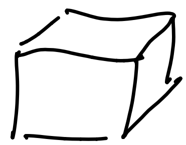
Statistical Physics, Monte Carlo, and Stochastic Neural Networks

Goal: derive a generative model for (quantum) state reconstruction.

Setting:  $\rightarrow$

$P(\vec{x}), \psi(\vec{x})$

$$\left.\begin{aligned} \vec{x}_1 &= (0,1,1,\cdots,1) \\ \vec{x}_2 &= (1,0,1,\cdots,0) \\ &\vdots \\ \vec{x}_{10,000} &= (1,1,0,\cdots,1) \end{aligned}\right\} \mathcal{D}$$

use it to determine (approximately) the unknown (underlying) $P(\vec{x})$ or $\psi(\vec{x})$

Generative modelling: given $p(\vec{x}) \approx P(\vec{x})$ use it to determine the liklihood of $\vec{x}$ not in the dataset $\vec{x} \in \mathcal{D}$

<u>or</u> use $p(\vec{x})$ to produce new $\vec{x}$.

Our approach: use a stochastic neural network called a restricted Boltzmann machine (RBM).

End result: calculating estimators $\langle O \rangle$ from the model $p(x)$.

Bored? MC_Tutorial/Tutorial_MC.pdf

$\rightarrow$ 2D Ising Model $H = -J \sum_{\langle ij \rangle} \sigma_i^z \sigma_j^z$

$\hookrightarrow \sigma^z = x$

Review: thermal distribution

$$P(x) = \frac{1}{Z} e^{-\frac{H}{k_B T}} = \frac{1}{Z} e^{-H\beta}$$

Consider it "ground truth".

- $x = \sigma^z$ is the "state" of the system
- E_x is its energy ($E = H$)
- P.F. $Z = \sum_x e^{-\beta E_x}$

$\uparrow$ trace is hard $\sim 2^N$

Recall the definition of an "expectation value"

$$E[\mathcal{O}] = \langle \mathcal{O} \rangle \underset{\text{qm}}{=} \frac{1}{Z} \langle \psi | \tilde{\mathcal{O}} | \psi \rangle$$

$Z = \langle \psi | \psi \rangle$

$$= \sum_x P(x)\, \mathcal{O}_x$$

e.g.) Boltzmann dist: $\langle \mathcal{O} \rangle = \frac{1}{Z} \sum_x \mathcal{O}_x e^{-\beta E_x}$

<u>Briefly</u>: what expectation values are interesting?

e.g.) energy $U \equiv \langle E \rangle = \frac{1}{Z} \sum_x E_x e^{-\beta E_x}$

note $U = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial \log Z}{\partial \beta}$

e.g.) specific heat $C \equiv \frac{\partial U}{\partial T}$

$$C = \frac{\partial \beta}{\partial T} \frac{\partial U}{\partial \beta} = -\frac{k_B}{(k_B T)^2} \frac{\partial U}{\partial \beta}$$

$$= k_B \beta^2 \frac{\partial^2 \log Z}{\partial \beta^2}$$

now

$$\langle (E - \langle E \rangle)^2 \rangle = \langle E^2 \rangle - \langle E \rangle^2$$

$$= \frac{1}{Z} \sum_x E_x^2 e^{-\beta E_x} - \left(-\frac{1}{Z} \frac{\partial Z}{\partial \beta} \right)^2$$

$$= \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2} - \left(\frac{1}{Z} \frac{\partial Z}{\partial \beta} \right)^2$$

$$= \frac{\partial^2 \log Z}{\partial \beta^2}$$ (confirm with chain rule)

ie. $$\langle E^2 \rangle - \langle E \rangle^2 = \frac{C}{k_B \beta^2}$$

in practice estimate this using $p(x) \simeq P(x)$

Basic "data driven" approach:

take $\mathcal{D} \equiv \{\vec{x}_k\}$

define an approximate $p(x) \simeq P(x)$

e.g.) simplest $p(\vec{x}) = \frac{1}{\|\mathcal{D}\|} \sum_{\vec{x}_k \in \mathcal{D}} \delta_{\vec{x}, \vec{x}_k}$

then from
$$\langle \mathcal{O} \rangle = \sum_x P(x) \mathcal{O}_x$$
$$\approx \frac{1}{\|\mathcal{D}\|} \sum_{\vec{x}_k \in \mathcal{D}} \mathcal{O}_{x_k}$$

exact only when $\|\mathcal{D}\| \to \infty$, however often we have a small finite set of samples (think 10,000)

We want our approximation to "generalize" well on unseen data $x_k \in \mathcal{D}$.

Machine learning ideas can help us here through a clever parametric approach (think interpolation)

Markov chain Monte carlo

Philosophy is to use MCMC to produce $\mathcal{D}$.
The goal of the technique is to produce $\mathcal{D}$ efficiently (and with the correct $P(x)$)

- important: $\mathcal{D} = \{\vec{x}_k\}$ isn't produced randomly but through "importance sampling"

idea: begin with a configuration $\vec{x}_1$ and modify to produce $\vec{x}_2$

$$\vec{x}_1 \to \vec{x}_2 \to \vec{x}_3 \cdots \to \vec{x}_{\|\mathcal{D}\|}$$

we simply need an appropriate "transition probability" from $\vec{x}_k \to \vec{x}_{k+1}$ (call $\mu \to \nu$)

$$T(\mu \to \nu) \geq 0 \qquad \sum_{\nu} T(\mu \to \nu) = 1$$

To approximate $P(\vec{x})$ correctly these must satisfy

1. detailed balance
2. ergodicity

Next: use d.b. to compare Metropolis updates with "Block Gibbs" updates of RBM.