MS-E2122 - Nonlinear optimization

Homework 5

Ari Viitala

432568

5.1 Conjugate Gradient Method

The code used for solving the conjugate gradient method and producing the plots can be found attached. The code has also feature to time the code execution. However, Julia seems to cache function values in some way that makes it possible to receive really fast times. This should be taken into account if time is used to measure the efficiency of the code.

Problem a

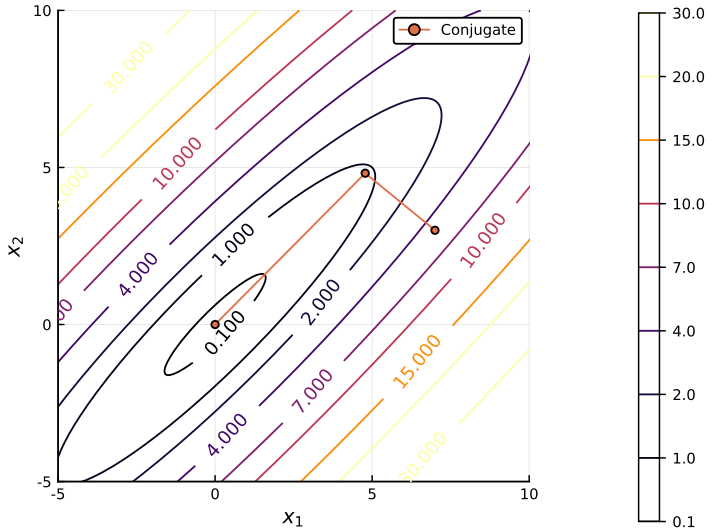
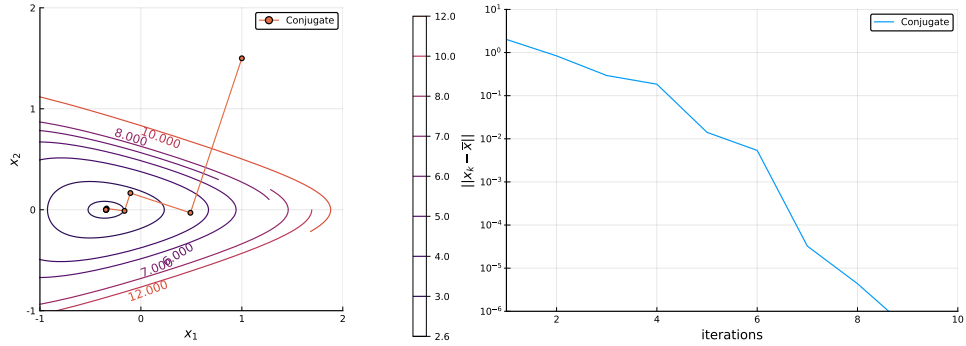


Figure 1: The path taken by conjugate gradient method in problem A

In figure 1 we can see the the path taken by conjugate gradient method whilst solving the objective function in problem a. As we can see the algorithm converges in two steps. This makes sense since the objective function is quadratic and based on the lecture slides of lecture 8 the conjugate gradient method should converge in at most n steps. Since this problem is 2 dimensional so convergence in 2 steps make sense.

Problem b

In figure 2 we can see the path conjugate gradient method uses whilst solving problem b as well as the convergence plot of the solution. All in all, the method takes 8 steps to reach the optimum but as we can see from plot 2a the method gets really close to the optimum already in 4 steps and the rest is just to satisfy the tolerance condition $\epsilon = 10^{-6}$.



(a) Path **(b)** Convergence

Figure 2: Path taken by the conjugate gradient method for problem b and the convergence plot of the solution.

In this case the problem is no longer quadratic so it is sensible that the method does not converge in two steps. From figure 2b we can see that the convergence is about linear for the non-quadratic problem. From figure 2a We can also see a behavior where at the second point the method clearly does not follow the direction on the most rapid descent but instead the conjugate direction of the first gradient.