# Fazinės erdvės rekonstrukcija ir laikinių sekų prognozė

# State space reconstruction issues and times series prediction

Feliksas Kuliešius, Andrius Lengvinas, Anželika Pavlova
Vilniaus universitetas, Fizikos fakultetas, Saulėtekio al. 9, LT-10222 Vilnius
feliksas.kuliesius@ff.vu.lt

Network flows [1] discussed as a complicated physical systems such as multi-filamentation in optical beams, fiber solitons and ocean rogue waves [2], financial markets [3] or social behaviour [4], are commonly subjected to chaotic regularities. Chaotic behaviour refers to a behaviour which, being irregular is generated by an underlying deterministic process and, therefore, is potentially controllable.

Usually, due to the experimental limitations only one dimensional data is available for chaotic physical systems which have higher dimensionality. Very often such phenomena can't be described analytically or even solved numerically. In such a case the only possibility of statistical analysis of experimental data – time series, remains. The powerful tool for time series prediction is state space reconstruction from one dimension experimental series. Suppose the observed scalar time series is $x(1),\ x(2),\ \ldots,\ x(N)$. According to Taken's theorem, an embedding of the chaotic attractor can be obtained by constructing the delay vectors $X(n) = \left[x(n),\ x(n-\tau),\ \ldots,\ x(n-(m-1)\tau)\right]^T$, where $n = (m-1)\tau + 1,\ (m-1)\tau + 2,\ \ldots,\ N$, and $m$ is an embedding dimension and $\tau$ is a time delay.

Deterministic predictions are based on continuous mapping between the current state and the future state, meaning that if the state at time $t$ is similar to the current state $X(n)$, then the state at time $t+1$ will also be close to the future state $X(n+1)$.

The local linear prediction model suggests that the prediction is a linear superposition of the $m$ elements of a delay vector, that is

$$X_{pr}(n+1) = a_0 + \sum_{i=1}^{m} a_i x(n-(i-1)\tau) = \vec{A}Y(n),$$

where $\vec{A} = \left[a_0,\ a_1,\ \ldots,\ a_m\right]$, $Y(n) = \left[1,\ X(n)^T\right]^T$.

$\vec{A}$ can be obtained from $\vec{A}\hat{B} = \vec{D}$, where $\hat{B}$ is a matrix, $i$ th column of which is composed of $Y(n_i)$, i.e. $Y(n_i) = \left[1,\ x(n_i),\ x(n_i-\tau),\ \ldots,\ x(n_i-(m-1)\tau\right]^T$, $\vec{D} = \left[x(n_1+1),\ x(n_2+1),\ \ldots,\ x(n_k+1)\right]$. Then the coefficient vector $\vec{A}$ can be found from $\vec{A} = \vec{D}\hat{B}^{-1}$.

The main problem related to the state space reconstruction is the uncertainty of parameters, namely time delay $\tau$ and reconstruction dimension $m$, necessary for it. The different techniques used to determine before mentioned parameters were analysed in the present work. Delay is obtained using the autocorrelation, mutual information and time window while embedding dimension is found via false nearest neighbors. Additionally, the new approach for evaluation the delay $\tau$ based on dimension variance analysis has been introduced.

Widely used way of assessing $\tau$ by means of autocorrelation $R_\tau = \frac{\sum_{j=\tau+1}^{n}(x_j-\bar{x})(x_{j-\tau}-\bar{x})}{\sum_{j=1}^{n}(x_j-\bar{x})^2}$ may result in bad value of $\tau$ for nonlinear systems. The delay, obtained by matching the first minimum of mutual information $I(x_t, x_{t+\tau}) = \sum_{t,\,t+\tau} P(x_t, x_{t+\tau}) \ln \frac{P(x_t, x_{t+\tau})}{P(x_t)P(x_{t+\tau})}$ is more prospective since it encounters higher correlation terms. Some authors suggest attitude named Time window, where $\tau$ is defined as $\tau = t_w/m$, where $t_w$ is time between time series peaks, interpreted as the mean time between two consecutive visits to a local neighborhood on the attractor. Examining the pointwise dimension $M_p(i,\tau) = \lim_{\epsilon\to 0} \frac{\log\frac{1}{N}\sum_{j=1}^{N}\theta(\epsilon - r_{i,j})}{\log(\epsilon)}$ [5] it is possible to conclude that the condition of the best attractor reconstruction in the state space can be defined as a minimum point of $M_p(i,\tau)$ variance: $f(\tau) = \frac{1}{N}\sum_{i=1}^{N}(M_p(i,\tau) - \bar{M}_p)^2$. Above, $\theta$ is the Heaviside function, $r_{i,j}$ – Euclidean distance between state space points, and $\epsilon$ – an arbitrarily small radius related to the point. Different approaches were evaluated by the possibility to reconstruct attractor adequate to original one ant to predict original time series. In this work, well known chaotic systems – Mackey-Glass, Lorenz, Rössler have been analyzed. For Mackey-Glass sequences the best results were found by Time windows approach – up to 2000 future point forecast succeeded with $\tau$=17. Other series exposed better when mutual information was used. It is rational to suppose, that the results reflect the model structure - Mackey-Glass system possesses a well-defined time-delay parameter, while other – are described by a sets of three nonlinear coupled differential equations without any explicit time delay. The dimension variance technique is suitable for all cases; it is more stable and yields acceptable results for all process under investigation, though not optimal (e.g., 600 point prediction for Mackey-Glass).

*Keywords: chaos, state space reconstruction, time series prediction*

**References**

[1] W, Zhang, Z. Wu, and G.Yang, Comp. Phys.Commun., **161**, 129, (2004)
[2] S. Birkholz, C. Bree, A. Demircan, and G. Steinmeyer, Phys. Rev. Lett., **114**, 213901 (2015)
[3] E.W. Saad, D.V. Prokhorov, D.C. Wunsch, IEEE Trans. Neural Netw., **9**, 1456 (1998)
[4] F. Shafee, xxx.lanl.gov, arXiv:cond-mat/0401191, p. 1-15, 2004
[5] E. Ott, Chaos in dynamical Systems, (Cambridge, 2002)