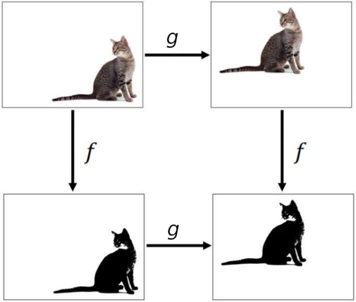
An introduction to groups, actions, and equivariance

Rob Cornish

Department of Statistics, University of Oxford

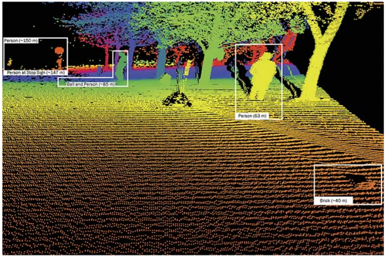
February 13, 2025

Motivation: “symmetry”



https://www.doc.ic.ac.uk/~bkainz/teaching/DL/notes/equivariance.pdf

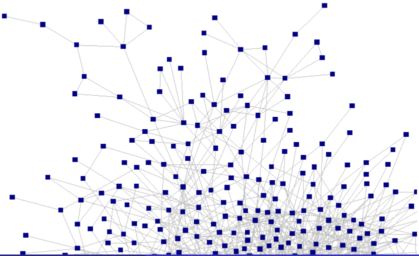
Many other examples



photonics.com

alphafold.ebi.ac.uk

orgnet.com



Question

Today's question: how can we talk precisely about such properties?

Question

Today's question: how can we talk precisely about such properties?

Three ingredients:

- **Groups** ("abstract symmetries")

Question

Today's question: how can we talk precisely about such properties?

Three ingredients:

- **Groups** ("abstract symmetries")
- **Actions** ("actual symmetries")

Question

Today's question: how can we talk precisely about such properties?

Three ingredients:

- **Groups** ("abstract symmetries")
- **Actions** ("actual symmetries")
- **Equivariance** ("mappings that respect symmetries")

Question

Today's question: how can we talk precisely about such properties?

Three ingredients:

- **Groups** ("abstract symmetries")
- **Actions** ("actual symmetries")
- **Equivariance** ("mappings that respect symmetries")

Intuitively, in terms of our example:

- Group actions describe things like "translate the cat"

Question

Today's question: how can we talk precisely about such properties?

Three ingredients:

- **Groups** ("abstract symmetries")
- **Actions** ("actual symmetries")
- **Equivariance** ("mappings that respect symmetries")

Intuitively, in terms of our example:

- Group actions describe things like "translate the cat"
- Equivariance says that the network "respects this translation"

Groups: intuition

Slogan: groups are an abstract way to talk about "symmetries"

Groups: intuition

Slogan: groups are an abstract way to talk about "symmetries"

But what *are* "symmetries"?

Groups: intuition

Slogan: groups are an abstract way to talk about "symmetries"

But what *are* "symmetries"?

Rough answer: a "symmetry" is an information-preserving transformation

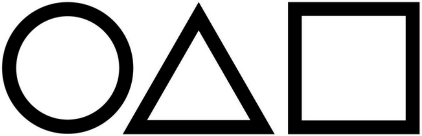
Groups: intuition

Slogan: groups are an abstract way to talk about "symmetries"

But what *are* "symmetries"?

Rough answer: a "symmetry" is an information-preserving transformation

Warning: many people use "symmetry" more specifically to mean a transformation that leaves an object unchanged



Towards a theory of symmetry

An easier question: how do "symmetries" behave?

Towards a theory of symmetry

An easier question: how do "symmetries" behave?

In many cases, the following seems reasonable to assume:

- We can always "combine" two symmetries to obtain another one

Towards a theory of symmetry

An easier question: how do "symmetries" behave?

In many cases, the following seems reasonable to assume:

- We can always "combine" two symmetries to obtain another one
- "Doing nothing" is always a symmetry

Towards a theory of symmetry

An easier question: how do "symmetries" behave?

In many cases, the following seems reasonable to assume:

- We can always "combine" two symmetries to obtain another one
- "Doing nothing" is always a symmetry
- Every symmetry gives rise to another that "undoes" it

Towards a theory of symmetry

An easier question: how do “symmetries” behave?

In many cases, the following seems reasonable to assume:

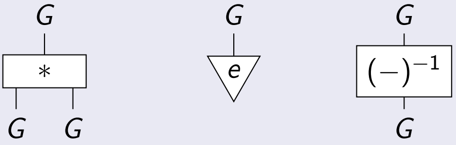
- We can always “combine” two symmetries to obtain another one
- “Doing nothing” is always a symmetry
- Every symmetry gives rise to another that “undoes” it

Basic idea: groups model collections of things that behave like this

Groups: formal definition

Definition

A group is a set G equipped with operations

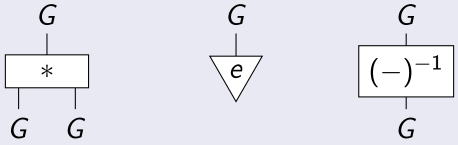


that satisfy some axioms (to come)

Groups: formal definition

Definition

A group is a set G equipped with operations

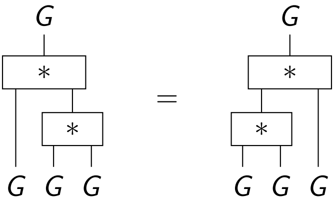


that satisfy some axioms (to come)

Refer to these as multiplication, unit, and inversion

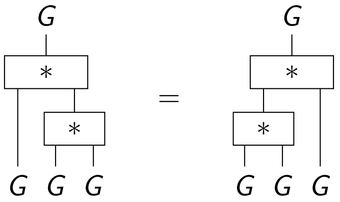
Associativity

Group multiplication must be associative



Associativity

Group multiplication must be associative

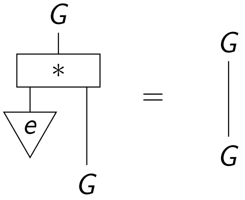


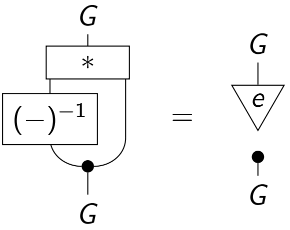
In classical notation, this just says:

$$g(hn) = (gh)n \qquad \text{for all } g, h, n \in G$$

Other group axioms

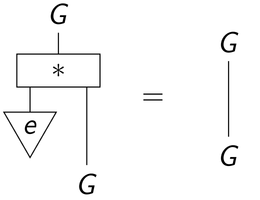
Group multiplication must also be unital and admit inverses

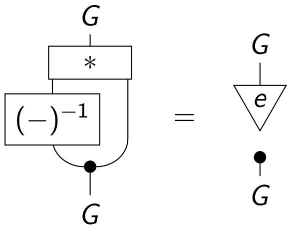




Other group axioms

Group multiplication must also be unital and admit inverses





In classical notation, this says

$$eg = g \qquad \text{and} \qquad g^{-1}g = e \qquad\qquad \text{for all } g \in G$$

Examples of groups

There are many examples of groups. For GDL, the following are especially relevant:

Examples of groups

There are many examples of groups. For GDL, the following are especially relevant:

- The translation group $\mathbb{T}_d$ of translations of $\mathbb{R}^d$

Examples of groups

There are many examples of groups. For GDL, the following are especially relevant:

- The translation group $\mathbb{T}_d$ of translations of $\mathbb{R}^d$
- The symmetric group S_n of permutations of $\{1, \dots, n\}$

Examples of groups

There are many examples of groups. For GDL, the following are especially relevant:

- The translation group $\mathbb{T}_d$ of translations of $\mathbb{R}^d$
- The symmetric group S_n of permutations of $\{1, \ldots, n\}$
- The special orthogonal group $\mathrm{SO}(d)$ of rotations of $\mathbb{R}^d$

Examples of groups

There are many examples of groups. For GDL, the following are especially relevant:

- The translation group $\mathbb{T}_d$ of translations of $\mathbb{R}^d$
- The symmetric group S_n of permutations of $\{1, \ldots, n\}$
- The special orthogonal group $\mathrm{SO}(d)$ of rotations of $\mathbb{R}^d$
- The orthogonal group $\mathrm{O}(d)$ of combinations of rotations and reflections of $\mathbb{R}^d$

Examples of groups

There are many examples of groups. For GDL, the following are especially relevant:

- The translation group $\mathbb{T}_d$ of translations of $\mathbb{R}^d$
- The symmetric group S_n of permutations of $\{1, \ldots, n\}$
- The special orthogonal group $\mathrm{SO}(d)$ of rotations of $\mathbb{R}^d$
- The orthogonal group $\mathrm{O}(d)$ of combinations of rotations and reflections of $\mathbb{R}^d$
- The special Euclidean group $\mathrm{SE}(d)$ of rotations and translations of $\mathbb{R}^d$

Examples of groups

There are many examples of groups. For GDL, the following are especially relevant:

- The translation group $\mathbb{T}_d$ of translations of $\mathbb{R}^d$
- The symmetric group S_n of permutations of $\{1, \ldots, n\}$
- The special orthogonal group $\mathrm{SO}(d)$ of rotations of $\mathbb{R}^d$
- The orthogonal group $\mathrm{O}(d)$ of combinations of rotations and reflections of $\mathbb{R}^d$
- The special Euclidean group $\mathrm{SE}(d)$ of rotations and translations of $\mathbb{R}^d$
- The Euclidean group $\mathrm{E}(d)$ of all combinations of rotations, reflections, and translations of $\mathbb{R}^d$
- The general linear group $\mathrm{GL}(d, \mathbb{R})$ of all invertible linear transformations of $\mathbb{R}^d$

Examples of groups

There are many examples of groups. For GDL, the following are especially relevant:

- The translation group $\mathbb{T}_d$ of translations of $\mathbb{R}^d$
- The symmetric group S_n of permutations of $\{1, \dots, n\}$
- The special orthogonal group $\mathrm{SO}(d)$ of rotations of $\mathbb{R}^d$
- The orthogonal group $\mathrm{O}(d)$ of combinations of rotations and reflections of $\mathbb{R}^d$
- The special Euclidean group $\mathrm{SE}(d)$ of rotations and translations of $\mathbb{R}^d$
- The Euclidean group $\mathrm{E}(d)$ of all combinations of rotations, reflections, and translations of $\mathbb{R}^d$
- The general linear group $\mathrm{GL}(d, \mathbb{R})$ of all invertible linear transformations of $\mathbb{R}^d$

Representation theory gives descriptions of these in terms of matrices, e.g.:

$$\mathrm{O}(d) \cong \{Q \in \mathbb{R}^{d \times d} \mid QQ^T = I\}$$

Actions

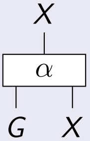
Key idea: whereas groups model abstract symmetries, group actions model concrete symmetries of a particular object X

Actions

Key idea: whereas groups model abstract symmetries, group actions model concrete symmetries of a particular object X

Definition

An action of a group G on a set X is a function of the form



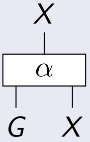
that satisfies some axioms (to come next)

Actions

Key idea: whereas groups model abstract symmetries, group actions model concrete symmetries of a particular object X

Definition

An action of a group G on a set X is a function of the form



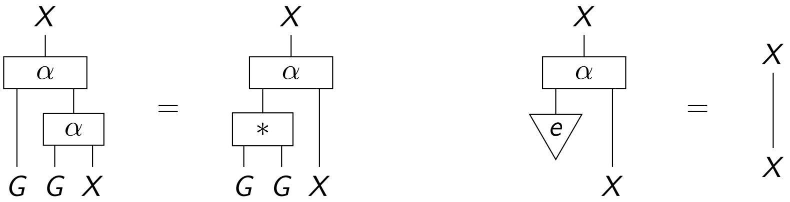
that satisfies some axioms (to come next)

In classical notation, actions are often written as

$$\alpha(g, x) = g \cdot x \qquad \text{where } g \in G \text{ and } x \in X$$

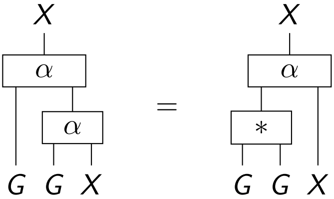
Action axioms

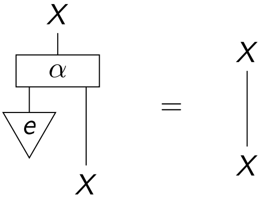
To be an action, α must satisfy the following:



Action axioms

To be an action, α must satisfy the following:



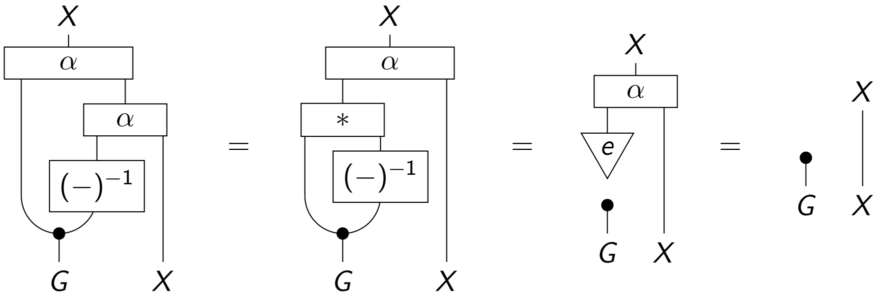


In classical notation, this says

$g \cdot (h \cdot x) = (gh) \cdot x$ and $e \cdot x = x$ for all $g, h \in G$ and $x \in X$

Invertibility

One consequence of this definition is the following:



This says that group actions are always invertible

Examples of actions

With actions, we can formalise the examples we had earlier:

- $\mathbb{T}_2$ acts on 2D images by translation

Examples of actions

With actions, we can formalise the examples we had earlier:

- $\mathbb{T}_2$ acts on 2D images by translation
- S_n acts on sequences X^n by

$$\sigma \cdot (x_1, \ldots, x_n) = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

Examples of actions

With actions, we can formalise the examples we had earlier:

- $\mathbb{T}_2$ acts on 2D images by translation
- S_n acts on sequences X^n by

$$\sigma \cdot (x_1, \ldots, x_n) = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

- $\mathrm{SE}(3)$ acts on 3D point cloud by rotation followed by a translation:

$$(t, R) \cdot x = Rx + t$$

- And many others

Equivariance

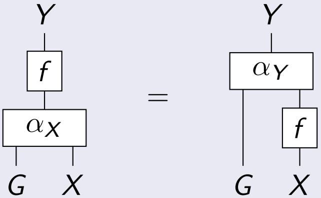
Key idea: Equivariant functions respect some choices of group actions

Equivariance

Key idea: Equivariant functions respect some choices of group actions

Definition

Given actions α_X and α_Y of a group G on some sets X and Y, a function $f : X \to Y$ is equivariant with respect to α_X and α_Y if

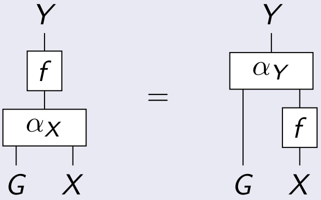


Equivariance

Key idea: Equivariant functions respect some choices of group actions

Definition

Given actions α_X and α_Y of a group G on some sets X and Y, a function $f : X \to Y$ is equivariant with respect to α_X and α_Y if

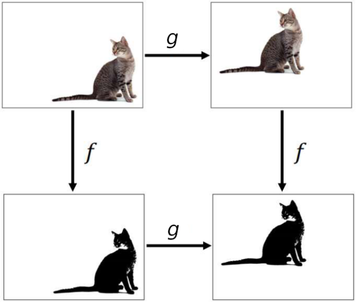


In classical notation, this says:

$$f(g \cdot x) = g \cdot f(x) \qquad \text{for all } g \in G \text{ and } x \in X$$

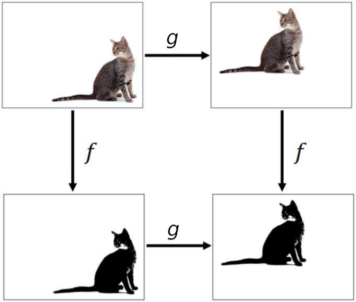
Examples of equivariance

We can now formalise the original example:



Examples of equivariance

We can now formalise the original example:



Here:

- G is $\mathbb{T}_2$, the group of 2D translations
- X is the set of colour images
- Y is the set of black-and-white images
- $\alpha_X(g, x)$ is the translation of x by g (with α_Y similar)

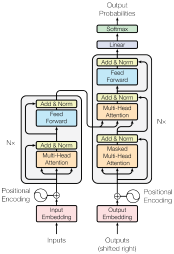


Figure 1: The Transformer - model architecture.

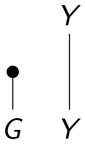
Attention is equivariant to the symmetric group: for $\sigma \in S_n$ we have

$$\begin{array}{ccc} (x_1, \ldots, x_n) & \stackrel{\sigma}{\longmapsto} & (x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \\ \downarrow \text{Attn.} & & \downarrow \text{Attn.} \\ (e_1, \ldots, e_n) & \stackrel{\sigma}{\longmapsto} & (e_{\sigma(1)}, \ldots, e_{\sigma(n)}) \end{array}$$

This constitutes a very elegant solution to catastrophic forgetting

Invariance

For every set Y, we can define the trivial action ε as

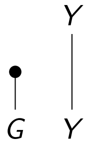


or in classical notation:

$$g \cdot x = x \qquad \text{for all } g \in G \text{ and } x \in X$$

Invariance

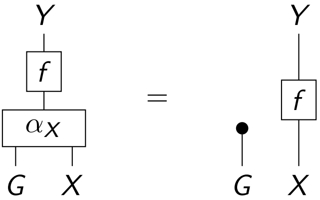
For every set Y, we can define the trivial action ε as



or in classical notation:

$$g \cdot x = x \qquad \text{for all } g \in G \text{ and } x \in X$$

Equivariance with respect to α_X and ε is called invariance:



Example of invariance

For processing sequences, with $X = \mathbb{R}^n$, often want:

$$f(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \qquad \text{for all permutations } \sigma$$

so f processes the sequence as if it were an unordered set

Example of invariance

For processing sequences, with $X = \mathbb{R}^n$, often want:

$$f(x_1, \dots, x_n) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) \qquad \text{for all permutations } \sigma$$

so f processes the sequence as if it were an unordered set

Can describe as invariance with respect to the following action of S_n:

$$\sigma \cdot (x_1, \dots, x_n) = (x_{\sigma(1)}, \dots, x_{\sigma(n)})$$

Example of invariance

For processing sequences, with $X = \mathbb{R}^n$, often want:

$$f(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \qquad \text{for all permutations } \sigma$$

so f processes the sequence as if it were an unordered set

Can describe as invariance with respect to the following action of S_n:

$$\sigma \cdot (x_1, \ldots, x_n) = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

DeepSets [Zaheer et al., 2017] is a well-known example of such an f

Enforcing equivariance

Fundamental problem of GDL

Suppose G is a group acting on X and Y. How can we parameterise a function $f : X \to Y$ that is equivariant with respect to these actions?

Enforcing equivariance

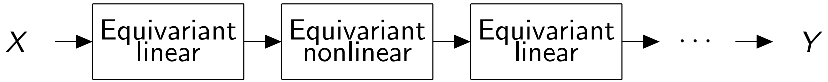
Fundamental problem of GDL

Suppose G is a group acting on X and Y. How can we parameterise a function $f : X \to Y$ that is equivariant with respect to these actions?

Two key approaches: intrinsic equivariance and symmetrisation

Intrinsic equivariance

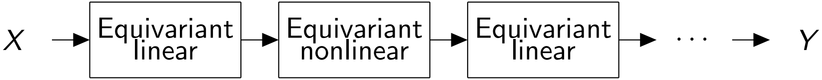
Overall model $f : X \to Y$ has form



where the individual layers are all equivariant via e.g. weight tying

Intrinsic equivariance

Overall model $f : X \to Y$ has form

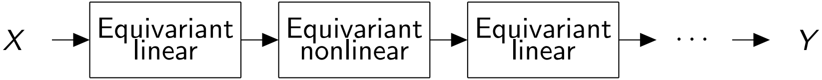


where the individual layers are all equivariant via e.g. weight tying

Many examples throughout GDL (CNNs, GNNs, etc.)

Intrinsic equivariance

Overall model $f : X \to Y$ has form



where the individual layers are all equivariant via e.g. weight tying

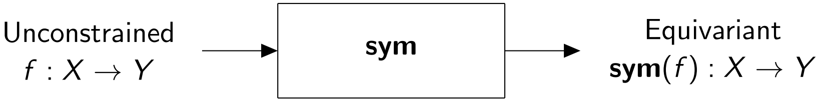
Many examples throughout GDL (CNNs, GNNs, etc.)

Works very well, but some caveats:

- Requires hand engineering for each case
- Can be somewhat brittle at scale (e.g. AlphaFold 2 vs. 3)

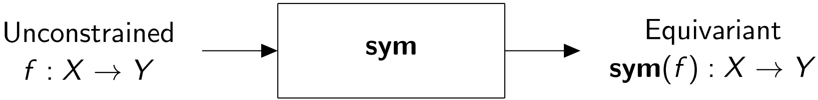
Symmetrisation

Recent interest instead in symmetrisation:



Symmetrisation

Recent interest instead in symmetrisation:



Here f is completely general and opaque

Symmetrisation: example

Early example is Janossy pooling [Murphy et al., 2019]: given

$$f : X^n \to \mathbb{R}^d,$$

the following function $X^n \to \mathbb{R}^n$ is always permutation invariant:

$$\frac{1}{n!} \sum_{\sigma \in S_n} f\big(x_{\sigma(1)}, \dots, x_{\sigma(n)}\big)$$

Symmetrisation: example

Early example is Janossy pooling [Murphy et al., 2019]: given

$$f : X^n \to \mathbb{R}^d,$$

the following function $X^n \to \mathbb{R}^n$ is always permutation invariant:

$$\frac{1}{n!} \sum_{\sigma \in S_n} f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

In turn, this generalises deep sets [Zaheer et al., 2017]

Symmetrisation: other examples

Other recent examples, given $f : X \to \mathbb{R}^d$ and a group G

$\frac{1}{\lvert\mathcal{F}(x)\rvert} \sum_{g \in \mathcal{F}(x)} g \cdot f(g^{-1} \cdot x)$	[Puny et al., 2022]
$h(x) \cdot f(h(x)^{-1} \cdot x)$	[Kaba et al., 2023]
$\mathbb{E}_{\boldsymbol{G} \sim p(g \mid x)}[\boldsymbol{G} \cdot f(\boldsymbol{G}^{-1} \cdot x)]$	[Kim et al., 2023]

Symmetrisation: other examples

Other recent examples, given $f : X \to \mathbb{R}^d$ and a group G

$\frac{1}{\lvert\mathcal{F}(x)\rvert} \sum_{g \in \mathcal{F}(x)} g \cdot f(g^{-1} \cdot x)$	[Puny et al., 2022]
$h(x) \cdot f(h(x)^{-1} \cdot x)$	[Kaba et al., 2023]
$\mathbb{E}_{\boldsymbol{G} \sim p(g \mid x)}[\boldsymbol{G} \cdot f(\boldsymbol{G}^{-1} \cdot x)]$	[Kim et al., 2023]

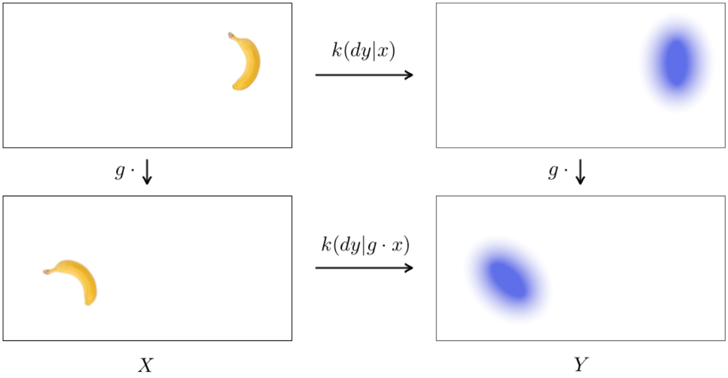
Under some conditions, each is equivariant in $x \in X$, even if f is arbitrarily complex

Stochastic equivariance: illustration

Equivariance can also be generalised to stochastic models

Stochastic equivariance: illustration

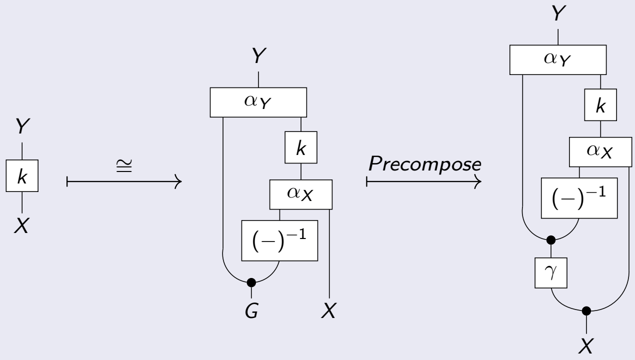
Equivariance can also be generalised to stochastic models



A general theory of symmetrisation

Theorem ([Cornish, 2024])

Given suitable $\gamma : X \to G$, can always symmetrise a general $k : X \to Y$ via:



Moreover, every (natural) symmetrisation procedure has this form.

Thank you!

References I

Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabas Poczos, Russ R Salakhutdinov, and Alexander J Smola. Deep sets. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL `https://proceedings.neurips.cc/paper_files/paper/2017/file/f22e4747da1aa27e363d86d40ff442fe-Paper.pdf`.

Ryan L. Murphy, Balasubramaniam Srinivasan, Vinayak Rao, and Bruno Ribeiro. Janossy pooling: Learning deep permutation-invariant functions for variable-size inputs. In *International Conference on Learning Representations*, 2019. URL `https://openreview.net/forum?id=BJluy2RcFm`.

Omri Puny, Matan Atzmon, Edward J. Smith, Ishan Misra, Aditya Grover, Heli Ben-Hamu, and Yaron Lipman. Frame averaging for invariant and equivariant network design. In *International Conference on Learning Representations*, 2022. URL `https://openreview.net/forum?id=zIUyj55nXR`.

References II

Sékou-Oumar Kaba, Arnab Kumar Mondal, Yan Zhang, Yoshua Bengio, and Siamak Ravanbakhsh. Equivariance with learned canonicalization functions. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pages 15546–15566. PMLR, 23–29 Jul 2023. URL https://proceedings.mlr.press/v202/kaba23a.html.

Jinwoo Kim, Dat Nguyen, Ayhan Suleymanzade, Hyeokjun An, and Seunghoon Hong. Learning probabilistic symmetrization for architecture agnostic equivariance. In A. Oh, T. Neumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, *Advances in Neural Information Processing Systems*, volume 36, pages 18582–18612. Curran Associates, Inc., 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/3b5c7c9c5c7bd77eb73d0baec7a07165-Paper-Conference.pdf.

Rob Cornish. Stochastic neural network symmetrisation in markov categories, 2024. URL https://arxiv.org/abs/2406.11814.