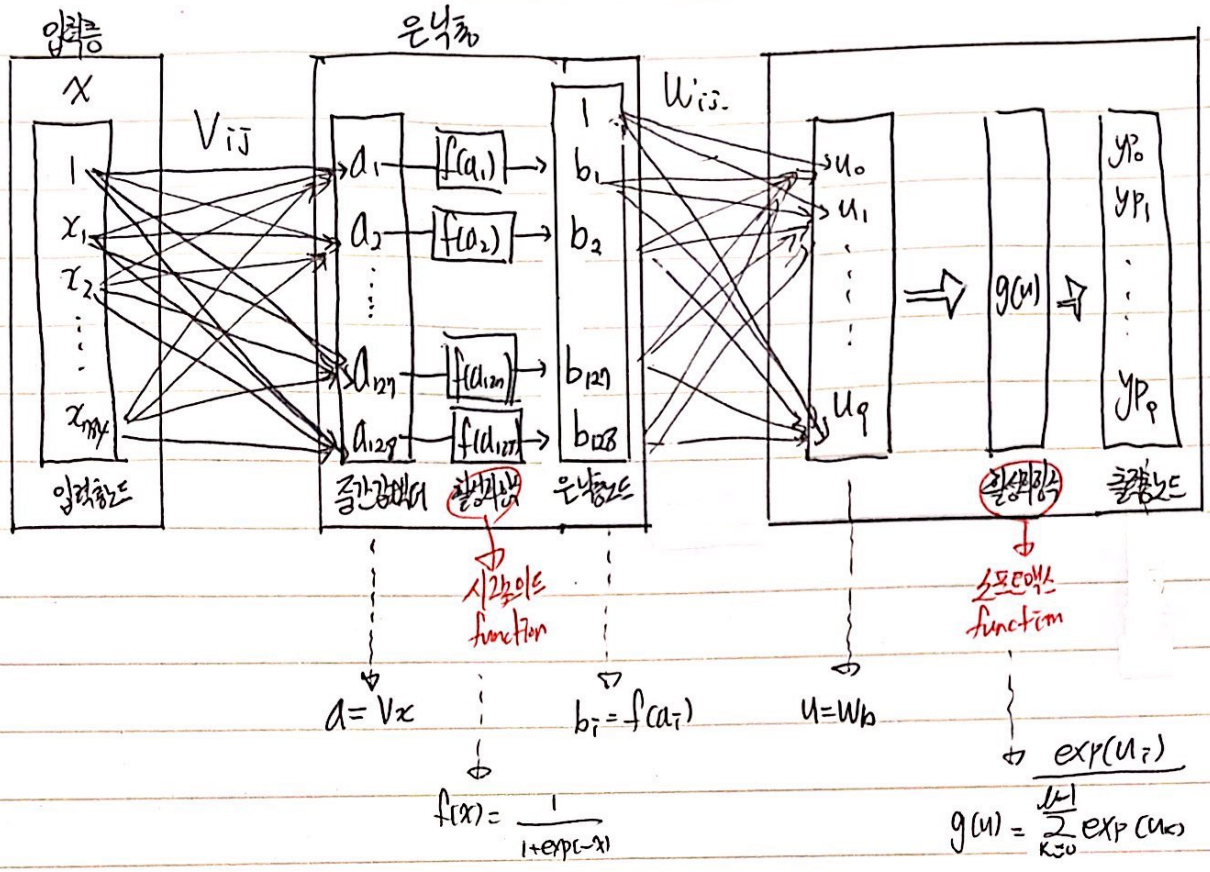
모델 구성과 예측함수



순전파 (forward propagation)

손실함수는 이전에 사용한 크로스엔트로피를 사용

$$L(w) = -\sum_{i=0} yt_i \log(yp_i)$$

① V_{12}에 대해 미분한다 (체인룰)

$$\frac{\partial L}{\partial V_{12}} = \frac{\partial a_1}{\partial V_{12}} \cdot \frac{\partial L}{\partial a_1}$$

$\frac{\partial a_1}{\partial V_{12}} = x_2$

$$\frac{\partial L}{\partial a_1} = \frac{db_1}{da_1} \cdot \frac{\partial L}{\partial b_1}$$

$\frac{db_1}{da_1} = f'(a_1)$

$\sum_{k=0}^{N-1} \frac{\partial u_k}{\partial b_1} \cdot \frac{\partial L}{\partial u_k}$

w_{21} $\quad$ yd_k

if $u_2 = w_{20}b_0 + w_{21}b_1 + w_{22}b_2 + \cdots$ $\quad \frac{\partial u_2}{\partial b_1} = w_{21}$

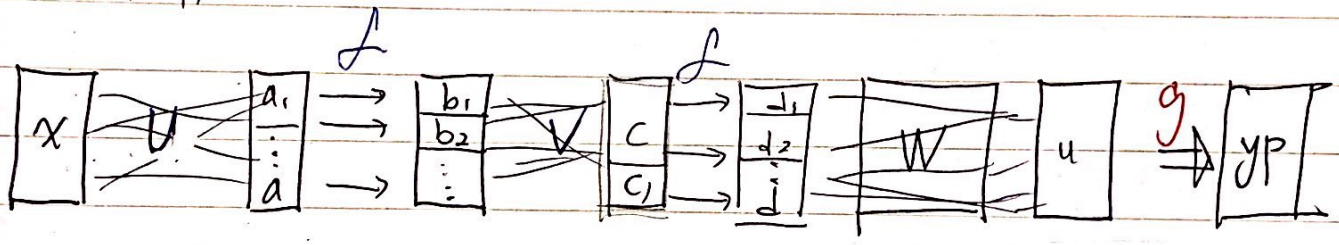
b_1와 곱해진 w_{k1}

따라서 $\sum_{k=0}^{N-1} w_{k1} \cdot yd_k$ 로 표현할 수 있다

$$\frac{\partial L}{\partial a_1} = f'(a) \times \sum_{k=0}^{N-1} w_{k1} \cdot yd_k$$

$$\frac{\partial L}{\partial v_{ij}} = x_j \cdot \frac{\partial L}{\partial a_i}$$

만약은닉층이 2개라면



u_{12}에 대해서보면

$$\frac{\partial L}{\partial u_{12}} = \frac{\partial L}{\partial b_1} \cdot \frac{db_1}{da_1} \cdot \frac{\partial a_1}{\partial u_{12}}$$

$f'(a_1)$ $\quad$ x_2

$\frac{\partial L}{\partial b_1}$ → $\sum_{k=1}^{H} \frac{\partial c_k}{\partial b_1} \frac{\partial L}{\partial c_k}$

v_{k1}

$\frac{dd}{dc} \cdot \frac{\partial L}{\partial d_q}$

$f'(c)$

$\sum_{p=1}^{H} \frac{\partial u_p}{\partial d_1} \cdot \frac{\partial L}{\partial u_p}$

w_{p1} $\quad$ yd_p

경사감법적용

은닉층 1개

$$W_{ij}^{(k+1)} = W_{ij}^{(k)} - \frac{\alpha}{n} \sum_{m=0}^{n-1} b_j \, yd_i$$

$$V_{ij}^{(k+1)} = V_{ij}^{(k)} - \frac{\alpha}{n} \sum_{m=0}^{n-1} x_j \, bd_i$$

은닉층 2개

$$U_{ij} = U_{ij} - \frac{\alpha}{n} \sum_{m=0}^{n-1} x_j \, bd_i$$

$$V_{ij} = V_{ij} - \frac{\alpha}{n} \sum_{m=0}^{n-1} b_j \, dd_i$$

$$W_{ij} = W_{ij} - \frac{\alpha}{n} \sum_{m=0}^{n-1} d_j \, yd_i$$