ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
TECHNICAL UNIVERSITY OF CRETE

Course “Optimization” Report

Exercise 3

Georgios Frangias 2018030086
`gfrangias@tuc.gr`

Course Professor:
Athanasios Liavas

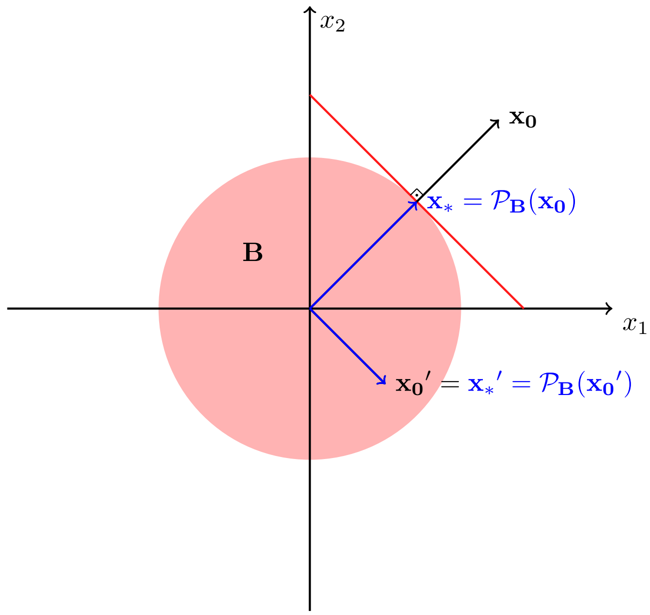
Technical University of Crete

School of Electrical & Computer Engineering

December 2022

1. 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) Draw a scheme of the problem.



(b) Write down the optimization problem you must solve, in terms of differentiable functions.

The optimization problem is the following

$$\begin{aligned} \text{minimize: } & f_0(\mathbf{x}) := ||\mathbf{x}_0 - \mathbf{x}||_2^2 \\ \text{subject to: } & f_i(\mathbf{x}) = x_i - r \leq 0 \quad i = 1, \dots, n \end{aligned}$$

The norms are squared because it doesn't affect the optimization problem and because it is easier to solve it this way.

(c) Write down the KKT conditions, in terms of the optimal parameters $\mathbf{x}_*$ and λ_*.

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

The first equation of the Langrangian multiplier can be analyzed as

$$\nabla f_0(\mathbf{x}_*) + \sum_{i=1}^{n} \lambda_{*,i} \nabla f_i(x_{*,i}) = 0 \Rightarrow$$

$$\Rightarrow 2(\mathbf{x}_* - \mathbf{x}_0) + \sum_{i=1}^{n} \lambda_{*,i} 2x_{*,i} = 0 \Rightarrow$$

$$\Rightarrow 2(\mathbf{x}_* - \mathbf{x}_0) + 2\lambda_* \mathbf{x}_* = 0 \Rightarrow \mathbf{x}_* = \frac{\mathbf{x}_0}{1+\lambda_*}$$

So the KKT conditions can be updated to

$$\left(1+\lambda_{*,i}\right) x_{*,i} = x_{0,i} \tag{1}$$
$$\lambda_{*,i}\left(x_{*,i} - r\right) = 0 \tag{2}$$
$$x_{*,i} - r \leq 0 \tag{3}$$
$$\lambda_{*,i} \geq 0 \tag{4}$$

(d) Consider the case $\lambda_* > 0$. What is the conclusion?

For $\lambda_* > 0$

$$(2) \Rightarrow ||\mathbf{x}_*||_2^2 = r^2 \tag{5}$$

$$(1) \xRightarrow{\cdot \mathbf{x}_*^T} \left(1+\lambda_*\right) \mathbf{x}_*^T \mathbf{x}_* = \mathbf{x}_*^T \mathbf{x}_0 \xRightarrow{(5)} \left(1+\lambda_*\right) r^2 = \mathbf{x}_*^T \mathbf{x}_0 \Rightarrow$$
$$\Rightarrow \lambda_* = \frac{\mathbf{x}_*^T \mathbf{x}_0 - r^2}{r^2} > 0 \Rightarrow \mathbf{x}_*^T \mathbf{x}_0 > r^2 \Rightarrow \left(\mathbf{x}_*^T \mathbf{x}_0\right)^2 > r^4 \Rightarrow$$
$$\xRightarrow{\text{Cauchy-Schwarz}} ||\mathbf{x}_*||_2^2 ||\mathbf{x}_0||_2^2 > r^4 \xRightarrow{(5)} ||\mathbf{x}_0||_2^2 > r^2$$

This means that for $\lambda_* > 0 \Rightarrow \mathbf{x}_0 \notin \mathbf{B}$

Let $\lambda_* = \frac{\mathbf{x}_*^T \mathbf{x}_0}{r^2} - 1$

$$(1) \Rightarrow \frac{\mathbf{x}_*^T \mathbf{x}_0}{r^2} \mathbf{x}_* = \mathbf{x}_0 \tag{6}$$

This means that $\mathbf{x}_*$ and $\mathbf{x}_0$ are co-linear and they can be described by the equation $\mathbf{x}_* = k\mathbf{x}_0$ with $k \in \mathbb{R}$. With this in mind

$$(6) \Rightarrow \frac{k^2}{r^2}\left(\mathbf{x}_0^T \mathbf{x}_0\right) \mathbf{x}_0 = \mathbf{x}_0 \Rightarrow k = \frac{r}{||\mathbf{x}_0||_2}$$

So,

$$\mathbf{x}_* = r \frac{\mathbf{x}_0}{||\mathbf{x}_0||_2}$$

(e) Consider the case $\lambda_* = 0$. What is the conclusion?

For $\lambda_* = 0$

$$(1) \Rightarrow \mathbf{x}_* = \mathbf{x}_0$$

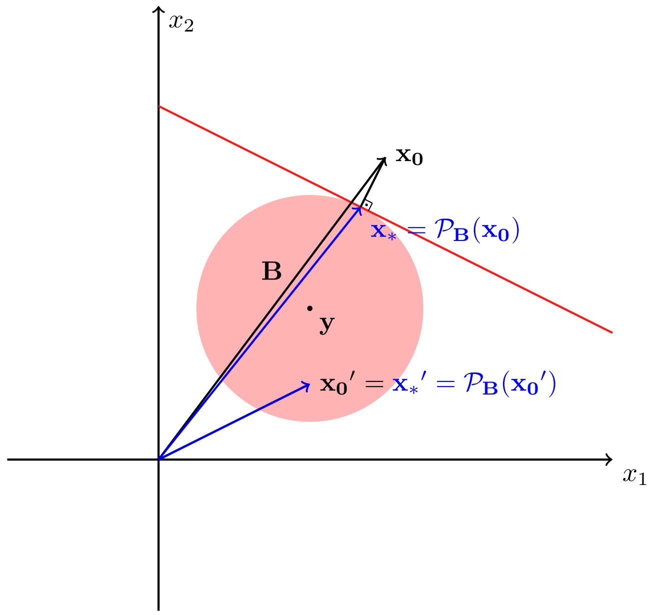
$$(3) \Rightarrow ||\mathbf{x}_*||_2^2 = ||\mathbf{x}_0||_2^2 \leq r^2$$

This means that for $\lambda_* = 0 \Rightarrow \mathbf{x}_0 \in \mathbf{B}$. So,

$$\mathbf{x}_* = \begin{cases} \mathbf{x}_0 & \lambda_* = 0 \\ r\frac{\mathbf{x}_0}{||\mathbf{x}_0||_2} & \lambda_* > 0 \end{cases}$$

2.Repeat the steps of the previous question and compute the projection of $\mathbf{x}_0 \in \mathbb{R}^n$ onto the set $\mathbf{B}(\mathbf{y}, r) := \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{y}\|_2 \leq r\}$ (for given $\mathbf{y} \in \mathbb{R}^n$ and $r \in \mathbb{R}_{++}$).

(a) Draw a scheme of the problem.



(b) Write down the optimization problem you must solve, in terms of differentiable functions.

The optimization problem is the following

$$\begin{aligned} \text{minimize: } & f_0(\mathbf{x}) := ||\mathbf{x}_0 - \mathbf{x}||_2^2 \\ \text{subject to: } & f_i(\mathbf{x}) = x_i - y_i - r \leq 0 \quad i = 1, \dots, n \end{aligned}$$

(c) Write down the KKT conditions, in terms of the optimal parameters $\mathbf{x}_*$ and λ_*.

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

The first equation of the Langrangian multiplier can be analyzed as

$$\nabla f_0(\mathbf{x}_*) + \sum_{i=1}^{n} \lambda_{*,i} \nabla f_i(x_{*,i}) = 0 \Rightarrow$$

$$\Rightarrow 2(\mathbf{x}_* - \mathbf{x}_0) + \sum_{i=1}^{n} \lambda_{*,i} 2\,(x_{*,i} - y_i) = 0 \Rightarrow$$

$$\Rightarrow 2(\mathbf{x}_* - \mathbf{x}_0) + 2\lambda_*\,(\mathbf{x}_* - \mathbf{y}) = 0 \Rightarrow \mathbf{x}_* = \frac{\mathbf{x}_0 + \lambda_* \mathbf{y}}{1 + \lambda_*}$$

So the KKT conditions can be updated to

$$\left(1+\lambda_{*,i}\right) x_{*,i} = x_{0,i} + \lambda_{*,i} y_i \tag{1}$$

$$\lambda_{*,i}\left(x_{*,i} - y_i - r\right) = 0 \tag{2}$$

$$x_{*,i} - y_i - r \leq 0 \tag{3}$$

$$\lambda_{*,i} \geq 0 \tag{4}$$

(d) Consider the case $\lambda_* > 0$. What is the conclusion?

For $\lambda_* > 0$

$$(2) \Rightarrow ||\mathbf{x}_* - \mathbf{y}||_2^2 = r^2 \tag{5}$$

$$(1) \Rightarrow \left(1+\lambda_*\right)\mathbf{x}_* = \mathbf{x}_0 + \lambda_*\mathbf{y} + \mathbf{y} - \mathbf{y} \Rightarrow \left(1+\lambda_*\right)\left(\mathbf{x}_* - \mathbf{y}\right) = \mathbf{x}_0 - \mathbf{y} \Rightarrow \tag{6}$$

$$\xRightarrow{\cdot\left(\mathbf{x}_* - \mathbf{y}\right)^T,(5)} \left(1+\lambda_*\right) r^2 = \left(\mathbf{x}_* - \mathbf{y}\right)^T\left(\mathbf{x}_0 - \mathbf{y}\right) \Rightarrow$$

$$\Rightarrow \lambda_* = \frac{\left(\mathbf{x}_* - \mathbf{y}\right)^T\left(\mathbf{x}_0 - \mathbf{y}\right) - r^2}{r^2} > 0 \Rightarrow \left(\mathbf{x}_* - \mathbf{y}\right)^T\left(\mathbf{x}_0 - \mathbf{y}\right) > r^2$$

$$\xRightarrow{\text{Cauchy-Schwarz}} ||\mathbf{x}_* - \mathbf{y}||_2^2 ||\mathbf{x}_0 - \mathbf{y}||_2^2 \geq \left(\left(\mathbf{x}_* - \mathbf{y}\right)^T\left(\mathbf{x}_0 - \mathbf{y}\right)\right)^2 > r^4 \xRightarrow{(5)} ||\mathbf{x}_0 - \mathbf{y}||_2^2 > r^2$$

This means that for $\lambda_* > 0 \Rightarrow \mathbf{x}_0 \notin \mathbf{B}$

Let $\lambda_* = \frac{\left(\mathbf{x}_* - \mathbf{y}\right)^T\left(\mathbf{x}_0 - \mathbf{y}\right)}{r^2} - 1$

$$(6) \Rightarrow \frac{\left(\mathbf{x}_* - \mathbf{y}\right)^T\left(\mathbf{x}_0 - \mathbf{y}\right)}{r^2}\left(\mathbf{x}_* - \mathbf{y}\right) = \mathbf{x}_0 - \mathbf{y} \tag{7}$$

This means that $\mathbf{x}_* - \mathbf{y}$ and $\mathbf{x}_0 - \mathbf{y}$ are co-linear and they can be described by the equation $\mathbf{x}_* - \mathbf{y} = k\left(\mathbf{x}_0 - \mathbf{y}\right)$ with $k \in \mathbb{R}$. With this in mind

$$(7) \Rightarrow \frac{k^2}{r^2}\left(\mathbf{x}_0 - \mathbf{y}\right)^T\left(\mathbf{x}_0 - \mathbf{y}\right)\left(\mathbf{x}_0 - \mathbf{y}\right) = \mathbf{x}_0 - \mathbf{y} \Rightarrow k = \frac{r}{||\mathbf{x}_0 - \mathbf{y}||_2}$$

So,

$$\mathbf{x}_* = r\frac{\mathbf{x}_0 - \mathbf{y}}{||\mathbf{x}_0 - \mathbf{y}||_2} + \mathbf{y}$$

(e) Consider the case $\lambda_* = 0$. What is the conclusion?

For $\lambda_* = 0$

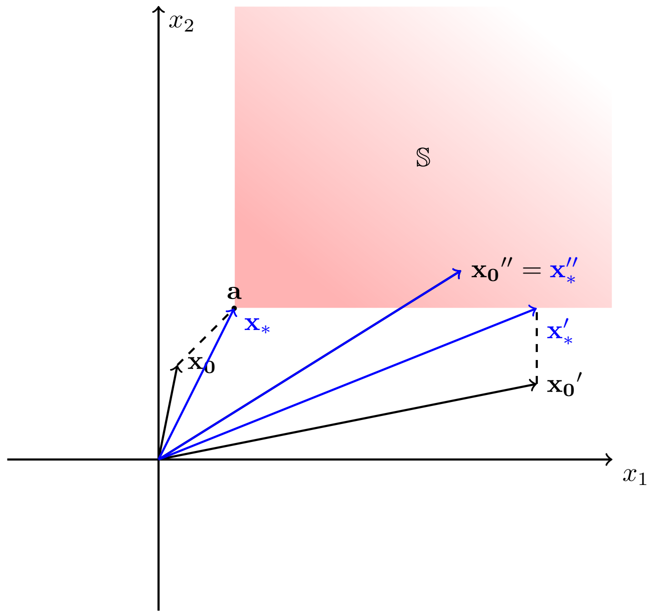
$$(1) \Rightarrow \mathbf{x}_* = \mathbf{x}_0$$

$$(3) \Rightarrow ||\mathbf{x}_* - \mathbf{y}||_2^2 = ||\mathbf{x}_0 - \mathbf{y}||_2^2 \leq r^2$$

This means that for $\lambda_* = 0 \Rightarrow \mathbf{x}_0 \in \mathbf{B}$. So,

$$\mathbf{x}_* = \begin{cases} \mathbf{x}_0 & \lambda_* = 0 \\ r\frac{\mathbf{x}_0 - \mathbf{y}}{||\mathbf{x}_0 - \mathbf{y}||_2} + \mathbf{y} & \lambda_* > 0 \end{cases}$$

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



$$\begin{aligned}\text{minimize: } & f_0(\mathbf{x}) := \tfrac{1}{2}||\mathbf{x}_0 - \mathbf{x}||_2^2 \\ \text{subject to: } & f_i(\mathbf{x}) = a_i - x_i \leq 0 \quad i = 1, \ldots, n\end{aligned}$$

The first equation of the Langrangian multiplier can be analyzed as

$$\nabla f_0(\mathbf{x}_*) + \sum_{i=1}^{n} \lambda_{*,i} \nabla f_i(x_{*,i}) = 0 \Rightarrow$$

$$\Rightarrow x_{*,i} - x_{0,i} - \lambda_{*,i} = 0 \Rightarrow x_{*,i} = x_{0,i} + \lambda_{*,i}$$

$$x_{*,i} = x_{0,i} + \lambda_{*,i} \tag{1}$$
$$\lambda_{*,i} \left(a_i - x_{*,i}\right) = 0 \tag{2}$$
$$a_i - x_{*,i} \leq 0 \tag{3}$$
$$\lambda_{*,i} \geq 0 \tag{4}$$

For $\lambda_{*,i} > 0$

$$(2) \Rightarrow x_{*,i} = a_i$$

$$(1) \Rightarrow a_i = x_{0,i} + \lambda_{*,i} \Rightarrow \lambda_{*,i} = a_i - x_{0,i} > 0$$

For $\lambda_{*,i} = 0$

$$(1) \Rightarrow x_{*,i} = x_{0,i}$$

$$(3) \Rightarrow a_i - x_{*,i} = a_i - x_{0,i} \leq 0$$

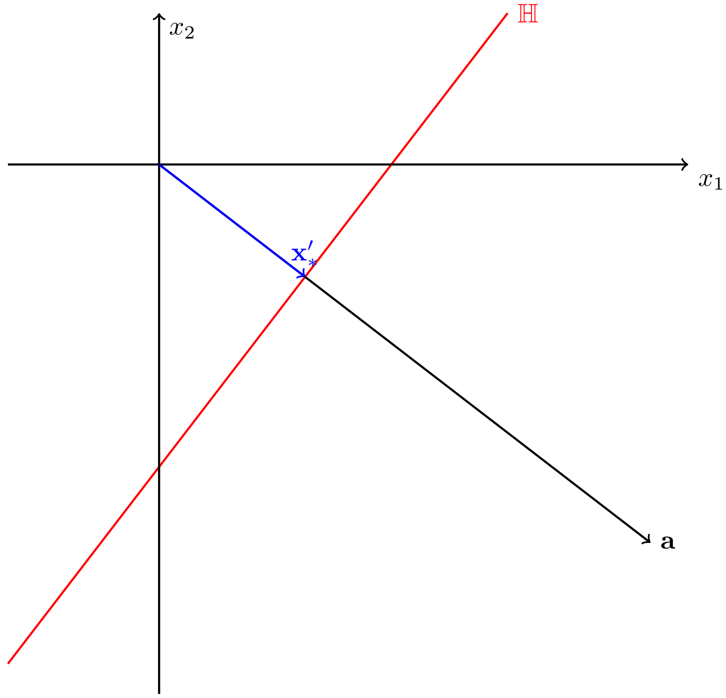
So if at least one $\lambda_{*,i} > 0$ then $x_{0,i} \notin \mathbb{R}$ and

$$x_{*,i} = \begin{cases} x_{0,i} & \lambda_{*,i} = 0 \\ a_i & \lambda_{*,i} > 0 \end{cases}$$

4. Let $\mathbf{0} \neq \mathbf{a} \in \mathbb{R}^n, b \in \mathbb{R}$ and consider the problem

$$\text{(P)} \qquad \min_{\mathbf{x}} f_0(\mathbf{x}) := \frac{1}{2}\|\mathbf{x}\|_2^2, \text{ subject to } \mathbf{x} \in \mathbb{H} := \left\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^T\mathbf{x} = b\right\} \tag{1}$$

(a) Write and solve the KKT for problem (P).



In problem (P) we actually need to prove the projection of the zero point $\mathbf{0}$ onto the hyperplane $\mathbb{H}$. The optimization problem (P) is subject to the equation $f_1(\mathbf{x}) = \mathbf{a}^T\mathbf{x} - b = 0$
The KKT conditions are

$$\nabla f_0(\mathbf{x}_*) + v_1 \nabla f_1(\mathbf{x}_*) = 0 \tag{2}$$

$$\mathbf{a}^T\mathbf{x}_* - b = 0 \tag{3}$$

The Langrangian is computed

$$(2) \Rightarrow \mathbf{x}_* + v_1\mathbf{a} = 0 \xRightarrow{\cdot \mathbf{a}^T} \mathbf{a}^T\mathbf{x}_* + v_1||\mathbf{a}||_2^2 = 0 \Rightarrow$$

$$\xRightarrow{(3)} v_1 = -\frac{b}{||\mathbf{a}||_2^2} \tag{4}$$

So,

$$(3) \xRightarrow{(4)} \mathbf{x}_* = \frac{b}{||\mathbf{a}||_2^2}\mathbf{a}$$

(b) Compute the solution of problem (P) using the projected gradient descent method

$$\mathbf{x}_{k+1} = \mathbf{P}_{\mathbb{H}}\left(\mathbf{x}_k - \frac{1}{L}\nabla f_0\left(\mathbf{x}_k\right)\right) \tag{5}$$

where $L := \max\left(\text{eig}\left(\nabla^2 f_0(\mathbf{x})\right)\right)$. What do you observe?
It is obvious that

$$\nabla f_0(\mathbf{x}_k) = \mathbf{x}_k$$

$$\nabla^2 f_0(\mathbf{x}_k) = \mathbf{I}_n$$

The projection of a point $\mathbf{x}_0$ onto the set $\mathbb{H}$ is already proven in the course notes' Example 6.7.2. as

$$\mathbf{x}_* = \mathbf{x}_0 - \frac{\mathbf{a}^T \mathbf{x}_0 - b}{||\mathbf{a}||_2^2}\mathbf{a}$$

Since, $\nabla f_0(\mathbf{x}_k) = \mathbf{x}_k$ and $L := \max\left(\mathrm{eig}\left(\nabla^2 f_0(\mathbf{x})\right)\right) = \max\left(\mathrm{eig}\left(\mathbf{I}_n\right)\right) = 1$

$$\mathbf{x}_{k+1} = \mathbf{P}_{\mathbb{H}}\left(\mathbf{x}_k - \mathbf{x}_k\right) = \mathbf{P}_{\mathbb{H}}\left(\mathbf{0}\right)$$

This means that the optimal point should be found on the first iteration since from (a)

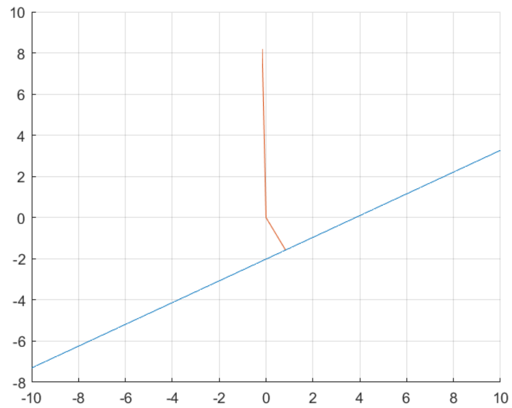
$$\mathbf{x}_1 = \mathbf{P}_{\mathbb{H}}\left(\mathbf{x}_0 - \mathbf{x}_0\right) = \mathbf{P}_{\mathbb{H}}\left(\mathbf{0}\right) = \mathbf{0} - \frac{\mathbf{a}^T \mathbf{0} - b}{||\mathbf{a}||_2^2}\mathbf{a} = \frac{b}{||\mathbf{a}||_2^2}\mathbf{a} = \mathbf{x}_*$$

This is verified computing using MATLAB in file `exercise4.m`. The projected gradient method is implemented in the file `projected_gradient_method.m`. The algorithm is implemented with the following steps

1. If $||\mathbf{x}_k - \mathbf{x}_{k-1}||_2 > \epsilon$
2. $\mathbf{x}_{k+1} = \mathbf{P}_{\mathbb{H}}\left(\mathbf{x}_k - \frac{1}{L}\nabla f_0\left(\mathbf{x}_k\right)\right)$

```
x=x0;
x_hist=x0;
x_prev = 10^10;
fun_val=f(x);
fun_val_hist=fun_val;
grad=g(x);
hessian=h(x);
L=max(eig(hessian));
iter=0;
while(norm(x-x_prev)>epsilon)
    iter=iter+1;
    x_prev=x;
    x_hist=[x_hist (x-grad/L)];
    x=p(x-grad/L);
    x_hist=[x_hist x];
    fun_val=f(x);
    fun_val_hist=[fun_val_hist fun_val];
    grad=g(x);
    hessian=h(x);
    L=max(eig(hessian));
    fprintf('iter_number = %3d norm_grad = %2.6f fun_val = %2.6f \n',...
    iter,norm(grad),fun_val);
end
```

As it is clear in the diagram below the projected gradient descent method starts at $\mathbf{x}_0$ and then moves at $\mathbf{0}$ and then at its projection on $\mathbb{H}$ which is the optimal point $\mathbf{x}_*$ according to `CVX`.



5. Let $\mathbf{A} \in \mathbb{R}^{p \times n}$, with $\operatorname{rank}(\mathbf{A}) = p$, and $\mathbf{b} \in \mathbb{R}^p$.

(a) Find the distance of a point $\mathrm{x}_0 \in \mathbb{R}^n$ from the set $\mathbb{S} := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{b}\}$ (you must compute the projection of x_0 onto $\mathbb{S}$).

Essentially, the optimization problem is the following,

$$
\begin{gathered}
\text{minimize: } f_0(\mathbf{x}) := \tfrac{1}{2}||\mathbf{x} - \mathbf{x}_0||_2^2 \\
\text{subject to: } f_1(\mathbf{x}) = \mathbf{Ax} - \mathbf{b} = 0 \\
\text{or} \\
\text{minimize: } f_0(\mathbf{x}) := \tfrac{1}{2}||\mathbf{x} - \mathbf{x}_0||_2^2 \\
\text{subject to: } f_i(\mathbf{x}) = \mathbf{a}_i^T\mathbf{x} - \mathbf{b} = 0 \quad i = 1, \ldots, n
\end{gathered}
$$

where $\mathbf{a}_i^T$ are the rows of the matrix $\mathbf{A}$. The KKT conditions are

$$\nabla f_0(\mathbf{x}_*) + \mathbf{v}_1 \nabla f_1(\mathbf{x}_*) = 0 \tag{1}$$

$$\mathbf{Ax}_* - \mathbf{b} = 0 \tag{2}$$

$$(1) \Rightarrow \mathbf{x}_* - \mathbf{x}_0 + \mathbf{A}^T\mathbf{v}_1 = 0 \Rightarrow \mathbf{x}_* = \mathbf{x}_0 - \mathbf{A}^T\mathbf{v}_1 \xRightarrow{\cdot\mathbf{A}} \mathbf{Ax}_* = \mathbf{Ax}_0 - \mathbf{AA}^T\mathbf{v}_1 \Rightarrow$$

$$\xRightarrow{(2)} \mathbf{v}_1 = \left(\mathbf{AA}^T\right)^{-1}(\mathbf{Ax}_0 - \mathbf{b}) \tag{3}$$

$$(1) \xRightarrow{(3)} \mathbf{x}_* - \mathbf{x}_0 = -\mathbf{A}\left(\mathbf{AA}^T\right)^{-1}(\mathbf{Ax}_0 - \mathbf{b}) \xRightarrow{(\mathbf{x}_* - \mathbf{x}_0)^T} ||\mathbf{x}_* - \mathbf{x}_0||_2^2 = (\mathbf{x}_0 - \mathbf{x}_*)^T\mathbf{A}\left(\mathbf{AA}^T\right)^{-1}(\mathbf{Ax}_0 - \mathbf{b})$$

$$\Rightarrow ||\mathbf{x}_* - \mathbf{x}_0||_2 = \sqrt{\left|(\mathbf{x}_0 - \mathbf{x}_*)^T\mathbf{A}(\mathbf{AA}^T)^{-1}(\mathbf{Ax}_0 - \mathbf{b})\right|}$$

So,

$$\mathbf{P}_{\mathbb{S}}(\mathbf{x}_0) = \sqrt{\left|(\mathbf{x}_0 - \mathbf{x}_*)^T\mathbf{A}(\mathbf{AA}^T)^{-1}(\mathbf{Ax}_0 - \mathbf{b})\right|}$$

where $\mathbf{x}_* = \mathbf{x}_0 - \mathbf{A}\left(\mathbf{AA}^T\right)^{-1}(\mathbf{Ax}_0 - \mathbf{b})$
In the MATLAB script `exercise5.m` the above is validated using `CVX`.

(b) Let the $(n \times n)$ positive definite matrix $\mathbf{P} = \mathbf{P}^T \succ \mathbf{0}, q \in \mathbb{R}^n$ and

$$f_0(\mathbf{x}) := \frac{1}{2}\mathbf{x}^T\mathbf{Px} + \mathbf{q}^T\mathbf{x}$$

Consider the problem

$$(\mathrm{Q}) \ \min_{\mathbf{x} \in \mathbb{S}} f_0(\mathbf{x})$$

i. Solve problem (Q) using `cvx`.

```
cvx_begin
    variable x(n)
    minimize( norm(x-x0) )
    subject to
        A*x-b==0;
cvx_end
```

ii. Write the KKT for problem (Q) and compute the optimal solution by solving them using `matlab` (no `cvx`).

$$\begin{cases} \nabla f_0(\mathbf{x}_*) + \mathbf{v}_1 \nabla f_1(\mathbf{x}_*) = 0 \\ \mathbf{A}\mathbf{x}_* - \mathbf{b} = 0 \end{cases} \Rightarrow \begin{cases} \mathbf{P}\mathbf{x}_* + \mathbf{q} + \mathbf{A}^T \mathbf{v}_1 = 0 \\ \mathbf{A}\mathbf{x}_* - \mathbf{b} = 0 \end{cases}$$

The KKT can be written as a system of equations

$$\begin{bmatrix} \mathbf{P} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x}_* \\ \mathbf{v}_1 \end{bmatrix} = \begin{bmatrix} -\mathbf{q} \\ \mathbf{b} \end{bmatrix}$$

Which can be easily computed using the MATLAB function `linsolve`.

```
% Solving the linear equations system
factor1 = [P,A';A,zeros(p)];
product = [-q;b];
factor2 = linsolve(factor1,product);
primal_optimal = factor2(1:n);
dual_optimal = factor2(n+1:n+p);
opt_val = 0.5*primal_optimal'*P*primal_optimal+q'*primal_optimal;
```

`CVX`'s optimal value is identical to the solution of the system.

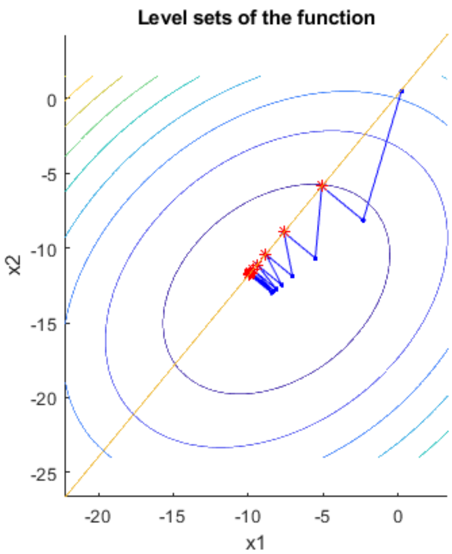
iii. Compute the optimal solution via the projected gradient method

$$\mathbf{x}_{k+1} = \mathbf{P}_{\mathbb{S}} \left(\mathbf{x}_k - \frac{1}{L} \nabla f_0 \left(\mathbf{x}_k \right) \right)$$

where $L := \max\left(\text{eig}\left(\nabla^2 f_0(\mathbf{x})\right)\right)$

The optimal solution is computed via the projected gradient method that was used on exercise 4. This solution is identical to the one found from `CVX` and the one that was calculated using the KKT conditions. It should be noted that for more dimensions (n, p) the projected gradient method required much fewer iterations than `CVX`. Then, the gradient descent points and the projections are plotted with the contour of f for $n = 2$ and $p = 1$.

On the diagram below, the orange line is $\mathbf{A}\mathbf{x} = \mathbf{b}$, the blue points are the gradient descent points and the red stars are the projection of those points onto $\mathbb{S}$. It can be clearly visualized that the algorithm is trying to find the point in $\mathbb{S}$ for which f_0 is minimized.


Level sets of the function
x2
x1
0
-5
-10
-15
-20
-25
-20
-15
-10
-5
0