A dimension-reduction AFN method for distinguishing chaos from noise

Ang Zhou¹ and Shilian Wang

School of Electronic Science and Engineering, National University of Defense Technology, Changsha, Hunan, 410073, P.R. China

¹Email: angzhou365@126.com

Abstract. Chaotic signal is stochastic in time domain and wide in frequency domain. Distinguishing chaos from noise is a key issue in some application scenarios involved with chaos detection. In this paper, a dimension-reduction method is proposed to distinguish chaos signal from noise. Combining the principle component analysis (PCA) and average false neighbour (AFN), this joint method depicts chaos features from the perspective of both energy and dynamic behaviour in phase space. It has the following advantages: (1) reduces the sensitivity towards noise intensity compared with former algorithms; (2) improves the precision of detection. Several numerical experiments are conducted to verify the availability of this method.

1. Introduction

Chaos is a stochastic-like phenomenon generated from deterministic system and cannot be predicted precisely in long term. It is random in waveform and wide in frequency domain, which is quite similar with white noise. Chaos has gained everlasting attention from researchers in convert communication and new radar system for decades, due to its noise-like feature caused by deterministic uncertainty. Several kinds of chaotic signals have been used in engineering projects. Detecting chaotic signals from noise correctly is necessary for the next signal processing. Therefore, distinguishing chaos from noise is a problem of practical significance.

Generally, distinguishing chaos from noise is conducted by calculating some characteristic quantities which reflect the inner characters of chaotic systems, such as Lyapunov exponent [1, 2] and Kolmogorov entropy [3]. However, calculating these quantities requires a great amount of observed data and thus the detection process is computationally intensive. Besides, it is quite sensitive to noise level and calculation errors cannot be ignored. Grassberger and Procaccia [5] proposed an algorithm (G-P algorithm) to determine the correlation dimension based on phase space reconstruction (PSR), which reflects obvious differences between chaos and noise. Nevertheless, it achieves no better performance than the method of calculating the characteristic quantity. Kennel [6] used the false nearest neighbors (FNN) to find the minimum embedded dimension. Noise is distributed randomly in phase space of different dimensions, which means it will never achieve a proper embedded dimension. On the contrary, chaos signal has a minimum embedded dimension. FNN method is sensitive towards noise so that it has deteriorated performance when the chaotic signal is contaminated by noise. Besides, FNN requires setting threshold artificially which lacks objectivity. Cao [7] enhanced FNN and proposed a method of calculating average false nearest (AFN) which curves the change of average varied quantity in noise and chaotic time series under different reconstruction dimensions. Ramdani
[8] proved the statistic characteristic of AFN in white gaussian noise. The highlight of this method is that it can be conducted with short time series and has strong anti-noise capability. However, this method is still calculated intensively. Aiming to solve this, Wei [9] proposed a new method based on AFN which decreases the scale of calculation. Although this method is efficient for certain chaos systems such as Lorenz, yet its performance deteriorates greatly in other chaos systems (Logistic, Henon, Tent). Hence, this method has little significance in practical use. Some innovative methods that are not based on PSR have been proposed in last decade. Gottwald and Melbourne [10, 11] introduced 0-1 test for chaos and this method needs only observed data. However, it is time consuming and requires abundant data. Theiler [12] proposed the method of surrogate data in testing for nonlinearity in time series. This statistical approach is based on preset null hypothesis and is easily realized by generating surrogate data. But its robustness in contaminated chaotic data has not been verified yet. Broomhead and King [13] introduced Principal Component Analysis (PCA) into PSR of chaos, aiming to decrease the linear dependence between coordinates in setting the embedded dimension. Meanwhile, PCA is also effective in distinguishing chaos from noise. PCA requires only short time series and is computationally efficient. Above all, it is robust in low signal-noise-ratio scenario and is applicable to known chaos system. Nevertheless, PCA is a classical method in detecting signal from noise and reflects no characteristics of chaos.

This paper analyses AFN and PCA based on phase space reconstruction. Combining these two typical chaos detection methods, a practical method is proposed to distinguish chaos from noise. Firstly, the experiment data is dealt with PCA. Then it is reconstructed after being filtered with a proper noise floor. Last, the rebuilt dimension-reduction data is dealt with AFN method. This new method reduces the impact of sensitivity to noise, compared with original AFN method.

The rest of this paper is organized as follows. In section 2, we introduce two former chaos detection methods, AFN and PCA respectively. And we analyze their performance in distinguishing chaos from noise. In section 3, a new joint method is described and analyzed explicitly. Several numerical experiments are demonstrated to prove the validity and robustness of this method. Finally, the conclusion is drawn in section 4.

2. Discussion on former chaos detection methods

2.1. Average false nearest method (AFN)

AFN method was initially proposed to determine a proper embedding dimension in phase space reconstruction. We assume that chaos signal is already sampled in fixed sampling rate. Suppose we have a fragment of discrete chaotic data \( \{x_1, x_2, \ldots, x_N\} \). This scalar time sequence is reconstructed in phase space with the embedding dimension \( m \) and delay time \( \tau \).

Then its trajectory matrix \( Y_{M \times m} \) can be described in equation (1), where \( M = N - (m - 1)\tau \).

\[
Y_{M \times m} = (Y_1, Y_2, \ldots, Y_M)^T
\]

where time-delay vector \( Y_i = (x_i, x_{i+\tau}, \ldots, x_{i+(m-1)\tau})^T, i = 1, 2, \ldots, M \).

Here, a variation \( E(m) \) with the embedding dimension \( m \) is defined in equation (2), where \( n(i,m)(1 \leq n(i,m) \leq N - m\tau) \) is an integer and \( x_{n(i,m)}(m) \) is the nearest neighbor of \( x_i(m) \). The measurement of distance \( || \cdot || \) here is the maximum norm. \( E(m) \) describes the average nearest distance between neighbors in different dimensions:

\[
E(m) = \frac{1}{N - m\tau} \sum_{j=1}^{N-m\tau} |x_{i+m\tau} - x_{n(i,m)+m\tau}|
\]

After calculating \( E(m) \) with different dimension \( m \), a quantity \( \gamma(m) \) is calculated to distinguish chaos from noise. It is defined as \( \gamma(m) = E(m+1) / E(m) \). For noise sequence, it is obviously known that the past value is independent with current value. If the noise model is white gaussian noise.
(WGN), then its mean and variance of \( \gamma(m) \) are easily calculated by statistical analysis tool. According to Ramdani [8], the mean of \( \gamma(m) \) equals 1, and the variance of \( \gamma(m) \) equals 0 for WGN. On the contrary, in the case of determined chaotic data, the value of \( \gamma(m) \) is related closely to embedding dimension \( m \) and has no fixed mean nor variance.

![Figure 1](image-url)

**Figure 1.** (a) \( \gamma(m)-m \) curves for Lorenz sequence and WGN (b) \( \gamma(m)-m \) curves for contaminated Lorenz sequence (SNR=20dB) and WGN.

We utilize a typical chaos system, Lorenz system [14], to produce a fragment of chaotic sequence. Here, we use its \( x \)-component as testing data. The Lorenz system is described in differential functions:

\[
\begin{align*}
\dot{x} &= -\sigma x + \sigma y \\
\dot{y} &= -xz + rx - y \\
\dot{z} &= xy - bz 
\end{align*}
\]

where the chaos dynamical system parameters are set as \( \sigma=10, r=28, b=8/3 \) with initial vector \((x_0, y_0, z_0)=(1.2,1.3,1.6)\). We calculate the \( E(m) \) and \( \gamma(m) \) using 1000 points of data sequence. The time delay \( \tau \) is set as 1. Figure 1(a) shows the changing trend of \( \gamma(m) \) with different dimension for both chaos sequence and WGN. For Lorenz sequence, \( \gamma(m) \) increases while \( m \) increases. And it fluctuates in a small range after reaching a saturation point (where \( m = 2 \)). Figure 1(b) shows that the contaminated Lorenz sequence has the almost identical changing trend as WGN, which proves that this method is sensitive to SNR and is not suitable for the detection of contaminated chaos signal.

2.2. Principal component analysis method (PCA)

The PCA method is based on phase space reconstruction. Unlike the first processing step in AFN, we produce a normalization matrix \( Y_{M \times n} \) with a factor \( M^{-1/2} \) added to the reconstruction matrix:

\[
Y_{M \times n} = M^{-1/2}(Y_1,Y_2,...,Y_M)^T
\]

where the time-delay vector \( Y_i \) has the same definition with that in equation (1). Then we have the covariance matrix \( \mathbf{A} = Y_{M \times n}^T Y_{M \times n} \). The eigenvalues of \( \mathbf{A} \) are easily calculated and are rearranged as \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \). The normalization eigenvalue spectrum is defined as \( \nu_i = \ln(\lambda_i / \sum_i \lambda_i) \). After having the normalized singular spectra plot, the changing trend of eigenvalues can be found out. To analyze the characteristics of the eigenvalue spectrum for contaminated chaotic signal, we firstly introduce the model of contaminated chaotic signal in a matrix form. Assume the chaotic sequence is \( x(t) = \{ x_1, x_2, ..., x_N \} \), and WGN (with its variance \( \sigma_n^2 \) related to SNR) is \( n(t) = \{ n_1, n_2, ..., n_M \} \). Then we have the practical observed data \( z(t) = x(t) + n(t) \). Its normalization matrix \( Y_{M \times n} \) is
\[ \mathbf{Y}_{M \times m} = M^{-1/2} (Y_1, Y_2, \ldots, Y_M)^T \]  

where time-delay vector \( Y_i = (x_{i+\tau}, x_{i+\tau} + n_{i+\tau}, \ldots, x_{i+(m-1)\tau} + n_{i+(m-1)\tau})^T, \) \( i = 1, 2, \ldots, M \).

Then we recalculate \( \mathbf{A} = \mathbf{Y}^T \mathbf{Y} = M^{-1} \sum_{i=1}^{M} Y_i Y_i^T \), and have its element \([\mathbf{A}_{ij}]:\)

\[ [\mathbf{A}_{ij}] = M^{-1} \sum_{l=1}^{M} z_{l(i+\tau)} z_{l(j+\tau)}, \quad l = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, m \]  

Consider that chaos signal is independent with WGN, then

\[ M^{-1} \sum_{j=1}^{M} z_{l(i+\tau)} z_{l(j+\tau)} = M^{-1} \sum_{i=1}^{M} \left( x_{i(i+\tau)} + n_{i(i+\tau)} \right) \left( x_{i(j+\tau)} + n_{i(j+\tau)} \right) \approx M^{-1} \sum_{i=1}^{M} \left( x_{i(i+\tau)} x_{i(j+\tau)} \right) + M^{-1} \sum_{i=1}^{M} \left( n_{i(i+\tau)} n_{i(j+\tau)} \right) \]  

\[ = M^{-1} \sum_{i=1}^{M} \left( x_{i(i+\tau)} x_{i(j+\tau)} \right) + \sigma_n^2 \delta(i(l-1)\tau, j(j-1)\tau) \]  

\[ \text{Figure 2. Normalized eigenvalue spectrums with different embedding dimension } m: \text{ (a) original Lorenz sequence; (b) contaminated Lorenz sequence (SNR=20dB); (c) contaminated Lorenz sequence (SNR=5dB); (d) WGN (variance } \sigma^2 = 0.2 \text{).} \]

where \( \delta(i,j) \) is Dirac function, and \( \sigma_n^2 \) is the covariance of WGN. Hence, the covariance matrix of \( \mathbf{Y}_{M \times m} \) can be rewritten as \( \mathbf{A} = \tilde{\mathbf{A}} + \sigma_n^2 \mathbf{E} \). Here, \( \tilde{\mathbf{A}} \) is the covariance matrix of \( \mathbf{Y}_{M \times m} \) for original chaotic signal and \( \mathbf{E} \) is an \( m \times m \) identity matrix. In terms of the additivity of eigenvalues, we have the
eigenvalues of $A$ as $\lambda_i = \tilde{\lambda}_i + \sigma^2_i (i = 1, 2, ..., m)$, where $\tilde{\lambda}_i$ is the eigenvalue of $\tilde{A}$. All the eigenvalues have a shift in contaminated chaotic signal. Note that $\tilde{A}$ has non-zero eigenvalue. Hence, its eigenvalue spectrum has limited number of points. On the contrary, eigenvalue spectrum of $A$ shows an approaching line that represents for several equal eigenvalues. This line is called “noise floor”. Then the detection can be conducted in terms of this obvious difference. Figure 2(a) shows a decreasing linear eigenvalue spectrum of original Lorenz sequence ($x$ component) with different embedding dimensions. The linearity of spectrum reflects the relativity between delayed vectors in phase space. Figure 2(b) and (c) depict the eigenvalue spectrums of contaminated Lorenz sequence (with SNR = 20dB and SNR = 5dB respectively). We can see a section of approximately horizontal line called “noise floor”. Because the noise intensity of SNR=5dB is larger than that of SNR=20dB, so the noise floor in (c) is higher than that in (b). Figure 2(d) depicts the eigenvalue spectrum of WGN. It shows a series of horizontal lines with different dimensions. As we discussed before, the covariance matrix of WGN has only the $\sigma^2_i E$ part which causes the fixed floor in spectrum. With this obvious difference, decision can be made by calculating the variance of eigenvalues to distinguish chaos from noise.

3. Dimension-reduction AFN method

From the former discussion of AFN and PCA method, we know that AFN is quite sensitive to noise while PCA is not. Note that PCA is a useful way to divide the contaminated chaotic sequence into two parts and produce a new trajectory matrix for dimension-reduced phase space, which can potentially counterbalance some bad effects caused by noise. On the other hand, PCA method is insufficient in reflecting the inner dynamic characteristics in chaotic signal. Hence, combining these two methods together will enhance the detection effect. Here, we proposed the dimension-reduction AFN method.

Before being analyzed by AFN method, original observed time sequence needs a process of dimension-reduction. Note that “noise floor” is clear and easy to find in eigenvalue spectrum. Dimension-reduction can be conducted by reconstructing the trajectory matrix by selecting the part that is above the noise floor. Suppose we have a $m \times m$ matrix $P^{(k)} = \delta_{ij} \delta_{jk} E$ with the $k$th row of its diagonal elements equals 1. Then the projection matrix is divided into two parts:

$$P = \sum P^{(k)}, \quad \sigma_1 > \sigma_2$$
$$Q = \sum P^{(k)}, \quad \sigma_1 = \sigma_2$$

According to Broomhead [13], the trajectory matrix of a time sequence can be dealt with singular value decomposition (SVD) as $Y = \Sigma \Sigma C^T$, where $C$ is the eigenvectors of $A$, $S$ is a orthogonal vector set related to $C$ and $\Sigma^2 = \text{diag} (\lambda_1, \lambda_2, ..., \lambda_m)$ is the eigen orthogonal matrix of $A = Y^{T}Y$. Then the SVD equation can be decomposed as following:

$$Y = \bar{Y} + \Delta Y = \Sigma \Sigma C^T + \Sigma \Sigma C^T$$

where $\bar{Y} = \Sigma \Sigma C^T$ is the trajectory matrix of chaos controlled part and $\Delta Y = \Sigma \Sigma C^T$ is the trajectory matrix of noise effected part. We select the first part and have the reconstructed trajectory matrix as

$$Y = \Sigma \Sigma C^T = (\sum Yc_i) c_i^T, \quad \sigma_1 > \sigma_2$$

By reconstructing this dimension-reduction trajectory matrix, this method can be described as follows:

- Reconstruct the phase space of original data(with embedding dimension $m$ and time delay $\tau$);
- Conduct the eigenvalue spectrum analysis and resort the eigenvalues as $\{\lambda_i | \lambda_1 \geq \lambda_2 \geq ... \geq \lambda_m\}$;
• Select \( m' \) eigenvalues from \( \{ \lambda_i \} (m' < m) \) with \( \sum_{i=1}^{m'} \lambda_i > 0.9 \sum_{i=1}^{m} \lambda_i \);

• Calculate the dimension-reduced phase space \( A' = [\sum_{i=1}^{m'} A_{ij}C_i^T] \);

• Conduct AFN method on newly reconstructed \( A' \) and draw \( \gamma(m) - m \) curves;

• Calculate \( |E[\gamma(m)] - 1| \) and \( D[\gamma(m)] \), then conduct the detection.

In the implementation of this method, we have a numerical test and take Lorenz sequence as the experiment object with the same parameter and initials in section 2. Different SNR is set to add corresponding intensive AWGN to chaotic signal. Then we utilize both original AFN method and dimension-reduction AFN method and calculate the \( \gamma(m) \). The results are shown in figure 3(a) and figure 3(b) respectively. Then we calculate \( |E[\gamma(m)] - 1| \) and probability of recognition \( P_{rec} \) with 100 times Monte-Carlo tests. The final results are listed in table 1. In figure 3(a), the contaminated Lorenz sequence shows no obvious fluctuation, making it quite difficult to recognize it from noise. Comparing figure 3(a) with figure 3(b), we see that the \( \gamma(m) \) of contaminated Lorenz sequence fluctuates much more intensively around 1 after dimension-reduction. Besides, noise stays unchanged after PCA dimension-reduction because no noise floor can be found to divide the trajectory matrix. From table 1 we observe that \( |E[\gamma(m)] - 1| \) increases after dimension-reduction which enhances the recognition rate. The recognition rate has been enhanced from 77\% to 91\% in the case of SNR=5dB. On the contrary, WGN shows no obvious difference after dimension-reduction, proving that PCA method has no positive effect on noise. Hence, dimension-reduction AFN method counterbalances the impact caused by noise and improves the precision of detection.

**Figure 3.** (a) \( \gamma(m) - m \) curves for Lorenz sequence (without dimension-reduction); (b) \( \gamma(m) - m \) curves for Lorenz sequence (with dimension-reduction).

**Table 1.** \( |E[\gamma(m)] - 1| \) and recognition rate \( P_{rec} \) of Lorenz sequence with different SNR and WGN (with/without PCA dimension-reduction).

| Type  | SNR (dB) | without PCA | with PCA |
|-------|----------|-------------|----------|
| Lorenz| 40       | 0.0102      | 0.2302   |
|       | 20       | 0.0135      | 0.0671   |
|       | 10       | 0.0142      | 0.0206   |
|       | 5        | 0.0056      | 0.0182   |
| WGN   | -        | 0.0031      | 0.0036   |
Figure 4. (a) $\gamma(m) - m$ curves for Henon map sequence with SNR=10dB (b) $\gamma(m) - m$ curves for Chen’s chaotic system (x-component) sequence with SNR=10dB.

Figure 4 shows that this method is also suitable for Henon map and Chen’s system. In the traditional AFN method, the contaminated chaotic signal shows an almost horizontal line near 1. After dimension-reduction with PCA method, the contaminated chaotic signal fluctuates much more intensively, making the average of $\gamma(m) - 1$ farther from 0. We also test other chaos systems, such as Logistic, Tent and Chua, and find that this method can enhance the precision of detection for these chaotic signals.

4. Conclusions

This paper proposed a dimension-reduction AFN method based on PCA to distinguish chaotic signal from noise. This method is based on an assumption that contaminated chaotic signal can be divided into two parts by selecting a proper noise floor. After dimension-reduction, the original data eliminates the irrelative part of AWGN. The processing can be described as following: observed data is firstly dealt with PCA and is reconstructed in phase space, then the rebuilt data is analysed by AFN, and finally the decision is made by calculating the variation $\gamma(m)$. We set different SNR to chaotic signal and conduct a series of numerical tests. The results show that this method decreases the sensitivity to noise compared with former AFN method.

Besides, the dimension-reduction AFN method does not require a long data sequence so that even a short chaotic sequence can be recognized correctly. However, this method is calculated intensively. This is because the nearest point of a high-dimension point needs to be searched out in each embedding dimension. It is necessary to optimize its time efficiency in the future study.

References

[1] Wolf A, Swift J B and Swinney H L 1985 *Physica D Nonlinear Phenomena* **16** 285-317
[2] Sano M and Sawada Y 1985 *Phys. rev. lett.* **55** 1082-5
[3] Eckmann J P and Ruelle D 1985 *Review of Modern Physics* **57** 617-56
[4] Wang J and Ma Q 2008 *Chinese Physics* B **17** 4424
[5] Grassberger P and Procaccia I 1983 *Physical Review Letters* **50** 346-9
[6] Kennel M B, Brown R and Abarbanel H D 1992 *Physical Review A* **45** 3403-11
[7] Cao L 1997 *Physica D Nonlinear Phenomena* **110** 43-50
[8] Ramdani S, Bouchara F and Casties J F 2007 *Physical Review E Statistical Nonlinear & Soft Matter Physics* **76** 036204
[9] Wei H D, Li L P and Guo J X 2010 *Chinese Physics* B **19** 98-103
[10] Gottwald G A and Melbourne I 2005 *Physica D Nonlinear Phenomena* **212** 100-10
[11] Gottwald G A and Melbourne I 2009 *Siam Journal on Applied Dynamical Systems* **8** 129-45
[12] Thiller J et al. 1992 *Physica D Nonlinear Phenomena* **58** 77-94
[13] Broomhead D S and King G P 1986 *Physica D Nonlinear Phenomena* **1986** **20** 217-36
[14] Lorenz E N 1963 *Journal of the atmospheric sciences* **20** 130-41