НИУ ИТМО
Факультет Информационных Технологий и Программирования
Направление "Прикладная Математика и Информатика"

Лабораторная работа 4

курса “Методы Оптимизации”

Выполнили Булкина Милена, Раков Николай

Санкт-Петербург, 2021

1 Постановка задания

Реализовать алгоритмы:

1. Метод Ньютона: а) классический, б) с одномерным поиском, в) с направлением спуска.

   (a) Продемонстрируйте работу методов на 2-3 функциях, в том числе на не квадратичных.

   - Для поиска ньютоновского направления спуска необходимо использовать прямой или итерационный метод решения СЛАУ (даже, если она размерности 2).
   - Результаты иллюстрируйте траекториями спуска.
   - Укажите количество итераций, необходимых для достижения заданной точности.
   - В случае одномерного поиска указывайте найденные значения параметра.
   - Проведите исследование влияние выбора начального приближения на результат (не менее трех).

   (b) Исследуйте работу методов на двух функциях с заданным начальным приближением:

   $$f(x) = x_1^2 + x_2^2 - 1.2x_1x_2,\, x^0 = (\mathbf{4}, \mathbf{1})^\intercal$$

   $$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2,\ x^0 = (\mathbf{1.2}, \mathbf{1})^\intercal$$

   (c) Для поиска ньютоновского направления спуска необходимо использовать прямой или итерационный метод решения СЛАУ (даже, если она размерности 2).

   (d) Сравните результаты с минимизацией методом наискорейшего спуска (из лаб. работы 2).

   (e) Постройте таблицу или график зависимости «метод : количество итераций».

2. Метод Бройдена-Флетчера-Шено и метод Пауэлла. Работу квазиньютоновских методов сравните с наилучшим методом Ньютона (по результатам 1.2) на функциях:

   $$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

   $$f(x) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$$

$$f(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$$

$$f(x) = 100 - \frac{2}{1 + \left(\frac{x_1 - 1}{2}\right)^2 + \left(\frac{x_2 - 1}{3}\right)^2} - \frac{1}{1 + \left(\frac{x_1 - 2}{2}\right)^2 + \left(\frac{x_2 - 1}{3}\right)^2}$$

- Для каждого метода приведите иллюстрации траекторий сходимости.
- Проведите исследование влияние выбора начального приближения на результат (не менее трех), оцените скорость сходимости.
- Постройте таблицу или график зависимости «метод : количество итераций».

3. Реализовать метод Марквардта двумя вариантами. Результаты работы продемонстрировать на минимизации многомерной функции Розенброка (n = 100) в сравнении с наилучшим методом Ньютона (по результатам 1.2):

$$f(x) = \sum_{i=1}^{n-1} 100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2$$

Для обоих методов построить график зависимости «итерация – параметр τ». В случае с разложением Холесского дополнительно построить график зависимости «итерация – число разложений Холесского».

2 Ньютон

2.1 Классический Ньютон

Разложим дважды дифференцируемую целевую функцию f (x) в ряд Тейлора в фиксированной точке x при произвольном приращении аргумента Δx , ограничиваясь слагаемыми второго порядка малости:

$$f(x) = f(x^{k-1}) + \nabla f(x^{k-1}, x - x^{k-1}) + \frac{1}{2}(H(x^{k-1})(x - x^{k-1}), x - x^{k-1}) + o(|x - x^{k-1}|)^2$$

Введем φ_k:

$$\varphi_k(x) = f(x^{k-1}) + \nabla f(x^{k-1}, x - x^{k-1}) + \frac{1}{2}(H(x^{k-1})(x - x^{k-1}), x - x^{k-1})$$

Условие минимума:

$$\nabla\varphi_k = \nabla f(x^{k-1}) + H(x^{k-1})(x - x^{k-1}) = 0$$

Получаем СЛАУ:

$$\nabla f(x^{k-1}) = -H(x^{k-1})(x - x^{k-1})$$

Решая систему, найдем вектор перемещения:

$$(x - x^{k-1}) = -H(x^{k-1})^{-1}\nabla f(x_{k-1})$$

<u>Алгоритм:</u>

1. Вычислить $\nabla f(x)$ и $H = \nabla^2 f(x)$
2. Решить СЛАУ: $Hp^k = -\nabla f(k)$
3. $x^k = x^{k-1} + p^k$
4. Если $\|p^k\| < \varepsilon$ остановиться, иначе перейти к пункту 1.

2.2 Метод Ньютона с одномерным поиском

Для повышения надежности метода Ньютона и обеспечения убывания значений целевой функции применяется одномерная минимизация функции.

<u>Алгоритм:</u>

1. Вычислить $\nabla f(x)$ и $H = \nabla^2 f(x)$
2. Решить СЛАУ: $Hp^k = -\nabla f(k)$
3. $\alpha_k = \min\limits_{\alpha} f(x^{k-1} + \alpha p^k)$
4. $x^k = x^{k-1} + \alpha_k p^k$
5. Если $\|p^k\| < \varepsilon$ остановиться, иначе перейти к пункту 1.

Метод Ньютона с одномерным поиском надежнее исходного ме- тода Ньютона. Однако его эффективность существенно зависит от то- го, является ли направление поиска направлением спуска.

2.3 Метод Ньютона с направлением спуска

Если p^k - направление спуска, то $\mathbf{p^k}^\intercal \nabla f(x) < 0$. Если $\mathbf{p^k}^\intercal \nabla f(x) > 0$, то p^k - не является направлением спуска, в этом случае нужно использовать антиградиент.

Получаем:

$$p^k = \begin{cases} p^k, (p^k)^\intercal \nabla f(x) < 0, \\ -\nabla f(x^k), (p^k)^\intercal \nabla f(x) > 0 \end{cases}$$

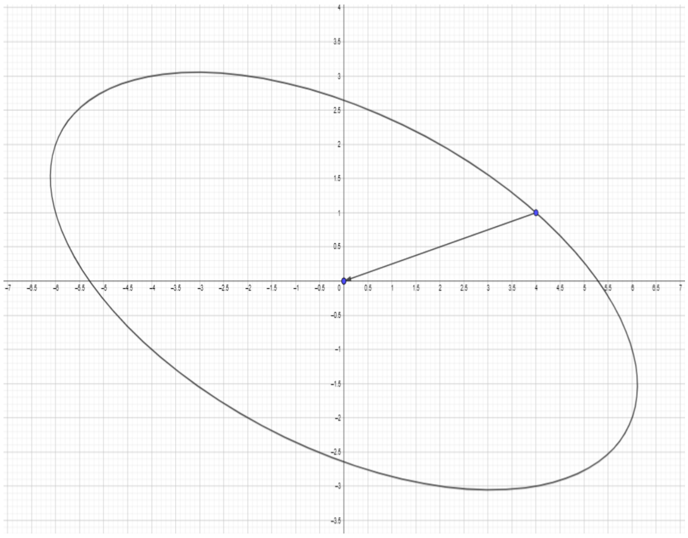
3 Работа Ньютоновских методов

В каждом из них $\varepsilon = 10^{-7}$

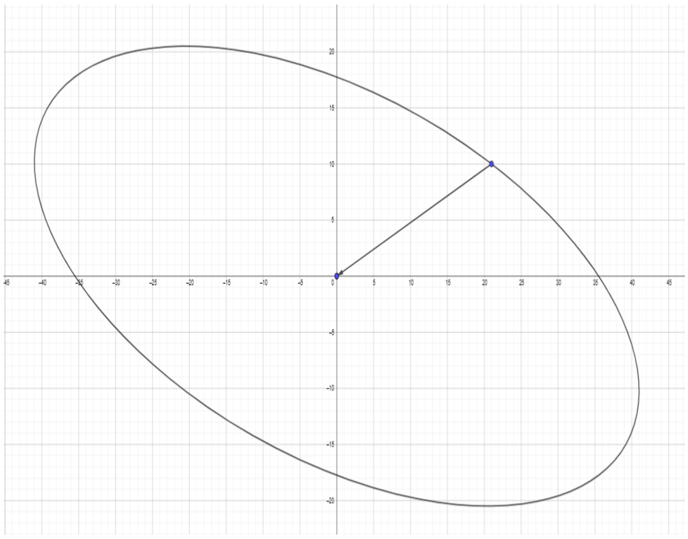
- $f_1(x) = x^2 + 4y^2 + 2xy$

  Искомая точка - (0, 0)

Классический Ньютон

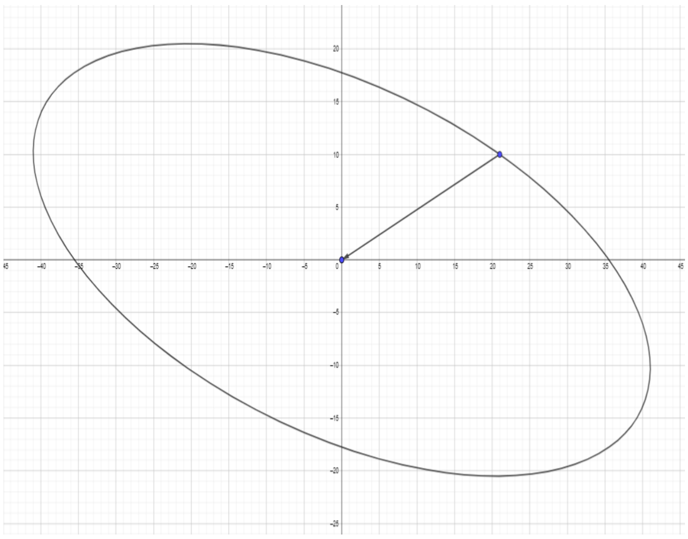


Начальное приближение (4, 1), 2 итерации

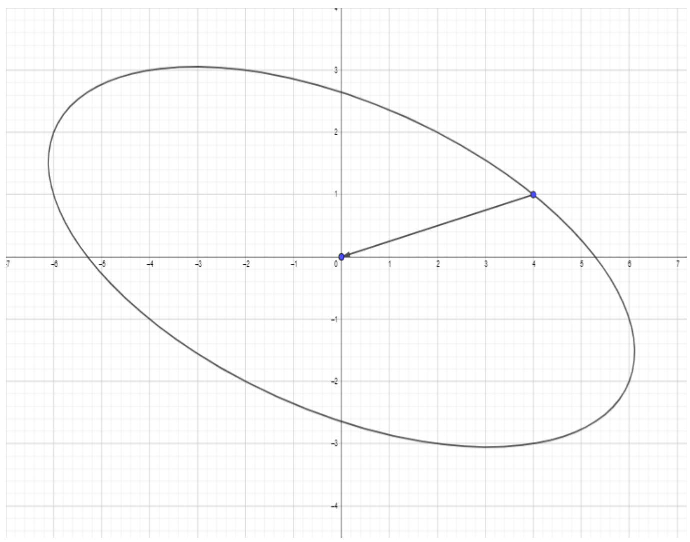


Начальное приближение (21, 10), 2 итерации

Ньютон С Линейным Поиском

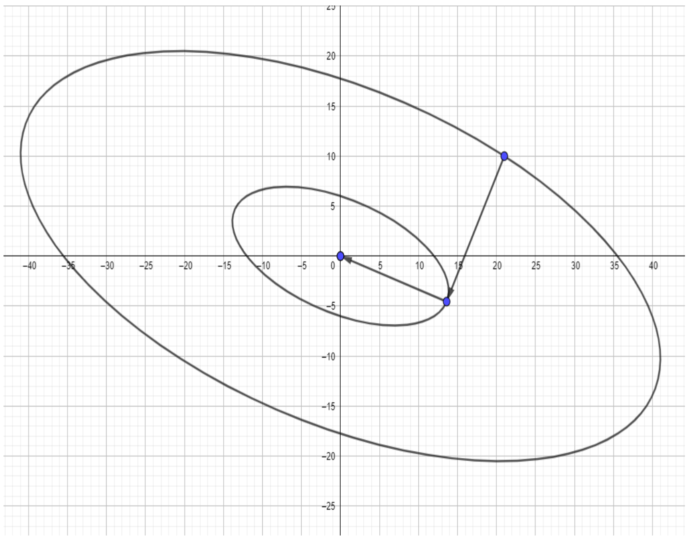


Начальное приближение (4, 1), 2 итерации

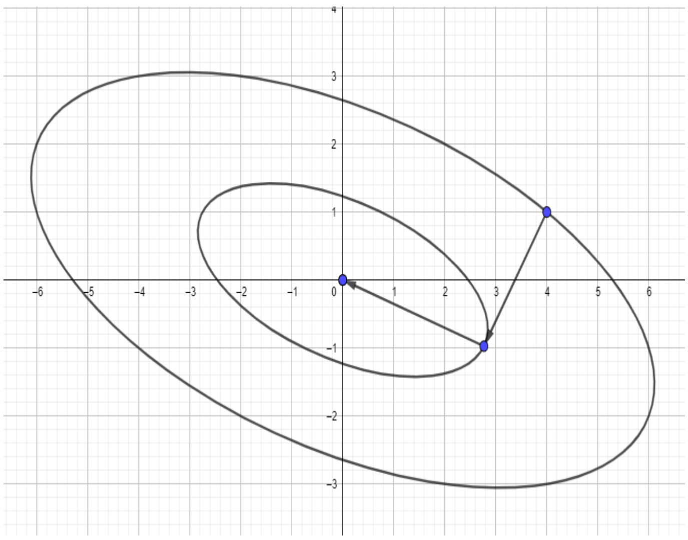


Начальное приближение (21, 10), 2 итерации

Ньютон С Направлением Спуска



Начальное приближение (4, 1), 3 итерации

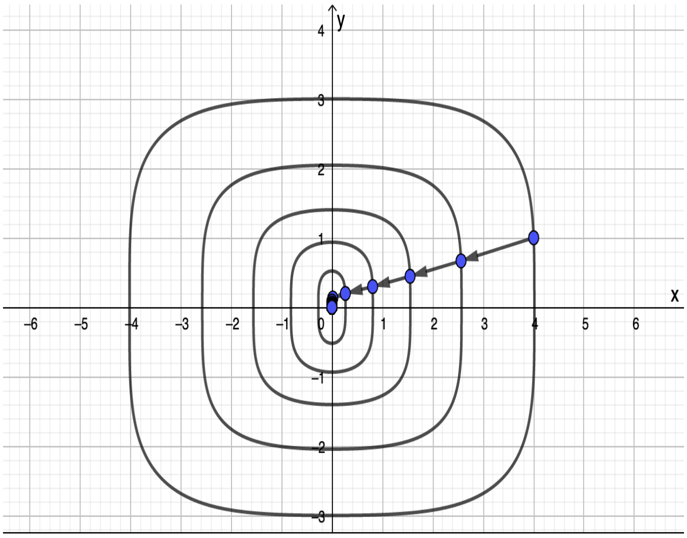


Начальное приближение (21, 10), 3 итерации

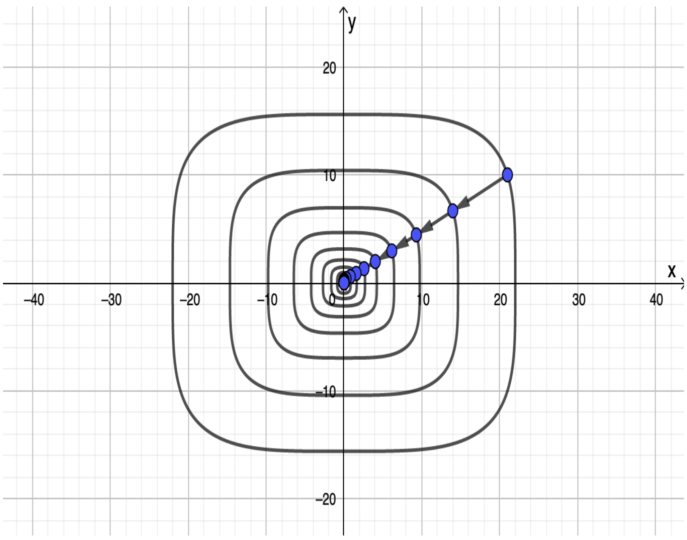
- $f_2(x) = \frac{1}{4}x^4 + y^4 + x^2$

  Искомая точка - (0, 0)

Классический Ньютон

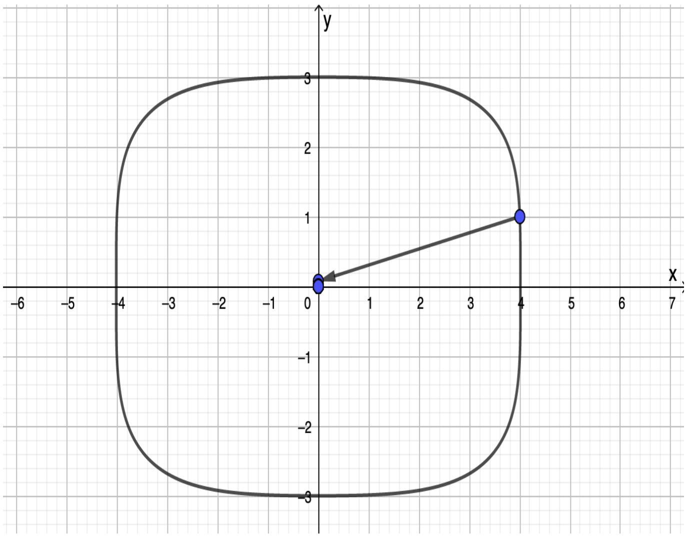


Начальное приближение (4, 1), 39 итераций

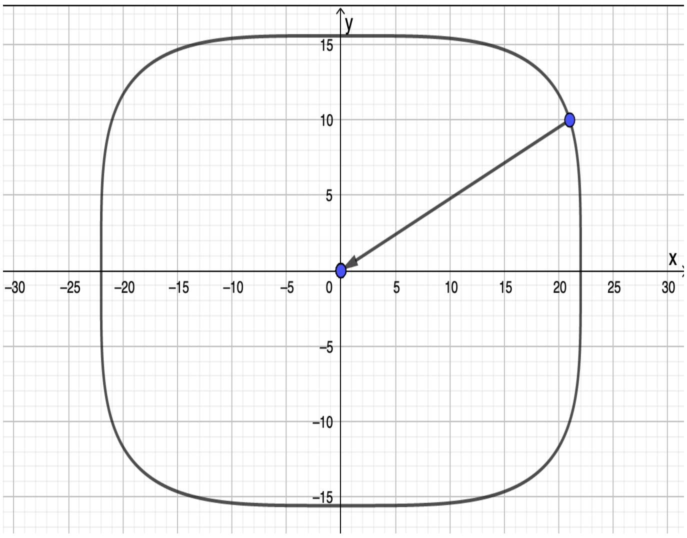


Начальное приближение (21, 10), 44 итерации

Ньютон С Линейным Поиском

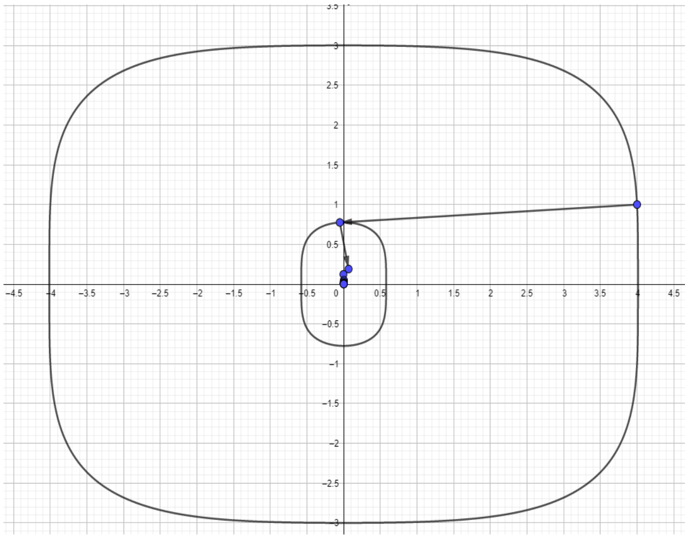


Начальное приближение (4, 1), 21 итераций

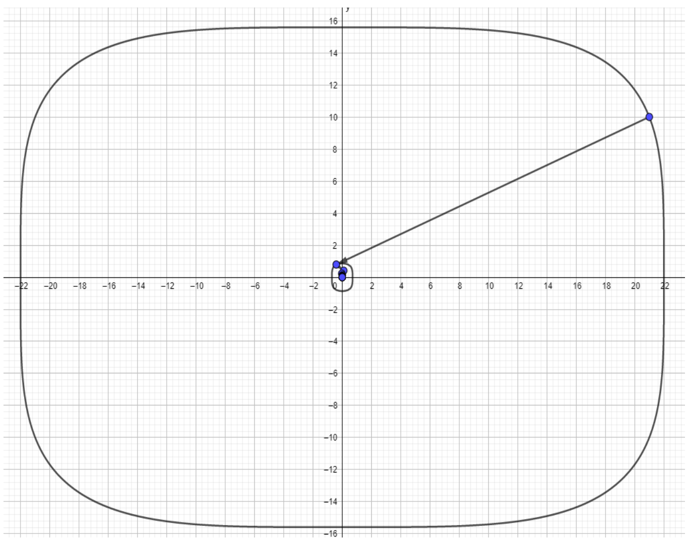


Начальное приближение (21, 10), 19 итераций

Ньютон С Направлением Спуска

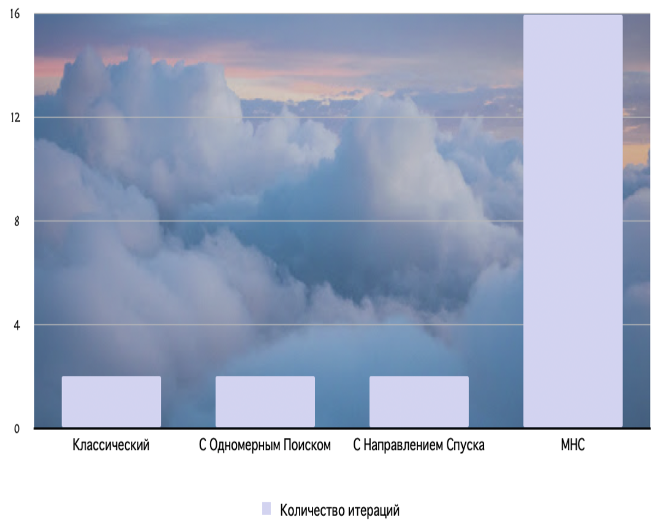


Начальное приближение (4, 1), 27 итераций

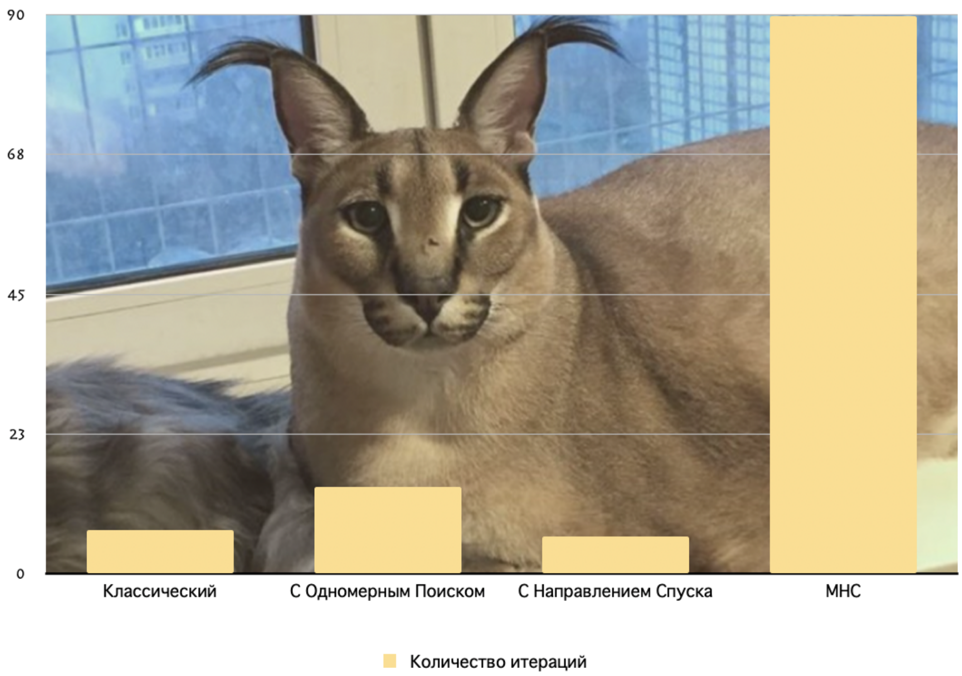


Начальное приближение (21, 10), 28 итераций

3.1 Сравнение методов



$$f(x) = x_1^2 + x_2^2 - 1.2x_1x_2,\ x^0 = (\mathbf{4}, \mathbf{1})^\intercal$$



$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2,\ x^0 = (\mathbf{1.2}, \mathbf{1})^\intercal$$

4 Квазиньютоновские методы

Опишем сначала общий алгоритм, затем для каждого метода опишем способ нахождения матрицы G.

Алгоритм:

1. Вычислить $-\nabla f(x^{k-1})$
2. Находим матрицу G размера $n \times n$
3. $p^k = -G_k \nabla f(x^{k-1})$
4. $\alpha_k = \min\limits_{\alpha} f(x^{k-1} + \alpha p^k)$
5. $x^k = x^{k-1} + \alpha_k p^k$
6. Если $\|x^k - x^{k-1}\| < \varepsilon$ остановиться, иначе перейти к пункту 1.

Введем обозначения: $\Delta x^k = x^k - x^{k-1}, \Delta\omega^k = \omega^k - \omega^{k-1}$

4.1 Метод Бройдена-Флетчера-Шено

На первой итерации $G_0 = I$, потом:

- $G_{k+1} = G_k - \dfrac{\Delta x^k (\Delta x^k)^\intercal}{(\Delta\omega^k, \Delta x^k)} - \dfrac{G_k \Delta\omega^k (\Delta\omega^k)^\intercal G_k^\intercal}{\rho_k} + \rho_k r^k (r^k)^\intercal$
- $r^k = \dfrac{G_k \Delta\omega^k}{\rho_k} - \dfrac{\Delta x^k}{(\Delta\omega^k, \Delta x^k)}$
- $\rho_k = (G_k \Delta\omega^k, \Delta\omega^k)$

4.2 Метод Пауэлла

На первой итерации $G_0 = I$, потом:

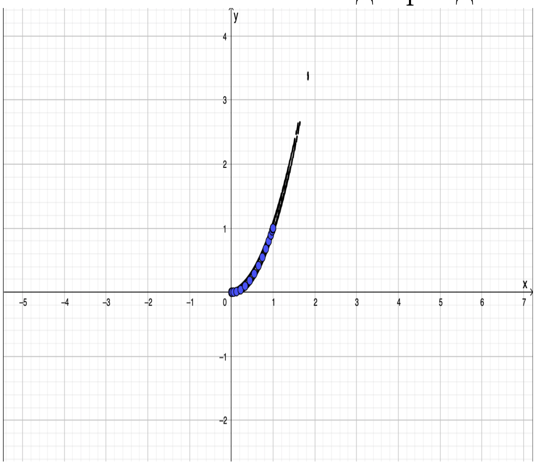
- $G_{k+1} = G_k - \dfrac{\Delta\tilde{x}^k (\Delta\tilde{x}^k)^\intercal}{(\Delta\omega^k, \Delta\tilde{x}^k)},\ k \in \mathbb{N}$
- $\Delta\tilde{x}^k = \Delta x^k + G_k \Delta\omega^k$

4.3 Сравнение методов

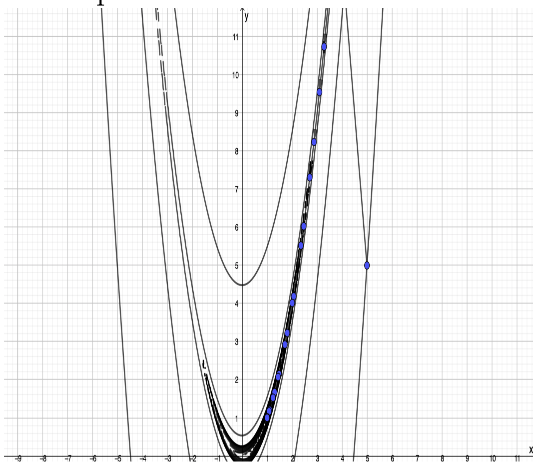
- $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$

  Минимум в точке $(1, -1)$

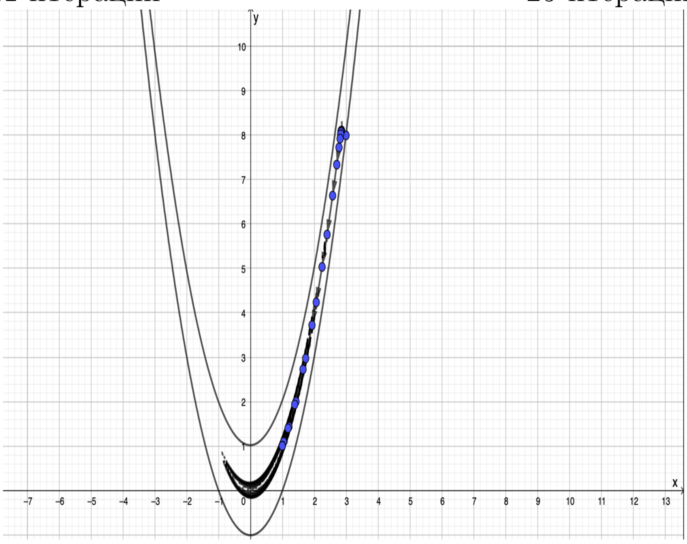
Метод Бройдена-Флетчера-Шено



Начальное приближение (0, 0)
11 итераций

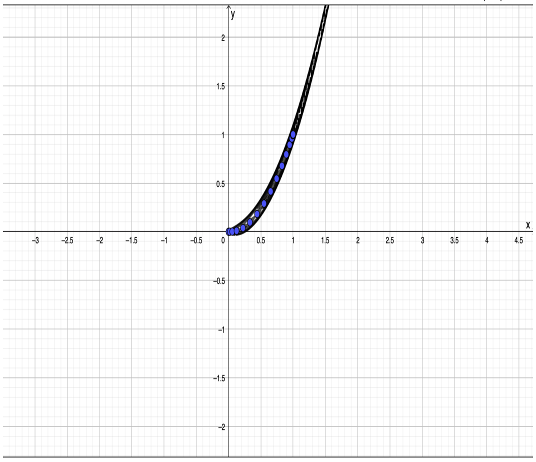


Начальное приближение (5, 5)
25 итераций

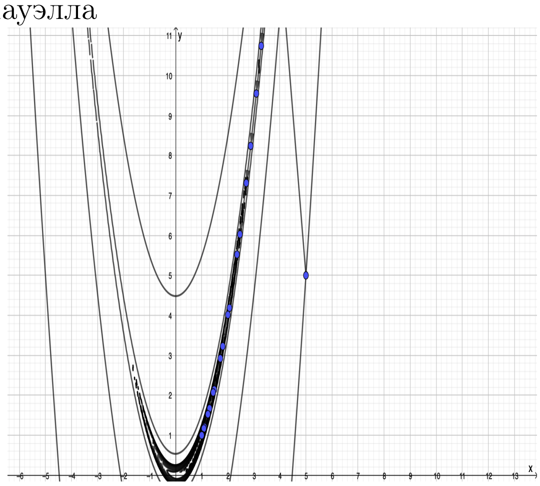


Начальное приближение (3, 8)
23 итерации

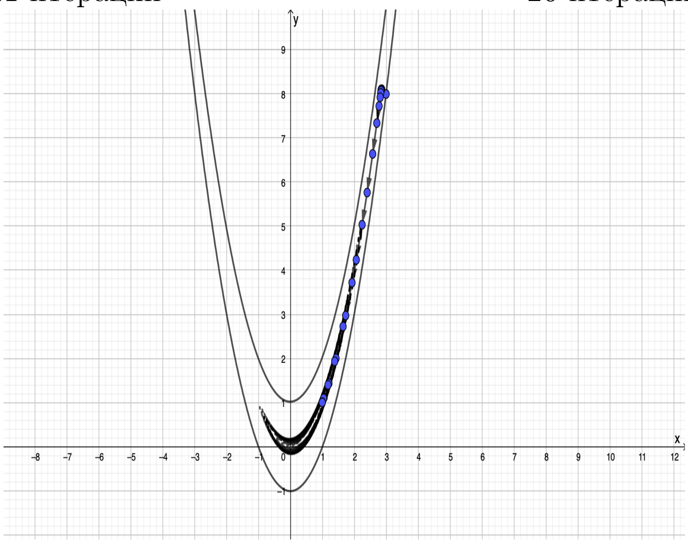
Метод Пауэлла



Начальное приближение (0, 0)
11 итераций

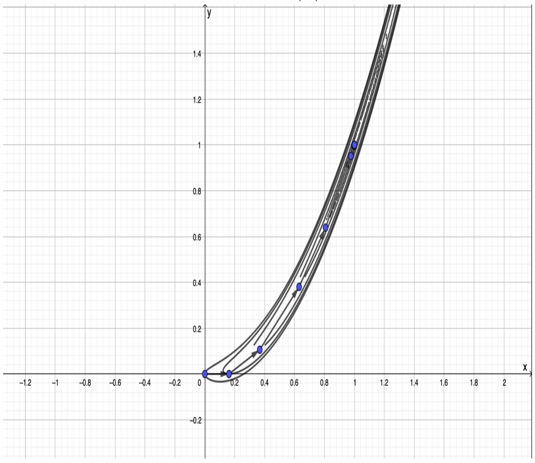


Начальное приближение (5, 5)
26 итераций

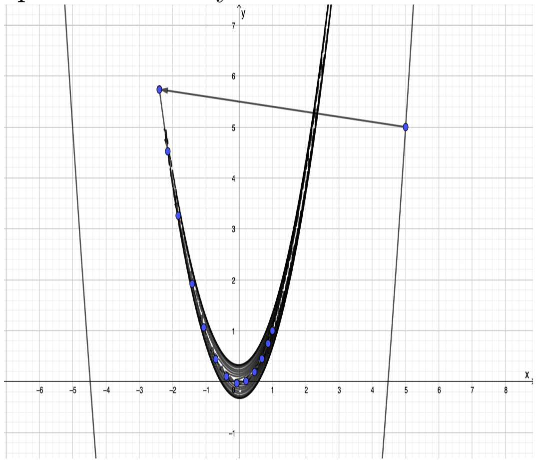


Начальное приближение (3, 8)
23 итерации

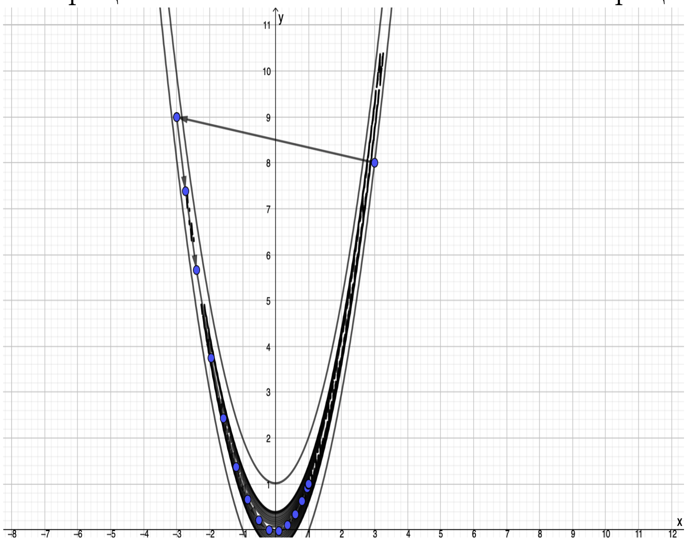
Метод Ньютона С Направлением Спуска



Начальное приближение (0, 0)
9 итераций



Начальное приближение (5, 5)
17 итераций



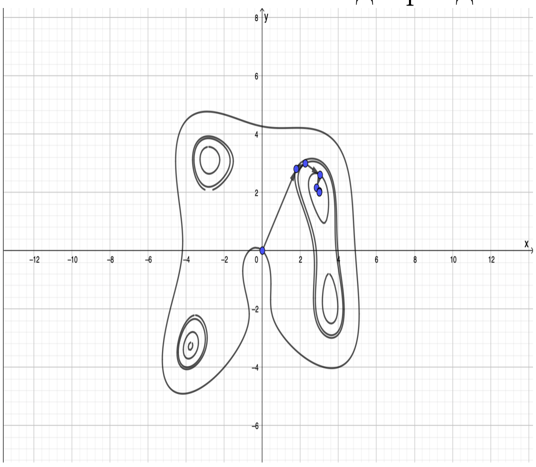
Начальное приближение (3, 8)
19 итераций

- $f(x) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$

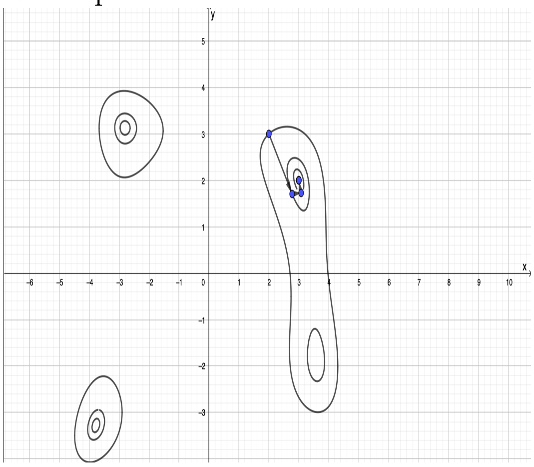
  Минимум в:

  - $(3, 2)$
  - $(-2.805118086, 3.131312518)$
  - $(3.584428340, -1.848126526)$
  - $(-3.779310253, -3.283185991)$

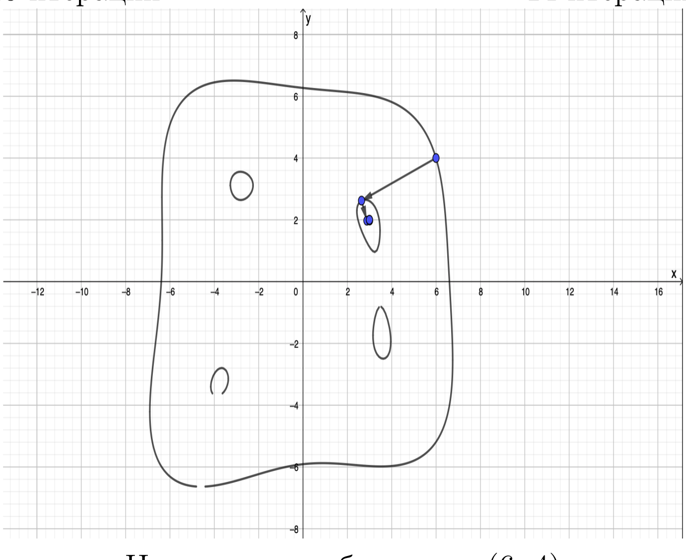
Метод Бройдена-Флетчера-Шено



Начальное приближение (0, 0)
10 итераций

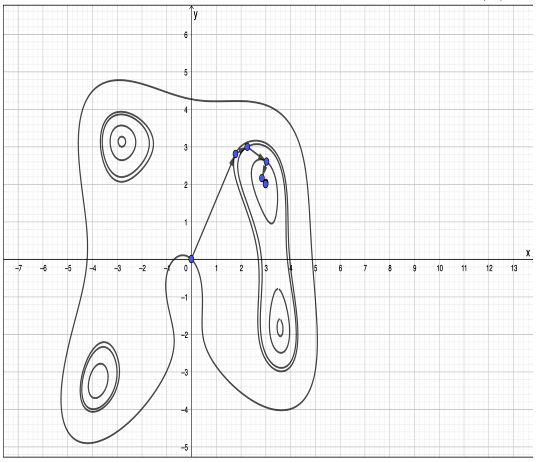


Начальное приближение (2, 3)
14 итераций

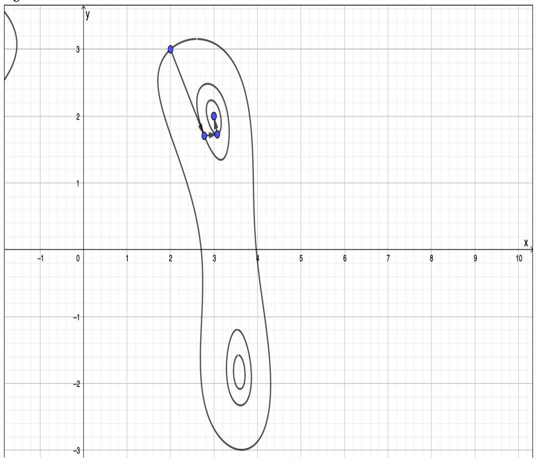


Начальное приближение (6, 4)
6 итерации

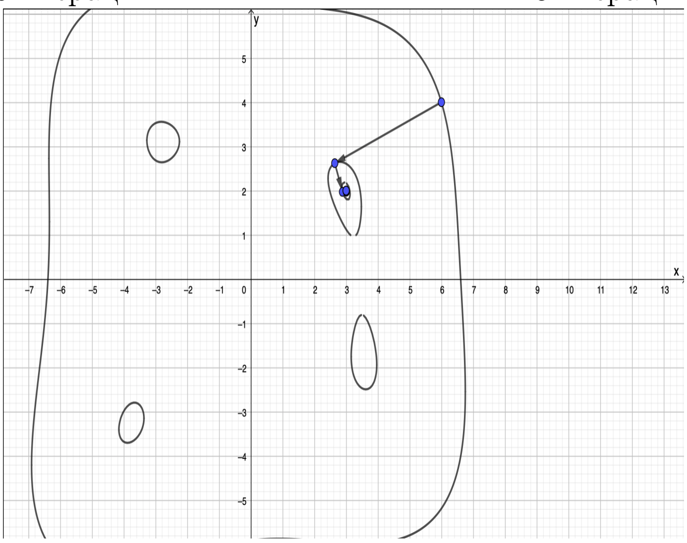
Метод Пауэлла



Начальное приближение (0, 0)
9 итераций

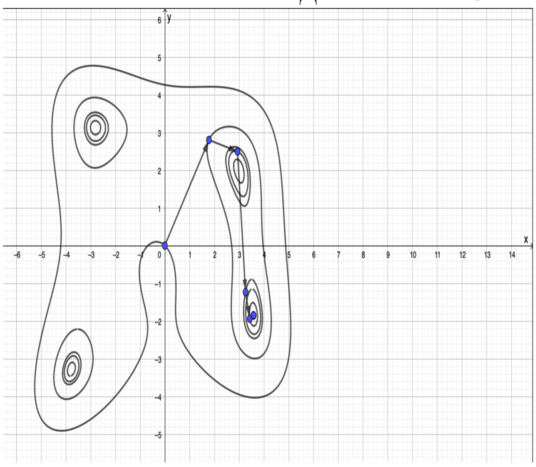


Начальное приближение (2, 3)
8 итераций

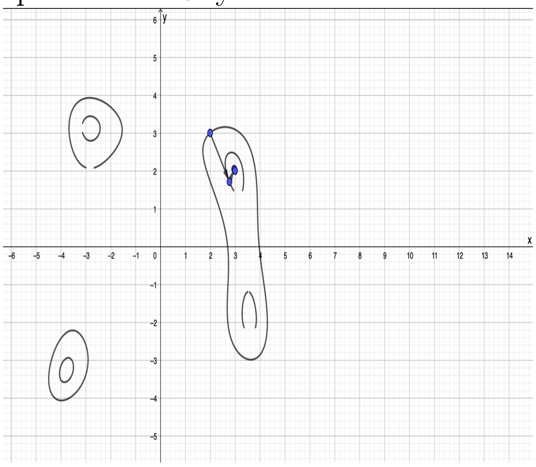


Начальное приближение (6, 4)
6 итерации

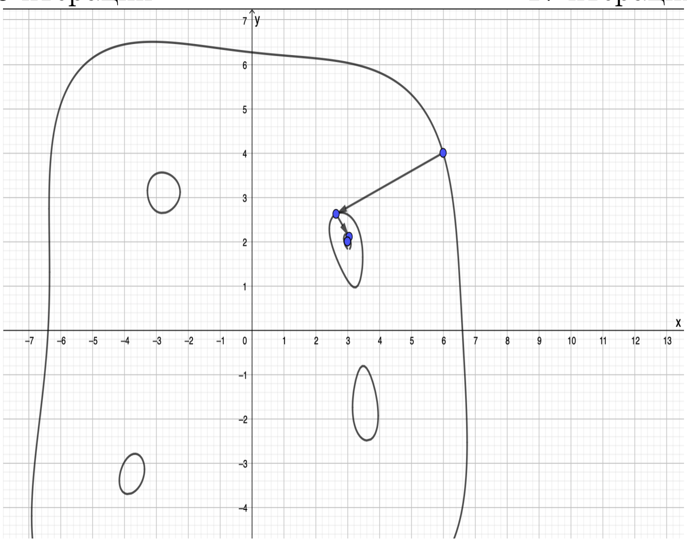
Метод Ньютона С Направлением Спуска



Начальное приближение (0, 0)
8 итераций



Начальное приближение (2, 3)
17 итераций

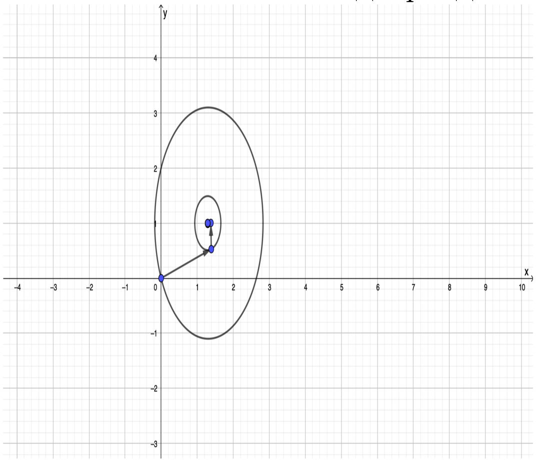


Начальное приближение (6, 4)
19 итераций

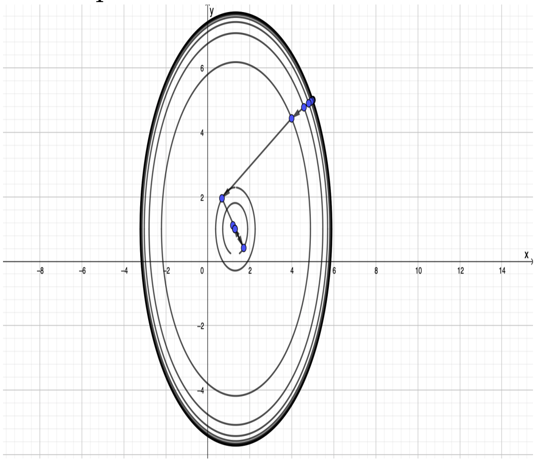
- $f(x) = 100 - \dfrac{2}{1 + \left(\dfrac{x_1 - 1}{2}\right)^2 + \left(\dfrac{x_2 - 1}{3}\right)^2} - \dfrac{1}{1 + \left(\dfrac{x_1 - 2}{2}\right)^2 + \left(\dfrac{x_2 - 1}{3}\right)^2}$

  Минимум в (1.291643031, 1)

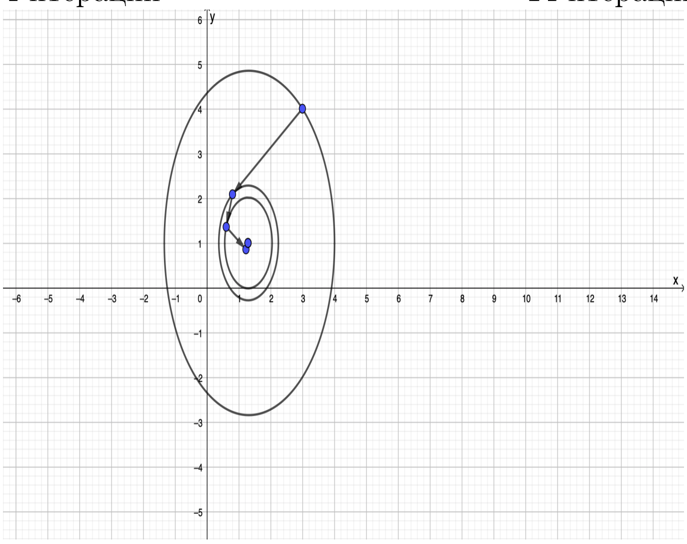
Метод Бройдена-Флетчера-Шено



Начальное приближение (0, 0)
34 итерации

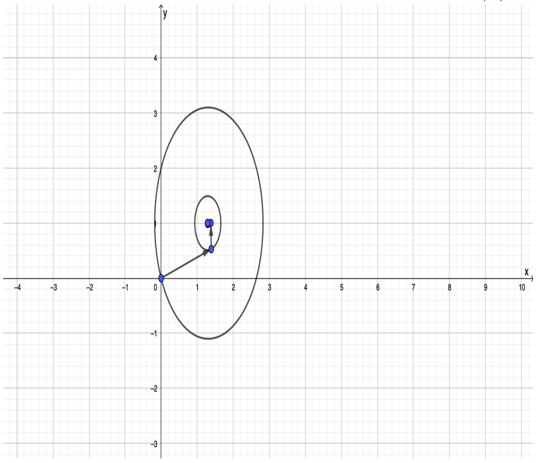


Начальное приближение (2, 3)
14 итераций

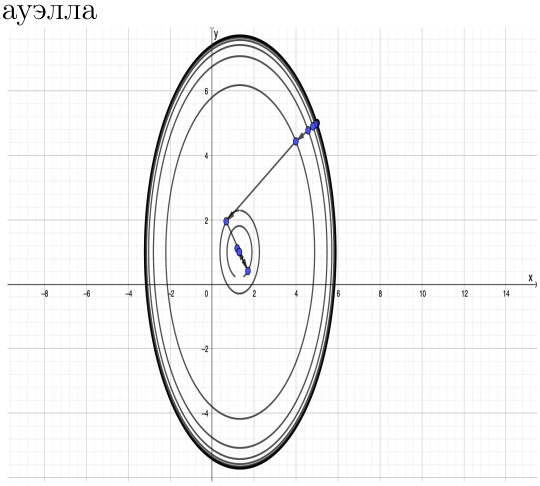


Начальное приближение (3, 4)
6 итераций

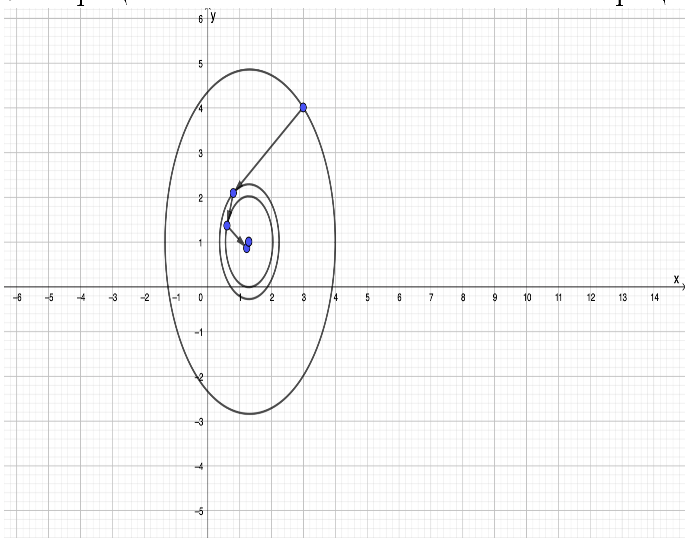
Метод Пауэлла



Начальное приближение (0, 0)
29 итераций

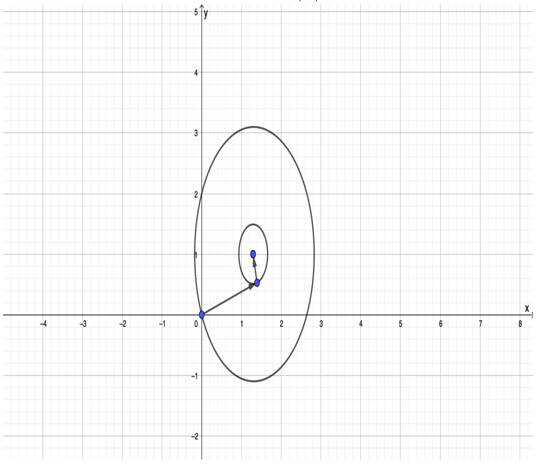


Начальное приближение (5, 5)
14 итераций

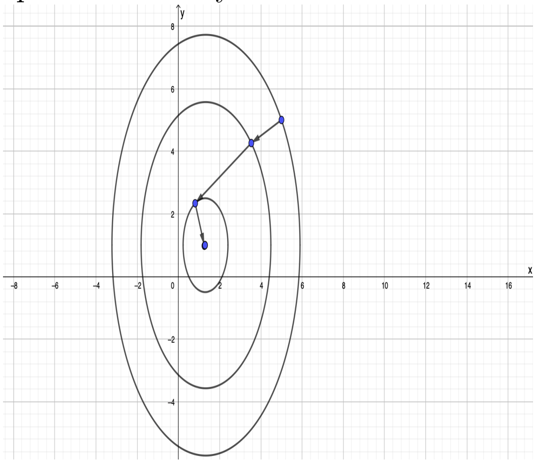


Начальное приближение (3, 4)
6 итераций

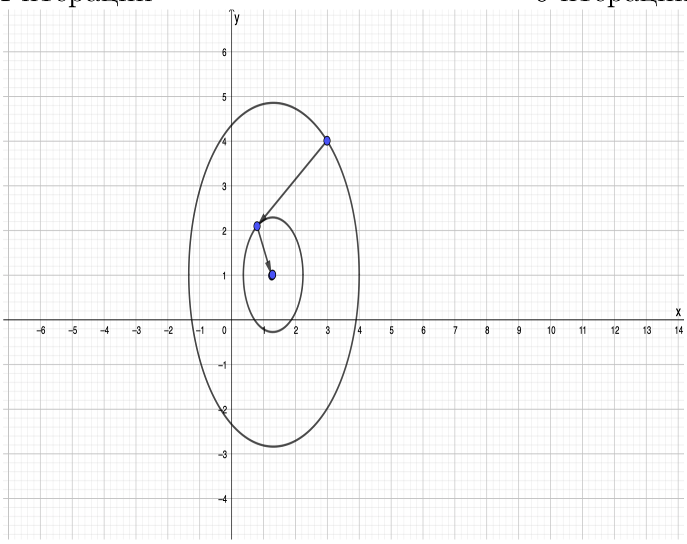
Метод Ньютона С Направлением Спуска



Начальное приближение (0, 0)
4 итерации



Начальное приближение (5, 5)
6 итераций



Начальное приближение (3, 4)
5 итераций

- $f(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$

  Минимум в $(0, 0, 0, 0)$

метод\приближение	(1, 1, 1, 1)	(5, 4, 3, 2)	(10, 10, 10, 10)
Бройдена-Флетчера-Шено	12	36	55
Пауэлла	16	22	15
Ньютон С Напр Спуска	30	31	30

5 Метод Марквардта

Комбинация методов наискорейшего спуска и Ньютона:

- движение из x^0 в направлении $-\nabla f(x)$ приводит к существенному уменьшению f(x),
- направление эффективного поиска в окрестности точки x^* определяется по методу Ньютона.

  Определить $x_0, \tau_0, \beta, \varepsilon$

  1. $x = x_0$, вычислить f(x)
  2. вычислить $\nabla f(x), H(x), \tau = \tau_0$
  3. СЛАУ: $(H(x) + \tau I)p = -\nabla f(x)$
  4. $y = x + p,\ f(y)$
  5. если $f(y) > f(x)$, $\tau = \dfrac{\tau}{\beta}$, переходим к (3)
  6. $x = y$, $f(x) = f(y)$, $\tau_0 = \tau_0\beta$
  7. если $\|p\| > \varepsilon$, то переходим к (2), иначе стоп

Метод Марквардта

итерация	τ
1	1.000000e+02
2	1.000000e+02
3	5.000000e+01
4	2.500000e+01
5	1.250000e+01
6	6.250000e+00
7	3.125000e+00
8	1.562500e+00
9	7.812500e-01
10	3.906250e-01
11	1.953125e-01
12	9.765625e-02
13	4.882813e-02
14	2.441406e-02
15	1.220703e-02
16	9.765625e-02
17	3.051758e-03
18	3.906250e-01
19	7.629395e-04
20	1.953125e-01
21	1.907349e-04
22	1.953125e-01
23	4.768372e-05
24	1.953125e-01
25	1.192093e-05
26	5.960464e-06
27	2.980232e-06
28	1.490116e-06
29	7.450581e-07
30	3.725290e-07

Метод Марквардта с разложением Холесского

итерация	τ
1	1.000000e+02
2	1.000000e+02
3	5.000000e+01
4	2.500000e+01
5	1.250000e+01
6	6.250000e+00
7	3.125000e+00
8	1.562500e+00
9	7.812500e-01
10	3.906250e-01
11	1.953125e-01
12	9.765625e-02
13	4.882813e-02
14	2.441406e-02
15	1.220703e-02
16	6.103516e-03
17	3.051758e-03
18	1.525879e-03
19	7.629395e-04
20	3.814697e-04
21	1.907349e-04
22	9.536743e-05
23	4.768372e-05
24	2.384186e-05
25	1.192093e-05

6 Выводы

Если матрица Гессе $\nabla^2 f(x)$ удовлетворяет условию Липшица, то для квазиньютоновских методов присуще квадратичная скорость сходимости. В этом преимущество квазиньютоновских методов перед методами сопряженных градиентов, которые требуют приблизительно в n раз больше итераций для одного и того же асимптотического поведения. Но минусом является, что память требуется пропорционально n^2 и матричные вычисления пропорционально n^2.

Метод Ньютона имеет квадратичную локальную скорость сходи- мости, если $\nabla^2 f(x)$ удовлетворяет в окрестности точки x^* условию Липшица. Следовательно, если использовать метод Ньютона в малой окрестности точки, он ходится быстрее остальных методов.

Все методы Ньютона достигают результата быстрее, чем МНС. Самым надежным является - метод Ньютона с направлением спуска, он также требует меньше итераций.