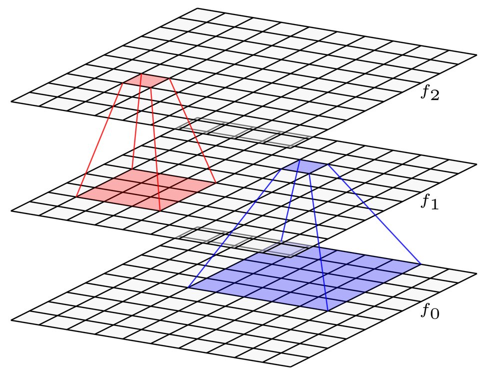
On the generalization of equivariance and convolution in neural nets to the action of compact groups

Risi Kondor* and Shubhendu Trivedi†

* The University of Chicago, †Toyota Technological Institute at Chicago

Convolutional Neural Networks



Definition 1. *Let $\mathcal{X}_0, \ldots, \mathcal{X}_L$ be a sequence of index sets, $\phi_1, \ldots, \phi_L$ linear maps*

$$\phi_\ell \colon L(\mathcal{X}_{\ell-1}) \longrightarrow L(\mathcal{X}_\ell),$$

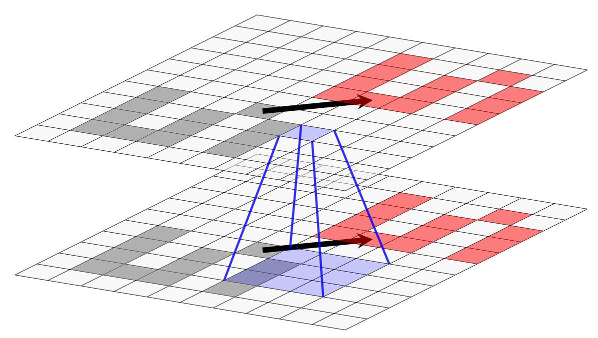
*and $\sigma_\ell \colon V_\ell \to V_\ell$ appropriate pointwise nonlinearities, such as the ReLU operator. The corresponding **multilayer feed-forward neural network (MFF-NN)** is then a sequence of maps $f_0 \mapsto f_1 \mapsto f_2 \mapsto \ldots \mapsto f_L$, where $f_\ell(x) = \sigma_\ell(\phi_\ell(f_{\ell-1})(x))$.*

In a **Convolutional Neural Network (CNN)** each ϕ_ℓ linear map is just a convolution with a corresponding filter g_ℓ:

$$\phi_\ell(f_{\ell-1}) = (f_{\ell-1} * g_\ell)(x) = \sum_{y \in \mathbb{Z}^2} f_{\ell-1}(x-y)\, g_\ell(y).$$

It is these filters that the CNN learns from the training data.

Equivariance



A traditional CNN is equivariant in the sense that if the input to the network is translated

$$f_0 \mapsto f_0' \qquad\qquad f_0(\mathbf{x}) = f_0(\mathbf{x} - \boldsymbol{t}),$$

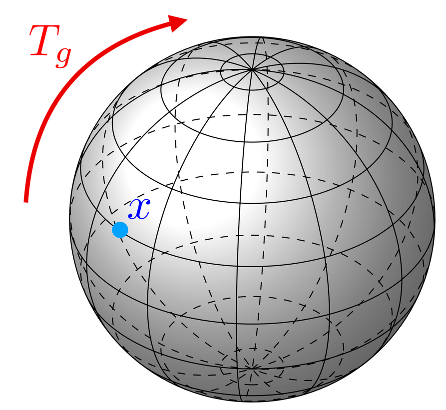
then the activations in higher layers transform in a corresponding way

$$f_\ell \mapsto f_\ell' \qquad\qquad f_\ell(\mathbf{x}) = f_\ell(\mathbf{x} - \boldsymbol{t}).$$

Equivariance is important for multiple reasons:
1. It reduces the number of parameters that the network needs to learn.
2. It ensures that the same filters are applied to every part of the image.
3. If we add a final translation invariant layer, then the entire neural network will be translation invariant.

Equivariance to groups

It many settings one wants to construct neural networks that are invariant to some group G other than translations, e.g., the group of 3D rotatations, $\mathrm{SO}(3)$. In these setting, however, the activations often live not on G itself, but on a space $\mathcal{X}$ that G **acts on** (technically, $\mathcal{X}$ is a **homogeneous space** or **quotient space** G/H).



The general setup is the following:
1. Each layer of the network corresponds to a homogeneous space $\mathcal{X}_\ell$.
2. G acts on $\mathcal{X}_\ell$ by $x \mapsto T_g^\ell(x)$ (with $g \in G$).
3. The activation of layer ℓ is a function $f_\ell \in L(\mathcal{X}_\ell)$.
4. The induced action of G on $L(\mathcal{X}_\ell)$ is

$$f \longmapsto \mathbb{T}_g^\ell(f) \qquad\qquad \mathbb{T}_g^\ell(f)(x) = f((T_g^\ell)^{-1}(x)).$$

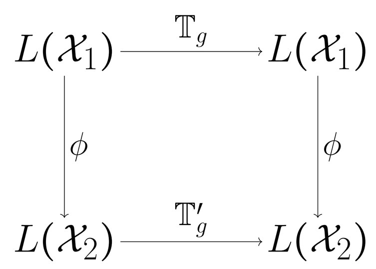
Definition 2. *Let G be a group and $\mathcal{X}_1, \mathcal{X}_2$ be two sets with corresponding G-actions*

$$T_g \colon \mathcal{X}_1 \to \mathcal{X}_1, \qquad\qquad T_g' \colon \mathcal{X}_2 \to \mathcal{X}_2.$$

*Let $\mathbb{T}$ and $\mathbb{T}'$ be the induced actions of G on $L\mathcal{X}_1$ and $L\mathcal{X}_2$. We say that a (linear or non-linear) map $\phi \colon L(\mathcal{X}_1) \to L(\mathcal{X}_2)$ is **equivariant** with the action of G (or G–**equivariant** for short) if*

$$\phi(\mathbb{T}_g(f)) = \mathbb{T}_g'(\phi(f)) \qquad\qquad \forall f \in L(\mathcal{X}_1)$$

for any group element $g \in G$.



Definition 3. *Let $\mathcal{N}$ be a feed-forward neural network with $L+1$ layers and G be a group that acts on each index space $\mathcal{X}_0, \ldots, \mathcal{X}_L$. Let $\mathbb{T}^0, \mathbb{T}^1, \ldots, \mathbb{T}^L$ be the corresponding actions on $L(\mathcal{X}_0), \ldots, L(\mathcal{X}_L)$. We say that $\mathcal{N}$ is a G–**equivariant feed-forward network** if, when the inputs are transformed $f_0 \mapsto \mathbb{T}_g^0(f_0)$ (for any $g \in G$), the activations of the other layers correspondingly transform as $f_\ell \mapsto \mathbb{T}_g^\ell(f_\ell)$.*

Convolution on groups

Given $f, g \colon G \to \mathbb{C}$, the **convolution** of f with g is defined

$$(f * g)(x) = \int f(xy^{-1})\, g(y)\, d\mu(y).$$

If $f \colon G/H \to \mathbb{C}$ and $g \colon G/K \to \mathbb{C}$, then

$$(f * g)(u) = \int_G f{\uparrow}^G(uv^{-1})\, g{\uparrow}^G(v)\, d\mu(v).$$

The **Fourier transform** of $f \colon G \to \mathbb{C}$ is defined as the collection of matrices

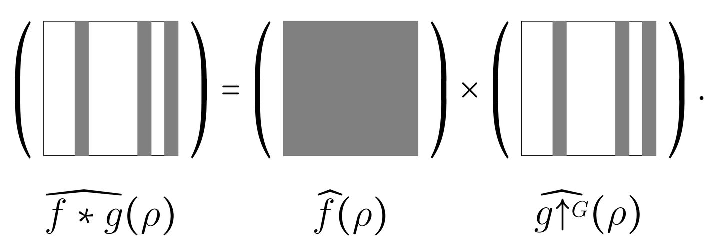
$$\widehat{f}(\rho_i) = \int_G f(u)\, \rho_i(u)\, d\mu(u),$$

where $\rho_0, \rho_1, \rho_2, \ldots$ are the **irredicible representations** of G. The **convolution theorem** on compact groups states that

$$\widehat{f * g}(\rho_i) = \widehat{f}(\rho_i) \cdot \widehat{g}(\rho_i).$$

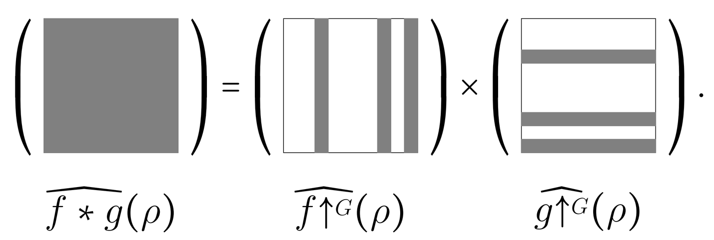
Case I: $f : G \to \mathbb{C}$ and $g : G/H \to \mathbb{C}$

$$f * g : G/H \to \mathbb{C} \qquad\qquad (f * g)(x) = \int_G f(\overline{x}v^{-1})\, g([v]_{G/H})\, d\mu(v).$$



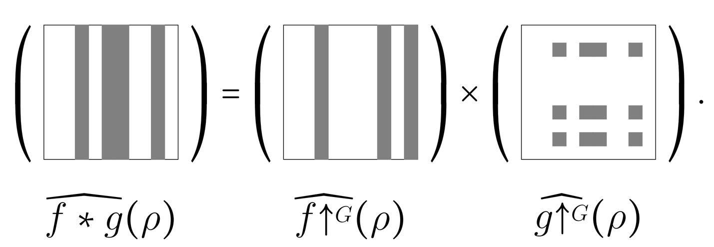
Case II: $f : G/H \to \mathbb{C}$ and $g : H\backslash G \to \mathbb{C}$

$$f * g \colon G \to \mathbb{C} \qquad\qquad (f * g)(u) = |H| \int_{H\backslash G} f([u\overline{y}^{-1}]_{G/H})\, g(y)\, d\mu(y)$$



Case III: $f : G/H \to \mathbb{C}$ and $g : H\backslash G/K \to \mathbb{C}$

$$f * g \colon G/K \to \mathbb{C} \qquad\qquad (f * g)(x) = |H| \int_{H\backslash G} f([\overline{x}\overline{y}^{-1}]_{\mathcal{X}})\, g([\overline{y}]_{H\backslash G/K})\, d\mu(y).$$



Main theorem

Theorem 1. *A feed-forward neural network is equivariant to the action of a compact group G if and only if the linear operation in each layer is of the form*

$$\phi_\ell(f_{\ell-1}) = f_{\ell-1} * g_\ell$$

for a learnable filter g_ℓ.

Applications

1. Spherical CNNs [Cohen et al., 2018] [Kondor et al., 2018]
2. CNNs on manifolds and steerability [Masci et al., 2015] [Marcos et al, 2017] [Worral et al, 2017]
3. CNNs for graphs [Gilmer et al., 2017] [Son et al., 2018]
4. Covariant networks for physical systems [Kondor, 2018]
5. ...

References

1. Cohen, T.S. and Welling, M. Group equivariant convolutional networks (ICML 2016)
2. Cohen, T.S., Geiger, M., K ohler J. and Welling, M. Spherical CNNs (ICLR 2018)
3. Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O., and Dahl, G. E. Neural message passing for quantum chemistry (ICML 2017)
4. Kondor, R. N-body networks: a covariant hierarchical neural network architecture for learning atomic potentials (preprint, 2018)
5. Kondor, R., Lin, Z. and Trivedi, S. Clebsch-Gordan Nets: a Fully Fourier Space Spherical Convolutional Neural Network (preprint, 2018)
6. Marcos, D., Volpi, M., Komodakis, N., and Tuia, D. Rotation equivariant vector field networks. (ICCV 2017)
7. Masci, J., Boscaini, D., Bronstein, M. M., and Vandergheynst, P. Geodesic convolutional neural networks on Riemannian manifolds (ICCVW 2015)
8. Ravanbakhsh, S., Schneider, J., and Poczos, B. Equivariance through parameter-sharing (ICML 2017)
9. Son, H-T, Trivedi, S., Pan, H, Anderson, B, M. and Kondor, R. Predicting Molecular Properties with Covariant Compositional Networks (JCP, 2018)
10. Worrall, D. E., Garbin, S. J., Turmukhambetov, D., and Brostow, G. J. Harmonic networks: Deep translation and rotation equivariance (CVPR 2017)