1. $\nabla f(x,y) = \langle a, b \rangle$

2. $\nabla f(x_1, x_2, x_3, \cdots x_N) = \langle a_1, a_2, a_3, a_4 \cdots, a_N \rangle$

3. $f_x(x,y) = 2x - 2Ax_0$

$f_y(x,y) = 2y - 2By_0$

4. $x^T = \begin{pmatrix} 3 & 1 & 4 \end{pmatrix}$ $[1 \times 3]$

$y^T = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}$ $[3 \times 1]$

$B^T = \begin{pmatrix} 3 & 5 & 1 \\ 5 & 2 & 4 \end{pmatrix}$ $[2 \times 3]$

$x \cdot x = 26$

$x \cdot y^T = 15$

$x \times y = \begin{pmatrix} -19 & 5 & 13 \end{pmatrix}$

$y \times x = \begin{pmatrix} 19 & -5 & -13 \end{pmatrix}$

$A \times X = \begin{pmatrix} 25 \\ 30 \\ 34 \end{pmatrix}$ $[3 \times 1]$

$A \times B = \begin{pmatrix} 39 & 38 \\ 19 & 37 \\ 41 & 50 \end{pmatrix}$ $[3 \times 2]$

B.reshape$(1,6) = \begin{pmatrix} 3 & 5 & 1 & 5 & 2 & 4 \end{pmatrix}$. $[1 \times 6]$

5. $L(p_1) = L(m,b) = \sum_{j=1}^{N} (\hat{y}_i - mx_i - b)^2 = \sum_{j=1}^{N} \hat{y}_i^2 + \sum_{j=1}^{N} m^2 x_i^2 + \sum_{j=1}^{N} b^2 - 2\sum_{j=1}^{N} \hat{y}_i m x_i - 2\sum_{j=1}^{N} y_i b$

$+2\sum_{j=1}^{N} m x_i b$

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

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

$$\Rightarrow m = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2} = \frac{cov(x,y)}{Var(x)}$$

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

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

$$= \bar{y} - \frac{cov(x,y)}{Var(x)} \cdot \bar{x}$$

Bonus: let $h_\theta(x) = \sum_{i=1}^{n} \theta_i x_i$

let $J(\theta) = \frac{1}{2}(\hat{y}_i - y_i)^2 = \frac{1}{2}\sum_{i=1}^{n}(h_\theta(x_i) - y_i)^2$.

$$= \frac{1}{2}(X\theta - Y)^2$$

$$= \frac{1}{2}(X\theta - Y)^T(X\theta - Y)$$

$$= \frac{1}{2}(X^T\theta^T - Y^T)(X\theta - Y)$$

$$= \frac{1}{2}(X^T\theta^T X\theta - X^T\theta^T Y - Y^T X\theta + Y^T Y$$

$$\therefore \partial J(\theta)/\partial\theta = \frac{1}{2}(2X^T X\theta - X^T Y - X^T Y)$$

$$= X^T X\theta - X^T Y$$

Let $\partial J(\theta)/\partial\theta = 0$

$$\therefore \theta = (X^T X)^{-1} X^T Y$$