Part 1: Analytic assignment

1. Gradient vector(s) of $z = f(x,y) = ax + by + c$

$$\nabla z = (\nabla f_x(x,y), \nabla f_y(x,y))$$
$$= (a, b)$$

2. Gradient Vector of $z = f(x) = f(x_1, \cdots, x_N) = \sum_{i=1}^{N} a_i(x_i - b) + S$

$$= a_1 x_1 + a_2 x_2 + \dots + a_N x_N + d$$

$$\nabla z = (\nabla f_{x_1}(x_1, x_2, \cdots, x_n), \cdots, \nabla f_{x_n}(x_1, x_2, \cdots, x_n))$$

$$= (a_1, a_2, \cdots, a_N)$$

3. Partial Derivative: $z = f(x,y) = A(x - x_0)^2 + B(y - y_0)^2 + C$

$$f_x(x,y) = \frac{\partial f(x,y)}{\partial x} = 2A \cdot (x - x_0) \cdot 1$$
$$= 2Ax - 2Ax_0$$

$$f_y(x,y) = \frac{\partial f(x,y)}{\partial y} = 2B(y - y_0) \cdot 1$$
$$= 2By - 2By_0.$$

4. Matrices multiplication:

$$x = \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}, \quad y = [2, 5, 1], \quad A = \begin{bmatrix} 4 & 5 & 2 \\ 3 & 1 & 5 \\ 6 & 4 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 5 \\ 5 & 2 \\ 1 & 4 \end{bmatrix}$$

$x^T = [3\ 1\ 4]$ $[1 \times 3]$; $y^T = \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix}$ $[3 \times 1]$

$x \cdot x$ ~~not defined~~ $= \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}^T \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}$; $x \cdot y^T$ not defined

$= 9 + 1 + 16 = 26$ $[1]$

$$x \times y = \begin{vmatrix} i & j & k \\ 3 & 1 & 4 \\ 2 & 5 & 1 \end{vmatrix} = -19i - (-5)j + 13j$$
$$= -19i + 5j + 13j$$
$$= (-19\ 5\ 13)\ [3 \times 1] \text{ or } [1 \times 3]$$

$$Y \times X = \begin{vmatrix} i & j & k \\ 2 & 5 & 1 \\ 3 & 1 & 4 \end{vmatrix} = \det\begin{pmatrix} 5 & 1 \\ 1 & 4 \end{pmatrix} i - \det\begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} j + \det\begin{pmatrix} 2 & 5 \\ 3 & 1 \end{pmatrix} k$$

$$= 19i - 5j - 13k$$

$$= (19 \; -5 \; -13) \qquad [3 \times 1] \text{ or } [1 \times 3]$$

$$A \times X = \begin{bmatrix} 4 & 5 & 2 \\ 3 & 1 & 5 \\ 6 & 4 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}$$

$$= \begin{bmatrix} 4\times3+5\times1+2\times4 & 3\times3+1\times1+5\times4 & 6\times3+4\times1+3\times4 \end{bmatrix}^T$$

$$= \begin{bmatrix} 25 & 30 & 34 \end{bmatrix}^T$$

$$= \begin{bmatrix} 25 \\ 30 \\ 34 \end{bmatrix} \qquad [3 \times 1]$$

$$A \times B = \begin{bmatrix} 4 & 5 & 2 \\ 3 & 1 & 5 \\ 6 & 4 & 3 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 5 & 2 \\ 1 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 4\times3+5\times5+2\times1 & 4\times5+2\times5+2\times4 \\ 3\times3+1\times5+5\times1 & 3\times5+1\times2+5\times4 \\ 6\times3+4\times5+3\times1 & 6\times5+4\times2+3\times4 \end{bmatrix}$$

$$= \begin{bmatrix} 39 & 38 \\ 19 & 37 \\ 41 & 50 \end{bmatrix} \qquad [3 \times 2]$$

$$B.\text{reshape}(1,6) = [3 \;\; 5 \;\; 5 \;\; 2 \;\; 1 \;\; 4] \qquad [1 \times 6]$$

LLS (Single variable):

1. Loss surface $L(p) = \sum_{i=1}^{N} (\hat{y_i} - M(\hat{x_i}, m, b)^2)$

$$= \sum_{i=1}^{N} (\hat{y_i} - m\hat{x_i} - b)^2$$

$$L(m,b) = \sum_{i=1}^{N} (\hat{y_i} - m\hat{x_i} - b)^2$$

By ②: $m\bar{x} + b = \bar{y}$:

$$b = \bar{y} - m\bar{x}$$

$$= \bar{y} - \frac{cov(X,Y)}{var(X,Y)} \bar{X}.$$

$$\frac{\partial L(m,b)}{\partial m} = \sum_{i=1}^{N} (2x_i(mx_i + b - y_i)) = 0 \quad ①$$

$$\frac{\partial L(m,b)}{\partial b} = \sum_{i=1}^{N} (2(mx_i + b - y_i)) = 0 \quad ②$$

so we get: $2\sum_{i=1}^{N} (x_i(mx_i + b - y_i)) = 2\sum_{i=1}^{N} (mx_i + b - y_i) = 0$

$$\Rightarrow m\sum_{i=1}^{N} x_i^2 + b\sum_{i=1}^{N} x_i - \sum_{i=1}^{N} x_i y_i = m\sum_{i=1}^{N} x_i + N \cdot b - \sum_{i=1}^{N} y_i$$

$$\Rightarrow m\sum_{i=1}^{N} x_i^2 - \sum_{i=1}^{N} x_i y_i + N\bar{x} \cdot b = m \cdot N\bar{x} + N \cdot b - N\bar{y}$$

$$\Rightarrow m\sum_{i=1}^{N} x_i^2 - \sum_{i=1}^{N} x_i y_i + \bar{x}(N\bar{y} - mN\bar{x}) = 0$$

$$\Rightarrow m\left(\sum_{i=1}^{N} x_i^2 - N\bar{x}^2\right) = \sum_{i=1}^{N} x_i y_i - N\bar{x}\bar{y}$$

Note: $cov(X,Y) = \frac{1}{N}\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{N}\sum_{i=1}^{n} (x_i y_i - \bar{x} y_i - \bar{y} x_i + \bar{x}\bar{y})$

$$= \frac{1}{N}\sum_{i=1}^{n} x_i y_i - 2\bar{x}\bar{y} \cdot N + N\bar{x}\bar{y}$$

$$= \frac{1}{N}\sum_{i=1}^{n} x_i y_i - \bar{x} \cdot \bar{y}$$

$$var(X,Y) = \frac{1}{N}\sum_{i=1}^{n} (x_i - \bar{x}) = \frac{1}{N}\left(\sum_{i=1}^{n} (x_i^2 + \bar{x}^2 - 2\bar{x}x_i)\right)$$

$$= \frac{1}{N}\left(\sum_{i=1}^{n} x_i^2 + N \cdot \bar{x}^2 - 2\bar{x}^2 \cdot N\right)$$

$$= \frac{1}{N}\sum_{i=1}^{n} x_i^2 - \bar{x}^2$$

so: $m = \frac{n\, cov(X,Y)}{n\, var(X,Y)} = \frac{cov(X,Y)}{var(X,Y)}$

2. Multivariate LLS :

$$\mathcal{L}(\beta^*) = \sum_{i=1}^{N} (\hat{y_i} - M(\hat{x_i}, \beta_0^*, \beta_i^*))^2$$, where β, y are vectors.

$$= (X\beta^* - y)^T (X\beta^* - y)$$

$$= \beta^{*T} X^T X \beta^* - y^T X \beta^* - \beta^{*T} X^T y - y^T y$$

since β, y are vectors, $y^T X \beta^*$, $\beta^{*T} X^T y$ are scalars.
& $(y^T X \beta^*)^T = \beta^{*T} X^T y$, so they are equal.

$$\frac{\partial \mathcal{L}(\beta)}{\partial \beta} = 2X^T X \beta^* - 2X^T y = 0$$

$$\Rightarrow \quad X^T X \beta^* = X^T y$$

& since $X^T X$ is square matrix it has inverse if $X^T X$ is invertible, but now we assume it is.

$$\Rightarrow (X^T X)^{-1} (X^T X) \beta^* = (X^T X)^{-1} X^T y$$

$$\underline{\beta^* = (X^T X)^{-1} X^T y}.$$