Optimization - Project 3

Projected Steepest Descent Method

Nikolaos Konstas

Department of Electrical & Computer Engineering
Aristotle University of Thessaloniki

December 2022

Contents

Introduction

In the present work, it is requested to implement the Projected Descent Method and to compare it with the classic Steepest Descent Method (without projection).

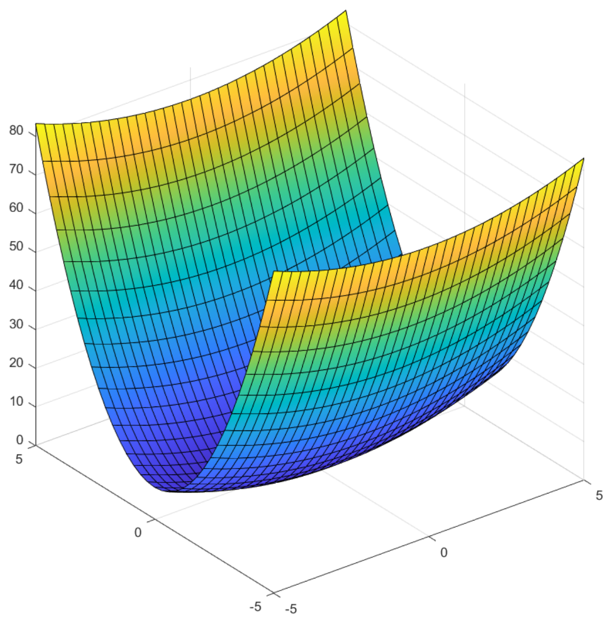
For the purposes of the paper, the objective function is minimized:

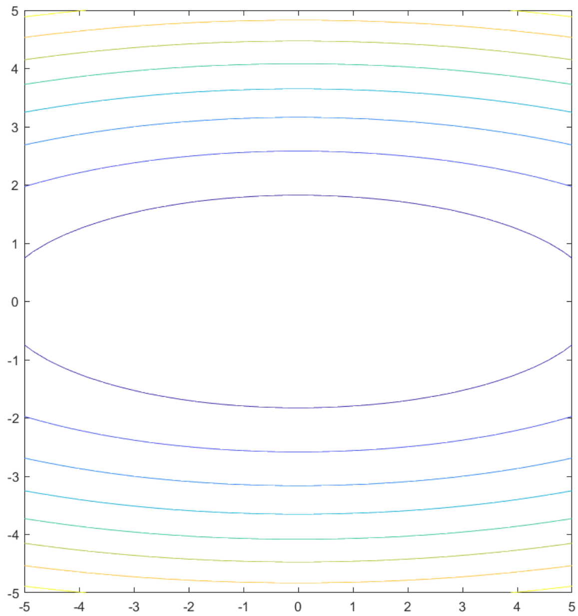
$$f : \mathbb{R}^2 \to \mathbb{R},\ f(x) = \frac{1}{3}x_1^2 + 3x_2^2,\ [x_1\ x_2]^T$$

, while for the projection method the constraints are additionally used:

$$-10 \leq x_1 \leq 5$$
$$-8 \leq x_2 \leq 12$$

Finally, to observe the performance of the algorithms and draw our conclusions, we use different values for the free parameters γ, s_k, but also for the starting point of the algorithm. Before proceeding to the implementation and application of the algorithms, with the help of matlab we observe the graph of f.

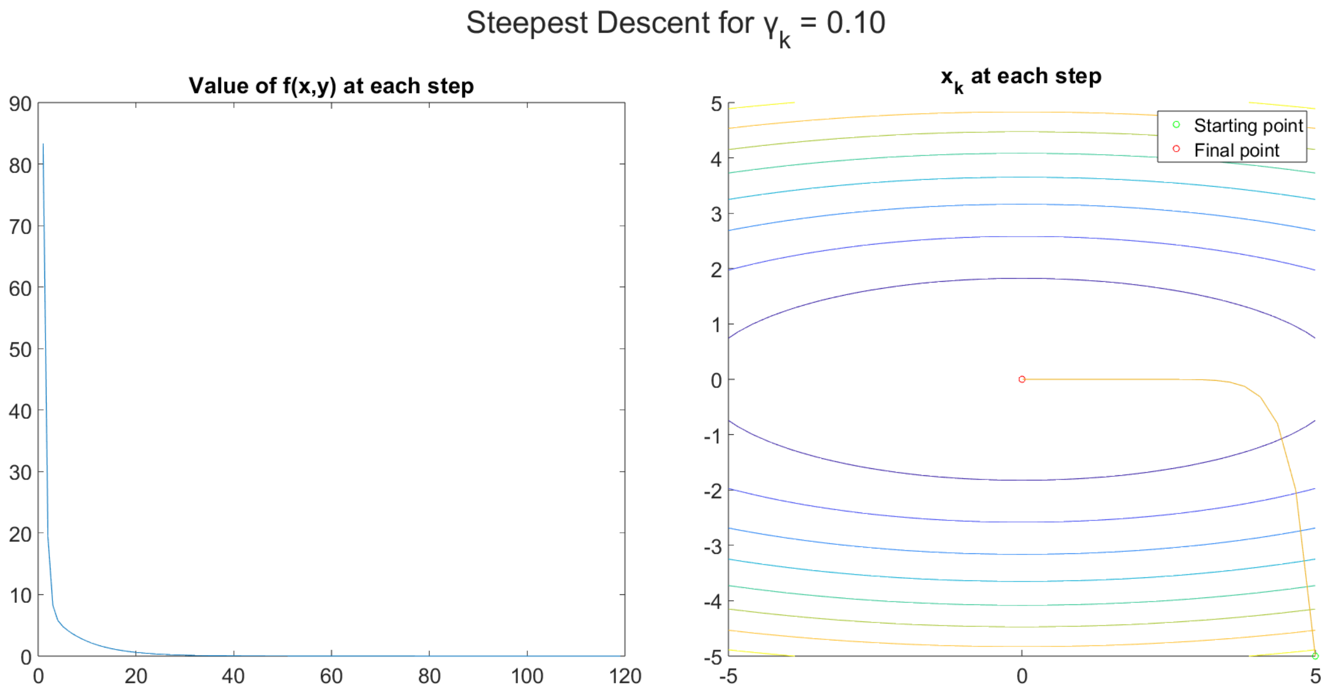


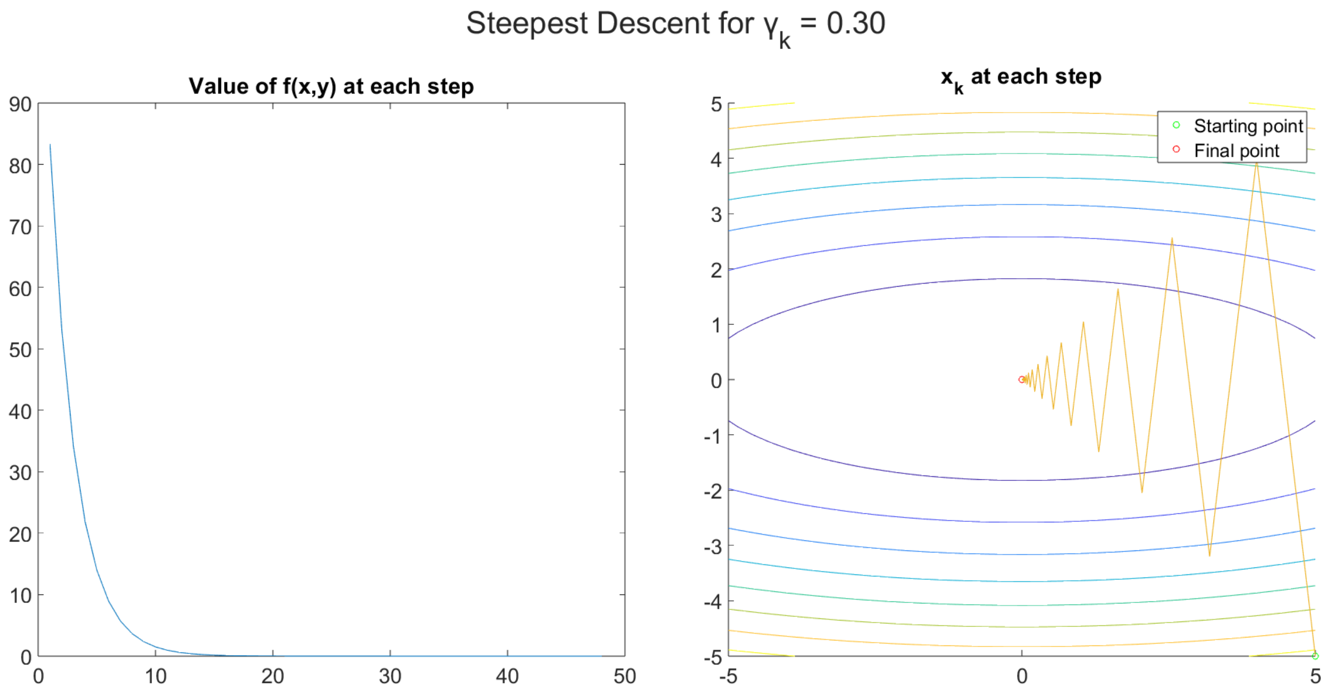


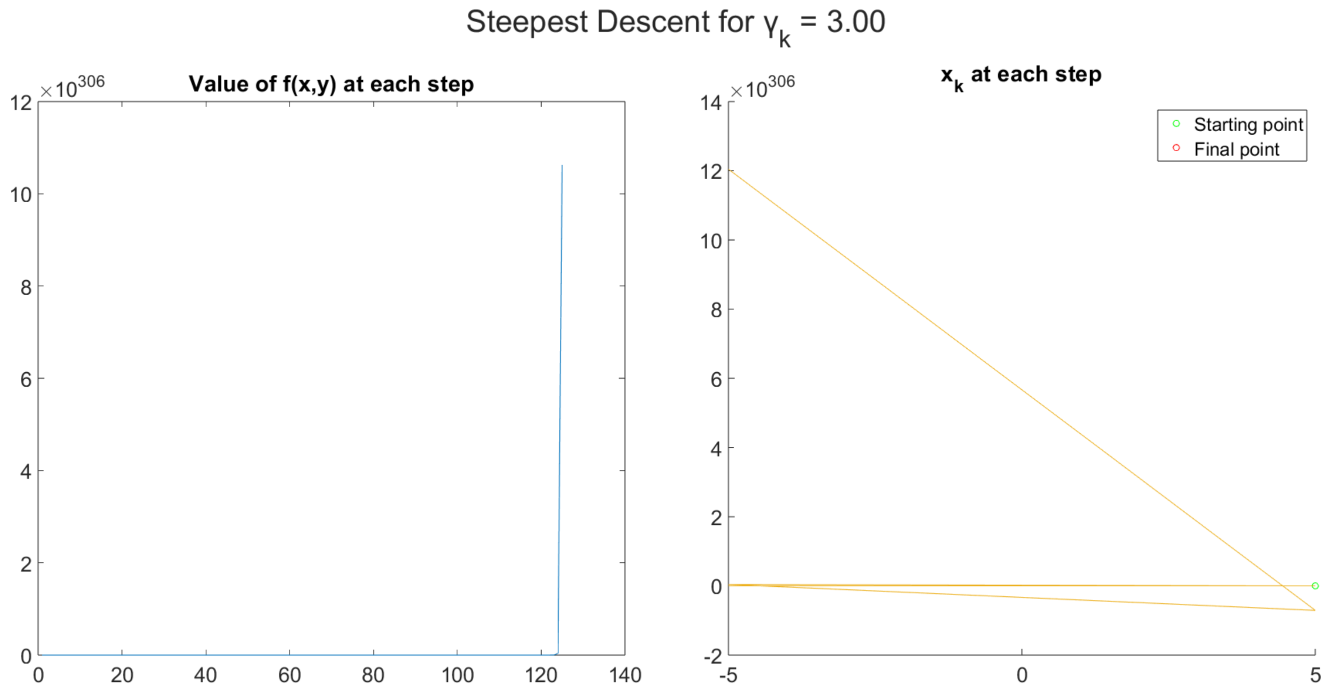
As is evident from the formula, f is a convex function and exhibits a global minimum at $(0, 0)$.

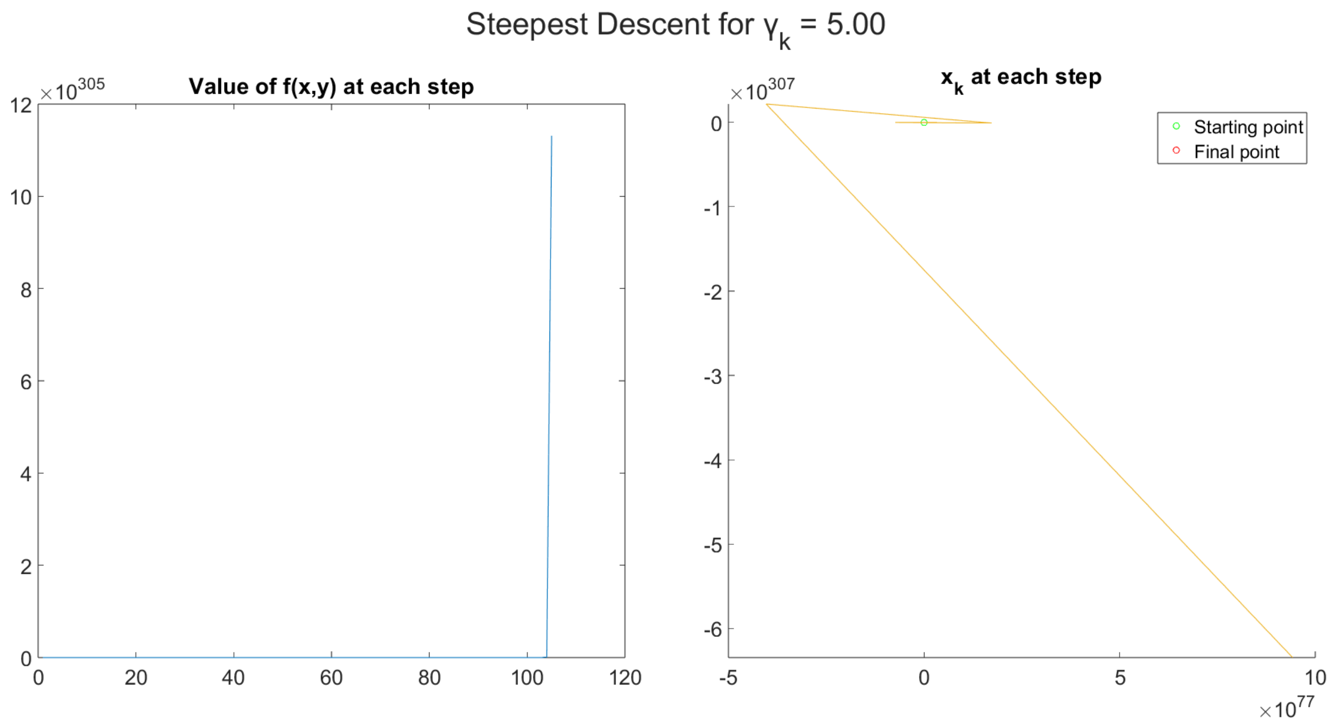
Subject 1

In the 1st Topic, we used the Steepest Descent Method with an accuracy of $\varepsilon = 0.001$, while different values were used for the step γ_k in order to observe the different behavior that results. In the specific tests $(5, -5)$ was used as the starting point. Here are the corresponding graphs:









Observing the graphs, we see that in the first 2 cases, for $\gamma_k = 0.1$ and $\gamma_k = 0.3$, the algorithm converges and results in the real minimum of f, while for $\gamma_k = 3$ and $\gamma_k = 5$ f goes to infinity and the algorithm is driven to instability. Let us now prove the above results mathematically.

First, we calculate the slope of f:

$$\nabla f(x) = \begin{bmatrix} \frac{2}{3}x_1 & 6x_2 \end{bmatrix}^T$$

So according to the calculation rule x_{k+1} will be:

$$x_{k+1} = x_k - \gamma_k \nabla f(x_k) \Rightarrow \begin{bmatrix} x_{1_{k+1}} \\ x_{2_{k+1}} \end{bmatrix} = \begin{bmatrix} x_{1_k} \\ x_{2_k} \end{bmatrix} - \gamma_k \begin{bmatrix} \frac{2}{3}x_{1_k} \\ 6x_{2_k} \end{bmatrix} \Rightarrow$$
$$\begin{bmatrix} x_{1_{k+1}} \\ x_{2_{k+1}} \end{bmatrix} = \begin{bmatrix} (1-\frac{2}{3}\gamma_k)x_{1_k} \\ (1-6\gamma_k)x_{2_k} \end{bmatrix}$$

As we noticed above, the minimum of f is at the point $(0, 0)$, so in order to have convergence, the following system of constraints for the values of γ_k results:

$$\begin{cases} |1-\frac{2}{3}\gamma_k| < 1 \\ |1-6\gamma_k| < 1 \end{cases} \Rightarrow \begin{cases} -1 < 1-\frac{2}{3}\gamma_k < 1 \\ -1 < 1-6\gamma_k < 1 \end{cases} \Rightarrow$$
$$\begin{cases} 0 < \gamma_k < 3 \\ 0 < \gamma_k < \frac{1}{3} \end{cases} \Rightarrow 0 < \gamma_k < \frac{1}{3}$$

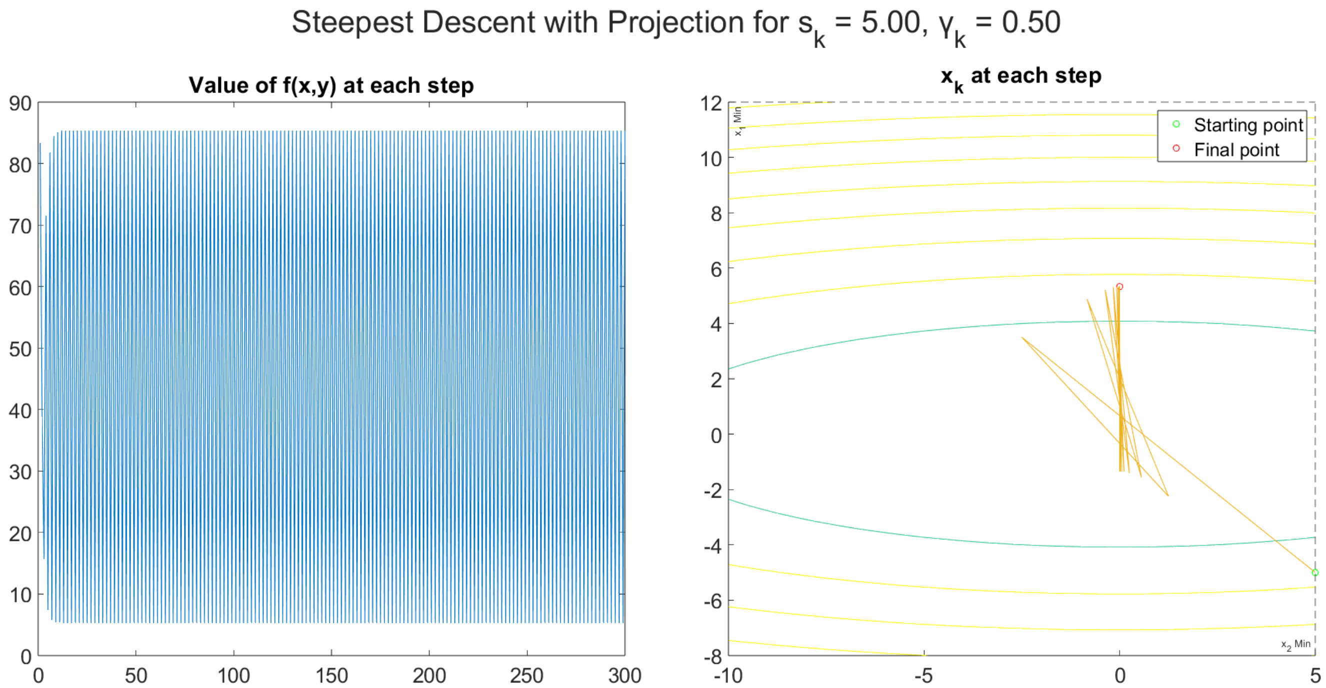
Now it is clear that in the first 2 cases we have convergence, as the values of γ_k belong to the interval $(0, \frac{1}{3})$, while in the other 2 it is $\gamma_k > \frac{1}{3}$ and thus we are led to instability. Moreover, notice that in the 3rd case, $\gamma_k = 3$, we are in the boundary condition for x_1 $(0 < \gamma_k < 3)$, so only x_2 goes to infinity, in contrast to the 4th case where both go to infinity variables x_1, x_2.

Subject 2

We now use the Projected Steepest Descent Method so that the constraints are met:

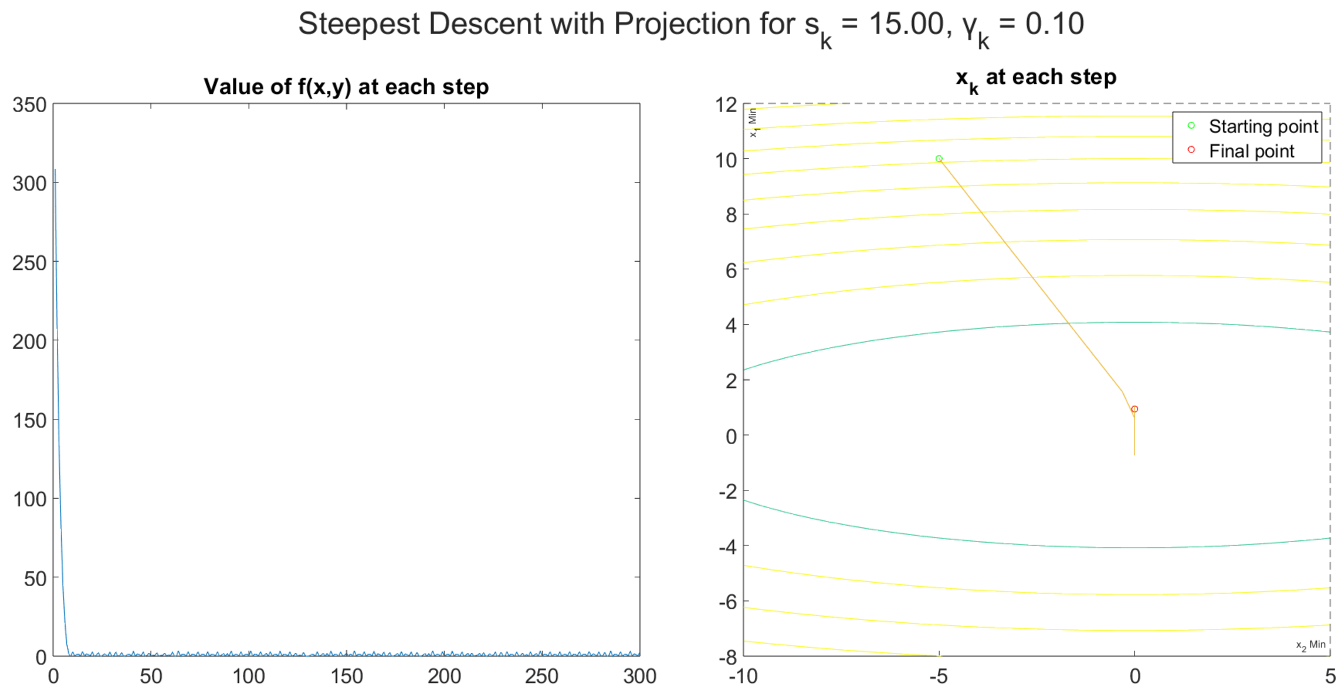
$$-10 \leq x_1 \leq 5$$
$$-8 \leq x_2 \leq 12$$

Corresponding to above, to have a successful convergence we want $s_k \cdot \gamma_k < \frac{1}{3}$ to hold, so in the specific case where we have $s_k = 5, \gamma_k = 0.5$, and therefore $s_k \cdot \gamma_k = 2.5$ we expect the algorithm to be unstable and not converge. Indeed, as we can see in the graphs below, the algorithm does not converge to the minimum, but it does not escape to infinity due to the restriction of x within the constraints, resulting in an oscillation. Because of this behavior, the termination criterion is never met and we terminate the algorithm when it reaches the 300th iteration.



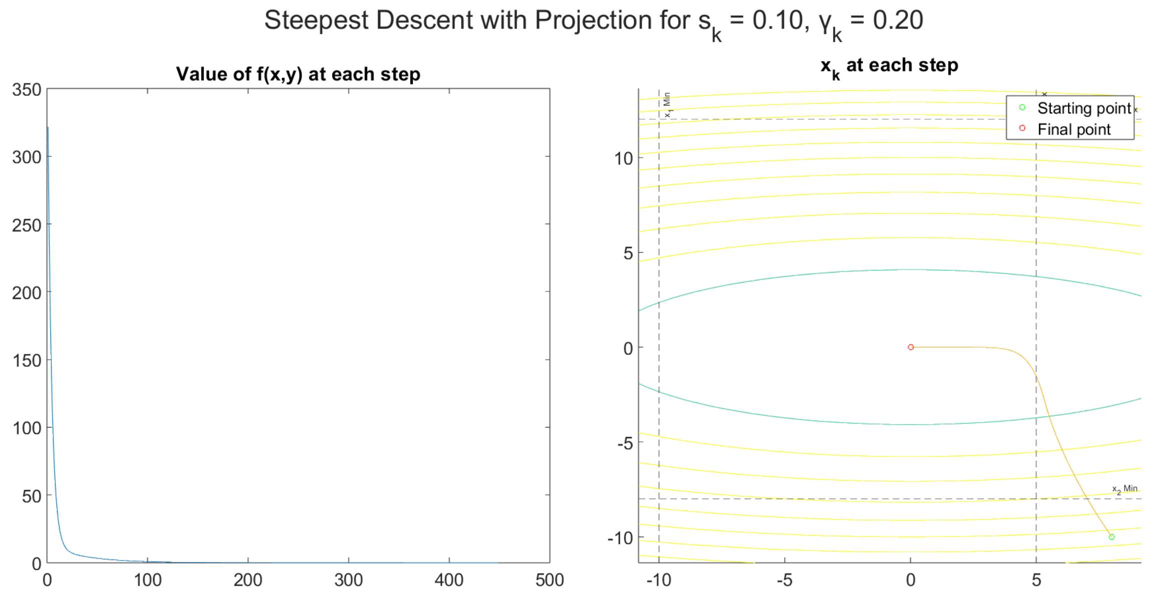
Subject 3

We again use the Projected Steepest Descent Method, starting from the point $(-5, 10)$ and with $s_k = 15, \gamma_k = 0.1$ this time. The product $s_k \cdot \gamma_k = 1.5 > \frac{1}{3}$, so in this case too we expect the algorithm not to converge. Indeed, as can be seen in the figures below, the algorithm does not terminate, except with an intervention at the 300th iteration, but this time the perturbation is, obviously, of a smaller amplitude with the end point being close to $(0, 0)$, as well as the product $s_k \cdot \gamma_k$ is smaller than Theorem 2.



Subject 4

We minimize f again, with Projection's Steepest Descent Method. In the parameters of the algorithm we give the values $s_k = 0.1, \gamma_k = 0.2$, so $s_k \cdot \gamma_k = 0.02 < \frac{1}{3}$ and we wait for the algorithm to converge to the minimum. As the product $s_k \cdot \gamma_k$ is very small, we spend the limit of 300 iterations. Furthermore, the starting point in this Topic is (8,-10), which is out of bounds. Although the algorithm converges to (0,0), as shown in the following figure, x_1, x_2 takes on multiple iterations out-of-bounds values due to the poor choice of starting point.



So we make a modification to the code, and since the origin point is outside the constraints, we use its view to convert it to a touchable point and have the desired behavior shown below: