

THOUGHTS AND THEORY

What Einstein Can Teach Us About Machine Learning

Harnessing symmetry in machine learning



Jason McEwen · Following

Published in Towards Data Science

6 min read · Mar 2, 2021



Listen



Share

... More

In many ways physics and machine learning share a common goal: to formulate models of observed phenomena. In achieving this goal physicists have long understood the importance of symmetry. In this post we look at how the ideas of symmetry from physics may be leveraged as guiding principles in machine learning.

This blog post was co-authored with Oliver Cobb from [Kagenova](#).



Symmetries of a snowflake. [Photo by [Damian McCoig](#) on [Unsplash](#)]

Rapid progress has been made in machine learning over the past decade, particularly for problems involving complex high dimensional data, such as those in computer vision or natural language processing. However, a common criticism of machine intelligence when compared to its biological counterpart is the inefficiency with which it learns from examples. Whereas a young child may learn to recognise a new animal from just a handful of examples, a modern machine learning system may require hundreds or even thousands of examples to achieve the same feat.

Symmetry in physics

As humans we form models of the world around us based on robust physical laws, many of which we learn subconsciously. Physicists explore how such laws and models can be formalised and discovered. Their aim is to formulate models of underlying processes that accurately describe and predict observed phenomena.

Physical systems may be modelled at various levels of abstraction. Models used to explain astronomical phenomena typically leverage different physical laws to those used

to explain subatomic particles. There is, however, a principle that pervades physical laws at all levels of abstraction: known symmetries of the natural world must be respected.

The notion of symmetry with respect to physical laws is slightly different to its more familiar use in describing symmetries of objects. An object is considered to have a symmetry if it remains unchanged (i.e. invariant) under some transformation. For example, the fact that a sphere remains a sphere under any arbitrary rotation means that it exhibits rotational symmetry.

On the other hand, a physical law governing the behaviour of a system is considered symmetric to some transformation if the law applies in the same way to the system before and after it has undergone the transformation.

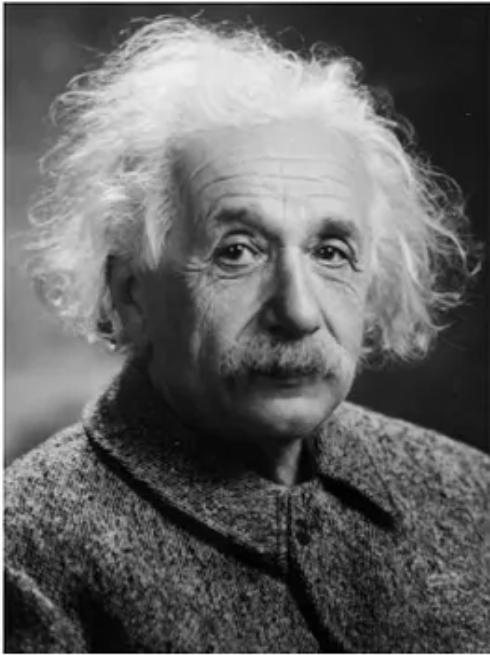
A simple example is translational symmetry, satisfied by laws that apply in the same way to a system regardless of the system's location. For example, a ball dropped in one room of a house behaves the same as a ball dropped in another room (ignoring any external factors like any slight breeze).

A second example is rotational symmetry, satisfied by laws that apply in the same way to a system regardless of the direction it is facing. A third example is time-translational symmetry, satisfied by laws that do not change with time.

Physicists have long been aware of the temporal and spatial symmetry properties of physical laws. However, in the early 20th century the significance of symmetry in physics underwent a paradigm shift.

Rather than starting with physical laws and deriving corresponding symmetry properties, in his famous 1905 paper on special relativity Einstein instead used principles of symmetry as a starting point to derive new physical laws.

A decade later German mathematician Emmy Noether, who made groundbreaking contributions to both abstract algebra and theoretical physics at a time when women were largely excluded from academic positions, elevated the role of symmetry within physics further. She proved that for every continuous symmetry of physical laws there exists a corresponding conservation law. For example the law of conservation of momentum may be derived from the translational symmetry of physical laws. Similarly the conservation of angular momentum follows from rotational symmetry and conservation of energy from time-translational symmetry.



Albert Einstein (left) and Emmy Noether (right). [Images sourced from Wikimedia Commons: [Einstein](#); [Noether](#)]

The fundamental laws of physics, like conservation of energy and momentum, actually follow from the symmetry of the Universe.

Leveraging symmetry as a guiding principle to discover corresponding laws and models to describe observed phenomena is not only of great use in physics, but may also be harnessed in machine learning.

Symmetry in machine learning

Machine learning practitioners are well aware of the importance of placing constraints on models to control the bias-variance tradeoff. When seeking a model of the relationship between explanatory and target variables, in machine learning we first specify a class of models that we hypothesise contains an adequately descriptive model. Within this class we search for the model that best describes the observed phenomena — i.e. that which maximises an empirical measure of fit.

It is important to specify a class that is broad enough that it contains a model that accurately describes the relationship, whilst also being restricted enough such that it is not outperformed by models that overfit to the data. This is typically difficult to achieve as machine learning is most useful when the relationship between explanatory and target variables is not well understood (it is, after all, what we are hoping to learn) and therefore it is not obvious how to set these boundaries. For example we know that the relationship between an image, i.e. an array of pixel intensities, and a category

corresponding to the image's semantic meaning is highly complex. How can we specify a model that allows for such complexity whilst also being relatively restricted?

One particularly effective way to introduce inductive biases into machine learning models to address this issue — which at this point should come as no surprise — is to leverage principles of symmetry!

Given a broad class of models we can immediately disregard the vast majority that don't adhere to notions of symmetry that the problem is known to exhibit. In the same spirit as Einstein in his discovery of special relativity, we start by noting the symmetry principles that should be satisfied and work backwards to find a model that best describes the observed data.

Symmetry in convolutional neural networks (CNNs)

The canonical example of how this principle has been leveraged in machine learning is in the design of convolutional neural networks (CNNs) for computer vision problems. As in any use of neural networks the aim is to hierarchically learn high-level features from low-level ones. The most important symmetry in computer vision is translational symmetry: a cat's eye is a cat's eye regardless of where it appears in the image.

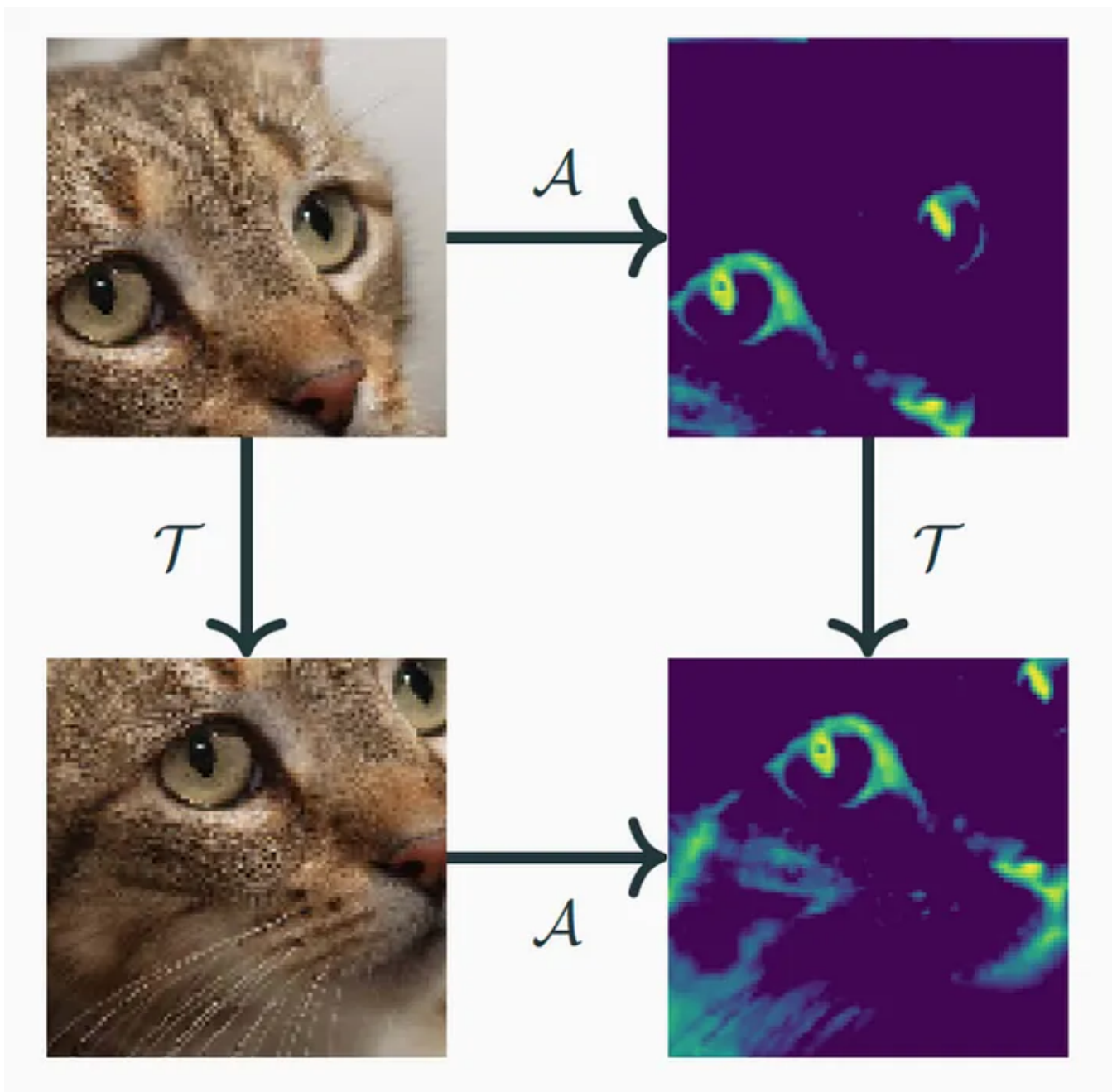


Illustration of translational equivariance. Given an image (top left), applying a convolutional kernel (\mathcal{A}) to obtain a feature map (top right) and then translating (\mathcal{T}) the feature map (bottom right) is equivalent to first translating the image (bottom left) and then applying the convolution kernel (bottom right). [Cat and feature map image source]

CNNs encode translational symmetry through the design of their architecture. Each neuron corresponds to a spatial region of the input and is connected only to a corresponding neighbourhood of neurons in the preceding layer. Crucially, every neuron is related to its corresponding neighbourhood in the preceding layer in exactly the same way. Therefore regardless of where a feature (e.g. a cat's eye) is located in an image, it stimulates the neurons in the corresponding location in an identical fashion. This property of the convolutional operator is called translational equivariance and is visualised in the diagram above — application of the operator to a feature followed by a translation is equivalent to translation followed by application of the operator.

By this careful architectural design we limit the space of models over which we search to only those adhering to this common-sense property of translational equivariance. Heuristically we may think of giving our learning algorithm a helping hand by assuring that a pattern need only be learnt once. Rather than having to learn the pattern in all possible locations, by encoding translational equivariance in the model itself we ensure the pattern can then be recognised in all locations.

Integrating symmetry into machine learning for planar images and beyond

The integration of translational symmetry into machine learning models is one of the key factors responsible for driving the revolutionary advances seen in computer vision over the past decade (combined with the proliferation of data and compute power).

It has certainly helped that 2D images have a simple planar form for which translational symmetry can be encoded in an intuitive and computationally efficient manner. For problems involving data with more complex (non-planar) geometry, respecting the desired symmetry principles can be more difficult. Dealing with complex geometry requires more advanced mathematical machinery, spawning the field of geometric deep learning. The geometric deep learning community has been making remarkable progress towards this goal, which we will consider further in upcoming posts.

References

- [1] Brading & Castellani, *Symmetries in Physics: Philosophical Reflections* (2018), [arXiv/0301097](https://arxiv.org/abs/0301097)
- [2] Higgins, Amos, Pfau, Racaniere, Matthey, Rezende, Lerchner, *Towards a Definition of Disentangled Representations* (2018), [arXiv:1812.02230](https://arxiv.org/abs/1812.02230)
- [3] Kunstatter, *Symmetry of Physical Laws* (1999), <https://theory.uwinnipeg.ca/users/gabor/symmetry/slide15.html>