

非线性模型的非线性强度及其在变形分析中的应用

李志伟¹, 朱建军¹, 王仁谦², 李陶¹

(1. 中南大学 资源环境与建筑工程学院, 湖南 长沙 410083;

2. 华侨大学 土木系, 福建 泉州 350029)

摘要: 介绍了 Bates 和 Watts 的非线性强度度量指标——最大固有曲率和最大参数效应曲率, 提出了当模型固有非线性不显著而参数效应非线性显著时, 利用 M. J. Box 提出的偏差计算和模拟研究重新参数化的方法. 结合某实测边坡的变形监测资料, 选用非线性模型 $l_i = \alpha - \beta \exp(-\gamma \delta^2) + \epsilon$ 描述边坡变形规律. 对该边坡上 73 个点的监测资料分别进行非线性回归计算, 发现其中大部分点不能收敛于 LS 估计, 该模型有较强的非线性特性; 研究该模型的非线性性态时, 发现其最大固有非线性曲率不显著, 而模型的最大参数效应曲率非常显著. 模拟结果表明, 估计量类似对数正态分布, 因此, 用新参数 $\phi = -\lg \gamma$ 对模型重新参数化. 用重新参数化后的模型对该边坡的实测资料进行回归计算时, 只有 4 个点未能收敛于 LS 估计, 而模型的最大固有非线性曲率和最大参数效应曲率则基本符合要求.

关键词: 非线性强度; 固有曲率; 参数效应曲率; 偏差; 模拟

中图分类号: TD854.6

文献标识码: A

文章编号: 1005-9792(2001)04-0339-05

测量中的函数模型大部分是非线性模型. 处理此类非线性模型时, 通常采用参数的近似值将其线性化. 实践结果表明, 对于这些非线性模型, 有些经线性近似后能得到令人满意的结果, 而有些只有在参数近似值的近似程度非常理想的情况下才能得到可靠的结果, 有些甚至不能进行线性近似. 这是因为不同的非线性模型具有不同的非线性强度.

非线性强度的概念是 Beale 于 1960 年首先提出来的, 它用数量指标来刻画非线性模型进行线性近似时在统计推断效果的优劣程度, 即用强度作为 1 种能否线性化的度量指标. 1980 年, Bates 和 Watts 根据几何上的曲率概念发展了一种度量非线性强度的新方法^[1]. 他们认为, 模型的非线性可分成固有非线性和参数效应非线性 2 部分, 据此, 提出了非线性强度的度量指标: 最大固有曲率 Γ^N 和最大参数效应曲率 Γ^T . 最大固有曲率是 1 个不依赖于坐标选择的量, 在模型的数据集给定时就确定了, 不能通过重新参数化而改变; 最大参数效应曲率依赖于所选择的坐标, 通过重新参数化可以改变.

近年来, 人们对非线性模型能否线性近似进行

了大量的研究, 提出了容许偏差^[2], 并研究了附有限制条件的非线性模型线性近似的条件; 王新洲提出了能否线性近似的容许曲率和 1 个实用判据^[3,4]. 这些判据有的只考虑了参数效应非线性; 有的虽然考虑了模型的固有非线性和参数效应非线性, 但没有指出当模型仅具有显著参数效应非线性时如何进行重新参数化的问题. 为此, 作者在 Bates 和 Watts 曲率度量 Γ^N 和 Γ^T 的计算方法上, 提出了当模型固有非线性不显著而参数效应非线性显著时, 用 M. J. Box 提出的偏差算法的和模拟研究进行重新参数化的方法^[5], 并结合边坡变形分析的实例, 对这些方法的可行性予以验证.

1 Bates 和 Watts 非线性强度指标 Γ^N 和 Γ^T 的计算

用立体阵计算曲率^[2], 若设 $\mathbf{I}_{(n-m) \times m \times m}$ 和 $\mathbf{P}_{m \times m \times m}$ 分别为非线性模型的固有曲率立体阵和参数效应立体阵^[6,7], 且 $\mathbf{d} = (d_1, d_2, \dots, d_m)^T$, 为单位

向量, 则非线性模型在 \mathbf{d} 方向的固有曲率和参数效应曲率分别为

$$K_d^N = \|\mathbf{d}^T \mathbf{I} \mathbf{d}\|, \quad K_d^T = \|\mathbf{d}^T \mathbf{P} \mathbf{d}\|.$$

最大固有非线性曲率和最大参数效应曲率为

$$\Gamma^N = \sigma \sqrt{\rho} \cdot \max_d K_d^N, \quad (1)$$

$$\Gamma^T = \sigma \sqrt{\rho} \cdot \max_d K_d^T. \quad (2)$$

其中:

$$\sigma = \frac{\sqrt{\mathbf{V}^T \mathbf{V}}}{\sqrt{n-m}}.$$

为了计算 Γ^N 和 Γ^T , 必须采用迭代解法. 迭代收敛的条件是 \mathbf{d} 的方向和 $\nabla(K_d^N)^2$ 的方向一致. 在实际计算时, 为了简便, 其主要算法描述如下:

a. 选择 1 个初始方向 \mathbf{d}_i ;

b. 计算 $\mathbf{g}_i = \nabla(K_d^N)^2|_{\mathbf{d}_i} = 4[\mathbf{d}_i^T \mathbf{I} \mathbf{d}_i][\mathbf{I} \mathbf{d}_i]$ 及 $\tilde{\mathbf{g}}_i = \mathbf{g}_i / \|\mathbf{g}_i\|$;

c. 若 $\tilde{\mathbf{g}}_i^T \mathbf{d}_i < 1 - \varepsilon$, 取 $\mathbf{d}_{i+1} = \tilde{\mathbf{g}}_i$, 重复 b., 否则 $\Gamma^N = \sigma \sqrt{\rho} \|\mathbf{d}_i^T \mathbf{I} \mathbf{d}_i\|$.

同样, 利用上述算法计算 Γ^T , 只需用 \mathbf{P} 代替 b. 和 c. 中的 \mathbf{I} 即可.

2 模型固有非线性和参数效应非线性的显著性判别

容许曲率 $\Gamma_{容} = 1/\sqrt{F}$. 其中,

$$F = F(m, n-m, \alpha)^{[3]}.$$

模型固有非线性和参数效应非线性的显著性通过 Γ^N , Γ^T 与 $\Gamma_{容}$ 的比较来判别. 当 $\Gamma^N < \Gamma_{容}$, $\Gamma^T < \Gamma_{容}$ 时, 二者都不显著, 这时对模型的线性近似是可以接受的; 当 $\Gamma^N > \Gamma_{容}$ 时, 模型的固有非线性显著, 这时不能对模型进行线性近似; 当 $\Gamma^N < \Gamma_{容}$, $\Gamma^T > \Gamma_{容}$ 时, 模型的参数效应非线性显著, 这时可以通过重新参数化来改善模型非线性性态, 使得模型在新参数下满足 $\Gamma^T < \Gamma_{容}$, 此时对模型的线性近似可以接受.

3 重新参数化的方法

当 Γ^N 可以接受而 Γ^T 不可接受时, 必须重新参数化. 应该对哪些参数参数化以及如何重新参数化, 需通过 M. J. BOX 偏差算法和模拟研究来解决.

3.1 M. J. BOX 提出的偏差计算公式

设有非线性模型:

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

其中, \mathbf{L} 为观测向量; $\Delta = (\Delta_1, \Delta_2, \dots, \Delta_n)^T$ 为随机向量, 相互独立且均值为零, 未知方差为 σ^2 ; $\mathbf{x}_{m,1}$ 为参数向量.

M. J. Box 导出了非线性模型参数的 LS 估计的偏差计算公式, 在单个因变量的情形下, 其计算公式为^[5]

$$\text{Bias}(\hat{\mathbf{x}}) = \frac{\sigma^2}{2} \left[\sum_{u=1}^n \mathbf{B}_u \mathbf{B}_u^T \right]^{-1} \times \sum_{i=1}^n \mathbf{B}_i \text{tr} \left[\left[\sum_{u=1}^n \mathbf{B}_u \mathbf{B}_u^T \right]^{-1} \mathbf{C}_i \right]. \quad (4)$$

其中: $\mathbf{B} = (\mathbf{B}_i^k)$; $\mathbf{C} = (\mathbf{C}_{ij}^k)$; $\mathbf{B}_i^k = \frac{\partial f_k}{\partial x_i}$ 和 $\mathbf{C}_{ij}^k = \frac{\partial^2 f_k}{\partial x_i \partial x_j}$ 是模型函数的一阶和二阶偏导数; $\text{Bias}(\hat{\mathbf{x}})$ 表示参数的估计量与真参数之间的差. Bates 和 Watts 认为, 参数估计量的偏差就是参数效应立体阵各分量的对角线元素的标量平均值^[1]. 因此, 也可以采用参数效应立体阵计算偏差^[2].

参数 \hat{x}_j 的偏差可以定义为

$$\text{Bias}(\hat{x}_j) = \frac{\text{Bias}(\hat{x}_j)}{\hat{x}_j} \times 100\% \quad (5)$$

3.2 模拟研究

参数的偏差可以作为非线性强度的 1 个重要指标^[2], 并定义 1% 为容许偏差, 但这是以模型的固有非线性很小为前提的. 利用参数的偏差可了解哪些参数有较大的偏差, 哪些参数对非线性的权重多大. 找到这些参数后, 就可以通过模拟研究对其参数化. 模拟研究是通过研究 LS 估计量的抽样性质来研究模型非线性性态的 1 种方法, 它可以充分揭示参数估计量的非正态程度, 并能建议有用的新参数化. 假如参数的估计量有较小的偏差, 分布接近于正态分布, 则可以认为模型在性态上是近似线性的; 反之, 若估计量严重偏离正态分布, 则有理由认为模型在性态上远离非线性性态. 但是, 模拟研究的结果可作为重新参数化的依据, 因为只有 1 个参数的函数其 LS 估计就是参数 LS 估计的函数^[5]. 因此, 根据新参数(由重新参数化产生的参数)与原参数的函数关系和其 LS 估计量可得到新参数的 LS 估计量, 从而容易得到新参数的抽样性质. 例如, 如果已经得到原参数 θ 的 1 000 个估计量, 假设新参数 ϕ 与原参数的函

数关系为

$$\phi = \exp(\theta),$$

那么, 通过 $\phi = \exp\theta$ 可得到 1 000 个新参数的估计量 ϕ .

模拟研究一般是通过计算机产生随机数据集, 然后把随机数据集作为样本进行参数估计. 随机数据集按以下公式产生:

$$l_i = f_i(\hat{x}) + \varepsilon_i. \quad (6)$$

其中: $\varepsilon_i \sim N(0, \sigma^2); i = 1, 2, \dots, n$.

一般地, 每 1 次模拟研究最少要产生 500 个随机数据集, 最好产生 1 000 个. 假设每 1 个数据集都能收敛到 LS 估计, 则 1 000 个数据集就可以获得 1 000 个 LS 估计量, 通过研究这些估计量的抽样性质, 就能了解参数的非线性性态. 研究的方法是把估计量减去样本均值除以样本标准差进行标准化, 然后绘制标准化后的直方图, 从直方图就可以大致看出估计量的分布特征, 从而用新参数代替原参数. 根据新参数与原参数的函数关系, 得到新参数的估计量. 如果新参数的估计量基本服从正态分布, 则可按此新参数重新参数化^[7, 8].

4 非线性强度理论在边坡变形分析中的应用

4.1 变形模型的参数估计和非线性强度

某大型高边坡, 边坡高差达 300 m, 为了监测边坡的变形, 在边坡的坡顶地面上和平硐内布设了 73 个监测点. 根据该边坡上监测点的实测变形曲线和边坡变形的蠕变规律, 选择非线性回归模型中的生物增长曲线来描述边坡的变形. 参照矿山岩移中的负指数函数模型, 选择与之类似的 Weibull 模型:

$$l_t = \alpha - \beta \exp(-\gamma t^\delta) + \varepsilon_t. \quad (7)$$

其中: l_t 为观测量; $\alpha, \beta, \gamma, \delta$ 为参数; t 为观测时间; ε_t 为随机误差项.

取边坡上相邻的 4 个监测点 g_1, h_1, i_1, j_1 的 1993~1994 年的实测资料(共计观测 24 次), 采用高斯-牛顿法进行回归计算^[5], 由于对参数的近似值没有先验信息, 故必须进行迭代计算. 模型的参数计算结果如表 1 所示.

监测点 g_1, j_1 的模型参数收敛于 LS 估计, 而监

测点 h_1, i_1 的模型参数未能收敛于 LS 估计. 由此, 可以初步判断该模型与上述监测点的实测数据组成的模型数据集具有较强的非线性性态, 因为同一边坡上相邻点的变形应具有一定的相似性. 对边坡其余 69 个监测点的实测资料进行回归计算, 有 42 个点不能收敛于 LS 估计. 考察 g_1, j_1 这 2 个监测点(h_1, i_1 2 点对应的参数无法计算)模型数据集的非线性性态, 发现其最大固有非线性曲率 Γ^N 均符合要求(显著性水平取 0.05, 样本容量为 24, 参数个数为 4, 所以 $F = F(4, 20, 0.05) = 2.87$, 从而 $1/\sqrt{F} = 0.590$), 而模型的最大参数效应曲率 Γ^T 不符合要求. 用 M. J. BOX 提出的偏差计算公式所得的结果表明, 模型非线性的大部分可能包含在参数 γ 中(见表 2). g_1 点和 j_1 点回归模型的方差估值分别为: $\sigma_{g_1} = 1.89$ 和 $\sigma_{j_1} = 1.13$.

表 1 式(7)参数计算结果

点名	α	β	γ	δ
g_1	- 79.352 4	- 78.001 6	0.009 1	1.966 1
h_1	-	-	-	-
i_1	-	-	-	-
j_1	- 45.870 9	- 44.048 2	0.006 1	2.345 7

表 2 g_1, j_1 2 点模型数据集的曲率和参数的偏差

点名	Γ^N	Γ^T	偏差/%			
			Bias(α)	Bias(β)	Bias(γ)	Bias(δ)
g_1	0.263	110.34	0.93	0.87	3.66E+9	0.85
j_1	0.294	70.65	0.89	0.79	- 18.34	- 0.84

4.2 变形模型的模拟

模拟研究所需的随机数据集按以下公式产生:

$$l_t = \alpha - \beta \exp(-\gamma t^\delta) + \varepsilon_t.$$

对 g_1 点, $\varepsilon_t \sim N(0, 1.89^2)$; 对 j_1 点, $\varepsilon_t \sim N(0, 1.13^2)$. g_1 点和 j_1 点参数 $\alpha, \beta, \gamma, \delta$ 按表 1 取值. 每个数据集有 24 个数据, g_1 和 j_1 2 点各产生 1 000 个数据集.

按高斯-牛顿法进行回归计算, g_1 点和 j_1 点分别有 983 和 991 个数据集收敛到 LS 估计. 对 g_1 和 j_1 2 点的估计量进行标准化, 模拟结果的直方图见图 1 和图 2.

从图 1 和图 2 可以看出, 估计量 γ 显示出明显的正偏、右长尾的类似图象, 这种形式类似于对数正态分布. 因此, 考虑用新参数 $\phi = \lg \gamma$. 由于估计量 γ 都小于 1(见表 1), 故 $\lg \gamma < 0$, 取 $\phi = -\lg \gamma$. 由函数关

系 $\phi = -\lg \gamma$ 可以得到估计量. 其标准化后的直方图如图 3 和图 4 所示.

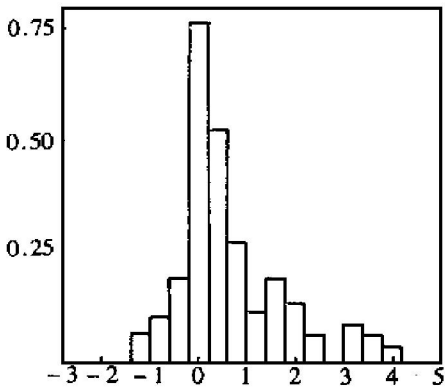


图 1 γ 模拟结果的直方图(g_1 点)

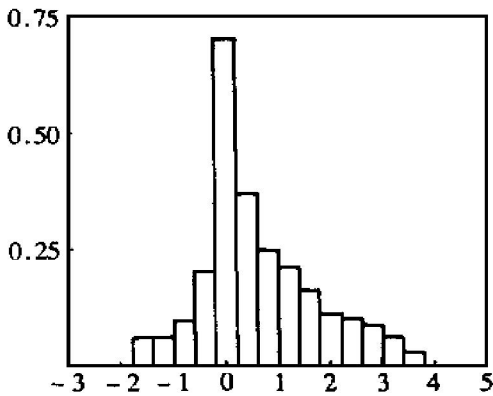


图 2 γ 模拟结果的直方图(i_1 点)

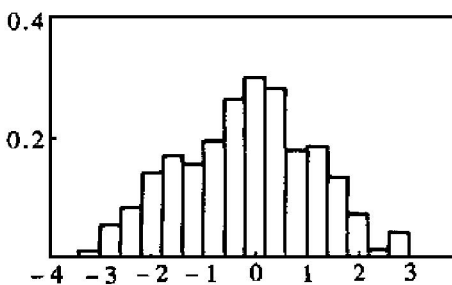


图 3 ϕ 模拟结果的直方图(g_1 点)

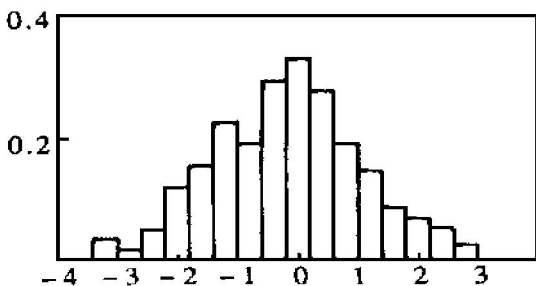


图 4 ϕ 模拟结果的直方图(i_1 点)

从图 3 和图 4 可以看出, 估计量 ϕ 无明显偏离, 基本服从正态分布. 因此, 可以预计参数化是成功

的. 由 $\phi = -\lg \gamma$ 可得: $\gamma = \exp(-\phi)$, 代入(7)式可得:

$$l_i = \alpha - \beta \exp[-\exp(-\phi) \cdot t^\delta] + \xi.$$

以 γ 代替 ϕ , 可得:

$$l_i = \alpha - \beta \exp[-\exp(-\gamma) \cdot t^\delta] + \xi. \quad (8)$$

对照式(7), 该过程相当于以 $\exp(-\gamma)$ 代替 γ , 即对 γ 重新参数化.

用式(8)对监测点 g_1, h_1, i_1, j_1 的实测资料回归, 都能收敛于 LS 估计, 其参数计算结果见表 3, 其最大非线性曲率和模型参数的偏差的比较结果见表 4.

表 3 式(8)的参数计算表

点名	参数估计值			
	α	β	γ	δ
g_1	-81.5607	-80.5348	3.974	1.9286
h_1	-48.9581	-48.9033	4.7008	2.1794
i_1	-48.0866	-48.2505	4.9612	2.2734
j_1	-47.1474	-47.2223	4.9490	2.2803

表 4 g_1, h_1, i_1, j_1 4 点构成的模型数据集的曲率和参数的百分偏差

点名	Γ^N	Γ^T	偏差/%			
			Bias(α)	Bias(β)	Bias(γ)	Bias(δ)
g_1	0.263	0.287	0.93	0.87	0.59	0.85
h_1	0.301	0.431	0.59	0.47	0.96	-0.89
i_1	0.312	0.551	0.98	-0.70	0.93	0.77
j_1	0.294	0.229	0.89	0.79	0.89	-0.84

从表 3, 4 可以看出, 经过重新参数化后, 模型参数稳定; 模型的参数效应曲率 Γ^T 有大幅度下降; 对于 g_1, h_1, i_1, j_1 4 个监测点的实测资料, 模型的最大固有非线性曲率 Γ^N 和参数效应曲率 Γ^T 均符合要求, 模型参数的偏差也都在 1% 以内. 因此, 可认为重新参数化是成功的. 对边坡其它点进行回归, 仅有 4 个点的实测资料未能收敛于 LS 估计, 因此, 式(8)适用于描述该边坡的变形规律.

参考文献:

[1] Bates D M, Wates D G. Relative curvature measures of nonlinearity [J]. J R Statist Soc B, 1980, 42: 1-25.
 [2] 王志忠, 阳可奇, 朱建军. 非线性模型的线性近似条件研究[J]. 中南工业大学学报(自然科学版), 1999, 30(3): 230-233.
 [3] 王新洲. 非线性模型线性近似的容许曲率[J]. 武汉测绘科技大学学报, 1997, 22(2): 119-121.

- [4] 王新洲. 非线性模型能否线性化的实用判据[J]. 武汉测绘科技大学学报, 1999, 24(2): 145-148.
- [5] Box M J. Bias in nonlinear estimation[J]. J R Statist Soc, 1971, 33: 171-201.
- [6] Ratkowsky D A. 非线性回归模型[M]. 洪再吉, 等译. 南京: 南京大学出版社, 1985.
- [7] 韦博成. 近代非线性回归模型[M]. 南京: 东南大学出版社, 1989.
- [8] 黄德荣. 尾矿坝变形监测[J]. 勘察科学技术, 1984, (6): 19-23.

Research on nonlinear strength of nonlinear model and its application in deformation analysis

LI Zhi-wei¹, ZHU Jian-jun¹, WANG Ren-qian², LI Tao¹

(1. College of Resources, Environment and Civil Engineering, Central South University, Changsha 410083 China;

2. Department of Civil Engineering, Huaqiao University, Quanzhou 350029, China)

Abstract: This paper introduces Bates and Watts' index of nonlinear strength, the maximal intrinsic curvature and the maximal parameter-effects curvature. It presents a re-parameterizing method which utilizes M. J. Box bias computation and simulating studies when the nonlinear model has significant parameter-effects curvature and insignificant intrinsic curvature. Combined with an example of slope deformation analysis, several points of the method are discussed. According to the mechanism of the slope deformation, the nonlinear model $l_t = \alpha - \beta \exp(-\gamma t^\delta) + \xi$ was recommended to describe the deformation. But when implementing the regression calculation for the seventy three points respectively, most of them could not converge to LS estimation. It can be inferred that the model has strong nonlinear nature. Study on the model shows that their intrinsic curvatures are acceptable, but the maximal parameter-effects curvature are not. It is shown that most of the parameter-effects curvature come from parameter γ . Simulating research reveals that estimations γ is approximately in a logarithm normal distribution, thus a new parameter $\phi = -\lg \gamma$ is used to re-parametrize the model. The regression calculation for the seventy three points respectively based on the revised model shows that only four of them can not converge to LS estimation, and both the maximal intrinsic curvature and the maximal parameter-effects curvature are acceptable.

Key words: nonlinear strength; intrinsic curvature; parameter-effects curvature; bias; simulating