神经网络拟合哈密顿量

> [Henry W Lin and Max Tegmark. Why does deep and cheap learning work so well? arXiv preprint arXiv:1608.08225 , 2016.]

神经网络有效性

理论目标

多层网络可写为

$$f(\vec{y}) = \hat{\sigma}_n \hat{W}_n \cdots \hat{\sigma}_1 \hat{W}_1 \vec{y}$$

令:

$$H_x(\vec{y}) \stackrel{def}{=} -\ln p(\vec{y}|x)$$

$$\mu_x \stackrel{def}{=} -\ln p(x)$$

则:

$$p(x|\vec{y}) = \frac{1}{Z(\vec{y})} e^{-(H_x(\vec{y})+\mu_x)} = \hat{\sigma}(-(H_x(\vec{y}) + \mu_x))$$
$$\left(Z(\vec{y}) = \int e^{-(H_x(\vec{y})+\mu_x)} dx \right)$$

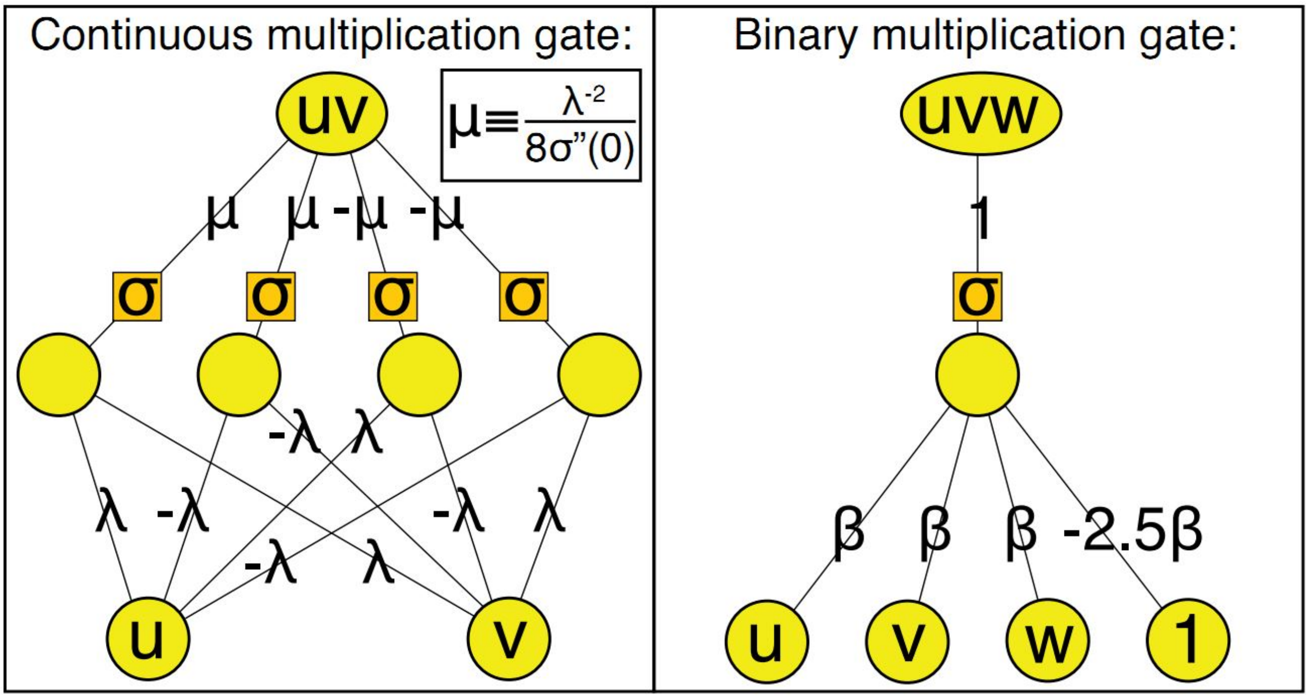
故只要多层网络能估计$-(H_x(\vec{y}) + \mu_x)$，则再加一层神经元为softmax即可

任意精度拟合

由 $\sigma(x) = \sigma + \sigma' x + \sigma'' x^2 + O(u^3)$ 得:

$$\frac{1}{8\sigma'' \lambda^2} [\sigma(\lambda u + \lambda v) + \sigma(-\lambda u - \lambda v) - \sigma(\lambda u - \lambda v) - \sigma(-\lambda u + \lambda v)] = uv(1 + O(\lambda^2 u^2 + \lambda^2 v^2))$$

故理论上可用神经网络以任意精度表示乘积



微扰项

$$H_x(\vec{y}) = H + \sum_i H' i y_i + \sum ijH'' ijy_iy_j + \sum ijkH'''_{ijk}y_iy_jy_k + \cdots$$

故理论上可用神经网络表示任意$H_x(\vec{y})$

高效

以下限制使得参数数量有限

- 低阶项：$H_x(\vec{y})$ 只需展开到有限项，无需太高阶项
- 局域性：只有短程作用，很多$H^{(k)}$ 为0
- 对称性：很多$H^{(k)}$ 不独立，相互依赖

分层

类似重整化群理论，从低层次信息中抽象出（近似）统计充分量作为高层次信息，逐层提取直至所需。