INFX 574 - Assignment II

1. a) the function is $f(x) = -x^2$
The maximum of this function is 0, which is produced when $x = 0$

b) $\nabla f = [\partial f / \partial x] = \left[\frac{\partial}{\partial x} -x^2\right] = [-2x]$

The ~~above~~ vector above is the gradient vector for this function. It is one-dimensional.

c) choosing x^0 as 1, $f(x^0) = -1$
$\nabla f(x^0) = [-2(1)] = [-2]$

d) learning rate $R = 0.1$
$x^1 = x^0 + R \cdot \nabla f(x^0) = 1 + 0.1 \cdot (-2)$
$= 0.8$

e) $f(x^1) = -0.64$, which is closer to the maximum value than $f(x^0)$.

4) $\cancel{x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}} \cancel{(x_1, x_2)}$ $\quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

$$f(x) = -x' A x$$

$$= \cancel{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}'} - (x_1, x_2) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= -(x_1, x_2) \begin{pmatrix} a_{11} x_1 + a_{12} x_2 \\ a_{21} x_1 + a_{22} x_2 \end{pmatrix}$$

$$= -\left(x_1 (a_{11} x_1 + a_{12} x_2) + x_2 (a_{21} x_1 + a_{22} x_2) \right)$$

$$= -(a_{11} x_1^2 + a_{12} x_2 x_1 + a_{21} x_1 x_2 + a_{22} x_2^2)$$

$$= -(a_{11} x_1^2 + a_{22} x_2^2 + \cancel{2} x_1 x_2 (a_{12} + a_{21}))$$

bivariate quadratic function:

$$ax^2 + by^2 + cxy + dx + ey + f$$

So our matrix equation describes a quadratic function where $a = -a_{11}$ $\quad b = -a_{22}$ $\quad c = -(a_{12} + a_{21})$

$d = e = f = 0$.

5) the problem has a single maximum (ie, $f(x) \le 0$ for all x)

iff A is nonnegative definite
or positive semidefinite

6) $\nabla f(x) = -2Ax = -2\begin{pmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{pmatrix}$

$$\frac{d}{dx_1} -\left(a_{11}x_1^2 + a_{22}x_2^2 + x_1 x_2 (a_{12}+a_{21})\right)$$

$$= -2a_{11}x_1 - x_2 a_{12} + x_2 a_{21}.$$

(if a is symmetric, $a_{12} = a_{21}$)

$$= -2a_{11}x_1 - 2a_{12}x_2$$

$$\boxed{= -2(a_{11}x_1 + a_{12}x_2)}$$

$$\frac{d}{dx_2} -\left(a_{11}x_1^2 + a_{22}x_2^2 + x_1 x_2 (a_{12}+a_{21})\right)$$

$$= -(2x_2 a_{22} + x_1 a_{12} + x_1 a_{21}).$$

(if a is symmetric, $a_{12} = a_{21}$)

$$\boxed{= -2(x_2 a_{22} + x_1 a_{21})}$$

THUS THIS IS TRUE FOR 2D CASE