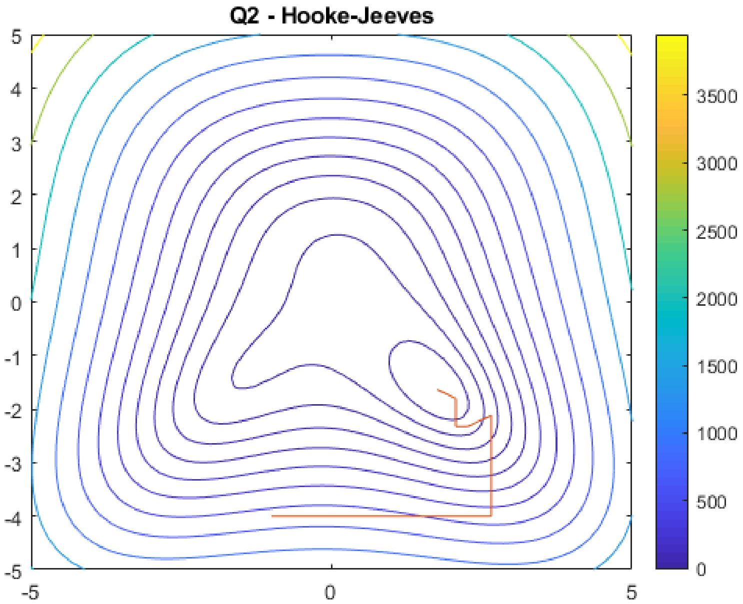
Question 1

Note: all functions take an additional optional argument that is used for question 2. **This additional argument must be left empty when testing question 1**.

Question 2

Hooke-Jeeves:



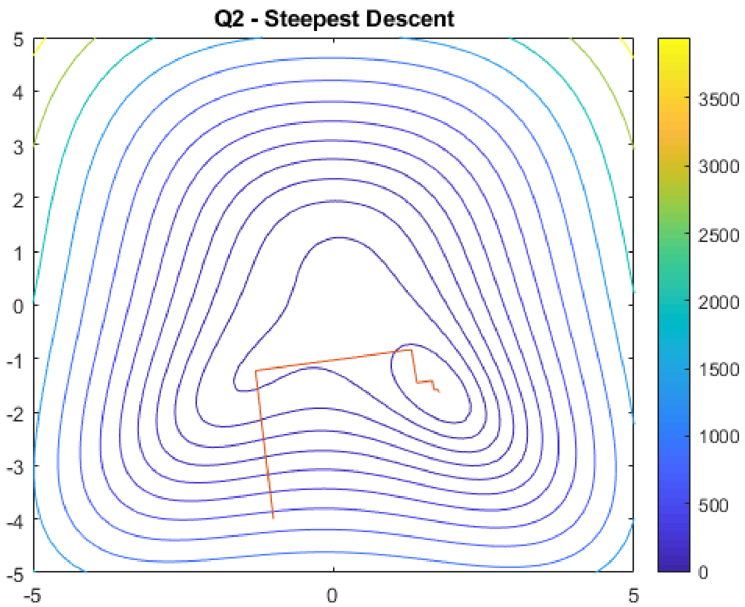
Iteration points:
Each column is an iteration point, row 1 containing x_1 and row 2 containing x_2.

-1	2.6597	2.6597	2.2564	2.0654	2.0654	1.7628	1.7676	1.7676	1.7622	1.7623	1.7623
-4	-4	-2.1185	-2.3258	-2.3258	-1.8017	-1.6405	-1.6405	-1.6321	-1.6290	-1.6290	-1.6290

Objective functions:
Each column is an iteration point.

488	344.5883	15.9791	-3.9047	-6.0348	-18.9607	-22.1212	-22.1221	-22.1243	-22.1252	-22.1252	-22.1252

Steepest descent:



Iteration points:
Each column is an iteration point, row 1 containing x_1 and row 2 containing x_2.

-1	-1.3029	1.2942	1.3873	1.6412	1.6675	1.7333	1.7397	1.7555	1.7570	1.7606	1.7610
-4	-1.2244	-0.8331	-1.4512	-1.4130	-1.5868	-1.5769	-1.6191	-1.6167	-1.6266	-1.6261	-1.6284

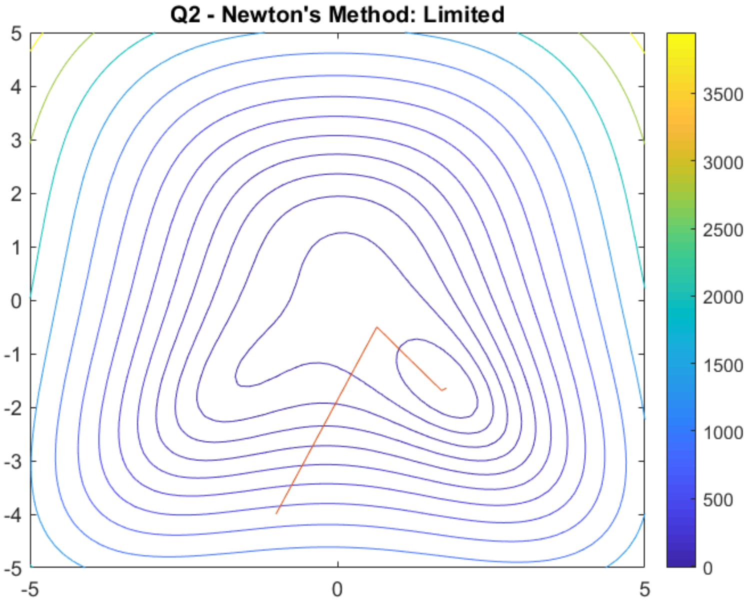
Gradients:
Each column is an iteration point, row 1 is in the direction of x_1 and row 2 is in the direction of x_2.

55	-2.0846	-2.2163	-16.1260	-1.1565	-5.4549	-0.3252	-1.3864	-0.0784	-0.3291	-0.0183	-0.0770
-504	-0.3141	14.7129	-2.4277	7.6506	-0.8233	2.1446	-0.2097	0.5208	-0.0495	0.1225	-0.0115

Objective functions:
Each column is an iteration point.

488	2.0602	-13.2015	-18.9107	-21.1655	-21.8768	-22.0637	-22.1105	-22.1217	-22.1244	-22.1250	-22.1251

Newton's Method:



Iteration points:
Each column is an iteration point, row 1 containing x_1 and row 2 containing x_2.

-1	0.6402	1.6933	1.7726
-4	-0.5030	-1.6892	-1.6430

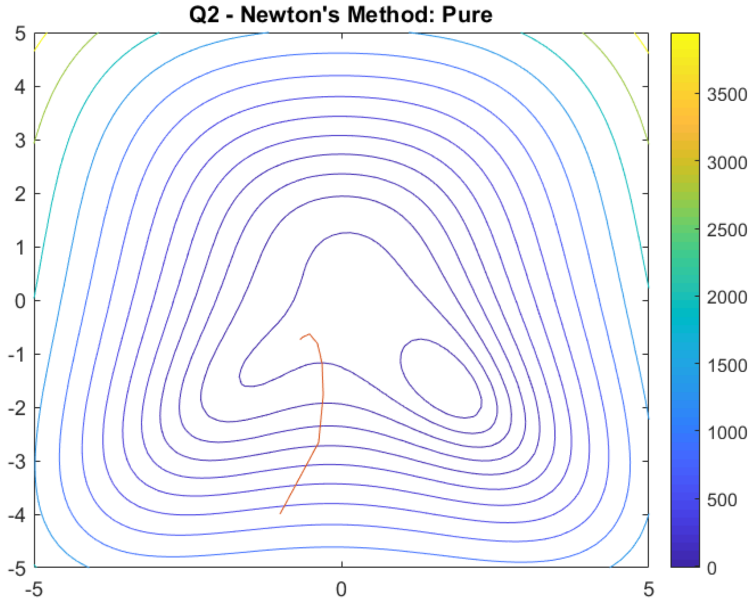
Gradients:
Each column is an iteration point, row 1 is in the direction of x_1 and row 2 is in the direction of x_2.

55	-9.2982	-7.3209	0.3014
-504	4.3612	-6.4995	-0.5174

Objective functions:
Each column is an iteration point.

488	-5.2753	-21.6738	-22.1200

Pure Newton's Method:

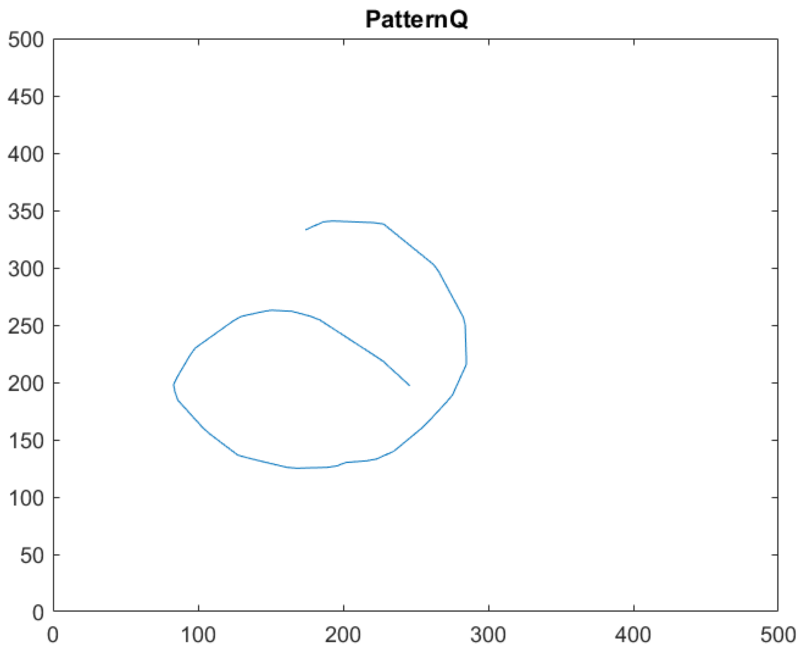


Pure Newton's method finishes with final objective value = 2.2243. It does not reach the global minima (f = -22.1200), instead finding some other local minima. This is because it does not travel far enough in the first step.

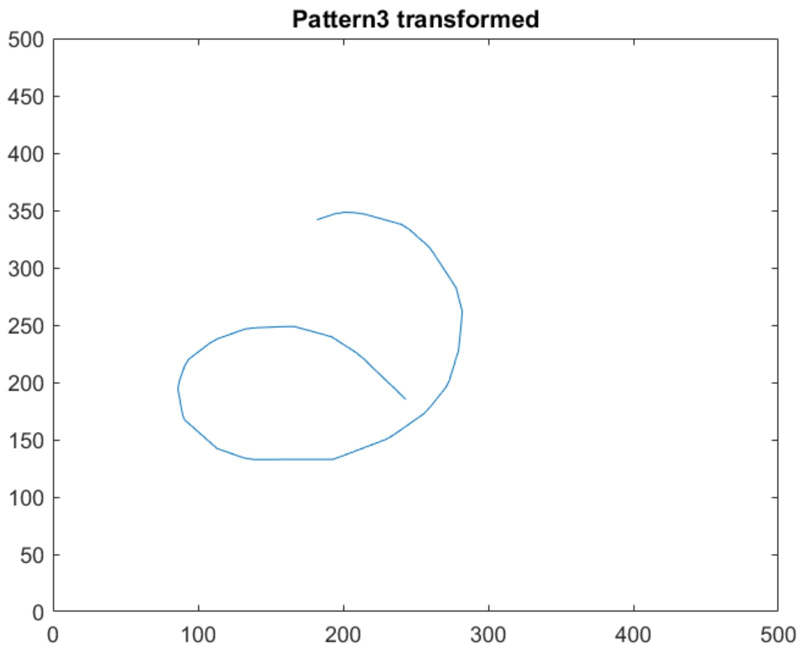
Question 3

Script to run question 3 is named Q3.m

Unknown pattern Q:



Best match - pattern 3:



Best match for all 3 patterns:

```
Pattern 1:
x = [-0.19455, -0.78707, -29.1891, 499.3757]
f = 125005.6479
Pattern 2:
x = [-0.16053, -0.43646, 75.2213, 367.1051]
f = 569742.881
Pattern 3:
x = [0.72766, 0.64803, 235.4454, -128.4313]
f = 7440.8567
```

Question 4

a) Sufficient conditions:

$$\nabla f(x^*) = 0 \tag{1}$$
$$d^T H(x^*)d > 0 \tag{2}$$

Consider (1):

$$f(x^*) = \sum_{k=1}^{n}(p_k \times (y_k - x^*)^2)$$
$$\nabla f(x^*) = \sum_{k=1}^{n}(p_k \times (y_k - x^*)(-2)) = 0$$

$$\sum_{k=1}^{n}(p_k \times (y_k - x^*)) = 0$$
$$\sum_{k=1}^{n}(p_k \times y_k - p_k \times x^*)) = 0$$
$$\sum_{k=1}^{n} p_k \times y_k = \sum_{k=1}^{n} p_k \times x^*$$
$$x^* = \sum_{k=1}^{n}(p_k \times y_k)/\sum_{k=1}^{n} p_k$$
$$x^* = \sum_{k=1}^{n}(p_k \times y_k)$$

Consider (2):

$$H(x^*) = \sum_{k=1}^{n}(2 \times p_k)$$
$$d^T H(x^*)d > 0$$
$$\sum_{k=1}^{n}(2 \times p_k \times d^2) > 0$$
$$\sum_{k=1}^{n}(p_k) > 0$$

(1) states x^* is optimal at $x^* = \sum_{k=1}^{n}(p_k \times y_k)$. (2) is satisfied by the definition of p_k. When $p_k = \frac{1}{n}$, x^* is the mean value of y_k.

b)

$$\min \mathrm{E}[\,|Y - x|\,]$$
$$= \min \sum_{k=1}^{n}(p_k \times |y_k - x|)$$
$$= \min \sum_{k=1}^{n}(\frac{1}{n} \times |y_k - x|)$$
$$= \min \sum_{k=1}^{n}(\frac{1}{n} \times \max(y_k - x, -(y_k - x))$$

Let $f(x) = \sum_{k=1}^{n} g_k(x)$ such that $g_k(x) = \frac{1}{n} \times \max(y_k - x, -(y_k - x))$.

Then, defining d^+, d^- as positive and negative direction respectively:

$$\begin{aligned} g_k'(x; d^+) = -\frac{1}{n} & \quad \text{if } x < y_k \\ \frac{1}{n} & \quad \text{if } x > y_k \\ \frac{1}{n} & \quad \text{if } x = y_k \end{aligned}$$

$$\begin{aligned} g_k'(x; d^-) = \frac{1}{n} & \quad \text{if } x < y_k \\ -\frac{1}{n} & \quad \text{if } x > y_k \\ \frac{1}{n} & \quad \text{if } x = y_k \end{aligned}$$

From page 52 of lecture notes we have:

$$f'(x) = \sum_{k=1}^{n} g_k'(x)$$

Note that there are 2 possible values for $g_k'(x)$ depending on the value of y_k relative to x_k. We denote $N_{(>)}$ as the number of $y_k > x$ and $N_{(<)}, N_{(=)}$ similarly. Then:

$$\begin{aligned} f'(x; d^+) &= -\frac{1}{n} \times N_{(<)} + \frac{1}{n} \times N_{(>)} + \frac{1}{n} \times N_{(=)} \\ f'(x; d^-) &= \frac{1}{n} \times N_{(<)} + -\frac{1}{n} \times N_{(>)} + \frac{1}{n} \times N_{(=)} \end{aligned}$$

Point x^* is at a global minima if $f'(x^*; d) \geq 0 \, \forall d$ as per page 52 of the lecture notes. We then have two conditions to meet:

$$\begin{aligned} N_{(>)} + N_{(=)} \geq N_{(<)} \\ N_{(<)} + N_{(=)} \geq N_{(>)} \end{aligned}$$

Re-ordering y_k such that $y_1 \leq y_2 \leq \cdots \leq y_n$, x^* is optimal between $[y_a, y_b]$ inclusive, where $a = \lfloor \frac{n}{2} \rfloor, b = \lceil \frac{n}{2} \rceil$. This is when x^* has half the y_k values above it, and half below it.