Commun Nonlinear Sci Numer Simulat 16 (2011) 2999–3004

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

journal homepage: www.elsevier.com/locate/cnsns

Short communication

Using information to generate derivative coordinates from noisy time series

B.P. Mann [a,*], F.A. Khasawneh [a], R. Fales [b]

[a] *Department of Mechanical Engineering & Material Science, Duke University, Durham, NC 27708, USA*
[b] *Department of Mechanical & Aerospace Engineering, University of Missouri, Columbia, MO 65211, USA*

ARTICLE INFO

Article history:
Received 25 August 2010
Received in revised form 16 November 2010
Accepted 21 November 2010
Available online 3 December 2010

Keywords:
Information theoretic
Signal derivatives
Derivative coordinates

ABSTRACT

This paper describes an approach for recovering a signal, along with the derivatives of the signal, from a noisy time series. To mimic an experimental setting, noise was superimposed onto a deterministic time series. Data smoothing was then used to successfully recover the derivative coordinates; however, the appropriate level of data smoothing must be determined. To investigate the level of smoothing, an information theoretic is applied to show a loss of information occurs for increased levels of noise; conversely, we have shown data smoothing can recover information by removing noise. An approximate criterion is then developed to balance the notion of information recovery through data smoothing with the observation that nearly negligible information changes occur for a sufficiently smoothed time series.

1. Introduction

It is rarely practical to measure each state variable in an experimental setting. To circumvent this issue, methods exist for reconstructing a pseudo-state space from a small number of measured observables [1–3]. A primary benefit of the reconstruction is that it produces a pseudo-state space with dynamics equivalent to those of the original state space. As a consequence of the equivalence, an attractor in the reconstructed state space has the same invariants, such as Lyapunov exponents and dimension, as the original attractor [4].

While delayed embedding is the predominant choice for attractor reconstruction [3,5–8], the use of derivative coordinates is appealing, given their obvious physical meaning. For instance, many physical systems can be described by a set of equations with states that are related through their derivatives. While the numerical derivatives of a signal can be used for noise-free data, the presence of noise renders the signal derivatives to be poor approximates – owing to noise amplification in the signal derivatives.

The inherent goal of signal analysis is to extract useful information about a system from the observed data. Consider a continuous and deterministic system that has produced the scalar time series $q(t)$. To mimic realistic data from an experiment, noise is superimposed onto the noise-free data as follows: $\eta(t) = q(t) + \sigma\alpha(t)$, where $\alpha(t)$ is a time series with normally distributed random noise, a zero mean, and a standard deviation equal to σ. The underlying goal of this investigation is to extract $q(t)$ and the derivatives of $q(t)$ from the noisy time series $\eta(t)$.

This work investigates smoothing $\eta(t)$ to recover $q(t)$ and its derivatives. While the process of data smoothing is well known, the question of how much smoothing yields accurate derivative coordinates is unclear. To answer this question, we explored the use of an information theoretic, known as the average mutual information, to develop a criterion for the appropriate amount of data smoothing.

* Corresponding author. Tel./fax: +1 919 660 5214.
E-mail address: brian.mann@duke.edu (B.P. Mann).

1007-5704/$ - see front matter
doi:10.1016/j.cnsns.2010.11.011

The work of this paper is organized as follows. The next section describes the data smoothing technique and average mutual information tools used in our analyses. These discussions are followed by a series of example results that apply the average mutual information to investigate the influence of noise and an approximate criterion for the level of data smoothing.

2. Example implementation

The investigations that follow use synthetic data generated from a Duffing oscillator,

$$q'' + \mu q' + \omega^2 q + \beta q^3 = \Gamma \cos \Omega t, \tag{1}$$

where a prime denotes a derivative with respect to time. Numerical simulation was used to generate a chaotic time series for $q(t), q'(t)$, and $q''(t)$ while using the following parameters $\mu = 0.2$, $\omega = 1$, $\beta = 1$, $\Gamma = 27$, and $\Omega = 1.33$; however, in an effort to mimic the realistic challenges of an experiment, we have assumed only the noisy time series is observable $\eta(t) = q(t) + \sigma\alpha(t)$. The remainder of this section describes the application of data smoothing for noise removal and the use of the average mutual information to determine the appropriate smoothing level to recover the derivatives of $q(t)$.

2.1. Noise removal with smoothing

Cubic splines are often applied to empirical data to estimate interim points or data points that lie between two measurements. The basic idea is to fit the data with a piecewise polynomial,

$$s(t) = b_{i0} + b_{i1}(t - t_i) + b_{i2}(t - t_i)^2 + b_{i3}(t - t_i)^3, \tag{2}$$

where the polynomial coefficients b_{i0}, b_{i1}, b_{i2}, and b_{i3} have subscripts that denote their validity between two neighboring data points - the time interval from t_i to t_{i+1}. Therefore, cubic splines provide a piecewise fit to the data while simultaneously giving a functional relationship for the derivatives of the data, i.e. after differentiation of Eq. (2). Furthermore, cubic splines invoke continuity in the signal and its first two derivatives at the intersection of the neighboring time steps [9].

Smoothing splines, which differ from the typical cubic spline fitting operation, provide a refinement to the idea using a polynomial to fit empirical data. The basic difference lies in the introduction of a smoothing parameter which reduces noise amplification in the signal derivatives by balancing a fit between the measured data $\eta(t)$ and the smoothness in the second derivative [10]. In the results that follow, cubic smoothing splines have been implemented; this approach obtains the polynomial coefficients for Eq. (2) by minimizing

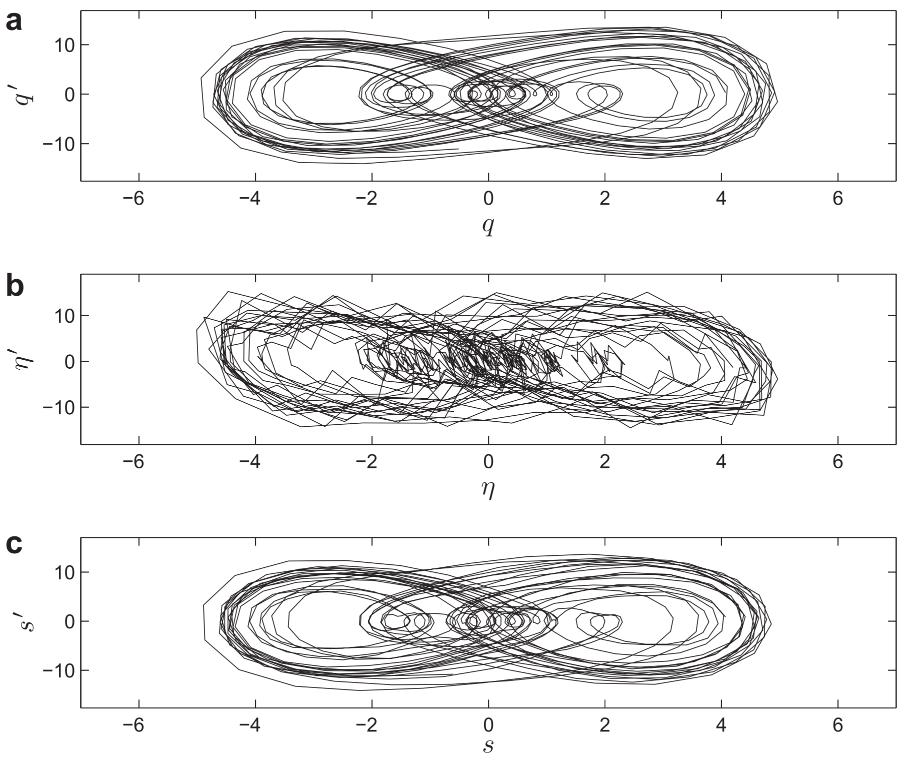

Fig. 1. Two dimensional state space with graphs showing: (a) the actual q and q'; (b) the noisy time series η and the numerical derivative η' when $\sigma = 0.1$; and (c) the smoothed version s of the noisy time series and the derivative of the smoothed signal s' obtained when $\gamma = 99.95 \times 10^{-2}$.