

文章编号: 1673-0062(2009)04-0009-05

确定岩体力学参数先验分布的随机加权 Bayes 方法

毕忠伟^{1,2}, 德馨²

(1. 清华大学 水利水电工程系, 北京 100084 2 南华大学 城市建设学院, 湖南 衡阳 421001)

摘要: 岩体力学参数在进行 Bayes 法统计时必须利用先验分布, 但经常出现先验信息少而不能确定先验分布的情况. 为了解决小样本条件下先验分布确定的问题, 采用随机加权重采样技术, 产生岩体力学参数再生样本来模拟先验信息的总体分布, 从而获得岩体力学参数先验信息统计分布的均值与方差, 将其与当前样本分布信息代入贝叶斯公式, 从而实现了岩体力学参数后验分布的确定. 仿真算例证明, 这种方法在进行岩体力学参数估计时比经典参数估计方法有更高的精确性.

关键词: 岩体力学; 贝叶斯; 随机加权; 小样本

中图分类号: TD322+.4 **文献标识码:** A

Random Weighting Bayes Method to Determine the Prior- information Distribution of Rock mass Mechanical Parameters

BI Zhong-wei^{1,2}, DING De-xin²

(1. Department of Hydraulic Engineering Tsinghua University, Beijing 100084, China
2. School of Urban Construction, University of South China Hengyang Hunan 421001, China)

Abstract The prior distribution must be used in the statistical process of mechanical parameters of rock mass by Bayes method but there often appears the situation that the prior information is frequently too little to determine the prior distribution. In order to determine the prior distribution under small samples, the random weighting sampling technique was used to simulate the population distribution of the prior information based on the regeneration sample of mechanical parameters of rock mass, thus the average value and the variance of the prior information statistical distribution of mechanical parameters of rock mass was obtained. Substitute the obtained average value and variance and the current sample distribution information to the Bayes formula, then there comes the determination of the posterior distribution of mechanical parameters of rock mass. The simulation example proved that this method had a higher accuracy compared to the classical parameter estimation method.

收稿日期: 2009-04-21

基金项目: 国家自然科学基金资助项目(50904036); 中国博士后科学基金资助项目(20090450421)

作者简介: 毕忠伟(1976-), 男, 山西晋城人, 南华大学城市建设学院副教授, 博士, 在站博士后. 主要研究方向: 岩石力学.

when estimating the mechanical parameters of rock mass

Key words rock mechanics bayes random weighting small samples

0 引言

由于岩体是一种参数变异性很大的非均质各项异性体,因此岩体力学参数(弹性模量、单轴抗压强度、抗剪强度等)统计规律是通过其分布形式(正态分布、极值分布等)和分布参数(均值、方差)来描述的。但由于经济或技术原因,许多指标实验资料(特别是现场原位实验资料)样本常常太小,难以进行单独的统计分析。有专家提议借鉴 Bayes 方法的思路,在工程勘查阶段数据不足时,依据小样本方法计算岩体参数的随机特征;待施工过程中积累的知识足够多时,利用收集的所有资料进行岩石的分布特征统计。近年来 Bayes 统计理论在岩土工程中得到了很大的应用。张广文^[1]认为: Bayes 方法将岩体参数作为随机变量,且在先验分布基础上根据样本确定参数的后验分布,和其它估计方法相比,具有优良性和无偏性,是一种与经典方法不同的、极其有效的统计方法,在小样本下,引用该方法使推断参数概率分布的误差较小。近年来,有很多学者纷纷针对应用 Bayes 估计对岩土工程材料参数估计中的超越概率作了初步的研究,并验证了 Bayes 统计估计方法对岩石抗剪强度指标统计分析的有效性^[2-3]。

但是贝叶斯方法是基于观测分布对总体进行参数估计的一种统计方法,其前提是任一未知参数都可以看作一个由先验分布描述的随机变量,通过计算先验分布与观测样本的联合分布得到参数的后验分布。贝叶斯方法成功的关键在于对先验信息的准确把握,如果未知变量的先验分布信息不够准确,容易导致结果的偏差^[4]。而岩体工程中通常只能获得有限的或者很少的力学参数样本,这样岩体力学参数的先验分布如何确定就成为一个关键问题。本文提出了一种将随机加权法与 Bayes 估计相结合的方法,即在 Bayes 评估框架下,利用随机加权重采样技术获得岩体力学参数的先验分布,再与当前样本结合从而进行岩体力学参数估计。

1 基于随机加权的 Bayes 方法的基本原理

1.1 Bayes 方法

Bayes 统计理论观点将未知的参数看作是随

机变量,而且在获得统计试验数据前就存在一个概率分布,称之为验前分布^[5]。样本数据下的未知参数的条件分布称为验后分布,验后分布是对未知参数进行统计推断的依据。

假设在一个随机试验中,有 n 个相互排斥的尽可能的事件 A_1, A_2, \dots, A_n 。如果以 $P(A_i)$ 表示事件 A_i 发生的概率,那么 $\sum_{i=1}^n P(A_i) = 1$; 记 B 为任一事件,则有:

$$P(A_i | B) = \frac{P(B | A_i)P(A_i)}{\sum_{i=1}^n P(B | A_i)P(A_i)}, \quad i = 1, 2, \dots, n \quad (1)$$

这就是离散情况下的 Bayes 公式。其中 $\{P(A_1), P(A_2), \dots, P(A_n)\}$ 称为验前分布; B 为实验发生的事件; $P(A_i | B), i = 1, 2, \dots, n$ 为验后分布,它综合了验前信息和试验所提供的新信息。这个由验前信息到验后信息的转化,是 Bayes 统计的特征^[6]。

连续随机变量场合下的 Bayes 公式为:

$$\pi(\theta | X) = \frac{f(X | \theta)\pi(\theta)}{\int_{\Theta} f(X | \theta)\pi(\theta) d\theta} \quad (2)$$

其中 $f(X | \theta)$ 为给定参数 θ 之下的 X 的概率密度函数, $\pi(\theta)$ 为 θ 的验前密度函数, $\pi(\theta | X)$ 为 θ 在给定 X 之下的条件密度函数, Θ 为参数空间。试验之中,如果将 X 看作试验样本, $f(X | \theta)$ 就是 θ 给定后样本的密度函数,常称为样本似然函数。 $\pi(\theta)$ 反映了试验之前对 θ 的认识,而 $\pi(\theta | X)$ 则为试验之后(在获得样本 X 之后)对 θ 概率分布特性的新的认识,称 $\pi(\theta | X)$ 为验后分布密度函数。 Bayes 统计推断以 $\pi(\theta | X)$ 作为出发点。

1.2 随机加权法

当岩体力学参数的先验样本数量较少时(例如只获得 10 个左右的数据),要想获得岩体力学参数先验分布的均值和方差,用经典的统计方法往往不能令人满意,因此如何利用小子样的样本信息做出比较满意的结果就非常重要了。随机加权法是郑忠国教授在 Bootstrap 方法基础上提出的一种新的统计分析方法^[7]。它直接利用样本数据,借助于抽样技术,通过对观测数据的重新抽样产生再生样本来模拟总体分布,从而获得小样本数据统计分布的均值与方差。其主要思想如下:

假设岩体力学参数的先验信息 (例如弹性模量) 服从未知分布 F , 未知分布的均值为 μ , 方差为 σ^2 . 如果我们通过实验或现场收集到服从独立同分布的岩体力学参数样本为:

$$X = (X_1, X_2, \dots, X_n) \quad (3)$$

这些参数样本的均值与方差分别为:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

未知分布 F 的均值 μ , 方差 σ^2 与收集到的样本数据获得的均值 \bar{X} , 方差 S^2 存在一定的误差. 记为:

$$T_n^{(1)} = \bar{X} - \mu \quad (4)$$

$$T_n^{(2)} = \frac{n}{n-1} S^2 - \sigma^2$$

对于 $T_n^{(1)}, T_n^{(2)}$ 分别构造随机加权统计量:

$$D_n^{(1)} = \sum_{i=1}^n V_i X_i - \bar{X}$$

$$D_n^{(2)} = \frac{n}{n-1} \sum_{i=1}^n V_i (X_i - \bar{X})^2 - \frac{n}{n-1} S^2 \quad (5)$$

其中 $V_i (V_1, V_2, \dots, V_n)$ 为服从 $D(1, 1, \dots, 1)$ 分布的狄里克莱随机向量. 该向量可以通过如下方法产生:

设 v_1, v_2, \dots, v_{n-1} 是 $[0, 1]$ 上均匀分布的随机变量 v 的独立同分布子样, 按从小到大的次序排列, 记为: $v_{(1)} \leq v_{(2)} \leq \dots \leq v_{(n-1)}$. 同时 $v_{(0)} = 0, v_{(n)} = 1$. 这样我们就可以得到: $V_i = v_{(i)} - v_{(i-1)}, i = 1, 2, \dots, n$

根据文献 [8], 对于 $D_n^{(1)}$ 和 $D_n^{(2)}$, 存在如下统计特性:

$$E[D_n^{(1)}] = E[\sum_{i=1}^n V_i X_i - \frac{1}{n} \sum_{i=1}^n X_i]$$

$$= \frac{1}{n} \sum_{i=1}^n \mu - \mu = 0 \quad (6)$$

$$E[D_n^{(2)}] = \frac{n}{n-1} [\sum_{i=1}^n E[V_i (X_i - \bar{X})^2] - E(S^2)]$$

$$= \frac{n}{n-1} [\sum_{i=1}^n E(X_i - \bar{X})^2 - E(S^2)]$$

$$= \frac{n}{n-1} [E(S^2) - E(S^2)] = 0 \quad (7)$$

因此可以得到:

$$E[D_n^{(1)}] = E[T_n^{(1)}]$$

$$E[D_n^{(2)}] = E[T_n^{(2)}] \quad (8)$$

由于上述期望值都是随机性得到的, 从均值的观点来看 $D_n^{(i)}$ 和 $T_n^{(i)} (i = 1, 2)$, 可以用 $D_n^{(i)}$ 来

近似 $T_n^{(i)}$. 因此对于收集到的岩体力学参数样本 (即 X_i) 来说, 只需通过计算机产生 N 组 $D(1, 1, \dots, 1)$ 分布的狄里克莱随机向量, 然后计算出 N 组狄里克莱随机向量取均值的随机加权子样, 从而就可以使用点估计或者区间估计法得到小样本下的岩体力学参数先验分布特征值的估计值:

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^N [X - D_n^{(1)}(i)] = X - D_n^{(1)} \quad (9)$$

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N [\frac{n}{n-1} S^2 - D_n^{(2)}(i)]$$

$$= \frac{n}{n-1} S^2 - D_n^{(2)} \quad (10)$$

这就是所谓的随机加权法. 作为一种重采样技术, 随机加权法非常有利于确定未知分布参数的验前分布, 因此本文将其与 Bayes 方法结合, 以使在估计分布函数的参数时有更高的精度.

随机加权 Bayes 法确定岩体力学参数的其基本流程如图 1 所示.

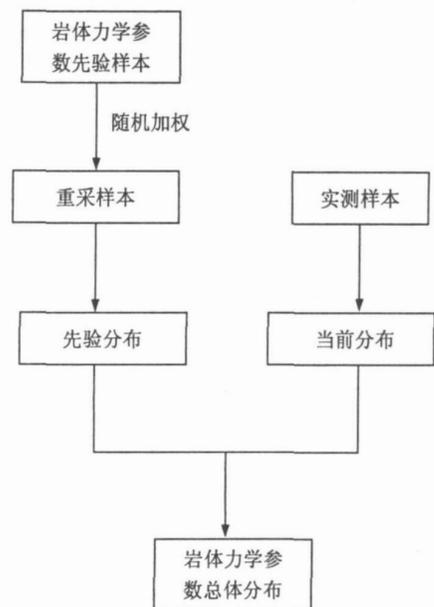


图 1 岩体力学参数随机加权 Bayes 法流程
 Fig 1 Process map of random weighting Bayes method of rock mass mechanical parameters

2 算例

岩体力学参数选择桃花江核电站粉砂岩的弹性模量. 桃花江核电站地址位于湖南省, 该处地质结构稳定, 地震烈度低, 是一个十分优越的内陆核电站址. 为了对该厂地进行合理的评估以及为施工等提供依据, 就需对该地区的岩体力学性质进行

统计. 本文收集到粉砂岩的先验弹性模量只有 6 个, 分别为 (GPa): 37. 704, 39. 268, 44. 991, 45. 030, 47. 013, 38. 742. 由于收集到的数据太少, 不能使用传统的统计方法, 因此利用上文推导出

的随机加权算法, 用 Matlab 编写了随机加权重采样的计算, 产生了 5 000 个弹性模量的新样本, 该样本分布的直方图如图 2

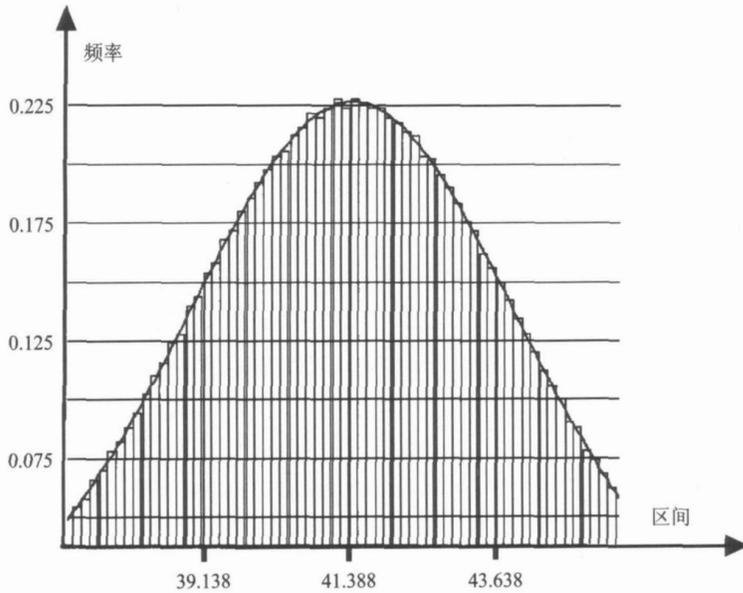


图 2 随机加权抽样直方图

Fig 2 Random weighting sample histogram

该抽样得到的弹性模量的先验分布服从正态分布, 均值为 41. 388 GPa, 方差为 2. 25

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \frac{e^{-\frac{(x-41.533)^2}{2 \times 0.614^2}}}{0.614} dx = \Phi\left(\frac{x-41.533}{0.614}\right)$$

通过实验获得 10 个弹性模量的实测数据 (GPa): 39. 010, 44. 417, 42. 758, 42. 175, 39. 858, 39. 268, 43. 397, 38. 503, 44. 143, 41. 194

为了与经验 Bayes 方法对比, 弹性模量的总体均值均取 40 GPa, 而方差取不同的值. 分别通过经典估计方法和基于随机加权的 Bayes 方法估计参数, 对比结果见表 1

通过随机加权重抽样得到弹性模量的先验分布函数, 然后与实验样本结合, 利用式 (2) 即可得到弹性模量的后验分布为:

表 1 弹性模量对比表

Table 1 Comparison table of elasticity model

| $N(\pi, \sigma^2)$ | 经典法 $N(\pi, \sigma^2)$ | 随机加权 Bayes 统计法 | | |
|--------------------|------------------------|--------------------|-------------------------|-----------------------|
| | | μ 的先验分布 | σ^2 的先验分布 | (π, σ^2) 估计值 |
| (40, 1) | (40, 181, 0.68) | $N(39.937, 1.232)$ | $N(0.74, 1.442)$ | (40, 118, 0.89) |
| (40, 3) | (40, 120, 6.13) | $N(40.015, 0.742)$ | $N(7.15, 2.012)$ | (40, 010, 7.35) |
| (40, 4) | (39.773, 11.08) | $N(39.768, 1.112)$ | $W(4.88, 15, 40, 2.30)$ | (39.832, 16.29) |

表 1 中, 符号 N 表示正态分布, W 表示三参数威布尔分布.

果则较经典方法更接近于真实值, 精度更高.

3 结论

由表 1 可以看出, 在估计样本总体均值时, 虽然经典方法可以和本文提供的方法一样取得较满意的结果, 但对于另一个参数方差, 本文的方法由于融入了较为科学的样本先验信息, 得到的计算结

本文在 Bayes 统计分析的框架下, 借助于近代计算机技术, 通过 Matlab 实现了随机加权方法, 利用重抽样技术, 对于小样本的先验分布进行

了合理的模拟, 得出了小样本条件下岩体力学参数的先验分布, 结合试验得到的岩体力学数据, 利用 Bayes 统计方法, 对岩体力学参数进行了估计. 仿真算例表明, 这种方法能较好地解决小样本条件下经典方法估计精度不高的问题.

参考文献:

- [1] 张广文, 刘令瑶. 确定随机变量概率分布参数的推广 Bayes 法 [J]. 岩土工程学报, 1995, 17(3): 91-94
- [2] 姜树海, 范子武. 应用 Bayes 方法对土石坝安全等级的动态概率评定 [J]. 水利学报, 2008, 39(8): 922-926
- [3] 李珍玉, 王永和, 丁加明, 等. 基于最大熵原理的地基变形分析 [J]. 铁道科学与工程学报, 2008, 5(2): 46-51.
- [4] 毕忠伟, 丁德馨, 饶龙, 等. 岩石力学参数 Bayes 推断的验前信息可信度研究 [J]. 水利学报, 2006, 37(8): 1000-1004
- [5] 毕忠伟. 岩体力学参数推断的 bayes 方法及截尾可靠度的研究与应用 [D]. 长沙: 中南大学, 2008
- [6] 郑桂兰, 王媛, 王飞. 广义 Bayes 法在裂隙岩体渗透系数随机反演中的应用 [J]. 水利学报, 2008, 39(4): 419-425
- [7] 郑忠国. 随机加权法 [J]. 应用数学学报, 1987, 10(2): 247-253
- [8] 张守玉, 封伟书. 基于随机加权法的装备平均维修时间验证研究 [J]. 装备指挥技术学院学报, 2009, 20(3): 100-103
- [4] 管井秀郎. 等离子体电子工程学 [M]. 张海波, 张丹, 译. 北京: 科学出版社, 2002: 119-121
- [5] 叶齐政, 孙敏. 电磁场 [M]. 武汉: 华中理工大学出版社, 2008: 229-232
- [6] 李宗谦, 余京兆, 高葆新. 微波工程基础 [M]. 北京: 清华大学出版社, 2004: 411-412
- [7] Ganachev I Sugai S Advanced large-area microwave plasmas for materials processing [J]. Surface and Coatings Technology, 2003, 174-175: 15-20

(上接第 8 页)