NORA Summer School 2024 on Geometric Deep Learning - Introduction

Nello Blaser

UNIVERSITY OF BERGEN

Nello Blaser

Erlend Grong

Raghavendra Selvan

Gabriele Cesa

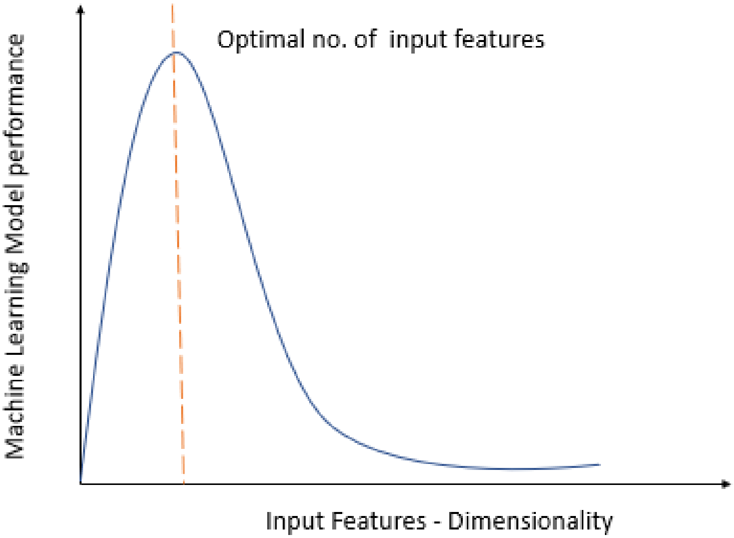
Time	June 10	June 11	June 12	June 13	June 14
9.00-10.30	Introduction (Nello Blaser)	Geometric blueprint (Erlend Grong)	Graphs II (Raghavendra Selvan)	Equivariance I (Gabriele Cesa)	Equivariance III (Gabriele Cesa)
10.30-11.00	Break	Break	Break	Break	Break
11.00-12.30	Geometry basics (Erlend Grong)	Graphs I (Raghavendra Selvan)	Graphs III (Raghavendra Selvan)	Equivariance II (Gabriele Cesa)	Summary (Nello Blaser)
12.30-14.00	Lunch	Lunch	Lunch	Lunch	Lunch
14.00-17.00	Practical	Practical	Practical	Practical	

Materials

- **Course material**
  https://github.com/blasern/NORA2024_GDL

- **Geometric deep learning** by *Michael M. Bronstein, Joan Bruna, Taco Cohen, Petar Veličković*
  https://geometricdeeplearning.com/lectures

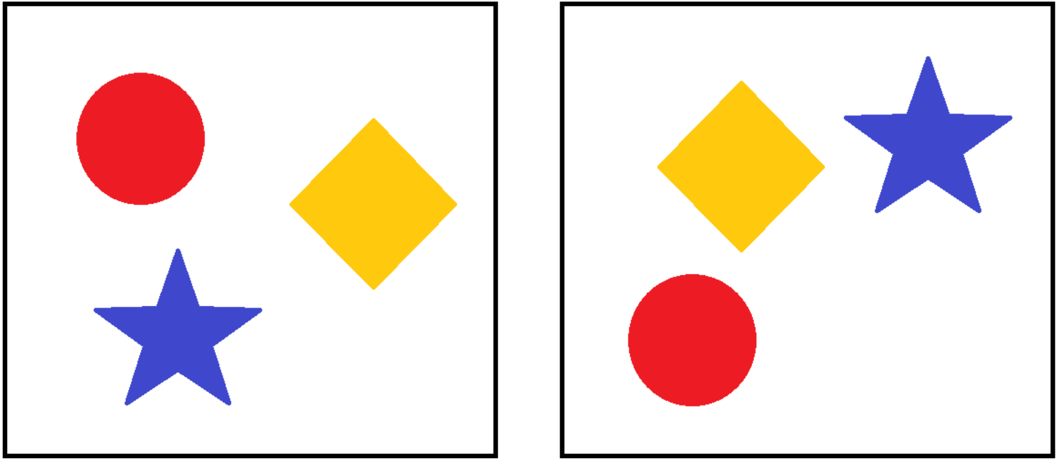
Course of dimensionality

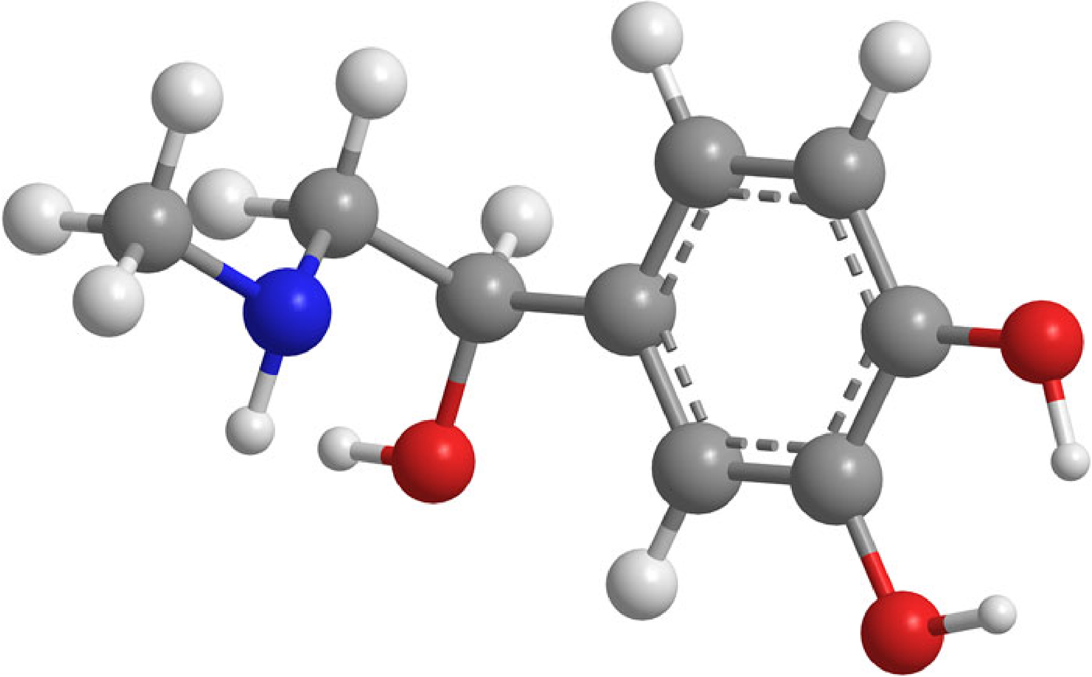


Error decomposition

- $y_i = f^*(x_i)$ for some unknown f^*
- hypothesis class $\mathcal{F} \subset \{f : \mathcal{X} \to \mathbb{R}\}$
- complexity measure $\gamma : \mathcal{F} \to \mathbb{R}$
- restricted hypothesis class $\mathcal{F}_\delta = \{f \in \mathcal{F} \,|\, \gamma(f) < \delta\}$
- loss $\ell : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$

$$\mathbb{E}\left[\ell(f(x), f^*(x))\right] \leq \varepsilon_{class} + \varepsilon_{opt} + \varepsilon_{stat} + \varepsilon_{approx}$$



Equivariance and invariance

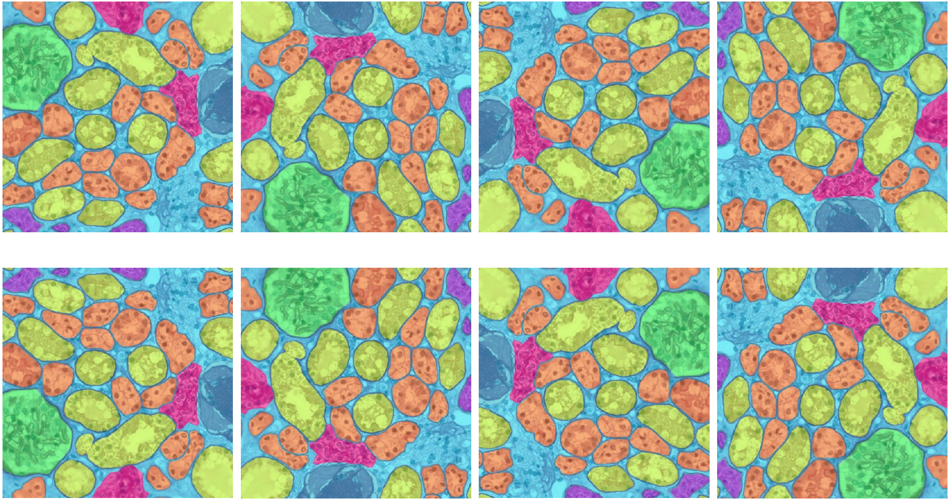
- **Invariance**:

$$f(x) = f(gx)$$

- **Equivariance**:

$$gf(x) = f(gx)$$

Example



Simple solution?

Simple solution?

Data augmentation

- Train-time data augmentation
- Test-time data augmentation

Equivariance and invariance

- **Invariance**:

$$f(x) = f(gx)$$

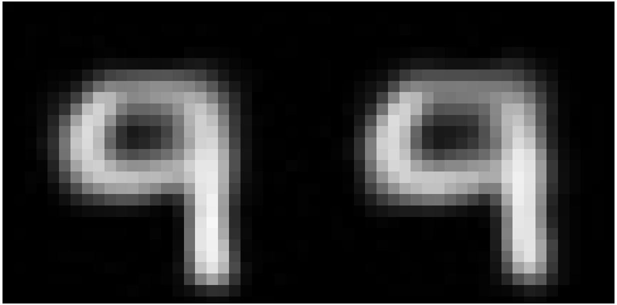
- **Equivariance**:

$$gf(x) = f(gx)$$

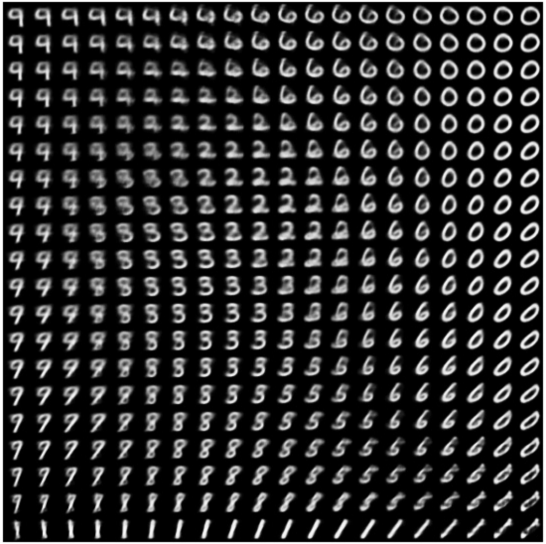
Example

Example

Example



Example



Group

A group $(G, \cdot)$ is a set of elements G and a binary operation $\cdot : G \times G \to G$, satisfying the four conditions below.

1. **Closure**: For all $a, b \in G$, also $a \cdot b \in G$.
2. **Identity**: There exists an element $e \in G$, such that for all $a \in G$ it holds that $e \cdot a = a \cdot e = a$.
3. **Inverses**: For all $a \in G$ there exists an element $a^{-1} \in G$ such that $a \cdot a^{-1} = a^{-1} \cdot a = e$.
4. **Associativity**: For all $(a, b, c) \in G$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

Equivariance and invariance

- **Invariance**:

$$f(x) = f(gx)$$

- **Equivariance**:

$$gf(x) = f(gx)$$

Geometric blueprint

- Linear equivariant layer
- Nonlinearity
- Local pooling
- Global pooling

Sets

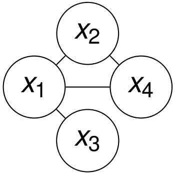
- **Data**: Set x with elements $x = (x_0, \ldots, x_k)$
- **Requirement**: $f(x_0, \ldots, x_k) = f(x_{\pi(0)}, \ldots, x_{\pi(k)})$ for any permutation π

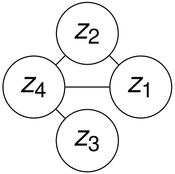
Sets

- **Data**: Set x with elements $x = (x_0, \ldots, x_k)$
- **Requirement**: $f(x_0, \ldots, x_k) = f(x_{\pi(0)}, \ldots, x_{\pi(k)})$ for any permutation π
- **Solution**:

$$f(x) = \varphi\left(\bigoplus_{u \in \mathcal{V}} \psi(x_u)\right)$$

Graphs





$$X_0 = (x_1, x_2, x_3, x_4)^T$$

$$A_0 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

$$X_1 = (z_1, z_2, z_3, z_4)^T$$

$$A_1 = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

Graphs

P permutation matrix, X node data, A adjacency matrix

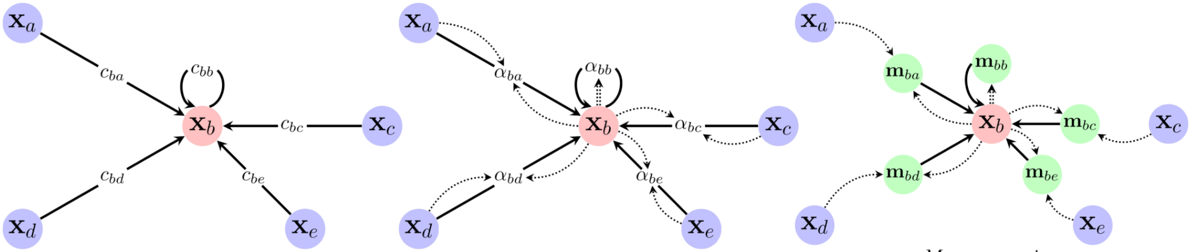
- **Invariance**:

$$f(PX, PAP^T) = f(X, A)$$

- **Equivariance**:

$$F(PX, PAP^T) = PF(X, A)$$

Graphs



Convolutional *Attentional* *Message-passing*

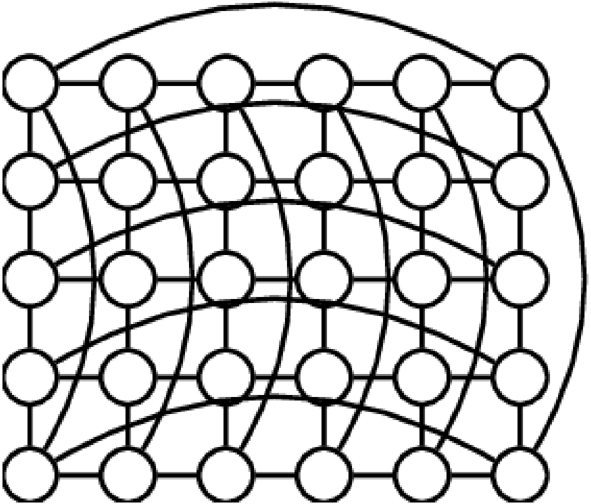
$$\mathbf{h}_i = \phi\left(\mathbf{x}_i, \bigoplus_{j \in \mathcal{N}_i} c_{ij}\psi(\mathbf{x}_j)\right) \qquad \mathbf{h}_i = \phi\left(\mathbf{x}_i, \bigoplus_{j \in \mathcal{N}_i} a(\mathbf{x}_i, \mathbf{x}_j)\psi(\mathbf{x}_j)\right) \qquad \mathbf{h}_i = \phi\left(\mathbf{x}_i, \bigoplus_{j \in \mathcal{N}_i} \psi(\mathbf{x}_i, \mathbf{x}_j)\right)$$

Grids

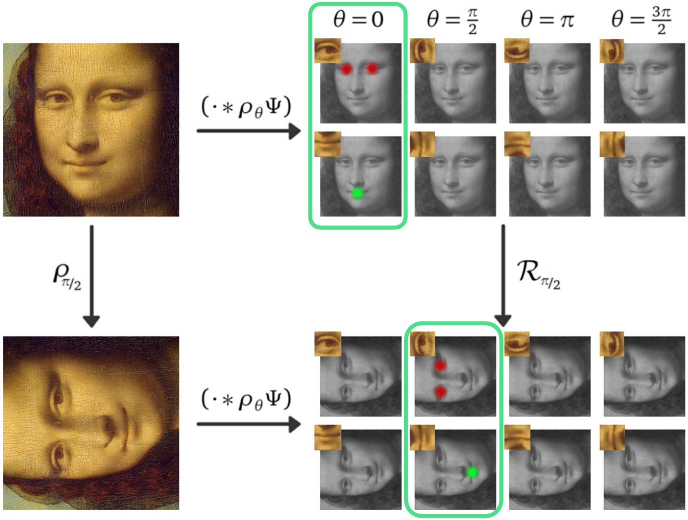
What are linear equivariant layers for grids?

$$\begin{tikzcd} x_1 \arrow[r] & x_2 \arrow[r] & x_3 \arrow[r] & x_4 \arrow[r] & x_5 \arrow[r] & x_6 \arrow[lllll, bend left] \end{tikzcd}$$

Grids

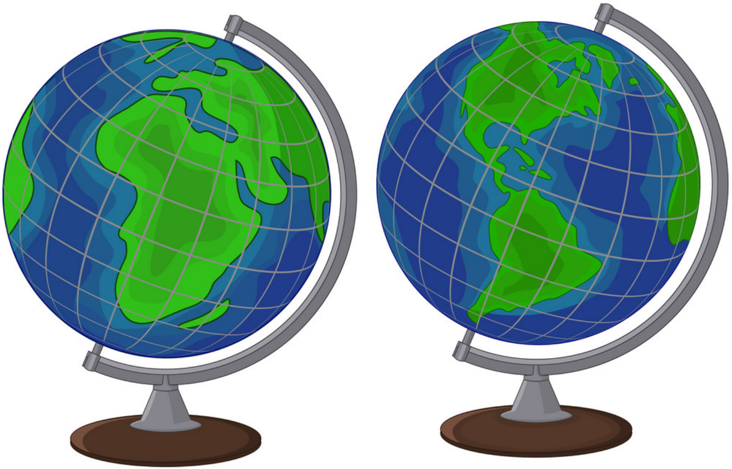


Groups

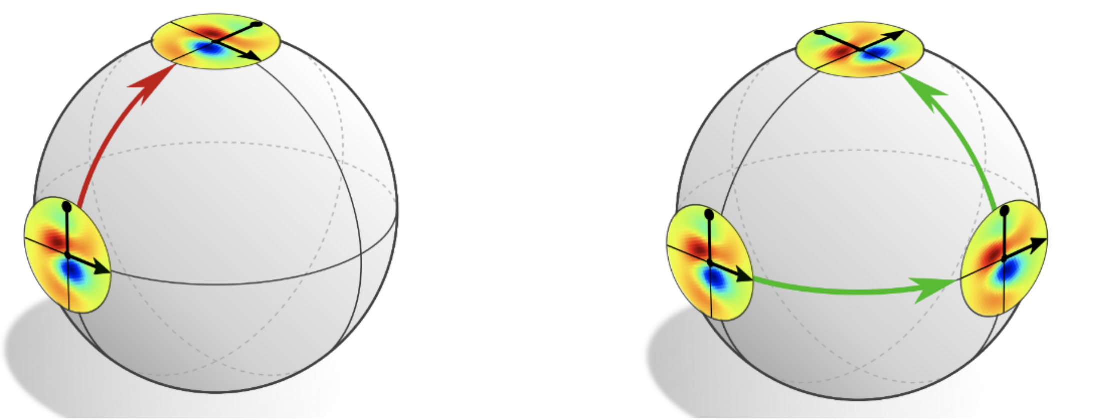


$\theta = 0$
$\theta = \frac{\pi}{2}$
$\theta = \pi$
$\theta = \frac{3\pi}{2}$
$(\cdot * \rho_\theta \Psi)$
$\rho_{\pi/2}$
$\mathcal{R}_{\pi/2}$
$(\cdot * \rho_\theta \Psi)$

Manifolds



Manifolds



Conclusions

- Geometric Deep Learning aims to take advantage of structural properties of learning tasks by providing a principled way to designing neural network architectures.
- Invariant and equivariant neural networks can be built from linear equivariant layers, non-linearities, local and global pooling.
- Examples of domains are sets, graphs, grids, groups and manifolds.

Time	June 10	June 11	June 12	June 13	June 14
9.00-10.30	Introduction (Nello Blaser)	Geometric blueprint (Erlend Grong)	Graphs II (Raghavendra Selvan)	Equivariance I (Gabriele Cesa)	Equivariance III (Gabriele Cesa)
10.30-11.00	Break	Break	Break	Break	Break
11.00-12.30	Geometry basics (Erlend Grong)	Graphs I (Raghavendra Selvan)	Graphs III (Raghavendra Selvan)	Equivariance II (Gabriele Cesa)	Summary (Nello Blaser)
12.30-14.00	Lunch	Lunch	Lunch	Lunch	Lunch
14.00-17.00	Practical	Practical	Practical	Practical	