A Method to Enhance Noise Reduction for Data Generated from a Known Differential Equation

P. G. Vaidya*) and Savita Angadi

Mathematical Modelling Unit, National Institute of Advanced Studies, Indian Institute of Science Campus, Bangalore 560 012, India

In this paper we propose a method to enhance noise reduction for data generated from a known differential equation. We develop a theoretical basis for the procedure and then illustrate it in the case of some data generated using Duffing's equation. This method consists of embedding the data in higher dimensions and then transforming the data into a lower dimension using a singular matrix. We show that the singular matrix squeezes out some of the noise and leaves the true signal intact. Finally, using a nonlinear function, we reverse the effect of the singular matrix to get closer to the original data.

§1. Introduction

In practice, it turns out that often noise contaminates chaotic signals. This leads to great difficulty in accessing the information carried by the chaotic signal. Thus the problem of cleaning a chaotic signal from external noise is of great interest in many applications. A number of linear and nonlinear noise reduction methods have been proposed.[1)–6)] In this paper we demonstrate a method to enhance the noise reduction for data generated from a known differential equation. We assume that the noise is additive. The method consists of embedding data in higher dimensions and then transforming the data into a lower dimension using a singular matrix. We show that the singular matrix squeezes out some of the noise and leaves the true signal intact. Then, using a nonlinear function, we reverse the effect of the singular matrix to get closer to the original data. This function is independent of the initial conditions and can be empirically determined via numerical experiments using the known differential equation.

§2. Overall strategy

2.1. *Statement of the problem*

A time series $s_0, s_1, s_2, \cdots$ and the sampling interval h between samples is given. Initial conditions are not given but we know the underlying equation of the dynamics. It is also known that the data are corrupted with additive noise. We need to find a portion of the time series with reduced noise.

2.2. *Method of solution*

Our method consists of the following steps.

1. Embedding the noisy data in higher dimensions. In this paper, we have used a special method of embedding which we call derivative-like embedding.[7)]

*) E-mail: pgvaidya@nias.iisc.ernet.in

2. Using a singular matrix, transforming the data into a lower dimension. This is a crucial stage. Our hope is that the singular matrix will "squeeze out" some of the noise but will leave the true signal more or less intact.
3. Using a nonlinear function (which we have to determine) to "reverse" the effect of the singular matrix to get closer to the original data.

2.3. *Embedding*

State space reconstruction is the creation of a multidimensional, deterministic state space from a lower dimensional time series. When the dimension is sufficiently high, a reconstruction is almost always an embedding: the process of finding a space in which smooth dynamics no longer has overlaps. There are three popular methods of embedding in popular use: delays,[8] derivatives[9] and principal components.[10]

From the data set we could choose N numbers next to one another and create a set of N dimensional vectors as follows, where $r(i)$ is an integer which depends on i and is distinct for each value of i.

$$x^i = \begin{bmatrix} x_{r(i)} & x_{r(i)+1} & \vdots & x_{r(i)+N-1} \end{bmatrix}^T.$$

For example if $r(i) = 15$, we would get

$$s = \begin{bmatrix} s_{15} & s_{16} & s_{17} & \vdots & s_{15+(N-1)} \end{bmatrix}^T,$$

where 15 was just arbitrarily chosen. The vectors x^i are like time delay embedding vectors, and the time delay τ is just h.

2.4. *Specific steps*

(1) Note down the dimension (D) of the equation. (2) Solve the differential equation using a Runge-Kutta or similar procedure. Using any initial conditions, generate a data set, sampled at the same sampling rate $h : x_0, x_1, x_2, \cdots$. (3) Choose a number N, much greater than $(2D+1)$. (4) Form a set of vectors x^p from various contiguous sets of data as described above. (5) Choose another integer M such that $N > M \geq 2D + 1$. (6) Choose a matrix J with M rows and N columns. Then find $y = Jx$. With $M < N, J$ is not of full rank N. In "most" cases the rank of $J \geq (2D + 1)$. It turns out that operation of the rows of J on x is analogous to the operation of an F.I.R filter.[10] If rank $\geq (2D + 1)$ then y bears some bijective relationship with x in the absence of noise.[11] In this paper we assumed that noise is a dynamics of order much greater than M and is additive. (7) Find the inverse function by fitting the first P components of the various x vectors to the P components of y vectors, ($P \leq M$) for this purpose we have used here a trilinear multivariable fit. (8) Now go back to the data set $(s_0, s_1, \cdots)$. Form the matrix.

$$T = \begin{bmatrix} s_0 & s_1 & s_2 & \cdot & \cdot \\ s_1 & s_2 & s_3 & \cdot & \cdot \\ s_2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ s_{N-1} & s_N & s_{N+1} & \cdot & \cdot \end{bmatrix}.$$

By partitioning the matrix T, we can verify $H = JT$, where the columns of H give various y vectors. (9) From each of the y vectors use f to find the value of the signal. We try to predict the middle row: $[s_{N/2},\ s_{(N/2)+1},\ s_{(N/2)+2},\ \cdots]$.

§3. Result of J on signal and noise

J does not destroy the D order dynamics but if we choose it well, it will damage the high order dynamics, $J(\text{data} + \text{noise}) = J(\text{data}) + J(\text{noise})$.

Here it turns out that only data get embedded but not noise. J is chosen such that it hurts noise and leaves data intact.

3.1. *Inverse of J*

We know that (in the absence of noise and $M \geq 2D + 1$) there exists some function f such that $x = f(y)$. The question then is, how do we find the nonlinear function? Since we know the data, we can empirically determine the nonlinear function.[10] This is the function, we use in step (3) above. In the absence of any further information, i.e., in the absence of any further reduction in entropy (Uncertainty) of information. We can show that:

The maximum entropy solution $\Leftrightarrow$ The least square solution $\Leftrightarrow$ $\bar{x}$, $\bar{x} = Ky$, where $K =$ generalized inverse of J.

3.2. *Effect of generalized inverse (K)*

$\bar{x}$ may not be same as x. For each J there exists a class of noise free data for which x is same as $\bar{x}$. But in general f has to be found so that $x = f(y)$. f is usually nonlinear.

3.3. *Choice of J for this paper*

For this paper, we choose K and find J as its generalized inverse using singular value decomposition. In this case, K is chosen to represent the terms of a truncated Taylor's series so that

$$K_{n,m} = \frac{\left(n - \frac{N}{2}\right)^m h^m}{m!}, \quad 0 \leq n \leq N-1, \quad 0 \leq m \leq M-1.$$

In case x can be represented as a polynomial of order less than or equal to M, then J is an exact inverse.

3.4. *Selection of M and N*

Some trial and error is useful in the selection of M and N. M should be larger than $2D$, as stated before. If N is much larger than 2D, then a large amount of noise reduction is possible. However, for very much larger N the inverse function becomes harder to find.

§4. Numerical results

We choose Duffing's equation under chaos:

$$\frac{d^2}{dt^2}x + c\frac{d}{dt}x + kx + \delta x^3 = F\cos(\omega t + \alpha),$$

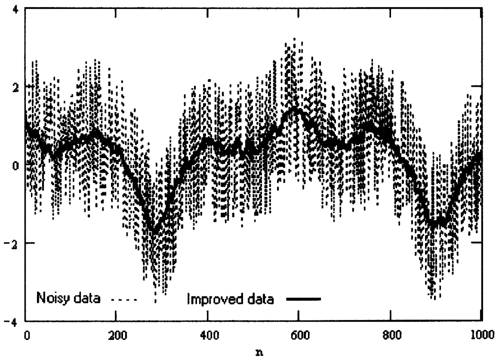

Fig. 1. Comparison of highly sampled noisy data and recovered data ($h = 0.0011179$).

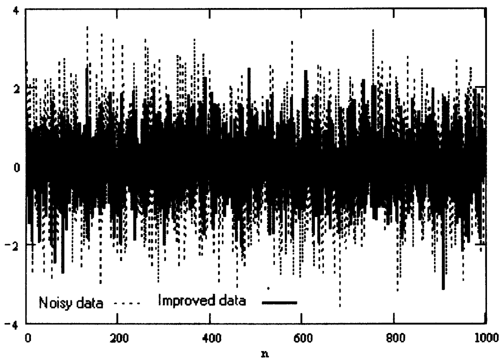

Fig. 2. Comparison of weakly sampled noisy data and recovered clean data ($h = 0.11179$).

where the parameters are $c = 0.044964$, $k = 0$, $\delta = 1$, $F = 1.02$, $\omega = 0.44964$, $\alpha = 0$.

There were two cases. For the first case the sampling interval $h = 0.0011179$, $N = 1000$, $M = 10$ and $P = 4$. In this case we added 160.23% noise to the data and then recovered the data with only 8.27% noise (Fig. 1). In the second case we consider the sampling time 100 times the before. $N = 100$, $M = 10$ and $P = 4$. In this case we added 150.63% noise to the data and recovered the data with 27.27% noise (Fig. 2). These figures are calculated by taking the ratio of standard deviation of noise and original data.

§5. Conclusions

We have developed a general procedure to eliminate noise from data generated by a known equation. Specifically, chaotic data generated from Duffing's equation was used. For the choice of the singular matrix, a generalized inverse of a truncated Taylor's series was chosen. Results are spectacular for very frequently sampled data and good for not very frequently sampled data.

References

1) E. J. Kostelich and J. A. Yorke, Phys. Rev. A **38** (1988), 1649; Physica D **41** (1990), 183.
2) J. D. Farmer and J. Sidorowich, Physica D **47** (1991), 373.
3) S. M. Hammeel, Phys. Lett. A **148** (1990), 421.
4) T. Schreiber and P. Grassberger, Phys. Lett. A **160** (1991), 411.
5) T. Saur, Physica D **58** (1992), 193.
6) R. Cawley and G. H. Hsu, Phys. Rev. A **46** (1992), 357; Phys. Lett. A **166** (1992), 188.
7) P. G. Vaidya and S. Angadi, Chaos, Solitons and Fractals **17** (2003), 379.
8) F. Takens, "Detecting strange attractors in fluid turbulence", in *Dynamical Systems and Turbulence*, ed. D. Rand and L. S. Young (Springer, Berlin, 1981).
9) N. H. Packard, J. P. Crutchfield, J. D. Farmer and R. S. Shaw, Phys. Rev. Lett. **45** (1980), 712.
10) D. S. Broomhead and G. P. King, Physica D **20** (1987), 217.
11) Nonlinear Inverse filtering in the presence of the noise 'Broomhead and Huke, AIP Conference Proceedings.
12) E. J. Kostelich and T. Schreiber, Phys. Rev. E **48** (1993), 1752