1. Newton's Method HW5 Chenrui Xu

$$F(X_{k+1}) = F(X_k + \Delta X_k) \approx F(X_k) + g_k^T \Delta X_k + \frac{1}{2} \Delta X_k^T A_k \Delta X_k$$

$$F(X) = (1 + (x_1 + x_2 - 5)^2)(1 + (3x_1 - 2x_2)^2)$$

$$\frac{\partial F}{\partial x_1} = 2(x_1+x_2-5)(1+(3x_1-2x_2)^2) + 2(3x_1-2x_2)\cdot 3(1+(x_1+x_2-5)^2)$$
$$= 2(x_1+x_2-5)(1+(3x_1-2x_2)^2) + 6(3x_1-2x_2)(1+(x_1+x_2-5)^2)$$

$$\frac{\partial F}{\partial x_2} = 2(x_1+x_2-5)(1+(3x_1-2x_2)^2) + (-2)\cdot 2(3x_1-2x_2)(1+(x_1+x_2-5)^2)$$ Let $x_1 = x_2 = 10$.

$$g_0 = \nabla F(x)\Big|_{x_1=x_2=10} = \begin{pmatrix} 30 \times 101 + 6 \times 10 \times 226 \\ 30 \times 101 - 40 \times 226 \end{pmatrix} = \begin{pmatrix} 16590 \\ -6010 \end{pmatrix}$$

$$\frac{\partial^2 F}{\partial x_1^2} = 2(1+(3x_1-2x_2)^2) + 2(x_1+x_2-5)(3x_1-2x_2)\cdot 2\cdot 3 + 6\cdot 3(1+(x_1+x_2-5)^2) + 6(3x_1-2x_2)\cdot 2(x_1+x_2-5)$$

When $x_1 = x_2 = 10$ $\frac{\partial^2 F}{\partial x_1^2} = 2\times 101 + 30 \times 6 \times 10 + 18 \times 226 + 60 \times 30 = 7870$

$$\frac{\partial^2 F}{\partial x_1 \partial x_2} = 2(1+(3x_1-2x_2)^2) + 2(x_1+x_2-5)\times 2\cdot(-2)\cdot(3x_1-2x_2) + (-12)(1+(x_1+x_2-5)^2) + 6(3x_1-2x_2)\cdot 2(x_1+x_2-5)$$
$$= 2\times 101 + 30 \times (-4) \times 10 + (-12)\times 226 + 60 \times 30 = -1910$$

$$\frac{\partial^2 F}{\partial x_2 \partial x_1} = 2(1+(3x_1-2x_2)^2) + 2(x_1+x_2-5)\cdot 2\cdot 3(3x_1-2x_2) - 12(1+(x_1+x_2-5)^2) - 4(3x_1-2x_2)\cdot 2(x_1+x_2-5)$$
$$= 2\times 101 + 30 \times 6 \times 10 - 12 \times 226 - 80 \times 15 = -1910$$

$$\frac{\partial^2 F}{\partial x_2^2} = 2(1+(3x_1-2x_2)^2) + 2\cdot(x_1+x_2-5)\cdot 2(-2)\cdot(3x_1-2x_2) + 8(1+(x_1+x_2-5)^2) - 4(3x_1-2x_2)\cdot 2\cdot(x_1+x_2-5)$$
$$= 2\times 101 + 30 \times (-4) \times 10 + 8 \times 226 - 80 \times 15 = -390$$

$$\nabla^2 F = \begin{pmatrix} 7870 & -1910 \\ -1910 & -390 \end{pmatrix} \Rightarrow A^{-1} = \begin{pmatrix} \frac{39}{671740} & -\frac{191}{671740} \\ -\frac{191}{671740} & -\frac{787}{671740} \end{pmatrix}$$

$$X_1 = X_0 - A^{-1} g_0 = \begin{pmatrix} 10 \\ 10 \end{pmatrix} - \begin{pmatrix} -\frac{89146}{33587} \\ \frac{78059}{33587} \end{pmatrix} = \begin{pmatrix} \frac{246126}{33587} \\ \frac{257811}{33587} \end{pmatrix} \approx \begin{pmatrix} 7.33 \\ 7.68 \end{pmatrix}$$

b-i) When $x_1 = x_2 = 2$. $\frac{\partial^2 F}{\partial x_1^2} = 10 + (-2)\times 6 \times 2 + 18 \times 2 + 24 \times (-1) = -2$ $\frac{\partial^2 F}{\partial x_1 \partial x_2} = 10 + (-2)\times(-4)\times 2 + (-12)\times 2 + 24(-1) = -22$

$\frac{\partial^2 F}{\partial x_2 \partial x_1} = 10 + 2\times(-1)\times 6 \times 2 - 12 \times 2 - 16 \times (-1) = -22$. $\frac{\partial^2 F}{\partial x_2^2} = 10 + (-2)\times(-4)\times 2 + 8 \times 2 - 16 \times (-1) = 58$

$$A = \begin{pmatrix} -2 & -22 \\ -22 & 58 \end{pmatrix} \Rightarrow A^{-1} = \begin{pmatrix} -\frac{29}{300} & -\frac{11}{300} \\ -\frac{11}{300} & \frac{1}{300} \end{pmatrix}$$ $X_1 = X_0 - A^{-1} g_0$

$$g_0 = \nabla F = \begin{pmatrix} (-2)\times 3 + 6\times 2 \\ (-2)\times 5 - 4 \times 2 \end{pmatrix} = \begin{pmatrix} 14 \\ -26 \end{pmatrix}$$ $X_1 = \begin{pmatrix} 2 \\ 2 \end{pmatrix} - \begin{pmatrix} -0.4 \\ -0.6 \end{pmatrix} = \begin{pmatrix} 2.4 \\ 2.6 \end{pmatrix}$

b-ii) $$\nabla F = \begin{pmatrix} 2(x_1+x_2-5)(1+(3x_1-2x_2)^2) + 6(3x_1-2x_2)(1+(x_1+x_2-5)^2) \\ 2(x_1+x_2-5)(1+(3x_1-2x_2)^2) - 4(3x_1-2x_2)(1+(x_1+x_2-5)^2) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{matrix} \cdots ① \\ \cdots ② \end{matrix}$$ ①-② ⇒

$\Rightarrow 10(3x_1-2x_2)(1+(x_1+x_2-5)^2) = 0$ $\begin{cases} 3x_1 - 2x_2 = 0 \\ 1 + (x_1+x_2-5)^2 = 0 \text{ not possible for real set} \end{cases}$

$3x_1 = 2x_2$. $x_1 = \frac{2}{3} x_2$ plug in ∇F.

$\Rightarrow x_1 = 2.\ x_2 = 3$

$F(x)|_{x_1=x_2=10} = 454685$. $F(x)|_{x_1=x_2=2} = 5$ $F(x)|_{x_1=2, x_2=3} = 1$ So when $x_1 = 2.\ x_2 = 3$. it has the minimum value. $(x_1, x_2) = (2,2)^T$ is closer to mini: $x_1 = x_2 = 10$ is largest one

if initial point is close to stationary point, it has more chance (easier, less step) to go min/max ①

2.(i) $F(x)=\frac{7}{2}x_1^2-6x_1x_2-x_2^2$. $\frac{\partial F}{\partial x_1}=7x_1-6x_2$ $\frac{\partial F}{\partial x_2}=-6x_1-2x_2$ $\nabla F=\begin{pmatrix}7x_1-6x_2\\-6x_1-2x_2\end{pmatrix}$

$x_1=x_2=1$ $\nabla F=\begin{pmatrix}1\\-8\end{pmatrix}$

$H=\begin{pmatrix}\frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1\partial x_2}\\ \frac{\partial^2 F}{\partial x_2\partial x_1} & \frac{\partial^2 F}{\partial x_2^2}\end{pmatrix}=\begin{pmatrix}7 & -6\\-6 & -2\end{pmatrix}$ $\quad \frac{(-1,1)\begin{pmatrix}1\\-8\end{pmatrix}}{\left\|\begin{pmatrix}-1\\1\end{pmatrix}\right\|}=\frac{-9}{\sqrt{2}}=-\frac{9\sqrt{2}}{2}$ $\quad \frac{(-1,1)\begin{pmatrix}7 & -6\\-6 & -2\end{pmatrix}\begin{pmatrix}-1\\1\end{pmatrix}}{\left\|\begin{pmatrix}-1\\1\end{pmatrix}\right\|^2}=\frac{17}{2}$

so first derivate along $(-1,1)$ is $-\frac{9\sqrt{2}}{2}$ second is $\frac{17}{2}$

(ii) $\nabla F(x)=\begin{pmatrix}\frac{\partial F}{\partial x_1}\\ \frac{\partial F}{\partial x_2}\end{pmatrix}=\begin{pmatrix}10x_1-6x_2+4\\-6x_1+10x_2+4\end{pmatrix}$ when $x_1=x_2=1$ $\nabla F(x)=\begin{pmatrix}8\\8\end{pmatrix}$

$\nabla^2 F(x)=\begin{pmatrix}\frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1\partial x_2}\\ \frac{\partial^2 F}{\partial x_2\partial x_1} & \frac{\partial^2 F}{\partial x_2^2}\end{pmatrix}=\begin{pmatrix}10 & -6\\-6 & 10\end{pmatrix}$ first derivate along $(-1,1)$ is $\frac{(-1,1)\begin{pmatrix}8\\8\end{pmatrix}}{\left\|\begin{pmatrix}-1\\1\end{pmatrix}\right\|}=0$

second is $\frac{(-1,1)\begin{pmatrix}10 & -6\\-6 & 10\end{pmatrix}\begin{pmatrix}-1\\1\end{pmatrix}}{\left\|\begin{pmatrix}-1\\1\end{pmatrix}\right\|^2}=16$

(iii) $\nabla F(x)=\begin{pmatrix}9x_1-2x_2+2\\-2x_1+6x_2-1\end{pmatrix}$ $x_1=x_2=1$ $\nabla F=\begin{pmatrix}9\\3\end{pmatrix}$

$\nabla^2 F(x)=\begin{pmatrix}9 & -2\\-2 & 6\end{pmatrix}$ first derivate along $(-1,1)$ $\frac{(-1,1)\begin{pmatrix}9\\3\end{pmatrix}}{\left\|\begin{pmatrix}-1\\1\end{pmatrix}\right\|}=-3\sqrt{2}$

second is $\frac{(-1,1)\begin{pmatrix}9 & -2\\-2 & 6\end{pmatrix}\begin{pmatrix}-1\\1\end{pmatrix}}{\left\|\begin{pmatrix}-1\\1\end{pmatrix}\right\|^2}=19/2$

(iv) $\nabla F=\begin{pmatrix}-7x_1-6x_2\\-6x_1+2x_2\end{pmatrix}$ $x_1=x_2=1$. $\nabla F=\begin{pmatrix}-13\\-4\end{pmatrix}$ $\nabla^2 F(x)=\begin{pmatrix}-7 & -6\\-6 & 2\end{pmatrix}$

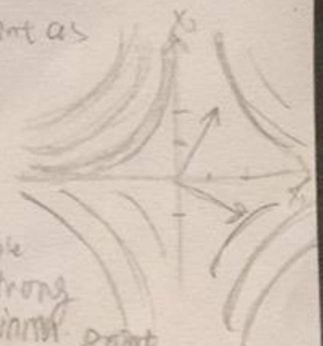
first derivate along $(-1,1)$ $\frac{(-1,1)\begin{pmatrix}-13\\-4\end{pmatrix}}{\left\|\begin{pmatrix}-1\\1\end{pmatrix}\right\|}=\frac{9}{\sqrt{2}}=\frac{9\sqrt{2}}{2}$

second: $\frac{(-1,1)\begin{pmatrix}-7 & -6\\-6 & 2\end{pmatrix}\begin{pmatrix}-1\\1\end{pmatrix}}{\left\|\begin{pmatrix}-1\\1\end{pmatrix}\right\|^2}=\frac{7}{2}$

3.① $\begin{vmatrix}7-\lambda & -6\\-6 & -2-\lambda\end{vmatrix}=0 \Rightarrow \begin{cases}\lambda_1=-5\\ \lambda_2=10\end{cases}$ $\lambda_1=-5$ $\begin{vmatrix}7+5 & -6\\-6 & -2+5\end{vmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix} \Rightarrow 2x_1=x_2 \Rightarrow \begin{pmatrix}1\\2\end{pmatrix}$

$\lambda_2=10$ $\begin{pmatrix}-3 & -6\\-6 & -12\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix} \Rightarrow x_1=-2x_2$ $\begin{pmatrix}2\\-1\end{pmatrix}$

Let $\nabla F=0$ $\begin{cases}7x_1-6x_2=0\\-6x_1-2x_2=0\end{cases} \Rightarrow \begin{cases}x_1=0\\x_2=0\end{cases}$ $(0,0)$ is single saddle point as $\lambda_1<0$ $\lambda_2>0$

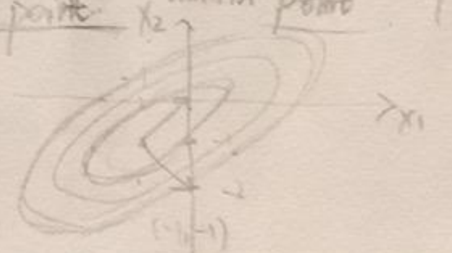


② $\begin{vmatrix}10-\lambda & -6\\-6 & 10-\lambda\end{vmatrix}=0 \Rightarrow \begin{cases}\lambda_1=4\\ \lambda_2=16\end{cases}$

$\lambda_1=4$ $\begin{pmatrix}6 & -6\\-6 & 6\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix} \Rightarrow x_1=x_2 \Rightarrow \begin{pmatrix}1\\1\end{pmatrix}$

$\lambda_2=16$ $\begin{pmatrix}-6 & -6\\-6 & -6\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix} \Rightarrow x_1=-x_2 \Rightarrow \begin{pmatrix}1\\-1\end{pmatrix}$ both $\lambda_1,\lambda_2>0$ so the point is single strong minima point

$\nabla F=0$ $\begin{pmatrix}10x_1-6x_2+4\\-6x_1+10x_2+4\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix} \Rightarrow x_1=x_2=-1$ $\begin{pmatrix}-1\\-1\end{pmatrix}$ ~~stationary point~~



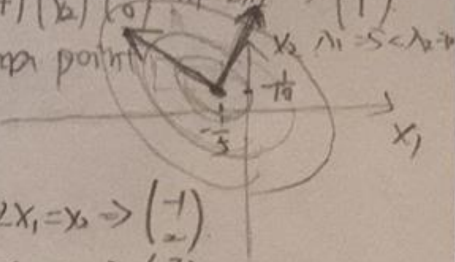
Chenrui Xu HW5

3. (iii) $\nabla F = \begin{pmatrix} 9x_1 - 2x_2 + 2 \\ -2x_1 + 6x_2 - 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow X = \begin{pmatrix} -\frac{1}{5} \\ \frac{1}{10} \end{pmatrix}$

$\begin{vmatrix} 9-\lambda & -2 \\ -2 & 6-\lambda \end{vmatrix} = 0 \quad (9-\lambda)(6-\lambda) - 4 = 0 \quad \Rightarrow \begin{cases} \lambda_1 = 5 \\ \lambda_2 = 10 \end{cases}$

$\lambda_1 = 5 \quad \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow 2x_1 = x_2 \quad \begin{pmatrix} 1 \\ 2 \end{pmatrix}$. $\lambda_2 = 10 \begin{pmatrix} -1 & -2 \\ -2 & -4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad x_1 = -2x_2 \Rightarrow \begin{pmatrix} -2 \\ 1 \end{pmatrix}$

As $\lambda_1, \lambda_2 > 0$ the point can be a single strong minima point



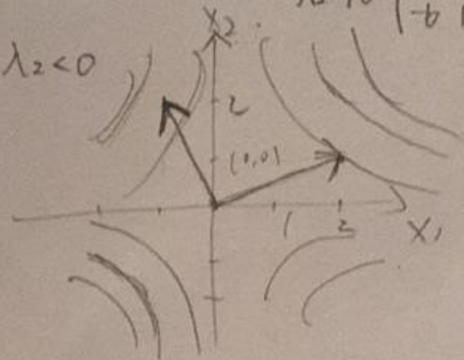
(iv). $\nabla F = \begin{pmatrix} -7x_1 - 6x_2 \\ -6x_1 + 2x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow X = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ stationary point

$\begin{vmatrix} -7-\lambda & -6 \\ -6 & 2-\lambda \end{vmatrix} = 0 \Rightarrow \begin{cases} \lambda_1 = 5 \\ \lambda_2 = -10 \end{cases}$

$\lambda_1 = 5 \quad \begin{pmatrix} -12 & -6 \\ -6 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad -2x_1 = x_2 \Rightarrow \begin{pmatrix} -1 \\ 2 \end{pmatrix}$

$\lambda_2 = -10 \begin{pmatrix} 3 & -6 \\ -6 & 12 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad x_1 = 2x_2 \Rightarrow \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

$\lambda_1 > 0 \quad \lambda_2 < 0$



it is a single saddle point