

单纯形法在微生物生长动力学和能学 参数估算中的应用

方柏山

(华侨大学化工与生化工程系, 泉州)

本文基于作者所导出的用于描述微生物分批和连续培养的数学模型^[6,7], 就单纯形法在微生物生长动力学和能学参数估算中的应用进行研究。根据模型的特点, 提出确定目标函数及单纯形初始值的方法, 并以大肠杆菌单罐连续培养为例介绍了估算的全过程。

关键词 微生物生长; 数学模型; 单纯形法; 动力学; 能学; 参数估算

无论是进行生化工程学还是生物工程学研究, 在确定了具有优良性状的菌种及其生理研究之后, 为了描述微生物的生长过程、提供放大及优化控制模型, 都必须对微生物生长动力学及其数学模型中有关参数的估算进行研究。

自从1942年 Monod 首次建立了单细胞生长动力学方程以来, 国外关于微生物生长动力学方面的研究报告颇多^[1-5]。借助于前人的工作, 作者先后导出了用于描述微生物分批及连续培养生长动力学模型的解析解, 并且用改正单纯形法, 估算了解析解中的微生物生长动力学和能学参数^[6,7]。为了进一步完善估算方法, 本文特根据所建立的数学模型就估算中碰到的诸如目标函数及单纯形初始值的选择等问题进行研究。并且通过实例具体介绍估算的全过程。

数学模型

用于由分批培养数据 X 、 S 及 t 估算微生物生长动力学及能学参数 μ_{max} 、 K_s 、 $Y_{x/s}^{m,x}$ 和 m 的数学模型^[8] (注: 符号及表达式有所更改);

$$X = A(S_0 - S - B \ln \frac{S_0 + B}{S + B}) + X_0 \quad (1)$$

$$t = C_3 \ln(C_1 + C_2 S - S^2) + C_4$$

$$\ln(ES + F) - C_5 \ln \frac{C_6 - S}{C_7 + S} + C_8 \quad (2)$$

式中:

$$A = Y_{x/s}^{m,x} \cdot \mu_{max} / (m_s \cdot Y_{x/s}^{m,x} + \mu_{max})$$

$$B = K_s m_s A / \mu_{max}$$

$$E = m_s + \mu_{max} / Y_{x/s}^{m,x}$$

$$F = m_s K_s$$

$$C_1 = (S_0 - 2B + X_0 / A) S_0$$

$$C_2 = 2B + X_0 / A$$

$$C_3 = (1 - C_4 \mu_{max}) / 2\mu_{max}$$

$$C_4 = (EK_s - F)(ES_0 - F) / (F^2 + C_2 EF - C_1 E^2) \cdot \mu_{max}$$

$$C_5 = (S_0 K_s + C_1 C_4 \mu_{max} + C_2 C_3 AF) / AF \sqrt{C_2^2 + 4C_1}$$

$$C_6 = (C_2 - \sqrt{C_2^2 + 4C_1}) / 2$$

$$C_7 = C_6 - C_2$$

$$C_8 = -C_3 \ln(C_1 + C_2 S_0 - S_0^2) -$$

本文于1990年4月13日收到。

本文曾于1989年7月在北京国际生物工程会议上交流。

$$C_4 \ln(ES_0 + F) + C_5 \ln \frac{C_6 - S_0}{C_7 + S_0};$$

μ_{max} 、 K_s 、 $Y_{x/s}^{max}$ 及 m_s 分别表示最大比生长速率、Monod 饱和常数、微生物关于限制性底物的最大宏观得率及维持系数； X 和 S 分别表示时间为 t 时的菌体浓度及限制性底物浓度； X_0 及 S_0 分别表示初始的 X 及 S 。

用于由连续培养数据 X 、 S 及 D 估算 μ_{max} 、 K_s 、 $Y_{x/s}^{max}$ 及 m_s 的数学模型^[7]：

$$X = A[S_0 - \frac{DK_s}{\mu_{max} - D} - B \cdot \ln \frac{S_0(\mu_{max} - D)}{DK_s + B(\mu_{max} - D)}] \quad (3)$$

$$S = \frac{DK_s}{\mu_{max} - D} \quad (4)$$

式中， X 和 S 分别表示稀释度为 D 时的流出发液中菌体及限制性底物浓度； S_0 为流入液中限制性底物的浓度； A 、 B 的含义同前述。

目标函数的选择

由式(1)一(4)可见，若以平方值为

优化准则，估算参数 μ_{max} 、 K_s 、 $Y_{x/s}^{max}$ 及 m_s 的问题实际上是如下所示的式(1')一(4')的极小化问题：

$$f_1 = \sum [A(S_0 - S - B \ln \frac{S_0 + B}{S + B} + X_0 - X)]^2 \quad (1')$$

$$f_2 = \sum [C_3 \ln(C_1 + C_2 S - S^2) + C_4 \ln(ES + F) - C_5 \cdot \ln \frac{C_6 - S}{C_7 + S} + C_5 - t]^2 \quad (2')$$

$$f_3 = \sum \{A[S_0 - \frac{DK_s}{\mu_{max} - D} - B \cdot \ln \frac{S_0(\mu_{max} - D)}{DK_s + B(\mu_{max} - D)}] - X\}^2 \quad (3')$$

$$f_4 = \sum (\frac{DK_s}{\mu_{max} - D} - S)^2 \quad (4')$$

如何较好地估算这四个参数，关键之一在于目标函数的选择。就本课题而言，目标函数的选择主要有如表1所示的三种方式：

方式1固然能够直接估算出 μ_{max} 、 K_s 、 $Y_{x/s}^{max}$ 及 m_s ，但对于整个分批(或连续)培养体系而言，所要解决的是双目标

表 1 选择目标函数的方式
Table 1 The pattern of selecting object function

方 式 Patterns	目标函数 Object function	估 算 参 数 Parameters estimated	数据来源 Source of data
1	f_1 or f_2	μ_{max} 、 K_s 、 $Y_{x/s}^{max}$ and m_s are estimated directly	bc
	f_3	Ditto	cc
2	First f_1 , then f_2	A 、 B then K_s and $Y_{x/s}^{max}$, finally μ_{max} and m_s are calculated	bc
	First f_4 then f_3	K_s and μ_{max} are calculated first, then $Y_{x/s}^{max}$ and m_s are estimated	cc
3	$\lambda_1 f_1 + \lambda_2 f_2$	μ_{max} 、 K_s 、 $Y_{x/s}^{max}$ and m_s are estimated directly	bc
	$\lambda_3 f_3 + \lambda_4 f_4$	Ditto	cc

Note: λ_1 、 λ_2 、 λ_3 and λ_4 stand for weighting factor(权因子)
bc, 分批培养 Batch cultures, cc, 连续培养 Continuous cultures

函数的优化问题。故由方式 1 所求得参数可能只适于对某一变量变化的描述, 而不适于同一体系中的另一变量。如由 $f_2 \rightarrow 0$ 所得到的四个参数对 S 随 t 变化的描述可能是适合的, 但对于 X 随 S (或 t) 变化的描述可能就不合适了。

方式 2 虽然同时考虑了双目标函数的优化问题, 而且对每个目标函数来说所要考虑的参数数目比方式 1 减少了一半, 即均为双参数函数。但 f_1 (或 f_4) 的优化固定了两个参数, 无形中限制了 f_2 (或 f_3) 的优化。作者在文献 [6, 7], 中所采用的目标函数就是这种方式。

方式 3 集方式 1 和 2 之优点。它既变双目标函数为单目标函数, 又同时考虑了四个参数, 使得寻求整个系统的最优解成为可能。据此本文采用这种方式进行估算。

SNM 单纯形法

目标函数 $\lambda_1 f_1 + \lambda_2 f_2$ 及 $\lambda_3 f_3 + \lambda_4 f_4$ 都是较为复杂的非线性多元函数。对于这类函数寻优问题目前虽然有许多方法, 但还没有一个被公认为最好的。

SNM 单纯形法是求解非线性多元函数、无约束极小化问题的有效方法之一, 该方法是 J. A. Nelder 和 R. Mead (1965 年) 在 Spendley 等 (1962 年) 所报道的正规单纯形法的基础上提出的一种改正的优化方法 [8]。由于该方法操作简单, 便于在计算机上实现, 故应用之较为普遍。

SNM 单纯形法的基本设想是: 1) 对于 n 元函数 $f(X)$ (X 为 n 维向量), 取 n 维空间中的 $n+1$ 个点, X^1, X^2, \dots, X^{n+1} 构成一个 $n+1$ 面的超棱面体, 求出各顶点所对应的函数值; 2) 求出除函数值最差点 X^k 以外的各顶点的中心 X^0 及其所对应的

函数 $f(X^0)$; 3) 以 X^0 为轴心求出 X^k 的反射点 X' 及 $f(X')$; 4) 通过“压缩”、“延伸”、或“收缩”原超棱面体寻找函数值较小的新点, 以取代 X^k 构成新的单纯形; 5) 重复前述的步骤, 直至满足给定的收敛要求为止。

本文, $n=4$, 向量 $X = (\mu_{0.1x}, K_1, Y_{x/f_2}^{m_2}, m_2)^T$, 反射、压缩及延伸因子分别取 1、0.5 及 2。收敛准则为: $\left\{ \frac{1}{n+1} \sum_{i=1}^{n+1} [f(X^i) - f(X^0)]^2 \right\}^{1/2} < \varepsilon$ (ε 为精度)。

单纯形初始点的确定

由于单纯形法是通过迭代过程中直接比较函数值的大小, 再根据一定的收敛判别条件来获取优化解的。它并没有利用目标函数的性态及其解析性质, 故所得的优化解一般只能认为是相对优化解 (或局部优化解)。为了使局部优化解尽可能地改善, 即尽可能避免择优过程陷入相近局部极值点的附近区域, 必须选择好单纯形的初始点。这一点对于参数较多的场合尤为重要。为了避免主观、片面和盲目地选择初始点。本文引用了实验优化设计中的正交设计法 [9]。从这种设计所合理安排的具有代表性的“实验”点中可望较为科学地确定单纯形的初始点。

应用实例

表 3 中的 D 、 X 及 S 列为大肠杆菌以葡萄糖为单一碳源进行单罐连续需氧培养的数据 [2]。为了确定用单纯形法进行参数估算的初始点, 选用正交表 $L_{16}(4^5)$ 进行“实验”。结果如表 2 所示。

由此表中 $K_1 - K_4$ 行的数据及极差 R 值可见: $\mu_{0.1x}$ 、 K_1 、 $Y_{x/f_2}^{m_2}$ 及 m_2 分别取

1.2、0.05、0.8、0.2或1.0, 0.05, 0.8及0.2有可能获得更小的 $f(X)$ 值, 但计算表明, 其 $f(X)$ 值分别为0.626及0.443均大于表2中所得到的最小 $f(X)$ (=0.11, 见序号3)。表2中的 R 值还表明: 各参数变化对 $f(X)$ 值的影响依 μ_{max} 、 K_s 、 $Y_{x/s}^{max}$ 及 m_s 的次序减弱。

据此, 单纯形初始点取: $\mu_{max} = 1.2$, $K_s = 0.1$, $Y_{x/s}^{max} = 0.4$ 及 $m_s = 0.2$ 。用BASIC语言编写程序, 在Apple-II型微电脑上运算, 取 $\epsilon = 10^{-6}$, 经118次迭代得: $\mu_{max} = 1.24$, $K_s = 0.138$, $Y_{x/s}^{max} = 0.543$ 及 $m_s = 0.211$ 。 $f(X)$ 值由原来的0.1067降至 8.0×10^{-3} , 降低了92.5%(历时约5min)。以所得到的参数值作为单纯形的新起点经过350多次迭代, $f(X)$ 值降至 6.29×10^{-3} , 即再次降低了22.3%。此时 $\mu_{max} = 1.11$ (h^{-1}), $K_s = 0.107$ ($g \cdot glucose/L$), $Y_{x/s}^{max} = 0.733$ ($g \cdot cell/g \cdot glucose$), $m_s = 0.62$ ($g \cdot glucose/g \cdot cell \cdot h$)。结果见表3及图1。

表2 确定 μ_{max} 、 K_s 、 $Y_{x/s}^{max}$ 及 m_s 估算初值的正交设计

Table 2 The orthogonal design of determining initial values of estimating μ_{max} , K_s , $Y_{x/s}^{max}$ and m_s

序号 No.	μ_{max}	K_s	$Y_{x/s}^{max}$	m_s	$f(X)$
1	1.2	0.4	0.8	0.6	0.90
2	1.2	0.25	0.6	0.4	0.21
3	1.2	0.1	0.4	0.2	0.11
4	1.2	0.05	0.2	0.05	0.68
5	1.0	0.4	0.6	0.2	3.36
6	1.0	0.25	0.8	0.05	1.08
7	1.0	0.1	0.2	0.6	0.82
8	1.0	0.05	0.4	0.4	0.17
9	0.8	0.4	0.4	0.05	40.9
10	0.8	0.25	0.2	0.2	13.4
11	0.8	0.1	0.8	0.4	1.71
12	0.8	0.05	0.6	0.6	0.29
13	0.75	0.4	0.2	0.4	298
14	0.75	0.25	0.4	0.6	120
15	0.75	0.1	0.6	0.05	20.4
16	0.75	0.05	0.8	0.2	4.42
K_1	1.89	343	8.11	122	
K_2	5.43	135	24.3	300	
K_3	56.3	23.0	161	21.3	
K_4	443	5.56	313	63.1	
R	110	84	76	70	

表3 大肠杆菌单罐连续培养的实验和计算值及数据分析

Table 3 Values of experiment and calculation of simple tank continuous cultures of *E.coli* and there data analysis ($S_0 = 0.968 g \cdot glucose/L$)

D (h^{-1})	X ($g \cdot dry\ cells/L$)	S ($g \cdot glucose/L$)	X_s			S_s		
			a	b	c	a'	b'	c'
0.73	0.352	0.210	0.222	0.367	0.372	0.508	0.227	0.205
0.71	0.391	0.221	0.266	0.376	0.380	0.418	0.208	0.190
0.69	0.420	0.170	0.297	0.384	0.387	0.353	0.191	0.175
0.66	0.422	0.153	0.330	0.395	0.395	0.282	0.169	0.157
0.60	0.434	0.122	0.372	0.411	0.408	0.193	0.133	0.125
0.53	0.427	0.102	0.399	0.426	0.420	0.132	0.102	0.097
0.43	0.422	0.064	0.419	0.440	0.431	0.082	0.069	0.067
0.31	0.438	0.040	0.426	0.452	0.440	0.047	0.042	0.041
0.24	0.417	0.033	0.423	0.457	0.443	0.033	0.030	0.027
0.12	0.434	0.013	0.394	0.462	0.446	0.015	0.013	0.012
0.06	0.427	0.006	0.339	0.463	0.447	0.008	0.006	0.006
离差平方和 Q Sum of squares of deviations			0.0728	0.00803	0.00517	0.184	0.00131	0.00109
参考文献 References			[2]	[7]		[2]	[7]	

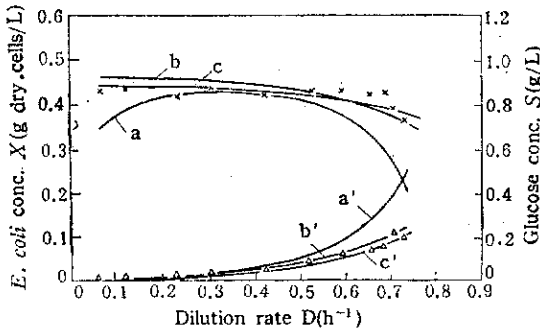


图 1 X及S的计算值与实测值

Fig.1 The calculated and experimental values of X and S

- a, a': $\mu_{max} = 0.84(\text{h}^{-1})$, $K_s = 0.074(\text{g}\cdot\text{glucose/L})$, $Y_{x/s}^{max} = 0.5(\text{g cell/g}\cdot\text{glucose})$, $m_s = 0.05(\text{g}\cdot\text{glucose/g}\cdot\text{cell}\cdot\text{h})$ [2];
- b, b': $\mu_{max} = 1.05(\text{h}^{-1})$, $K_s = 0.0997(\text{g}\cdot\text{glucose/L})$, $Y_{x/s}^{max} = 0.573(\text{g}\cdot\text{cell/g}\cdot\text{glucose})$, $m_s = 0.241(\text{g}\cdot\text{glucose/g}\cdot\text{cell}\cdot\text{h})$ [7];
- c, c': $\mu_{max} = 1.11(\text{h}^{-1})$, $K_s = 0.107(\text{g}\cdot\text{glucose/L})$, $Y_{x/s}^{max} = 0.733(\text{g}\cdot\text{cell/g}\cdot\text{glucose})$, $m_s = 0.62(\text{g}\cdot\text{glucose/g}\cdot\text{cell}\cdot\text{h})$

若文献〔6〕中的目标函数由表1中的方式2换成方式3, 取该文中的估算结果作为单纯形的起始点, 即 $\mu_{max} = 0.719(\text{h}^{-1})$; $K_s = 0.190(\text{kg}/\text{m}^3)$, $Y_{x/s}^{max} = 0.549(\text{kg}/\text{kg})$ 和 $m_s = 0.0666(\text{h}^{-1})$, 经270多次迭代得 $\mu_{max} = 0.709$, $K_s = 0.108$, $Y_{x/s}^{max} = 0.54$, 和 $m_s = 0.0523$ 。总离差平方和(即 $f(X)$)由 6.61×10^{-3} 降至 5.1×10^{-3} 。

讨 论

就作者所导出的分别用于描述微生物分批及连续培养的数学模型〔参见式(1)一(4)〕而言, SNM单纯形法是一种估算模型中微生物生长动力学和能学参数行之有效的办法。

用作者所导出的数学模型进行参数估算, 对于分批和连续培养体系来说都是一个双目标函数的极小化问题。为了得到整个体系的优化解, 本文提出了通过加权求和使多目标函数转化为单目标函数的方法(见表1中方式3), 权因子的取值建议视同一体系中各目标函数的重要程度及其值的数量级而定。实例表明(见表3), 用这种方法能够更好地求得优化解。

为了尽量避免SNM单纯形法在择优过程中可能陷入某些局部极值点, 本文提出了用正交设计法于单纯形初始点选择的方法。实践表明这种方法有利于避免选择单纯形初始点的盲目性, 从而能够在较短的时间内较为科学地确定单纯形的初始点。而且通过正交设计的数据分析还能够就各参数对目标函数影响的大小作出判断, 这对于模型的简化及参数值的评价具有指导意义。

符 号 说 明

- A 复合参数(kg/kg)
 B 复合参数(kg/m³或g/L)
 C₁—C₈ 常数
 D 稀释度(h⁻¹)
 E 复合参数(h⁻¹)
 F 复合参数(kg/m³·h或g/L·h)
 f(X) 目标函数
 K_s Monod饱和常数(kg/m³或g/L)
 m_s 维持常数(h⁻¹)
 n 向量维数
 Q 离差平方和
 R 极差
 S 底物浓度(kg/m³或g/L)
 t 时间(h)
 X 菌体浓度(kg/m³或g/L)
 Y_{x/s} 微生物关于底物的宏观得率(kg/kg)
 μ_{max} 最大比生长速率(h⁻¹)
 上标符号: *h* 单纯形的最差点, *l* 单纯形的最佳点,
r 单纯形的反射点, *o* 单纯形的中心点
 下标符号: *o* 初始(值), *c* 计算(值)

参 考 文 献

- [1] Roels, J. A.: *Energetics and Kinetics in Biotechnology*, Elsevier Science Publishers, 1983.
- [2] 合叶修一, 永井史郎(胡章助等译): *生物化学工程——反应动力学*, 化学工业出版社, 1984.
- [3] Atkinson, B. and Mavituna, F.: *Biochemical Engineering and Biotechnology Handbook*, Macmillan Publisher Ltd., 1983.
- [4] Heijnen, J. J. and Roels, J. A.: *Biotechnol. Bioeng.*, 23:739, 1981.
- [5] Takamatsu, T. et al.: *J. Ferment. Technol.*, 59(2):131, 1981.
- [6] 方柏山: *生物工程学报*, 2(4):54—60, 1986.
- [7] 方柏山: *华侨大学学报*, 8(4):458, 1987.
- [8] 邓乃扬: *无约束最优化计算方法*, 科学出版社, 1982.
- [9] 白新桂: *数据分析及试验优化设计*, 清华大学出版社, 1986.

The Application of Simplex Method to the Estimation of Kinetic and Energetic Parameters of Microbial Growth

Fang Baishan

(Huaqiao University, Quanzhou)

Based on the mathematical models derived by the author for describing both batch microbial cultivation and continuous one, the application of simplex method to the estimation of K_s , μ_{max} , $Y_{x/i}^m$ and m_i is investigated. Methods of determination of objective function and initial value of simplex method as well as the introduction of the whole process of estimation with single tank continuous culture of *E. coli* as an example is presented.

Key words

Microbial growth; mathematical model; simplex method; kinetics; energetics; parameter estimation