

基于神经网络的生化过程预估优化控制

杨一兵 王树青 张泉灵 王骥程

(浙江大学工业控制技术研究所 杭州 310027)

综观历史与现状,生化过程数学模型的建立是很困难的。不多的一些数学模型也往往由于精度低,应用范围窄而无法在实际中应用。这是由于生化过程的机理非常复杂,具有高度非线性和时变特性。并且不同于一般物理过程的是生化过程是个物理上不可逆的过程。近几年来,人工神经网络(ANN)得到了迅速发展,并被广泛应用到各个领域。同样,ANN也为生化过程控制提供了一种新方法。本文以工业生产中发酵过程的补料控制为例,叙述了ANN如何用于生化过程预估和优化控制。相应地,对结果进行了分析讨论。

1 问题描述

工业生产中发酵过程有一种间歇补料生产方式,即在生产中每隔一定时间,根据发酵状况进行补料。这类生产过程的优化控制问题就是求取最优补料曲线,使最终产率为最大。描述如下:

$$\text{目标函数: } \max y(f), \quad (1)$$

$$\text{状态方程: } X(k+1) = F[X(k), U(k)] \quad (2)$$

$$\text{输出方程: } y(k) = CX(k) \quad (3)$$

$$\text{控制约束: } U_{\min} \leq U(k) \leq U_{\max} \quad (4)$$

其中, $k(=0,1,\dots,k_f-1)$ 为发酵时刻(k_f 为发酵终止时刻), $X(k), U(k), y(k)$ 分别为状态变量、控制变量和输出变量, C 是个常数矩阵, $F[\cdot]$ 为非线性隐式函数,无法用经典数学公式描述。为了解决这一优化问题,我们必须首先解决状态预估问题。

2 状态预估

ANN的一个显著优点就是能够通过样本进行训练,学习输入、输出之间的映射关系。BP(Back-propagation,反向传播)网络是其中一种简单、有效的网络。理论上只要提供足够数量的输入、输出样本,BP就可以无限逼近输入样本与输出样本之间的实际映射关系,而不管这种关系是否是线性关系。另一方面,实际生产过程积累了大量的数据,也就是可以提供足够数量的输入、输出样本。因此,通过BP对这些样本的学习而建模是一种切实可行的办法,我们就是选用BP作为某一生化过程的预估器。

2.1 预估器

图1所示的是链霉素间歇发酵过程的预估器结构。各符号的具体含义见符号说明。

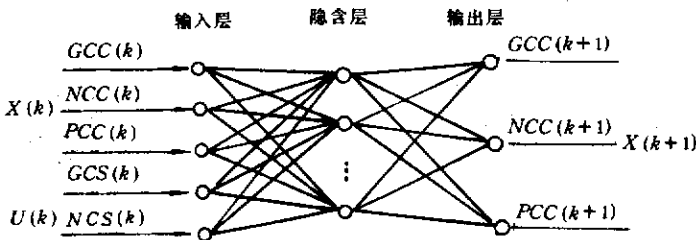


图1 链霉素发酵过程的预估器结构

图中, k ($k=0, 1, \dots, k_f-1$) 是发酵期间补料时刻。网络输入为 k 时刻的发酵状态和控制量, 而输出是下一时刻的状态变量。

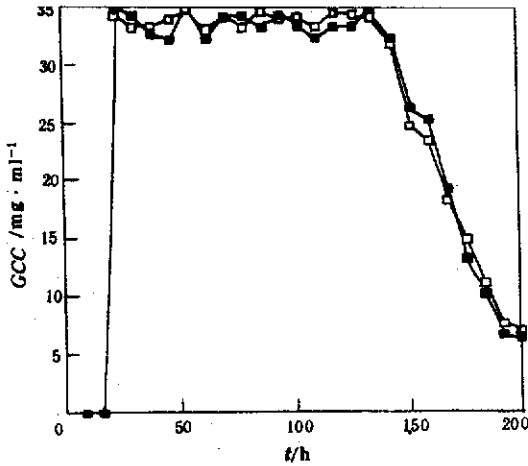


图2 糖浓度预估效果

□ 预估数据, □ 实际数据

具体算法, 假设

$$h_i = \sum_{i=0}^{nin} w_{hi} x_i + \theta_{h_i} \quad (i \in \text{NIN}, 1 \in \text{NHID}) \quad (5)$$

$$\text{hid}_i = \frac{1}{1 + e^{-h_i}} \quad (1 \in \text{NHID}) \quad (6)$$

$$o_j = \sum_{i=0}^{nhid} w_{oj} \text{hid}_i + \theta_{o_j} \quad (1 \in \text{NHID}, j \in \text{NOUT}) \quad (7)$$

$$y_j = \frac{1}{1 + e^{-o_j}} \quad (j \in \text{NOUT}) \quad (8)$$

那么, 输出对权重的灵敏度为:

$$\frac{\partial y_j}{\partial w_{oj}} = \begin{cases} y_j(1 - y_j)\text{hid}_i & (j = u) \\ 0 & (j \neq u) \end{cases} \quad (j, u \in \text{NOUT}, 1 \in \text{NHID}) \quad (9)$$

$$\frac{\partial y_j}{\partial w_{oi}} = y_j(1 - y_j)w_{oj}\text{hid}_i(1 - \text{hid}_i)x_i \quad (j \in \text{NOUT}, 1 \in \text{NHID} \quad i \in \text{NIN}) \quad (10)$$

输出对输入的灵敏度为:

$$\frac{\partial y_j}{\partial x_i} = y_j(1 - y_j) \sum_{i=0}^{nhid} [w_{oj}w_{hi}\text{hid}_i(1 - \text{hid}_i)] \quad (j \in \text{NOUT}, i \in \text{NIN} \quad 1 \in \text{NHID}) \quad (11)$$

分析(9)、(10)、(11)三式, 可以得到以下几个结论:

(1) 三式都有公因式 $y_j(1 - y_j)$, 因此若 y_j 接近 1 或 0, 则三式都接近 0. 因此, 在 1 或 0 附近灵敏度很低. 此时权重和输入的变化对输出影响不大, 具有良好的鲁棒性。

(2) 如果 hid_i 较大, 则 y_j 对相应权重 w_{oj} 的变化比较灵敏. 因此, 若能在训练中优先调整较大 hid_i 相关的权重 w_{oj} 则可以提高收敛速度。

3 优化控制

有了上述足够精度的预估器, 我们给出间歇发酵过程的优化控制方法. 优化目标是发酵终止时产率 $y(k_f)$ 最高. 实际生产中控制手段为各物料的补料速率. 由于选择了发酵罐中各基质浓度为状态变量, 我们把各物料的消耗速率作为控制变量. 得到最优物料消耗速率曲线后, 只需根据物质平衡原理, 可以方便得出各物料补料速率曲线。

预估器通过大量的样本进行训练. 训练初期, 采用梯度下降法, 目标函数是各个样本训练误差平方之和. 梯度下降法在训练初期具有良好的收敛速度, 但在接近稳态时收敛速度急剧下降. 考虑到在稳态附近函数可以近似为二次函数, 采用共轭梯度法可明显提高收敛速度. 另外, 时变会影响预估精度, 因此预估器需要新鲜数据不断刷新. 训练中, 我们采用 50 批数据作为样本. 先采用梯度下降法进行训练. 训练次数达到 22965 次时误差为 0.01. 此时切换到共轭梯度法, 训练次数为 31527 次时误差为 0.001. 认为训练完毕。

作为例子, 图 2 显示了糖浓度的预估效果. 可以认为预估结果令人满意, 方差约为 0.01。

2.2 灵敏度分析

在分析预估器的灵敏度之前, 必须给出 BP 的

首先,我们给出这一优化问题的拉格朗日函数:

$$\mathcal{L} = -y(k_f) + \lambda_1^* [U_{\min} - U(k)] + \lambda_2^* [U^T(k)U_{\max}] \quad (12)$$

这里, λ_1, λ_2 是拉格朗日乘子, $y(k_f)$ 是 $U(k)$ 的高度非线性函数, 并且 $U(k)$ 为 $y(k_f)$ 的一个非显式变量, 因此这一优化问题是一个非奇异问题。

优化控制系统结构如图 3 所示。

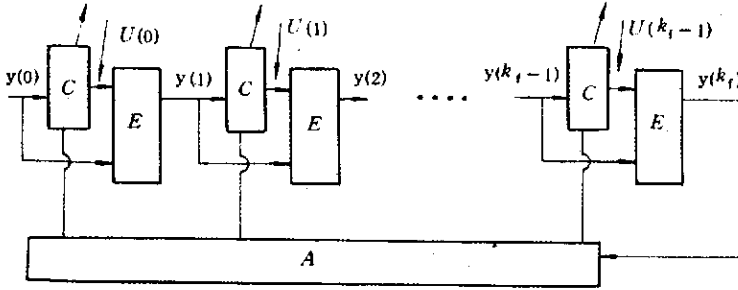


图 3 控制系统结构

图中, C, E 分别为控制器和预估器, 预估器的实现如前所述, 而控制器的实现叙述如下。

3.1 控制器训练算法

如图 3 所示, $U(k)$ 取决于 $X(k)$, 两者之间的非线性关系由一个 BP 网络来实现, 令 $W^i(k)$ 为 C_i 的权重, 则控制器训练算法为:

$$W^{i+1}(k) = W^i(k) - \eta(k) \frac{\partial \mathcal{L}}{\partial W^i(k)} \quad k = 0, 1, \dots, k_f - 1 \quad (13)$$

这里, i 是迭代步骤, $\eta(k)$ 为迭代步长。

由 (12) 式得:

$$\frac{\partial \mathcal{L}}{\partial W^i(k)} = - \frac{\partial y(k_f)}{\partial W^i(k)} \quad (14)$$

再由 (3) 式可知:

$$\frac{\partial y(k_f)}{\partial W^i(k)} = C \frac{\partial X(k_f)}{\partial W^i(k)} \quad (15)$$

$\frac{\partial X(k_f)}{\partial W^i(k)}$ 按下列步骤求取:

$$\frac{\partial X(k+1)}{\partial W^i(k)} = \frac{\partial X(k+1)}{\partial U(k)} \frac{\partial U(k)}{\partial W^i(k)} \quad (16)$$

$$\frac{\partial X(k+2)}{\partial W^i(k)} = \frac{\partial X(k+2)}{\partial X(k+1)} \frac{\partial X(k+1)}{\partial W^i(k)} + \frac{\partial X(k+2)}{\partial U(k+1)} \frac{\partial U(k+1)}{\partial X(k+1)} \frac{\partial X(k+1)}{\partial W^i(k)} \quad (17)$$

⋮

$$\frac{\partial X(k_f)}{\partial W^i(k)} = \frac{\partial X(k_f)}{\partial X(k_f-1)} \frac{\partial X(k_f-1)}{\partial W^i(k)} + \frac{\partial X(k_f)}{\partial U(k_f-1)} \frac{\partial U(k_f-1)}{\partial X(k_f-1)} \frac{\partial X(k_f-1)}{\partial W^i(k)} \quad (18)$$

3.2 优化控制步骤

整个控制过程按以下步骤实现:

步骤 1: 已知 $X(0)$, 令 $k = 0$, 初始化 $\{U(k), U(k+1), \dots, U(k_f-1)\}$

步骤 2: 得到各个发酵时期状态估计值, $\{\hat{X}(k+1), \dots, \hat{X}(k_f)\}$

步骤 3: 训练 $k, k+1, \dots, k_f-1$ 时的 C, E

步骤 4: 求取 $\{U(k), U(k+1), \dots, U(k_f-1)\}$

步骤 5: 把 $U(k)$ 加到实际生产过程中, 进行生产

步骤 6: $k = k+1$, 测量 $X(k)$, 若 $k = k_f$, 结束; 否则转到步骤 2。

4 应用实例

为了检验这一控制策略的有效性,我们考虑某厂 36t 发酵罐中进行的链霉素间歇发酵过程。这个例子中, $U(k) = [u_1(k), u_2(k)]^T$, 这里 $u_1(k), u_2(k)$ 分别是糖、氮消耗率。 $X(k) = [x_1(k), x_2(k)]^T$, 这里 $x_1(k), x_2(k)$ 是发酵罐中的糖浓度, 氮浓度。图 4 显示了最优物料消耗曲线和发酵罐物料浓度曲线。发酵终止时的效价为 26 000u/ml, 而普通产率为 24 000u/ml。

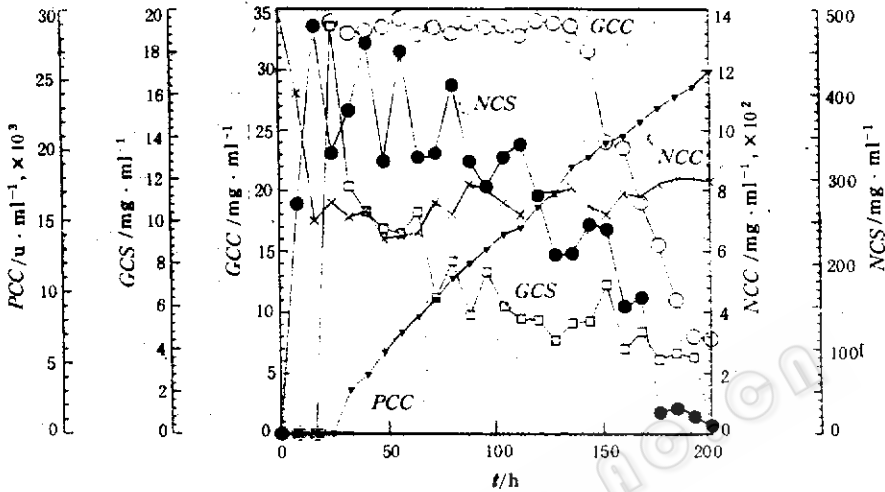


图 4 链霉素发酵过程的最优补料曲线

○ GCC, ● NCS, × NCC, □ GCS, ▼ PCC.

5 结果讨论

综上所述,我们认为 ANN 在生化过程控制领域的应用是可行的。ANN 不需要专家系统必需的先验知识,也不需常规控制所需的数学模型。ANN 不需要系统是线性的,它适用于任何复杂非线性系统,另外,某一发酵过程往往需要一两百个小时,在如此长的时间内,时变是不可避免的。因此,如本文所采取的每个补料周期(8h)单独建模是克服时变的有效方法。

符号说明

k ——补料时刻

$GCC(k)$ —— k 时刻发酵罐中糖浓度/ $\text{mg} \cdot \text{ml}^{-1}$

$NCC(k)$ —— k 时刻发酵罐中氮浓度/ $\text{mg} \cdot \text{ml}^{-1} \cdot 10^2$

$GCS(k)$ —— k 时刻糖消耗率/ $\text{mg} \cdot \text{ml}^{-1}$

$NCS(k)$ —— k 时刻氮消耗率/ $\text{mg} \cdot \text{ml}^{-1}$

$PCC(k), y(k)$ —— k 时刻的产率即效价/ $\text{u} \cdot \text{ml}^{-1} \cdot 10^5$

n_{in} ——输入节点数

n_{hid} ——隐含节点数

n_{out} ——输出节点数

w_{h1i} ——隐含节点 1 与输入节点 i 之间的权重

w_{o1j} ——输出节点 j 与隐含节点 1 之间的权重

θ_{h1} ——隐含节点 1 的阈值

θ_{o1j} ——输出节点 j 的阈值

h_1 ——隐含节点 1 的输入总和

hid_1 ——隐含节点 1 的输出

o_j ——输出节点 j 的输入总和

- y_j ——输出节点 j 的输出
 NIN —— {1, 2, ..., nin}
 $NOUT$ —— {1, 2, ..., nout}
 $NHID$ —— {1, 2, ..., nhid}
 $X(k)$ ——发酵罐基质浓度/ $\text{mg} \cdot \text{ml}^{-1}$
 $U(k)$ ——补料消耗率/ $\text{mg} \cdot \text{ml}^{-1} \cdot \text{s}^{-1}$
 C ——控制器
 E ——预估器

参 考 文 献

- [1] Derrick H N, Bernard W. *Int J Control*, 1991, 54 (6), 1439~1451.
 [2] Modak J M, Lim H C. *Biotechnol Bioeng*, 1987, 30: 528~540.
 [3] Modak J M, Lim H C. *Biotechnol Bioeng*, 1989, 33: 11~15.
 [4] Pekka L, Yihong Z. *Journal of Biotechnology*, 1991, 21: 253~270.

Prediction and Optimal Control for Bioprocess Based on ANNs

Yang Yibing Wang Shuqing Zhang Quanling Wang Jicheng
 (Institute of Industrial Process Control, Zhejiang University, Hangzhou 310027)

Abstract Bioprocess is highly complex process with high variability and little knowledge about its mechanisms. It is very difficult to get mathematical models that are accurate enough for state estimation, modelling and optimization. In this paper, the approach for prediction of a fed- batch fermentation based on artificial neural networks is introduced in first. Then, the method of feedback optimization of this process is presented. Finally, an example of industrial application is introduced, and the result is discussed.

Key words Bioprocess, prediction, optimization, artificial neural networks (ANNs)