Ordinal Analysis of Time Series

K. Keller, M. Sinn

Mathematical Institute, Wallstraße 40, 23552 Lübeck

Abstract

In order to develop fast and robust methods for extracting qualitative information from non-linear time series, Bandt and Pompe have proposed to consider time series from the pure ordinal viewpoint. On the base of counting ordinal patterns, which describe the up-and-down in a time series, they have introduced the concept of permutation entropy for quantifying the complexity of a system behind a time series. The permutation entropy only provides one detail of the ordinal structure of a time series. Here we present a method for extracting the whole ordinal information.

Key words: time series, complexity, ordinal patterns, permutation entropy
PACS: 05.45.Tp

1 Ordinal Patterns

The quantification of the complexity of a system is one of the aims of non-linear time series analysis. Complexity is related to complicated intrinsic patterns hidden in the dynamics of the system; if however there is no recognizable structure in the system, it is considered to be stochastic. Because of the occurrence of noise and artefacts in various forms, it is often not easy to get reliable information from a series of measurements. In order to overcome this problem, Bandt and Pompe (1) have proposed an interesting robust approach to time series analysis. They consider the order relation between the values of a time series instead of the values themselves. Their *permutation entropy*, which is strongly related to the Kolmogorov-Sinai entropy in the case of one-dimensional dynamical systems (see (2)), is based on the distribution of ordinal patterns. This paper provides a method for a closer look at this distribution.

Consider a one-dimensional time series $(x_t)_{t \in \mathbb{Z}}$ of real values. For simplicity we take time domain $\mathbb{Z}$, in the case of real-world time series, however, the following considerations must be adapted to a finite time domain. Further, fix some time delay τ. By the *ordinal pattern* of *order* d at time t we understand

Preprint submitted to Elsevier Science *7 March 2005*

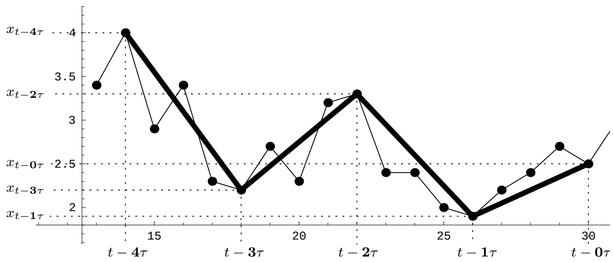

Fig. 1. Ordinal pattern

the permutation $\pi_d^\tau(t) = (r_0, r_1, \ldots, r_d)$ of $(0, 1, \ldots, d)$ satisfying

$$x_{t-r_0\tau} \geq x_{t-r_1\tau} \geq \ldots \geq x_{t-r_{d-1}\tau} \geq x_{t-r_d\tau}.$$

In order to get a unique result, we set $r_{l-1} > r_l$ in the case $x_{t-r_{l-1}\tau} = x_{t-r_l\tau}$.

Example. Fig. 1 illustrates the definition of ordinal patterns for a fictive time series. To get $\pi_4^4(30)$ one has to compare the values $x_{t-\mathbf{0}\tau} = x_{30} = 2.5, x_{t-\mathbf{1}\tau} = x_{26} = 1.9, x_{t-\mathbf{2}\tau} = x_{22} = 3.3, x_{t-\mathbf{3}\tau} = x_{18} = 2.2$ and $x_{t-\mathbf{4}\tau} = x_{14} = 4$. Clearly,

$$x_{t-\mathbf{4}\tau} > x_{t-\mathbf{2}\tau} > x_{t-\mathbf{0}\tau} > x_{t-\mathbf{3}\tau} > x_{t-\mathbf{1}\tau},$$

implying $\pi_4^4(30) = (\mathbf{4}, \mathbf{2}, \mathbf{0}, \mathbf{3}, \mathbf{1})$.

2 Efficient Coding

For reasons becoming clear later, it is useful to code ordinal patterns other than above. Here we omit the mathematical proofs, but we refer to (6).

For $l = 1, 2, 3, \ldots$, let

$$i_l^\tau(t) = \#\{r \in \{0, 1, \ldots, l-1\} \mid x_{t-r\tau} \leq x_{t-l\tau}\},$$

i.e. $i_l^\tau(t)$ counts the inversions (see (7; 4)) of $<$ in time to $\geq$ for the corresponding amplitudes. Now $\pi_d^\tau(t)$ is uniquely coded by the sequence $(i_1^\tau(t), i_2^\tau(t), \ldots, i_d^\tau(t))$. The way from from $(i_1^\tau(t), i_2^\tau(t), \ldots, i_d^\tau(t))$ to $\pi_d^\tau(t)$ is provided via a sequence of permutations $\pi_0, \pi_1, \ldots, \pi_d = \pi_d^\tau(t)$ of $\{0\}, \{0,1\}, \ldots, \{0, 1, \ldots, d\}$:

(1) $\pi_0 = (0)$ is the trivial permutation of the single set $\{0\}$.

(2) When $\pi_{l-1} = (\rho_0, \rho_1, \ldots, \rho_{l-1}); 0 < l \leq d$ is already given, π_l is obtained from π_{l-1} by inserting l into $(\rho_0, \rho_1, \ldots, \rho_{l-1})$ right to ρ_{l-1} if $i_l^\tau(t) = 0$, and left to ρ_{l-i_l} else.

In the above example, $(i_1^4(30), i_2^4(30), i_3^4(30), i_4^4(30)) = (0, 2, 1, 4)$, and the insertion process $\pi_0 \to \pi_1 \to \pi_2 \to \pi_3 \to \pi_4$ is

$$(0) \to (0, 1) \to (2, 0, 1) \to (2, 0, 3, 1) \to (4, 2, 0, 3, 1).$$

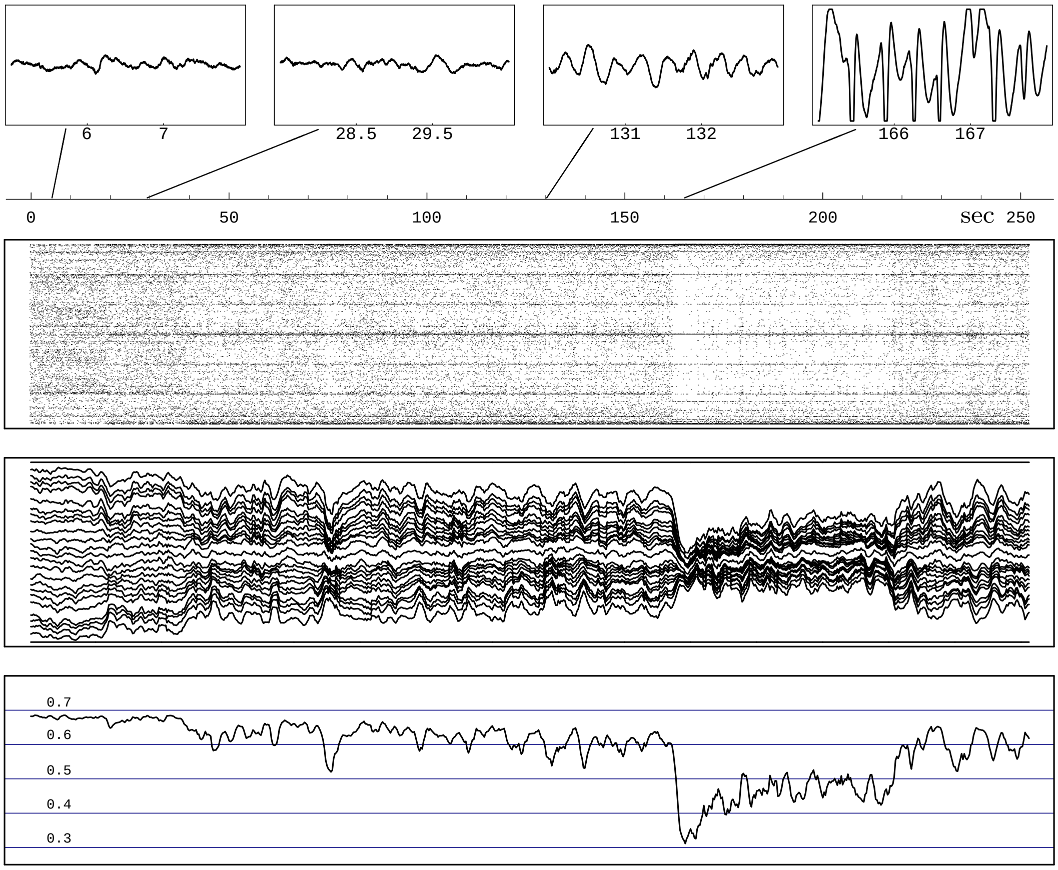

Fig. 2. Ordinal EEG analysis

Remark: According to

$$i_l^\tau(t) = \begin{cases} i_{l-1}^\tau(t-\tau)+1 \text{ if } x_t \leq x_{t-l\tau} \\ \quad i_{l-1}^\tau(t-\tau) \text{ else} \end{cases}$$

$(i_1^\tau(t), i_2^\tau(t), \ldots, i_d^\tau(t))$ can be computed by d comparisons and (with $i_0^\tau(t) := 0$) by at most d incrementations when $(i_1^\tau(t-\tau), i_2^\tau(t-\tau), \ldots, i_{d-\tau}^\tau(t-\tau))$ is given. (In the above example $(i_1^4(30), i_2^4(30), i_3^4(30), i_4^4(30)) = (0, i_1^4(26)+1, i_2^4(26)+0, i_3^4(26)+1) = (0, 1+1, 1+0, 3+1) = (0, 2, 1, 4)$.) With this, computing the ordinal patterns for a time series of length l and given order d needs approximately ld comparisons and ld incrementations.

3 The ordinal transformation

Clearly, there are $(d+1)!$ ordinal patterns of some given order d. On the base of the following statement, the inversion representation $(i_1^\tau(t), i_2^\tau(t), \ldots, i_d^\tau(t))$ of $\pi_d^\tau(t)$ can be used for enumerating all patterns (see (6)):

Background I: The map $(i_1, i_2, \ldots i_d) \mapsto n_d = \sum_{l=1}^{d} i_l \frac{(d+1)!}{(l+1)!}$ is a bijection from $\{0,1\} \times \{0,1,2\} \times \ldots \times \{0,1,\ldots,d\}$ onto $\{0,1,\ldots,(d+1)!-1\}$ turning the lexicographic order into the usual one.

An ordinal pattern is now coded by the number

$$n_d^\tau(t) = \sum_{l=1}^{d} i_l^\tau(t) \frac{(d+1)!}{(l+1)!}$$

being $0,1,2,\ldots$ or $(d+1)-1$. The higher d is, the better the obtained number anticipates the past at time t, the scale of of the numbers however is different. We overcome this disadvantage by a linear scaling to the interval $[0,1]$. So let

$$\nu_d^\tau(t) = \frac{n_d^\tau(t)}{(d+1)!} = \sum_{l=1}^{d} \frac{i_l^\tau(t)}{(l+1)!}.$$

The interesting point is that $\nu_d^\tau(t)$ (theoretically) converges for $d \to \infty$ due to the following statement (see (6)):

Background II: The map $(i_1, i_2, i_3, \ldots) \mapsto \lim_{d\to\infty} \frac{n_d}{(d+1)!} = \sum_{l=1}^{\infty} \frac{i_l}{(l+1)!}$ is a surjection ('near' to a bijection) from $\{0,1\} \times \{0,1,2\} \times \{0,1,2,3\} \times \ldots$ onto the interval $[0,1]$ turning the lexicographic order into the usual one.

We call the assignment of a $(x_t)_{t\in\mathbb{Z}}$ to the time series $(\nu_d^\tau(t))_{t\in\mathbb{Z}}$ with

$$\nu_d^\tau(t) = \sum_{l=1}^{d} \frac{i_l^\tau(t)}{(l+1)!} \in [0,1]$$

ordinal transformation of *order* $d \in \{1,2,3,\ldots,\infty\}$. Roughly speaking, the ordinal transformation extracts the ordinal information contained in a time series, the more of it the higher the order is, and in the (theoretical) case of infinite order all information is extracted. Note that it preserves a part of the geometry of the 'ordinal patterns space' (see (6)).

For the above example, we have $n_4^4(30) = 0 \cdot 60 + 2 \cdot 20 + 1 \cdot 5 + 4 \cdot 1 = 49$ and $\nu_4^4(30) = \frac{0}{2} + \frac{2}{2\cdot3} + \frac{1}{2\cdot3\cdot4} + \frac{4}{2\cdot3\cdot4\cdot5} = \frac{49}{120}$.

Remark: By $n_1^\tau(t) = i_1^\tau(t)$ and the obvious equality

$$n_{d+1}^\tau(t) = (d+1)n_d^\tau(t) + i_{d+1}^\tau(t),$$

the computation of n_d from $(i_1^\tau(t), i_2^\tau(t), \ldots, i_d^\tau(t))$ can be done by $d-1$ multiplications and $d-1$ additions. ν_d is obtained by multiplying n_d with $1/(d+1)!$. So, according to the above remark, the ordinal transformation of order d of a (long) time series of length l needs about $5dl$ logic-arithmetic operations.

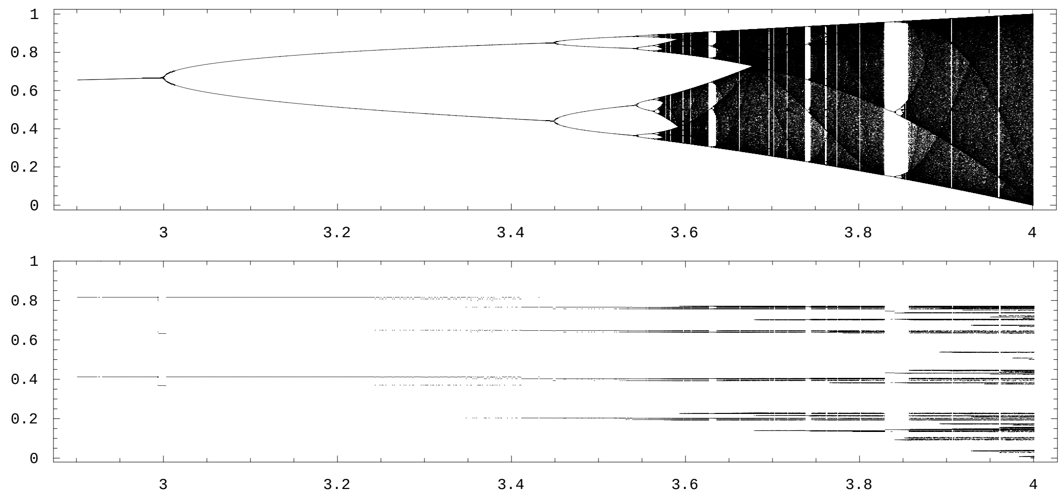

Fig. 3. Feigenbaum and 'ordinal' Feigenbaum diagram

4 Demonstration

We illustrate the ordinal transformation by an EEG example and by considering the celebrated family of quadratic maps $f_r(x) = rx(1-x)$ on $[0,1]$ (see (3)) for $r \in [2.8, 4]$. In both cases $d = 7$, giving $8! = 40320$ possible ordinal patterns, and $\tau = 1$. Other delays τ provide other details of a the time series.

The EEG data set behind Fig. 2, recorded with a sampling rate of $256Hz$ from a scalp electrode, reflects 252 seconds of brain activity. There is an epileptic seizure beginning at 152 sec and ending at 217 sec. The upper plot accompanied with four 3 sec long parts of the original EEG data shows the transformed data. The extremely truncated representation in time direction allows to distinguish different parts. In particular, the extremely 'thin' (attractor-like) part is related to the epileptic seizure. There are however parts where the first inspection of the truncated transformed data gives more information than a close look at the original data, even for different time-scales. (For example, compare the data between 5 and 8 seconds and 27.5 and 30.5 seconds.)

In order to show details of the transformed data, we have added two plots based on a sliding time window analysis (window length 512) for order 3. The first shows the time-dependent pattern distribution in the following way: The spaces between succeeding curves represent the relative frequencies of ordinal patterns, where the pattern $n_3 = 0$ is associated to the bottom space, $n_3 = 1$ to the space between the first and the second curve from below,..., and $n_3 = 23$ to the top space. Differences between distributions at two times coincide with qualitative differences between the original signals at these times (compare the four EEG parts). In particular, a vaster occurrence of the 'bottom' and 'top' patterns $n_3 = 0$ and $n_3 = 23$, respectively, indicates a vaster occurrence of monotone parts within the original signal. The second plot provides the permutation entropy (normalized to maximum one). Here compare also (5).

Fig. 3 shows the *Feigenbaum diagram* and an ordinal version of it. In the first a set of successive large iterates of a random point are drawn in vertical direction for each r, giving an impression how a 'typical' orbit of f_r looks like. The *'ordinal' Feigenbaum diagram* contains the analogue sets for the transformed data. The degree of complexity of the sets correspond, but the 'ordinal' diagram usually shows thinner structures. For example, 'typical' orbits for f_4 are dense in the interval $[0, 1]$, and the ordinal variant fills only a rare part of it. Roughly speaking, the ordinal transformation is able to decide between determinism and stochasticity. Here note that the 'patterns' obtained from discrete white noise by ordinal transformation of order ∞ are equidistributed.

5 Conclusions

Ordinal time series analysis seems to be a promising approach for investigating complex systems. As a base for data analysis on the ordinal level, we have introduced the ordinal transformation. The given examples illustrate that this method allows to recognize structure and to discriminate and classify different states. In order to get reliable statements, it is necessary to develop models and 'ordinal' statistical characteristics beyond the permutation entropy.

Acknowledgement. The authors would like to thank Heinz Lauffer from the Department of Pediatric Medicine of the University Greifswald for the data and the fruitful discussion behind Fig. 2.

References

[1] C. Bandt, B. Pompe, Permutation entropy: A natural complexity measure for time series, *Phys. Rev. Lett.* 88 (2002), 174102.

[2] C. Bandt, G. Keller, B. Pompe, Entropy of interval maps via permutations, *Nonlinearity* 15 (2002), 1595 – 1602.

[3] R. L. Devaney, *An introduction to chaotical dynamical systems*, Addison Wesley, Reading, Massachusetts, 1989.

[4] P. Diaconis, *Group representations in probability and statistics*, IMS, Hayward, California, 1988.

[5] K. Keller, H. Lauffer, Symbolic analysis of high-dimensional time series, *Int. J. Bifurcation Chaos* 13 (2003), 2657 – 2668.

[6] K. Keller, M. Sinn, Ordinal symbolic dynamics, Preprint, Lübeck, 2005.

[7] D. Knuth, *The art of computer programming. Volume 3: Sorting and searching*, Addison Wesley, Reading, Massachusetts, 1973.