Equivariant Deep Dynamical Model for Motion Prediction: Supplementary Materials

A Background on Group Theory

In this paper, we use numerous concepts in abstract algebra, group theory, and representation theory. We provide a wider range of details on the notations and definitions we employed in the paper.

Symmetry: A symmetry is a set of The transformations should preserve the properties of the structure. Generally, it is presumed that the transformations must be invertible, i.e., for each transformation there is another transformation, called its inverse, which reverses its effect. Symmetry is thus can be stated mathematically as an operator acting on an object, are modeled by **Groups**.

Group: Let G be a non-empty set with a binary operation defined as $\circ : G \times G \mapsto G$. We call the pair $(G; \circ)$ a group if it has the following properties: G is closed under its binary operation (Closure), the group operation is associative –i.e., $(g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3)$ for $g_1, g_2, g_3 \in G$ (Associativity axiom), there exists an identity $e \in G$ such that $g \circ e = e \circ g = g$ for all $g \in G$ (Identity axiom), every element $g \in G$ has an inverse $g^{-1} \in G$, such that $g \circ g^{-1} = g^{-1} \circ g = e$ (Inverse axiom).

Subgroup: A non-empty subset H of G is called a subgroup, if H is a group equipped with the same binary operation of as in G. We show this as $H \leq G$. H is called a proper subgroup of if $H \neq G$ and we show it as $H < G$.

Group action: We say a group G *acts* on a set $\mathcal{X}$ if there exist a map $\phi : G \times \mathcal{X} \to \mathcal{X}$ such that: (1) $\phi(e, x) = x$, where e is the identity element of G, and (ii) $\phi(g, \phi(h, x)) = \phi(gh, x)$ for all $g, h \in G$ and $x \in \mathcal{X}$. In this case, G is called a transformation group, $\mathcal{X}$ is a called a G-set, and ϕ is called the group action.

Lie group and infinitesimal generator: A Lie group G is a smooth manifold equipped with the structure of a group such that the group operation and inverse-assigning operation are smooth functions (Gilmore, 2006). This means that the group operation and inverse operation are continuous on the manifold, and they can be expressed in terms of the coordinates. The manifold is locally represented by a chart mapping to an underlying Euclidean space $\mathbb{R}^D$, where D is the dimensionality of the manifold. Furthermore, the chart map is defined in such a way that it associates the identity element in the group with the origin of Euclidean space. Elements of the Lie group can *act* as a transformation on the basis an n-dimensional vector space known as the *geometric space*, and change the coordinates of elements accordingly. We analyze Lie groups in terms of their infinitesimal generators which are the derivative of the group elements with respect to its D underlying parameters at the identity. These infinitesimal generators are the basis for a new vector space, called the *Lie Algebra*.

Lie algebra: Lie algebra, denoted as $\mathfrak{g}$, is the first order infinitesimal approximation to a Lie group, and can be interpreted as a tangent space at the identity $\mathfrak{g} \coloneqq T_{\mathrm{id}}G \subseteq \mathbb{R}^{n \times n}$. In general, an algebra over a field is a vector space equipped with a bilinear product. Thus, it consists of a set together with operations of multiplication and addition and scalar multiplication by elements of a field and satisfies the axioms of vector space and bilinear form (Hazewinkel et al., 2004).

Lie algebra representation: Each element of the Lie group can also be understood as a transformation on some other vector space such as what we call a Lie group of *transformations* (A N dimensional vector space, geometric space G_N). Every point has its own coordinate. We have some basis e_μ, and each point in the vector space has its coordinate which will just be its component relative to the e_μ basis. Each element of the Lie group represent a transformation of the basis which changes the coordinate of every point. The effect of the group elements on the elements of the underlying geometric space is describe by a function denoted as $y = f(\mu, x)$. every Lie group is its own Lie group of transformation – geometric space G_N.

B Useful Tensor Manipulations

If A is an $n \times p$ matrix and B an $m \times q$ matrix, the $mn \times pq$ matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1p}B \\ a_{21}B & a_{22}B & \dots & a_{2p}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \dots & a_{np}B \end{bmatrix}$$

is the Kronecker product of A and B. It is also called the tensor product. The $\mathsf{vec}(\cdot)$ operator creates a column vector from a matrix A by stacking the column vectors of $A = [a_1, a_2, \dots, a_n]$ below one another as:

$$\mathsf{vec}(A) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Kronecker Product and the $\mathsf{vec}(\cdot)$ Operator: For given matrices A, B, and X we have:

$$\mathsf{vec}(AXB) = (B^\top \otimes A)\mathsf{vec}(X). \tag{1}$$

C Network architecture and experimental settings

The network architectures for $\boldsymbol{\pi}^s_{\boldsymbol{\theta}}$, $\boldsymbol{\mu}^s_{\boldsymbol{\theta}}$, $\boldsymbol{\sigma}^s_{\boldsymbol{\theta}}$, and $\boldsymbol{\mu}^{\mathbf{x}}_{\boldsymbol{\theta}}$ are provided in table 1. For the pendulum experiment, we set the number of states $S = 2$, latent dimension $K = 3$, latent representation $U_{\mathrm{z}} = T_1$, state representation $U_{\mathrm{s}} = 2T_0$, observation representation $U_{\mathrm{x}} = T_1$, $\boldsymbol{\pi}^s_{\boldsymbol{\theta}}$ hidden representation $U_{h_s} = 3T_0 \oplus 2T_1$, $\boldsymbol{\mu}^s_{\boldsymbol{\theta}}$ and $\boldsymbol{\sigma}^s_{\boldsymbol{\theta}}$ hidden representation $U_{h_z} = 3T_0 \oplus T_1 \oplus T_2$, and $\boldsymbol{\mu}^{\mathbf{x}}_{\boldsymbol{\theta}}$ hidden representation $U_{h_x} = 3T_0 \oplus T_1$. For other experiments, we set the number of states $S = 2$, latent dimension $K = 6$, latent representation $U_{\mathrm{z}} = 3T_0 \oplus T_1$, state representation $U_{\mathrm{s}} = 2T_0$, observation representation $U_{\mathrm{x}} = D\,T_1$, $\boldsymbol{\pi}^s_{\boldsymbol{\theta}}$ hidden representation $U_{h_s} = 3T_0 \oplus 2T_1 \oplus T_2$, $\boldsymbol{\mu}^s_{\boldsymbol{\theta}}$ and $\boldsymbol{\sigma}^s_{\boldsymbol{\theta}}$ hidden representation $U_{h_z} = 3T_0 \oplus 3T_1 \oplus 2T_2$, and $\boldsymbol{\mu}^{\mathbf{x}}_{\boldsymbol{\theta}}$ hidden representation $U_{h_x} = 3T_0 \oplus 3T_1$. This setting of hidden representations and dimensions roughly match the generative parameter count of EqDDM with that of the baselines for a fair comparison. We set the latent dimension and number of states (if applicable) for the baselines accordingly.

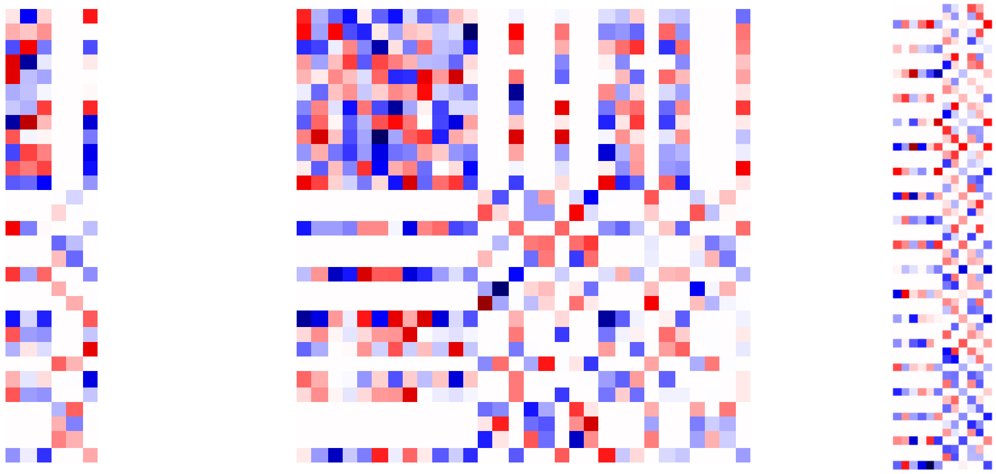
We have visualized the three equivariant linear layers with dimensions $K \times 5K$, $5K \times 5K$, and $2K \times 3D$, respectively in figure 1 for $K = 6$ and $D = 19$.

Table 1: Network architectures for the nonlinear mappings in EqDDM. ELL: Equivariant Linear Layer, ILL: Invariant Linear Layer, ENL: Equivariant Nonlinearity

Network	$\boldsymbol{\pi}^s_{\boldsymbol{\theta}} : \mathbb{R}^K \to \mathbb{R}^S$	$\boldsymbol{\mu}^s_{\boldsymbol{\theta}}$, $\boldsymbol{\sigma}^s_{\boldsymbol{\theta}} : \mathbb{R}^{\lvert\ell\rvert \times K} \to \mathbb{R}^{K,K}$	$\boldsymbol{\mu}^{\mathbf{x}}_{\boldsymbol{\theta}} : \mathbb{R}^K \to \mathbb{R}^{3D}$
Input	$z_{t-1} \in \mathbb{R}^K$	$z_{t-\ell} \in \mathbb{R}^{\lvert\ell\rvert \times K}$	$z_t \in \mathbb{R}^K$
1	ELL $K \times 3K$ ENL	ELL $\lvert\ell\rvert \times K \times 5K$ ENL	ELL $K \times 2K$ ENL
2	ELL $3K \times 3K$ ENL	ELL $\lvert\ell\rvert \times 5K \times 5K$ ENL	ELL $2K \times 2K$ ENL
3	ILL $3K \times S$ Softmax	AvgPool($\lvert\ell\rvert$)	ELL $2K \times 2K$ ENL
4		ELL $5K \times 5K$ ENL	ELL $2K \times 3D$
5		ELL $5K \times K$ ILL $5K \times K$	

D Computational Resources

We implemented EqDDM with PyTorch v1.8 (Paszke et al., 2017) and used the Adam optimizer (Kingma and Ba, 2014) with learning rate of 0.01. We initialized all the parameters randomly. We performed all the experiments



(a) 6×30 (b) 30×30 (c) $12 \times (19 \times 3)$

Figure 1: Visual representations of three equivariant linear layers for dimensions (a) 6×30, (b) 30×30, and (c) $12 \times (19 \times 3)$. For each representation, similar colors denote shared parameters.

on an Intel Core i9 CPU@3.6GHz with 32 GB of RAM. Per-epoch training time varied from 200 msec in smaller datasets to 1 sec in larger experiments and 300 epochs sufficed for all the experiments.

E Impact

The goal of this paper is to design a deep structured architecture for generative modeling of dynamic data by adopting the formalism of group theory. We do not expect that the developed model has an immediate societal impact or poses any direct risks. However, because the problem of deep generative modeling concerns designing a probabilistic model that can generate realistic data, it may be misused in producing fake realistic data.

References

Gilmore, R. (2006). *Lie Groups, Lie Algebras, and Some of Their Applications*. Courier Corporation.

Hazewinkel, M., Gubareni, N., and Kirichenko, V. V. (2004). *Algebras, rings and modules*, volume 1. Springer Science & Business Media.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., and Lerer, A. (2017). Automatic differentiation in pytorch.