Topological dimension and local coordinates from time series data

1987 J. Phys. A: Math. Gen. 20 L563

(http://iopscience.iop.org/0305-4470/20/9/003)

J. Phys. A: Math. Gen. **20** (1987) L563-L569. Printed in the UK

LETTER TO THE EDITOR

Topological dimension and local coordinates from time series data

D S Broomhead, R Jones and Gregory P King†

Royal Signals and Radar Establishment, St Andrews Rd, Gt Malvern, Worcestershire WR14 3PS, UK

Received 7 April 1987

Abstract. A method for the estimation of the topological dimension of a manifold from time series data is presented. It is based on the approximation of the manifold near a point x by its tangent space at x. The dimension of the tangent space is estimated by constructing a maximal set of linearly independent vectors from the data near x using the method of singular value decomposition. The method is used to analyse experimental data obtained from a non-linear electronic oscillator in a chaotic state.

Recently techniques have appeared in the literature for analysing time series data for geometric information relating to the attractor of the underlying dynamical system [1-3]. These techniques assume that the attractor lies within a smooth, finite dimensional manifold M ($\dim(M) = m$). An embedding of M is then constructed using delays or derivatives of time series data. An embedding preserves the topological structure of the manifold and important information about the restriction of the dynamical system to the manifold. These embedding techniques are used as the basis for the calculation of, for example, Lyapunov exponents and various fractal dimensions and entropies [4].

Rigorous justification for this approach was given by Takens [2], following Whitney [5] who showed that M can always be embedded in the Euclidean space $\mathbb{R}^n$, where $n \geq 2m+1$ (although fewer dimensions may be sufficient for particular manifolds). In practice, however, the absence of a simple and effective estimate for m limits the usefulness of this result. We present a resolution of this difficulty whereby m, the topological dimension of M, can be estimated from time series data. This leads naturally to a description of M, obtained from experimental data, in terms of its basic mathematical structure, i.e. a collection of locally flat neighbourhoods related globally by a set of charts. Topological dimension is an integer and is easy to compute. The calculation uses standard algorithms whose convergence properties are well known and are robust in the presence of noise.

The subject of this letter is an extension of earlier work on the singular system analysis of time series data [3]. The method of delays [2] is used to construct the sequence, Ψ, of n-dimensional vectors $\{x_1, x_2, \ldots, x_N\}$ where $x_j^{\mathrm{T}} = (v_j, v_{j-1}, \ldots, v_{j+n-1})$ from the measurements $\{v_1, v_2, \ldots, v_{N+n-1}\}$. The fundamental entity in the singular system analysis is the rectangular trajectory matrix X whose rows are the $\{x_j^{\mathrm{T}}\}$. The

† Also at: Department of Mathematics, Imperial College, London SW7 2BZ, UK.

right singular vectors of X are a natural orthonormal basis for the embedding space. The corresponding singular values (the global singular spectrum), which are the RMS projections of the trajectory onto the set of singular vectors, give a crude indication of the distribution of the entire attractor in the embedding space.

In general, the global singular spectrum has d significant, deterministic components and $n-d$ components dominated by noise. Thus the attractor is effectively contained in the d-dimensional subspace spanned by the corresponding d singular vectors. This partitioning implies that, for noisy finite precision data, $d \geqslant 2m+1$ should replace $n \geqslant 2m+1$ in Takens' theorem. The value of d says nothing about the value of m since d is sensitive to arbitrary curvature in the embedding of M and hence is not an invariant of the embedding process [3]. To use a singular spectrum to obtain an estimate $\hat{m}$ for m such curvature effects must be eliminated.

A manifold M is a topological space which locally looks like $\mathbb{R}^m$ [6]. Thus, for a small enough neighbourhood of any point $x \in M$, the effects of curvature become unimportant and M is well approximated by its tangent space at x.

A local analysis about a point $x \in \Psi$ uses those elements of Ψ which lie within an open ball of radius ε centred on x. These points are represented by the $N_b \times d$ ε-neighbourhood matrix $B_\varepsilon(x)$ whose rows consist of the vectors $(x_j - x)^T$ where $\{x_j\} \in \Psi$ such that $|x_j - x| < \varepsilon$. For sufficiently small ε the rows of $B_\varepsilon(x)$ are approximately tangent vectors to M at x. Then if $N_b > m$, we make the generic assumption that the rows of $B_\varepsilon(x)$ span the tangent space of M at x. Since the rank measures the number of linearly independent rows of a matrix, the rank of $B_\varepsilon(x)$ is, generically, the dimension of M.

The most effective numerical method for obtaining the rank of a general matrix is through its singular spectrum [7]. Thus a local analysis of the experimental trajectory leads to a local singular system analysis using the neighbourhood matrices $B_\varepsilon(x)$. In this case the singular vectors give a local coordinate system centred at x, while the local singular spectrum characterises the disposition of sample points within the ε ball centred at x.

In practice, ε cannot be made arbitrarily small since both the quantity and precision of the data are finite. We therefore estimate the effects of curvature with a model calculation. Let y denote a point in the neighbourhood of x. In suitable local coordinates $y^T = (y_1, y_2, \ldots, y_m, f_{m+1}(\{y_j\}), \ldots, f_d(\{y_j\}))$, where $f: \mathbb{R}^m \to \mathbb{R}^{d-m}$ is such that $f_i(\{0\}) = 0$ and $\partial f_i(\{0\})/\partial y_j = 0$ for $m+1 \leqslant i \leqslant d$ and $1 \leqslant j \leqslant m$. We assume that $m = 2$, $d = 3$ and $f_3(y_1, y_2) = ay_1^2 + by_1y_2 + cy_2^2 + O(y^3)$, where a, b and c determine the local curvature of the manifold. Formally, the singular spectrum of $B_\varepsilon(x)$ is the square root of the eigenvalue spectrum of $B_\varepsilon^T B_\varepsilon$, and the singular vectors of $B_\varepsilon(x)$ are the eigenvectors of $B_\varepsilon^T B_\varepsilon$. Each matrix element of $B_\varepsilon^T B_\varepsilon$ is a covariance of components of y averaged over the points in the neighbourhood. Thus

$$B_\varepsilon^T B_\varepsilon = N_b \begin{bmatrix} \langle y_1^2 \rangle & \langle y_1 y_2 \rangle & \langle y_1 f_3 \rangle \\ \langle y_1 y_2 \rangle & \langle y_2^2 \rangle & \langle y_2 f_3 \rangle \\ \langle y_1 f_3 \rangle & \langle y_2 f_3 \rangle & \langle f_3^2 \rangle \end{bmatrix}$$

where $\langle\ \rangle \equiv (1/N_b)\Sigma_k (\ )$, $k = 1, \ldots, N_b$. Assuming an isotropic, randomly sampled neighbourhood in the continuum limit this matrix reduces to a form that can be readily diagonalised to yield the following singular values of $B_\varepsilon(x)$: $\{\frac{1}{2} N_b^{1/2}\varepsilon, \frac{1}{2} N_b^{1/2}\varepsilon + O(\varepsilon^5), (\beta N_b)^{1/2}\varepsilon^2 + O(\varepsilon^4)\}$, where β is a numerical factor depending on a, b and c.

In the absence of curvature or for small enough ε, $\text{rank}(B_\varepsilon(\boldsymbol{x})) = m = 2$. This is an example of the general result anticipated above. The two singular vectors of $B_\varepsilon(\boldsymbol{x})$, corresponding to the significant singular values, span the tangent space at $\boldsymbol{x}$. Since the significant singular values, normalised by $N_b^{1/2}$, measure the RMS radius of the neighbourhood projected onto the tangent plane, they scale linearly with ε. This model ignores the possibility of the manifold being 'thin' in some dimension. If the 2-manifold is a narrow ribbon, one of the linearly growing singular values will saturate as ε increases beyond the width of the ribbon. The third singular vector is normal to the manifold at $\boldsymbol{x}$, and its corresponding singular value measures the deviation of the manifold from the tangent plane. In this example the deviation is due solely to local curvature and scales as ε^2. At degenerate points on the manifold $f_3 = O(y^k)$, and the third singular value scales as ε^k.

In summary, we expect for small enough ε that the local singular spectrum for a neighbourhood of an m-manifold in $\mathbb{R}^d$ will have m approximately equal singular values above a 'noise floor'. As ε is increased these will grow linearly until saturation or the effects of curvature in the manifold become noticeable. The remaining $d-m$ singular values, which are dominated by noise, are independent of ε until the effects of curvature become important. These singular values will then grow as ε^2 or faster.

In chaotic systems a thin direction may be associated with the corresponding cross section of the attractor being a Cantor set. A model calculation using a Cantor set constructed by removing the 'middle rth' from a straight line segment of length L embedded in some $\mathbb{R}^d$, and assuming a uniform distribution of N_b points on this set, shows that the neighbourhood matrix has unit rank and a corresponding singular value $\sigma = N_b^{1/2} L/(3-r)^{1/2}$ where $r \in (0,1)$. Thus for $\varepsilon > L$, $\sigma/N_b^{1/2}$ will appear as a saturated singular value which must exceed the noise floor if any aspect of the Cantor set is to be experimentally observable. As ε decreases from L, $\sigma/N_b^{1/2}$ decreases in the form of a devil's staircase [8]. Typically, the observability of this in an experiment will be limited by noise and finite amounts of data.

These ideas are used to analyse experimental data taken from a non-linear electronic oscillator with natural frequency ≈ 70 Hz [9]†. The output was quantised with a 14 bit A/D converter sampling at 10 KHz. Approximately 2.5×10^5 samples of the time series were taken with the oscillator in a chaotic state following a sequence of period doubling transitions. After removal of its mean, the time series was normalised by its standard deviation of 2441 A/D levels. The global singular spectrum, calculated using $n = 50$, yields $d = 5$. Figure 1 shows a projection of the trajectory, and also shows the points about which we analyse the manifold locally: at $\boldsymbol{x}_{23}$ the flow is compressive and densely samples the manifold, at $\boldsymbol{x}_{92}$ the flow is divergent, strongly curved, and sparsely samples the manifold, and at $\boldsymbol{x}_{209}$ the flow is transitional between the other two cases. These points serve as examples of the types of local structure that we have observed in this system. The results of the calculations are summarised in figures 2-4 which show plots of the local singular spectrum, scaled by $2N_b^{-1/2}$, as a function of ε. For computational convenience we restricted N_b to a maximum of 250 points, although at the smallest values of ε used, values of $N_b \simeq d$ (but greater than $\hat{m}$) were found to be adequate.

We implemented our algorithm on a 16 bit minicomputer with double precision floating point hardware and the singular value decomposition was done directly on the neighbourhood matrix using the Golub-Reinsch algorithm (NAG-F02WAF). The

† Details of this oscillator and its analysis will be discussed in a future publication.

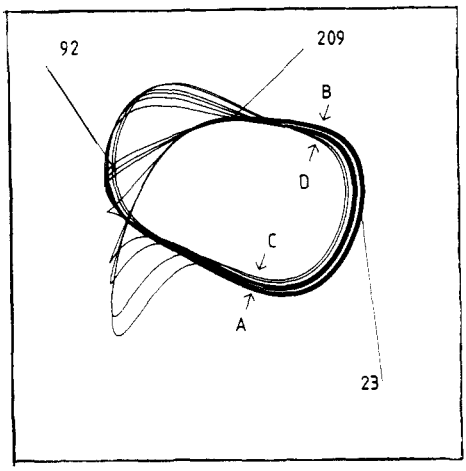

Figure 1. A projection of 2000 samples of the trajectory reconstructed from the time series. The numbers indicate the positions x_{23}, x_{92} and x_{209} of the origins of the neighbourhoods used in the analysis. The direction of the flow is anticlockwise. The point B maps to C, while D maps to a region near A.

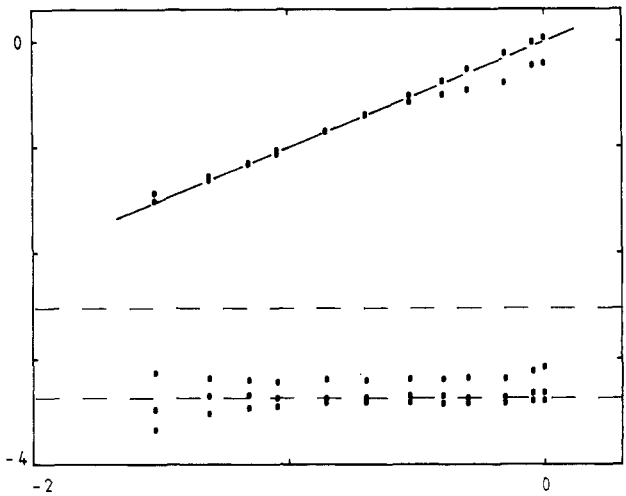

Figure 2. A log-log plot of the local singular spectrum scaled by $2N_b^{-1/2}$ against the radius, ε, of neighbourhoods taken about the point x_{23}. The horizontal broken lines mark the range of the global singular values found in the noise. For comparison of the data with the model calculation, the diagonal line is a plot of ε.

total time taken for the search for the neighbourhoods and the decomposition of the neighbourhood matrices for the three reference points was about 100 min. This time was dominated by the search for the neighbourhoods.

Figure 2 shows the results of the local analysis for x_{23}: there are two singular values which scale linearly with ε and three which are independent of ε and are in the noise (estimated from the global singular spectrum). We conclude that $\hat{m} = 2$.

The results of the local analysis for x_{209}, shown in figure 3, are similar to those for x_{23}. However, for $\varepsilon > 10^{-0.5}$ a new feature appears: singular values grow from the

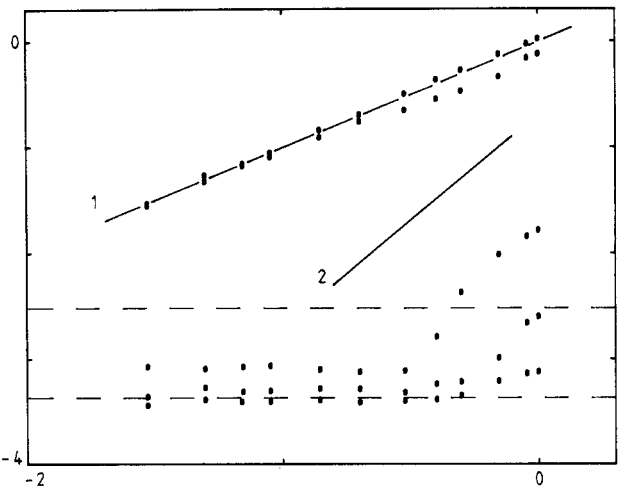

Figure 3. A plot similar to figure 2, for neighbourhoods taken about x_{209}. The line marked 1 is a plot of ε. The line marked 2 has slope 2 and is for comparison with the plot of the singular value growing from the noise floor. Here the comparison is inconclusive.

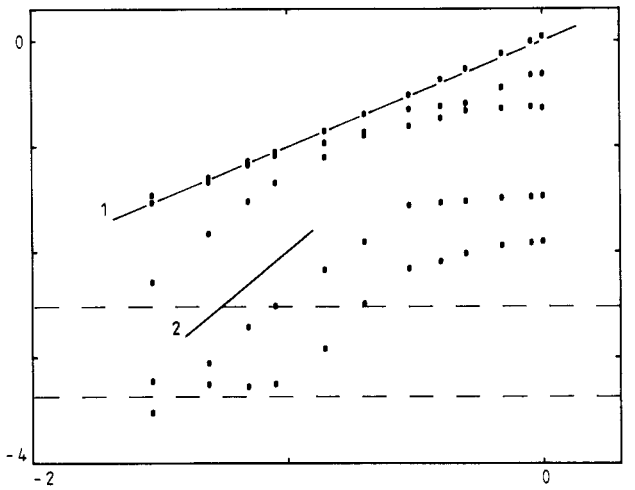

Figure 4. A plot similar to figures 2 and 3 for neighbourhoods taken about x_{92}. Note the smaller singular values which begin to scale as ε^2 as they grow out of the noise.

noise faster than ε indicating curvature effects on the scale of the neighbourhood. Note that this means that the curvature of the manifold is greater here than in the region near x_{23}.

The results for x_{92} are shown in figure 4. Here two singular values scale linearly with ε for $\varepsilon < 10^{-1}$. At greater ε saturation and curvature effects dominate. In addition, curvature effects are apparent in the three smallest singular values over the entire range of ε. At the smallest accessible values of ε, two of them scale as ε^2. In order to bring all three into the noise, we estimate from the figure that values of $\varepsilon < 10^{-2}$ will be required. Below this limit the manifold should appear flat. A covering dense enough to resolve the required length scales in this region is impractical since the divergent flow implies that here the natural measure [10] is small.

The calculations were repeated in a ten-dimensional subspace spanned by the first ten singular vectors. This produced no change other than an increase in the number of local singular values in the noise floor.

We conclude from our results that the flow constructed from the data is confined to a manifold whose topological dimension is two. However, to avoid self-intersection of the trajectory the flow must explore a third dimension. Indeed, the folding associated with the re-injection apparent in figure 1 implies a Cantor set structure in this third dimension. As discussed above, there is nothing intrinsic to a Cantor set which makes it invisible to a rank calculation. In the present case the thickness of the Cantor set is masked by experimental noise.

More detailed information about the manifold may be obtained from the projection shown in figure 1. In this figure the flow is anticlockwise: the portion from AC to BD (which contains x_{23}) is the weakly curved, compressive region already referred to, while the portion from BD to AC is the strongly curved region containing x_{92} and x_{209}. Despite the distortion of the latter region, our analysis indicates that the whole manifold consists of a two-dimensional 'ribbon'. Moreover, by following the trajectory, we see that B is mapped to C while D is mapped to a region near A. Thus the ribbon contains a half-twist and is, therefore, a Möbius band.

This result is connected with the fact that the oscillator was observed in a state reached through a period doubling sequence. Theoretically, it is known that close to a period doubling bifurcation the unstable manifold of a limit cycle is an attracting Möbius band [11]. We conjecture, therefore, that the attractor observed in this experiment is close to the unstable manifold of the limit cycle from which the period doubling sequence developed.

Our approach to the definition and calculation of topological dimension using finite sets of data points is an extension of earlier work on the singular system analysis of time series. Recently, however, we have been made aware of previous work in this area. Of this the work found in the psychometrics literature [12] and more recent work in the signal analysis literature [13] is based on the concept of iterative perturbation of the data points. The purpose of this is to deform the data set so that it fits into a minimum linear subspace whose dimension may be readily estimated. The motivation for this approach is to make more efficient use of the data in making the dimension estimate. The penalty is the loss of detailed local geometric information (e.g. local curvature). Moreover, the distortions must in general destroy the topological structure inherent in the data.

The work of Trunk [14] and of Froehling *et al* [15] adopted a statistical approach based on a goodness of fit of local linear models of the data taken over the whole data set. The paper by Froehling *et al* was the first to apply these methods to the attractor of a dynamical system. They set out to estimate an 'approximate fractal dimension' for strange attractors represented as branched manifolds. A problem with their approach is that the χ^2 statistic employed suffers distortion due to the orientation of the linear subspace containing the local data. This necessitated that the fit be performed twice for each linear model in each neighbourhood making the proceudre computationally inefficient. On the other hand, the singular value method orients the singular vectors so that the mean square error due to assuming that the data is fitted by any k-dimensional plane is minimised by choosing the plane spanned by the first k singular vectors. Note that for one calculation this method allows the goodness of fit to be checked for the whole range of possible values of k since the squared error is the sum of the squares of the remaining $d-k$ singular values. A further difficulty with the

approach taken by Froehling *et al* arises with the histogramming technique employed to display the distribution of the χ^2 statistic for the whole attractor. This emphasises the first large fall in the error and obscures effects due to the attractor being thin in some directions.

The work most directly connected with ours is that of Fukunaga and Olsen [16]. Their approach is mathematically equivalent to ours in that they derive a local Karhunen–Loeve basis by diagonalising the local covariance matrix. The distinctions that arise are due to differences in the intended application. The starting model of Fukunaga and Olsen is a randomly sampled manifold. In our case the manifold is not randomly sampled since the data is generated by a deterministic dynamical system and is thus found correlated within the attractor. It is therefore common to find thin directions, boundaries to the data, and correlated gaps (as in Cantor sets), all of which give discernable features in the scaling behaviour of the local singular spectra. Estimates of rank using *ad hoc* criteria, such as the number of eigenvalues exceeding a specified percentage of the largest eigenvalue or the number of eigenvalues before the first large decrease, cannot be expected to take account of these effects and can thus lead to erroneous estimates of m in the current application.

In this letter we have presented a method for the calculation of the topological dimension of a manifold constructed from time series data and applied it to the analysis of experimental data taken from a non-linear electronic oscillator. This technique is based on the approximation of the manifold near a point $\boldsymbol{x}$ by its tangent space at $\boldsymbol{x}$. The dimension of the tangent space is estimated by constructing a maximal set of linearly independent vectors from the local distribution of points near $\boldsymbol{x}$ using the method of singular value decomposition. We suggest that the estimates of d from the global analysis and m from the local analyses can be used to check the embedding criterion (modified for noisy, finite precision data) $d \geqslant 2m+1$.

References

[1] Packard N H, Crutchfield J P, Farmer J D and Shaw R S 1980 *Phys. Rev. Lett.* **45** 712

[2] Takens F 1981 *Lecture Notes in Mathematics* vol 898, ed D A Rand and L-S Young (Berlin: Springer) p 366

[3] Broomhead D S and King G P 1986 *Physica* **20D** 217; 1986 *Nonlinear Phenomena and Chaos* ed S Sarkar (Bristol: Adam Hilger) p 113

[4] Eckmann J-P and Ruelle D 1985 *Rev. Mod. Phys.* **57** 617

[5] Whitney H 1936 *Ann. Math.* **37** 645

[6] Bröcker Th and Jänich K 1982 *Introduction to Differential Topology* (Cambridge: Cambridge University Press)

[7] Stewart G W 1982 *Introduction to Matrix Computations* (New York: Academic)

[8] Mandelbrot B B 1982 *The Fractal Geometry of Nature* (San Francisco: Freeman)

[9] Freire E, Franquelo L G and Aracil J 1984 *IEEE Trans. Circuits Syst.* **CAS-31** 237

[10] Farmer J D, Ott E and Yorke J A 1983 *Physica* **7D** 153

[11] Alligood K T, Mallet-Paret J and Yorke J A 1981 *J. Diff. Geom.* **16** 483
Holmes P and Williams R F 1985 *Arch. Ration. Mech. Anal.* **90** 115

[12] Shepard R N and Carroll J D 1966 *Proc. Int. Symp. on Multivariate Analysis* ed P R Krishnaiah (New York: Academic) p 98

[13] Bennett R S 1969 *IEEE Trans. Inform. Theory* **IT-15** 517
Schwartzmann D H and Vidal J J 1975 *IEEE Trans. Comput.* **C-24** 1175

[14] Trunk G V 1976 *IEEE Trans. Comput.* **C-25** 165

[15] Froehling, H, Crutchfield J P, Farmer D, Packard N H and Shaw R 1981 *Physica* **3D** 605

[16] Fukunaga K and Olsen D R 1971 *IEEE Trans. Comput.* **C-20** 176