

神经网络在生化过程动态建模中的应用*

胡泽新

邵惠鹤

(华东化工学院自动化所·上海,200237) (上海交通大学自动化所,200030)

摘要: 本文将多层前馈神经网络用于螺旋霉素发酵过程的动态建模,获得了满意的结果. 为高度复杂的非线性生化过程的模型化提供了一条新的途径. 该方法只需一些量测值,实施方便.

关键词: 神经网络; 非线性系统; 动态建模; 螺旋霉素发酵

1 引言

微生物发酵过程的模型化是实现生产过程优化、控制的关键一步. 对于具有悠久历史,而生产水平又相对落后的发酵工业,对象数学模型的开发尤其显得重要.

微生物发酵的模型可分为两大类,即非结构模型和结构模型. 非结构模型是建立在工程水平上的,它把培养液中各物质作为“黑箱”,不考虑菌体代谢过程中内部结构和组分的变化. 换句话说,非结构模型仅假设用单变量(如菌丝数量或干重)来描述生物相. 结构模型是通过菌体内部结构物质的变化来研究菌体代谢规律,它把生物相分成两到多个化学组分来进行描述. 这些组分可以是非规范组分(如合成组分、结构性组分等). 因此,这类模型是建立在细胞分子水平上的,它考虑到代谢过程中酶活性的变化和营养成分的变异.

微生物发酵过程的建模方法亦可分为两大类,即机理建模方法和系统辨识与参数估计方法. 发酵过程的机理复杂,通过机理建模难度大. 另外,由于发酵过程是一个时变、非线性多变量输入输出关联系统,因此给辨识和参数估计带来了困难. 特别是对于一些高度复杂的非线性生化过程,以及在缺乏系统有关足够先验知识的情况下,这两种传统的建模方法均难以奏效.

针对以上方法的不足,本文将多层前馈神经网络理论^[1]用于螺旋霉素发酵过程动态模型的建立,为高度复杂的非线性生化过程的建模提供了一条新的途径.

2 建模

考虑如下一般形式的非线性系统

$$x_{k+1} = f(x_k, u_k). \quad (1)$$

其中 $x_k \in R^r$ 为 k 时刻的状态向量, $f(\cdot) \in R^r$ 为非线性函数向量,它把 k 时刻的输入映射为 $k+1$ 时刻的状态.

假设(1)式的内部结构未知,仅能获得 x_k, u_k 和 x_{k+1} 的一系列量测值, (1)式的非线性系统动态特性可以通过神经网络的输入输出样本之间的非线性映射得以反映. (1)式的

* 国家教委博士点基金资助项目.

本文于1991年8月17日收到. 1992年4月12日收到修改稿.

神经网络描述形式为

$$x_{k+1} = f(x_k, u_k, w, t, \theta, \beta). \quad (2)$$

式中 w, t, θ, β 为神经网络参数, 即模型参数.

把已知的输入输出测量值做为学习样本, 训练网络的突触联系强度, 阈值等学习参数. 直至网络输出对预定输出的误差小到符合要求, 方可结束学习训练过程.

3 学习算法

在这一小节里, 我们给出神经网络的学习算法, 以求得学习参数 w, t, θ, β . 在神经网络系统中用广义 S 型函数描述神经元的输入-输出特性

$$O_{i,k} = \frac{\theta_{i,k}}{1 + e^{-\beta_{i,k} \text{net}_{i,k}}}. \quad (3)$$

式中 $\theta_{i,k}$ 控制节点输出的饱和值; $\beta_{i,k}$ 则调节 S 型函数的斜率; $O_{i,k}$ 为第 k 层第 i 节点的输出; $\text{net}_{i,k}$ 为第 k 层第 i 节点的内部状态或称输入, 由下式给出

$$\text{net}_{i,k} = \sum_j [w_{i,j,k} O_{j,k-1}] + t_{i,k}. \quad (4)$$

式中 $w_{i,j,k}$ 为第 $(k-1)$ 层第 j 节点对第 k 层第 i 节点的突触联系强度, $t_{i,k}$ 为第 k 层第 i 节点的阈值.

对于给定的输入输出对, 误差定义为

$$E = \frac{1}{2} \sum_{j=1}^q (d_j - O_{j,n})^2. \quad (5)$$

式中 n 为神经网络总的层数, q 为神经网络输出节点数, $d_j = x_{d_j}$ 为给定的输出. 误差利用目前的学习参数值, 在前向通道中计算之.

学习过程是通过适当调整学习参数 $w_{i,j,k}, t_{i,k}, \theta_{i,k}$ 和 $\beta_{i,k}$ 使 E 趋于极小实现的, 这就要求关于学习参数的偏导, 学习参数的修正公式为

$$\begin{aligned} w_{i,j,k}^{(r)} &= w_{i,j,k}^{(r-1)} - \eta_1 \left(\frac{\partial E}{\partial w_{i,j,k}} \right)^{(r)}, \\ t_{i,k}^{(r)} &= t_{i,k}^{(r-1)} - \eta_2 \left(\frac{\partial E}{\partial t_{i,k}} \right)^{(r)}, \\ \theta_{i,k}^{(r)} &= \theta_{i,k}^{(r-1)} - \eta_3 \left(\frac{\partial E}{\partial \theta_{i,k}} \right)^{(r)}, \\ \beta_{i,k}^{(r)} &= \beta_{i,k}^{(r-1)} - \eta_4 \left(\frac{\partial E}{\partial \beta_{i,k}} \right)^{(r)}. \end{aligned} \quad (6)$$

其中 r 为迭代步数, $\eta_1, \eta_2, \eta_3, \eta_4$ 为学习步长.

利用(3)式和(4)式可得到如下偏导计算公式

$$\begin{aligned} \frac{\partial \text{net}_{i,k}}{\partial w_{i,j,k}} &= O_{j,k-1}, \\ \frac{\partial \text{net}_{i,k}}{\partial t_{i,k}} &= 1, \\ \frac{\partial O_{i,k}}{\partial \theta_{i,k}} &= \frac{O_{i,k}}{\theta_{i,k}}, \\ \frac{\partial O_{i,k}}{\partial \beta_{i,k}} &= O_{i,k} \left(1 - \frac{O_{i,k}}{\theta_{i,k}} \right) \text{net}_{i,k}, \end{aligned} \quad (7)$$

$$\frac{\partial O_{i,k}}{\partial \text{net}_{i,k}} = O_{i,k} \left(1 - \frac{O_{i,k}}{\theta_{i,k}} \right) \beta_{i,k}$$

注意(7)式中所有偏导均可在前向通道上由各个神经元局部计算之。设

$$\begin{aligned} \frac{\partial E}{\partial \text{net}_{i,k}} &= -\delta_{i,k}, \\ \frac{\partial E}{\partial O_{i,k}} &= -\xi_{i,k}. \end{aligned} \quad (8)$$

E 有关学习参数的偏导可以表示为

$$\begin{aligned} \frac{\partial E}{\partial w_{i,j,k}} &= \left(\frac{\partial E}{\partial \text{net}_{i,k}} \right) \left(\frac{\partial \text{net}_{i,k}}{\partial w_{i,j,k}} \right) = -\delta_{i,k} O_{j,k-1}, \\ \frac{\partial E}{\partial \theta_{i,k}} &= \left(\frac{\partial E}{\partial \text{net}_{i,k}} \right) \left(\frac{\partial \text{net}_{i,k}}{\partial \theta_{i,k}} \right) = -\delta_{i,k}, \\ \frac{\partial E}{\partial O_{i,k}} &= \left(\frac{\partial E}{\partial O_{i,k}} \right) \left(\frac{\partial O_{i,k}}{\partial \theta_{i,k}} \right) = -\xi_{i,k} \left(\frac{O_{i,k}}{\theta_{i,k}} \right), \\ \frac{\partial E}{\partial \beta_{i,k}} &= \left(\frac{\partial E}{\partial O_{i,k}} \right) \left(\frac{\partial O_{i,k}}{\partial \beta_{i,k}} \right) = -\xi_{i,k} \left[O_{i,k} \left(1 - \frac{O_{i,k}}{\theta_{i,k}} \right) \text{net}_{i,k} \right]. \end{aligned} \quad (9)$$

(9)式中除 $\delta_{i,k}$ 和 $\xi_{i,k}$ 之外的所有量均在前向通道计算得到, δ 和 ξ 通过误差在神经网络上反向传播计算之。

考虑输出层

$$-\delta_{i,n} = \frac{\partial E}{\partial \text{net}_{i,n}} = \left(\frac{\partial E}{\partial O_{i,n}} \right) \left(\frac{\partial O_{i,n}}{\partial \text{net}_{i,n}} \right). \quad (10)$$

利用(5)式和(7)式有
$$\delta_{i,n} = (d_i - O_{i,n}) O_{i,n} \left(1 - \frac{O_{i,n}}{\theta_{i,n}} \right) \beta_{i,n}. \quad (11)$$

$\xi_{i,n}$ 由下式决定
$$-\xi_{i,n} = \frac{\partial E}{\partial O_{i,n}}.$$

利用(5)式和(7)式有
$$-\xi_{i,n} = (d_i - O_{i,n}). \quad (12)$$

对于较低层
$$-\xi_{i,k} = \frac{\partial E}{\partial O_{i,k}} = \sum_j \left(\frac{\partial E}{\partial \text{net}_{j,k+1}} \right) \left(\frac{\partial \text{net}_{j,k+1}}{\partial O_{i,k}} \right).$$

这里对第 $(k+1)$ 层所有节点求和, 利用(4)式和(8)式有简化结果如下

$$\xi_{i,k} = \sum_j [\delta_{j,k+1} w_{j,i,k+1}]. \quad (13)$$

$\delta_{i,k}$ 由下式决定
$$\delta_{i,k} = \xi_{i,k} O_{i,k} \left(1 - \frac{O_{i,k}}{\theta_{i,k}} \right) \beta_{i,k}. \quad (14)$$

注意 $\xi_{i,k}$ 仅取决于第 $(k+1)$ 层的 δ . 因此, 在同一层里 ξ 可以并行计算, 并且这种计算是逐层进行的. 多层前馈神经网络误差回传并行运算的意义就体现在这里, 它大大加快了学习算法的运行速度, 简化了学习算法的复杂程度.

4 螺旋霉素发酵过程的动态建模

4.1 菌体生长模型

发酵过程模型通常由菌体生长模型、基质消耗模型和产物形成模型等组成. 其核心部分为菌体生长模型, 它又有结构模型与非结构模型之分. 前者通过菌体的内部结构物质的变化来研究菌体代谢规律, 这种模型反映代谢过程的机理; 而后者只依据菌体的外部特

征,描述菌体代谢的宏观规律.由于目前对许多发酵过程的机理尚未研究透彻,对这些过程的生物合成途径也仅知大概,加之缺乏有效的检测手段,所以工程上常采用非结构模型.对于高度复杂的非线性生化过程,以上两种模型均无能为力,这时可用神经网络模型描述之.根据 SPM 分批发酵的实际情况,菌体生长模型可写成

$$\frac{dx}{dt} = f(x). \quad (15)$$

其中 x 为菌体浓度(mg 干重/ml), $f(\cdot)$ 为一非线性函数.

4.2 产物形成模型

SPM 发酵是在一定的细胞成熟时间之后开始分泌抗生素的.因此,SPM 分批发酵产物形成模型可写成

$$\frac{dp}{dt} = g(x_{t-t_m}). \quad (16)$$

其中 p 为次级代谢产物浓度(mg/ml), x 细胞浓度(mg/ml), t 发酵时间(h), t_m 细胞成熟时间(h), $g(\cdot)$ 非线性函数.

考虑到 SPM 发酵有产物抑制这一因素,故对模型加以修正得

$$\frac{dp}{dt} = g(x_{t-t_m}, p). \quad (17)$$

综合式(15)和式(17)便得到 SPM 间歇发酵过程菌体生长和产物生成的动态模型

$$\begin{cases} \frac{dx}{dt} = f(x), \\ \frac{dp}{dt} = g(x_{t-t_m}, p). \end{cases} \quad (18)$$

其中 x 为螺旋霉素菌体浓度(mg 干重/ml), p 为产物螺旋霉素浓度(unit/ml) $\times 10^{-2}$, t_m 为 SPM 产生菌的细胞成熟时间,由实验研究取作 24 小时.实验测得数据如表 1.

表 1 实验测得数据

$t(h)$	0	6	12	18	24	30	36	42	48	54	60	66	72	78	84	90	96
x	3.90	5.86	7.30	12.37	16.45	20.81	22.68	23.71	24.53	24.90	24.80	24.10	24.10	23.91	23.75	23.75	23.52
$P(\times 10^{-2})$						1.60	2.22	4.03	4.63	5.16	8.66	10.32	11.97	14.20	15.88	18.94	18.00

考虑到状态方程中的时滞,在动态建模中,以 $x=5.86$, $p=1.60$ 作为初始状态,且取 x 在 6~72 时间的值分别与 p 在 30~96 时间的值相对应计算.

对(18)式进行离散化,采样间隔为 $T_1=T_2=6(h)$,有

$$\begin{cases} x_{k+1} = f(x_k), \\ p_{k+1} = g(x_{k-4}, p_k). \end{cases} \quad (19)$$

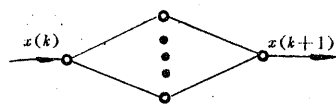
4.3 神经网络模型

假设螺旋霉素发酵过程动态数学模型(19)式的内部结构未知,仅能获得该系统输入输出的一些量测值,用多层前馈神经网络描述其动态特性如下

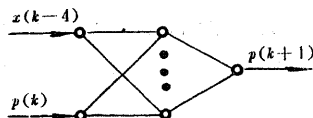
$$\begin{cases} x_{k+1} = f(x_k, w_1, t_1, \theta_1, \beta_1), \\ p_{k+1} = g(x_{k-4}, p_k, w_2, t_2, \theta_2, \beta_2). \end{cases} \quad (20)$$

式中 w, t, θ, β 为神经网络的学习参数。

选择图 1 所示的网络结构来建立螺旋霉素发酵过程的动态模型。



(a) 菌体生长动态模型



(b) 产物形成动态模型

图 1 SPM 发酵过程动态模型的 N-N 描述

图中 $x(k)$ 为菌体生长动态模型的输入, $x(k+1)$ 为输出, $x(k-4)$ 和 $p(k)$ 为产物形成动态模型的输入, $p(k+1)$ 为输出。所有这些输入输出值均可由实验测得。

4.4 仿真结果

首先产生输入输出训练样本, 根据实验数据, 模型(a)有 16 组训练样本, 模型(b)有 11 组训练样本。网络结构为 $\{1, 15, 1\}$, 即三层网络, 输入层一个神经元, 中间层 15 个神经元, 输出层一个神经元。学习参数 w, t, θ, β 的初始值由随机函数产生。学习步长取为 $\eta_1 = \eta_2 = \eta_3 = \eta_4 = 0.3$ 。对于第一组训练样本(a): $x(0) = 3.90, x(1) = 5.86$; (b): $x(1) = 5.86, p(1) = 1.60, p(2) = 2.20$ 。模型(a)学习 16 组数据总计运行时间为 45 分钟, 每一组学习次数为 20 次左右, 时间为 2.5 分钟左右。模型(b)学习 11 组数据总计运行时间为 30 分钟左右, 每一组学习次数为 20 次左右。至此, 神经网络模型建立完毕。

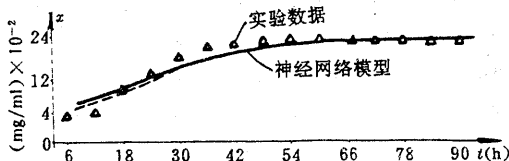


图 2 SPM 菌体生长模型

任取与训练样本不同的数据(a) 70 组, (b) 35 组对所建模型进行检验, 将神经网络模型数据与实验数据放在同一张图上, 如图 2 和图 3。可以看出两者较为接近, 也就验证了本文方法的有效性。

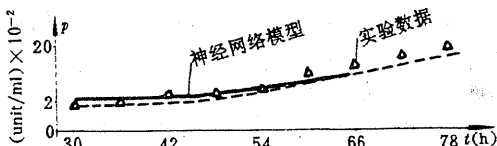


图 3 SPM 产物生成模型

5 结 论

本文方法对多输入多输出系统也适用。

模型(a)中第一组训练样本学习误差 E 的变化过程如图 4 所示, 误差由 26.4 经过 18 次学习趋于 0, 时间为 2.5 分。学习步长 η 的取值直接影响学习的快慢。 η 越小, 学习次数越多, 反之亦然。 η 过小学习时间太长, η 过大会导致学习算法发散。因此, 要选一适当的 η 值。当 η 取值确定后, 每一组训练样本的学习就基本在某一固定次数上下波动。如本文取 $\eta = 0.3$, 每一组训练样本的学习次数均在 20 次左右。

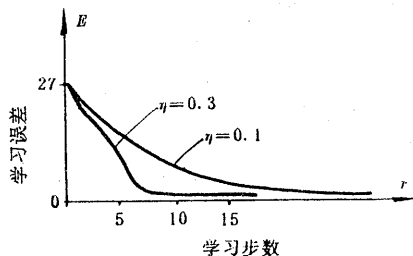


图 4 误差随学习步数的变化情况

6 结束语

本文用多层前馈神经网络建立了螺旋霉素(SPM)发酵过程的动态模型, 获得了满意的结果。该方法无需知道生化反应过程的复杂机理和模型结构, 只需一些输入输出量测

值,简便易行.它为生化过程建模提供了一条新的途径.

参 考 文 献

- [1] Bhat, N. and McAvoy, T. J. . Use of Neural Nets for Dynamic Modeling and Control of Chemical Process Systems. Computers Chem. Engng, 1990, 14(4);573—583

Neural Networks for Dynamic Modelling of Biotechnical Processes

HU Zexin

(East China Institute of Chemical Technology • Shanghai, 200237, PRC)

SHAO Huihe

(Research Institute of Automatic Control, Shanghai Jiaotong University • Shanghai, 200030, PRC)

Abstract: Neural networks are applied to dynamic modelling of spiramycin fermentation. Satisfactory results were obtained. A new way of modelling in very complex nonlinear biotechnical processes is given. Only are some off-line or on-line measurements needed, so it is convenient to be carried out.

Key words: neural networks; nonlinear systems; dynamic modelling; spiramycin fermentation

本文作者简介

胡泽新 1961年生. 1990年获工学博士学位. 后在华东化工学院自动化所任教至今. 现为副教授. 主要研究兴趣是工业过程模型化和控制, 智能控制. 目前研究领域为生化过程建模和控制, 神经网络在过程控制中的应用, 故障诊断和容错控制.

邵惠鹤 1936年生. 教授, 博士生导师. 上海交通大学自动化研究所副所长, 中国自动化学会计算机应用委员会副主任, 中国微生物学会生化过程模型化与控制学会副理事长, 上海市自动化学会理事, 自动化应用委员会和化工计算机应用委员会主任. “化工自动化及仪表”杂志编委. 1960年毕业于华东化工学院化工机械系. 至今已发表60多篇中外文学术论文和出版7本著作. 当前研究的领域是工业过程模型化与计算机优化控制, 智能控制, 生产调度推断估计理论与控制, 稳态优化理论, 人工神经网络理论在建模、估计和控制中的应用. 特别着重于精馏过程和生化过程的建模与计算机优化控制的研究.