Convex Optimization

Exercise 1 (100/500)

Report Delivery Date: 20 March 2018

Instructor: Athanasios P. Liavas

Student: Konidaris Vissarion 2011030123

---

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

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

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

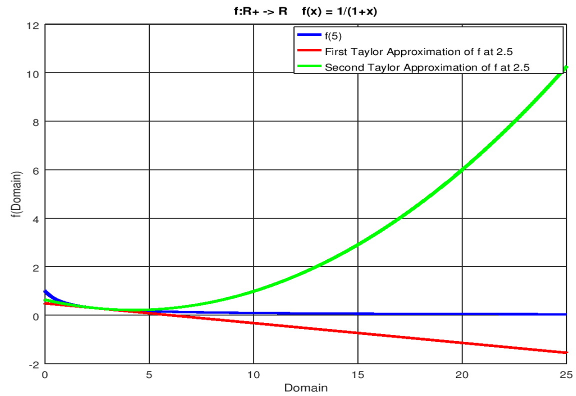
$$f''(x) = \frac{\partial(\frac{-1}{(1+x)^2})}{\partial x} = \frac{\frac{\partial((1+x)^2)}{\partial x}}{(1+x)^4} = \frac{2(1+x)}{(1+x)^4} = \frac{2}{(1+x)^3}$$

$$f_{(2)}(x) = f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2}f''(x_0)(x-x_0)^2 = \frac{1-x+2x_0}{(1+x_0)^2}+\frac{1}{2}\frac{2}{(1+x)^3}(x-x_0)^2 =$$

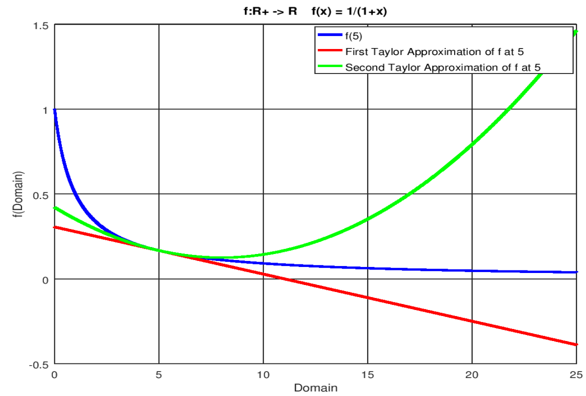
$$= \frac{1-x+2x_0+x_0-x_0x+2x_0^2+x^2-2x_0x+x_0^2}{(1+x_0)^3} =$$

$$= \frac{1+3x_0^2+3x_0-3x_0x-x+x^2}{(1+x_0)^3}$$

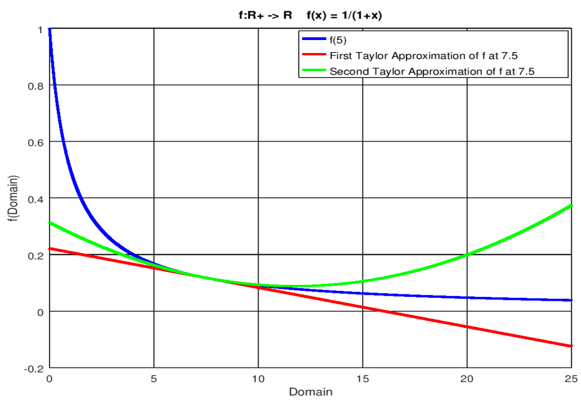
(b) Figure 1 shows some common plots of $f(x)$, $f_{(1)}(x)$ and $f_{(2)}(x)$ for various x_0. It is evident just by looking at the plots that the first order Taylor approximation of the convex function f at x_0 is a global underestimator of f. Moreover close to x_0 the first order Taylor approximation is very similar to f. The second order Taylor approximation of f at the same point is not a global underestimator of f but it approximates f more accurately in the vicinity of x_0.



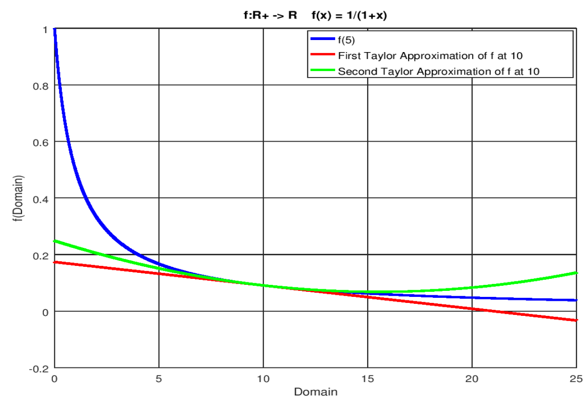
(a) $x_0 = 2.5$



(b) $x_0 = 5$



(c) $x_0 = 7.5$



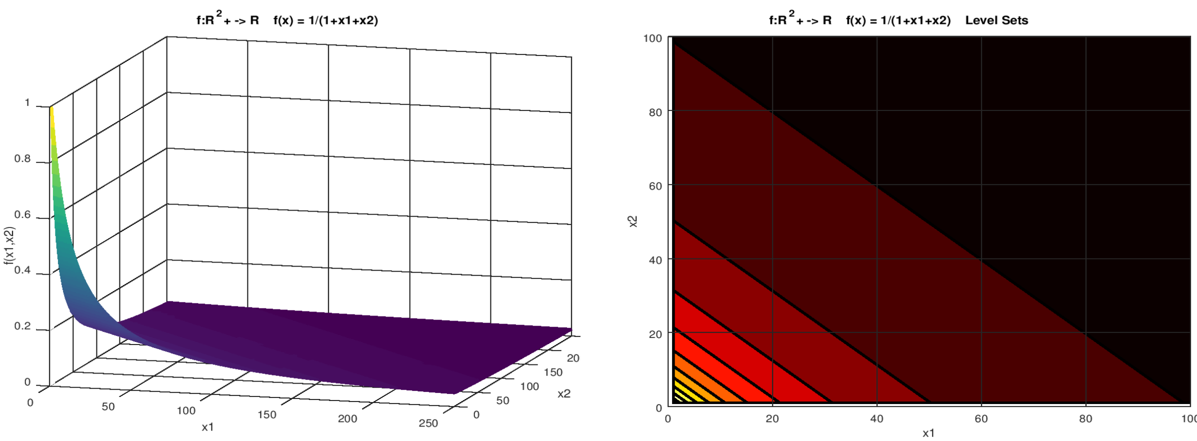
(d) $x_0 = 10$

Figure 1: Plots of f, $f_{(1)}(x_0)$ and $f_{(2)}(x_0)$ at various x_0.

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

   (a) The plot of the convex function f for $x1, x2 \in [0, x_*]$, with $x_* > 0$ is shown in Figure 2 (a).

   (b) The level sets of f are shown in figure 2 (b). The brighter the regions are the higher the value of the function. In each contour the value of the function is the same.



(a) Plot of f.

(b) Level sets of f.

Figure 2: Plot and level sets of f.

(c)

$$\frac{\partial f(x1,x2)}{\partial x1} = \frac{\partial(\frac{1}{1+x1+x2})}{\partial x1} = \frac{-\frac{\partial(1+x1+x2)}{\partial x1}}{(1+x1+x2)^2} = -\frac{1}{1+x1+x2}$$

$$\frac{\partial f(x1,x2)}{\partial x1} == \frac{\partial f(x1,x2)}{\partial x2}$$

$$\nabla f(x1,x2) = \begin{bmatrix} \frac{\partial f(x1,x2)}{\partial x1} \\ \frac{\partial f(x1,x2)}{\partial x2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{1+x1+x2} \\ -\frac{1}{1+x1+x2} \end{bmatrix}$$

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

$$= \frac{1}{1+x_{0,1}+x_{0,2}} + \begin{bmatrix} -\frac{1}{1+x_{0,1}+x_{0,2}} & -\frac{1}{1+x_{0,1}+x_{0,2}} \end{bmatrix} \begin{bmatrix} x1 - x_{0,1} \\ x2 - x_{0,2} \end{bmatrix} =$$

$$\frac{1}{1+x_{0,1}+x_{0,2}} - \frac{x1 - x_{0,1}}{(1+x_{0,1}+x_{0,2})^2} - \frac{x2 - x_{0,2}}{(1+x_{0,1}+x_{0,2})^2} = 1 - \frac{(x_{0,1}+x_{0,2})^2 + x1 + x2}{(1+x_{0,1}+x_{0,2})^2}$$

$$\nabla f^2(x1,x2) = \begin{bmatrix} \frac{\partial^2 f(x1,x2)}{\partial x1^2} & \frac{\partial^2 f(x1,x2)}{\partial x1 x2} \\ \frac{\partial^2 f(x1,x2)}{\partial x2 x1} & \frac{\partial^2 f(x1,x2)}{\partial x2^2} \end{bmatrix} = \begin{bmatrix} \frac{\partial \frac{-1}{(1+x1+x2)^2}}{\partial x1} & \frac{\partial \frac{-1}{(1+x1+x2)^2}}{\partial x2} \\ \frac{\partial \frac{-1}{(1+x1+x2)^2}}{\partial x1} & \frac{\partial \frac{-1}{(1+x1+x2)^2}}{\partial x2} \end{bmatrix} = \begin{bmatrix} \frac{2}{(1+x1+x2)^3} & \frac{2}{(1+x1+x2)^3} \\ \frac{2}{(1+x1+x2)^3} & \frac{2}{(1+x1+x2)^3} \end{bmatrix}$$

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

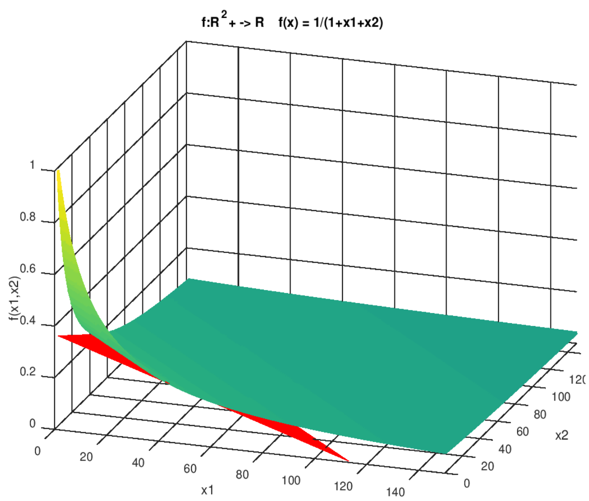
$$= 1 - \frac{(x_{0,1} + x_{0,2})^2 + x1 + x2}{(1 + x_{0,1} + x_{0,2})^2} +$$

$$+\frac{1}{2}\begin{bmatrix} x1 - x_{0,1} & x2 - x_{0,2} \end{bmatrix}\begin{bmatrix} \frac{2}{(1+x_{0,1}+x_{0,2})^3} & \frac{2}{(1+x_{0,1}+x_{0,2})^3} \\ \frac{2}{(1+x_{0,1}+x_{0,2})^3} & \frac{2}{(1+x_{0,1}+x_{0,2})^3} \end{bmatrix}\begin{bmatrix} x1 - x_{0,1} \\ x1 - x_{0,1} \end{bmatrix} =$$

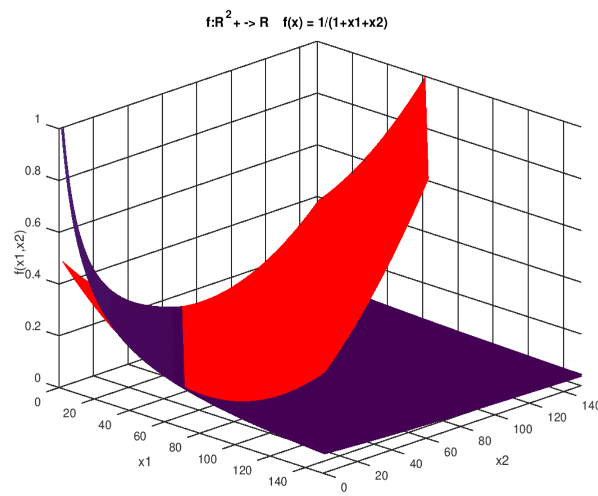
$$= 1 - \frac{(x_{0,1} + x_{0,2})^2 + x1 + x2}{(1 + x_{0,1} + x_{0,2})^2} + \frac{(x1 - x_{0,1} + x2 - x_{0,2})^2}{(1 + x_{0,1} + x_{0,2})^3}$$

(d) The plot of f along $f_{(1)}$ can be seen in Figure 3 (a).

(e) The plot of f along $f_{(2)}$ can be seen in Figure 3 (b).



(a) Plot of f and $f_{(1)}$.



(b) Plot of f and $f_{(2)}$.

Figure 3: Plots of f along with its approximations $f_{(1)}$ and $f_{(2)}$.

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

   (a) Suppose $x1$, $x2 \in S_{a,b}$ and $0 \leq \theta \leq 1$.

$$a^T x_1 \leq b \Leftrightarrow \theta a^T x_1 \leq \theta b$$

$$a^T x_2 \leq b \Leftrightarrow (1-\theta) a^T x_2 \leq (1-\theta) b$$

By adding the two inequalities we get

$$\theta a^T x_1 + (1-\theta) a^T x_2 \leq \theta b + (1-\theta) b$$

$$a^T (\theta x_1 + (1-\theta) x_2) \leq b$$

$$a^T x_* \leq b$$

   Where x_* is any convex combination of any x_1, $x_2 \in S_{a,b}$. Hence $S_{a,b}$ is convex.

   (b) In order for a set to be affine it must contain all the points of a line that connects any two of its points. Suppose x_1, $x_2 \in S_{a,b}$. The line connecting these two points is $\theta x_1 + (1-\theta)x_2$. Now suppose that $\theta > 1$ and x_1 lies on the boundary of the halfspace(i.e $a^T x_1 = b$). Then

$$a^T x_1 = b \Leftrightarrow \theta a^T x_1 = \theta b$$

$$a^T x_2 \leq b \Leftrightarrow (1-\theta) a^T x_2 \geq (1-\theta) b$$

By adding the two inequalities we get

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

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

Hence $S_{a,b}$ is not affine.

4. For a hyperplane $H_{a,b} := \{x \in \mathbb{R}^n | a^T x = b\}$ we want to find a point x_* that is co-linear with a and lies on the hyperplane $H_{a,b}$.

$$x = ca, c \in \mathbb{R}$$

$$a^T x_* = b$$

Combining the above relations we get

$$ca^T a = b$$

$$c||a||_2^2 = b$$

$$c = \frac{b}{||a||_2^2}$$

Hence the point x_* is the

$$x_* = \frac{b}{||a||_2^2} a$$

5. (a) $f : \mathbb{R}_+ \to \mathbb{R}$, $f(x) = \frac{1}{1+x}$. The domain of f is a convex set. Moreover f is differentiable within that domain. The second derivative of f, that was calculated in the first exercise, is $f''(x) = \frac{2}{(1+x)^3}$. $f''(x) \geq 0$ for $x \in \mathbb{R}_+$ hence f is convex.

   (b) $f : \mathbb{R}_+^2 \to \mathbb{R}$, $f(x_1, x_2) = \frac{1}{1+x_1+x_2}$. The domain of f is a convex set. Moreover f is differentiable within that domain. The second derivative of f, that was calculated in the first exercise, is $\begin{bmatrix} \frac{2}{(1+x_1+x_2)^3} & \frac{2}{(1+x_1+x_2)^3} \\ \frac{2}{(1+x_1+x_2)^3} & \frac{2}{(1+x_1+x_2)^3} \end{bmatrix}$. Suppose $h \in \mathbb{R}^2$.

$$h^T \begin{bmatrix} \frac{2}{(1+x_1+x_2)^3} & \frac{2}{(1+x_1+x_2)^3} \\ \frac{2}{(1+x_1+x_2)^3} & \frac{2}{(1+x_1+x_2)^3} \end{bmatrix} h = \begin{bmatrix} h_1 & h_2 \end{bmatrix} \begin{bmatrix} \frac{2(h_1+h_2)}{(1+x_1+x_2)^3} \\ \frac{2(h_1+h_2)}{(1+x_1+x_2)^3} \end{bmatrix} =$$

$$= \frac{2h_1^2 + 4h_1h_2 + 2h_2^2}{(1+x_1+x_2)^3} = \frac{(2h_1 + 2h_2)^2}{(1+x_1+x_2)^3} \geq 0$$

$$\forall x_1, x_2 \in dom f$$

$$\forall h \in \mathbb{R}^2$$

Hence f is convex.

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

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

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

For $a \geq 1$ or $a \leq 0$ and $x \in \mathbb{R}_{++}$, $f''(x) \geq 0$ hence f is convex.
For $0 \leq a \leq 1$ and $x \in \mathbb{R}_{++}$, $f''(x) \leq 0$ hence f is concave.

(d) $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x) = ||x||_2$.

Suppose x_1, $x_2 \in dom f$ and $\theta \in \mathbb{R}$ with $0 \leq \theta \leq 1$.

$$f(x) = ||\theta x_1 + (1-\theta)x_2||_2$$

$$\leq ||\theta x_1||_2 + ||(1-\theta)x_2||_2 = |\theta| ||x_1||_2 + |(1-\theta)| ||x_2||_2 =$$

$$= \theta ||x_1||_2 + (1-\theta)||x_2||_2 = \theta f(x_1) + (1-\theta) f(x_2)$$

So f is convex. We used the triangle inequality and the homogeneity of a norm.

6. Consider the composition of scalar functions $f(x_0) := (hog)(x) = h(g(x))$ with $f, g, h : \mathbb{R} \to \mathbb{R}$, where h and g are smooth functions.

$$f'(x) = h'(g(x))g'(x)$$

$$f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x)$$

(a) If h is convex and non-decreasing, and g is convex, then f is convex. This is because $h''(g(x)) \geq 0$ as h is convex, $g'(x)^2 \geq 0$, $h'(g(x)) \geq 0$ as h is non-decreasing and $g''(x) \geq 0$ as g is convex. An example of h is e^x and of g the $||x||_2$.

(b) If h is convex and non-increasing, and g is concave, then f is convex. This is because $h''(g(x)) \geq 0$ as h is convex, $g'(x)^2 \geq 0$, $h'(g(x)) \leq 0$ as h is non-increasing and $g''(x) \leq 0$ as g is concave. An example of h is $-logx$ and of g the $logx$.

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

(a) Assume that the columns of A are linearly independent.

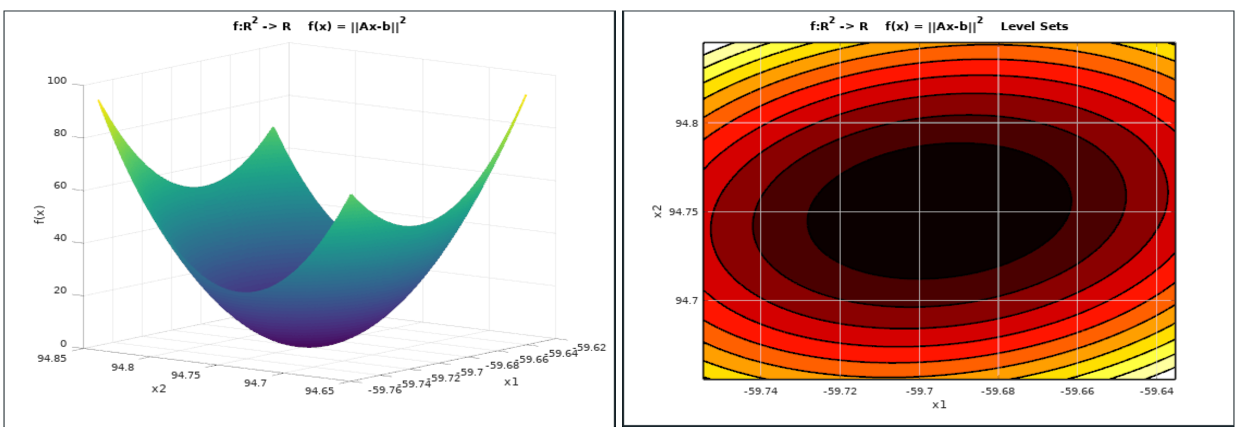
$$f'(x) = \frac{\partial((Ax-b)^T(Ax-b))}{\partial x} = \frac{\partial(x^TA^TAx - x^TA^Tb - b^TAx + b^Tb)}{\partial x} =$$

$$= \frac{1}{2}A^TAx - 2A^Tb$$

$$f''(x) = A^TA$$

A^TA is called the Gram matrix witch is always positive semi-definite. Because all the columns of A are linearly independent(equivalently, it is invertible with $detA \neq 0$) the matrix A^TA is positive definite. Hence f is strictly convex.

(b) The function f takes its minimal value, which is zero, at the solution x. This can also be seen from the plot and level sets of f near its solution.

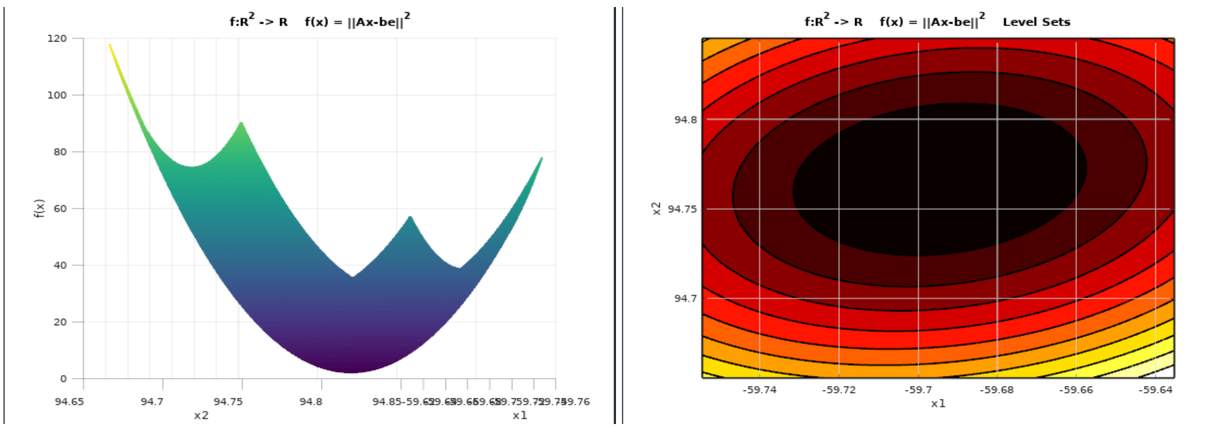


(a) Plot of f in the vicinity of its solution.

(b) Level sets of f.

Figure 4: Plot and level sets of f.

By adding some noise to the vector b, f is no longer minimized at x. The minimum is attained at a different point. Observe that f does not attain the zero value at its minimum.



(a) Plot of f in the vicinity of its solution.

(b) Level sets of f.

Figure 5: Plot and level sets of f.

main.m

```
#Vissarion Konidaris
#Convex Optimization Course
# 9/3/2018

clc;
clear;
clc;

%%%%%%%%%%%%%%%%%%% First Exercise %%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

start = 0.;
finish = 25.;
```

```
precision = 1000000;
domain = linspace(start,finish,precision);
func1 = 1./(1.+domain);

[FirstTaylor,SecondTaylor] = F1_Taylor_Appr(2.5,domain);
% Plot the function along with its first and
% second taylor approximations at point 2.5
figure(1);
pl = plot(domain,func1,'linewidth',2,'b',
          domain,FirstTaylor,'linewidth',2,'r',
          domain,SecondTaylor,'linewidth',2,'g');
xlabel('Domain');
ylabel('f(Domain)');
title('f:R+ -> R     f(x) = 1/(1+x)');
legend('f(5)','First Taylor Approximation of f at 2.5',
       'Second Taylor Approximation of f at 2.5');
grid on

[FirstTaylor,SecondTaylor] = F1_Taylor_Appr(5,domain);
% Plot the function along with its first and
% second taylor approximations at point 5
figure(2);
pl = plot(domain,func1,'linewidth',2,'b',
          domain,FirstTaylor,'linewidth',2,'r',
          domain,SecondTaylor,'linewidth',2,'g');
xlabel('Domain');
ylabel('f(Domain)');
title('f:R+ -> R     f(x) = 1/(1+x)');
legend('f(5)','First Taylor Approximation of f at 5',
       'Second Taylor Approximation of f at 5');
grid on

[FirstTaylor,SecondTaylor] = F1_Taylor_Appr(7.5,domain);
```

```
% Plot the function along with its first and
% second taylor approximations at point 7.5
figure(3);
pl = plot(domain,func1,'linewidth',2,'b',
          domain,FirstTaylor,'linewidth',2,'r',
          domain,SecondTaylor,'linewidth',2,'g');
xlabel('Domain');
ylabel('f(Domain)');
title('f:R+ -> R     f(x) = 1/(1+x)');
legend('f(5)','First Taylor Approximation of f at 7.5',
       'Second Taylor Approximation of f at 7.5');
grid on

[FirstTaylor,SecondTaylor] = F1_Taylor_Appr(10,domain);
% Plot the function along with its first and
% second taylor approximations at point 10
figure(4);
pl = plot(domain,func1,'linewidth',2,'b',
          domain,FirstTaylor,'linewidth',2,'r',
          domain,SecondTaylor,'linewidth',2,'g');
xlabel('Domain');
ylabel('f(Domain)');
title('f:R+ -> R     f(x) = 1/(1+x)');
legend('f(5)','First Taylor Approximation of f at 10',
       'Second Taylor Approximation of f at 10');
grid on

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%% Second Exercise %%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
start = 0.;
finish = 20.;
precision = 250;
x1 = linspace(start, finish, precision);
x2 = x1;
[X,Y] = meshgrid(x1,x2);
func2 = 1./(1.+X.+Y);

% Plot the 3d plot of the second function using mesh
figure(5);
mesh(func2,'LineWidth',2);
rotate3d on;
axis([0 100 0 100 0 1]);
title('f:R^2+ -> R    f(x) = 1/(1+x1+x2)');
zlabel('f(x1,x2)');
xlabel('x1');
ylabel('x2');
grid on

% Plot the level sets of the second function using contour
figure(6);
contourf(func2,'LineWidth',2);
colormap hot
axis([0 100 0 100 0 1]);
title('f:R^2+ -> R    f(x) = 1/(1+x1+x2)    Level Sets');
xlabel('x1');
ylabel('x2');
grid on

[FirstTaylor,SecondTaylor] = F2_Taylor_Appr(2,2,X,Y);
% Common 3d plot of the second function with its
% first Taylor Approximation at point (2,2)
figure(7);hold on
```

```
mesh(func2,'LineWidth',2);
surf(FirstTaylor,'FaceColor','red','EdgeColor','none');
rotate3d on;
axis([0 150 0 150 0 1]);
title('f:R^2+ -> R      f(x) = 1/(1+x1+x2)');
zlabel('f(x1,x2)');
xlabel('x1');
ylabel('x2');
grid on

% Common 3d plot of the second function with its
% second Taylor Approximation at point (2,2)
figure(8);hold on
mesh(func2,'LineWidth',2);
surf(SecondTaylor,'FaceColor','red','EdgeColor','none');
rotate3d on;
axis([0 150 0 150 0 1]);
title('f:R^2+ -> R      f(x) = 1/(1+x1+x2)');
zlabel('f(x1,x2)');
xlabel('x1');
ylabel('x2');
grid on

[FirstTaylor,SecondTaylor] = F2_Taylor_Appr(5,5,X,Y);
% Common 3d plot of the second function with its
% first Taylor Approximation at point (5,5)
figure(9);hold on
mesh(func2,'LineWidth',2);
surf(FirstTaylor,'FaceColor','red','EdgeColor','none');
rotate3d on;
axis([0 150 0 150 0 1]);
title('f:R^2+ -> R      f(x) = 1/(1+x1+x2)');
zlabel('f(x1,x2)');
```

```
xlabel('x1');
ylabel('x2');
grid on

% Common 3d plot of the second function with its
% second Taylor Approximation at point (5,5)
figure(10);hold on
mesh(func2,'LineWidth',2);
surf(SecondTaylor,'FaceColor','red','EdgeColor','none');
rotate3d on;
axis([0 150 0 150 0 1]);
title('f:R^2+ -> R      f(x) = 1/(1+x1+x2)');
zlabel('f(x1,x2)');
xlabel('x1');
ylabel('x2');
grid on

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%% Fifth Exercise %%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

start = 0.01;
finish = 25.;
precision = 1000000;
domain = linspace(start,finish,precision);
func1 = domain.^3;
func2 = domain.^0.25;
func3 = domain.^(-2);

% Plot the function along with its first and
% second taylor approximations at point 2.5
```

```
figure(11);
pl = plot(domain,func1,'linewidth',2,'b',
           domain,func2,'linewidth',2,'r',
           domain,func3,'linewidth',2,'g');
xlabel('Domain');
ylabel('f(Domain)');
title('f:R++ -> R      f(x) = x^a');
legend('a=3','a=0.25','a=-2');
axis([-0.5 10 -2.5 25]);
grid on

start = -20.;
finish = 20.;
precision = 500;
x1 = linspace(start,finish,precision);
x2 = x1;
[X,Y] = meshgrid(x1,x2);
norm2 = (X.^2.+Y.^2).^(1/2);

% Plot the 3d plot of the second function using mesh
figure(12);
mesh(norm2,'LineWidth',2);
rotate3d on;
title('f:R^2 -> R      f(x) = ||X||');
zlabel('f(x1,x2)');
xlabel('x1');
ylabel('x2');
grid on

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%% Exercise 7 %%%%%%%%%%%%%%%%%%%%%%%%%
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

A = 200*rand(3,2)-100;
x = 200*rand(2,1)-100;
b = A*x;

precision = 15;
x1 = linspace(x(1,1)-abs(x(1,1))/1000,x(1,1)+abs(x(1,1))/1000,precision);
x2 = linspace(x(2,1)-abs(x(2,1))/1000,x(2,1)+abs(x(2,1))/1000,precision);
[X,Y] = meshgrid(x1,x2);

LeastSquares = (A(1,1)*X+A(1,2)*Y-b(1,1)).^2 ...
              +(A(2,1)*X+A(2,2)*Y-b(2,1)).^2 ...
              +(A(3,1)*X+A(3,2)*Y-b(3,1)).^2;

% Plot the 3d plot of the Least Squares Equation
figure(1);
mesh(X,Y,LeastSquares,'LineWidth',2);
rotate3d on;
title('f:R^2 -> R     f(x) = ||Ax-b||^2');
zlabel('f(x)');
xlabel('x1');
ylabel('x2');
grid on

% Plot the level sets of the Least Squares Equation
figure(2);
contourf(X,Y,LeastSquares,'LineWidth',2);
colormap hot
title('f:R^2 -> R     f(x) = ||Ax-b||^2     Level Sets');
xlabel('x1');
ylabel('x2');
grid on
```

```
b+=randn(3,1); % Adding Gaussian noise to the b vector
LeastSquaresE = (A(1,1)*X+A(1,2)*Y-b(1,1)).^2 ...
                +(A(2,1)*X+A(2,2)*Y-b(2,1)).^2 ...
                +(A(3,1)*X+A(3,2)*Y-b(3,1)).^2;

% Plot the 3d plot of the Least Squares Equation
figure(3);
mesh(X,Y,LeastSquaresE,'LineWidth',2);
rotate3d on;
title('f:R^2 -> R      f(x) = ||Ax-be||^2');
zlabel('f(x)');
xlabel('x1');
ylabel('x2');
%xlim([X(1,1)-10 X(1,size(X,2))+10]);
%ylim([Y(1,1)-10 Y(size(Y,1),1)+10]);

grid on

% Plot the level sets of the Least Squares Equation
figure(4);
contourf(X,Y,LeastSquaresE,'LineWidth',2);
colormap hot
title('f:R^2 -> R      f(x) = ||Ax-be||^2      Level Sets');
xlabel('x1');
ylabel('x2');
grid on

disp('x')
disp(x)
disp('X')
disp(X)
disp('Y')
```

```
disp(Y)
disp('LeastSquares')
disp(LeastSquares)
disp('LeastSquaresE')
disp(LeastSquaresE)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [FirstTaylor, SecondTaylor] = F1_Taylor_Appr(Num,Dom)
  FirstTaylor = (1-Dom+2*Num)./(1+Num)^2;
  SecondTaylor = (1+3*Num^2+3*Num-3*Num.*Dom.-Dom.+Dom.^2)...
                 ./(1+Num)^3;
  return
endfuncion

function [FirstTaylor, SecondTaylor] = F2_Taylor_Appr(Num1,Num2,X,Y)
  FirstTaylor = 1.-((Num1+Num2)^2.+X.+Y)./(1+Num1+Num2)^2;
  SecondTaylor = FirstTaylor.+((X.-Num1.+Y.-Num2).^2)...
                 ./(1+Num1+Num2)^3;
  return
endfuncion
```