

文章编号:1006-0456(2008)02-0134-04

基于潜周期模型的两种群食饵-捕食者模型的参数估计

马新生,王来群,胡文玉

(南昌大学数学系,江西 南昌 330031)

摘要:给出了一种两种群食饵-捕食者模型参数估计的两阶段方法。考虑到食饵-捕食者模型解的周期性,首先用潜周期模型来逼近观测序列,然后将模型参数的非线性最小二乘估计问题转化为一个线性最小二乘估计。两阶段法求解过程中不需要数值求解微分方程,具有计算方便,能同时得出参数的估计值及其置信区间的优点,改进了现有算法只能得到参数点估计的不足。算例结果表明,该算法误差很小,比现有算法的精度大大提高。

关键词:潜周期模型;参数估计;食饵-捕食者模型

中图分类号:0212.1

文献标识码:A

Parameter Estimation in Two Species Predator-prey Equations by Using Hidden Periodic Model

MA Xin-sheng, WANG Lai-qun, HU Wen-yu

(Department of Mathematics, Nanchang University, Nanchang 330031, China)

Abstract: In this paper, the method of Hidden Periodic Model was used to study the parameters estimation in two species predator-prey model. Since the two species predator-prey model has periodic solutions, the method of Hidden Periodic Model was first used to approximate the original data. And then the non-linear least square was transformed into a simple least square method. This estimating method was called as two-stage estimation (TSE). TSE does not need numerical solution of the ODE and is easy to be carried out. Especially, we may have the interval estimation of the parameters that modified the existing algorithm. The simulation result showed that the errors were much smaller than the published study (eg, see [4]).

Key Words: hidden periodic model; estimation of parameters; predator-prey model

微分方程模型参数估计是目前国际上概率统计领域的研究热点^[1-5]。微分方程模型是描述实际对象的某些特性随时间或空间演变的过程,进而分析它的变化规律、预测它的未来性态、研究它的控制手段的重要方法。但是,诸如生物科学、物理学等领域中建立的微分方程通常没有解析解,实际应用中一般使用它们的数值解。数值解的精度直接依赖于方程中的各参数以及初始值(或边界值)的精度。Himmelblau^[6]在1967年首次利用统计方法进行了微分方程参数估计的研究,Bock^[7]于1983年综述了自20世纪60年代以后20年间常微分方程参数辨

识的研究进展,指出现有的利用观测数据估计非线性微分方程模型参数的方法效果较差,如计算速度慢、拟合精度差、得不到有较好统计解释的区间估计等,无法满足实际工作的需要。Jerome^[4]在2005年直接考虑求解参数的非线性最小二乘估计,利用信赖域方法求解非线性优化问题,其结果虽然较为精确,但计算过程十分复杂。2006年Jiguo Cao^[2]和2007年Ramsay^[4]采用函数型数据分析法研究了一类微分方程的参数估计方法。这些方法主要存在三点不足,一是,在计算过程中需要数值求解微分方程;二是,没有充分利用微分方程的大量定性研究结

收稿日期:2008-01-20

基金项目:江西省自然科学基金资助项目(2007GZS2398);南昌大学研究生创新专项资金资助

作者简介:马新生(1966-),男,教授,博士,maxinsheng@ncu.edu.cn。

果,如其相轨线的周期性^[8]等;三是,不能得出具有统计意义的置信区间,这在参数估计理论中具有十分重要的作用^[9]。本文将以一类简单的食饵-捕食者模型为例,以其解的周期型为基础,导出一种微分方程参数估计的两阶段方法。该方法将非线性最小二乘问题转化为时间序列分析的潜周期估计和一个线性最小二乘问题,求解过程中不需要数值求解微分方程,能同时得出参数的估计值及其置信区间。

1 两种群食饵-捕食者模型参数的非线性最小二乘估计

在自然界中,许多种群竞争同一有限食物,从而抑制了相互间种群的增长,这是一个有趣而普遍的现象。Lotka-Volterra 竞争系统是著名的种群模型之一,以两种群生物为例,记A生物是捕食者,B生物是被捕食者,假设*t*时刻捕食者A的数量为*x(t)*,被捕食者B的数量为*y(t)*,它们之间满足以下变化规律^[7]:

$$\begin{cases} x'(t) = \alpha_1 x(t) - \alpha_2 x(t)y(t) \\ y'(t) = \alpha_3 y(t) - \alpha_4 x(t)y(t) \end{cases} \quad (1)$$

初始条件为:

$$\begin{cases} x(t_0) = X_0 \\ y(t_0) = Y_0 \end{cases} \quad (2)$$

方程(1)是非线性方程,没有解析解,给定参数和初始条件后可采用数值求解。

式(1)中的参数加上初始条件共6个未知参数的估计问题可表示为如下无约束最优化问题:

$$\min_x f(x) \quad (3)$$

其中,

$x = (X_0, Y_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4)$ 为待估参数,

$$f(x) = \frac{1}{2} (\|X - \hat{X}\|_2^2 + \|Y - \hat{Y}\|_2^2)$$

X, *Y* 为观测序列, \hat{X} , \hat{Y} 为估计值, N_{obs} 为观测数据总数。

由于 \hat{X} , \hat{Y} 的计算需要通过数值求解方程(1)(2),故采用通常的非线性优化算法求解(3)是比较复杂的^[4]。

2 潜周期模型

潜周期模型是一种处理实际问题中周期性数据的常用方法^[10-11]。设观测序列为 $\{z(t), t \in N_+\}$, 且均值为0(否则,作零均值化处理),则采用:

$$z(t) = \sum_{j=1}^k A_j \cos(\omega_j t + \varphi_j) + \xi_t, t \in N_+ \quad (4)$$

进行逼近^[8]。其中, $0 < \omega_1 < \omega_2 < \dots < \omega_k \leq \pi$, 正数 A_j 是相应于第 j 个角频率 ω_j 的振幅, $\varphi_j \in [0, 2\pi)$ 是相应于角频率 ω_j 的初相位,且 A_j 和 φ_j 均为常数。 $\{\xi_t\}$ 是一个零均值的线性平稳序列,为有色噪声。满足模型(4)的时间序列 $\{z(t), t \in N_+\}$ 称为潜频率序列。

对(4)式中周期个数的估计 \hat{k} , 角频率的估计 $\hat{\omega}_j$ ($1 \leq j \leq \hat{k}$), 振幅的估计 \hat{A}_j ($1 \leq j \leq \hat{k}$) 和初相位的估计 $\hat{\varphi}_j$ ($1 \leq j \leq \hat{k}$) 的算法如下(详见文献[10])。

首先通过

$$S_N(\lambda) = \sum_{t=1}^N z_t e^{-i\lambda t} \quad (5)$$

求出 $\hat{\omega}_j$ ($1 \leq j \leq \hat{k}$) 的初估计。

然后采用周期图最大估计, 求出 $\hat{\lambda} = 2\pi/\hat{\omega}_j$, 并得出振幅 \hat{A}_j 的估计,

$$\hat{A}_j = \frac{1}{N} \sum_{t=1}^N z_t e^{-i\hat{\omega}_j t}, 1 \leq j \leq \hat{q} \quad (6)$$

进而求出 φ_j 的估计。可以证明^[10], 以上估计具有相合性, 且收敛速度为 $o((N^3/\ln N)^{-1/2})$ 。

拟合模型的统计检验通过残差

$$\hat{\xi}_t = z_t - \sum_{j=1}^{\hat{k}} \hat{A}_j \cos(\hat{\omega}_j t + \hat{\varphi}_j), t = 1, 2, \dots, N$$

的样本自协方差函数

$$\hat{\gamma}_k = \frac{1}{N} \sum_{j=1}^{N-k} (\hat{\xi}_j - \hat{\mu})(\hat{\xi}_{j+k} - \hat{\mu}), k = 1, 2, \dots, [\sqrt{N}] \quad (7)$$

来判定, 其中, $\hat{\mu}$ 是 $\hat{\xi}_t$ 的样本均值。如果 $\hat{\gamma}_k$ 收敛到零, 则通过检验, 认为模型是合适的。

3 两种群食饵-捕食者模型参数的两阶段估计法

考虑到方程(1)的解具有周期性^[7], 以及(1)可化为线性方程:

$$\begin{cases} z_1(t) = \alpha_1 - \alpha_2 y(t) \\ z_2(t) = \alpha_3 - \alpha_4 x(t) \end{cases} \quad (8)$$

其中

$$\begin{cases} z_1(t) = \frac{x'(t)}{x(t)} \\ z_2(t) = \frac{y'(t)}{y(t)} \end{cases} \quad (9)$$

故可采用如下两阶段法进行参数估计。

第一阶段, 将方程(1)的解 $x(t)$, $y(t)$ 用潜周期模型(4)来近似表示, 即求解

$$\min f_1(x) = \|X - \hat{X}\| \text{ 和 } \min f_2(x) = \|Y - \hat{Y}\|$$

(10)

其中向量 X, Y 为观测序列, 向量 \hat{X}, \hat{Y} 为由 (4) 式计算出的估计值。利用 (9) 式得出 $z_1(t), z_2(t)$ 的估计, 并得到方程 (1) 初始值的估计。

第二阶段, 由式 (8) 中两方程估计其余参数, 并求出置信区间。此时, 即为求解两个线性最小二乘问题

$$\min f_3(x) = \|\hat{Z}_1 - \hat{Z}_2\| \text{ 和 } \min f_4(x) = \|\hat{Z}_2 - \hat{Z}_2\|$$

(11)

其中, 向量 \hat{Z}_1, \hat{Z}_2 为第一阶段通过潜周期模型得到的 z_1, z_2 的估计值, 向量 \hat{Z}_1, \hat{Z}_2 为由方程 (8) 得出的 z_1, z_2 的估计值。

由此, 非线性最优化问题 (3) 式转化为潜周期估计问题 (10) 和线性最小二乘问题 (11)。

4 数值算例

以文献 [4]、[5] 中数据产生方法构造模拟序列来验证两阶段方法的准确性。参数的真值取为 $x = (13, 72, 2, 0.1, 10, 1)$, 观测序列采用四阶 Runge-Kutta 法数值求解方程 (1), 时间区间取 $T = [0, 15]$, 采样间隔 $\Delta T = 0.1, N_{\text{obs}} = 151$ 。不妨假设, 这样得出的观测序列没有误差, 为更准确模拟实际观测序列, 我们给数据再叠加噪声 $N(0, \sigma^2)$, 取 $\sigma = 0.5$ 。对产生的序列采用两阶段法估计参数。

首先, 分别对 $x(t)$ 和 $y(t)$ 用潜周期模型逼近。结果分别为,

$$x(t) = 9.8931 + \sum_{j=1}^4 \hat{A}_j \cos(\hat{\omega}_j t + \hat{\varphi}_j), t \in N_+$$

(12)

和

$$y(t) = 20.8207 + \sum_{j=1}^4 \hat{A}_j \cos(\hat{\omega}_j t + \hat{\varphi}_j), t \in N_+$$

(13)

(12) 和 (13) 中各系数估计值见表 1 和表 2。

表 1 式 (12) 中各系数的值
Tab.1 Coefficients in the equation (12)

| k | 1 | 2 | 3 | 4 |
|-------------------|---------|---------|----------|----------|
| $\hat{\omega}$ | 3.867 7 | 7.744 8 | 11.619 0 | 15.606 |
| $\hat{\varphi}_k$ | 0.792 5 | 0.633 1 | 0.327 5 | -0.698 3 |
| \hat{A}_k | 6.930 5 | 2.538 6 | 0.920 6 | 0.377 6 |

表 2 式 (13) 中各系数的值
Tab.2 Coefficients in the equation (13)

| K | 1 | 2 | 3 | 4 |
|-------------------|----------|----------|----------|----------|
| $\hat{\omega}$ | 3.864 8 | 7.747 9 | 11.618 8 | 15.489 6 |
| $\hat{\varphi}_k$ | -0.667 8 | -1.342 1 | -1.834 1 | -2.262 2 |
| \hat{A}_k | 29.213 5 | 16.414 1 | 8.400 6 | 4.090 3 |

$x(t)$ 和 $y(t)$ 模型误差的自协方差序列如图 1 所示。从图 1 看出, 自协方差函数接近于零, 表明模型通过统计检验, 方程 (12) 和 (13) 可作为 $x(t)$ 、 $y(t)$ 的逼近。作出 $x(t)$ 、 $y(t)$ 的观测值和利用潜周期模型得出的估计值, 见图 2, 近似效果很好。

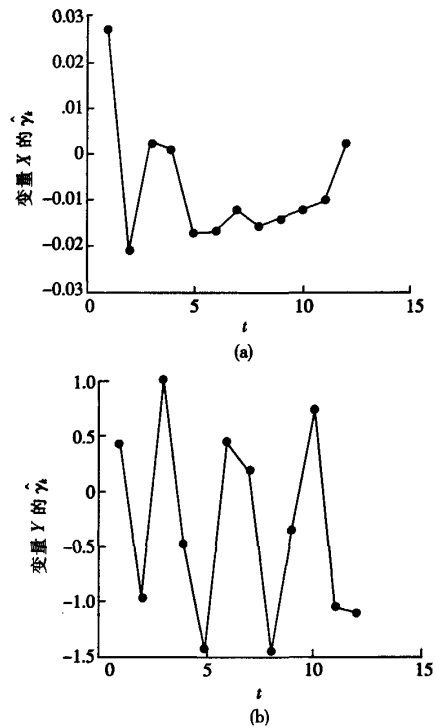


图 1 潜周期模型自协方差函数统计检验图

Fig.1 Diagram of auto-covariance of errors for statistical testing

然后, 利用线性最小二乘法, 求出各参数估计值。综合 $X(0)$ 、 $Y(0)$ 的估计值, 结果列于表 3。

表 3 参数估计值及其误差
Tab.3 Parameter values obtained and their error

| 待估参数 x | X(0) | Y(0) | α_1 | α_2 | α_3 | α_4 |
|-----------------------------|---------|---------|--------------------|--------------------|--------------------|--------------------|
| 真实值 | 13 | 72 | 2 | 0.1 | 10 | 1 |
| 估计值 | 13.025 | 72.494 | 2.129 5 | 0.101 8 | 9.989 0 | 0.982 3 |
| 相对误差/% | 0.192 3 | 0.686 1 | 6.475 | 1.800 0 | 0.110 0 | 1.770 0 |
| 置信区间 ($\alpha = 0.95$) | | | (1.969 6, 2.323 0) | (0.098 3, 0.105 4) | (7.492 0, 12.4658) | (0.749 2, 1.215 3) |

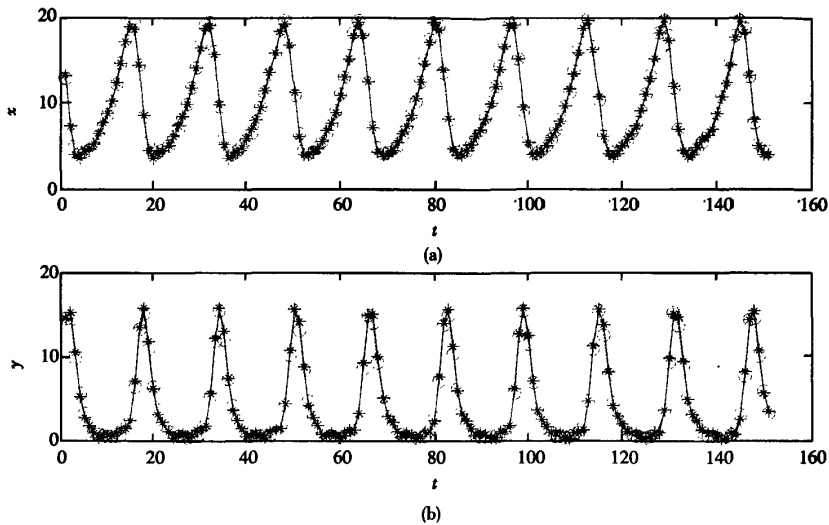


图2 观测序列和潜周期模型逼近序列(o 为原始数据,*为预测数据)

Fig.2 Reference and predicted data versus time

从表3中可知,各参数的估计精度很高,相对误差的平均值只有1.8389%,远小于文献[4]中近10%的平均误差,而且最大误差 $\|e\|_{\infty} = 6.475\%$ 也小于[4]中的 $\|e\|_{\infty} = 11.6\%$ 。置信区间(置信度0.95)不含0点,且均包含真值,结果是可靠的。

本方法的一个缺点是方程(1)的初始值的区间估计无法得出,这也是潜周期模型需要研究的一个内容^[10-11],有待今后进一步研究。

参考文献:

- [1] Fussmann G F, Ellner S P, Shertzer K W, et al. Crossing the hopf Bifurcation in a Live Predator-prey system[J]. Science, 2000, 290: 1358 - 1360.
- [2] Cao Jiguo. Generalized Profiling Method and the Applications to Adaptive Penalized Smoothing, Generalized Semiparametric Additive Models and Estimating Differential Equations[D]. Montreal: McGill University, 2006.
- [3] Ramsay J O, Hooker G, Cao J, et al. Estimating Differential Equations[J]. Journal of the Royal Statistical Society, Series B, 2007, 69(5): 741 - 796
- [4] Jerome M B Walmag, Eric J M Delhez. A Trust-region

Method Applied to Parameter Identification of a Simple Prey-predator model[J]. Applied Mathematical Modeling, 2005, 29: 289 - 307.

- [5] 陈务深, 戴风, 甘泉. 确定高精度参数问题的评注[J]. 数学的实践与认识, 2007, 37(14): 90 - 94.
- [6] Himmelblau D, Jones C, Bischoff K B. Determination of rate Constants for Complex Kinetics Models[J]. Industrial Engineering Chemistry Fundamentals, 1967, 6(2): 539 - 552.
- [7] Bock H G. Recent advances in Parameter Identification Techniques for ode. [M]//Deuffhard P, Harrier E. Numerical Treatment of Inverse Problems in Differential and Integral Equations. Basel: Birkhauser, 1983: 95 - 121.
- [8] 姜启源, 谢金星, 叶俊. 数学模型[M]. 北京: 高等教育出版社, 2004: 192 - 197.
- [9] 陈希孺, 倪国熙. 数理统计学教程[M]. 上海: 上海科学技术出版社, 1988: 250 - 268.
- [10] 何书元. 应用时间序列分析[M]. 北京: 北京大学出版社, 2005: 230 - 246.
- [11] Brockwell P J, Davis R A. Time Series: Theory and Method[M]. New York: Springer-Verlag, 1991: 324 - 332.