Nonlinear Optimization - Homework 3

Christian Segercrantz 481056

November 27, 2021

3.1 FJ and KKT Conditions at Optimal Point

a)

$$\begin{aligned} \text{min.} \quad & -x_1 && (1)\\ \text{subject to:} \quad & x_2 \le (1-x_1)^3 && (2)\\ & x_1 \ge 0 && (3)\\ & x_2 \ge 0 && (4) \end{aligned}$$

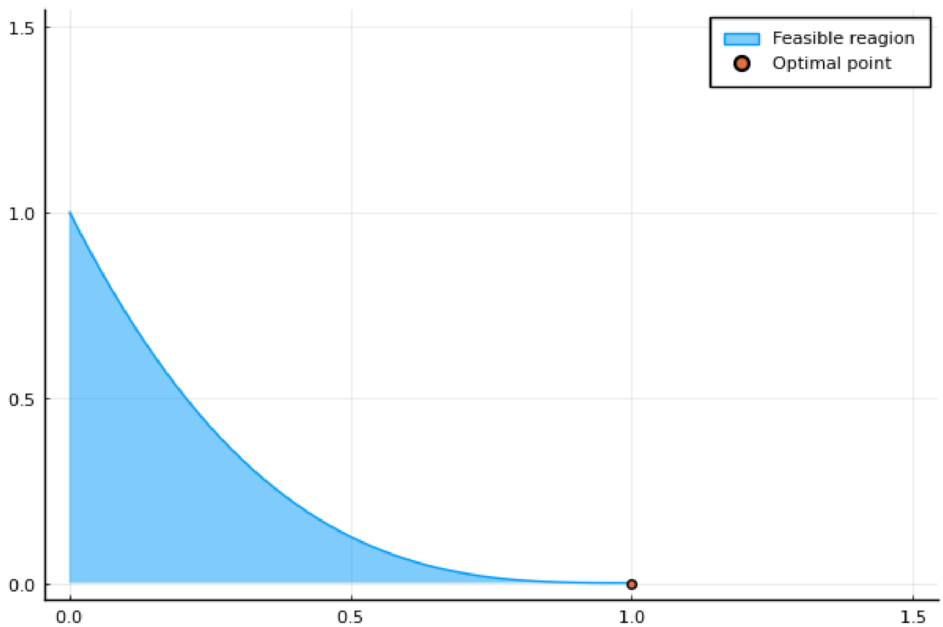

Figure 1: The feasible region of the problem of exercise 3.1. The condition of x_1,$x_2 \ge 0$ is implemented by the limits of the plot.

Figure 1 shows the feasible region for the problem above. Since minimizing $-x_1$ is the same as maximizing x_1, we can identify the optimal point as $\bar{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

b)

We will change around Equation 2 to be $(1-x_1)^3 - x_2 \ge 0$ for it to fit into the FJ conditions. We know that $u_i g_i(\bar{x}) = 0$ for all $i = 1, ..., m$. Hence we can calculate u_i for all $i = 1, ..., m$ as

$$\begin{aligned} u_1 g_1(\bar{x}) &= u_1 \cdot (1-1)^3 - 0 = 0 \implies 0 = 0 && (5)\\ u_2 g_2(\bar{x}) &= u_2 \cdot 1 = 0 \implies u_2 = 0 && (6)\\ u_3 g_3(\bar{x}) &= u_3 \cdot 0 = 0 \implies 0 = 0 && (7) \end{aligned}$$

$$0 = u_0 \nabla f(\bar{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) \tag{8}$$

$$0 = u_0(-1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}) + u_1 \begin{bmatrix} 3(1-x_1)^2 \\ -1 \end{bmatrix} + u_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + u_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{9}$$

$$0 = u_0(\begin{bmatrix} -1 \\ 0 \end{bmatrix}) + u_1 \begin{bmatrix} 0 \\ -1 \end{bmatrix} + u_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + u_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{10}$$

$$0 = \begin{bmatrix} -u_0 + u_2 \\ -u_1 + u_3 \end{bmatrix} \tag{11}$$

$$\implies \begin{cases} u_0 = u_2 = 0 \\ u_1 = u_3 \end{cases} \tag{12}$$

The point $\bar{x}$ is a FJ point, since we can choose u_1 and u_3 such that FJ conditions are satisfied. U is thus $u = \begin{bmatrix} 0 \\ t \\ 0 \\ t \end{bmatrix}, \quad t > 0.$

c)

The KKT conditions are

$$\nabla f(\bar{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) = 0, \tag{13}$$

$$u_i g_i(x) = 0, \qquad \forall i \tag{14}$$

$$u_i \geq 0, \qquad \forall i \tag{15}$$

which gives us

$$\begin{bmatrix} -1 \\ 0 \end{bmatrix} + u_1 \begin{bmatrix} 0 \\ -1 \end{bmatrix} + u_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{16}$$

$$= \begin{bmatrix} -1 \\ u_3 - u_1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{17}$$

Hence, KKT conditons are not satisfied for any u.

In order for LIQC to hold, the gradient of all active inequality constraints and all equality constraints needs to be linearly indepdendent. We can clearly see that, at $\bar{x}$, $\nabla g_1(\bar{x}) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ and $\nabla g_3(\bar{x}) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are linearly dependent.

For Slater's QC to hold, all inequality costraints needs to be convex in the feasible region. Since g_2 and g_3 are linear functions, we know that they are convex. We will examine the Hessian for g_1 to determine it's convexity.

$$H(g_1(x)) = \begin{bmatrix} -6(1-x_1) & 0 \\ 0 & 0 \end{bmatrix} \tag{18}$$

Since the Hessian for g_1 is not positive semi-definite in all of the feasible reagion, e.g. at $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, Slater's QC are not satisfied.

3.2 KKT Conditions for a Quadratic Problem

a)

$$\begin{aligned}
\text{min. } & (x_1 + \frac{9}{4})^2 + (x_2 - 2)^2 && (19)\\
\text{subject to: } & x_2 - x_1^2 \geq 0 \iff x_1^2 - x_2 \leq 0 && (20)\\
& x_1 + x_2 \leq 6 \iff x_1 + x_2 - 6 \leq 0 && (21)\\
& x_1 \geq 0 \iff -x_1 \leq 0 && (22)\\
& x_2 \geq 0 \iff -x_2 \leq 0 && (23)
\end{aligned}$$

The KKT conditions for the problem is

$$\begin{aligned}
& \nabla f(\bar{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) = 0 && (24)\\
& u_i g_i(x) = 0, \qquad \forall i && (25)\\
& u_i \geq 0, \qquad \forall i && (26)
\end{aligned}$$

From 25 we get the following at $\bar{x}$:

$$\begin{aligned}
u_1 \left(\left(\frac{3}{2}\right)^2 - 9/4 \right) &= u_1 \cdot 0 = 0 && (27)\\
u_2 \left(\frac{3}{2} + 9/4 - 6 \right) &= -\frac{9}{4} u_2 \implies u_2 = 0 && (28)\\
u_3 \left(-\frac{3}{2} \right) &= -\frac{3}{2} u_3 \implies u_3 = 0 && (29)\\
u_4 \left(-\frac{9}{4} \right) &= -\frac{9}{4} u_4 \implies u_4 = 0 && (30)
\end{aligned}$$

We can see that $u_2, u_3, u_4 = 0$ and the condition holds for $u_1 \geq 0$. Using this we can calculate the condition 24.

$$\begin{aligned}
0 &= \begin{bmatrix} 2(\bar{x}_1 - \frac{9}{4}) \\ 2(\bar{x}_2 - 2) \end{bmatrix} + u_1 \begin{bmatrix} 2\bar{x}_1 \\ -1 \end{bmatrix} + u_2 \begin{bmatrix} -1 \\ -1 \end{bmatrix} + u_3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + u_4 \begin{bmatrix} 0 \\ 1 \end{bmatrix} && (31)\\
0 &= \begin{bmatrix} 2(\frac{3}{2} - \frac{9}{4}) \\ 2(\frac{9}{4} - 2) \end{bmatrix} + u_1 \begin{bmatrix} 2\frac{3}{2} \\ -1 \end{bmatrix} && (32)\\
0 &= \begin{bmatrix} 2(\frac{3}{2} + \frac{9}{4}) + 3u_1 \\ 2(\frac{9}{4} - 2) - u_1 \end{bmatrix} && (33)
\end{aligned}$$

Solving the system of equations above we see that $u_1 = \frac{1}{2}$ solves the problem.

b)

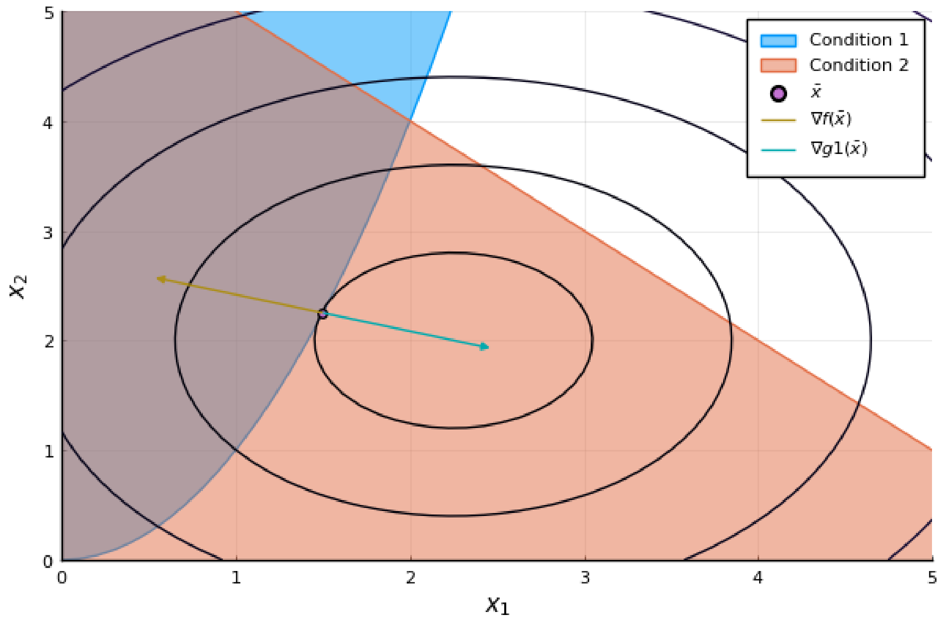

Figure 2

From the picture we can see that the optimal point is in the feasible region. The only active constraint is g_1. The gradient of the objective function and the constraint is plotted at the optimal point and we can see that the of equal size and opposite directions. Thus, we can graphically conclude, that, at the point, KKT conditions hold.

c)

For KKT conditions to be sufficient for global optimality, the feasible region and the objective function needs to be convex and Slater's QC needs to hold. The hessian of the objective function is $H(f(x)) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$. Since it is positive semi-definite everywhere, the objective function is convex. The same can be seen for the first condition. Since the other conditions are linear, they are also convex.

For SQC to be satisfied, in addition to what is already shown above, we need to show that there exists a x such that $g_i(x) < 0$. We can see that for example $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ fulfills the needed condition for all conditions.

With this information we can conclude that the point is globally optimal.

3.3 Lagrangian Dual of a Least-Squares Problem

a)

$$\text{min. } x^\top x \tag{34}$$

$$\text{subject to: } Ax = b \tag{35}$$

For our problem, The Lagrangian function and dual function becomes, since it has no inequality constraints,

$$\phi(x, v) = f(x) + v^\top h(x) \tag{36}$$
$$\theta(v) = \inf_x\{\phi(x, v) : x \in \mathbb{R}^n\} \tag{37}$$

The Lagrangian dual problem thus becomes $\sup_v \theta(v)$ subject to $v \geq 0$ or

$$\sup_v \inf_x\{x^\top x + v^\top(-Ax + b)\} \tag{38}$$
$$= \sup_v\{\inf_x\{x^\top x - v^\top Ax)\} + v^\top b\} \tag{39}$$
$$\text{s.t.} \quad v \geq 0 \tag{40}$$

Inside the infimum, we have $(x^\top - v^\top A)x$. This can be made arbitrarily small. In order for us to have a solution, we need to make $x^\top - v^\top A = 0$. Thus can the dual problem be written as:

$$\sup \quad v^\top b \tag{41}$$
$$\text{s.t.} \quad x^\top - v^\top A = 0 \tag{42}$$
$$v \geq 0 \tag{43}$$
$$v, x \in \mathbb{R}^n \tag{44}$$

b)

Let us start by solving the dual variable v by looking at the first constraint.

$$x^\top - v^\top A = 0 \tag{45}$$
$$v^\top A = x^\top \quad |\text{transpose} \tag{46}$$
$$A^\top v = x \tag{47}$$
$$AA^\top v = Ax \tag{48}$$
$$AA^\top v = b \tag{49}$$
$$v = (AA^\top)^{-1} b \tag{50}$$

We assumed, since $AA^\top$ is a n-by-n square matrix, that it is invertible. From this we also get that $x = (AA^\top)^{-1}bA^\top$ from the original constraint equation.

For strong duality to hold we need to check a number of conditions. Firstly, if the set is a nonempty set. We know $x \in \mathbb{R}$, so the set is nonempty. Secondly, the objective function and the inequality constraints need to be convex. Since the objective function, $x^\top x$, is a quadratic function do we know that it's convex. Since there are no inequality constraints, we do not need to check that. Next, we need to check weather all equality constraints are affine. Since the only equality constraint is $b - Ax$, can we clearly see that it is affine as well. Finally, Slater's QC needs to be satisfied. We have already shown that most parts of Slater's QC are fulfilled, we only need to know weather there exists such a x in the set that $g_i(x) < 0$. Since we have no inequality constraints, the condition is satisfied. Thus, Slater's QC holds, all the conditions are satisfied and we know that the problem has strong duality.

3.4 Concavity of Lagrangian Dual Functions

$$\theta(w) := \inf\{f(x) + w^\top \beta(x) : x \in X\} \tag{51}$$

By the definition of a concave (or convex) function, the $f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y)$ where $\lambda \in [0, 1]$. Let us start by calculating the left-hand side

$$\theta(\lambda w_1 + (1 - \lambda)w_2) = \inf\{f(x) + (\lambda w_1 + (1 - \lambda)w_2)^\top \beta\}. \tag{52}$$

We split $f(x) = \lambda f(x) + (1-\lambda) f(x)$, multiply in the β-term and continue our calculation

$$\theta(\lambda w_1 + (1-\lambda) w_2) = \inf\{\lambda f(x) + (1-\lambda) f(x) + \lambda w_1^\top \beta + (1-\lambda) w_2^\top \beta\} \tag{53}$$

$$\theta(\lambda w_1 + (1-\lambda) w_2) = \inf\{\lambda (f(x) + w_1^\top \beta) + (1-\lambda)(f(x) + w_2^\top \beta)\}. \tag{54}$$

Since we know that infimum is a concave function we get

$$\inf\{\lambda (f(x) + w_1^\top \beta) + (1-\lambda)(f(x) + w_2^\top \beta)\} \geq \lambda \inf\{(f(x) + w_1^\top \beta)\} + (1-\lambda) \inf\{(f(x) + w_2^\top \beta)\} \tag{55}$$

$$= \lambda \theta(w_1) + (1-\lambda)\theta(w_2). \tag{56}$$

This implies that

$$\theta(\lambda w_1 + (1-\lambda) w_2) \geq \lambda \theta(w_1) + (1-\lambda)\theta(w_2), \tag{57}$$

which shows that, according to the definition above, the function is concave.