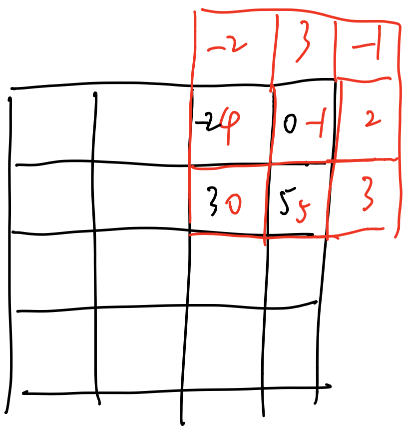
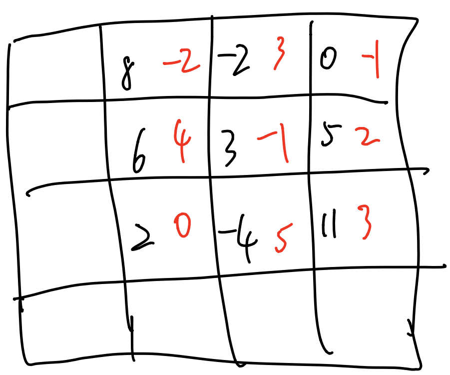
Q_2:

A:



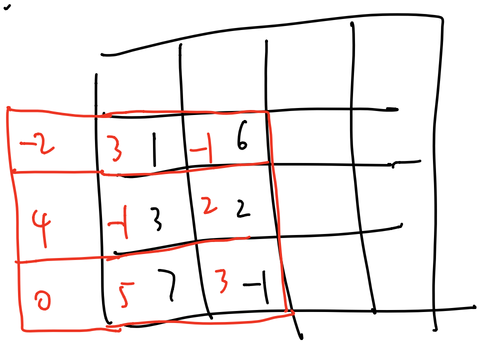
$$A = 4\cdot(-2) + 0\cdot(-1) + 0\cdot 3 + 5\cdot 5$$
$$+ 2\cdot 0 + (-1)\cdot 0 + 3\cdot 0 + (-2)\cdot(-2) + 3\cdot 5$$
$$= 36$$

B:



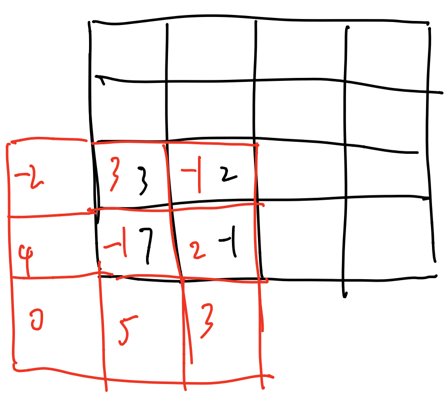
$$B = (-2)\cdot 8 + (-2)\cdot 3 + 0\cdot(-1)$$
$$+ 6\cdot 4 + 3\cdot(-1) + 5\cdot 2$$
$$+ 2\cdot 0 + (-4)\cdot 5 + 11\cdot 3 \quad = 22$$

C:



$$C = 3\cdot 1 + (-1)\cdot 6 + (-1)\cdot 3 + 2\cdot 2 + 5\cdot 7 + 3\cdot(-1)$$
$$+ (-2)\cdot 1 + 4\cdot 3 + 0\cdot 7$$
$$= 40$$

D:



$$D = 3\cdot 3 + (-1)\cdot 2 + (-1)\cdot 7 + 2\cdot(-1)$$
$$+ (-2)\cdot 3 + 4\cdot 7 + 0\cdot 7 + 5\cdot 7 + 3\cdot(-1)$$
$$= 52$$

Bonus Question:

(a) proof Bp1: $\delta_j^L = \frac{\partial C}{\partial z_j^L}$

$$\delta_j^L = \sum_k \frac{\partial c}{\partial a_k^2} \frac{\partial a_k^2}{\partial z_j^L}$$

$$\delta_j^L = \frac{\partial c}{\partial a_j^L} \frac{\partial a_j^L}{\partial z_j^L}$$

$$\delta_j^L = \frac{\partial c}{\partial a_j^L} \sigma'(z_j^L)$$

proof Bp2:

$$\delta_j^L = \frac{\partial c}{\partial z_j^L} = \sum_k \frac{\partial c}{\partial z_k^{l+1}} \frac{\partial z_k^{l+1}}{\partial z_j^L} = \sum_k \frac{\partial z_k^{L+1}}{\partial z_j^L} \delta_k^{L+1}$$

$$z_k^{L+1} = \sum_j W_{kj}^{L+1} a_j^L + b_k^{L+1} = \sum_j W_{kj}^{L+1} \sigma(z_j^L) + b_k^{L+1}$$

$$\frac{\partial z_k^{L+1}}{\partial z_j^L} = W_{kj}^{L+1} \sigma'(z_j^L)$$

$$\delta_j^L = \sum_k W_{kj}^{L+1} \delta_k^{L+1} \sigma'(z_j^L)$$

proof Bp3: $\frac{\partial c}{\partial b_j^L} = \sum_k \left(\frac{\partial c}{\partial z_k^L} \frac{\partial z_k^L}{\partial b_j^L} \right) = \frac{\partial c}{\partial z_j^L} \frac{\partial z_j^L}{\partial b_j^L}$

$$= \delta_j^L \frac{\partial (\Sigma_k (w_{jk} a_k^{L-1} + b_j^L))}{\partial b_j}$$

$$= \delta_j^L$$

proof BP4 : $\frac{\partial C}{\partial w_{jk}^L} = \sum_m \frac{\partial C}{\partial z_m^L} \frac{\partial z_m^L}{\partial w_{jk}^L} = \frac{\partial C}{\partial z_j^L} \frac{\partial z_j^L}{\partial w_{jk}}$

$$= \delta_j^L \frac{\partial}{\partial w_{jk}} \left(\sum_k w_{jk}^L a_k^{L-1} + b_j^L \right)$$

$$= a_k^{L-1} \delta_j^L$$

(b) $\frac{\partial C}{\partial b^2} = \begin{bmatrix} \frac{\partial C}{\partial b_1^L} \\ \vdots \\ \frac{\partial C}{\partial b_{d_1}^L} \end{bmatrix}$

From BP3, we know that $\frac{\partial C}{\partial b_j^L} = \delta_j^L$

$$\frac{\partial C}{\partial b^2} = \begin{bmatrix} \delta_1^L \\ \vdots \\ \delta_d^L \end{bmatrix} = \delta^2$$

$$\frac{\partial C}{\partial w^L} = \begin{bmatrix} \frac{\partial C}{\partial w_{11}^L} & \cdots & \frac{\partial C}{\partial w_{1 d_{L-1}}^L} \\ \vdots & \ddots & \vdots \\ \frac{\partial C}{\partial w_{d_L 1}^2} & \cdots & \frac{\partial C}{\partial w_{d_L d_{L-1}}^L} \end{bmatrix}$$

from BP4 we know $\frac{\partial L}{\partial w_{jk}^L} = a_k^{L-1} \delta_j^L$

$$= \begin{bmatrix} a_1^{L-1} \delta_1^L & \cdots & a_{d_{L-1}}^{L-1} \delta_1^L \\ \vdots & \ddots & \vdots \\ a_1^{L-1} \delta_{d_1}^L & \cdots & a_{d_{L-1}}^{L-1} \delta_{d_L}^L \end{bmatrix} = \delta^L \cdot (a^{L-1})^T$$

(c) Two assumptions: $C = \frac{1}{2n} \sum_x \| y(x) - a^L(x) \|^2$

$$C = \frac{1}{2} \| y - a^L \|^2 = \frac{1}{2} \sum_j (y_j - a_j^L)^2$$

$$\delta^2 = \nabla_{\hat{y}_1} L \odot \sigma'(z^L) = \frac{\partial}{\partial \hat{y}_1} \frac{1}{2} \| y_1 - \hat{y}_1 \|_2^2 \odot \sigma'(z^L)$$

$$\delta^2 = (\hat{y}_1 - y_1) \odot \sigma'(z^3)$$

$$\delta^2 = ((w^3)^T \cdot \delta^3) \odot \sigma'(z^2)$$

$$\delta^1 ((w^2)^T \delta^2) \odot \sigma'(z^1)$$

Bp (3-4) gives

$$\frac{\partial L}{\partial b^1} = \delta^1$$

$$\frac{\partial L}{\partial w^1} = \delta^1 (a^0)^T = \delta^1 \cdot x_1^T$$