Choice of Measurement for Phase-Space Reconstruction: Decision Based on False Nearest Neighbors Method

K. Mezeiová, A. Krakovská

Institute of Measurement Science SAS, Bratislava, Slovak Republic
Email: Kristina.Mezeiova@savba.sk

Abstract. *False nearest neighbors (FNN) method is examinated with respect to equivariance of individual observables. The aim is to reveal the most appropriate observable for phase space reconstruction. Results calculated for benchmark systems are compared with symbolic observability degrees. The FNN method resulted in different values of embedding dimensions when calculated for various observables of the same dynamical system. The results roughly corresponded to the symbolic observability degrees; however, in some details disagreed with them.*

Keywords: False Nearest Neighbors Method, Observability, Equivariance of Observables

1. Introduction

According to the Takens' theorem [1] a sufficient embedding dimension to reconstruct a phase space topologically equivalent to the original space is 2m+1, where m is the dimension of the attractor of the system, and the embedding can be done by delayed coordinates of a single observable. Takens' theorem states the sufficient condition of the embedding dimension, however, for many systems the embedding can be reached in fewer dimensions. The second problem not solved is the fact, that if we have more observables from the same system, they are not equivalent with respect to the phase space reconstruction – for one observable the embedding can be reached in less dimensions than for the another one.

Letellier et. al. [2] extended the theory of observability to nonlinear systems and later proposed the symbolic observability coefficients η, which quantify the level to which the system is observable based on the measurement of some variable. Details of the computation are explained in [2].

What to do in practice when the equations are unknown? The aim of this paper was to compare the method of False nearest neighbors, which is traditionally used to estimate the embedding dimension, with the results of the symbolic observability coefficients.

2. Methods

False nearest neighbors algorithms

False nearest neighbors method [3] is an iterative algorithm which estimates the embedding dimension of the system. In its n^{th} step the phase space is reconstructed into n dimensions by taking the time delayed coordinates of the measured time series: $y_i = x_i, x_{i-\tau}, \ldots, x_{i-(n-1)\tau}$, where x is the measured time series and τ is the time delay. In this paper the time delay was set as the first minimum of the mutual information function [4]. For each point **y(n)** the distance R(n) to its k^{th} nearest neighbor $\mathbf{y}^{\mathbf{k}}\mathbf{(n)}$ is calculated. If $\mathbf{y}^{\mathbf{k}}\mathbf{(n)}$ is close to **y(n)** not due to the dynamics of the system, but due to a projection of the trajectory from the natural phase space to the lower dimensional space, in the n+1 dimensions the distance R(n+1) between these two points becomes large and $\mathbf{y}^{\mathbf{k}}\mathbf{(n)}$ is called a false nearest neighbor. Criterion for a neighbor to be false can be evaluated by the formula: $\sqrt{\frac{R^2(r,n+1)-R^2(r,n)}{R^2(r,n)}} > R_{tol}$, where R_{tol} is some threshold.

There is also a second criterion for the nearest neighbor to be false – if it is not a close point to **y(n)**. E. g. if the distance R(n) is half the size of the attractor, then the iterated distance R(n+1) can be maximally 2R(n) if **y(n+1)** and $\mathbf{y}^{\mathbf{k}}\mathbf{(n)}$ are located at the extremes of the attractor. Such points are considered to be false neighbors and the second formula for the point to be a false neighbor is: $\frac{R(n+1)}{R_A} > A_{tol},$ where R_A is some attractor's size and A_{tol} is the second threshold. Here R_A was set to the standard deviation of x.

In each dimension the percentage of the false nearest neighbors is calculated and algorithm terminates when the percentage drops to zero. In this paper the thresholds were adjusted to R_{tol} = 15 and A_{tol} = 2 and only the first nearest neighbor was taken into account.

Symbolic observability degrees

The symbolic observability degrees η calculated at the dimensions of the dynamical systems were taken from [2].

Systems

Following systems were integrated by the means of the 4th order Runge-Kutta formula with integration step 0.01:

Rössler system:

$$\dot{x} = -y - z, \qquad \dot{y} = x + ay, \qquad \dot{z} = b + z(x - c),$$

where [a, b, c] are bifurcation parameters; in this study the values were [0.398, 2, 4]; initial conditions were [0, 0, 0.4].

Lorenz system:

$$\dot{x} = \sigma(y - x), \qquad \dot{y} = Rx - y - xz, \qquad \dot{z} = -bz + xz,$$

with the parameters [σ, R, b] = [10, 28, 8/3] and the initial conditions [0.3, 0.3, 0.3].

Sprott F system:

$$\dot{x} = y + z, \qquad \dot{y} = -x + ay, \quad \dot{z} = -bz + x^2,$$

with the parameters [a, b] = [0.5, 1] and the initial conditions [0.05, 0.05, 0.05].

Hyperchaotic Rössler system:

$$\dot{x} = -y - z, \qquad \dot{y} = x + ay - w, \qquad \dot{z} = b + xz, \qquad \dot{w} = -cz + dw,$$

with parameters [a, b, c, d] = [0.25, 3, 0.5, 0.05] and the initial conditions [-10, -6, 0, 10.1].

3. Results

For the selected dynamical systems the false nearest neighbors method was computed with the time delay set to the minimum of the mutual information function (see Fig. 1). Results and parameters of the calculations of the FNN method as well as the values of the symbolic observability degrees η can be found in Table 1.

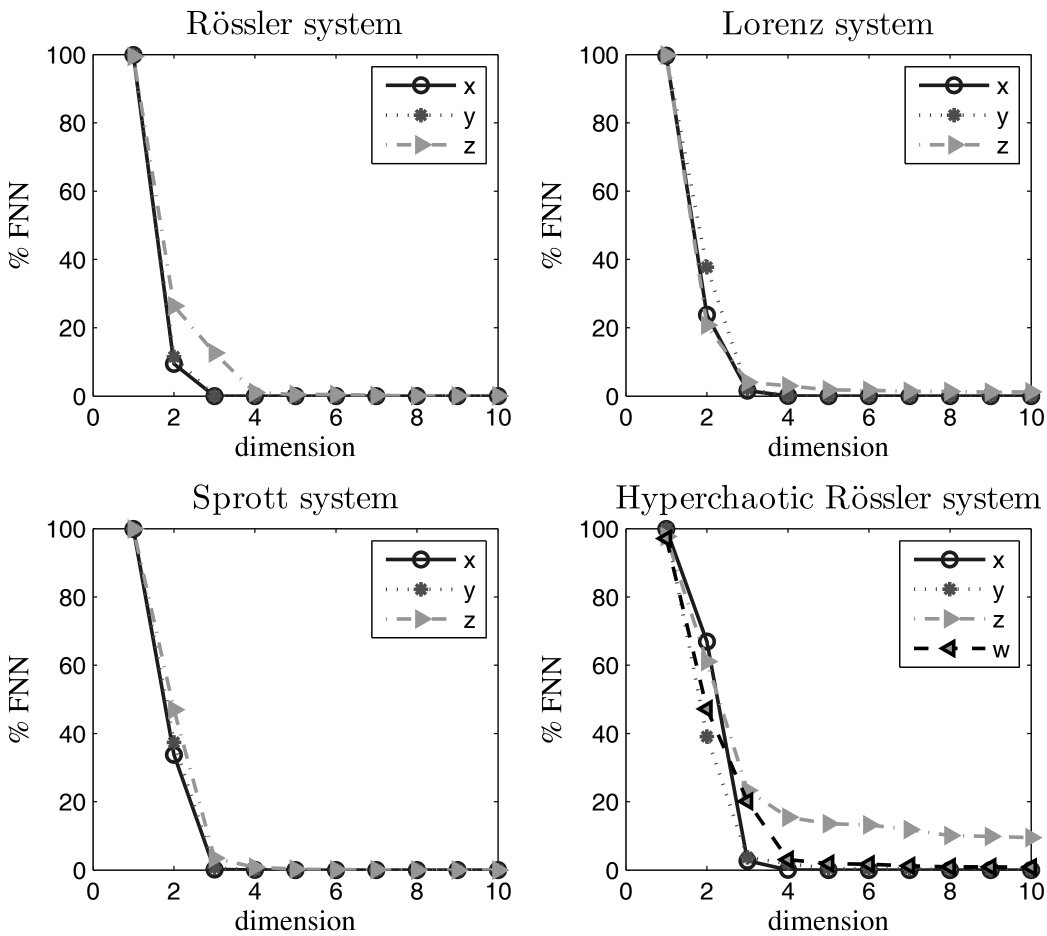

Fig. 1. False nearest neighbors method calculated with the parameters R_{tol} =15 and A_{tol} =2.

The FNN method can be evaluated in two ways. The result is either the embedding dimension – the dimension at which the percentage of false neighbors approaches zero - or directly the percentage of false neighbors in the dimension, at which also the symbolic observability degrees were calculated.

Both ways implied the same consequences: the FNN method resulted in different values of the embedding dimensions when calculated for various observables from the same dynamical system. The worst observable (with the lowest η) resulted with the highest embedding dimension as well as achieved the highest percentage of false neigbors at the dimension of the dynamical system. However, some details were not in according with the observability coefficients. The FNN algorithm did not distinguish variables with slightly different observability degrees or even resulted with higher embedding dimension for slightly more observable variable.

E.g. for the Rössler system the η values can be sorted from the highest value to the lowest one as $\eta_y > \eta_x > \eta_z$; Rössler system with the value η_y =1 is observable in 3 dimensions. The results of FNN for the Rössler system implied that the z observable appeared to be the less appropriate observable for the reconstruction of the phase space with the embedding dimension 5. However, there was no difference between variables x and y, which both resulted in the embedding dimension 3. Similar findings can be written about all benchmark systems.

Table 1. Parameters and results of the FNN method and comparison with the symbolic observability degrees. Min MI – index, at which the first minimum of the mutual information function appeared; d(FNN ~ 0) – dimension, at which the % of false nearest neighbors approached zero (decreased below 1%, resp.); η – symbolic observability degree; %FNN(d =d_{DS}) - % of FNN at the dimension of the original dynamical system

* % FNN did not decrease below 1 % even in the highest dimension calculated

System	Observable	min MI	d(FNN ~ 0)	η	%FNN(d =d_{DS}) [%]
Rössler	x	160	3	0.88	0
	y	171	3	1	0
	z	159	5	0.44	12.62
Lorenz	x	24	4	0.89	1.5571
	y	24	4	0.46	2.0185
	z	19	>10*	0.35	4.0493
Sprott	x	187	3	1	0.1317
	y	195	3	1	0.6176
	z	229	4	0.44	3.4729
Hyperchaotic Rössler	x	181	4	0.85	0.03
	y	168	5	0.92	1.89
	z	178	>10*	0.56	15.43
	w	203	8	0.69	3.02

4. Discussion and Conclusion

The results of the FNN method roughly corresponded to the values of the symbolic observability degrees; however, in some details disagreed with them. The FNN algorithm is dependent on the value of the time delay and on the parameter $R_{tol;}$ the results can differ with various values of these two parameters.

Although the FNN method mirrors to some extent various ability of the variables to reconstruct the phase space, it is not ideal for the choice of the measurement. As the next plan other measures will be examined, e.g. the predictability of the time series based on the reconstructed phase space from different observables.

Acknowledgements

This work was supported by Slovak Grant Agency for Science (grant No 2/0019/10).

References

[1] Takens F. Detecting strange attractors in fluid turbulence. In Dynamical systems and turbulence. Rand D, Young LS, editors. Springer, Berlin, 1981.

[2] Letellier C, Aguirre LA. Symbolic observability coefficients for univariate and multivariate analysis. *Physical Review E*, 79: 066210, 2009.

[3] Kennel MB, Brown R, Abarbanel HDI. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A*, 45(6): 3403-3411, 1992.

[4] Fraser AM, Swinney HL. Independent coordinates for strange attractors from mutual information. *Physical Review A*, 33(2): 1134-1140, 1986.