

双倒数法和优化法求算 Michaelis-Menten 方程 参数的对比研究^①

龙怀玉^① 李韵珠

(中国农业大学资源与环境学院)

摘要 从数理上考察了双倒数法求算 Michaelis-Menten 方程参数的误差后,提出了一种“约束条件下 n 维极值复形调优法”的优化求算方法。运用一个具体算例证明了双倒数法对实验误差相当敏感,它要求实验误差控制在 0.1% 之内,才能保证求算出的 Michaelis-Menten 方程参数不出现错误结果,因而是一种不可靠的求算法。而优化法即使在实验数据具有 10% 的误差的情况下仍能保证求算结果的误差不超过 10%。认为优化法是一种很好的求算 Michaelis-Menten 方程参数的方法。

关键词 Michaelis-Menten 动力学; 双倒数法; 优化法

分类号 Q945.12

Comparison of Double-Reciprocal Method and Optimization Method in Solving the Parameters in Michaelis-Menten Equation

Long Huaiyu Li Yunzhu

(College of Resource & Environment Sciences, CAU)

Abstract The error of the double-reciprocal method in solving the parameters in Michaelis-Menten equation was investigated, and a optimization method was established. With a example analysis, it was clear that the optimization method was better than the double-reciprocal method.

Key words Michaelis-Menten equation; double-reciprocal method; optimization method

Michaelis-Menten 动力学方程(1)经常用来描述生化反应和根系吸收动力学过程,根据实验数据求算其参数目前常用的是双倒数法^[1],然而双倒数法有时会得到一些令人费解的结果,例如得到的 V_{\max} 和 K_m 为负数,使得一些实际上符合 Michaelis-Menten 动力学方程的数据却得出错误的结果,特别是根系吸收动力学,增加参数 C_{\min} 后,不便直接用双倒数法求算参数,往往是在双倒数法的基础上对 C_{\min} 进行迭代求解。但在迭代过程中,有时会遇到 $C_{\min} - C \approx 0$ 的情况,使迭代失败,为了避免迭代失败,不得不减少数据,却增大了误差。因此有必要寻找一种新的求算米氏动力学参数的方法。

$$\begin{cases} V = \frac{V_{\max} \cdot C}{K_m + C} & (1a) \\ V = \frac{V_{\max} \cdot (C - C_{\min})}{K_m + C - C_{\min}} & (1b) \end{cases}$$

收稿日期: 1999-06-01

①龙怀玉,北京林业大学博士后流动站,100083

式中 V 为生化反应或根系吸收速率, V_{\max} 为最大生化反应或根系吸收速率, K_m 为米氏常数, C 为底物或离子的浓度, C_{\min} 为根系吸收速率为零的离子浓度。

1 关于 Michaelis-Menten 方程参数的双倒数求算法

1.1 双倒数求算过程中误差将被恶性传播

双倒数求算法就是将方程的两边同时取倒数, 得: $\frac{1}{V} = \frac{K_m}{V_{\max}} \cdot \frac{1}{C} + \frac{1}{V_{\max}}$ 根据实验数据作 $\frac{1}{V} \sim \frac{1}{C}$ 的直线相关方程得出: 斜率 $\alpha = \frac{K_m}{V_{\max}}$, 截距 $b = \frac{1}{V_{\max}}$ 从而计算出 V_{\max} 与 K_m 。实验过程不可能没有误差, 设 V 为速率测定值, $V_{\text{真}}$ 为真实速率值, e 为速率误差, $e_{\text{倒}}$ 为速率倒数的误差, 则有:

$$V = V_{\text{真}} + e$$

$$e_{\text{倒}} = \frac{1}{V_{\text{真}}} - \frac{1}{V_{\text{真}} + e} = \frac{e}{V_{\text{真}}(V_{\text{真}} + e)} \approx \frac{e}{V_{\text{真}}^2} \quad (2)$$

由(2)式可以看出, 倒数误差是原误差的 $V_{\text{真}}^{-2}$ 倍, 对于 C 亦然。由于在低浓度部分是速率变化最为剧烈的地方, 是体现一个生化反应或根系吸收动力学是否符合 Michaelis-Menten 动力学方程的最重要的部分, 因此不能为了回避参数求算中的误差问题, 而不用低浓度部分的数据。当 $C \rightarrow 0, V \rightarrow 0, e_{\text{倒}} \rightarrow \infty$, 因此误差严重地干扰了真实值, 以致得出错误结果。

1.2 双倒数法对误差的敏感性考察

本研究不从理论上对双倒数法对误差的敏感性作出定量分析, 而是通过一个具体算例考察双倒数法对误差的敏感性, 为此在方程 $V = \frac{30 \cdot C}{2 + C}$ 上按(3)式取 40 个点, 分以下 3 种情况考察双倒数法对误差的敏感性:

$$\begin{cases} C_n = 0.05 \cdot (1+n) & 0 \leq n \leq 19 \\ C_n = 1.0 + (n-19) & 20 \leq n \leq 39 \end{cases} \quad (3)$$

- ① 维持 C 不变, 给 V 随机加一个误差 e_{λ} , 即 $V = V \cdot (1 + e_{\lambda} \cdot \text{rand}())$
- ② 维持 V 不变, 给 C 随机加一个误差 e_{λ} , 即 $C = C \cdot (1 + e_{\lambda} \cdot \text{rand}())$
- ③ 同时给 V, C 加一个误差 e_{λ} , 即 $V = V \cdot (1 + e_{\lambda} \cdot \text{rand}()), C = C \cdot (1 + e_{\lambda} \cdot \text{rand}())$

以上各式中的 $\text{rand}()$ 为随机函数, 其值在 $(-1, 1)$ 上。 e_{λ} 取值为 $10^{-3} \sim 10^{-1}$ 。

根据概率论知识, 当将某一随机实验进行的次数很大时, 事件的频率近似地等于事件的概率。因此将以上三种情况分别计算 10^4 次, 记录不同输出误差的出现频率, 其结果(表 1)表明:

(1) 在每种误差输入情况下, 在同一输出误差范围内, V_{\max} 与 K_m 出现的频率几乎相等。这说明: 双倒数法求算动力学参数时, V_{\max} 与 K_m 对相同输入误差的敏感性是一致的。

(2) 只有当输入误差小于 10^{-3} 时, 3 种误差输入情况才能使参数求算结果出现错误的频率才为零。这说明为了防止出现错误结果, 必须把输入误差控制在 10^{-3} 之内。

(3) 3 种情况下均在 10^{-2} 开始出现错误结果, 但①出现错误的概率(13.15%)小于②出现错误的概率(27.67%), ②出现错误的概率小于③出现错误的概率(33.75%); 使输出误差不超过 5% 的输入误差, ①(10^{-4})大于②(10^{-5}); 使输出误差不超过 10% 的输入误差, ②(10^{-4})大于③(10^{-5}); 这说明: 双倒数法求算动力学参数对浓度输入误差比速率输入误差要敏感; 同时具有 2 种误差时的求算结果比只有单独 1 种误差时的求算结果更不可靠。

表 1 不同误差来源下双倒数法的输出误差的频率(%)分布

e_{λ}	输出误差 $e_{出} = \text{理论值} - \text{求算值} / \text{理论值} \times 100\%$									
	0~5		5~10		10~50		>50		错误	
	V_{max}	Km	V_{max}	Km	V_{max}	Km	V_{max}	Km	V_{max}	Km
①维持 C 不变、给 V 随机加上误差 e_{λ} 时的输出误差频率分布										
10^{-5}	100	100	0	0	0	0	0	0	0	0
10^{-4}	100	100	0	0	0	0	0	0	0	0
10^{-3}	42.26	42.2	31.05	30.96	26.69	26.84	0	0	0	0
10^{-2}	4.26	4.28	4.51	4.42	41.37	41.22	36.71	37.93	13.15	13.15
10^{-1}	0.38	0.37	0.51	0.51	4.92	4.94	48.72	48.71	45.47	45.47
②维持 V 不变、给 C 随机加上误差 e_{λ} 时的输出误差频率分布										
10^{-5}	100	100	0	0	0	0	0	0	0	0
10^{-4}	99.86	99.86	0.14	0.14	0	0	0	0	0	0
10^{-3}	22.16	22.14	19.99	19.94	55.00	55.07	2.85	2.85	0	0
10^{-2}	1.89	1.91	2.14	2.12	23.65	23.58	44.65	44.72	27.67	27.67
10^{-1}	0.21	0.22	0.23	0.21	2.55	2.56	47.98	47.98	49.03	~49.03
③同时给 C, V 随机加上误差 e_{λ} 时的输出误差频率分布										
10^{-5}	100	100	0	0	0	0	0	0	0	0
10^{-4}	95.76	95.71	4.23	4.28	0.01	0.01	0	0	0	0
10^{-3}	15.84	15.83	14.96	14.97	59.79	59.76	9.41	9.44	0	0
10^{-2}	1.75	1.70	1.60	1.61	17.60	17.64	45.27	45.30	33.75	33.75
10^{-1}	0.12	0.11	0.15	0.16	1.84	1.84	49.85	49.85	48.04	48.04

以上结果说明双倒数法对实验数据的误差要求是十分苛刻的,即使宽容地允许求算出的参数有 50% 的误差,要使计算结果具有 90% 的可靠性,也要将实验误差控制在 10^{-3} 之内;如仅仅要求不得出错误结果,也要把实验误差控制在 10^{-2} 之内;由于实验误差不可避免,特别是在低浓度部分,要将误差控制在 10^{-2} 之内几乎不可能,因此双倒数法求算出的结果是很不可靠的,甚至是错误的。

2 用优化法求算 Michaelis-Menten 动力学方程

2.1 优化方法的选取及其主要计算步骤

为了防止产生象双倒数法那样的误差扩大问题,可以用优化的方法直接求算 Michaelis-Menten 方程参数。对于生化反应,根据生物学意义,要求 $V_{max} > 0, Km > 0$, 对于根系吸收动力学过程,还要求 $C_{min} \geq 0, Km > C_{min}$, 因此 Michaelis-Menten 动力学方程的优化求参,是一个既有常数约束条件,又有函数约束条件的优化问题,可以选用“约束条件下 n 维极值的复形调优法”^[2],其基本原理是:在 n 维空间中,构造 $2n$ 个顶点,比较各个顶点的函数值大小,并以去掉最大点后的中心点为参照,在满足约束条件下,使最大方向收缩或反向延伸,使最小方向维持或正向延伸,反复计算,直到各个顶点之间的距离小于预定精度。其计算步骤如下^[2]:

①随机给定 $2n$ 个满足约束条件的顶点,计算目标函数值: $f(j) = f(X_j), j = 1, 2, 3, \dots, 2n$ 。



②确定: $F(R) = f(X_R) \max_{1 \leq j \leq 2n} f(i)$

$$f(G) = f(X_G) = \max_{1 \leq i \leq 2n, i \neq R} f(i)$$

其中 X_R 称为最坏点, X_G 称为次坏点。

③计算最坏点的对称点: $X_T = (1 + \alpha)X_F - \alpha X_R$

其中
$$X_F = \frac{1}{2n-1} \sum_{i=1, i \neq R}^{i=2n} X_i$$

α 称为反射系数, 一般取 1.3 左右。

④确定一个新顶点以代替最坏点 X_R , 构成新的复形。

如果 $f(X_T) > f(G)$, 则 $X_T = (X_F + X_T)/2$, 直到 $f(X_T) \leq f(G)$ 为止。然后检查 X_T 是否满足约束条件, 如果某个分量 $X_T(j)$ 不满足常量约束条件, 即 $X_T(j) < a_j$, 或 $X_T(j) > b_j$, 则 $X_T(j) = a_j + \delta$, 或 $X_T(j) = b_j - \delta$, δ 是一个非常小的量。如果 X_T 不满足函数约束条件, 则用下式修改 X_T :

$$X_T = (X_F + X_T)/2$$

重复④, 直到 $f(X_T) \leq f(G)$, 且 X_T 满足所有的约束条件为止, 此时令

$$X_R = X_T, f(R) = f(X_T)$$

重复②~④, 直到复形中各顶点距离小于预先给定的精度为止。

将上述“约束条件下 n 维极值的复形调优法”应用到 Michaelis-Menten 动力学方程则:

目标函数为:
$$\begin{cases} f(V_{\max}, Km, C_m) = \sum_{i=1}^n (V_i - \bar{V}_i)^2 \\ \bar{V}_i = \frac{V_{\max} \cdot (C_i - C_{\min})}{Km + C_i - C_{\min}} \end{cases}$$

常量约束条件为:
$$\begin{aligned} 0 < V_{\max} < +\infty \\ 0 < Km < +\infty \\ 0 < C_{\min} < +\infty \end{aligned}$$

函数约束条件为:
$$C_{\min} < Km$$

在实际计算中 $+\infty$ 为各参数理论值的 20 倍。

2.2 优化法对误差的敏感性考查

由于在理论上难以定量考察优化法对输入误差的敏感性, 同时也为了与双倒数法进行对比, 应用(3)式的数据和同样的误差输入法, 计算 10 000 次, 记录不同大小误差的发生频率, 结果说明(表 2):

(1)在每种误差输入情况下, V_{\max} 的误差不超过 5%。这说明: 当输入误差小于 10% 时, 用优化法求算动力学参数, V_{\max} 对输入误差是不敏感的, 即使在浓度和速率均存在 10% 的误差的情况下, V_{\max} 的误差也不超过 5%。

(2)当①、③的输入误差小于 1/100、②的输入误差小于 1/10 时, Km 的误差不超过 5%。当输入误差为 10% 时, ①、③ Km 的输出误差频率分布差不多, ②的 Km 输出误差不超过 5%, 这说明优化法求参时, Km 对速率误差比对浓度误差要敏感。但即使同时给浓度和速率输入 10% 的误差, Km 的输出误差超过 10% 的频率也不超过 4%。

表 2 不同误差来源下优化法的输出误差频率(%)分布

e_{λ}	输出误差 $e_{出} = \text{理论值} - \text{求算值} / \text{理论值} \times 100\%$									
	0~5		5~10		10~50		>50		错误	
	V_{max}	Km	V_{max}	Km	V_{max}	Km	V_{max}	Km	V_{max}	Km
①维持 C 不变、给 V 随机加上误差 e_{λ} 时的输出误差频率分布										
1/100	100	100	0	0	0	0	0	0	0	0
1/10	100	70.05	0	26.95	0	2.15	0	0.85	0	0
②维持 V 不变、给 C 随机加上误差 e_{λ} 时的输出误差频率分布										
1/100	100	100	0	0	0	0	0	0	0	0
1/10	100	100	0	0	0	0	0	0	0	0
③同时给 C, V 随机加上误差 e_{λ} 时的输出误差的分布频率										
1/100	100	100	0	0	0	0	0	0	0	0
1/10	100	69.49	0	27.21	0	3.30	0	0	0	0

注: 由于当输入误差小于 1/1 000 时, 3 种误差输入的输出误差均小于 5%, 因此只列出输入误差为 1/100 和 1/10 的结果。

以上是当 C_{min} 为零, 用优化法求参时, 输入误差对 V_{max} 和 Km 的影响。为了考察当 C_{min} 不为零, 且用优化法求参时, 输入误差对 V_{max} 、 Km 和 C_{min} 的影响, 在根系吸收动力学方程 $V = \frac{30 \cdot (C - 0.2)}{2 + C - 0.2}$ 取如(3)式的 40 个点。按前述的方法考察输入误差对参数的影响, 结果表明(表 3):

表 3 优化法求算根系吸收动力学时不同误差来源下输出误差频率(%)分布

e_{λ}	输出误差 $e_{出} = \text{理论值} - \text{求算值} / \text{理论值} \times 100\%$											
	0~5			5~10			10~50			>50		
	V_{max}	Km	C_{min}	V_{max}	Km	C_{min}	V_{max}	Km	C_{min}	V_{max}	Km	C_{min}
①维持 C 不变、给 V 随机加上误差 e_{λ} 时的输出误差的分布频率												
1/100	100	100	100	0	0	0	0	0	0	0	0	0
1/10	100	92.34	92.34	0	3.05	3.05	0	4.61	4.61	0	0	0
②维持 V 不变、给 C 随机加上误差 e_{λ} 时的输出误差分布频率												
1/100	100	100	100	0	0	0	0	0	0	0	0	0
1/10	100	96.48	96.48	0	2.30	2.30	0	1.22	1.22	0	0	0
③同时给 C, V 随机加上误差 e_{λ} 时的输出误差分布频率												
1/100	100	98.12	98.12	0	1.88	1.88	0	0	0	0	0	0
1/10	100	82.44	82.44	0	8.56	8.56	0	9.00	9.00	0	0	0

(1) 在 3 种误差输入情况下, Km 和 C_{min} 的输出误差频率分布相同, 说明输入误差对 Km 和 C_{min} 的影响相同。当输入误差小于 10% 时, V_{max} 的输出误差不超过 5%, 而 Km 和 C_{min} 的输出误差不超过 5% 的频率小于 100%, 说明在根系吸收动力学中输入误差对 V_{max} 的影响要比对 Km 和 C_{min} 的影响要小。

(2) 当输入误差为 10% 时, Km 、 C_{min} 输出误差分布在 5% 以内的频率②比①要大, 说明 Km 、 C_{min} 对速率误差要比对浓度误差敏感。在两种不同大小的误差输入下, Km 、 C_{min} 分布在

5%的误差以内的频率③比①、②要小,这说明浓度误差和速率误差共同影响了 K_m 、 C_{min} 的输出误差。

3 讨论

前面的结果充分说明,在求算 Michaelis-Menten 方程参数时,优化法的求算结果要比双倒数法的求算结果可靠。文中所引用的算例虽是 1 组随机从根据实验可能得到的数据中取来的特例,但其说明了以上结论。从(2)式可知,只有当 C 或 V 小于 1 时, e 在双倒数法中才会被恶性传播。(3)式的数据只有 50% 小于 1,双倒数法对输入误差的要求就已非常苛刻了,在实际中的浓度,特别是速率常常以 10 的 $-n$ 次方出现(n 为正整数),因此在实践中双倒数法更不可取。

本文得到任理副教授、冀元石副教授的指导,在此致谢。

参 考 文 献

- 1 Epstein E, Hagen C E. A kinetic study of the absorption of alkali cations by barley roots. *Plant Physiology*, 1952, 27: 457~474
- 2 徐士良. C 常用算法程序集. 北京:清华大学出版社,1996, 351~359