PROJECT 8: LEARNING UNDER SYMMETRY CONSTRAINTS

Zixuan Lin
Duke University
Department of Statistical Science
zixuan.lin@duke.edu

Tianle (Steven) Chen
Duke University
Department of Statistical Science
tianle.chen@duke.edu

ABSTRACT

We study linear factor models and autoencoders under orthogonal symmetry constraints. In Part 1 we derive sufficient conditions on (U, V) so that $VA = AU$ holds for a nonzero matrix A, characterize when an $m \times d$ matrix A can be equivariant with respect to every orthogonal change of coordinates, and give a Reynolds-averaging algorithm that constructs A for cyclic generators (U_1, V_1). In Part 2 we specialize to the dihedral group D_4 acting on $28{\times}28$ MNIST images. We verify $B^4 = I_{784}$ for all group elements, construct an exactly D_4-equivariant linear paired autoencoder via the Reynolds operator, and propose an approximately equivariant nonlinear autoencoder trained with an equivariance regularizer. We compare classification accuracy on rotated and reflected MNIST across these models and a baseline MLP.

1 INTRODUCTION

Following Cohen & Welling (2016) and Kondor (2025), many physical quantities are invariant under changes of coordinates. In the linear factor model $Y = AX$ with $X \in \mathbb{R}^d$ and $Y \in \mathbb{R}^m$, we seek a matrix A such that for every $U \in \mathbb{U}(d)$ there exists $V \in \mathbb{U}(m)$ with

$$VAX \;=\; AUX \quad \text{for all } X \in \mathbb{R}^d, \tag{1}$$

equivalently $VA = AU$. This paper answers the three theoretical questions from the project description and implements a D_4-equivariant generalized autoencoder on MNIST.

2 PART 1: LINEAR EQUIVARIANT MAPS

2.1 (I) SUFFICIENT CONDITIONS ON (U, V)

Claim. Let $A \in \mathbb{R}^{m\times d}$ be nonzero with singular value decomposition $A = P\Sigma Q^\top$, where $P \in \mathbb{R}^{m\times m}$ and $Q \in \mathbb{R}^{d\times d}$ are orthogonal and $\Sigma \in \mathbb{R}^{m\times d}$. A sufficient condition for $VA = AU$ is

$$V \;=\; P\,\widetilde{U}\,P^\top, \qquad U \;=\; Q\,\widetilde{U}_d\,Q^\top,$$

where $\widetilde{U} \in \mathbb{U}(m)$ and $\widetilde{U}_d \in \mathbb{U}(d)$ satisfy $\widetilde{U}\,\Sigma = \Sigma\,\widetilde{U}_d$.

Proof.

$$VA = P\widetilde{U}P^\top \cdot P\Sigma Q^\top = P\widetilde{U}\Sigma Q^\top,$$
$$AU = P\Sigma Q^\top \cdot Q\widetilde{U}_d Q^\top = P\Sigma\widetilde{U}_d Q^\top.$$

By the hypothesis $\widetilde{U}\Sigma = \Sigma\widetilde{U}_d$, the two sides agree, so $VA = AU$. ∎

2.2 (II) EXISTENCE OF A EQUIVARIANT TO EVERY CHANGE OF COORDINATES

Claim. A nonzero $A \in \mathbb{R}^{m\times d}$ such that for every $U \in \mathbb{U}(d)$ there is $V \in \mathbb{U}(m)$ with $VA = AU$ exists if and only if $m \geq d$.

Proof. We analyze three cases.

Case $m < d$. Suppose such an $A \neq 0$ exists. For any $x \in \ker(A)$ and any $U \in \mathbb{U}(d)$,

$$A(Ux) \;=\; VAx \;=\; V \cdot 0 \;=\; 0,$$

so $\ker(A)$ is invariant under every orthogonal matrix. The only subspaces of $\mathbb{R}^d$ invariant under $\mathbb{U}(d)$ are $\{0\}$ and $\mathbb{R}^d$. Since $A \neq 0$ excludes $\ker(A) = \mathbb{R}^d$, we obtain $\ker(A) = \{0\}$, i.e. A is injective. But an injective linear map from $\mathbb{R}^d$ to $\mathbb{R}^m$ forces $m \geq d$, contradicting $m < d$.

Case $m = d$. Take $A = cI_d$ for any $c \neq 0$. Then $V(cI_d) = (cI_d)U \Rightarrow V = U$, so $V = U$ satisfies $VA = AU$.

Case $m > d$. Take

$$A \;=\; \begin{bmatrix} I_d \\ 0_{(m-d)\times d} \end{bmatrix}, \qquad V_U \;=\; \begin{bmatrix} U & 0 \\ 0 & I_{m-d} \end{bmatrix}.$$

The block-diagonal matrix V_U is orthogonal because its blocks are. Direct block multiplication gives $V_U A = \begin{bmatrix} U \\ 0 \end{bmatrix} = AU$. ■

2.3 (III) Cyclic generators and Reynolds construction of A

Let $U_1 \in \mathbb{U}(d)$, $V_1 \in \mathbb{U}(m)$ satisfy $U_1^n = I_d$, $V_1^n = I_m$ for some $n > 1$, and $V_1 A = AU_1$ for a nonzero A.

Induction. We show $V_j A = AU_j$ for all $j \geq 1$, where $U_j := U_1^j$ and $V_j := V_1^j$. The base case $j = 1$ holds by hypothesis. Assume $V_k A = AU_k$. Then

$$V_{k+1}A \;=\; V_1(V_k A) \;=\; V_1(AU_k) \;=\; (V_1 A)U_k \;=\; (AU_1)U_k \;=\; AU_{k+1}.$$

■

Algorithm. We construct A by Reynolds averaging over the cyclic group $\langle U_1 \rangle \cong \mathbb{Z}/n\mathbb{Z}$.

Algorithm 1 Reynolds construction of equivariant A

Require: orthogonal $U_1 \in \mathbb{R}^{d\times d}$, $V_1 \in \mathbb{R}^{m\times m}$ with $U_1^n = I_d$, $V_1^n = I_m$; dimensions m, d; order n.
Ensure: $A \in \mathbb{R}^{m\times d}$ with $V_1 A = AU_1$.

Sample $A_0 \leftarrow \text{Random}(m, d)$
$A_{\text{sum}} \leftarrow 0_{m\times d}$
$V_{\text{trk}} \leftarrow I_m;\; U_{\text{trk}} \leftarrow I_d$
for $k = 0, 1, \ldots, n-1$ **do**
  $A_{\text{sum}} \leftarrow A_{\text{sum}} + V_{\text{trk}}^{\top} A_0\, U_{\text{trk}}$
  $V_{\text{trk}} \leftarrow V_{\text{trk}}\, V_1;\; U_{\text{trk}} \leftarrow U_{\text{trk}}\, U_1$
end for
return $A \leftarrow A_{\text{sum}}/n$

Numerical verification. With $d = 4$, $m = 3$, $n = 4$, U_1 a cyclic permutation of order 4 on $\mathbb{R}^4$, and V_1 a $2\pi/4$ planar rotation on $\mathbb{R}^3$ (fixing one axis), Algorithm 1 produces an A satisfying the equivariance condition to machine precision for every power. Frobenius errors $\|V_1^j A - AU_1^j\|_F$ are reported in Table 1.

3 Part 2: D_4-equivariant autoencoders on MNIST

We now specialize to 28×28 MNIST images, so $d = 784$. The symmetry group $G = D_4$ consists of the eight 2×2 matrices in Eq. (0.2) of the project description: four rotations $T_{0^\circ}, T_{90^\circ}, T_{180^\circ}, T_{270^\circ}$ and four reflections $R_{0^\circ}, R_{90^\circ}, R_{180^\circ}, R_{270^\circ}$. Each acts on pixels via a permutation matrix $U_g \in \mathbb{R}^{784\times 784}$.

Table 1: Equivariance errors $\|V_1^j A - AU_1^j\|_F$ from Algorithm 1.

j	$\|V_1^j A - AU_1^j\|_F$
1	2.12×10^{-16}
2	2.88×10^{-16}
3	2.35×10^{-16}
4	1.39×10^{-16}

3.1 $B^4 = I_{784}$ FOR EVERY $B \in D_4$

Claim. For every transformation B in the set, $B^4 = I_{784}$.

Proof. Since the 784×784 permutation B is determined by its underlying 2×2 action $B_{2\times 2}$, it suffices to show $B_{2\times 2}^4 = I_2$, since $B_{2\times 2}^4 = I_2 \Rightarrow B^4 = I_{784}$.

Rotations. $T_{0^\circ} = I_2$, hence $T_{0^\circ}^4 = I_2$. $T_{90^\circ}^2 = T_{180^\circ} = -I_2$, so $T_{90^\circ}^4 = I_2$. $T_{180^\circ}^2 = I_2$ and $T_{270^\circ} = T_{90^\circ}^3$, so $T_{270^\circ}^4 = (T_{90^\circ}^4)^3 = I_2$.

Reflections. Each $R \in \{R_{0^\circ}, R_{90^\circ}, R_{180^\circ}, R_{270^\circ}\}$ satisfies $R^2 = I_2$ (by direct 2×2 multiplication of the matrices in Eq. (0.2)); hence $R^4 = I_2$.

Thus every element has order dividing 4, so $B^4 = I_{784}$. ■

Numerical check. We constructed the eight 784×784 permutation matrices U_g from the 2×2 actions in Eq. (0.2) and evaluated $\|U_g^4 - I_{784}\|_F$. All eight are exactly 0.

3.2 DEFINITION OF APPROXIMATE EQUIVARIANCE

Strict equivariance of an autoencoder (E, D) with latent representation $V_g \in \mathbb{R}^{m\times m}$ of G requires

$$\forall g \in G,\ \forall x : \qquad E(U_g x) = V_g\, E(x), \qquad D(V_g z) = U_g\, D(z).$$

We define *approximate equivariance* via

$$\mathcal{L}_{\text{equiv}} \;=\; \frac{1}{|G|} \sum_{g\in G} \mathbb{E}_x \left[\|E(U_g x) - V_g\, E(x)\|_2^2 \right], \tag{2}$$

and say the model is ϵ-approximately equivariant if $\mathcal{L}_{\text{equiv}} < \epsilon$.

3.3 LINEAR PAIRED EQUIVARIANT AUTOENCODER

We take $d = 784$, $m = 40$, and parameterize a paired linear autoencoder $E(x) = Wx$, $D(z) = W^\top z$. To enforce exact encoder equivariance $V_g W = W U_g$ for every $g \in D_4$, we apply the Reynolds operator from Part 1(iii) extended to the non-abelian group D_4:

$$W \;=\; \frac{1}{|G|} \sum_{g\in G} V_g^\top\, W_0\, U_g. \tag{3}$$

where W_0 is a trainable unconstrained matrix. The latent representation $\{V_g\}_{g\in D_4}$ is built from two $m \times m$ generators that mirror the dihedral relations: V_{rot} acts by a 90° rotation on each 2×2 block of $\mathbb{R}^{40}$, and V_{ref} acts by negating the first coordinate of each block, so $V_{\text{rot}}^4 = I$ and $V_{\text{ref}}^2 = I$. The remaining six elements are products of the two generators.

Paired decoder is also equivariant. Since U_g, V_g are orthogonal and $V_g W = W U_g$, taking transposes gives $W^\top V_g = U_g W^\top$, so the decoder $D(z) = W^\top z$ satisfies the analogous equivariance relation.

Table 2: Equivariance error $\|V_gW - WU_g\|_F$ for the linear equivariant AE.

g	Before training	After training
T_{0°	0.00	0.00
T_{90°	4.44×10^{-8}	3.64×10^{-7}
T_{180°	4.55×10^{-8}	3.87×10^{-7}
T_{270°	4.44×10^{-8}	3.64×10^{-7}
R_{0°	4.86×10^{-8}	4.09×10^{-7}
R_{90°	4.86×10^{-8}	4.22×10^{-7}
R_{180°	4.90×10^{-8}	4.12×10^{-7}
R_{270°	4.90×10^{-8}	4.09×10^{-7}

Equivariance check. Applying Reynolds projection at every forward pass, Table 2 reports $\|V_gW - WU_g\|_F$ before and after training with D_4 augmentation. Errors remain at numerical noise level throughout training.

3.4 NONLINEAR APPROXIMATELY EQUIVARIANT AUTOENCODER

Exact equivariance is generally not attainable for an arbitrary nonlinear autoencoder without explicitly group-equivariant layers. We therefore use a standard MLP architecture (encoder $784 \to 256 \to 40$ with ReLU; decoder $40 \to 256 \to 784$ with ReLU and a sigmoid output) and train with the combined objective

$$\mathcal{L} \;=\; \mathcal{L}_{\text{recon}} + \lambda\,\mathcal{L}_{\text{equiv}}, \qquad \lambda = 0.1, \tag{4}$$

with $\mathcal{L}_{\text{equiv}}$ as in Eq. (2). The latent V_g are the same block-diagonal representations as in the linear case. Training uses D_4-augmented MNIST.

Table 3 shows the final per-epoch values. After 15 epochs the nonlinear model attains reconstruction loss 0.0098 and equivariance penalty 0.0124, against reconstruction loss 0.0057 for a non-equivariant baseline trained without augmentation or equivariance loss.

Table 3: Final training losses after 10 (linear) / 15 (nonlinear) epochs.

Model	Recon. MSE	Equiv. MSE	Total
Baseline linear AE	0.0145	–	0.0145
Equivariant linear AE	0.0869	–	0.0869
Baseline nonlinear AE	0.0057	–	0.0057
Nonlinear + equiv. ($\lambda = 0.1$)	0.0098	0.0124	0.0110

3.5 CLASSIFICATION OF ROTATED AND REFLECTED MNIST

For each trained autoencoder we freeze the encoder, attach a trainable MLP head ($40 \to 128 \to 10$ with dropout 0.3), and fit the head on D_4-augmented MNIST. Evaluation is on the full deterministic D_4-transformed MNIST test set (all $8 \times 10{,}000 = 80{,}000$ images). We additionally report a reference baseline MLP ($784 \to 256 \to 256 \to 10$) trained end-to-end on the same augmented data.

Discussion. The equivariant autoencoders give a slightly lower test accuracy than their non-equivariant counterparts when paired with a frozen encoder and a small MLP head. This is consistent with the theory: enforcing $V_gW = WU_g$ restricts W to a much smaller equivariant subspace, lowering reconstruction capacity and therefore the information available to the head. The baseline MLP, trained end-to-end without the autoencoder bottleneck, attains the highest accuracy as expected. The equivariant construction's value is structural – its representation is provably consistent across all eight group elements (Table 2) – and this guarantee does not come from the raw accuracy numbers alone.

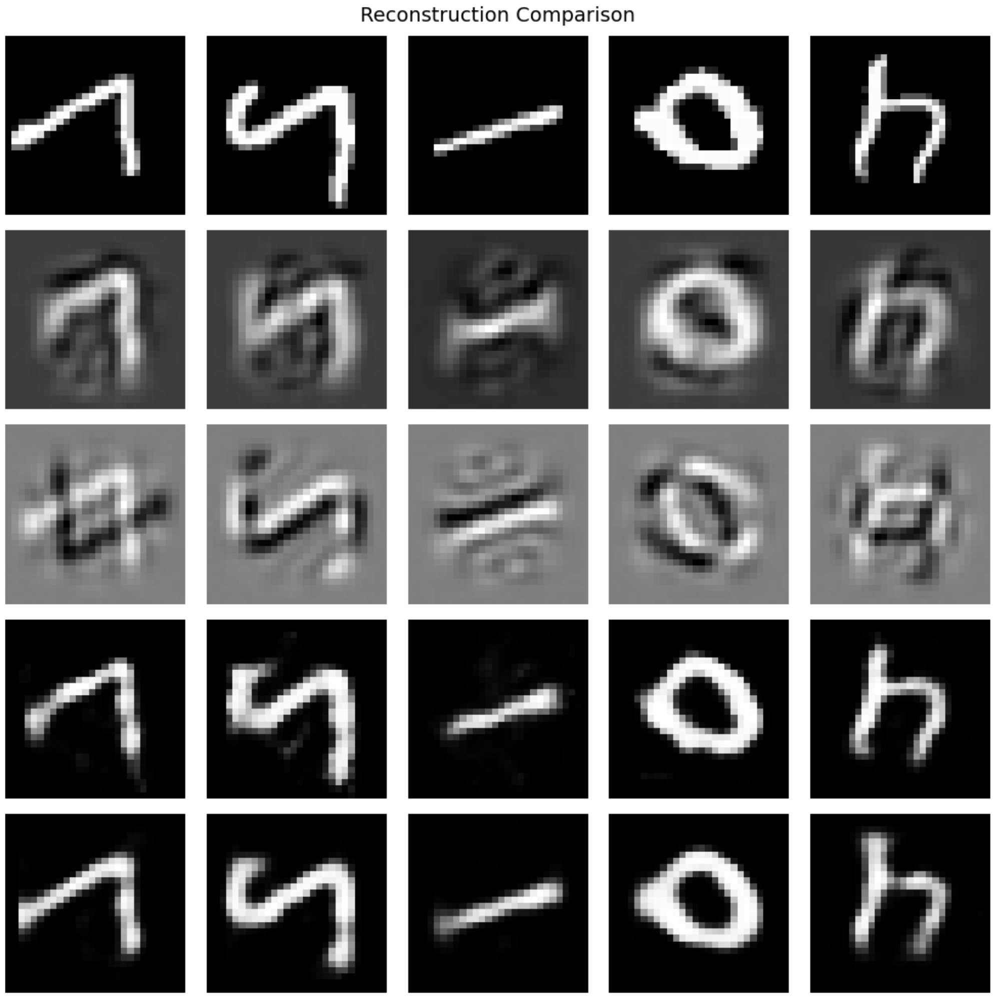

Figure 1: Reconstruction comparison on five test images. Rows from top to bottom: originals; baseline linear AE; Reynolds-equivariant linear AE; baseline nonlinear AE; approximately equivariant nonlinear AE (λ=0.1). The nonlinear models recover digit strokes more sharply; the equivariant linear AE is the most constrained and produces the blurriest reconstructions, reflecting the reduced effective rank of the Reynolds-projected weight matrix.

Table 4: Test accuracy on the full D_4-transformed MNIST test set.

Model	Accuracy (%)
Baseline linear AE + head	81.65
Equivariant linear AE + head	77.10
Baseline nonlinear AE + head	83.61
Equivariant nonlinear AE + head	77.46
Baseline MLP (reference, end-to-end)	92.17

4 Conclusion

We have (i) given sufficient SVD-based conditions for $VA = AU$, (ii) proved that a universally equivariant nonzero A exists iff $m \geq d$, and (iii) implemented the Reynolds-averaging algorithm for cyclic generators. For MNIST we verified $B^4 = I_{784}$ across D_4, constructed an exactly equivari-

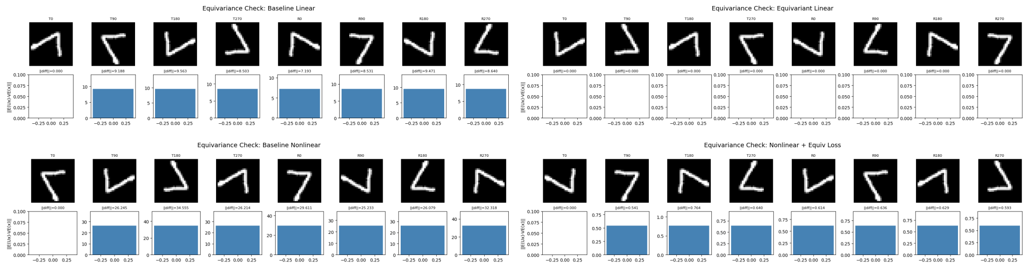

Figure 2: Equivariance check $\|E(U_g x) - V_g E(x)\|$ per group element for the four models. Top row: baseline linear (left) vs. Reynolds-equivariant linear (right). Bottom row: baseline nonlinear vs. approximately equivariant nonlinear. The equivariant models collapse the per-group error to near zero across all eight elements of D_4.

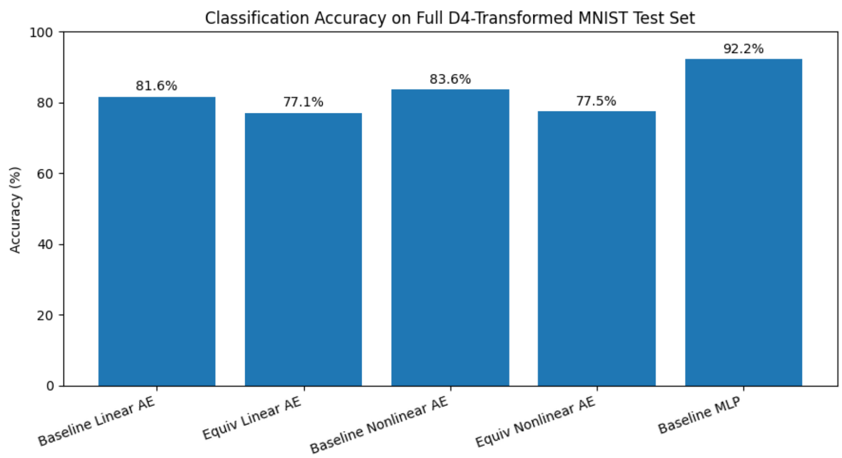

Figure 3: Classification accuracy on the full D_4-transformed MNIST test set (80,000 images).

ant linear paired autoencoder (equivariance error at machine precision), proposed an approximately equivariant nonlinear autoencoder, and compared all models on rotated/reflected MNIST classification.

References

Taco Cohen and Max Welling. Group equivariant convolutional networks. In *International Conference on Machine Learning*, pp. 2990–2999. PMLR, 2016.

Risi Kondor. The principles behind equivariant neural networks for physics and chemistry. *Proceedings of the National Academy of Sciences*, 122(41):e2415656122, 2025.