Санкт-Петербургский политехнический университет имени Петра Великого

Институт прикладной математики и механики
Высшая школа прикладной математики и физики

Интервальный анализ
Отчёт по лабораторной работе №4

Выполнил:

Студент: Ахметшин Рауль
Группа: 5030102/90201

Принял:

к. ф.-м. н., доцент
Баженов Александр Николаевич

2023 г.

Содержание

Список иллюстраций

1. Постановка задачи

Дан набор интервальных данных. Считая что они задают линейно распределенную величину, требуется построить информационное множество параметров, корридор совместности и произвести "предсказание значений":

1. для значения между имеющимися данными (интерполяция).
2. для значений вне имеющихся данных (экстраполяция).

2. Теория

2.1. Точечная оценка параметров регрессии

Пусть x - номер измерения в выборке, а y - получившийся результат. Тогда мы можем представить линейную регрессию как

$$y = b_0 + b_1 * x$$

Для получения точечной оценки можно поставить задачу оптимизации

$$\begin{cases} \text{mid}(\text{y}_i) - w_i * rad(y_i) \leq X * \beta \leq mid(y_i) + w_i * rad(y_i) & i = 1, m \\ \sum_{i=1}^{m} w_i \to min & \\ \text{w}_i \geq 0 & i = 1, m \\ \text{w}, \beta =? & \end{cases}$$

Здесь X — матрица $m \times 2$, в первом столбце которой элементы, равные 1, во втором — значения x_i. В качестве значений середины и радиуса возьмем $mid(y_i) = y_i$ и $rad(y_i) = 1$.

2.2. Интервальная оценка параметров регрессии

В ходе вычисления точечной оценки мы получили вектор w_i, котрые являются минимальными радиусами, необходимыми для того чтобы выборка была накрывающей. Для устранения избыточной информации, примем радиусы каждого измерения равными между собой и равными величине $\epsilon = max(w_i)$.

2.3. Информационное множество параметров

Построим визуальное представление информационного множества параметров b0 и b1. Для этого воспользуемся следующим алгоритмом:

1. Для индекса i от 0 до m:
   (a) Для индекса j от $i + 1$ до m:
      i. По $(x_i, y_i \pm \epsilon)$ и $(x_j, y_j \pm \epsilon)$ построим 4 прямые.

ii. Для каждой прямой проверим, попадает ли она во все интервалы нашей выборки

iii. Если да - сохраняем параметры прямой как вершину нашего информационного множества.

3. Реализация

Лабораторная работа выполнена на языке программирования Python(3.7) с использованием следующих библиотек: Numpy, Scipy, Tabulate, Statsmodels, Matplotlib.
Отчет написан в онлайн редакторе LaTeX - Overleaf.

4. Результаты

4.1. Графики

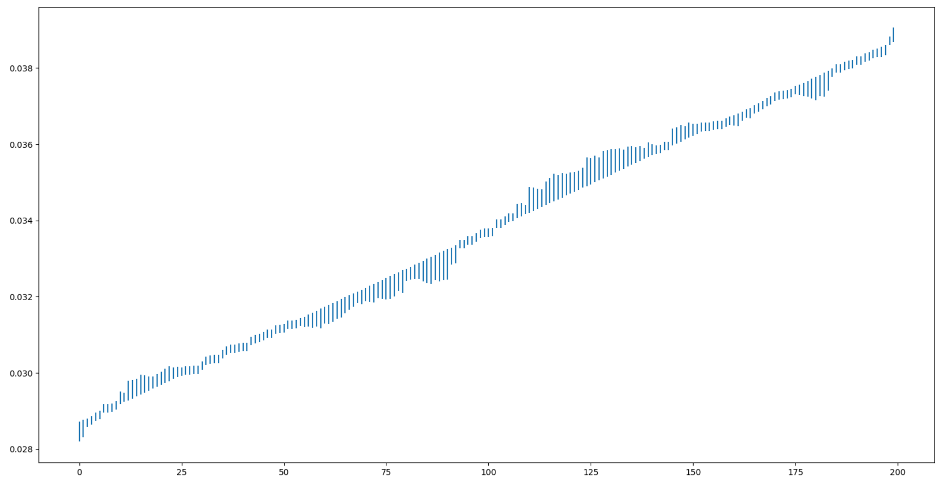

Рис. 1. График входных интервальных данных

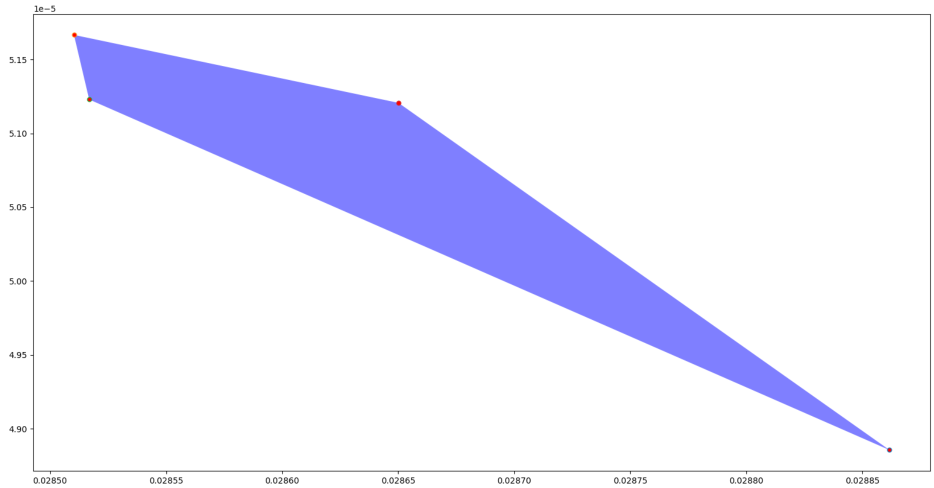

Рис. 2. Информационное множество

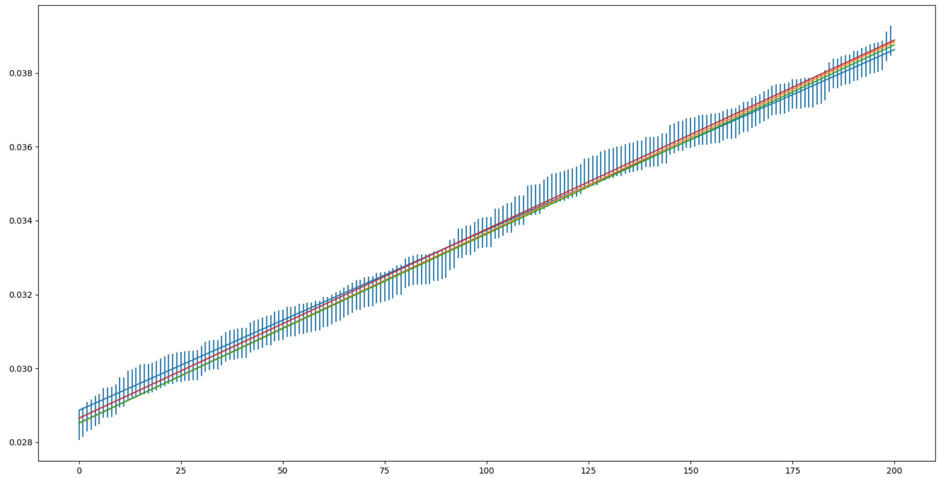

Рис. 3. Допусковый корридор

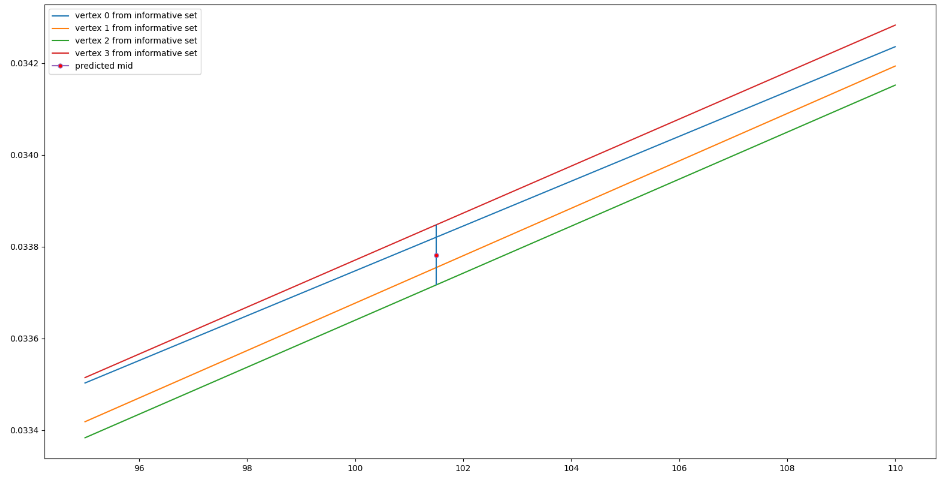

Рис. 4. Предсказание значения при аргументе 101.5

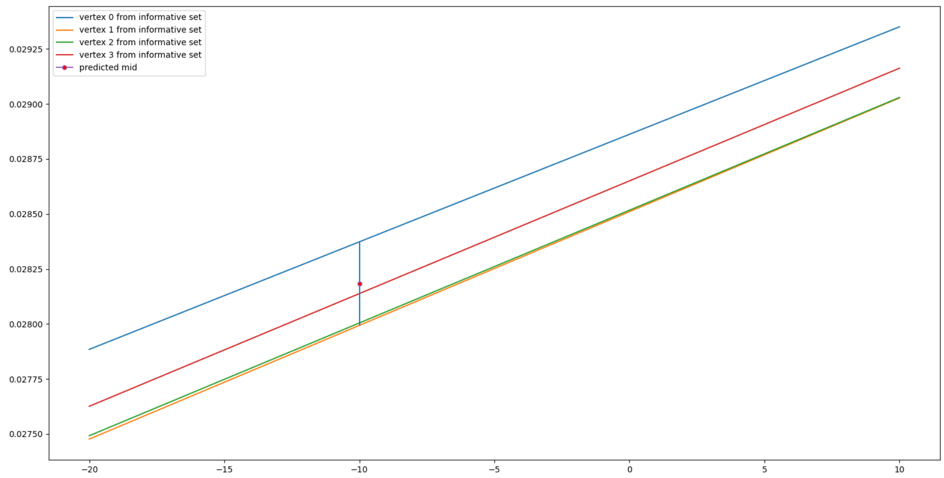

Рис. 5. Предсказание значения при аргументе -10

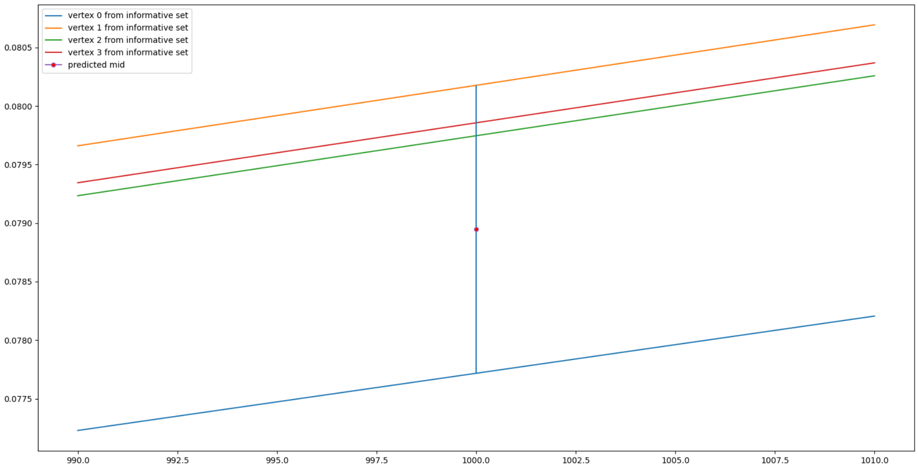

Рис. 6. Предсказание значения при аргументе -10

4.2. Числовые значения

Уравнение вершин информационного множества

$$y = 0.028553543633027520 + 5.122935779816516e05x$$
$$y = 0.028548538011976047 + 5.156306586826348e05x$$
$$y = 0.028764453647058822 + 4.952847058823532e05x$$
$$y = 0.028613429043478262 + 5.120652173913043e05x$$

Предсказанные значения

$$y(-10) == [0.02803290735329341, 0.028269168941176467]$$
$$mid = 0.028151038147234944, rad = 0.00011813079394152812$$
$$y(101.5) = [0.033753323449541274, 0.033810891]$$
$$mid = 0.033782107224770644, rad = 2.8783775229364317e05$$
$$y(1000) = [0.07829292423529415, 0.08011160388023954]$$
$$mid = 0.07920226405776684, rad = 0.0009093398224726962$$

5. Обсуждение

- Исходя из 4 можно заметить что в районе 100-ого испытания у нас наблюдается излом.
- Исходя из предсказанных значений можно заметить, что при экстаполяции погрешность гораздо больше чем при интерполяции.

- Также из предсказанных значений можно заметить, что при экстаполяции погрешность увеличивается по мере удаления от имеющихся данных.
- Прямая, полученная как центр масс информационного множества, почти совпадает с прямой, полученной в результате решения задачи оптимизации.

6. Ссылка на репозиторий

https://github.com/Enoras/IntervalLabs