Can neural networks emulate physics?

Towards a Cosmology Emulator using Generative Adversarial Networks

Mustafa Mustafa
Data Analytics Services, NERSC
Berkeley Lab.

AI@SLAC, SLAC, Stanford University
10/03/2017
The inverse problem of cosmology

The sky surveys collected by observatory experiments pose an inverse problem:

given images of the sky and the “standard model” of cosmology ($\Lambda$CDM), can we extract the cosmological parameters of our universe?
Cosmo Convergence Maps

Weak lensing convergence maps $\kappa(\nu)$ for a $\Lambda$CDM cosmological model.
The central problem of generative models is that given a data distribution $\mathbb{P}_{data}$, can one devise a generator $G$ such that the generated distribution $\mathbb{P}_{model} = \mathbb{P}_{data}$?
Generative Models

The central problem of generative models is that given a data distribution $\mathbb{P}_{data}$ can one devise a generator $G$ such that the generated distribution

$$\mathbb{P}_{model} = \mathbb{P}_{data} ?$$

Our information about $\mathbb{P}_{data}$ comes from an independent and identically distributed sample $x_1, x_2, \ldots, x_n$ which is assumed to have the same distribution as $\mathbb{P}_{data}$. 
Generative Models

Unit Gaussian $Z$ is generated by a generative model (neural net) $\theta$, resulting in a generated distribution $P_{\text{model}}$. The true data distribution $P_{\text{data}}$ is compared to the generated distribution through a loss function, as shown on the blog.openai.com/generative-models.
Density Estimation

Achieving a high fidelity generation scheme amounts to the construction of a density estimator of the training data.
Generative Adversarial Networks

GANs, Goodfellow et al. arXiv:1406.2661
Generative Adversarial Networks

Update discriminator parameters $\mathcal{W}$

$$\bar{z} \xrightarrow[G_\theta]{\longrightarrow} \begin{array}{c} x_{\text{real}} \\ x_{\text{gen}} \end{array} \xrightarrow[D_w]{\longrightarrow} \begin{array}{c} 1 \\ 0 \end{array}$$

Target

GANs, Goodfellow et al. arXiv:1406.2661
Generative Adversarial Networks

Update discriminator parameters $\mathcal{W}$

\[ \tilde{z} \rightarrow G_\theta \rightarrow x_{real} \rightarrow D_w \rightarrow 1 \]

Update generator parameters $\theta$

\[ \tilde{z} \rightarrow G_\theta \rightarrow x_{gen} \rightarrow D_w \rightarrow 1 \]
Generative Adversarial Networks

Update discriminator parameters $\mathcal{W}$

\[ \bar{z} \xrightarrow{G_{\theta}} x_{\text{real}} \xrightarrow{D_w} 1 \quad \bar{z} \xrightarrow{G_{\theta}} x_{\text{gen}} \xrightarrow{D_w} 0 \]

Target

Update generator parameters $\theta$

\[ \bar{z} \sim [\mathcal{N}_0(0, 1), \ldots, \mathcal{N}_{63}(0, 1)] \]

\[ G_{\theta} : \bar{z} \rightarrow x \in \mathbb{R}^{256 \times 256} \]
Generative Adversarial Networks – Loss function

Minimax game formulation (saturating):

\[ J^{(D)} = -\frac{1}{2} \mathbb{E}_{x \sim \mathbb{P}_{data}} \log D(x) - \frac{1}{2} \mathbb{E}_{z \sim p_z} \log(1 - D((G(z)))) \]

\[ J^{(G)} = -J^{(D)} \]
Generative Adversarial Networks – Loss function

Minimax game formulation (saturating):

\[
J^{(D)} = -\frac{1}{2} \mathbb{E}_{x \sim P_{\text{data}}} \log D(x) - \frac{1}{2} \mathbb{E}_{z \sim p_z} \log(1 - D(G(z)))
\]

\[
J^{(G)} = -J^{(D)}
\]
Generative Adversarial Networks – Loss function

Minimax game formulation (saturating):

\[
J^{(D)} = -\frac{1}{2} \mathbb{E}_{x \sim \mathcal{P}_{data}} \log D(x) - \frac{1}{2} \mathbb{E}_{z \sim p_z} \log (1 - D((G(z))))
\]

\[
J^{(G)} = -J^{(D)}
\]

Heuristic loss function (non-saturating):

\[
J^{(G)} = -\frac{1}{2} \mathbb{E}_{z \sim p_z} \log D(G(z))
\]
Deep Convolutional Generative Adversarial Networks (DCGAN)

DCGAN architecture, Radford, Metz and Chintala arXiv:1511.06434
Deep Convolutional Generative Adversarial Networks (DCGAN)

DCGAN architecture, Radford, Metz and Chintala arXiv:1511.06434

DCGAN generated celebrity face images
Deep Convolutional Generative Adversarial Networks (DCGAN)

DCGAN architecture, Radford, Metz and Chintala arXiv:1511.06434

Interpolation in the latent space. Rotations are linear!
Deep Convolutional Generative Adversarial Networks (DCGAN)

DCGAN architecture, Radford, Metz and Chintala arXiv:1511.06434

Arithmetics in the latent space
Generative Adversarial Networks
Evaluation of Generative Models

How close is \( P_{model} \) to \( P_{data} \)?
Evaluation of Generative Models

How close is $P_{model}$ to $P_{data}$?

We think that when it comes to practical applications of generative models, such as in the case of emulating scientific data, the criterion to evaluate generative models is to study their ability to reproduce the characteristic statistics which we can measure on the original dataset.
Kolmogorov-Smirnov two tailed test yields p-value >0.999
Fourier Spectral Analysis

quora.com/Whats-the-use-of-Fast-Fourier-Transform
Fourier Spectral Analysis

https://xkcd.com/26/

Hi, Dr. Elizabeth?
Yeah, uh... I accidentally took the Fourier transform of my cat...

Meow!
Fourier Spectral Analysis: Power Spectrum

\[ \langle \tilde{\kappa}(l)\tilde{\kappa}^*(l') \rangle = (2\pi)^2 \delta_D(l - l') P_\kappa(l) \]

Bands are \( \mu(l) \pm \sigma(l) \)
Fourier Spectral Analysis: Power Spectrum

(b) Validation vs GAN for different $l$ values.

(a) Log-log plot showing the power spectrum for $l(l+1)P_l/2\pi$.

Mustafa Mustafa (2017)
Fourier Spectral Analysis: Power Spectrum

Kolmogorov-Smirnov

<table>
<thead>
<tr>
<th># moments</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>242</td>
<td>&gt; 0.995</td>
</tr>
<tr>
<td>6</td>
<td>&gt; 0.93</td>
</tr>
</tbody>
</table>
Non-Gaussian Corrections

The power spectrum captures the Gaussian structures in the images. However, gravity produces non-Gaussian structures.
Non-Gaussian Corrections

The power spectrum captures the Gaussian structures in the images. However, gravity produces non-Gaussian structures.

The three Minkowski Functionals are sensitive to the higher order correlations.
Non-Gaussian Corrections
## Non-Gaussian Corrections

<table>
<thead>
<tr>
<th># thresholds</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>34</td>
<td>&gt;0.999</td>
</tr>
<tr>
<td>16</td>
<td>&gt;0.97</td>
</tr>
<tr>
<td>6</td>
<td>&gt;0.6</td>
</tr>
<tr>
<td>1</td>
<td>0.32</td>
</tr>
</tbody>
</table>
Generative Models for Emulating Scientific Data

Model parameters

\[
\begin{bmatrix}
\sigma_1 \\
\sigma_2 \\
\vdots \\
\sigma_m
\end{bmatrix}
\]

Physical Model

\[S(\bar{\sigma}, r)\]

Data of interest, images or otherwise.
Generative Models for Emulating Scientific Data

Random seed(s) \( r \)

Model parameters

\[
\begin{bmatrix}
\sigma_1 \\
\sigma_2 \\
\vdots \\
\sigma_m
\end{bmatrix}
\]

Physical Model

\( S(\bar{\sigma}, r) \)

Data of interest, images or otherwise.

\( \mathbb{P}_{data} \)
Generative Models for Emulating Scientific Data

Random seed(s) $r$ → Physical Model $S(\bar{\sigma}, r)$ → How close can it get?

Model parameters $\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_m \end{bmatrix}$

Random vector $z \sim \text{prior}$ → Generative Model $G(z)$ → $P_{\text{data}}$ → $P_{\text{model}}$

Mustafa Mustafa (2017)
Generative Models for Emulating Scientific Data

Can we use GANs to build parametric density estimators $G(\bar{\sigma}, z)$ of scientific data?

Random seed(s) $r$

Model parameters $\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_m \end{bmatrix}$

Physical Model $S(\bar{\sigma}, r)$

Generative Model $G(\bar{\sigma}, z)$

Random vector $z \sim \text{prior}$
Can we use GANs to build parametric density estimators $G(\bar{\sigma}, z)$ of scientific data?
Generative Models for Emulating Scientific Data

Can we use GANs to build parametric density estimators $G(\bar{\sigma}, z)$ of scientific data?

Such generators would likely exclude regions in parameter space where the physical model $S(\bar{\sigma}, r)$ exhibits critical behavior.
CaloGAN: Simulating 3D Calorimeter Showers using GANs

Paganini, de Oliveira and Nachman arXiv:1705.02355
Summary and Outlook

- We have shown with statistical confidence that GANs can emulate ΛCDM cosmological model convergence maps
  - Fourier spectrum of generated maps match that of a validation dataset
  - Non-Gaussian structures are discovered and emulated by the generator

- Deep generative models have the potential of creating high-fidelity computationally inexpensive emulators of scientific data.
Summary and Outlook

- We have shown with statistical confidence that GANs can emulate ΛCDM cosmological model convergence maps
  - Fourier spectrum of generated maps match that of a validation dataset
  - Non-Gaussian structures are discovered and emulated by the generator

- Deep generative models have the potential of creating high-fidelity computationally inexpensive emulators of scientific data.

Summary and Outlook

- We have shown with statistical confidence that GANs can emulate ΛCDM cosmological model convergence maps
  - Fourier spectrum of generated maps match that of a validation dataset
  - Non-Gaussian structures are discovered and emulated by the generator

- Deep generative models have the potential of creating high-fidelity computationally inexpensive emulators of scientific data.


- Future lines of work:
  - investigate the ability of GANs to interpolate in the parameter space of physical models
  - multi dimensional data: 1 & 3D (see CaloGAN), time dependence, “sequential data”
  - using NN interpretation techniques to gain insight in what these networks are learning