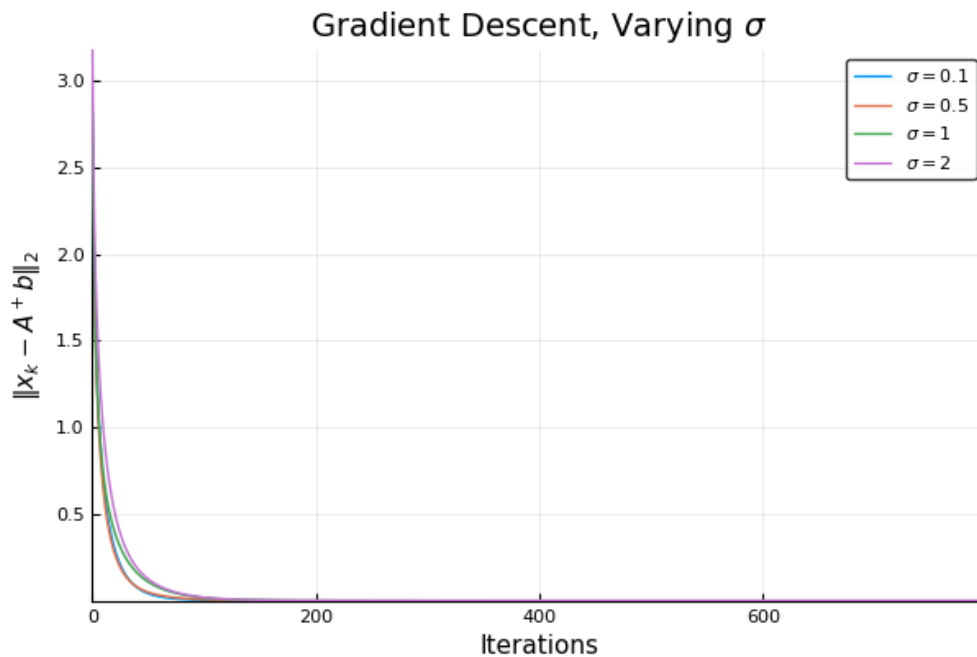


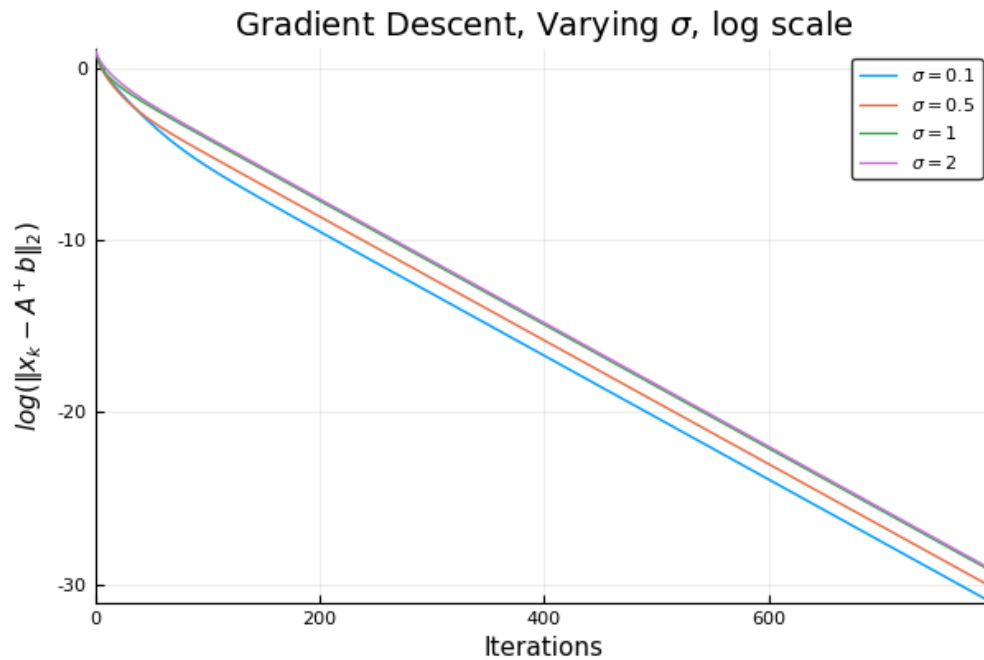
18.065 PSet 4

Dimitris Koutentakis

Problem 1

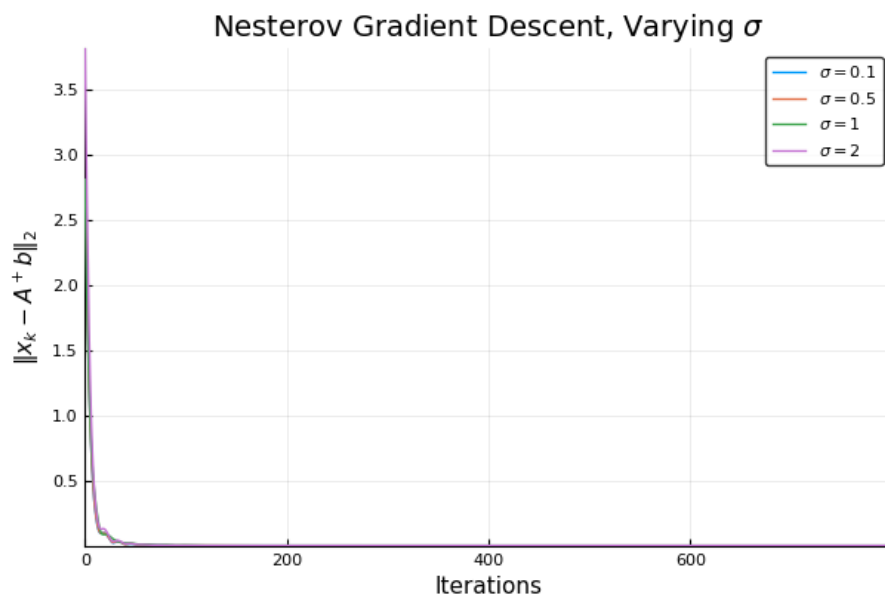
Based on the two plots below (in log and normal scale), we can see that $\|x_{ls} - s_k\|$ decreases monotonically. The Gradient Descent algorithm seems to perform better for lower values of σ . However, the algorithm has relatively similar performance for different values of σ .

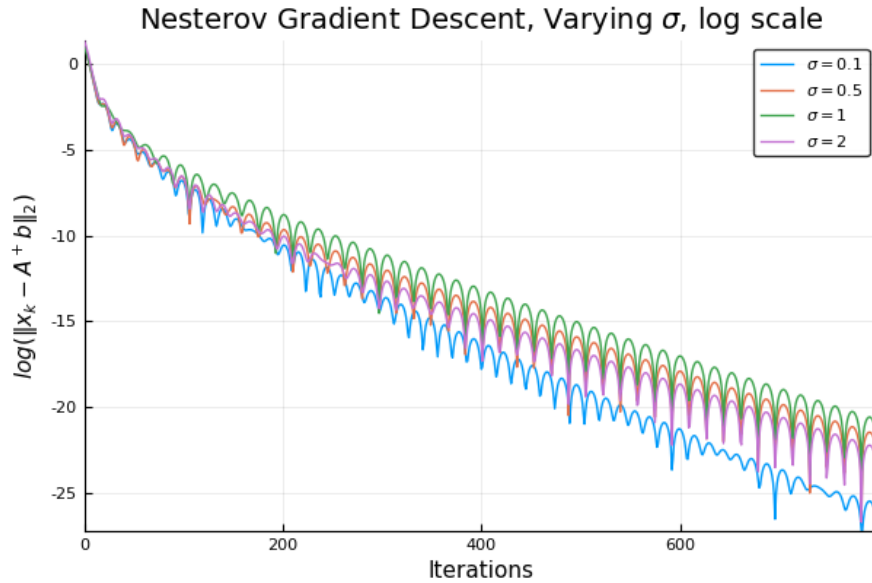




Problem 2

Attached below are the graphs of Nesterov-accelerated Gradient descent for varying values of σ . The first plot is in linear scale and the second in logarithmic, which shows more clearly that lower values of σ have slightly better performance.





For $\sigma = 1$ and a given step-size of μ , we can see that in fact the Nesterov-accelerated Gradient Descent converges much faster than the normal Gradient Descent. However, it is important to note that after a certain number of iterations, the Nesterov-accelerated Gradient Descent does not outperform the normal Gradient Descent algorithm.

As the step-size μ increases, both Nesterov-accelerated and normal Gradient descent converge faster. However, the number of iterations for which Nesterov-accelerated outperforms the normal gradient descent decreases. This is obvious in the plots attached below.

