Technical University of Crete

School of Electrical and Computer Engineering

Course: **Optimization**

Exercise 3 (100/1000)

Angelopoulos Dimitris, 2020030038

1. In this problem we will compute the projection of $\mathbf{x}_0 \in \mathbb{R}^n$ onto the set $\mathbf{B}(\mathbf{0}, r) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 \leq r\}$

   (a) First of all, we will draw a scheme of the problem in MATLAB.

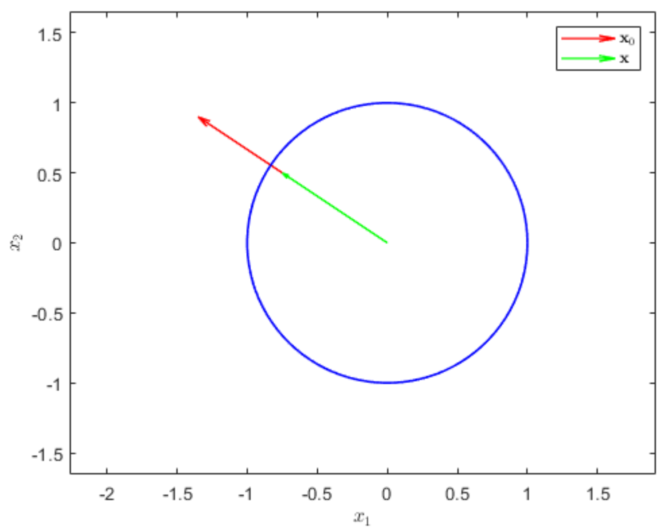

Fig. 1

   (b) The equations we must solve are :

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) = \|\mathbf{x}_0 - \mathbf{x}\|_2^2 \\ \text{s.t.} & f_1(\mathbf{x}) = \|\mathbf{x}\|_2^2 - r^2 \leq 0 \end{array}$$

(c) To write down the KKT conditions we first need to compute the gradients of f_0 and f_1.

$$\nabla f_0(\mathbf{x}) = 2(\mathbf{x} - \mathbf{x}_0)$$

$$\nabla f_1(\mathbf{x}) = 2\mathbf{x}$$

The KKT conditions are :

- $\nabla f_0(\mathbf{x}_*) + \lambda_* \nabla f_1(\mathbf{x}_*) = \mathbf{0}$
- $\lambda_* \geq 0$
- $f_1(\mathbf{x}_*) \leq 0$
- $\lambda_* f_1(\mathbf{x}_*) = 0$

(d) We consider the case where $\lambda_* > 0$. Then using the first condition and solving for $\mathbf{x}_*$ we get :

$$\mathbf{x}_* = \frac{1}{1 + \lambda_*} \mathbf{x}_0 \tag{1}$$

Then substituting to the fourth condition and solving for λ_* we get :

$$\lambda_* = \frac{\|\mathbf{x}_0\|}{r} - 1 \tag{2}$$

Using (2) equation (1) becomes :

$$\mathbf{x}_* = \frac{r}{\|\mathbf{x}_0\|} \mathbf{x}_0$$

This is the projection of $\mathbf{x}_0$ to the set $\mathbf{B}(\mathbf{0}, r)$ for $\lambda_* > 0$. We can see the projection in Fig. 1. We observe that if $\mathbf{x}_0 \notin \mathbf{B}(\mathbf{0}, r)$ the projection of $\mathbf{x}_0$ is the vector scaled in a way that makes it to belong to $\mathbf{B}(\mathbf{0}, r)$.

(e) Now we consider the case where $\lambda_* = 0$. For this case the we solve $\nabla f_0(\mathbf{x}_*) = 0$ to get :

$$\mathbf{x}_* = \mathbf{x}_0$$

In this case it is obvious that $\mathbf{x}_0 \in \mathbf{B}(\mathbf{0}, r)$ and that is the reason the projection of $\mathbf{x}_0$ is itself.

2. In this problem we will repeat the previous steps to compute the projection of $\mathbf{x}_0 \in \mathbb{R}^n$ onto the set $\mathbf{B}(\mathbf{y}, r) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{y}\|_2 \leq r\}$

(a) As in the previous problem we draw a scheme of the problem as shown below.

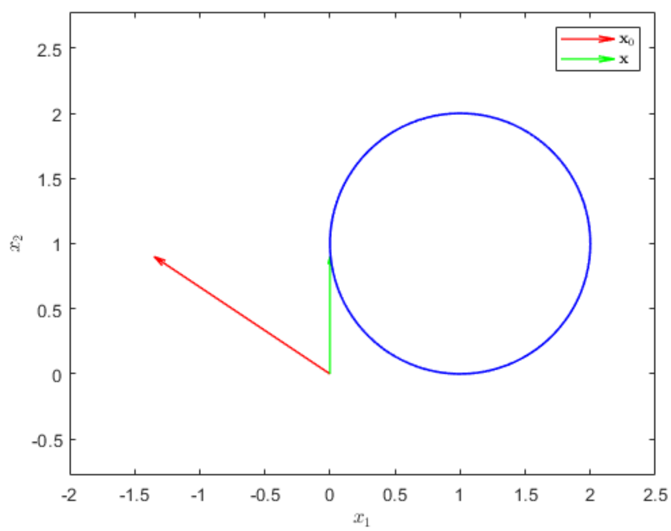

Fig. 2

(b) The optimization problem is :

$$\begin{aligned} \text{minimize} \quad & f_0(\mathbf{x}) = \|\mathbf{x}_0 - \mathbf{x}\|_2^2 \\ \text{s.t.} \quad & f_1(\mathbf{x}) = \|\mathbf{x} - \mathbf{y}\|_2^2 - r^2 \leq 0 \end{aligned}$$

(c) The KKT conditions are the same with the ones in the previous problem for :

$$\nabla f_0(\mathbf{x}) = 2(\mathbf{x} - \mathbf{x}_0)$$

$$\nabla f_1(\mathbf{x}) = 2(\mathbf{x} - \mathbf{y})$$

(d) We consider the case where $\lambda_* > 0$. Then using the first condition and solving for $\mathbf{x}_*$ we get :

$$\mathbf{x}_* = \frac{\mathbf{x}_0 + \lambda_* \mathbf{y}}{1 + \lambda_*} \tag{3}$$

Now substituting (3) to the fourth KKT condition and doing some calculations we can find λ_* in closed form. That is :

$$\lambda_* = \frac{\|\mathbf{x}_0 - \mathbf{y}\|}{r} - 1 \tag{4}$$

Using (4), equation (3) becomes :

$$\mathbf{x}_* = \frac{r}{\|\mathbf{x}_0 - \mathbf{y}\|}(\mathbf{x}_0 - \mathbf{y}) + \mathbf{y}$$

(e) Now we consider the case where $\lambda_* = 0$. For this case the we solve $\nabla f_0(\mathbf{x}_*) = 0$ to get :

$$\mathbf{x}_* = \mathbf{x}_0$$

As in the previous problem, $\mathbf{x}_0 \in \mathbf{B}(\mathbf{y}, r)$ and thus the projection of $\mathbf{x}_0$ is itself.

3. Let $\mathbf{a} \in \mathbb{R}^n$. We will compute the projection of $\mathbf{x}_0 \in \mathbb{R}^n$ onto the set $\mathbb{S} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a} \leq \mathbf{x}\}$.

   The optimization problem that we have to solve becomes :

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\|\mathbf{x}_0 - \mathbf{x}\|_2^2 \\ \text{s.t.} \quad & a_i - x_i \leq 0 \end{aligned}$$

   The KKT conditions for this case are :

   - $\nabla f_0(\mathbf{x}_*) + \sum_{i=1}^{n} \lambda_i \nabla f_i(\mathbf{x}_*) = \mathbf{0}$
   - $\lambda_i \geq 0,\ i \in \{1, ..., n\}$
   - $f_i(\mathbf{x}_*) \leq 0,\ i \in \{1, ..., n\}$
   - $\lambda_i f_i(\mathbf{x}_*) = 0,\ i \in \{1, ..., n\}$

   To start of, we will compute the gradients as follows :

$$\nabla f_0(\mathbf{x}) = \mathbf{x} - \mathbf{x}_0$$

$$\nabla f_i(\mathbf{x}) = \nabla(\mathbf{c}_i^T(\mathbf{a} - \mathbf{x})) = -\mathbf{c}_i$$

   Where $\mathbf{c}_i$ is the vector that its inner product with another vector $\mathbf{v}$ gives us the i-th element of the last, i.e $\mathbf{c}_i^T \mathbf{v} = v_i$.

   Substituting to the first KKT condition we get :

$$\mathbf{x}_* - \mathbf{x}_0 - \sum_{i=1}^{n} \lambda_i \mathbf{c}_i = \mathbf{0} \Rightarrow$$

$$\mathbf{x}_* - \mathbf{x}_0 - \lambda = 0 \Rightarrow$$

$$x_{*,i} = x_{0,i} + \lambda_i \tag{5}$$

Now we distinguish the following cases :

(i) $\lambda_i = 0$

Substituting to (5) we get $x_{*,i} = x_{0,i}$ which means that $\mathbf{x}_0 \in \mathbb{S}$.

(ii) $\lambda_i > 0$

Using the fourth KKT condition we find that :

$$\lambda_i = a_i - x_{o,i} \tag{6}$$

Substituting equation (6) to (5) we get that $x_{*,i} = a_i$.

The projection can be written in a more compact form using the function *max* as :

$$x_{*,i} = max\{a_i, x_{0,i}\}$$

or

$$\mathbf{x}_* = max\{\mathbf{a}, \mathbf{x}_0\}$$

4. Let $\mathbf{A} \in \mathbb{R}^{p \times n}$, $\mathbf{b} \in \mathbb{R}^p$ and $\mathbf{x}_0 \in \mathbb{R}^n$. We will solve the following optimization problem with affine constraints.

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\|\mathbf{x}_0 - \mathbf{x}\|_2^2 \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned}$$

The KKT conditions are :

- $\mathbf{x}_* - \mathbf{x}_0 + \mathbf{A}^T\mathbf{v} = \mathbf{0}$
- $\mathbf{A}\mathbf{x}_* - \mathbf{b} = \mathbf{0}$

We can re-write the conditions above in the following form :

$$\begin{bmatrix} \mathbb{I}_{n\times n} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{O} \end{bmatrix} \begin{bmatrix} \mathbf{x}_* \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{b} \end{bmatrix} \tag{7}$$

The matrix $\mathbf{F} = \begin{bmatrix} \mathbb{I}_{n\times n} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{O} \end{bmatrix}$ has dimensions $(n+p)\times(n+p)$ and is invertible. Thus equation (7) has a closed form solution of the form :

$$\begin{bmatrix} \mathbf{x}_* \\ \mathbf{v} \end{bmatrix} = \mathbf{F}^{-1} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{b} \end{bmatrix} \Rightarrow$$

$$\mathbf{x}_* = \mathbf{F}^{-1}(1:n,:) \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{b} \end{bmatrix}$$

Notice that the projection is given by the multiplication of the first n rows of the inverse of $\mathbf{F}$ which is an $n \times (n+p)$ matrix, times the vector $\begin{bmatrix} \mathbf{x}_0 \\ \mathbf{b} \end{bmatrix} \in \mathbb{R}^{n+p}$.

5. Let $\mathbb{S} = \{\mathbf{x} \in \mathbb{R}^2 \mid \|\mathbf{x}\| \leq 1, x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 1\}$. We will solve the problem :

$$(P) \ \underset{\mathbf{x} \in \mathbb{S}}{\text{minimize}} \ f_0(\mathbf{x})$$

First of all we will draw a figure with every constraint to check if there exists a redundant constraint.

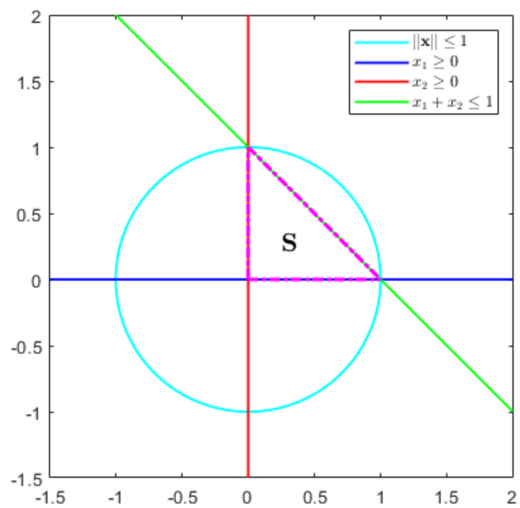

Fig. 3

We observe that the constraint $\|\mathbf{x}\| \leq 1$ is redundant. In that case we can simplify $\mathbb{S}$ to be :

$$\mathbb{S} = \{\mathbf{x} \in \mathbb{R}^2 | x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 1\}$$

Problem (P) can be re written in the following form.

$$\begin{aligned} \text{minimize} \quad & f_0(\mathbf{x}) \\ \text{s.t.} \quad & -x_1 \leq 0 \\ & -x_2 \leq 0 \\ & x_1 + x_2 \leq 1 \end{aligned}$$

(a) $f_0(\mathbf{x}) = (x_1 - 2)^2 + (x_2 - 2)^2$

The KKT conditions are :

- $\nabla f_0(\mathbf{x}_*) + \sum_{i=1}^{3} \lambda_i \nabla f_i(\mathbf{x}_*) = \mathbf{0}$
- $\lambda_i \geq 0,\ i \in \{1, ..., n\}$
- $f_i(\mathbf{x}_*) \leq 0,\ i \in \{1, ..., n\}$
- $\lambda_i f_i(\mathbf{x}_*) = 0,\ i \in \{1, ..., n\}$

Where the gradients from above are :

$$\nabla f_0(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_0}{\partial x_1} \\ \frac{\partial f_0}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 - 4 \\ 2x_2 - 4 \end{bmatrix} \tag{8}$$

All 3 constraints are of the form $f_i(\mathbf{x}) = \mathbf{c}_i^T \mathbf{x} + b_i \Rightarrow \nabla f_i(\mathbf{x}) = \mathbf{c}_i$. Thus the gradients of the constraints are :

$$\nabla f_1(\mathbf{x}) = \begin{bmatrix} -1 \\ 0 \end{bmatrix},\ \nabla f_2(\mathbf{x}) = \begin{bmatrix} 0 \\ -1 \end{bmatrix},\ \nabla f_3(\mathbf{x}) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{9}$$

Using equations (8), (9) and substituting to the first KKT condition, we get the following system, of equations :

$$\begin{aligned} 2x_1 - 4 - \lambda_1 + \lambda_3 = 0 \\ 2x_2 - 4 - \lambda_2 + \lambda_3 = 0 \end{aligned} \Rightarrow \begin{aligned} x_1 = \tfrac{\lambda_1 - \lambda_3 + 4}{2} \\ x_2 = \tfrac{\lambda_2 - \lambda_3 + 4}{2} \end{aligned} \tag{10}$$

The second pair of equations comes from the fourth KKT condition

$$\begin{aligned} \lambda_1 x_1 = 0 \\ \lambda_2 x_2 = 0 \end{aligned} \tag{11}$$

Now we distinct the following cases.

(1) For $\lambda_3 = 0$ ($f_3(\mathbf{x}) \leq 0$ is not an active constraint)

That means that our values x_1 and x_2 from equation (10) become :

$$x_1 = \frac{\lambda_1 + 4}{2}, \quad x_2 = \frac{\lambda_2 + 4}{2}$$

Now we will examine the different scenarios for λ_1 and λ_2.

i. $\lambda_1 = \lambda_2 = 0$

That means that $x_1 = x_2 = 2$. From the third KKT condition we have :

$$x_1 + x_2 - 1 \leq 0 \Rightarrow 2 + 2 - 1 = 3 \leq 0$$

That means that the constraints can not be all inactive at the same time, by contradiction.

ii. $\lambda_1 > 0$ and $\lambda_2 = 0$

From equation (11) $x_1 = 0 \Rightarrow \lambda_1 = -4 < 0$, contradiction. This can not be the case.

iii. $\lambda_1 = 0$ and $\lambda_2 > 0$

Likewise $x_2 = 0 \Rightarrow \lambda_2 = -4 < 0$. This is not the case either.

iv. $\lambda_1 > 0$ and $\lambda_2 > 0$

For the same reason as (ii) and (iii) this case is not acceptable.

That means that $f_3(\mathbf{x}) \leq 0$ is an active constraint, meaning that the optimal lies on this line.

(2) For $\lambda_3 > 0$ we distinguish the same cases.

i. $\lambda_1 = \lambda_2 = 0$

From the fourth KKT condition we derive that $x_1 + x_2 = 1$. The values in this case are :

$$x_1 = \frac{\lambda_3 + 4}{2}, \quad x_2 = \frac{\lambda_3 + 4}{2}$$

The above equations allow us to find that $x_1 = x_2 = \frac{1}{2}$ for $\lambda_1 = \lambda_2 = 0$, $\lambda_3 = 3$, which all satisfy the KKT conditions meaning that the optimal point is :

$$\mathbf{x}_* = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

ii. $\lambda_1 = 0$ and $\lambda_2 > 0$

$$x_1 = \frac{-\lambda_3 + 4}{2}, \quad x_2 = \frac{\lambda_2 - \lambda_3 + 4}{2} = 0$$

and

$$x_1 = 1 - x_2 = 1 \Rightarrow \lambda_3 = 2$$

This means that $\lambda_2 = \lambda_3 - 4 = -2$ which is a contradiction with the second KKT condition.

iii. $\lambda_1 > 0$ and $\lambda_2 = 0$

Likewise with (ii) $\lambda_3 = 2 \Rightarrow \lambda_1 = -2$ we result in a contradiction.

iv. $\lambda_1 > 0$ and $\lambda_2 > 0$

This means that $x_1 = x_2 = 0$. Substituting to $x_1 + x_2 = 1 \Rightarrow 0 = 1$ which is a contradiction.

(b) $f_0(\mathbf{x}) = (x_1+2)^2 + (x_2+2)^2$

For this cost function the KKT conditions are the same with the previous one but here the gradient of f_0 is:

$$\nabla f_0(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_0}{\partial x_1} \\ \frac{\partial f_0}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_1+4 \\ 2x_2+4 \end{bmatrix} \tag{12}$$

Using the KKT conditions we get the following system of equations

$$\begin{aligned} x_1 &= \tfrac{\lambda_1-\lambda_3-4}{2} \\ x_2 &= \tfrac{\lambda_2-\lambda_3-4}{2} \\ \lambda_1 x_1 &= 0 \\ \lambda_2 x_2 &= 0 \\ \lambda_3(x_1+x_2-1) &= 0 \end{aligned} \tag{13}$$

Following the same methodology we distinguish 2 cases for λ_3.

(1) For $\lambda_3 > 0$ we have :

$$x_1 + x_2 = 1$$

i. $\lambda_1 = \lambda_2 = 0$

$$x_1 = x_2 = \frac{-\lambda_3 - 4}{2} \Rightarrow -\lambda_3 - 4 = 1 \Rightarrow \lambda_3 = -3$$

This can not be the case where the optimal point occurs since λ_3 must be positive.

ii. $\lambda_1 > 0$ and $\lambda_2 = 0$

$$x_1 = 0 \Rightarrow x_2 = \frac{-\lambda_3 - 4}{2} = 1 \Rightarrow \lambda_3 = -2$$

For the same reason as in case (i) this case is not acceptable.

iii. $\lambda_1 = 0$ and $\lambda_2 > 0$

$$x_2 = 0 \Rightarrow x_1 = \frac{-\lambda_3 - 4}{2} = 1 \Rightarrow \lambda_3 = -2$$

iv. $\lambda_1 > 0$ and $\lambda_2 > 0$

$$x_1 = x_2 = 0$$

The equation $x_1 + x_2 = 1$ is not satisfied.

That means that the optimal point occurs when $\lambda_3 = 0$ which means that the third constraint is inactive.

(2) For $\lambda_3 = 0$ we have :

$$x_1 = \frac{\lambda_1 - 4}{2},\ x_2 = \frac{\lambda_2 - 4}{2}$$

i. $\lambda_1 = \lambda_2 = 0$

$$x_1 = x_2 = -2$$

That can not happen since x_1 and x_2 must be positive.

ii. $\lambda_1 > 0$ and $\lambda_2 = 0$

$$x_2 = -2$$

Contradiction due to the third KKT condition.

iii. $\lambda_1 = 0$ and $\lambda_2 > 0$

$$x_1 = -2$$

Likewise with (ii).

iv. $\lambda_1 > 0$ and $\lambda_2 > 0$

$$x_1 = x_2 = 0 \Rightarrow \lambda_1 = \lambda_2 = 4$$

In the last case every KKT condition is satisfied. The optimal point and the corresponding lagrangian multipliers are :

$$\mathbf{x}_* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \ \lambda_1 = \lambda_2 = 4, \ \lambda_3 = 0$$

We can also observe that the active constraints are both the first and the second. Looking at Fig. 3, the only point where both of these constraints are active is the origin. Having said that, it is logical to compute that minimum point.