

CONF 901209-4

LA-UR -90-3328

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36

LA-UR--90-3328

DE91 000184

TITLE: State Space Forecasting and Noise Reduction

AUTHOR(S): Stephen Eubank, Martin Casdagli,
J. Doyne Farmer, John Gibson
Los Alamos National Laboratory

SUBMITTED TO: IEEE Conference on Decision & Control
Honolulu, Hawaii, 1990

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Received by
OCT 04 1990

By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes

The Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER

Los Alamos Los Alamos National Laboratory
Los Alamos, New Mexico 87545

State Space Forecasting and Noise Reduction

S. Eubank, M. Casdagli, J. D. Farmer, J. Gibson
Los Alamos National Laboratory
Los Alamos, NM 87545

Abstract

We discuss the effects of nonlinearity and noise on the problem of finding states and dynamics. We point out problems which arise in the construction of state space models for scalar time series. We describe solutions for the case when error due to noise dominates parameter estimation error and outline a framework for the general case.

1. States and Dynamics

A deterministic view of the world suggests that the best approach to modeling systems is to identify underlying states and their dynamics. The two are closely related, since a state can be defined as that set of quantities in terms of which the dynamics is completely deterministic. The problem of finding states can be expressed as the problem of embedding time series data in a reconstructed state space. Thus recognizing the state variables for a system depends on finding good deterministic models for the dynamics in the embedding space.

Takens has proven that for generic dynamical systems, there is a diffeomorphism between the true state space and one reconstructed from delayed values of a scalar time series of observations on that system, as long as the dimension of the embedding space is large enough. Given noiseless data from a deterministic system and a method for modeling functions in high dimensional spaces, the determination of a good reconstruction for the state is not difficult: simply increase the dimension m of the reconstruction and fit models for the dynamics. When a good model is obtained, the corresponding delay space is by definition a good embedding. A dimension algorithm such as Grassberger and Procaccia's gives a reasonable first estimate of m . [1] There are two problems with applying this strategy to real data: nonlinearity and noise.

2. Nonlinearity And Noise

If our models are drawn from a class of functions, *e.g.* linear functions, which does not include the true dynamics, forecasting accuracy will be limited. We find that the best way to fit the graph of a nonlinear function in high dimensional space is with locally low order maps, such as splines rather than high order polynomials might be used to fit a one dimensional curve. The locality of the fits allows us to model a large class of dynamics without overparameterizing the models. [2,3]

In the presence of noise our definition of a state can be readily generalized by replacing "completely deterministic" with "as deterministic as possible". In this view, complete determinism is a limiting case of a Markov process in which the state transition probabilities are delta functions. The *degree* of determinism is measured by the width of the transition probability density. Since the error in a deterministic model is bounded below by this width, an optimal embedding is one which minimizes the width. The problems of noise reduction and optimal embedding are thus closely related.

In classical systems, we expect even noise to be deterministically generated. The distinction between noise and determinism is better thought of as a distinction between low and high dimensional dynamics. In this context, high means "higher than we are willing to model."

Even a small contamination by noise can invalidate Takens's theorem. For example, if the sampling rate $1/\Delta$ is too large the change between consecutive measurements will be obscured by noise. In this case not all delay reconstructions in $2d + 1$ dimensions will be sufficient for determining the state, since the coordinates $x_0(t)$ and $x_0(t - \Delta)$ are redundant. Noise in a chaotic system makes the past seem causally disconnected from the future. That is, the measured values $x_0(t)$ and $x_0(t + k\Delta)$ are statistically independent variables. In this case there is nothing to be gained by making the embedding dimension larger than $k - 1$.

3. Nonlinear Embeddings

When the noise level is low and the amount of data large enough that distances between nearby points are smaller than the local radius of curvature of the dynamics, local principal components embeddings are optimal. These are local analogues of Broomhead and King embeddings, which project neighborhoods from a high dimensional delay space onto lower-dimensional hyperplanes.[4] The appropriate dimension for the hyperplanes is determined by a singular value analysis of the correlation matrix on a local neighborhood. The projection averages over noisy degrees of freedom to yield a better embedding than delay reconstruction in the same dimension. Figure 1 compares the accuracy of forecasts \hat{x} for noisy data from the Lorenz attractor using global and local principal component embeddings in three dimensions. Error is measured relative to the true (non-noisy) value

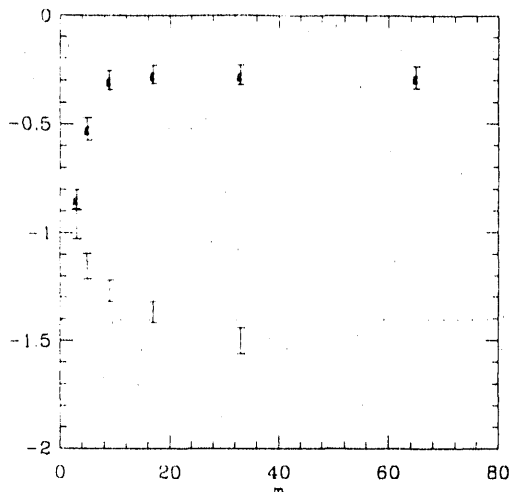


Figure 1: Forecast error using global (upper points) and local principal component embeddings *vs.* the number of dimensions in the original delay space. The dotted line shows the noise level in the original data.

of the time series $x(t)$:

$$E \equiv \langle \log(|x(t+T) - \hat{x}(t+T)|/\sigma_x) \rangle$$

at $T = 10$. The forecast error for 3 dimensional delay coordinates is -0.9 for this example. [5] For details of the technique and the conditions under which it performs well see the reference.

An alternative procedure is to use the estimated dynamics to push data measured at different times forward or backward to a common time, effectively yielding many measurements of a single variable. In addition, estimates of the relative precision of each measurement can be determined from the approximate dynamics. These estimates give the weights with which the measurements should be averaged together.[2] Figure 2 shows the noise reduction achieved using this procedure when the true dynamics is known.

For higher noise levels, a nonlinear generalization of factor analysis would be useful. We outline below some of the steps in such a method. Estimates of probability densities based on trial embeddings are used to build dynamical models which in turn help refine the embeddings.

Consider the problem of finding a transformation from (x_1, x_2) to (y_1, y_2) , such that y_1 predicts a third variable z as well as possible. A typical linear solution uses Gram-Schmidt orthogonalization to enforce the following two conditions: y_1 and y_2 are not redundant, and y_1 is maximally relevant to z . The second (variational) condition is equivalent to the (algebraic) condition that y_2 be irrelevant to z . In this framework, redundancy and relevance are measured by correlations.

These two conditions and the method of solution can be generalized to the nonlinear case.

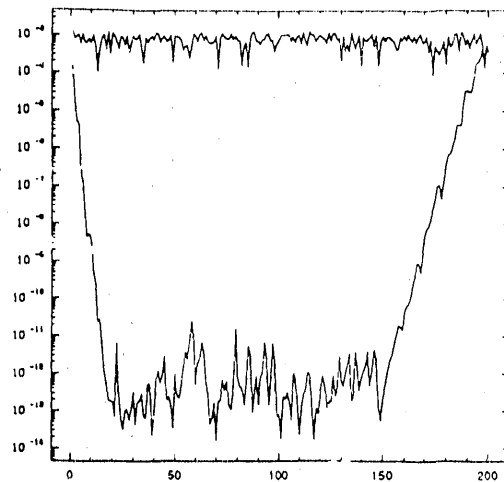


Figure 2: Observational noise added to a time series is removed using knowledge of the dynamics. The noise in the original signal is shown in the upper curve; that of the cleaned signal in the lower curve.

Redundance is the amount of information about z which is common to both y_1 and y_2 . Variables which are redundant can be combined to reduce noise. Redundance is best quantified not by a covariance as above, but by the independence of y_2 and z given y_1 . An information based statistic which measures redundancy is $R \equiv I(z; y_1)/I(z; y_1, y_2)$, where the semicolon denotes mutual information. The second condition translates to maximizing the mutual information $I(z; y_1)$ or minimizing $I(z; y_2)$. A nonlinear generalization of Gram-Schmidt is likely to be useful for enforcing these conditions: given two variables x_1 and x_2 which are not statistically independent, they can be made so using the invertible transformation:

$$y_1 = x_1, \quad (1)$$

$$y_2 = \int_{-\infty}^{x_2} p_{x_2|x_1}(x) dx. \quad (2)$$

References

- [1] P. Grassberger and I. Procaccia. *Physical Review Letters*, 50:346, 1983.
- [2] J.D. Farmer and J.J. Sidorowich. Exploiting chaos to predict the future and reduce noise. In Y.C. Lee, editor, *Evolution, Learning and Cognition*, World Scientific, 1988.
- [3] M. Casdagli. Nonlinear prediction of chaotic time series. *Physica D*, 35:335, 1989.
- [4] D.S. Broomhead and G.P. King. Extracting qualitative dynamics from experimental data. *Physica D*, 20:217, 1987.
- [5] M. Casdagli, S. Eubank, J.D. Farmer, and J. Gibson. *to appear in Physica D*, 1990.

END

DATE FILMED

11 / 08 / 90

