

0.618 法在两种常用 生长曲线拟合中的应用

袁志发 刘光祖 肖俊璋

(西北农学院基础部) (西北农学院农化系)

Mitsherlich 生长曲线

$$y = A + ae^{-kt} \quad (a < 0) \quad (1)$$

和 Logistic 生长曲线

$$y = \frac{A}{1 + ae^{-kt}} \quad (a > 0) \quad (2)$$

是应用比较广泛的两种限制性生长曲线, 其中 A 为生长上限。

在生长曲线 (1) 或 (2) 中, 若 A 为已知常数, 经过变换, 可以使曲线成为未知参数 a 和 k 的线性函数:

$$\ln(A - y) = \ln(-a) - kt \quad (3)$$

和

$$\ln\left(\frac{A - y}{y}\right) = \ln a - kt \quad (4)$$

这时, 可以用最小二乘法通过解正则方程组的方法将 a 和 k 的值求出来。

当 A 未知时, 生长曲线 (1) 或 (2) 为待定参数 A, a, k 的非线性函数, 可表示为:

$$y = f(t, A, a, k) \quad (5)$$

拟合 N 个观察点 ($N > 3$)

$$(t_1, y_1), (t_2, y_2), \dots, (t_N, y_N)$$

则最小二乘解 $\hat{A}, \hat{a}, \hat{k}$ 应满足准则

$$Q = \sum_{i=1}^N (y_i - \hat{y}_i)^2 = \min \quad (6)$$

一般情况下, 无法求出 A, a, k 的解析解。往往利用台劳展开使生长曲线线性化, 然后用逐次迭代法求解。由于上述曲线与线性关系差别很大, 只有在最小二乘解 $\hat{A}, \hat{a}, \hat{k}$ 的附近, 迭代过程才是收敛的。因此, 用迭代法解最小二乘问题应使参数初值充分地接近参数的最小二乘解, 否则迭代过程就不能收敛到最小二乘解 $\hat{A}, \hat{a}, \hat{k}$ 。另外, 迭代过程若用计算器完成, 不但麻烦而且往往办不到。为此, 人们常用三点法, 费歇 (Fisher) 法, 尤拉 (Jull) 法, 罗茨 (Rhodes) 法, 奈尔 (Nair) 法等方法来估计待定参数。在这

些方法中, 由于三点法最简单, 所以是人们经常采用的方法。这些方法所估计的参数对于同一组观察点, 由于选点不同或用的方法不同而又不同, 所选点又非理论值, 又往往不能满足准则 (6)。这样, 对一组观察点来说, 上述方法所估计的参数不能客观的反映生长过程的特征; 对于需要比较的两组观察点来讲, 就失去了可比的条件。因此, 我们提出用0.618法来搜索满足准则 (6) 的最小二乘解, 其步骤如下:

首先确定A的搜索区间 $[a_0, b_0]$, 这可由A的生物学意义来确定, 如由观察点中y的最大值来确定A的搜索区间。

求出区间 $[a_0, b_0]$ 中的两点

$$A_1 = a_0 + 0.382 (b_0 - a_0) \quad A_2 = a_0 + 0.618 (b_0 - a_0)$$

则

$$a_0 < A_1 < A_2 < b_0$$

以 A_1 和 A_2 作为A的估计值, 利用线性化公式 (3) 和 (4), 待定a和 k_i 对于 (3) 来讲, 有

$$\begin{cases} K_i = - \sum_{j=1}^{n_i} (t_j - \bar{t}_i) [\ln (A_i - y_j) - \ln (A_i - \bar{y}_i)] / \sum_{j=1}^{n_i} (t_j - \bar{t}_i)^2 \\ a_i = - \exp [\ln (A_i - \bar{y}_i) + k_i \bar{t}_i] \end{cases} \quad (7)$$

对于 (4) 来讲, 有

$$\begin{cases} K_i = - \sum_{j=1}^{n_i} (t_j - \bar{t}_i) [\ln (\frac{A_i - y_j}{y_j}) - \ln (\frac{A_i - \bar{y}_i}{\bar{y}_i})] / \sum_{j=1}^{n_i} (t_j - \bar{t}_i)^2 \\ a_i = \exp [\ln (\frac{A_i - \bar{y}_i}{\bar{y}_i}) + k_i \bar{t}_i] \end{cases} \quad (8)$$

其中 a_i, k_i 是对应于 A_i ($i=1, 2$) 的最小二乘解, 大于 A_i 的y值不参加拟合运算。

计算 A_1 和 A_2 两点上的Q值, 此时用所有N个观察点来计算, 由于拟合时只用了前 n_i 个点, 故 $Q_i = Q(A_i, n_i, N)$ ($n_i \leq N$)。比较它们的大小:

若 $Q_1 < Q_2$, 则A的搜索区间缩小为 $[a_0, A_2]$, 如图1 (a)。

若 $Q_1 > Q_2$, 则A的搜索区间缩小为 $[A_1, b_0]$, 如图1 (b)。

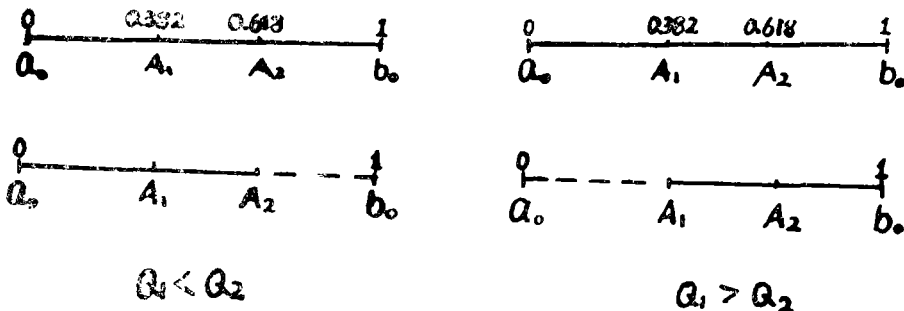


图 1

如果A在 $[a_0, A_2]$ 内, 则第三点取为

$$A_3 = a_0 + A_2 - A_1 \quad (9)$$

其最小二乘解为 $a_3, k_3, Q_3 = Q(A_3, n_3, N)$; 比较 Q_1 与 Q_3 的大小:

若 $Q_3 < Q_1$, 则A的搜索区间缩小为 $[a_0, A_1]$, 如图2(a)。

若 $Q_3 > Q_1$, 则A的搜索区间缩小为 $[A_3, A_2]$, 如图2(b)。

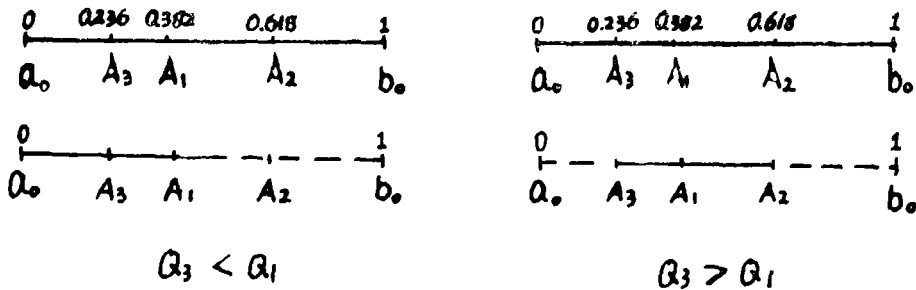


图2

每次缩小了的A的搜索区间都包含有三个点, 用区间的两头点相加减去中间保留点的办法决定新的搜索区间, 如此反复地搜索下去, 直到搜索区间满足预定精度为止。显然, m 次搜索缩小区间之后, 搜索区间长度为:

$$\Delta = (0.618)^m (b_0 - a_0) \quad (10)$$

根据(10)式, 由预定的精度可确定搜索次数。

下面用实例来说明这一方法。

小黄金一号大豆叶面积系数增长动态资料(丁希泉, 1978年引用)如下表:

(t 为从6月1日开始的经历天数)

t	10	20	30	40	50	60	70	80
y	0.1	0.4	1.0	2.4	4.6	5.3	5.5	5.6

试拟合生长曲线(2), 使 $\Delta < 1\%$ 。

由所给资料及A的生物学意义, A的搜索区间我们确定为 $[5, 6]$,

$$\Delta = (0.618)^m (b_0 - a_0) \Big|_{m=10} = 0.008 < 1\%$$

即需要搜索10次, 也就是说要计算到 A_{11} 为止。

$$A_1 = 5 + 0.382(6 - 5) = 5.382 \quad A_2 = 5 + 0.618(6 - 5) = 5.618$$

则 $n_1 = 6$, $N = 8$, $a_1 = 349.29215$, $K_1 = 0.15663$, $Q_1(A_1, 6, 8) = 0.87780$;

$n_2 = 8$, $N = 8$, $a_2 = 239.95959$, $K_2 = 0.13707$, $Q_2(A_2, 8, 8) = 0.20185$ 。

由于 $Q_1 > Q_2$, 故A的搜索区间缩小为 $[5.382, 6]$, A_2 在其内。

$$A_3 = 5.382 + 6 - 5.618 = 5.764$$

则 $n_3 = N = 8$, $a_3 = 128.81756$, $K_3 = 0.11301$, $Q_3(A_3, 8, 8) = 0.49387$ 。由于 $Q_3 >$

Q_2 ，故A的搜索区间缩小为 $[5.382, 5.764]$ ， A_2 在其内。

$$A_4 = 5.382 + 5.764 - 5.618 = 5.528$$

则 $n_4 = 7$ ， $N = 8$ ， $a_4 = 331.33360$ ， $K_4 = 0.15117$ ， $Q_4(A_4, 7, 8) = 0.56164$ 。由于 $Q_4 > Q_2$ ，故A的搜索区间缩小为 $[5.528, 5.764]$ ， A_2 在其内。仿此作到 A_{11} ，其结果列表如下。

搜索过程	搜索点	观察 点数 N	拟合 点数 n_i	a_i	k_i	Q_i	Q大小比较	缩小的搜 索 区 间	精度 Δ
1	$A_1 = 5.382$ $A_2 = 5.618$	8 8	6 8	349.29219 239.95959	0.15663 0.13707	0.87780 0.20185	$Q_1 > Q_2$	5.382—6 A_2 在其内	0.618
2	$A_3 = 5.764$	8	8	128.81756	0.11301	0.49387	$Q_3 > Q_2$	5.382—5.764 A_2 在其内	0.382
3	$A_4 = 5.528$	8	7	331.33360	0.15117	0.56164	$Q_4 > Q_2$	5.528—5.764 A_2 在其内	0.236
4	$A_5 = 5.674$	8	8	162.89231	0.12267	0.23483	$Q_5 > Q_2$	5.528—5.674 A_2 在其内	0.146
5	$A_6 = 5.584$	8	7	238.51311	0.13678	0.24769	$Q_6 > Q_2$	5.584—5.674 A_2 在其内	0.09
6	$A_7 = 5.640$	8	8	192.81512	0.12922	0.18106	$Q_7 < Q_2$	5.618—5.674 A_7 在其内	0.056
7	$A_8 = 5.652$	8	8	182.3991	0.12676	0.189921	$Q_8 > Q_7$	5.618—5.652 A_7 在其内	0.034
8	$A_9 = 5.630$	8	8	208.88912	0.13213	0.18065	$Q_9 < Q_7$	5.618—5.640 A_9 在其内	0.022
9	$A_{10} = 5.628$	8	8	212.89109	0.13282	0.18269	$Q_{10} > Q_9$	5.628—5.640 A_9 在其内	0.012
10	$A_{11} = 5.638$	8	8	195.61237	0.12975	0.18002	$Q_{11} < Q_9$	5.630—5.640 A_{11} 在其内	0.01

经过10次搜索，取 $A = 5.638$ ， $K = 0.12975$ ， $a = 195.61237$ ，则可得拟合生长曲线

$$\hat{y} = \frac{5.638}{1 + 195.61237e^{-0.12975t}}$$

A的精度为1%，剩余平方和为 $Q = 0.18002$ ，Q的精度为

$$Q_7 - Q_9 = 0.18106 - 0.18065 = 0.00041 = 0.04\%$$

在搜索过程中，为加速收敛速度，还可采取抛物线加速法，如搜索到某一步后，待估参数 θ 的搜索区间为 $[\theta_1, \theta_3]$ ， θ_2 在其中，下一个点可取

$$\theta_4 = \frac{1}{2} \frac{Q_1(\theta_2^2 - \theta_3^2) + Q_2(\theta_3^2 - \theta_1^2) + Q_3(\theta_1^2 - \theta_2^2)}{Q_1(\theta_2 - \theta_3) + Q_2(\theta_3 - \theta_1) + Q_3(\theta_1 - \theta_2)}$$

如上例，搜索到第6步，A的搜索区间为 $[5.618, 5.674]$ ， $A_7 = 5.640$ 在其内。令

$$\theta_1 = 5.618, Q_1 = 0.20185; \theta_2 = 5.640, Q_2 = 0.18106; \theta_3 = 5.674, Q_3 = 0.23483, \text{则}$$

$$\theta_4 = 5.639$$

它与 $A_{11} = 5.638$ 很接近。

图3 为0618算法框图。

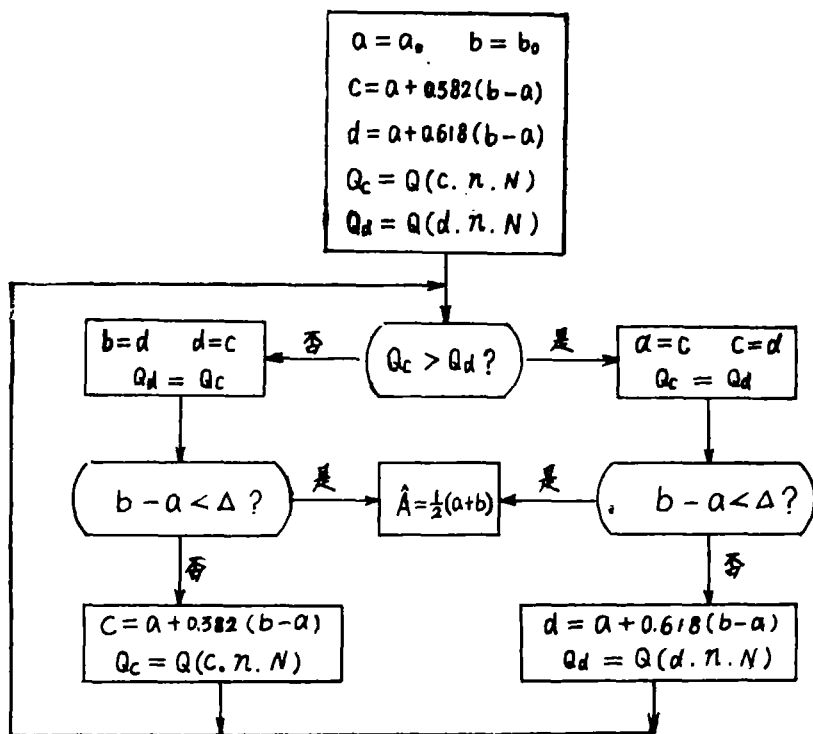


图3

参考文献

- [1] E. C. 皮诺:《数学生态引论》, 科学出版社, 1978。
 [2] 杨纪柯、齐翔林、陈霖:《生物数学概论》, 科学出版社, 1982。
 [3] 中国科学院数学研究所运筹室优选法小组:《优选法》, 科学出版社, 1975。

Application of 0.618 Method To Fitting Growth Curve of two kinds

Yuan Zhi—fa Liu Guang—zu Xiao Jun—zhang

Abstract

0.618 method is suggested to fit Mitecherlich growth curve with $y=A+ae^{-kt}$ and Logistic curve with $y=A/(1+ae^{-kt})$. In this paper, if the traced interval of A begins with (a_0, b_0) , precison be $\Delta=(0.618)^m(b_0-a_0)$ after having traced m orders.