Contents

First we only consider **one play**

$$G(P_n, w^1) = \sum_{i=1}^{S} \tilde{\theta_i} \tanh(\theta_i P_n + \theta_{i0}) + \tilde{\theta_0} \tag{1}$$

Where $P_n = [p_1, p_2, p_3, ..., p_n]$, $G(P_n, w^1) = [y_1, y_2, y_3, ..., y_n]$
$\forall i \in [1, ..., S]$, $\theta_i P_n = (\theta_i p_1, \theta_i p_2, \theta_i p_3, ..., \theta_i p_n)$,
So $\tanh(\theta_i P_n + \theta_{i0}) = [\tanh(\theta_i p_1 + \theta_{i0}), \tanh(\theta_i p_2 + \theta_{i0}), \tanh(\theta_i p_3 + \theta_{i0}), ..., \tanh(\theta_i p_n + \theta_{i0})]$
So $\sum_{i=1}^{S} \tilde{\theta_i} \tanh(\theta_i P_n + \theta_{i0}) + \tilde{\theta_0} = [\sum_{i=1}^{S} \tilde{\theta_i} \tanh(\theta_i p_1 + \theta_{i0}) + \tilde{\theta_0}, \sum_{i=1}^{S} \tilde{\theta_i} \tanh(\theta_i p_2 + \theta_{i0}) + \tilde{\theta_0}, ..., \sum_{i=1}^{S} \tilde{\theta_i} \tanh(\theta_i p_n + \theta_{i0}) + \tilde{\theta_0}] = [y_1, y_2, ..., y_n]$.
Take y_j, where $j \in [1, ..., n]$ for example

$$y_j = \sum_{i=1}^{S} \tilde{\theta_i} \tanh(\theta_i p_j + \theta_{i0}) + \theta_{i0} \tag{2}$$

Calculate derivation for y_j

$$\frac{\partial y_j}{\partial p_j} = \sum_{i=1}^{S} \tilde{\theta_i} \frac{\partial \tanh(\theta_i p_j + \theta_{i0})}{\partial p_j} \tag{3}$$

$$\tag{4}$$

Now let's consider the mapping between p_j and x_j.

$$p_j = \Phi(w_1 x_j - p_{j-1}) + p_{j-1} \tag{5}$$

and

$$\Phi(x) = \begin{cases} x, x > 0 \\ 0, -1 < x < 0 \\ x - 1, x < -1 \end{cases}$$

(6)

Using Chain Rule, we obtain

$$\begin{aligned}
\frac{\partial y_j}{\partial x_j} &= \frac{\partial y_j}{\partial p_j}\frac{\partial p_j}{\partial x_j} && (7)\\
&= \sum_{i=1}^{S} \tilde{\theta}_i \frac{\partial \tanh(\theta_i p_j + \theta_{i0})}{\partial p_j}\frac{\partial \Phi(w^1 x_j - p_{j-1})}{\partial x_j} && (8)
\end{aligned}$$

To consider **multiple plays** case, we reformulate the derivation as following:

$$\begin{aligned}
\frac{\partial y_j^1}{\partial x_j} &= \frac{\partial y_j^1}{\partial p_j^1}\frac{\partial {p_j}^1}{\partial x_j} && (9)\\
&= \sum_{i=1}^{S} \tilde{\theta}_i^1 \frac{\partial \tanh(\theta_i^1 p_j^1 + \theta_{i0})}{\partial p_j^1}\frac{\partial \Phi(w^1 x_j - p_{j-1}^1)}{\partial x_j} && (10)\\
& && (11)
\end{aligned}$$

Now from the architecture, we know that if we have P plays,

$$F = \frac{1}{P}\sum_{i=1}^{P} G^i \qquad (12)$$

Where $F = [f_1, f_2, ..., f_n]$, and

$$f_j = \frac{1}{P}\sum_{i=1}^{P} y_j^i \qquad (13)$$

our derivation is:

$$\begin{aligned}
\frac{\partial f_j}{\partial x_j} &= \frac{1}{P}\sum_{k=1}^{P}\frac{\partial y_j^k}{\partial x_j} && (14)\\
&= \frac{1}{P}\sum_{k=1}^{P}\frac{\partial y_j^k}{\partial p_j^k}\frac{\partial p_j^k}{\partial x_j} && (15)\\
&= \frac{1}{P}\sum_{k=1}^{P}\sum_{i=1}^{S}\tilde{\theta}_i^k \frac{\partial \tanh(\theta_i^k p_j^k + \theta_{i0})}{\partial p_j^k}\frac{\partial \Phi(w^k x_j - p_{j-1}^k)}{\partial x_j} && (16)\\
& && (17)
\end{aligned}$$