Technical University of Crete
School of Electrical and Computer Engineering
Course: **Convex Optimization**
Exercise 1 (100/1000)
Report Delivery Date: 2 November 2021
Instructor: Athanasios P. Liavas

Student: Alevrakis Dimitrios 2017030001

---

1. Let $f : \mathbb{R}_+ \to \mathbb{R}$, with $f(x) = \frac{1}{1+x}$. Let $x_0 \in \mathbb{R}_+$, and define the first- and second-order Taylor approximations of f at x_0 as

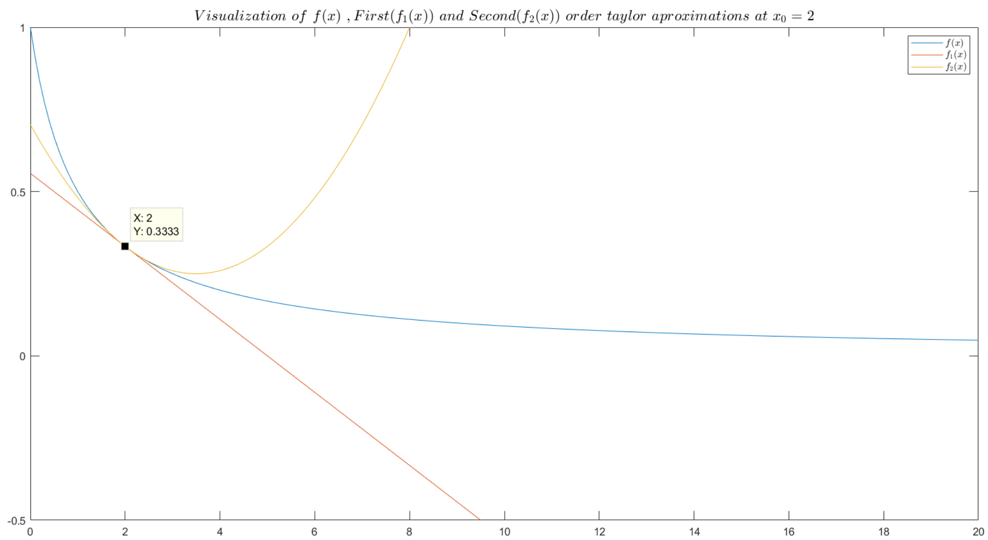
$$\begin{aligned} f_{(1)}(x) &= f(x_0) + f'(x_0)(x - x_0), \\ f_{(2)}(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2. \end{aligned} \tag{1}$$

   (a) Analytic expressions for functions f' and f'';

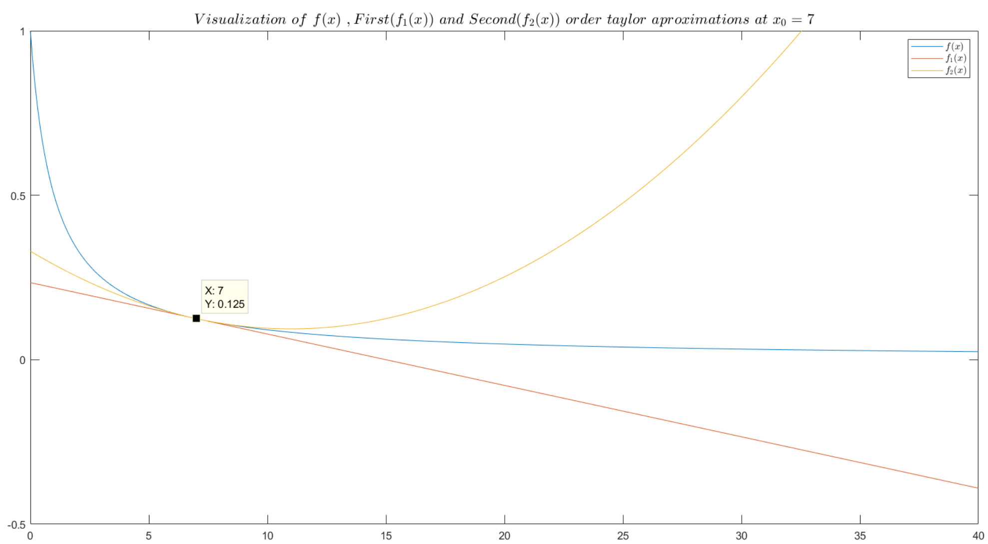
$$\begin{aligned} f'(x) &= -\frac{1}{(1+x)^2}, \\ f''(x) &= \frac{2}{(1+x)^3} \end{aligned} \tag{2}$$

   (b) Plots of $f(x), f_{(1)}(x)$ and $f_{(2)}(x)$ for various x_0.

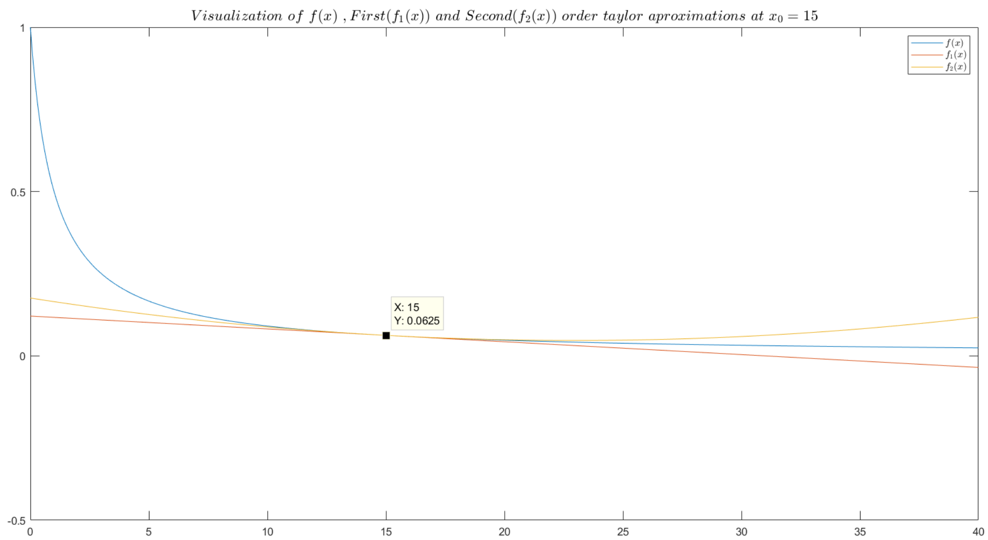
As is confirmed from the above plots, since f is convex, for every x_0,the first order taylor estimation is an underestimation of f and the second derivative is always non-negative for every x.


$Visualization\ of\ f(x)\ , First(f_1(x))\ and\ Second(f_2(x))\ order\ taylor\ aproximations\ at\ x_0 = 2$
$f(x)$
$f_1(x)$
$f_2(x)$
X: 2
Y: 0.3333
1
0.5
0
-0.5
0
2
4
6
8
10
12
14
16
18
20

(a) $x_0 = 2$

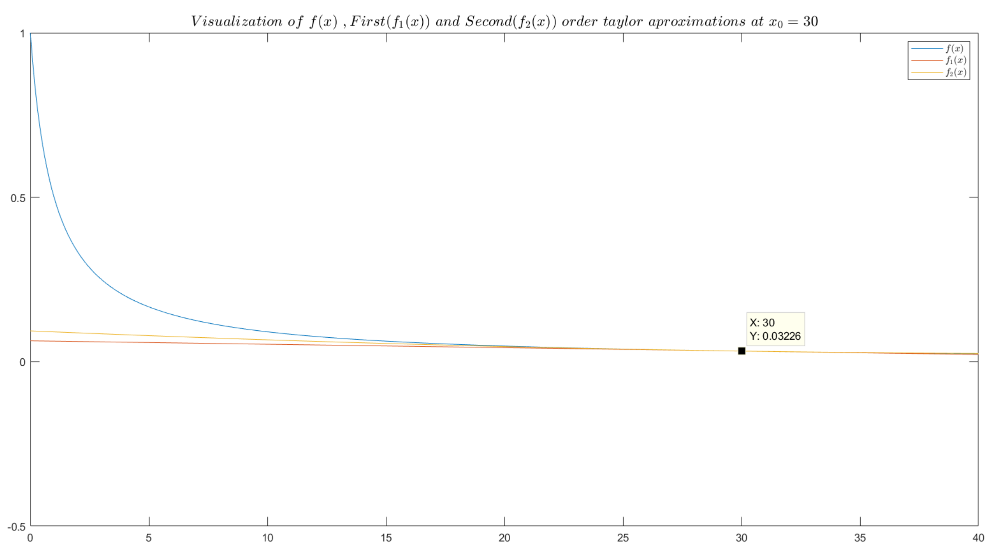

$Visualization\ of\ f(x)\ , First(f_1(x))\ and\ Second(f_2(x))\ order\ taylor\ aproximations\ at\ x_0 = 7$
$f(x)$
$f_1(x)$
$f_2(x)$
X: 7
Y: 0.125
1
0.5
0
-0.5
0
5
10
15
20
25
30
35
40

(b) $x_0 = 7$



$Visualization\ of\ f(x)\ , First(f_1(x))\ and\ Second(f_2(x))\ order\ taylor\ aproximations\ at\ x_0 = 15$
$f(x)$
$f_1(x)$
$f_2(x)$
X: 15
Y: 0.0625

(a) $x_0 = 15$



$Visualization\ of\ f(x)\ , First(f_1(x))\ and\ Second(f_2(x))\ order\ taylor\ aproximations\ at\ x_0 = 30$
$f(x)$
$f_1(x)$
$f_2(x)$
X: 30
Y: 0.03226

(b) $x_0 = 30$

2. Let $f : \mathbb{R}_+^2 \to \mathbb{R}$, with $f(x_1, x_2) = \frac{1}{1+x_1+x_2}$.

   (a) Plotted f using mesh, for $x* = 25$

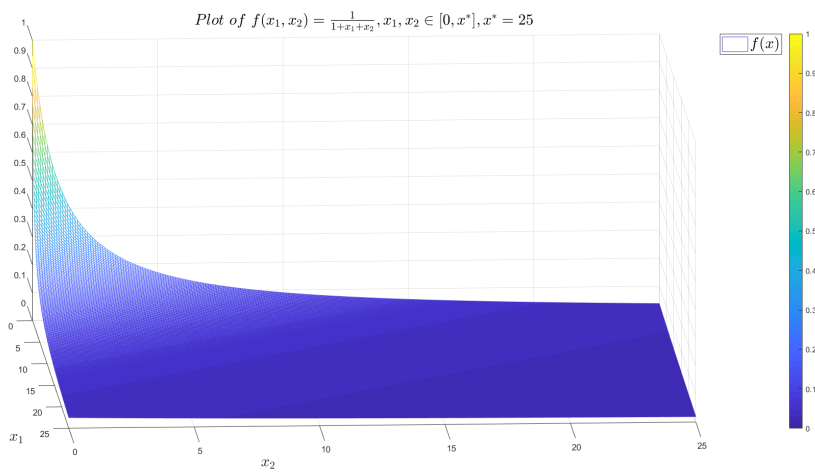


Figure 3: $f(x_1, x_2) = \frac{1}{1+x_1+x_2},\ x, 1x_2 \in [0, x^*], x^* = 25$

   (b) Level Sets of f:

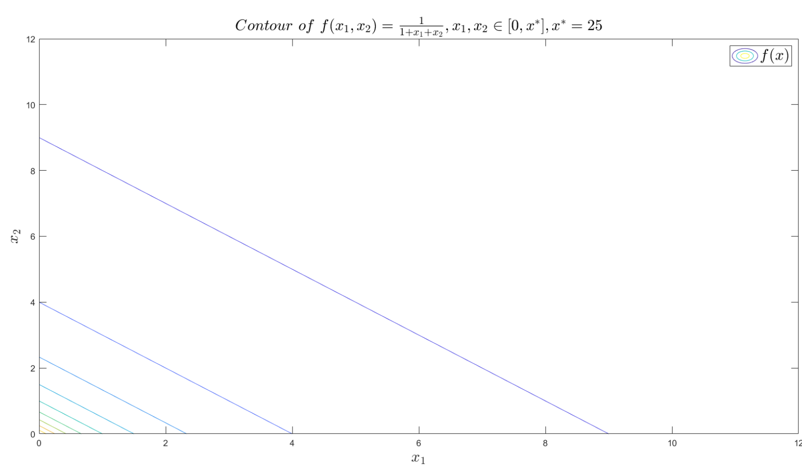


Figure 4: Level Sets of f using contour

We observe that the level sets of f:

$S_c = \{x_1, x_2 \in [0, x^*] : x_1 + x_2 = k | f(x_1, x_2) = c\}$. Since the level sets are described by a linear equation of x_1 and x_2, they are line segments parallel to each other.

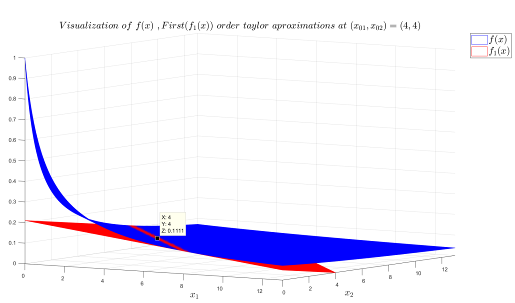
(c) First and Second Order Taylor approximations of f at $x_0 = (x_{01}, x_{02})$:
Computation of the Gradient and the Hessian matrix;

$$\begin{aligned} \nabla f(x) &= \begin{bmatrix} \frac{df}{dx_1} \\ \frac{df}{dx_2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{(1+x_1+x_2)^2} \\ -\frac{1}{(1+x_1+x_2)^2} \end{bmatrix} \\ Hf(x) &= \begin{bmatrix} \frac{d^2}{d^2x_1} & \frac{d^2}{dx_1dx_2} \\ \frac{d^2}{dx_2dx_1} & \frac{d^2}{d^2x_2} \end{bmatrix} = \begin{bmatrix} \frac{2}{(1+x_1+x_2)^3} & \frac{2}{(1+x_1+x_2)^3} \\ \frac{2}{(1+x_1+x_2)^3} & \frac{2}{(1+x_1+x_2)^3} \end{bmatrix} \end{aligned} \tag{3}$$

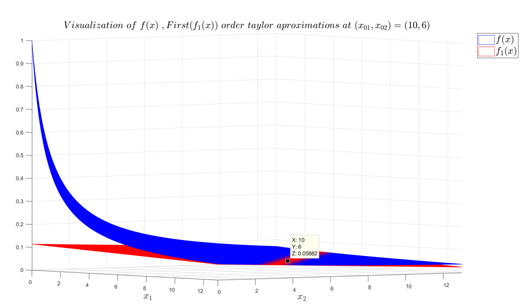
Computation of Taylor Approximations at $x_0 = (x_{01}, x_{02})$:

$$\begin{aligned} f_1(x) &= f(x_0) + \nabla f(x_0)'(x - x_0) = f(x_0) + \begin{bmatrix} -\frac{1}{(1+x_{01}+x_{02})^2} \\ -\frac{1}{(1+x_{01}+x_{02})^2} \end{bmatrix}' \begin{bmatrix} x_1 - x_{01} \\ x_2 - x_{02} \end{bmatrix} \\ &= \frac{1}{1 + x_{01} + x_{02}} - \frac{1}{(1 + x_{01} + x_{02})^2}(x_1 + x_2 - x_{01} - x_{02}) \\ f_2(x) &= f(x_0) + \nabla f(x_0)'(x - x_0) + \frac{1}{2}(x - x_0)'Hf(x)(x - x_0) \\ &= f_1(x) + \frac{1}{2} \begin{bmatrix} x_1 - x_{01} \\ x_2 - x_{02} \end{bmatrix}' \begin{bmatrix} \frac{2}{(1+x_1+x_2)^3} & \frac{2}{(1+x_1+x_2)^3} \\ \frac{2}{(1+x_1+x_2)^3} & \frac{2}{(1+x_1+x_2)^3} \end{bmatrix} \begin{bmatrix} x_1 - x_{01} \\ x_2 - x_{02} \end{bmatrix} \\ &= \frac{1}{1 + x_{01} + x_{02}} - \frac{(x_1 + x_2 - x_{01} - x_{02})}{(1 + x_{01} + x_{02})^2} + \frac{(x_1 + x_2 - x_{01} - x_{02})^2}{(1 + x_1 + x_2)^3} \end{aligned} \tag{4}$$

(d) Common plot f and its first-order Taylor approximation at various $x_0 = (x_{01}, x_{02})$.
We notice that $f_1(x)$ is an underestimation of f for every $x \in dom f$ thus f is



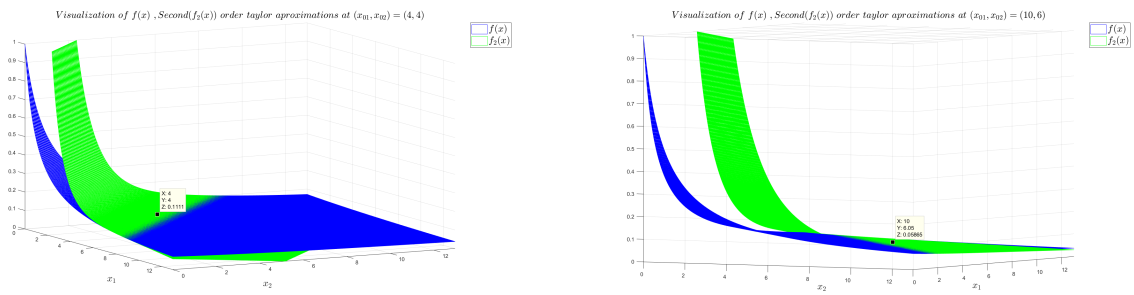
(a) $x_0 = (4, 4)$



(b) $x_0 = (10, 6)$

convex.

(e) Common plot f and its second-order Taylor approximation at various $x_0 = (x_{01}, x_{02})$.



(a) $x_0 = (4, 4)$

(b) $x_0 = (10, 6)$

3. Let $\mathbb{S}_{\mathbf{a},b} = \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{a}^T\mathbf{x} \leq b\}$.

   (a) Proof that $\mathbb{S}_{\mathbf{a},b}$ is convex.

   Let $x_1, x_2 \in S_{\mathbf{a},b}$ then for $0 \leq \theta \leq 1$:

$$
\begin{aligned}
&\left.\begin{aligned}
\mathbf{a}^T\mathbf{x_1} \leq b &\iff \theta\mathbf{a}^T\mathbf{x_1} \leq \theta b \\
\mathbf{a}^T\mathbf{x_2} \leq b &\iff (1-\theta)\mathbf{a}^T\mathbf{x_2} \leq (1-\theta)b
\end{aligned}\right\} \iff \\
&\theta\mathbf{a}^T\mathbf{x_1} + (1-\theta)\mathbf{a}^T\mathbf{x_2} \leq \theta b + (1-\theta)b \iff \\
&\mathbf{a}^T(\theta\mathbf{x_1} + (1-\theta)\mathbf{x_2}) \leq b
\end{aligned}
\tag{5}
$$

   Therefore the convex combination of any $x_1, x_2 \in \mathbb{S}_{\mathbf{a},b}$ belongs in $\mathbb{S}_{\mathbf{a},b}$. Therefore $\mathbb{S}_{\mathbf{a},b}$ is convex.

   (b) In order to prove that $\mathbb{S}_{\mathbf{a},b}$ is not affine we need to show that there exists an affine combination of $\mathbf{x} \in \mathbb{S}_{\mathbf{a},b}$ which is not in $\mathbb{S}_{\mathbf{a},b}$.

   Let $\mathbf{a}^T = [2, 3]$ and $b = 20$.

   We choose:

$$
\begin{aligned}
\mathbf{x_1}^T = [2, 4] &\;:\; \mathbf{a}^T\mathbf{x_1} = 16 < b \\
\mathbf{x_2}^T = [6, 2] &\;:\; \mathbf{a}^T\mathbf{x_1} = 18 < b \\
\mathbf{x_3}^T = [1, 3] &\;:\; \mathbf{a}^T\mathbf{x_1} = 11 < b
\end{aligned}
\tag{6}
$$

   And let $\theta_1 = 0.2$, $\theta_2 = 1.5$ $\theta_2 = -0.7$, for which is true that $\theta_1 + \theta_2 + \theta_3 = 1$

   So:$\theta_1\mathbf{a}^T\mathbf{x_1} + \theta_2\mathbf{a}^T\mathbf{x_2} + \theta_3\mathbf{a}^T\mathbf{x_3} = 0.2 * 16 + 1.5 * 18 - 0.7 * 11 = 22.5 > b$

   Therefore $\mathbb{S}_{\mathbf{a},b}$ is not affine.

4. Let $\mathbf{z} \in \mathbb{R}^\mathbf{n}$ co-linear to $\mathbf{a}$. Then $\mathbf{z} = n\mathbf{a}, n \in R*$.

   In order for $\mathbf{z}$ to lie in $\mathbb{H}_\mathbf{a}$ it must satisfy the equality:

   $\mathbf{a}^T\mathbf{z} = b \iff n\mathbf{a}^T\mathbf{a} = b \iff n = \frac{b}{||\mathbf{a}||_2^2}$.

   Therefore $\mathbf{z} = \frac{b}{||\mathbf{a}||_2^2}\mathbf{a}$

5. Check whether the following functions are convex or not.

   (a) $f : \mathbb{R}_+ \to \mathbb{R}$, with $f(x) = \frac{1}{1+x}$;

$$
\begin{aligned}
f'(x) &= -\frac{1}{(1+x)^2}, \\
f''(x) &= \frac{2}{(1+x)^3}
\end{aligned}
\tag{7}
$$

Since $f''(x) \succcurlyeq 0 \forall x \in \mathbb{R}_+$, f convex.

(b) $f : \mathbb{R}^2_+ \to \mathbb{R}$, with $f(x_1, x_2) = \frac{1}{1+x_1+x_2}$;

$$\begin{aligned}
\nabla f(x) &= \begin{bmatrix} \frac{df}{dx_1} \\ \frac{df}{dx_2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{(1+x_1+x_2)^2} \\ -\frac{1}{(1+x_1+x_2)^2} \end{bmatrix} \\
Hf(x) &= \begin{bmatrix} \frac{d^2}{d^2x_1} & \frac{d^2}{dx_1dx_2} \\ \frac{d^2}{dx_2dx_1} & \frac{d^2}{d^2x_2} \end{bmatrix} = \begin{bmatrix} \frac{2}{(1+x_1+x_2)^3} & \frac{2}{(1+x_1+x_2)^3} \\ \frac{2}{(1+x_1+x_2)^3} & \frac{2}{(1+x_1+x_2)^3} \end{bmatrix}
\end{aligned} \tag{8}$$

In order for f to be convex, $Hf(x)$ must be positive definite $\Rightarrow$ $\mathbf{z}^T\mathbf{H}f(x)\mathbf{z} \geq 0 \forall z \in \mathbb{R}^2 - \{\mathbf{0}\}$.
Let $\mathbf{z} = (a, b)^T \in \mathbb{R}^2$.

$\mathbf{z}^T\mathbf{H}f(x)\mathbf{z} =$

$$\begin{aligned}
& a^2\frac{2}{(1+x_1+x_2)^3} + ab\frac{2}{(1+x_1+x_2)^3} + ab\frac{2}{(1+x_1+x_2)^3} + b^2\frac{2}{(1+x_1+x_2)^3} \\
&= 2\frac{a^2+2ab+b^2}{(1+x_1+x_2)^3} = 2\frac{(a+b)^2}{(1+x_1+x_2)^3} \geq 0 \forall \mathbf{z} \in \mathbb{R}^2,\ \mathbf{x} \in \mathbb{R}_{++}
\end{aligned} \tag{9}$$

Therefore $Hf(x) \succcurlyeq 0 \forall x \in \mathbb{R}_{++}$ and f is convex.

(c) $f : \mathbb{R}_{++} \to \mathbb{R}$, with $f(x) = x^a$.

$$\begin{aligned}
f'(x) &= ax^{a-1}, \\
f''(x) &= a(a-1)x^{a-2}
\end{aligned} \tag{10}$$

i. if $a \leq 0 \iff a - 1 \leq -1 \leq 0 \implies a(a-1) \geq 0$
Since $f''(x) \succcurlyeq 0 \forall x \in \mathbb{R}_{++}$, f is convex.

ii. if $a \geq 1 \iff a - 1 \geq 0 \implies a(a-1) \geq 0$
Since $f''(x) \succcurlyeq 0 \forall x \in \mathbb{R}_{++}$, f is convex.

iii. if $0 \leq a \leq 1 \iff a - 1 \leq 0 \implies a(a-1) \leq 0$
Since $f''(x) \preccurlyeq 0 \forall x \in \mathbb{R}_{++}$, f is not convex.

(d) $f : \mathbb{R}^2 \to \mathbb{R}$, with $f(\mathbf{x}) = \|\mathbf{x}\|_2$.
Since $dom f = \mathbb{R}^2$ is convex, then for any $\mathbf{x_1}, \mathbf{x_2} \in \mathbb{R}^2$, f is convex as long as $f(\theta\mathbf{x_1} + (1-\theta)\mathbf{x_2}) \leq \theta f(\mathbf{x_1}) + (1-\theta)f(\mathbf{x_2})$.
So knowing that $||a+b||_2 \leq ||a||_2 + ||b||_2$:

$$
\begin{aligned}
f(\theta\mathbf{x_1} + (1-\theta)\mathbf{x_2}) &= ||\theta\mathbf{x_1} + (1-\theta)\mathbf{x_2}||_2 \\
&\leq ||\theta\mathbf{x_1}||_2 + ||(1-\theta)\mathbf{x_2}||_2 \\
&= \theta||\mathbf{x_1}||_2 + (1-\theta)||\mathbf{x_2}||_2 \\
&= \theta f(\mathbf{x_1}) + (1-\theta) f(\mathbf{x_2})
\end{aligned} \tag{11}
$$

Therefore f is convex.

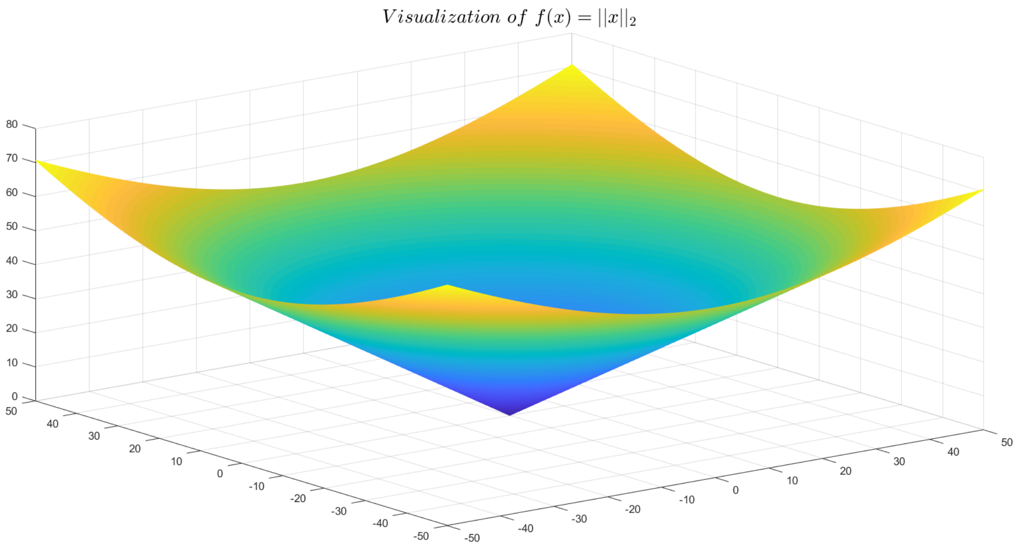


Figure 7: $f(\mathbf{x}) = ||\mathbf{x}||_2$

6. Let $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, and $f : \mathbb{R}^n \to \mathbb{R}$, with $f(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_2^2$.

   (a) Assuming that the columns of $\mathbf{A}$ are linearly independent and we will prove that f is strictly convex.

   f can be also be described as:

$$
\begin{aligned}
f(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{b}\|_2^2 &= (\mathbf{Ax} - \mathbf{b})^T(\mathbf{Ax} - \mathbf{b}) \\
&= (\mathbf{x}^T\mathbf{A}^T - \mathbf{b}^T)(\mathbf{Ax} - \mathbf{b}) \\
&= \mathbf{x}^T\mathbf{A}^T\mathbf{Ax} - \mathbf{x}^T\mathbf{A}^T\mathbf{b} - \mathbf{b}^T\mathbf{Ax} + \mathbf{b}^T\mathbf{b} \\
&\overset{*}{=} \mathbf{x}^T\mathbf{A}^T\mathbf{Ax} - \mathbf{b}^T\mathbf{Ax} - \mathbf{b}^T\mathbf{Ax} + \mathbf{b}^T\mathbf{b} \\
&= \mathbf{x}^T\mathbf{A}^T\mathbf{Ax} - 2\mathbf{b}^T\mathbf{Ax} + \mathbf{b}^T\mathbf{b}
\end{aligned}
\tag{12}
$$

   (*) Since $\mathbf{x}^T\mathbf{A}^T\mathbf{b}$ is a scalar, then we can say that:
   $\mathbf{x}^T\mathbf{A}^T\mathbf{b} = (\mathbf{x}^T\mathbf{A}^T\mathbf{b})^T = (\mathbf{A}^T\mathbf{b})^T\mathbf{x} = \mathbf{b}^T\mathbf{Ax}$

   We can see that f is quadratic, therefore:

$$
\begin{aligned}
\nabla f(\mathbf{x}) &= 2\mathbf{A}^T\mathbf{Ax} - 2\mathbf{b}^T\mathbf{A} \\
\nabla^2(\mathbf{x}) &= 2\mathbf{A}^T\mathbf{A}
\end{aligned}
\tag{13}
$$

   We will prove the first and second order derivatives.

   - For the first order derivative we need to calculate the partial derivatives of the terms $\mathbf{x}^T\mathbf{A}^T\mathbf{Ax}$, $2\mathbf{b}^T\mathbf{Ax}$ and $\mathbf{b}^T\mathbf{b}$.

     In order to prove the term $\mathbf{x}^T\mathbf{A}^T\mathbf{Ax}$, let $\mathbf{P} = \mathbf{A}^T\mathbf{A}$. It is true that $\mathbf{x}^T\mathbf{Px} = \sum_{i=1}^n x_i\Big(\sum_{j=1}^n P_{ji}x_j\Big) =$

Therefore the partial derivatives:

$$
\begin{aligned}
\frac{(d\mathbf{x}^T\mathbf{P}\mathbf{x})}{dx_k} &= \frac{d}{dx_k}\sum_{i=1}^{n} x_i\Big(\sum_{j=1}^{n} P_{ji}x_j\Big) \\
&= \frac{d}{dx_k}\sum_{i=1}^{n}\Big(x_i^2P_{ii} + \sum_{j=1,j\neq i}^{n} x_iP_{ji}x_j\Big) \\
&= \frac{d}{dx_k}\sum_{i=1}^{n}\Big(x_i^2P_{ii}\Big) + \frac{d}{dx_k}\sum_{i=1}^{n}\Big(\sum_{j=1,j\neq i}^{n} x_iP_{ji}x_j\Big) \\
&= 2x_kP_{kk} + \sum_{i=1,i\neq k}^{n} P_{ik}x_i + \sum_{i=1,i\neq k}^{n} x_iP_{ki} \\
&= 2\sum_{i=1}^{n} P_{ki}x_i = 2\mathbf{P_k}^T\mathbf{x},\ \textit{where } P_k \textit{ the } k-th \textit{ column}
\end{aligned}
\tag{14}
$$

After re-constructing the partial derivatives in a column vector we get $2\mathbf{P}\mathbf{x} = 2\mathbf{A}^T\mathbf{A}\mathbf{x}$.

For the calculation of the partial derivatives for the term $2\mathbf{b}^T\mathbf{A}\mathbf{x}$:
Let $\mathbf{q} = 2\mathbf{b}^T\mathbf{A}\mathbf{x} = 2\sum_{i=1}^{n}\sum_{j=1}^{m} x_ib_jAji$

$$
\begin{aligned}
\frac{(d\mathbf{b}^T\mathbf{A}\mathbf{x})}{dx_k} &= \frac{d}{dx_k}\Big(2\sum_{i=1}^{n}\sum_{j=1}^{m} x_ib_jAji\Big) \\
&= \sum_{j=1}^{m} b_jAjk = \mathbf{b}^T\mathbf{A_k},\ \textit{where } A_k \textit{ the } k-th \textit{ row}
\end{aligned}
\tag{15}
$$

After re-constructing the partial derivatives in a column vector we get $\mathbf{b}^T\mathbf{A}$.

The derivative of last term is obviously $\mathbf{0}$ because its independent of $\mathbf{x}$.

Therefore $\nabla f(\mathbf{x}) = 2\mathbf{A}^T\mathbf{A}\mathbf{x} - 2\mathbf{b}^T\mathbf{A}$.

- The second order differentiation of f.

$$
\nabla^2 f(\mathbf{x})_{ij} = \frac{d^2 f}{dx_id_xj} = \frac{d}{x_j}\Big(2\sum_{k=1}^{n} P_{ik}x_k - \sum_{k=1}^{m} b_kA_{ki}\Big) = 2P_{ij} \tag{16}
$$

Therefore $\nabla^2 f(\mathbf{x}) = 2\mathbf{P} = 2\mathbf{A}^T\mathbf{A}$

In order to prove that f is strictly convex, it is suffice to show that the hessian matrix is positive definite.

Firstly since the columns of $\mathbf{A}$ are linearly independent then the equation $\mathbf{Ax} = 0$ has only the trivial solution $\mathbf{x} = \mathbf{0}$.

By definition a matrix $\mathbf{M}$ is positive definite if and only the inequality $\mathbf{z}^T\mathbf{Mz} > 0 \forall z \in \mathbb{R}^n - \{\mathbf{0}\}$.

Replacing with the hessian:

$$\mathbf{z}^T\nabla^2\mathbf{f}(\mathbf{x})\mathbf{z} = 2\mathbf{z}^T\mathbf{A}^T\mathbf{Az} = 2(\mathbf{Az})^T\mathbf{Az} = 2||\mathbf{Az}||^2 \overset{*}{>} 0 \forall \mathbf{z} \in \mathbb{R}^n - \{\mathbf{0}\} \tag{17}$$

(*) The equality cannot be true because of the linear independency of $\mathbf{A}$.

Therefore $\nabla^2\mathbf{f}(\mathbf{x})$ is positive definite and f is strictly convex.

(b) Plots and contours of f for $m = 3$ and $n = 2$. We generate a random $\mathbf{A}$ and $\mathbf{x_{seed}}$ using the function **rand** of Matlab and we calculate $\mathbf{b} = \mathbf{Ax_{seed}}$.
Also by generating a random error vector $\mathbf{e}$ using the function **normrnd** with deviation 5 and median 3 we calculate $\mathbf{b_noise} = \mathbf{Ax_{seed}} - \mathbf{e}$.
Using $\mathbf{A}$, $\mathbf{x_{seed}}$, $\mathbf{b}$ and $\mathbf{A}$, $\mathbf{x_{seed}}$, $\mathbf{b_{noise}}$ we calculate f and f_{noise} respectively.

From the graph and the contours of f observe that it is strictly convex as proved before.
Also adding the noise vector causes f to shift and change shape slightly but still remaining strictly convex. This explains the differences of f and the contour lines.

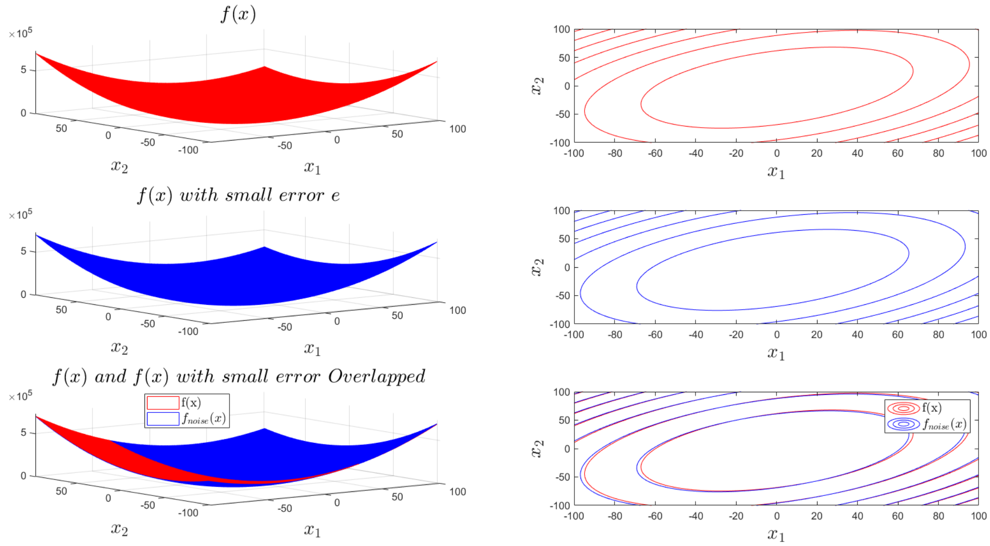


Figure 8: f and f_{noise} for $x_{seed} = (0.2853, -3.3435)^T$

The Matlab code used:

```
%% 1
clear; clc; close all;
%function definitions
f = @(x) 1./(1+x);
f_prime = @(x) -1./(1+x).^2; %%f'
f_prime2 =@(x) 2./(1+x).^3; %%f''


f_1 = @(x,x_0) f(x_0) + f_prime(x_0)*(x-x_0); %%First order Taylor
    aproximation of f at x_0
f_2 = @(x,x_0) f(x_0) + f_prime(x_0)*(x-x_0) + 1/2*f_prime2(x_0)*(x-x_0)
    .^2; %%Second order Taylor aproximation of f at x_0

x = 0:0.1:40; %%x axis
x_0 = 30;
figure;
plot(x,f(x));
hold on;
plot(x,f_1(x,x_0));
plot(x,f_2(x,x_0));
hold off;
```

```
ylim([-0.5,1]);
legend({'$f(x)$','$f_1(x)$','$f_2(x)$'},'Interpreter','latex');
fontSize=14;


%% 2
clear; clc; close all;
f = @(x_1, x_2) 1./(1+x_1+x_2);
x_star = 25;

%a
figure;
[x_1, x_2] = meshgrid(0:0.05:x_star);
mesh(x_1, x_2, f(x_1,x_2));
fontSize = 18;
title('$Plot\ of\ f(x_1,x_2)=\frac{1}{1+x_1+x_2},x_1,x_2\in[0,x^{*}],x^{*}=25$','Interpreter','latex','fontSize',fontSize);
legend({'$f(x)$','$f_1(x)$','$f_2(x)$'},'Interpreter','latex','fontSize',fontSize);
xlabel('$x_1$','Interpreter','latex','fontSize',fontSize);
ylabel('$x_2$','Interpreter','latex','fontSize',fontSize);
colorbar;
caxis([0,1]);

%b
figure;
contour(x_1, x_2, f(x_1,x_2));
xlim([0,12]);
ylim([0,12]);
title('$Contour\ of\ f(x_1,x_2)=\frac{1}{1+x_1+x_2},x_1,x_2\in[0,x^{*}],x^{*}=25$','Interpreter','latex','fontSize',fontSize);
legend({'$f(x)$','$f_1(x)$','$f_2(x)$'},'Interpreter','latex','fontSize',fontSize);
xlabel('$x_1$','Interpreter','latex','fontSize',fontSize);
ylabel('$x_2$','Interpreter','latex','fontSize',fontSize);
colorbar;
caxis([0,1]);
%c
Grad =@(x_1,x_2) [-1./(1+x_1+x_2).^2 ; -1./(1+x_1+x_2).^2];
Hessian =@(x_1,x_2) [ 2./(1+x_1+x_2).^3 , 2./(1+x_1+x_2).^3 ; 2./(1+x_1+
```

```
    x_2).^3 , 2./(1+x_1+x_2).^3];

f_1 = @(x_1,x_2,x_01,x_02)  f(x_01,x_02) + (-1./(1+x_01+x_02).^2)*(x_1-
    x_01) + (-1./(1+x_01+x_02).^2)*(x_2-x_02);
f_2 = @(x_1,x_2,x_01,x_02) f_1(x_1,x_2,x_01,x_02) +1/2*(x_1-x_01+x_2-x_02
    ).^2*2./(1+x_1+x_2).^3;

x_01=10;
x_02=6;

figure;
mesh(x_1, x_2, f(x_1,x_2),'edgecolor','b');
hold on;
mesh(x_1, x_2, f_1(x_1,x_2,x_01,x_02),'edgecolor','r');
hold off;
xlim([0,13]);
ylim([0,13]);
zlim([0,1]);
legend({'$f(x)$','$f_1(x)$'},'Interpreter','latex','fontSize',fontSize);
xlabel('$x_1$','Interpreter','latex','fontSize',fontSize);
ylabel('$x_2$','Interpreter','latex','fontSize',fontSize);
title('$$Visualization\ of\ f(x)\ ,First(f_1(x))\ order\ taylor\
    aproximations\ at\ (x_{01},x_{02})=(10,6)$$','Interpreter','latex','
    fontSize',fontSize);


figure;
mesh(x_1, x_2, f(x_1,x_2),'edgecolor','b');
hold on;
mesh(x_1, x_2, f_2(x_1,x_2,x_01,x_02),'edgecolor','g');
hold off;
xlim([0,13]);
ylim([0,13]);
zlim([0,1]);
legend({'$f(x)$','$f_2(x)$'},'Interpreter','latex','fontSize',fontSize);
xlabel('$x_1$','Interpreter','latex','fontSize',fontSize);
ylabel('$x_2$','Interpreter','latex','fontSize',fontSize);
title('$$Visualization\ of\ f(x)\ ,Second(f_2(x))\ order\ taylor\
    aproximations\ at\ (x_{01},x_{02})=(10,6)$$','Interpreter','latex','
    fontSize',fontSize);
```

```
%% 5d
clear; clc; close all;
f=@(x_1,x_2) sqrt(x_1.^2+x_2.^2);
[x_1, x_2] = meshgrid(-50:0.1:50);
mesh(x_1, x_2, f(x_1, x_2));
fontSize=18;
title('$Visualization\ of\ f(x)=||x||_2$','Interpreter','latex','fontSize',fontSize);
%% 6b
clear; clc; close all;
m=3;
n=2;

A=rand(m,n).*10 - 5;
x=rand(n,1).*10 - 5;
b=A*x;

f = @(x_1,x_2) (A(1,1)*x_1+A(1,2)*x_2-b(1)).^2 + (A(2,1)*x_1+A(2,2)*x_2-b(2)).^2 + (A(3,1)*x_1+A(3,2)*x_2-b(3)).^2;

figure;
subplot(3,2,1);
[x_1, x_2] = meshgrid(-100:1:100);
mesh(x_1, x_2, f(x_1, x_2), 'edgecolor','r' );
xlim([x(1)-100,x(1)+100]);
ylim([x(2)-100,x(2)+100]);
fontSize=18;
title('$f(x)$','Interpreter','latex','fontSize',fontSize);
xlabel('$x_1$','Interpreter','latex','fontSize',fontSize);
ylabel('$x_2$','Interpreter','latex','fontSize',fontSize);

subplot(3,2,2);
contour(x_1,x_2,f(x_1,x_2), 'edgecolor','r');
xlabel('$x_1$','Interpreter','latex','fontSize',fontSize);
ylabel('$x_2$','Interpreter','latex','fontSize',fontSize);

e=normrnd(5,3);
b_noise=A*x+e;
f_noise = @(x_1,x_2) (A(1,1)*x_1+A(1,2)*x_2-b_noise(1)).^2 + (A(2,1)*x_1+
```

```
    A(2,2)*x_2-b_noise(2)).^2 + (A(3,1)*x_1+A(3,2)*x_2-b_noise(3)).^2;

subplot(3,2,3);
mesh(x_1, x_2, f_noise(x_1, x_2), 'edgecolor','b' );
xlim([x(1)-100,x(1)+100]);
ylim([x(2)-100,x(2)+100]);
title('$f(x)\ with\ small\ error\ e$','Interpreter','latex','fontSize',
    fontSize);
xlabel('$x_1$','Interpreter','latex','fontSize',fontSize);
ylabel('$x_2$','Interpreter','latex','fontSize',fontSize);

subplot(3,2,4);
contour(x_1,x_2,f_noise(x_1,x_2), 'edgecolor','b');
xlabel('$x_1$','Interpreter','latex','fontSize',fontSize);
ylabel('$x_2$','Interpreter','latex','fontSize',fontSize);

subplot(3,2,5);
mesh(x_1, x_2, f(x_1, x_2),'edgecolor','r' );
hold on;
mesh(x_1, x_2, f_noise(x_1, x_2),'edgecolor','b' );
xlim([x(1)-100,x(1)+100]);
ylim([x(2)-100,x(2)+100]);
title('$f(x)\ and\ f(x)\ with\ small\ error\ Overlapped$','Interpreter','
    latex','fontSize',fontSize);
xlabel('$x_1$','Interpreter','latex','fontSize',fontSize);
ylabel('$x_2$','Interpreter','latex','fontSize',fontSize);
legend({'f(x)','$f_{noise}(x)$'},'Interpreter','latex','fontSize',
    fontSize);

subplot(3,2,6);
contour(x_1, x_2, f(x_1, x_2),'edgecolor','r' );
hold on;
contour(x_1, x_2, f_noise(x_1, x_2),'edgecolor','b' );
xlabel('$x_1$','Interpreter','latex','fontSize',fontSize);
ylabel('$x_2$','Interpreter','latex','fontSize',fontSize);
legend({'f(x)','$f_{noise}(x)$'},'Interpreter','latex','fontSize',
    fontSize);
```