УДК 519.21

В. Н. Михайлов, А. В. Харламов

ЛИНЕЙНАЯ РЕГРЕССИЯ В СЛОЖНЫХ ОБЛАСТЯХ НА ПЛОСКОСТИ

Рассмотрим в плоскости прямоугольной системы координат x, y, в общем случае, многосвязную область D, границы которой состоят из гладких участков $L_1, L_2, ..., L_M$, представимых уравнениями в параметрическом виде

$$x = x_k(\sigma), y = y_k(\sigma),\ d_{k-1} \le \sigma \le d_k,\ \ k = 1,2,...,M. \qquad (1)$$

В области D задана функция $z(x, y)$, которую приближенно можно представить в виде

$$z = \theta_0 q_0(x, y) + \theta_1 q_1(x, y) + ... + \theta_m q_m(x, y), \qquad (2)$$

коэффициенты $\theta_0, \theta_1, ..., \theta_m$ не известны.

В точках (x_i, y_i), $i = \overline{1, N}$ известны значения $\tilde{z}_i$ функции $z(x, y)$, определенные со случайной погрешностью ε_i, т.е. $\tilde{z}_i = z(x_i, y_i) + \varepsilon_i$. По этим данным необходимо оценить коэффициенты $\theta_0, \theta_1, ..., \theta_m$. Если зависимость (2) точная, то эти коэффициенты можно определить методом наименьших квадратов. В рассматриваемом случае приближенной зависимости оценки метода наименьших квадратов будут смещенными. Если функции $q_0, q_1, ..., q_m$ – ортонормированные, то с помощью несмещенного планирования [1] можно найти методом наименьших квадратов несмещенные оценки коэффициентов (2). Поэтому возникает задача построения систем ортонормированных в области D функций, которая для сложных областей не является тривиальной. Рассмотрим эту задачу.

Пусть C – множество действительных функций $g(x, y)$. Введем в C скалярное произведение

$$(f, g) = \iint_D p(x, y) f(x, y) g(x, y) dxdy, \qquad (3)$$

где $p(x, y) \in C$, и $p(x, y) > 0$ в области D.

Для определения скалярного произведения необходимо вычислять интеграл по области D, что для сложных областей можно осуществить, используя следующую формулу [2, 3]:

$$(f, g) = -\sum_{k=1}^{N} \int_{d_{k-1}}^{d_k} \Big(\int_{y_0}^{y_k(\sigma)} p(x_k(\sigma), y) f(x_k(\sigma), y) g(x_k(\sigma), y) dy \Big) x'_k(\sigma) d\sigma. \qquad (4)$$

Интеграл по области представлен, как сумма двукратных интегралов с известными пределами интегрирования, которые могут быть вычислены путем повторного применения известных формул численного интегрирования однократных интегралов.

Рассмотрим в C последовательность линейно независимых функций

$$f_0(x,y), f_1(x,y), \dots, f_n(x,y), \dots \quad (5)$$

Методом ортогонализации из (2) можно построить систему ортонормированных функций

$$q_n(x,y) = \sum_{k=0}^{n} a_{nk} f_k(x,y), \ n = 0,1,2,\dots, \quad (6)$$

где $a_{nn} = \frac{1}{\beta_{nn}}$, $a_{nk} = -\frac{1}{\beta_{nn}} \sum_{j=k}^{n-1} \beta_{ni} a_{jk}$, $k = \overline{0, n-1}$, и коэффициенты β_{nk} находятся из рекуррентных соотношений

$$\beta_{nk} = \frac{1}{\beta_{kk}}((f_n, f_k) - \sum_{i=0}^{n-1} \beta_{ki}\beta_{ni}), \ k = \overline{0, n-1}. \quad (7)$$

$$\beta_{nn} = \sqrt{(f_n, f_n) - \sum_{i=0}^{n-1} \beta_{ni}^2}, \ n = 0,1,2,\dots, \ \beta_{00} = \sqrt{(f_0, f_0)}.$$

Возьмем в качестве системы (5) систему степенных функций $1, x, y, x^2, xy, y^2, x^3, \dots$ или

$$f_n(x,y) = x^{\alpha_n} y^{\beta_n}, \ n = 0,1,2,\dots \quad (8)$$

$$p_n = \left[\frac{\sqrt{1+8n} - 1}{2}\right], \ \beta_n = n - \frac{p_n(p_n + 1)}{2}, \ \alpha_n = p_n - \beta_n.$$

Для этих функций скалярные произведения (4) имеют вид

$$(f_i, f_j) = -\sum_{k=1}^{M} \int_{d_{k-1}}^{d_k} \left(\int_{y_0}^{y_k(\sigma)} x_k^{\alpha_i + \alpha_j}(\sigma) y^{\beta_i + \beta_j} dy \right) x'_k(\sigma) d\sigma.$$

Внутренний интеграл по у вычисляется, в результате получаем

$$(f_i, f_j) = -\sum_{k=1}^{N} \frac{1}{\beta_i + \beta_j + 1} \int_{d_{k-1}}^{d_k} (x_k^{\alpha_i + \alpha_j}(\sigma) y_k^{\beta_i + \beta_j + 1}(\sigma)) x'_k(\sigma) d\sigma.$$

Следовательно, для степенных функций скалярные произведения являются одномерным интегралом.

Приведем пример расчета. Область D ограничена эллипсом с центром в начале координат и полуосям $a = 1.2$, $b = 0.8$ и кругом радиуса $r = 0.5$. Система линейно независимых функций имеет вид (8). Приведем несколько элементов построенной ортонормированной системы функций $q_0, q_1, \dots$

$q_0 = 0.66867$,

$q_1 = 0.00052 + 0.97998x$,

$q_2 = -0.00002 + 1.51888y$,

$q_3 = -0.85691 + 0.00374x - 0.00004y + 1.84055x^2$,

$q_4 = -0.00004 - 0.00006x + 0.00010x^2 + 2.96478xy$,

$q_5 = -1.70626 + 0.00196x + 0.0004y + 1.52973x^2 + 0.00009xy + 5.12905y^2$,

$q_6 = 0.00073 - 2.1353x + 0.00006y + 0.0061x^2 - 0.0001xy - 0.0068y^2 + 2.8506x^3$.

Пусть система функций в (2) – ортонормированная, тогда коэффициенты θ_j можно выразить через функции $z(x,y), q_j(x,y)$ по формуле

$$\theta_j = (z, q_j) = \iint_D z(x,y) q_j(x,y) dxdy. \tag{9}$$

Предположим, что точки (x_i, y_i), $i = \overline{1,N}$, проведения эксперимента расположены в узлах квадратурной формулы, которая имеет вид

$$\iint_D f(x,y)dxdy = \sum_{i=1}^{N} c_i f(x_i, y_i),\ c_i > 0, \sum_{i=1}^{N} c_i = \iint_D dxdy = S. \tag{10}$$

Пусть эта формула будет точной для степенных функций (8) до $n = n_1$. Возьмем в качестве оценки $\hat{\theta}_j$ коэффициента θ_j формулу, следующую из (9) и (10):

$$\hat{\theta}_j = \sum_{i=1}^{N} c_i q_j(x_i, y_i) \tilde{z}_i. \tag{11}$$

Будем считать, что ортонормированная система функций $q_0, q_1, \ldots$ получена из системы степенных функций (8). Свойства оценки (11) определяются следующей теоремой.

ТЕОРЕМА. Если ε_j – независимые случайные величины, $M\varepsilon_i = 0$, $D\varepsilon_i = \sigma^2$ и $\alpha_j + \beta_j < \alpha_m + \beta_m < \alpha_{n_1} + \beta_{n_1}$, то $\hat{\theta}_j$ является несмещенной и состоятельной оценкой коэффициента θ_j и ее дисперсия $D\hat{\theta}_j$ имеет вид

$$D\hat{\theta}_j = \sigma^2 \sum_{i=1}^{N} c_i^2 q_j^2(x_i, y_i).$$

Доказательство. Так как $M\varepsilon_i = 0$, то $M\tilde{z}_i = z_i$, где z_i – истинное значение функции $z(x,y)$ в точке (x_i, y_i), поэтому, используя (11), (10) и (9), получаем

$$M\hat{\theta}_j = \sum_{i=1}^{N} c_i q_j(x_i, y_i) M\tilde{z}_i = \sum_{i=1}^{N} c_i q_j(x_i, y_i) z(x_i, y_i) = \iint_D z(x,y) q_j(x,y) dxdy = \theta_j$$

т.е. оценка $\hat{\theta}_j$ является несмещенной.

Так как $\tilde{y}_i$, $i = \overline{1,N}$ – независимые случайные величины, то из (11) следует формула для дисперсии оценки $\hat{\theta}_j$:

$$D\hat{\theta}_j = \sum_{i=1}^{N} c_i^2 q_j^2(x_i, y_i) D\tilde{z}_i = \sigma^2 \sum_{i=1}^{N} c_i^2 q_j^2(x_i, y_i).$$

Пусть $c = \max\limits_{1 \le i \le N} c_i$, тогда, учитывая (10) и ортонормированность функций $q_0, q_1, \ldots$, будем иметь

$$\sum_{i=1}^{N} c_i^2 q_j^2(x_i, y_i) \le c \sum_{i=1}^{N} c_i q_j^2(x_i, y_i) = c \iint\limits_D q_j^2(x, y) dx dy = c,$$

поэтому $D\hat{\theta}_j \le \sigma^2 c$. Но $c \to 0$ при $N \to \infty$, следовательно, оценка $\hat{\theta}_j$ является состоятельной.

СПИСОК ЛИТЕРАТУРЫ

1. *Ермаков С. М., Жиглявский А. А.* Математическая теория оптимального эксперимента. М.: Наука,1987.

2. *Федик И. И., Михайлов В. Н., Кожуховский В. И.* Формализация вариационного метода решения некоторых краевых задач в сложных двумерных областях // Докл. АН УССР. Сер. А. 1982. № 10. С. 51 – 54.

3. *Михайлов В. Н.* Автоматизация вычисления интегралов по плоским областям на ЭВМ // Вычислительные методы и программирование: Межвуз. сб. науч. тр. Вып. 9. Саратов: Изд-во Сарат. ун-та, 1999. С. 109 – 112.

УДК 519.2, 681.3

В. В. Мозжилкин

АНАЛИЗ МАТЕМАТИЧЕСКОЙ МОДЕЛИ СИМОНОВИЧА РЫНКА ТРУДА

Целью данной статьи является математический анализ модели рынка труда, предложенной Симоновичем [1]. Эта модель представляет собой нелинейную систему двух разностных уравнений, определяющих зависимость от времени функций $L(t)$ и $I(t)$, характеризующих соответственно действительное число сделок на рынке труда и разницу между предложением товаров и числом реально совершенных сделок на рынке товаров.

Введем следующие функции:

$L^D(t)$ – функция спроса на единицу труда за время t;

$L^S(t)$ – функция предложения труда за время t;

$L(t)$ –функция, характеризующая действительное число сделок на рынке труда;

$Y^D(t)$ – функция спроса на рынке товаров за время t;

$Y^S(t)$ – функция предложения на рынке товаров за время t;

$Y(t)$ – функция, характеризующая действительное число сделок на рынке товаров;

$I(t)$ – разница между предложением товаров и числом реально совершенных сделок на рынке товаров за время t: