5 A Category Theory Perspective

In this section, we reformulate equivariant neural networks using category theory, providing a general language for the principle that "an equivariant map preserves structure". This abstraction allows for the uniform treatment of different types of symmetry (groups, monoids, categories) through the lens of functors and natural transformations.

We assume the input space is a topological space, and the latent space, via a network (encoder), is a manifold. For a monoid (Definition 2.50) (or group) S, denote $\mathbf{B}S$ as the one-object category (Definition 2.49) with endomorphisms (Definition 2.47) S.

5.1 Categorical S-map

Definition 5.1 (S-object)**.** *Let* $\mathbf{C}$ *be a category and S be a monoid. An S-object in* $\mathbf{C}$ *is a functor:*

$$F : \mathbf{B}S \longrightarrow \mathbf{C},$$

where F maps the unique object $* \in \mathbf{B}S$ to some object $F(*) \in \mathbf{C}$; maps each morphism $s : * \to *$ in $\mathbf{B}S$ to a morphism $F(s) : F(*) \to F(*)$ in $\mathbf{C}$.

An S-object is simply "data equipped with an S-action": the functor packages both the data (the object $F(*)$) and how symmetry transformations act on it (the morphisms $F(s)$). For example if S is a group, the functoriality conditions $F(e) = \mathrm{id}$ and $F(s_1 s_2) = F(s_1) \circ F(s_2)$ are precisely the group action axioms.

The category of all S-objects in $\mathbf{C}$ is the functor category $[\mathbf{B}S, \mathbf{C}]$. Intuitively, S specifies the "shape of constraints" inside $\mathbf{C}$; an S-object is a diagram of that shape inside $\mathbf{C}$.

Definition 5.2 (S-map)**.** *Given two S-objects $F, E \in [\mathbf{B}S, \mathbf{C}]$, an S-map is a natural transformation*

$$\alpha : F \Rightarrow E.$$

Explicitly, α consists of a morphism $\alpha_* : F(*) \to E(*)$ in $\mathbf{C}$ *satisfying*

$$\alpha_* \circ F(s) = E(s) \circ \alpha_*.$$

Equivalently, the following diagram:

$$\begin{tikzcd}
F(*) \arrow[r, "\alpha_*"] \arrow[d, "F(s)"'] & E(*) \arrow[d, "E(s)"] \\
F(*) \arrow[r, "\alpha_*"'] & E(*)
\end{tikzcd}$$

commutes for all $s \in S$.

The naturality square is exactly the equivariance condition. Translate the diagram: "transform input then apply α" equals "apply α then transform output", this is exactly what we mean when we say a function respects the symmetry. In this context, equivariance is not an arbitrary design choice but a fundamental coherence condition for maps (neural networks) between structured objects.

Definition 5.3 (S-equivariant learner)**.** *An S-equivariant learner is an S-map in a machine learning context:*

- *Input is data equipped with S-symmetry (an S-object F).*
- *Output respects this symmetry (an S-object E).*
- *Model is a natural transformation $\alpha : F \Rightarrow E$.*

So the learner preserves the symmetry structure by design.

This gives us a suggestion when designing an equivariant architecture. If we know what the symmetry is, we need to specify three things: (1) the symmetry monoid/group S, (2) how S acts on inputs (the functor F), and (3) how S acts on outputs (the functor E). The architecture then consists of all parameterized maps satisfying the naturality condition.

Proposition 5.4 (Closure)**.** *The composition of S-maps is again an S-map. In particular, S-equivariant learners are closed under composition.*

Proof. Given S-maps $\alpha : F \Rightarrow E$ and $\beta : E \Rightarrow H$, the composition $\beta \circ \alpha$ satisfies:

$$(\beta_* \circ \alpha_*) \circ F(s) = \beta_* \circ (E(s) \circ \alpha_*) = H(s) \circ (\beta_* \circ \alpha_*)$$

by applying naturality of α then β. This is exactly the functoriality of vertical composition of natural transformations. □

This proposition is why deep equivariant networks work. If layer 1 is equivariant and layer 2 is equivariant, their composition is automatically equivariant. We can stack as many equivariant layers as we want without breaking the symmetry.

5.2 Categorical G-map

In the general S-map framework, S can be any monoid or category. A particularly important case in machine learning is when $S = G$ is a Lie group, as we discuss in Section 4.

We now place Lie groups within the categorical framework. Let **Top** be the category of topological spaces and **Man** be the category of smooth manifolds.

Proposition 5.5. *A Lie group is a group object in the category* **Man** *of smooth manifolds.*

A group object internalizes the group axioms within a category. Just as a group in **Set** is an ordinary group, a group in **Man** is a Lie group: the smoothness of the manifold structure is automatically inherited by the group operations.

Definition 5.6 (G-space and G-manifold)**.** *Let G be a Lie group.*

A G-space, denoted as (X, ρ_X), is a topological space X together with a continuous action

$$\rho_X : G \times X \to X$$

satisfying identity $e \cdot x = x$ and compatibility $(g_1 g_2) \cdot x = g_1 \cdot (g_2 \cdot x)$ for all $g_1, g_2 \in G, x \in X$.

Similarly, a G-manifold, denoted as (M, ρ_M), is a manifold M together with a smooth action

$$\rho_M : G \times M \to M$$

satisfying the same identity and compatibility conditions.

We write $\mathbf{Top}^G$ for the category of G-spaces with G-equivariant continuous maps as morphisms; $\mathbf{Man}^G$ for the category of G-manifolds with G-equivariant smooth maps as morphisms.

Intuitively, a G-space is the concrete geometric realization of a G-object. The action ρ_X tells us how G moves points around in X.

Definition 5.7 (G-equivariant map)**.** *Given a G-space (X, ρ_X) and a G-manifold (M, ρ_M), a map $f : X \to M$ is a G-equivariant map (G-map) if*

$$f(g \cdot x) = g \cdot f(x) \quad \forall g \in G, x \in X.$$

Equivalently, this diagram

$$\begin{tikzcd}
G \times X \arrow[r, "\rho_X"] \arrow[d, "id_G \times f"'] & X \arrow[d, "f"] \\
G \times M \arrow[r, "\rho_M"'] & M
\end{tikzcd}$$

commutes.

In this case, we say that f commutes with the group action.

Definition 5.8 (G-equivariant learner)**.** *Same as Definition 5.3, a G-equivariant learner is a G-map in a machine learning context processing data equipped with G-symmetry while preserving this symmetry structure.*

Conjecture 5.9 (Categorical Equivalence)**.** *Let G be a discrete group (equipped with discrete topology) and $\mathbf{B}G$ be the one-object category with endomorphisms G.*

Then there is an isomorphism of categories:

$$\mathbf{Top}^G \cong [\mathbf{B}G, \mathbf{Top}].$$

Similarly for manifolds (with discrete G):

$$\mathbf{Man}^G \cong [\mathbf{B}G, \mathbf{Man}].$$

If G is not discrete, the isomorphism of categories degrades to an equivalence of categories:

$$\mathbf{Top}^G \simeq [\mathbf{B}G, \mathbf{Top}].$$

This conjecture asserts that the two ways of thinking about "spaces with G-action" are essentially the same:

1. The *geometric* view: a space X with an action map $\rho : G \times X \to X$ (Definition 5.6).

2. The *categorical* view: a functor $F : \mathbf{B}G \to \mathbf{Top}$ (Definition 5.1).

The idea of this conjecture is that, for discrete groups, these perspectives are isomorphic, but for continuous groups, the functor category does not automatically encode continuity of the action, so we only get an equivalence.

Proposition 5.10 (Equivariance as Naturality). *Let $F : \mathbf{B}G \to \mathbf{Top}$ and $E : \mathbf{B}G \to \mathbf{Man}$ be functors, with $U : \mathbf{Man} \to \mathbf{Top}$ the forgetful functor.*

A G-equivariant map $f : F(*) \to E(*)$ (from the G-space $F(*)$ to the underlying topological space of G-manifold $E(*)$) is precisely a natural transformation:

$$\alpha : F \Rightarrow U \circ E$$

The naturality condition for each $g \in \mathrm{Mor}(\mathbf{B}G)$ is $f \circ F(g) = UE(g) \circ f$. Equivalently the commutative diagram:

$$\begin{tikzcd}
F(*) \arrow[r, "f"] \arrow[d, "F(g)"'] & UE(*) \arrow[d, "UE(g)"] \\
F(*) \arrow[r, "f"'] & UE(*)
\end{tikzcd}$$

Reading this diagram element-wise:

$$f(g \cdot x) = g \cdot f(x) \quad \forall g \in G, x \in F(*)$$

This is exactly the G-equivariance condition.

This is the central observation of the categorical perspective: equivariance is naturality. The equivariance condition imposed on neural network layers is precisely the condition that makes a morphism "natural" in the categorical sense.

Conjecture 5.11 (Total symmetry group)**.** *Real data typically carries multiple symmetries. But according to our definition, a G-space, or a functor category $[\mathbf{B}G, \mathbf{Top}]$ specifies one group. To apply the categorical framework, we must assemble these into a single "total" group G_{tot}. The construction depends on how the individual symmetries interact.*

Suppose X carries actions of groups $G_1, \ldots, G_k$.

- *Independent symmetries (direct product): If the actions commute on X, i.e., $g_i \cdot (g_j \cdot x) = g_j \cdot (g_i \cdot x)$ for all $i \neq j$, they assemble into*

  $$G_{\mathrm{tot}} = G_1 \times \cdots \times G_k.$$

  Intuition: The symmetries don't interfere, and the order of application doesn't matter.

- *Relabeling symmetries (semidirect product): If H acts on N by automorphisms $\varphi : H \to \mathrm{Aut}(N)$, capturing how H "rotates" or "relabels" the N-action, we use*

  $$G_{\mathrm{tot}} = N \rtimes_{\varphi} H.$$

  Intuition: One symmetry changes the meaning of another.

- *Replicated structure (wreath product): If $X = Y^n$ consists of n copies of a G-space Y, with S_n permuting copies, we use*

  $$G_{\mathrm{tot}} = G \wr S_n = G^n \rtimes S_n.$$

  Intuition: Each component has its own symmetry, components can be exchanged.

- *No relations (free product): If the actions satisfy no relations beyond those internal to each G_i, the free product $G_{\mathrm{tot}} = *_i G_i$ acts on X. This is rare for finite-dimensional data, as the ambient space typically induces additional relations.*

- *Beyond groups: If the structure of interest is not captured by any group (e.g., path-dependent or partially-defined symmetries), we can employ the framework of Section 4.1.*