

# 基于ReLU神经网络近似的全局优化方法与应用

张子轩<sup>1,2</sup>, 曾玉娇<sup>1</sup>, 肖焯<sup>1</sup>

(<sup>1</sup>中国科学院过程工程研究所, <sup>2</sup>中国科学院大学)

## 引言

数学模型在复杂系统优化中扮演关键角色。传统基于模拟的优化方法成本高, 代理模型如机器学习成为替代, 但需要使用无导数算法求解; MINLP模型的全局最优解也存在挑战。研究显示, 解析形式的ReLU激活的神经网络(ReLU NN)可通过混合整数线性规划求解。

本文通过换热网络和乙酸乙酯生产系统案例, 探讨了模型近似和求解算法对优化的影响, 旨在为工程优化提供新思路。

## ReLU NN的解析形式

### A 基于Big-M松弛的ReLU函数

$$u = \max\{0, x\} \Leftrightarrow \begin{cases} x \leq u \leq x + M(1-\gamma) \\ 0 \leq u \leq M\gamma \\ \gamma \in \{0, 1\} \end{cases}$$

式中:  $M$  为一个很大的正数;  $\gamma$  为二元松弛变量;  $x$  为激活前函数值;  $u$  为激活后的函数值。

### B 神经网络的解析形式

$$y_{MLP} = W_{out}x_{out}^{h_n} + b_{out}$$

$$x_{out}^{h_i} = \Phi(x_{in}^{h_i}) \quad i = 1, \dots, n$$

$$x_{in}^{h_{i+1}} = W_{i,i+1}x_{out}^{h_i} + b_{i,i+1} \quad i = 1, \dots, n-1$$

$$x_{in}^{h_1} = W_{in}x + b_{in} \quad b_{i,i+1}$$

式中:  $y_{MLP}$  为多层感知机的预测值;  $W_{in}$ 、 $W_{out}$  和  $W_{i,i+1}$  分别为输入层到第1个隐藏层、第  $n$  个隐藏层到输出层、第  $i$  个隐藏层到第  $i+1$  个隐藏层的权重;  $b_{in}$  和  $b_{out}$  分别为输入层到第1个隐藏层、第  $n$  个隐藏层到输出层、第  $i$  个隐藏层到第  $i+1$  个隐藏层的偏置;  $x$  为输入特征;  $x_{in}^{h_i}$  和  $x_{out}^{h_i}$  分别为第  $i$  个隐藏层激活前后的值;  $\Phi(\cdot)$  为激活函数。

### C 不等式约束

$$y_{MLP} \leq 0$$

## 案例分析

ReLU神经网络(NN)用于两类场景:

- 对已知非线性模型进行混合整数线性近似
- 实现数据驱动模型的解析化。

ReLU NN的结构对解的质量和求解速度有显著影响。通过粒子群算法优化ReLU NN的超参数, 包括网络层数和隐藏层神经元数量。为简化优化模型, 限制隐藏层不超过3层, 每层神经元不超过100个。

## A 换热网络优化

### 1 部分近似的MIQCP模型

MINLP模型的非线性特性主要来源于对数平均温差公式, 本研究采用ReLU神经网络进行混合整数线性近似。该网络结构包含1个隐藏层, 共100个神经元, 测试集的均方误差达到0.0119。

$$LMTD_{i,j,k} = W_{out}x_{i,j,k}^{out} + b_{out}$$

$$x_{i,j,k}^{in} = W_{in} \begin{bmatrix} \Delta T_{i,j,k} \\ \Delta T_{i,j,k+1} \end{bmatrix} + b_{in}$$

$$x_{i,j,k}^{in} \leq x_{i,j,k}^{out} \leq x_{i,j,k}^{in} + M(1-\gamma_{i,j,k})$$

$$0 \leq x_{i,j,k}^{out} \leq M\gamma_{i,j,k}$$

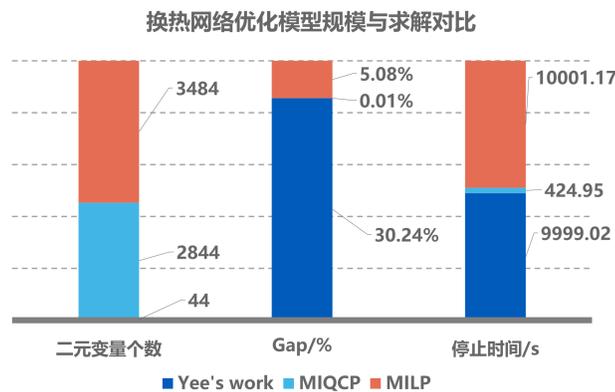
$$\gamma_{i,j,k} \in \{0, 1\}$$

$$r = (\text{area})^{0.6} \Leftrightarrow r^5 = \text{area}^3$$

$$\Rightarrow \begin{cases} r_1 = \text{area}^2 & r_4 = r_5^2 \\ r_2 = r_1 \text{area} & r_5 = r_4 \\ r_3 = r^2 & r_2 = r_5 \end{cases}$$

### 3 优化结果

所有优化结果都得到相同的换热网络结构, 如右图所示, 但每个换热器的换热量不同。



### 2 完全近似的MILP模型

对于计算目标函数中换热网络面积费用的幂函数一项进行混合整数线性近似。所建立的ReLU NN共2个隐藏层, 每个隐藏层含有20个神经元, 隐藏层使用ReLU函数激活, 测试集均方误差为0.334。

$$\left( \frac{y_{i,j,k} - y_{i,j,k}^{mean}}{y_{i,j,k}^{var}} \right)^{0.5} = W_{out}x_{i,j,k}^{h_2} + b_{out}$$

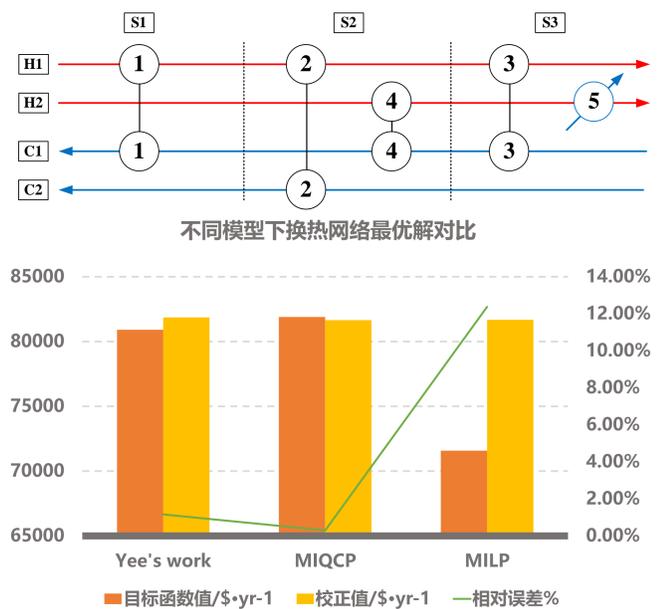
$$x_{i,j,k}^{h_1} = W_{in} \left[ \begin{matrix} (LMDT_{i,j,k} - LMDT_{i,j,k}^{mean}) / (LMDT_{i,j,k}^{var})^{0.5} \\ (Q_{i,j,k} - Q_{i,j,k}^{mean}) / (Q_{i,j,k}^{var})^{0.5} \end{matrix} \right] + b_{in}$$

$$x_{i,j,k}^{h_2} = W_{i_2}x_{i,j,k}^{h_1} + b_{i_2}$$

$$x_{i,j,k}^{h_1} \leq x_{i,j,k}^{h_2} \leq x_{i,j,k}^{h_1} + M(1-\gamma_{i,j,k}^h)$$

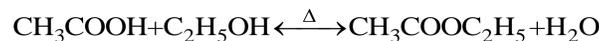
$$0 \leq x_{i,j,k}^{h_2} \leq M\gamma_{i,j,k}^h$$

$$\gamma_{i,j,k}^h \in \{0, 1\}$$



## B 乙酸乙酯生产优化

现考虑一个简单的乙酸乙醇制乙酸乙酯的反应器如右图, 存在如下反应:



假设进料摩尔流量一定, 进料摩尔分率及反应温度为操作变量, 乙酸乙酯的质量流量和进料热负荷为状态变量。数据来源为Aspen模拟。ReLU NN含有3个隐藏层, 每个隐藏层含有20个神经元, 测试集相对误差为0.019。

优化模型如下:

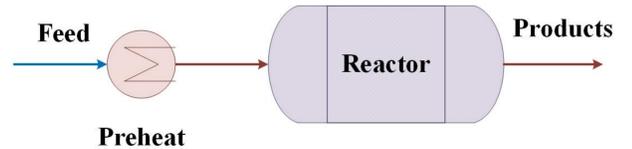
$$\min \sum_{i \in R} \text{cost}_i^r x_i F + \text{cost}^u q - \sum_{j \in P} \text{price}_j m_j$$

$$= 74511.744x_{\text{acetic}} + 124962.404x_{\text{ethanol}} + 0.0029q - 10m_{\text{ethyl}}$$

$$\text{s.t. } ML(x_i, T) \longrightarrow q, m_j$$

$$\sum_{i \in R} x_i = 1$$

$$T \in [0, 200] \quad x_i \in [0, 1]$$



参数	进料流量 /kmol·h <sup>-1</sup>	反应器类型	反应器直径/m	反应器长度/m	正反应指前因子	正反应活化能 /J·kmol <sup>-1</sup>	逆反应指前因子	逆反应活化能 /J·kmol <sup>-1</sup>
值	388	绝热	0.3	2	1.9×10 <sup>8</sup>	5.95×10 <sup>7</sup>	5×10 <sup>7</sup>	5.95×10 <sup>7</sup>

	解析表达		非解析表达	
	种群数量	收敛时间/s	种群数量	收敛时间/s
最优化	-	5.76	500	26.60
		0.44	0.44	0.44
		0.56	0.56	0.56
	/K	30.7	30.8	30.7
	/kW	124.62	126.12	124.62
	/ton·hr <sup>-1</sup>	10.86	10.85	10.86
最优值/CNY·hr <sup>-1</sup>	-11838.33	-11835.79	-11838.33	-11838.33

## 结论

本文以换热网络和生产工况优化为例, 探讨了ReLU NN在全局优化中的应用:

- 优化模型有效性: 使用ReLU NN的解析表达优化模型在换热网络和生产工况优化中有效;
- 换热网络优化案例: 部分近似的MIQCP模型在保证小误差的同时快速收敛;
- 乙酸乙酯生产案例: 模型在乙酸乙酯生产中不仅收敛到最优值, 求解速度提升78.35%;
- 线性化与误差: 对MINLP问题非线性项的过度线性化可能增加误差和求解难度;
- 数据驱动优化模型求解算法: 需要进一步开发高效求解算法。