УДК 004.416.6

ПРОГРАММНЫЙ АЛГОРИТМ ПОСТРОЕНИЯ ОЦЕНКИ РЕГРЕССИИ СТОХАСТИЧЕСКОГО ОБЪЕКТА С МНОГОМЕРНЫМ ВХОДОМ

Краснобровкин П.С.

Научный руководитель - д.т.н, проф., Рубан А. И.

Сибирский федеральный университет

Предложен алгоритм построения непараметрической оценки регрессии объекта с одним выходом и несколькими входами. Приведён тестовый пример, показывающий работоспособность алгоритма.

Постановка задачи

Рассматриваем стохастический объект, имеющий несколько входов X и один выход Y. Y и X являются случайными величинами. Идеализированными моделями стохастических объектов являются статистические характеристики, особое место среди которых занимает регрессия

$$M\{Y \mid x\} \equiv \eta(x) = \int_{-\infty}^{\infty} y\, f(y \mid x)\, dy,$$

где $M\{Y \mid x\}$- условное математическое ожидание, показывающее связь между выходом Y и фиксированным значением входов X = x.

После проведения измерений объекта была получена выборка, которая в общем случае не является упорядоченной ни по одной из входных координат X. Также считаем, что в измерительном устройстве присутствует аддитивная помеха, распределённая по нормальному закону с нулевым математическим ожиданием и единичной дисперсией. Для получения универсальной оценки, а также для сокращения объёмов вычислений, оценка регрессии будет рассматриваться в безразмерных координатах. По полученной выборке строится непараметрическая оценка регрессии

$$\eta(x) = \sum_{i=1}^{n} y_i \times \frac{\prod_{j=1}^{m} K(\beta_j \, \frac{\overrightarrow{x_j} - x_{ji}}{\Delta_j})}{\sum_{l=1}^{n} \prod_{j=1}^{m} K(\beta_j \, \frac{\overrightarrow{x_j} - x_{jl}}{\Delta_j})},$$

$$K(z) = \begin{matrix} 0.75(1 - z^2), & |z| \le 1 \\ 0, & 1 < |z| \end{matrix}.$$

Оптимальная величина для каждого коэффициента β_j подбирается минимизацией показателя качества (путём разделения выборки на две части – обучающую и экзаменующую, по обучающей строится непараметрическая оценка, по экзаменующей – показатель качества, либо методом скользящего экзамена).

Неупорядоченность полученной выборки создаёт проблему при поиске достоверного значения Δ_j. Идея предложенного алгоритма заключается в параллельном нахождении Δ_j для каждой входной оси. Каждому входу X_j сопоставляется поток выполнения, в котором копия исходной выборки сначала упорядочивается по j-ой оси, а затем в упорядоченной выборке находится максимальная разность $\Delta_j = \max\{(x_{ji+1} - x_{ji}), i = \overline{1, n-1}\}$.

Пример

В качестве примера рассмотрим эллиптический параболоид с радиусом равным 10, находящийся в начале координат. Сигнальная часть объекта имеет вид

$$y(x1, x2) = \frac{x_1^2 + x_2^2}{100}$$

Считаем, что на сигнальную часть объекта накладывается аддитивная, нормально распределённая помеха с нулевым математическим ожиданием и дисперсией равной единице. В результате работы алгоритма были получены значения безразмерных коэффициентов: $\beta = (0.016630; 0.018604)$. На рисунках ниже представлены графики:

- сигнальной части объекта;
- измерений объекта (сигнальная часть + аддитивная помеха);
- непараметрической оценки регрессии

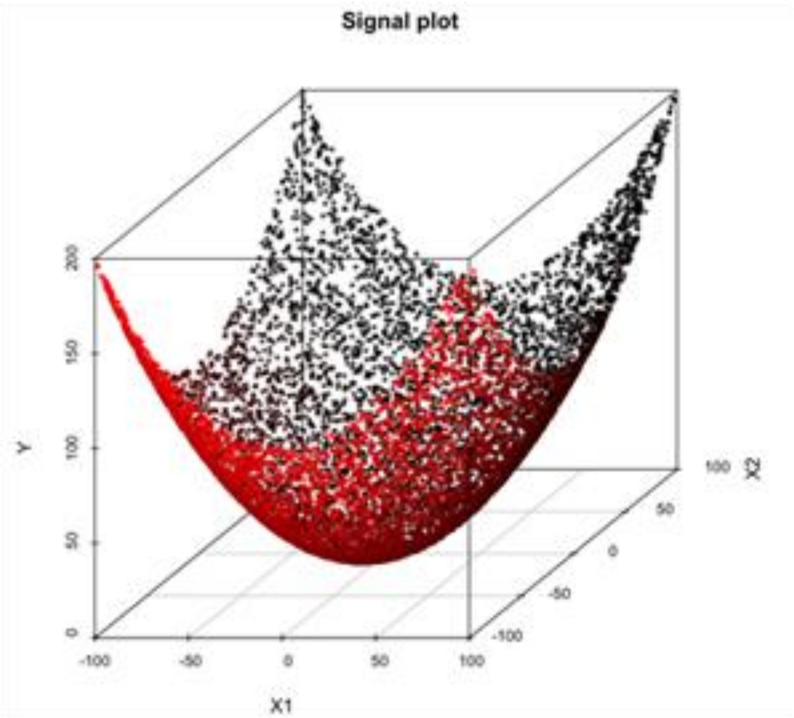

Рисунок 1 – График сигнальной части объекта в точках выборки

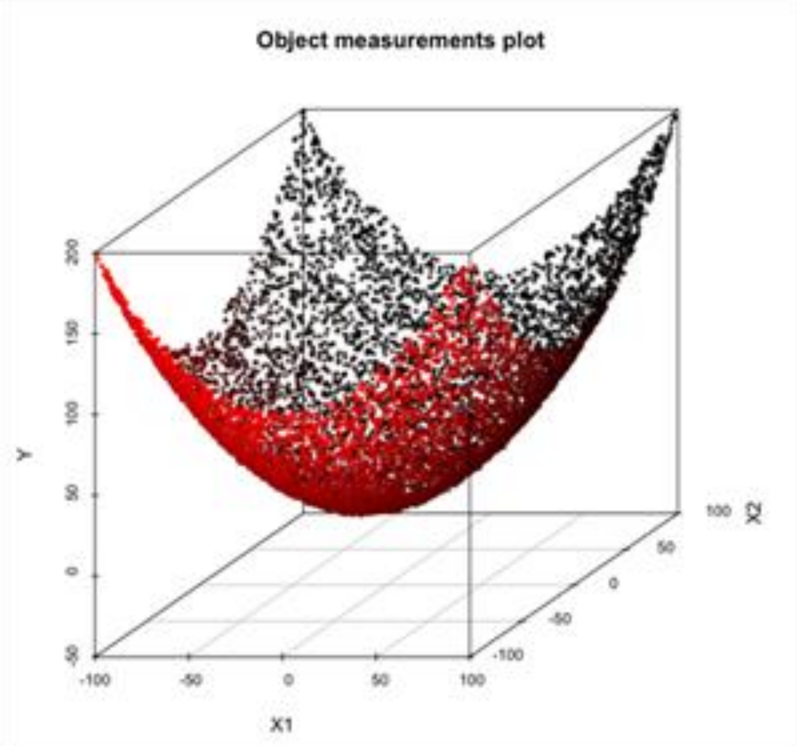

Рисунок 2 – Полученные точки выборки

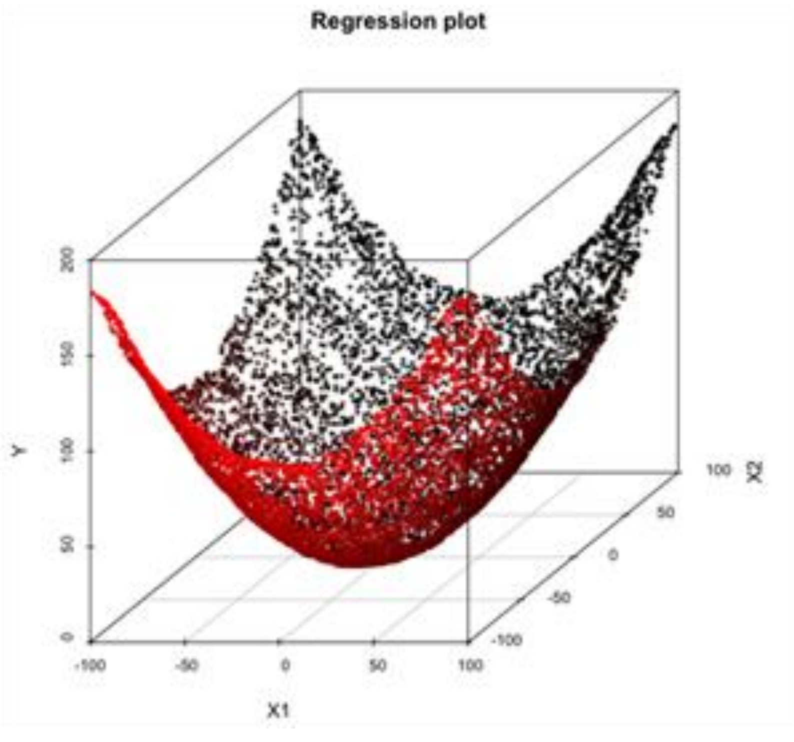

Рисунок 3 – График непараметрической оценки регрессии