On the Way to a Significance Assessment for Recurrence Plot Based Measures

N. Marwan, M. C. Romano, M. Thiel, J. Kurths
Institute of Physics, University of Potsdam, D-14415 Potsdam, Germany

The method of recurrence plots (RP) was firstly introduced to visualize the time dependent behavior of the dynamics of systems x_i, which can be pictured as a trajectory in the phase space (Eckmann 1987). It represents the recurrence of the phase space trajectory to a certain state. The main step of this visualization is the calculation of the $N \times N$-matrix

$$\boldsymbol{RP}(i,j) = \Theta(\varepsilon - \|\vec{x}_i - \vec{x}_j\|)$$

where ε_i is a cut-off distance and $\Theta(\cdot)$ is the Heaviside function.

The recurrence plot exhibits characteristic large-scale and small-scale patterns which are caused by typical dynamical behavior (Eckmann 1987, Webber & Zbilut 1994), e. g. diagonals (similar local evolution of different parts of the trajectory) or horizontal and vertical black lines (state does not change for some time).

Zbilut and Webber have recently developed the recurrence quantification analysis (RQA) to quantify a RP (Zbilut & Webber 1992, Webber & Zbilut 1994). They defined measures using the recurrence point density and diagonal structures in the recurrence plot, the *recurrence rate* RR (density of recurrence points), the *determinism* DET (ratio of recurrence points forming diagonal structures to all recurrence points), the *maximal length of diagonal structures* L_{max} (or the their *averaged length* $\langle L \rangle$), the *entropy* ENT (of the distribution of the diagonal lengths) and the *trend* TR (paling in the RP). A computation of these measures in small windows moving along the main diagonal of the RP obtains time dependent behavior of these variables and thus, transitions in the time series (e. g. Trulla et al. 1996).

These RQA measures became rather popular and are broadly applied in various scientific disciplines. However, it seems that nobody has a critical look at the reliability or significance of his results. This is due to the fact, that the structures occuring in recurrence plots are not yet understood to the explicit detail. Finally, a significance theory for the RQA is not trivial and not yet developed. The development of the significance criterias for the RQA should be the major task in the future. Here we present a first step in this direction.

In the following, we show a snapshot of our last week work, where we tried to get the theoretical distributions for the RQA measures. These distributions are necessary for the significance assessment of the RQA. This work was dedicated to Gaussian white noise, maximum norm and embeddings of m=1. The first approach uses the assumption, that the recurrence points are independent, which is in deed not the case. This invalid assumption leads to uncorrect distributions for RR and DET. Now we are looking for a better approach.

Recurrence Rate (RR)

The probability of the occurence of a state at i is

$$p(x_i = x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-x^2}{2\sigma^2}}.$$

The probability of the occurence of a recurrence, i. e. a black point, anywhere is

$$p_\bullet(\varepsilon) = p(\|x_i - x_{i+n}\| \le \varepsilon) = \int\limits_{x=-\infty}^{\infty} \int\limits_{y=x-\varepsilon}^{x+\varepsilon} p(x_i = x; x_{i+n} = y)\, dy\, dx$$

and, thus, of a white point

$$p_\circ(\varepsilon) = p(\|x_i - x_{i+n}\| > \varepsilon) = 1 - p_\bullet.$$

For Gaussian white noise with standard deviation σ, $p_\bullet$ is:

$$p_\bullet(\varepsilon) = \operatorname{erf}\left(\frac{\varepsilon}{2\sigma}\right),$$

which is instantly the recurrence rate for Gaussian white noise:

$$RR(\varepsilon) = \operatorname{erf}\left(\frac{\varepsilon}{2\sigma}\right).$$

The distribution of an observed recurrence rate ϱ can be determined with the Binomial distribution:

$$p_{RR}(\varrho) = \binom{\tilde{N}}{\varrho\,\tilde{N}} p_\bullet^{\varrho\,\tilde{N}}\, p_\circ^{(1-\varrho)\tilde{N}},$$

where $\tilde{N}$ is the maximal amount of possible recurrence points, i. e. N^2, when $N \times N$ is the size of the recurrence plot, and $\varrho\,\tilde{N}$ is the amount of observed recurrence points. For large $\tilde{N}$ one can approximate this distribution by the Normal distribution:

$$p_{RR}(\varrho) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\hat{z}^2}{2}} \quad \text{with} \quad \hat{z} = \frac{(\varrho - p_\bullet)\tilde{N}}{\sqrt{p_\bullet\, p_\circ\, \tilde{N}}}.$$

Diagonal Length (L)

The distribution of lengths λ of the diagonal lines for the uncorrelated limit (e. g. white noise) is

$$p_L(\lambda) = p_\bullet^\lambda\, p_\circ^2 = p_\bullet^\lambda (1 - p_\bullet)^2.$$

Note, this is not the distribution of the mean or maximum of the diagonal line lengths ($\langle\lambda\rangle$, $\max\{\lambda\}$).

Determinism (DET)

Applying $p_\bullet$ to the definition of DET leads to

$$DET(\varepsilon) = 2p_\bullet(\varepsilon) - p_\bullet^2(\varepsilon).$$

With the distribution of ϱ and λ we can write the distribution of the determinism δ as

$$p_{DET}(\delta) = \binom{p_\bullet\,\tilde{N}}{\delta\, p_\bullet\,\tilde{N}} (1 - \tilde{p}_L(1))^{\delta\, p_\bullet\,\tilde{N}}\; \tilde{p}_L(1)^{(1-\delta)p_\bullet\,\tilde{N}},$$

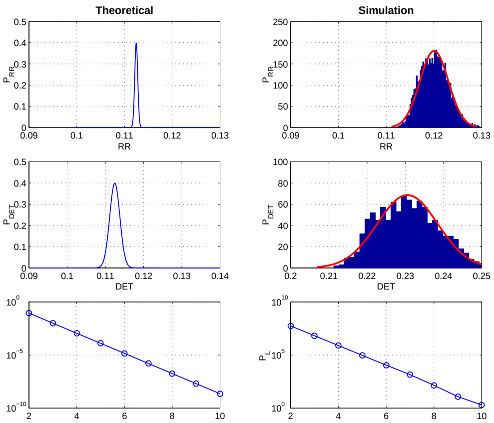
where $\delta\, p_\bullet\,\tilde{N}$ is the amount of recurrence points forming diagonal lines, $p_\bullet\,\tilde{N}$ is the amount of all recurrence points and $\tilde{p}_\Lambda(\lambda)$ is the normalized probability of the diagonal line length λ (with $\sum_\lambda \tilde{p}_\Lambda(\lambda) = 1$ and $p_\Lambda(1) = p_\bullet(1-p_\bullet)^2$). Again, when $\tilde{N}$ is large, we approximate this distribution:

$$p_{DET}(\delta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad \text{with} \quad z = \frac{(\delta - (1 - \tilde{p}_L(1)))\, p_\bullet\,\tilde{N}}{\sqrt{\tilde{p}_L(1)\,(1 - \tilde{p}_L(1))\, p_\bullet\,\tilde{N}}}.$$

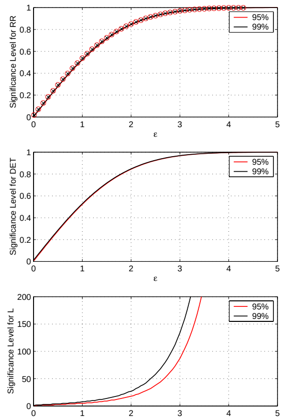
Analogismen

$$RR = p_\bullet = \sum_{\lambda=0}^{N} p_L(\lambda) = \int\limits_{\varrho=0}^{1} \varrho\, p_{RR}(\varrho) d\varrho$$

$$DET = \frac{\sum_{\lambda=\lambda_{min}}^N \lambda\, p_L(\lambda)}{N p_\bullet} = \frac{\sum_{\lambda=\lambda_{min}}^N \lambda\, p_L(\lambda)}{N\, RR} = \frac{\sum_{\lambda=\lambda_{min}}^N \lambda\, p_L(\lambda)}{\sum_{\lambda=0}^N \lambda\, p_L(\lambda)}$$



Distributions for the RQA measures RR and DET and for the length L of single diagonal structures for white, Gaussian noise. Left the theoretical result and right the result of a simulation with 5000 (DET: 1000) realisations. Used embedding was m=1, τ=1; data length 1000. The discrepancy between the theoretical and computed distributions, especially for DET, is due to the erroneous assumption, that the recurrence points are independent.



Estimated one-side significance levels for RR, DET and the length L of single diagonal structures for different thresholds ε. For RR the result of a simulation with 5000 realisations is shown (circles and crosses).

References:

Eckmann, J.-P., S. O. Kamphorst, D. Ruelle: Recurrence Plots of Dynamical Systems, Europhysics Letters 4, 1987

Marwan, N., Kurths, J.: Nonlinear analysis of bivariate data with cross recurrence plots, subm. to Physics Letters A

Thiel, M., Romano, M. C., Kurths, J.: Theoretical Fundamentals of Recurrence Plots and Recurrence Quantification Analysis, Physica D, in press

Trulla, L. L., A. Giuliani, J. P. Zbilut, C. L. Webber Jr.: Recurrence quantification

analysis of the logistic equation with transients, Physics Letters A 223, 1996

Zbilut, J. P., C. L. Webber Jr.: Embeddings and delays as derived from quantification of recurrence plots, Physics Letters A 171, 1992

Webber Jr. C. L., J. P. Zbilut: Dynamical assessment of physiological systems and states using recurrence plot strategies. Journal of Applied Physiology 76, 1994

Suggestions/ Comments: