ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
TECHNICAL UNIVERSITY OF CRETE

# Optimization

## Taylor approximations, Convex sets, Convex functions, Least-Squares

School of Electrical and Computer Engineering

Michailidis Stergios - 2020030080

November 2023

# 1 Question 1

## 1.1 a.

Let $f : \mathbb{R}_+ \to \mathbb{R}$ with $f(x) = \frac{1}{1+x}$. We will explore the first 2 Taylor approximations around $x_0 \in \mathbb{R}_+$:

$$f_{(1)}(x) = f(x_0) + f'(x_0)(x - x_0)$$

$$f_{(2)}(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2$$

Computing the derivatives:

$$f'(x) = \frac{d}{dx} \frac{1}{1+x} = -\frac{1}{(1+x)^2}$$

$$f''(x) = \frac{d^2}{dx^2} \frac{1}{1+x} = \frac{1}{(1+x)^3}$$

In the end we have:

$$f_{(1)}(x) = \frac{1}{1+x_0} - \frac{x - x_0}{(1+x_0)^2} = \frac{1 - x + 2x_0}{(1+x_0)^2}$$

and

$$\begin{aligned} f_{(2)}(x) &= \frac{1 - x + 2x_0}{(1+x_0)^2} + \frac{1}{2} \frac{2}{(1+x_0)^3} (x - x_0)^2 \\ &= \frac{(1 - x + 2x_0)(1 + x_0) + (x - x_0)^2}{(1+x_0)^3} \\ &= \frac{1 - x + 2x_0 + x_0 - xx_0 + 2x_0^2 + (x^2 + 2xx_0 + x_0^2)}{(1+x_0)^3} \\ &= \frac{x^2 - (3x_0 + 1)x + 3x_0^2 + 3x_0 + 1}{(1+x_0)^3} \end{aligned}$$

### 1.2 b.

Plotting the functions in MATLAB for different values of $x_0$:

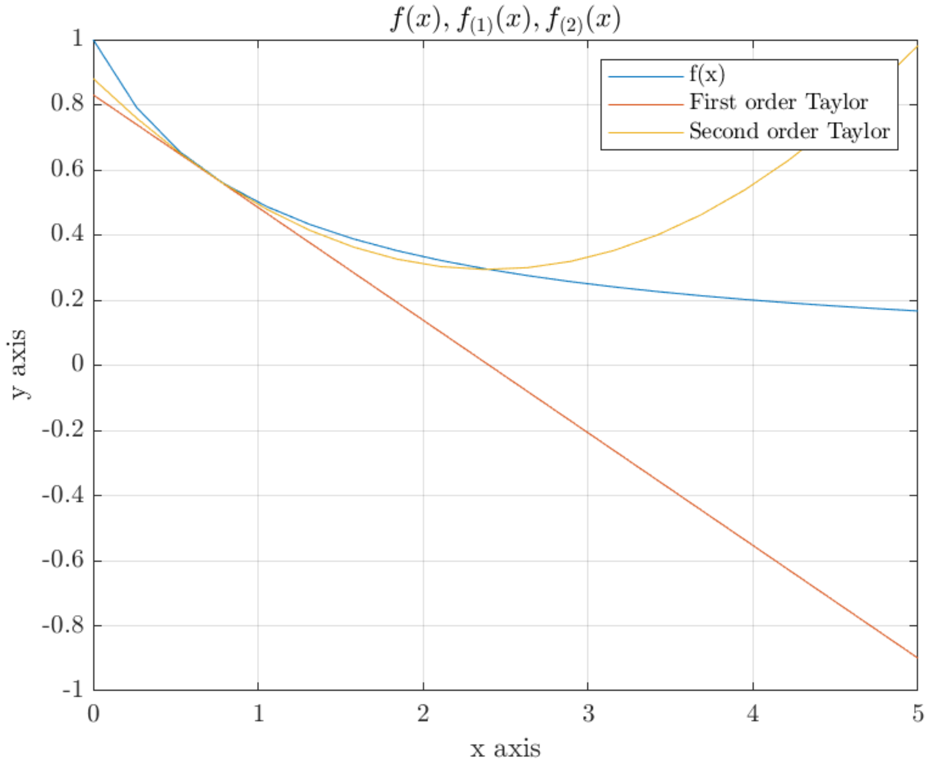


Figure 1: $x_0 = 0.7$

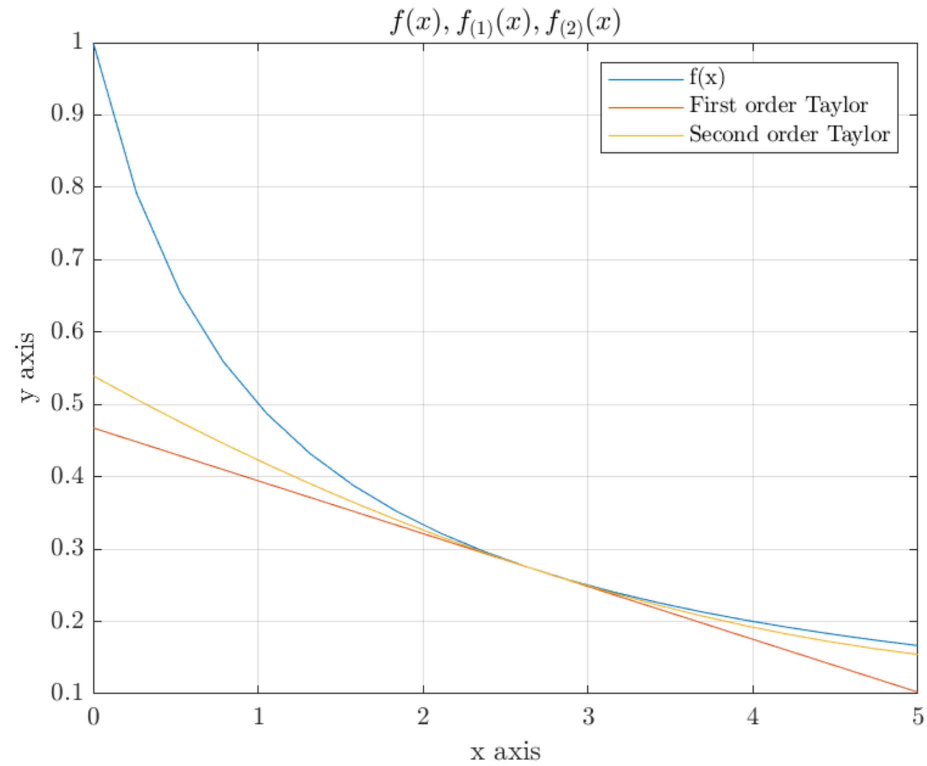


Figure 2: $x_0 = 2.7$

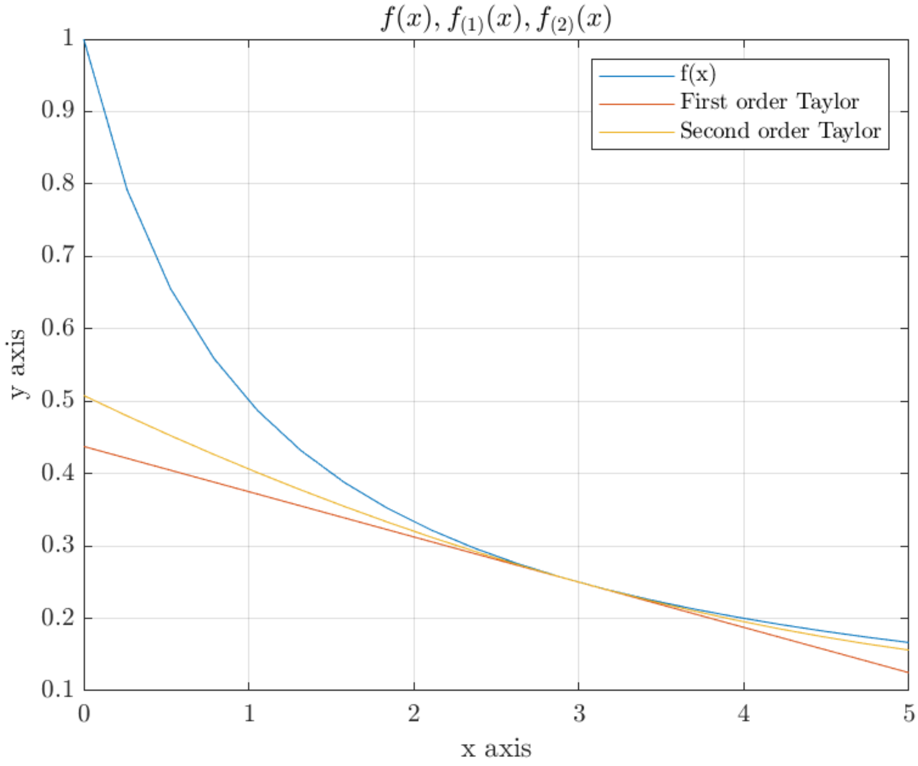


Figure 3: $x_0 = 3$

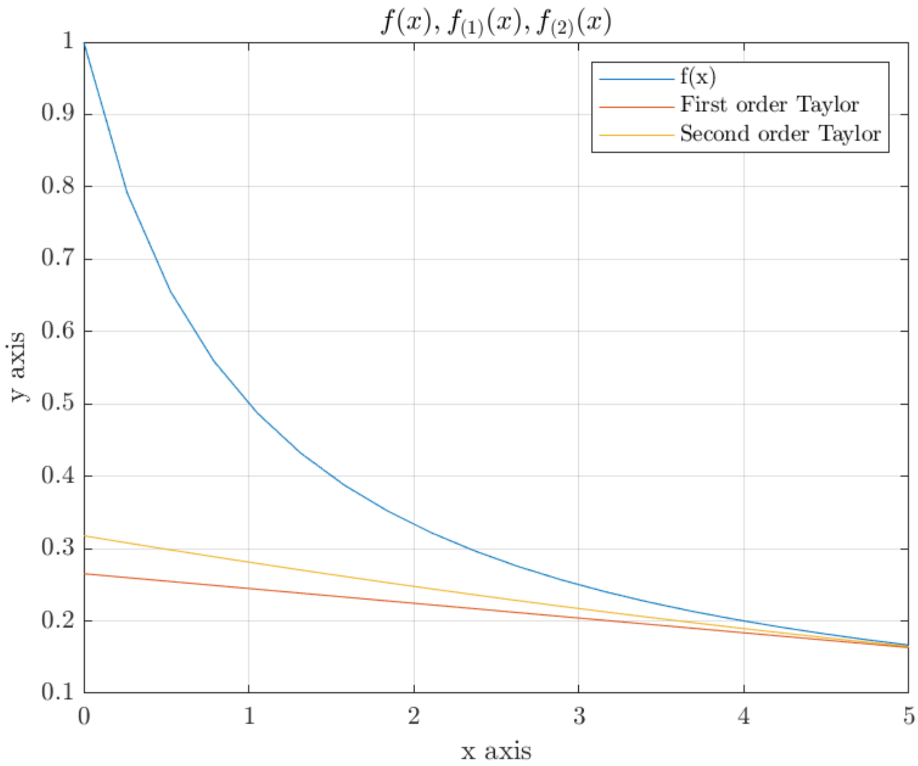


Figure 4: $x_0 = 6$

## 2 Question 2

### 2.1 a.

Let $f : \mathbb{R}_+^2 \to \mathbb{R}$ with $f(x_1, x_2) = \frac{1}{1+x_1+x_2}$.Similarly to Question 1 We will explore the first 2 Taylor approximations around $\vec{x_0} \in \mathbb{R}_+^2$:
Pick $x_* = 5 > 0$

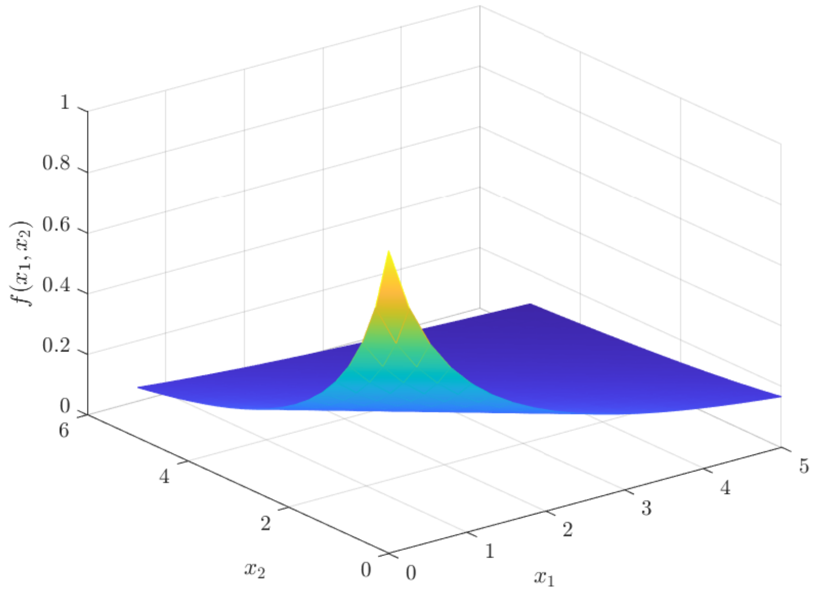


Figure 5: Plot of $f$

### 2.2 b.

The levels sets of $f$ are shown in the countour plot:

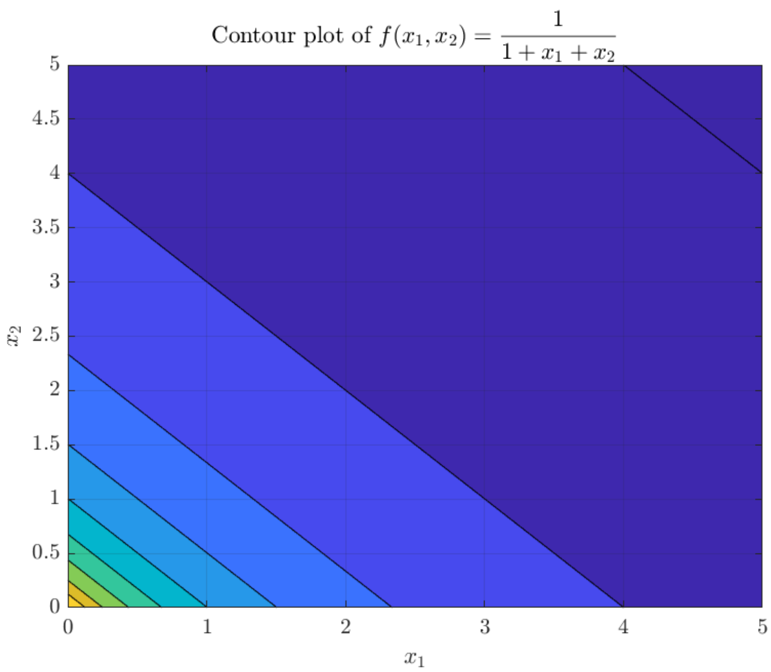


Figure 6: Contour plot of $f$

From the contour plot we can make the remark that as $x_1, x_2 \to 0$ the value of the function increases, as shown by the colour map.

### 2.3 c.

We have chosen point $\vec{x_0} = (x_{0,1}, x_{0,2})^T = (3,3)^T$

$$f_{(1)}(\vec{x}) = f(\vec{x_0}) + \nabla f(x_0)^T(\vec{x} - \vec{x_0})$$

$$f_{(2)}(\vec{x}) = f_{(1)}(\vec{x_0}) + \frac{1}{2}(\vec{x} - \vec{x_0})^T \mathbf{H_f}(\vec{x} - \vec{x_0})$$

Where the $\nabla f$ is defined as:

$$\nabla f(x_1, x_2) = \begin{bmatrix} \frac{\partial}{\partial x_1} f(x_1, x_2) \\ \frac{\partial}{\partial x_2} f(x_1, x_2) \end{bmatrix} = \begin{bmatrix} -\frac{1}{1+x_1+x_2} \\ -\frac{1}{1+x_1+x_2} \end{bmatrix}$$

and $\mathbf{H_f}$ is the hessian matrix:

$$\begin{aligned} \mathbf{H_f} &= \begin{bmatrix} \frac{\partial^2}{\partial x_1^2} f(x_1, x_2) & \frac{\partial^2}{\partial x_1 x_2} f(x_1, x_2) \\ \frac{\partial^2}{\partial x_2 x_1} f(x_1, x_2) & \frac{\partial^2}{\partial x_2^2} f(x_1, x_2) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial}{\partial x_1} \frac{-1}{(1+x_1+x_2)^2} & \frac{\partial}{\partial x_2} \frac{-1}{(1+x_1+x_2)^2} \\ \frac{\partial}{\partial x_1} \frac{-1}{(1+x_1+x_2)^2} & \frac{\partial}{\partial x_2} \frac{-1}{(1+x_1+x_2)^2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{2}{(1+x_1+x_2)^3} & \frac{2}{(1+x_1+x_2)^3} \\ \frac{2}{(1+x_1+x_2)^3} & \frac{2}{(1+x_1+x_2)^3} \end{bmatrix} \end{aligned}$$

using the above equations, we compute the taylor approximations of $f$:

$$f_{(1)}(\vec{x}) = f(\vec{x_0}) + \nabla f(x_0)^T(\vec{x} - \vec{x_0}) = \frac{-(x_1 + x_2) + 2(x_{0,1} + x_{0,2}) + 1}{(1 + x_{0,1} + x_{0,2})^2}$$

$$\begin{aligned} f_{(2)}(\vec{x}) &= \frac{-(x_1 + x_2) + 2(x_{0,1} + x_{0,2}) + 1}{(1 + x_{0,1} + x_{0,2})^2} + \frac{(x_1 + x_{0,1})^2 + (x_2 - x_{0,2})^2 + 2(x_1 + x_{0,1})(x_2 + x_{0,2})}{(1 + x_1 + x_2)^3} \\ &= \frac{-(x_1 + x_2) + 2(x_{0,1} + x_{0,2}) + 1}{(1 + x_{0,1} + x_{0,2})^2} + \frac{((x_1 - x_{0,1} + x_2 + x_{0,2})^2}{(1 + x_1 + x_2)^3} \end{aligned}$$

### 2.4 c. d.

Lets now plot all the functions together:

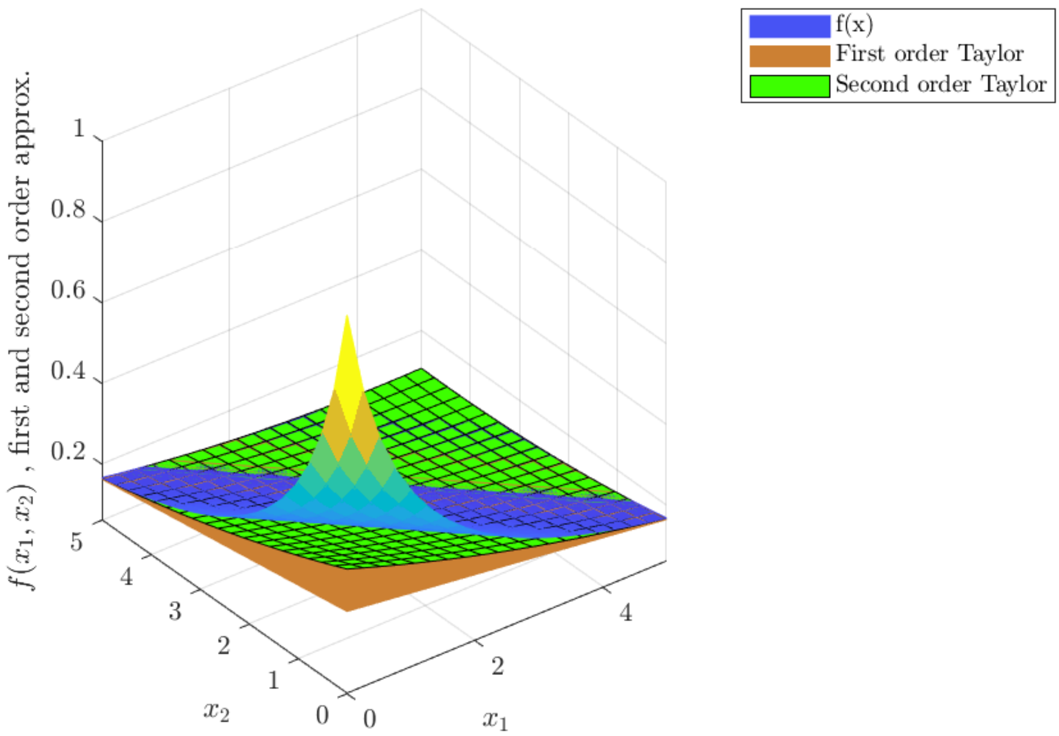


Figure 7: Plot of $f$,$f_{(1)}$,$f_{(2)}$

# 3 Question 3

## 3.1 a.

Lets use the direct method of proof. Let $x_1, x_2 \in \mathbb{S}_{\vec{a},b} = \{\mathbf{x} \in \mathbb{R}^n \,|\, \vec{a}^T\mathbf{x} \leq b\}$ and $0 \leq \theta \leq 1$

Multiplying each side by $\theta$ and $\theta - 1$ respectively:

$$\vec{a}^T x_1 \leq b \iff \theta(\vec{a}^T x_1) \leq \theta b \tag{1}$$

$$\vec{a}^T x_2 \leq b \iff (1-\theta)(\vec{a}^T x_2) \leq (1-\theta) b \tag{2}$$

Performing addition by parts on the above inequalities (1) (2), we end up having:

$$\begin{aligned} &\theta(\vec{a}^T x_1) + (1-\theta)(\vec{a}^T x_2) \leq \theta b + (1-\theta)b \\ \iff &\vec{a}^T(\theta x_1 + (1-\theta)x_2) \leq b \\ \iff &\vec{a}^T x_s \leq b \end{aligned}$$

where $x_s$ is any convex combination of $x_1, x_2 \in \mathbb{S}_{\vec{a},b}$. In conclusion, $\mathbb{S}_{\vec{a},b}$ is convex.

## 3.2 b.

We shall prove this using a counterexample. Let $\mathbb{S}_{\vec{a},b}$ b affine. This means $\forall x_1, x_2 \in \mathbb{S}_{\vec{a},b}$ and for random $\theta \in \mathbb{R}$ the point $\theta x_1 + (1-\theta)x_2 \in \mathbb{S}_{\vec{a},b}$.

Choose random $\theta > 1$ and $x_1$ such that $\vec{a}^T x_1 = b$. Following the same procedure as in question a., we will end up with:

$$\vec{a}^T(\theta x_1 + (1-\theta)x_2) \geq b$$

Concluding that $\mathbb{S}_{\vec{a},b} = \{\mathbf{x} \in \mathbb{R}^n \,|\, \vec{a}^T\mathbf{x} \leq b\}$ is not an affine set.

# 4 Question 4

$\mathbb{H}_{\mathbf{a},b} := \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{a}^T\mathbf{x} = b\}$.
We have that $a^T x_* = b$ since $x_* \in \mathbb{H}_{\mathbf{a},b}$ . Also $x_* = \lambda a$, since a and $x_*$ are co-linear.
For non-zero $a$ we have: $\lambda a^T a = b \iff \lambda = \frac{b}{||a||_2^2}$
In conclusion, $x_* = \frac{ba}{||a||_2^2}$

# 5 Question 5

## 5.1 a.

$$f''(x) = \frac{d^2}{dx^2}\frac{1}{1+x} = \frac{1}{(1+x)^3} \geq 0$$

$\forall x \in \mathbb{R}_+$. Thus $f$ is convex.

## 5.2 b.

Compute the Hessian matrix $\mathbf{H_f}$:

$$\begin{aligned}\mathbf{H_f} &= \begin{bmatrix} \frac{\partial^2}{\partial x_1^2} f(x_1, x_2) & \frac{\partial^2}{\partial x_1 x_2} f(x_1, x_2) \\ \frac{\partial^2}{\partial x_2 x_1} f(x_1, x_2) & \frac{\partial^2}{\partial x_2^2} f(x_1, x_2) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial}{\partial x_1} \frac{-1}{(1+x_1+x_2)^2} & \frac{\partial}{\partial x_2} \frac{-1}{(1+x_1+x_2)^2} \\ \frac{\partial}{\partial x_1} \frac{-1}{(1+x_1+x_2)^2} & \frac{\partial}{\partial x_2} \frac{-1}{(1+x_1+x_2)^2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{2}{(1+x_1+x_2)^3} & \frac{2}{(1+x_1+x_2)^3} \\ \frac{2}{(1+x_1+x_2)^3} & \frac{2}{(1+x_1+x_2)^3} \end{bmatrix}\end{aligned}$$

We can compute the 2 minor determinants and clearly see that they are positive and 0 respectively:
$det_1 = \frac{2}{(1+x_1+x_2)^3}$
$det_2 = 0$ This is equivalent to the Hessian matrix being p.d.

$$\mathbf{H_f} \succeq 0 \tag{3}$$

hence $f$ : convex.

### 5.3 c.

Computing the derivative we have:

$f''(x) = a(a-1)x^{a-2}$

For every $a \geq 1 \vee a \leq 0$ we have that $f''(x) \geq 0$ so $f$ is convex.
For every $0 \leq a \leq 1$ we have that $f''(x) \leq 0$ so $f$ is concave.

We can plot the different curves to verify our findings:

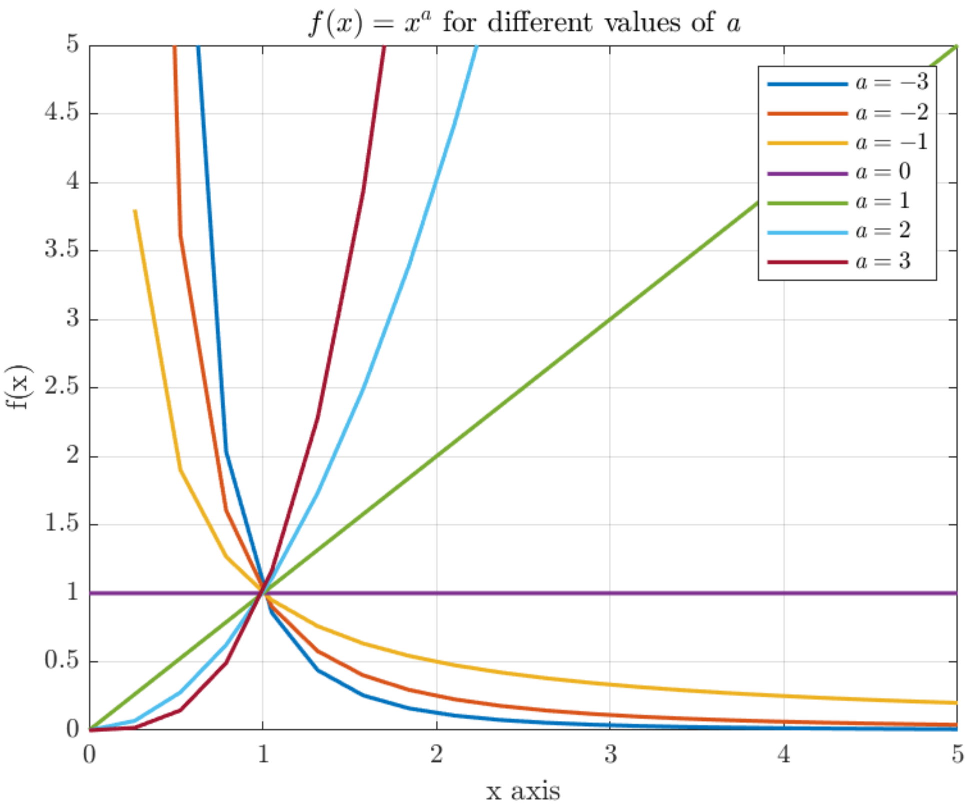


Figure 8: Plot of $f$,for different values of $a$

### 5.4 d.

To prove the convexity of the norm function, we cannot use the normal method of applying the gradient operator, since $f$ is not diff/able everywhere in its domain. Instead we are going to use the **Definition**. But first let us observe the plots in $\mathbb{R}^3$:

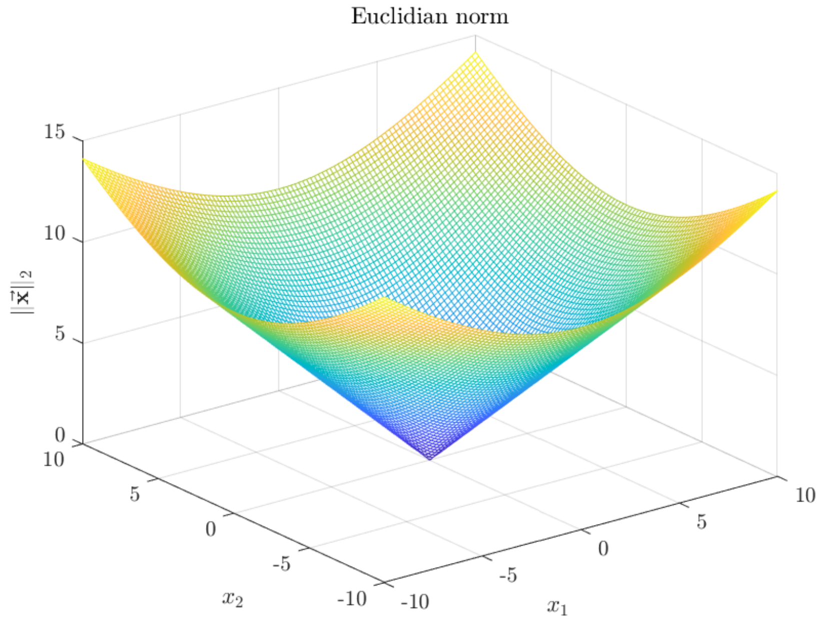


Figure 9: Plot of euclidean norm

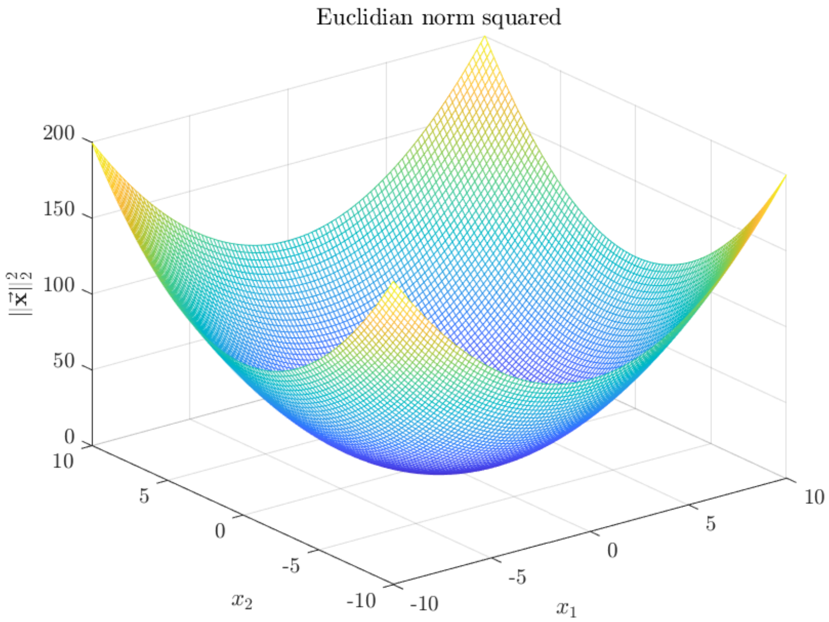


Figure 10: Plot of euclidean norm squared

It is apparent that they are both convex. Lets proove it:

Norm: we shall prove that $\forall x, y \in \mathbb{D}_f$ the following is true:

$$
\begin{aligned}
& f(\theta x + (1-\theta)y) \leq \theta f(X) + (1-\theta)f(y) \\
\iff & ||\theta x + (1-\theta)y||_2 \leq \theta||x||_2 + (1-\theta)||y||_2 \\
\iff & ||\theta x + (1-\theta)y||_2 \leq ||\theta x||_2 + ||(1-\theta)y||_2
\end{aligned}
$$

which is indeed true $\forall x, y \in \mathbb{D}_f$from the triangle inequality. Thus the equivalent starting proposition is also true, so the norm function is convex.
The proof for the norm squared is much simpler:

$\nabla^2 = 2\mathbb{I} \succ 0$

so in conclusion the norm function squared is convex.

# 6 Question 6

Plot in $\mathbb{R}^3$:

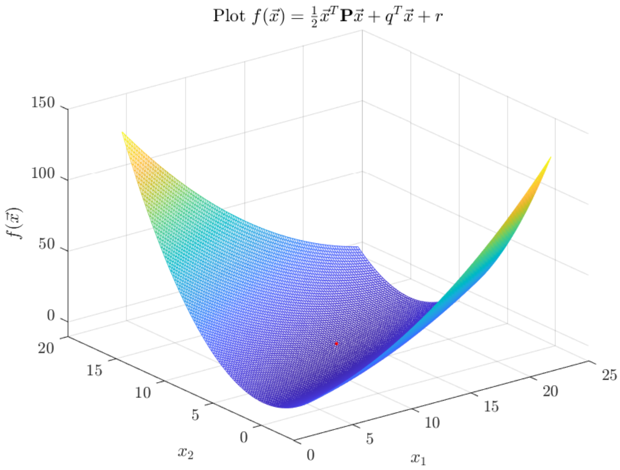


Figure 11: Plot of quadratic function

Contour of quadratic function:

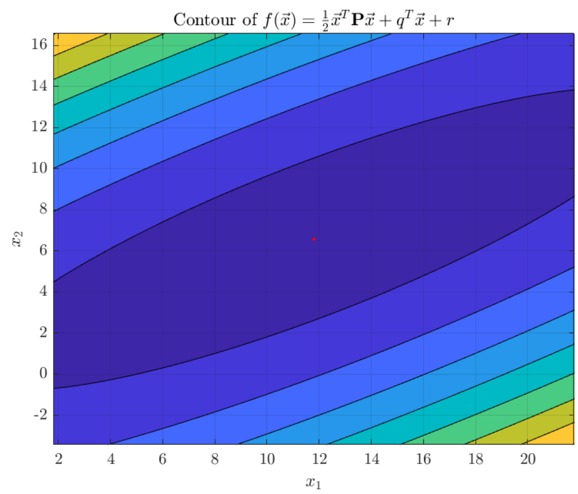


Figure 12: Plot of quadratic function

The marked point (in red) is indeed the minimal.