VPS VIBRATIONS IN PHYSICAL SYSTEMS

DOI: 10.21008/j.0860-6897.2023.1.21
Vibrations in Physical Systems, 2023, 34(1), 2023121

Application of unsupervised learning algorithms for analysis the vibrations of an oscillator forced by a random series of impulses

Agnieszka OZGA, Marek SULEWSKI

AGH University of Science and Technology, Department of Mechanic Engineering and Robotics, Mickiewicza Av. 30, 30-059 Krakow, Poland

Corresponding author: Marek Sulewski, email: sulewski@agh.edu.pl

Abstract Paper discusses a mathematical model describing the vibrations of a linear oscillator forced by a random series of impulses. The study aims at checking how precisely the distributions of values of the impulses forcing the vibrations of an oscillator can be differentiated. The analysis was carried out in the MatLab environment with the use of hierarchical clustering algorithms of unsupervised machine learning, for samples generated from computer simulation. The time series are non-stationary. The studies showed that high precision could be achieved in distinguishing two very similar distributions forcing the vibrations, on the basis of an analysis of the two first moments calculated from the movement.

Keywords: machine learning, stochastic series of impulses, unsupervised machine learning, hierarchical clustering.

1. Introduction

Research into the systems whose vibrations are forced by random series of impulses was carried out in several centres in Poland and abroad [1-8]. The investigations were based on creating of mathematical models for linear and non-linear systems. With the development of computers and check charts, simulation studies [9] and experimental ones [10] started to develop.

Introducing the non-supervised algorithms of machine learning in the analysis of dynamic mechanical systems, we are starting a new level of studies on random issues. This paper discusses a mathematical model describing the vibrations of a linear oscillator. The vibrations are forced by a random series of impulses.

The study aims at checking how precisely it is possible to differentiate the distributions of values of the impulses forcing the vibrations of an oscillator with the use of algorithms of unsupervised machine learning.

2. Mathematical model for determining the impulse distributions forcing oscillator oscillations

The vibrations of a discrete, linear system, excited by a random series of impulses (1) in the form of a random series of impulses

$$f(t) = \sum_{t_i<t} \eta_i \delta(t - t_i) \tag{1}$$

described by the equation (2)

$$x(t) = \frac{1}{c} \sum_{0<t_i<t} \eta_i e^{-b(t-t_i)} sin(c(t - t_i)), \tag{2}$$

where i =1, 2, 3, ... are the impulse numbers, $\delta(t - t_i)$ are Dirac distributions at the time t_i, $\{\eta_i\}_{-\infty}^{+\infty}$ is a sequence of independent identically random values of amplitude of i-th impulse with finite expectation, t_i are random instants of time at which the impulses occur, x is the deflection of the system from its equilibrium position, b the damping coefficient and the frequency c are parameters of the vibrating system.

The time intervals (3)

$$\tau_i = (t_i - t_{i-1}), \tag{3}$$

between impulses $\{\tau_i\}_{-\infty}^{+\infty}$ are independent continuous random variables for which the function of probability density assumes the form of exponential distribution (4):

$$f(\tau) = \begin{cases} \lambda e^{-\lambda t} & for\ \tau \geq 0 \\ 0 & for\ \tau < 0 \end{cases}. \tag{4}$$

The constant λ is the impulse rate. The intervals between the impulses $\{\tau_i\}_{-\infty}^{+\infty}$ and value of the impulses $\{\eta_i\}_{-\infty}^{+\infty}$ are independent random variables.

A mathematical model the inverse identification problem (5-6) that allows for determining the distribution of value of impulses forcing the vibrations of the system was developed [11-14] in several stages and was constructed on the basis of linear differential equations using the ergodic theory together with the basics of the theory of dynamic systems, measure theory, group theory, probability calculus and the theory of stochastic processes based on it, and when $t \to \infty$.

$$\sum_{i=1}^{k} p_i \left[(m_n m_1 - m_{n+1})\eta_i + \sum_{j=1}^{n} \binom{n}{j} m_{n-j} m_1 \eta_i^{j+1} \frac{C(j+1)}{C(1)c^j} \right] = 0, \tag{5}$$

$$\sum_{i=1}^{k} p_i = 1, \tag{6}$$

where k is the number of the sought values of random variable η_k, p_i is the probability of η_i determined from the vibrations $x(t)$, m_n are n-th stochastic raw moments of the random variable $x(t)$ and $C(j)$ are constants depending on the parameters of the oscillator.
For $j>1$ and even j

$$C(j) = \frac{j!}{\prod_{r=0}^{j/2-1}((jb/c)^2 + (2r)^2)} \frac{c}{jb}, \tag{7}$$

where as for odd $j > 0$,

$$C(j) = \frac{j!}{\prod_{r=0}^{(j-1)/2-1}((jb/c)^2 + (2r+1)^2)}. \tag{8}$$

3. Simulation studies

Application of a model for t <3600 seconds is limited [15-19]. Estimators of the k-th stochastic raw moments of the random variable $x(t)$ calculated using the equation (8) change with the passage of time presented on the Fig. 1A and Fig. 1B

$$\overline{m}_k(t) = \frac{1}{[t/h]} \sum_{n<t/h} x^k(nh), \tag{9}$$

where h is the time period of sampling, t is time. Hence the distributions of probabilities η_i determined with the help of equations (5-8) are burdened with uncertainty what presented on the Fig. 1C.

At the present level of knowledge, execution of an experiment in which Dirac Delta would occur at the impulse and the restitution coefficient would not be necessarily taken into account in the model is practically impossible. However, it is possible to execute simulated studies, which are an approximation of the modelled phenomenon, and in these studies the qualitative analysis is used to prepare the experiments in the proper way. In this article we introduce application of unsupervised learning algorithms to solve the problem of recognizing the distribution of impulses generated for two different distributions values of impulses:

$$\Phi_1: p(\eta_1=130) = 0.5,\ p(\eta_2=20) = 0.5.$$
$$\Phi_2: p(\eta_1=140) = 0.5,\ p(\eta_2=10) = 0.5. \tag{10}$$

The distributions Φ_1 and Φ_2 were characterized by two events of different force of impact. The value of η_1 symbolizes an impulse of a great force of impact while the value of η_2 an impulse of a little force of impact on the oscillator. Distributions were selected so that the mean value was the same in all two cases, stochastic raw moments of the second order (and subsequent orders) are different.

The study is supposed to check how precisely two distributions with the same mean value and similar subsequent stochastic moments can be differentiated. Fig. 2 shows estimators of the first two raw moments

(8) for the distributions Φ_1 and Φ_2 for 100 different samples. In the distribution Φ_1, 48 samples were generated while the second distribution gave 52 samples.

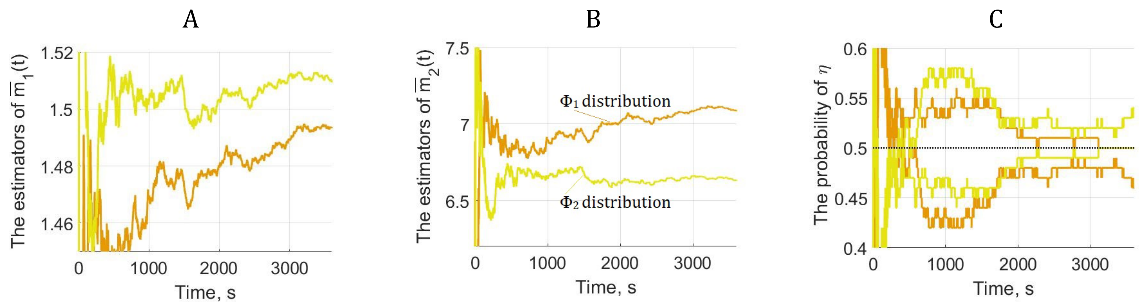

Figure 1. The first stochastic raw moment (A) and the second one (B) computed on the basis of $x(t)$ for 3600 seconds for one sample from Φ_1 distribution and one sample from Φ_2 distribution. (C) The probability calculated from the model of impulses.

For further study that is supposed to achieve the intended goal , the method of unsupervised learning, connected with data clustering, was selected. Similarities between elements – expressed with the help of metrics like Manhattan, Euclidean, Chebyshev or Minkowski distance, etc., were adopted as the basis for clustering. Apart from the metric, the algorithm requires also selection of the method of clustering. In MatLab environment, in which the study was carried out, agglomerative analysis can use the following methods: average, centroid, complete, median, single, ward and weighted.

For the samples generated between the 600th and the 1800th seconds, the estimators of the first moment have different statistical parameters describing non-stationary time series for both considered distributions even though the distributions forcing the vibrations of the oscillator are of the same mean value presented on the Fig. 2. In turn, the estimators of the second moment start differing in statistical values of time series as late as after the 1800th second. The differences between mean values for the samples obtained from the distributions Φ_1 and Φ_2 are smaller than 0.1 for the estimators of the first moment and five times as big, amounting to 0.5, for the estimators of the second moment.

In cluster analysis an algorithms works as follows: at the beginning, each observation generates a one-element cluster. Then, pairs of clusters are merged – a graphic representation of the executed analysis has the form of a dendrogram presented on Fig. 3.

The study investigated various connections between the metric and the method in the context of division of the set into two groups. The following statistical parameters of time series were used in the study: mean value, median, mode, standard deviation, maximum, minimum, and skewness. All variables were standardized. Thanks to the fact that the samples were generated in simulation investigations, it was possible to label them and verification of the applied algorithms could be executed with the use of confusion matrix [20] for the accuracy calculation purposes.

Table 1 shows conclusion of the study. It presents the combinations of metrics and methods which, both in the time series from the 800th to the 1400th second and in that between the 2400th and 3000th second revealed that unsupervised clustering is fully compatible with the labels of particular samples. It should be remembered, however, that the results are compatible only when in the segmentations algorithm it is assumed that two clusters are searched for Fig. 3.

Table 1. Methods and metrics that ensured 100% precision in clustering in both time intervals – the one from 800 to 1400 sec and the other from 2400 to 3000 sec.

method	metrics	method	metrics	method	metrics
	Euclidean		average		average
	squaredEuclidean		centroid		centroid
	seuclidean		complete		complete
ward	mahalanobis	Chebychev	median	cosine	median
	cityblock		single		single
			ward		ward
			weighted		weighted

For these and only these samples the differences between the estimators of the first moment and the second one were distributed in this time series. In order to check with what precision the cluster algorithms will differentiate between the distributions Φ_1 and Φ_2 and investigations should be executed also in other time intervals. For the next step, for the classification purposes set ward method and Euclidean metrics. From 300s, the statistics for the five-minute time interval are calculated every minute. It can be noticed that after 900s archived 100% correct classification in each time period as presented on the Fig. 6.

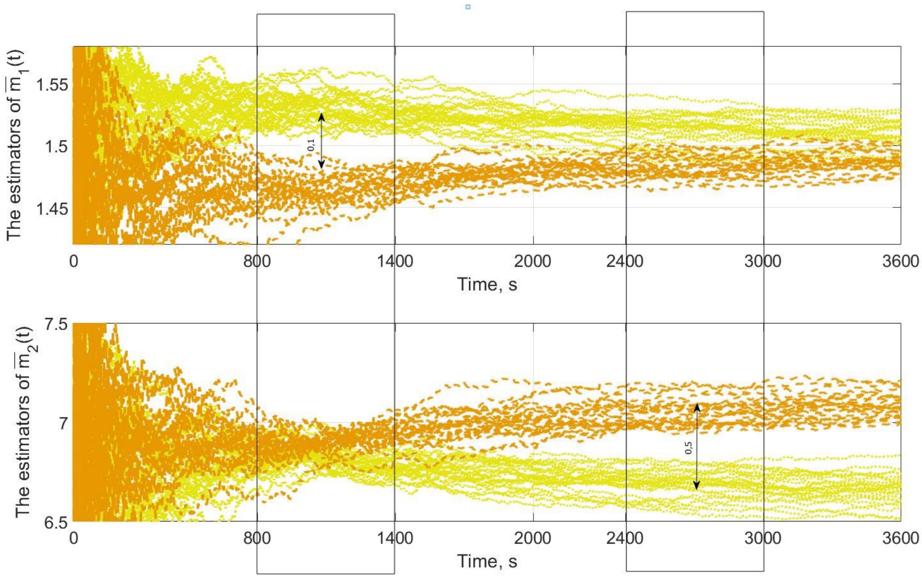

Figure 2. The first stochastic raw moment (9) and the second one computed on the basis of $x(t)$ for 3600 seconds for 48 samples from Φ_1 distribution and 52 samples from Φ_2 distribution.

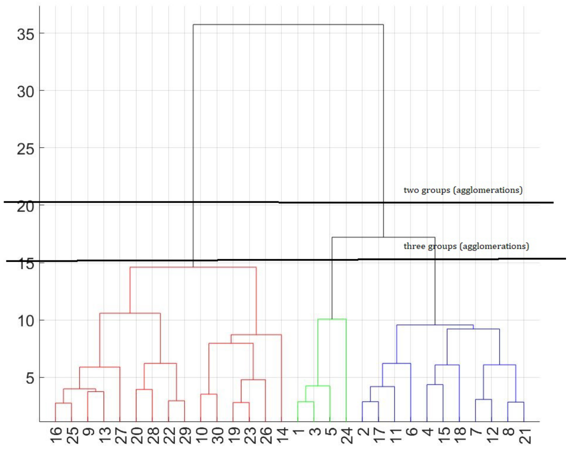

Figure 3. Dendrogram for ward method combined with Euclidean metric defined for the time period between 800 seconds to 1400 seconds.

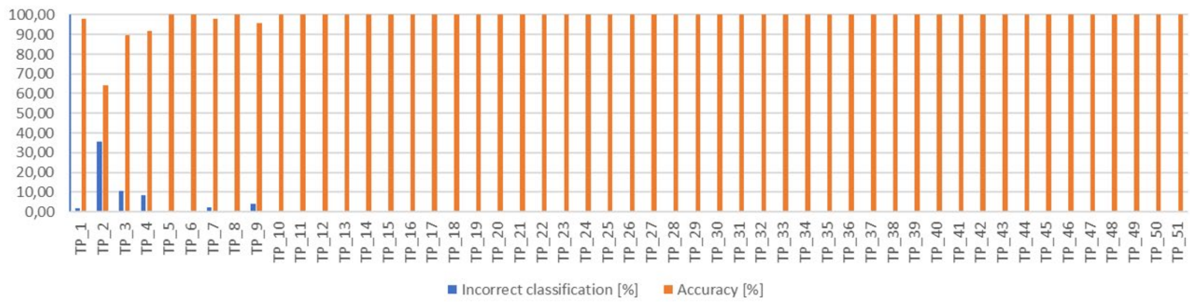

Figure 4. Summary after data classification using data clustering algorithm for interval set to 300 seconds.

4. Conclusion and summary

The paper discusses the first attempts at application of machine learning algorithms in the analysis of random vibrations of an oscillator. These algorithms have connections in the sets and allow for decision making on the basis of the shared data. Time series used in the work, which contain two first raw moments calculated from the movements of vibrations of the oscillator are non-stationary. Cluster analysis were selected for the study. The investigations showed that the distributions forcing the vibrations of the oscillator can be differentiated highly precisely on the basis of an analysis of the two first moments. The only weakness of this method is determining at the final stage of the analysis how many groups are sought. Removal of this limitation and conduction investigations for a larger number of distributions with more complex structures will be the object of further analyses.

Acknowledgement

The article was financed by a research subsidy of the number: 16.16.130.942.

Additional information

The author(s) declare: no competing financial interests and that all material taken from other sources (including their own published works) is clearly cited and that appropriate permits are obtained.

References

1. R. Iwankiewicz, S. R. K. Nielsen; Dynamic-response of nonlinear-systems to poisson-distributed random impulses; Journal of Sound and Vibration, 1992, 156, 407-423
2. K. Mazur-Śniady, P. Śniady; Dynamic response of linear structures to random streams of arbitrary impulses in time and space; Journal of Sound and Vibration, 1986, 110(1), 59-68
3. J. B. Roberts; System response to random impulses; Journal of Sound and Vibration, 1972, 24(1), 23-34
4. J. B. Roberts, P.D. Spanos; Random vibration and statistical linearization; New York Dover Publication, 2003
5. B. Skalmierski, A. Tylikowski; Procesy stochastyczne w dynamice; PWN, Warszawa, 1972
6. K. Sobczyk; Metody stochastyczne w mechanice: stan i tendencje rozwojowe (in Polish); Mechanika Teoretyczna i Stosowana, 1983
7. L. Socha, T.T. Soong; Linearization in analysis of nonlinear stochastic systems, 1991
8. A. Tylikowski; Drgania oscylatora harmonicznego wywołane ciągiem przypadkowych zderzeń (in Polish); Prace Instytutu Podstaw Budowy Maszyn, 1982, z.13
9. R. Iwankiewicz; Metody stochastyczne w zagadnieniach układów dynamicznych poddanych losowym seriom impulsów (in Polish); Prace Naukowe Instytutu Materiałoznawstwa i Mechaniki Technicznej Politechniki Wrocławskiej, Wrocław, 1993
10. M. Jablonski, A. Ozga, T. Korbiel, P. Pawlik; Determining the distribution of stochastic impulses acting on a high frequency system through an analysis of its vibrations; Acta Physica Polonica A, 2011, 119, 6A, 977-980
11. M. Jabłoński, A. Ozga; Statistical characteristics of vibrations of a string forced by stochastic forces; Mechanics/AGH University of Science and Technology, 2008, 27(1), 1-7
12. M. Jablonski, A. Ozga; Distribution of Stochastic Impulses Acting on an Oscillator as a Function of Its Motion; Acta Physica Polonica A, 2010, 118, 74-77
13. M. Jablonski, A. Ozga; Determining the Distribution of Values of Stochastic Impulses Acting on a Discrete System in Relation to Their Intensity; Acta Physica Polonica A, 2012, 121, A174-A178
14. M. Jablonski, A. Ozga; Distribution of Random Pulses Acting on a Vibrating System as a Function of Its Motion; Agh-Univ Sci & Technol, Krakow, 2013
15. A. Ozga; The effect of pulse amplitudes on quality of determining distribution of pulses forcing vibration of an damped oscillator, Acta Physica Polonica A, 2015, 128(1A), A-67; DOI: 10.12693/APhysPolA.128.A-67
16. A. Ozga; Identyfikacja rozkładu losowych obciążeń impulsowych w liniowych dyskretnych układach dynamicznych (in Polish); Wydawnictwa AGH, 2019; ISSN 0867-6631
17. T. Smolnicki, M. Stańco, D. Pietrusiak; Distribution of loads in the large size bearing–problems of identification; Tehnički vjesnik, 2013, 20(5), 831-836
18. H. Weber, S. Kaczmarczyk, R. Iwankiewicz; Non-Linear Response of Cable-Mass-Spring System in High-Rise Buildings under Stochastic Seismic Excitation; Materials, 2021, 14 (22), 6858

19. Z. Zembaty; Tutorial on surface rotations from wave passage effects: stochastic spectral approach; Bulletin of the Seismological Society of America, 2009, 99(2B), 1040-1049
20. M. Szeliga; Praktyczne uczenie maszynowe (in Polish); Wydawnictwo naukowe PWN SA, Warszawa, 2019