

## 青霉素半连续发酵过程状态与参数在线辨识

金沙 张嗣良 俞俊棠

(华东化工学院生化工程研究所, 上海)

本文建立了带有噪声的适合于工业规模青霉素(丝状菌)半连续发酵数学模型,以排气CO<sub>2</sub>释放速率(CER)作为在线观测量,应用推广Kalman滤波器(EKF)进行状态与参数在线辨识,为过程的自适应优化控制奠定了基础。试验表明:滤波得到的状态值与实测值相当吻合,证明滤波器是一种较好的发酵状态及动力学参数变化的在线观察器。

**关键词** 青霉素发酵; 状态估计; 参数辨识; 推广卡尔曼滤波

青霉素发酵与其他生化反应一样,存在着明显的不确定性,具体表现在原材料成分不稳定、菌体本身遗传变异性、设备变化以及微生物对环境多因子变化的敏感性等。这些因素都是不可预知和不可控的<sup>[1]</sup>。同时,由于测量条件的苛刻,使过程的重要状态量的在线测定非常困难,因此长期以来,菌量和产物量等的测定是依靠离线手工分析得到,然而,这些状态变量是过程优化控制的重要参量,状态量的离线测定将造成信息延迟,影响操作人员的正确判断和决策,同时也限制了过程优化控制方案的实施。而过程的随机性又使已有的确定性模型无法推广到随机干扰不可忽略的工业规模发酵中。

因而,需建立能反映过程动态特性及噪声的不确定性模型,以使模型能适应动力学参数慢时变特性,并通过滤波技术,从有限的在线可测量信息中感知不可测状态量信息,为过程优化控制提供依据。

正是基于上述考虑,本文以上海第三制药厂新近从日本引进的青霉素丝状菌(JS-8)为对象,旨在建立带噪声的过程微分模型,并利用EKF,以排气CO<sub>2</sub>浓度为在线观测量,递推估计在线不可测的菌量、残糖浓度及产物浓度等,还对模型中

各参数进行辨识,观察其波动情况,加深对过程理解。该方面的研究,为许多过去难以定量研究的问题开拓了新的研究途径,也为过程的自适应控制提供了必要前提。

## 发酵噪声模型的建立

一般青霉素发酵前40h为菌体快速生长期,产物量极少,发酵第40—180h是产物大量形成期。在其半连续发酵建模中,较为成功的是 Bajpai<sup>[2]</sup>等提出的机理性物料衡算模型:

1. 菌体生长与基质浓度的关系:

$$\frac{dx}{dt} = \frac{\mu_{xm} \cdot x \cdot s}{K_x \cdot x + s} - x \cdot \frac{F}{s_f} \quad (1)$$

2. 产物浓度与基质浓度的关系:

$$\frac{dp}{dt} = \frac{\mu_{pm} \cdot x \cdot s}{K_p + s(1 + s/K_i)} - K \cdot p - p \cdot \frac{F}{s_f} \quad (2)$$

3. 基质消耗速率:

本文于1988年10月4日收到。

本工作得到上海第三制药厂青霉素车间主任诺茂华工程师的大力协助,特此致谢。

$$\begin{aligned} \frac{ds}{dt} = & -\frac{\mu_{xm} \cdot x \cdot s}{Y_{x/s}(K_x \cdot x + s)} - \\ & -\frac{\mu_{pm} \cdot x \cdot s}{Y_{p/s}(K_p + s(1 + s/K_i))} - m \cdot x + \\ & + F - s \cdot \frac{F}{s_f} \end{aligned} \quad (3)$$

式中,  $K_x$  为 Contois 饱和系数,  $K_p$  为 Monod 饱和系数,  $K$  为考虑到青霉素水解的衰退速率系数,  $F$  为基质进料速率,  $\mu_{xm}$ 、 $\mu_{pm}$  分别是最大比生长速率和比产物形成速率,  $Y_{x/s}$  为单位基质消耗量上的菌体得率系数,  $Y_{p/s}$  为产物得率系数,  $m$  是维持所需消耗基质的维持系数。

Calam<sup>[3]</sup> 等研究了青霉素发酵时期  $CO_2$  释放速率 (CER) 与菌体生长间的关系, 提出下述模型:

$$\begin{aligned} \frac{d(CER)}{dt} = & \frac{\mu_{xm} \cdot x \cdot s}{Y_{x/CO_2}(K_x \cdot x + s)} + \\ & + M \cdot x + k_{p,c} \end{aligned} \quad (4)$$

其中,  $Y_{x/CO_2}$  为菌体生长产生的  $CO_2$  得率系数,  $M$  为菌体维持项所产生的  $CO_2$  系数,  $k_{p,c}$  是产物形成所释放的  $CO_2$  量。

上述机理模型与实验室试验结果较为吻合, 但一旦用于描述工业规模发酵过程就存在较大偏差。原因是: Baipai 模型只是基本上反映了青霉素发酵过程的一般结构和特点, 未考虑随机扰动对过程的影响。而在工业规模发酵时, 扰动影响较为显著, 致使一方面扰动引起模型精度下降, 另一方面, 过程动态特性又引起动力学参数波动, 因此, 用确定性模型描述具有不可忽视的随机扰动的发酵过程是不合适的, 必须考虑扰动, 对模型进行修正。本文运用噪声理论, 用一组零均值白噪声序列标度扰动, 提出了下述不确定性模型:

$$\frac{dx}{dt} = \frac{\mu_{xm} \cdot x \cdot s}{K_x \cdot x + s} - x \cdot \frac{F}{s_f} + w_1 \quad (5)$$

$$\begin{aligned} \frac{dp}{dt} = & \frac{\mu_{pm} \cdot x \cdot s}{K_p + s(1 + s/K_i)} - K \cdot p - \\ & - p \cdot \frac{F}{s_f} + w_2 \end{aligned} \quad (6)$$

$$\begin{aligned} \frac{ds}{dt} = & -\frac{\mu_{xm} \cdot x \cdot s}{Y_{x/s}(K_x \cdot x + s)} - \\ & -\frac{\mu_{pm} \cdot x \cdot s}{Y_{p/s}(K_p + s(1 + s/K_i))} - m \cdot x + \\ & + F - s \cdot \frac{F}{s_f} + w_3 \end{aligned} \quad (7)$$

$$\begin{aligned} \frac{d(CER)}{dt} = & \frac{\mu_{xm} \cdot x \cdot s}{Y_{x/CO_2}(K_x \cdot x + s)} + \\ & + M \cdot x + k_{p,c} + w_4 \end{aligned} \quad (8)$$

其中  $w_i$  ( $i = 1, \dots, 4$ ) 代表随机噪声, 它实质上是对模型本身不精确性和过程中各种扰动的在线校正。

## 状态估计与参数辨识

### (一) 推广的状态方程

方程 (5) — (8) 是描述过程动态特性的噪声模型, 其中包含了 12 个待估动力学参数, 必须通过实验数据进行辨识。由于参数的波动性, 可以将参数看作为拟状态量建立推广的状态方程, 并采用 EKF 将参数与状态量  $x$ 、 $p$ 、 $s$ 、 $CER$  一起估计。推广的状态方程为:

$$\begin{aligned} \frac{dx}{dt} = & \frac{\mu_{xm} \cdot x \cdot s}{K_x \cdot x + s} - x \cdot \frac{F}{s_f} + w_1 = f_1 + \\ & + w_1 \end{aligned} \quad (9)$$

$$\begin{aligned} \frac{dp}{dt} = & \frac{\mu_{pm} \cdot x \cdot s}{K_p + s(1 + s/K_i)} - K \cdot p - \\ & - p \cdot \frac{F}{s_f} + w_2 = f_2 + w_2 \end{aligned} \quad (10)$$

$$\frac{ds}{dt} = \frac{\mu_{xm} \cdot x \cdot s}{Y_{x/s}(K_x \cdot x + s)} -$$

$$-\frac{\mu_{xm} \cdot x \cdot s}{Y_{x/s}(K_p + s(1 + s/K_i))} - m \cdot x + F - s \cdot \frac{F}{s_f} + w_3 = f_3 + w_3 \quad (11)$$

$$\frac{dCER}{dt} = \frac{\mu_{xm} \cdot x \cdot s}{Y_{x/CO_2}(K_x \cdot x + s)} + M \cdot x + k_{pe} + w_4 = f_4 + w_4 \quad (12)$$

$$\frac{d\mu_{xm}}{dt} = 0 + w_5 = f_5 + w_5 \quad (13)$$

⋮

$$\frac{dk_{pe}}{dt} = 0 + w_{16} = f_{16} + w_{16} \quad (14)$$

$$\text{令 } \vec{f} = [f_1, f_2, \dots, f_{16}]^T,$$

$$\vec{W} = [w_1, w_2, \dots, w_{16}]^T$$

$$\text{有 } \frac{d\vec{X}}{dt} = \vec{f}(\vec{X}(t)) + \vec{W} \quad (15)$$

其中  $\vec{W}$  为动态噪声, 具有如下性质:

$$E[\vec{W}(k)] = 0, E[\vec{W}(k) \cdot \vec{W}^T(l)] = \vec{Q}(k) \cdot \delta_{kl}$$

$$\vec{Q} \text{ 是非负定矩阵, } \delta_{kl} = \begin{cases} 1 & k=l \\ 0 & k \neq l \end{cases}$$

## (二) 观测方程

前面已提到, 在状态量中仅有尾气  $CO_2$  排放速率是可在在线测量, 因此建立如下观测方程:

$$\vec{Z} = \vec{H} \cdot \vec{X} + \vec{V}co_2, \vec{H} = [0, 0, 0, 1, \dots, 0]_{1 \times 16} \quad (16)$$

$$\text{令 } \vec{X} = [x, p, s, CER, \mu_{xm}, \dots, k_{pe}]^T$$

$$\vec{Z} = [CER], \text{ 是排气 } CO_2 \text{ 在线检测信号, } \vec{V}co_2 \text{ 和 } \vec{W} \text{ 一样为零均值白噪声序列,}$$

$$E[\vec{V}co_2(k)] = 0, E[\vec{V}co_2(k) \cdot \vec{V}co_2(l)] = \vec{R}(k) \delta_{kl} \text{ 式中 } \vec{R} \text{ 是非负定阵, } \vec{V}co_2 \text{ 又$$

称观测噪声。

可以看到,  $\vec{W}$ 、 $\vec{V}co_2$  反映了模型与过程偏差的校正, 其方差的确定对状态估计与参数辨识精度起着十分重要的作用。

## (三) EKF原理与算法

有了推广状态方程与观测方程, 就可以进行状态估计与参数辨识。由于发酵过程具有时变、不确定性, 状态方程又是非线性的, 所以采用EKF技术辨识。为此必须首先将方程(9)~(16)离散化, 即将  $\vec{X}(k)$  在  $\vec{X}(k-1)$  处展开成一阶 Taylor 级数。

$$\vec{X}(k) = \vec{X}(k-1) + \dot{\vec{X}}(k-1) \cdot \Delta k + \dots$$

$$\text{而 } \dot{\vec{X}}(k-1) = \vec{f}(\vec{X}(k-1))$$

$$\therefore \vec{X}(k) = \vec{X}(k-1) + \vec{f}(\vec{X}(k-1)) \cdot \Delta k + \vec{W}$$

$$\text{又若令: } \vec{\varphi} = [\varphi_1, \varphi_2, \dots, \varphi_{16}]^T$$

$$\varphi_i = X_i(k-1) + f_i(\vec{X}(k-1))$$

$$\vec{X}(k) = \vec{\varphi}(k-1) + \vec{W}$$

又因  $\vec{\varphi}(k-1)$  是非线性函数向量, 所以, 将  $\vec{\varphi}(\vec{X}(k-1))$  在  $\vec{X}(k-1)$  处展开成一阶 Taylor 级数而线性化, 得:

$$\vec{X}(k) = \vec{X}(k-1) + \left. \frac{\partial \vec{\varphi}}{\partial \vec{X}} \right|_{\vec{X}(k-1)} [\vec{X}(k-1) - \vec{X}(k-1)] + \vec{W} \quad (17)$$

其中,  $\vec{X}(k-1)$  是  $(k-1)$  处的预估值。

EKF滤波原理是: 首先计算状态量  $\vec{X}$  的预估值, 然后利用观测方程, 通过增益矩阵  $\vec{K}$  根据输出信息进行校正, 得到精确

的 $\vec{X}$ 值。其具体算法如下<sup>[6]</sup>：

$$\hat{\vec{X}}(k|k-1) = \vec{\Phi}(\vec{X}(k-1)) \quad (18)$$

$$\begin{aligned} \hat{\vec{X}}(k) = & \hat{\vec{X}}(k|k-1) + \vec{K}(k) [\vec{Z}(k) - \\ & - \vec{H}\hat{\vec{X}}(k|k-1)] \quad (19) \end{aligned}$$

$$\begin{aligned} \vec{K}(k) = & \vec{P}(k|k-1)\vec{H}^T(k) \\ & \cdot [\vec{H}(k)\vec{P}(k|k-1)\vec{H}^T(k) + \\ & + \vec{R}(k)]^{-1} \quad (20) \end{aligned}$$

$$\begin{aligned} \vec{P}(k|k-1) = & \vec{\Phi}(k-1)\vec{P}(k-1) \\ & \cdot \vec{\Phi}^T(k-1) + \vec{\Gamma}\vec{Q}(k-1) \\ & \cdot \vec{\Gamma}^T \quad (21) \end{aligned}$$

$$\vec{P}(k) = [\vec{I} - \vec{K}(k)\vec{H}(k)]\vec{P}(k|k-1) \quad (22)$$

初始条件： $\vec{X}(0) = E\{\vec{X}(0)\}$   
 $\vec{P}_x(0|0) = \vec{P}_x(0) = V_{x, x}\{\vec{X}(0)\}$

其中： $\vec{X} = [x, p, s, CER, \mu_{xm}, \dots, k_p]^T$

$$\vec{Z} = \vec{H}\vec{X} + \vec{V} \quad (23)$$

$$\vec{H} = [0, 0, 0, 1, 0, \dots, 0]_{1 \times 16}$$

$$\vec{\Phi}(k-1) = \vec{I} + \left. \frac{\partial f}{\partial X} \right|_{x(k-1)} \cdot \Delta t =$$

$$= \left. \frac{\partial \Phi}{\partial X} \right|_{x(k-1)}$$

$$\text{Cov}\{\vec{W}(k), \vec{W}(j)\} = \vec{Q}\delta_{kj}$$

$$\text{Cov}\{\vec{V}_{CO_2}(k), \vec{V}_{CO_2}(j)\} = \vec{R}(k)\delta_{kj}$$

$(k \geq j)$

$$\vec{\Gamma} = \begin{pmatrix} 1 & & & & & & & & & & & & & & & & & & \\ & 1 & & & & & & & & & & & & & & & & & & \\ & & 1 & & & & & & & & & & & & & & & & & \\ & & & 1 & & & & & & & & & & & & & & & & \\ & & & & 1 & & & & & & & & & & & & & & & \\ & & & & & 0 & & & & & & & & & & & & & & \\ & & & & & & 0 & & & & & & & & & & & & & \\ & & & & & & & 0 & & & & & & & & & & & & \\ & & & & & & & & 0 & & & & & & & & & & & \\ & & & & & & & & & 0 & & & & & & & & & & \end{pmatrix}_{16 \times 16}$$

前面已经提到， $\vec{Q}$ 、 $\vec{R}$ 的选择极为关键，然而，通常难于从理论上精确给出，只能凭经验确定。一旦 $\vec{Q}/\vec{R}$ 比值不合适就会引起发散。

## 结 果 与 讨 论

本文对青霉素发酵过程进行了滤波估计。考虑到产物大量形成期始于发酵的40h之后，所以滤波从40h开始，并取初始条件为：

$x(0) = 20, p(0) = 1.94, s(0) = 1,$   
 $CER(0) = 0.7434, \mu_{xm} = 0.132,$   
 $K_x = 0.15, \mu_{pm} = 0.045,$   
 $K_p = 0.0002, K_i = 0.1, K = 0.01,$   
 $Y_{x/f} = 0.45, Y_{p/f} = 0.9,$   
 $m = 0.005, M = 0.0014,$   
 $Y_{CO_2/x} = 30.8, k_{ps} = 0.0014$

$$\vec{Q}_{11} = [1.5, 0.09, 0.015, 0.01306, 0, 0, \dots, 0]^T$$

$$\vec{P}_{11} = [1.4, 0.13, 0.1, 0.0139, 0.009, 0.01, 0.0005, 0.00002, 0.01, 0.009, 0.05, 0.09, 0.002, 0.00005, 3.5, 0.00021]^T$$

$$\vec{R} = 0.05$$

$\vec{Q}_{11}$ 、 $\vec{P}_{11}$ 数值代表矩阵主对角线上元素，其取值是凭经验确定的。

图1是EKF计算得到的状态量 $x$ 、 $p$ 、 $s$ 值，可见，残糖浓度 $s$ 在产物大量形成期始终维持在某一低水平上，这符合工艺分

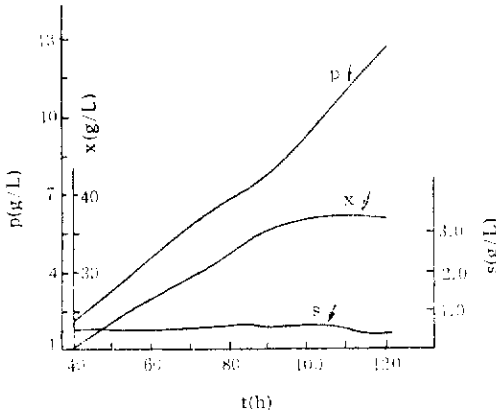


图 1 x、p、s 的滤波估计值变化曲线

Fig.1 Filter calculation of x、p、s against fermentation time

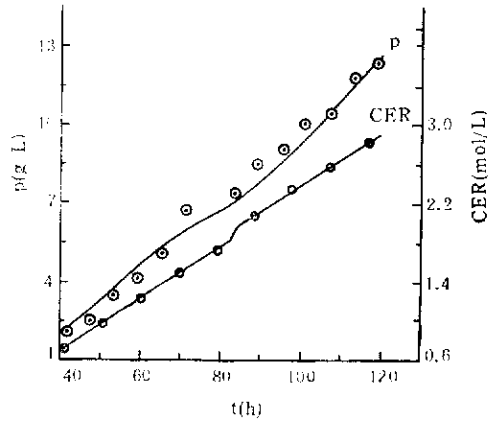


图 2 滤波值与实测值比较曲线

Fig.2 Comparison curve of estimation value with experimental data

○ 实测值 Experimental data — 滤波值 model calculation

析结果。而发酵进行到一定时期，菌量已不再明显上升，产物则大量产生，表明发酵朝着合成产物的方向进行。

图 2 给出了产物的手工分析值、CO<sub>2</sub> 在线检测值与其估计值的比较曲线。可以看到，CER 的估计精度较高，一般计算的相对误差在 0—0.6% 之间，这主要是正确选择了信噪比 ( $\overline{Q/R}$ )，使 EKF 具有良好的滤波品质。产物的估计精度在 3.00% 左右 (平均相对误差)。表明 EKF 在目前还缺乏可以在线检测的传感器的情况下，能用于工业生产实际中。

图 3、4、5、6 显示了发酵过程中

的动力学参数变化情况。在图 3 中，菌体的比生长速率在发酵第 40—50h 内下降较快，从发酵第 50—85h  $\mu_x$  变化较为平稳，这是因为菌体正处于其快速生长期的后期，而开始其过渡期，随后即处在一个较为稳定的生长环境之中，所以  $\mu_x$  变化不大。但是当发酵至一定阶段，可能是由于某种干扰，正如本文前面曾提到过的那样：生化发酵过程中存在着不可预知和不可控的随机干扰，因而造成了如图 3、4、5、6 中的在发酵 85h 左右所反映出来的比较明显的转折变化。随着发酵过程的进行，到了发酵后期 (110—120h)，菌体

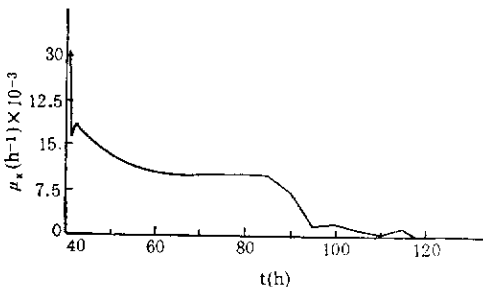


图 3 瞬时比生长速率变化曲线

Fig.3 Comparison curve of specific growth rate against time

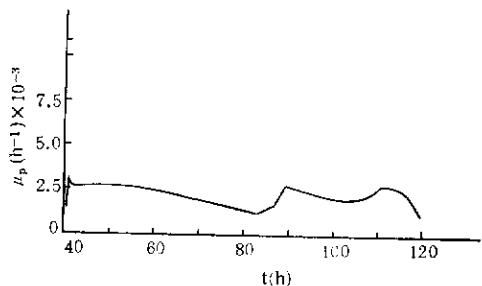


图 4 瞬时比产物形成速率变化曲线

Fig.4 Comparison curve of specific rate of product formation against time

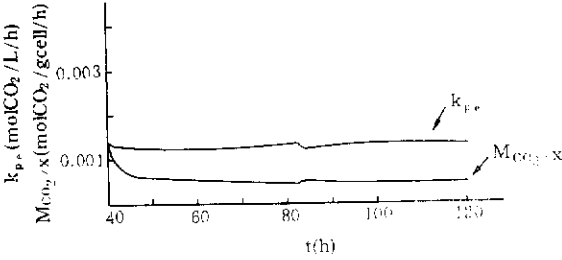


图 5  $k_{p e}$ 、 $M$ 随过程波动曲线

Fig.5 The curves of  $k_{p e}$ 、 $M$  during fermentation process

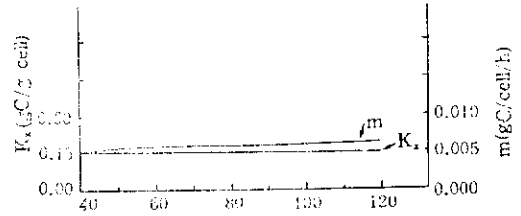


图 6、 $m$ 、 $k_x$ 随过程变化曲线

Fig.6 The curves of  $m$ 、 $k_x$  during fermentation process

的比生长速率 $\mu_x$ 逐渐下降至零，相应地，从图4可以发现，比产物形成速率也很快下跌，致使生产终止。所以，在青霉素半连续发酵中，控制好菌体生长是十分重要的，因为菌体生长的好坏，直接影响到产物青霉素的生产。

从图5可见， $k_{p e}$ 随过程的进行而缓慢上升，这是因为产物大量形成时CER相应增加的缘故。而 $M$ 稍有下降，表明菌体细胞增殖到一定程度，活力下降，形成的

$CO_2$ 量稍有减小，这与文献[4]的结论一致。发酵后期因菌体内源呼吸加强， $m$ 缓慢上升，而其参数，例如图6中的 $k_x$ ，变化波动就不明显，表明外界扰动对其影响可以忽略。

值得指出的是，EKF技术克服了目前生化反应过程传感器缺乏这一难点，利用EKF方法，就可以通过有限的状态在线检测信息推算过程中其他变量的变化，这就为优化控制的实施提供了前提条件。

### 符 号 说 明

(一) 青霉素发酵模型

$CO_2$	二氧化碳	
$F$	补料速率	(g/L/h)
$K$	青霉素一级衰减常数	(h <sup>-1</sup> )
$K_x$	基质限制时菌体形成的Contois饱和系数	
$K_p$	基质限制时产物形成的Monod饱和系数	(g/L)
$K_i$	产物形成时基质抑制常数	(g/L)
$k_{p e}$	产物形成所产生的 $CO_2$ 系数	(mol/L/h)
$M$	菌体维持所产生的 $CO_2$ 系数	(mol/g/h)
$m$	维持所需基质系数	(gc/gcell/h)
$p$	青霉素浓度	(g/L)
$s$	基质浓度	(g/L)
$s_f$	进料中基质浓度	(g/L)
$\mu_{x m}$	最大比生长速率	(h <sup>-1</sup> )
$\mu_{p m}$	最大比产物形成速率	(h <sup>-1</sup> )
$\mu_x$	比生长速率	(h <sup>-1</sup> )
$\mu_p$	比产物形成速率	(h <sup>-1</sup> )

$x$	菌体浓度	(g/L)
$Y_{x/s}$	单位基质上的菌体得率系数	(gcell/gC)
$Y_{p/s}$	单位基质上的产物得率系数	(gpeni/gC)
CER	$CO_2$ 释放速率	(mol $CO_2$ /L)

(二) EKF

$X$	状态向量
$Z$	观察向量
$\varphi$	系统状态变量函数
$H$	观察函数在X处的雅可比阵
$K$	增益阵
$P$	估计误差协方差阵
$Q$ 、 $R$	具有零均值高斯白噪声序列的协方差阵
$I$	单位阵
$\hat{\cdot}$	估计值
$(k k-1)$	$(k-1)$ 时刻对 $k$ 时刻的估计值
$\cdot^T$	转置
$\Gamma$	函数阵
$\Phi$	系统状态变量函数的导数

## 参 考 文 献

- [1] Yu Juntang, Zhang Siliang, Horizons of Biochemical Engineering, Shuichi Aiba, Tokyo University press. 1987.
- [2] Bajpai, R.K. and Reub, M.: *J. Chem. Tech. Biotechnol.*, 33:332, 1980.
- [3] Calam, C.T. and Ismail, B.A.K.: *J. Chem. Tech. Biotechnol.*, 30:249, 1980.
- [4] 张嗣良等: 华东化工学院内部资料, 1987.
- [5] 刘 豹编: 现代控制理论, 机械工业出版社, 1983.

## ONLINE IDENTIFICATION OF STATE AND PARAMETERS FOR FED-BATCH PENICILLIN FERMENTATION PROCESS

Jin Sha Zhang Siliang Yu Juntang

(*Research Institute of Biochemical Engineering, East China University of Chemical Technology, Shanghai*)

In this paper, a mathematical model with noises which is suitable for penicillin fed-batch fermentation in industrial scale has been presented. By taking the carbon dioxide production rate as an online measured variable, an extended Kalman filter (EKF) is used to realtime-identify the state and parameters. And this work lays the foundation of the adaptive optimal control for fermentation processes. It is shown that the state estimation values by EKF is rather coincide with experimental data. The results prove that the filter is a better online observer of fermentation process state and dynamical parameters.

### Key words

Penicillin fermentation; state estimation; parameters identification; extended Kalman filter