Choice of Measurement for Phase-Space Analysis:
Review of the Actual Findings

A. Krakovská, K. Mezeiová

Institute of Measurement Science, Slovak Academy of Sciences,
Bratislava, Slovak Republic
Email: krakovska@savba.sk

Abstract. *If data are generated by a d-dimensional system, but only one observable is known, Takens' theorem guarantees that reconstruction diffeomorphic to the original dynamics can be build from the single time series in (2d+1)-dimensional phase space. However, some recent results show that, under certain conditions reconstruction in lower dimension is possible, whereby the choice of the variable used for the embedding may be critical. A short review of the latest results and methods regarding transition from single time series to multi-dimensional phase portrait is given. The choice of the lowest possible dimension for the manifold reconstruction and the choice of the variable with the highest observability index are discussed.*

Keywords: Observability, Takens' theorem, Rössler, Lorenz

1. Introduction

Measured data have often a form of scalar time series - sequence of numbers, typically spaced at uniform time intervals. There are two main approaches to time series analysis. The first one views time series as a manifestation of a stochastic process, and works with the traditional statistical tools. The second - younger - approach sees time series as produced by a deterministic dynamical system. The second type of data handling, often called non-linear, originates in physics and the theory of nonlinear dynamical systems. It is not older than 25-30 years and initially its development was quite independent from the established statistical methodology. However, the two schools are still less and less separated. Statisticians discover the power of the state-space idea, while the "dynamicists" realized that taking advantage of rich resources of statistics accelerates further progress in their field.

In 1981 Takens mathematically established that, if we know a single variable of a dynamical system with sufficient accuracy, then it is possible to reconstruct a state portrait, topologically equivalent to the attractor of the original system [1]. Assume that the dynamics has an attractor of box-counting dimension d. According the Takens' delay theorem, the attractor can be embedded (by means of diffeomorphism - one to one, everywhere differentiable with a differentiable inverse) in D-dimensional Euclidean space whereby $D \geq 2d+1$. The theorem has a powerful implication: under certain conditions one single observable is enough for reproducing and studying of complex multi-dimensional underlying dynamics.

Diffeomorphism between the original manifold and its reconstructed image preserves relevant geometrical and dynamical invariants like the dimension of the attractor or the Lyapunov exponents (measure of the sensitivity to initial conditions). It also holds that trajectories which are close in the original state space are also close in the embedding space. This leads to a successful technique of noise reduction and prediction based on the so called method of analogs: in order to predict the next step of the series, take the last point in the reconstructed space, find the nearest point from the past with a known successor, and assume that the current point will make the analogous move.

Experimental researchers unexpectedly got a valuable tool to extract black-box models for complicated systems directly from one or a few measurements. Takens' theorem gave rise to thousands of interesting publications since its statement.

However, some new results suggest that, the investigation of time series based on the multi-dimensional reconstruction can be even more enhanced.

These results and their perspectives will be introduced in the present paper.

2. Review of Methods and Results

The lowest possible dimension for multi-dimensional reconstruction

Let us remind that, if the data are generated by a *d*-dimensional system, Takens' theorem guarantees that *(2d+1)*-dimensional embedding is equivalent to the original dynamics.

While this is satisfactory for theoretical work, in practice an embedding of lowest possible dimension is preferred, ideally a *d*-dimensional embedding. But can we use lower than *(2d+1)*-dimensional space? Or we may ask: is the condition of *(2d+1)*-dimensional space sufficient or necessary? As far as almost 30 years after appearance of the Takens' theorem this question has not been theoretically treated. But lately, Cross and Gilmore contributed to the issue, when they analyzed differential mappings of the rotationally equivariant Lorenz dynamical system in some detail [2]. They showed that, while the differential reconstruction based on the *x* coordinate is an embedding of the attractor in three dimensions, it does not yield an embedding of the entire manifold. The projection of the manifold into R^3 possesses singularities. However it is possible to embed the manifold into a 3-dimensional twisted submanifold of R^4. Then not only diffeomorphism invariants (fractal dimensions, Lyapunov exponents, and so on) but also information about the mechanism responsible for generating chaotic behaviour is preserved. The authors say that the failure to achieve an embedding in R^3 is related to the different symmetry properties of original and the reconstructed attractor. They showed that the two systems are actually equivalent under topological deformation (isotopic) in R^4. Nonisotopic embeddings provide distinct or inequivalent representations of an attractor, as one may not be deformed into another without selfintersection. In a sufficiently high dimension, not greater than *2d+1* for *d*-dimensional systems, and 5 for 3-dimensional dynamical systems, all embeddings are equivalent [3]. However, so far little is known about lower than *(2d+1)*-dimensional embeddings of dynamical systems with *d>3*.

Choice of variable for reconstruction: univariate case

Theoretically, the variable used for the reconstruction of the attractor can be chosen arbitrarily. In practice, however, some variables seem to be more convenient for analysis than others. For instance, it is much easier to obtain a global model from variable *y* of the Rössler system than from variable *z*. When we are facing an unknown system with single known time series, it would be useful to be able to estimate if it is a good variable or not. In other words: we are looking for some index, usually called „observability" that enables ranking of the observables according to their effectiveness in the reconstruction process.

The concept of observability in control theory is standard and well defined. However, it is a type of "yes or no" measure, that is, the system is evaluated as either observable or not for a given output data series. If a system is observable, it is possible to determine the behaviour of the entire system. If it is not observable, then the output data disallow to estimate (and control) the states of the system completely. To check if a linear system with *n* states is observable, the rank of the so-called observability matrix is calculated. If it is equal to *n*, then the *n* rows are linearly independent, each of the *n* states of the system is given through linear combinations of the output variables and the system is observable.

As concerns nonlinear systems, the notion of observability is not firmly established yet. Since 1998, Letellier et al. have introduced several measures that rank the variables of the system according their observability - assuming that the reconstruction space has the same dimension as the original one [4], [5], [6].

In [5] the authors proposed a definition of observability for nonlinear systems, related to the Jacobian matrix of the coordinate transformation between the original phase space and the differential embedding induced by the given variable. The system is fully observable when the determinant of the Jacobian never vanishes.

In [6] the so-called symbolic observability coefficient has been introduced. The computation of symbolic observability coefficients is based on the so-called fluency matrix of the system, which emphasizes constant and nonconstant elements of the Jacobian matrix, corresponding to linear and nonlinear terms in the vector field of the system. The symbolic observability coefficients are greater than one when the dimension of the reconstructed state space is too large. For some systems the symbolic observability indicates the sufficient dimension to be smaller than provided by the Takens' criterion.

Disadvantage of the above indices is that their estimation is limited to the cases in which the equations of the system are known.

In [7] a procedure is put forward by which it is possible to compare two observables of the same system without the need of the system equations. The proposed time-series approach is based on a recently defined omnidirectional nonlinear correlation functions and it agrees very well with the older indices with respect to observability order of benchmark systems variables.

For example, for 3-dimensional chaotic Rössler system the results were in total agreement, indicating that y variable is the best and z is the worst choice for the reconstruction.
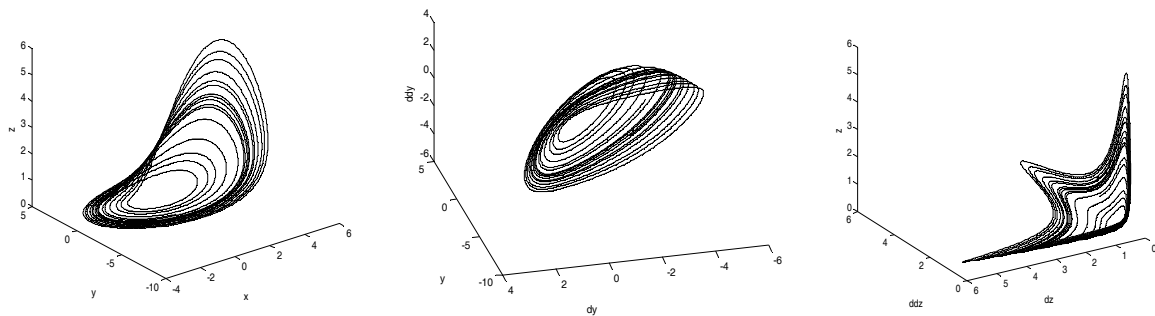

Fig. 1. Attractor of the Rössler system. From right to left: original attractor, reconstruction from y variable and reconstruction from z variable.

Choice of variable for reconstruction: multivariate case

In practice, sometimes more than one physical variable is recorded simultaneously. For example, physiological or economic data are often multivariate.

According to Takens' theorem, one variable suffices for reconstructing a space equivalent to the original phase space - multivariate time series are not required. However, in real world systems, in particular when the dimension of the original phase space is large, there may be advantageous to use more of the measurable variables. As an artificial example Lorenz system can be mentioned. If one variable is used for 3-dimensional reconstruction, the system is not observable. But the combination of x-variable, delayed x-variable and z-variable leads to system that is topologically equivalent to the original [6].

It might seem to be evident that use of more different observables is always preferable to univariate embedding. But it is not true in general. Univariate embeddings are often evenly successful as the best combinations of several measurements.

Moreover, looking for the right multivariate combination is somewhat more complicated than use of a single series, mainly because there is a large choice of possible embeddings. One of a very few tests for optimal choices of observables from a multivariate set is based on

eliminating linear dependence using singular value decomposition. Some observability based test would be very helpful in multivariate analysis.

3. Discussion

The concepts discussed in this paper are of great relevance in phase-space reconstruction problems. We summarized the latest results evidencing that variables of a dynamical system are not equally effective in reconstructing the dynamics from a scalar time series.

On the present we have observability indices available for evaluating the variables in the cases when the system equations are known. Let us use this knowledge as a benchmark and try to find a method how to identify the "good" and "bad" observables in situations, when we are facing an unknown system with one or more scalar time series recorded.

As a first step we can return to the methods that have been used for the estimation of the embedding dimension during past 25-30 years. They are based on step by step expansion of the reconstruction space with simultaneous following of some proper invariants (correlation dimension, largest Lyapunov exponent, predictability, percentage of false nearest neighbours, and so on), which are expected to stay constant after reaching the sufficient embedding dimension. Now, having the benchmark systems of known observability, we can evaluate the worthiness of these methods with accuracy that was not possible before.

Let us begin with false nearest neighbours method (FNN), which is the most popular tool for selection of minimal embedding dimension. The idea of FNN is that if the current embedding dimension d is sufficient to resolve the dynamics, then the images reconstructed in higher dimensions will no longer change considerably. In particular, points which were close in the d-dimensional space should remain close in the $(d+1)$-dimensional space. But if the embedding dimension is too small, points, which are in reality far apart, may seem as neighbours (consequence of projecting into a smaller space).

As the next presentation shows, the results of the FNN method do not contradict to observability indices, though its ability to detect the best observables is limited.

Acknowledgements

References

[1] Takens F. Detecting Strange Attractors in Turbulence. *Lecture Notes in Mathematics*, 898, 366-381, 1981.

[2] Cross DJ, Gilmore R. Differential embedding of the Lorenz attractor. *Physical Review*, E 81, 066220-9, 2010.

[3] Cross DJ, Gilmore R. Representation theory for strange attractors. *Physical Review*, E 80, 056207-6, 2009.

[4] Letellier C, Maquet J, Le Sceller L, Gouesbet G, Aguirre LA. On the non-equivalence of observables in phase-space reconstructions from recorded time series. *J. Phys. A: Math. Gen.*, 31, 7913–7927, 1998.

[5] Letellier C, Aguirre LA, Maquet J. How the choice of the observable may influence the analysis of nonlinear dynamical systems. *Communications in Nonlinear Science and Numerical Simulation* 11, 555–576, 2006.

[6] Letellier C, Aguirre LA. Symbolic observability coefficients for univariate and multivariate analysis. *Physical Review*, E 79, 066210-11, 2009.

[7] Aguirre LA, Bastos SB, Alves MA, Letellier C. Observability of nonlinear dynamics: Normalized results and a time-series approach. *Chaos* 18, 013123-9, 2008.