Nicholas Wilson
nbwilson
931044347

Below are my answers to parts 1, 2, and 3 for homework 1, as well as to question 1 of part 5. Code is in gd.py and excercise_1.ipynb. This also includes my answer to the optional question in part 6 (why don't you normalize the column of all 1's in X?).

Question 1:

HW1 Q1 Nicholas Wilson

$$f(u,v) = 8u^2v^4 + 4v^3 + 6u$$
$$g(u,v,w) = x\log(u) + yuvw^3 + 13x^3$$
$$h(u,v) = \sum_{i=1}^{m} \frac{1}{2}\left(x^{(i)}u + y^{(i)}v\right)^2$$

1. $\frac{\partial}{\partial u} f(u,v) = 16uv^4 + 6$
2. $\frac{\partial}{\partial v} f(u,v) = 32u^2v^3 + 12v^2$
3. $\frac{\partial}{\partial u} g(u,v,w) = \frac{x}{u} + yvw^3$
4. $\frac{\partial}{\partial v} g(u,v,w) = yuw^3$
5. $\frac{\partial}{\partial w} g(u,v,w) = 3yuvw^2$
6. $\frac{\partial}{\partial u} h(u,v) = \sum_{i=1}^{m} \left(x^{(i)}u + y^{(i)}v\right) \cdot x^{(i)}$
7. $\frac{\partial}{\partial v} h(u,v) = \sum_{i=1}^{m} \left(x^{(i)}u + y^{(i)}v\right) \cdot y^{(i)}$

Question 2:

HW1 Q2

1. $\frac{\partial}{\partial u} f(-2,-2)$ is negative because as u increases from this point, f decreases.

2. $\frac{\partial}{\partial v} f(-2,-2)$ is negative, because as v increases from this point, f decreases.

3. $\frac{\partial}{\partial u} f(3,-3)$ is positive because as u increases from this point, f increases.

4. $\frac{\partial}{\partial v} f(3,-3)$ is negative because as v increases from this point, f decreases.

5. Function is minimized at $u=1, v=2$.

Question 3:

HW1 Q3

1. $$\begin{bmatrix} 3 & -1 \\ 2 & 5 \\ -2 & 2 \end{bmatrix} \cdot \begin{bmatrix} u & a \\ v & b \end{bmatrix} = \begin{bmatrix} 3u-v & 3a-b \\ 2u+5v & 2a+5b \\ -2u+2v & -2a+2b \end{bmatrix}$$

2. Yes, AB exists and is 2×4.

3. No, AB does not exist.

4. $y^T A$ $1\times 3 \cdot 3\times 2 \Rightarrow 1 \times 2$, so it's a row vector.

5. $Ax \Rightarrow 3\times 2 \cdot 2\times 1 \Rightarrow 3\times 1$, so a column vector.

6.
$$(Bx+y)^T A^T = 0$$
$$[(Bx+y)^T A^T]^T = 0^T$$
$$A(Bx+y) = 0^T$$
$$A^{-1}A(Bx+y) = A^{-1}0^T$$
$$Bx+y = 0^T$$
$$Bx = 0^T - y = -y$$
$$B^{-1}Bx = -B^{-1}y$$
$$x = -B^{-1}y$$

QED

$$\left.\begin{array}{l} A, A^T, B, B^T: n \times n \\ x, y: n \times 1 \\ x^T, y^T: 1 \times n \\ Bx: n \times 1 \\ 0: 1 \times n \\ 0^T: n \times 1 \end{array}\right\}$$

Dimensions of various matrices

Question 5:

HW1 Q5

1. $h_\theta(x_1) = \theta_0 + 2\theta_1$ $\theta = \begin{bmatrix}\theta_0 \\ \theta_1\end{bmatrix}$ $X = \begin{bmatrix}2 \\ -1 \\ 1\end{bmatrix}$ $y = \begin{bmatrix}5 \\ -1 \\ 3\end{bmatrix}$

$h_\theta(x_2) = \theta_0 - \theta_1$

$h_\theta(x_3) = \theta_0 + \theta_1$

$$J(\theta_0, \theta_1) = \frac{1}{2}\left[(\theta_0 + 2\theta_1 - 5)^2 + (\theta_0 - \theta_1 + 1)^2 + (\theta_0 + \theta_1 - 3)^2\right]$$

$$\frac{\partial J(\theta_0, \theta_1)}{\partial \theta_0} = \frac{1}{2}\left[2(\theta_0 + 2\theta_1 - 5) + 2(\theta_0 - \theta_1 + 1) + 2(\theta_0 + \theta_1 - 3)\right]$$

$$= 3\theta_0 + 2\theta_1 - 7$$

$$\frac{\partial J(\theta_0, \theta_1)}{\partial \theta_1} = 2(\theta_0 + 2\theta_1 - 5) - (\theta_0 - \theta_1 + 1) + (\theta_0 + \theta_1 - 3)$$

$$= 2\theta_0 + 6\theta_1 - 14$$

$3\theta_0 + 2\theta_1 - 7 = 0$

$2\theta_0 + 6\theta_1 - 14 = 0 \Rightarrow \theta_0 + 3\theta_1 - 7 = 0$

$\theta_0 = 7 - 3\theta_1$

$21 - 9\theta_1 + 2\theta_1 - 7 = 0$

$14 - 7\theta_1 = 0$

$\theta_1 = 2$ $\theta_0 = 1$

$\theta_0 = 1$ $\theta_1 = 2$

Question 6 optional:

You don't normalize the column of 1's in X because that column is only there to serve as a placeholder so you can matrix multiply with theta and get the offset of theta_zero. Changing those 1's to anything else would stop that. Furthermore, the mean is obviously one, and if you subtracted that from the column, you'd get a column of zeros which would remove the offset. Standard deviation is also zero, so you'd get an error if you tried to divide it by its standard deviation.