Nonlinear Optimization - Homework 4

Christian Segercrantz 481056

December 16, 2021

4.1 Frank-Wolfe Method

The completed code can be seen from the attached file. Compared to the pseudo code found in the lecture slides, the algorithm did not result in the correct answer when using $\nabla f(x)^\top d^k$ but when using $\nabla f(x)^\top x$ as the objective value. At the first iteration $d^k = x$ since $x^k = 0$, but from there this result seems odd. Using d^k results in way faster convergence, but another optimal value.

4.2 Interior-Point Method for Quadratic Problems

The problem

$$\text{min. } c^\top x + \frac{1}{2}x^\top Q x \tag{1}$$

$$\text{subject to: } Ax = b \tag{2}$$

$$x \geq 0 \tag{3}$$

and it's dual

$$\text{min. } b^\top v \frac{1}{2}x^\top Q x \tag{4}$$

$$\text{subject to: } A^\top v + u - Qx = c \tag{5}$$

$$x \geq 0. \tag{6}$$

a)

The KKT conditions for the problem are

$$0 = \nabla f(x) + \sum_j u_j \nabla g(x) + \sum_i v_i \nabla h(x) \tag{7}$$

$$u_j g(x) = 0 \quad j = 1...m \tag{8}$$

$$v_i h(x) = 0 \quad i = 1...l \tag{9}$$

$$x \in X, g_j(x) \leq 0 \quad j = 1...m \tag{10}$$

$$u_j \geq 0 \quad j = 1...m \tag{11}$$

$$\tag{12}$$

which for our problem becomes

$$\nabla(c^\top x + \frac{1}{2}x^\top Q x) - u_1 + v_1 \nabla(Ax - b) = 0 \tag{13}$$

$$c^\top + Qx - u_1 + v_1 A = 0 \tag{14}$$

$$u_1^\top x = 0 \tag{15}$$

$$v_1(Ax - b) = 0 \tag{16}$$

$$x \geq 0 \tag{17}$$

$$u_i \geq 0 \tag{18}$$

$$\tag{19}$$

b)

The general form of the Barrier problem looks as

$$\text{(BP): } \inf_{\mu} \theta(\mu) \tag{20}$$

$$\text{subject to: } \mu > 0 \tag{21}$$

$$\tag{22}$$

where $\theta(\mu)$ is, in our case,

$$\theta(\mu) = \inf_{x}\{f(x) + \mu B(x) : -x \leq 0\} \tag{23}$$

$$= \inf_{x}\{f(x) + \mu(-\sum_{j=1}^{m}\ln(-g_j(x))) : -x \leq 0\} \tag{24}$$

$$= \inf_{x}\{f(x) + \mu(-\ln(x)) : -x \leq 0\} \tag{25}$$

$$= \inf_{x}\{f(x) - \mu\ln(x) : -x \leq 0\}. \tag{26}$$

The problem can thus be formulated as

$$\text{(BP): } \inf_{\mu}\{\inf_{x}\{c^{\top}x + \frac{1}{2}x^{\top}Qx - \mu\ln(x) : -x \leq 0\}\} \tag{27}$$

$$\text{subject to: } \mu > 0. \tag{28}$$

c)

The first part of the KKT conditions of part b) is

$$c^{\top} + Qx - u_1 + v_1 A - \frac{\mu}{x} = 0 \tag{29}$$

and the rest follows as from the a) part.

From the lecture slides and using our problem we get the following Newton system

$$\begin{bmatrix} A & 0 & 0 \\ -Q & A^{\top} & I \\ \bar{U}^k & 0 & \bar{X}^k \end{bmatrix} \begin{bmatrix} d_x^{k+1} \\ d_v^{k+1} \\ d_u^{k+1} \end{bmatrix} = \begin{bmatrix} Ax^k - b \\ A^{\top}v^k + u^k - Qx^k - c \\ X^kU^ke - \mu^{k+1}e \end{bmatrix} = -\begin{bmatrix} r_p \\ r_d \\ r_c \end{bmatrix}, \tag{30}$$

where the first row of the matrix after the first equality sign is the condition of the primal problem and the second is the condition of the dual problem. From this we get the following system of equations

$$\begin{cases} Ad_x^{k+1} = Ax^k - b = -r_p \\ -Qd_x^{k+1} + A^{\top}d_v^{k+1} + Id_u = A^{\top}v^k + u^k - Qx^k - c = -r_d \\ \bar{U}^k d_x + \bar{X}^k d_u = X^kU^ke - \mu^{k+1}e = -r_c \end{cases}. \tag{31}$$

From the second expression we can solve d_u as

$$A^{\top}d_v^{k+1} + Id_u = -r_d \tag{32}$$

$$d_u = -r_d + Qd_x^{k+1} - A^{\top}d_v^{k+1}. \tag{33}$$

By substituting this into third expression we get

$$\bar{U}^k d_x + \bar{X}^k d_u = -r_c \tag{34}$$

$$\bar{U}^k d_x - \bar{X}^k(r_d - Qd_x^{k+1} + A^{\top}d_v^{k+1}) = -r_c \tag{35}$$

$$\bar{U}^k d_x + \bar{X}^k Qd_x^{k+1} - \bar{X}^k r_d - \bar{X}^k A^{\top}d_v^{k+1} = -r_c \tag{36}$$

$$(\bar{U}^k + \bar{X}^k Q)d_x^{k+1} = -r_c + \bar{X}^k r_d + \bar{X}^k A^{\top}d_v^{k+1} \tag{37}$$

$$d_x^{k+1} = (\bar{U}^k - \bar{X}^k Q)^{-1}(-r_c + \bar{X}^k r_d + \bar{X}^k A^{\top}d_v^{k+1}) \tag{38}$$

From the first expression we know that $Ad_x^{k+1} = -r_p$, thus by multiplying the Equation 38 by A we get that

$$Ad_x^{k+1} = A(\bar{U}^k + \bar{X}^k Q)^{-1}(-r_c + \bar{X}^k r_d + \bar{X}^k A^\top d_v^{k+1}) = -r_p \tag{39}$$

from which we can solve d_v^{k+1} as

$$-r_p = A(\bar{U}^k + \bar{X}^k Q)^{-1}(-r_c + \bar{X}^k r_d + \bar{X}^k A^\top d_v^{k+1}) \tag{40}$$

$$-r_p = -A(\bar{U}^k + \bar{X}^k Q)^{-1} r_c + A(\bar{U}^k + \bar{X}^k Q)^{-1}\bar{X}^k r_d + A(\bar{U}^k + \bar{X}^k Q)^{-1}\bar{X}^k A^\top d_v^{k+1} \tag{41}$$

$$A(\bar{U}^k + \bar{X}^k Q)^{-1}\bar{X}^k A^\top d_v^{k+1} = -r_p + A(\bar{U}^k + \bar{X}^k Q)^{-1} r_c - A(\bar{U}^k + \bar{X}^k Q)^{-1}\bar{X}^k r_d \tag{42}$$

$$d_v^{k+1} = (A(\bar{U}^k + \bar{X}^k Q)^{-1}\bar{X}^k A^\top)^{-1}(-r_p + A(\bar{U}^k + \bar{X}^k Q)^{-1} r_c - A(\bar{U}^k + \bar{X}^k Q)^{-1}\bar{X}^k r_d) \tag{43}$$

$$d_v^{k+1} = (A(\bar{U}^k + \bar{X}^k Q)^{-1}\bar{X}^k A^\top)^{-1}(-r_p + A(\bar{U}^k + \bar{X}^k Q)^{-1}(r_c - X r_d)) \tag{44}$$

We know how a solution for d_v^{k+1} which we can substitute into the earlier expressions for d_x^{k+1} and d_u^{k+1}. The trajectory of the algorithm can be seen in Figure 1.

d)

Using the provided skeleton and earlier results, we obtain the solution $x \approx \begin{bmatrix} 0.905 & 0.819 \end{bmatrix}^\top$ after 6 iterations.
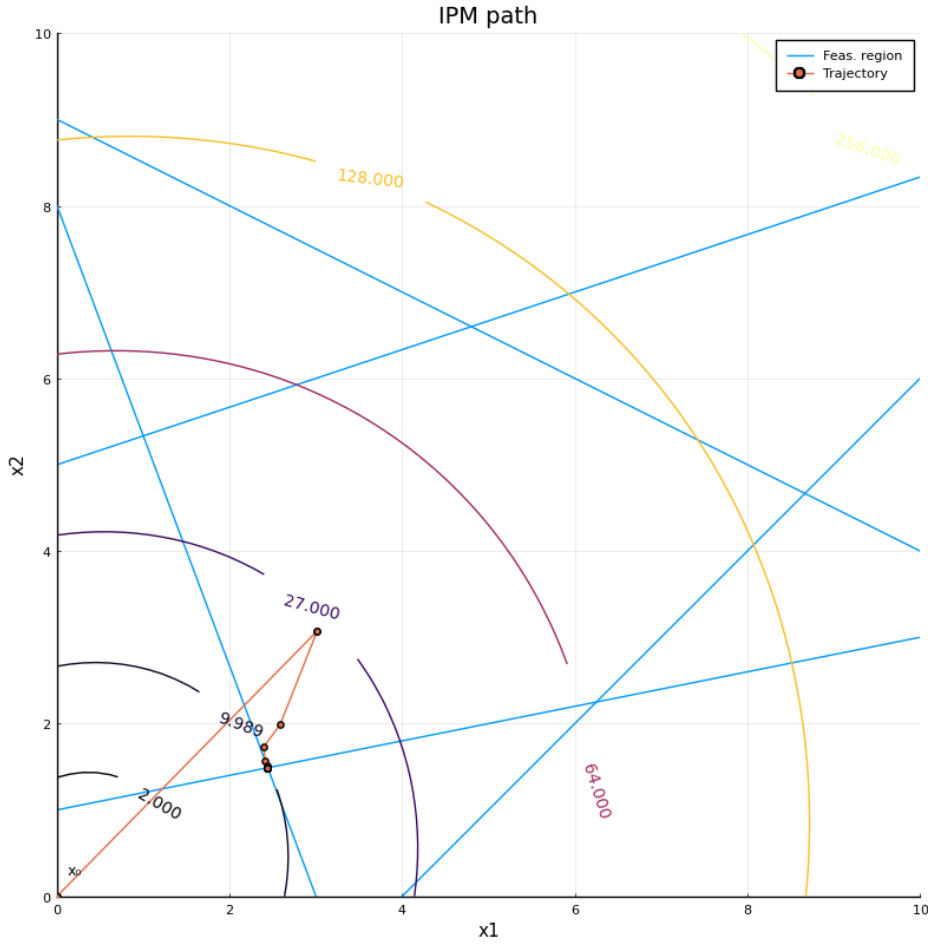
Figure 1: The trajectory the x-values gotten from the interior-point method algorithm used.

4.3 Sequential Quadratic Programming

a)

Using the skeleton and lecture notes 11 we can create the algorithm. The algorithm can be seen in the attached file. It seems using the default values for the trust region Δ and the penalty term μ seem to be enough for the algorithm to converge. The plot is made using the default values. Some additional values were tested, e.g. $\Delta = 0.5$ and $\mu = 1$ did not result in convergence while very large values $\Delta = 1000$ and $\mu = 1000$ did result in convergence. Plots of the aforementioned examples can be seen in Figure 3. We can see that the default values seem to provide the best result.
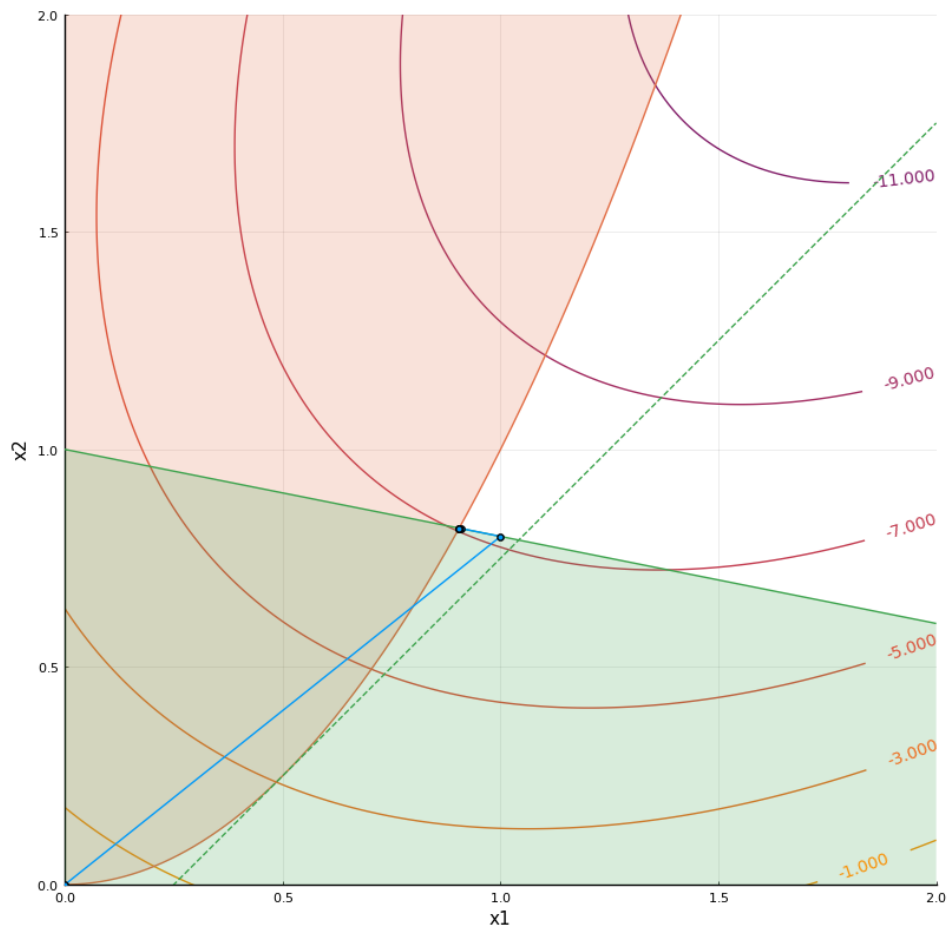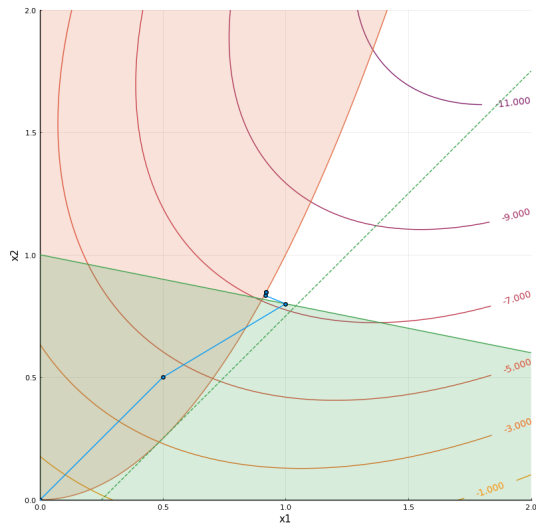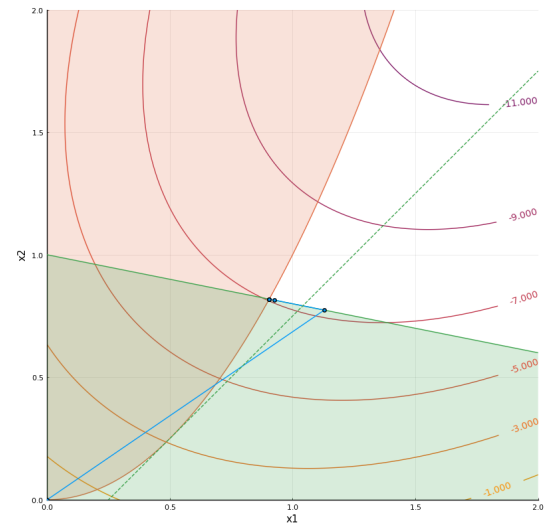
b)

Figure 2: The trajectory the x-values take, gotten from the interior-point method algorithm used. The feasible area is also highlighted

(a) $\Delta = 0.5$, $\mu = 1$

(b) $\Delta = 1000$, $\mu = 1000$

Figure 3: Trejectory plots of alternative trust region Δ and the penalty term μ values.