$$f_W(X) = ABX$$
$$\underset{1 \times n}{} \quad \underset{L \times k}{} \quad \underset{k \times d \times d \times n}{}$$

$$= [A_1 A_2 \cdots A_k] \begin{bmatrix} B_{11} & . & B_{1d} \\ B_{21} & & \\ & \ddots & \\ B_{k1} & & B_{kd} \end{bmatrix} \begin{bmatrix} x^{(1)} & x^{(2)} & \cdots & x^{(n)} \\ & & & \end{bmatrix}$$

$$= \underbrace{[A_1 B_{11} + \cdots, \; A_1 B_{12} + \cdots, \cdots, A_1 B_{1d} + \cdots]}_{\in 1 \times d} \begin{bmatrix} x^{(1)} & \cdots & x^{(n)} \\ & & \end{bmatrix} \in d \times n \qquad x^{(b)}.$$

$$= \left[A_1 B_{11} x_1^{(1)} + A_1 B_{12} x_2^{(1)} \cdots + A_1 B_{1d} x_d^{(1)}, \quad \cdots \cdots, \qquad \right]$$

$$\frac{d f_W(x)}{d A_1} = \begin{bmatrix} B_{11} x_1^{(1)} + B_{12} x_2^{(1)} + \cdots B_{1d} x_d^{(1)} \\ \\ B_{11} x_1^{(2)} + B_{12} x_2^{(2)} + \cdots B_{1d} x_d^{(2)} \\ \\ \vdots \end{bmatrix} \in n \times 1$$

$$= \underset{n \times d}{X^T} \cdot \underset{d \times 1}{B_1^T}$$

$$\Downarrow$$

$$\text{stack} \quad A_1 \cdots A_k \implies \frac{d f_W(x)}{d A} = X^T B^T \in n \times k.$$

$$\implies A^{t+1} = A^t - \eta \, \nabla_A \mathcal{L}(w)$$

$$\mathcal{L}(w) = \frac{1}{2} (y - f_W(x))^2$$

$$\nabla_{A} \mathcal{L}(w) = - \underbrace{(y - f_W(x))}_{1 \times n} \cdot \underbrace{\nabla_A f_W(x)}_{}$$

$$\boxed{= (y - f_W(x)) \cdot X^T B^T \quad \text{(a)}}$$