

Санкт-Петербургский политехнический университет  
Высшая школа прикладной математики и  
вычислительной физики,  
Физико-механический институт

Отчет по лабораторной работе №1  
по дисциплине «Интервальный анализ»

Выполнил студент гр. 5030102/80201  
Дойников И. Д.  
Руководитель  
Баженов А. Н.

Санкт-Петербург  
2021

# Содержание

	Страница
<b>1 Постановка задачи</b>	<b>4</b>
1.1 Линейная и полиномиальная регрессия . . . . .	4
1.2 Решение задач томографии . . . . .	4
1.3 Глобальная оптимизация . . . . .	4
<b>2 Теория</b>	<b>4</b>
2.1 Выяснение радиуса элементов матрицы, при котором она становится особенной . . . . .	4
2.1.1 Критерий Баумана . . . . .	4
2.1.2 Признак Румпа . . . . .	5
2.2 Глобальная оптимизация . . . . .	5
<b>3 Реализация</b>	<b>5</b>
<b>4 Результаты</b>	<b>5</b>
4.1 Линейная и полиномиальная регрессия . . . . .	5
4.2 Задача томографии . . . . .	6
4.3 Глобальная оптимизация . . . . .	6
4.3.1 Функция с одним экстремумом . . . . .	6
4.3.2 Функция с несколькими экстремумами . . . . .	9
<b>5 Обсуждение</b>	<b>12</b>

## Список иллюстраций

	<b>Страница</b>
1    График функции МакКормика (4) . . . . .	7
2    Иллюстрация работы алгоритма для функции МакКормика (4) . . . . .	7
3    Расстояние до точки экстремума для функции МакКормика (4) . . . . .	8
4    Радиусы рабочих брусков для функции МакКормика (4) . . . . .	9
5    График функции Химмельблау (5) . . . . .	9
6    Иллюстрация работы алгоритма для функции Химмельблау (5) . . . . .	10
7    Расстояние до точки экстремума для функции Химмельблау (5) . . . . .	11
8    Радиусы рабочих брусков для функции Химмельблау (5) . . . . .	11

# 1 Постановка задачи

## 1.1 Линейная и полиномиальная регрессия

Дана задача вида  $\mathbf{y} = \mathbf{X}\beta$ , в общей регрессионной постановке необходимо найти вектор  $\beta$  и восстановить зависимости между рассматриваемыми величинами. В данной работе рассматривается вопрос установления особенности матрицы  $\mathbf{X}$ , так как в этом случае рассматриваемая ИСЛАУ становится неразрешимой.

Дано

$$\text{mid } \mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1.1 & 1 \end{pmatrix} \quad (1)$$

Необходимо рассмотреть матрицу вида

$$\mathbf{X} = \begin{pmatrix} [1 - \varepsilon, 1 + \varepsilon] & 1 \\ [1.1 - \varepsilon, 1.1 + \varepsilon] & 1 \end{pmatrix} \quad (2)$$

и определить, при каком значении  $\varepsilon$  она содержит особенные точечные матрицы.

## 1.2 Решение задач томографии

При решении задач томографии появляются системы типа  $\mathbf{A}x = \mathbf{b}$ . В данной работе необходимо рассмотреть матрицу

$$\mathbf{A} = \begin{pmatrix} [1 - \varepsilon, 1 + \varepsilon] & [1 - \varepsilon, 1 + \varepsilon] \\ [1.1 - \varepsilon, 1.1 + \varepsilon] & [1 - \varepsilon, 1 + \varepsilon] \end{pmatrix} \quad (3)$$

и определить, при каком радиусе она содержит особенную матрицу.

## 1.3 Глобальная оптимизация

Для функции МакКормика

$$f(x, y) = \sin(x + y) + (x - y)^2 - 1.5x + 2.5y + 1, \quad (4)$$

имеющей один глобальный экстремум, и функции Химмельблау

$$f(x, y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2, \quad (5)$$

имеющей 4 равнозначных глобальных экстремума, необходимо провести вычисления по поиску глобального минимума с помощью простейшего интервального адаптивного алгоритма глобальной оптимизации.

# 2 Теория

## 2.1 Выяснение радиуса элементов матрицы, при котором она становится особенной

Интервальная матрица  $\mathbf{A}$  - особенная, если  $\exists A \in \mathbf{A} : \det(A) = 0$ , иначе - неособенная.

$\text{rad } \mathbf{A} = \{A \mid a_{ij} = \text{rad } \mathbf{a}_{ij}\}$  - радиус интервальной матрицы  $\mathbf{A}$ ,  $\text{mid } \mathbf{A} = \{A \mid a_{ij} = \text{mid } \mathbf{a}_{ij}\}$  - середина интервальной матрицы  $\mathbf{A}$ .

$\text{vert } \mathbf{A} = \{A \in \mathbf{A} \mid A = \{a_{ij}\}, a_{ij} \in \{\underline{\mathbf{a}_{ij}}, \overline{\mathbf{a}_{ij}}\}\}$  - множество вершин интервальной матрицы  $\mathbf{A}$ .

$\sigma(A)$  - множество сингулярных чисел матрицы, вычисляемых как арифметический квадратный корень из собственных значений матрицы  $AA^T$ .

### 2.1.1 Критерий Баумана

Матрица  $\mathbf{A} \in \mathbb{IR}^{n \times n}$  неособенна  $\Leftrightarrow \forall A', A'' \in \text{vert } \mathbf{A} \det(A') \cdot \det(A'') > 0$

### 2.1.2 Признак Румпа

$\mathbf{A} \in \mathbb{R}^{n \times n}$ ,  $\sigma_{\max}(\text{rad } \mathbf{A}) < \sigma_{\min}(\text{mid } \mathbf{A}) \Rightarrow \mathbf{A}$  неособенна.

## 2.2 Глобальная оптимизация

Суть простейшего интервального адаптивного алгоритма глобальной оптимизации похожа на алгоритм дихотомии, только для многомерного случая. Имеется рабочий список рассматриваемых брусьев, для каждого из которых вычислено целевое значение функции (в интервальном смысле). На каждой итерации метод выбирает из этого списка брус, на котором нижняя оценка значения функции наименьшая. Этот брус удаляется из списка, после чего туда добавляются два новых, которые получились из исходного путем дробления его самой длинной компоненты пополам (от нижней границы до середины и от середины до верхней границы). На этих брусках вычисляется интервальная оценка целевой функции, выполняется переход на новую итерацию.

## 3 Реализация

Для осуществления вычислений и визуализации результатов использовалась среда Matlab с библиотекой интервальной арифметики IntLab.

## 4 Результаты

### 4.1 Линейная и полиномиальная регрессия

Для того, чтобы точечная матрица была особенной, достаточно линейной зависимости ее строк. В случае рассматриваемой матрицы (2) это означает, что нам достаточно добиться непустого пересечения интервалов  $\mathbf{x}_{11}$  и  $\mathbf{x}_{21}$ . Очевидно, что это достижимо при  $\varepsilon \geq 0.05$ . Проверим, что при  $\varepsilon = 0.05$  найдется особенная точечная матрица:

$$\mathbf{X} = \begin{pmatrix} [0.95, 1.05] & 1 \\ [1.05, 1.15] & 1 \end{pmatrix}, X = \begin{pmatrix} 1.05 & 1 \\ 1.05 & 1 \end{pmatrix} \in \mathbf{X}, \det(X) = 0$$

Воспользуемся признаком Румпа.

$$\text{rad } \mathbf{X} = \begin{pmatrix} \varepsilon & 0 \\ \varepsilon & 0 \end{pmatrix}$$

$$\sigma(\text{rad } \mathbf{X}) = \{0, \varepsilon\sqrt{2}\}, \varepsilon > 0 \Rightarrow \sigma_{\max}(\text{rad } \mathbf{A}) = \varepsilon\sqrt{2}$$
$$\sigma(\text{mid } \mathbf{X}) = \left\{ \sqrt{\frac{421 + 21\sqrt{401}}{200}}, \sqrt{\frac{421 - 21\sqrt{401}}{200}} \right\} \approx \{2.0512, 0.0488\} \Rightarrow \sigma_{\min}(\text{mid } \mathbf{X}) = 0.0488$$

Можно сделать вывод, что при  $\varepsilon\sqrt{2} < 0.0488 \Rightarrow \varepsilon < 0.0345$  матрица  $\mathbf{X}$  не будет особенной.

Также воспользуемся критерием Баумана. Множество  $\text{vert } \mathbf{X}$  содержит 4 точечные матрицы. Обозначим их определители за  $\Delta_i$ ,  $i = \overline{1, 4}$  и считаем.

$$\Delta_1 = -0.1 \quad \Delta_2 = 2\varepsilon - 0.1 \quad \Delta_3 = -2\varepsilon - 0.1 \quad \Delta_4 = -0.1$$

Из этих значений уникальными являются  $\Delta_i$ ,  $i = \{1, 2, 3\}$ . С помощью критерия Баумана можно найти значения  $\varepsilon$ , для которых рассматриваемая матрица будет неособенной, а так как критерий является необходимым и достаточным условием, то дополнение найденного множества будет описывать все особенные матрицы  $\mathbf{A}$ .

$\Delta_1 < 0 \forall \varepsilon$ . Следовательно, для выполнения критерия необходимо потребовать, чтобы все остальные определители тоже были меньше 0. Учитывая  $\varepsilon > 0$ , получаем множество  $\varepsilon < 0.05$ . Таким образом, матрица  $\mathbf{X}$  особенна  $\forall \varepsilon \geq 0.05$ , что согласуется с предыдущими расчетами.

## 4.2 Задача томографии

Воспользуемся признаком Румпа.

$$\text{rad } \mathbf{A} = \begin{pmatrix} \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \end{pmatrix} \quad \text{mid } \mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1.1 & 1 \end{pmatrix}$$

$$\sigma(\text{rad } \mathbf{A}) = \{0, 2\varepsilon\}, \varepsilon > 0 \Rightarrow \sigma_{\max}(\text{rad } \mathbf{A}) = 2\varepsilon$$

$$\sigma(\text{mid } \mathbf{A}) = \left\{ \sqrt{\frac{421 + 21\sqrt{401}}{200}}, \sqrt{\frac{421 - 21\sqrt{401}}{200}} \right\} \approx \{2.0512, 0.0488\} \Rightarrow \sigma_{\min}(\text{mid } \mathbf{A}) = 0.0488$$

Можно сделать вывод, что при  $2\varepsilon < 0.0488 \Rightarrow \varepsilon < 0.0244$  матрица  $\mathbf{A}$  не будет особенной.

Для решения задачи воспользуемся критерием Баумана. Множество  $\text{vert } \mathbf{A}$  содержит 16 точечных матриц. Обозначим их определители за  $\Delta_i$ ,  $i = \overline{1, 16}$  и сосчитаем.

$$\Delta_1 = 0.1\varepsilon - 0.1 \quad \Delta_2 = -2\varepsilon^2 + 2.1\varepsilon - 0.1 \quad \Delta_3 = 2\varepsilon^2 - 2.1\varepsilon - 0.1 \quad \Delta_4 = 2\varepsilon^2 - 1.9\varepsilon - 0.1$$

$$\Delta_5 = -2\varepsilon^2 + 2.1\varepsilon - 0.1 \quad \Delta_6 = -0.1\varepsilon - 0.1 \quad \Delta_7 = 0.1\varepsilon - 0.1 \quad \Delta_8 = 4.1\varepsilon - 0.1$$

$$\Delta_9 = -4.1\varepsilon - 0.1 \quad \Delta_{10} = 0.1\varepsilon - 0.1 \quad \Delta_{11} = -0.1\varepsilon - 0.1 \quad \Delta_{12} = -2\varepsilon^2 - 2.1\varepsilon - 0.1$$

$$\Delta_{13} = 2\varepsilon^2 + 2.1\varepsilon - 0.1 \quad \Delta_{14} = 2\varepsilon^2 + 1.9\varepsilon - 0.1 \quad \Delta_{15} = -2\varepsilon^2 - 2.1\varepsilon - 0.1 \quad \Delta_{16} = -0.1\varepsilon - 0.1$$

Из этих значений уникальными являются  $\Delta_i$ ,  $i = \{1, 2, 3, 4, 6, 8, 9, 12, 13, 14\}$ . С помощью критерия Баумана можно найти значения  $\varepsilon$ , для которых рассматриваемая матрица будет неособенной, а так как критерий является необходимым и достаточным условием, то дополнение найденного множества будет описывать все особенные матрицы  $\mathbf{A}$ .

Учитывая, что  $\varepsilon > 0$ , получаем  $\Delta_6 < 0 \forall \varepsilon$ . Следовательно, для выполнения критерия необходимо потребовать, чтобы все остальные определители тоже были меньше 0. Получаем следующее множество решений:

$$(0, 1) \cap (0, 0.0456) \cap (0, 1.095) \cap (0, 1) \cap (0, \infty) \cap (0, 0.0244) \cap (0, \infty) \cap (0, \infty) \cap (0, 0.0456) \cap (0, 0.05)$$

То есть матрица  $\mathbf{A}$  неособенна тогда и только тогда, когда  $\varepsilon < \frac{1}{41}$ . То есть  $\forall \varepsilon \geq \frac{1}{41}$  матрица  $\mathbf{A}$  будет содержать особенные точечные матрицы. Рассмотрим  $\varepsilon = \frac{1}{41}$ :

$$\mathbf{A} = \begin{pmatrix} \left[ \frac{40}{41}, \frac{42}{41} \right] & \left[ \frac{40}{41}, \frac{42}{41} \right] \\ \left[ 1, \frac{31}{410}, 1, \frac{51}{410} \right] & \left[ \frac{40}{41}, \frac{42}{41} \right] \end{pmatrix}$$

$$A = \begin{pmatrix} 1, \frac{1}{41} & \frac{40}{41} \\ 1, \frac{31}{410} & 1, \frac{1}{41} \end{pmatrix} \in \mathbf{A}, \quad \det(A) = 0$$

## 4.3 Глобальная оптимизация

### 4.3.1 Функция с одним экстремумом

Для функции (4) в качестве начального приближения был выбран брус  $\begin{pmatrix} [-1.5, 4] \\ [-3, 4] \end{pmatrix}$ .

Для наглядности иллюстраций в этом и следующем численном эксперименте число итераций ограничено 200.

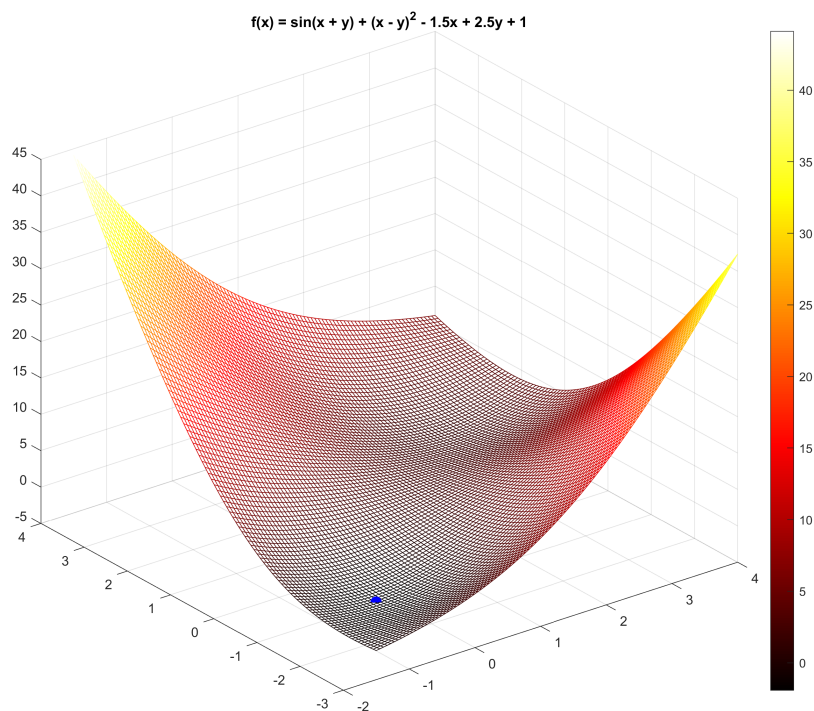


Рис. 1: График функции МакКормика (4)

Минимум изображен на графике синей точкой. Получены следующие результаты:

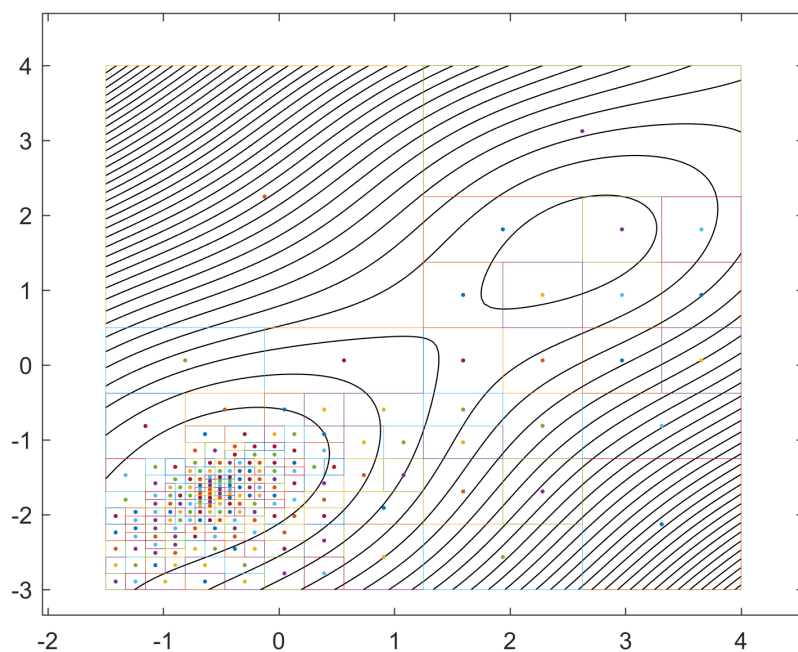


Рис. 2: Иллюстрация работы алгоритма для функции МакКормика (4)

Как видим, число брусьев сгущается по мере приближения к экстремуму. Также это явление наблюдается и около локального экстремума в окрестности точки  $(2.5, 1.5)$ .

В качестве ответа программа предлагает наименьшую нижнюю оценку значения функции за 200 итераций, поэтому в качестве итоговой оценки положения экстремума

будет выбран брус, отвечающий этому значению  $\mathbf{x} = \begin{pmatrix} [-0.3829, -0.2968] \\ [-1.5782, -1.4687] \end{pmatrix}$ . Для него получаем значение функции  $f(\mathbf{x}) = [-2.3019, -1.3809]$ . Точное значение минимума -  $f(-0.5471, -1.5471) = -1.9133$ .

Проиллюстрируем работу алгоритма еще двумя графиками. Для их построения были взяты брусы из рабочего списка алгоритма. Для построения графика радиусов брусков бралась наибольшая компонента радиуса, так как в предложенном методе последующее дробление бруса производится по измерению, отвечающему наибольшей компоненте радиуса.

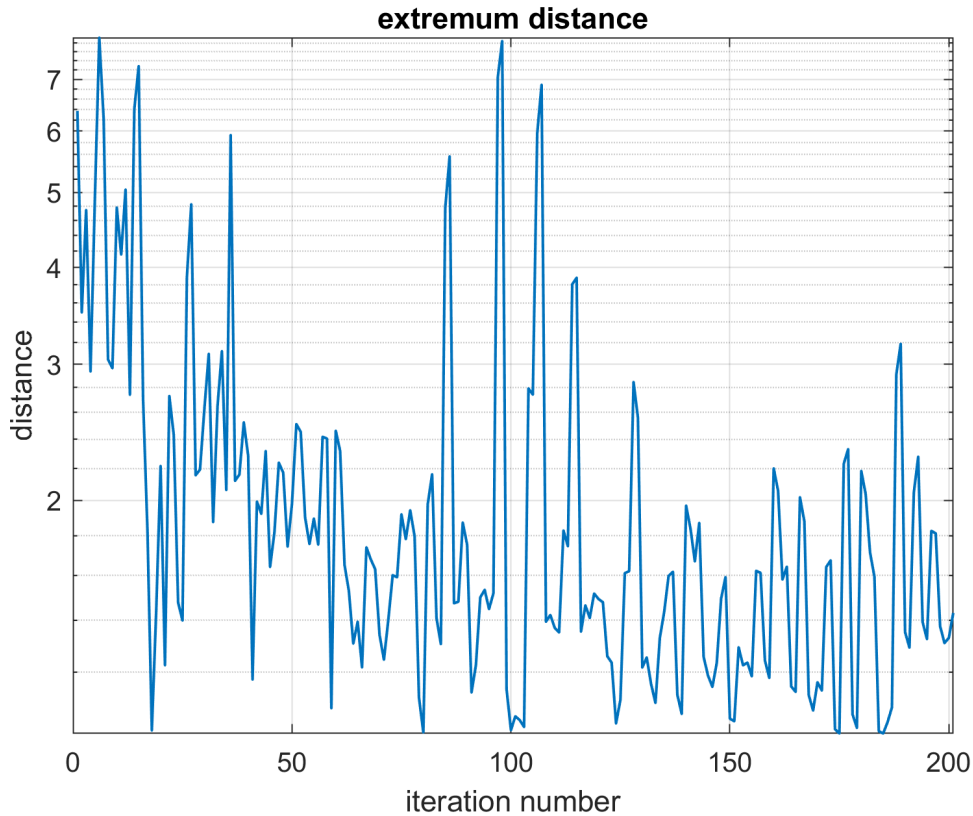


Рис. 3: Расстояние до точки экстремума для функции МакКормика (4)



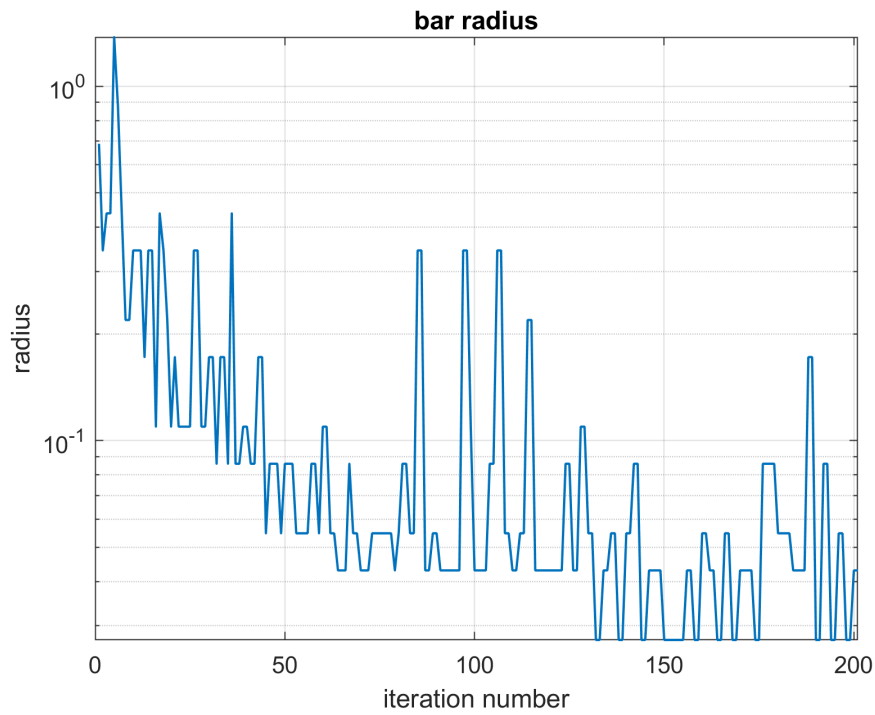


Рис. 4: Радиусы рабочих брусов для функции МакКормика (4)

#### 4.3.2 Функция с несколькими экстремумами

Для функции (5) в качестве начального приближения был выбран брус  $\begin{pmatrix} [-5, 5] \\ [-5, 5] \end{pmatrix}$ .

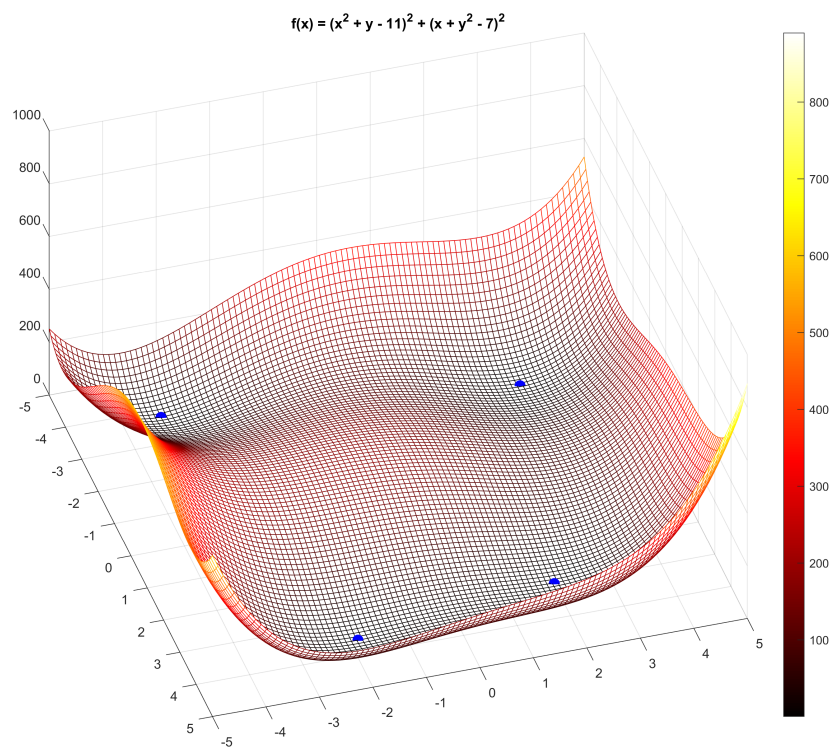


Рис. 5: График функции Химмельблау (5)

Функция содержит 4 глобальных экстремума в точках  $(3, 2)$ ,  $(-2.8051, 3.1313)$ ,  $(-3.7793, -3.2831)$ ,  $(3.5844, -1.8481)$ , отмечены на графике синими точками. Получены следующие результаты:

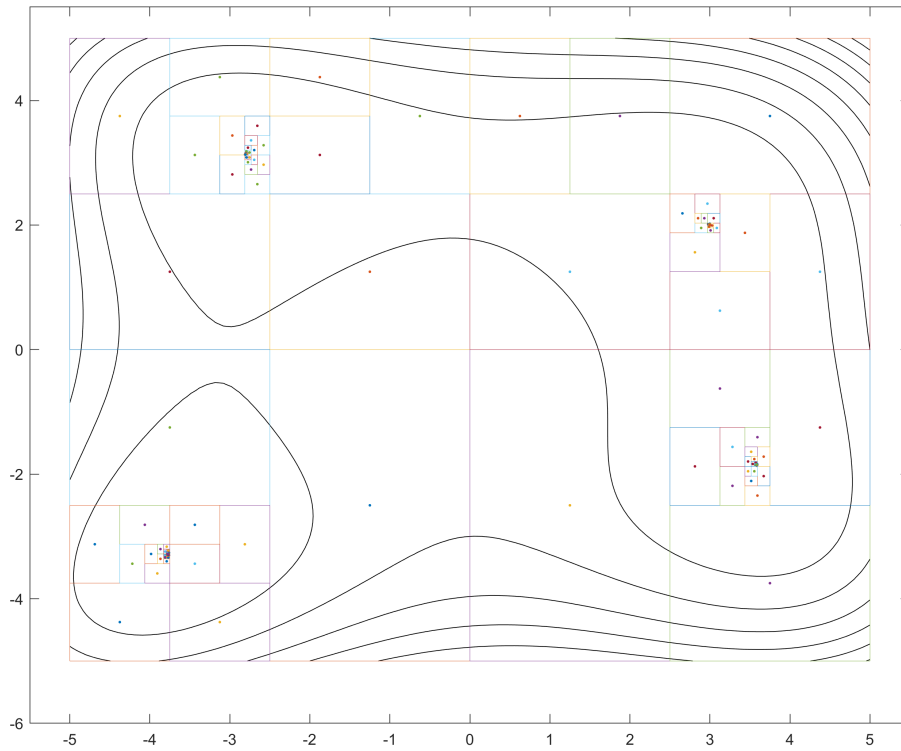


Рис. 6: Иллюстрация работы алгоритма для функции Химмельблау (5)

Возле всех 4 минимумов наблюдается сгущение брусев. В качестве решения получен вырожденный брус  $\mathbf{x} \approx \begin{pmatrix} [3, 3] \\ [2, 2] \end{pmatrix}$ , являющийся точным (считая до  $10^{-5}$ ) расположением одного из экстремумов,  $f(\mathbf{x}) \approx 0$ . График расстояния до экстремума построен для ближайшего на каждой итерации экстремума.

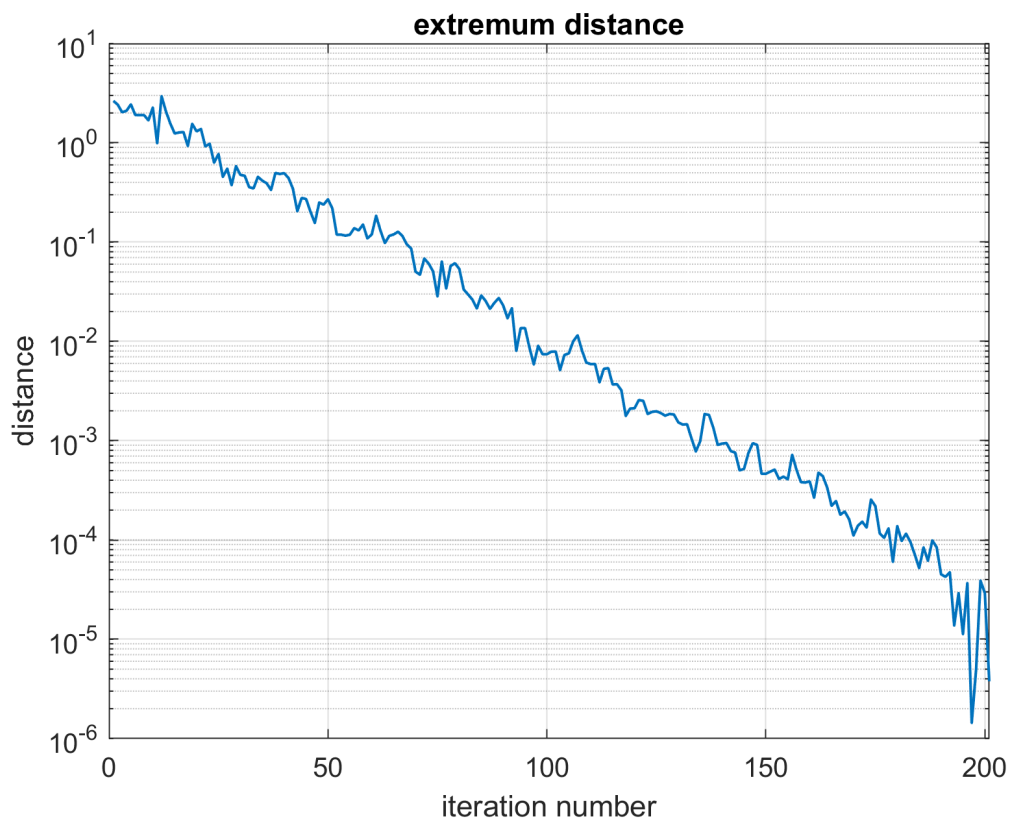


Рис. 7: Расстояние до точки экстремума для функции Химмельблау (5)

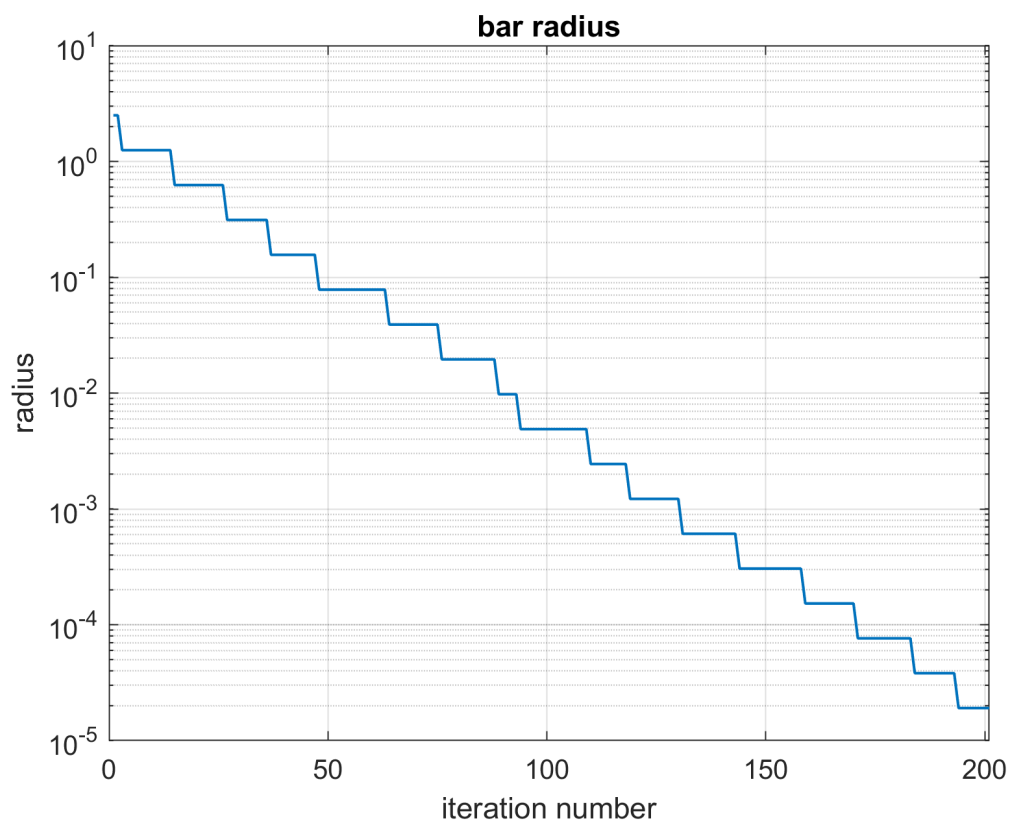


Рис. 8: Радиусы рабочих брусов для функции Химмельблау (5)

## 5 Обсуждение

1. Говоря о переходе от задачи линейной регрессии к полиномиальной, мы подразумеваем рассмотрение квадратных матриц  $\mathbf{X}_n \in \mathbb{IR}^{n \times n}$  и увеличение параметра  $n$ . Если обсуждать задачи общего вида и не налагать ограничений на входящие в эти матрицы величины (сохранив при этом  $\forall i = \overline{1, n}, j = \overline{1, n-1} \text{ rad } \mathbf{x}_{ij} = \varepsilon$ ), то можно предположить, что множество значений  $\varepsilon$ , при которых матрица будет оставаться неособенной, зависит от значения  $\text{mid } \mathbf{X}_n$ . Чем ближе друг к другу будут середины вновь добавляемых в матрицу регрессии величин, тем меньшее значение  $\varepsilon$  потребуется для того, чтобы добиться попарного пересечения интервальных величин, стоящих в одинаковых столбцах в каких-либо двух строках матрицы  $\mathbf{X}_n$ , и тем самым получить вырожденную точечную матрицу в составе рассматриваемой матрицы регрессии. Это рассуждение можно подтвердить, рассмотрев признак Румпа. Чем меньше будут величины, входящие в  $\text{mid } \mathbf{X}_n$ , тем меньше будет норма этой матрицы, следовательно, норма Фробениуса также будет уменьшаться. Это приведет к тому, что  $\sigma_{\min}(\text{mid } \mathbf{X}_n)$  будет приближаться к 0, и даже для малых  $\varepsilon$  признак не будет выполняться. Провести схожий анализ с помощью критерия Баумана намного сложнее. Можно лишь предположить, что при увеличении размерности матрицы множество ее вершин также будет увеличиваться, и для выполнения критерия придется наложить намного больше ограничений на  $\varepsilon$ , следовательно, множество значений этой величины, при которых регрессионная матрица будет неособенной, должно уменьшиться.
2. Переход от задачи (2) к задаче (3) сопряжен с введением дополнительных интервальных величин в матрицу, что резко увеличивает сложность определения ее особенности. Более того, из полученных результатов можно сделать вывод, что матрицу  $\mathbf{A}$  проще сделать особенной, чем  $\mathbf{X}$ , так как допуск на значения  $\varepsilon$ , при которых матрица сохраняет свою неособенность, стал меньше.
3. В задаче (3) результаты, полученные с помощью критерия Баумана, не противоречат результатам, посчитанным с помощью признака Румпа. Более того, в данной задаче по результатам признака Румпа можно дать очень точную оценку значениям  $\varepsilon$ , при которых матрица  $\mathbf{A}$  может стать особенной. Хотя вообще говоря, признак Румпа задает только достаточные условия неособенности, что становится видно после его применения на задаче регрессии.
4. Для функции МакКормика рассматриваемый алгоритм нахождения минимума функции дал правильную оценку самого значения минимума функции:  $-1.9133 \in [-2.3019, -1.3809]$ , но при этом не смог обеспечить сходимости к его точке:  $\begin{pmatrix} -0.5471 \\ -1.5471 \end{pmatrix} \notin \begin{pmatrix} [-0.3829, -0.2968] \\ [-1.5782, -1.4687] \end{pmatrix}$ . Для функции Химмельблау была получена точечная оценка положения одного из экстремумов, причем довольно точная.
5. Опираясь на графики (3) и (4), можно сделать вывод, что на функции (4) метод действительно не обеспечивает сходимость к точке экстремума (во всяком случае за конечное число итераций): радиус брусков имеет более высокий порядок малости нежели расстояние до экстремума. Скачкообразное поведение графиков объясняется самим алгоритмом: ведущий брус, который будет далее дробиться, выбирается на каждой итерации путем полного итерирования по рабочему списку, то есть сам автор алгоритма предусматривал, что последний рассмотренный брус может не быть оптимальным, и что необходимо проводить по сути полный перебор рабочего списка в поисках лучшей оценки значения минимума функции. По этой причине монотонного убывания не наблюдается ни на одном из данных графиков.
6. На функции Химмельблау (5) метод показал монотонное убывание радиусов брусков (8) и сошелся к вырожденному брусу, довольно точно описывающему реальное

положение точки минимума (примерно до  $10^{-5}$ ). График (7) не показывает строгого убывания, что описывается ранее приведенными рассуждениями, однако ведет себя намного лучше, чем аналогичный для функции МакКормика (3). По нему видно, что остальные положения экстремумов тоже находятся довольно точно.

7. Исходя из полученных результатов, можно утверждать, что в общем случае метод пригоден лишь для довольно грубой (причем вряд ли можно провести аналитическую оценку того, насколько грубой) оценки самого значения минимума функции, а не положения точки экстремума, что также подтверждается сигнатурой его функции: на выход выдается лишь нижняя оценка значения минимума и рабочий список брусов, для получения остальных оценок необходимо видоизменять код процедуры или проводить анализ этого списка.
8. Тем не менее, для функции (5) были получены отличные результаты, что дает повод предположить, что плохое поведение метода обусловлено характером функции МакКормика: рассматриваемая область очень пологая и в ней присутствует локальный минимум в окрестности точки (2.5, 1.5), в то время, как у функции Химмельблау глобальные минимумы более ярко выражены.

## Исходный код

С исходным кодом программы и отчета можно ознакомиться в репозитории <https://github.com/ilich/Interval>.