Symmetry Discovery with Deep Learning

Krish Desai, Benjamin Nachman, Jesse Thaler   UC Berkeley, LBNL, MIT

Abstract

Symmetries are a fundamental property of functions associated with data. We:

- give a **rigorous statistical notion** of symmetries for data
- construct a **novel Generative Adversarial Network (GAN) method** to learn symmetries from data
- **test** our method on particle physics simulations *vidilicet* the LHC Olympics dataset.

Symmetry discovery may lead to **new insights** and can **reduce effective dimensionality** of datasets to increase its statistics.

Statistics of Symmetries

- **Naive definition** $g : \mathbb{R}^n \to \mathbb{R}^n$ such that $p = p \circ g \cdot |g'|$
- **CDF Mapping Theorem** Let F be the CDF of X. Then $F(X) \sim \mathcal{U}(0,1)$.
- **Problem** $\mathcal{U}(0,1)$ has $\mathbb{Z}_2$ symmetry $\tilde{g} : x \mapsto 1 - x$. Hence, every dataset X has $\mathbb{Z}_2$ symmetry $g = F^{-1} \circ \tilde{g} \circ F$ (!)
- **Solution** Stronger definition of symmetry required. Seek inspiration from inertial reference frames of classical mechanics.
- **Definition** A symmetry of p relative to a reference *inertial* distribution p_I is a map g which preserves both p and p_I. This eliminates 'fake' symmetries like the quantile map above.

Machine Learning Approach

Neural Nets

- **Generator**, $g : \mathbb{R}^n \to \mathbb{R}^n$ represents the symmetry map, tries to maximize the loss
- **Discriminator**, $d : \mathbb{R}^n \to [0,1]$ tries to distinguish the input data $\{x_i\}$ from the transformed data $\{g(x_i)\}$.

Methods to Enforce Inertial Distribution Restriction

- **Simultaneously apply classifier** to samples from p and p_I
- Identify all PDF preserving maps g p. Then **post hoc select** the ones that preserve p_I
- **Restrict search space** of g to $\mathbb{A}SL_n^{\pm}(\mathbb{R})$, the group of linear symmetries of $\mathcal{U}(\mathbb{R}^n)$. This is the method we use.

Deep Learning Implementation

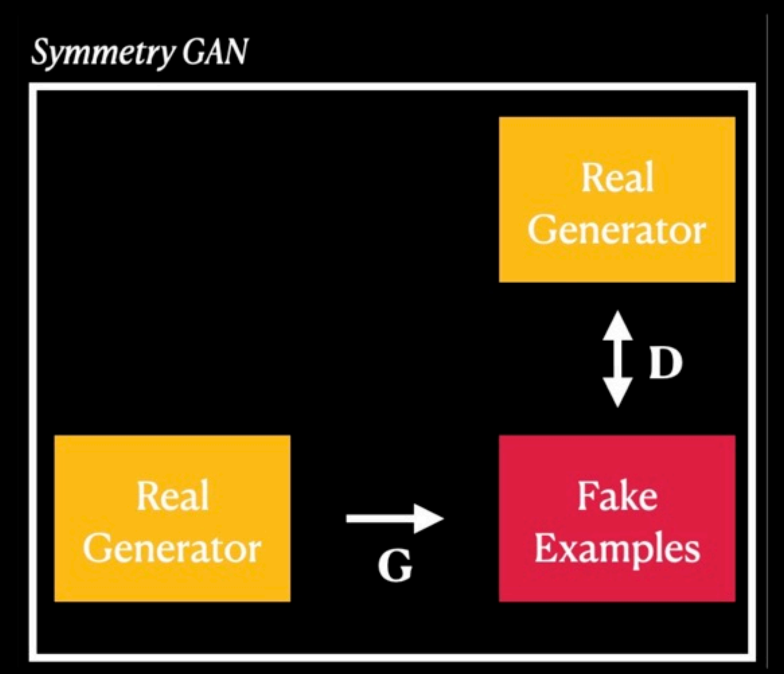
Loss Function

Modified Binary Cross Entropy

$$L[g,d] = -\sum_{x \in \{x_i\}} [\log(d(x)) + \log(1 - d(g(x)))]$$

Same samples in both terms, like a neural resampler.

- For Symmetry Discovery **latent probability** density is the **same as** the **target probability** density.
- The SymmetryGAN **generator** is **bijective** while usual GAN generators are not injective



Particle Physics Example

LHC Olympics Dataset

- Synthetic **proton-proton collisions** at the **Large Hadron Collider** at CERN.
- Focused on **jets**, collimated sprays of particles.
- Result from high energy **quark and gluon** interactions.
- Leading two jets of each event considered

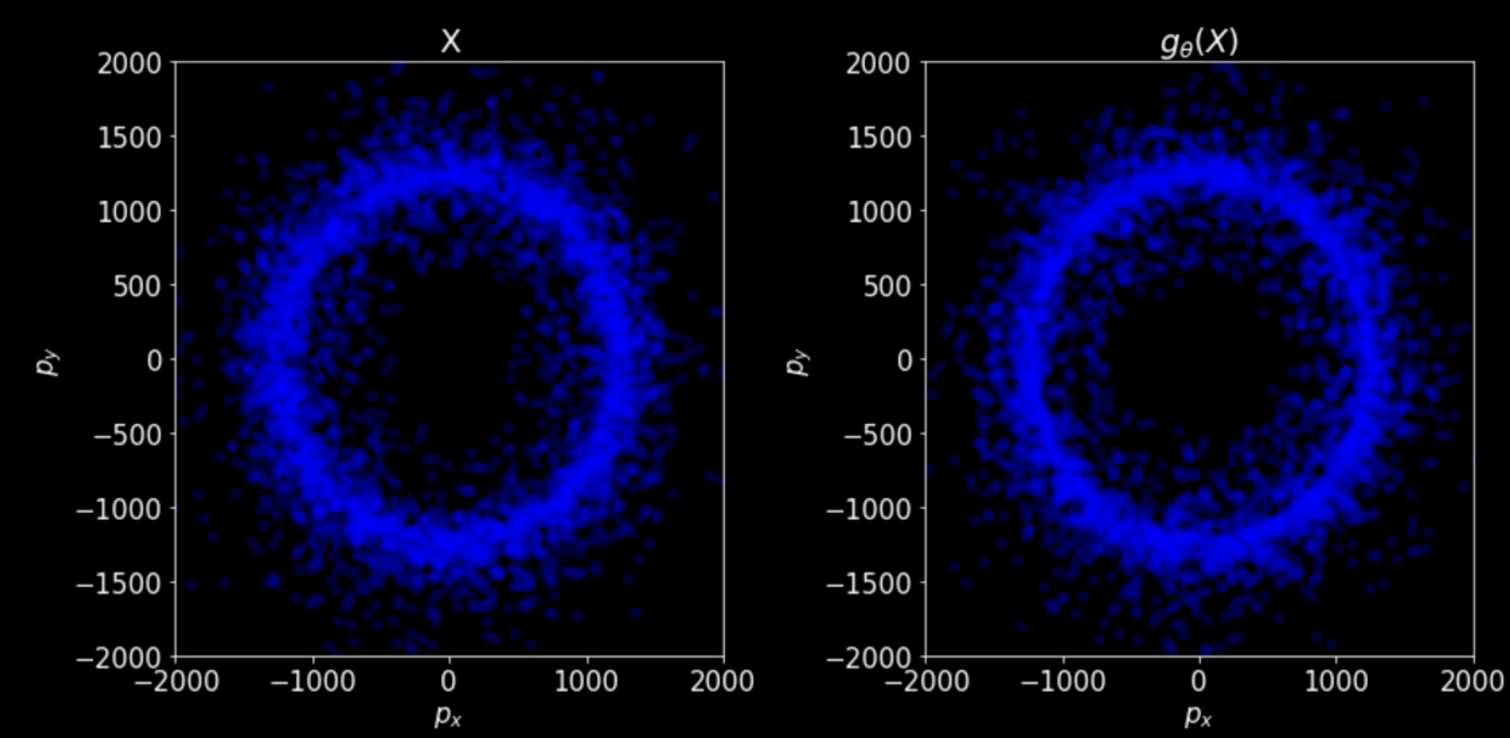
- Each event – four dimensional vector $(p_{1x}, p_{1y}, p_{2x}, p_{2y})$ – p_1 leading jet momentum – p_2 subleading jet momentum
- Focus on transverse plane – jets back-to-back – p conservation.
- Longitudinal momentum of parton-parton interaction not known – no conservation law for p_z.
- Natural search space, SO(4) – parameterized by 6 rotation angles.
- Generator $g_{\boldsymbol{\theta}}(X) = \left(\prod_{i=1}^{6} R_i(\theta)\right) X$

- Symmetry group not in any single plane of $SO(4)$
- Six dimensional group – hard to visualize – must seek alternate ways to verify discovered maps are symmetries

Analysis

Method 1

Plot X and $g_{\boldsymbol{\theta}}(X)$ and compare their features. The similarity suggests g is a symmetry.



Method 2

- $\phi_i = \arctan 2(p_y, p_x) \sim \mathcal{U}(-\pi, \pi)$.
- For arbitrary $SO(4)$ rotations on X, $\phi_f \not\sim \mathcal{U}$
- Applying a symmetry rotation, $\phi_f \sim \mathcal{U}$.
- Verify using KL Divergence.
- KLD $= 0 \implies$ analytic symmetry so smaller is better