Optimalisering og regulering TTK4135 - Assignment 1

Ola Kristoffer Hoff

26. January 2022

Problem 1

Go through Example 12.3 in detail in Nocedal & Wright. Replace the objective function with $x_1 + 2x_2$.

a Find the optimal point by inspecting the feasible area and the objective function.

As in the example it is clear to see that the solution is at $\begin{bmatrix} -\sqrt{2} \\ 0 \end{bmatrix}$, as both constraints are active and no direction d in the feasible area is in a decent direction.

Given that the criteria for a direction d is in a feasible decent direction is:

$$\nabla c_i(x)^T d \geq 0, \quad i \in \mathcal{I} = \{1, 2\}, \quad \nabla f(x)^T d < 0, \tag{1}$$
$$c_1(x) = 2 - x_1^2 - x_2^2 \geq 0, \tag{2}$$
$$c_2(x) = x_2 \geq 0 \tag{3}$$

We see that both constraints are active. This gives us the feasible area for d being a vector between $\nabla c_1 = \begin{bmatrix} -2x_1 \\ -2x_2 \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix}$ and $\nabla c_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

However, we can see that $\nabla f(x) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ also lays in the feasible area, and hence d is not a decent direction since $\nabla f(x)^T d \geq 0$.

b Check the KKT conditions at the optimal point.

To check the KKT conditions at the optimal point, $x^* = \begin{bmatrix} -\sqrt{2} \\ 0 \end{bmatrix}$, we use the Lagrangian function.

$$\mathcal{L}(x, \lambda) = f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x) \tag{4}$$

In our case it is:

$$\mathcal{L}(x, \lambda) = f(x) - \lambda_1 c_1(x) - \lambda_2 c_2(x), \quad \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \tag{5}$$

The conditions are:

$$\nabla_x \mathcal{L}(x^*, \lambda^*)) = 0, \tag{6}$$
$$c_i(x^*) = 0, \quad for\ all\ i \in \mathcal{E}, \tag{7}$$
$$c_i(x^*) \geq 0, \quad for\ all\ i \in \mathcal{I}, \tag{8}$$
$$\lambda_i^* \geq 0, \quad for\ all\ i \in \mathcal{I}, \tag{9}$$
$$\lambda_i^* c_i(x^*) = 0, \quad for\ all\ i \in \mathcal{E} \cup \mathcal{I}. \tag{10}$$
$$\tag{11}$$

Since $\mathcal{E} = \emptyset$ condition 2 is fulfilled. And since all conditions are active, $\mathcal{A} = \mathcal{I}$, condition 3 and 5 is fulfilled, since all conditions are equal to 0. This leaves condition 1 and 4 to be checked.

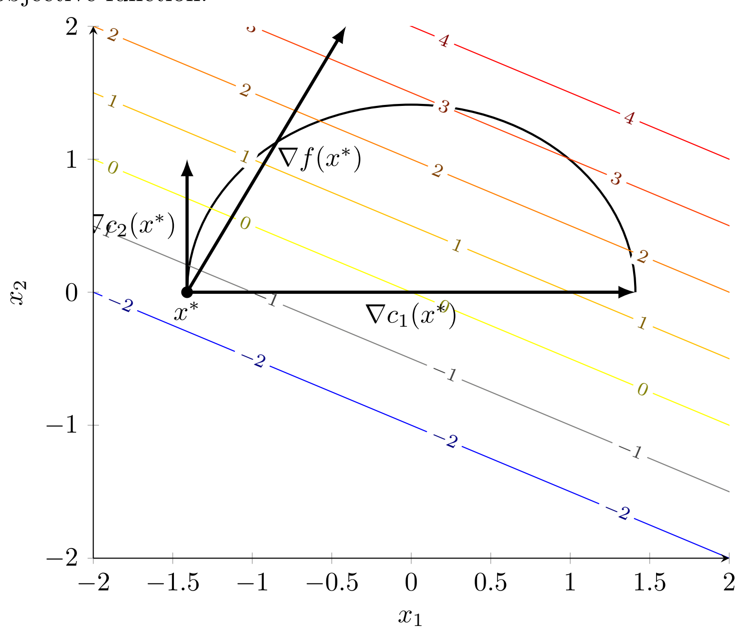
Let's solve for λ^* to fulfil condition 1.

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0 \tag{12}$$
$$\nabla_x \mathcal{L}(x^*, \lambda^*) = \nabla f(x^*) - \lambda_1^* \nabla c_1(x^*) - \lambda_2^* \nabla c_2(x^*) = 0 \tag{13}$$
$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} - \lambda_1^* \begin{bmatrix} -2 * -\sqrt{2} \\ -2 * 0 \end{bmatrix} - \lambda_2^* \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \tag{14}$$
$$\begin{bmatrix} 1 - \lambda_1^*(2\sqrt{2}) \\ 2 - \lambda_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{15}$$
$$\begin{bmatrix} \frac{1}{2\sqrt{2}} \\ 2 \end{bmatrix} = \begin{bmatrix} \lambda_1^* \\ \lambda_2^* \end{bmatrix} = \lambda^* \tag{16}$$
$$\lambda^* = \begin{bmatrix} \frac{1}{2\sqrt{2}} \\ 2 \end{bmatrix} \tag{17}$$

Given this λ^* condition 1 is fulfilled, and we also see that condition 4 is fulfilled. All conditions are fulfilled.

c Illustrate the gradients of the active constraints and the objective function at the solution.

Figure 1: Illustration of the gradients of the constraints and objective function at the solution point, x^*. Also a contour plot of the objective function.



d Explain why the Lagrange multipliers are positive.

First of we see from the KKT conditions that all Lagrange multipliers for inequality constraints must be positive. However, to explain why the Lagrange multipliers, λ, are positive we need the criteria:

$$\nabla c_i(x)^T d \geq 0, \quad i \in \mathcal{I}, \quad \nabla c_i(x)^T d = 0, \quad i \in \mathcal{E}, \quad \nabla f(x)^T d < 0, \tag{18}$$

These, as stated earlier, simply mean that direction d must be in the feasible area and in a decent direction.

Using this knowledge we can see that all gradients of constraints, ∇c, form a vector space. A subset of this space is the feasible area. More precisely we remove the vector space where a vector is a linear combination of a vector which has a negative scalar.

In our problem the gradient constraint vectors at the optimal point are: $\nabla c_1 = \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix}$ and $\nabla c_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

These are trivial vectors to see form a vector space. c_1 describes the x_1 direction and c_2 the x_2 direction. However, to get the vector $d = \begin{bmatrix} -2\sqrt{2} \\ -2 \end{bmatrix}$, we must have $d = -\nabla c_1 - 2\nabla c_2$. Since d is a created by the negative of the gradient constraints its obvious that its not in the feasible area defined by $\nabla c_i(x)^T d \geq 0$.

Then we can see that in this function:

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = \nabla f(x^*) - \lambda_1^* \nabla c_1(x^*) - \lambda_2^* \nabla c_2(x^*) = 0 \tag{19}$$

$$\nabla f(x^*) = \lambda_1^* \nabla c_1(x^*) + \lambda_2^* \nabla c_2(x^*) \tag{20}$$

The gradient of the objective function, f, must be equal to a linear combination, of only positive scalars, of the gradient constraints. This is only possible if the objective function is in the feasible area, which defines where d is. If this is the case, then d must be in the same direction as f and hence not in a decent direction, meaning no further decent is possible.

To sum up, the Lagrange multipliers must be positive because the direction d is defined as a linear combination of the gradients of the constraints, with only positive scalars, to give the feasible area. If we solve for λ and any of them are negative, that means f is not in the feasible area and we can have a d that are in a decent direction, and hence the optimal point might not be found yet.

e Is this problem a convex problem? Substantiate your answer.

Yes, this is a convex problem. To start with we can look at a complete circle, ignoring c_2 for now. It is quite easy to see that any two points chosen on or inside the circle has a line between them that is completely within the circle. Then if we add c_2 we simply split the circle in two and have the top half. Now for any pair of points where one is in the top half and the other in the bottom half, observe that the line crosses the constraint line of c_2. That means that if we move the point in the bottom half to where the line intercept with c_2. The rest of the line from that point to the point in the top half is completely within the constraint area. Thus, all points on the line c_2 and any point in the top half have a line between them inside the half circle. Hence, the problem is a convex problem.

A bit more formally we can say that since the objective function, f, and all the active constraints are concave functions, the problem is convex. f is a linear function, which is always convex, the same with c_2. However, c_1 describes a circle, a parabolic function, but since it is an inequality the interior of the circle is also in the feasible set. It's easy to see that two points on or in a circle has a line between them completly contained in the circle too.

Problem 2

Go through Example 12.1 in detail in Nocedal & Wright. Replace the objective function with $2x_1 + x_2$.

a Find all extreme points.

$$min\ x_1 + 2x_2 \quad s.t. \quad x_1^2 + x_2^2 - 2 = 0 \tag{21}$$

The extreme points can be found where $\nabla f(x^*) \parallel \nabla c_1(x^*)$.

$$\nabla f(x) = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \tag{22}$$
$$\nabla c_1(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} \tag{23}$$
$$\nabla f(x^*) = \lambda_1 c_1(x^*) \tag{24}$$
$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \lambda_1 \begin{bmatrix} 2x_1^* \\ 2x_2^* \end{bmatrix} \tag{25}$$
$$\begin{bmatrix} 2 - \lambda_1 2x_1^* \\ 1 - \lambda_1 2x_2^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{26}$$
$$x_1^* = 2x_2^* \tag{27}$$
$$(2x_2^*)^2 + x_2^{*2} - 2 = 0 \tag{28}$$
$$5x_2^{*2} - 2 = 0 \tag{29}$$
$$x_2^{*2} = \frac{2}{5} \tag{30}$$
$$x_2^* = \pm\sqrt{\frac{2}{5}} \tag{31}$$
$$(x_1^* = 2\sqrt{\frac{2}{5}} \wedge x_2^* = \sqrt{\frac{2}{5}}) \vee (x_1^* = 2\sqrt{\frac{2}{5}} \wedge x_2^* = -\sqrt{\frac{2}{5}}) \vee \tag{32}$$
$$(x_1^* = -2\sqrt{\frac{2}{5}} \wedge x_2^* = \sqrt{\frac{2}{5}}) \vee (x_1^* = -2\sqrt{\frac{2}{5}} \wedge x_2^* = -\sqrt{\frac{2}{5}}) \tag{33}$$
$$\tag{34}$$

Since λ_1 is a single value and $\nabla f(x^*)$ goes in the positive direction in both x_1 and x_2, the only solutions are:

$$(x_1^* = 2\sqrt{\frac{2}{5}} \wedge x_2^* = \sqrt{\frac{2}{5}}) \vee (x_1^* = -2\sqrt{\frac{2}{5}} \wedge x_2^* = -\sqrt{\frac{2}{5}}) \tag{35}$$
$$\tag{36}$$

We don't need to find λ_1, we know that we can scale $\nabla c_1(x^*)$ at these points to be equal to $\nabla f(x^*)$, hence they are parallel and we have found the extreme points.

b Check the KKT conditions at these points. Are the KKT conditions satisfied at these points? Explain why.

In our case it is:

$$\mathcal{L}(x, \lambda) = f(x) - \lambda_1 c_1(x), \quad \lambda = \begin{bmatrix} \lambda_1 \end{bmatrix} \tag{37}$$

The conditions are:

$$\nabla_x \mathcal{L}(x^*, \lambda^*)) = 0, \tag{38}$$
$$c_i(x^*) = 0, \quad for\ all\ i \in \mathcal{E}, \tag{39}$$
$$c_i(x^*) \geq 0, \quad for\ all\ i \in \mathcal{I}, \tag{40}$$
$$\lambda_i^* \geq 0, \quad for\ all\ i \in \mathcal{I}, \tag{41}$$
$$\lambda_i^* c_i(x^*) = 0, \quad for\ all\ i \in \mathcal{E} \cup \mathcal{I}. \tag{42}$$
$$\tag{43}$$

For both points we can see that condition 2 is fulfilled by checking that the points are in fact on the circle. $\mathcal{I} = \emptyset$ so condition 3 and 4 is also fulfilled. It also follows that condition 5 is fulfilled since condition 2 is fulfilled and $\mathcal{I} = \emptyset$. The only thing remaining is to check if condition 1 is fulfilled.

We check the point:

$$
\begin{aligned}
(x_1^* = 2\sqrt{\frac{2}{5}} \wedge x_2^* = \sqrt{\frac{2}{5}}) && (44)\\
\nabla_x \mathcal{L}(x^*, \lambda^*) = 0 && (45)\\
\nabla_x \mathcal{L}(x^*, \lambda^*) = \nabla f(x^*) - \lambda_1^* \nabla c_1(x^*) = 0 && (46)\\
\begin{bmatrix} 2 \\ 1 \end{bmatrix} - \lambda_1^* \begin{bmatrix} 2 * 2\sqrt{\frac{2}{5}} \\ 2 * \sqrt{\frac{2}{5}} \end{bmatrix} = 0 && (47)\\
\begin{bmatrix} 2 - \lambda_1^*(4\sqrt{\frac{2}{5}}) \\ 1 - \lambda_1^* 2\sqrt{\frac{2}{5}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} && (48)\\
\begin{bmatrix} 1 - \lambda_1^* 2\sqrt{\frac{2}{5}} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} && (49)\\
\begin{bmatrix} 1 - \lambda_1^* 2\sqrt{\frac{2}{5}} \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix} && (50)\\
\begin{bmatrix} \frac{1}{2\sqrt{\frac{2}{5}}} \end{bmatrix} = \begin{bmatrix} \lambda_1^* \end{bmatrix} && (51)\\
\lambda^* = \begin{bmatrix} \frac{1}{2\sqrt{\frac{2}{5}}} \end{bmatrix} && (52)
\end{aligned}
$$

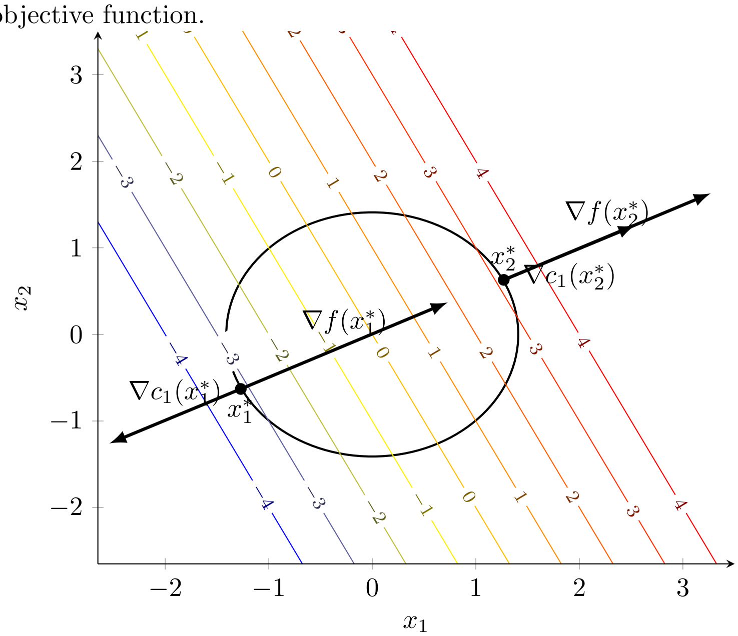
For this point the conditions are satisfied. Now to check the other point.

$$
\begin{aligned}
(x_1^* = -2\sqrt{\frac{2}{5}} \wedge x_2^* = -\sqrt{\frac{2}{5}}) && (53)\\
\nabla_x \mathcal{L}(x^*, \lambda^*) = 0 && (54)\\
\nabla_x \mathcal{L}(x^*, \lambda^*) = \nabla f(x^*) - \lambda_1^* \nabla c_1(x^*) = 0 && (55)\\
\begin{bmatrix} 2 \\ 1 \end{bmatrix} - \lambda_1^* \begin{bmatrix} 2 * -2\sqrt{\frac{2}{5}} \\ 2 * -\sqrt{\frac{2}{5}} \end{bmatrix} = 0 && (56)\\
\begin{bmatrix} 2 + \lambda_1^*(4\sqrt{\frac{2}{5}}) \\ 1 + \lambda_1^* 2\sqrt{\frac{2}{5}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} && (57)\\
\begin{bmatrix} 1 + \lambda_1^* 2\sqrt{\frac{2}{5}} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} && (58)\\
\begin{bmatrix} 1 + \lambda_1^* 2\sqrt{\frac{2}{5}} \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix} && (59)\\
\begin{bmatrix} \frac{1}{2\sqrt{\frac{2}{5}}} \end{bmatrix} = \begin{bmatrix} -\lambda_1^* \end{bmatrix} && (60)\\
\lambda^* = \begin{bmatrix} -\frac{1}{2\sqrt{\frac{2}{5}}} \end{bmatrix} && (61)\\
\Rightarrow && (62)\\
\lambda^* = \pm\frac{\sqrt{10}}{4} && (63)
\end{aligned}
$$

This point is satisfying the KKT conditions. The λ-values can be not be negative for inequalities, but this is equality constraints.

c Illustrate the gradients of the active constraint and the objective function at the optimal point(s).

Figure 2: Illustration of the gradients of the constraints and objective function at the solution point, x^*. Also a contour plot of the objective function.



By looking at the figure it is obvious that x_1 is the optimal solution point.

d What is the value of the Lagrange multiplier? Is this consistent with the KKT conditions?

As shown in subsection b, the Lagrange multiplier is: $\lambda^* = \pm\frac{\sqrt{10}}{4}$. Since we only have equality constraints this holds for the KKT conditions.

e Check the 2nd order conditions for the extreme points. How does this relate to the theory in Chapter 12.5?

The theory in 12.5 explains that from the first derivative we cannot tell if a point will increase or decrease in value if it moves in any direction in the feasible area. We know it's exclusively doing either one. This is what the second derivative can check for us, and we can separate maximums from minimums.

We have from earlier:

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = \begin{bmatrix} 2 - 2\lambda^* x_1^* \\ 1 - 2\lambda^* x_2^* \end{bmatrix} \tag{64}$$

$$\nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) = \begin{bmatrix} -2\lambda^* & 0 \\ 0 & -2\lambda^* \end{bmatrix} > 0 \tag{65}$$

The greater than zero comes from the fact that if the second derivatives is greater than zero all feasible directions from x^* must be increasing, hence we are in a minimum.

Plugging in our λ-values from earlier we see that only $\lambda = -\frac{\sqrt{10}}{4}$ fulfils the second order condition, hence x_1^* must be the minimum point, as is obvious to confirm from the illustration.

f Is this problem a convex problem? Substantiate your answer.

No, this problem is not convex. The same logic as for the previous problem still stands. The objective function, f, is still concave, since it's linear. And c_1 is a circle. The difference here is that c_1 is now an equality constraint, and the interior of the circle is then not included. Thus, the problem is not convex. An easy proof by example is that points $\begin{bmatrix} -\sqrt{2} \\ 0 \end{bmatrix}$ and $\begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix}$

have the point (center of circle) $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ between them. Since the center of the circle is not a part of the feasible set, it's not a convex problem.

Problem 3

Solve Problem 12.19 a), b), c) and d) in Nocedal & Wright.

$$\underset{x \in \mathbb{R}^2}{min} \ f(x) = -2x_1 + x_2 \quad s.t. \begin{cases} (1-x_1)^3 - x_2 & \geq 0 \\ x_2 + 0.25x_1^2 - 1 & \geq 0 \end{cases} \tag{66}$$

The optimal solution is $x^* = (0,1)^T$, where both constraints are active.

a Do the LICQ hold at this point?

LICQ holds if the set of active constraints gradients are linearly independent.

$$\nabla c_1(x^*) = \begin{bmatrix} -3(1-x_1^*)^2 \\ -1 \end{bmatrix} = \begin{bmatrix} -3 \\ -1 \end{bmatrix} \tag{67}$$

$$\nabla c_2(x^*) = \begin{bmatrix} 0.50x_1^* \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{68}$$

$$\tag{69}$$

$$\nabla c_1(x^*) \neq \lambda \nabla c_2(x^*) \tag{70}$$

Since there is no constant λ that can transform c_2 in to c_1 they are linearly independent, and hence the LICQ condition holds at point x^*.

b Are the KKT conditions satisfied?

In our case the Lagrangian function is:

$$\mathcal{L}(x,\lambda) = f(x) - \lambda_1 c_1(x) - \lambda_2 c_2(x), \quad \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \tag{71}$$

The KKT conditions are:

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0, \tag{72}$$

$$c_i(x^*) = 0, \quad for\ all\ i \in \mathcal{E}, \tag{73}$$

$$c_i(x^*) \geq 0, \quad for\ all\ i \in \mathcal{I}, \tag{74}$$

$$\lambda_i^* \geq 0, \quad for\ all\ i \in \mathcal{I}, \tag{75}$$

$$\lambda_i^* c_i(x^*) = 0, \quad for\ all\ i \in \mathcal{E} \cup \mathcal{I}. \tag{76}$$

$$\tag{77}$$

We only have inequality constraints in this problem so $\mathcal{E} = \emptyset$. Then we see that condition 2 is satisfied. We can also easly see that condition 3 is satisfied by inserting x^* in the inequality constraints.

$$c_1(x^*) = (1-1)^3 - 1 = 0 \tag{78}$$

$$c_2(x^*) = 1 + 0.25 * 0^2 - 1 = 0 \tag{79}$$

Since both constraints are equal to zero condition 5 is also satisfied. We can now check condition 1, and find a λ value to check condition 4.

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = \nabla f(x) - \lambda_1 \nabla c_1(x) - \lambda_2 \nabla c_2(x) = 0 \tag{80}$$

$$\begin{bmatrix} -2 \\ 1 \end{bmatrix} - \lambda_1 \begin{bmatrix} -3 \\ -1 \end{bmatrix} - \lambda_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \tag{81}$$

$$\begin{bmatrix} -2 + 3\lambda_1 \\ 1 + \lambda_1 - \lambda_2 \end{bmatrix} = 0 \tag{82}$$

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ 1 + \lambda_1) \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ 1 + \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{5}{3} \end{bmatrix} \tag{83}$$

$$\lambda^* = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{5}{3} \end{bmatrix} \tag{84}$$

As we can see, we found a value for λ, so condition 1 is satisfied. And the λ-values are all positive, meaning condition 4 is satisfied. All KKT conditions are satisfied.

c Write down the sets $\mathcal{F}(x^*)$ and $\mathcal{C}(x^*, \lambda^*)$.

The feasible set $\mathcal{F}(x)$ is defined as (12.3 N&W):

$$\mathcal{F}(x) = \begin{cases} d| \ d^T \nabla c_i(x) = 0, & \forall i \in \mathcal{E} \\ d| \ d^T \nabla c_i(x) \geq 0, & \forall i \in \mathcal{A}(x) \cap \mathcal{I} \end{cases} \tag{85}$$

Since $\mathcal{E} = \emptyset$ the first one is ruled out. We know from earlier that both our inequality constraints are zero at x^* which means they are in the active set, $\mathcal{A}$.

The feasible set will be the vector filed of the non-negative linear combination of the constraints.

Let's look at it more formally, we use the gradient values for the constraints from the previous subsections.

$$d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \tag{86}$$

$$\tag{87}$$

$$d^T \nabla c_1(x^*) \geq 0 \tag{88}$$

$$\begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \end{bmatrix} \geq 0 \tag{89}$$

$$-3d_1 - d_2 \geq 0 \tag{90}$$

$$\tag{91}$$

$$d^T \nabla c_2(x^*) \geq 0 \tag{92}$$

$$\begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \geq 0 \tag{93}$$

$$d_2 \geq 0 \tag{94}$$

$$\tag{95}$$

$$\tag{96}$$

This shows us that as long as d_2 is non-negative and d_1 is greater (absolute value) then a third of d_2 and negative, we have a d in the feasible set.

$$\mathcal{F}(x^*) = \Big\{ d| \ d_1 \leq -\tfrac{d_2}{3} \wedge d_2 \geq 0 \tag{97}$$

Now to find the set $\mathcal{C}(c^*, \lambda^*)$. We have the definition of the critical cone (N&W p. 330):

$$\mathcal{C}(x^*, \lambda^*) = \{ w \in \mathcal{F}(x^*) | \nabla c_i(x^*)^T w = 0, \text{ all } i \in \mathcal{A}(x^*) \cap \mathcal{I} \text{ with } \lambda_i^* > 0 \} \tag{98}$$

Now we do a very similar process as we did above with d.

$$w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \tag{99}$$

$$\tag{100}$$

$$\nabla c_1(x^*)^T w = 0 \tag{101}$$

$$\begin{bmatrix} -3 & -1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0 \tag{102}$$

$$-3w_1 - w_2 = 0 \tag{103}$$

$$\tag{104}$$

$$\nabla c_2(x^*)^T w = 0 \tag{105}$$

$$\begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = 0 \tag{106}$$

$$w_2 = 0 \tag{107}$$

$$\tag{108}$$

$$\tag{109}$$

We check both constraints since they are both inequalities and in the active set (equal to zero). We also had a positive λ^*. Since w_2 is equal to zero, we just insert that in the first equation we get and find that w_1 also must be zero. Then we check that it is indeed in the feasible set too, by the rules above. d_2 is greater than or equal to zero, so that's OK. d_1 must be less than or equal to $-d_2$ divided by three, which is zero, and hence this is too OK. This is the only vector in $\mathcal{C}$, at the point x^*.

$$\mathcal{C}(x^*, \lambda^*) = \{(w_1, w_2)^T | w_1 = 0 \cap w_2 = 0\} = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \tag{110}$$

d Are the second-order necessary conditions satisfied? Are the second-order sufficient conditions satisfied?

First we look at the necessary conditions. This uses the second derivatives, and state that: "If x^* is a local solution, then the Hessian of the Lagrangian has nonnegative curvature along cirtical directions (that is, the directions in $\mathcal{C}(x^*, \lambda^*)$)." (N&W, p.332).

We have the Lagrangian from earlier:

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = \begin{bmatrix} -2 + 3\lambda_1^*(1 - x_1^*)^2 - 0.50\lambda_2^* x_1^* \\ 1 + \lambda_1^* - \lambda_2^* \end{bmatrix} \tag{111}$$

$$\nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) = \begin{bmatrix} -6\lambda_1^*(1 - x_1^*) - 0.50\lambda_2^* & 0 \\ 0 & 0 \end{bmatrix} \tag{112}$$

$$\nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) = \begin{bmatrix} -\frac{29}{6} & 0 \\ 0 & 0 \end{bmatrix} \tag{113}$$

The theorem states that:

$$w^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) w \geq 0, \qquad \forall w \in \mathcal{C}(x^*, \lambda^*) \tag{114}$$

Using the w from the critical cone above:

$$w^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) w \geq 0, \qquad \forall w \in \mathcal{C}(x^*, \lambda^*) \tag{115}$$

$$\begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} -\frac{29}{6} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} \geq 0 \tag{116}$$

$$0 \geq 0 \tag{117}$$

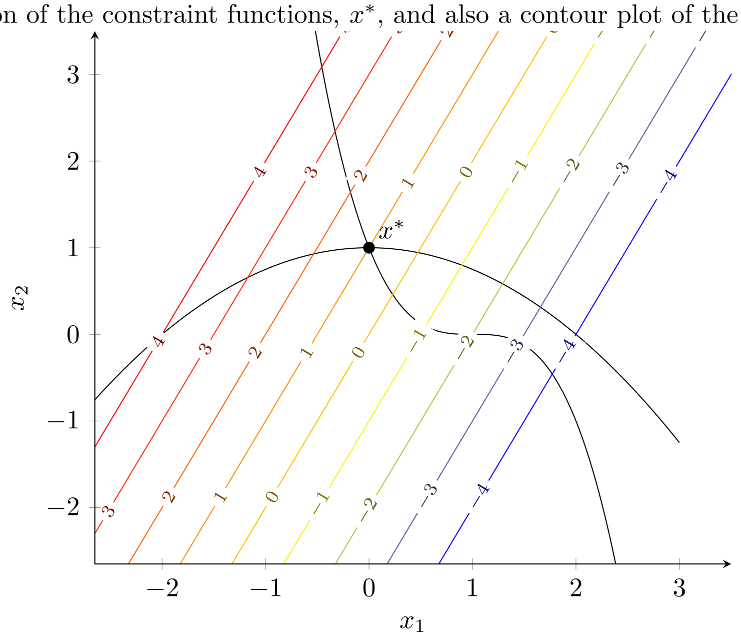
Thus it is quite clear to see that the necessary condition is met. Now to look at the sufficient condition. This condition check if the local solution, x^*, is a strict local solution. This meaning that all feasible points around it is in an incline. The maths is almost the same, we just say that it must be strictly greater than zero.

The theorem states that:

$$w^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*) w > 0, \qquad \forall w \in \mathcal{C}(x^*, \lambda^*), w \neq 0 \tag{118}$$

Since we only have one possible w which is zero, which will not be checked by definition here. There are no other w's to check, and we can thus conclude that x^* is a strict local solution. This is also clear to see from the illustration I have added below.

Figure 3: Illustration of the constraint functions, x^*, and also a contour plot of the objective function.

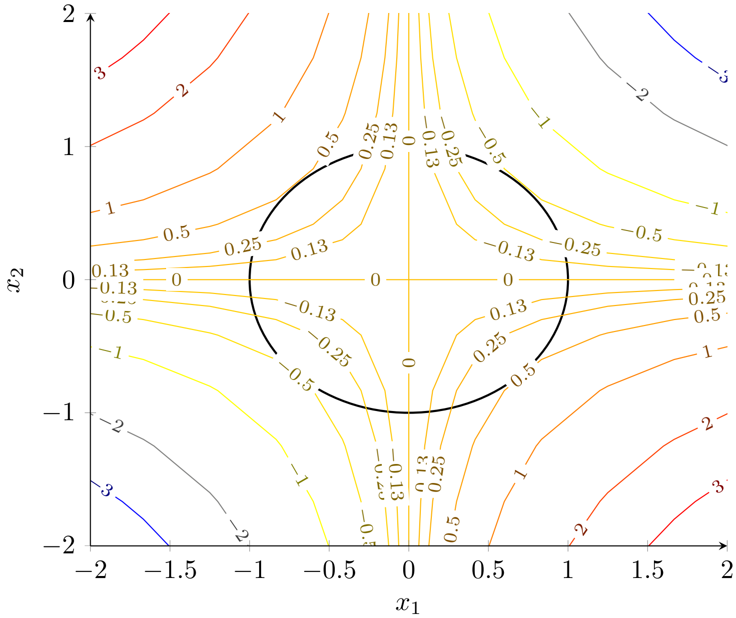


Problem 4

Find the maxima of $f(x) = x_1 x_2$ over the unit disk defined by the inequality constraint $1 - x_1^2 - x_2^2 \geq 0$.

First of we will convert the maximum problem into a minimum problem. This due to the fact that is easy and we can then apply the same techniques as earlier. We then solve for the minimum of: $f(x) = -x_1 x_2$. Let's first visualise the problem:

Figure 4: Illustration of the constraint functions and also a contour plot of the objective function.



It is from the visuals very clear to see that it's a saddle point around the origin and that the minimum will be found in the "corners" (45 degrees) of the circle in the top-right and bottom-left. Now let's prove it.

So we think that the optimal points are $1-x_1^2-x_2^2=0,$ where $x_1=x_2$. This solves to be $x_1=x_2=\pm\sqrt{\frac{1}{2}}$. Since it's in the top-right or bottom-left it's either both negative or both positive coordinates, giving us: $x_1^* = \begin{bmatrix} -\sqrt{\frac{1}{2}} \\ -\sqrt{\frac{1}{2}} \end{bmatrix}$ and $x_2^* = \begin{bmatrix} \sqrt{\frac{1}{2}} \\ \sqrt{\frac{1}{2}} \end{bmatrix}$.

In our case the Lagrangian function is:

$$\mathcal{L}(x,\lambda) = f(x) - \lambda_1 c_1(x), \quad \lambda = \begin{bmatrix}\lambda_1\end{bmatrix} \tag{119}$$

$$\mathcal{L}(x,\lambda) = -x_1x_2 - \lambda_1(1 - x_1^2 - x_2^2) \tag{120}$$

The KKT conditions are:

$$\nabla_x \mathcal{L}(x^*,\lambda^*) = 0, \tag{121}$$

$$c_i(x^*) = 0, \quad for\ all\ i \in \mathcal{E}, \tag{122}$$

$$c_i(x^*) \geq 0, \quad for\ all\ i \in \mathcal{I}, \tag{123}$$

$$\lambda_i^* \geq 0, \quad for\ all\ i \in \mathcal{I}, \tag{124}$$

$$\lambda_i^* c_i(x^*) = 0, \quad for\ all\ i \in \mathcal{E} \cup \mathcal{I}. \tag{125}$$

$$\tag{126}$$

The 2 condition is met because $\mathcal{E} = \emptyset$. Condition 3 and 5 is fulfilled since we choose points on the circle (constraint is equal to zero). Then we only have the 1 and 4 condition left.

Let's check the first point, x_1^*, with the first condition:

$$\nabla f(x_1^*) = \begin{bmatrix} -x_{12}^* \\ -x_{11}^* \end{bmatrix} = \begin{bmatrix} \sqrt{\frac{1}{2}} \\ \sqrt{\frac{1}{2}} \end{bmatrix} \tag{127}$$

$$\nabla c_1(x_1^*) = \begin{bmatrix} -2x_{11}^* \\ -2x_{12}^* \end{bmatrix} = \begin{bmatrix} 2\sqrt{\frac{1}{2}} \\ 2\sqrt{\frac{1}{2}} \end{bmatrix} = \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \end{bmatrix} \tag{128}$$

$$\nabla \mathcal{L}(x_1^*,\lambda^*) = \nabla f(x_1^*) - \lambda^* \nabla c_1(x^*) = 0 \tag{129}$$

$$\begin{bmatrix} \sqrt{\frac{1}{2}} \\ \sqrt{\frac{1}{2}} \end{bmatrix} - \lambda^* \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \end{bmatrix} = \begin{bmatrix} \sqrt{\frac{1}{2}} - \lambda^*\sqrt{2} \\ \sqrt{\frac{1}{2}} - \lambda^*\sqrt{2} \end{bmatrix} = 0 \Rightarrow \lambda^* = \frac{1}{2} \tag{130}$$

$$\tag{131}$$

The same line of logic applies to the other point, x_2^*, only negating the terms, but its still the same answer because of the double negative from using both f and c. This means that condition 1 and 4 of the KKT conditions are fulfilled as well.

Hence, the points x_1^* and x_2^* are both optimal solutions to the problem. I have illustrated the gradient of the objective function and the constraint for the two points below.

Figure 5: Illustration of the constraint functions and also a contour plot of the objective function, as well as the gradient of the functions in point x_1^*.

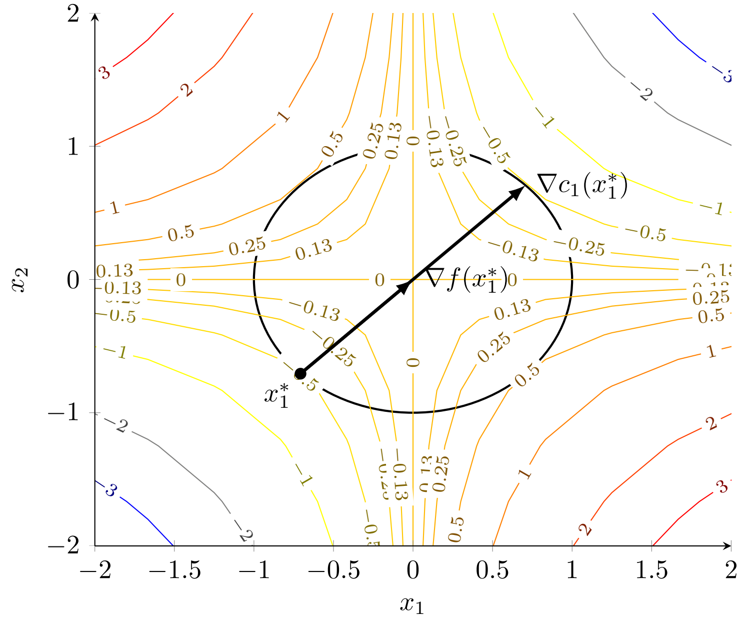


Figure 6: Illustration of the constraint functions and also a contour plot of the objective function, as well as the gradient of the functions in point x_2^*.