NLID algorithm

1. One-dimensional time series → Multidimensional phase space

For a time series X[i], by setting a time delay and an embedding dimension, it can be transformed into multi-dimensional phase space vectors. This transformation is based on the theory of phase space reconstruction, and the corresponding formula is as follows:

$$X[i] = (X[i], X[i + \tau], X[i + 2\tau], \dots, , X[i + (m - 1)\tau] \tag{1}$$

Where:

- X[i] is the current value of the time series.
- X[i+τ] is the value obtained by shifting the series forward by τ time units.
- X[i+2τ] is the value obtained by shifting the series forward by 2τ time units.

For example, in Figure 1, when τ=5 and m=3, X[i]=0.75 after time delays of τ and 2τ, we obtain the corresponding points X[i+τ] = -0.25 and X[i+2τ] = -0.75. These three values form a three-dimensional coordinate point X = (X = 0.75, Y = -0.25, Z = -0.75). As i increases, a series of multidimensional points are formed, and the connection of these points constitutes the phase space trajectory, used to describe the structure and dynamic behavior of time series data.
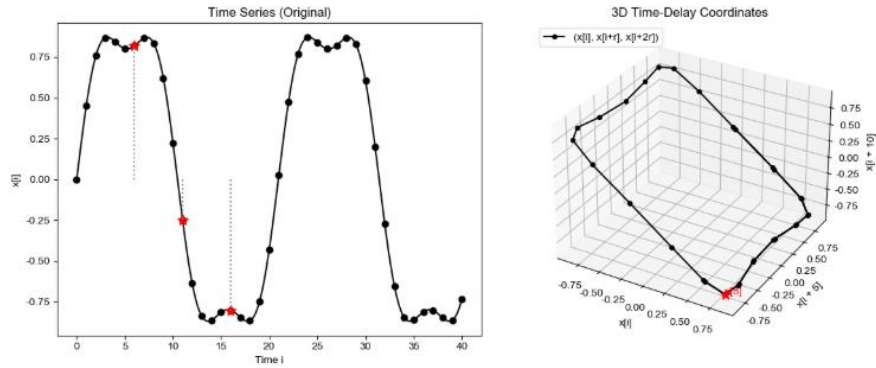

Figure 1 Schematic diagram of converting a one-dimensional time series to a multidimensional phase space

Constructing the Recurrence Plot

Definition:

$$R(i,j) = \Theta\big(d_{Th} - ||X_i - X_j||\big) \tag{2}$$

If the distance between Xi and Xj   is smaller than the threshold, then R(i,j)=1; otherwise, R(i,j)=0.

Example:

- Suppose $|| X1{-}X2 || < d_{Th}$, thenR(1,2)=1.
- Suppose $|| X1{-}X3 || > d_{Th}$ then R(1,3)=0.

As a result, we obtain an N×N binary matrix that characterizes the recurrence structure of the system. The threshold d_{Th} was determined as 10% of the maximum phase space distance, which is a commonly adopted criterion to ensure a balanced recurrence density and to avoid overly sparse or overly dense recurrence plots.

(一) $R(i,j)$
- Elements of the recurrence matrix
- Indicate whether a recurrence event occurs between point i and point j
- Values are either 0 or 1

(二) $\Theta()$
- Heaviside step function
- Definition:

$$\Theta(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

- Here, it is used to determine whether the distance is less than the threshold.

(三) d_{Th}

- Threshold
- The dividing line between "similar" and "dissimilar"
- In reconstruction applications, it is often set to a certain percentage (e.g., 10%) of the maximum distance.

(四) $||X_i - X_j||$

- Euclidean distance
- Calculate the distance between the i-th point and the j-th point in the feature space:

$$||X_i - X_j|| = \sqrt{\sum_{k=1}^{m}(X_{i,k} - X_{j,k})^2}$$

where m is the embedding dimension

Calculated NLID factor

The nonlinear inter-driven (NLID) factor at time point j is defined as:

$$\Gamma(Y \mid X)(j) = \frac{\sum_{i=1}^{N} R_X(i,j)\cdot R_Y(i,j)}{\sum_{i=1}^{N} R_X(i,j)} \tag{3}$$

where $R_X(i,j)$ and $R_Y(i,j)$ enote the recurrence matrices of signals X and Y, respectively. The numerator represents the number of coincident recurrence points in column j of both matrices, i.e., the extent to which recurrences in X at time j are also observed in Y. The denominator corresponds to the total recurrence points in column j of X, thereby serving as a normalization factor. As a result, $\Gamma(Y \mid X)(j)$ quantifies the proportion of recurrences in X that are simultaneously present in Y, given time point j.

Numerator:

$$\sum_{i=1}^{N} R_X(i,j) \cdot R_Y(i,j)$$

When $R_X(i,j)=1$ and $R_Y(i,j)=1$, it indicates that both X and Y exhibit recurrence at the same time point. This corresponds to detecting the event of overlap.

Denominator:

$$\sum_{i=1}^{N} R_X(i,j)$$

This represents the total number of recurrence events of X at time point j. It serves as a normalization factor to ensure that the result lies between 0 and 1.

Global average

To quantify the overall nonlinear interdependence between two signals across the entire time span, the local NLID factor $\Gamma(Y \mid X)(J)$ is averaged over all time indices j:

$$\Gamma(Y \mid X) = \frac{1}{N}\sum_{j=1}^{N} \Gamma(Y \mid X)(j) \tag{4}$$

Here, N denotes the total number of embedded states in the reconstructed phase space. The measure $\Gamma(Y \mid X)$ therefore represents the global driving strength of system X on system Y, expressed as the mean proportion of recurrence events in X that are simultaneously observed in Y.

- $\Gamma(Y \mid X)(j)$ =1: The behavior of the two systems is completely identical.
- $\Gamma(Y \mid X)(j)$ =0: The two systems are completely independent, with no correlation or mutual influence.
- $\Gamma(Y \mid X)(j)$ is a value between 0 and 1: System X has a partial influence on system Y, and there is a certain degree of nonlinear interaction.