

Computational Probabilistic Analysis and Transformation Method in the Earth Remote Research

Boris Dobronets^{1,*}, *Olga Popova*¹, and *Alexey Merko*¹

¹Institute of Space and Information Technology, Siberian Federal University, Kirenskogo 26, Krasnoyarsk, 660074 Russia

Abstract. The article deals with the issues of numerical modeling of problems with random input data. Finding the joint probability density function of the vector of output parameters is considered. It is proposed to use computational probabilistic analysis and the transformation method. A numerical example of the joint probability density function of the vector of a solution of a system of nonlinear equations with random input data is given.

1 Introduction

To solve many practical problems, including using Earth remote sensing (ERS) data, an approach based on numerical modeling is used. The issue becomes much more complicated when the input parameters of the mathematical model have various types of uncertainty, including random ones. In the case of random uncertainty of model parameters, the researcher must have knowledge of the probability density functions of random characteristics, the joint distribution function and the form of functional dependencies that establish model relationships between the input and output parameters.

The available knowledge about the properties of random parameters significantly affects the choice of methods and approaches for solving both particular problems of numerical modeling and the general problem as a whole, dictated by a practical problem.

One of the main ways to solve such problems is a numerical simulation by the Monte Carlo method. For example, this approach is used to simulate the interaction of solar radiation with the plant environment [1].

An important stage of modeling is the generation of sufficient statistics (usually tens, hundreds or more thousands) of trajectories of random parameters. Working with such a volume of information is characterized by computational complexity and low speed of convergence of the computational process.

An alternative to the Monte Carlo method is Computational Probabilistic Analysis (CPA), which offers methods for solving linear and nonlinear systems of equations with random parameters. The article discusses a mathematical model in the form of a system of functions with random arguments. A modified method is used to calculate the joint probability density function of the vector of output parameters. The approach is based on the use of CPA and the transformation method. Using the example of solving a system of nonlinear equations, an algorithm for constructing the joint density function of the solution vector is demonstrated. It uses a new concept of CPA called probabilistic extension.

*e-mail: BDobronets@yandex.ru

2 Transformation method

In this section, we will consider the application of the transformation method for calculating the joint probability density function of the vector of output parameters of a system of functions with random arguments [9].

Let the mathematical model be described by a system of functions

$$\begin{aligned} z_1 &= f_1(x_1, \dots, x_n), \\ &\vdots \\ z_m &= f_m(x_1, \dots, x_n), \end{aligned} \tag{1}$$

where $Z = (z_1, \dots, z_m)$, $X = (x_1, \dots, x_n)$, $m \leq n$.

The joint probability density function p_Z of a random vector Z can be found directly in terms of the joint probability density function p_X of a random vector X . This will require defining the equivalent events of infinitesimal rectangles.

For example, consider the linear transformation of random vectors for $m = n$

$$Z = AX.$$

Let us denote by A the matrix of dimension $n \times n$. Suppose that the matrix A has an inverse, therefore, each vector X corresponds to a unique Z obtained from

$$X = A^{-1}Z \tag{2}$$

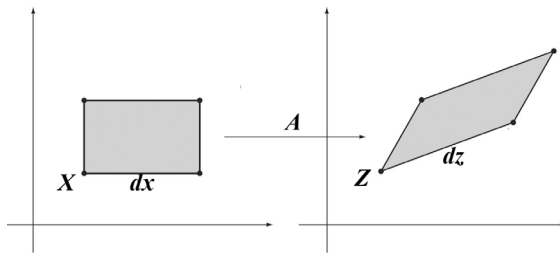


Figure 1. Transformation of an Infinitesimal Parallelepiped

Consider the infinitesimal rectangular parallelepiped dx shown in figure 1. Points in this rectangular parallelepiped are mapped to the parallelepiped dz . The infinitesimal rectangular parallelepiped and the parallelepiped are equivalent events, so their probabilities must be equal

$$p_X(X)|dX| = p_Z(Z)|dZ|,$$

where $|dX|$, $|dZ|$ are the areas of parallelepipeds. Thus, the joint probability density function for Z is defined as

$$p_Z(Z) = p_X(X)/(|dZ|/|dX|), \tag{3}$$

where random vectors X and Z are related to formula (2).

Equation (3) states that the joint probability density function for the random vector X is the joint probability density function for the random vector Z at the corresponding point, but

it is recalculated using the "stretch factor" $|dP|/|dX|$. Next, we will consider the general case $m < n$. For this purpose, we represent model (1) in the form

$$\begin{aligned}
 z_1 &= f_1(x_1, \dots, x_n), \\
 &\vdots \\
 z_m &= f_m(x_1, \dots, x_n), \\
 z_{m+1} &= f_{m+1}(x_1, \dots, x_n), \\
 &\vdots \\
 z_n &= f_n(x_1, \dots, x_n).
 \end{aligned}
 \tag{4}$$

Suppose that the functions f_1, \dots, f_m are one-to-one continuously differentiable and there is an inverse transformation

$$X = F^{-1}Z_1.$$

The differential of system (4) can be represented as

$$dZ_1 = F'(X)dX,$$

where $F'(X)$ is matrix of first derivatives. Then the joint probability density function $f_{Z_1}(Z_1)$ for a random vector Z_1

$$f_{Z_1}(Z_1) = f_X(F^{-1}Z_1)/|F'(F^{-1}Z_1)|.$$

Since, in the general case, the random vector $Z_1 = (z_1, \dots, z_m, \dots, z_n)$ contains fictitious components z_{m+1}, \dots, z_n , then to calculate the joint probability density function $f_Z(Z)$ it is necessary to integrate over fictitious variables.

In the next section we will illustrate this statement by the example of solving a system of nonlinear equations with random parameters.

3 System of nonlinear equations

Let us investigate the problem of calculating the solution set for a system of nonlinear equations

$$f_i(x, k) = 0, i = 1, \dots, n,$$

where $x \in R^n$ is the vector of solutions, $k \in R^m$ is the vector of parameters. Relatively $k \in R^m$, we will assume that the probability density functions are known. We will consider the case $m \geq n$. Note that the case of strict inequality $m > n$ can be reduced to the case $m = n$ of using the results of probabilistic extensions [8].

Let's consider a special case $m = n$. Differentiating the original system of nonlinear equations, we obtain

$$\begin{aligned}
 F'_x(x, k)dx + F'_k(x, k)dk &= 0, \\
 dx &= -(F'_x(x, k))^{-1}F'_k(x, k)dk.
 \end{aligned}$$

where $F(x, k) = (f_1(x, k), \dots, f_n(x, k))$.

Thus, knowing the solution x for some parameter values k , it is possible to obtain the relationship dx and dk .

This fact allows us to calculate the value of the joint probability density function of the solution. Consider an example of a system of nonlinear equations

$$\begin{aligned}
 x^2 + y^2 - r^2 &= 0, \\
 xy - c &= 0,
 \end{aligned}$$

where r, c are uniform random variables whose probability densities have supports $[1, 1.1], [0.4, 0.5]$.

Let x_0, y_0 be the solution of the system for some values r_0, c_0 , then it is true

$$\begin{pmatrix} 2x_0 & 2y_0 \\ y_0 & x_0 \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} 2r_0 dr \\ dc \end{pmatrix}.$$

In this case, the rectangle S_0 sides dr, dc goes into a quadrangle S_1 with sides dx, dy . Having solved the system, we get the values dx, dy . The probability density of the set of solutions at a point (x_0, y_0) is proportional to the ratio of the areas $|S_0|/|S_1|$:

$$p(x_0, y_0) = p_1(r_0, c_0)|S_0|/|S_1|.$$

Thus, in order to construct a joint density function of random variables, we construct grids in the region of carriers of random variables $r, c: \{r_i, i := 0, \dots, m\}, \{c_i, i := 0, \dots, m\}$.

Let us solve $(m + 1)^2$ systems of nonlinear equations:

$$x^2 + y^2 - r_i^2 = 0,$$

$$xy - c_j = 0, i, j = 0, \dots, m.$$

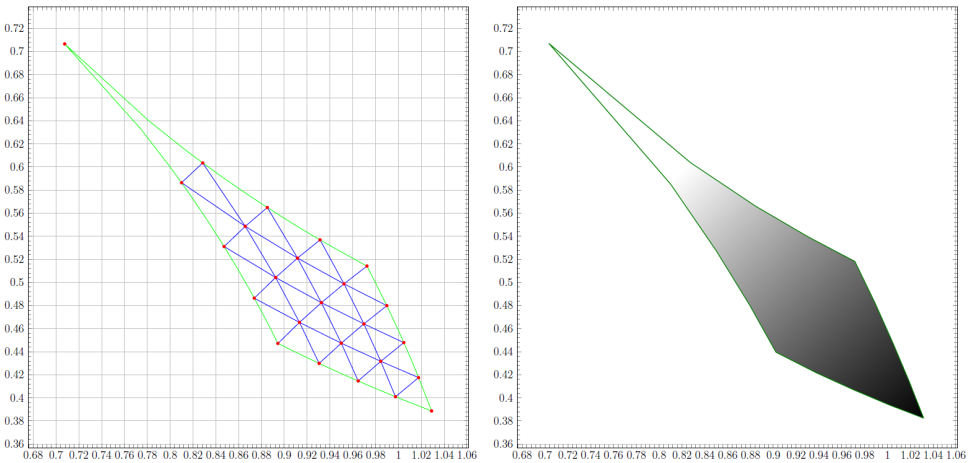


Figure 2. Triangulation of the domain and joint density function of the solution set

Figure 2 shows the triangulation of the solution domain for a system of nonlinear equations. The vertices of the triangles are the decision points with the calculated values of the density of the joint probability function of the solution $p(x_i, y_j)$. Thus, using linear interpolation, on each triangle, the value of the joint probability density function can be plotted. Shades of gray represent the joint density function of the solution to a system of nonlinear equations.

Note, to construct the joint density function of the solution, it is required to solve the system of nonlinear equations only 25 times; for the same purpose, the Monte Carlo method will require testing 10^6 . The general case of constructing probability densities can be realized by constructing probabilistic extensions.

4 Conclusion

The case is considered when the modeling relationships of input and output variables are presented in the form of systems of equations, both linear and nonlinear, with random uncertainties. The use of CPA and the transformation method makes it possible to find joint probability density functions of the output parameters of remote sensing problems. The method for constructing the joint probability density function of the solution removes restrictions on the requirement of independence of the input parameters, provides an analytical and graphical opportunity to construct the probabilistic region of possible values of the output parameters.

The proposed modified approach provides the results of numerical simulation with the required accuracy of calculations, which is guaranteed by the properties of numerical arithmetic procedures developed within the framework of the CPA.

References

- [1] S. Bartalev, V. Egorov, V. Zharko, E. Lupyán, D. Plotnikov, S. Khvostikov, N. Shabanov, Satellite mapping of the vegetation cover of Russia. Moscow: IKI RAN, (2016)
- [2] B. Dobronets, O. Popova. Computational Probabilistic Analysis: models and methods. Krasnoyarsk: Siberian Federal University, (2020)
- [3] B. Dobronets, O. Popova, Scientific Computing, Computer Arithmetic, and Validated Numerics. Lecture Notes in Computer Science 9553. Springer International Publishing, Cham, 43–53 (2016)
- [4] B. Dobronets, O. Popova, Journal of Siberian Federal University, Engineering and Technologies, 9(7), (2016), 960–971.
- [5] B. Dobronets, O. Popova, IOP Conf. Series: Journal of Physics: Conf. Series 1015 (2018) 032028 doi :10.1088/1742-6596/1015/3/032028
- [6] B. Dobronets, O. Popova, IOP Conf. Series: Materials Science and Engineering 354 (2018) 012006 doi:10.1088/1757-899X/354/1/012006
- [7] B. Dobronets, O. Popova, Journal of Siberian Federal University Mathematics and Physics, (2017) 10(1), pp. 16–21.
- [8] B. Dobronets, O. Popova, Tomsk State University Journal of Control and Computer Science 47 (2019) 41–48.
- [9] A. Leon-Garcia, Probability, Statistics, and Random Processes for Electrical Engineering. University of Toronto: Pearson Education, Inc. (2008)