# Group Theory & ML

## 0. Summary of abstract

- [Aslan, Platt, Sheard, 2022]
  Group Invariant Machine Learning by Fundamental Domain Projections.

- Problem: We have a problem we're trying to solve that is invariant under some gp action. We are using some ML to solve the problem, so we want the ML output to be group invariant. This isn't often the case.

- Solution: project input data into a geometric space that parametrises the orbits of the symmetry group.

- They give an algorithm to compute the projection, and found an increase in accuracy compared to other algorithms.

ML

Problem          ML Solution

gp-invariant $\longrightarrow$ might not be gp-inv.
$\Rightarrow$ lower accuracy.

Use symmetry gp to fix
$\Rightarrow$ higher accuracy results

# 1. Introduction

- Many tasks in machine learning are essentially approximating a function $\alpha : X \to Y$ where $X$ is a feature space, $Y$ an output space.
- We consider the problem in the presence of symmetries.

- e.g.: Recognising a single handwritten digit which might have been rotated by factor of $90^\circ$.

: "Defn:" A group acts on a set, $G \curvearrowright X$, if it does something to the set. Denoted $g \cdot x$ for $g \in G$, $x \in X$.

- e.g. $\mathbb{Z}/4 = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}\} \curvearrowright \boxed{9}$ by $90^\circ$ clockwise rotation.
  i.e.: $\{\bar{0}\} \curvearrowright \boxed{9} = \boxed{9}$, $\{\bar{1}\} \curvearrowright \boxed{9} = \boxed{\rotatebox{-90}{9}}$, etc.

- Defn: Let $G \curvearrowright X$, $\alpha : X \to Y$. $\alpha$ satisfies the invariance property if
$$\alpha(g \cdot x) = \alpha(x) \quad \forall x \in X, g \in G.$$

- Let $\beta : X \to Y$ be the ML approximant of $\alpha$. If $\alpha$ $G$-inv., we want $\beta$ to be $G$-inv.
- e.g. since $\alpha : \{\text{handwritten digits}\} \to \{\text{digits } 0-9\}$ is $\mathbb{Z}_4$ invariant, $\beta : X \to Y$ should be too.

- We will restrict to the case that feature space $X$ is a Riemannian mfld, $G \curvearrowright X$ discretely via isom.

# §1.1: Methods

- There are two approaches to the problem of group invariant MLMs.
  1. Symmetrisation based: averaging a non-G-invariant model over the action of G.
  2. Intrinsic: designing the model to be G-invariant by imposing conditions coming from group action.

- E.g.: [symmetrisation] Data augmentation
  - Increase size of training data $D_{train} = \{(x, y) \mid x \in X_{train} \subset X,\ y = \alpha(x) \in Y\}$.
  - By applying sample element $G_0 \subset G$, to get $D_{train}^{aug} = \{(g \cdot x, y) \mid (x, y) \in D_{train}, g \in G_0\}$.

- E.g. [intrinsic] This paper. They suggest a G-invariant pre-processing step to be applied to the input data that can be composed with any ML architecture
  - Done in general by imposing restrictions on weights

## §2.1: Fundamental domains

- A fundamental domain represents $G \curvearrowright X$.
- "<u>Defn</u>": $F$ is a <u>fundamental domain</u> for $G \curvearrowright X$ if
$$X = \bigcup_{g \in G} g \cdot F$$
- i.e. $F$ is the building block of the action.
- e.g. for $\mathbb{Z}_4 \curvearrowright \{\text{handwritten digits}\}$,
$$F = \{\boxed{\cdots}, \ldots, \boxed{\cdots} \mid \boxed{\cdots} \text{ right way up}\}$$

- $F$ is all the info we need for $G \curvearrowright X$ & it's $G$-invariant.

- We need to find $\pi : G \to F$ s.t. $\pi$ is $G$-invariant.

- Then can apply MLMs to $\alpha|_F : F \to Y$ trained on $D^{\pi}_{train} = \{(\pi(x), y) \mid (x,y) \in D_{train}\}$
Then $\beta = \bar{\beta} \circ \pi$ gives required approx

$$\begin{tikzcd} X \arrow[r, "\beta"] \arrow[d, "\pi"'] & Y \\ F \arrow[ur, "\bar{\beta}"'] & \end{tikzcd}$$

# § 2.2: Dirichlet projections

- $F_{Dir}$:
  - Start with a base pt $x_0 \in X$ (only fixed by $g \in G$ fixing $X$ pointwise)
  - $F = \{x \in X \mid d(x, x_0) < d(x, g \cdot x_0) \; \forall g \in G\}$
  - $F$ is all points closer to $x_0$ than any other pts of $G \cdot x_0$.

- Finding $\Pi_{Dir}: X \to F_{Dir}$ is a minimisation problem for the metric on $X$. [want to find $g \in G$ minimising $d(g \cdot x, x_0)$]
  - Can be approximated using a discrete gradient descent algorithm.

- This alg. cannot be guaranteed to find the minimum, so can only approximate $\Pi$.

## §2-3: Combinatorial Projections

- We would like to have an explicit and easy to compute description of $\pi$.

- The paper gives an algorithm to do this for $G < S_n \curvearrowright X = \mathbb{R}^n$ by co-ord. permutation

as then can't get from $f_1 \in F$ to $f_2 \in F$ via action of $S_n$, fits defn (1 before).

- $F$ corresponds to a consistent choice of reordering. e.g. $F = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_i < x_{i+1} \ \forall i < n\}$

- Then our projection to $F$ is applying group elements which order $x_i$ with our choice

- <u>e.g.</u> there is the <u>ascending projection map</u> $\pi_\uparrow$.

- Once we've applied the projection, we conduct ML on $\pi(X)$ to get $\bar{\beta} : \pi(X) \to Y$ and $\bar{\beta} \circ \pi : X \to Y$ is gp inv. by design, which is what we wanted.

# Summary of results

- Augmentation is a data pre-processing step: computationally impractical for large groups.

- Intrinsic group invariant neural networks are model-specific.

- Fund. dom. projections are easy to use, maintain original dim. and geom. of data, and are compatible with any MLM.
- Resulting $\beta$ is $G$-invariant.