An Algorithm for Clustering of the Phase Trajectory of a Dynamic System

Vadim Strijov, Vsevolod Shakin
Computer Center of the Russian Academy of Sciences
E-mail: strijov@ccas.ru, shakin@ccas.ru

Abstract: This paper describes an approach to quantitative analysis of multivariate dynamic system in phase space. The system is used as mathematical model for various living systems. The model is used in various applications. One of the related problems is to represent a phase trajectory as a sequence of clusters to classify the system's state.

The algorithm for partitioning a phase trajectory into clusters is presented. Input data for the algorithm is a data matrix which is corresponds to a set of sequential samples of the given phase trajectory. Optional parameters are dimension of the space in which the clusters lie, and phase trajectory noise variance. The algorithm results in a tree-like graph. The graph nodes contain given phase trajectory clusters and might be used for system's state classification.

Phase trajectory of a dynamic system with Lorenz attractor is considered as a test problem to demonstrate the approach. The initial phase trajectory lies in 3D-space. It was projected into N-dimensional space and distorted with non-correlated additive Gaussian noise. The given phase trajectory was partitioned into clusters using the described algorithm. The clusters make a tree T. The root of the tree corresponds to the phase trajectory that lies in r-dimensional space R^r. The next level of the tree consists of cluster nodes that lie in $(r-1)$-dimensional space, etc., up to the last level that corresponds to one-dimensional cluster nodes. An example of the tree is presented.

The algorithm was examined with various test trajectories. It is currently being applied to assess temporal dynamics of social and economical systems under extreme conditions.

Keywords: Phase trajectory, dynamic system, cluster.

1 Introduction

Life of a biosystem can be described as a vector function of time or a phase trajectory. Examine an autonomous dynamic system that is described as a set of n ordinary differential equations

$$\mathbf{x}' = F(\mathbf{x})$$

with initial conditions $\mathbf{x}(0) = \mathbf{x_0}$, where $\mathbf{x} = \mathbf{x}(t)$, $t \in [0, T]$, $\mathbf{x} \in R^n$ and R^n is phase space of the system. Trace of the vector function $\mathbf{x} = \mathbf{x}(t)$ is a line in the phase space, and it is dubbed as dynamic system's phase trajectory.

The model is used in various applications [1], [3], [4]. One of the related problems is to represent a phase trajectory as a sequence of clusters to classify the system's state. Clustering procedure depends on shape of the trajectory and on requirements such as 1) dimensions of the subspace $R^r \subset R^n$, which contains a trajectory cluster or 2) size of a parallelepiped or an ellipse, which contains a cluster and lies in the subspace R^r, $r < n$ [1].

2 Problem

Let us consider a phase trajectory of a dynamic system with Lorenz attractor as a test problem. The phase trajectory corresponds to solution of this ordinary differential equations set

$$\begin{cases} x_1'(t) = -3(x_1(t) - x_2(t)) \\ x_2'(t) = -x_1(t)x_3(t) + 26.5x_1(t) - x_2(t) \\ x_3'(t) = x_1(t)x_2(t) - x_3(t) \end{cases}$$

with initial conditions $x_1(0) = x_3(0) = 0$, $x_2(0) = 1$.
The phase trajectory lies in R^3 space. It was projected into an R^n space and non-correlated Gaussian noise with variance of σ^2 was added.

Values of the function $\mathbf{x} = \mathbf{x}(t)$ were represented as phase vector sequence $\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_m}$, where $\mathbf{x_i} =< x_1, x_2, ..., x_n >\in R^n$ and $t = 0, \tau, 2\tau, 3\tau, ..., (m-1)\tau$. Let

$$A = (x_{i,j})_{i,j=1}^{m,n}$$

be $(m \times n)$-matrix, n being the number of parameters, or dimension of a space, which contains vector $\mathbf{x}$ and m being the number of the vectors.

Let S^p — image of A by linear projection f

$$f : (A - B) \to S^p,$$

where:
A is finite well-ordered set of vectors $A = \{\mathbf{x_i} : \mathbf{x_i} \in R^n, i = 1, ..., m\}$, S^p is finite well-ordered set of vectors $S^p = \{\mathbf{s_i} : \mathbf{s_i} \in R^p, i = 1, ..., m\}$, $B = (b_{i,j})_{i,j=1}^{m,n}$ is residue matrix. The matrix B approximates S^p to $f^{-1}(A)$.

With linear projection f each vector $\mathbf{x_i}$ is projected accordingly to the vector $\mathbf{s_i}$, $i = 1, ..., m$. Let the set S^p consists of well-ordered subsets S^p_ξ, where

$$S^p = \{S^p_\xi : S^p_\xi \subseteq R^{k \times p}, \xi = 1, ..., k\}$$

under conditions $\bigcup\limits_{\xi=1}^{k} S^p_\xi = S^p$ and $\bigcap\limits_{\xi=1}^{k} S^p_\xi = \emptyset$. S^p is called the set of the phase trajectory A clusters that lies in a space R^p.

The following problems one to be solved:
1) Representing the matrix A as:

$$A \to S^p + B$$

where
$S^p = \bigcup\limits_{\xi=1}^{k} S^p_\xi$ is union of the k clusters of the phase trajectory A, p is given value ($p \in Z$), under condition $p \leq r$, where $r = \text{rank}(A - B)$.
2) Constructing:

$$T = \bigcup_{p=1}^{r} \left(\bigcup_{\xi=1}^{k} S^p_\xi + B \right)$$

where T is tree that contains all the sets S^p, which are placed in subsets $R^1, ..., R^r$, $r = \text{rank}(A - B)$.

3 Algorithm

1. Dimension evaluation: To obtain the effective dimension[2] or rank of the matrix $(A - B)$, we are using the singular value decomposition

$$A = UWV^T.$$

The matrixes U and V are orthogonal; the matrix W contains decrescent singular values on its diagonal

$$W = \text{diag}(w_1 \geq w_2 \geq ... \geq w_n \geq 0).$$

Rank of the matrix $(A - B)$ corresponds to the index r of the singular value w_r under condition

$$\frac{w_r^2}{\sqrt{m}} \leq \delta$$

where δ is a given value, related to the additive Gaussian noise. Two statements are held:
1. For the space dimension $R^p : p = r$, $r = \text{rank}(A - B)$, the phase trajectory, which is described with the matrix A will be represented as a singleton cluster

$$A = S_1^{p=r} + B.$$

2. A phase trajectory cluster that is lies in p-dimensional space can be represented as union of clusters that are lie in q-dimensional space, $1 \leq q \leq p \leq r$, $r = \text{rank}(A - B)$.

2. Clusterisation: Consider clusterisation algorithm in space R^1. For the first iteration let each cluster S^1_ξ be containing only one vector $\mathbf{x_i}$, $\mathbf{x_i} \in A$ and $\xi = i = 1, ..., m$. The next iterations are done in the following way:
Let $S^1_\xi = \{\mathbf{x}_\zeta, \zeta = 1, ..., l\}$. The median of the cluster S^1_ξ will be vector $\mathbf{x}_{\frac{\mathbf{l}}{\mathbf{2}}}$ if l is even or $\mathbf{x}_{\frac{\mathbf{l+1}}{\mathbf{2}}}$ if l is odd.

$$\begin{cases} \text{med}(S^1_\xi) = \mathbf{x}_{\frac{\mathbf{l}}{\mathbf{2}}}, & \text{if } \text{mod}_2(l) = 0 \\ \text{med}(S^1_\xi) = \mathbf{x}_{\frac{\mathbf{l+1}}{\mathbf{2}}}, & \text{if } \text{mod}_2(l) = 1 \end{cases}$$

For $\xi = 2, ..., k-1$ clusters $S^1_{\xi-1}, S^1_\xi, S^1_{\xi+1}$ are taken. If the distance (in given metric) between $\mathrm{med}(S^1_{\xi-1})$ and $\mathrm{med}(S^1_\xi)$ is less than the distance between $\mathrm{med}(S^1_\xi)$ and $\mathrm{med}(S^1_{\xi+1})$, then the cluster S^1_ξ is joined to the cluster $S^1_{\xi-1}$ right side, otherwise the cluster S^1_ξ is joined to the cluster $S^1_{\xi+1}$ left side. We can join the cluster to its left of right neighbours if and only if the dimension of the space that contain both clusters is not greater then 1. Iteration process is terminated when there will be no clusters S^1_ξ, $\xi = 1, ..., k$ to join to their right or left neighborhood clusters.

Consider clusterisation algorithm in space R^n. Let the set of the clusters S^{p-1} consists of well-ordered subsets S^{p-1}_ξ

$$S^p = \{S^p_\xi : S^p_\xi \subseteq R^{k \times p}, \xi = 1, ..., k\}.$$

Then we can join the cluster S^{p-1}_ξ to its left or right neighbour cluster in p-dimensional space like we did it 1-dimensional space.

4 Programming

1. Solve the Lorenz ordinary differential equations system:

```
la=evaluate((x(t), y(t), z(t))
        odeSolve((x'(t) == -3 (x(t) - y(t)),
                y'(t) == -x(t) z(t) + 26.5 x(t) - y(t),
                z'(t) == x(t) y(t) - z(t),
                x(0) == z(0) == 0, y(0) == 1),
                (x, y, z),
                (t, 0, 40)
        )
)
```

2. Represent the given phase trajectory as a matrix:

```
tau = 0.01;
A=tble(la,(t,0,40,tau));
```

3. Project the matrix to n-dimensional space and add a noise:

```
n = rnd(integer,(4,10));
C= rndTable((n),(n));
B= makeNoise((n,length(A),sigma2);
A=freeRotate(A,C)+B;
```

4. Evaluate the factual dimensions of the matrix A:

```
(u,w,v)=singularValues(A);
r=findItemIndex(min(w, w2>sigma2));
r=3
```

5. Evaluate clusters for the sets $S^1, ..., S^r$ and construct the tree T:

```
T=();
for(i=1,i<=r,i++,
        T=append(
                T,
                evaluateClusters(A,i)
        );
);
```

The results are placed in structured array T.

5 Results

The initial phase trajectory projection to a plane surface is shown on Fig. 1:

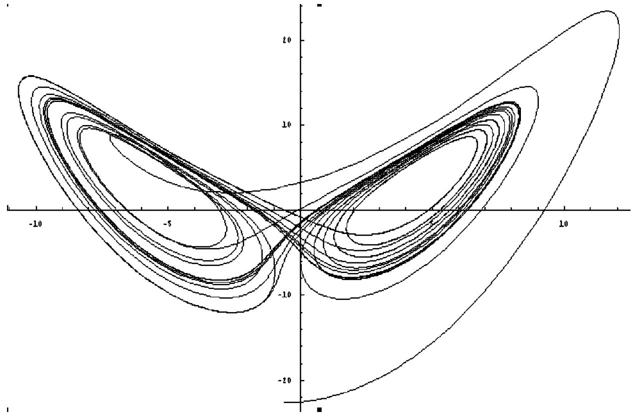

Fig 1. The initial phase trajectory projection to a plane surface

The given phase trajectory was partitioned into clusters by the described algorithm. Some clusters of the phase trajectory are shown on the Fig. 2.

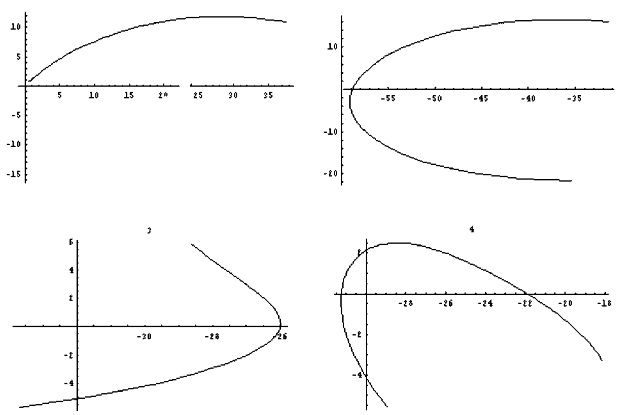

Fig. 2 Some clusters of the phase trajectory

To display cluster relations, we make set of the clusters if they lie in the same plane with exactness to $\pm\frac{\sigma^2}{2}$. (on the Fig. 3 only one set is shown).

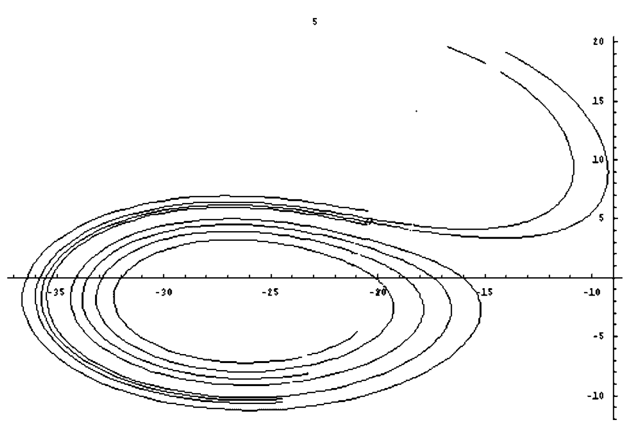

Fig. 3 Set of the clusters that lie in the same plane

The clusters make a tree T. The root of the tree corresponds to the phase trajectory that lies in r-dimensional space R^r. The next level of the tree consists of cluster nodes that lie in $(r-1)$-dimensional space, et cetera until the last level that corresponds to 1-dimensional cluster nodes. An example of the tree is shown on the Fig . 4.

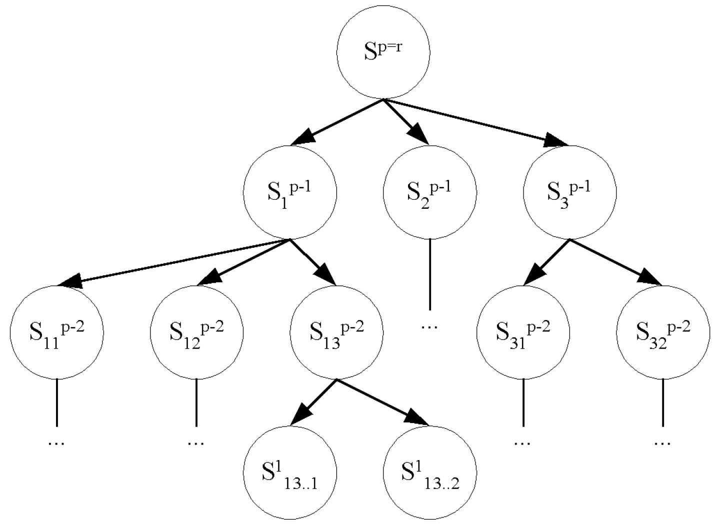

Fig . 4 Example of the tree T

6 Conclusion

In this paper the algorithm for partitioning a phase trajectory into clusters was developed. Input data for the algorithm is matrix A, which is corresponds to given phase trajectory. Optional parameters are dimension of the space in which the clusters lie, and phase trajectory noise variance. The algorithm result is tree-like graph. The graph nodes contain given phase trajectory clusters and might be used for system's state classification problem. The algorithm was examined with various test trajectories. It is currently being applied to assess temporal dynamics of social and economical systems under extreme conditions.

7 Acknowledgement

Results presented in this contribution have been obtained with financial support of The Russian Foundation for Basic Research (RFBR), within Grant 00-01-00197 "Criteria of Life Quality and Sustainable Development for Social-Economical Systems under Extreme Conditions". Authors are greatly indebted to Prof. Tihomir Hunjak from University of Zagreb, and to Dr. Vlasta Molak, GAIA UNLIMITED, Inc., Cincinnati, Ohio, USA for the very possibility to attend the meeting.

References

[1] Strijov V.V., Shakin V.V. Software for phase trajectory analysis. Mathematical Methods of Pattern Optimization (MMPO-9), Abstracts of the Russian conference. Scientific edition - RAS, CC, RFBR, Moscow, 1999. - P. 227-230.

[2] Classification and Space Dimensionality Reduction /ed. by Aivasian S.A., M.: Finance and statistics, 1989 - P. 332-419.

[3] Shakin V.V. The models, the measurements and identification for the biosystems under the extreme conditions // Proceedings of the 7th International IMEKO TC-13 Conference on Measurements in Clinical Medicine "MODEL BASED BIOMEASUREMENTS"/Stara Lesna, Slovak Republic, Sept.6-9, 1995, pp.37-39.

[4] Shakin V.V. Quality of Life and Sustainable Development for Living Systems under Extreme Conditions. Abstracts. 8th International Conference on Operational Research KOI 2000 Rovinj, Croatia, September 27 - 29, 2000.