Mutual Information for Quaternion Time Series

Michal Piórek

HAL Id: hal-01637476

https://hal.inria.fr/hal-01637476

Submitted on 17 Nov 2017

Mutual Information for Quaternion Time Series

Michał Piórek

Department of Computer Engineering
Wrocław University of Science and Technology
Wybrzeźe Stanisława Wyspiańskiego 27, 50-370 Wrocław, Poland
`michal.piorek@pwr.edu.pl`

Abstract. Mutual Information method is a widely used method for estimation of time delay value in the process of time delay embedding. It's designed for a univariate scalar time series. In the real systems often many outputs of investigated system are available. In this case a multivariate time delay estimation method is necessary if one may require to perform the uniform time delay embedding. The special case of multivariate data is a kinematic time series(e.g quaternion time series or Euler angles time series). The main goal of this paper is to provide a method for this case: Mutual Information method's extension for quaternion time series. The results are also compared with previously presented quaternion's angle method. The method was tested on the real kinematic data - the recordings of human gait.

Keywords: mutual information, quaternions, deterministic chaos, nonlinear time series analysis,human motion analysis, human gait data

1 Introduction

The time delay embedding is heavily explored area of nonlinear time series analysis and nonlinear dynamical systems areas. Usually the most common application of this method is when we want to discover the dynamics of underlying system from the univariate scalar time series created from the measurements of one of the investigated system's outputs. Taken's Embedding Theorem[18] implies, that one can reconstruct an equivalent dynamics from univariate time series using it's time delays. To carry out the embedding procedure, which should result in the reconstructed attractor in the output, two parameters need to be estimated: time delay and embedding dimension.

The time delay is a integer value describing which samples from the investigated time series we need to incorporate to time-lagged embedding vector - reconstructing the underlying phase space. There are few approaches of estimation of time delay value T_d[9]. One group of methods is series correlation approaches (autocorrelation, mutual information[8] or high order correlations[2]). Second grop are approaches of phase space extension (fillfactor[5], wavering product[4] or average displacement[17]). There are available also multiple autocorrelation and non-bias multiple autocorrelation methods[11]. The embedding dimension is

an equivalent of the real underlying phase space dimension. It could be estimated using the false nearest neighbors method[3] or it's extension - Cao's method[7]. Another methods are also the saturation of system invariants method[1] or neural network approaches[13].

The popularity of the univariate embedding may be caused by the fact, that according to the embedding theorem, for recovering dynamics only a univariate time series is needed. In fact, often many time series measured in the output of the test process are available. Since in multivariate case more data is available, it helps to establish more accurate embedding - in the sense of further predictions or in the presence of data noise. However, it brings a new dilemma: which quantities from multivariate time series to use and whether is better to use constant or non-constant embedding parameters for all quantities selected to embedding vector[6].

The problem of multivariate time series embedding can be seen in terms of suitable conditioned embedding of the considered set of time series[19]. In the related work there are two approaches of multivariate time series embedding: Uniform embedding and Non-uniform embedding. The uniform embedding scheme is more popular approach and assumes that embedding parameters: the time delay and embedding dimension are selected a priori and separately for each time series. The non-uniform embedding is based on the progressive selection of time delayed values from a set of candidate values(e.g X, Y, Z) and incorporation them to the embedding vector. In each step the most informative time delayed variables are chosen and then added to the time delay vector. As a selection criteria the mutual information between constructed embedding vector and the future state of the system is used[14][19].

A particular case of multivariate time series is a rotational data time series. There are three main parametrization of rotations: matrix of rotation, Euler angles and Quaternions. Basing on this fact one may record and construct rotational time series according to the one of the above parameterizations. In this paper quaternion rotational time series is considered.

The main goal of this work is to propose time delay estimation method for uniform time delay embedding of multivariate quaternion rotational data. The proposed approach bases on mutual information approach and it's re-designed for quaternion kinematic time series. The presented method could be used in the further time delay embedding and nonlinear analysis aimed to detect deterministic chaos properties in the investigated data. The author would like to underline that the considered method allows to estimate the time delay value staying in quaternion domain, which should help to keep physical sense of the kinematic data.

The paper is organized as following: in the second section the applicability of mutual information method for quaternion data is discussed and proposed approach is described. It also includes the information about investigated quaternion time series and how K-Means algorithm is applied. The third section presents the numerical results. The conclusions are presented in section four.

2   Mutual Information approach for Quaternions

Quaternions are computationally efficient parametrization of rotational data. They are an extension of complex numbers defined as following:

$$q = [w, (x, y, z)] = w + ix + jy + kz \tag{1}$$

where: w represents a real part and $\boldsymbol{v} = (x, y, z)$ is called a vector part (i, j and k are equivalents of imaginary unit).

The details of Quaternions algebra widely used in the parametrization of rotations is well described in the related work(e.g [10]). In the scope of our interests are unit quaternions which describe the rotation in 3D space:

$$\|q\| = 1 \tag{2}$$

Where quaternion norm is defined by:

$$\|q\| = \sqrt{w^2 + x^2 + y^2 + z^2} \tag{3}$$

We assume that the method is designed for the following quaternion time series formed by unit quaternions as following:

$$Q(n) = (q_1, q_2, ..., q_N) = (w_1 + ix_1 + jy_1 + kz_1, ..., w_N + ix_N + jy_N + kz_N) \tag{4}$$

2.1   Mutual Information - existing approach

The mutual information is a measure which describes the general dependence of two variables. The definition comes from Shannon's information theory, which gives the formalism of measuring information spreading. Frasser proposed to use this approach in time delay estimation process[8].

Let's assume that there are two nonlinear systems: A and B. The outputs of these systems are denoted as a and b, while the values of these outputs are represented by a_i and b_k. The mutual information factor describes how many bits of b_k could be predicted where a_i is known.

$$I_{AB}(a_i, b_k) = log_2\left(\frac{P_{AB}(a_i, b_k)}{P_A(a_i)P_B(b_k)}\right), \tag{5}$$

where $P_A(a_i)$ is the probability that $a = a_i$ and $P_B(b_k)$ is the probability that $b = b_k$ and $P_{AB}(a_i, b_k)$ is the joint probability that $a = a_i$ and $b = b_k$.

The average mutual information factor can be described by:

$$I_{AB}(T) = \sum_{a_i, b_k} P_{AB}(a_i, b_k)I_{AB}(a_i, b_k). \tag{6}$$

In order to use this method to assess the correlation between different samples in the same time series, the Average mutual information factor is finally described by the equation:

$$I(T) = \sum_{n=1}^{N} P(S(n), S(n+T))$$
$$log_2\left(\frac{P(S(n), S(n+T))}{P(S(n))P(S(n+T))}\right). \tag{7}$$

Fraser and Swinney ([8]) propose that T_m where the first minimum of $I(T)$ occurs as a useful selection of time lag T_d. This selection guarantees that the measurements are somewhat independent, but not statistically independent. In case of absence of the average mutual information clear minimum, this criterion needs to be replaced by choosing T_d as the time for which the average mutual information reaches four-fifths of its initial value:

$$\frac{I(T_d)}{I(0)} \approx \frac{4}{5}. \tag{8}$$

2.2   Mutual Information extension for quaternion time series

The average mutual information method for univariate time series consists of 2-dimensional adaptive histogram and that is the problem in it's application to multivariate(quaternion's case). The empirical histogram is straightforward to estimate for univariate time series, however for quaternion's time series it's not trivial. Computation of multivariate histogram is exhaustive process and in the result one may obtain the histogram empty in some places.

In the current approach instead of multidimensional histogram for quaternions we propose here to use histogram based on clusters. The whole quaternion time series is initially clustered into k-groups (where k is defined a priori). The obtained clusters are treated as an equivalent of histogram bins. Further in empirical histogram estimation, instead of computing the probability of belonging to the histogram's bins, the probability of belonging to the clusters is being computed.

Algorithm 1 k Means algorithm

 1: **procedure** K MEANS
 2:      $X \leftarrow$ Instance set
 3:      $k \leftarrow$ Number of clusters
 4:      $C \leftarrow$ *Initialize clusters randomly*
 5:      $i \leftarrow 0$ *Number of iterations*
 6:      **while** $C_i! = C_{i+1}$ **do**
 7:          Assign instances to the closest cluster center
 8:          Update cluster centers based on the assignment
 9:      **end while**
10: **end procedure**

Data clustering as a part of machine learning and data sciences is an actively investigated field of science. There are many available clustering techniques. The review of clustering methods is presented in the related work e.g [16]. In the presented approach K-means clustering algorithm was selected as a the simplest and commonly used method. The main goal of this work is not to examine the efficiency of clustering approaches but to provide the mutual information estimation technique used clusters based histogram. In the oder hand, the author

see the underlying potential in investigation of the impact of clustering method selection on the general algorithm's performance.

The K-means algorithm was described by MacQueen[12]. It partitions the data into K clusters (C_1, C_2, ..., C_k), represented by their centers. The center of each cluster, until converge, is calculated as the mean of all the samples belonging to that cluster. Initially the centers are selected randomly. Then, in each iteration each sample is assigned to the closest cluster center according to the Euclidean distance. Then the centers are re-calculated. The whole procedure is repeated until the convergence criteria is fulfilled (e.g there is no relocation of the centers in new iteration)[16]. The whole procedure is presented in the pseudo-code 1.

Algorithm 2 Mutual information algorithm for quaternion time series

---

1: **procedure** MUTUALINFORAMTIONQUAT(QVect,T)   ▷ Mutual information for QVec-quaternion vector and T-time delays vector
2:     $k \leftarrow$ Number of clusters
3:     $n \leftarrow$ length(QVec)
4:     $[labels, centroids] \leftarrow KMeans(QVect,k)$
5:     **for each** t in T **do**
6:         $labelsQ \leftarrow$ labels(1:n-t)
7:         $labelsQD \leftarrow$ labels(1+t:n)
8:         $mI(t) \leftarrow 0$
9:         **for** i=1:k **do**
10:            **for** j=1:k **do**
11:                $c1 \leftarrow$ centroids(i)
12:                $c2 \leftarrow$ centroids(k)
13:                $jointP \leftarrow$ find(c1==labelsQ and c2==labelsQD)
14:                $jointP \leftarrow$ length(jointP)
15:                $c1P \leftarrow$ find(c1==labelsQ)
16:                $c2P \leftarrow$ find(c2==labelsQD)
17:                **if** $jointP > 0$ **then**
18:                    $jointP \leftarrow$ jointP/(n-t)
19:                    $c1P \leftarrow$ length(c1P)/(n-t)
20:                    $c2P \leftarrow$ length(c2P)/(n-t)
21:                    $mI(t) \leftarrow$ mI(t)+ jointP*log2(jointP/(c1P*c2P))
22:                **end if**
23:            **end for**
24:        **end for**
25:     **end for**
26: **end procedure**

Finally, the mutual information algorithm for a quaternion time series is described by the pseudo code[2]. Initially it treats quaternion time series as a 4-dimensional time series and partitions it into K-clusters. The next step is an estimation of probability of belonging samples to the each cluster which is an equivalent of estimation of probability of belonging samples to histogram bins in a standard version of the algorithm.

The author sees the potential advance in the method using an algorithm of clustering which partitions quaternion time series staying in quaternion domain. It will be the subject of further research. It is also worth to investigate how the number of the clusters impacts the performance of the algorithm.

3   Numerical results

The method was tested on live kinematic data recorded in the Human Motion Laboratory (HML) of the Polish-Japanese Institute of Information Technology. The recordings were performed using the Vicon Motion Kinematics Acquisition and Analysis System equipped with 10 Near InfraRed cameras. The cameras were attached to a suit which was worn by a subject.
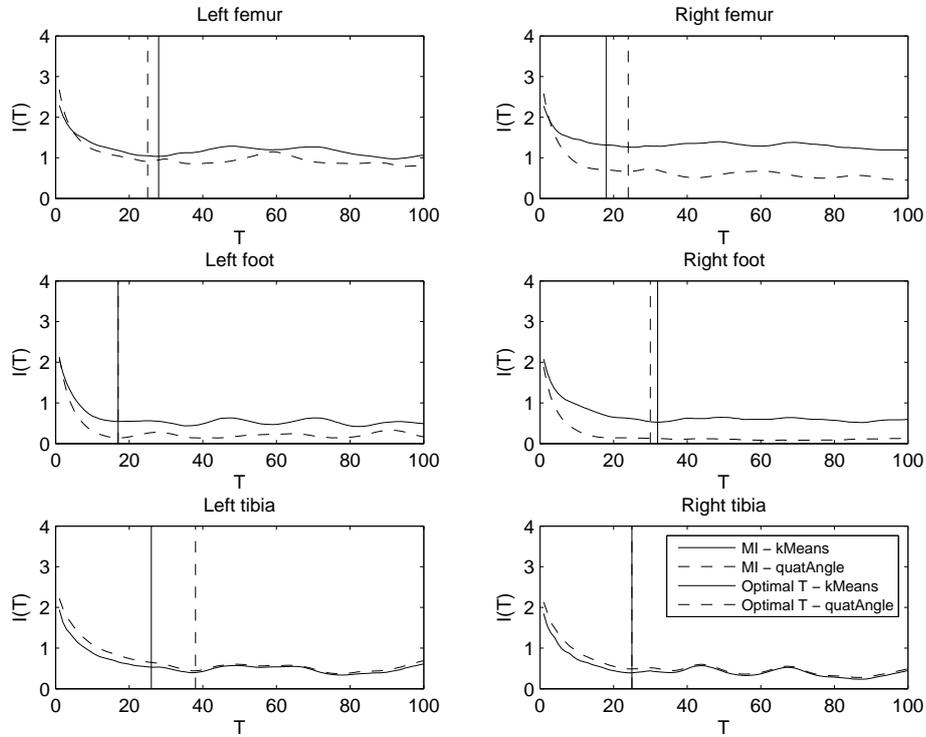
Fig. 1. Mutual information dependency for the patient A

Gait sequences were recorded in Euler angles and then converted to quaternions. Six kinds of time series were recorded - movements of femurs, tibias and feet (left and right). The method was tested on the data recorded from the treadmill walking of two healthy patients. The number of clusters in K-means

algorithm was set to 7. The designed method was additionally compared with the quaternion angle method presented by the author in the same conference last year[15]. The results for the designed method are presented using solid line, where the results performed using quaternion angle method are presented using dashed line. Constantly the first local minima of the mutual information functions(time delay selection criteria) was marked by vertical lines. All estimated time delays are gathered in the table1.
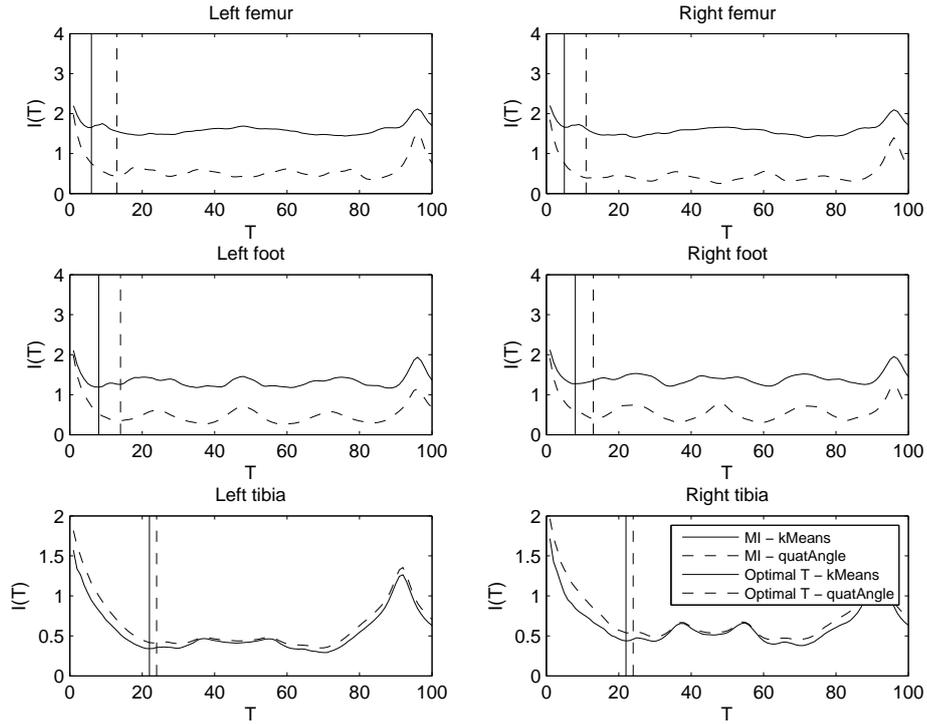
Fig. 2. Mutual information dependency for the patient B

4   Conclusion

The main goal of this article was to present a time delay estimation method for a quaternion time series. The approach extends the existing mutual information approach for quaternion time series by incorporation of K-means clustering for multivariate data instead of the empirical histogram. The method might be a first step to perform time delay embedding staying only in the quaternion domain, which will be a field of author's further research.

body part	T_d MI-angle(Patient A)	T_d MI-quat(Patient A)	T_d MI-angle(Patient B)	T_d MI-quat(Patient B)
left femur	25	28	13	6
right femur	24	18	11	5
left foot	17	17	14	8
right foot	30	32	13	8
left tibia	38	26	24	22
right tibia	25	25	24	22

Table 1. Time delay estimation comparison for subject A and subject B

From visual inspection one can see that the results from the proposed method are in the same range as the results coming from previously investigated quaternion angle method. It is worth to underline that the differences in the result are expected, since quaternion angle method bases only on a part of quaternion's information where the new method utilize the whole information carried by a quaternion.

The field of further interests should be also the impact of clustering parameters on the methods performance and the analysis of the quality of embedding using the proposed approach.

Bibliography

[1] H. Abarbanel. Analysis of observed chaotic data. *Springer-Verlag New York*, 1996.

[2] A. M. et al Albano. Using high-order correlations to define an embedding window. *Physica D*, 54:85–97, 1991.

[3] R. Brown, M. Kennel, and H. Abarbanel. Determinig embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A*, 45(6), 1992.

[4] T. Buzug and G. Pfister. Comparison of algorithms calculating optimal embedding parameters for delay time coordinates. *Physica D*, 58:127–137, 1992.

[5] T. Buzug and G. Pfister. Optimal delay time and embedding dimension for delay-time coordinates by analysis of the global static and local dynamical behaviour of strange attractors. *Phys. Rev. A*, 45:7073–7084, 1992.

[6] L. Cao, A Mees, and K. Judd. Dynamics from multivariate time series. *Physica D*, 121:75–88, 1998.

[7] Liangyue Cao. Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D: Nonlinear Phenomena*, 110(1):43–50, 1997.

[8] A.M. Fraser and H.L. Swinney. Independent coordinates for strange attractors from mutual information. *Physical Review A*, 1986.

[9] Chong-zhao H. Hong-guang M. Selection of embedding dimension and delay time in phase space reconstruction. *Front. Electr. Electron. Eng.*, 1:111–114, 2006.

[10] Bartosz Jablonski. Quaternion dynamic time warping. *Signal Processing, IEEE Transactions on*, 60(3):1174–1183, 2012.

[11] Huang Z. Lin J., Wang Y. and Shen Z. Selection of proper time-delay in phase space reconstruction of speech signals. *Signal Process.*, 15:220–225, 1999.

[12] James MacQueen et al. Some methods for classification and analysis of multivariate observations. 1(14):281–297, 1967.

[13] A. Maus and J.C. Sprott. Neural network method for determining embedding dimension of a time series. *Commun Nonlinear Sci Numer Simula*, 16:3294–3302, 2011.

[14] Alessandro Montalto, Luca Faes, and Daniele Marinazzo. Mute: a matlab toolbox to compare established and novel estimators of the multivariate transfer entropy. *PloS one*, 9(10):e109462, 2014.

[15] Michal Piorek. *Computer Information Systems and Industrial Management: 14th IFIP TC 8 International Conference, CISIM 2015, Warsaw, Poland, September 24-26, 2015, Proceedings*, chapter Chaotic Properties of Gait Kinematic Data, pages 111–119. Springer International Publishing, Cham, 2015.

[16] Lior Rokach and Oded Maimon. Clustering methods, 2005.

[17] Colins J. J. Rossenstein M. T. and de Luca C. J. Reconstruction expansion as a geometry-based framework for choosing proper delay times. *Physica D*, 73:82–98, 1994.

[18] F. Takens. *Detecting strange attractors in turbulence*. Springer-Verlag, Berlin, 1981.

[19] Ioannis Vlachos and Dimitris Kugiumtzis. Nonuniform state-space reconstruction and coupling detection. *Physical Review E*, 82(1):016207, 2010.