Machine Learning in the String Landscape

Ben Campbell, Per Berglund, University of New Hampshire

String Theory

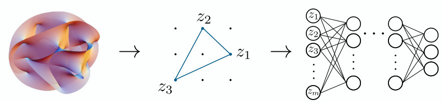
- String theory is currently the leading contender for a physical framework that unifies quantum mechanics and general relativity.
- There are different formulations of the theory, but each formulation requires at least a (9+1)-dimensional spacetime structure.
- In order for string theory to agree with the experimentally verified (3+1)-dimensional spacetime structure, 6 of 9 space dimensions must be "compactified" – these compactified dimensions take the form of a Calabi-Yau 3-fold, a three complex dimensional manifold.
- There is a lower bound of 10^{500} choices for compactification.

Machine Learning

- Used to classify data according to certain known/unknown patterns (relationships between data).
- Data is split into two groups: a training and testing set
- Many different types of models / network architectures
- We use a fully-connected network.
  - Each data point is passed forward through the network individually producing a prediciton.
  - A cost function measures how accuracte the networks prediciton was.
  - An optimization algorithm alters the network to minimize the cost function.
- Preparation of data is crucially important.
  - Number of data points must be as close to uniform across classes (uniform information density).

Compactification as a Machine Learning Problem

A Calabi-Yau n-fold can be uniquely mapped to an $(n+1)$-dimensional reflexive polytope, Δ^*, representing a particular type of compact $(n+1)$-dimensional manifold known as a toric variety:



$$(z_1)^3 + (z_2)^3 + (z_3)^3 = 0$$
$$\Rightarrow \; (z_1, z_2, z_3) \sim (\lambda z_1, \lambda z_2, \lambda z_3)$$

Our model:

- Fully-connected feed forward neural network
  - 2 hidden layers
  - Stochastic Gradient Decent Optimization
- Hyperparameters are optimized by a brute force method
- Data must be regularized to avoid overfitting

2-dimensional polytopes:

- There are only 16 reflexive 2D polytopes, including the dual polytopes $(\Delta^*)^\circ$
- Polytopes are grouped by their volume using the relation

$$\mathrm{Vol}(\Delta^*) + \mathrm{Vol}((\Delta^*)^\circ) = 12$$

- Reflections, rotations, addition of zero vectors and permutations of the vertices expand dataset from 16 to 6912
- Model achieves an accuracy of around 90%

3-dimensional polytopes:

- There are 4319 reflexive 3-dimensional polytopes.
- The problem of grouping the 3D polytopes is not as straight forward as for the 2D case.
- Each Calabi-Yau manifold has a topological invariant (not-unique) called its Picard number.
- The Picard number can be calculated from the polytope data by the equation

$$\mathrm{Pic} = l(\Delta^*) - \sum_{facets\ \theta^* \in \Delta^*} l^*(\theta^*) + C - 4$$

where l is the number of integer points of a polytope and l^* is the number of interior integer points of a facet or an edge, with the correction term

$$C = \sum_{edges\ \theta^* \in \Delta^*} l^*(\theta^*) l^*(\theta)$$

- A relation analogous to the volume equation for 2D polytopes is

$$\mathrm{Pic}(\Delta^*) + \mathrm{Pic}((\Delta^*)^\circ) - C = 20$$

- The 3D polytopes are then grouped into three classes:

$$\text{Class } 0 : 2 \cdot \mathrm{Pic} < 20 + C$$
$$\text{Class } 1 : 2 \cdot \mathrm{Pic} > 20 + C$$
$$\text{Class } 2 : 2 \cdot \mathrm{Pic} = 20 + C$$

- An accuracy of around 80% has been achieved with current models.

Future work

- The phenomenologically relevant case of the 4-dimensional toric varieties, corresponding to the Calabi-Yau 3-folds of string theory, are still to be worked on and contain 473,800,776 reflexive polytopes.
- The question remains of how to group the 4-dimensional polytopes, and what grouping methods can be carried over from the 3-dimensional case.

References

[1] Poonen, Bjorn, and Fernando Rodriguez-Villegas. "Lattice Polygons and the Number 12." The American Mathematical Monthly 107, no. 3 (2000): 238-50.
[2] Kreuzer, Maximilian and Harald Skarke. Classification of Reflexive Polyhedra in Three Dimensions. (1998).
[3] Bao, Jiakang et al. Lectures on the Calabi-Yau Landscape. (2020).
[4] Mehta, Pankaj et al. A high-bias, low-variance introduction to Machine Learning for physicists. Physics reports 810 (2019): 1-124 .
[5] He, Yang-Hui. The Calabi-Yau Landscape: from Geometry, to Physics, to Machine-Learning. (2018).
[6] Douglas, Michael R.. Statistics of string vacua. (2003).

University of New Hampshire