Доклады Академии Наук СССР
1947. Том LVIII, № 5

МАТЕМАТИКА

О. В. САРМАНОВ

О ВЫПРЯМЛЕНИИ СИММЕТРИЧНОЙ КОРРЕЛЯЦИИ

(Представлено академиком С. Н. Бернштейном 12 V 1947)

§ 1. Пусть $F(x,y)=F(y,x)$ — симметрическая функция распределения, определяющая корреляцию во всей плоскости.

Если корреляция криволинейна, т. е. м.о.$_x y=\psi(x)$ нелинейная функция, требуется найти такие преобразования переменных $X=\varphi(x)$, $Y=\varphi(y)$, что корреляция между X и Y прямолинейна.

Если такие преобразования существуют, корреляция называется в ы п р я м л я е м о й.

Задача о выпрямлении корреляции эквивалентна нахождению монотонного решения корреляционного интегрального уравнения

$$\varphi(x)=\lambda\int_{-\infty}^{\infty}\varphi(y)\,\frac{F(x,y)}{p(x)}\,dy\,, \tag{1}$$

где $p(x)=\int\limits_{-\infty}^{\infty}F(x,y)\,dy$.

Если $\varphi(x)$ монотонное решение уравнения, замена переменных $X=\varphi(x)$, $Y=\varphi(y)$ возможна и корреляция между X и Y прямолинейна.

Ниже приводится достаточное условие существования монотонного решения уравнения (1) и показывается его необходимость для широкого класса поверхностей корреляции.

§ 2. Ядро уравнения (1) симметризуется. На симметрическое ядро $D(x,\ y)=\frac{F(x,y)}{\sqrt{p(x)\,p(y)}}$ наложим ограничение: $\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}D^2(x,y)\,dxdy$ существует.

Тогда можно написать разложение по фундаментальным функциям для k-й итерации ядра $D^k(x,y)=\frac{F^{(k)}(x,y)}{\sqrt{p(x)\,p(y)}}$, справедливое для $k\geqslant 2$, где $F^{(k)}(x,y)=\int\limits_{-\infty}^{\infty}\frac{F^{(k-1)}(x,t)\,F(y,t)}{p(t)}$ — k-я „итерированная поверхность корреляции“.

Предположим, что первое характеристическое число λ_1 является

m-кратным, $m \geqslant 1$. В этом случае указанное разложение будет иметь вид:

$$\frac{F^{(k)}(x,y)}{p(x)} = p(y) + \frac{1}{\lambda_1^k}\left[\varphi_1(x)\varphi_1(y)p(y) + \cdots + \varphi_m(x)\varphi_m(y)p(y)\right] + \\ + \frac{\varphi_{m+1}(x)\varphi_{m+1}(y)p(y)}{\lambda_{m+1}^k} + \cdots \tag{2}$$

Используя неравенства Бесселя и Шварца, легко доказать, что существует предел конечной разности

$$\lim_{k\to\infty} \lambda_1^k \left[\frac{F^{(k)}(x+h,y)}{p(x+h)} - \frac{F^k(x,y)}{p(x)}\right] = F(x,y,h),$$

причем

$$\Phi(x,y,h) = [\varphi_1(x+h) - \varphi_1(x)]\varphi_1(y)p(y) + \cdots \\ \cdots + [\varphi_m(x+h) - \varphi_m(x)]\varphi_m(y)p(y). \tag{3}$$

Тогда справедлива следующая теорема:

Теорема 1. *Если существует хоть одно значение $y = y_0$, при котором функция*

$$\int_{-\infty}^{y} \Phi(x,t,h)\,dt = \Psi(x,y,h) \tag{4}$$

сохраняет знак при любом x и любом $h > 0$, то уравнение (1) *имеет в своем спектре единственное монотонное решение.*

В самом деле, пусть условие теоремы выполнено, тогда

$$\Psi(x,y_0,h) = [\varphi_1(x+h) - \varphi_1(x)]\int_{-\infty}^{y_0} \varphi_1(y)p(y)\,dy + \cdots \\ \cdots + [\varphi_m(x+h) - \varphi_m(x)]\int_{-\infty}^{y_0} \varphi_m(y)p(y)\,dy. \tag{5}$$

и правая часть равенства (5) сохраняет знак при всех x и $h > 0$, т. е. характеристическая функция $c_1\varphi_1(x) + c_2\varphi_2(x) + \cdots + c_m\varphi_m(x)$, где $c_i = \int_{-\infty}^{y_0} \varphi_i(y)p(y)\,dy$ $(i = 1, 2, \ldots, m)$, принадлежащая числу λ_1, монотонна при всех x. Единственность вытекает из того обстоятельства, что в спектре не может быть двух монотонных ортогональных функций (см. ([1])).

§ 3. Если λ_1 — простое характеристическое число, т. е. $m = 1$, то из знакопостоянства функции (4) при одном y вытекает ее знакопостоянство при всех y, причем это последнее условие является необходимым.

Теорема 2. *Если λ_1 — простое характеристическое число, то для монотонности первой фундаментальной функции $\varphi_1(x)$ необходимо и достаточно, чтобы функция* (4) *была знакопостоянна при всех x, y и $h > 0$.*

Достаточность вытекает из теоремы 1, необходимость сразу получается из равенства (5), которое теперь имеет вид

$$\Psi(x,y,h) = [\varphi_1(x+h) - \varphi_1(x)]\int_{-\infty}^{y} \varphi_1(t)p(t)\,dt. \tag{5'}$$

Если $\varphi_1(t)$ монотонна, $\int\limits_{-\infty}^{y} \varphi_1(t)\, p(t)\, dt$ знакопостоянна при всех y, а $\varphi_1(x+h) - \varphi_1(x)$ знакопостоянна при всех x и всех $h > 0$, поэтому $\Psi(x, y, h)$ знакопостоянна при всех x и y и всех $h > 0$.

§ 4. В моей предыдущей заметке ([1]) в предположении дифференцируемости $\frac{F(x, y)}{p(x)}$ по x указано, что для монотонности одного из решений уравнения (1) достаточно знакопостоянство функции $\Psi_1(x, y) = \int\limits_{-\infty}^{y} \frac{\partial}{\partial x} \frac{F(x, t)}{p(x)}\, dt$ при всех x и y.

Легко проверить, пользуясь определением функций $F^{(k)}(x, y)$, что из этого условия вытекает знакопостоянство конечных разностей $\int\limits_{-\infty}^{y} \left[\frac{F^{(k)}(x+h, t)}{p(x+h)} - \frac{F^{k}(x, t)}{p(x)} \right] dt$ при всех $k > 1$ и, следовательно, знакопостоянство $\Psi(x, y, h)$ при всех x, y и $h > 0$.

Вообще, если написать последовательность функций

$$\Psi_k(x, y) = \int\limits_{-\infty}^{y} \frac{\partial}{\partial x} \frac{F^{k}(x, t)}{p(x)}\, dt, \quad k = 1, 2, \ldots, \tag{6}$$

то из знакопостоянства при всех x, y одной из них вытекает знакопостоянство функции (4) при всех x, y, $h > 0$. Мы получаем таким образом бесконечную последовательность все более и более общих условий, достаточных для существования монотонного решения уравнения (1).

Заметим, что первое условие при $k = 1$ легко проверяется по эмпирической корреляционной таблице и допускает простую геометрическую интерпретацию: графики интегральных условных функций распределения y при различных x не должны пересекаться.

Ленинградское отделение
Математического института
им. В. А. Стеклова
Академии Наук СССР

Поступило
12 V 1947

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] О. В. Сарманов, ДАН, 53, 9 (1946).