered heavy losses. Since 2011 the blue warning states have gradually diminished, yellow warning states have gradually increased and red warning states have begun to appear, indicating China faces a financial crisis. In 2014 and especially in 2015, a large number of red warning states appeared, showing China on the verge of crisis or having entered a crisis.

As Fig. 6(c) shows, the red early warning ranges are mainly concentrated between 2007 and 2009. The global financial crisis starting in 2007 began with an initially well-defined flashpoint focused on mortgage-backed securities and cascaded into a global economic collapse, the increasing severity and uncertain duration of which generated massive losses and damage for billions of people (Sornette & Woodard, 2010, chap. 6). Its outbreak came as a blow with a long-term build-up during the economic bubble. In the U.S. between 2005 and 2006, rapid expansion of the real estate bubble led to an accompanying bubble of subprime mortgage-backed securities and complex packages of associated financial derivatives (Yuliya & Iftekhar, 2010). These spilled over into the stock market, and the bubble appeared as super-exponential growth. At that time, the financial crisis could have been triggered at any moment. Our method detects numerous yellow warning data points during the period.

Investors would not need early warnings as soon as the bubble economy appeared, for the bubble was inevitable. Nor would they need them at the moment of critical transition, as whatever triggers the crisis would be unpredictable. However, when the probability of the crisis breaking out has accumulated to a discernible extent, there is a need to warn investors, to give them time to prepare for its arrival. Our method provides an effective early warning signal for that purpose.

All of the results above constitute a crisis diagnosis obtained by our method. As shown in Fig. 6(b), red or yellow warning ranges appear during 2014–2016. In the face of high risk, our method signals an approaching systemic financial crisis, especially for China.

4.6. Crisis prognosis

Our IMML algorithm detects informational geometric structures in the reconstructed phase space of the financial time series, as shown in Fig. 7(a) and (b). In differential geometry, mean curvature represents the degree of a curved surface, its intrinsic property. According to the differential manifold definition, a manifold is a generalization of the Euclidean space; every local manifold can be seen as Euclidean space, and the various local manifolds are "bonded" properly. Therefore, we can calculate the mean curvature of the minimal surface of each point; then, through the integral method, all mean curvatures of the minimal surface can be bonded to obtain the total curvature of the surface. Let c represent the mean curvature of the minimal surface and C the total curvature. Then $C = \iint_{\Omega} c d\sigma$, where $d\sigma$ is the minimal surface and Ω the observed surface. In our experiment, for the two manifold surfaces, $C_{\mathrm{CSI800}} = 6.7175 > C_{\mathrm{S\&P500}} = 2.9954$. The significance of the curvature is shown in Fig. 7(c) and (d).

According to Fig. 7, the curvature of the manifold surface for CSI800 is higher than for the S&P500. As seen in Fig. 7(c) and (d), when a small ball falls on a "high curvature" or "low curvature" surface, the same force is applied to the ball on the different surfaces. The ball that falls on the high-curvature surface will have difficulty falling to another place, whereas the ball on the low-curvature surface can easily fall elsewhere. In other words, the ball on the high-curvature surface is less susceptible to external disturbance and alterations in its state. Thus, we can say, to some extent, the robustness and resistance to disturbance of the Chinese market are higher than those of the U.S. market.

CSI800 reflects the overall situation of the Chinese stock market, which offers a glimpse into the Chinese economy. Examining this more closely reveals, first, that the Chinese government has strong macroeconomic regulations and market control. Macroeconomic policies are a prerequisite for financial stability (Schaller & Van Norden, 1997); therefore, they subjectively enhance the resilience of the economic system and improve the curvature of its intrinsic manifold. Second, through reform and the opening up of

Please cite this article as: Y. Huang et al., Nonlinear manifold learning for early warnings in financial markets, European Journal of Operational Research (2016), http://dx.doi.org/10.1016/j.ejor.2016.08.058

the system, especially after the 2008 financial crisis, the Chinese government has vigorously adjusted and optimized the economic structure. Its actions make the system more reasonable and objectively increase its capacity to handle external disturbances and accept risks.

July 2015 brought great volatility to China's stock market. Although its fluctuations greatly affected Chinese financial markets and investors, their impact was far from resembling a systemic financial crisis (Sevim, Oztekin, Bali, Gumus, & Guresen, 2014). The Chinese government has acted to deter financial crimes and corruption and to rectify and standardize market order. These actions purify the market environment. With continuous improvement in the macro environment, China's stock market has gradually stabilized since September 2015 even though short-term high volatility has appeared because of the circuit breaker mechanism in January 2016. Therefore, since the Chinese economy is characterized by strong macroeconomic regulations despite its high financial risks, the Chinese government can find a way to release its potential, and the country's long-term economic expectations encourage confidence.

No matter the outcome, the early warning signal remains an important decision tool for investors, one that would help avoid crises or limit their negative effects (Lang & Schmidt, 2016; Kou & Lin, 2014; Kou & Ergu, 2014).

5. Conclusion

This study has proposed an information metric-based manifold learning (IMML) algorithm to discover the attractor manifold of a dynamic financial system. In contrast to traditional manifold learning methods, IMML employs its information metric to measure relationships between financial vectors in the reconstructed phase space and yields a reasonable and accurate low-dimensional embedding manifold. Our method identifies the periods indicating critical market transitions and provides reliable early warnings for practitioners. We further deduce the differential curvature of the financial system through the attractor manifold and find that China's financial system is highly resilient. In the future, we will extend this analysis to other fields, such as mechanical engineering and biometrics, and investigate a more flexible algorithm to discover underlying manifold structures.

Acknowledgments

We would like to thank two reviewers and the editor for their comments, which are helpful for us to improve the paper. This research was supported in part by grants from the National Natural Science Foundation of China (#71222108, #71325001 and #71471149) and Major project of the National Social Science Foundation of China (#15ZDB153).

References

Andersen, T. G., & Bondarenko, O. (2015). Assessing measures of order flow toxicity and early warning signals for market turbulence. *Review of Finance, 19*, 1–54.
Ang, A., & Timmermann, A. (2012). Regime changes and financial markets. *Annual Review of Financial Economics, 4*, 313–337.
Ausín, M. C, Galeano, P., & Ghosh, P. (2014). A semiparametric Bayesian approach to the analysis of financial time series with applications to value at risk estimation. *European Journal of Operational Research, 232*, 350–358.
Battiston, S., Farmer, J. D., Flache, A., Garlaschelli, D., Haldane, A. G., Heesterbeek, H., et al. (2016). Complexity theory and financial regulation. *Science, 351*, 818–819.
Bennett, C. H., Gács, P., Li, M., Vitányi, P. M. B., & Zurek, W. H. (1998). Information distance. *IEEE Transactions on Information Theory, 44*, 1407–1423.
Brock, W. A., & Carpenter, S. R. (2012). Early warnings of regime shift when the ecosystem structure is unknown. *PLoS One, 7*, e45586.
Cao, L. (1997). Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D, 110*, 43–50.
Carter, K. M., Raich, R., Finn, W. G., & Hero, A. O. (2011). Information-geometric dimensionality reduction. *IEEE Signal Processing Magazine, 28*, 89–99.
Christofides, C., Eicher, T. S., & Papageorgiou, C. (2016). Did established early warning signals predict the 2008 crises? *European Economic Review, 81*, 103–114.
De Angelis, L., & Dias, J. G. (2014). Mining categorical sequences from data using a hybrid clustering method. *European Journal of Operational Research, 234*, 720–730.
Faranda, D., Emanuele Pons, F. M., Giachino, E., Vaienti, S., & Dubrulle, B. (2015). Early warnings indicators of financial crises via auto regressive moving average models. *Communications in Nonlinear Science and Numerical Simulation, 29*, 233–239.
Figueiredo, M., & Jain, A. K. (2002). Unsupervised learning of finite mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24*, 381–396.
Han, L. Y., & Christopher, J. R. (2011). Stable Takens' embeddings for linear dynamical systems. *IEEE Transactions on Signal Processing, 59*, 4781–4794.
He, Q., Liu, Y., Long, Q., & Wang, J. (2012). Time-frequency manifold as a signature for machine health diagnosis. *IEEE Transactions on Instrumentation and Measurement, 61*, 1218–1230.
Holger, B., & Sławomir, S. (2006). The Kullback–Leibler divergence and nonnegative matrices. *IEEE Transactions on Information Theory, 52*, 5539–5545.
Huang, Y., & Kou, G. (2014). A kernel entropy manifold learning approach for financial data analysis. *Decision Support Systems, 64*, 31–42.
Jamshidi, A. A., Kirby, M. J., & Broomhead, D. S. (2011). Geometric manifold learning. *IEEE Signal Processing Magazine, 28*, 69–76.
Jenssen, R. (2010). Kernel entropy component analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*, 847–860.
Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problems of Information Transmission, 1*, 1–7.
Kou, G., & Ergu, D. (2014). Jennifer Shang, Enhancing data consistency in decision matrix: adapting hadamard model to mitigate judgment contradiction. *European Journal of Operational Research, 236*, 261–271.
Kou, G., & Lin, C. (2014). A cosine maximization method for the priority vector derivation in AHP. *European Journal of Operational Research, 235*, 225–232.
Kou, G., Peng, Y., & Wang, G. (2014). Evaluation of clustering algorithms for financial risk analysis using MCDM methods. *Information Sciences, 275*, 1–12.
Lang, M., & Schmidt, P. G. (2016). The early warnings of banking crises: Interaction of broad liquidity and demand deposits. *Journal of International Money and Finance, 61*, 1–29.
Leiva-Murillo, J. M., & Artés-Rodríguez, A. (2012). Information-theoretic linear feature extraction based on kernel density estimators: A review. *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews, 42*(6), 1180–1189.
Lin, T., & Zha, H. (2008). Riemannian manifold learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 30*, 796–809.
Maheu, J., & McCurdy, T. (2009). How useful are historical data for forecasting the long-run equity return distribution? *Journal of Business and Economic Statistics, 27*, 95–112.
Moon, T. K. (1996). The expectation-maximization algorithm. *IEEE Signal Processing Magazine, 13*, 47–59.
Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE, 77*, 257–286.
Renyi, A. (1959). On the dimension and entropy of probability distributions. *Acta Mathematica Academiae Scientiarum Hungarica, 10*, 193–215.
Richard, J. P., Michael, T. J., Andrew, C. L., & Ye, J. (2004). Time series classification using Gaussian mixture models of reconstructed phase spaces. *IEEE Transactions on Knowledge and Engineering, 16*, 779–783.
Roweis, S. T., & Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science, 29*, 2323–2326.
Rydén, T., Terasvirta, T., & Asbrink, S. (1998). Stylized facts of daily return series and the hidden markov model. *Journal of Applied Economics, 13*, 217–244.
Schaller, H., & Van Norden, S. (1997). Regime switching in stock market returns. *Applied Financial Economics, 7*, 177–191.
Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S., Dakos, V., et al. (2009). Early-warning signals for critical transitions. *Nature, 461*, 53–59.
Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W., Dakos, V., et al. (2012). Anticipating critical transitions. *Science, 338*, 344–348.
Seber, G. A. F. (2004). *Multivariate observations*. Hoboken, NJ: Wiley.
Seung, H. S., & Lee, D. D. (2000). The manifold ways of perception. *Science, 290*, 2268–2269.
Sevim, C., Oztekin, A., Bali, O., Gumus, S., & Guresen, E. (2014). Developing an early warning system to predict currency crises. *European Journal of Operational Research, 237*, 1095–1104.
Sornette, D., & Woodard, R. (2010). *Econophysics approaches to large-scale business data and financial crisis*. Tokyo: Springer.
Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., et al. (2012). Detecting causality in complex ecosystems. *Science, 338*, 496–500.
Takens, F. (1981). *Dynamical systems and turbulence*. Berlin: Springer.
Tenenbaum, J. B., Sivlar, V., & Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science, 290*, 2319–2323.
Whitney, H. (1936). Differentiable manifolds. *Annals of Mathematics, 37*, 645–680.
Yang, L. (2008). Alignment of overlapping locally scaled patches for multidimensional scaling and dimensionality reduction. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 30*, 438–450.
Yuliya, D., & Iftekhar, H. (2010). Financial crises and bank failures: A review of prediction methods. *Omega, 38*, 315–324.
Zhang, Z., & Zha, H. (2004). Principal manifolds and nonlinear dimension reduction via local tangent space alignment. *SIAM Journal on Scientific Computing, 26*, 313–338.

Please cite this article as: Y. Huang et al., Nonlinear manifold learning for early warnings in financial markets, European Journal of Operational Research (2016), http://dx.doi.org/10.1016/j.ejor.2016.08.058