

Министерство науки и высшего образования Российской Федерации
федеральное государственное автономное образовательное учреждение высшего образования
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО»

Отчет
по лабораторной работе

Изучение алгоритмов метода Ньютона и его модификаций

по дисциплине **«Методы оптимизации»**

Авторы: Герасимов Михаил, Пушкарев Глеб, Асмирко Антон

Группа: М3235

Факультет: ФИТиП

Преподаватель: Корсун Мария Михайловна,
Михаил Свинцов Викторович



УНИВЕРСИТЕТ ИТМО

Санкт-Петербург 2021

Лабораторная работа №4

Тема. Изучение алгоритмов метода Ньютона и его модификаций, в том числе квазиньютоновских методов.

Цель. Разработать программы для безусловной минимизации функций многих переменных.

Реализовать алгоритмы

1. Метод Ньютона: **а)** классический, **б)** с одномерным поиском (одномерный метод на выбор студентов), **в)** с направлением спуска.

1.1. 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).

1.2. Исследуйте работу методов на двух функциях с заданным начальным приближением:

$$f(x) = x_1^2 + x_2^2 - 1.2x_1x_2, x^0 = (4, 1)^T;$$

$$f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2, x^0 = (-1.2, 1)^T.$$

- Для поиска ньютоновского направления спуска необходимо использовать прямой или итерационный метод решения СЛАУ (даже, если она размерности 2).
- Сравните результаты с минимизацией методом наискорейшего спуска (из лаб. работы 2).
- Постройте таблицу или график зависимости «метод : количество итераций».

- Для каждого метода приведите иллюстрации траекторий сходимости.

2. Квазиньютоновский метод (вариант выдает преподаватель):

вариант 1: метод Давидона-Флетчера-Пауэлла и метод Пауэлла;

вариант 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».

Для отображения траектории поиска точки используйте ранее реализованную графическую систему или сторонние программы.

Бонусное задание. 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 Холецкого дополнительно построить график зависимости «итерация – число разложений Холецкого».

Отчет. По результатам выполнения лабораторной работы необходимо подготовить отчет: титульный лист с указанием организации, названия учебной

дисциплины, темы работы, номера варианта, исполнителя и принимающего, города, года.

Отчет должен содержать все исследования, проведенные в п.п. 1-2, а также соответствующие выводы. Для разработанного программного кода в отчете привести код основных модулей, диаграмму классов, сделать текстовое описание.

Требования к программному коду (вычислительные алгоритмы)

1. Рекомендуется использовать языки программирования: C++, C#, Java.
2. Рекомендуется придерживаться основных положений ООП при разработке.
3. Рекомендуется выполнять документирование программного кода.

Оценка результатов

<i>Задание</i>	<i>Результат (в виде коэффициента)</i>
Сдача в срок	0.15
Основные численные результаты, выводы, защита	0.0 – 0.55
Программная реализация и индивидуальный код	0.0 – 0.25
Грамотность изложения и общее качество отчета	0.0 – 0.05
Бонусное задание	0.0 – 0.3

Срок сдачи четвертой лабораторной работы – до 3.06.2021 (включительно).

После указанного срока лабораторные приниматься не будут!

Описание структур и методов

1. Метод Ньютона

5.7. Метод Ньютона

Пусть функция $f(x)$ дважды дифференцируема в E_n . Тогда, с помощью градиента и матрицы Гессе, для нее можно записать разложение в ряд по формуле Тейлора в окрестности точки x^k

$$f(x) = f(x^k) + \nabla f(x^k)^T \Delta x + \frac{1}{2} \Delta x^T H(x^k) \Delta x + o(\|\Delta x\|^2), \quad (5.27)$$

где $o(\|\Delta x\|^2)$ – сумма всех членов разложения, имеющих порядок выше второго; $\Delta x^T H(x^k) \Delta x$ – квадратичная форма.

Отсюда следует, что поведение функции $f(x)$ с точностью до величины порядка $o(\|\Delta x\|^2)$ может быть описано квадратичной функцией

$$\Phi_k(x) = f(x^k) + \nabla f(x^k)^T \Delta x + \frac{1}{2} \Delta x^T H(x^k) \Delta x. \quad (5.28)$$

Минимизируем функцию $\Phi_k(x)$ вместо $f(x)$. Найдем ее точку минимума x^{k+1} из условия стационарности точки $\nabla \Phi_k(x) = 0$

$$\nabla \Phi_k(x) = H(x^k)(x - x^k) + \nabla f(x^k) = 0. \quad (5.29)$$

Пусть матрица Гессе $H(x)$ положительно определена при всех $x \in E_n$, следовательно, она невырождена ($\det H(x) > 0$). Тогда существует обратная матрица $H^{-1}(x)$. Отметим, что квадратичная функция (5.28) с положительно определенной матрицей $H(x^k)$ сильно выпукла и уравнение (5.29) определяет единственную точку глобального минимума функции $\Phi_k(x)$. Умножим слева обе части равенства (5.29) на матрицу $H^{-1}(x^k)$ и найдем точку минимума x^{k+1} квадратичной функции $\Phi_k(x)$, аппроксимирующей $f(x)$ в окрестности точки $x = x^k$

$$x^{k+1} = x^k - H^{-1}(x^k) \nabla f(x^k), \quad k = 0, 1, \dots \quad (5.30)$$

Итерационный процесс минимизации функции многих переменных (5.30), начатый из произвольной точки $x^0 \in E_n$, называется *методом Ньютона* и является обобщением метода Ньютона в одномерном случае.

Очевидно, для квадратичной функции с положительно определенной матрицей A применение метода Ньютона обеспечивает получение точки глобального минимума ровно за один шаг из любой точки $x^0 \in E_n$.

Для выпуклой функции, отличной от квадратичной, применение этого метода обеспечивает, как правило, быструю сходимость. Дело в том, что на каждом шаге итерационного процесса используется информация о поведении функции $f(x)$ в окрестности точки x^k , содержащаяся в значениях не только первых, но и вторых ее частных производных. Поэтому при прочих равных условиях следует ожидать более быструю сходимость метода Ньютона по сравнению с градиентными методами.

При выборе достаточно хорошего начального приближения $x^0 \in E_n$ минимизирующая последовательность $\{x^k\}$ для сильно выпуклой дважды дифференцируемой функции $f(x)$ сходится к точке минимума с квадратичной скоростью $\rho(x^k, x^*) \leq [c \rho(x^0, x^*)]^2$. Если точка x^0 выбрана недостаточно близко к точке x^* , то последовательность (5.30) может расходиться.

Отметим, что даже сходящаяся последовательность $\{x^k\}$ метода Ньютона не всегда обеспечивает монотонное убывание $f(x)$, т.е. неравенство $f(x^{k+1}) < f(x^k)$ для некоторых $k = 0, 1, \dots$ может нарушаться. Этот недостаток устранен в *обобщенном методе Ньютона*

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

где величина $\alpha_k > 0$ находится на каждом шаге из условия исчерпывающего спуска по направлению $p^k = -H^{-1}(x^k) \nabla f(x^k)$. Можно показать, что если целевая функция является сильновыпуклой и ее матрица Гессе $H(x)$ для любых точек $x, y \in E_n$ удовлетворяет неравенству $\|H(x) - H(y)\| \leq L \|x - y\|$, то при произвольном выборе начальной точки обобщенный метод Ньютона обладает квадратичной скоростью сходимости.

Основным недостатком метода Ньютона можно считать необходимость вычисления и обращения матрицы Гессе на каждой итерации.

2. Квазиньютоновские методы

Среди алгоритмов многомерной минимизации можно выделить группу алгоритмов, которые объединяют достоинства метода наискорейшего спуска и метода Ньютона. Такие алгоритмы называют *квазиньютоновскими методами*. Особенность этих алгоритмов состоит в том, что при их применении нет необходимости обращать матрицу Гессе целевой функции $f(x)$ и в то же время удается сохранить высокую скорость сходимости, присущую методу Ньютона и его модификациям.

Элементы релаксационной последовательности $\{x^k\}$ в алгоритмах квазиньютоновских методов минимизации непрерывно дифференцируемой в E_n целевой функции $f(x)$ строят в соответствии с рекуррентным соотношением $x^{k+1} = x^k + \alpha_k p^k$, $k = 0, 1, \dots$, но направление спуска на каждой k -й итерации задают в виде

$$p^k = -A_k \nabla f(x_k) = A_k w^k, \quad (5.31)$$

где $w^k = -\nabla f(x^k)$ – антиградиент целевой функции, а A_k – положительно определенная матрица порядка n , обновляемая на каждой итерации. Отметим, что направление, задаваемое на каждой k -й итерации вектором p^k , является направлением спуска, так как с учетом (5.31)

$$(\nabla f(x^k), p^k) = -(w^k, A_k w^k) < 0.$$

Нетрудно заметить, что при $A_k = H^{-1}(x^k)$ метод переходит в классический метод Ньютона, а при $A_k = \alpha_k$ – в один из градиентных методов. Матрицы $\{A_k\}$ определяют таким образом, чтобы последовательность $\{A_k\}$ при $k \rightarrow \infty$ имела предел, равный $H^{-1}(x^*)$, где $H(x^*)$ – матрица Гессе целевой функции, вычисленная в точке минимума x^* .

Известно, что если целевая функция является сильно выпуклой, то алгоритмы метода Ньютона обладают квадратичной скоростью сходимости. Поэтому можно ожидать, что алгоритмы квазиньютоновских методов будут иметь достаточно высокую скорость сходимости, если на каждой k -й итерации матрица A_k выбрана близкой к матрице $H^{-1}(x^k)$. Используя при конструировании A_k аппроксимацию матрицы $H^{-1}(x^k)$ с учетом информации о градиенте целевой функции в точке x^k , можно существенно упростить процедуру нахождения направления спуска на k -й итерации. Именно эти соображения лежат в основе построения алгоритмов квазиньютоновских методов.

Квазиньютоновские методы чрезвычайно популярны среди пользователей методов оптимизации в связи с тем, что они сочетают высокую скорость сходимости с невысокой трудоемкостью итерации. Практическое значение этих методов трудно переоценить, хотя с теоретической точки зрения любой квазиньютоновский метод есть не более чем умная реализация фундаментальной идеи метода Ньютона.

3. Алгоритм Бroyдена — Флетчера — Шанно

Пусть решается задача оптимизации функционала: $\arg \min_x f(x)$

Методы второго порядка решают данную задачу итерационно, с помощью разложения функции в полином второй степени:

$$f(x_k + p) = f(x_k) + \nabla f^T(x_k)p + \frac{1}{2}p^T H(x_k)p$$

где H — гессиан функционала f в точке x . Зачастую вычисление гессиана трудоемки, поэтому BFGS алгоритм вместо настоящего значения $H(x)$ вычисляет приближенное значение B_k , после чего находит минимум полученной квадратичной задачи:

$$p_k = -B_k^{-1} \nabla f(x_k)$$

В качестве начального приближения гессиана можно брать любую невырожденную, хорошо обусловленную матрицу. Часто берут единичную матрицу. Приближенное значение гессиана на следующем шаге вычисляется по формуле:

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k^T}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}$$

где I — единичная матрица, $s_k = x_{k+1} - x_k$ — шаг алгоритма на итерации, $y_k = \nabla f_{k+1} - \nabla f_k$ — изменение градиента на итерации.

Поскольку вычисление обратной матрицы вычислительно сложно, вместо того, чтобы вычислять B_k^{-1} , обновляется обратная к B_k матрица $C_k = B_k^{-1}$

$$C_{k+1} = (I - \rho_k s_k y_k^T) C_k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T$$

где

$$\rho_k = \frac{1}{y_k^T s_k}$$

4. Метод Пауэлла

Матрица B_k определяется следующим образом:

$$B_1 = I$$

$$B_{k+1} = B_k - \frac{y_k y_k^T}{(\Delta\omega^k, y^k)}$$

где

$$y^k = \Delta x^k + B_k \Delta\omega^k$$

Результаты исследований 1

$$f(x, y) = -x\sqrt{y} + 2x^2 - 7x + y$$

$$f(x, y) \rightarrow \min = M(2, 1); f(2, 1) = -7$$

$$eps = 10^{-6}$$

Newton			Newton				Newton		
Start:	(0; 2)		Start:	(0; 1)			Start:	(1; 1)	
Answer x:	(2; 1)		Answer x:	(-nan(ind); -nan(ind))			Answer x:	(2; 1)	
Answer y:	-7		Answer y:	-nan(ind)			Answer y:	-7	
Iterations:	3		Iterations:	2			Iterations:	7	
-----			-----				-----		
NewtonDirectionDescent			NewtonDirectionDescent				NewtonDirectionDescent		
Start:	(0; 2)		Start:	(0; 1)			Start:	(1; 1)	
Answer x:	(2; 1)		Answer x:	(2; 0.999997)			Answer x:	(2; 1)	
Answer y:	-7		Answer y:	-7			Answer y:	-7	
Iterations:	4		Iterations:	4			Iterations:	4	
-----			-----				-----		
NewtonOneDimension			NewtonOneDimension				NewtonOneDimension		
Start:	(0; 2)		Start:	(0; 1)			Start:	(1; 1)	
Answer x:	(2; 1)		Answer x:	(2; 0.999997)			Answer x:	(2; 1)	
Answer y:	-7		Answer y:	-7			Answer y:	-7	
Iterations:	4		Iterations:	4			Iterations:	4	
-----			-----				-----		

Как видно из результатов неквадратичные функции плохо подходят для Ньютоновских методов. При некоторых начальных значениях результат работы даже не смог сойтись и стал равен NaN.

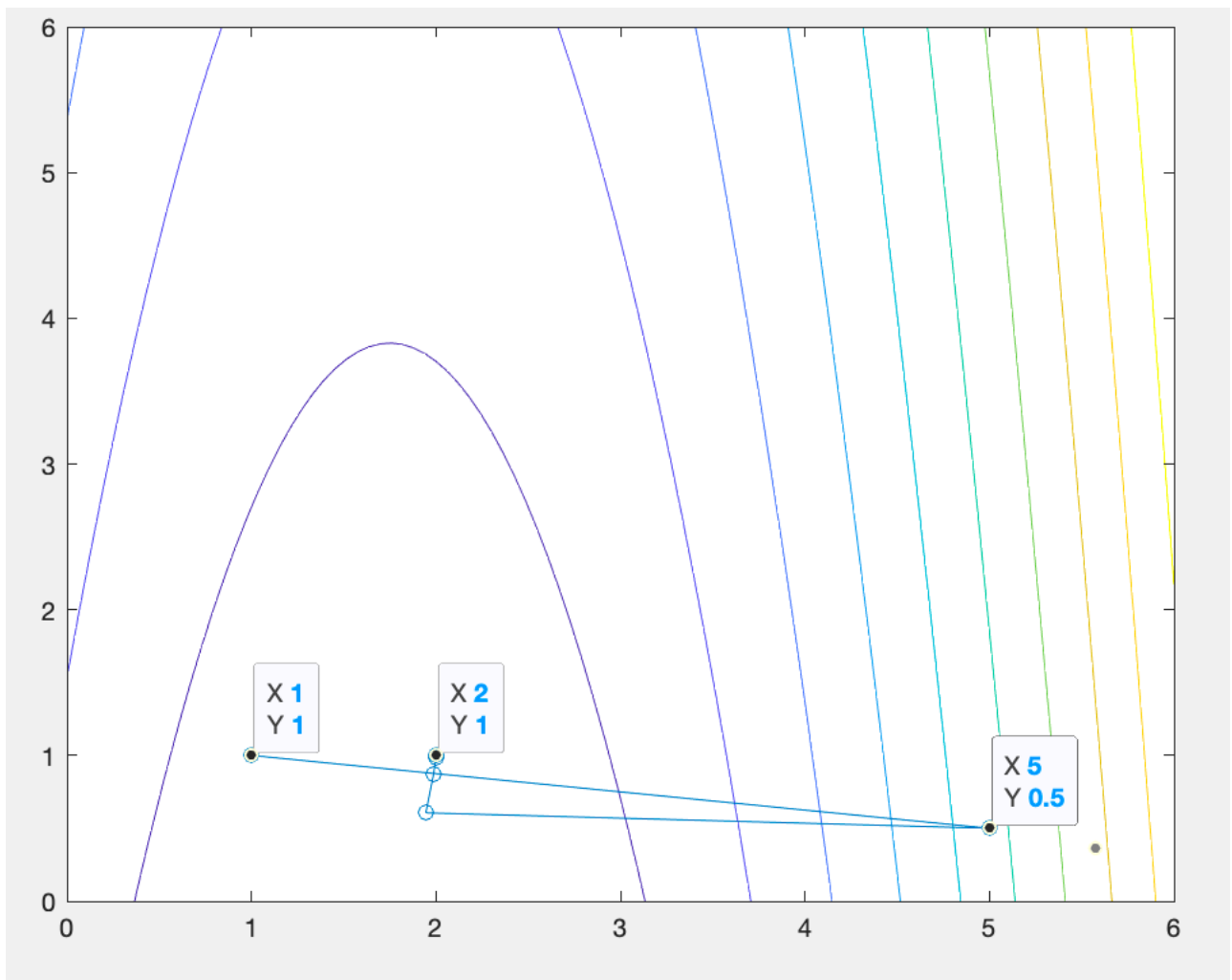


Figure 1: Newton Start (1;1)

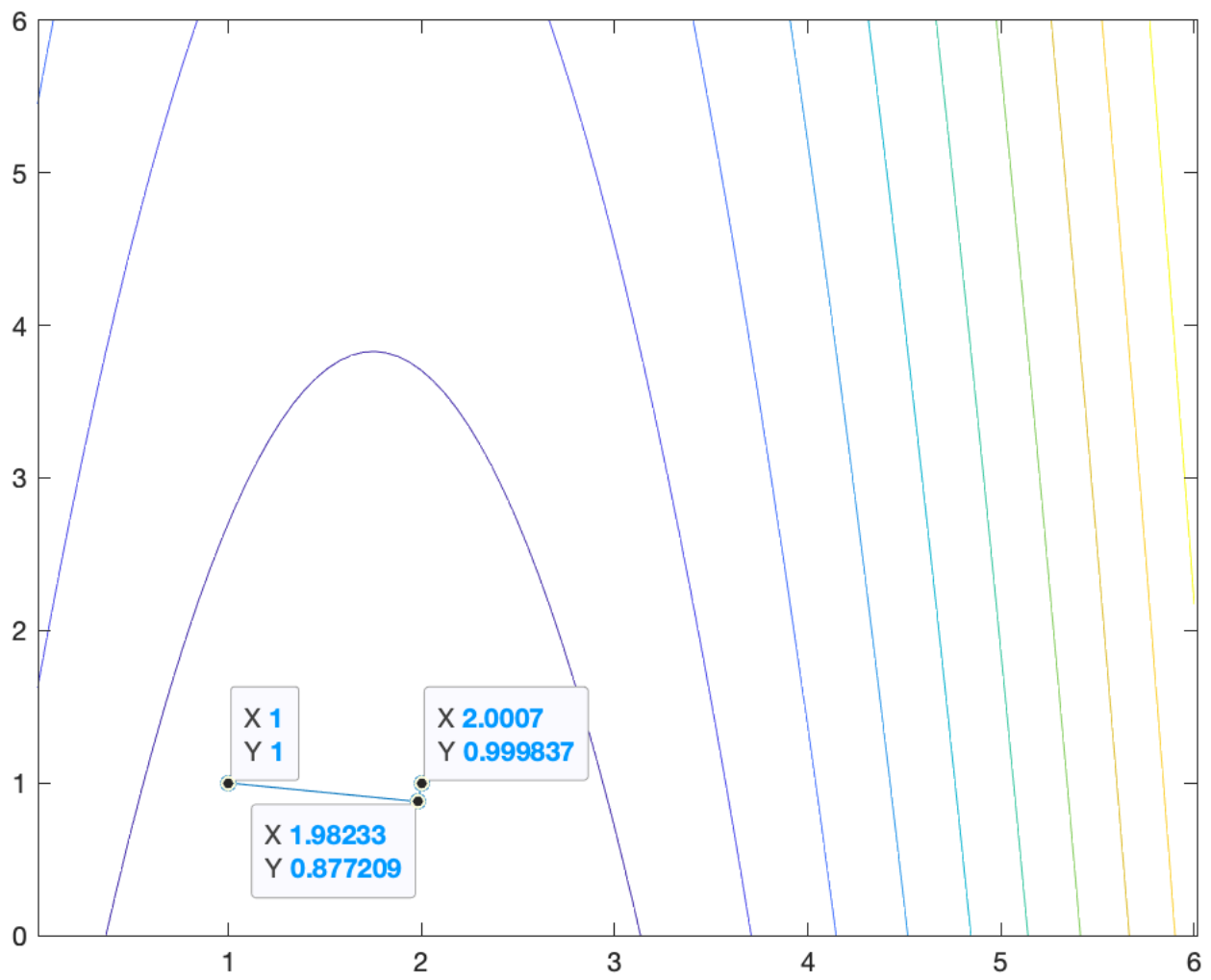


Figure 2: NewtonOneDimension Start (1;1)

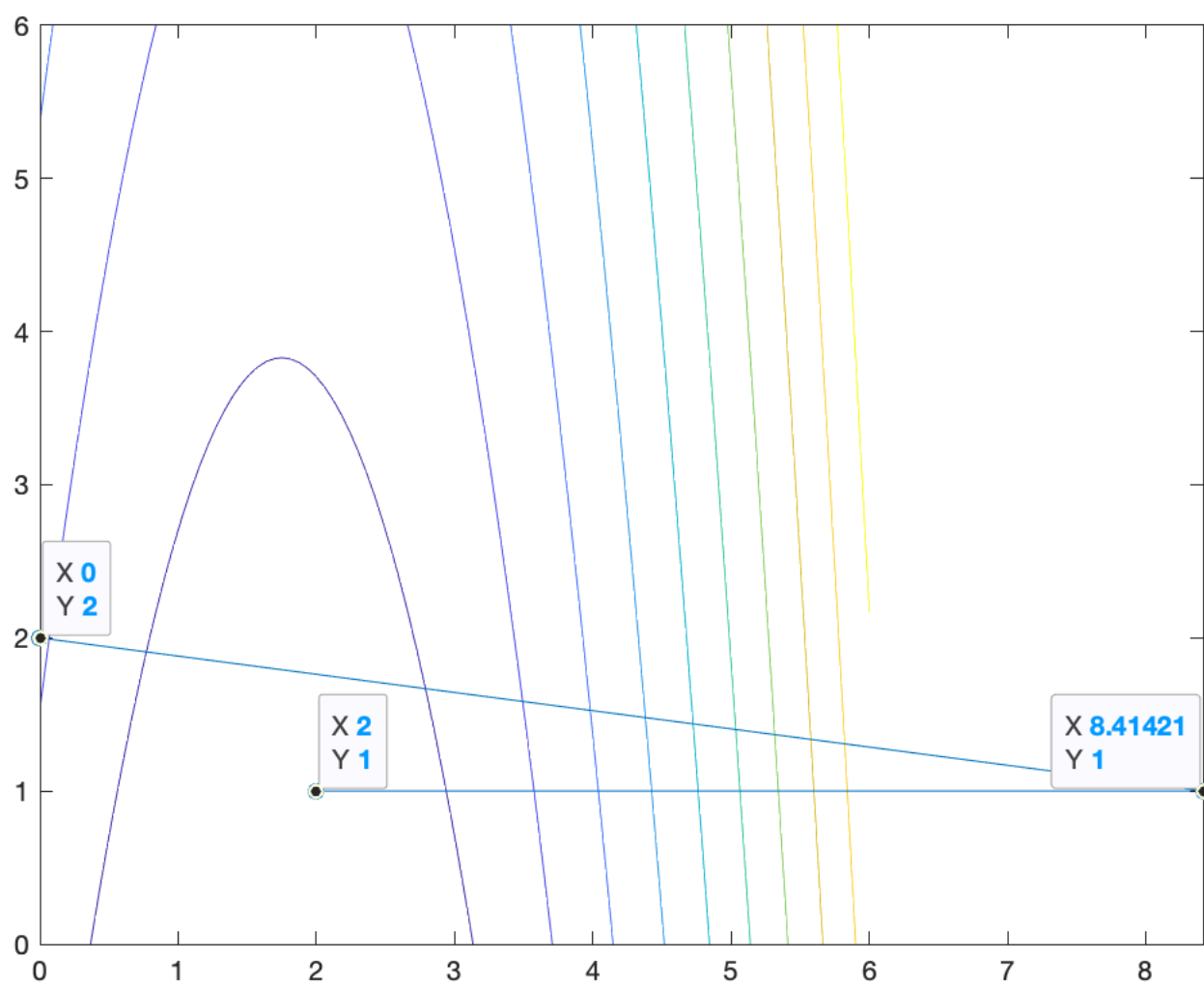


Figure 3: Newton Start (0;2)

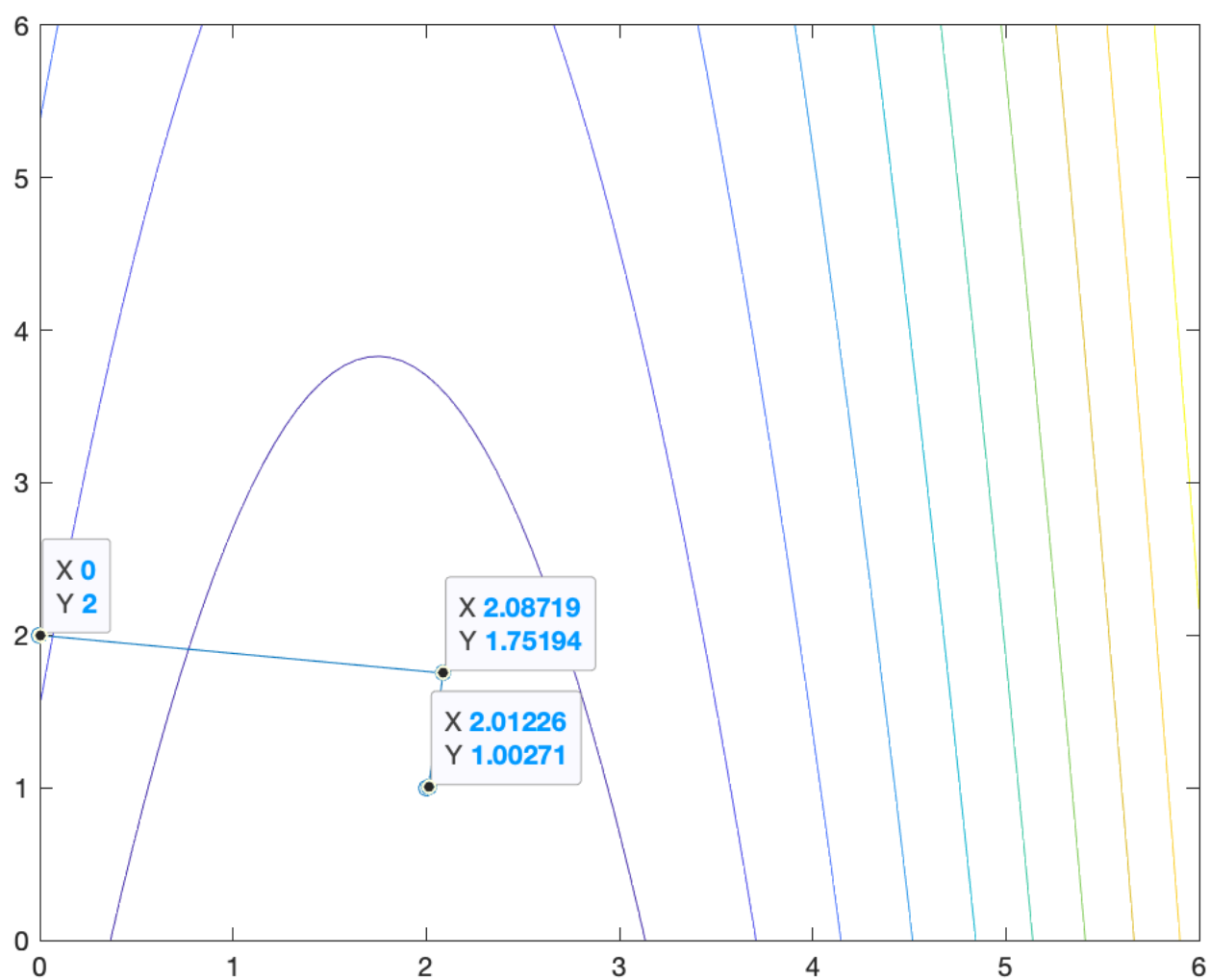


Figure 4: NewtonOneDimension Start (0;2)

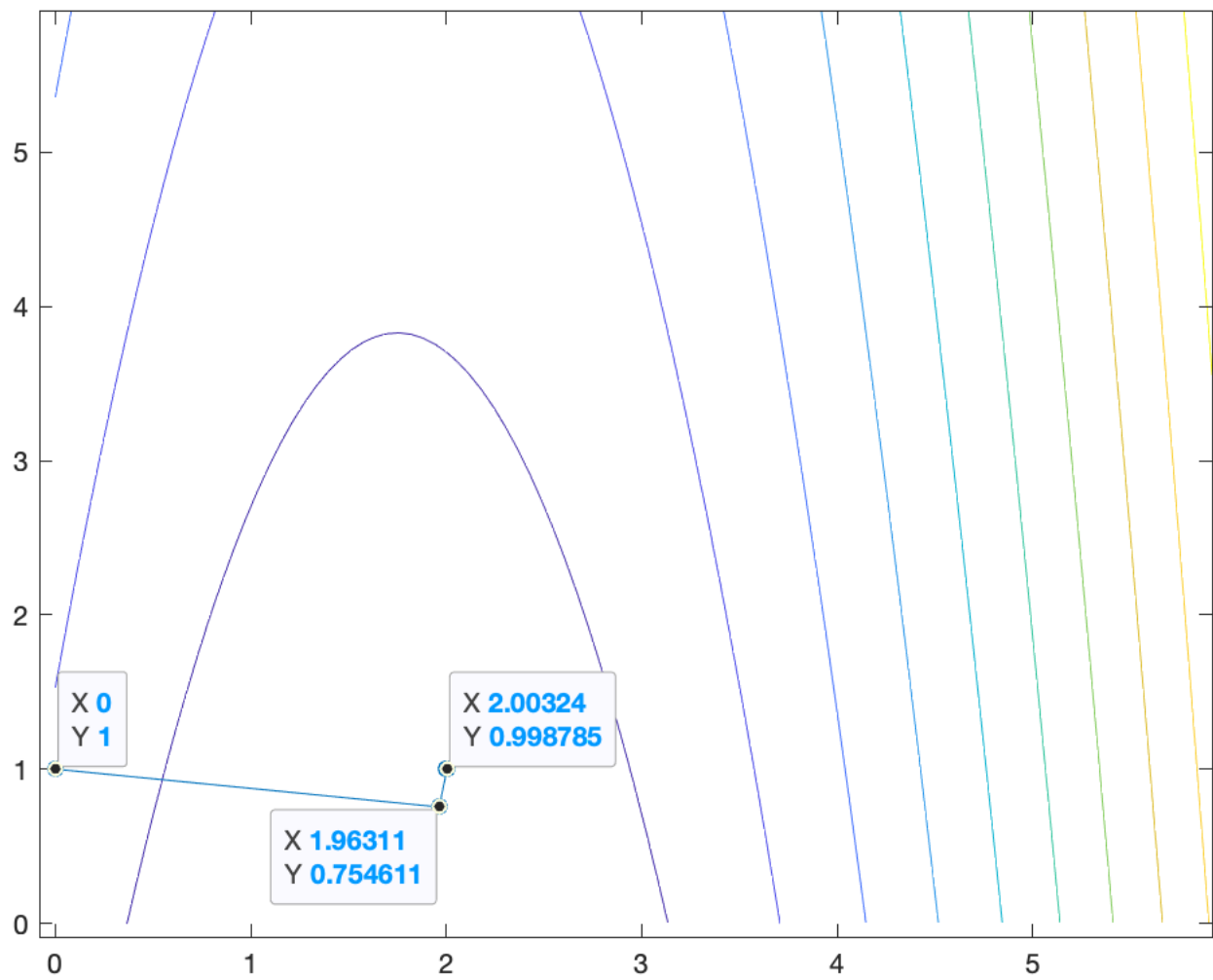


Figure 5: NewtonOneDimension Start (0;1)

$$f(x, y) = x^2 + y^2 + xy + x + y$$

$$f(x, y) \rightarrow \min = M\left(-\frac{1}{3}, -\frac{1}{3}\right); f\left(-\frac{1}{3}, -\frac{1}{3}\right) = -\frac{1}{3}$$

$$eps = 10^{-6}$$

Newton			Newton			Newton		
Start:	(-5; 5)		Start:	(1; 1)		Start:	(100; 200)	
Answer:	(-0.333333; -0.333333)		Answer:	(-0.333333; -0.333333)		Answer:	(-0.333333; -0.333333)	
Iterations:	2		Iterations:	2		Iterations:	2	
-----			-----			-----		
NewtonDirectionDescent			NewtonDirectionDescent			NewtonDirectionDescent		
Start:	(-5; 5)		Start:	(1; 1)		Start:	(100; 200)	
Answer:	(-0.333333; -0.333333)		Answer:	(-0.333333; -0.333333)		Answer:	(-0.333333; -0.333333)	
Iterations:	3		Iterations:	2		Iterations:	3	
-----			-----			-----		
NewtonOneDimension			NewtonOneDimension			NewtonOneDimension		
Start:	(-5; 5)		Start:	(1; 1)		Start:	(100; 200)	
one dimension	1		one dimension	1		one dimension	1	
one dimension	0,5683585		one dimension	0,00640181		one dimension	0,99576411	
Answer:	(-0.333333; -0.333333)		Answer:	(-0.333333; -0.333333)		Answer:	(-0.333333; -0.333333)	
Iterations:	2		Iterations:	2		Iterations:	2	
-----			-----			-----		

На квадратичной функции методы работают без нареканий. Каждый метод дал правильный ответ при самых разных начальных значениях. Кол-во итераций не зависит от начальных значений и всегда примерно равно одному и тому же значению.

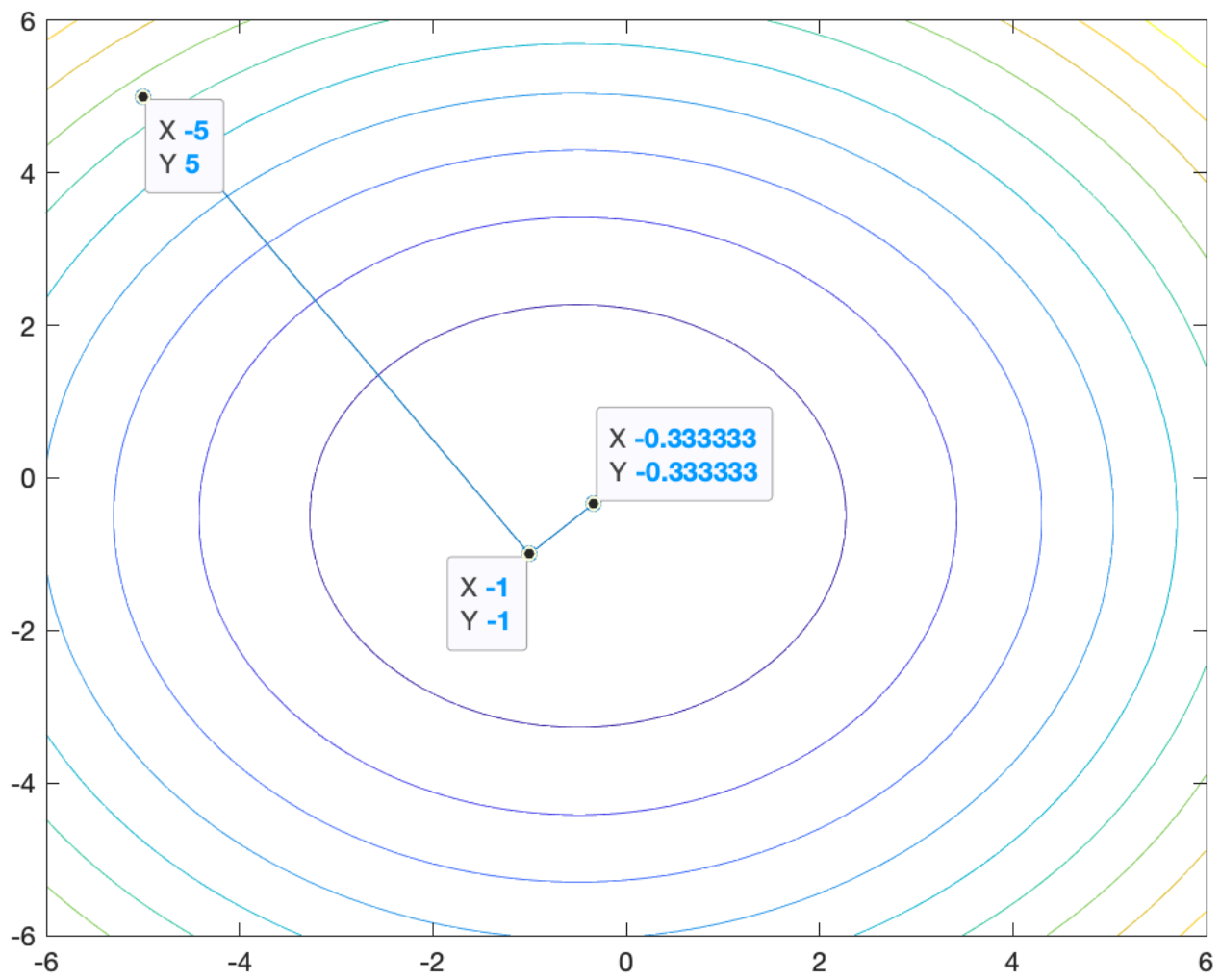


Figure 6: Newton Start $(-5;5)$

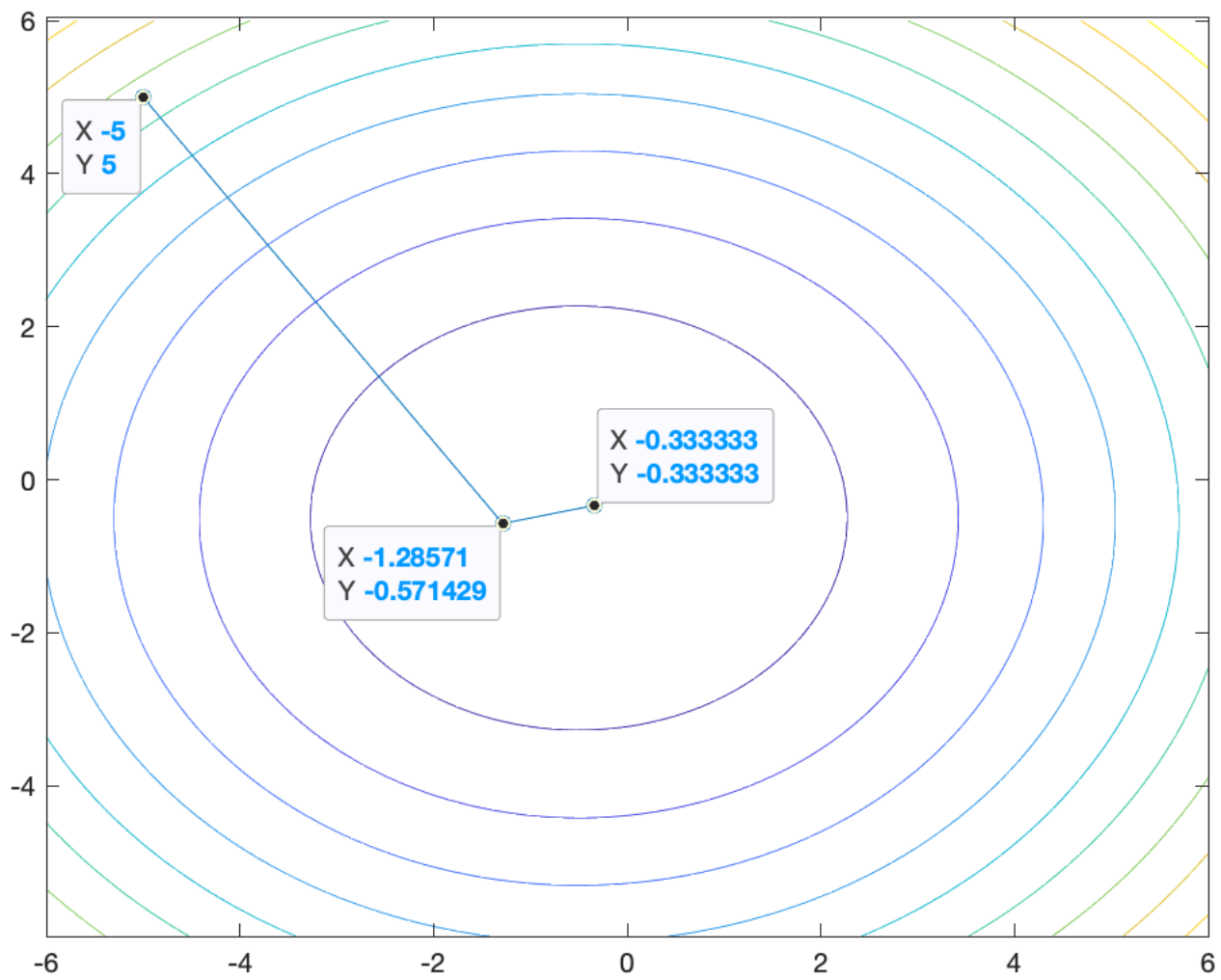


Figure 7: *NewtonOneDimension* & *NewtonDirectionDescent* Start (-5;5)

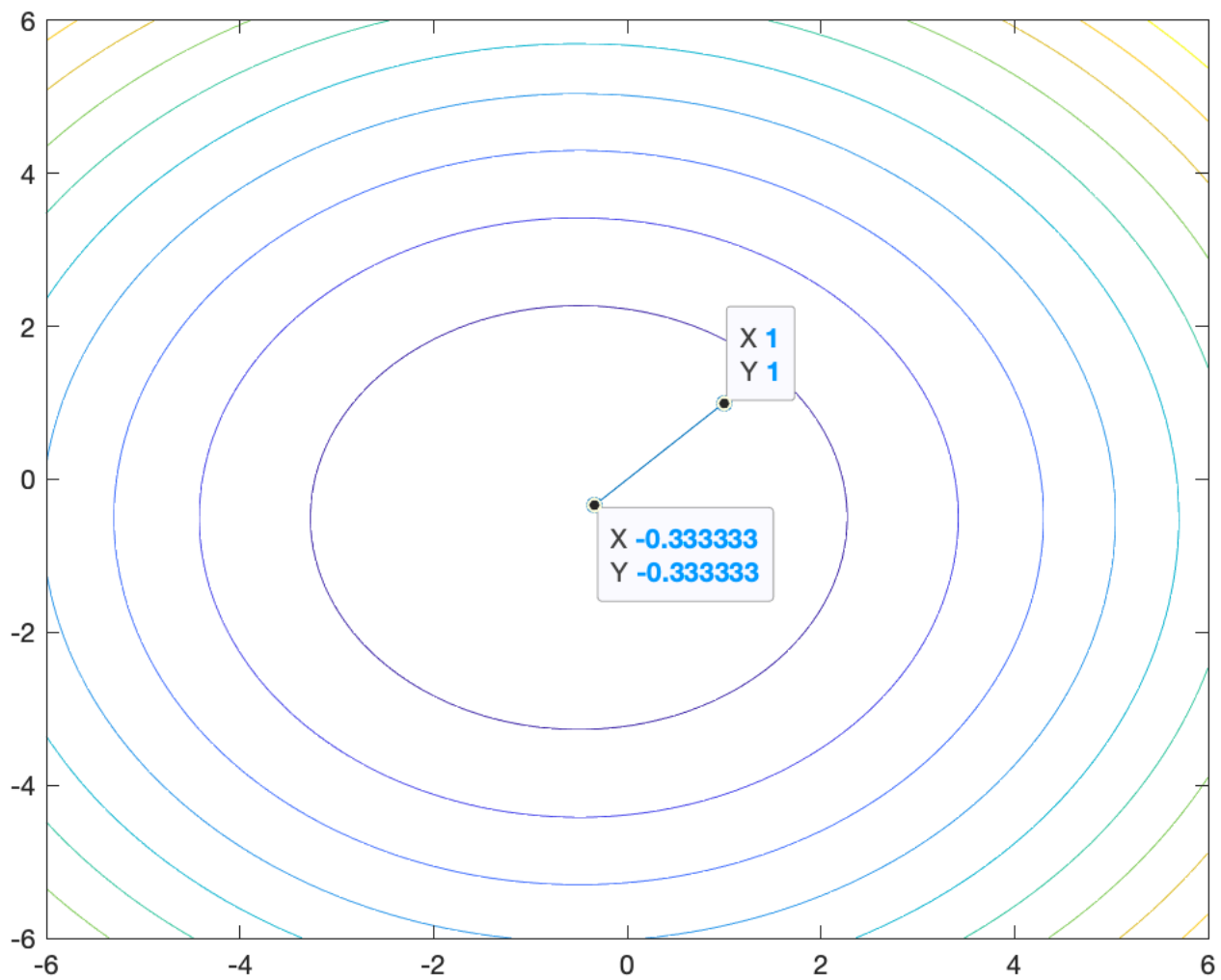


Figure 8: *Newton & NewtonOneDimension & NewtonDirectionDescent Start (-1;1)*

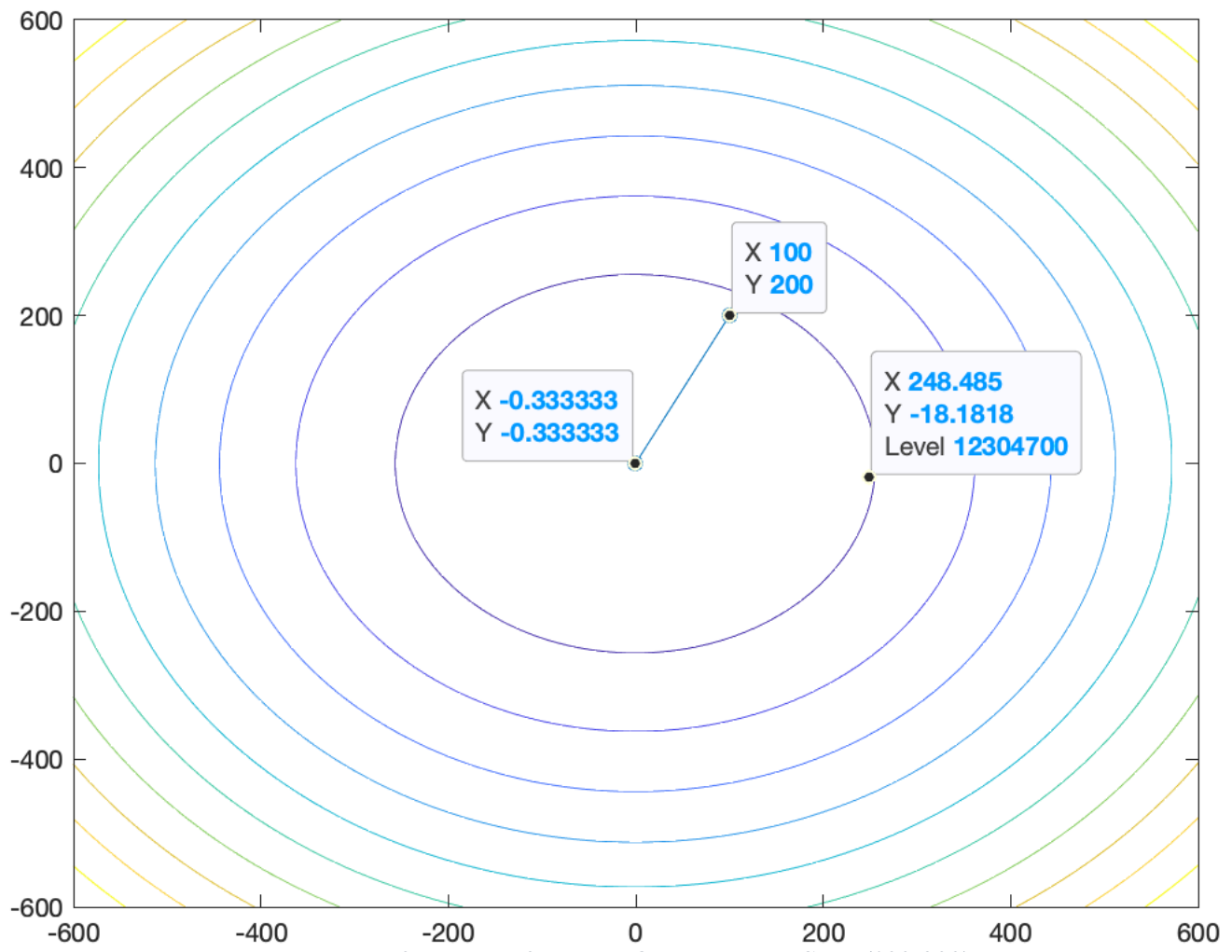


Figure 9: Newton & NewtonOneDimension Start (100;200)

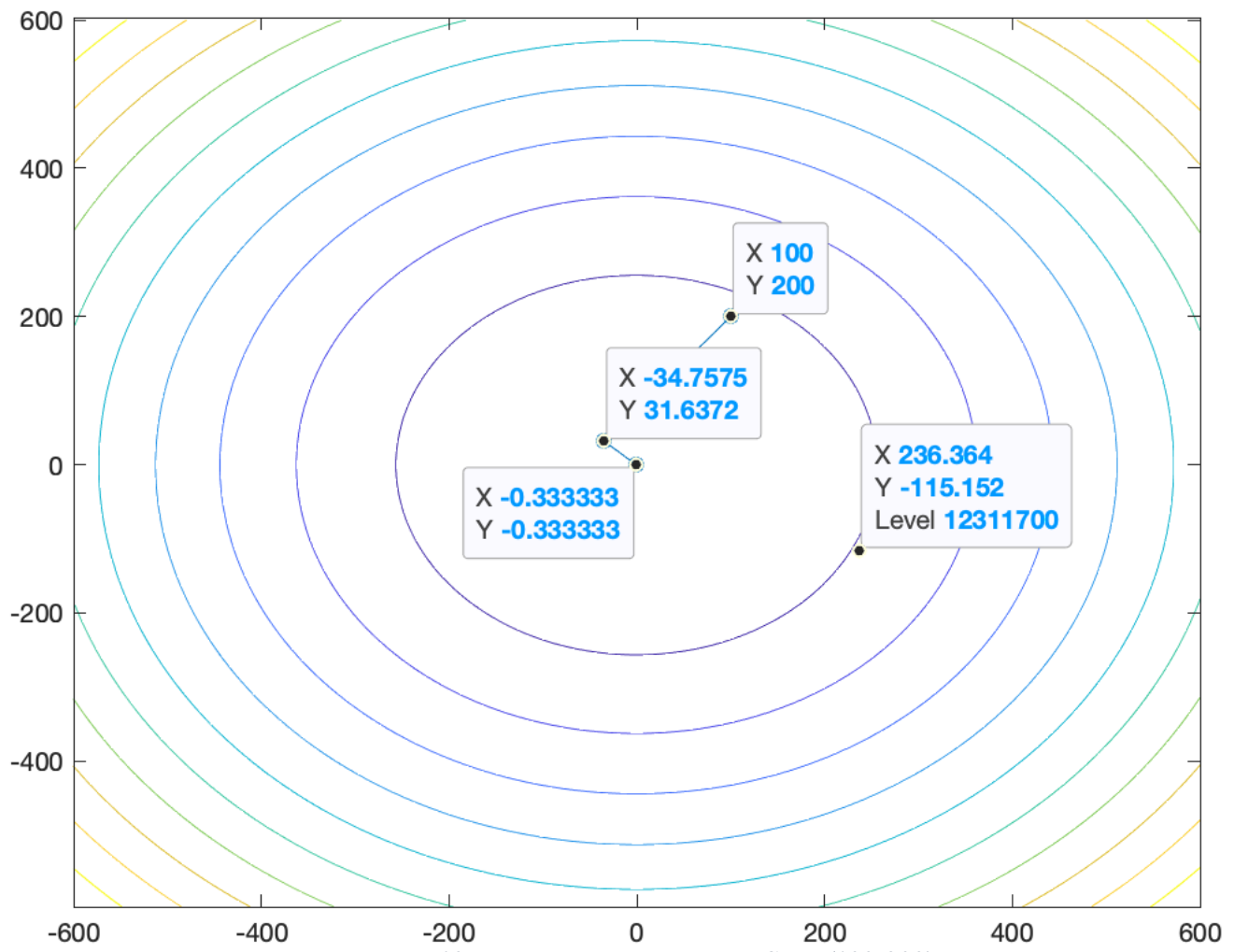


Figure 10: NewtonDirectionDescent Start (100;200)

$$f(x, y) = x^2 + y^2 - 1.2xy$$

$$f(x, y) \rightarrow \min = M(0,0); f(0,0) = 0$$

$$eps = 10^{-6}$$

$$f(x, y) = 100(y - x^2)^2 + (1 - x)^2$$

$$f(x, y) \rightarrow \min = M(1,1); f(1,1) = 0$$

$$eps = 10^{-6}$$

Method	f1	f2
Newton	2	7
NewtonDirectionDescent	3	24
NewtonOneDimension	2	39
Steepest Descent	25	69

Newton				Newton		
Start:	(4; 1)			Start:	(-1.2; 1)	
Answer:	(0; -2.22045e-16)			Answer:	(1; 1)	
Iterations:	2			Iterations:	7	
-----				-----		
NewtonDirectionDescent				NewtonDirectionDescent		
Start:	(4; 1)			Start:	(-1.2; 1)	
Answer:	(1.62628e-17; 2.00448e-17)			Answer:	(1; 1)	
Iterations:	3			Iterations:	24	
-----				-----		
NewtonOneDimension				NewtonOneDimension		
Start:	(4; 1)			Start:	(-1.2; 1)	
one dimension:	1			one dimension: a =	719,069511	
one dimension:	1			one dimension: a =	-0,0208679	
Answer:	(4.0409e-17; 1.01022e-17)			one dimension: a =	1,25885112	
Iterations:	2			one dimension: a =	2,3053139	
-----				one dimension: a =	1,90430532	
				one dimension: a =	2,12172273	
				one dimension: a =	2,01026127	
				one dimension: a =	2,07268785	
				one dimension: a =	2,04221656	
				one dimension: a =	2,06156082	
				one dimension: a =	2,05442034	
				one dimension: a =	2,0617812	
				one dimension: a =	2,06158528	
				one dimension: a =	2,0658466	
				one dimension: a =	2,06806551	
				one dimension: a =	2,07180673	
				one dimension: a =	2,07519902	
				one dimension: a =	2,07938277	
				one dimension: a =	2,08380967	
				one dimension: a =	2,08894446	
				one dimension: a =	2,09462272	
				one dimension: a =	2,1011333	
				one dimension: a =	2,10862861	
				one dimension: a =	2,11732205	
				one dimension: a =	2,12751728	
				one dimension: a =	2,1397349	
				one dimension: a =	2,15456073	
				one dimension: a =	2,17293901	
				one dimension: a =	2,19638059	
				one dimension: a =	2,22730254	
				one dimension: a =	2,26998938	
				one dimension: a =	2,33275842	
				one dimension: a =	2,43368603	
				one dimension: a =	2,61394628	
				one dimension: a =	2,03911429	
				one dimension: a =	0,88488423	
				one dimension: a =	0,97513353	
				one dimension: a =	1,0018741	
				one dimension: a =	-3,1100769	
				Answer:	(1; 1)	
				Iterations:	39	
-----				-----		

Все методы нашли верное решение задачи. Стоит заметить, что методы работают лучше на квадратичных функциях. Метод наискорейшего спуска показал себя с наихудшей стороны.

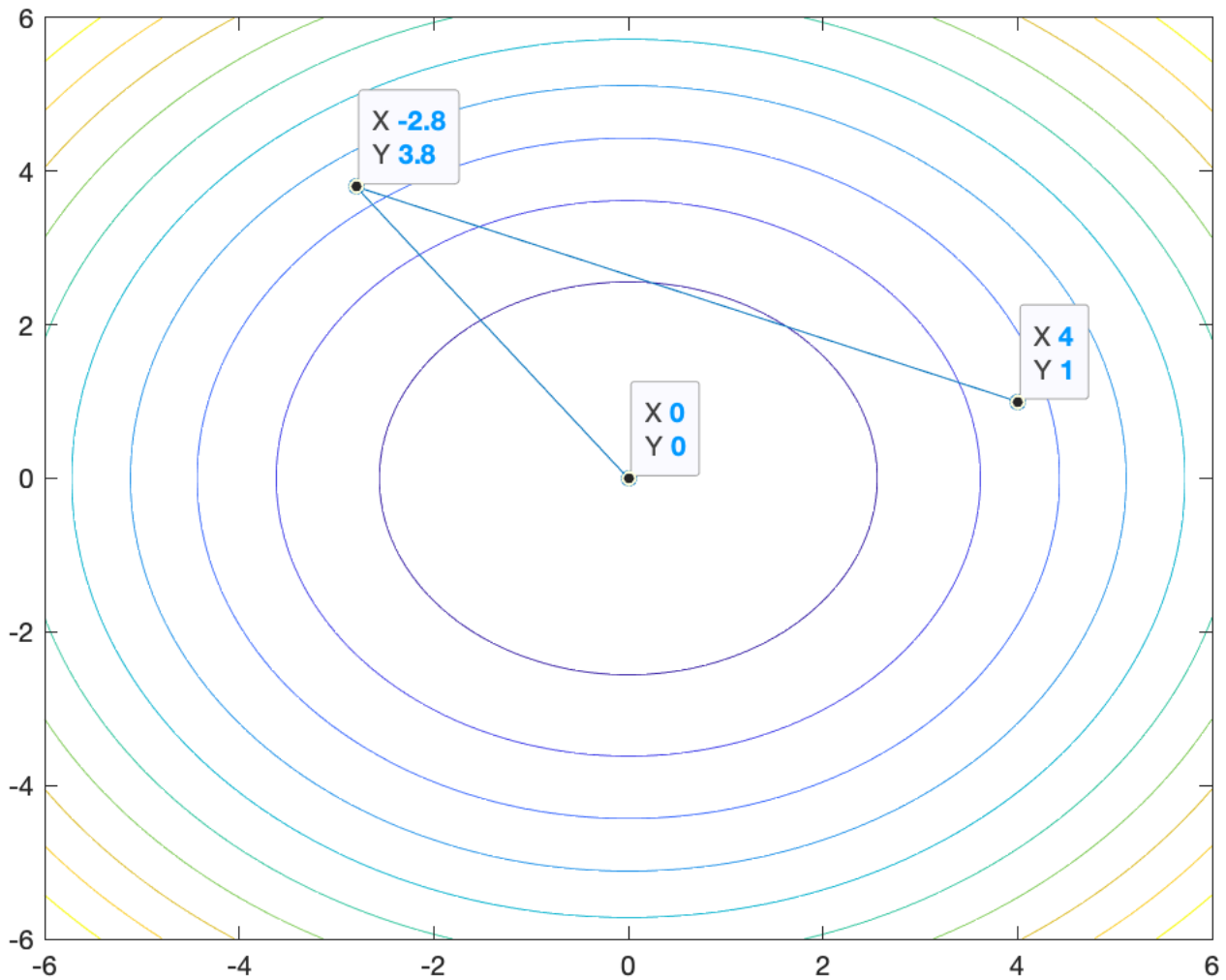


Figure 11: *Func1 Newton Start (4;1)*

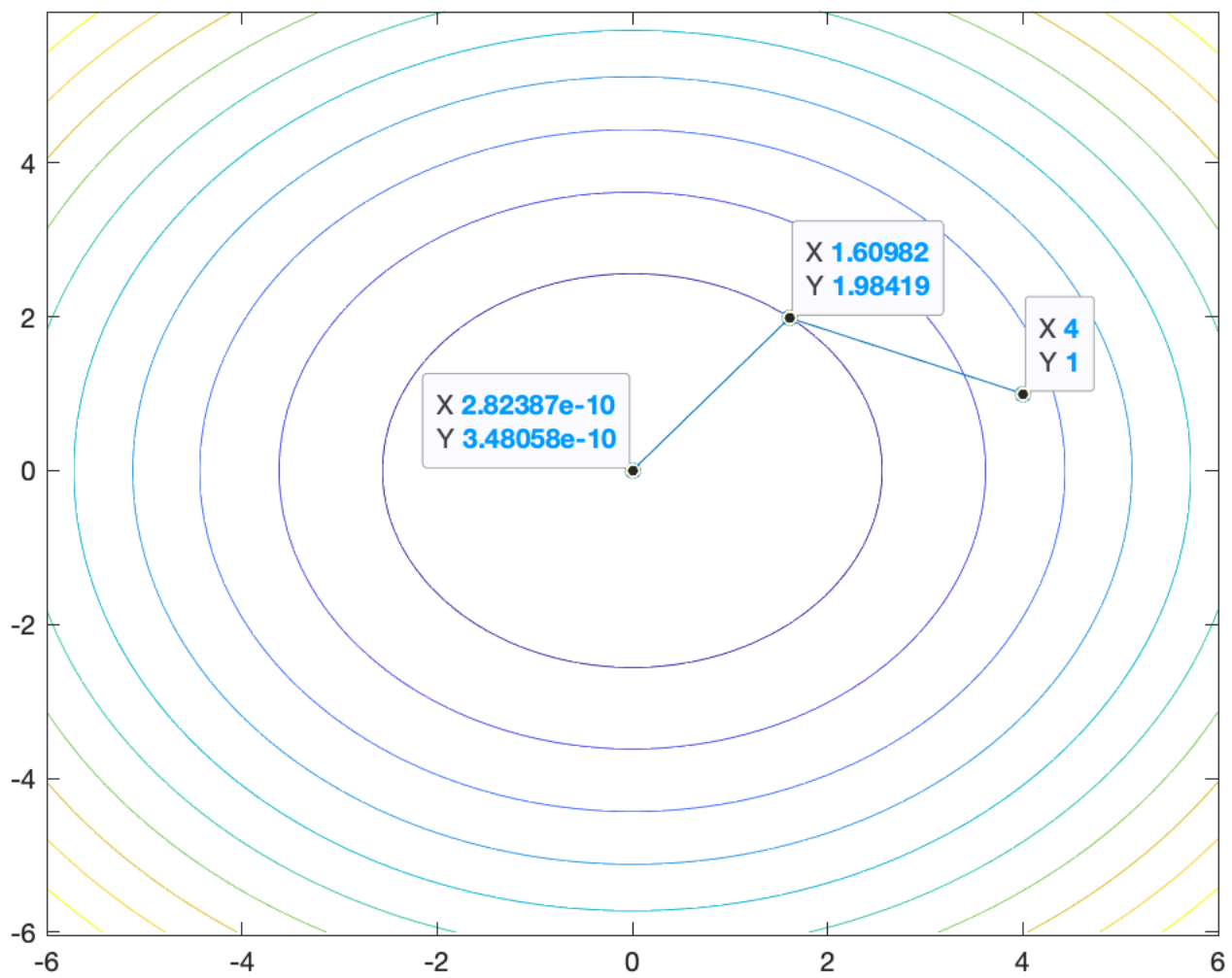


Figure 12: *Func1* NewtonDirectionDescent Start (4;1)

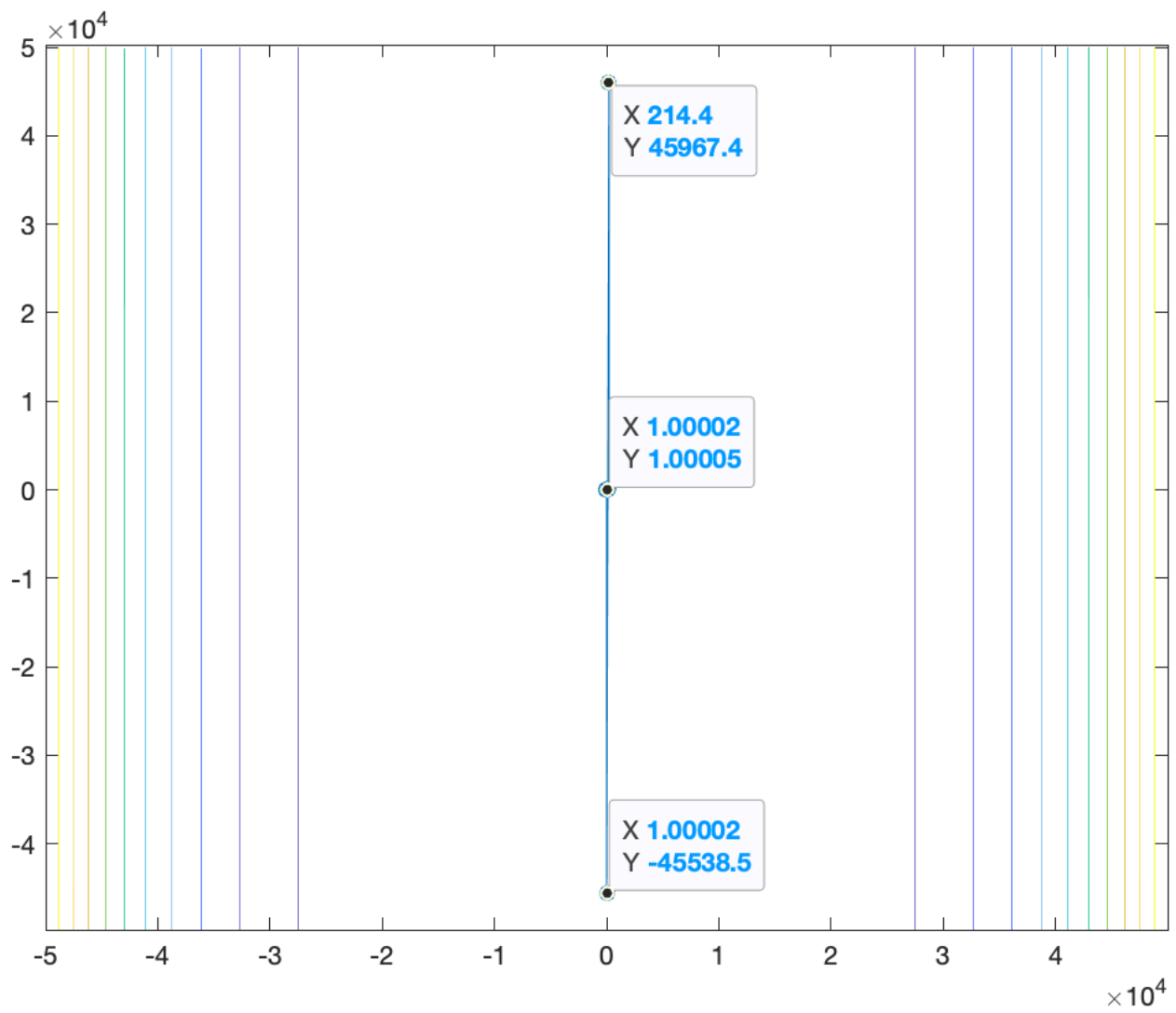


Figure 13: *Func2* Newton Start (-1.2;1)

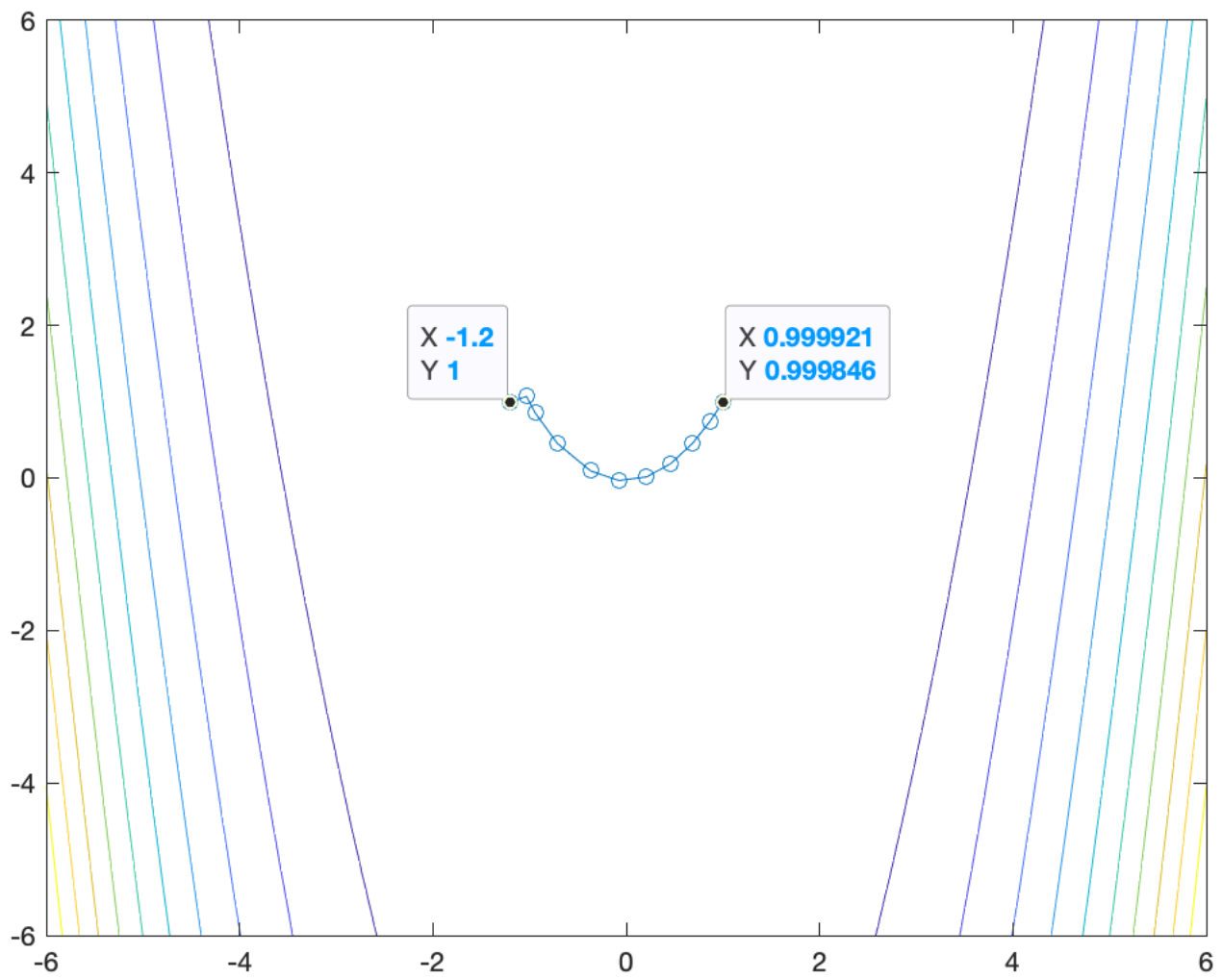


Figure 14: Func2 NewtonDirectionDescent Start (-1.2;1)

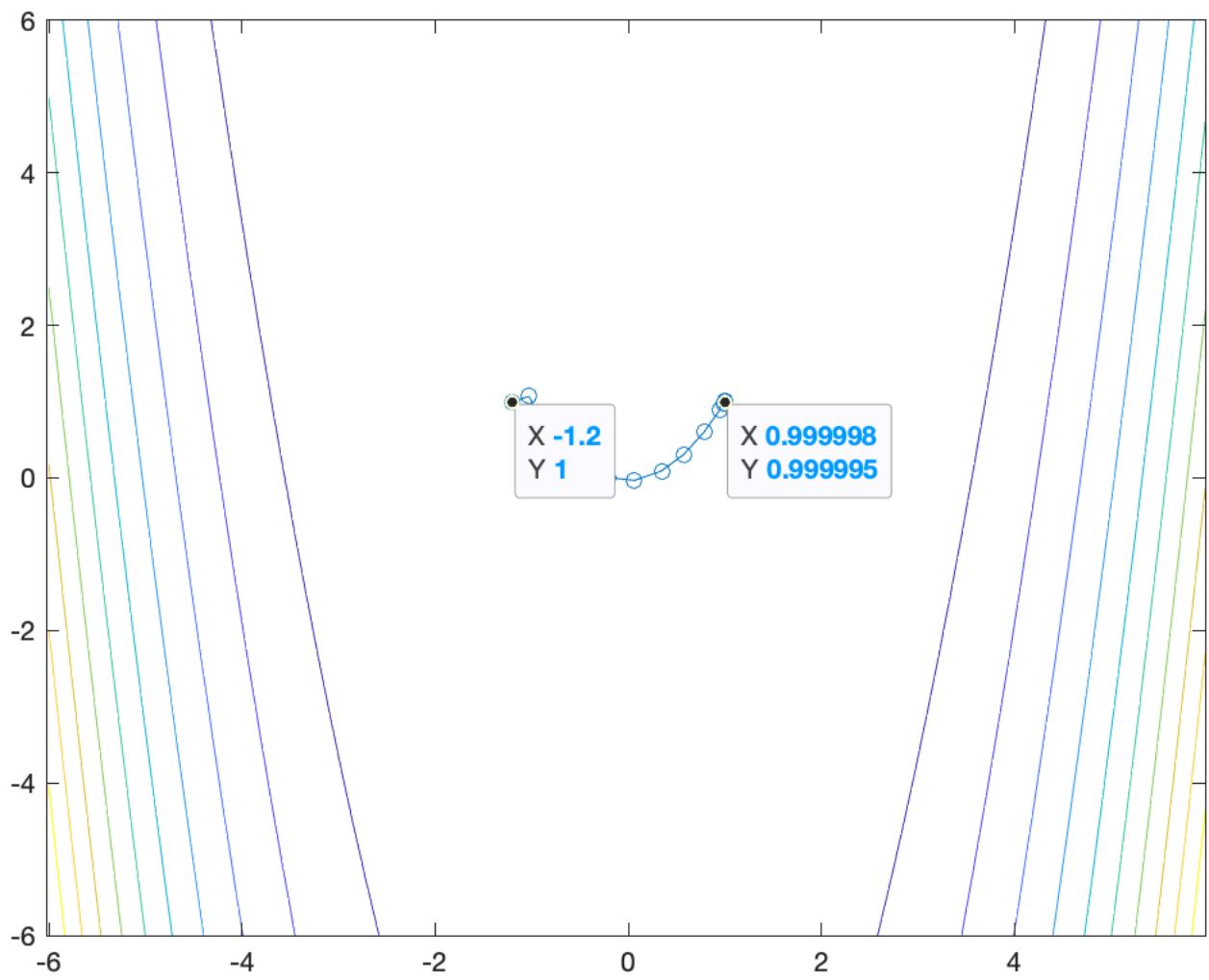


Figure 15: Func2 NewtonOneDimension Start (-1.2;1)

Результаты исследований 2

$f(x,y) = 100(y - x^2)^2 + (1 - x)^2$
 $f(x,y) \rightarrow \min = M(1,1); f(1,1) = 0$
 $eps = 10^{-6}$

BroadenFletcherChen			Newton	
Start:	(-1.2; 1)		Start:	(0.5; 0.5)
Answer:	(1; 1)		Answer:	(1; 1)
Iterations:	13		Iterations:	6
-----			-----	
Powell				
Start:	(-1.2; 1)			
Answer:	(1; 1)			
Iterations:	19			

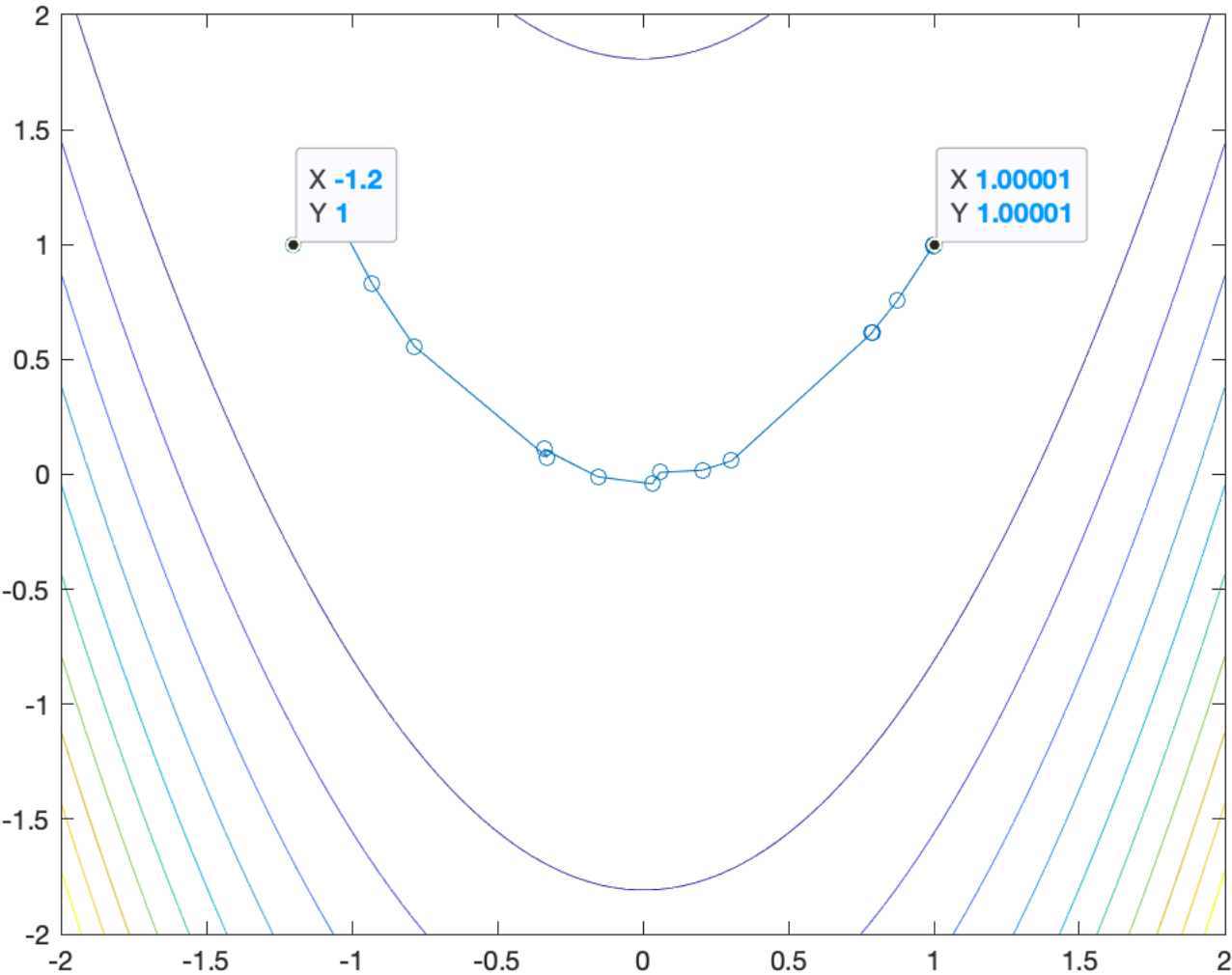


Figure 16: BroadenFletcherChen

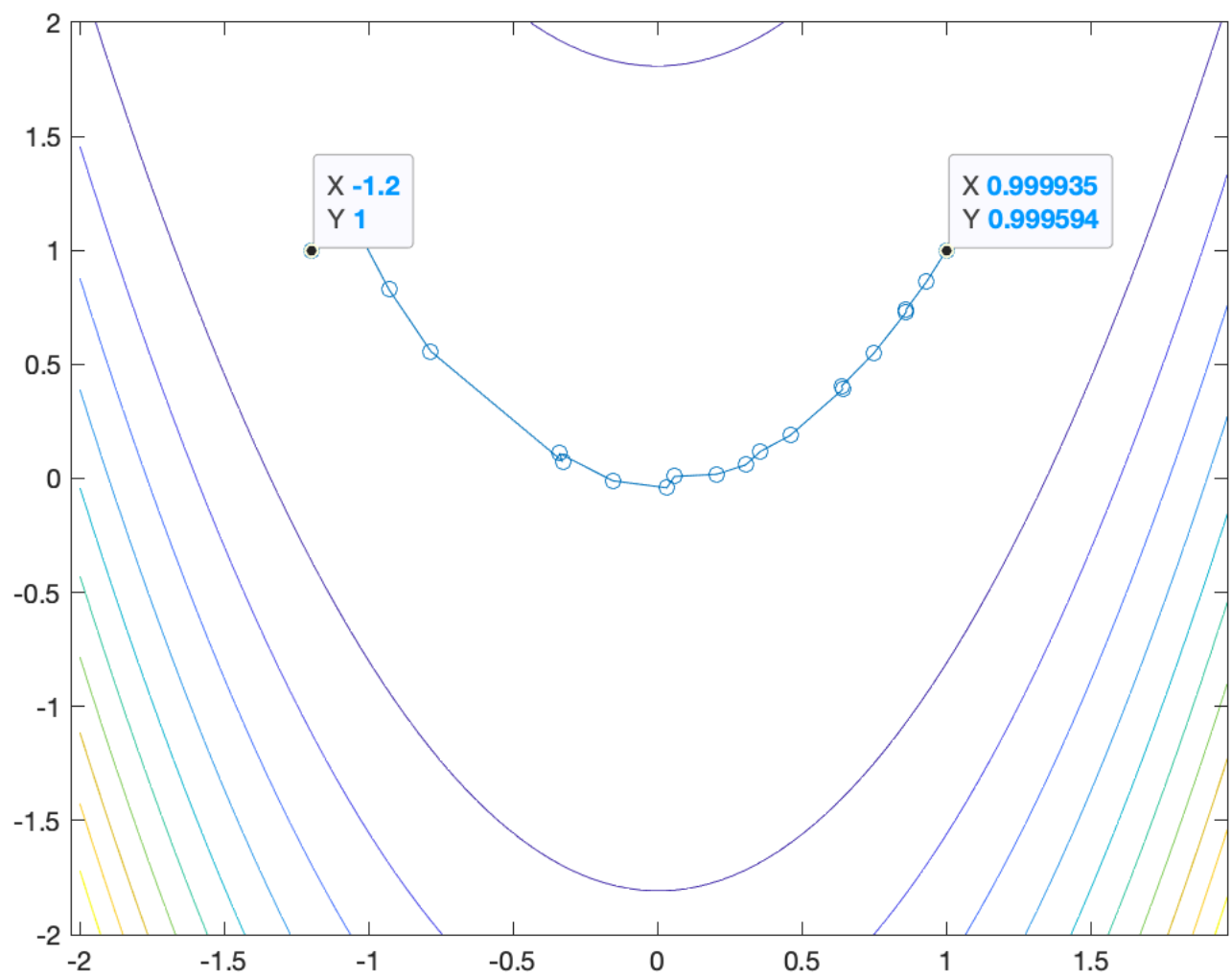


Figure 17: Powell

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

глобальных экстремумов нет
 $eps = 10^{-6}$

BroadenFletcherChen			Newton		
Start:	(1; 1)		Start:	(1; 1)	
Answer:	(3; 2)		Answer:	(-3.77931; -3.28319)	
Iterations:	6		Iterations:	8	
-----			-----		
Powell					
Start:	(1; 1)				
Answer:	(3; 2)				
Iterations:	6				

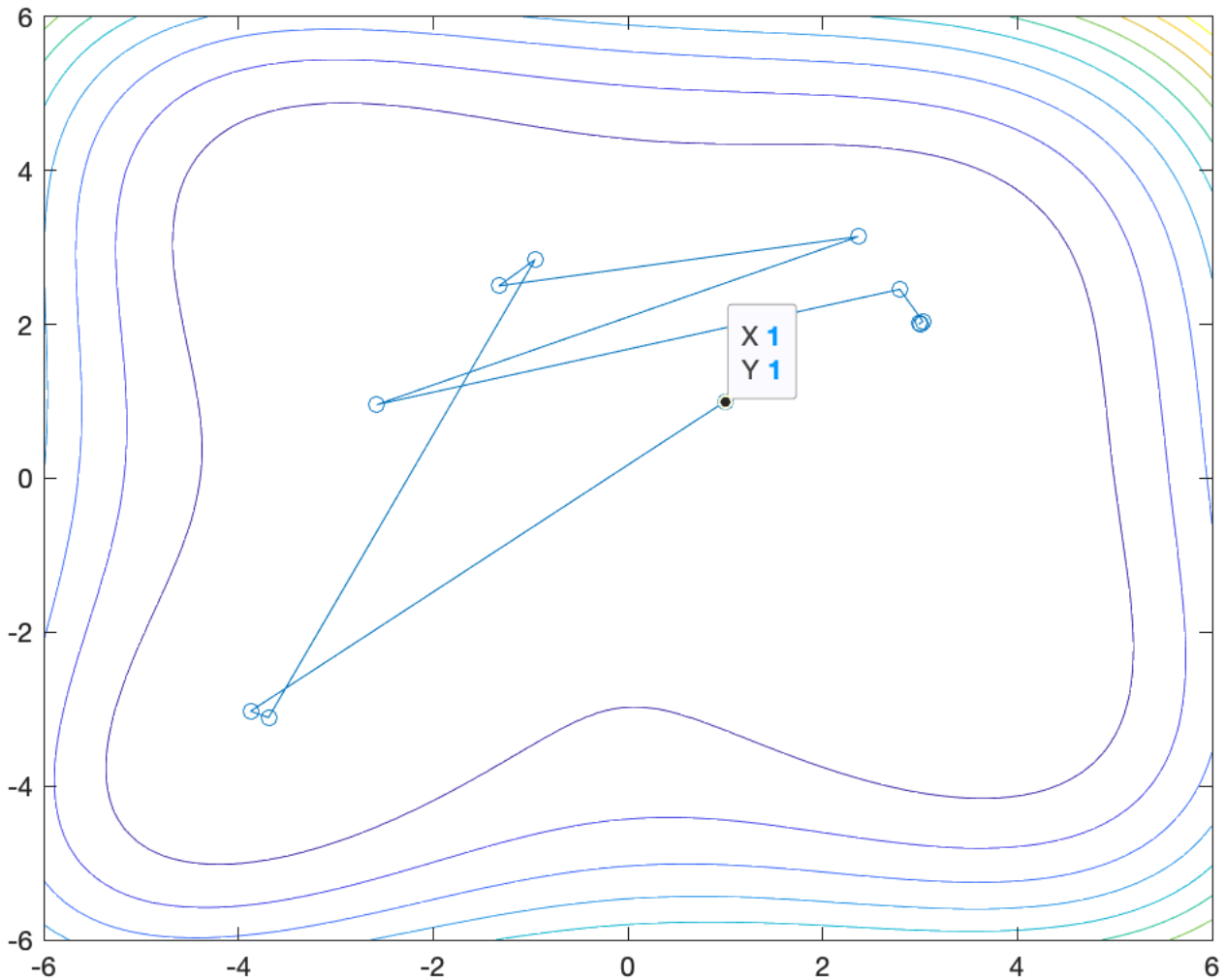


Figure 18: BroadenFletcherChen

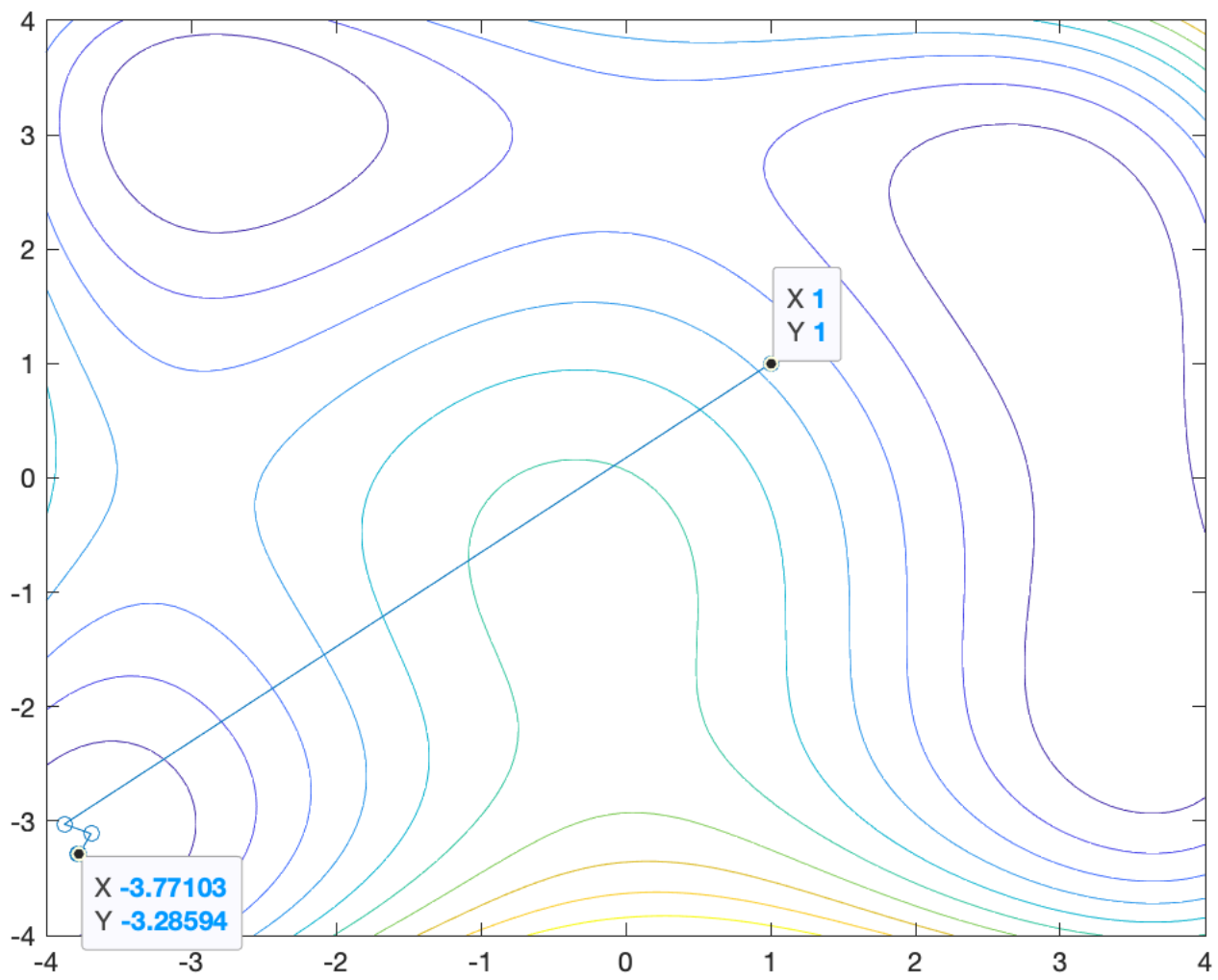


Figure 19: Powell

$$f(x, y, z, w) = (x + 10y)^2 + 5(z - w)^2 + (y - 2z)^4 + 10(x - w)^4$$

$$f(x, y) \rightarrow \min = M(0,0,0,0); f(0,0) = 0$$

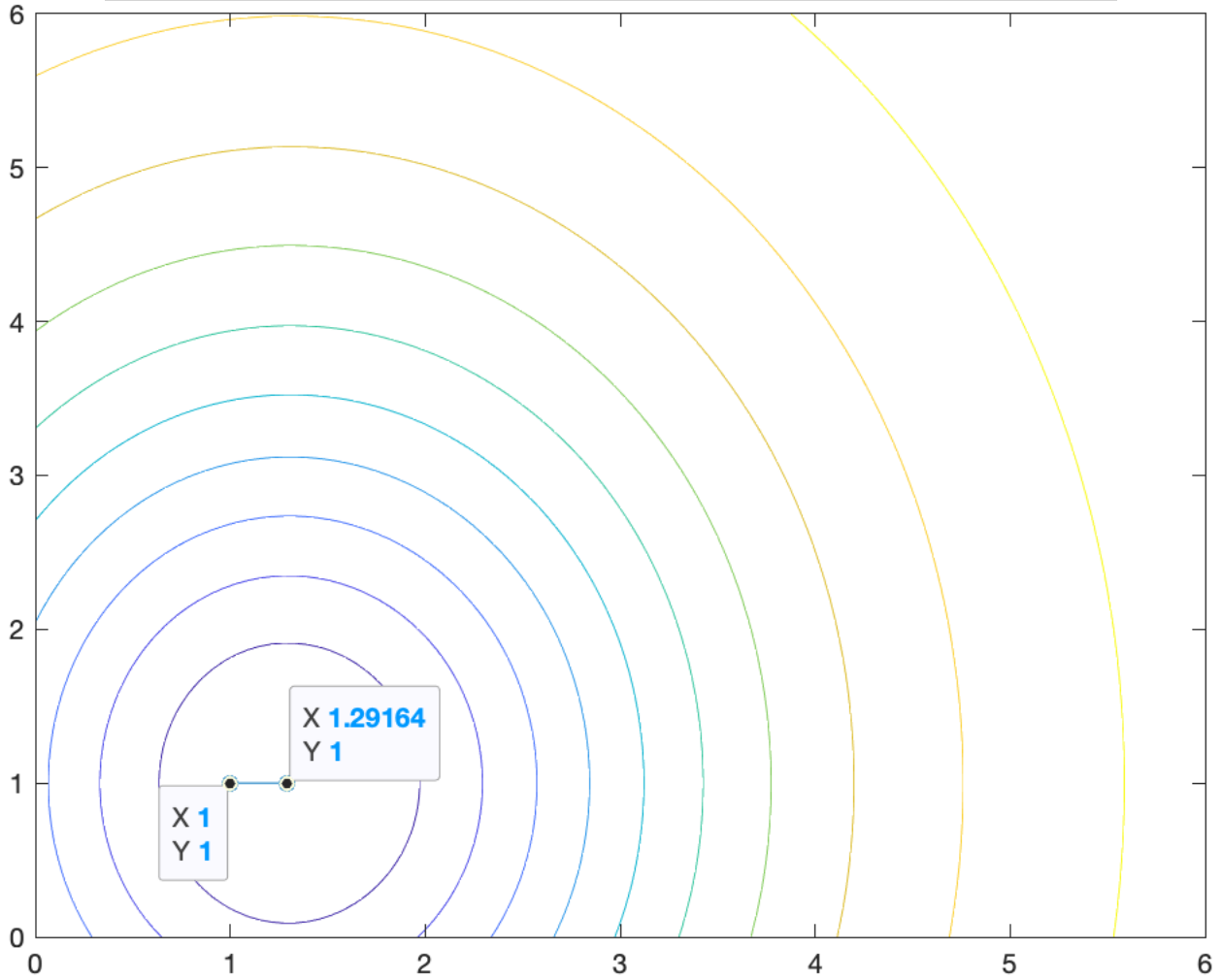
$$eps = 10^{-6}$$

BroadenFletcherChen				Newton			
Answer x:	(0.000229535; -2.29568e-05; -5.24201e-05; -5.24197e-05)			Answer x:	(7.70154e-05; -7.70154e-06; 9.96785e-05; 9.96785e-05)		
Answer y:	6,433E-14			Answer y:	1,8408E-15		
Iterations:	23			Iterations:	35		
-----				-----			
Powell							
Answer x:	(-0.00860751; 0.000860726; -0.00323133; -0.00323133)						
Answer y:	1,123E-08						
Iterations:	13						

$$f(x,y,z,w) = 100 - \frac{1}{1 + \frac{(-2+x)^2}{4} + \frac{(-1+y)^2}{9}} - \frac{2}{1 + \frac{(-1+x)^2}{4} + \frac{(-1+y)^2}{9}}$$

$f(x,y) \rightarrow \min = M(1,29,1); f(1,1) = 97.1531029$
 $eps = 10^{-6}$

BroadenFletcherChen			Newton		
Start:	(1; 1)		Start:	(1; 1)	
Answer:	(1.29164; 1)		Answer:	(1.29164; 1)	
Iterations:	1		Iterations:	4	
-----			-----		
Powell					
Start:	(1; 1)				
Answer:	(1.29164; 1)				
Iterations:	1				



Зависимость от начальной точки:

func1

BroadenFletcherChen			BroadenFletcherChen			BroadenFletcherChen		
Start:	(-1.2; 1)		Start:	(-50; 50)		Start:	(-0.1; 0.1)	
Answer:	(1; 1)		Answer:	(1; 1)		Answer:	(1; 1)	
Iterations:	13		Iterations:	9		Iterations:	39	
-----			-----			-----		
Powell			Powell			Powell		
Start:	(-1.2; 1)		Start:	(-50; 50)		Start:	(-0.1; 0.1)	
Answer:	(1; 1)		Answer:	(1; 1)		Answer:	(1; 1)	
Iterations:	19		Iterations:	9		Iterations:	250	
-----			-----			-----		

func3

BroadenFletcherChen			BroadenFletcherChen			BroadenFletcherChen		
Answer x:	(0.000229535; -2.29568e-05; -5.24201e-05; -5.24197e-05)		Start:	(-50; 50; -50; 50)		Start:	(-0.1; 0.1; -0.1; 0.1)	
Answer y:	6.433E-14		Answer:	(3.27491e-06; -3.27489e-07; -2.05985e-05; -2.05985e-05)		Answer:	(1.52352e-05; -1.52352e-06; -7.97688e-05; -7.97688e-05)	
Iterations:	23		Iterations:	37		Iterations:	18	
-----			-----			-----		
Powell			Powell			Powell		
Answer x:	(-0.00860751; 0.000860726; -0.00323133; -0.00323133)		Start:	(-50; 50; -50; 50)		Start:	(-0.1; 0.1; -0.1; 0.1)	
Answer y:	1.123E-08		Answer:	(3.83142e-07; -3.83144e-08; 9.79578e-06; 9.79578e-06)		Answer:	(1.97263e-05; -1.97264e-06; 8.31483e-06; 8.31483e-06)	
Iterations:	13		Iterations:	43		Iterations:	31	
-----			-----			-----		

func4

BroadenFletcherChen			BroadenFletcherChen			BroadenFletcherChen		
Start:	(1; 1)		Start:	(-50; 50)		Start:	(-0.1; 0.1)	
Answer:	(1.29164; 1)		Answer:	(-1.26627e+08; -7.20058e+08)		Answer:	(1.29164; 1)	
Iterations:	1		Iterations:	5		Iterations:	4	
-----			-----			-----		
Powell			Powell			Powell		
Start:	(1; 1)		Start:	(-50; 50)		Start:	(-0.1; 0.1)	
Answer:	(1.29164; 1)		Answer:	(1.29164; 1)		Answer:	(1.29164; 1)	
Iterations:	1		Iterations:	8		Iterations:	4	
-----			-----			-----		

Выводы

В ходе исследования были рассмотрены различные ньютоновские и квазиньютоновские методы, проведены сравнения и анализ методов на различных функциях в результате чего были выделены следующие преимущества и недостатки каждого из них:

Преимущества методов Ньютона:

1. Квадратичная скорость сходимости
2. Глобальная сходимость гарантируется в методе с одномерным поиском и направлением спуска

Недостатки методов Ньютона:

1. Серьезные ограничения на функцию
2. Сложность вычисления Гессиана и решения СЛАУ
3. Зависимость от начальной точки для неквадратичных функций

Преимущества квазиньютоновских методов:

1. Быстрота вычисления направления спуска благодаря приближению, а не точному вычислению обратного гессиана
2. Сверхлинейная скорость сходимости хуже, чем у методов Ньютона, однако, если $H(x)$ удовлетворяет условию Липшица, то достигается также квадратичная сходимость

Недостатки квазиньютоновских методов:

1. $O(n^2)$ дополнительной памяти на каждой итерации

Исходный код: <https://github.com/NelosG/Optimization-Methods>