EAS6939 Homework #5

1. Minimize $f(\mathbf{x}) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ by means of Newton method. Use the initial estimate $\mathbf{x}^0 = [0,0]$.
2. Apply two steps of the steepest descent method to the minimization of $f(\mathbf{x}) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$. Use the initial estimate $\mathbf{x}^0 = [0,0]$.
3. Apply the conjugate gradient method to the minimization of $f(\mathbf{x}) = (2x_1 - x_2)^2 + (x_2 + 1)^2$ with $\mathbf{x}^0 = [\frac{5}{2}, 2]$.
4. Determine the first updated matrix $\mathbf{A}_1$ when applying DFP method to the minimization of $f(\mathbf{x}) = 3x_1^2 - 2x_1x_2 + x_2^2 + x_1$ with $\mathbf{x}^0 = [1,1]$.

1. $\nabla f(\underline{x}) = \begin{bmatrix} 1+4x_1+2x_2 \\ -1+2x_1+2x_2 \end{bmatrix} = \underline{c}(\underline{x})$; $\underline{H}(\underline{x}) = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}$ $\underline{H}^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}$

$\underline{x}^{(1)} = \underline{x}^{(0)} - \underline{H}^{-1} \cdot \underline{c} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix}$

$\nabla f(\underline{x}^{(1)}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\underline{H}$ is always P.D. $\therefore \underline{x}^* = \underline{x}^1 = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix}$ //

2. $\nabla f(\underline{x}) = \begin{bmatrix} 1+4x_1+2x_2 \\ -1+2x_1+2x_2 \end{bmatrix}$, $\underline{x}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\underline{d}^{(1)} = -\nabla f(\underline{x}^{(0)}) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

a) $\underline{x}^{(1)} = \underline{x}^{(0)} + \alpha_1 \underline{d}^{(1)} = \begin{bmatrix} -\alpha_1 \\ \alpha_1 \end{bmatrix}$

α_1 minimizes $\Phi(\alpha_1) = f(x^{(1)}) = -\alpha_1 - \alpha_1 + 2\alpha_1^2 - 2\alpha_1^2 + \alpha_1^2$

$\frac{\partial \Phi}{\partial \alpha_1} = 2\alpha_1 - 2 = 0$ $\therefore \alpha_1 = 1$ $\therefore x^{(1)} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. //

b) $x^{(2)} = x^{(1)} + \alpha_2 d^{(2)} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} +1 \\ +1 \end{bmatrix} = \begin{bmatrix} -1+\alpha_2 \\ 1+\alpha_2 \end{bmatrix}$

$\Phi(\alpha_2) = (-1+\alpha_2) - (1+\alpha_2) + 2(\alpha_2-1)^2 + 2(\alpha_2-1)(\alpha_2+1) + (\alpha_2+1)^2$

$\frac{\partial \Phi}{\partial \alpha_2} = 10\alpha_2 - 2 = 0$ $\therefore \alpha_2 = \frac{1}{5}$ $\underline{x}^{(2)} = \begin{bmatrix} -0.8 \\ 1.2 \end{bmatrix}$. //

3. First step, $d^{(1)} = -C^{(0)} = \begin{bmatrix} -12 \\ 0 \end{bmatrix}$ $\quad \nabla f(x) = \begin{bmatrix} 4(2x_1 - x_2) \\ -4x_1 + 4x_2 + 2 \end{bmatrix}$

$x^{(1)} = x^{(0)} + \alpha_1 d^{(1)}$

$\frac{\partial \Phi}{\partial \alpha_1} = 0 \Rightarrow x^{(1)} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ $\quad \nabla f(\underline{x}^{(1)}) = \begin{bmatrix} 0 \\ 6 \end{bmatrix} = C^{(1)}$

$d^{(2)} = -C^{(1)} + \beta_2 d^{(1)}$ $\quad \beta_2 = \left(\frac{\|C^{(1)}\|}{\|C^{(0)}\|}\right)^2 = \left(\frac{6}{12}\right)^2 = \frac{1}{4}.$

$= \begin{bmatrix} 0 \\ -6 \end{bmatrix} + \frac{1}{4}\begin{bmatrix} -12 \\ 0 \end{bmatrix} = \begin{bmatrix} -3 \\ -6 \end{bmatrix}$

$\underline{x}^{(2)} = \underline{x}^{(1)} + \alpha_2 d^{(2)} = \begin{bmatrix} 1 - 3\alpha_2 \\ 2 - 6\alpha_2 \end{bmatrix}$

$\Phi(\alpha_2) = (2(1-3\alpha_2) - (2-6\alpha_2))^2 + (2 - 6\alpha_2 + 1)^2$

$\frac{\partial \Phi}{\partial \alpha_2} = -12(3 - 6\alpha_2) = 0 \quad \therefore \alpha_2 = \frac{1}{2}$

$\Rightarrow \underline{x}^{(2)} = \begin{bmatrix} -\frac{1}{2} \\ -1 \end{bmatrix}$ $\quad \nabla f(x^{(2)}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. and $H = \begin{bmatrix} 8 & -4 \\ -4 & 4 \end{bmatrix}$ is P.D.

$\therefore \underline{x}^* = \underline{x}^{(2)}.$

4. $\nabla f(\underline{x}) = \begin{bmatrix} 6x_1 - 2x_2 + 1 \\ -2x_1 + 2x_2 \end{bmatrix}$ $\quad \nabla f(\underline{x}^0) = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$, $\underline{A}_0 = \underline{I}$.

$\underline{x}^{(1)} = \underline{x}^{(0)} + \alpha_1 \underline{d}^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \alpha_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -5 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 - 5\lambda_1 \\ 1 \end{bmatrix}$

$\Phi(\alpha_1) = 3(1-5\alpha_1)^2 - 2(1-5\alpha_1) + 1 + (1-5\alpha_1)$

$\frac{\partial \Phi}{\partial \alpha_1} = 6(1-5\alpha_1)(-5) + 10 - 5 = 0 \Rightarrow \alpha_1 = \frac{1}{6}.$

$\Rightarrow \underline{x}^{(1)} = \begin{bmatrix} \frac{1}{6} \\ 1 \end{bmatrix}$ $\quad \underline{s}_0 = \underline{x}^{(1)} - \underline{x}^{(0)} = \begin{bmatrix} -\frac{5}{6} \\ 1 \end{bmatrix}$

$\nabla f(x^{(1)}) = \begin{bmatrix} 0 \\ \frac{5}{3} \end{bmatrix}$ $\quad \underline{y}_0 = C^{(1)} - C^{(0)} = \begin{bmatrix} -5 \\ \frac{5}{3} \end{bmatrix}$

$\underline{A}_1 = \underline{A}_0 + \frac{s_0 s_0^T}{s_0^T y_0} - \frac{(A_0 y_0)(y_0^T A_0)}{y_0^T A_0 y_0} = \begin{bmatrix} \frac{4}{15} & \frac{3}{10} \\ \frac{3}{10} & \frac{9}{10} \end{bmatrix}$ //