Sensitivity Analysis of Neural Network Models: Applying Methods of Analysis of Finite Fluctuations

Anton Sysoev[1*], Alessandro Ciurlia[2], Roman Sheglevatych[3], Semen Blyumin[1]

[1] Department of Applied Mathematics, Faculty of Automation and Computer Science, Lipetsk State Technical University, RU-398055 Lipetsk, 30 Moskovskaya str., Russia
[2] Department of Information Engineering, Faculty of Engineering, Marche Polytechnic University, IT-60131 Ancona, 12 Via Brecce Bianche, Italy
[3] Lipetsk Region Health Department, RU-398050 Lipetsk, 6 Ulitsa Zegelya str., Russia
* Corresponding author, e-mail: anton_syssoyev@mail.ru

Abstract

As an initial stage prior to Mathematical Modeling, the information processing should provide qualitative data preparation for the construction of consistent models of technical, economic, social systems and technological processes. The question, concerning choosing the most significant input factors affecting the function of the system, is a very actual and important. This problem could be solved with the application of methods of Sensitivity Analysis. The presented paper has the purpose to show a possible approach to this problem through the method of the Analysis of Finite Fluctuations, based on Lagrange mean value theorem, to study the sensitivity of the model under consideration. The numerical example of comparing the results obtained by Sobol sensitivity coefficients, Garson algorithm and proposed approach showed the sustainability of the introduced method. There is shown, that the proposed approach is stable in the sense of applying different input datasets. In particular, the proposed approach has been applied to the construction of a neural network model identifying any anomalies present in certain medical insurances, in order to define the most significant input factors in the anomaly's detecting, discard the others and get a slim and efficient model.

Keywords

Sensitivity Analysis, neural network, Analysis of Finite Fluctuations

1 Introduction

A mathematical model can be descripted as:

$$y = f(X), \tag{1}$$

where $X = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ represents an input factors vector, where $x_i \in \mathbb{R}$ for $i = 1, \ldots, n$ are its components and input values of the model, $y \in \mathbb{R}$ is an output value, and $f(\cdot)$ is a model, that can be a function, a system of differential equations, etc., even a program code. In this particular study a neural network model with continuous and differentiable activation functions is investigated. In many applied problems it is very necessary to know, which input factors are the most significant in order, for example, to reduce the model or to transform it to the model from the other class. Commonly, the answer to such kind of question could be obtained after applying approaches of Sensitivity Analysis [1], which is based on statistical and probabilistic techniques and allows to estimate the influence of each model input value (independent variable, argument of the function, factor of the system, etc.) on the output value (dependent variable, function value, index of the system, etc.). This procedure is made through the individuation of some indicators, called impact factors, determining quantitatively the influence, that each input has on the output, and consequently, allowing to understand which of these inputs have to be changed the least possible, so that the output of the model does not change too much [2]. There are many well-known techniques of Sensitivity Analysis, some of them will be presented further. However, some of them have drawbacks (like stochastic nature or high computational costs, etc.). In contrast to these techniques it is possible to use the approach based on applying Lagrange mean value theorem and called Analysis of Finite Fluctuations [3]. This analysis can be seen as an approach of Mathematical Remodelling [4, 5],

which aim is to transform a model, connecting the output value with the input values, into another model that puts in relation fluctuation of the output value with fluctuations of input values.

The purpose of the reported study is to present Analysis of Finite Fluctuations as a powerful tool of Sensitivity Analysis. Neural networks were studied as a class of model to which this type of analysis was applied. Results of numerical experiments under the abstract data set are presented and after the approach is widespread in application of dataset on medical insurance information with the main purpose to construct a neural network classifier to predict anomalies in such datasets.

2 State-of-art in Sensitivity Analysis

Considering Eq. (1), it is possible to apply either a deterministic approach, where X has specific values and the space of uncertain input can be explored with statistical methods, or a probabilistic approach, where $X = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ could be studied as a vector of random variables, and consequently, Y is a random variable with an unknown distribution. The problem of Sensitivity Analysis is in finding a relationship between the input values and the output value, and understanding how the firsts influence on the second [2].

To implement Sensitivity Analysis to a model, one can use a local approach or a global strategy. The local approach is based on calculating or estimating the partial derivatives of the model in a specific point of the input space, and it is very effective to elaborate models with a high number of inputs. The global approach, instead, considers the model in the entire domain of the input values. This fact is important, because it allows to study the model in its entirety rather than on a local domain, and therefore, allows to get overall info on it [2].

All existed approaches of Sensitivity Analysis can be divided into several strategies.

The first strategy is in applying graphical instruments, which is very easy to implement. It consists of analyzing inputs/output sample $(x_1^i, x_2^i, \ldots, x_n^i, Y^i)$ through scatterplots, which describe the existing linearity between the output and every single input. If 2-dimensional scatterplots are not able to give information about such relation, parallel coordinates analysis is used. This relationship can be described through a linear (Pearson) correlation coefficient, or through a rank (Spearman) correlation coefficient.

In case of using linear regression model, standardized regression coefficients

$$\text{SRC}_i = \beta_i \sqrt{\frac{\sigma_{x_i}^2}{\sigma_Y^2}},$$

show proportion of a single input x_i variance on the variance of the output Y, (but only in case if the linearity hypothesis is confirmed).

The second strategy is the utilization of screening techniques [6-8], which have the objective to give a qualitative ranking of the input factors with the less possible number of model evaluations, exploiting the design of experiments for screening and generally using a surrogate model to estimate the output. Here are also used variational methods, aiming at the obtaining of information from the model through the utilization of partial derivatives. A very simple idea is to describe the sensitivity of output to the variations of input factors through the partial derivatives of the model, where this model is described by a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ with continue derivatives and gradient

$$\nabla_x f(X) = \left(\frac{\partial f}{\partial x_1}(X), \frac{\partial f}{\partial x_2}(X), \ldots, \frac{\partial f}{\partial x_n}(X) \right)^T.$$

The third strategy is in applying a sampling techniques [9, 10]. Here are mainly methods using composite indicators, namely indicators which have the purpose to aggregate input factors and to find a relationship between them and the output. Generally, these indicators are obtained as a linear combination of input factors with some scalars called weights, indicating their influence on the output:

$$Y_i = \sum_{i=1}^{d} (w_i x_{ji}), \quad j = 1, \ldots, n$$

where x_{ji} are values of input factors, w_i are weights and Y_i is the output. Mostly known examples of such techniques are Garson algorithm [11], used to assess the sensitivity of neural network models, and the variance based methods as Sobol algorithm [12].

In conclusion, the last, but not the least strategy is using metamodels, namely substitutive models, which are applied instead of originals, when the latter need too much resources in terms of CPU to be processed during their calculations. In particular, two classes of metamodels are widely used. They are Polynomial Chaos Expansion and Gaussian Process [13]. When the model is replaced with its metamodel, the others strategies described above could be adopted.

3 Basics of Analysis of Finite Fluctuations

Fluctuations of some variable can be described in several ways. Let us have a value of variable v at an initial instant of time $t_0 (v^{t_0})$, and a value of the same variable at a final

instant of time $t_1 \left(v^{t_1} \right)$. The most used forms to represent finite fluctuation $\mu(v)$ of the variable v are [4]:

- absolute increment

$$\mu(v) = \Delta v = v^{t_1} - v^{t_0};$$

- index

$$\mu(v) = i(v) = \frac{v^{t_1}}{v^{t_0}};$$

- relative increment

$$\mu(v) = \delta(v) = \frac{v^{t_1} - v^{t_0}}{v^{t_0}} = \frac{\Delta v}{v^{t_0}} = i(v) - 1.$$

These mentioned fluctuations must be finite, namely also small but not infinitesimal. Being a branch of Mathematical Remodelling, Analysis of Finite Fluctuations [14] is aimed at solving the problem of constructing based on the model $y = f(X)$ the new model $\mu(Y) = \phi(\mu(X))$, showing the connection between the fluctuation of the output $\mu(Y)$ and the fluctuations $\mu(X)$ of its factors $X \in \mathbb{R}^n$. In case of small increments Mathematical Analysis gives an example of the approximative Remodelling. If the function $y = f(X)$ describing the model under consideration, is defined and continuous in a closed domain and has continuous partial derivatives within this domain, the approximate connection between the response small fluctuation and small fluctuations of its arguments is

$$\Delta y = f \left(X^{(0)} + \Delta X \right) - f \left(X^{(0)} \right) = f \left(x_1^{(0)} + \Delta x_1, \ldots, x_n^{(0)} + \Delta x_n \right)$$

$$- f \left(x_1^{(0)}, \ldots, x_n^{(0)} \right) \approx \sum_{i=1}^{n} \frac{\partial f \left(X^{(0)} \right)}{\partial x_i} \cdot \Delta x_i.$$

On the other hand, in some applied problems the fluctuations could not be considered as small values, but could be interpreted as finite values.

Mathematical Analysis gives in case of finite increments the model, which allows to represent the model (Eq. (1)) to an exact connection between the finite fluctuation of the model output and its factors fluctuations. This is Lagrange mean value theorem (the formula of finite increments, intermediate value theorem of Differential Calculus) for multivariable functions, defined and continuous in a closed domain and having continuous partial derivatives inside this domain. It is formulated by Eq. (2):

$$\Delta y = \sum_{i=1}^{n} \frac{\partial f \left(x^{(m)} \right)}{\partial x_i} \cdot \Delta x_i, \tag{2}$$

$$x^{(m)} = \left(x_1^{(m)}, \ldots, x_n^{(m)} \right),$$

$$x_i^{(m)} = x_i^{(0)} + \alpha \cdot \Delta x_i, \quad 0 < \alpha < 1.$$

Here the mean (or intermediate) values of arguments (factors) $x_i^{(m)}$ are defined by the value of α.

4 Sensitivity Analysis of neural network models: applying Analysis of Finite Fluctuations

This section presents the implementation of Analysis of Finite Fluctuations to study sensitivity of neural network models.

Let us have a neural network model with n hidden layers, which describes studied technical, economic or social system or technological process

$$y = f(X) = \Psi^{(n)} \Psi^{(n-1)} \ldots \Psi^{(1)} X,$$

where $\Psi^{(n)}, \ldots, \Psi^{(1)}$ are activation functions applied to hidden and output layers.

According to the idea of Analysis of Finite Fluctuations let us take the initial instant of time t_0, where the input factors vector is $X^{(t_0)} = \left(x_1^{(t_0)}, \ldots, x_n^{(t_0)} \right)$ and respectively output is

$$y^{(t_0)} = f \left(X^{(t_0)} \right) = f \left(x_1^{(t_0)}, \ldots, x_n^{(t_0)} \right).$$

After a while, at the final instant of time t_1, there are

$$X^{(t_1)} = \left(x_1^{(t_1)}, \ldots, x_n^{(t_1)} \right)$$

and

$$y^{(t_1)} = f \left(X^{(t_1)} \right) = f \left(x_1^{(t_1)}, \ldots, x_n^{(t_1)} \right),$$

where $x_i^{(t_1)} = x_i^{(t_0)} + \Delta x_i, \quad i = 1, \ldots, n.$

It follows, that

$$\Delta y = y^{(t_1)} - y^{(t_0)} = f \left(X^{(t_1)} \right) - f \left(X^{(t_0)} \right)$$

$$= f \left(\ldots, x_i^{(t_0)} + \Delta x_i, \ldots \right) - f \left(\ldots, x_i^{(t_0)}, \ldots \right). \tag{3}$$

From the other hand, it is also possible to estimate Δy using Eq. (2) as

$$\Delta y = \sum_{i=1}^{n} \left(\frac{\delta f}{\delta x_i} \left(\ldots, x_i^{(t_0)} + \alpha \cdot \Delta x_i, \ldots \right) \cdot \Delta x_i \right) \tag{4}$$

and, using notations $A_i = \frac{\delta f}{\delta x_i} \left(\ldots, x_i^{(t_0)} + \alpha \cdot \Delta x_i, \ldots \right)$, the formula becomes

$$\Delta y = \sum_{i=1}^{n} \left(A_i \cdot \Delta x_i \right) = A_1 \Delta x_1 + A_2 \Delta x_2 + \ldots + A_n \Delta x_n. \tag{5}$$

Equating Eq. (3) and Eq. (4) and obtaining Eq. (5), it is resolved the resultant equation according to the unknown parameter α, finding which and, respectively, estimating impact indexes A_i, determining the influence, that each input factor fluctuation has on the output fluctuation.

5 Numerical experiments

5.1 Scope of experiment

To implement the proposed approach, the R language was used through the free and open-source integrated development environment RStudio. The neural network model with 15 input factors, 1 hidden layer with 3 neurons, and 1 output was investigated. This mentioned model has the following structure

$$y_k = \psi_1\left(b_2 + \sum_{j=1}^{3} w_j \psi_2\left(b_1 + \sum_{m=1}^{15} w_{jm} x_m\right)\right), \qquad (6)$$

where y_k are fitted output values; $x_m = \{x_1, x_2, \ldots, x_{15}\}$ are input factors; w_j and w_{jm} are the weights between the hidden layer and the output, and between the inputs and the hidden layer respectively; b_2 and b_1 are the bias values on the output and on the hidden layer respectively; $\psi_1(net) = \psi_2(net) = 1/(1 + \exp(-net))$ are the logistic activation functions.

For the experimental part 2000 cases from the dataset "neuraldat" from NeuralNetTools R package was used. The model was fitted through the specific function "nnet" from the homonymous R package. It should be also mentioned, that the obtained model has demonstrated the accuracy of 99.82 %.

5.2 Studying approach based on applying Analysis of Finite Fluctuations

Three approaches were tested (Sobol sensitivity coefficients, Garson algorithm and the proposed approach). Table 1 and Fig. 1 show the results of estimation of the importance of

Table 1 Importance of studied input factors

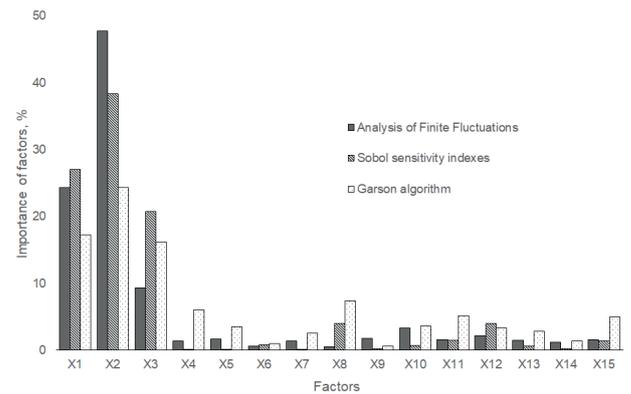| Input factors	Importance of input factors, %		
	Sobol sensitivity coefficients	Garson algorithm	Analysis of Finite Fluctuations
x_1	27.07	17.23	24.31
x_2	38.34	24.29	47.71
x_3	20.77	16.15	9.36
x_4	0.02	6.03	1.41
x_5	0.16	3.51	1.72
x_6	0.81	0.97	0.61
x_7	0.17	2.53	1.42
x_8	3.96	7.38	0.51
x_9	0.25	0.57	1.73
x_{10}	0.74	3.66	3.28
x_{11}	1.51	5.17	1.56
x_{12}	4.01	3.36	2.14
x_{13}	0.61	2.81	1.46
x_{14}	0.20	1.34	1.21
x_{15}	1.36	5.00	1.57

Fig. 1 Studying sensitivity of the model according to input factors

model factors. Due to the different scales used in these approaches, all results were aggregated to be further compared (reduced to percentages). It should be mentioned, that for the proposed approach there were obtained 1999 estimates for each input factor (because of existing 1999 finite fluctuations for input values dataset) and their median values were taken as a sensitivity measure.

Analyzing Table 1 and Fig. 1 it is possible to conclude, that all compared strategies gave similar results with a slight variation, which proves, that Analysis of Finite Fluctuations is not contradictory and could be applied in such kind of problems. But proposed approach has an undeniable advantage. In contrast to Sobol sensitivity coefficients it does not use an approximation procedure to model statistical parameters of the studied structure and in contrast to Garson strategy, it operates with both parameters and factors of the studied model.

The next studied step was the stability of the proposed approach. According to [15], a numerical algorithm is stable, if considerable changes in input variables do not produce considerable changes in the output. For testing this property of the proposed approach there were constructed 500 random samples presenting different fluctuations of input variables. There was studied sensibility of the model (Eq. (6)) for each input dataset; results are presented as boxplots of calculated medians for each variable on Fig. 2.

Based on the results of conducted calculations we can conclude, that proposed approach is stable under different input variables values.

5.3 Reduction of neural network classifier with Analysis of Finite Fluctuations

In this section is presented the application of the proposed approach to the problem of reduction of neural network model which is namely a classifier.
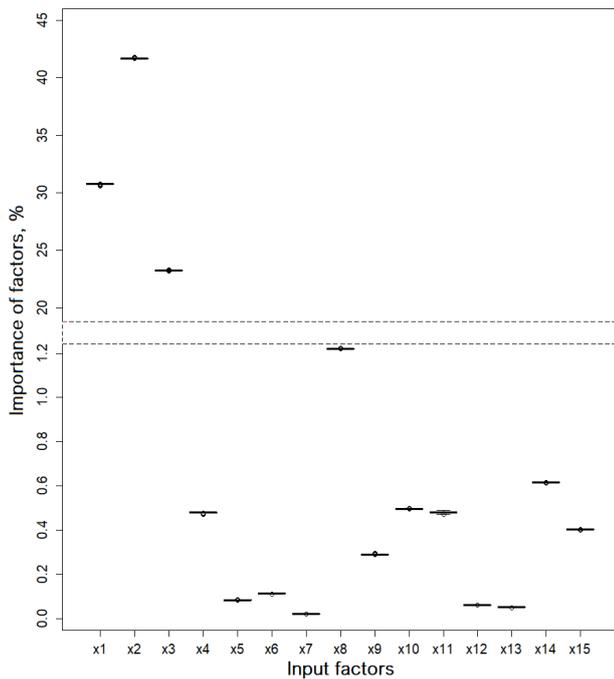
Fig. 2 Studying stability of the proposed approach

Fig. 3 Applying Analysis of Finite Fluctuations to the reduction of a neural network model

Fig. 3 presents the results of applying the proposed approach to the problem. Taking, for example, a 5 % importance rate as the threshold, the initial model could be reduced to the model with 5 input factors to accelerate the calculation time for fitting.

6 Conclusion and outlook

The reported study introduces the approach to Sensitivity Analysis of neural network models based on Analysis of Finite Fluctuations. It should be mentioned, that this approach is not limited only to these class of network. The model under consideration has to be differentiable function to apply the key theorem of the approach (Lagrange mean value theorem). In contrast to studied approach operating Sobol sensitivity coefficients the proposed approach does not use an approximation procedure to model statistical parameters of the studied structure and in contrast to Garson strategy, it operates with both parameters and factors of the studied model.

An open question concerning the proposed approach and the reduction of the model using it, is to study the decreasing of a simplified model accuracy. This rate has to be described in connection with the number of input factors of the studied model. The following researches also will be devoted to the extension of studied models and using another types of theorem connecting finite fluctuations of factors (for example, the first and the second Bonnet theorems).

Selected statistical indicators from the Lipetsk Regional Compulsory Medical Insurance Fund information system of personified medical care registration (Lipetsk region, Russia) were taken as input factors. The period of fixation was from 01.04.2018 to 30.04.2018 (570 111 cases). Based on the expert analysis there were chosen indicators characterizing uniquely each case of medical care. They are 14 different values, such as indicators uniquely identifying patients (e.g., the number of obligatory medical insurance policy and ages), medical care organization (e.g., the code used to identify which health care organization provides a medical help, the specialization of the medical staff provided analyzed type of care), specific features of a particular case (e.g., major diagnosis on the first, second, third stages of hospitalization, minor diagnosis, the length of the hospitalization) and some statistical information. Due to technical errors while filling databases with analytical information caused by human factors or various kinds of malicious intent, relevant (primarily from the point of view of insurance medicine) is the task of finding anomalies among the array of big data. This applied problem is solved by applying consequently Isolation forest algorithm to estimate the rate of anomaly for each observation and then take this rate as input factor and based on expert training dataset to fit neural network model which is able to predict anomalies in a new dataset. The model of structure (Eq. (6)) was applied. The accuracy of the fitted model is 89.01 %.
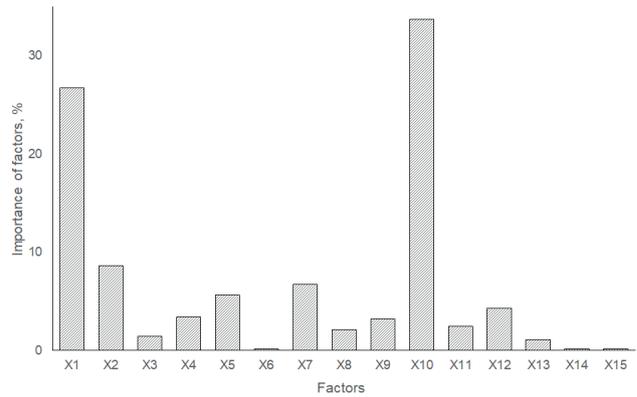
References

[1] Iooss, B., Saltelli, A. "Introduction to Sensitivity Analysis", In: Ghanem, R., Higdon, D., Owhadi, H. (eds.) Handbook of Uncertainty Quantification, Springer, Cham, Switzerland, 2017, pp. 1103–1122.
https://doi.org/10.1007/978-3-319-12385-1_31

[2] Saltelli, A. "Sensitivity Analysis for Importance Assessment", Risk Analysis, 22(3), pp. 579–590, 2002.
https://doi.org/10.1111/0272-4332.00040

[3] Blyumin, S. L., Borovkova, G. S., Serova, K. V., Sysoev, A. S. "Analysis of finite fluctuations for solving big data management problems", In: 2015 9th International Conference on Application of Information and Communication Technologies (AICT), Rostov on Don, Russia, 2015, pp. 48–51.
https://doi.org/10.1109/ICAICT.2015.7338514

[4] Blyumin, S., Galkin, A., Saraev, P., Sysoev, A. "Analysis of finite fluctuations as an approach to mathematical remodeling", Journal of Physics: Conference Series, 1202, article ID: 012025, 2019.
https://doi.org/10.1088/1742-6596/1202/1/012025

[5] Saraev, P. V., Blyumin, S. L., Galkin, A. V., Sysoev, A. S. "Neural Remodelling of Objects with Variable Structures", In: Abraham, A., Kovalev, S., Tarassov, V., Snasel, V., Vasileva, M., Sukhanov, A. (eds.) Proceedings of the Second International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'17), Advances in Intelligent Systems and Computing, vol. 679, Springer, Cham, Switzerland, 2018, pp. 141–149.
https://doi.org/10.1007/978-3-319-68321-8_15

[6] Woods, D. C., Lewis, S. M. "Design of Experiments for Screening", In: Ghanem, R., Higdon, D., Owhadi, H. (eds.) Handbook of Uncertainty Quantification, Springer, Cham, Switzerland, 2017, pp. 1143–1185.
https://doi.org/10.1007/978-3-319-12385-1_33

[7] Kurowicka, D., Cooke, R. "Uncertainty Analysis with High Dimensional Dependence Modelling", John Wiley & Sons Ltd, Chichester, England, 2006.
https://doi.org/10.1002/0470863072

[8] Cacuci, D. G., Ionescu-Bujor, M., Navon, I. M. "Sensitivity and Uncertainty Analysis: Applications to Large-Scale Systems", Vol. 2, CRC Press, Boca Raton, USA, 2005.
https://doi.org/10.1201/9780203483572

[9] Overstall, A. M., Woods, D. C. "Multivariate emulation of computer simulators: model selection and diagnostics with application to a humanitarian relief model", Journal of the Royal Statistical Society: Series C Applied Statistics, 65(4), pp. 483–505, 2016.
https://doi.org/10.1111/rssc.12141

[10] Saltelli, A., Tarantola, S., Campolongo, F. "Sensitivity Analysis as an Ingredient of Modeling", Statistical Science, 15(4), pp. 377–395, 2000.
https://doi.org/10.1214/ss/1009213004

[11] Maozhun, S., Ji, L. "Improved Garson algorithm based on neural network model", In: 2017 29th Chinese Control and Decision Conference (CCDC), Chongqing, China, 2017, pp. 4307–4312.
https://doi.org/10.1109/CCDC.2017.7979255

[12] Becker, W., Paruolo, P., Saisana, M., Saltelli, A. "Weights and Importance in Composite Indicators: Mind the Gap", In: Ghanem, R., Higdon, D., Owhadi, H. (eds.) Handbook of Uncertainty Quantification, Springer, Cham, Switzerland, 2017, pp. 1187–1216.
https://doi.org/10.1007/978-3-319-12385-1_40

[13] Le Gratiet, L., Marelli, S., Sudret, B. "Metamodel-Based Sensitivity Analysis: Polynomial Chaos Expansions and Gaussian Processes", In: Ghanem, R., Higdon, D., Owhadi, H. (eds.) Handbook of Uncertainty Quantification, Springer, Cham, Switzerland, 2017, pp. 1289–1325.
https://doi.org/10.1007/978-3-319-12385-1_38

[14] Blyumin, S. L., Borovkova, G. S., Serova, K. V., Sysoev, A. S. "Основы лагранжева анализа конечных изменний" (Basics of Lagrange Analysis of Finite Fluctuations), Lipetsk State Technical University Press, Lipetsk, Russia, 2016. (in Russian)

[15] Higham, N. J. "Accuracy and Stability of Numerical Algorithms", Society for Industrial and Applied Mathematics (SIAM), Philadelphia, USA, 2002.
https://doi.org/10.1137/1.9780898718027