第22卷 第2期 DOI:10.13203/j.whugis1997.02.007 武汉测绘科技大学学报 Vol.22 No.2
1997年6月 Journal of Wuhan Tehcnnical University of Surveying and Mapping June 1997

非线性模型线性近似的容许曲率*

王新洲

(武汉测绘科技大学地学测量工程学院,武汉市珞喻路 39号, 430070)

摘 要 由于不同的非线性模型具有不同的非线性强度,使得一些非线性模型可以线性近似,而另一些则不能。本文介绍度量非线性强度的方法,提出判断非线性模型能否线性近似的数值标准—— 容许曲率。

关键词 非线性模型;非线性强度;容许曲率;线性近似

分类号 P 207

大地测量中的函数模型大部分都是非线性模型。处理这类非线性模型的传统方法,通常是利用参数的近似值将其线性化。实践表明,有些非线性模型线性近似后能得到令人满意的结果。而有些非线性模型对参数的近似值十分敏感,只有在参数近似值的近似程度非常好的情况下才能得到可用的结果。还有些非线性模型甚至不能线性近似。之所以会产生这种现象,是因为不同的非线性模型具有不同的非线性强度。对于某一具体的非线性模型,如果其非线性强度较弱,就可以进行线性近似;如果其非线性强度很强,就不能进行线性近似。那么,如何描述非线性强度?如何根据非线性强度来判断非线性模型能否线性近似?这些问题就是本文所要解决的核心问题。

1 非线性强度的度量

设非线性模型为:

$$L = f(x) + \Delta \tag{1}$$

式中,L 为 $n\times 1$的观测向量;Δ 为 $n\times 1$的随机误差向量;x 为 $t\times 1$的参数向量 $(t<n)$;$f(x)$ 是 $n\times 1$的向量,其中各元素 $f_i(x)$ 为参数 x 的非线性函数。为了描述模型 (1) 线性近似后所得结果的优劣, Beale于 1960年提出用曲率来度量模型 (1) 的非线性强度[1]。1980年 Bates和 Watts从微分几何的观点出发,定义了非线性模型的固有曲率和参数效应曲率[2]。下面扼要介绍非线性模型 (1) 的固有曲率和参数效应曲率的定义。

假设模型 (1) 关于 x 存在二阶以上的连续导数。现考虑参数空间 $\mathscr{X}$ 中过 x_0以 h 为方向的一条直线:

$$l_h:\quad x(b) = x_0 + bh \tag{2}$$

式中,b 为实数;h 为固定方向。l 通过 $\eta = f(x)$ 映射到样本空间 R^n中解轨迹 c 上的一条曲线为:

$$C_h:\quad \eta = \eta_h(b) = f(x_0 + bh) \tag{3}$$

曲线 C_h称为提升线,见图 1。

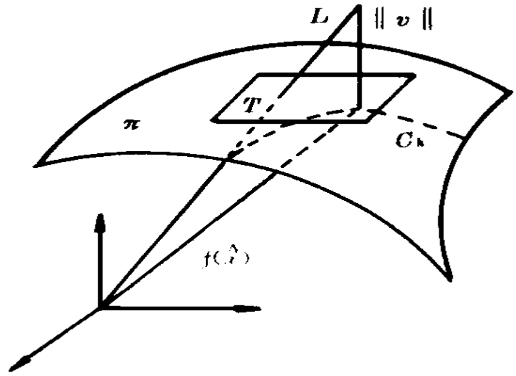

图 1 解轨迹及 LS 估计量

Fig. 1 Solution Locus and Least Square Estimation

由于

$$\begin{cases} \dfrac{d\eta_k}{dt} = \sum\limits_{i=1}^{t} \dfrac{\partial f_k}{\partial x_i}\dfrac{dx_i}{db} = \sum\limits_{i=1}^{t} B_{ki} h_i \\ \dfrac{d^2\eta_k}{db^2} = \sum\limits_{i=1}^{t}\sum\limits_{j=1}^{t} \dfrac{\partial^2 f_k}{\partial x_i \partial x_j}\dfrac{dx_i}{db}\dfrac{dx_j}{db} \\ \qquad = \sum\limits_{i=1}^{t}\sum\limits_{j=1}^{t} C_{kij} h_i h_j \\ \qquad k = 1, 2, \cdots, n \end{cases} \tag{4}$$

因为 $f(x)$ 关于 x 的前二阶导数分别为:

$$B(x) = (B_{ki}),\quad C(x) = (C_{kij})$$

C_{kij}是立阵 C 的第 k 层中第 i 行第 j 列的元素。所以提升线 C_h沿 h 方向的前二阶导数分别为:

$$\dot{\eta}_h = Bh,\quad \ddot{\eta}_h = h' Ch \tag{5}$$

(5) 式中的加速度向量 $\ddot{\eta}_h$可分解为垂直于切平面

收稿日期: 1996-04-12。王新洲,男, 42岁,博士,教授,现从事数据处理研究。

* 国家自然科学基金资助项目,编号 49474204

的法分量 $\ddot{Z}_h^N$ 和切平面上的切分量 $\ddot{Z}_h^T$,即

$$\ddot{Z}_h = \ddot{Z}_h^N + \ddot{Z}_h^T \tag{6}$$

法分量 $\ddot{Z}_h^N$ 是由于解轨迹沿着法方向和弯曲而引起的,切分量 $\ddot{Z}_h^T$ 则是由于切平面上沿着 h 方向及其垂直方向的不均匀性而引起的。于是定义模型 (1) 沿 h 方向在 x_0 处的固有曲率 K_h^N 和参数效应曲率 K_h^T 为:

$$\begin{cases} K_h^N = \|\ddot{Z}_h^N\| / \|\dot{Z}_h\|^2 = \|P_N \ddot{Z}_h\| / \|\dot{Z}_h\|^2 \\ K_h^T = \|\ddot{Z}_h^T\| / \|\dot{Z}_h\|^2 = \|P_T \ddot{Z}_h\| / \|\dot{Z}_h\|^2 \end{cases} \tag{7}$$

式中,

$$P_T = B(B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}, \quad P_N = E - P_T \tag{8}$$

可以证明[3], K_h^N 是一个不依赖于坐标选择的不变量,即 K_h^N 与参数的选取无关,它只是由模型本身固有的性质所决定的,所以称固有曲率。而 K_h^T 不仅由模型本身决定,而且还强烈地依赖于参数的选择,所以 K_h^T 称为参数效应曲率。

另外, x_0 处的最大固有曲率和最大参数效应曲率定义为:

$$K^N = \max_h K_h^N; \quad K^T = \max_h K_h^T$$

式中, h 为所有可能的方向.

2 线性近似的容许曲率

为了研究模型 (1) 线性近似的容许曲率,先将模型 (1) 线性近似,得:

$$L \doteq f(x_0) + B(x_0)(x - x_0) + \Delta \tag{9}$$

在假定 Δ 相互独立,且服从正态分布的前提下, (9) 式在置信水平 (1- T) 下的置信域为:

$$\|f(x) - f(\hat{x})\|^2 \doteq (x - \hat{x})' B(\hat{x})' B(\hat{x})(x - \hat{x})$$
$$\leqslant t\hat{e}^2 F(t, n-t, 1-\mathrm{T}) \tag{10}$$

式中, t 为参数的个数; $\hat{e}^2 = V^{\mathrm{T}}V/(n-t)$ 为方差的估值。

置信域 (10) 式的边界可以看成是一个以 $f(\hat{x})$ 为球心,以 $R = \mathrm{d}\sqrt{F}$ 为半径的球面,其中 $\mathrm{d} = \hat{e}\sqrt{t}$。显然这个球面上任一点的曲率为:

$$K_F = 1/(\mathrm{d}\sqrt{F}) \tag{11}$$

(11) 式是线性近似后,在置信水平 (1- T) 下置信域的曲率。显然,当模型 (1) 的解轨迹 $Z = f(x)$ 上 x_0 处沿 h 方向的固有曲率 K_h^N 和参数效应曲率 K_h^T 都小于 K_F 时,说明解轨迹接近于线性。即当

$$\begin{matrix} K_h^N < K_F = 1/(\mathrm{d}\sqrt{F}) \\ K_h^T < K_F = 1/(\mathrm{d}\sqrt{F}) \end{matrix} \tag{12}$$

成立时,就认为解轨迹接近线性。

令

$$V_h^N = \mathrm{d}K_h^N, \quad V_h^T = \mathrm{d}K_h^T \tag{13}$$

(13) 式称为相对固有曲率和相对参数效应曲率,简称相对曲率。顾及 (8) 式和 (13) 式,模型 (1) 的最大相对曲率为:

$$\Gamma^N = \mathrm{d}K^N, \quad \Gamma^T = \mathrm{d}K^T \tag{14}$$

显然,当

$$\Gamma^N < 1/\sqrt{F}, \quad \Gamma^T < 1/\sqrt{F} \tag{15}$$

成立时, (12) 式一定成立,则模型 (1) 的解轨迹就接近线性,于是模型 (1) 就可以线性近似。否则将不能线性近似。因此,

$$\Gamma_{容} = 1/\sqrt{F} \tag{16}$$

可定义为模型 (1) 线性近似时的容许曲率。显然容许曲率 $\Gamma_{容}$ 也是一个相对曲率。

3 判断非线性模型能否线性近似的方法

有了线性近似的容许曲率后,就能对任一非线性模型按如下方法判断能否线性近似。

1) 首先对具体的非线性模型计算最大相对固有曲率 Γ^N 和最大相对参数效应曲率 Γ^T。

2) 根据自由度 $t, n-t$ 和置信水平 1- T 查 F 分布表,得临界值 F。并按 (16) 式计算容许曲率 $\Gamma_{容}$。

3) 检查 (15) 式是否成立。如果 (15) 式成立,即 Γ^N 和 Γ^T 都小于 $\Gamma_{容}$,则该非线性模型可以线性近似。如果 (15) 式不成立,这时可能出现 3 种情况: 第一种情况, Γ^N 和 Γ^T 都大于 $\Gamma_{容}$,说明模型的非线性强度很强,不能线性近似。第二种情况, $\Gamma^N > \Gamma_{容}$,而 $\Gamma^T < \Gamma_{容}$,说明模型的固有曲率很大,只是参数选择较好。由于固有曲率大,表明模型的非线性强度很强,所以仍不能线性近似。第三种情况, $\Gamma^N < \Gamma_{容}$, $\Gamma^T > \Gamma_{容}$,表明模型的非线性强度较弱,只要设法对参数进行变换,使得在新参数下的参数效应曲率很小,则在此新参数下可对模型进行线性近似。

4 举 例

为了估计某圆半径 R,测得其面积 $L_1 = 3.138\ \mathrm{m}^2$,周长 $L_2 = 6.285\ \mathrm{m}$,试分析函数模型

$$L_1 = \pi R^2 + \Delta_1, \quad L_2 = 2\pi R + \Delta_2$$

可否线性近似。

取 $R_0 = 1\ \mathrm{m}$,则得误差方程为:

$$v_1 = W_R + 0.003\,59$$

$$v_2 = 6.283\,2\hat{R} - 0.001\,82$$

由最小二乘原理解得:

$$\hat{R} = 1.000\,19 \text{ m}$$

用 $\hat{R}$ 可求得:

$$B = (6.284\,4 \quad 6.283\,2)^{\mathrm{T}}$$

$$C = (6,283\,2 \quad 0)^{\mathrm{T}}$$

于是由 (5) 式得:

$$\dot{Z}_h = Bh = (6.284\,4 \quad 6.283\,2)^{\mathrm{T}} h$$

$$\ddot{Z}_h = Ch^2 = (6.283\,2 \quad 0)^{\mathrm{T}} h^2$$

由 (8) 式得:

$$P_T = \begin{pmatrix} 0.500\,1 & 0.500\,0 \\ 0.500\,0 & 0.499\,9 \end{pmatrix}$$

$$P_N = \begin{pmatrix} 0.499\,9 & -0.500\,0 \\ -0.500\,0 & 0.500\,1 \end{pmatrix}$$

故由 (7) 式得:

$$K_h^T = K^T = \|P_T \ddot{Z}_h\| / \|\dot{Z}_h\|^2$$

$$= 4.443\,3h^2 / 78.972\,3h^2 = 0.056\,3$$

$$K_h^N = K^N = \|P_N \ddot{Z}_h\| / \|\dot{Z}_h\|^2$$

$$= 4.442\,5h^2 / 78.972\,3h^2 = 0.056\,3$$

因为 $t = 1$,所以 $\hat{e} = \sqrt{v_1^2 + v_2^2} = 0.003\,8$,故

$$d = \hat{e} \times \sqrt{t} = 0.003\,8$$

于是 $\Gamma^T = 0.000\,2$; $\Gamma^N = 0.000\,2$

而 $F(1, 1, 0.05) = 161$,所以

$$\Gamma_{容} = 1/\sqrt{F} = 1/\sqrt{161} = 0.0788$$

因 $\Gamma^T < \Gamma_{容}$, $\Gamma^N < \Gamma_{容}$,所以函数模型

$$L_1 = \pi R^2 + \Delta_1, \quad L_2 = 2\pi R + \Delta_2$$

可以线性近似。

本文扼要介绍了度量非线性模型的非线性强度的几种曲率,并根据线性近似的置信域,导出了非线性模型能否线性近似的数值标准—— 容许曲率 $\Gamma_{容}$。结合一个简单的算例介绍了仅含一个参数的非线性模型的几种曲率的计算和非线性随型能否线性近似的方法。对于多个参数的非线性模型,其最大相对固有曲率 Γ^N和最大相对参数效应曲率 Γ^T必须采用专门的程序进行迭代计算。

参 考 文 献

1 Beale E M L. Confidence Regions in Nonlinear Estimation. In Statist J R. Soc. B22, 1960.

2 Bates D M, Watts D G. Relative Curvature Measures of Nonlinearity. In Statist J R. Soc. B42, 1980.

3 韦博成.近代非线性回归分析.南京: 东南大学出版社, 1989.

4 Ratkowsky D H.非线性回归模型.洪再吉等译.南京: 南京大学出版社, 1986.

Acceptable Curvature of Nonlinear Model for Linear Approximation

Wang Xinzhou

(School of Geo-science and Surveying Engineering, WTUSM, 39 Luoyu Road, Wuhan, China, 430070)

Abstract In geodesy many models are nonlinear models. The classical method dealing with these nonlinear models is linear approximation using the approximate value of parameter. Due to the difference of nonlinearity between different nonliear models, some nonlinear models can be linear approximation and the others can be not. How to measure nonlinearity of nonlinear models? This paper introduces a method to measure nonlinearity of nonlinear mdoel and presents a numerical standard for linear approximation- acceptable curvature.

Key words nonlinar model; nonlinear strenghth; acceptable curvature; linear approximation