DSA3102 HW1
Ethan Keck Jun Wei (March 8, 2024)

---

**Problem 1.**

*Answer:* Given

$$f(a,b) = \sum_{i=1}^{n}(a + bx_i - y_i)^2$$
$$\Rightarrow \nabla f(a.b) = \begin{pmatrix} \sum_{i=1}^{n} 2(a + bx_i - y_i) \\ \sum_{i=1}^{n} 2x_i(a + bx_i - y_i) \end{pmatrix}$$

Let $\nabla f(a,b) = \mathbf{0}$

$$\begin{pmatrix} 2\sum_{i=1}^{n}(a + bx_i - y_i) \\ 2\sum_{i=1}^{n} x_i(a + bx_i - y_i) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Then,

$$\sum_{i=1}^{n}(a + bx_i - y_i) = 0$$
$$na + b\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i = 0$$
$$a = \frac{1}{n}\sum_{i=1}^{n} y_i + \frac{1}{n}b\sum_{i=1}^{n} x_i$$
$$a = \bar{y} - b\bar{x}$$

and,

$$\sum_{i=1}^{n} x_i(a + bx_i - y_i) = 0$$
$$a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i y_i = 0$$
$$an\bar{x} + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i$$
$$n\bar{x}\bar{y} - bn\bar{x}^2 + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i$$
$$b\left(n\bar{x}^2 - \sum_{i=1}^{n} x_i^2\right) = n\bar{x}\bar{y} - \sum_{i=1}^{n} x_i y_i$$
$$b = \frac{n\bar{x}\bar{y} - \sum_{i=1}^{n} x_i y_i}{n\bar{x}^2 - \sum_{i=1}^{n} x_i^2}$$

For these solutions to be optimal, $f(a, b)$ must be convex. Let $f_i(a, b) = (a + bx_i - y_i)^2$, such that $f(a, b) = \sum_{i=1}^{n} f_i(a, b)$ . We show $f_j(a, b) = (a + bx_i - y_i)^2$ is convex;

$$
\begin{gathered}
\nabla f_i(a,b) = \begin{pmatrix} 2(a + bx_i - y_i) \\ 2x_i(a + bx_i - y_i) \end{pmatrix}, \\
H_{f_i}(a,b) = \begin{pmatrix} 2 & 2x_i \\ 2x_i & 2x_i^2 \end{pmatrix} \\
\Rightarrow \det \begin{pmatrix} \lambda - 2 & -2x_i \\ -2x_i & \lambda - 2x_i^2 \end{pmatrix} \\
= (\lambda - 2)(\lambda - 2x_i^2) - 4x_i^2 \\
= \lambda^2 - 2\lambda - 2\lambda x_i^2 + 4x_i^2 - 4x_i^2 \\
= \lambda(\lambda - 2(1 + x_i^2)) \\
\Rightarrow \lambda = 0, \lambda = 2(1 + x_i^2)
\end{gathered}
$$

Since $\lambda = 2(1 + x_i^2) > 0$ for all $x_i \in \mathbb{R}$, the eigenvalues of $H_{f_i}(a, b)$ are non-negative. Then $H_{f_i}(a, b)$ is **Positive Semi-definite**, and thus $f_i(a, b)$ is convex. By Corollary 3.1, $f(a, b) = \sum_{i=1}^{n} f_i(a, b)$ is also convex. Since $\left(a = \bar{y} - b\bar{x}, b = \frac{n\bar{x}\bar{y} - \sum_{i=1}^{n} x_i y_i}{n\bar{x}^2 - \sum_{i=1}^{n} x_i^2}\right)$ is a local min, it is thus an optimal solution. $\square$

**Problem 2.**

*Answer:* i) Given

$$f(\mathbf{x}) = \ln\left(x_1^2 + 1\right) + x_1^2 - 2x_1x_2 + 4x_2^2 + x_3^4 - 8x_3^3 + 16x_3^2$$

We find

$$\begin{aligned}
\nabla f(\mathbf{x}) &= \begin{pmatrix} \frac{2x_1}{x_1^2+1} + 2x_1 - 2x_2 \\ 8x_2 - 2x_1 \\ 4x_3^3 - 24x_3^2 + 32x_3 \end{pmatrix} \\
&= \begin{pmatrix} \frac{2x_1 + 2(x_1 - x_2)(x_1^2+1)}{x_1^2+1} \\ 8x_2 - 2x_1 \\ 4x_3(x_3 - 4)(x_3 - 2) \end{pmatrix}
\end{aligned}$$

Let $\nabla f(\mathbf{x}) = \mathbf{0}$

$$\begin{gathered}
\begin{pmatrix} \frac{2x_1 + 2(x_1 - x_2)(x_1^2+1)}{x_1^2+1} \\ 8x_2 - 2x_1 \\ 4x_3(x_3 - 4)(x_3 - 2) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\Rightarrow 4x_3(x_3 - 4)(x_3 - 2) = 0 \\
x_3 = 0, x_3 = 4, x_3 = 2 \\
\Rightarrow 8x_2 - 2x_1 = 0 \\
x_1 = 4x_2 \\
\Rightarrow \frac{2x_1 + 2(x_1 - x_2)(x_1^2 + 1)}{x_1^2 + 1} = 0 \\
\Rightarrow 2(4x_2) + 2(4x_2 - x_2)((4x_2)^2 + 1) = 0 \\
\Rightarrow 8x_2 + 6x_2(16x_2^2 + 1) = 96x_2^3 + 6x_2 + 8x_2 = 0 \\
\Rightarrow 96x_2^3 + 14x_2 = 0 \\
\Rightarrow x_2 = 0, \text{then } x_1 = 0
\end{gathered}$$

Then the stationary points are: $(x_1, x_2, x_3) : (0,0,0), (0,0,2), (0,0,4)$

ii) We compute:

$$H_f(\mathbf{x}) = \begin{pmatrix} \frac{2}{x_1^2+1} - \frac{4x_1^2}{(x_1^2+1)^2} + 2 & -2 & 0 \\ -2 & 8 & 0 \\ 0 & 0 & 12x_3^2 - 48x_3 + 32 \end{pmatrix}$$

For $\mathbf{x} = (0,0,0)$ and $\mathbf{x} = (0,0,4)$,

$$\Rightarrow H_f(0,0,0) = \begin{pmatrix} 4 & -2 & 0 \\ -2 & 8 & 0 \\ 0 & 0 & 32 \end{pmatrix} = H_f(0,0,4)$$

$$\Rightarrow \det \begin{pmatrix} \lambda - 4 & 2 & 0 \\ 2 & \lambda - 8 & 0 \\ 0 & 0 & \lambda - 32 \end{pmatrix}$$

$$\Rightarrow (\lambda - 32)(\lambda^2 - 12\lambda + 28) = 0$$

$$\Rightarrow (\lambda - 32)(\lambda - (6 - 2\sqrt{2}))(\lambda - (6 + 2\sqrt{2})) = 0$$

With eigenvalues: $\lambda = 32, \lambda = 6 - 2\sqrt{2}, \lambda = 6 + 2\sqrt{2}$, that are greater than 0, thus $H_f(0,0,0)$ and $H_f(0,0,4)$ are positive definite and thus $\mathbf{x} = (0,0,0)$ and $\mathbf{x} = (0,0,4)$ are local minimizer.
For $\mathbf{x} = (0,0,2)$,

$$\Rightarrow H_f(0,0,2) = \begin{pmatrix} 4 & -2 & 0 \\ -2 & 8 & 0 \\ 0 & 0 & -16 \end{pmatrix}$$

$$\Rightarrow \det \begin{pmatrix} \lambda - 4 & 2 & 0 \\ 2 & \lambda - 8 & 0 \\ 0 & 0 & \lambda + 16 \end{pmatrix}$$

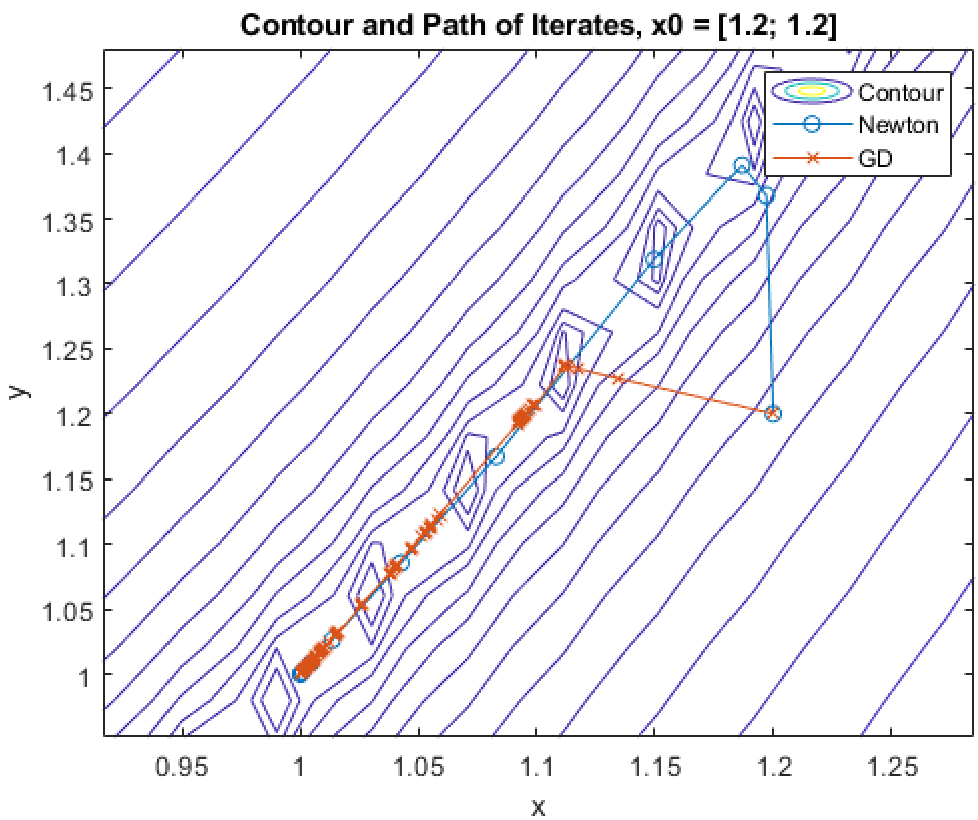
$$\Rightarrow (\lambda + 16)(\lambda^2 - 12\lambda + 28) = 0$$

$$\Rightarrow (\lambda + 16)(\lambda - (6 - 2\sqrt{2}))(\lambda - (6 + 2\sqrt{2})) = 0$$

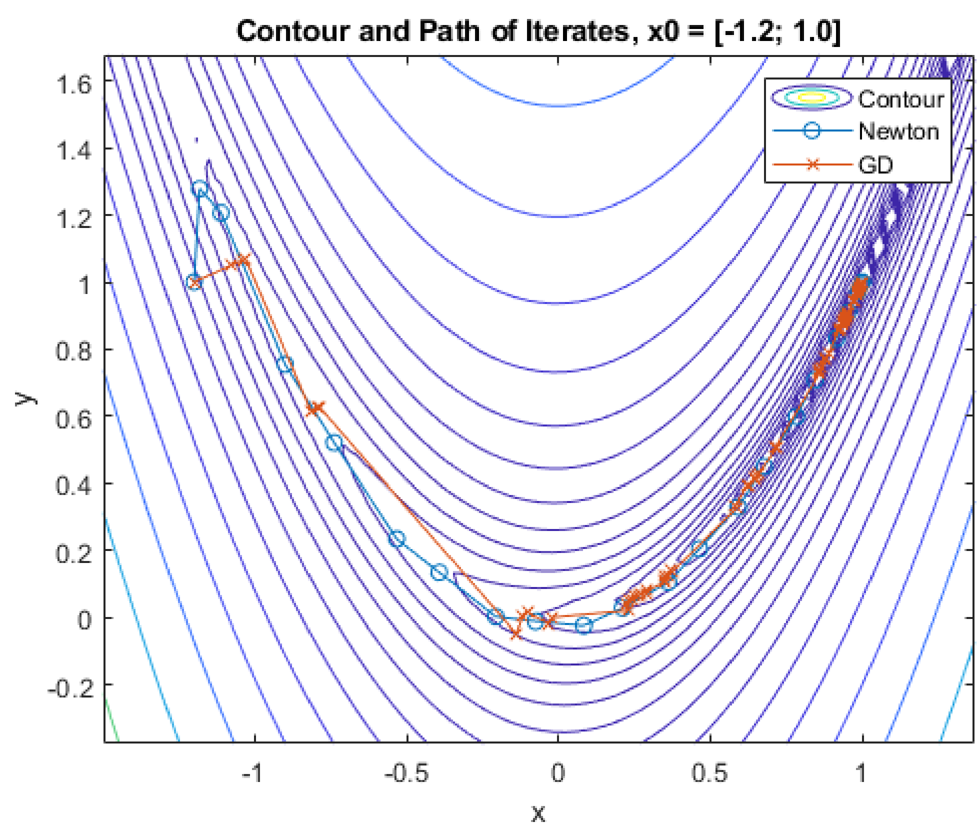
With eigenvalues: $\lambda = -16, \lambda = 6 - 2\sqrt{2}, \lambda = 6 + 2\sqrt{2}$, thus $H_f(0,0,2)$ is indefinite and thus $\mathbf{x} = (0,0,2)$ is a saddle point.
Since $f(0,0,0) = 0 = f(0,0,4)$, the global minimizers are: $\mathbf{x} = (0,0,0)$ and $\mathbf{x} = (0,0,4)$ □

**Problem 3.**

*Answer:* □



Contour and Path of Iterates, x0 = [1.2; 1.2]
Contour
Newton
GD
1.45
1.4
1.35
1.3
1.25
1.2
1.15
1.1
1.05
1
y
0.95
1
1.05
1.1
1.15
1.2
1.25
x

Contour and Path of Iterates, x0 = [-1.2; 1.0]
Contour
Newton
GD
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
y
-1
-0.5
0
0.5
1
x