УДК. 62-50

ПАВЛОВ А. А.,
ЧЕХОВСКИЙ А.В.

СВЕДЕНИЕ ЗАДАЧИ ПОСТРОЕНИЯ МНОГОМЕРНОЙ РЕГРЕССИИ К ПОСЛЕДОВАТЕЛЬНОСТИ ОДНОМЕРНЫХ ЗАДАЧ

Обосновывается возможность сведения задачи построения многомерной полиномиальной регрессии к последовательности одномерных регрессионных задач и решению соответствующих систем алгебраических линейных уравнений с постоянными коэффициентами.

The possibility of bringing of the multidimensional polynomial regression building problem to the sequence of one-dimensional regression problems is grounded in the article. Also the possibility of solving of corresponding systems of algebraic linear equations with constant coefficients is also described in the article.

Проблема эффективного восстановления многомерной полиномиальной регрессии по настоящее время является актуальной [1–9]. Использование нормированных ортогональных полиномов Форсайта [9,10] позволяет предложить эффективный метод построения многомерной полиномиальной регрессии. Он базируется на следующем анализе построения одномерной полиномиальной регрессии с помощью нормированных ортогональных полиномов Форсайта [9]: рассмотрим одномерную модель

$$Y(x) = \theta_0 + \theta_1 x + \ldots + \theta_r x^r + \mathrm{E}, \qquad (1)$$

где x – детерминированная переменная, значение которой в эксперименте исследователь может задавать произвольно. $\theta_i,\ i=\overline{0,r}$ – неизвестные коэффициенты, r может быть избыточно большим. Е – случайная величина с произвольным распределением. δ_E^2 неизвестна, либо существует ее верхняя оценка δ^2. По результатам экспериментов ($x_i, n y_i,\ i=\overline{1,\ }$) нужно найти истинное значение r и оценить значение коэффициентов $\theta_i,\ \ i=\overline{0,r}$.

Были построены таблицы для оценки дисперсий $\theta_j,\ \ j=\overline{0,r}$ для случая, когда значения детерминированного аргумента $x_i,\ i=\overline{1,n}$ равномерно распределены внутри отрезка с концами (–*a*, *a*), *a* >0. Для *a* >1 анализ таблиц позволил сделать следующие выводы:

Для известной δ^2, можно определить необходимое число испытаний *n*, чтобы с заданной вероятностью оценка коэффициентов определялась с заданной погрешностью. Например для $a=50,\ j\geq 2,\ \ n\geq 10,\qquad D\hat{\theta}_j \leq \delta^2\cdot 7{,}55\cdot 10^{-6}$. При этом чем больше *j* тем меньшее число экспериментов гарантирует практически точную оценку неизвестных коэффициентов θ_j. Из этого следует возможность достоверного нахождения истинной степени полинома одномерной регрессии (1) по минимальному числу экспериментов.

Построение многомерной полиномиальной регрессии

Модель имеет вид:

$$\overline{y}(x) = \sum_{\forall(i_1,\ldots i_l)\in k}\ \sum_{\forall(j_1,\ldots j_l)\in k(i_1\ldots i_l)}^{n} b_{i_1\ldots i_l}^{j_1\ldots j_l}\left(x_{i_1}\right)^{j_1}\times \times\left(x_{i_2}\right)^{j_2}\ldots\times\left(x_{i_l}\right)^{j_l}+E, \qquad (2)$$

где $\overline{x}=(x_1\ldots x_n)^T$ – детерминированный вектор входных переменных, $b_{i_1\ldots i_l}^{j_1\ldots j_l}$ – неизвестные коэффициенты l – натуральные числа, i_l – натуральные индексы из множества $\{1,\ldots,n\}$. Модель (2) является избыточной. Если возможен активный эксперимент, то задача восстановления многомерной регрессии сводится к последовательности построения одномерных регрессий и решению систем линейных равенств. Порядок вычислений заключается в следующем: последовательно для каждой переменной $x_j\,(j=\overline{1,n})$ фиксируется необходимое количество T_j различных значений переменных x_l, $\forall l(l\neq j)$. Тогда (2) превращается в T_j одномерных регрессий от переменной x_j, найденные коэффициенты которых составляют вектор – столбец правой части *j*-той $\left(j=\overline{1,n}\right)$ системы линейных уравнений. Решение *j*-той системы линейных уравнений определяет все коэффициенты $b_{i_1\ldots i_l}^{j_1\ldots j_l}$ членов (2), которые содержат x_j и не были определены на предыдущих l этапах $l=\overline{1,\ j_{-1}}$.

Выводы

В статье предложен и обоснован эффективный метод построения многомерной полиномиальной регрессии. Метод предусматривает реализацию активного эксперимента и базируется на свойствах нормированных ортогональных полиномов Форсайта.

Список литературы

1. Адлер Ю.П., Маркова Е.В., Грановский Ю.В. Планирование эксперимента при поиске оптимальных условий. – 2-е изд., перераб. и доп. – М.: Наука, 1976. –280 С.
2. Айвазян С.А. Многомерный статистический анализ // Математическая энциклопедия / Гл. ред. И.М. Виноградов. – М., 1982. – Т.3. – Стб. 732-738.
3. Андерсон Т. Введение в многомерный статистический анализ / Пер. с англ.
4. Ю.Ф. Кичатова; Под ред. Б.В. Гнеденко. – М.: Физматтиз, 1963. – 500 С.
5. Ершов А. А. Стабильные методы оценки параметров: (Обзор) // Автоматика и телемеханика. – 1978. – № 8. – С. 66-100.
6. Колмогоров А.Н. К обоснованию метода наименьших квадратов // Успехи математических наук. – 1946. – Т. 1, Вып. 1. – С.57-70.
7. Форсайт Дж., Малькольм М., Моулер К. Машинные методы математических вычислений / Пер. с англ. Х.Д. Икрамова. – М.: Мир, 1980. – 280 С.
8. Радченко С.Г. Устойчивые методы оценивания статистических моделей: Монография – К.: ПП «Санспарель», 2005. – 504 С.
9. Норманн Р. Дрейпер, Гарри Смит. Прикладной регрессионный анализ, 3-е изд.: Пер. с англ. – М.: Издательский дом «Вильямс», 2007. – 912 С
10. Д. Худсон. Статистика для физиков. Москва, Мир, 1970.
11. Forsythe G, journ. Sos.Ind. Appl. Math 5,74 (1957).