Equivariance in Deep Learning

$$g \cdot f(x) = f(g \cdot x)$$

Classification is Translation-Invariant

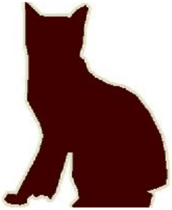
$$\text{class} : L^2(\mathbb{Z}^2, \mathbb{R}^3) \to \{1, \ldots, n\}$$

 $\mapsto$ Cat

 $\mapsto$ Cat

Segmentation is Translation-Equivariant

$$\mathrm{mask} : L^2(\mathbb{Z}^2, \mathbb{R}^3) \to L^2(\mathbb{Z}^2, \{0, 1\})$$

 $\mapsto$ 

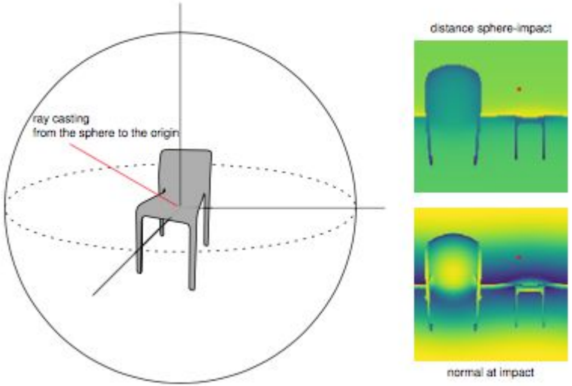
 $\mapsto$ 

Definition: Let X and Y be spaces admitting an action of a group G. Then $f : X \to Y$ is G-equivariant if $f(g \cdot x) = g \cdot f(x)$ for all x.

If G acts trivially on Y, then f is G-invariant.

3D Classification is SO(3)-Invariant

$$\text{class} : L^2(S^2, \mathbb{R}^3) \to \{1, \ldots, n\}$$



$\mapsto$

Chair

Parameter Estimation is Permutation-Invariant

Draw $X_i \sim \mathcal{P}_\theta$. Learn

$$\{X_1, \ldots, X_n\} \to f(\theta)$$

Outlier Detection is Permutation-Equivariant

A permutation-invariant construction

Let X be a set of k elements. Let $\phi : X \to \mathbb{C}^n$ and $\rho : \mathbb{C}^n \to \mathbb{C}^m$. Then

$$x \mapsto \rho(\sum_{x \in X} \phi(x))$$

is invariant under the action of S_k on X.

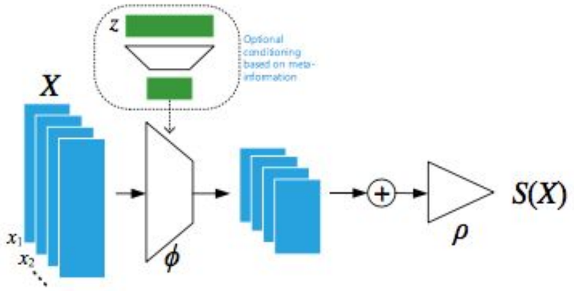


Figure 2. Architecture of deep sets

It’s universal

Theorem: Every permutation-invariant function can be approximated arbitrarily well using such a ρ and ϕ.

Theorem: All symmetric polynomials of $x_1, \ldots, x_k$ are polynomials of the homogeneous symmetric monomials $x_1^p + \cdots x_k^p$ for $0 \leq p \leq k$.

Example: $\max(x_1, \ldots, x_n) \approx \sqrt[p]{x_1^p + \cdots + x_n^p}$ for p large.

Equivariant Layers

$$f(\mathbf{x}) = \sigma(\lambda \mathbf{I}\mathbf{x} + \gamma(\mathbf{1}^{\mathbf{T}}\mathbf{x})\mathbf{1})$$

$$f(\mathbf{x}) = \sigma(\lambda \mathbf{I}\mathbf{x} + \gamma \max(\mathbf{x})\mathbf{1})$$

S2 CNN

SO(3)-convolutions

Let $\psi, f \in L^2(S^2, \mathbb{C})$. The $SO(3)$ convolution of ψ and f is a function in $L^2(SO(3), \mathbb{C})$ given by

$$[\psi \star f](R) = \langle L_R \psi, f \rangle = \int_{S^2} \sum_{k=1}^{K} \psi_k(R^{-1}x) f_k(x) dx$$

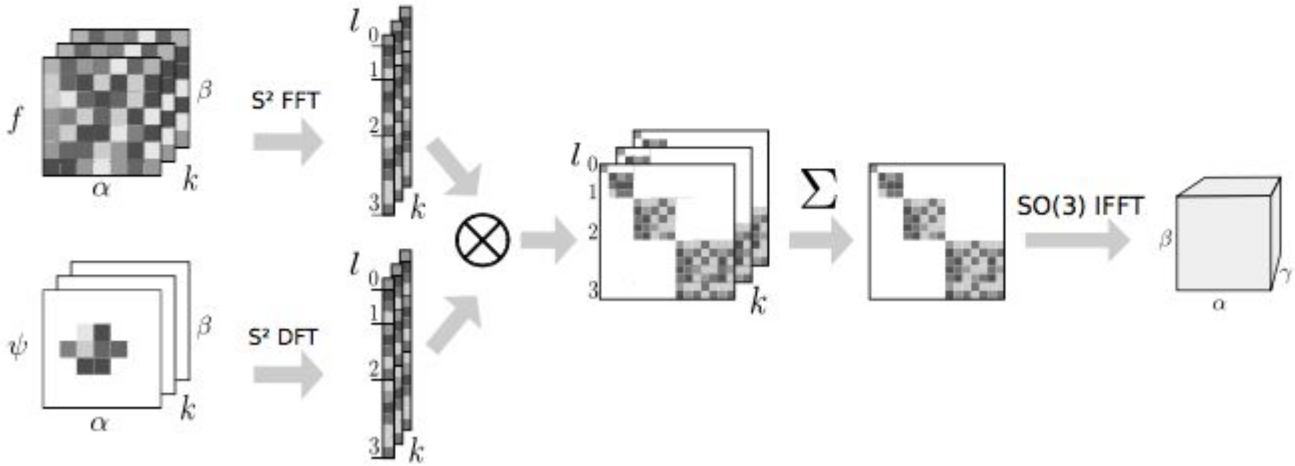
SO(3)-convolutions

The inverse SO(3) Fourier transform is defined as:

$$f(R) = \sum_{l=0}^{b} (2l+1) \sum_{m=-l}^{l} \sum_{n=-l}^{l} \hat{f}^l_{mn} D^l_{mn}(R),$$

$$\widehat{\psi \star f} = \hat{f} \cdot \hat{\psi}^\dagger$$

SO(3)-convolutions



G-convolutions

Let X be a G-space, and let $f, g \to \mathbb{C}$ be functions. The **convolution** of f and g is a function on G defined by

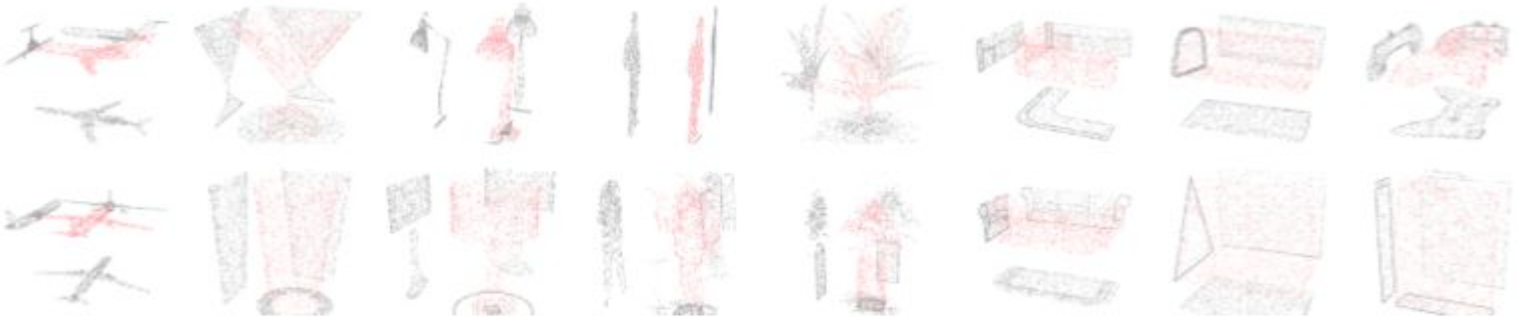
$$f * g(R) = \int_X f(R^{-1} \cdot x) g(x) dx.$$

G-convolutions

Theorem (Kondor-Trivedi): If a neural network connecting layers of the form $L^2(X_i, \mathbb{C}^{n_i})$ for a series of G-spaces X_i is G-equivariant, then it is a composition G-convolutions on the X_i and nonlinearities applied to the $\mathbb{C}^{n_i}$.

Questions

- What does this mean for permutation-equivariance? Schur representations?
- What does this mean for graph convolutions?
- Point clouds?!



References

- *Spherical CNNs*, Cohen et al.
  - SO(3)-equivariant convolutions
- *Deep Sets*, Zaheer et al.
  - Permutation invariant functions
- *On the Generalization of Equivariance and Convolution in Neural Networks to the Action of Compact Groups*, Kondor and Trivedi
  - Convolutions between functions on G-spaces