Numerical Analysis for Problems of Remote Sensing with Random Input Data

Boris Dobronets[1,*] and *Olga* Popova[1,**]

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

Abstract. The study is devoted to the remote sensing data processing using the models with random input data. In this article we propose a new approach to calculation of functions with random arguments, which is a technique of fast computations, based on the idea of parallel computations and the application of numerical probability analysis. To calculate a function with random arguments we apply one of the basic concepts of numerical probabilistic analysis as the probabilistic extension. To implement the technique of fast computations, a new method based on parallel recursive calculations is proposed.

1 Introduction

Remote sensing data analysis often includes numerical study of models with random input data. Such models are traditionally studied by Monte Carlo methods. Monte Carlo method is a powerful approach, but it has some serious shortcomings. First of all, this is an extremely low rate of convergence. Non-Monte Carlo methods have been developed since the 1960s. A major non-Monte Carlo approach is interval analysis. However, interval analysis is computed only on the boundaries of random processes without examining their internal distributions. In our work, we develop a technique that uses Numerical Probabilistic Analysis (NPA) to solve various problems with stochastic data uncertainty.

The basis of numerical probabilistic analysis is numerical operations on probability density functions of the random values. These are operations "+", "−", "·", "/", "↑", "max", "min", as well as binary relations "≤", "≥" and some others. The numerical operations of the piecewise polynomial function arithmetic constitute the major component of NPA.

Using the arithmetic of probability density functions and probabilistic extensions, we can construct numerical methods that enable us solving systems of linear and nonlinear algebraic equations with random parameter [6].

In our approach we use the piecewise polynomial models to represent probability density functions:

- piecewise constant functions (histograms);

- piecewise linear functions (frequency polygons);

- piecewise polynomial functions (splines).

---

*e-mail: BDobronets@yandex.ru
**e-mail: OlgaArc@yandex.ru

In many cases problems of distance sensing can be reduced to the numerical analysis of functions

$$z = f(x_1, x_2, \ldots, x_n).$$

When the probability density of $x_1, x_2, \ldots, x_n$ is known, Monte Carlo method is used. Using NPA for these problems is a thousand times more effective than Monte Carlo method.

2 Probabilistic extensions

One of the most important problems that NPA deals with is to construct probability density functions of random variables.

Let us start with the general case when $(x_1, \ldots, x_n)$ is a system of continuous random variables with joint probability density function $p(x_1, \ldots, x_n)$ and the random variable z is a function $f(x_1, \ldots, x_n)$

$$z = f(x_1, \ldots, x_n).$$

By *probabilistic extension* of the function f we mean a probability density function of the random variable z.

Definition. The support of $\boldsymbol{f}$, written supp($\boldsymbol{f}$), is the set of points $x \in R$ where $\boldsymbol{f}$ is non-zero

$$\text{supp}(\boldsymbol{f}) = \{x \in R | \boldsymbol{f}(x) > 0\}.$$

Consider the procedure for computing the probabilistic extensions for the one-dimensional case. Let there be given a functional dependence

$$z = f(x),$$

where x is a random variable. Let $\boldsymbol{x}$ be the probability density function of a random variable x with support $[\underline{x}, \overline{x}]$. Further $\{x_i(z) \in [\underline{x}, \overline{x}] | i = 1, \ldots, n\}$ are the roots of the equation $z = f(x)$.

It is necessary to find the probability density function $\boldsymbol{f}$ of the random variable z as probabilistic extension $\boldsymbol{f}(\cdot, \boldsymbol{x})$ of function $f(x)$. To calculate probabilistic extension, we use

$$\boldsymbol{f}(\xi, \boldsymbol{x}) = \sum_{i=1}^{n} \frac{\boldsymbol{x}(x_i(\xi))}{|f'(x_i(\xi))|}.$$

Example. Consider function $z = x + t$, where t is real number. Then probabilistic extensions can be represented as

$$\boldsymbol{f}(\xi, \boldsymbol{x}, t) = \boldsymbol{x}(\xi - t).$$

If $z = tx$, and $t > 0$, then probabilistic extensions can be represented as

$$\boldsymbol{f}(\xi, \boldsymbol{x}, t) = \frac{1}{t}\boldsymbol{x}(\frac{\xi}{t}).$$

Theorem 1.[7] Let $\boldsymbol{f}(\cdot, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ be Probabilistic Extensions of function $f(x_1, x_2, \ldots, x_n)$ and for each real t function $\boldsymbol{f}(\cdot, t, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ be probabilistic extensions of the function $f(t, x_2, \ldots, x_n)$. Then

$$\boldsymbol{f}(z, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) = \int \boldsymbol{x}_1(t)\boldsymbol{f}(z, t, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)dt \tag{1}$$

Corollary. Theorem 1 infers the possibility of recursive computations for the general form of probability extensions and reduction to the calculation of the one-dimensional case.

Let us consider the computing the integral (1). For simplicity, we represent (1) as a quadrature

$$\int \boldsymbol{x}_1(t)\boldsymbol{f}(z, t, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)dt \approx \sum_{l=1}^{m} \gamma_l \boldsymbol{x}_1(t_l)\boldsymbol{f}(z, t_l, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$$

Further, we can also use numerical quadratures for computing $\boldsymbol{f}(z, t_l, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$. In general, it is NP-hard problem with actual parallelization.
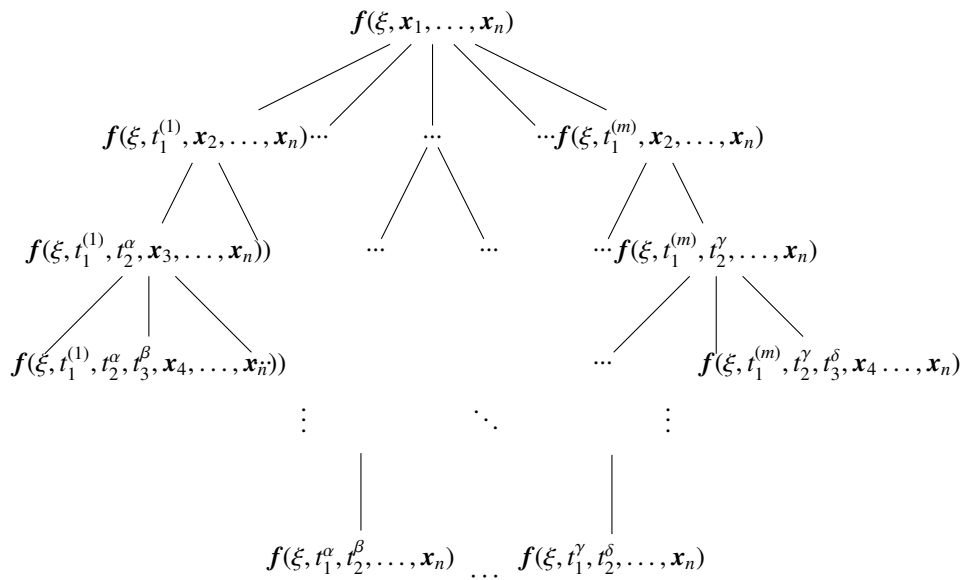
Figure 1. The tree of parallel recursive programming

In Figure 1 shows the tree of parallel recursive organization of the computational process. Thus, on the lower layer, it is necessary to compute the probabilistic extensions only for one variable. Note that all computations on each layer are independent and can be computed simultaneously.

Let x, y be system of two continuous random variables with probability density functions $\boldsymbol{x}, \boldsymbol{y}$.

Densities arising as a result of arithmetic operations on random variables have the form

$$(\boldsymbol{x} + \boldsymbol{y})(\xi) = \int_{\underline{x}}^{\overline{x}} \boldsymbol{x}(t)\boldsymbol{y}(\xi - t)dt;$$

if $\operatorname{supp}(\boldsymbol{x}) = [\underline{x}, \overline{x}]$ and $\underline{x} > 0$ then

$$(\boldsymbol{xy})(\xi) = \int_{\underline{x}}^{\overline{x}} \frac{1}{t}\boldsymbol{x}(t)\boldsymbol{y}(\frac{\xi}{t})dt,$$

$$(\boldsymbol{y}/\boldsymbol{x})(\xi) = \int_{\underline{x}}^{\overline{x}} t\boldsymbol{x}(t)\boldsymbol{y}(t\xi)dt.$$

These arithmetic operations completely coincide with the classical operations on the probability density functions.

Let $f(x_1, \ldots, x_n)$ be a rational function. To construct a probabilistic extensions of f, we replace the arithmetic operations by the arithmetic operation on random variables, while the variables x_1, x_2, $\ldots$, x_n are replaced by probability density functions. It makes sense to call the resulting probabilistic extensions of f as *natural probabilistic extension*.

Case 1. [7] Let $x_1, \ldots, x_n$ be independent random variables. If $f(x_1, \ldots, x_n)$ is a rational expression where each variable x_i occurs not more than once, then the natural probabilistic extension approximates a probabilistic extension.

Case 2. [7] Let the function $f(x_1, \ldots, x_n)$ can be a change of variables, so that $f(z_1, \ldots, z_k)$ is a rational function of the variables $z_1, \ldots, z_k$ satisfying the conditions of Case 1. The variable z_i is a function of x_i, $i \in Ind_i$. and Ind_i be mutually disjoint. Suppose for each z_i is possible to construct a probabilistic extension. Then the natural extension $\boldsymbol{f}(z_1, \ldots, z_k)$ would be approximated by a probabilistic extension $\boldsymbol{f}(x_1, \ldots, x_n)$.

Example. Let $f(x_1, x_2) = (-x_1^2 + x_1) \sin x_2$, and $z_1 = (-x_1^2 + x_1)$, $z_2 = \sin x_2$. Notice that it is possible to construct probabilistic extensions for the functions z_1, z_2, and then compute $f = z_1 * z_2$, which is a rational function satisfying the conditions of Case 1. So, natural extension will approximate probabilistic extension for the function $f(x_1, x_2)$.

3 Systems of linear algebraic equations

As one example of numerical simulation, let us consider solution of a system of linear algebraic equations

$$Ax = b, \tag{2}$$

where $A = (a_{ij})$ a random matrix and $b = (b_i)$ a random right-hand side vector respectively. Suppose that the random matrix A and the vector b have independent components with probability densities $\boldsymbol{A} = (\boldsymbol{a}_{ij})$, $\boldsymbol{b} = (\boldsymbol{b}_i)$ respectively and

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_{11} & \boldsymbol{a}_{12} & \ldots & \boldsymbol{a}_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_{n1} & \boldsymbol{a}_{n2} & \ldots & \boldsymbol{a}_{nn} \end{pmatrix}.$$

The support of the solution set can be represented as follows [6]

$$X = \{x | Ax = b, A \in \mathrm{supp}(\boldsymbol{A}), b \in \mathrm{supp}(\boldsymbol{b})\}.$$

Construct the probabilistic extension of the solution vector $\boldsymbol{x}(\cdot, \boldsymbol{A}, \boldsymbol{b})$

$$\boldsymbol{x}_1(\cdot, \boldsymbol{A}, \boldsymbol{b}) = \frac{\begin{vmatrix} \boldsymbol{b}_1 & \boldsymbol{a}_{12} & \ldots & \boldsymbol{a}_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{b}_n & \boldsymbol{a}_{n2} & \ldots & \boldsymbol{a}_{nn} \end{vmatrix}}{\begin{vmatrix} \boldsymbol{a}_{11} & \boldsymbol{a}_{12} & \ldots & \boldsymbol{a}_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_{n1} & \boldsymbol{a}_{n2} & \ldots & \boldsymbol{a}_{nn} \end{vmatrix}}$$

or

$$\boldsymbol{x}_1(\xi, \boldsymbol{A}, \boldsymbol{b}) = \iint \boldsymbol{a}_{12}(t_{12}) \ldots \boldsymbol{a}_{nn}(t_{nn}) \frac{\sum \boldsymbol{b}_i \Delta_i(t_{12}, \ldots, t_{nn})}{\sum \boldsymbol{a}_{1i} \Delta_i(t_{12}, \ldots, t_{nn})}(\xi) dt_{12} \ldots dt_{nn}, \tag{3}$$

where $\Delta_i(t_{12}, \ldots, t_{nn}) \in R$ are minors from the Cramer method for solving SLAE, $t_{ij} \in \mathrm{supp}(\boldsymbol{a}_{ij})$. The expression

$$\left(\frac{\sum \boldsymbol{b}_i \Delta_i(t_{12}, \ldots, t_{nn})}{\sum \boldsymbol{a}_{1i} \Delta_i(t_{12}, \ldots, t_{nn})} \right)(\xi)$$

is computed using probabilistic arithmetic.

Example. Consider a system of linear algebraic equations

$$Ax = b, \tag{4}$$

Let $A = (a_{ij})$ be random matrix $n = 2$. The elements of the matrix A are independent and distributed on the triangular law, a_{11}, a_{22} distributed on the interval $[2, 4]$, a_{21}, a_{12} distributed on the interval $[-1, 1]$. The vector b consists of independent components b_1, b_2, they are distributed according to a triangular law on the interval $[0, 2]$.
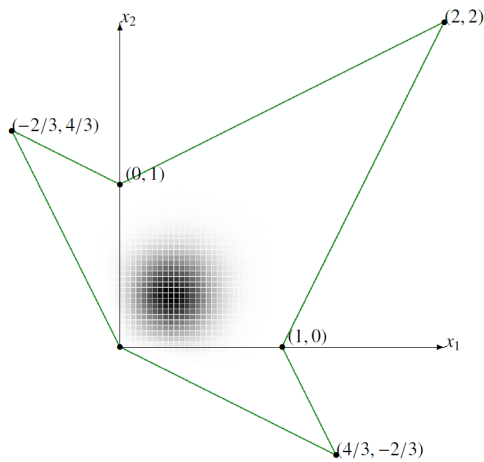
Figure 2. The boundary of the solution set and the joint probability density function of (x_1, x_2)
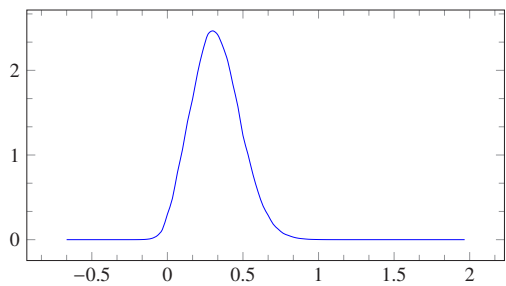
Figure 3. The probability density function of random variable x_1

The vector (x_1, x_2) solution of (4)

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}},$$

$$x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}}.$$

Figure 2 shows the joint probability density of the vectors (x_1, x_2). The value of the probability is represented by shades of gray. A solid line is the boundary of the solution set.

In this case, to compute x_1, we replace random variables a_{22}, a_{12} to the real t_{22}, t_{12}. Computing the integral numerically for different ξ

$$x_1(\xi) = \iint a_{22}(t_{22}) a_{12}(t_{12}) \left(\frac{t_{22} b_1 - t_{12} b_2}{a_{11} t_{22} - t_{12} a_{21}} \right)(\xi) dt_{22} dt_{12}.$$

we obtain probability density function of x_1.

Figure 3 shows the probability density function of random variable x_1. The random variable x_1 is distributed on interval $[-2/3, 2]$. However, significant values are on the interval $[0, 1]$.

4 Conclusion

The proposed approach makes it possible to solve the problem of computing the probability density function in the modeling processes with random input data. For these purposes we propose using a parallel-recursive organization of the computational process. Thus, the important problem of computing probability extensions can be solved within parallel recursive programming. This opens multifold possibilities for studying issues of Earth remote sensing with random input data. Fast and accurate calculations are based on the properties of numerical arithmetic procedures over piecewise polynomial models developed within the framework of numerical probabilistic analysis.

References

[1] Hamideh Ebrahimi, Shadi Aslebagh, W. Linwood Jones Proceedings of IEEE Southeast-con (2013). DOI: 10.1109/SECON.2013.6567506

[2] Dobronets B.S., Popova O.A. Journal of Siberian Federal University. Engineering & Technologies, **9**(7), 960–971, (2016)

[3] Dobronets B.S., Popova O.A. Journal of Siberian Federal University, Mathematics and Physics, **10** (1), 16–21 (2017).

[4] Dobronets B.S., Popova O.A. IOP Conf. Series: Journal of Physics: Conf. Series **1015** (2018) 032028 doi :10.1088/1742-6596/1015/3/032028

[5] Dobronets B.S., Popova O.A. IOP Conf. Series: Materials Science and Engineering **354** (2018) 012006 doi:10.1088/1757-899X/354/1/012006

[6] Dobronets B., Popova O. *Scientific Computing, Computer Arithmetic, and Validated Numerics. Lecture Notes in Computer Science* **9553**. Springer International Publishing, 43–53 (2016)

[7] Dobronets B.S., Popova O.A. Reliable Computing. **19**, 274–289 (2014).