

САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ИМЕНИ
ПЕТРА ВЕЛИКОГО

ФИЗИКО-МЕХАНИЧЕСКИЙ ИНСТИТУТ
КАФЕДРА ПРИКЛАДНОЙ МАТЕМАТИКИ

Интервальный анализ
Отчёт по лабораторной работе №5

Выполнил:

Студент: Бочкарев И. А.

Группа: 3630102/80201

Принял:

к. ф.-м. н., доцент

Баженов Александр Николаевич

2022 г.

Содержание

1	Постановка задачи	2
2	Теория	2
2.1	Точечная оценка параметров регрессии	3
3	Реализация	3
4	Результаты	3
4.1	Генерация данных	3
4.2	Результаты	4
5	Обсуждения	9

1 Постановка задачи

Задать набор входных точечных переменных x и выходных интервальных переменных y с зависимостью близкой к линейной.

Провести вычисления и привести иллюстрации:

- Построить интервальное множество решений β , сделать точечные оценки параметров.
- Построить коридор совместных зависимостей
- Задать набор точек предсказания внутри и вне x , построить набор значений выходной переменной y

2 Теория

Пусть величина y является функцией от независимых аргументов x_1, \dots, x_m .

$$y = f(x, \beta) \tag{1}$$

где $x = (x_1, \dots, x_m)$ - вектор независимых переменных, $\beta = (\beta_1, \dots, \beta_l)$ - вектор параметров функции.

Имея набор значений x и y , нужно найти вектор β , который соответствует конкретной функции из семейства 1. Эта задача называется задачей восстановления зависимости.

Будем рассматривать задачу, в которой x - точечный вектор, а y - интервальный вектор. Информационным множеством задачи восстановления зависимости называется множество значений параметров зависимости, совместных с данными. Для случая линейной зависимости информационное множество - это выпуклое множество, ограниченное гиперплоскостями в пространстве R^n

Коридором совместных зависимостей называется многозначное отображение Υ , которое сопоставляет каждому значению аргумента x множество

$$\Upsilon(x) = \bigcup_{\beta \in \Omega} f(x, \beta) \tag{2}$$

где $y = f(x, \beta)$ - зависимость из задачи восстановления зависимости, Ω - непустое информационное множество параметров.

2.1 Точечная оценка параметров регрессии

Пусть модель задаётся в классе линейных функций $y = \beta_0 + \beta_1 x$. Для нахождения точечной оценки параметров регрессии поставим задачу линейной оптимизации и решим её:

$$\sum_{i=1}^m w_i \rightarrow \min \quad (3)$$

$$\text{mid} \mathbf{y}_i - w_i \cdot \text{rad} \mathbf{y}_i \leq X\beta \leq \text{mid} \mathbf{y}_i + w_i \cdot \text{rad} \mathbf{y}_i \quad (4)$$

$$w_i \geq 0, i = 1, \dots, m \quad (5)$$

$$w, \beta - ? \quad (6)$$

где m - число входных значений, X - матрица линейной регрессии, w - вектор весов.

Другие варианты точечной оценки параметров регрессии:
Середина наибольшей диагонали информационного множества:

$$\beta = \frac{(b_1 + b_2)}{2} \quad (7)$$

где b_1, b_2 - вершины информационного множества, находящиеся на максимальном расстоянии друг от друга.

Центр тяжести информационного множества:

$$\beta = \text{mean} V \quad (8)$$

где V - множество вершин информационного множества.

3 Реализация

Язык программирования: Python. Среда разработки: PyCharm. Ссылка на реализацию: <https://github.com/BoIlAl/Interval-Analysis/labs/reports/lab5>

4 Результаты

4.1 Генерация данных

Рассматривается модель $y = kx + b$. Для заданного набора значений входных значений x считается точечный вектор выходных значений y .

Затем строится интервальный вектор выходных значений следующим образом: точечное значение y_j заменяется интервалом $[y_j - |\delta_1|, y_j + |\delta_2|]$ ($j = 1, \dots, m$), где δ_1, δ_2 - случайные величины, $\delta_1, \delta_2 \in N(0, 5)$.

4.2 Результаты

Параметры модели: $k = 1, b = 0$, набор входных значений $x = (1.0, 2.0, \dots, 25.0)$.
График $y = kx + b$ модели и исходная выбока:

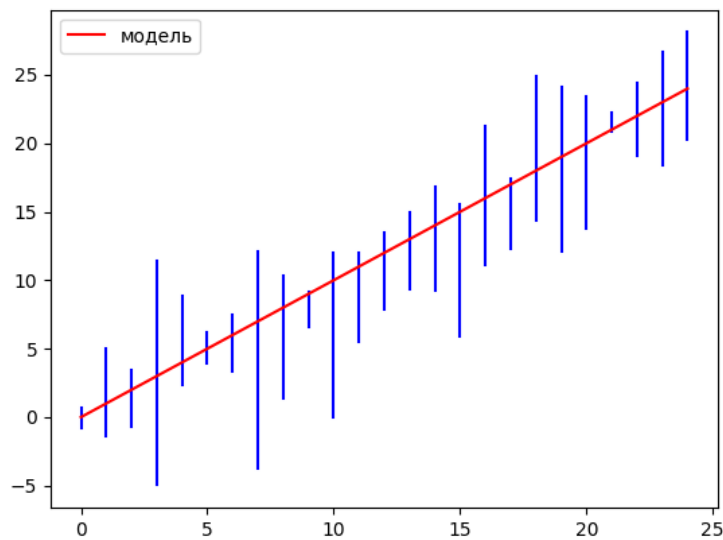


Рис. 1: Исходная выборка

Будем решать задачу восстановления зависимости для класса функций $y = \beta_0 + \beta_1 x$. Построим информационное множество и точечные оценки параметров зависимости.

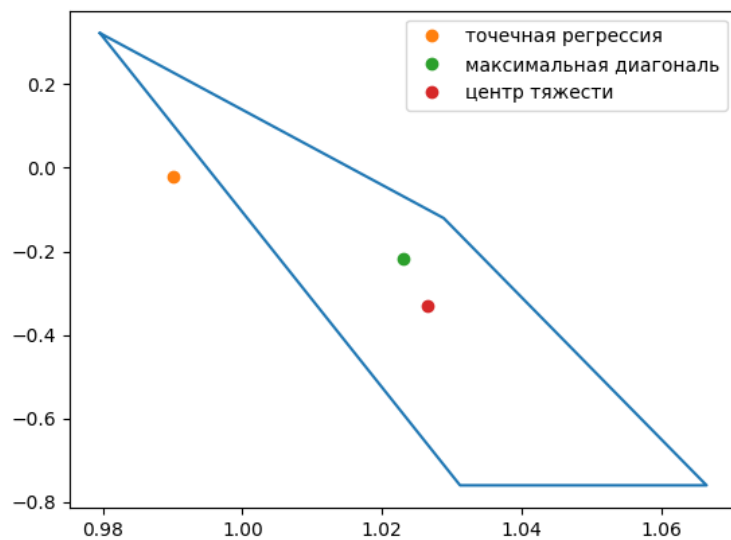


Рис. 2: Интервальная оценка параметров зависимости

Графики функций с полученными точечными оценками параметров β_0, β_1 будут иметь следующий вид:

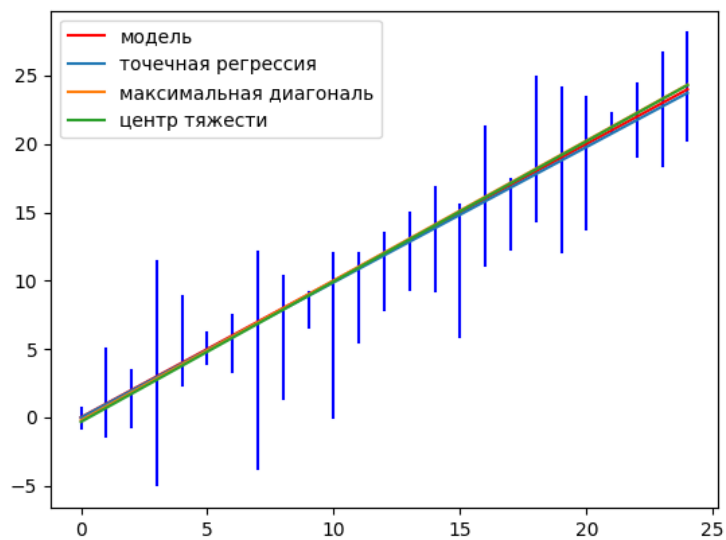


Рис. 3: Графики функций с точечными оценками параметров зависимости

Теперь построим коридор совместных зависимостей.

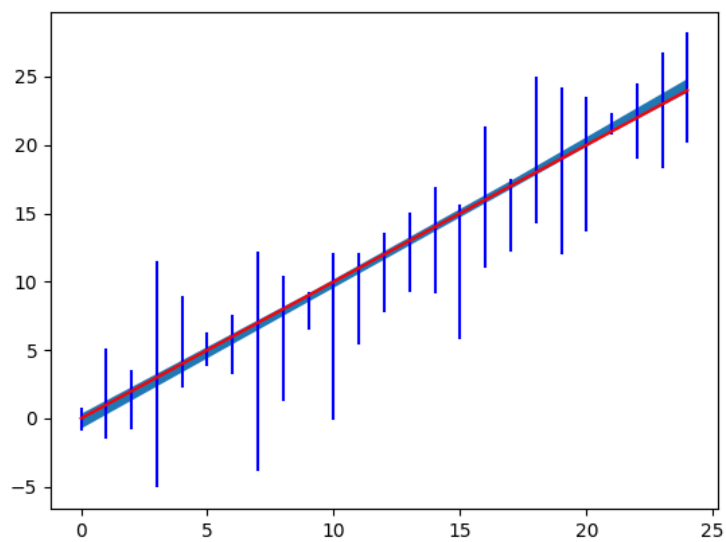


Рис. 4: Коридор совместных зависимостей

Зададим набор точек предсказания внутри и вне x и построим набор выходных значений y .

Набор внутри x : $x_1 = (5.5, \dots, 14.5)$.

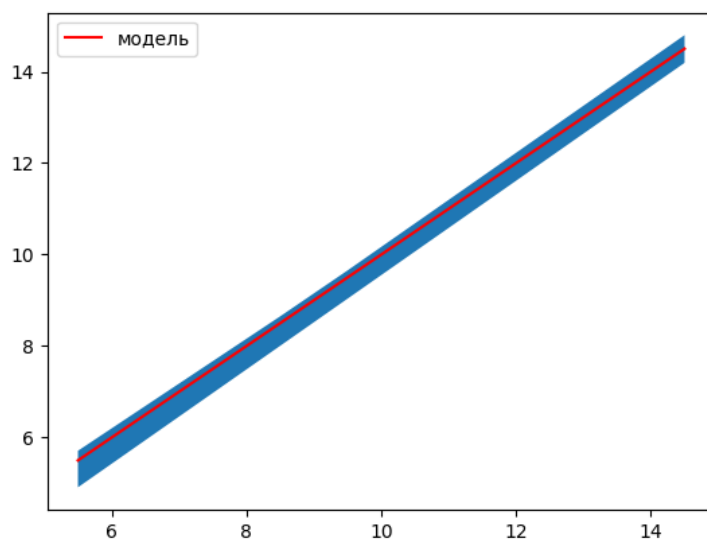


Рис. 5: Значения выходной переменной y для точек предсказания внутри x

Набор вне x : $x_2 = (50, \dots, 59)$.

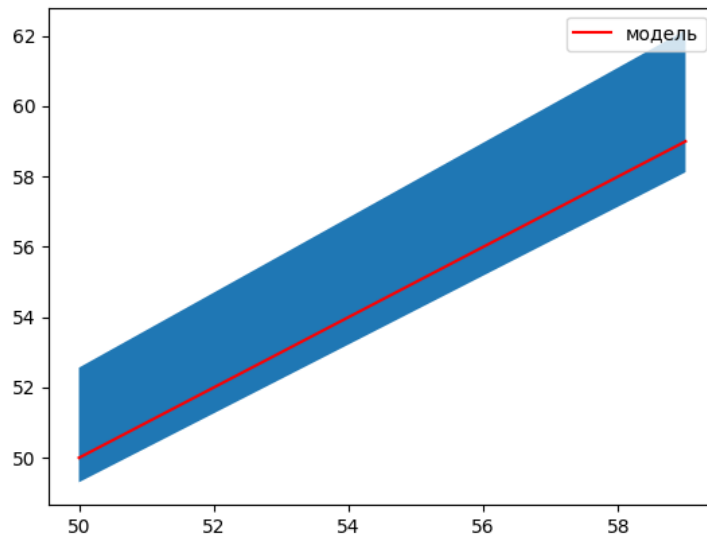


Рис. 6: Значения выходной переменной y для точек предсказания вне x

Набор $x_3 = (0.0, 0.5, \dots, 59, 59.5)$.

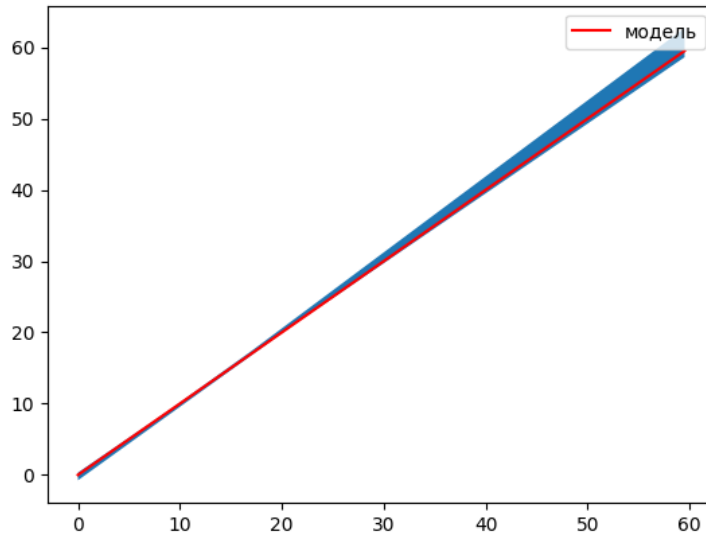


Рис. 7: Значения выходной переменной y для большого набора точек предсказания

5 Обсуждения

На рисунке 2 видно, что построенное информационное множество содержит истинные значения параметров β_0, β_1 ($\beta_0 = 0, \beta_1 = 1$). А лучшую точечную оценку параметров зависимости в этом случае даёт центр тяжести. Также стоит отметить, что точечная регрессия попала в информационное множество. На рисунке 3 более наглядно видно, что построенные точечные оценки близки к исходной модели.

Построенный коридор совместных зависимостей содержит исходную модель, что видно на рисунке 4. Коридор совместных зависимостей имеет непостоянную ширину, наименьшую неопределённость построенная модель имеет в центре исходной выборки, а при приближении к концам неопределённость выходных значений возрастает.

Тоже самое наблюдается на рисунках 5 - 6, для набора значений предсказания внутри x неопределённость выходных значений y значительно меньше, чем при значениях предсказания вне x . Также стоит отметить, что для обоих наборов предсказания полученные интервальные выходные значения y содержат исходную модель. На рисунке 7 представлена общая картина для наборов предсказания внутри и вне x .