Краткие теоретические сведения

Регрессией называют первый начальный условный момент:

$$M\{Y|x\} = \int_{-\infty}^{\infty} yf(y|x)dy = \eta(x) \ (1)$$

Оценка $\eta_n(x)$ регрессии строится на основе серии измерений выхода и входа объекта: x_i, y_i, где i изменяется от *1* до n:

$$\widehat{M}\{Y|x\} \equiv \eta_n(x) = \sum_{i=1}^{n} K_N(\frac{x-x_i}{h}) y_i \ (2)$$

В формуле 2 нормированное ядро определяется как:

$$K_N\left(\frac{x-x_i}{h}\right) = \frac{K(\frac{x-x_i}{h})}{\sum_{j=1}^{n} K(\frac{x-x_j}{h})} \ (3)$$

Из 3 формулы следует, что ядро $K_N(\frac{x-x_i}{h})$ нормировано на 1 на системе экспериментальных точек. Нормированность приводит к условию:

$$\min\{y_i, i = \overline{1,n}\} \le \eta_n(x) \le \max\{y_i, i = \overline{1,n}\} \ (4)$$

Усечённость нормированных ядер (в силу усечённости ядра K(·)) позволяет при построении оценки $\eta_n(x)$ в каждой фиксированной точке x учитывать только несколько близлежащих значений x_i.

Основное влияние на оценку регрессии оказывает положительная константа c, но зависимость от c при возрастании n ослабевает. Форма ядра, усечённая параболическая. Константа c, определяющая коэффициент размытости, вычисляется по выборке путём минимизации эмпирических показателей (характеризующих наилучшее сглаживание экспериментальных данных).

Считаем, что выборка $(x_i,\ y_i)$, $i = \overline{1,n}$ измерениях входа находятся на равных расстояниях друг от друга $\Delta = x_{i+1} - x_i$ ($i = \overline{1, n-1}$), а объем выборки n фиксирован. Перейдем от размерного параметра c (его размерность, обратная размерности x) к безразмерному β ($0 \le \beta \le 1$):

$$\beta = c^{-1} \Delta n^{1/5} \ (5)$$

Формулы для расчётов

При расчётах использовались безразмерные переменные, поэтому формулы для расчёта приобрели вид:

$$\eta_n(x) = \sum_{i=1}^{n} K_N(\beta \frac{x-x_i}{\Delta}) y_i \quad (6)$$

$$K_N\left(\beta \frac{x-x_i}{\Delta}\right) = \frac{K(\beta \frac{x-x_i}{\Delta})}{\sum_{j=1}^{n} K(\beta \frac{x-x_j}{\Delta})} \quad (7)$$

При β=0 оценка регрессии $\eta_n(x)$ не зависит от x . Такой вариант, хотя и редко, но возможен. Выбранный вход объекта не оказывает влияния на выход объекта. При $\beta = 1$ оценка регрессии $\eta_n(x)$ точно проходит через экспериментальные точки, т. е. оценка не осуществляет сглаживания экспериментальных данных. Такой вариант тоже возможен, если сигнальная часть выхода объекта не зашумлена помехой.

В процессе выполнения программы осуществляется подбор оптимального коэффициента β по выбранному критерию оптимальности.

В данной самостоятельной работе был применён метод «скользящего экзамена», где по β минимизируется результирующее значение формулы 8:

$$I_{1n} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \overline{\eta_n}(x_i))^2 \quad (8)$$

$$\overline{\eta_n}(x_i) = \sum_{k=1;k\neq i}^{n} K_N(\beta \frac{x-x_i}{\Delta}) y_k \quad (9)$$

Выборка x_i, y_i ($i = \overline{1,n}$) при этом своеобразно разбивается на две части: одна используется для построения модернизированной модели $\overline{\eta_n}(x)$, вторая – для ее проверки (по вышеуказанному критерию). Например, первое слагаемое в I_{1n} равно квадрату невязки между выходом объекта y_1 и выходом модели $\overline{\eta_n}(x_1)$ в первой экзаменующей точке $(x_1,\ y_1)$. Эта экзаменующая точка не участвует в построении (в обучении) модели $\overline{\eta_n}(x_1)$.

Формирование имитации объекта состоит из трёх стадий:

- выбор вида имитационной модели;
- формирование сигнальной части объекта (все ближайшие точки имеют константу в качестве расстояния);

- добавка к результату имитационной модели аддитивной помехи.

В качестве способа подбора коэффициента β используется метод деления отрезка пополам.