

# Unsupervised learning on causal sets with variational autencoders

Anton van Niekerk

April 10, 2018

## Abstract

I apply a variational autencoder to the causal matrix for causal sets embedded in 2, 3, and 4 dimensional causal diamonds, and find that the latent variables successfully separate the causal sets by dimension.

## 1 Introduction

### 1.1 Causal Sets

When one typically thinks of the spacetime (time being thought of as a distinct “direction” to the three spatial dimensions we live in), the tendency in physics is to represent it as a continuous space, i.e., as a four dimensional real space where one of the dimensions is associated to the time direction. In that case, for every point in the spacetime, there are infinitely many other points arbitrarily close to that point (in other words the spacetime is continuous).

When one tries to combine gravity, which in Einstein’s theory of general relativity is the geometric curvature of spacetime, with quantum mechanics, the theory which describes elementary particles such as electrons and quarks, any quantity one cares about, such as the energy at which particles interact with gravity, gives the answer of infinity. While physicists have found ways to deal with infinite answers in other theories, such as quantum electrodynamics and find out the correct finite answer to their questions, none of these methods work for quantum gravity, and the infinities remain. Clearly, some new assumptions about nature are needed to successfully combine gravity and quantum mechanics.

One attempt to do this was through *causal sets* [1]. A causal set is a discrete version of spacetime, meaning that there are finite gaps between any spacetime point and the nearest other point to it. Because the distances between points in a causal set is finite, rather than infinitesimal as in the usual description of spacetime, quantum gravity in causal sets are free of the infinite answers mentioned above.

One way to create a causal set that approximates a continuous spacetime, is to “sprinkle” points at random into the continuum spacetime, see figure 1 for a two-dimensional example. Because the points are randomly sprinkled, there is approximately the same number of points per small volume of the spacetime. Also, this random sprinkling ensures that the causal set satisfies the same symmetries as the continuum spacetime. As an example, if a sphere is sprinkled with random points, since the sphere looks the same under any rotation, so will the random sprinkling under the same rotation.

A final property of causal sets that will be important is that the points in the set only “know” about points that can communicate with them. From special relativity we know that no information can travel faster than the speed of light, so only points that can send a signal at or slower than the speed of light to some future point will be “connected” to the future point. In physics this property is called *causality*, hence the name “causal sets”.

Figure 1: The graph for the coding problem, with the vertices numbered as implemented in Java.

The causal connections of the points in the causal set are encoded in the *causal matrix*  $c_{mn}$ , where

$$c_{mn} = \begin{cases} 1 & \text{if point } m \text{ is to the past of point } n, \text{ and can send a message to point } n \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

It turns out that all the properties of the continuum spacetime<sup>1</sup> (such as its symmetries and shape) can be derived from the causal matrix. One can therefore view a causal set as a discrete representation of a continuum spacetime, which can be a useful tool to do certain kinds of calculations more easily.

Certain calculations on causal sets can be extremely costly, requiring operations such as finding the eigenvalues of matrices of the same order as the causal matrix. This has to some extent hampered research in this field by requiring the sizes of causal sets to be not more than a few thousand points. The motivation of this project is for it to be a first step in using machine learning in aiding causal set research. If a neural network can be trained to recognize details or estimate quantities in causal sets, it could potentially help accelerate certain computational research programs in the field.

If the causal connections between the causal set points are drawn as lines, the causal set can be viewed as a graph. This graphlike nature means that one might want to look for things like clusters of vertices in the graph, or perhaps more relevantly, the number of connections per vertex.

## 1.2 Variational Autoencoders

A variational autoencoder [2] (VAE) is a type of neural network, which is trained to approximately map the input back to itself. As such, VAE’s are an example of unsupervised learning, since it doesn’t require labels to train the network parameters.

When represented by the size of the hidden layers of the neural network, a VAE typically has an hour-glass shape. That is, starting with the input layer, each subsequent hidden layer has smaller dimensions than the previous layer, until one reaches a layer with the smallest dimension. The elements of this hidden layer are known as the *latent variables*, and represent an *encoding* of the input that is of reduced dimension. The hidden layers following the layer of latent variables each have greater dimension than the previous layers, until one reaches the output layer which has the same dimension as the input layer.

The layers of the VAE up to and including the latent layer are known collectively as the *encoder*, and the layers after the latent variables up to the output are known as the *decoder*. This separation is useful, as the encoder can be used for a low-dimensional representation of the input data and used as an unsupervised clustering algorithm, while the decoder can be used to generate artificial data that nonetheless looks realistic (see e.g. [3] for examples of encoders and decoders using the MNIST [4] dataset). I will be primarily interested in the encoder part of the VAE for now.

One final property of a variational autoencoder is that it implements *variational Bayes*, which means that when training the neural network, the input data is assumed to be from a random distribution. This is represented by the fact that on training the network, a random error is added to the latent variable [2, 5]. The latent variables are represented by the vector  $z_{mean}$ , and the variance of random error is represented by the vector  $z_{logvar}$ . Both of these vectors are (independently) densely connected to the previous layer in the network, and are then combined as

$$z = z_{mean} + \epsilon \exp(z_{logvar}/2), \quad (2)$$

where  $\epsilon$  is a random number drawn from a normal distribution with variance 1 and we are therefore adding a random number to  $z_{mean}$  with variance  $\exp(z_{logvar})$ . This new vector  $z$  is then fully connected to the following layer in the network. The motivation behind this is to find an encoding for each “different” kind of input data (e.g. each different integer digit in MNIST). We find a distribution for each type of data, and then constrain the width of the distribution, to minimize overlap of the different distributions in the latent space. Finally, to generate the encoding of the input data, one needs only to consider  $z_{mean}$ , since the role of  $z_{logvar}$  was only to localize the distribution. Whereas to generate “new” examples with the decoder, one can pick random points in the latent space’s domain, and use the decoder to map it to an output.

In order to train the network, we use the loss function of [5] which is the binary cross-entropy loss plus the KL-loss. The binary cross-entropy is computed on the error between the input and output of the

---

<sup>1</sup>except for volume and distance, since scale is an independent parameter.

network, and while this loss function is normally appropriate for classification problems rather, the nature of the matrix, i.e., consisting of one's and zeros actually makes it a good measure of the loss. The KL-loss is calculated on the variables  $z_{mean}$  and  $z_{logvar}$ , and has the effect of constraining the distributions of the encoded inputs to be standard normals, as explained in [6].

In the next section I will explain what each of the important files in this project does, followed by the results in section 3. Finally, in section 4, I outline next steps in the project.

## 2 Explanation of the Code

## 3 Results

## 4 Future Sections

This is a work in progress, and using causal matrices for causal sets of different dimensions in flat spacetime was the first proof of concept of using VAE's to distinguish between types of causal sets. For my next step, I intend to use causal sets with curvature in addition to the flat ones, to see if two latent variables are still enough to distinguish between them.

This will be done by adding causal diamonds on de Sitter (dS) (with positive curvature) and anti-de Sitter (AdS) (with negative curvature) in various dimensions. The main technical obstacle is calculating the shapes of the statistical distributions for the sprinklings on these spacetimes, while the autoencoder may need extra depth added in order to accommodate the additional complexity of the spaces.

dS and AdS are the simplest examples of curved spacetimes, since like flat spacetime, their curvature is constant at all points. Later on, I could try to use spacetimes with nonconstant curvature, such as black hole or cosmological spacetimes. A potentially exciting result would be if the VAE could encode all these different spacetimes using two, or at least very few, latent variables, in a way that could make it resemble a phase diagram of spacetimes.

Other future projects could include making the neural network more "clever". Right now it will work for a causal set of fixed size only. Ideally it should be able to work for causets of variable sizes.

## 5 Acknowledgements

I am grateful to Jonah Miller for suggesting this project, his helpful discussions, and recommending excellent resources. The variational autoencoder of Keras in [5] was very valuable for this project, and was used with very little modification.

## References

- [1] L. Bombelli, J. Lee, D. Meyer and R. Sorkin, "Space-Time as a Causal Set," *Phys. Rev. Lett.* **59**, 521 (1987).
- [2] D. P. Kingma, M. Welling, "Auto-Encoding Variational Bayes," <https://arxiv.org/abs/1312.6114>,  
C. Doersch, "Tutorial on Variational Autoencoders," <https://arxiv.org/abs/1606.05908>.
- [3] F. Cholet, "Building Autoencoders in Keras," <https://blog.keras.io/building-autoencoders-in-keras.html>,  
M. Shiffman, "Introducing Variational Autoencoders (in Prose and Code)," <http://blog.fastforwardlabs.com/2016/08/12/introducing-variational-autoencoders-in-prose-and.html>
- [4] The MNIST Database of Handwritten Digits, <http://yann.lecun.com/exdb/mnist/>.
- [5] Variational Autoencoder created by Keras team. Available at [https://github.com/keras-team/keras/blob/master/examples/variational\\_autoencoder.py](https://github.com/keras-team/keras/blob/master/examples/variational_autoencoder.py)

- [6] Comment 5 by user radek on <http://forums.fast.ai/t/intuition-behind-kl-divergence-regularization-in-vaes/1650/4>