Global Optimization through Normalizing Flow

Longli Zheng

Suppose that we are trying to solve the global optimization problem of the form

$$\boldsymbol{x}^* = \operatorname{argmin}_{\boldsymbol{x}} V(\boldsymbol{x})$$
$$\text{s.t. } \boldsymbol{x} \in \mathbb{R}^n$$

which is equivalent to sampling the distribution

$$\rho(\boldsymbol{x}) = \frac{e^{-\beta V(\boldsymbol{x})}}{Z},$$

where Z is the normalization factor, which is usually intractable.

Following the spirit of normalizing flow, we assume the complicated distribution can be derived from the Gaussian distribution by a bijection f, namely,

$$\rho^*(\boldsymbol{x}) = \mathcal{N} \circ f(\boldsymbol{x}) \left| \det \frac{\partial f}{\partial \boldsymbol{x}} \right|,$$

where $\mathcal{N}(\boldsymbol{z})$ is the standard n-dimensional normal distribution

$$\mathcal{N}(\boldsymbol{z}) = \frac{1}{\sqrt{2\pi}} e^{-\boldsymbol{z}^2/2}.$$

Also notice that the distribution $\rho^*(\boldsymbol{x})$ is already normalized.

$$\int \rho^*(\boldsymbol{x}) d\boldsymbol{x} = \int \mathcal{N} \circ f(\boldsymbol{x}) \left| \det \frac{\partial f}{\partial \boldsymbol{x}} \right| d\boldsymbol{x} = \int \mathcal{N}(\boldsymbol{z}) d\boldsymbol{z}$$

The loss function chosen to train the flow can be the KL divergence.

$$\begin{aligned}
\mathcal{L} = \mathrm{KL}(\rho^* \parallel \rho) &= \int d\boldsymbol{x} \rho^*(\boldsymbol{x}) [\log \rho^*(\boldsymbol{x}) - \log \rho(\boldsymbol{x})] \\
&= \int d\boldsymbol{x} \rho^*(\boldsymbol{x}) \left[\beta V(\boldsymbol{x}) + \log Z - \frac{1}{2} f(\boldsymbol{x})^2 + \log \left| \det \frac{\partial f}{\partial z} \right| - \frac{1}{2} \log 2\pi \right] \\
&= \mathbb{E}_{\tilde{\boldsymbol{x}} \sim \rho^*(\boldsymbol{x})} \left[\beta V(\boldsymbol{x}) - \frac{1}{2} f(\boldsymbol{x})^2 + \log \left| \det \frac{\partial f}{\partial z} \right| \right] + \text{const.}
\end{aligned}$$

and the sample $\tilde{\boldsymbol{x}}$ is easy to get.

$$\tilde{\boldsymbol{z}} \sim \mathcal{N}(\boldsymbol{z}), \tilde{\boldsymbol{x}} = f^{-1}(\tilde{\boldsymbol{z}})$$