2005 International Symposium on Nonlinear
Theory and its Applications (NOLTA2005)
Bruges, Belgium, October 18-21, 2005

Reconstruction of chaotic systems with hidden variables using modified Bock's algorithm

Boris P. Bezruchko[†,‡], Dmitry A. Smirnov[‡] and Ilya V. Sysoev[†]

†Saratov State University, Department of Electronics, Oscillations and Waves
Astrakhanskaya str., 83, Saratov, Russia
‡Saratov Branch, Institute of RadioEngineering and Electronics of Russian Academy of Sciences
Zelenaya str., 20, Saratov, Russia,
Email: bbp@sgu.ru, smirnovda@info.sgu.ru, sysoevi_v@info.sgu.ru

Abstract—We consider the problem of estimating parameters of chaotic dynamical systems from a time series in a situation when some of state variables cannot be observed. Using specially developed quantitative criteria, we compare the efficiency of original multiple shooting approach (Bock's algorithm) and its modified version that is shown to be significantly superior for long chaotic time series.

1. Introduction

The problem of complex systems modelling from experimental time series is well-known and has multiple names such as "reconstruction of dynamical systems" in nonlinear science [1] and "system identification" in statistics and control theory [2]. It has many aspects and can be formulated in different ways. In our paper, we consider the case when the model equation structure is known a priori from "the first principles". It reads

$$d\mathbf{y}/dt = \mathbf{f}(\mathbf{y}, \mathbf{c}), \quad (1)$$

where $\mathbf{y}$ is D-dimensional state vector, $\mathbf{c}$ is P-dimensional parameter vector. The task is *to estimate the unknown parameters* $c_1, \ldots, c_P$ from a time series of observable $\{\eta_1, \ldots, \eta_N\}$, where η is assumed to be a function of state vector $\mathbf{y}$, N is a time series length. Let us consider the case of scalar η, which is most typical and complicated one. Such a formulation has been considered in a number of works for differential equations , and maps. In practice, it is encountered in chemical kinetics, laser physics [3], electric engineering, cell biology, etc.

One must reconstruct all D components of $\{\mathbf{y}_i\}$ from a scalar time series $\{\eta_i\}$ to construct the so-called standard model [1]. However, for a model structure specified from the first principles, some of state variables cannot often be measured or obtained from observed data. Such variables are usually called "hidden". The presence of hidden variables requires more sophisticated approaches for parameter estimation. Usually, maximal likelihood principle is appealed to, but practically it reduces to the least-squares method. In the case we consider, the problem is formalised as follows. One searches for initial conditions $\mathbf{s}$ and parameters $\mathbf{c}$ which provide the smallest least-squares difference between the appropriate components of a model orbit $\mathbf{y}(t)$ and observed data $\overline{\mathbf{y}}^l$. The sum of errors (2) involves only l non-hidden variables:

$$S(\mathbf{s}, \mathbf{c}) = \sum_{i=1}^{N} \left[\mathbf{y}^l(t_i, \mathbf{s}, \mathbf{c}) - \overline{\mathbf{y}}_i^l\right]^2 = \min, \quad (2)$$

where $\overline{\mathbf{y}}_i^l$ are observed vectors, $\mathbf{y}^l(t_i, \mathbf{s}, \mathbf{c})$ are l-dimensional vectors consisting of the corresponding model state variables. (2) is minimised with the aid of iterative algorithms for some "starting guesses" for $\mathbf{s}$ and $\mathbf{c}$.

For chaotic time series, a model trajectory is very sensitive to initial conditions. Therefore, the cost function (2) is very complex for large N and has a lot of local minima. Thus, the global minimum is unlikely to be found with arbitrary starting guesses. In order to overcome this difficulty, a special method — multiple shooting approach (Bock's algorithm) — was proposed [4]. Later, it was noticed [5] that it also encounters significant difficulties and additional efforts are necessary to succeed, although systematic investigation of this problem has not been still performed. In this work, we develop special measures to quantify the performance of different parameter estimation techniques. With their aid, we compare different versions of multiple shooting approach (sec. 2). By considering noisy time series of exemplary chaotic systems, we demonstrate that a modified Bock's algorithm allowing discontinuity of a model trajectory is the most efficient.

2. Parameter estimation methods for hidden variable case

2.1. Bock's algorithm

The idea of the multiple shooting approach is to divide the entire series N into L segments (n is the length of a segment, $N = Ln$) and initial conditions for each of them $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_L$ are considered as additional arguments of S (as quantities to be estimated):

$$S(\mathbf{s}_1, \ldots, \mathbf{s}_L, \mathbf{c}) = \sum_{i=1}^{L} \sum_{j=1}^{n} \left[\mathbf{y}^l\left(t_j, \mathbf{s}_i, \mathbf{c}\right) - \overline{\mathbf{y}}^l_{(i-1)n+j}\right]^2 = \min, \quad (3)$$

To avoid a great number of free estimated quantities, that increases variances of the estimates, one imposes a constraint of model trajectory continuity over the entire observed interval:

$$\mathbf{y}\left(t_{(i-1)n+1}, \mathbf{s}_i, \mathbf{c}\right) = s_{i+1},\ i = 1, \ldots, L-1. \tag{4}$$

Minimisation of (3) under the constraints (4) is the problem of constrained multidimensional optimisation. For arbitrarily chosen starting guesses for parameters and initial conditions, the model trajectory consists of L "disconnected" pieces. However, it becomes "more continuous" gradually, after each iteration of the minimisation procedure.

It was claimed [4] that Bock's algorithm does not require "genuine" starting guesses. Meanwhile, experience shows that *this is not typically the case*, since the condition (4) is very strong. Therefore, often only local minima of (3) can be found. One way to overcome this disadvantage is to divide the original series into a number of shorter series and to apply the algorithm to each of them separately, with the parameter estimators being obtained as average values. Such an approach is called "piecewise" or "segmentation" technique [5]. Though it gives some advantages, we suppose that the better approach is possible.

2.2. Modified Bock's algorithm

It is known from statistical theory, that the use of the entire time series in maximum likelihood estimation is preferable for obtaining unbiased estimators than segmentation approach. So, we suggest to pay attention to a modification of Bock's algorithm that has been already applied in [3] for non-chaotic signals consisting of a number of independent shot realisations as a technique for "multiple experiment approach" problem solution. It was also briefly mentioned in [5]. The idea is to refuse the constraints (4) for several $(\nu - 1)$ time instants holding the same parameter values $\mathbf{c}$ for the entire time series. So, the initial conditions for the ν time instants, including the first one, become independent quantities to be estimated. We choose these instants equidistantly within the time series. Such an approach involves two adjustable parameters: the number of segments ν and the number of subsegments within each segment L, $(N = \nu L n)$. Subsegments are required to apply Bock's algorithm within each of the ν segments.

The modified approach is not widely applied so far, even though it should have a number of advantages. The fact that a final model trajectory is discontinuous is not an indication that the model is "bad" but weakening of the constraints (4) may help to find global minimum and reasonable model when "strict" Bock's algorithm is not feasible.

3. Comparative study in numerical experiment

3.1. Comparison technique

We normalise starting guesses so that the centre of a diagram corresponds to genuine guesses, i.e. to the true values of parameters c_i^0. The normalised starting guesses are denoted $b_i = \left(c_i - c_i^0\right)/c_i^0$. We compare the methods using gray-scale "convergence diagrams" on the planes of starting guesses for parameters b_{i_1}, b_{i_2} (Fig. 1). White points denote starting guesses for which the global minimum is achieved, i.e. quite accurate estimates are obtained. Gray colour means starting guesses from which minimisation procedure converges to any local minimum, with darker one corresponding to local minimum situated further from the true values. The size of white area on the diagrams quantifies the estimation method's performance. The broader this area, the better the method. Also we suggest an integral measure which is relative number μ of white points within a circle of radius r. The larger μ (for a given r), the better the method. We denote r_μ the maximum value of the circle radius corresponding to the relative ratio of white points equal to μ. Here, we use mainly the value of r_{100}, which is the radius of "100% convergence" to global minimum.

Below, we consider the case of three unknown parameters. Therefore, 3-dimensional diagrams for all three starting guesses for parameters would contain complete information about the method. We use 2-dimensional projections for simplicity of illustration taking into account that they lead to the same qualitative conclusions about the methods' inferiority/superiority.

3.2. Identification of the Lorenz system

As the first test system for investigation of the performance of different parameter estimation techniques in case of long chaotic series and different starting guesses, we choose the Lorenz system

$$\begin{aligned} \dot{y}_1 &= c_1\left(y_2 - y_1\right) \\ \dot{y}_2 &= -y_2 + y_1\left(c_3 - y_3\right) \\ \dot{y}_3 &= -c_2 y_3 + y_1 y_2 \end{aligned} \tag{5}$$

with parameters $c_1 = 10$, $c_2 = 8/3$, $c_3 = 46$ corresponding to a chaotic regime. The equations are integrated with the fourth-order Runge-Kutta technique with sampling interval 0.002. An observed scalar time series is a realisation of the variable y_1 corrupted with additive Gaussian noise: $\eta = y_1 + \xi$. The variables y_2 and y_3 are regarded hidden.

Since the choice of genuine starting guesses for the values of y_2 and y_3 is unrealistic, we use the observable values as starting guesses for all state variables $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_L, \ldots, \mathbf{s}_{\nu L}$. Even though such a choice is not the best possible, it is simple and efficient [4]. To minimise the function (3) the generalised Gauss-Newton method is used [4].

Convergence of the original Bock's algorithm and the modified method to global minimum is illustrated in fig. 1a,b. These results correspond to the time series length for which the Bock's approach exhibits the best performance (the broadest convergence region). Only the section

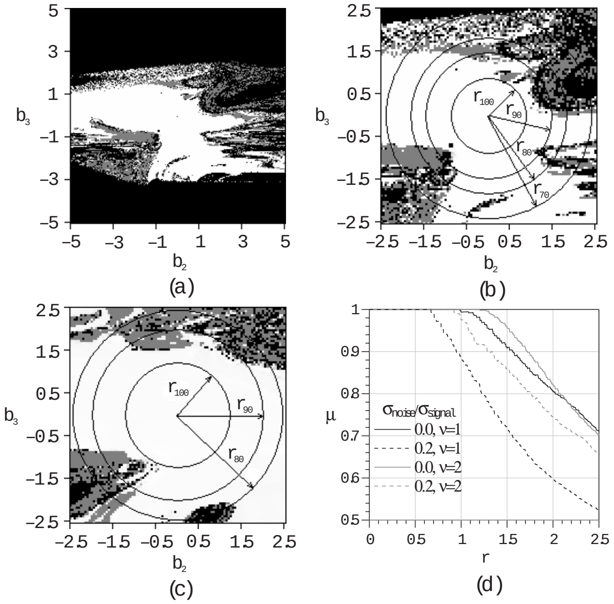

Figure 1: The plane of normalised starting guesses for parameters of the Lorenz system (section with the plane $b_1 = 0$). (a) Bock's algorithm with $L = 30$, $n = 35$; (b) is a magnification of (a); (c) the modified method with $L = 15$, $n = 35$, $\nu = 2$; (d) the dependence $\mu(r)$ for Bock's algorithm (black) and the modified method (gray) at different noise levels.

of starting guesses space with the plane $b_1 = 0$ is shown since unlucky choice of b_1 is not so crucial as the choice of b_2, b_3. It can be seen that the area of 100% convergence of Bock's algorithm is broad and the radius r_{100} is greater than 1.0, so relative deviations of starting guesses from true values (let us call them errors in starting guesses) may exceed 100%. There is also a wide area which is very distant from global minimum but allows to find global minimum (Fig. 1a). However, the modified method allows larger errors in starting guesses as it can be seen from comparison of Fig. 1b and Fig. 1c. The values of r_{100}, r_{90}, and r_{80} are greater for the modified method and the white area is wider.

The value of $\mu(r)$ for different noise levels is shown in Fig. 1d. The performance of both methods remains almost unchanged for moderate noise. The horizontal line of 100% convergence ($\mu = 1$) becomes shorter but not significantly: in a noise-free setting its length is 1.2 for the modified approach and 1.1 for Bock's method, while for 20% noise-to-signal ratio (ratio of rms amplitudes) it is 0.9 and 0.7, respectively.

From Fig. 2a where the 100% convergence radius is shown versus time series length N for different number of segments ν is obvious that modified approach has advantages for longer time series. The number of subsegments L has been selected to make r_{100} as large as possible. For small N, the amount of data is insufficient to "average out" the noise influence, while for large N, the small initial per-

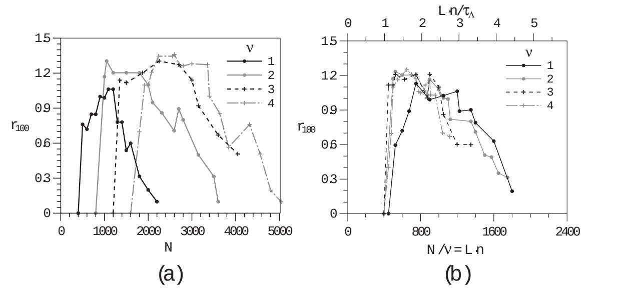

Figure 2: The dependence of 100% convergence radius r_{100} on different factors: (a) on the entire time series length N for different numbers of discontinuity points allowed; (b) on the continuity segment length Ln.

turbations reaches the magnitude comparable to the size of the attractor during time interval $\tau_\Lambda = 1/\lambda_1$ that leads to complication of the cost function "relief". The curves for larger ν attain larger values of r_{100}, i.e. the modified method is more efficient than the original Bock's algorithm. Those curves correspond also to larger values of N, therefore they are located closer to the right-hand side of the panel. Furthermore, the range of time lengths within which the modified method is "100% convergent" increases with the number of discontinuity points ν, so the curves for greater ν are "wider".

The investigation reveals (Fig. 2b) that the optimal value of segment length Ln is connected with Lyapunov time τ_Λ. Optimal time series lengths correspond to 1–2 Lyapunov times, see the upper horizontal axis in Fig. 2b. It is explained as follows. The success of estimation depends on the segment length Ln (over which small initial perturbations of the model orbit should not increase too strongly, so Ln should not be very large) and also on the number $P + \nu D$ of free parameters to be estimated which should not be very large since in very high-dimensional space relief of the cost function may become very complicated also, i.e. Ln should not be very small. As a consequence, there exists some intermediate optimal value of Ln related via a certain proportionality constant to the characteristic time scale τ_Λ of the divergence of nearby model trajectories.

Similar results have been obtained from time series generated at different initials conditions, from time series of the variable y_2, and from time series of y_1 generated at a different set of "true" parameter values $c_1 = 10$, $c_2 = 8/3$, $c_3 = 28$ that is known as a "classical" chaotical set for the Lorenz system.

3.3. Identification of Rössler system

In order to check whether our results hold for other systems, we perform the same investigation for the Rössler's system.

$$\dot{y}_1 = -y_2 - y_3$$
$$\dot{y}_2 = y_1 + c_1 y_2 \qquad (6)$$

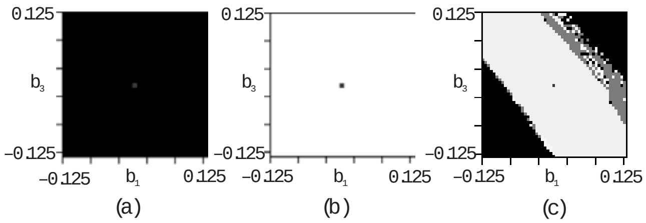

Figure 3: The plane of normalised starting guesses for parameters of the Rössler system (section with the plane $b_2 = 0$) illustrating convergence of the original Bock's algorithm: (a) all starting guesses for the hidden variables are equal to simultaneous observable values; (b) genuine starting guesses; (c) starting guesses are obtained via the time shift of the observed time series by a quarter of basic period.

$$\dot{y}_3 = c_2 + y_3 (y_1 - c_3),$$

with parameters $c_1 = 0.2$, $c_2 = 0.15$, $c_3 = 10$, that corresponds to a chaotic regime. The equations (6) are integrated with 4-th order Runge-Kutta technique with sampling interval 0.01. The variable y_1 is used as an observable both in a noise free setting and corrupted with additive Gaussian white noise.

We have chosen this system as an object since the "shape" of its attractor differs from the Lorenz one. The simultaneous values of Lorenz system y_1 and y_2 variables are relatively close to each other. The dynamics on the Rössler attractor is a rotation about a single unstable fixed point (in projection onto the plane $y_3 = 0$). So that the variables y_1 and y_2 are shifted in time by a quarter of the rotation period which is the main time scale here.

Due to such relationships between the state variables, the choice of starting guesses for the hidden variables equal to the simultaneous observable value is more or less appropriate for the Lorenz system (as we have shown above) but leads to unsuccessful results of parameter estimation in the Rössler system using any of the estimation techniques considered. In Fig. 3a it is shown that $r_{100} = 0$, i.e. one cannot find the global minimum for such a choice of starting guesses for the hidden variables at all. Quite good results are achieved if one uses genuine starting guesses for the hidden variables Fig. 3b). To develop "good" and realistic starting guesses is also possible if one takes into account the knowledge about character of the original dynamics which can be gained by studying model dynamics. Namely, for the Rössler system it is relevant to take the observed time series shifted by a quarter of basic period as a starting guess for the variable y_2 and zero as a starting guess for y_3 because due to attractor features this variable is close to zero most of the time (Fig. 3c).

For starting guesses we proposed, the results of investigation are similar to that presented above for the Lorenz system.

4. Conclusions

We performed comparison of performance of two methods for estimation of parameters (identification) of dynamical systems from chaotic time series in the case of hidden variables using specially developed quantitative measure. Both methods rely upon the multiple shooting idea. The original Bock's algorithm is shown to be less efficient than its modified version which allows a model orbit to be discontinuous in several points within an observation interval.

The length of a time series and the number of its segments are shown to have significant influence upon the estimation results, and the choice of starting guesses for the hidden variables is quite important too. The chances for accurate estimation rise with time series length if the number of allowable points of model trajectory discontinuity is also increased. The optimal length of a continuity segment is close to Lyapunov time for long chaotic time series.

The effect of measurement noise is shown to be not dramatical for both methods, even if noise-to-signal ratio is as high as 20% in rms amplitude.

Acknowledgments

The work is supported by Russian Foundation for Basic Research (grant 05-02-16305), CRDF (REC-006), the President of Russia (MK–1067.2004.2).

References

[1] Chaos and Its Reconstructions, eds. Gouesbet G., Meunier-Guttin-Cluzel S., Menard O. (Nova Science Publishers, New York, 2003).

[2] L. Ljung. System identification. Theory for the User. Moscow, 1991.

[3] W. Horbelt and J. Timmer, M. J. Bunner, R. Meucci and M. Ciofini. Identifying physical properties of a CO2 laser by dynamical modeling of measured time series. Phys. Rev. E, 2001, vol.64, 016222.

[4] Baake E., Baake M., Bock H.G., and Briggs K.M. Fitting ordinary differential equations to chaotic data // Phys. Rev. A, 1992, V. 45, No. 8, P. 5524-5529.

[5] V.F. Pisarenko, D. Sornette, Statistical methods of parameter estimation for deterministically chaotic time series. Phys. Rev. E, 2004. V. 69. 036122. J.P. Schloeder, Numerische Methoden zur