Symmetric neural networks

YKY

May 16, 2019

1 General case for $y = Ax$

$$\boxed{\text{original}} \quad y_j = \sum_i a_{ij} x_i. \tag{1}$$

Equivariance implies:

$$\begin{aligned} \boxed{\text{swapped}} \quad y_j \cdot \sigma(x_j\ x_k) &= y_k \quad \boxed{\text{original}} \\ \sum_i \square + a_{kj} x_j + a_{jj} x_k &= \sum_i \square + a_{jk} x_j + a_{kk} x_k \end{aligned} \tag{2}$$

where $\square$ denotes "the rest of the elements".

This gives us a set of equations:

$$a_{kj} x_j + a_{jj} x_k = a_{jk} x_j + a_{kk} x_k \qquad \forall x_j, x_k \tag{3}$$

for $(j\ k) \in \mathfrak{S}_n$.

By setting x_j and x_k to zero, this yields:

$$\begin{aligned} a_{kj} &= a_{jk} \qquad \forall j, k \\ a_{jj} &= a_{kk} \qquad \forall j, k. \end{aligned} \tag{4}$$

In other words, the matrix A is of the form:

$$A = \alpha I + \beta 1 1^T. \tag{5}$$

2 Case for $y_k = A_k x \cdot x$

The general form of a "quadratic" vector function is:

$$y = (Ax) \cdot x + Bx + C. \tag{6}$$

We just focus on the quadratic term $(Ax) \cdot x$:

$$\boxed{\text{original}} \quad y_k = \sum_j \left[\sum_i a_{ij}^k x_i \right] x_j. \tag{7}$$

Equivariance implies:

$$\begin{aligned} \boxed{\text{swapped}} \quad y_k \cdot \sigma(x_k \; x_h) &= y_h \quad \boxed{\text{original}} \\ \sum_{j \neq h,k} \sum_{i \neq h,k} a_{ij}^k x_i x_j + a_{hh}^k x_k^2 + a_{kh}^k x_h x_k + a_{hk}^k x_k x_h + a_{kk}^k x_h^2 &= \sum_{j \neq h,k} \sum_{i \neq h,k} a_{ij}^h x_i x_j + a_{hh}^h x_h^2 + a_{kh}^h x_k x_h + a_{hk}^h x_h x_k + a_{kk}^h x_k^2 \end{aligned} \tag{8}$$

which yields:

$$\begin{aligned} a_{ij}^h &= a_{ij}^k \qquad \forall h, k, (i, j) \neq (h, k) \\ a_{hh}^h &= a_{kk}^k \qquad \forall h, k \\ a_{hh}^k &= a_{kk}^h \qquad \forall h, k \\ a_{kh}^k + a_{hk}^k &= a_{kh}^h + a_{hk}^h \qquad \forall h, k. \end{aligned} \tag{9}$$