**МИНОБРНАУКИ РОССИИ**

**САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ «ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)**

**Кафедра АМ**

# ОТЧЕТ

**по лабораторной работе №1**

**по дисциплине «Методы оптимизации»**

**ТЕМА: Исследование методов безусловной оптимизации.**


| | |
|---|---|
| Студент гр. 1384 | Бобков В. Д. |
| Студентка гр. 1384 | Пчелинцева К. Р. |
| Студентка гр. 1384 | Усачева Д. В. |
| Преподаватель | Балтрашевич В. Э. |


Санкт-Петербург

2023 г.

## Цель работы.

Провести исследование методов безусловной минимизации, сравнивая процессы минимизации и результаты испытаний методов.

## Выполнение работы.

1) Градиентный метод с постоянным шагом

Итерационный метод с постоянным шагом.

Формула: $x_{k+1} = x_k - t \cdot \nabla f(x_k)$

Для начала запустим программу со следующими параметрами: a = 1.0, x0 = 3, 8, t = 0.2. Результат представлен на рисунке 1.

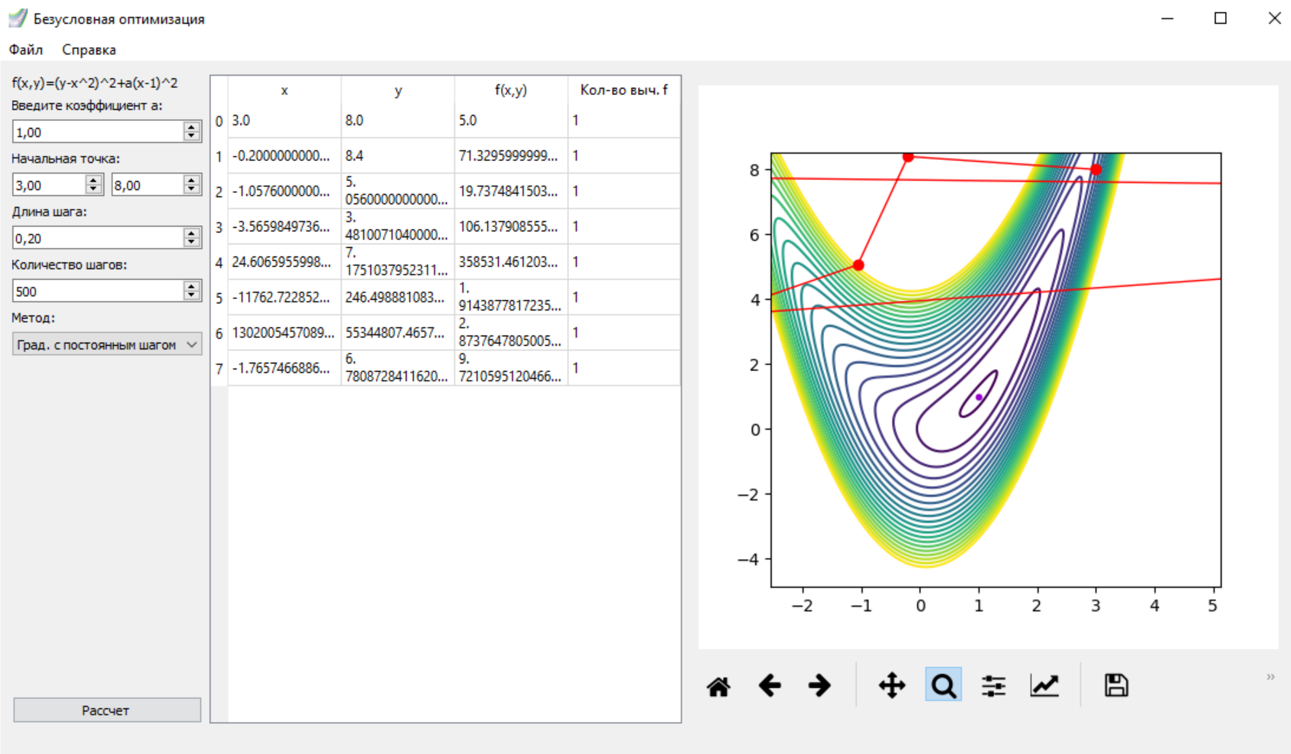


*Рис.1 - Параметры (1.0 / (3, 8) / 0.2)*

При данных параметрах метод разошёлся, поэтому подберём оптимальные для минимизации целевой функции. Для следующего испытания шаг уменьшен в 10 раз.

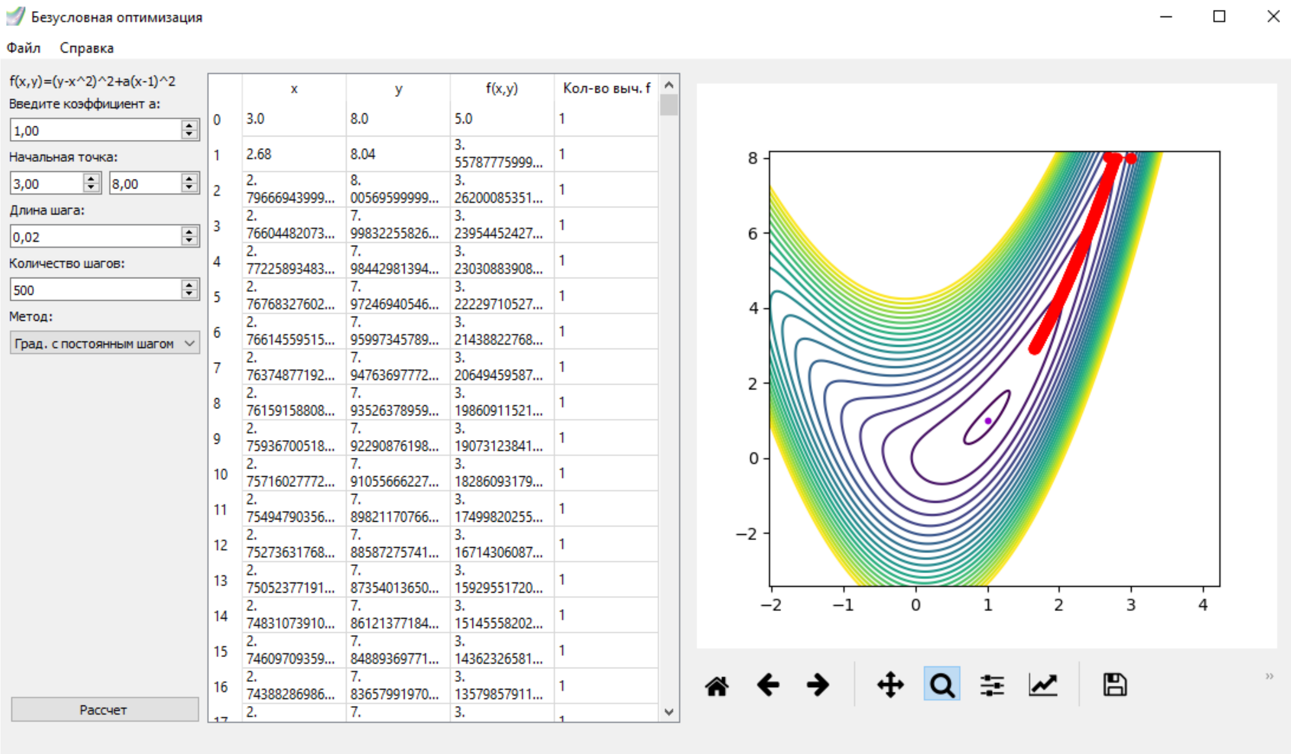


*Рис. 2 - Параметры (1 / (3, 8) / 0.02)*

Метод чувствителен к длине шага, при неправильном выборе – может легко разойтись.

Преимущества: начальная точка может располагаться достаточно далеко от x*; целевая функция должна принадлежать классу $C^1$; вычисления относительно простые. Однако метод имеет медленную скорость сходимости. При втором испытании было сделано 500 шагов, однако метод сошелся не до конца.

2) Градиентный метод с дроблением шага

Итерационный метод с изменяющимся по условию шагом

Условия выбора шага: $f\left(x_{k+1}\right) - f\left(x_k\right) \leq -\varepsilon t_k \left\|\nabla f\left(x_k\right)\right\|$, где $0 < \varepsilon < 1$

(если условие не выполняется, длину шага делим пополам)

Итерационная формула: $x_{k+1} = x_k - t_k \cdot \nabla f(x_k)$

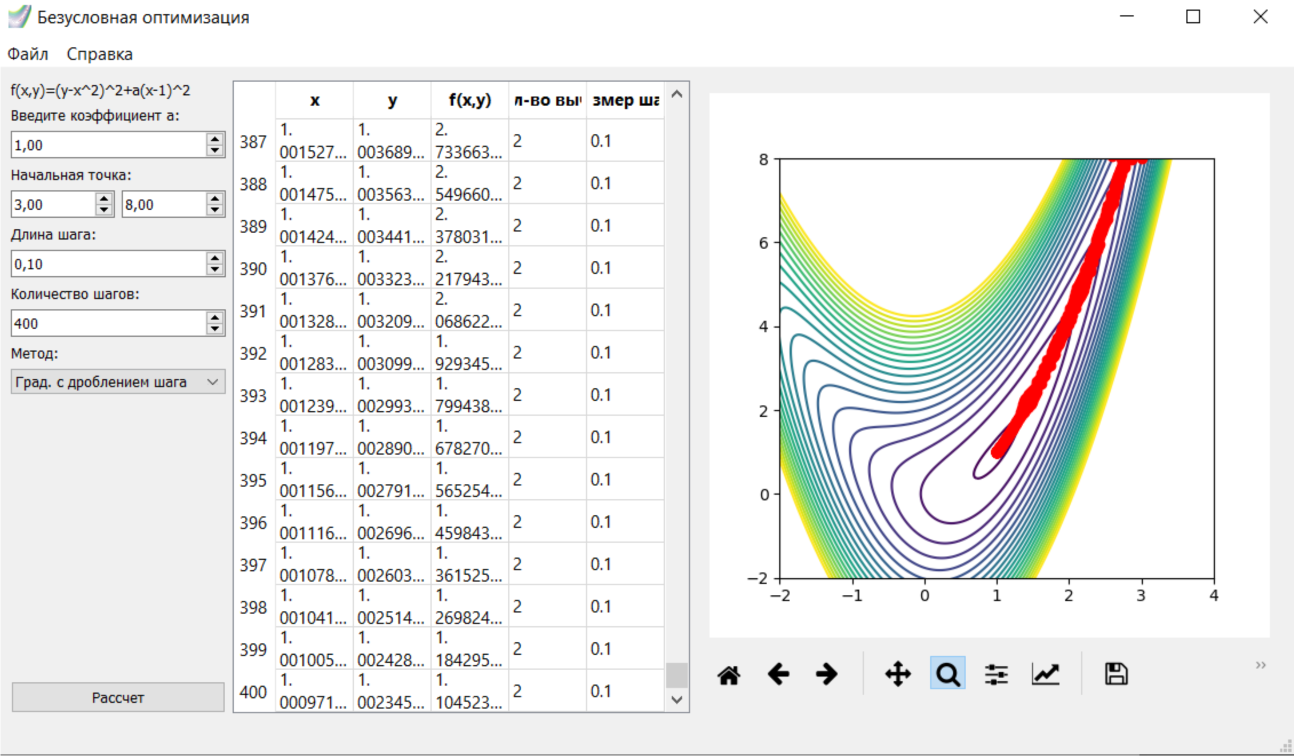


*Рис. 3 - Параметры (1 / (3, 8) / 0,1)*

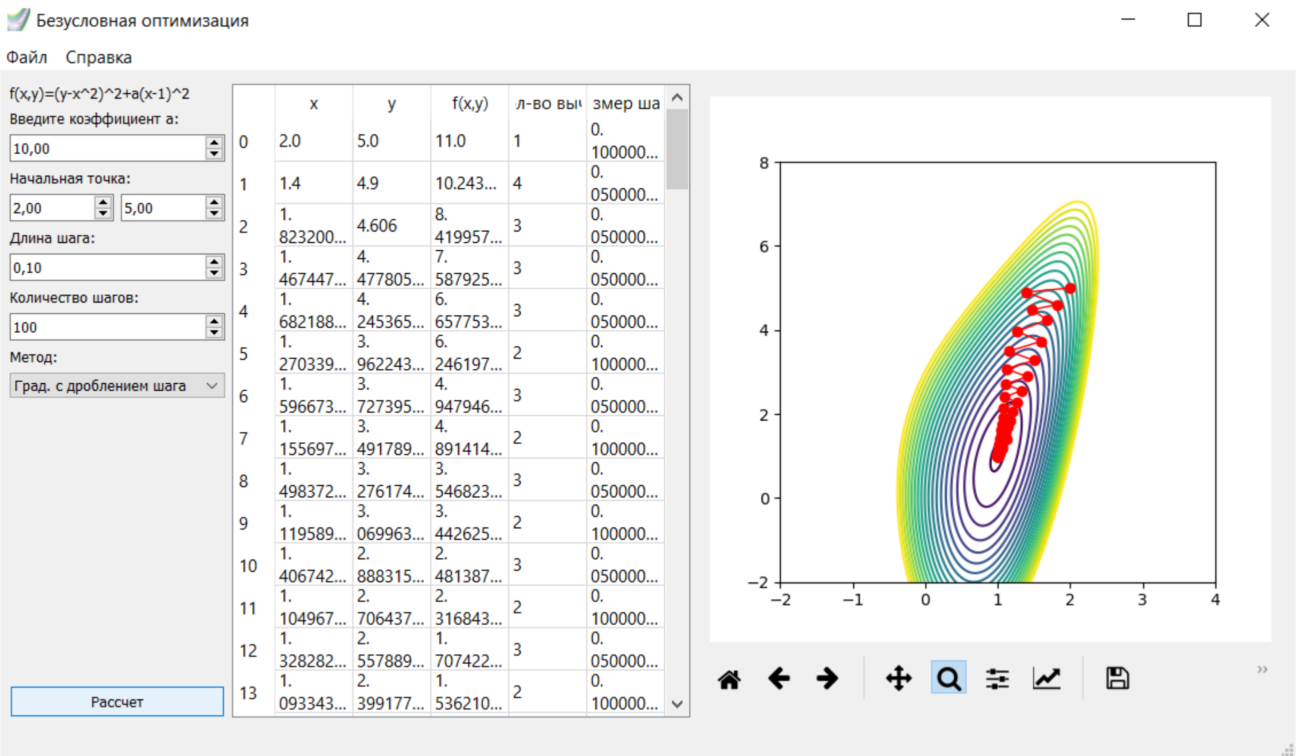


*Рис. 4 - Параметры (10 / (2, 5) / 0,1)*

Данный метод похож на предыдущий, но теперь длина шага регулируется условием, что влияет на сходимость метода. Из-за такого подхода сокращение шага приводит к большему количеству шагов.

3) Градиентный метод с убыванием шага как 1/k.

Итерационный метод, шаг которого убывает с каждой новой  итерацией по закону 1/k.

Итерационная формула: $x_{k+1} = x_k - t_k \cdot \nabla f\left(x_k\right)$, где $t_k = \frac{1}{k}$

```
параметр a   1.000000  - 1
длину шага   0.000100  - 2
координаты начальной точки   2.995358   8.000579  - 3
количество шагов 10  - 4
интервал для печати 1  - 5
номер анализируемого алгоритма:  3  - 6
  1 - градиентный с постоянным шагом
  2 - градиентный с дроблением шага
  3 - градиентный с убыванием шага как 1/k
  4 - наискорейшего спуска
  5 - Ньютона-Рафсона
  6 - овражный алгоритм
  7 - алгоритм Полака-Ривьера
  8 - алгоритм Давидона-Флетчера-Пауэлла
  9 - алгоритм Бройдена-Флетчера-Шанно
изменений не требуется - 0
иначе - конец работы программы
```

```
0
ном.шага    x1          x2          f(x1,x2)     число выч f на 1 шаг
   1     2.993795    8.000773     4.9007284235       1
   2     2.993019    8.000869     4.8885424339       1
   3     2.992505    8.000933     4.8804773950       1
   4     2.992119    8.000981     4.8744585415       1
   5     2.991812    8.001019     4.8696615611       1
   6     2.991556    8.001051     4.8656762040       1
   7     2.991337    8.001078     4.8622688716       1
   8     2.991146    8.001102     4.8592939845       1
   9     2.990976    8.001123     4.8566547247       1
  10     2.990823    8.001141     4.8542834616       1
```

Рис. 5 - Параметры (1 / (3, 8) / 0,0001)

4) Метод наискорейшего спуска

Итерационный метод с постоянным пересчётом оптимального шага.

Итерационная формула: $x_{k+1} = x_k - t_k \cdot \nabla f(x_k)$

Формула выбора длины шага: $t_k = argmin_{t>0} f(x_k - t_k \cdot \nabla f\left(x_k\right))$

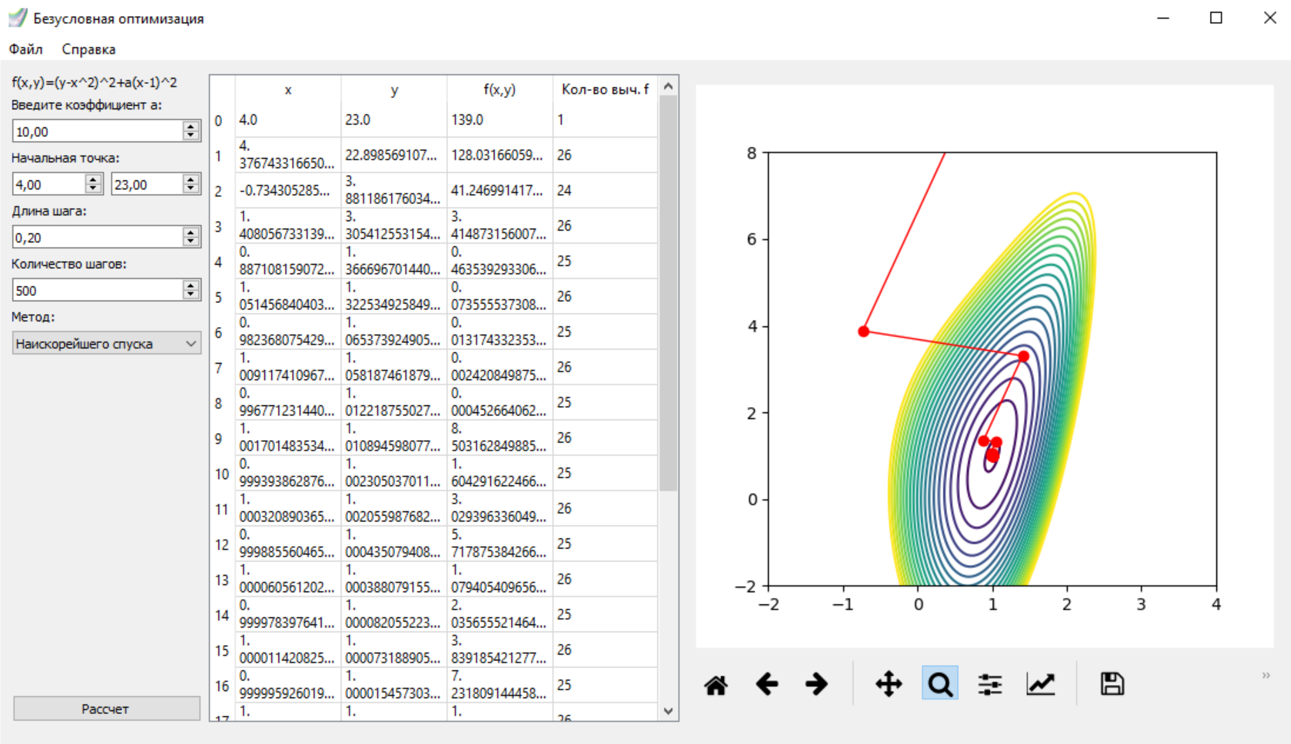


Рис. 6 - Параметры (10 / (4, 23) / 0,2)

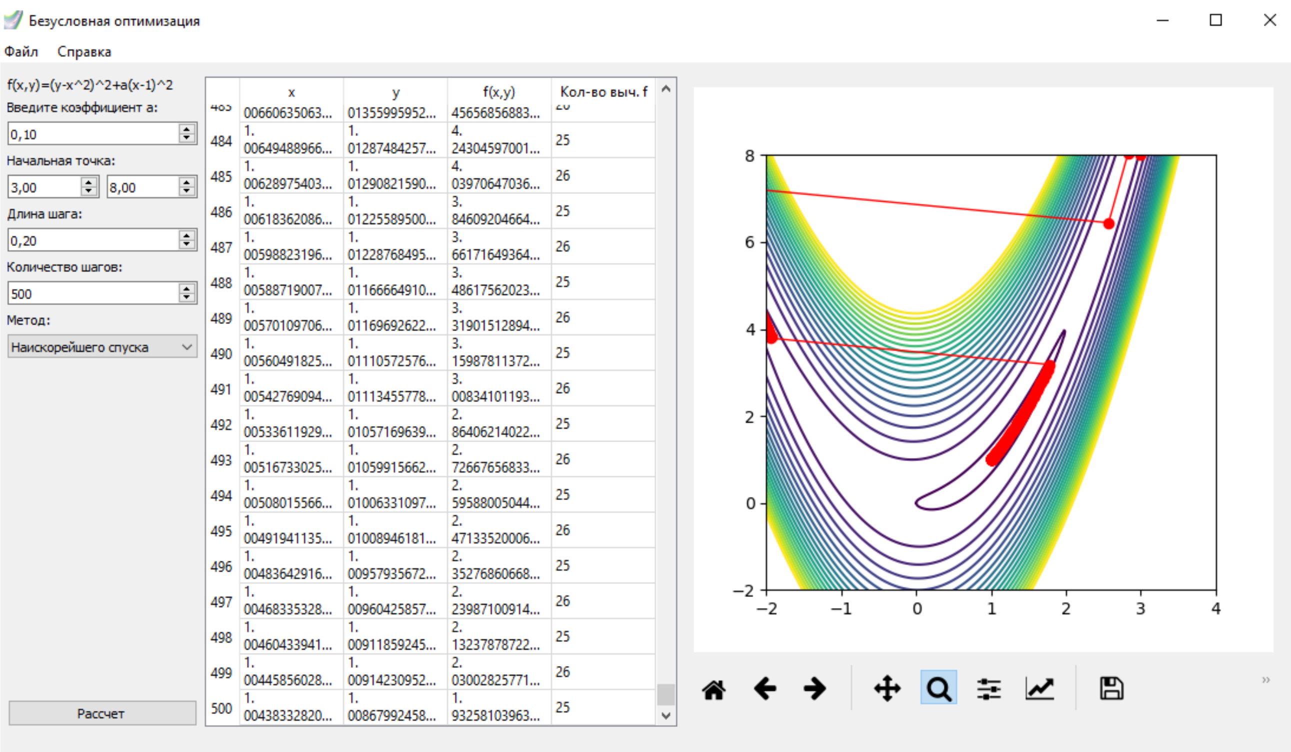


Рис. 7 - Параметры (0.1 / (3, 8) / 0,2)

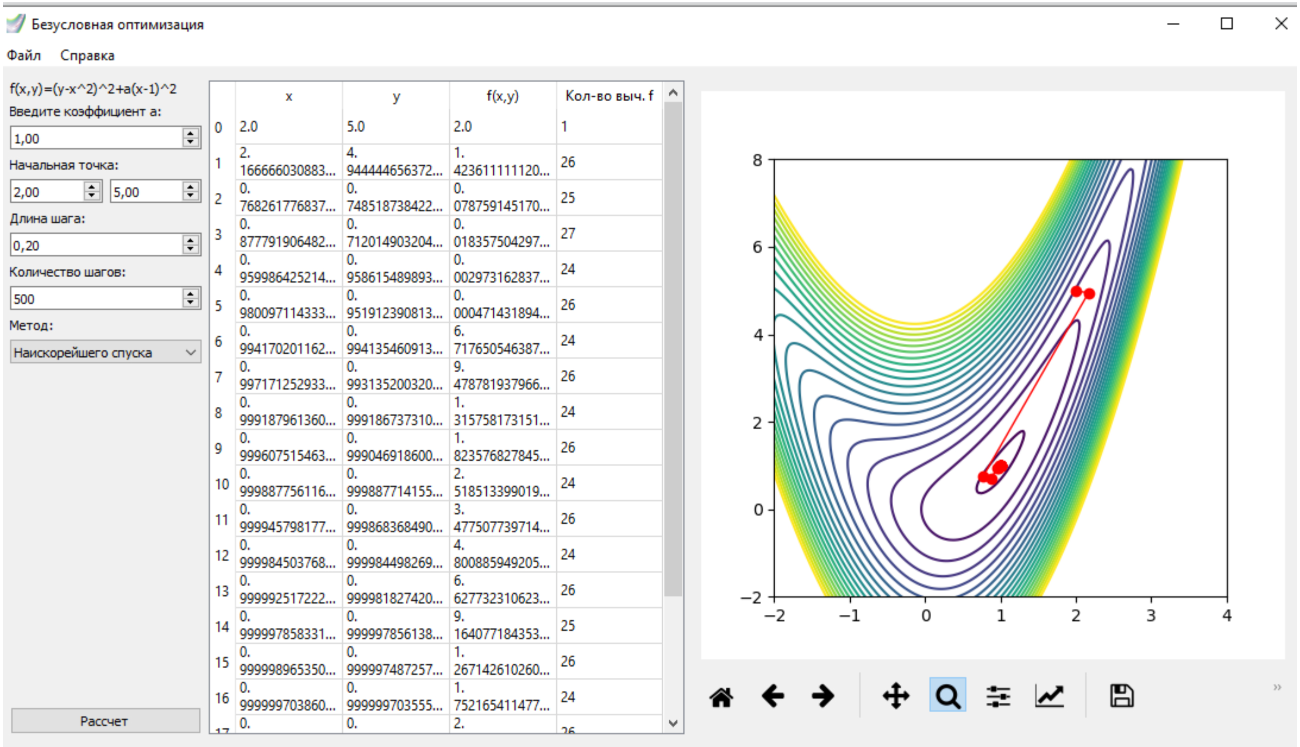


Рис. 8 - Параметры (1 / (2, 5) / 0,2)

Характерной чертой метода является постоянная ортогональность “соседних” градиентов. По сравнению с градиентным спуском с фиксированным шагом, метод сходится быстрее, однако может задерживаться в «оврагах», являющиеся локальными минимумами (рис. 6 и 8)

5) Метод Ньютона-Рафсона

Итерационный метод с вычислением обратной матрицы Гессе. Метод является модифицированной версией метода Ньютона.

Итерационная формула: $x_{k+1} = x_k - \alpha_k \left[H\left(x_k\right)\right]^{-1} \cdot \nabla f\left(x_k\right)$, где $\alpha_k > 0$

Решение одномерной минимизации:

$$\alpha_k = min_{\alpha>0} g_k(\alpha), \text{ где } g_k(\alpha) = f(x_k - \alpha\left[H\left(x_k\right)\right]^{-1} \cdot \nabla f\left(x_k\right))$$

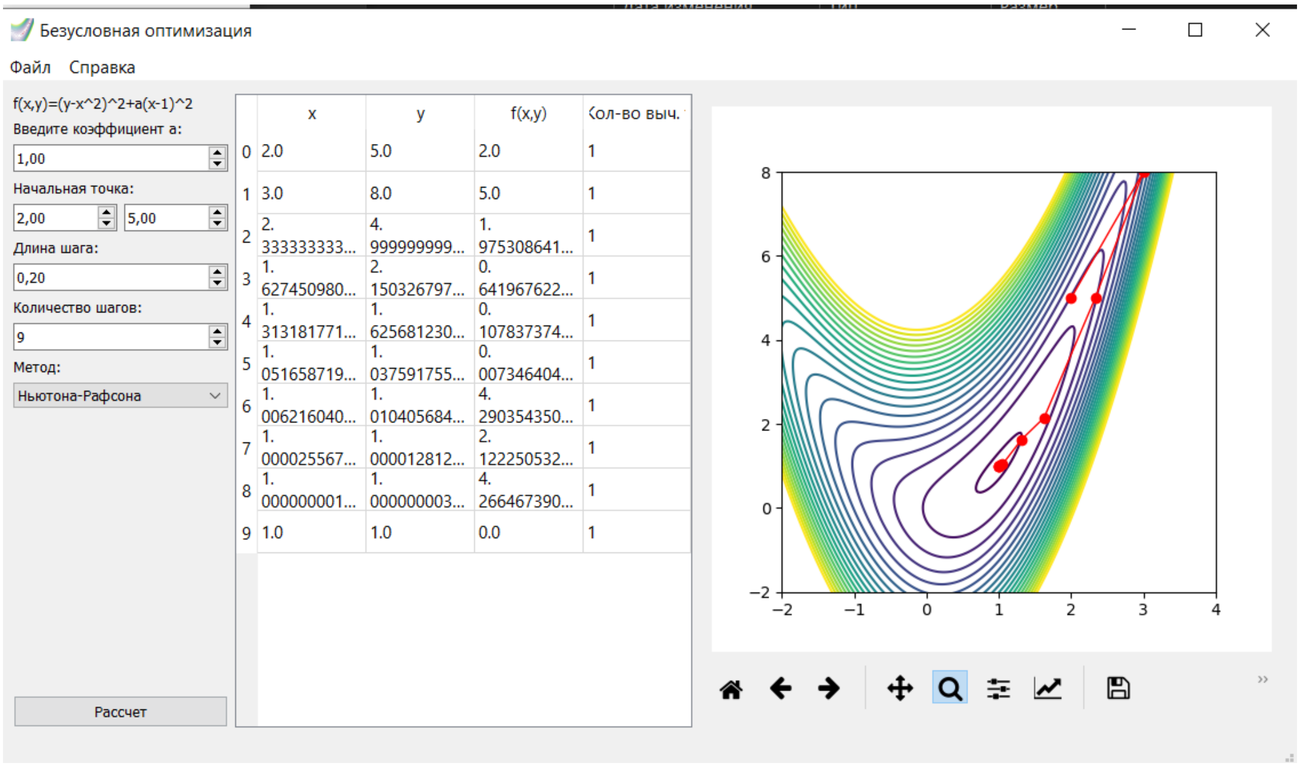


Рис. 9 - Параметры (1 / (2, 5) / 0,2)

Данный метод примечателен тем, что имеет квадратичную скорость сходимости, которая отражается в значительно меньшем количеством шагов, по сравнению с предыдущими методами.

К минусам данного метода можно отнести то, что он вычислительно затратен; имеет только локальную сходимость (хотя иногда получается минимизировать целевую функцию и из точки достаточно удаленной от глобального минимума – рис. 10); целевая функция должна быть дважды непрерывно дифференцируемой.

Модификацией метода Ньютона является возможность устранения несходимости последовательности $\{x_k\}$ к точке x*.

6) Алгоритм Полака-Рибьера

Двухшаговый итерационный метод с оптимальным шагом.

Итерационная схема метода имеет вид:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k, \text{ где } \alpha_k = \arg\min f(\mathbf{x}_k + \alpha \mathbf{p}_k), \ \alpha > 0;$$

$$\mathbf{p}_k = -\nabla f(\mathbf{x}_k) + \beta_k \mathbf{p}_{k-1}; \ \beta_k = \frac{((\nabla f(\mathbf{x}_k) - \nabla f(\mathbf{x}_{k-1})), \nabla f(\mathbf{x}_k))}{\left\|\nabla f(\mathbf{x}_{k-1})\right\|^2}, \quad \beta_0 = 0.$$

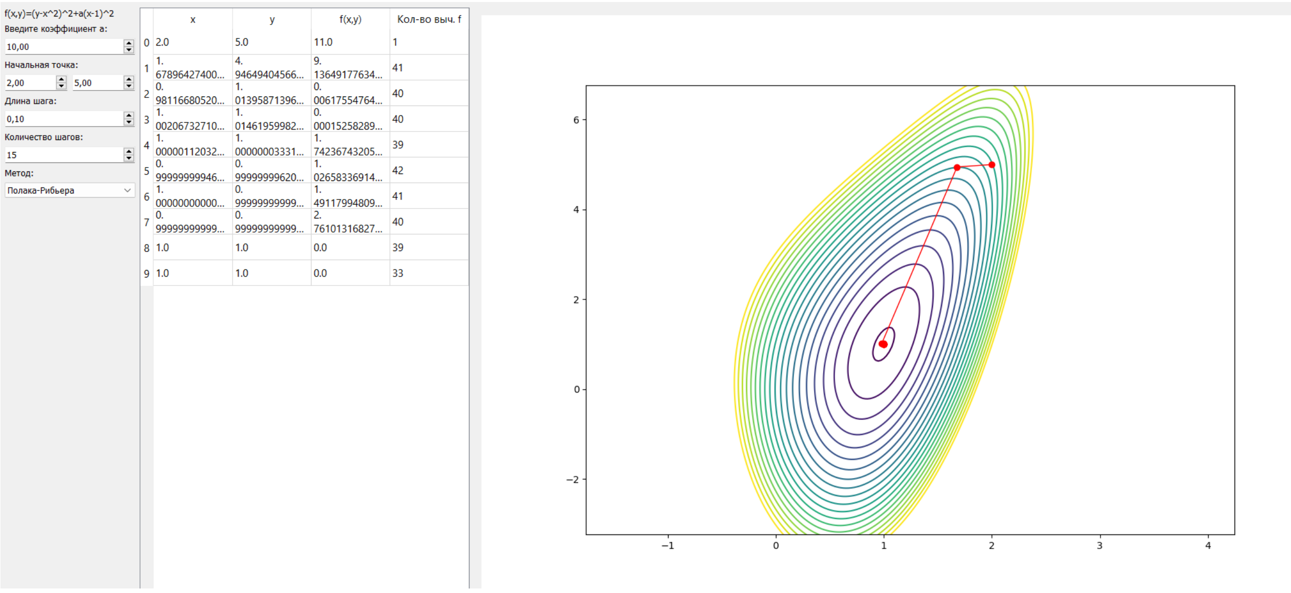


Рис. 11 - Параметры (10 / (2, 5) / 0,1)

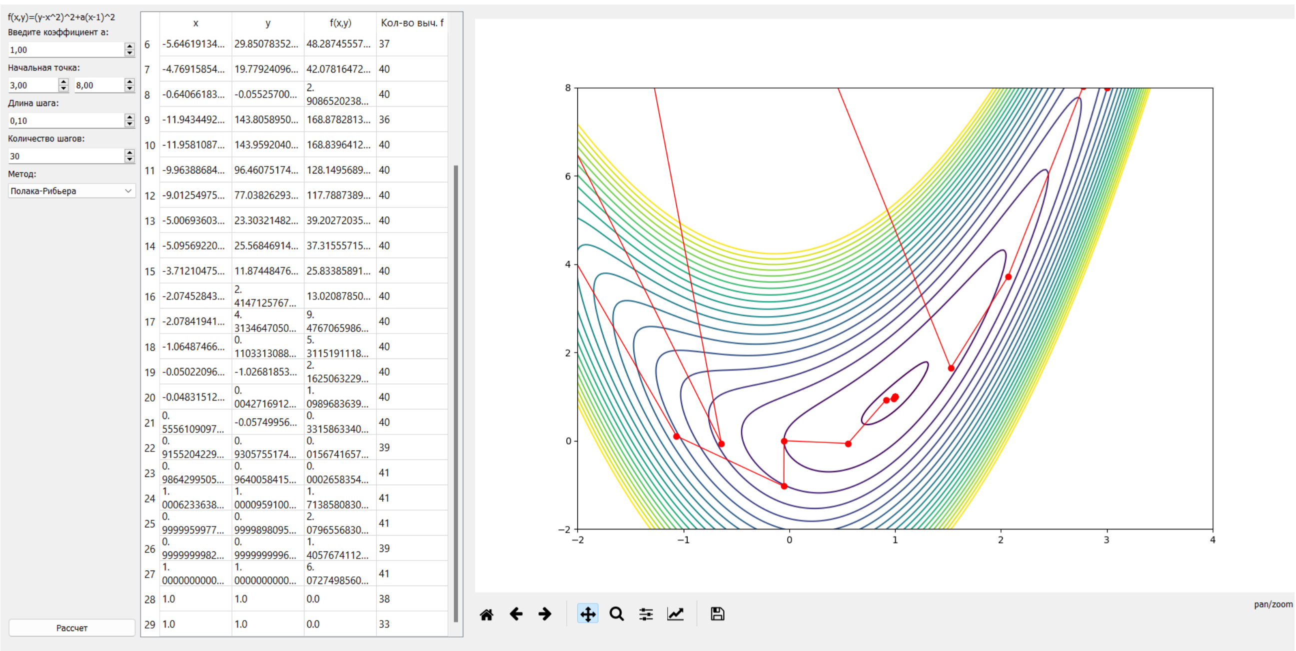


Рис 12 - Параметры (1 / (3, 8) / 0,1)

Как мы можем заметить на рис. 11 и 12, скорость сходимости метода сопряженных градиентов выше, чем для метода наискорейшего спуска, но ниже, чем для метода Ньютона.

7) Алгоритм Давидона-Флетчера-Пауэлла

Квазиньютоновская итерационная схема с перевычислением матриц.

Итерационная схема: $x_{k+1} = x_k - \gamma_k H_k \nabla f(x_k)$

Формулы для перевычисления матриц Гессе:

$$\mathbf{H}_{k+1} = \mathbf{H}_k - \frac{\mathbf{H}_k \mathbf{y}_k (\mathbf{H}_k \mathbf{y}_k)^{\mathrm{T}}}{(\mathbf{H}_k \mathbf{y}_k, \mathbf{y}_k)} + \gamma_k \frac{\mathbf{p}_k \mathbf{p}_k^{\mathrm{T}}}{(\mathbf{p}_k, \mathbf{y}_k)}; \quad \mathbf{H}_0 > 0$$

$$p_k = - H_k \cdot \nabla f(x_k), \;\; y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$$

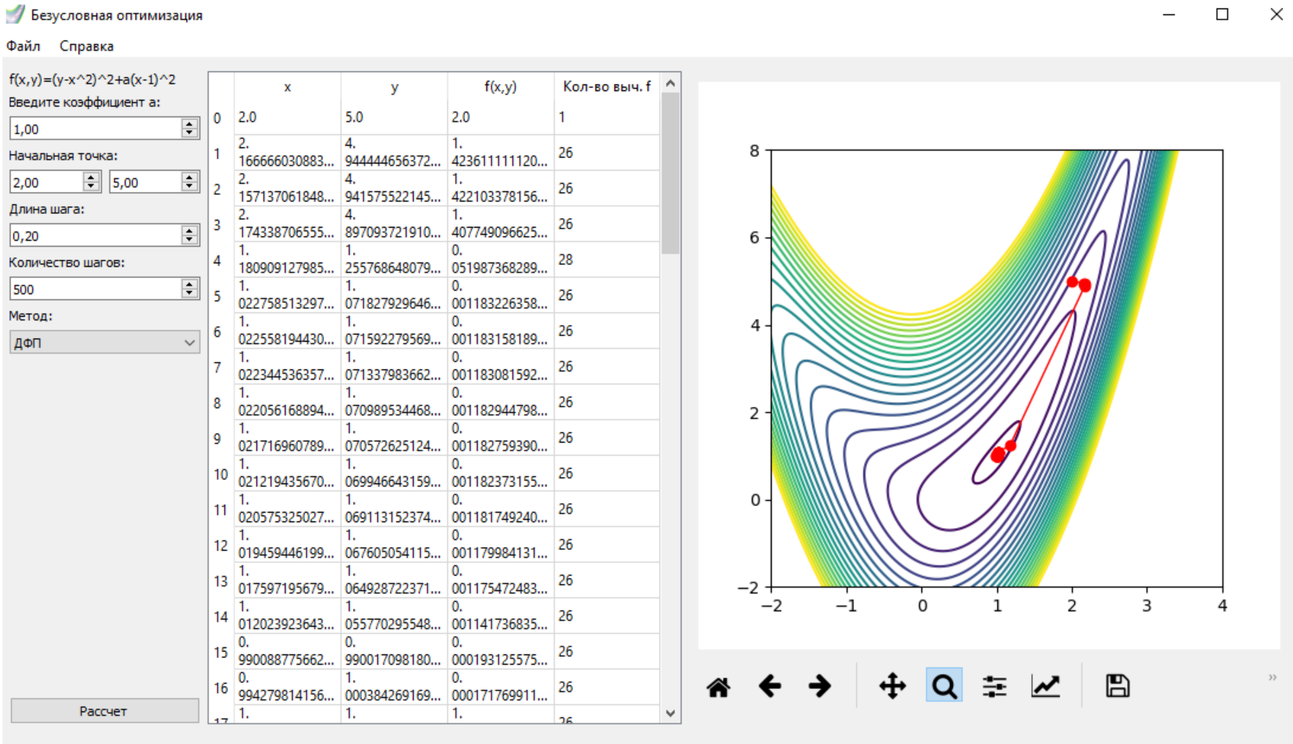


Рис 13 - Параметры (1 / (2, 5) / 0,2)

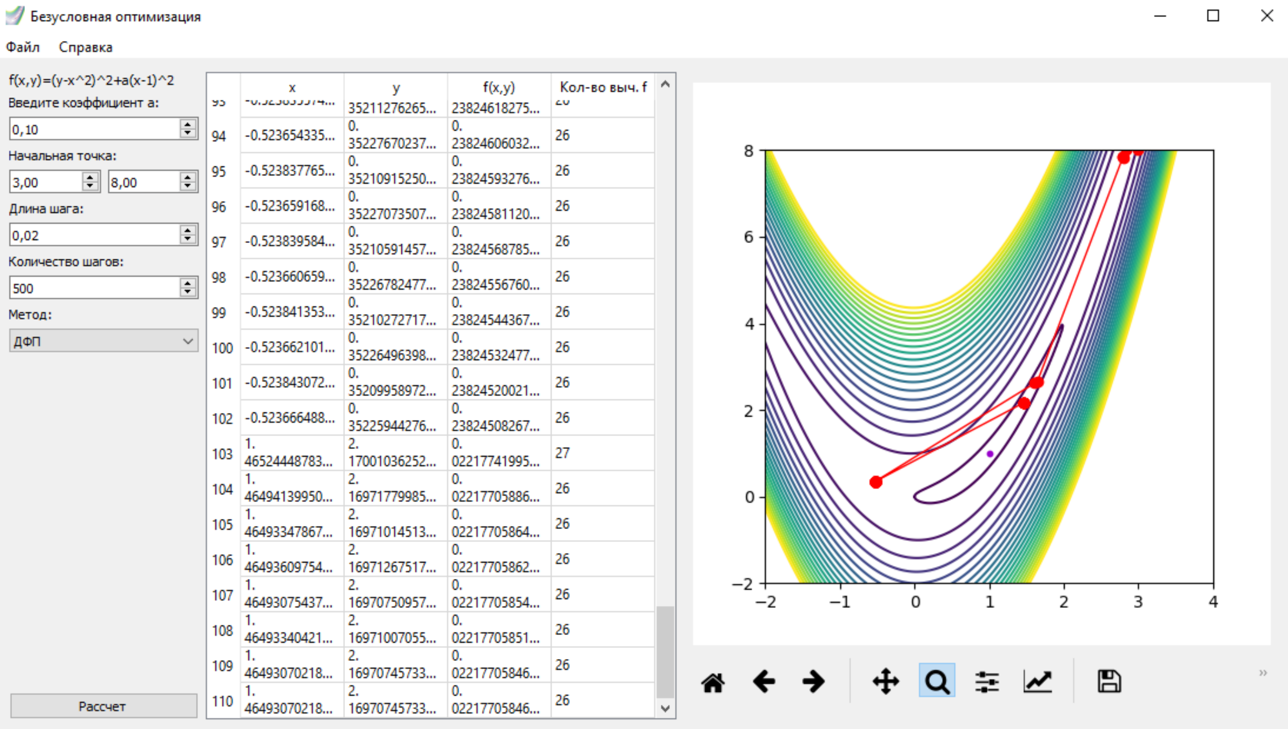


Рис 14 - Параметры (0.1 / (3, 8) / 0,02)

Преимущество квазиньютонных методов заключается в аппроксимации обратной матрицы Гесса и пересчитывается рекуррентным способом на основе информации, полученной на k-й итерации.

Данный метод имеет сверхлинейную (быстрее геометрической прогрессии) сходимость (т.е. работает быстрее градиентного спуска, но медленнее метода Ньютона). Метод имеет глобальную сходимость. Получается, он берет от градиентных методов глобальную сходимость, а от метода Ньютона – увеличенную скорость сходимости.

8) Алгоритм Бройдена-Флетчера-Шенно

Квазиньютоновская итерационная схема с перевычислением матриц.

Метод перевычисления матриц $\mathbf{H}_k$ имеет вид

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \frac{\rho_k \mathbf{p}_k (\mathbf{p}_k)^{\mathrm{T}} - \mathbf{p}_k (\mathbf{y}_k)^{\mathrm{T}} \mathbf{H}_k - \mathbf{H}_k \mathbf{y}_k (\mathbf{p}_k)^{\mathrm{T}}}{(\mathbf{y}_k, \mathbf{p}_k)},$$

где $\rho_k = \gamma_k + \dfrac{(\mathbf{H}_k \mathbf{y}_k, \mathbf{y}_k)}{(\mathbf{y}_k, \mathbf{p}_k)}$, $\mathbf{H}_0 > 0.$

| ном.шага | x1 | x2 | f(x1,x2) | число выч f на 1 шаг |
|---|---|---|---|---|
| 1 | 2.166692 | 4.944436 | 1.4236111259 | 11 |
| 2 | 1.344567 | 1.565382 | 0.1775216773 | 23 |
| 3 | 1.228406 | 1.593644 | 0.0593370741 | 11 |
| 4 | 1.042945 | 1.039985 | 0.0041244086 | 21 |
| 5 | 1.019585 | 1.047810 | 0.0004517415 | 14 |
| 6 | 1.000436 | 0.999993 | 0.0000009635 | 22 |
| 7 | 1.000060 | 1.000144 | 0.0000000042 | 12 |
| 8 | 1.000000 | 1.000000 | 0.0000000000 | 22 |
| 9 | 1.000000 | 1.000000 | 0.0000000000 | 22 |
| 10 | 1.000000 | 1.000000 | 0.0000000000 | 22 |
| 11 | 1.000000 | 1.000000 | 0.0000000000 | 22 |
| 12 | 1.000000 | 1.000000 | 0.0000000000 | 22 |
| 13 | 1.000000 | 1.000000 | 0.0000000000 | 22 |
| 14 | 1.000000 | 1.000000 | 0.0000000000 | 22 |
| 15 | 1.000000 | 1.000000 | 0.0000000000 | 22 |

всего вычислено 136 значений функции f

Рис. 15 - Параметры (1 / (2, 5) / 0,1)

**Выводы:**

В ходе выполнения лабораторной работы были рассмотрены и протестированы на практике несколько типов безусловных методов оптимизации: градиентные, многошаговые (двухшаговые), квазиньютонные.

Метод наискорейшего спуска сходится быстрее градиентных методов (1-3), но скорость его сходимости сильно варьируется в зависимости от исходных данных, и он требует больших вычислений (т.к. предполагает решение на каждом шаге задачи минимизации - вычисление оптимального шага). Часто этот метод требует меньшего числа операций, чем градиентный метод с постоянным шагом.

Градиентные методы (1) и (3) критичны в отношении выбора исходных данных (длины шага $t_k$). Но зато для этих методов характерны очень малые вычислительные затраты на каждом шаге.

Модифицированный метод Ньютона-Рафсона (5) выгодно использовать, когда начальная точка приближения достаточна близкка к минимуму, т.к. сходимость квадратичная. Метод требует большого числа вычислений для расчёта обратной матрицы Гессе.

Двухшаговый метод (6) используется при плохой обусловленности начальных данных. В таком случае его выгоднее использовать, чем методы градиентного спуска, т.к. он требует меньшего количества шагов (при большем количестве вычислений). Сходимость глобальная.

Остальные – квазиньютонные методы (7-8) являются чем-то средним между градиентными методами и методом Ньютона (и его модификаций): имеют среднюю скорость сходимости, а также сходятся глобально.