Computer Exercise 1

Edoardo Mangia

In this assignment, different methods are used to minimize some given functions.

The methods used are:

- Steepest Descent method
- Newton's method
- Newton's method (Levenberg-Marquardt)
- Newton's method (modified)

Along with the implementation of the methods, some considerations will be highlighted.

Contents

1 Function 1 **2**
1.1 a) 2
1.2 b) 2
1.3 c) 3

2 Function 2 (Rosenbrock's function) **4**
2.1 a) 4
2.2 b) 4
2.3 c) 5

3 Function 4 **7**
3.1 a) 7
3.2 b) 7

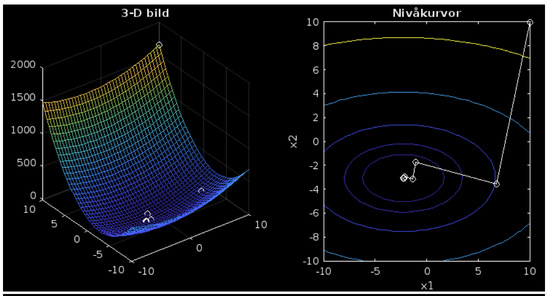
1 Function 1

1.1 a)

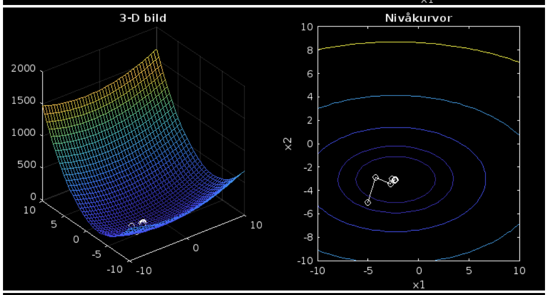
The function considered is

$$f(x_1, x_2) = 2(x_1 + 1)^2 + 8(x_2 + 3)^2 + 5x_1 + x_2$$

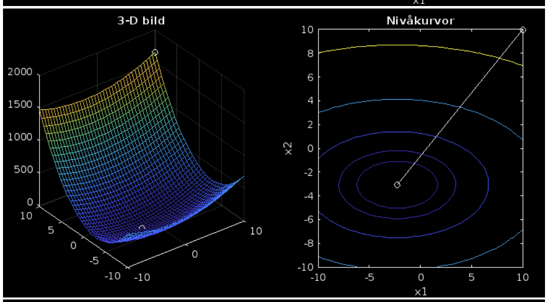
These are the results after studying the function:



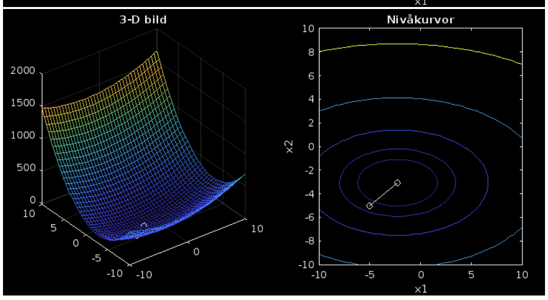
Steepest Descend method
Starting point: (10, 10)
Function value: -11.1562
Solution point: (-2.2491, -3.06255)
No. of iterations: 9



Steepest Descend method
Starting point: (-5, -5)
Function value: -11.1562
Solution point: (-2.25225, -3.06234)
No. of iterations: 9



Newton's (unit step) method
Starting point: (10, 10)
Function value: -11.1562
Solution point: (-2.25, -3.0625)
No. of iterations: 9



Newton's (unit step) method
Starting point: (-5, -5)
Function value: -11.1562
Solution point: (-2.25, -3.0625)
No. of iterations: 9

1.2 b)

The point obtained is locally optimal.
Since the function is convex, for the Fundamental Theorem of Global Optimality, the obtained point is also globally optimal.

1.3 c)

The Steepest Descent method uses the gradient of the function to determine the search direction and a fixed coefficient α, chosen arbitrarily, to determine the step length.
The update rule for a new point x_k, starting from a point x_k, is defined as

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$$

The Newton's method instead, uses the Hessian of the function to determine the step length.
The update rule for a new point x_k, starting from a point x_k, is defined as

$$x_{k+1} = x_k - \left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k)$$

The curvature information from the Hessian, allows the Newton's method to take bigger steps when the function is steeper. On the contrary, it will take smaller steps when the function is less steep.

This method converges to the optimal point quadratically with each iteration.
In other words, the error, or difference between the current point x_n and the optimal point $x*$, is reduced quadratically with each iteration.

Because of this, when the function considered is quadratic, the method reaches a optimal point in just 1 iteration.

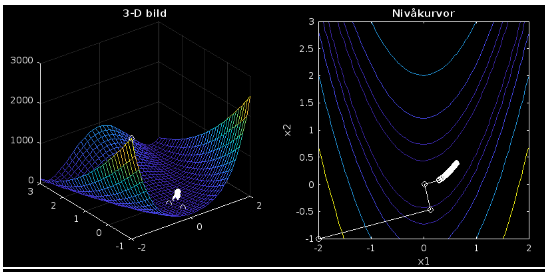
2 Function 2 (Rosenbrock's function)

2.1 a)

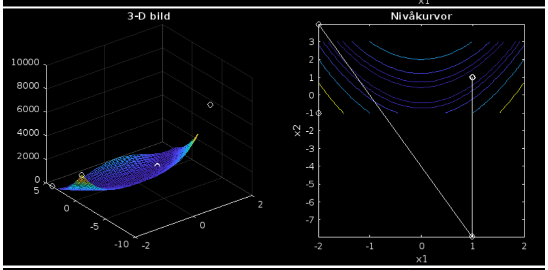
The function considered is

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

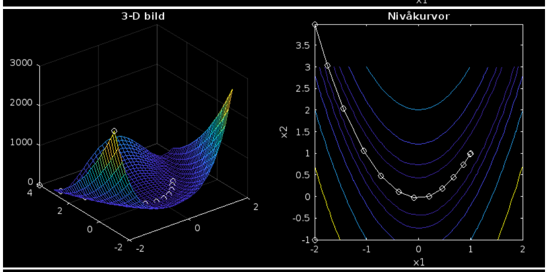
These are the results after studying the function:



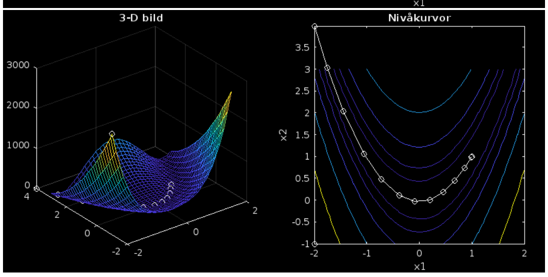
Steepest Descend method
Starting point: (-2, -1)
Function value: 0.13932
Solution point: (0.6269, 0.3919)
No. of iterations: 100



Newton's (unit step) method
Starting point: (-2, -1)
Function value: 1.9932e-27
Solution point: (1, 1)
No. of iterations: 5



Newton's (Marquardt's) method
Starting point: (-2, -1)
Function value: 2.6568e-11
Solution point: (0.9999, 0.9999)
No. of iterations: 14



Newton's (modified) method
Starting point: (-2, -1)
Function value: 2.6568e-11
Solution point: (0.9999, 0.9999)
No. of iterations: 14

2.2 b)

The Rosenbrock's function is not convex and the global minimum is found in $(x_1, x_2) = (1, 1)$, where $f(1, 1) = 0$.

By using the Newton's method, in it's different iterations, it was possible to converge to the minimum in significantly less iterations, compared to using the steepest descent method.

The function is not convex.

To prove that this function is not convex, we can use the definition of convexity.
A function f is convex if for all $\lambda \in [0,1]$ and for all points $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$, the following inequality holds:

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}).$$

We will find two points $\mathbf{x}$ and $\mathbf{y}$, and a value of $\lambda \in (0,1)$, such that this inequality does not hold. Let us choose the points:

$$\mathbf{x} = (0,1), \quad \mathbf{y} = (2,1).$$

Let $\lambda = \frac{1}{2}$. Then:

$$\lambda \mathbf{x} + (1-\lambda)\mathbf{y} = \frac{1}{2}(0,1) + \frac{1}{2}(2,1) = (1,1).$$

Now, calculate the function values:

$$f(\mathbf{x}) = f(0,1) = 100(1-0^2)^2 + (1-0)^2 = 100 + 1 = 101,$$

$$f(\mathbf{y}) = f(2,1) = 100(1-2^2)^2 + (1-2)^2 = 100(-3)^2 + (-1)^2 = 900 + 1 = 901,$$

$$f\left(\frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{y}\right) = f(1,1) = 100(1-1^2)^2 + (1-1)^2 = 0.$$

Now, check the convexity inequality:

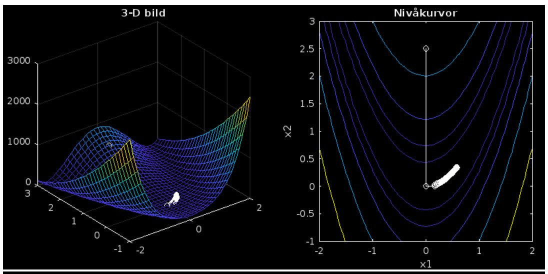
$$\lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}) = \frac{1}{2} \cdot 101 + \frac{1}{2} \cdot 901 = \frac{101 + 901}{2} = 501.$$

Clearly:

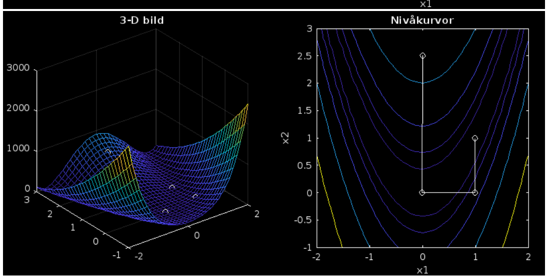
$$f\left(\frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{y}\right) = 0 \quad \text{and} \quad \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}) = 501.$$

Since $0 \not\leq 501$, the inequality fails, and hence the function $f(x_1, x_2)$ is **not convex**.

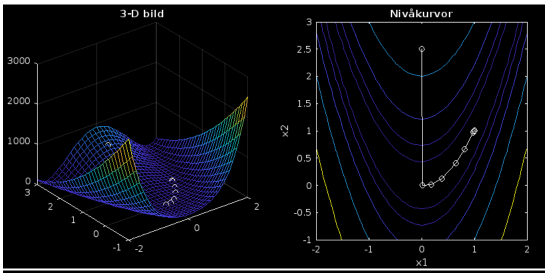
2.3 c)



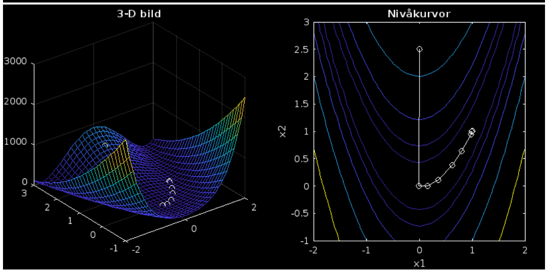
Steepest Descent method
Starting point: (0, 2.5)
Function value: 0.17365
Solution point: (0.5846, 0.3385)
No. of iterations: 100



Newton's (unit step) method
Starting point: (0, 2.5)
Function value: 6.4029e-07
Solution point: (0.9992, 0.9984)
No. of iterations: 3



Newton's (Marquardt's) method
Starting point: (0, 2.5)
Function value: 9.5434e-11
Solution point: (1.000, 1.000)
No. of iterations: 8



Newton's (modified) method
Starting point: (0, 2.5)
Function value: 2.3718e-09
Solution point: (1.000, 1.000)
No. of iterations: 8

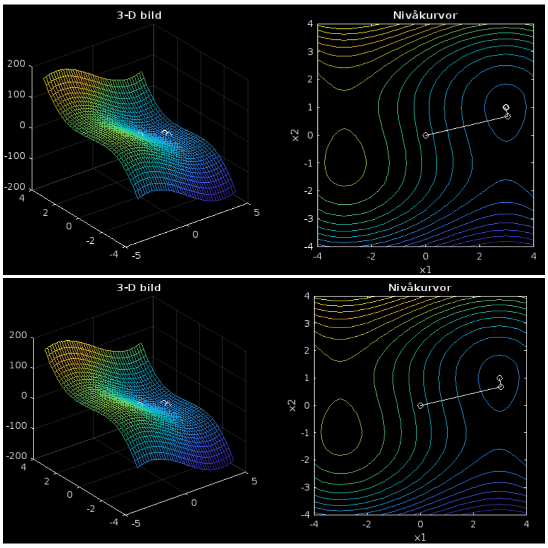
3 Function 4

3.1 a)

The function considered is

$$f(x_1, x_2) := x_1^3 + 2x_2^3 - 27x_1 - 6x_2$$

These are the results after studying the function:



Steepest Descent method
Starting point: (0, 0)
Function value: -58
Solution point: (2.999, 0.999)
No. of iterations: 4

Newton's (Marquardt's) method
Starting point: (0, 0)
Function value: -58
Solution point: (3, 0.999)
No. of iterations: 3

Choosing $x_1, x_2 = (0, 0)$ as the starting point, we find that the Newton's method (unit step) doesn't behave well and can't find an optimal point in the first iteration.
The reason for this, goes back to how the method rely on the information of the Hessian of the function to search for new points.
In $x_1, x_2 = (0, 0)$, the Hessian of the function is 0 and therefore the computation results in an error.

For the same starting point, the other methods, do not rise errors during the computation.
This is why:

- The Steepest Descent method, does not rely on the Hessian of the function to calculate the next optimal point in each iteration.
- The Levenberg-Marquardt modification of the Newton's method, doesn't use the Hessian of the function, but rather a first-order derivative approximation of the latter, the Jacobian matrix.

3.2 b)

Studying the function, we can find some stationary points.
In particular:

- Local minimum in $(-3, -1)$ (Function Value: 58).
- Saddle point in $(-3, 1)$ (Function Value: 50).

- Saddle point in $(3, -1)$ (Function Value: -50).
- Local minimum in $(3, 1)$ (Function Value: -58).