

САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
ПЕТРА ВЕЛИКОГО

ФИЗИКО-МЕХАНИЧЕСКИЙ ИНСТИТУТ
КАФЕДРА «ПРИКЛАДНАЯ МАТЕМАТИКА»

Отчёт
по лабораторной работе №10
по дисциплине
«Математическая статистика»

Выполнил:
Анищенко Михаил Денисович
группа:
5030102/00201

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

Санкт-Петербург
2023 г.

Содержание

1	Постановка задачи	4
2	Реализация	5
3	Результаты	5
3.1	Диаграмма рассеяния	5
3.2	Варьирование неопределенности измерений	6
3.3	Варьирование неопределённости измерений с расширением и сужением интервалов . . .	7
3.4	Анализ регрессионных остатков	8
3.5	Информационное множество задачи	9
3.6	Коридор совместных зависимостей	10
3.7	Построение прогноза внутри и вне области данных	10
4	Обсуждение	11
4.1	Варьирование неопределенности измерений	11
4.2	Варьирование неопределенности измерений с расширением и сужением интервалов . . .	11
4.3	Анализ регрессионных остатков	11
4.4	Информационное множество задачи	11
4.5	Коридор совместных зависимостей	11
4.6	Построение прогноза внутри и вне области данных	11

Список иллюстраций

1	Диаграмма рассеяния выборки \mathbf{X}_1 с уравновешенным интервалом погрешности	5
2	Диаграмма рассеяния выборки \mathbf{X}_1 и регрессионная прямая по модели (2.35) и (2.36) . .	6
3	Диаграмма рассеяния выборки \mathbf{X}_1 и регрессионная прямая по модели (2.41) и (2.42) . .	7
4	Векторы ω_1 и ω_2	7
5	Диаграмма рассеяния по модели (2.35) и (2.36)	8
6	Диаграмма рассеяния регрессионных остатков выборки \mathbf{X}_1 по (2.41) и (2.42)	8
7	Частоты элементарных подинтервалов регрессионных остатков выборки \mathbf{X}_1 по модели (2.35) и (2.36) — красный график, и (2.41) и (2.42) — синий график	9
8	Информационное множество по модели (2.41) и (2.42), интервальная оболочка — красный брус	9
9	Коридор совместных зависимостей (2.54)	10
10	Коридор совместных зависимостей (2.54). Построение прогноза	10

Список таблиц

1 Постановка задачи

Дадим общую формулировку задачи восстановления функциональной зависимости. Пусть некоторая величина y является функцией от независимых переменных x_1, x_2, \dots, x_m :

$$y = f(\beta, x) \quad (1)$$

где $x = (x_1, x_2, \dots, x_m)$ является вектором независимых переменных, $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ — вектор параметров функции. Заметим, что переменные x_1, x_2, \dots, x_m также называются входными, а переменные y_1 — выходной.

Задача восстановления функциональной зависимости заключается в том, чтобы, располагая набором значений x и y , найти такие $\beta_1, \beta_2, \dots, \beta_p$ в выражении (1), которые соответствуют конкретной функции f из параметрического семейства.

Если функция f является линейной, то можно записать

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_m x_m \quad (2)$$

В общем случае результаты измерений величин x_1, x_2, \dots, x_m и y являются интервальнозначными

$$x_1^{(k)}, x_2^{(k)}, \dots, x_m^{(k)}, y^k.$$

Индекс k пробегает значения от 1 до n , равного полному числу измерений.

Определение 2.2.1 Брусом неопределенности k -го измерения функциональной зависимости будем называть интервальный вектор-брус, образованный интервальными результатами измерений с одинаковыми значениями индекса k [1]:

$$(x_{k1}, x_{k2}, \dots, x_{km}, y_k) \subset \mathbb{R}^{m+1}, k = 1, 2, \dots, n. \quad (3)$$

Брус неопределенности измерения является прямым декартовым произведением интервалов неопределенности независимых переменных и зависимой переменной.

2 Реализация

Лабораторная работа выполнена с помощью пакета GNU Octave с использованием следующих сторонних файлов:

- <https://github.com/AlexanderBazhenov/Solar-Data>
- <https://github.com/szhilin/octave-interval-examples>

Отчет подготовлен с помощью Overleaf.

3 Результаты

3.1 Диаграмма рассеяния

Данные для выборки взяты из файла `octave/Channel_1_500nm_0_23mm.csv`, погрешность прибора $\epsilon = 10^{-4}$.

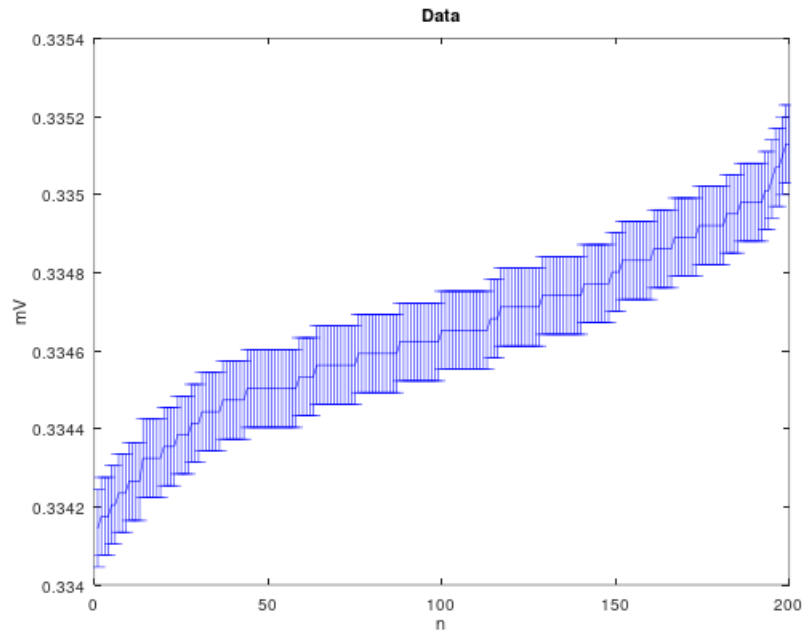


Рис. 1: Диаграмма рассеяния выборки \mathbf{X}_1 с уравновешенным интервалом погрешности

3.2 Варьирование неопределенности измерений

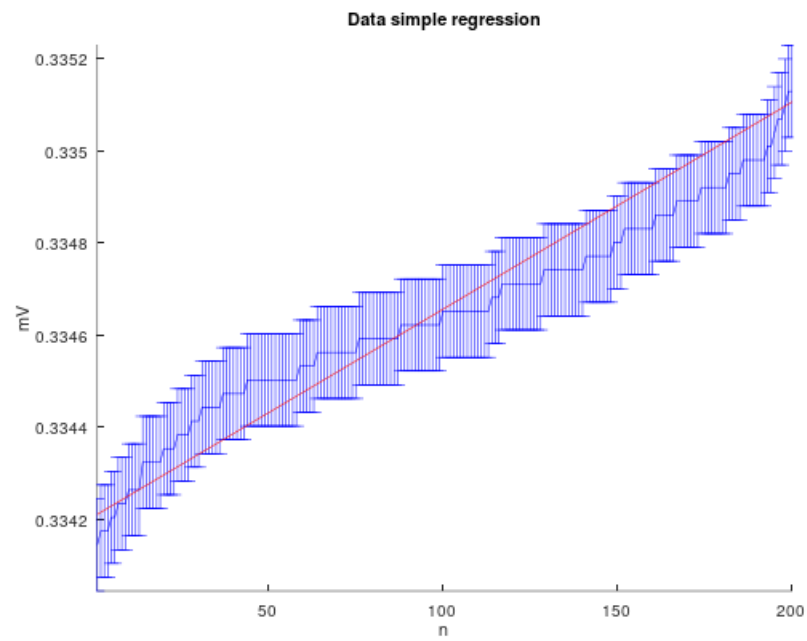


Рис. 2: Диаграмма рассеяния выборки \mathbf{X}_1 и регрессионная прямая по модели (2.35) и (2.36)

$$\sum_{i=1}^n \omega_i = 200, \beta_0 = 0.334207, \beta_1 = 4.5 \cdot 10^{-6}$$

3.3 Варьирование неопределённости измерений с расширением и сужением интервалов

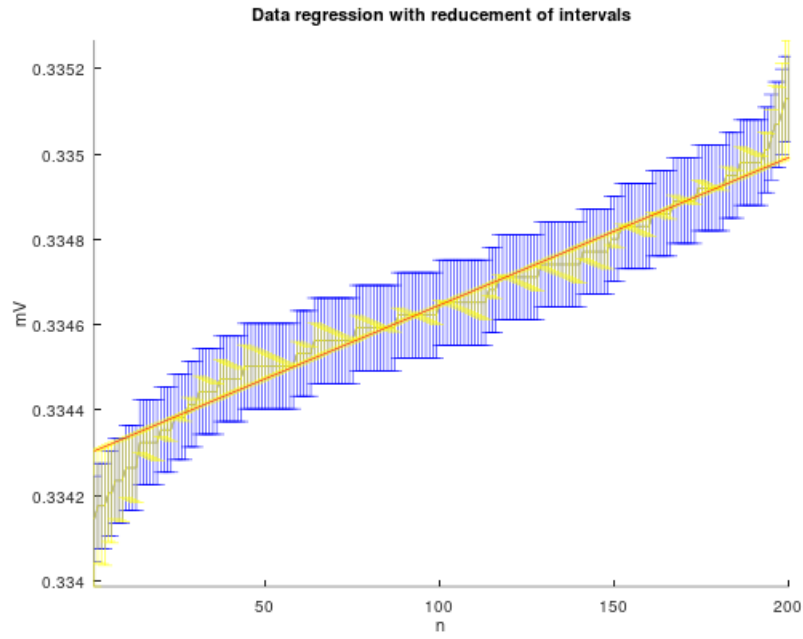


Рис. 3: Диаграмма рассеяния выборки \mathbf{X}_1 и регрессионная прямая по модели (2.41) и (2.42)

$$\sum_{i=1}^n \omega_i = 54.021, \beta_0 = 0.334301, \beta_1 = 3.5 \cdot 10^{-6}$$

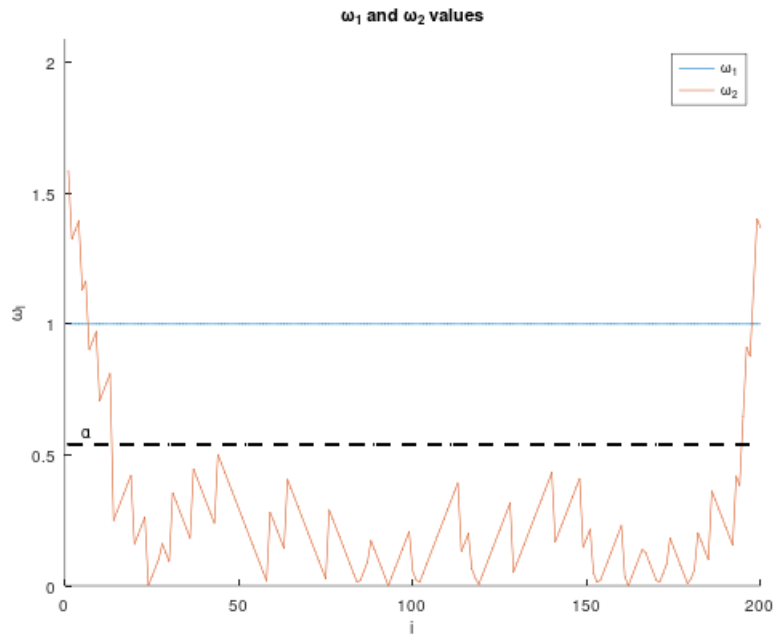


Рис. 4: Векторы ω_1 и ω_2

3.4 Анализ регрессионных остатков

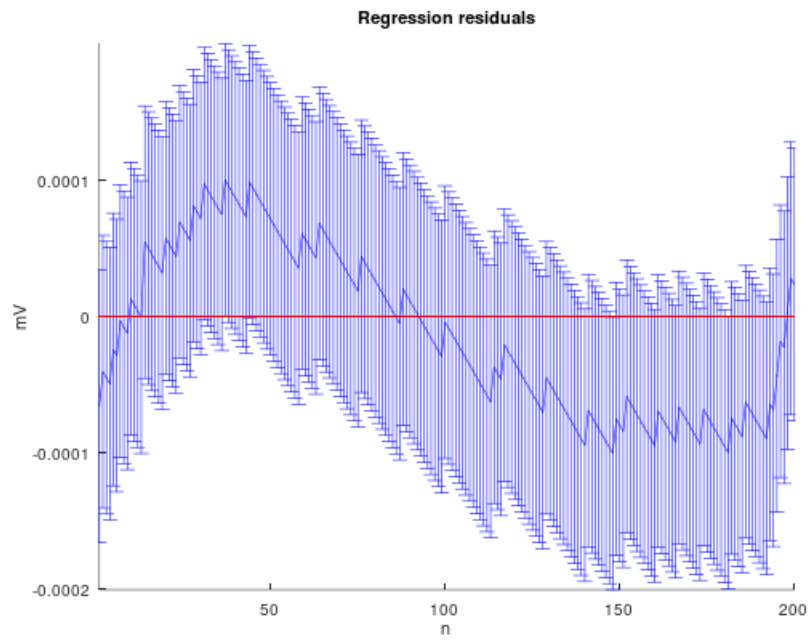


Рис. 5: Диаграмма рассеяния по модели (2.35) и (2.36)

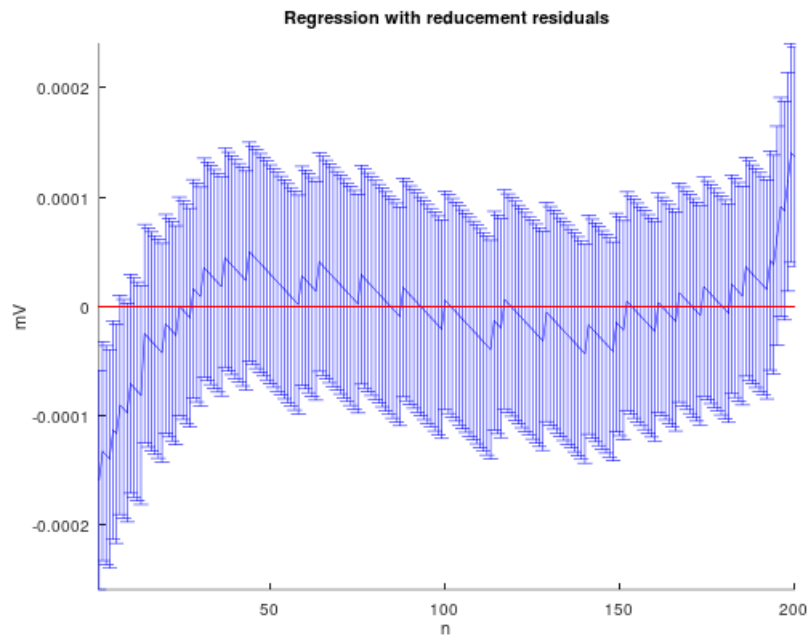


Рис. 6: Диаграмма рассеяния регрессионных остатков выборки \mathbf{X}_1 по (2.41) и (2.42)

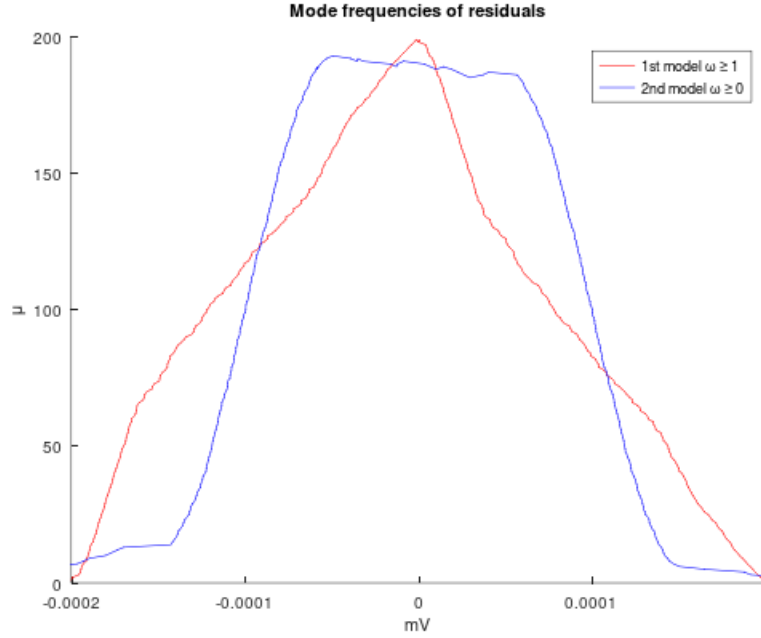


Рис. 7: Частоты элементарных подинтервалов регрессионных остатков выборки \mathbf{X}_1 по модели (2.35) и (2.36) — красный график, и (2.41) и (2.42) — синий график

3.5 Информационное множество задачи

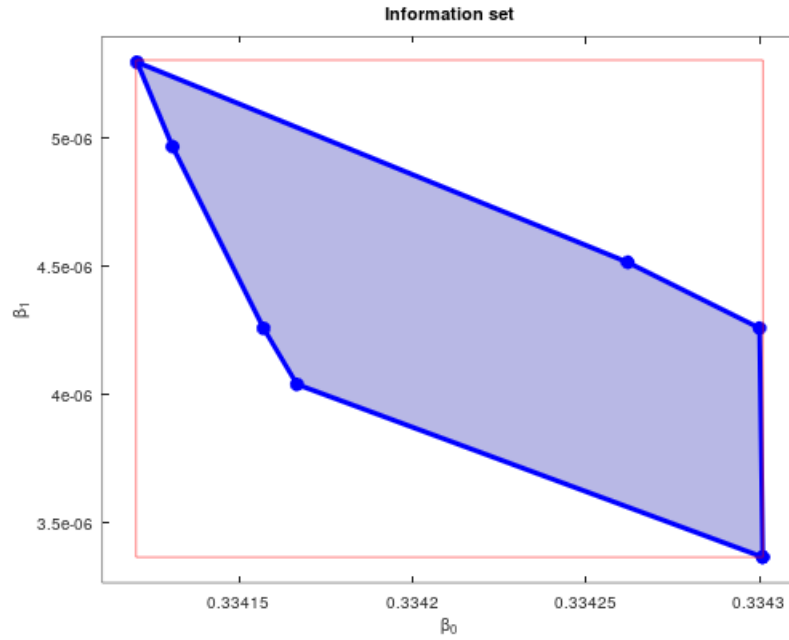


Рис. 8: Информационное множество по модели (2.41) и (2.42), интервальная оболочка — красный брус

$$\begin{aligned} \text{mid } \beta_0 &= [0.334120, 0.334301] \\ \text{mid } \beta_1 &= [3.4 \cdot 10^{-6}, 5.3 \cdot 10^{-6}] \end{aligned}$$

3.6 Коридор совместных зависимостей

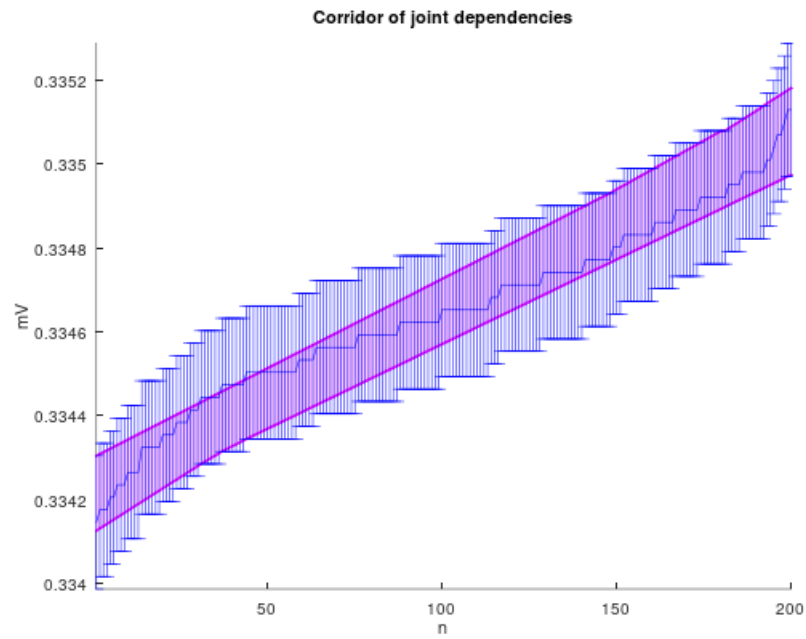


Рис. 9: Коридор совместных зависимостей (2.54)

3.7 Построение прогноза внутри и вне области данных

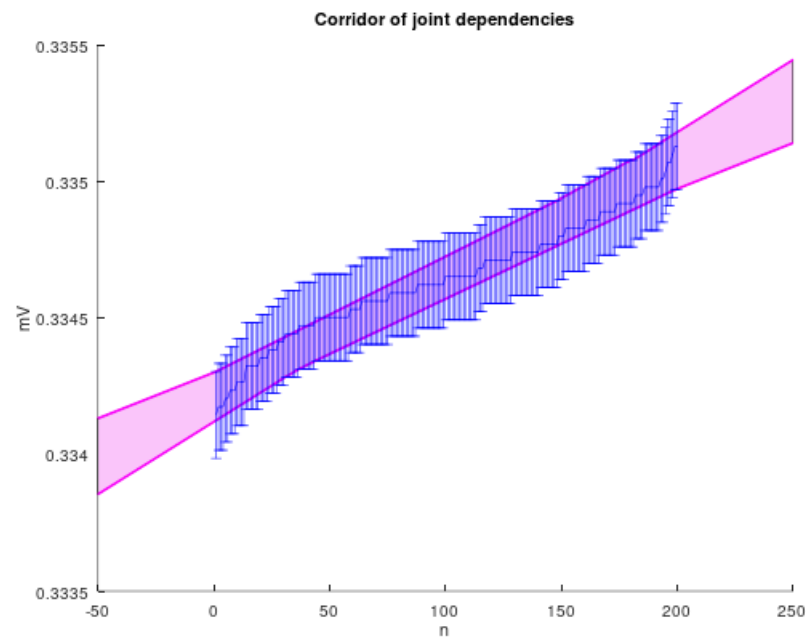


Рис. 10: Коридор совместных зависимостей (2.54). Построение прогноза

4 Обсуждение

4.1 Варьирование неопределенности измерений

Для модели регрессии с $\omega_i \geq 1$ видим, что все $\omega_i = 1$, а регрессионная прямая пересекает каждый отрезок без необходимости увеличения какого-либо из них.

4.2 Варьирование неопределенности измерений с расширением и сужением интервалов

Для модели регрессии с $\omega_i \geq 0$ видим, что для большинства интервалов $\omega_i < 1$, однако в начале и конце имеются выбросы $\omega_i \approx 1.5$. Кроме того из рисунка видно, что регрессионная прямая пересекает не все интервалы. Это объясняется тем, что некоторые из них были увеличены, и регрессионная прямая пересекает измененные (желтые) интервалы, притом пересекая увеличенный интервал она не обязана пересечь исходный.

4.3 Анализ регрессионных остатков

По результатам вычислений для регрессионных остатков можно сделать вывод, что мода регрессионных остатков по модели с $\omega_i \geq 0$ представляет собой более широкую окрестность нуля. Это означает, что регрессия по этой модели качественнее, чем по модели $\omega_i \geq 1$.

4.4 Информационное множество задачи

Из графика для информационного множества задачи видим, что решение классических численных задач в интервальных постановках является не интервалом, а многогранным множеством. При этом можно дать независимые оценки для компонент решения, построив интервальную оболочку.

4.5 Коридор совместных зависимостей

По результатам построения коридора совместных зависимостей получено множество, любая прямая, лежащая в котором, будет являться совместной регрессионной зависимостью для данной интервальной выборки.

4.6 Построение прогноза внутри и вне области данных

На уменьшенном рисунке для коридора совместных зависимостей видим, что он сужается ближе к центру выборки по n и расширяется при отдалении от центра. Такие результаты весьма ожидаемы.