Q No 9

(a) $f(x) = x^T A x + b^T x + c$ where x = vector R^N

A = $N \times N$ symetric i.e $A = A^T$

b = $N \times 1$ vector

c = Scalar.

$$f(x) = [x_1 \; x_2 \dots x_N] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & \\ \vdots & & \ddots & \\ a_{N1} & & & a_{NN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} + [b_1 \; b_2 \dots b_N] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} + c$$

$$= [x_1 a_{11} + x_2 a_{21} + \dots, x_N a_{N1} \quad x_1 a_{12} + x_2 a_{22} + \dots + x_N a_{N2} \cdots x_1 a_{1N} + x_2 a_{2N} + \dots + x_N a_{NN}] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

$$+ [b_1 x_1 + b_2 x_2 + b_N x_n + c]$$

$$= [x_1^2 a_{11} + x_1 x_2 a_{21} + \dots + x_1 x_N a_{N1} + x_1 x_2 a_{12} + x_2^2 a_{22} + \dots x_2 x_N a_{N2} + x_N x_1 a_{1N} + x_N x_2 a_{2N} + \dots + x_N^2 a_{NN}]$$

$$+ [b_1 x_1 + b_2 x_2 + b_N x_N + c]$$

Now the Hessian matrix is

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 x_N} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_N \partial x_1} & \frac{\partial^2 f}{\partial x_N \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_N^2} \end{bmatrix}$$

$$\frac{\partial f}{\partial x_1} = 2x_1a_{11} + x_2a_{21} + \dots + x_Na_{N1} + x_2a_{12} + x_3a_{13} + \dots x_Na_{1N} + b_1$$

$$\frac{\partial^2 f}{\partial x_1^2} = 2a_{11}$$

$$\frac{\partial f}{\partial x_2} = x_1a_{21} + 2x_2a_{22} + x_1a_{12} + x_Na_{N2} + x_Na_{2N}$$

$$\frac{\partial^2 f}{\partial x_2^2} = 2a_{22}$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = a_{21} + a_{12}$$ as A is symetric i.e. $a_{21} = a_{12}$ so

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = 2a_{12}$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = a_{21} + a_{12}$$ here also $a_{12} = a_{21}$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = 2a_{21}$$

From this pattern we can write Hessian matrics As

$$H = \begin{bmatrix} 2a_{11} & 2a_{12} & \dots & 2a_{1N} \\ 2a_{21} & 2a_{22} & \dots & 2a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 2a_{N1} & 2a_{N2} & \dots & 2a_{NN} \end{bmatrix} = 2A$$

(b)

$x^T x = 1$

$f(x) = -\cos(2\pi x^T x) + x^T x$

$$[x_1 \; x_2 \cdots x_N] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

$$\frac{\partial f}{\partial x_1} = \frac{\partial}{\partial x_1}\left[-\cos\left(2\pi(x_1^2 + x_2^2 + \dots x_N^2)\right) + x_1^2 + x_2^2 + \dots + x_N^2\right]$$

$$= \sin\left(2\pi(x_1^2 + x_2^2 + \dots + x_N^2)\right) 4\pi x_1 + 2x_1$$

$$\frac{\partial^2 f}{\partial x_1^2} = 4\pi \sin(2\pi x^T x) + 4\pi x_1 \cos(2\pi x^T x) \cdot 4\pi x_1 + 2$$

$$= 4\pi \sin(2\pi x^T x) + 16\pi^2 x_1^2 \cos(2\pi x^T x) + 2.$$

$$\frac{\partial f}{\partial x_2} = \sin\left(2\pi(x_1^2 + x_2^2 + \dots + x_N^2)\right) 4\pi x_2 + 2x_2$$

$$\frac{\partial^2 f}{\partial x_2^2} = 4\pi \sin(2\pi x^T x) + 16\pi^2 x_2^2 \cos(2\pi x^T x) + 2$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \cos(2\pi x^T x) \cdot 16\pi^2 x_1 x_2$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = \cos(2\pi x^T x)\, 16\pi^2 x_2 x_1$$

The pattern is recognized Hence Hessian Matrix can easily be constructed.

(Equations to long to fit in matrix on paper)

(c) $$f(x) = \sum_{m=1}^{M} \log(1+e^{-a_m^T x})$$ $$[a_1\ a_2 \dots a_N]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

$$= \sum_{m=1}^{M} \log(1+e^{-a_1x_1-a_2x_2\dots-a_Nx_N})$$

$$\frac{\partial f}{\partial x_1} = \sum_{m=1}^{M} \frac{-a_1 e^{-a_m^T x}}{(1+e^{-a_m^T x})\ln(10)}$$

$$\frac{\partial^2 f}{\partial x_1^2} = \sum_{m=1}^{M} \frac{(1+e^{-a_m^T x})(\ln 10)a_1^2 e^{-a_m^T x} + a_1 e^{-a_m^T x}(-\ln 10\, a_1 e^{-a_m^T x})}{((1+e^{-a_m^T x})\ln(10))^2}$$

$$= \sum_{m=1}^{M} \frac{\ln 10\, a_1^2 e^{-a_m^T x} + \cancel{\ln 10\, a_1^2 e^{-2a_m^T x}} - \cancel{\ln 10\, a_1^2 e^{-2a_m^T x}}}{(1+e^{-a_m^T x})^2(\ln(10))^2}$$

$$= \sum_{m=1}^{M} \frac{a_1^2 e^{-a_m^T x}}{\ln 10(1+e^{-a_m^T x})^2}$$

similarly

$$\frac{\partial f}{\partial x_2} = \sum_{m=1}^{M} \frac{-a_2 e^{-a_m^T x}}{(1+e^{-a_m^T x})\ln 10}$$

$$\frac{\partial^2 f}{\partial x_2^2} = \frac{a_2^2 e^{-a_m^T x}}{\ln 10 \cdot (1+e^{-a_m^T x})^2}$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \sum_{m=1}^{M} \frac{a_1 a_2 e^{-a_m^T x}}{\ln 10 \left(1 + e^{-a_m^T x}\right)^2}$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = \sum_{m=1}^{M} \frac{a_2 a_1 e^{-a_m^T x}}{\ln 10 \left(1 + e^{-a_m^T x}\right)^2}$$

Similarly all other derivatives can be calculated

Hence Hessian matrix can be constructed.