Санкт-Петербургский политехнический университет
Петра Великого

Физико-механический институт
Высшая школа прикладной математики и вычислительной физики

Лабораторная работа №4
по курсу
«Интервальный анализ»

Выполнил:
Романчук Евгений Вячеславович
Группа: 5030102/10201
Проверил:
Баженов Александр Николаевич

Санкт-Петербург
2024г.-2025г.

Содержание

1 Теоретическое обоснование **2**

2 Постановка задачи **2**

3 Процесс **4**

4 Результаты **4**

5 Заключение **10**

1 Теоретическое обоснование

Перед использованием измерительной системы её необходимо откалибровать. В конечном счёте калибровка сводится к определению параметров линейной регрессии

$$\mathbf{y} = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \cdot \mathbf{x}, \tag{1}$$

где $\mathbf{x}$ — входные данные, $\mathbf{y}$ — выходные данные, $\boldsymbol{\beta}_0$, $\boldsymbol{\beta}_1$ — параметры регрессии.

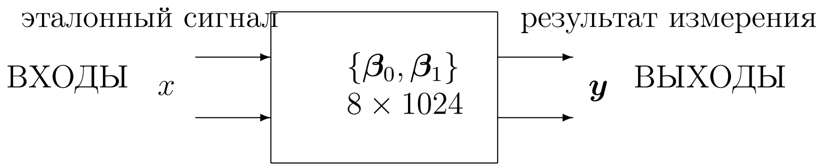

Рис. 1: Структурная схема калибровки сигнала

Параметры регрессии определяются из решения *интервальной системы линейных алгебраических уравнений.*

$$\mathbf{Y} = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \cdot \mathbf{X}. \tag{2}$$

где $\mathbf{X}$ — вектор входных калибровочных данных, $\mathbf{Y}$ — вектор выходных данных.

В общем случае входных данных получение оценок параметров регрессии является нетривиальной задачей. Внешние оценки часто получаются очень грубыми. Внутренние оценки не являются однозначными, и их вычисление является математически трудной задачей. Одновременное получение внутренних и внешних оценок мотивирует использование твинной арифметики. На этом пути получено еще не много результатов.

2 Постановка задачи

Необходимо определить параметры линейной регрессии:

$$\boldsymbol{y} = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \cdot \boldsymbol{x},$$

где:

- $\boldsymbol{x}$ — входные данные;

- $\boldsymbol{y}$ — выходные данные;
- $\boldsymbol{\beta}_0, \boldsymbol{\beta}_1$ — параметры регрессии.

Для калибровки измерителя на вход подаётся набор постоянных напряжений:

$$X = \{x_i\}.$$

Для повышения надёжности для каждого значения x_i выполняется 100 измерений, что формирует набор интервальных выборок:

$$\boldsymbol{Y} = \{y_k\}_{k=1}^{100}.$$

Разрядность $\boldsymbol{y}$ определяется как:

$$\mathrm{rad}(\boldsymbol{y}) = \frac{1}{2^N}\,\text{В}, \quad N = 14.$$

Файлы данных

- Имя архива: `27_08_2024ADC_rawData.zip`.
- Формат файлов: см. описание в документе `Save to BIN.pdf`.

Связь кодов данных с напряжением

Напряжение V вычисляется по формуле:

$$V = \frac{\text{Code}}{16384} - 0.5.$$

Задание:

1. Сделать оценки значений $\boldsymbol{Y}$ двумя способами:
   - in: как интервал между первым и третьим квартилем
   - ex: как границы бокс-плота
2. Решить ИСЛАУ (1) для внутренних и внешних оценок y
3. Построить множество решений $\boldsymbol{\beta}_0$, $\boldsymbol{\beta}_1$
4. Построить коридор совместных зависимостей
5. Пример — https://github.com/szhilin/octave-interval-examples/blob/master/SteamGenerator.ipynb

3 Процесс

- Язык программирования - python.
- Сторонние библиотеки: numpy, intvalpy
- Репозиторий на github: `https://github.com/UUyy-Geniy/Interval_analysis`

4 Результаты

0. Вид данных

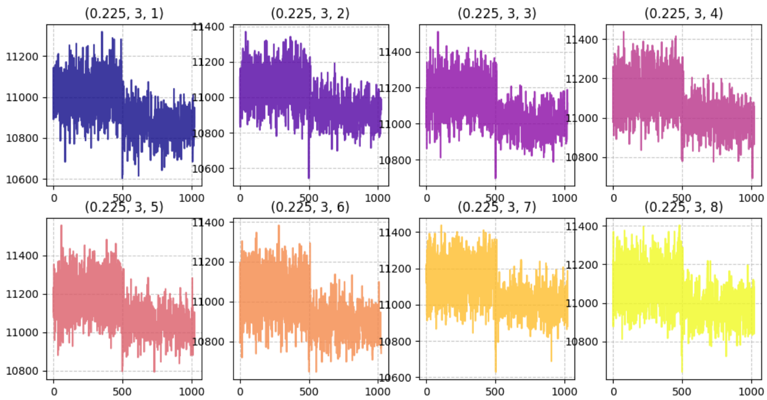

Рис. 2: Данные для уровня 0.225, 3 ячейки

1. Оценки значений Y

Для внешних оценок возьмем границы боксплота (после отсеивания выбросов берем минимум и максимум) с параметром k=1.5, а для внутренних 1-й и 3-й квартили:

ys_int	ys_ext
[1583; 1607]	[1542; 1638]
[1826; 1855]	[1784; 1885]
[5291; 5313]	[5257; 5338]
[7640; 7663]	[7606; 7680]
[7989; 8007]	[7963; 8028]
[8791; 8810]	[8765; 8833]
[10938; 10958]	[10908; 10974]
[13653; 13673]	[13626; 13701]

Таблица 1: Сравнение интервалов `ys_int` и `ys_ext`

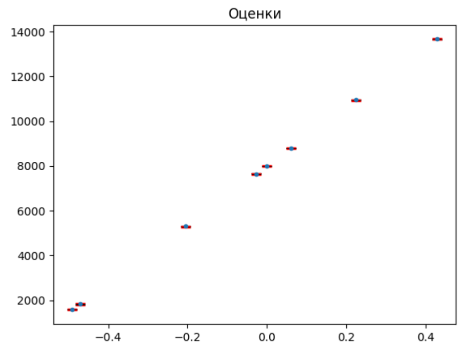

Рис. 3: Внешние и внутренние оценки для каждого файла

1. Построение допускового множества

Выберем для демонстрации работы алгоритма канал и ячейку - channel = 1, cell = 1000

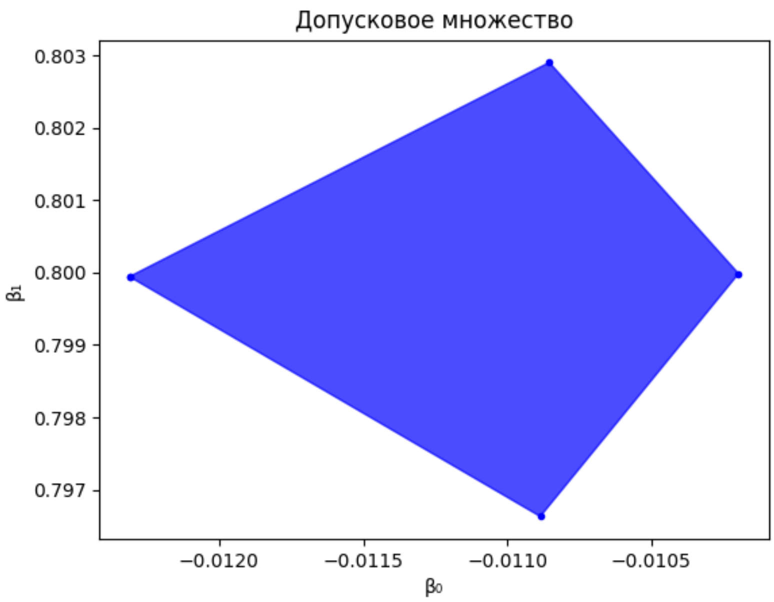

Рис. 4: Допусковое множество до расширения внутренних оценок до внешних

Видно, что на графике отображается только множество внешних оценок. Это указывает на то, что интервальная система линейных алгебраических уравнений (ИСЛАУ) для неизвестных коэффициентов регрессии несовместная. Чтобы получить множество внутренних оценок, необходимо заменить строки системы с отрицательной невязкой. Номера таких строк соответствуют номерам образующих распознающего функционала с отрицательным значением Tol.

Допусковое множество решений интервальной линейной системы пусто.

$$\textbf{tolmax:} \; -0.00083$$

$$\textbf{argmax:} \begin{bmatrix} -0.011 \\ 0.799 \end{bmatrix}$$

$$\textbf{env:} \begin{bmatrix} 8 & -0.00083 \\ 3 & -0.00083 \\ 1 & -0.00083 \\ 5 & -0.00020 \\ 6 & 0.00007 \\ 7 & 0.00014 \\ 4 & 0.00038 \\ 2 & 0.00065 \end{bmatrix}$$

Расширяем внутренние оценки до внешних:

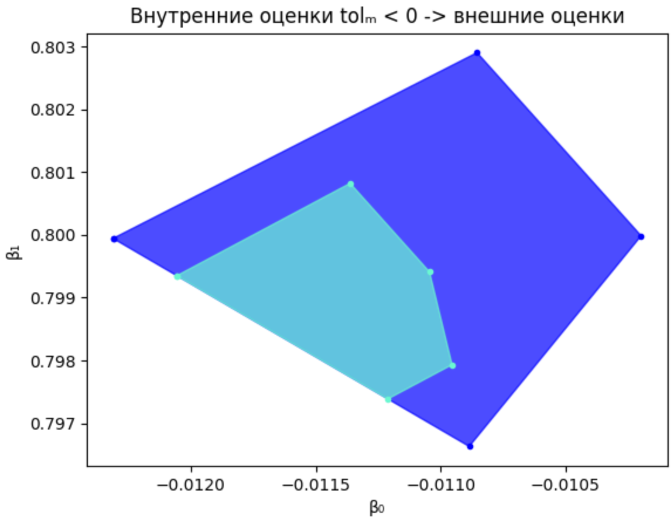

Рис. 5: Допусковое множество до расширения внутренних оценок до внешних

Допусковое множество решений интервальной линейной системы непусто.

tolmax: 0.00051

$$\textbf{argmax:} \begin{bmatrix} -0.0115 \\ 0.799 \end{bmatrix}$$

$$\textbf{env:} \begin{bmatrix} 6 & 5.067 \cdot 10^{-4} \\ 8 & 5.079 \cdot 10^{-4} \\ 2 & 5.089 \cdot 10^{-4} \\ 7 & 5.169 \cdot 10^{-4} \\ 4 & 5.617 \cdot 10^{-4} \\ 1 & 1.057 \cdot 10^{-3} \\ 3 & 1.220 \cdot 10^{-3} \\ 5 & 1.538 \cdot 10^{-3} \end{bmatrix}$$

	$b_{\textbf{int}}$	$b_{\textbf{out}}$
$\boldsymbol{\beta}_0$	$[-0.0121; -0.0110]$	$[-0.0123; -0.0102]$
$\boldsymbol{\beta}_1$	$[0.7974; 0.8008]$	$[0.7966; 0.8029]$

Таблица 2: Интервальные (b_{int}) и внешние (b_{out}) оценки параметров регрессии

Ограничения	**Значения**
Активные нижние ограничения	$[1, 2, 3, 4, 5, 6, 7]$
Активные верхние ограничения	$[0, 1, 2, 3, 4, 5, 6, 7, 8]$

Таблица 3: Активные ограничения для интервальных и внешних оценок

X_i	x_1	x_2
-0.492	$[-0.492; -0.487]$	$[-0.495; -0.484]$
-0.471	$[-0.474; -0.468]$	$[-0.477; -0.466]$
-0.205	$[-0.208; -0.204]$	$[-0.211; -0.202]$
-0.027	$[-0.029; -0.025]$	$[-0.031; -0.024]$
0.0	$[-0.0018; 0.0010]$	$[-0.0038; 0.0026]$
0.061	$[0.059; 0.062]$	$[0.057; 0.064]$
0.225	$[0.223; 0.227]$	$[0.221; 0.228]$
0.43	$[0.430; 0.435]$	$[0.428; 0.437]$

Таблица 4: Результаты применения вычисленных коэффициентов к тренировочным данным

Коридор значений

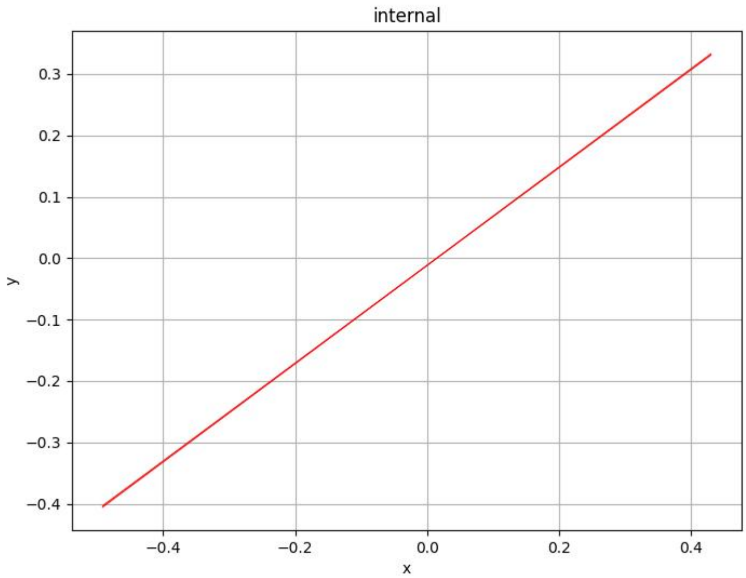

Рис. 6: Коридор значений

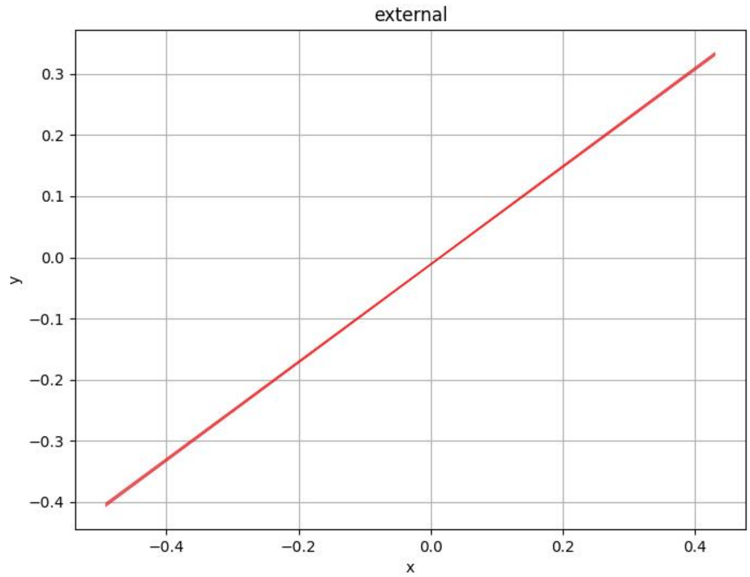

Рис. 7: Коридор значений (правый край)

5 Заключение

В рамках выполнения лабораторной работы была представлена методика интервальной оценки параметров линейной регрессии. Проведённые исследования и вычисления позволили достичь следующих результатов:

- Реализован алгоритм, обеспечивающий расчёт внутренних и внешних границ параметров регрессии, что делает возможным учёт неопределённости в данных.
- Выполнен анализ коэффициентов β_0 и β_1 в формате интервалов, что позволяет отразить диапазон их вероятных значений.
- Построены визуальные представления зависимости параметров, которые предоставляют наглядное отображение полученных решений и способствуют изучению устойчивости модели.

Предложенный метод позволяет эффективно работать с данными, содержащими ошибки и неопределённости, и предлагает инструменты для получения надёжных оценок параметров. Это делает его полезным в широком спектре практических задач, включая анализ данных, прогнозирование и построение регрессионных моделей в условиях неполной информации.