On the Universality of Invariant Networks: Supplementary Material

Haggai Maron [1] **Ethan Fetaya** [2] **Nimrod Segol** [1] **Yaron Lipman** [1]

Lemma 2. *There exists a G-invariant network in the sense of definition 3 that realizes the sum of G-invariant networks $F=\sum_{k=0}^{d}\sum_{j=1}^{n_k}\alpha_{kj}F^{kj}$.*

Proof. We need to show that $F=\sum_{k=0}^{d}\sum_{j=1}^{n_k}\alpha_{kj}F^{kj}$ can indeed be realized as a *single, unified* G-invariant network. As we already saw, each network F^{kj} has the structure

$$\mathbb{R}^n \xrightarrow{L^\tau} \mathbb{R}^{n^k\times k} \xrightarrow{M^k} \mathbb{R}^{n^k} \xrightarrow{s} \mathbb{R},$$

with a suitable k-class τ. To create the unified G-invariant network we first lift each F^{kj} to the maximal dimension d. That is, $\widetilde{F}^{kj}$ with the structure

$$\mathbb{R}^n \xrightarrow{\widetilde{L}^{kj}} \mathbb{R}^{n^d\times k} \xrightarrow{\widetilde{M}^k} \mathbb{R}^{n^d} \xrightarrow{s} \mathbb{R}.$$

This is done by composing each equivariant layer $L:\mathbb{R}^{n^k\times a}\rightarrow\mathbb{R}^{n^l\times b}$ with two linear equivariant operators $U^b:\mathbb{R}^{n^k\times b}\rightarrow\mathbb{R}^{n^d\times b}$ and $D^a:\mathbb{R}^{n^d\times a}\rightarrow\mathbb{R}^{n^k\times a}$,

$$U^bLD^a:\mathbb{R}^{n^d\times a}\rightarrow\mathbb{R}^{n^d\times b}, \tag{1}$$

where

$$U^b(x)_{i_1\dots i_d,j}=x_{i_1\dots i_k,j}$$

and

$$D^a(y)_{i_1\dots i_k,j}=n^{k-d}\sum_{i_{k+1}\dots i_d=1}^{n} y_{i_1\dots i_k i_{k+1}\dots i_d,j}\ .$$

Since U^b, D^a are equivariant, U^bLD^a in equation 1 is equivariant. Furthermore $D^a\circ\sigma\circ U^a=\sigma$, where σ is the pointwise activation function. Lastly, given two G-invariant networks with the same tensor order d they can be combined to a single G-invariant network by concatenating their features. That is, if $L_1:\mathbb{R}^{n^d\times a}\rightarrow\mathbb{R}^{n^d\times b}$, and $L_2:\mathbb{R}^{n^d\times a'}\rightarrow\mathbb{R}^{n^d\times b'}$, then their concatenation would yield $L_{1,2}:\mathbb{R}^{n^d\times(a+a')}\rightarrow\mathbb{R}^{n^d\times(b+b')}$. Applying this concatenation to all $\widetilde{F}^{kj}$ we get our unified G-invariant network. □

[1]Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel [2]Department of Computer Science, University of Toronto, Toronto, Canada. Correspondence to: Haggai Maron <haggai.maron@weizmann.ac.il>.

Proceedings of the 36^{th} International Conference on Machine Learning, Long Beach, California, PMLR 97, 2019.

Fixed-point equation for equivariant layers. We have an affine operator $L:\mathbb{R}^{n^k\times a}\rightarrow\mathbb{R}^{n^l\times b}$ satisfying

$$g^{-1}\cdot L(g\cdot\mathbf{X})=L(\mathbf{X}), \tag{2}$$

for all $g\in G$, $\mathbf{X}\in\mathbb{R}^{n^k\times a}$. The purely linear part of L can be written using a tensor $\mathbf{L}\in\mathbb{R}^{n^{k+l}\times a\times b}$: Write

$$L(\mathbf{X})_{j_1\dots j_l,j}=\sum_{i_1\dots i_k,i}\mathbf{L}_{j_1\dots j_l,i_1\dots i_k,i,j}\mathbf{X}_{i_1\dots i_k,i}.$$

Writing equation 2 using this notation gives:

$$\begin{aligned}&\sum_{i_1\dots i_k,i}\mathbf{L}_{g(j_1)\dots g(j_l),i_1\dots i_k,i,j}\mathbf{X}_{g^{-1}(i_1)\dots g^{-1}(i_k),i}\\ &=\sum_{i_1\dots i_k,i}\mathbf{L}_{g(j_1)\dots g(j_l),g(i_1)\dots g(i_k),i,j}\mathbf{X}_{i_1\dots i_k,i}\\ &=\sum_{i_1\dots i_k,i}\mathbf{L}_{j_1\dots j_l,i_1\dots i_k,i,j}\mathbf{X}_{i_1\dots i_k,i},\end{aligned}$$

for all $g\in G$ and $\mathbf{X}\in\mathbb{R}^{n^k\times a}$. This implies equation 9, namely

$$g\cdot\mathbf{L}=\mathbf{L},\ g\in G.$$

The constant part of L is done similarly.