Імітаційне моделювання систем – мистецтво і наука. – М., 1978. 6. Затраты на топливно-энергетические ресурсы в структуре ВВП выросли. (21.11.2000), 17:18.: http:// www. liga. kiev. ua/lenta/get.html?id=3998, Copyright by Liga Online © 2000. 7. Нечеткие множества в моделях управления и искусственного интеллекта / Под ред. Д.А. Поспелова. – М., 1986. 8. Fishman G.S., Kiviat P.J., The Analisys of Simulation-generated Time Series, Management Science. – Mar. 1967. – Vol. 13. № 7. 9. Козубовский С.Ф., Куприянов В.В. Методы векторной оптимизации при управлении сложными системами // Автоматика. – 1993. – № 4. – С. 72–77. 10. Зайченко Ю.П. Исследование операций. – К., 1988.

УДК 621.3.011:31.004

О.В. Денисюк, Б.Г. Марченко
Інститут електродинаміки НАН України

ПРО ПОБУДОВУ СТОХАСТИЧНОЇ ОЦІНКИ БАГАТОВИМІРНИХ ФУНКЦІЙ РОЗПОДІЛІВ І ЩІЛЬНОСТЕЙ РОЗПОДІЛІВ ІЗ ВИКОРИСТАННЯМ ТРИГОНОМЕТРИЧНИХ СПЛАЙНІВ

У роботі розглядається метод оцінювання багатовимірних функцій розподілів і щільностей розподілів, заснований на застосуванні двовимірних систем фундаментальних тригонометричних сплайнів.

In work a method evaluation of multidimensional functions of distribution and density of distribution, based on application multidimensional systems of fundamental trigonometrically splines.

У багатьох задачах ймовірнісно-статистичного аналізу стану і керування об'єктами електроенергетики часто доводиться будувати оцінки багатовимірних функцій розподілів ймовірностей і багатовимірних щільностей розподілу ймовірностей за вибірковими даними. Так, наприклад, ця задача виникає при обробці статистичних даних, отриманих під час дослідження характеристик дизель-електричних генераторів, при розв'язанні задач, пов'язаних з проблемою енергозбереження тощо.

Сплайн-функції в задачах одновимірного ймовірнісно-статистичного аналізу застосовуються здавна; достатньо нагадати про полігони відносних частот і гістограми, які є не чим іншим, як поліноміальними сплайнами відповідно першого і нульового порядку. До доцільності застосування багатовимірних сплайнів ми приходимо, розглядаючи задачі про наближення багатовимірних функцій розподілу і щільностей розподілу.

Поліноміальні сплайни, як апарат наближення, мають вади, серед яких відзначимо зростаючу складність перетворень, необхідних для їх побудови, з зростанням степеня і розмірності використовуваних сплайнів. Крім того, важливим недоліком поліноміальних сплайнів є те, що вони є кусково-складеними функціями, тобто не мають єдиного аналітичного виразу на всьому інтервалі наближення [1]. Проте поліноміальні сплайни широко використовують у задачах оцінювання функцій розподілів і щільностей розподілів [2, 3].

У цій роботі розглянуто метод побудови наближень багатовимірних функцій розподілів ймовірностей і багатовимірних щільностей розподілу ймовірностей за експериментальними даними, із застосуванням фундаментальних тригонометричних сплайнів [4], які є вільними від вказаних недоліків поліноміальних сплайнів.

Доцільність застосування тригонометричних сплайнів у задачах моделювання багатовимірних функцій розподілу та щільності розподілу пояснюється тим, що ці сплайни дають можливість отримати єдиний аналітичний вираз цих функцій, що має неабияке значення у багатьох задачах теорії ймовірності та математичної статистики [5]. Розглянемо побудову оцінок багатовимірних функцій розподілу і щільностей розподілу за допомогою фундаментальних тригонометричних сплайнів детальніше.

Сплайни, що наближують багатовимірну функцію розподілу будемо називати багатовимірними сплайн-розподілами, а сплайни, що наближують багатовимірну щільність розподілу – багатовимірними гістосплайнами. Надалі, не втрачаючи загальності, обмежимося розглядом наближення лише двовимірних функцій розподілу і щільностей розподілу.

Розглянемо таку задачу. Нехай $D = D[0,1;\ 0,1]$ – двовимірна прямокутна область. Якщо область $D_1 = D[a,b;\ c,d]$, де a, b, c, d – дійсні числа, то в цьому разі лінійною заміною змінних нескладно перейти до області D.

На області D задамо рівномірну сітку $\Delta_{N,M} = \{x_i, y_j\}_{i=0\ j=0}^{N\ \ M}$, де N і M непарні; $0 = x_0 < x_1 < \ldots < x_N = 1$; $0 = y_0 < y_1 < \ldots < y_M = 1$; ця сітка має однакові кроки по x:

$$h_x = x_n - x_{n-1}, \quad n = 1,2,\ldots, N$$

і однакові кроки по y:

$$h_y = y_m - y_{m-1}, \quad m = 1,2,\ldots, M$$

проте ці кроки є різними. Тому таку сітку ми будемо називати квазідистантною.

Нехай також на області задано деяку функцію розподілу $F(x, y)$ двох випадкових величин, причому є відомими значення цієї функції $\{F(x_i, y_j)\}_{i=0\ j=0}^{N\ \ M} = \{F_{i,j}\}_{i=0\ j=0}^{N\ \ M}$ у вузлах сітки $\Delta_{N,M}$. Ставиться задача побудови функції двох змінних $G(x,y)$, яка у вузлах сітки $\Delta_{N,M}$ збігалася б із значеннями функції розподілу $F(x,y)$, тобто, щоб виконувалася умова

$$G(x_i, y_j) = F_{i,j}.$$

Функцію $G(x,y)$ будемо шукати у вигляді

$$G(x, y) = \sum_{i=0}^{N}\sum_{j=0}^{M} F_{i,j} \cdot u_{i,j}(x, y), \tag{1}$$

де $u_{i,j}(x, y)$ – система двовимірних фундаментальних тригонометричних сплайнів.

Розглянемо побудову функції $G(x, y)$, що інтерполює функцію розподілу $F(x, y)$ на області D.

Крок по вісі x дорівнює $h_x = \frac{1}{N}$, а по вісі y дорівнює $h_y = \frac{1}{M}$, тоді $x_i = h_x \cdot i$; $y_j = h_y \cdot j$. Кожному вузлу сітки відповідає значення функції розподілу $F(x_i, y_j) = F_{i,j} = \beta_{i,j}$. З властивостей неперервних багатовимірних функцій розподілу випливає, що скінченний набір значень $\beta_{i,j}$; $i = 0,1,\dots,N$, $j = 0,1,\dots,M$ повинен задовольняти такі умови [5, 6]:

а) $\beta_{i,j} \ge 0 \quad (i = 0,1,\dots,N;\ j = 0,1,\dots,M)$;

б) при фіксованому значенні j, $j = 0,1,\dots,M$, $F_{i,j} \le F_{i+1,j}$, $(i = 0,1,\dots,N-1)$;

в) при фіксованому значенні i, $i = 0,1,\dots,N$, $F_{i,j} \le F_{i,j+1}$, $(j = 0,1,\dots,M-1)$;

г) набір значень $\beta_{i,j}$; $i = 0,1,\dots,N$ має бути оцінкою функції розподілу, тобто

$$\beta_{0,0} = \beta_{0,j} = \beta_{i,0} = 0; \quad \beta_{N,M} = 1; \quad \beta_{N,j} = F_1(y_j); \quad \beta_{i,M} = F_2(x_i),$$

де $F_1(y)$, $F_2(x)$ – одновимірні маргінальні функції розподілу для $F(x,y)$, тобто

$$F_1(y) = F(\infty,\ y); \quad F_2(x) = F(x,\ \infty).$$

На області D задамо також сітку $\Delta^*_{N,M} = \{\xi_i, \eta_j\}_{i=1}^{N}{}_{j=1}^{M}$, вузли якої обчислюються за формулами

$$\xi_i = \frac{1}{2}(x_i + x_{i-1}); \quad \eta_j = \frac{1}{2}(y_j + y_{j-1}). \qquad (2)$$

Якщо у вузлах сітки $\Delta^*_{N,M}$ є відомими значення щільності розподілу $p(x, y)$, то ставиться задача про побудову функції двох змінних $g(x, y)$, яка у вузлах сітки $\Delta^*_{N,M}$ збігалася б із значеннями щільності розподілу $p(x, y)$, тобто, щоб виконувалася умова

$$g(\xi_i, \eta_j) = p_{i,j}.$$

Кожному вузлу сітки відповідає значення функції щільності розподілу $p(\xi_i, \eta_j) = p_{i,j} = \alpha_{i,j}$. Значення $\alpha_{i,j}; i = 0,1,\dots,N; j = 0,1,\dots,M$, – це скінченний набір чисел [6].

В одновимірному випадку числа β_i, пов'язані з числами α_i, співвідношенням

$$\beta_i = \sum_{k=0}^{i} \alpha_k, \quad i = 0,1,\dots,N.$$

У багатовимірному випадку цей зв'язок значно ускладнюється. Так, для двовимірного випадку цей зв'язок набуває форми

$$\beta_{i,j} = h_x h_y \sum_{k=1}^{i} \sum_{m=1}^{j} \alpha_{k,m}; \qquad \alpha_{i,j} = h_x h_y (\beta_{i,j} + \beta_{i-1,j-1} - \beta_{i-1,j} - \beta_{i,j-1});$$

При наближенні функцій розподілу і щільності розподілу сплайнами виходять з таких умов:

а) оцінка функції розподілу повинна мати мішану похідну другого порядку;

в) ці наближення повинні зберігати міру на сітках по змінних x і y, яка для функцій розподілу при фіксованому значенні i (j) визначається так:

$$Gs(x_i, y_j) - Gs(x_i, y_{j-1}) = \beta_{i,j} - \beta_{i,j-1} = h_x h_y \sum_{k=1}^{j} \alpha_{i,k}$$

$$\left(Gs(x_i, y_j) - Gs(x_{i-1}, y_j) = \beta_{i,j} - \beta_{i-1,j} = h_x h_y \sum_{k=1}^{i} \alpha_{k,j} \right) \qquad (4)$$

а для гістосплайнів – вигляд

$$\int_{x_{i-1}}^{x_i} gs(x,\, y_j)dx = \beta_{i,j} - \beta_{i-1,j}; \qquad \left(\int_{y_{j-1}}^{y_j} gs(x_i,\, y)dy = \beta_{i,j} - \beta_{i,j-1} \right), \tag{5}$$

де відповідно через $Gs(x,y)$ позначено сплайн-розподіл, а через $gs(x,y)$ – гістосплайн.

Для наближення щільності розподілу можна застосувати два методи. При одному з них інтерполюються значення функції щільності розподілу на сітці $\Delta^*_{N,M}$, а функція розподілу оцінюється інтегруванням отриманої оцінки. При другому інтерполюються значення функції розподілу на сітці $\Delta_{N,M}$, а функція щільності розподілу оцінюється диференціюванням отриманої оцінки по кожній змінній.

У цій роботі, ми будемо наближувати функцію розподілу, а наближення щільності розподілу будемо отримувати за допомогою диференціювання функції розподілу по обох змінних.

Будемо наближувати двовимірну функцію розподілу виразом

$$G(x,\, y) = \sum_{i=0}^{N}\sum_{j=0}^{M} F_{i,j} St_{i,\,j}(r_1,\, r_2,\, N,\, M,\, x,\, y) = \sum_{i=0}^{N}\sum_{j=0}^{M} F_{i,j} St_i(r_1,\, N,\, x) St_j(r_2,\, M,\, y),$$

де $St_k(r, L, t)$ фундаментальний тригонометричний сплайн порядку r, який визначається за формулами

$$St_k(r,\, L,\, t) = \frac{1}{L}\left\{ 1 + 2\sum_{p=1}^{m} g_p^{-1}(r,L)\left[C_p(r,t)\cos 2\pi p t_k + S_p(r,t)\sin 2\pi p t_k \right] \right\}, \tag{6}$$

де

$$C_p(r,t) = \frac{\cos 2\pi pt}{(2\pi p)^{r+1}} + \sum_{m=1}^{\infty}\left[\frac{\cos 2\pi(mL+p)t}{[2\pi(mL+p)]^{r+1}} + \frac{\cos 2\pi(mL-p)t}{[2\pi(mL-p)]^{r+1}} \right];$$

$$S_p(r,t) = \frac{\sin 2\pi pt}{(2\pi p)^{r+1}} + \sum_{m=1}^{\infty}\left[\frac{\sin 2\pi(mL+p)t}{[2\pi(mL+p)]^{r+1}} - \frac{\sin 2\pi(mL-p)t}{[2\pi(mL-p)]^{r+1}} \right];$$

$$g_p(r,L) = -(2\pi p)^{-(r+1)} + (2\pi L)^{-(r+1)}\left[\zeta\left(r+1, pL^{-1}\right) + \zeta\left(r+1, -pL^{-1}\right) \right];$$

$\zeta(.,.)$ – узагальнена дзета-функція.

Побудований таким способом фундаментальний тригонометричний сплайн наближує двовимірну функцію розподілу і має неперервні частинні похідні $r_1 - 1$ і $r_2 - 1$ порядку по змінних відповідно x і y. Оскільки, цей сплайн зберігає міру (4) по обох змінних, то він є сплайн-розподілом.

Для отримання гістосплайна знайдемо другу мішану похідну від функції $G(x, y)$. Внаслідок диференціювання отримуємо

$$g(x, y) = \frac{\partial^2 G(x, y)}{\partial x \, \partial y} = \sum_{i=0}^{N} \sum_{j=0}^{M} \beta_{i,\,j} \frac{\partial^2 St_{i,\,j}(r_1, r_2, N, M, x, y)}{\partial x \, \partial y} =$$

$$= \sum_{i=0}^{N} \sum_{j=0}^{M} \beta_{i,\,j} \frac{dSt_i(r_1, N, x)}{dx} \frac{dSt_j(r_2, M, y)}{dy},$$

де $\dfrac{\partial St_k(r, L, t)}{\partial t}$ неважко одержати частинним диференціюванням по відповідній змінній виразу (6).

Висновки. Запропоновано метод наближення гістограм функцій розподілу і щільностей цих розподілів, Цей метод має переваги порівняно з аналогічними методами на базі поліноміальних сплайнів, а саме:

а) застосування фундаментальних тригонометричних сплайнів дозволяє досить просто забезпечувати бажані диференціальні властивості отримуваних наближень функцій розподілів і щільностей розподілів.

б) метод наближення функцій розподілів і щільностей розподілів фундаментальними тригонометричними сплайнами є більш гнучким, ніж метод, заснований на наближеннях поліноміальними сплайнами, оскільки клас розглядуваних нами фундаментальних тригонометричних сплайнів є значно ширшим і включає, як підмножину, клас поліноміальних сплайнів.

в) метод наближення розподілів і щільностей розподілів фундаментальними тригонометричними сплайнами припускає уніфікацію розроблених на його основі прикладних пакетів алгоритмів і програм обробки статистичної інформації під час розв’язання задач енергозбереження.

1. Завьялов Ю.С., Квасов Б.И., Мирошниченко В.Л. Методы сплайн-функций. – М., 1980. – 350 с. 2. Приставка А.Ф. Сплайн-распределения в статистическом анализе. – Днепропетровск, 1990. – 160 с. 3. Boneva L., Kendall D., Stefanov J. Spline transformations the new diagnostic for the statistical data analyses // J.Roy. Stistic. Soc. Ser. B. – 1971. – 33 (B33). – № 1. – P. 1–36. 4. Денисюк В.П., Денисюк О.В., Марченко Б.Г. Моделювання вимірювальних сигналів класами фундаментальних функцій // Актуальні проблеми автоматизації та інформаційних технологій: Зб. наукових праць. – Дніпропетровськ, 2000. – С. 31–39. 5. Колмогоров А.М. Теорія ймовірностей і математична статистика. – М., 1986. – 535 с. 6. Глівенко В.І. Курс теорії ймовірностей. – М., 1939. – 220 с.