آرین تشکر - ۴۰۰۲۳۴۹۴ تکلیف سوم درس شبکه های عصبی

E8.1. $F(x) = (x^3 - \frac{3}{4}x - \frac{1}{2})^{-1}$

$F(x) \approx F(x^*) + \nabla F(x)^T\Big|_{x=x^*}(x - x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 F(x)\Big|_{x=x^*}(x - x^*)$

(second order Taylor expansion)

i. First will calculate $\nabla F(x)$ and $\nabla^2 F(x)$:

$$\nabla F(x) = -1(3x^2 - \frac{3}{4})(x^3 - \frac{3}{4}x - \frac{1}{2})^{-2}$$

$$\nabla^2 F(x) = -1\left[6x(x^3 - \frac{3}{4}x - \frac{1}{2}) - 2(x^3 - \frac{3}{4}x - \frac{1}{2})^{-3}(3x^2 - \frac{3}{4})^2\right]$$

$$\rightarrow F_1(x) = \underbrace{F(-0.5)}_{-4} + (x+0.5)\underbrace{\cancel{\nabla f(x)}}_{0}\Big|_{x=-0.5} + \frac{1}{2}(x+0.5)^2\underbrace{\cancel{\nabla^2 F(x)}}_{48}\Big|_{x=-0.5}$$

$$= -4 + \frac{1}{2}(x+0.5)^2(48) = -4 + 24(x+0.5)^2$$

ii. $F_2(x) = 114925000(x - 11/10)^2/3 - 80000x + \frac{264500}{3}$

iii. Refer to the Matlab script.

E8.5. $F(x) = (x_1 + x_2)^4 - 12x_1x_2 + x_1 + x_2 + 1$

i. $\nabla F(x) = \begin{bmatrix} \frac{\partial}{\partial x_1} F(x) \\ \frac{\partial}{\partial x_2} F(x) \end{bmatrix} = \begin{bmatrix} 4(x_1 + x_2)^3 - 12x_2 + 1 \\ 4(x_1 + x_2)^3 - 12x_1 + 1 \end{bmatrix}$

for the rest refer to the matlab script.

E8.8.

Refer to the MATLAB Script.

E8.10. $F(x) = \frac{3}{2}x_1^2 + 2x_1x_2 + x_2^3 + 4x_1 + 4x_2$

$$\nabla F(x) = \begin{bmatrix} \frac{\partial F(x)}{\partial x_1} \\ \frac{\partial F(x)}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 3x_1 + 2x_2 + 4 \\ 2x_1 + 3x_2^2 + 4 \end{bmatrix}$$

$$\nabla^2 F(x) = \begin{bmatrix} \frac{\partial^2 F(x)}{\partial x_1^2} & \frac{\partial^2 F(x)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 F(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 F(x)}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 6x_2 \end{bmatrix}$$

i.

$x^* = [1 \quad 0]^T$

$$F_2(x) \approx F(x^*) + \nabla F(x)^T \Big|_{x=x^*} (x - x^*) + \frac{1}{2}(x - x^*)^T \nabla^2 F(x) \Big|_{x=x^*} (x - x^*)$$

$$= \frac{11}{2} + [7 \quad 6]\begin{bmatrix} x_1 - 1 \\ x_2 \end{bmatrix} + \frac{1}{2}[x_1 - 1 \quad x_2]\begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} x_1 - 1 \\ x_2 \end{bmatrix}$$

$$= 4x_1 + 4x_2 + 2x_1x_2 + \frac{3}{2}x_1^2$$

ii. $\nabla F_2(x) = \begin{bmatrix} 4 + 2x_2 + 3x_1 \\ 4 + 2x_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightarrow \begin{cases} 4 + 2x_2 + 3x_1 = 0 \rightarrow 2x_2 - 2 = 0 \\ 4 + 2x_1 = 0 \rightarrow x_1 = -2 \end{cases}$ $\quad x_2 = 1$

$\rightarrow$ stationary point: $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$, $\nabla^2 F_2(x) = \begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix}$, $\text{eig}(\nabla^2 F_2(x)) = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$ $\rightarrow$ saddle point of F_2

iii. $h = \nabla^2 F(x)$ @ $x = \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}$

$\text{eig}(h) = \begin{bmatrix} 2 \\ 7 \end{bmatrix} \longrightarrow$ positive definite

$g = \nabla F(x)$ @ $x = \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix} \longrightarrow$ non-zero $\rightarrow$ not an extrema

E8.12. $F(x) = x_1^2 + 2x_1x_2 + x_2^2 + (x_1 - x_2)^3$

i. Following the same method as last problem:

$$\hat{F}_2(x) = 4x_1^2 - 4x_1x_2 - 3x_1 + 4x_2^2 + 3x_2 + 1$$

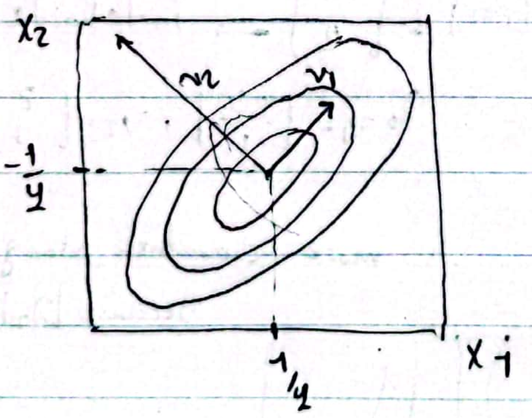
$$\nabla \hat{F}_2(x) = \begin{bmatrix} 8x_1 - 4x_2 - 3 \\ 8x_2 - 4x_1 + 3 \end{bmatrix} = 0 \longrightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1/4 \\ -1/4 \end{bmatrix}$$

$$\nabla^2 \hat{F}_2(x) \underset{@\begin{bmatrix} 1/4 \\ -1/4 \end{bmatrix}}{=} \begin{bmatrix} 8 & -4 \\ -4 & 8 \end{bmatrix}, \quad eig(\nabla^2 \hat{F}_2(x)) = \begin{bmatrix} 4 \\ 12 \end{bmatrix} \rightarrow \text{pos. definite}$$

$$, v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$\rightarrow$ strong minima @ $\begin{bmatrix} 1/4 \\ -1/4 \end{bmatrix}$

ii. sketch of the contour plot: <u>max curvature towards v_2</u>



E9.3. $F(x) = x_1^2 + 2x_2^2$ line: $x = \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\text{initial guess } x_0} + \alpha \underbrace{\begin{bmatrix} -1 \\ -2 \end{bmatrix}}_{\text{direction } p_0}$

i. $\alpha_k = \dfrac{-g_k^T p_k}{p_k^T A p_k}$, $g = \nabla F(x) = \begin{bmatrix} 2x_1 \\ 4x_2 \end{bmatrix}$, $A = \nabla^2 F(x) = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$

$\rightarrow \alpha_1 = \dfrac{-[2\ \ 4]\begin{bmatrix}-1\\-2\end{bmatrix}}{[-1\ \ -2]\begin{bmatrix}2&0\\0&4\end{bmatrix}\begin{bmatrix}-1\\-2\end{bmatrix}} \approx 0.5556$

$- x_1 = \begin{bmatrix}1\\1\end{bmatrix} + 0.5556\begin{bmatrix}-1\\-2\end{bmatrix} = \begin{bmatrix}0.4444\\-0.1112\end{bmatrix}$

ii. Have to confirm that $g_1^T P_0 = 0 \Longrightarrow$

$g_1 = \nabla F(x)\big|_{x=x_1} = \begin{bmatrix}0.8888\\-0.4448\end{bmatrix} \rightarrow g_1^T P_0 \approx 0.0008$

very close to zero ?

(precision error)

$\rightarrow$ new direction is orthogonal to the previous direction.

E9.6. $F(x) = \frac{1}{2} x^T \begin{bmatrix}3&2\\2&0\end{bmatrix} x + [4\ \ 4]x$

i. $\nabla F(x) = Ax + d = \begin{bmatrix}3&2\\2&0\end{bmatrix} x + \begin{bmatrix}4\\4\end{bmatrix} = 0 \Rightarrow x = \begin{bmatrix}3&2\\2&0\end{bmatrix}^{-1} \begin{bmatrix}-4\\-4\end{bmatrix} = \begin{bmatrix}-2\\1\end{bmatrix}$

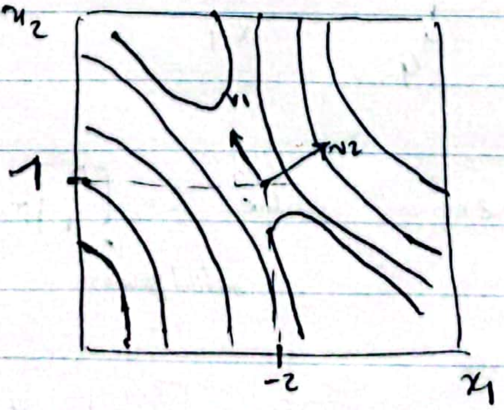
$\nabla^2 F(x) = A = \begin{bmatrix}3&2\\2&0\end{bmatrix} \rightarrow eig(\nabla^2 F(x)) = \begin{bmatrix}-1\\4\end{bmatrix} \rightarrow$ indefinite

$\hookrightarrow v_1 = \begin{bmatrix}-1/2\\1\end{bmatrix},\ v_2 = \begin{bmatrix}2\\1\end{bmatrix}$

$\rightarrow$ saddle point at $\begin{bmatrix}-2\\1\end{bmatrix}$

more curvature along v_2 because $|\lambda_2| > |\lambda_1|$

sketch of contour.



ii. Newton's Method:

$$x_{k+1} = x_k - A_k^{-1} g_k$$

$$A_k = A = \begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix}, \quad A^{-1} = \begin{bmatrix} 0 & 1/2 \\ 1/2 & -3/4 \end{bmatrix}$$

$$g_k = \nabla F(x)\Big|_{x=x_k} = \begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 14 \\ 8 \end{bmatrix}$$

$$\rightarrow x_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & 1/2 \\ 1/2 & -3/4 \end{bmatrix}\begin{bmatrix} 14 \\ 8 \end{bmatrix} = \begin{bmatrix} -5 \\ 0 \end{bmatrix}$$

iii. No. F(x), as proven in part (i), does not have a minimum.

It's only stationary point is a saddle point which occurs at $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

E9.8. Because it is difficult show sketches here, a complete explanation is provided in the Matlab live script.

E9.9. Refer to the MATLAB script for the same reason as above.