SCAN — 2014

16th GAMM-IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics

Numerical probabilistic approach for optimization problems

Boris S. Dobronets & Olga A. Popova

Sibirian Federal University
E-mail: BDobronets@yandex.ru

Würzburg, 2014

Numerical Probabilistic Analysis (NPA)

- NPA is the section of Computing Mathematics.
- Subject of NPA is a decision of the problems with stochastic uncertainty in data.
- Methods of NPA use numerical operations under probability density functions of random variables and their functions.
- Numerical operations of histogram arithmetic is one of NPA components.
- The first idea histogram arithmetic was published in the article

V.A. Gerasimov, B.S. Dobronets, and M.Yu. Shustrov

Numerical operations of histogram arithmetic and their applications. Automation and Remote Control, (Feb 1991), 52(2), pp. 208–212.

- Arithmetic on probability density function uses operations as $* \in \{+, -, \cdot, /, \uparrow, \mathsf{max}, \mathsf{min}\}$, and binary relations as $\{\leq, \geq\}$.

The probability density function types

- Discrete random variables.
- Histograms.
- Second order histograms
- Splines.
- Analytically given probability density.

Probabilistic extensions

Definition 1.

Let $(x_1, ..., x_n)$ be a system of continuous random variables with joint probability density function $p(x_1, ..., x_n)$ and random variable z is the function $f(x_1, ..., x_n)$

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

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

Histogram probabilistic extensions

Suppose the histogram F is defined mesh $\{z_i | i = 0, ..., n\}$.
The region is denoted as $\Omega_i = \{(x_1, ..., x_n) | z_i < f(x_1, ..., x_n) < z_{i+1}\}$.
Then the histogram value F_i on the interval $[z_i, z_{i+1}]$ is defined as

$$F_i = \int_{\Omega_i} p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n / (z_{i+1} - z_i). \quad (1)$$

Definition 2.

By histogram probabilistic extension of the function f we mean an histogram F constructed from (1).

Natural histogram extensions

Let $f(x_1, ..., x_n)$ be rational function.
To construct of histogram of F replaced by the arithmetic operation on the histogram operation, and variables x_1, x_2,..., x_n replaced by histogram of values.

Definition 3.

The resulting histogram of F is called a natural histogram extension.

Histogram probabilistic extensions and arithmetic operations

Let P be a histogram of the probability density function $z = x * y$, and $* \in \{+, -, \cdot, /, \uparrow\}$. Then the value of P_i on the interval $[z_i, z_{i+1}]$ is defined by formula

$$P_i = \int_{\Omega_i} p(x, y) dxdy / (z_{i+1} - z_i), \tag{2}$$

where $\Omega_i = \{(x, y) | z_i \leq x * y \leq z_{i+1}\}$.

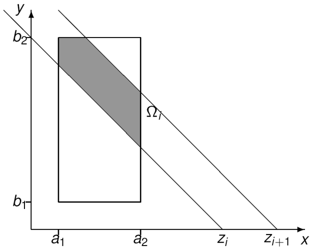
The histogram of the sum for two random variables

$$z = x + y,$$

then P_z is a histogram of the probability density function of z and

$$p_{zi} = \int_{\Omega_i} p(x, y) dx dy / (z_{i+1} - z_i). \quad (3)$$

Support of $p(x, y)$ is a rectangle $[a_1, a_2] \times [b_1, b_2]$ and
$\Omega_i = \{(x, y) | z_i \le x + y \le z_{i+1}\}$.



Theorem 1.

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 histogram extension approximates a probabilistic extension to $O(h^\alpha)$, $\alpha \geq 1$.

Theorem 2.

Let the function $f(x_1, ..., x_n)$ 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 Theorem 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 $f(z_1, \ldots, z_k)$ would be approximated by a probabilistic extension $f(x_1, ..., x_n)$.

$$f(x, y) = xy + x + y + 1 = (x + 1)(y + 1).$$

Example

Let $f(x_1, x_2) = (-x_1^2 + x_1)sin(x_2)$.
Then $z_1 = (-x_1^2 + x_1)$ and $z_2 = sin(x_2)$.
We shell notice that possible to construct a probabilistic extension for functions z_1, z_2 and $f = z_1 * z_2$ be a rational function satisfying the conditions of Theorem 1.
So natural extension will approximate probabilistic extension to function $f(x_1, x_2)$.

General case

Consider case when necessary to find probabilistic extension for function $f(x_1, x_2 ..., x_n)$ but conditions of Theorem 2 are not fulfilled.
Suppose for definiteness that only x_1 occurs a few times.
If instead of random variable x_1 to substitute determinate value t then possible construct natural probabilistic extension to function $f(t, x_2 ..., x_n)$.
Suppose t is an discrete random value approximating x_1 the following let t takes values t_i with probability P_i and each one function $f(t_i, x_2 ..., x_n)$ possible to construct natural probabilistic extension.
Then a probabilistic extension f of the function $f(x_1, ..., x_n)$ can be approximated by a probability density φ as follows:

$$\varphi(\xi) = \sum_{i=1}^{n} P_i \varphi_i(\xi).$$

Example

Let $f(x, y) = x^2 y + x$ and x, y are uniformly distributed on $[0, 1]$ interval random values.
We shell change x to discrete random value t, $\{t_i | t_i = (i - 0.5)/n, i = 1, 2, ..., n\}$, $P_i = 1/n$ and shell calculate natural probabilistic extensions φ_i.

Table 1. Approximating error of the probabilistic extensions

n	$\|\|f - \varphi\|\|_2$
10	1.2887825282E-03
20	4.5592973952E-04
40	1.6120775967E-04
80	5.6996092139E-05
160	2.0151185588E-05

Analysis of calculated results has shown that φ approximates f with $\alpha = 1.4998$, here α is approximation order.

Comparison of NPA and Monte Carlo Methods

Monte Carlo method displays convergence $1/\sqrt{N}$. Monte Carlo Errors reduce by a factor of $1/\sqrt{N}$. Where N is the number of sampled points.
Error of histogram extension is $O(1/n^{\alpha})$, $\alpha \geq 1$.
In practice using the histogram extensions is more efficient than Monte Carlo Methods more than $10^2 - 10^3$ times.

Necessary to find p the sum of four standard uniformly distributed random variables.

$$p(x) = \begin{cases} \frac{1}{6}x^3, & \text{in } 0 \le x \le 1; \\ -\frac{1}{2}x^3 + 2x^2 - 2x + \frac{2}{3}, & \text{in } 1 \le x \le 2; \\ \frac{1}{2}x^3 - 4x^2 + 10x - \frac{22}{3}, & \text{in } 2 \le x \le 3. \\ -\frac{1}{6}x^3 + 2x^2 - 8x + \frac{32}{3}, & \text{in } 3 \le x \le 4. \end{cases}$$

Let N be the number of sampled and n be dimension of mesh.
H_n is histogram probabilistic extension of p for n (exact histogram). P_n is natural histogram extension of p for n, $MC_{n,N}$ is histogram approximation of Monte Carlo method of p for n, N

Table 2. Errors of histogram arithmetic and Monte Carlo Methods

n	$N = 10^4$	$N = 10^5$	$N = 10^6$	$\|\|H_n - P_n\|\|_2$.
10	0.0059	0.00168	0.00037	4.16e-3
20	0.0055	0.00198	0.00041	5.39e-4
50	0.0026	0.00103	0.00026	3.47e-5
100	0.0023	0.00062	0.00018	4.35e-6
150	0.0016	0.00055	0.00016	1.28e-6
200	0.0014	0.00044	0.00014	5.44e-7

This table represents the approximation errors $||H_n - P_n||_2$ and $||H_n - MC_{n,N}||_2$. We can see that for a fixed n error of the Monte Carlo method decreases as $\approx 1/\sqrt{N}$, order of convergence natural histogram extension is $\alpha \approx 3.5$.

Second Order Histogram (SOH)

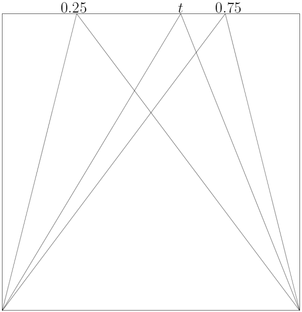
Definition 4.

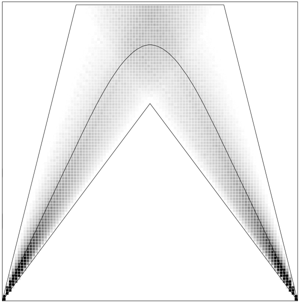
Second-order histogram is piecewise histogram function.

SOH is determined by the mesh $\{z_i | i = 1, 2, ..., n\}$ and set of histogram $\{P_i | i = 1, 2, ..., n\}$.
On each interval $[z_i, z_{i+1}]$ SOH is a histogram P_i.

Example 5. Second Order Histogram

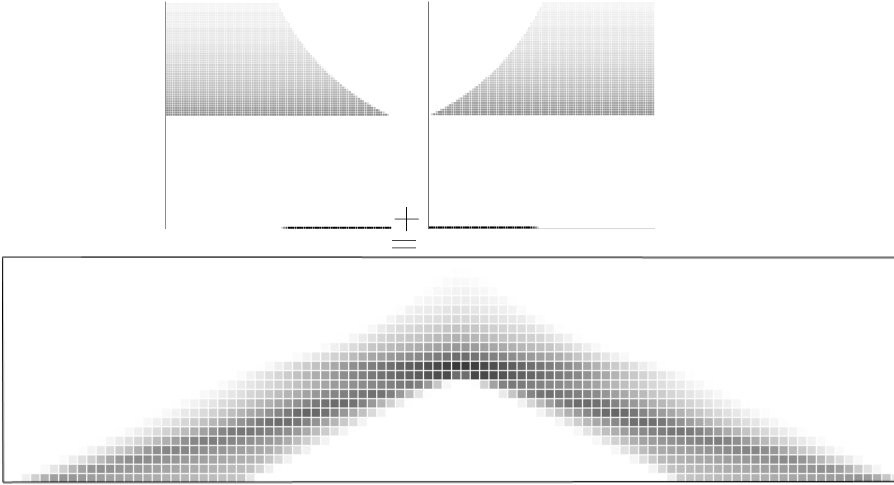
Let P_t be triangular distributed on $[0,1]$ random variable with height $h=2$ and top $(t,2)$. Let t be triangular distributed on $[0.25,0.75]$ random variable with top $(0.5,4)$.





The top and bottom lines corresponds to the interval histogram and the middle line correspond to the mean SOH.
Values probability densities are shades of gray.

The sum of two second order histograms



Values of probability densities are shades of gray.

The nonlinear equations

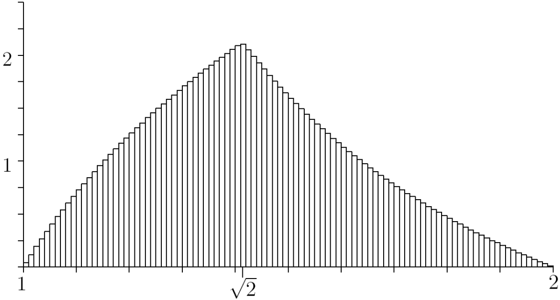
$$f(x, k) = 0,$$

where k — vector of random parameters, $x \in [a, b]$.
Let ϕ_z be probabilistic extensions of $f(z, \boldsymbol{k})$ and $z \in [a, b]$.
Then $P(z)$ is a probability that the root x is to the left (right) point z:

$$P(z) = \int_{-\infty}^{0} \phi_z(\xi) d\xi.$$

The nonlinear equations

$ax^2 - b = 0$, where a, b — random variable with uniform distribution on $[1, 2]$, $[2, 4]$.



Histogram of the root of the square equation

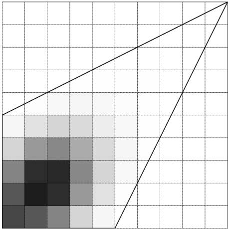
Solution of linear system equations

$$Ax = b.$$

Let $\boldsymbol{b}$ be random vector, and $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ be independing uniformly distributed components on $[0, 1]$ interval.
Suppose that matrix A is

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_{11} & \boldsymbol{a}_{12} \\ \boldsymbol{a}_{21} & \boldsymbol{a}_{22} \end{pmatrix}.$$

and component $\boldsymbol{a}_{11}$, $\boldsymbol{a}_{22}$ is independent random value uniformly distributed on $[2, 4]$ interval, a_{21}, a_{12} — $[-1, 0]$.



Piecewise constant with step 0.1 approaching the joint density probability $\boldsymbol{x}$. The solid line shows the boundary the set of solutions of the original system.

Random Programming

Let us formulate problem of random programming as follows:

$$f(x, \xi) \to \min, \tag{4}$$

$$g_i(x, \xi) \leq 0, \quad i = 1, ..., m. \tag{5}$$

where x is the solution vector, ξ is vector of parameters, $f(x, \xi)$ is objective function, $g_i(x, \xi)$ are constraint functions.
Relative to ξ is known that

$$\xi \in \boldsymbol{\xi}, \tag{6}$$

where $\boldsymbol{\xi}$ is random vector.
Vector x^* is the solution of problem (4) – (6), if

$$f(x^*, \xi) = \inf_U f(x, \xi),$$

где

$$U = \{x | g_i(x, \xi) \leq 0, \quad i = 1, ..., m.\}$$

The solution set of (4)– (6) is defined as follows

$$\mathcal{X} = \{x | f(x, \xi) \to \min, g_i(x, \xi) \leq 0, \quad i = 1, ..., m, \xi \in \boldsymbol{\xi}\}$$

Example

$$(c, x) \to \min, \tag{7}$$

$$Ax = b, x \geq 0. \tag{8}$$

$$A \in \boldsymbol{A}, b \in \boldsymbol{b}, c \in \boldsymbol{c}, \tag{9}$$

where $\boldsymbol{A} = (\boldsymbol{a}_{ij})$ is uniform random matrix, each element are the uniform random variable with support $[\underline{a}_{ij}, \overline{a}_{ij}]$, the same $\boldsymbol{b}, \boldsymbol{c}$ is random vector with elements in a uniform random variables.

Supports

$$A = \left(\begin{array}{cc} [1-r, 1+r] & [1-r, 1+r] \\ [1-r, 1+r] & [-1-r, -1+r] \end{array} \right.$$

$$\left. \begin{array}{cc} [3-r, 3+r] & [1-r, 1+r] \\ [1-r, 1+r] & [2-r, 2+r] \end{array} \right),$$

$$b = \left(\begin{array}{c} [3-r, 3+r] \\ [1-r, 1+r] \end{array} \right),$$

$$c = ([-1-r, -1+r], [-1-r, -1+r], [-r, +r], [-r, +r]).$$

If $r = 0$, which corresponds to the deterministic case, the solution $x^* = (2, 1, 0, 0)$, columns of the matrix A_1, A_2 correspond to the angular point.

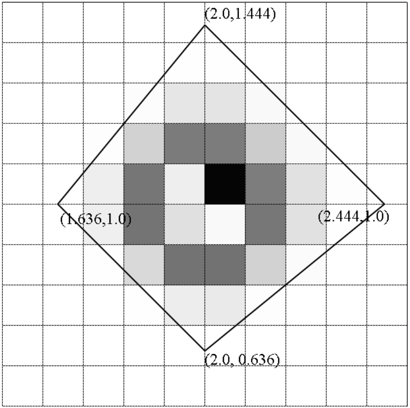


Рис.: Joint density of the vector $\boldsymbol{x}_1, \boldsymbol{x}_2$

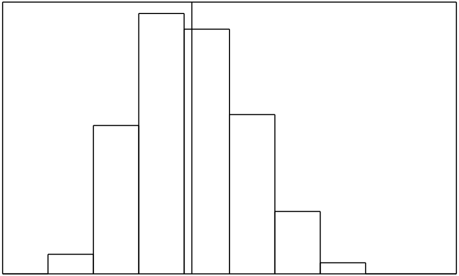


Рис.: Histogram $c_1 x_1 + c_2 x_2$

Fig. 2 is a graph of the objective function $c_1 x_1 + c_2 x_2$, expectation at - 2.834.

Random nonlinear programming

$$\frac{1}{2}(Ax, x) - (b, x) \to \min. \tag{10}$$

$$A \in \boldsymbol{A}, b \in \boldsymbol{b}, \tag{11}$$

where $\boldsymbol{A}$ is random matrix, $\boldsymbol{b}$ is random vector. Problem (10),(11) in the case of symmetric positive definite matrix $\boldsymbol{A}$ is reduced to solving a random system of linear algebraic equations

$$\boldsymbol{Ax} = \boldsymbol{b}. \tag{12}$$

Numerical examples

In the problem (10) $\boldsymbol{A}$ is uniform random matrix

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 \\ \boldsymbol{a}_2 & \boldsymbol{a}_1 \end{pmatrix},$$

$\boldsymbol{b}$ is uniform random vector. Supports $\boldsymbol{a}_1 = [2,4]$, $\boldsymbol{a}_2 = [-1,0]$, $\boldsymbol{b}_1 = \boldsymbol{b}_2 = [0.5,1]$.

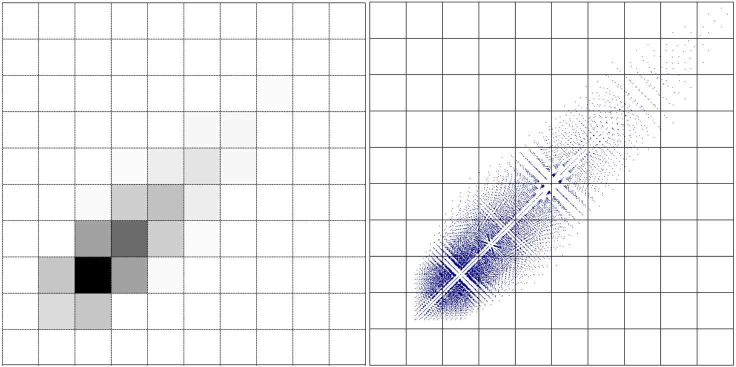


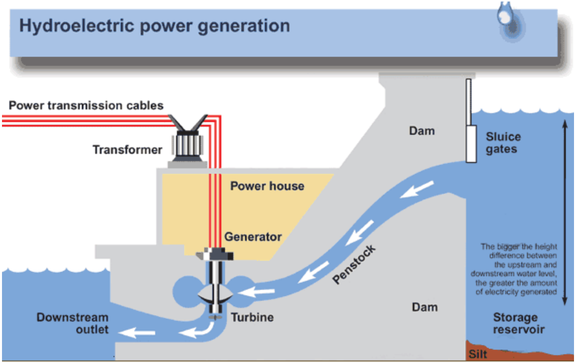
Рис.: Joint density of the vector $\boldsymbol{x}$. Samples of solutions (12)

Optimization of hydroelectric power generation

Power generating electricity p can be represented

$$p = Chu,$$

where
C — constant;
h — height of the water level, $h \in [h_{min}, h_{max}]$,
u — water passing through the turbine, $u \in [u_{min}, u_{max}]$.

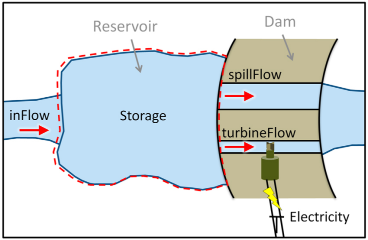


Statement of the problem

Height h depends on the amount of water in the reservoir of V:

$$h = h(V).$$

$$V(t) = V_0 + \int_0^t q(\xi) - u(\xi) - u_x(\xi) d\xi.$$



$q(t)$ — inflow;
$u_x(t)$ — water passing through the spillway (known and is determined by plant personnel);
u — water passing through the turbine, $u \in [u_{min}, u_{max}]$.

Statement of the problem

Suppose we want to maximize the generation of electricity in the time interval $[0, T]$. The task of optimal control

$$P(u) = \int_0^T C\ h\left(V_0 + \int_0^T q(t) - u(t) - u_x(t)dt\right)\ u(t)dt \rightarrow \max,$$

where u — control.

Simplify the problem

Volume of the reservoir

$$V = V_0 + S(h - h_0),$$

$$h(t) = h_0 + (V(t) - V_0)/S = h_0 + (\int_0^t q(\xi) - u(\xi) - u_x(\xi)d\xi)/S.$$

$$P(u) = C\int_0^T \left(h_0 + (\int_0^t q(\xi) - u(\xi) - u_x(\xi)d\xi)/S\right) u(t)dt \to \max.$$

$q(t)$ — inflow;
$u_x(t)$ — water passing through the spillway;
u — water passing through the turbine, $u \in [u_{min}, u_{max}]$.

The inflow of water in the reservoir

$$q = f(x_1, x_2, \ldots, x_n),$$

где x_i — input parameters (rainfall, moisture content in soil, humidity, temperature, etc.)

Data acquisition

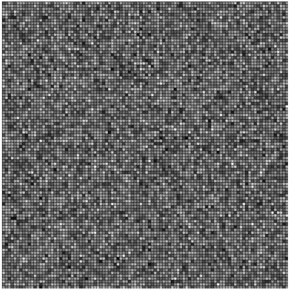
- Ground-based detection methods, observation and data collection.
- Space means of recording, monitoring and data collection
- Earth observation satellite.
- Remote sensing of the earth.

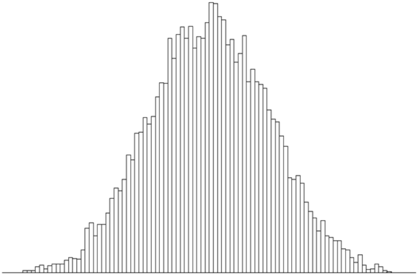
Monitoring the Earth from Space

- as a way of recording, data collection, transmission, processing, storage and storage of the information received;

Example of data aggregation

Histogram of P is constructed from the values of Ω.
Region Ω consists of 100×100 pixels, each pixel is mapped value x_i.
For clarity, in Fig. values are shades of gray. Lighter colors correspond to higher temperatures. Thus, the histogram describes the frequency distribution of temperature in the region Ω.





Model

$$P(u) = C \int_0^T \left(h_0 + \left(\int_0^t q(\xi) - u(\xi) - u_x(\xi) d\xi \right) / S \right) \, u(t) dt \to \max .$$

$q(t)$ — inflow;
$u_x(t)$ — water passing through the spillway;
u — water passing through the turbine.

Discrete model

Mesh $\omega = \{t_0 < t_1 < \ldots < t_n\}$,
$\{\boldsymbol{q}_i | t \in [t_{i-1}, t_i]\}$ — histogram of water inflow during $[t_{i-1}, t_i]$,
$\{u_{xi} | t \in [t_{i-1}, t_i]\}$ — water passing through the spillway during $[t_{i-1}, t_i]$,
$\boldsymbol{U} = \{\boldsymbol{u}_i | t \in [t_{i-1}, t_i]\}$ — histogram of water passing through the turbine during $[t_{i-1}, t_i]$.

$$P(\boldsymbol{U}) = C \sum_{i=1}^{n} \left(h_0 + (\sum_{j=1}^{i} q_j - u_j - u_{xj})/S \right) \ u_i \to \max .$$

Random system of linear algebraic equations

Problem in some cases can be reduced to the solution of random system of linear algebraic equations In our case, random only right-hand side of the system

$$2u_1 + u_2 + \ldots + u_n = h_0 + \boldsymbol{q}_1,$$

$$u_1 + 2u_2 + \ldots + 2u_i + \ldots + u_n = h_0 + \boldsymbol{q}_1 + \boldsymbol{q}_2,$$

$$u_1 + u_2 + \ldots + 2u_i + \ldots + u_n = h_0 + \sum_{j=1}^{i} \boldsymbol{q}_j,$$

$$u_1 + u_2 + \ldots + 2u_n = h_0 + \sum_{j=1}^{n} \boldsymbol{q}_j.$$

Random system of linear algebraic equations

Can be expressed in $\boldsymbol{u}_i$, $i = 1, \ldots, n$ as a linear combination of $\boldsymbol{q}_i$, $i = 1, \ldots, n$.
For $n = 3$ optimal control:

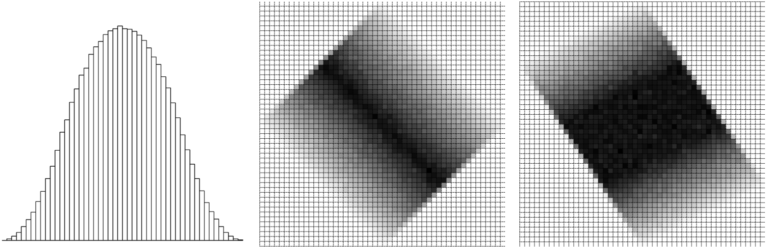
$$\boldsymbol{u}_1 = \frac{-\boldsymbol{q}_3 - 2\boldsymbol{q}_2 + \boldsymbol{q}_1 + h_0}{4},$$

$$\boldsymbol{u}_2 = \frac{-\boldsymbol{q}_3 + 2\boldsymbol{q}_2 + \boldsymbol{q}_1 + h_0}{4},$$

$$\boldsymbol{u}_3 = \frac{3\boldsymbol{q}_3 + 2\boldsymbol{q}_2 + \boldsymbol{q}_1 + h_0}{4}.$$

Numerical example

Let $\boldsymbol{q}_i \in [\underline{q}_i, \overline{q}_i]$ be uniform random variables, $n = 3$, $S = 1$, supports $q_1 = [0.1, 0.2]$, $q_2 = [0.2, 0.3]$, $q_3 = [0.3, 0.4]$ and $h_0 = 0.9$.



Histogram $\boldsymbol{u}_1$ and joint probability density (u_1, u_2), (u_2, u_3).
Supports $u_1 = [0.0, 0.1]$, $u_2 = [0.25, 0.35]$, $u_3 = [0.575, 0.725]$.

Olga A. Popova Optimization Problems with Random Data // Journal of Siberian Federal University. Mathematics & Physics 2013, 6(4), 506–515

Boris S. Dobronets and Olga A. Popova Numerical Probabilistic Analysis under Aleatory and Epistemic Uncertainty // Reliable Computing. 2014. Vol. 19.pp. 274—289.