Invariant and Equivariant Graph Networks

Angelo Rajendram

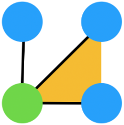
14 Feb 2024

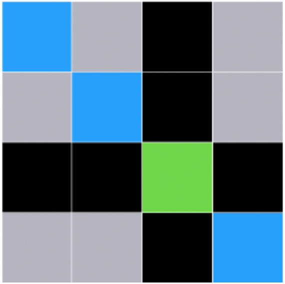
Question:

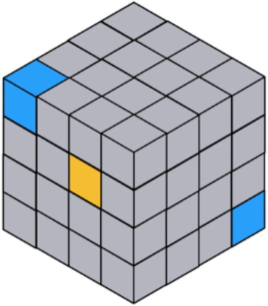
How do we construct operators that respect arbitrary symmetries over graph data?

Definitions: Hyper-graphs And Tensors

$$\mathcal{G} = (\mathbb{V}, \mathbf{A})$$

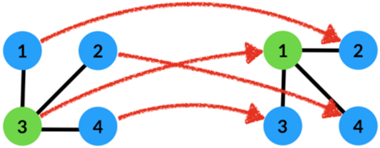




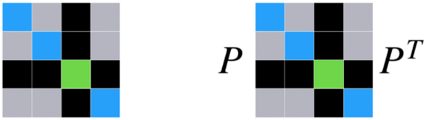


$\mathbf{A} \in \mathbb{R}^{n} \equiv \mathbb{R}^{n^1}$ $\mathbf{A} \in \mathbb{R}^{n \times n} \equiv \mathbb{R}^{n^2}$ $\mathbf{A} \in \mathbb{R}^{n \times n \times n} \equiv \mathbb{R}^{n^3}$

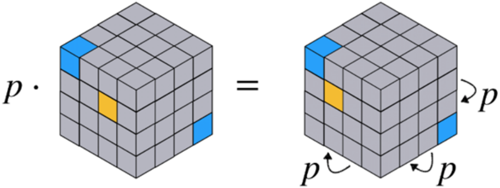
Definitions: Permutations



Permutation Invariance



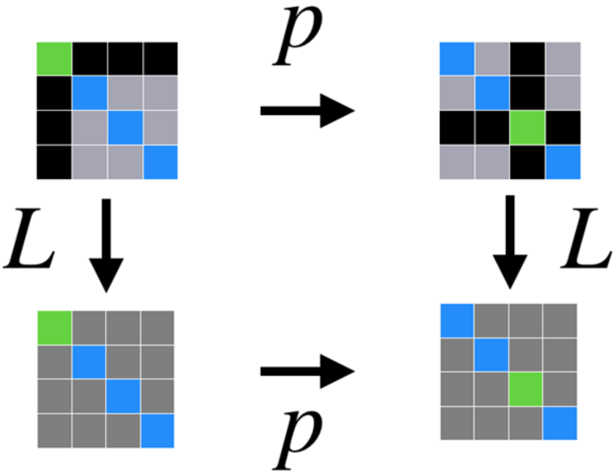
$$f(P^T \mathbf{A} P) = f(\mathbf{A})$$



$$f(p \cdot \mathbf{A}) = f(\mathbf{A})$$

Definitions: Permutation Equivariance

$$f(P\mathbf{A}P^T) = P^T f(\mathbf{A})P \qquad f(p \cdot \mathbf{A}) = p \cdot f(\mathbf{A})$$

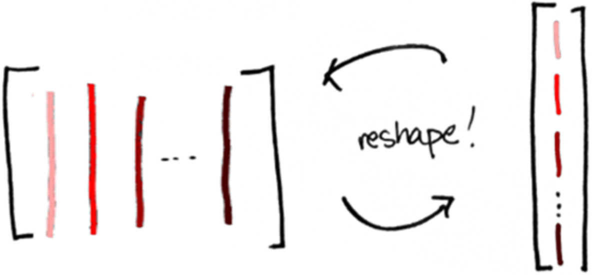


Definitions: Vectorization

$$\text{Given } \mathbf{X} \in \mathbb{R}^{a \times b}, \textit{ denote}$$

$$\mathrm{vec}(\mathbf{X}) \in \mathbb{R}^{ab \times 1}$$

$$[\mathrm{vec}(\mathbf{X})] = \mathbf{X} \in \mathbb{R}^{a \times b}$$



Example for vectors,

$$\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \ \vec{w} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$$\vec{v} \otimes \vec{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 4 & 5 \end{bmatrix} = \begin{bmatrix} 4 \cdot 1 & 5 \cdot 1 \\ 4 \cdot 2 & 5 \cdot 2 \\ 4 \cdot 3 & 5 \cdot 3 \end{bmatrix} = \begin{bmatrix} 4 & 5 \\ 8 & 10 \\ 12 & 15 \end{bmatrix}$$

Definitions: Kronecker Product for Matrices

Consider $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$,

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{mm}B & \dots & a_{mn}B \end{bmatrix}$$

$$A \otimes B \in \mathbb{R}^{(mp) \times (nq)}$$

Definitions: Linear Operators as Tensors

Matrix multiplication is a linear operation of vectors,

$L : \mathbb{R}^n \to \mathbb{R}^m$ is encoded as $\mathbf{L} \in \mathbb{R}^{m \times n}$,

Definitions: Linear Operators as Tensors

Matrix multiplication is a linear operation of vectors,

$L : \mathbb{R}^n \to \mathbb{R}^m$ is encoded as $\mathbf{L} \in \mathbb{R}^{m \times n}$, similarly,

$L_I : \mathbb{R}^{n \times n} \to \mathbb{R}$ is encoded as $\mathbf{L}_I \in \mathbb{R}^{1 \times n^2}$ (invariance)

Definitions: Linear Operators as Tensors

Matrix multiplication is a linear operation of vectors,

$L : \mathbb{R}^n \to \mathbb{R}^m$ is encoded as $\mathbf{L} \in \mathbb{R}^{m \times n}$, similarly,

$L_I : \mathbb{R}^{n \times n} \to \mathbb{R}$ is encoded as $\mathbf{L}_I \in \mathbb{R}^{1 \times n^2}$ (invariance)

$L_E : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$ is encoded as $\mathbf{L}_E \in \mathbb{R}^{n^2 \times n^2}$ (equivariance)

Definitions: Linear Operators as Tensors

Matrix multiplication is a linear operation of vectors,

$L : \mathbb{R}^n \to \mathbb{R}^m$ is encoded as $\mathbf{L} \in \mathbb{R}^{m \times n}$, similarly,

$L_I : \mathbb{R}^{n \times n} \to \mathbb{R}$ is encoded as $\mathbf{L}_I \in \mathbb{R}^{1 \times n^2}$ (invariance)

$L_E : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$ is encoded as $\mathbf{L}_E \in \mathbb{R}^{n^2 \times n^2}$ (equivariance)

$L_M : \mathbb{R}^{n^k} \to \mathbb{R}^{n^l}$ is encoded as $\mathbf{L}_M \in \mathbb{R}^{n^l \times n^k}$ (mixed-order)

Definitions: Linear Operators as Tensors

Matrix multiplication is a linear operation of vectors,

$L : \mathbb{R}^n \to \mathbb{R}^m$ is encoded as $\mathbf{L} \in \mathbb{R}^{m \times n}$, similarly,

$L_I : \mathbb{R}^{n \times n} \to \mathbb{R}$ is encoded as $\mathbf{L}_I \in \mathbb{R}^{1 \times n^2}$ (invariance)

$L_E : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$ is encoded as $\mathbf{L}_E \in \mathbb{R}^{n^2 \times n^2}$ (equivariance)

$L_M : \mathbb{R}^{n^k} \to \mathbb{R}^{n^l}$ is encoded as $\mathbf{L}_M \in \mathbb{R}^{n^l \times n^k}$ (mixed-order)

The exponent of the index size n indicates the tensor order

Invariant and Equivariant Linear Operators

What are the set of all invariant and equivariant linear operators?

Paper Sketch:

- Determine conditions for invariant/equivariant linear operators: Fixed-point Equations

Invariant and Equivariant Linear Operators

What are the set of all invariant and equivariant linear operators?
Paper Sketch:

- Determine conditions for invariant/equivariant linear operators: Fixed-point Equations
- Identify a basis for solutions to the Fixed-point Equations

Invariant and Equivariant Linear Operators

What are the set of all invariant and equivariant linear operators?
Paper Sketch:

- Determine conditions for invariant/equivariant linear operators: Fixed-point Equations
- Identify a basis for solutions to the Fixed-point Equations
- Incorporate biases (e.g. a linear layer in an ANN is given as $Ax + b$, where $A \in \mathbb{R}^{n \times n}$, and $x, b \in \mathbb{R}^n$)

Invariant and Equivariant Linear Operators

What are the set of all invariant and equivariant linear operators?
Paper Sketch:

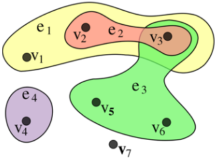
- Determine conditions for invariant/equivariant linear operators: Fixed-point Equations
- Identify a basis for solutions to the Fixed-point Equations
- Incorporate biases (e.g. a linear layer in an ANN is given as $Ax + b$, where $A \in \mathbb{R}^{n \times n}$, and $x, b \in \mathbb{R}^n$)
- Incorporate node/edge/hyper-edge vector-valued features

Invariant and Equivariant Linear Operators

What is the set of all invariant and equivariant linear operators?
Paper Sketch:

- Determine conditions for invariant/equivariant linear operators: Fixed-point Equations
- Identify a basis for solutions to the Fixed-point Equations
- Incorporate biases (e.g. a linear layer in an ANN is given as $Ax + b$, where $A \in \mathbb{R}^{n \times n}$, and $x, b \in \mathbb{R}^n$)
- Incorporate node/edge/hyper-edge vector-valued features
- Generalize to Mixed-order Equivariant layers and Multi-node sets

$$L_M : \mathbb{R}^{n^k} \to \mathbb{R}^{n^l}$$



Fixed-Point Equations (Invariant Layer)

Consider order-2 tensors with edge-value data (adjacency matrix) $\mathbf{A} = A \in \mathbb{R}^{n \times n} \equiv \mathbb{R}^{n^2}$
Operators of interest for invariance are $L_I : \mathbb{R}^{n^2} \to \mathbb{R}$, given as $\mathbf{L}_I \in \mathbb{R}^{1 \times n^2}$

L_I is order invariant iff $\mathbf{L}_I \, \mathrm{vec}(p \cdot \mathbf{A}) = \mathbf{L}_I \, \mathrm{vec}(\mathbf{P}^T \mathbf{A} \mathbf{P}) = \mathbf{L}_I \, \mathrm{vec}(\mathbf{A})$

Fixed-Point Equations (Invariant Layer)

Consider order-2 tensors with edge-value data (adjacency matrix) $\mathbf{A} = A \in \mathbb{R}^{n \times n} \equiv \mathbb{R}^{n^2}$

Operators of interest for invariance are $L_I : \mathbb{R}^{n^2} \to \mathbb{R}$, given as $\mathbf{L}_I \in \mathbb{R}^{1 \times n^2}$

L_I is order invariant iff $\mathbf{L}_I \,\mathrm{vec}(p \cdot \mathbf{A}) = \mathbf{L}_I \,\mathrm{vec}(\mathbf{P}^T \mathbf{A} \mathbf{P}) = \mathbf{L}_I \,\mathrm{vec}(\mathbf{A})$

Property of Kronecker Product

$\mathrm{vec}(\mathbf{X}\mathbf{A}\mathbf{Y}) = \mathbf{Y}^T \otimes \mathbf{X} \,\mathrm{vec}(\mathbf{A})$

$\mathbf{L}_I \mathbf{P}^T \otimes \mathbf{P}^T \,\mathrm{vec}(\mathbf{A}) = \mathbf{L}_I \,\mathrm{vec}(\mathbf{A})$

Fixed-Point Equations (Invariant Layer)

Consider order-2 tensors with edge-value data (adjacency matrix) $\mathbf{A} = A \in \mathbb{R}^{n \times n} \equiv \mathbb{R}^{n^2}$

Operators of interest for invariance are $L_I : \mathbb{R}^{n^2} \rightarrow \mathbb{R}$, given as $\mathbf{L}_I \in \mathbb{R}^{1 \times n^2}$

L_I is order invariant iff $\mathbf{L}_I \operatorname{vec}(p \cdot \mathbf{A}) = \mathbf{L}_I \operatorname{vec}(\mathbf{P}^T \mathbf{A} \mathbf{P}) = \mathbf{L}_I \operatorname{vec}(\mathbf{A})$

Property of Kronecker Product

$\operatorname{vec}(\mathbf{XAY}) = \mathbf{Y}^T \otimes \mathbf{X} \operatorname{vec}(\mathbf{A})$

$\mathbf{L}_I \mathbf{P}^T \otimes \mathbf{P}^T \operatorname{vec}(\mathbf{A}) = \mathbf{L}_I \operatorname{vec}(\mathbf{A})$

Noting that $\mathbf{L}_I^T = \operatorname{vec}(\mathbf{L}_I)$, and transposing we have,

Fixed Point Equation (Invariance) for order-2 tensors

$\mathbf{P} \otimes \mathbf{P} \operatorname{vec}(\mathbf{L}_I) = \operatorname{vec}(\mathbf{L}_I)$

Fixed-Point Equations (Equivariant Layer)

Consider order-2 tensors with edge-value data (adjacency matrix) $\mathbf{A} = A \in \mathbb{R}^{n \times n} \equiv \mathbb{R}^{n^2}$
Operators of interest for equivariance are $L_E : \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$, given as $\mathbf{L}_E \in \mathbb{R}^{n^2 \times n^2}$

L_E is order equivariant iff
$[\mathbf{L}_E \operatorname{vec}(p \cdot \mathbf{A})] = [\mathbf{L}_E \operatorname{vec}(\mathbf{P}^T \mathbf{A} \mathbf{P})] = \mathbf{P}^T [\mathbf{L}_E \operatorname{vec}(\mathbf{A})] \mathbf{P}$

Fixed-Point Equations (Equivariant Layer)

Consider order-2 tensors with edge-value data (adjacency matrix)
$\mathbf{A} = A \in \mathbb{R}^{n \times n} \equiv \mathbb{R}^{n^2}$
Operators of interest for equivariance are $L_E : \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$, given as
$\mathbf{L}_E \in \mathbb{R}^{n^2 \times n^2}$

L_E is order equivariant iff
$[\mathbf{L}_E \, \mathrm{vec}(p \cdot \mathbf{A})] = [\mathbf{L}_E \, \mathrm{vec}(\mathbf{P}^T \mathbf{A} \mathbf{P})] = \mathbf{P}^T [\mathbf{L}_E \, \mathrm{vec}(\mathbf{A})] \mathbf{P}$

Using properties of the Kronecker product we get,

Fixed-Point Equation (Equivariance) for order-2 tensors

$$\mathbf{P} \otimes \mathbf{P} \otimes \mathbf{P} \otimes \mathbf{P} \, \mathrm{vec}(\mathbf{L}_E) = \mathrm{vec}(\mathbf{L}_E)$$

(Note: $\mathbf{P} \in \mathbb{R}^{n^2}$ and $\mathbf{P} \otimes \mathbf{P} \in \mathbb{R}^{n^2 \times n^2}$)

Fixed Point Equations (Invariant and Equivariant)

In general we have,

(Note, $\mathbf{P}^{\otimes \ell} = \overbrace{\mathbf{P} \otimes \cdots \otimes \mathbf{P}}^{\ell}$)

Invariant $\mathbf{L}_I : \mathbf{P}^{\otimes k}\,\mathrm{vec}(\mathbf{L}_I) = \mathrm{vec}(\mathbf{L}_I)$
Equivariant $\mathbf{L}_E : \mathbf{P}^{\otimes 2k}\,\mathrm{vec}(\mathbf{L}_E) = \mathrm{vec}(\mathbf{L}_E)$

Fixed Point Equations (Invariant and Equivariant)

In general we have,

$$\text{Invariant } \mathbf{L}_I : \mathbf{P}^{\otimes k}\,\mathrm{vec}(\mathbf{L}_I) = \mathrm{vec}(\mathbf{L}_I) \quad \text{(Note, } \mathbf{P}^{\otimes \ell} = \overbrace{\mathbf{P} \otimes \cdots \otimes \mathbf{P}}^{\ell}\text{)}$$

$$\text{Equivariant } \mathbf{L}_E : \mathbf{P}^{\otimes 2k}\,\mathrm{vec}(\mathbf{L}_E) = \mathrm{vec}(\mathbf{L}_E)$$

Key Identity

$$\mathbf{P}^{\otimes \ell}\,\mathrm{vec}(\mathbf{L}) = \mathrm{vec}(p \cdot \mathbf{L})$$

Fixed Point Equations (Invariant and Equivariant)

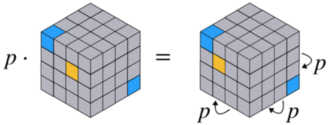
In general we have,

$$\text{(Note, } \mathbf{P}^{\otimes \ell} = \overbrace{\mathbf{P} \otimes \cdots \otimes \mathbf{P}}^{\ell}\text{)}$$

Invariant $\mathbf{L}_I : \mathbf{P}^{\otimes k}\,\mathrm{vec}(\mathbf{L}_I) = \mathrm{vec}(\mathbf{L}_I)$
Equivariant $\mathbf{L}_E : \mathbf{P}^{\otimes 2k}\,\mathrm{vec}(\mathbf{L}_E) = \mathrm{vec}(\mathbf{L}_E)$

Key Identity

$$\mathbf{P}^{\otimes \ell}\,\mathrm{vec}(\mathbf{L}) = \mathrm{vec}(p \cdot \mathbf{L})$$



Fixed Point Equations (Invariant and Equivariant)

In general we have,

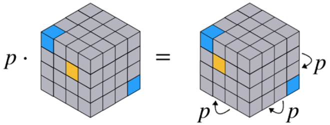
(Note, $\mathbf{P}^{\otimes \ell} = \overbrace{\mathbf{P} \otimes \cdots \otimes \mathbf{P}}^{\ell}$)

Invariant $\mathbf{L}_I : \mathbf{P}^{\otimes k} \operatorname{vec}(\mathbf{L}_I) = \operatorname{vec}(\mathbf{L}_I)$

Equivariant $\mathbf{L}_E : \mathbf{P}^{\otimes 2k} \operatorname{vec}(\mathbf{L}_E) = \operatorname{vec}(\mathbf{L}_E)$

Key Identity

$$\mathbf{P}^{\otimes \ell} \operatorname{vec}(\mathbf{L}) = \operatorname{vec}(p \cdot \mathbf{L})$$



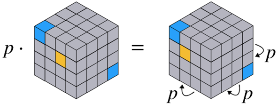
Question: What are the fixed points under the action $\mathbf{L} \to p \cdot \mathbf{L}$?

Basis of the Solution Space to the Fixed-Point Equations

Question: What are the fixed points under the action $\mathbf{L} \to p \cdot \mathbf{L}$?

Consider $L_E : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, $\mathbf{L}_E \in \mathbb{R}^{n^2\times n^2}$, $\mathbb{V} = \{1, 2, ..., n\}$
Permutation $p(i,j) : v_i \leftrightarrow v_j$

On the diagonal, require $\mathbf{L}_{E,(i,i,i,i)} = a \; \forall i$

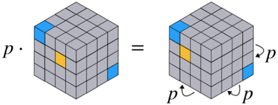


Basis of the Solution Space to the Fixed-Point Equations

Question: What are the fixed points under the action $\mathbf{L} \to p \cdot \mathbf{L}$?

Consider $L_E : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, $\mathbf{L}_E \in \mathbb{R}^{n^2\times n^2}$, $\mathbb{V} = \{1, 2, ..., n\}$
Permutation $p(i,j) : v_i \leftrightarrow v_j$

On the diagonal, require $\mathbf{L}_{E,(i,i,i,i)} = a \ \forall i$



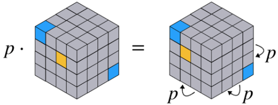
For the off-diagonal elements,
Ex. Consider $\mathbf{L}_{E,(i,i,j,s)}$ with $i \neq j \neq s$, require $\mathbf{L}_{E,(i,i,j,s)} = b \ \forall i, j, s$

Basis of the Solution Space to the Fixed-Point Equations

Question: What are the fixed points under the action $\mathbf{L} \to p \cdot \mathbf{L}$?

Consider $L_E : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$, $\mathbf{L}_E \in \mathbb{R}^{n^2 \times n^2}$, $\mathbb{V} = \{1, 2, ..., n\}$
Permutation $p(i, j) : v_i \leftrightarrow v_j$

On the diagonal, require $\mathbf{L}_{E,(i,i,i,i)} = a \; \forall i$



For the off-diagonal elements,
Ex. Consider $\mathbf{L}_{E,(i,i,j,s)}$ with $i \neq j \neq s$, require $\mathbf{L}_{E,(i,i,j,s)} = b \; \forall i, j, s$

In general require $\mathbf{L}_{index} = \mathbf{L}_{index'}$, where *index* and *index*$'$ have the same equality pattern.

Aside: Bell Numbers and Partitions

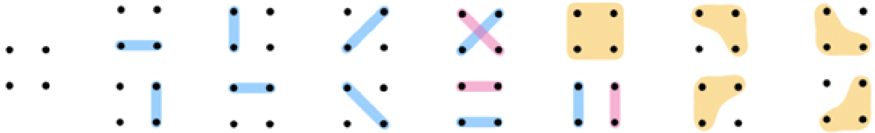
We had $L_E : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$, $\mathbf{L_E} \in \mathbb{R}^{n^2 \times n^2} \equiv \mathbb{R}^{n^4}$,
$\mathbf{L_E}$ is indexed by a tuple of size 4, (i, j, s, t)

Question: How many ways are there to partition sets of size 4 (and generally of size ℓ)?

Aside: Bell Numbers and Partitions

We had $L_E : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$, $\mathbf{L_E} \in \mathbb{R}^{n^2 \times n^2} \equiv \mathbb{R}^{n^4}$,
$\mathbf{L_E}$ is indexed by a tuple of size 4, (i, j, s, t)

Question: How many ways are there to partition sets of size 4 (and generally of size ℓ)?

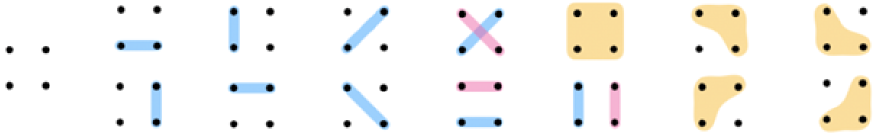


Aside: Bell Numbers and Partitions

We had $L_E : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$, $\mathbf{L_E} \in \mathbb{R}^{n^2 \times n^2} \equiv \mathbb{R}^{n^4}$,
$\mathbf{L_E}$ is indexed by a tuple of size 4, (i, j, s, t)

Question: How many ways are there to partition sets of size 4 (and generally of size ℓ)?

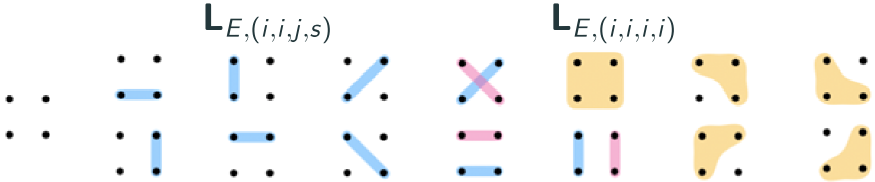
$\mathbf{L}_{E,(i,i,j,s)}$ $\mathbf{L}_{E,(i,i,i,i)}$



Aside: Bell Numbers and Partitions

We had $L_E : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$, $\mathbf{L_E} \in \mathbb{R}^{n^2\times n^2} \equiv \mathbb{R}^{n^4}$,
$\mathbf{L_E}$ is indexed by a tuple of size 4, (i, j, s, t)

Question: How many ways are there to partition sets of size 4 (and generally of size ℓ)?



In general there are bell(ℓ) equality patterns
Grows combinatorially with ℓ

Basis of the Solution Space to the Fixed-Point Equations

Question: What are the fixed points under the action $\mathbf{L} \to p \cdot \mathbf{L}$? (for $\mathbf{L} \in \mathbb{R}^{n^\ell}$)

Ans: There are bell(ℓ) fixed points.

Basis of the Solution Space to the Fixed-Point Equations

Question: What are the fixed points under the action $\mathbf{L} \rightarrow p \cdot \mathbf{L}$? (for $\mathbf{L} \in \mathbb{R}^{n^{\ell}}$)

Ans: There are bell(ℓ) fixed points.

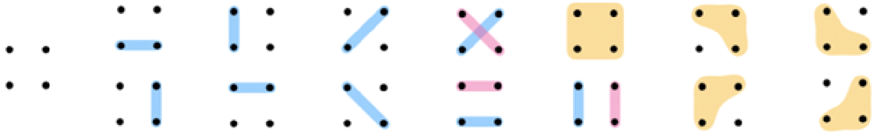
Given an equality pattern, define the indicator tensor, $\mathbf{B}^{\alpha}_{i,j,s,t} = 1$ iff $(i, j, s, t) \in \alpha$ and 0 otherwise.

Basis of the Solution Space to the Fixed-Point Equations

Question: What are the fixed points under the action $\mathbf{L} \to p \cdot \mathbf{L}$? (for $\mathbf{L} \in \mathbb{R}^{n^\ell}$)

Ans: There are bell(ℓ) fixed points.

Given an equality pattern, define the indicator tensor, $\mathbf{B}^{\alpha}_{i,j,s,t} = 1$ iff $(i,j,s,t) \in \alpha$ and 0 otherwise.



e.g. $\mathbf{B}^{\alpha_1}_{i,j,s,t} = 1$ for equality pattern $\alpha_1 = \{(i,i,i,i)\}$, and $\mathbf{B}^{\alpha_2}_{i,j,s,t} = 1$ for equality pattern $\alpha_2 = \{(i,i,j,s)\}$

Basis of the Solution Space to the Fixed-Point Equations

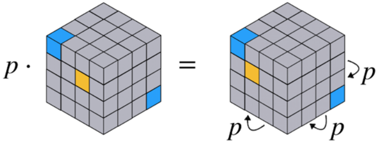
So far,

Fixed Point Equation (Invariance)

$$\mathbf{P}^{\otimes k}\,\mathrm{vec}(\mathbf{L}_I) = \mathrm{vec}(\mathbf{L}_I)$$

Fixed-Point Equation (Equivariance)

$$\mathbf{P}^{\otimes 2k}\,\mathrm{vec}(\mathbf{L}_E) = \mathrm{vec}(\mathbf{L}_E)$$



Basis of the Solution Space to the Fixed-Point Equations

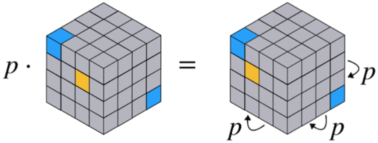
So far,

Fixed Point Equation (Invariance)

$$\mathbf{P}^{\otimes k}\, \text{vec}(\mathbf{L}_I) = \text{vec}(\mathbf{L}_I)$$

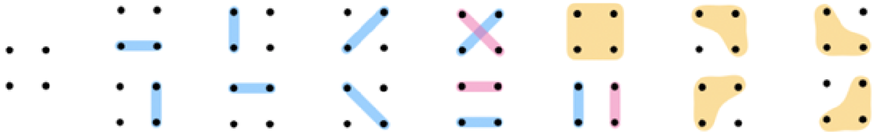
Fixed-Point Equation (Equivariance)

$$\mathbf{P}^{\otimes 2k}\, \text{vec}(\mathbf{L}_E) = \text{vec}(\mathbf{L}_E)$$



Basis

$\mathbf{B}^{\alpha}_{i,j,s,t} = 1$ iff $(i, j, s, t) \in \alpha$
$\mathbf{B}^{\alpha}_{i,j,s,t}$ is a complete orthogonal basis
bell(k) bases for $\mathbf{L}_I$ and bell($2k$) bases for $\mathbf{L}_E$



Examples

Invariant operators on $\mathbf{a} \in \mathbb{R}^n$,
e.g. $L_I : \mathbb{R}^n \to \mathbb{R}$, $\qquad \mathbf{L}_I \in \mathbb{R}^{1\times n}$

We have the sum operator $L(\mathbf{a}) = \gamma \mathbf{1}^T \mathbf{a}$
$\mathbf{B}_i^\alpha = \mathbf{1}^T$

Examples

Invariant operators on $\mathbf{a} \in \mathbb{R}^n$,
e.g. $L_I : \mathbb{R}^n \to \mathbb{R}$, $\mathbf{L}_I \in \mathbb{R}^{1 \times n}$

We have the sum operator $L(\mathbf{a}) = \gamma \mathbf{1}^T \mathbf{a}$
$\mathbf{B}_i^{\alpha} = \mathbf{1}^T$

Equivariant operators on $\mathbf{a} \in \mathbb{R}^n$,
e.g. $L_E : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{L}_E \in \mathbb{R}^{n \times n}$
$\mathbf{B}_{ij}^{\alpha_1} = \mathbf{I}$ for equality pattern $\alpha_1 = \{(i,i)\}$
$\mathbf{B}_{ij}^{\alpha_2} = \mathbf{1}\mathbf{1}^T - \mathbf{I}$ for equality pattern $\alpha_2 = \{(i,j)\}, i \neq j$

Extensions: Incorporating bias terms

Incorporating bias terms:

- For invariant layers ($\mathbf{L}_I \in \mathbb{R}^{1\times n^k}$) use $c \in \mathbb{R}$, i.e. $\mathbf{L}_I\mathbf{a} + c$
- For equivariant layers ($\mathbf{L}_E \in \mathbb{R}^{n^k\times n^k}$), use $\mathbf{B}^\beta$ where $|\beta| = \text{bell}(k)$

Extensions: Incorporating bias terms

Incorporating bias terms:

- For invariant layers ($\mathbf{L}_I \in \mathbb{R}^{1\times n^k}$) use $c \in \mathbb{R}$, i.e. $\mathbf{L}_I\mathbf{a} + c$
- For equivariant layers ($\mathbf{L}_E \in \mathbb{R}^{n^k\times n^k}$), use $\mathbf{B}^\beta$ where $|\beta| = \text{bell}(k)$

e.g. for $L : \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$, $\mathbf{L}_E \in \mathbb{R}^{n^2\times n^2} \equiv \mathbb{R}^{n^4}$,
we have $\mathbf{La} + \mathbf{B}^\beta$, where $k = 4$ and $|\beta| = 2$

Extensions: Incorporating bias terms

Incorporating bias terms:

- For invariant layers ($\mathbf{L}_I \in \mathbb{R}^{1\times n^k}$) use $c \in \mathbb{R}$, i.e. $\mathbf{L}_I\mathbf{a} + c$
- For equivariant layers ($\mathbf{L}_E \in \mathbb{R}^{n^k\times n^k}$), use $\mathbf{B}^\beta$ where $|\beta| = \text{bell}(k)$

e.g. for $L : \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$, $\mathbf{L}_E \in \mathbb{R}^{n^2\times n^2} \equiv \mathbb{R}^{n^4}$,
we have $\mathbf{La} + \mathbf{B}^\beta$, where $k = 4$ and $|\beta| = 2$

Figure 1: The full basis for equivariant linear layers for edge-value data $\mathbf{A} \in \mathbb{R}^{n\times n}$, for $n = 5$. The purely linear 15 basis elements, $\mathbf{B}^\mu$, are represented by matrices $n^2 \times n^2$, and the 2 bias basis elements (right), $\mathbf{C}^\lambda$, by matrices $n \times n$, see equation 9.

Note: $\text{bell}(4) = 15$, $\text{bell}(2) = 2$

Extensions: Multi-order Equivariance, Multi-node sets

Straightforward generalization of **equivariant** operators
$L_E : \mathbb{R}^{n^k} \to \mathbb{R}^{n^k}$
to **mixed-order equivariant** operators
$L_M : \mathbb{R}^{n^k} \to \mathbb{R}^{n^l}$

Generalization to multi-node sets

Extensions: Multi-order Equivariance, Multi-node sets

Straightforward generalization of **equivariant** operators
$L_E : \mathbb{R}^{n^k} \to \mathbb{R}^{n^k}$
to **mixed-order equivariant** operators
$L_M : \mathbb{R}^{n^k} \to \mathbb{R}^{n^l}$

Generalization to multi-node sets

Tuples of nodes on subsets of the nodes

Invariance:

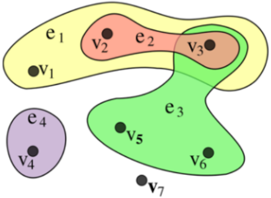
$L_I : \mathbb{R}^{n_1^{k_1} \times n_2^{k_2} \times \cdots \times n_m^{k_m}} \to \mathbb{R}$

dimension $\prod_{i=1}^m \text{bell}(k_i)$

Equivariance:

$L_E : \mathbb{R}^{n_1^{k_1} \times n_2^{k_2} \times \cdots \times n_m^{k_m}} \to \mathbb{R}^{n_1^{l_1} \times n_2^{l_2} \times \cdots \times n_m^{l_m}}$

dimension $\prod_{i=1}^m \text{bell}(k_i + l_i)$



Extensions: Vector-valued features

When we have vector-valued features instead of scalars on node tuples:

Invariance $L_I : \mathbb{R}^{n^k \times d} \to \mathbb{R}^{1 \times d'}$

dimension $dd' \,\mathrm{bell}(k) + d'$

Equivariance $L_E : \mathbb{R}^{n^k \times d} \to \mathbb{R}^{n^k \times d'}$

dimension $dd' \,\mathrm{bell}(2k) + d' \,\mathrm{bell}(k)$

Message Passing

Step 1: Compute Messages

$$m_u^{t+1} = \sum_{v \in N(u)} M_t(h_u^t, h_v^t, e_{uv})$$

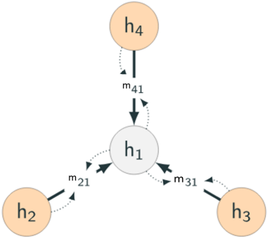
Step 2: Update feature vectors

$$h_u^{t+1} = U_t(h_u^t, m_u^{t+1})$$

$$\mathbf{H} = (h_u) \in \mathbb{R}^{n \times d}$$

$$\mathbf{A} = (a_{uv}) \in \mathbb{R}^{n \times n}$$

$$\mathbf{E} = (e_{uv}) \in \mathbb{R}^{(n \times n) \times l}$$

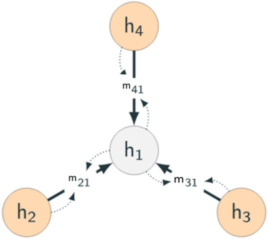


Connection to Message-Passing

$$\mathbf{H} = (h_u) \in \mathbb{R}^{n \times d}$$

$$\mathbf{A} = (a_{uv}) \in \mathbb{R}^{n \times n}$$

$$\mathbf{E} = (e_{uv}) \in \mathbb{R}^{(n \times n) \times l}$$



Message Passing

Step 1: Compute Messages

$$m_u^{t+1} = \sum_{v \in N(u)} M_t(h_u^t, h_v^t, e_{uv})$$

Step 2: Update feature vectors

$$h_u^{t+1} = U_t(h_u^t, m_u^{t+1})$$

New Formulation

Input data: $\mathbf{Y} \in \mathbb{R}^{(n \times n) \times (1+l+d)}$

Theorem 4: The proposed model can represent message passing layers to an arbitrary precision on compact sets.

Idea: Combine Linear operators (tensors) with MLPs to mimic multiplication of features by the adjacency matrix to allow summing over local neighbourhoods.

Synthetic Experiments

Table 1: Comparison to baseline methods on synthetic experiments.

	Symmetric projection			Diagonal extraction			Max singular vector				Trace		
# Layers	1	2	3	1	2	3	1	2	3	4	1	2	3
Trivial predictor	4.17	4.17	4.17	0.21	0.21	0.21	0.025	0.025	0.025	0.025	333.33	333.33	333.33
Hartford et al.	2.09	2.09	2.09	0.81	0.81	0.81	0.043	0.044	0.043	0.043	316.22	311.55	307.97
Ours	**1E-05**	**7E-06**	**2E-05**	**8E-06**	**7E-06**	**1E-04**	**0.015**	**0.0084**	**0.0054**	**0.0016**	**0.005**	**0.001**	**0.003**

Graph Classification

Table 3: Graph Classification Results.

dataset	MUTAG	PTC	PROTEINS	NCI1	NCI109	COLLAB	IMDB-B	IMDB-M
size	188	344	1113	4110	4127	5000	1000	1500
classes	2	2	2	2	2	3	2	3
avg node #	17.9	25.5	39.1	29.8	29.6	74.4	19.7	13
Results								
DGCNN	85.83±1.7	58.59±2.5	75.54±0.9	74.44±0.5	NA	73.76±0.5	70.03±0.9	47.83±0.9
PSCN (k=10)	88.95±4.4	62.29±5.7	75±2.5	76.34±1.7	NA	72.6±2.2	71±2.3	45.23±2.8
DCNN	NA	NA	61.29±1.6	56.61± 1.0	NA	52.11±0.7	49.06±1.4	33.49±1.4
ECC	76.11	NA	NA	76.82	75.03	NA	NA	NA
DGK	87.44±2.7	60.08±2.6	75.68±0.5	80.31±0.5	80.32±0.3	73.09±0.3	66.96±0.6	44.55±0.5
DiffPool	NA	NA	78.1	NA	NA	75.5	NA	NA
CCN	91.64±7.2	70.62±7.0	NA	76.27±4.1	75.54±3.4	NA	NA	NA
GK	81.39±1.7	55.65±0.5	71.39±0.3	62.49±0.3	62.35±0.3	NA	NA	NA
RW	79.17±2.1	55.91±0.3	59.57±0.1	> 3 days	NA	NA	NA	NA
PK	76±2.7	59.5±2.4	73.68±0.7	82.54±0.5	NA	NA	NA	NA
WL	84.11±1.9	57.97±2.5	74.68±0.5	84.46±0.5	85.12±0.3	NA	NA	NA
FGSD	92.12	62.80	73.42	79.80	78.84	80.02	73.62	52.41
AWE-DD	NA	NA	NA	NA	NA	73.93±1.9	74.45 ±5.8	51.54 ± 3.6
AWE-FB	87.87±9.7	NA	NA	NA	NA	70.99 ± 1.4	73.13 ±3.2	51.58 ± 4.6
ours	84.61±10	59.47±7.3	75.19±4.3	73.71±2.6	72.48±2.5	77.92±1.7	71.27±4.5	48.55±3.9