ECE 595 **Homework 3** **Due: Feb. 28, 4 PM**

1. For the following function of two variables, find the stationary points, i.e., where the gradient is zero.

$$F(\mathbf{X}) = (x_1 + x_2)^4 - 12x_1x_2 + x_1 + x_2 + 1$$

Which of the stationary points, if any, is the minimum?

Gradient: $\frac{\partial f(\underline{x})}{\partial x_1} = 4(x_1+x_2)^3 - 12x_2 + 1$

$\frac{\partial f(\underline{x})}{\partial x_2} = 4(x_1+x_2)^3 - 12x_1 + 1$

$\therefore \nabla_x f(\underline{x}) = \begin{bmatrix} 4(x_1+x_2)^3 - 12x_2 + 1 \\ 4(x_1+x_2)^3 - 12x_1 + 1 \end{bmatrix}$

① to get stationary points: let $\begin{cases} 4(x_1+x_2)^3 - 12x_2 + 1 = 0 \\ 4(x_1+x_2)^3 - 12x_1 + 1 = 0 \end{cases}$

$p_1 \begin{cases} x_1 = 0.08497 \\ x_2 = 0.08497 \end{cases}$ $p_2 \begin{cases} x_1 = 0.5655 \\ x_2 = 0.5655 \end{cases}$ $p_3 \begin{cases} x_1 = -0.6504 \\ x_2 = -0.6504 \end{cases}$

② to judge if these points are minimum:

Second order / Hessian:

$$H = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1^2} & \frac{\partial f(x)}{\partial x_1 \partial x_2} \\ \frac{\partial f(x)}{\partial x_2 \partial x_1} & \frac{\partial f(x)}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 12(x_1+x_2)^2 & 12(x_1+x_2)^2 - 12 \\ 12(x_1+x_2)^2 - 12 & 12(x_1+x_2)^2 \end{bmatrix}$$

$\therefore |H(p_1)| = \begin{vmatrix} 0.3466 & -11.6534 \\ -11.6534 & 0.3466 \end{vmatrix} < 0$

$|H(p_2)| = \begin{vmatrix} 15.3499 & 3.3499 \\ 3.3499 & 15.3499 \end{vmatrix} > 0$

$|H(p_3)| = \begin{vmatrix} 20.3049 & 8.3049 \\ 8.3049 & 20.3049 \end{vmatrix} > 0$

$\therefore$ p_2 and p_3 is minimum

p_1 is not minimum.

2. For the function $F(\mathbf{X}) = \frac{7}{2}x_1^2 - x_2^2 - 6x_1x_2$, find (a) the stationary points, and (b) the gradient at $\mathbf{x} = [1\ 1]^T$ in the direction of $\mathbf{p} = [-1\ 1]^T$.

(a) Gradient: $\frac{f(x)}{\partial x_1} = 7x_1 - 6x_2$ $\quad \frac{f(x)}{\partial x_2} = -2x_2 - 6x_1$

$$\nabla_x f(x) = \begin{bmatrix} 7x_1 - 6x_2 \\ -2x_2 - 6x_1 \end{bmatrix}$$

let $\begin{cases} 7x_1 - 6x_2 = 0 \\ -2x_2 - 6x_1 = 0 \end{cases}$ get: $\begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$

$\therefore$ The stationary point is $\begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$

(b) $\underline{p} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ $\quad \|p\|_2 = \sqrt{2}$

$\nabla f(\underline{x}^*) = \begin{bmatrix} 1 \\ -8 \end{bmatrix}$

Directional derivative $= \frac{\underline{p}^T \nabla f(\underline{x}^*)}{\|p\|_2} = \frac{[-1\ \ 1]\begin{bmatrix} 1 \\ -8 \end{bmatrix}}{\sqrt{2}} = -\frac{9}{\sqrt{2}} = -\frac{9\sqrt{2}}{2}$

3. Find the minimum point for the function, $F(\mathbf{X}) = \frac{1}{2}\mathbf{x}^T \begin{bmatrix} 6 & -2 \\ -2 & 6 \end{bmatrix} \mathbf{x} + \begin{bmatrix} -1 & -1 \end{bmatrix} \mathbf{x}$,

using the steepest descent algorithm with the initial guess of $\mathbf{x}_0 = [0\ 0]^T$ and a learning rate of 0.1. Show a few iterations by hand calculations. Use MATLAB to verify your calculations. Show your code and final results.

Gradient: $\nabla f(\underline{x}) = \text{"}\ \underline{A}\,\underline{x} + \underline{b}\ \text{"} = \begin{bmatrix} 6 & -2 \\ -2 & 6 \end{bmatrix} \underline{x} + \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

let $\underline{x}_0 = [0\ \ 0]^T$ & $\alpha = 0.1$

$\underline{x}_1 := \underline{x}_0 - \alpha \nabla f(\underline{x}_0)$

$= \begin{bmatrix} 0 \\ 0 \end{bmatrix} - 0.1 * \left(\begin{bmatrix} 6 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right) = \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix}$

$\underline{x}_2 := \underline{x}_1 - \alpha \nabla f(\underline{x}_1)$

$= \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix} - 0.1 * \left(\begin{bmatrix} 6 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right) = \begin{bmatrix} 0.16 \\ 0.16 \end{bmatrix}$

$\underline{x}_3 := \underline{x}_2 - \alpha \nabla f(\underline{x}_2)$

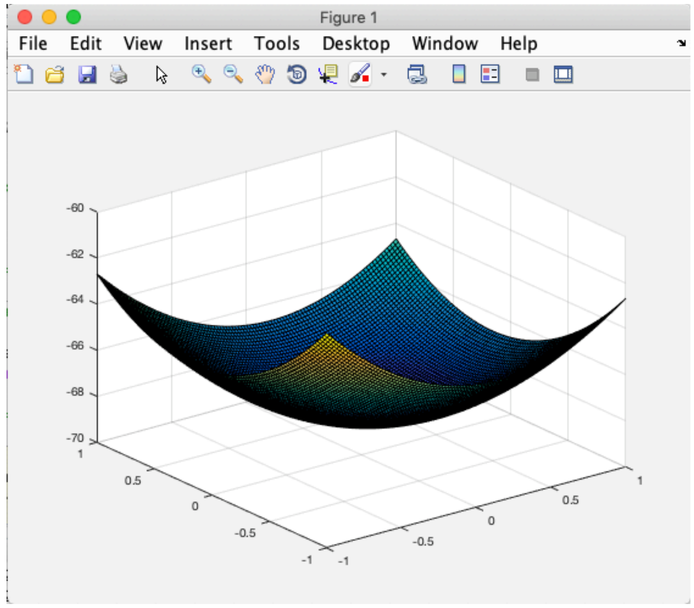
$= \begin{bmatrix} 0.16 \\ 0.16 \end{bmatrix} - 0.1 * \left(\begin{bmatrix} 6 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} 0.16 \\ 0.16 \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right) = \begin{bmatrix} 0.196 \\ 0.196 \end{bmatrix}$

For MATLAB coding, I got the result like:

```
Command Window
>> question_3
iteration time: 1, x1 = [0.100000; 0.100000].
iteration time: 2, x2 = [0.160000; 0.160000].
iteration time: 3, x3 = [0.196000; 0.196000].
iteration time: 4, x4 = [0.217600; 0.217600].
iteration time: 5, x5 = [0.230560; 0.230560].
```

```
Command Window
iteration time: 22, x22 = [0.249997; 0.249997].
iteration time: 23, x23 = [0.249998; 0.249998].
iteration time: 24, x24 = [0.249999; 0.249999].
iteration time: 25, x25 = [0.249999; 0.249999].
iteration time: 26, x26 = [0.250000; 0.250000].
iteration time: 27, x27 = [0.250000; 0.250000].
iteration time: 28, x28 = [0.250000; 0.250000].
iteration time: 29, x29 = [0.250000; 0.250000].
iteration time: 30, x30 = [0.250000; 0.250000].
```

The function will converge after around 26 iterations, also I plot the function to check if there is a global minimum.



The shape of the function is a perfect bowl shape and its global minimum is at the point $[0.25\ 0.25]^T$.

4. For the function $F(\mathbf{X}) = \frac{7}{2}x_1^{\ 2} - x_2^{\ 2} - 6x_1x_2$, perform two iterations of the **steepest descent** algorithm, starting at the initial guess of $\mathbf{x}_0 = [1\ 1]^T$.

$$F(X) = \frac{7}{2}x_1^2 - x_2^2 - 6x_1x_2 = [x_1 \ \ x_2]\begin{bmatrix} \frac{7}{2} & -3 \\ -3 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Gradient : $\nabla F(\underline{X}) = 2\underline{A}\,\underline{X} = \begin{bmatrix} 7 & -6 \\ -6 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

let $X_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\alpha = 0.1$

$X_1 := X_0 - \alpha \nabla F(X_0)$

$= \begin{bmatrix} 1 \\ 1 \end{bmatrix} - 0.1 * \left(\begin{bmatrix} 7 & -6 \\ -6 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0.9 \\ 1.8 \end{bmatrix}$

$X_2 := X_1 - \alpha \nabla F(X_1)$

$= \begin{bmatrix} 0.9 \\ 1.8 \end{bmatrix} - 0.1 * \left(\begin{bmatrix} 7 & -6 \\ -6 & -2 \end{bmatrix}\begin{bmatrix} 0.9 \\ 1.8 \end{bmatrix}\right) = \begin{bmatrix} 1.35 \\ 2.7 \end{bmatrix}$

5. Determine the eigenvalues and eigenvectors for the following matrices.

(a) $\begin{bmatrix} 3 & 2 & -2 \\ -3 & -1 & 3 \\ 1 & 2 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ -2 & 0 & 3 \end{bmatrix}$ (c) $\begin{bmatrix} 15 & 7 & -7 \\ -1 & 1 & 1 \\ 13 & 7 & -5 \end{bmatrix}$

(a) $\underline{A} = \begin{bmatrix} 3 & 2 & -2 \\ -3 & -1 & 3 \\ 1 & 2 & 0 \end{bmatrix}$ let $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$|\underline{A} - \lambda I| = \begin{vmatrix} 3-\lambda & 2 & -2 \\ -3 & -1-\lambda & 3 \\ 1 & 2 & -\lambda \end{vmatrix} = \overset{-\lambda^3+2\lambda^2+3\lambda}{(3-\lambda)(-1-\lambda)(-\lambda)} + \overset{6}{2\times3\times1} + \overset{12}{(-3)\times2\times(-2)}$$

$$- \underset{2+2\lambda}{(-2)\times(-1-\lambda)\times1} - \underset{6\lambda}{2\times(-3)\times(-\lambda)} - \underset{18-6\lambda}{2\times3\times(3-\lambda)}$$

$$= -\lambda^3 + 2\lambda^2 + 3\lambda + \cancel{6} + \cancel{1}2 - 2 - 2\lambda - \cancel{6\lambda} - \cancel{1}8 + \cancel{6\lambda}$$

$$= -\lambda^3 + 2\lambda^2 + \lambda - 2 = 0$$

$$-\lambda^3+2\lambda^2+\lambda-2 = 0 \implies \begin{cases} \lambda_1 = 2 \\ \lambda_2 = -1 \\ \lambda_3 = 1 \end{cases}$$

when $\lambda_1 = 1$ let $\underline{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

$(A - \lambda I)\underline{X} = 0$

$$\left(\begin{bmatrix} 3 & 2 & -2 \\ -3 & -1 & 3 \\ 1 & 2 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{bmatrix} 2 & 2 & -2 \\ -3 & -2 & 3 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{cases} 2x_1 + 2x_2 - 2x_3 = 0 \\ -3x_1 - 2x_2 + 3x_3 = 0 \\ x_1 + 2x_2 - x_3 = 0 \end{cases}$$

when $x_2 = 0$, $x_1 = x_3$

$\therefore$ get $p_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

when $\lambda_2 = -1$

$$\left(\begin{bmatrix} 3 & 2 & -2 \\ -3 & -1 & 3 \\ 1 & 2 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{bmatrix} 4 & 2 & -2 \\ -3 & 0 & 3 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{cases} 4x_1 + 2x_2 - 2x_3 = 0 \quad (1) \\ -3x_1 + 3x_3 = 0 \quad (2) \\ x_1 + 2x_2 + x_3 = 0 \quad (3) \end{cases}$$

According (2) : $x_1 = x_3$

$\therefore$ (1) $\rightarrow$ $2x_1 + 2x_2 = 0$ $\quad x_1 = -x_2$

$\therefore$ If $x_1 = 1$ then $x_2 = -1$, $x_3 = 1$

$\therefore$ $p_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

when $\lambda_3 = 2$

$$\left(\begin{bmatrix} 3 & 2 & -2 \\ -3 & -1 & 3 \\ 1 & 2 & 0 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{bmatrix} 1 & 2 & -2 \\ -3 & -3 & 3 \\ 1 & 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{cases} x_1 + 2x_2 - 2x_3 = 0 \\ -x_1 - x_2 + x_3 = 0 \\ x_1 + 2x_2 - 2x_3 = 0 \end{cases}$$

$x_2 - x_3 = 0$ $\quad x_2 = x_3$ $\quad$ let $x_1 = 0$

$p_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

(b) $\underline{A} = \begin{bmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ -2 & 0 & 3 \end{bmatrix}$, let $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & -2 & 2 \\ -2 & 1-\lambda & 2 \\ -2 & 0 & 3-\lambda \end{vmatrix} = \overset{-\lambda^3+5\lambda^2-7\lambda+3}{(1-\lambda)^2(3-\lambda)} + 0 + \overset{8}{(-2)\times 2\times(-2)}$$

$$- \underset{4\lambda-4}{2\times(1-\lambda)\times(-2)} - \underset{12-4\lambda}{(-2)\times(-2)\times(3-\lambda)} - 0$$

$$= -\lambda^3+5\lambda^2-7\lambda+3+\cancel{8}-\cancel{4\lambda}+\cancel{4}-\cancel{12}+\cancel{4\lambda}$$

$$= -\lambda^3+5\lambda^2-7\lambda+3 = 0$$

$$-\lambda^3+5\lambda^2-7\lambda+3=0 \Rightarrow \begin{cases} \lambda_1 = 3 \\ \lambda_2 = 1 \end{cases}$$

when $\lambda_1 = 3$

$(A - \lambda I)X = 0$

$$\left(\begin{bmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ -2 & 0 & 3 \end{bmatrix} - 3\times\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{bmatrix} -2 & -2 & 2 \\ -2 & -2 & 2 \\ -2 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{cases} -x_1 - x_2 + x_3 = 0 \\ -x_1 - x_2 + x_3 = 0 \\ -2x_1 = 0 \end{cases} \qquad x_1 = 0,\ x_2 = x_3 \qquad \therefore p_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

when $\lambda_2 = 1$

$(A - \lambda I)X = 0$

$$\left(\begin{bmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ -2 & 0 & 3 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = 0$$

$$\begin{bmatrix} 0 & -2 & 2 \\ -2 & 0 & 2 \\ -2 & 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{cases} -2x_2 + 2x_3 = 0 \\ -2x_1 + 2x_3 = 0 \\ -2x_1 + 2x_3 = 0 \end{cases} \qquad x_1 = x_2 = x_3 \qquad \therefore \underline{p_2} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

(c) $\underline{A} = \begin{bmatrix} 15 & 7 & -7 \\ -1 & 1 & 1 \\ 13 & 7 & -5 \end{bmatrix}$ let $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$$|A - \lambda I| = \begin{vmatrix} 15-\lambda & 7 & -7 \\ -1 & 1-\lambda & 1 \\ 13 & 7 & -5-\lambda \end{vmatrix} = \overset{-\lambda^3+11\lambda^2+65\lambda-75}{(15-\lambda)(1-\lambda)(-5-\lambda)} + \overset{49}{(-1)\times 7\times(-7)} + \overset{91}{7\times 1\times 13}$$

$$- \underset{91\lambda-91}{(-7)\times(1-\lambda)\times 13} - \underset{35+7\lambda}{(-1)\times 7\times(-5-\lambda)} - \underset{105-7\lambda}{7\times 1\times(15-\lambda)}$$

$$= -\lambda^3 + 11\lambda^2 + 6\cancel{5}\lambda - 75 + 49 + 91 - \cancel{9}1\lambda + 91 - 35 - \cancel{7}\lambda - 105 + \cancel{7}\lambda$$

$$= -\lambda^3 + 11\lambda^2 - 26\lambda + 16 = 0$$

$$-\lambda^3 + 11\lambda^2 - 26\lambda + 16 = 0 \Rightarrow \begin{cases} \lambda_1 = 2 \\ \lambda_2 = 8 \\ \lambda_3 = 1 \end{cases}$$

when $\lambda_1 = 2$

$(A - \lambda I)X = 0$

$$\left(\begin{bmatrix} 15 & 7 & -7 \\ -1 & 1 & 1 \\ 13 & 7 & -5 \end{bmatrix} - 2\times\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{bmatrix} 13 & 7 & -7 \\ -1 & -1 & 1 \\ 13 & 7 & -7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{cases} 13x_1 + 7x_2 - 7x_3 = 0 \\ -x_1 - x_2 + x_3 = 0 \\ 13x_1 + 7x_2 - 7x_3 = 0 \end{cases}$$

let $x_2 = 1$

$$p_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

when $\lambda_2 = 8$

$$\left(\begin{bmatrix} 15 & 7 & -7 \\ -1 & 1 & 1 \\ 13 & 7 & -5 \end{bmatrix} - 8\times\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{bmatrix} 7 & 7 & -7 \\ -1 & -7 & 1 \\ 13 & 7 & -13 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{cases} x_1 + x_2 - x_3 = 0 \\ -x_1 - 7x_2 + x_3 = 0 \\ 13x_1 + 7x_2 - 13x_3 = 0 \end{cases}$$

let $x_1 = 1$,

$$p_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

when $\lambda_3 = 1$

$$\left(\begin{bmatrix} 15 & 7 & -7 \\ -1 & 1 & 1 \\ 13 & 7 & -5 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right)\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{bmatrix} 14 & 7 & -7 \\ -1 & 0 & 1 \\ 13 & 7 & -6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\begin{cases} 14x_1 + 7x_2 - 7x_3 = 0 \\ -x_1 + x_3 = 0 \\ 13x_1 + 7x_2 - 6x_3 = 0 \end{cases}$$

let $x_1 = x_3 = 1$

$$p_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

6. For the function given in Problem 4, show an iteration of **Newton's method** using the initial guess of $\mathbf{x}_0 = [1\ 1]^T$. Complete the solution in MATLAB. Compare the two algorithms for this problem.

$$F(\underline{x}) = \frac{7}{2}x_1^2 - x_2^2 - 6x_1x_2 = [x_1\ \ x_2]\begin{bmatrix} \frac{7}{2} & -3 \\ -3 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Gradient $\nabla F(\underline{x}) = 2\underline{A}\,\underline{x} = \begin{bmatrix} 7 & -6 \\ -6 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

Hessian $\underline{H} = 2\underline{A} = \begin{bmatrix} 7 & -6 \\ -6 & -2 \end{bmatrix}$ $\quad \underline{H}^{-1} = \begin{bmatrix} 0.04 & -0.12 \\ -0.12 & 0.14 \end{bmatrix}$

let $\underline{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$\underline{x}_1 := \underline{x}_0 - H^{-1}\nabla F(\underline{x}_0)$$

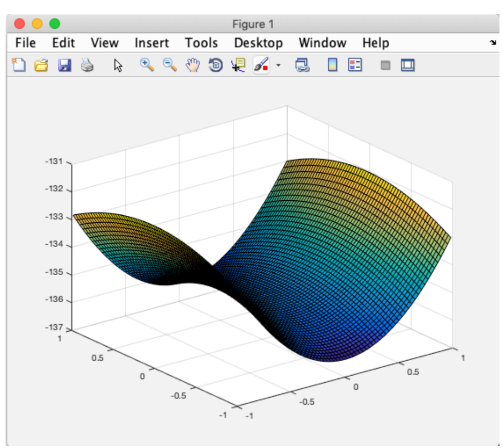
$$= \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0.04 & -0.12 \\ -0.12 & 0.14 \end{bmatrix}\left(\begin{bmatrix} 7 & -6 \\ -6 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$$

$$= \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

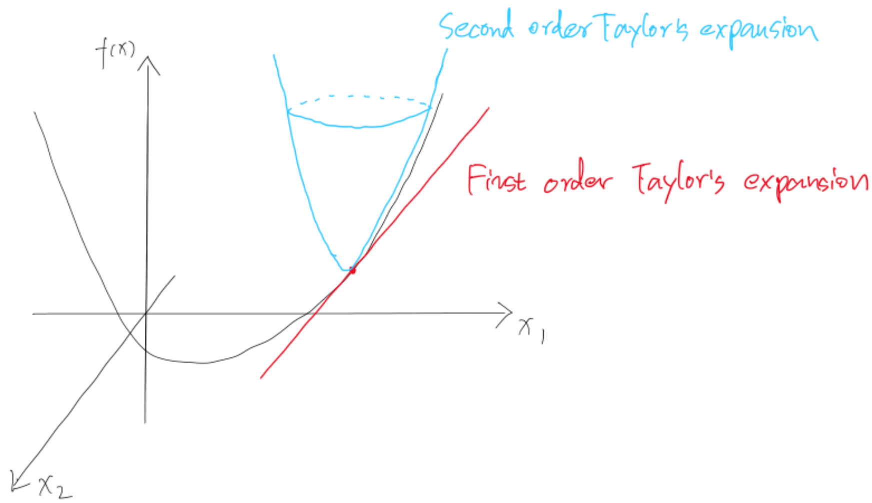
For MATLAB coding:

```
>> question_6
x found by Newtons Method is: [0.000000; -0.000000]
```

Also, I plot the function and find that the point $[0\ 0]^T$ is a saddle point.



To compare the steepest descent and Newton's method, in my opinion, according to Taylor's series, steepest descent uses a straight line to fit the curve, while Newton's method is using a curve to fit the curve.



Therefore, the Newton's method is much faster than steepest descent.

Code:

Question_3.m

```
%% Machine Learning Homework 3 Question 3
% Author: Xinrun Zhang
% Time: 02/28/2019 13:42
% ======================================================================

%% initializting
% gradient = A * x[i] + b;
A = [6 -2; -2 6];
b = [-1; -1];

% initial guess x, learning rate alpha and iteration
x = [0; 0];
alpha = 0.1;
itr = 30;
% ======================================================================

%% steepest descent
for i = 1:itr
    x = x - alpha * (A * x + b);
    fprintf('iteration time: %d, x%d = [%f; %f].\n',i,i,x(1),x(2));
end
% ======================================================================

%% plot
[x1, x2] = meshgrid(-1:0.02:1);
z = 3 * (x1.^2) + 3 * (x2.^2) - 2 * x1 * x2 - x1 - x2;
surf(x1, x2, z);
% ======================================================================
```

Question_4.m

```
%% Machine Learning Homework 3 Question 6
% Author: Xinrun Zhang
% Time: 02/28/2019 14:03
% ======================================================================

%% initializting
% gradient = 2 * A * x
% Hessian = 2 * A
A = [7/2 -3; -3 1];

H = 2 * A;
H_inv = (H)^-1;

% initial guess x
x = [1; 1];
% ======================================================================

%% Newton's Method
x = x - H_inv * (2 * A * x);
fprintf('x found by Newtons Method is: [%f; %f]\n',x(1), x(2));
% ======================================================================

%% plot
[x1, x2] = meshgrid(-1:0.03:1);
z = (7/2) * (x1.^2) - (x2.^2) - 6 * x1 * x2;
surf(x1, x2, z);
% ======================================================================
```