第2期	水 利 水 运 工 程 学 报	No. 2
2020年4月	HYDRO-SCIENCE AND ENGINEERING	Apr. 2020

DOI:10.12170/20181218001

程井,李培聪,李同春,等. 基于 MCMC 法的混凝土坝坝体坝基变形模量随机反演 [J]. 水利水运工程学报, 2020(2): 99-106. (CHENG Jing, LI Peicong, LI Tongchun, et al. Stochastic inversion of deformation modulus of concrete dam body and foundation based on Markov chain Monte Carlo method[J]. Hydro-Science and Engineering, 2020(2): 99-106. (in Chinese))

基于 MCMC 法的混凝土坝坝体坝基变形模量 随机反演

程 井1,李培聪1,李同春1,袁 平2

(1. 河海大学水利水电学院, 江苏南京 210098; 2. 中冶长天国际工程有限责任公司, 湖南长沙 410007)

摘要:针对混凝土坝材料力学参数反演中存在大量不确定性问题,提出了混凝土重力坝坝体弹性模量与坝基 变形模量的 MCMC 随机反演法。将坝体及坝基变形模量参数视为随机变量,基于 Bayesian 理论,利用无似然 函数的马尔可夫链蒙特卡罗方法(Markov chain Monte Carlo (MCMC) without likelihoods)进行随机参数后验分 布抽样。通过平稳后的马尔可夫链得到参数后验分布的随机样本,进而得到对应的期望值和标准差。以龙滩高 混凝土重力坝为例,结合典型断面的二维平面有限元模型,采用无似然函数的 MCMC 算法对坝体、坝基变形模 量进行了随机反演,得出所需反演参数(坝体弹性模量、坝基变形模量)的分布;分析了坝体、坝基变形模量分布 的统计特性与观测值波动之间的关系,得出后验分布变异性与观测值离散性呈正相关关系。

关键 词:重力坝;变形模量; MCMC法; 随机反演

中图分类号: TV642.2 文献标志码: A 文章编号: 1009-640X(2020)02-0099-08

坝体、坝基物理力学参数的选定是坝工设计和分析坝体应力、变形以及裂缝形成机理的基础。传统反 演方法主要建立在确定性分析基础上。然而,由于混凝土坝工作环境的复杂多变性,坝体变形是一个随机 变量,变形过程是一个随机过程,观测值只是一个样本序列或样本值的实现,即观测的变形具有不确定性, 只能从概率意义来研究变形的均值与方差^[1]。因此,考虑不确定性的随机反演方法更符合实际情况。

目前, 混凝土重力坝主要通过对原型观测资料分析, 建立确定性数学模型, 反求大坝的材料参数以及某些结构特性^[2]。近年来, 部分学者开始考虑水工结构反演中的不确定性问题。杨杰等^[3] 提出基于最大熵原理的贝叶斯不确定性反分析法; 苏怀智等^[4] 在传统的确定性模型中引入区间数因子, 考虑了不确定性因素对反演结果的影响; 杜永峰等^[5] 基于模糊理论提出考虑测量数据不确定性的结构物理参数识别方法; 雷鹏等^[6-7] 进一步研究了参数不确定性区间反演。贝叶斯方法是一种考虑不确定性因素的有效反演方法, 马尔可夫链蒙特卡罗方法 (MCMC) 是目前贝叶斯方法的标准抽样方法, 在岩土工程、水资源优化以及土壤学等学科已有学者进行相关研究^[8-10]。然而, 对于许多复杂问题, 似然函数难以获得, 因此, Marjoram等^[11] 首次提出了无似然函数的马尔可夫链蒙特卡罗方法, 在此基础上, Daniel等^[12] 通过调节马尔可夫链的容差, 提高该方法的计算效率, 简化了许多复杂的工程问题。

本文基于 Bayesian 理论,考虑参数先验信息分布及观测值的不确定性,采用无似然函数的马尔可夫链 蒙特卡罗(MCMC without likelihoods)方法进行弹性模量的随机反演,研究参数后验分布的统计特性,并分 别对观测值出现波动情况(人为观测误差及拟合中出现误差)与观测值进行比较,评价模型误差和预测精度。

收稿日期: 2018-12-18

基金项目: 国家自然科学基金资助项目(51009056);水利部堤防安全与病害防治工程技术研究中心开放课题基金资助项目(2018003);贵州省水利厅科技专项经费资助项目(KT201812)

作者简介:程 井(1982—),男,湖北黄石人,副教授,博士,主要从事水工结构计算力学研究。E-mail:cj042@126.com

1 基于现代贝叶斯理论的参数随机反演

贝叶斯原理中, 令 X 为一连续型分布的随机变量, 其先验概率密度函数为 $\pi(X)$, 则后验概率密度函数 $f(X \mid D)$ 为:

$$f(X|D) \propto f(D|X)(X) \tag{1}$$

式中:*X*为随机变量(待反演参数);*D*为观测样本; $f(X \mid D)$ 为随机变量的后验分布; $f(D \mid X)$ 为似然函数,数 值上等于 $D \in X$ 上的条件分布; $\pi(X)$ 为随机变量(参数 X)的先验分布。

1.1 马尔可夫链蒙特卡罗方法(MCMC)

对于复杂问题,若先验分布和后验分布不满足共轭分布条件,常规贝叶斯方法难以得到后验分布估计值(均值、标准差)的解析式,因此,需要借助于数值方法或近似方法进行模拟^[8]。为解决该问题,采用马尔可夫链蒙特卡罗方法计算后验分布。而 MCMC 方法的基本思想则是通过随机抽样建立一个最终平稳分布为所求后验分布的马尔可夫链,通过马尔可夫链得到后验分布的样本,进而得到对应的期望值和标准差。常用的 MCMC 方法有 Gibbs 抽样、Metropolis 抽样等基本算法。采用的无似然函数 MCMC 算法具体步骤^[11] 如下:

(1) 初始化马尔可夫链初始状态 X₀=x₀;

(2) 在 *i* 时刻(第 *i* 次循环), 马尔可夫链状态 X_i=x_i, 由转移核 q(X_i→X_{i+1}) 采样 X_{i+1};

(3) 根据 X_{i+1} 及已知先验计算模型 M, 产生随机样本 D';

(4) 如果 D'=D(已知抽样样本), 进入下一步, 否则从第(2)步重新开始;

(5) 计算接受率 h:

$$h = \min\left(1, \frac{\pi(X_{i+1})q(X_{i+1} \to X_i)}{\pi(X_i)q(X_i \to X_{i+1})}\right)$$
(2)

(6) 若满足接收条件,则接收 Xi+1, 否则从第 (2) 步重新开始。

抽样过程中,抽样终止以马尔可夫链收敛为前提,因此,马尔可夫链收敛与否对模型参数后验估计具有 重要影响。BGR 诊断法^[13] 是马尔可夫链收敛性诊断常用方法之一,该方法基于区间统计特性及长度进行 诊断,定义诊断指标为:

$$\hat{R}_{\text{interval}} = L/l \tag{3}$$

式中: *L* 为 马尔可夫链总的序列区间长度或统计特性; *l* 为马尔可夫链单链区间长度均值或统计特性均值。 若马尔可夫链收敛, *R*_{interval}值应接近于 1。

1.2 观测数据(样本)不确定性分析

坝体位移是一个随机变量,变形过程是一个随机过程,观测值只是一个样本序列或样本值的实现,因此 具有不确定性。根据文献 [2] 大坝位移统计模型表达式为:

$$\delta = \delta_{\rm H} + \delta_{\rm T} + \delta_{\theta} \tag{4}$$

式中: $\delta_{\rm H}, \delta_{\rm T}, \delta_{\theta}$ 分别为水压分量、温度分量、时效分量。考虑到混凝土坝工作环境的复杂多变性以及各种不确定因素(如测量误差、仪器精度、拟合误差等),假设对于某个区间内的每个值有:

$$\delta \sim N \left(\delta_{\rm H} + \delta_{\rm T} + \delta, \sigma^2 \right) \tag{5}$$

记 $\varepsilon = \delta - (\delta_H + \delta_T + \delta_\theta)$, 并假设:

$$\delta = \delta_{\rm H} + \delta_{\rm T} + \delta_{\theta} + \varepsilon, \varepsilon \sim N(0, \sigma^2) \tag{6}$$

式(6)表明,大坝位移由两部分组成,一部分是由水压分量 $\delta_{\rm H}$ 、温度分量 $\delta_{\rm T}$ 和时效分量 δ_{θ} 等组成,另一部分 $\varepsilon \sim N(0, \sigma^2)$ 为不确定性因素造成的随机误差。则对于水压分量有:

$$\delta_{\rm H} = \sum_{i=1}^{3} a_i H^i + \varepsilon_{\rm H}, \varepsilon_{\rm H} \sim N(0, \sigma_{\rm H}^{2})$$
⁽⁷⁾

式中:H为上游水位测值与坝底高程之差;ai为水压因子回归系数;EH为水压分量中不确定性因素造成的随机误差。

1.3 坝体坝基变形模量随机反演

采用 MCMC without likelihoods 算法进行参数随机反演,具体步骤:

(1)根据观测坝体位移变形数据统计回归分析^[2],得到不同水位下实测位移统计模型水压分量 **D** 以及不确定性因素造成的随机误差 ε_H~N(0,σ_H²);

(2)确定坝体、坝基变形模量 X 的先验分布 π(X);

(3)由于观测的变形具有不确定性,由统计模型分析得到的坝体位变形与实际观测值必然存在差距,因此,抽样过程中必须考虑不确定性因素造成的随机误差 ε_H~N(0,σ_H²);

(4)以后验概率密度函数 $f(X \mid D)$ 为目标函数, 采用 MCMC without likelihoods 算法产生坝体、坝基变 形模量随机样本,其中通过随机误差 $\varepsilon_{H} \sim N(0, \sigma_{H}^2)$ 筛选相应的有效样本, 2.4 节通过人为调整随机误差,进一 步研究了随机误差对坝体、坝基变形模量分布的影响;

(5)将有效的马尔可夫链样本作为后验概率密度函数 *f*(*X* | *D*)的样本, 计算后验分布 *f*(*X* | *D*)的估计 值(均值, 标准差)。

2 工程算例

2.1 工程概况

龙滩水电站前期工程于 2006 年 10 月开始下闸蓄水, 2008 年 12 月初步达到正常蓄水位。大坝坝顶高程 382 m, 建基面高程 216.43 m, 坝高 165.57 m, 坝顶宽度 14 m。以 11 号坝段为典型坝段, 根据实测坝体水平位移反演坝体、坝基变形模量。如图 1 所示, 二维有限元模型范围为:向上、下游各延伸 250 m、坝基建基面以下延伸 235 m。有限元节点布置时, 尽可能考虑将位移测点安排在单元节点上, 采用四节点单元, 共2 383 个节点, 2 244 个单元。



图 1 11 号坝段有限元模型及垂线测点 Fig. 1 FEM mesh and vertical monitoring points of 11# dam section

参数反演时,选取正垂线测点 PL11-1(高程 342.00~379.00 m), PL11-2-2(高程 270.00~342.00 m)与 PL11-3(高程 222.75~270.00 m)2010 年 5 月 28 日至 2011 年 8 月 15 日的实测顺河向位移进行分析,依据式 (4)分离坝体位移水压、温度及时变等主要分量,各垂线的复相关系数 R 分别为 0.964, 0.951 和 0.977,反演 分析中输入不同水位下实测位移统计模型水压分量 D,即对模型依次施加水位为 335, 340, 345, 355 和 360 m 的上游静水压力,结果如表 1 所示。

表1

不同水位下实测位移统计模型水压分量 D

Tab. 1 Water pressure component D given by statistical model at different water levels						
上游北台(顺河向位移/mm					
上研水山/m	高程222.75 m	高程270.00 m	高程342.00 m			
335	2.67	4.85	5.79			
340	2.86	5.32	6.61			
345	3.06	5.79	7.43			
355	3.45	6.73	9.08			
360	3.65	7.20	9.90			

观测值只是一个样本序列或样本值的实现,也就是说观测的变形具有不确定性。因此,观测值必须考虑各种不确定因素。由图 2(a) 所示, 拟合值与实测值存在误差, 误差来源于上述各种不确定因素。根据 1.2 节所述引入水压分量的随机误差 $\varepsilon_{H} \sim N(0, \sigma_{H}^{2})$, 图 2(b), (c) 和 (d) 分别为 222.75, 270.00 和 342.00 m 高程 拟合总体位移及其分量。



图 2 龙滩观测位移拟合值与实测值

Fig. 2 Fitted values and measured values of observation displacements of Longtan Dam

2.2 单参数反演分析

假定坝基变形模量为确定量,坝体弹性模量为待反演随机变量。采用无似然函数的 MCMC 算法来反 演单个参数,初步认为坝体弹性模量服从区间为 [10,60] 的均匀分布,模拟时马氏链样本数取 20 000 个。

图 3 为无似然函数 MCMC 抽样样本。图 4 为正态分布 Q-Q 检验图,其中横坐标为分位数,纵坐标为 样本值;图中点近似为一条直线,说明样本基本服从正态分布。图 5 为 BGR 诊断法马氏链收敛性诊断图, 通过比较可以看出样本数量达到 7 500 时,诊断指标稳定在 1 附近,所以在以下分析中有效链长度取为 10 000~20 000。图 6 比较了参数的先验分布、后验分布概率密度曲线及其正态拟合曲线,结果表明:(1)坝 体弹性模量均值由先验分布的 45 GPa 变为后验分布均值 32.09 GPa, 基本符合龙滩重力坝实际坝体弹性模量;(2)弹性模量后验分布区间比先验分布明显收窄, 坝体弹性模量反演结果的标准差为 3.81 GPa, 变异系数为 0.119。





2.3 多参数反演分析

假定坝体弹性模量及坝基变形模量均为待反演随机变量,采用无似然函数的 MCMC 算法反演多个参数。反演时,初步认为两参数均服从区间为 [10,60] 的均匀分布,模拟时马氏链样本数取 20 000 个。

图 7 为坝体弹性模量、坝基变形模量 MCMC 抽样样本,其 Q-Q 检验图表明两个变量的马氏链样本均 服从正态分布。图 8 比较了两个参数的先验分布、后验分布概率密度曲线及其正态拟合曲线,结果表明: (1)坝体弹性模量均值由先验分布的 45 GPa 变为后验分布均值 33.49 GPa,坝基变形模量的均值由先验分布 35 GPa 变为后验分布均值 29.53 GPa,基本符合龙滩重力坝实际情况;(2)坝体弹性模量、坝基变形模量标准 差分别为 3.78 和 1.90 GPa,变异系数分别为 0.11 和 0.06,标准差与变异系数均明显缩小,说明参数的不确 定性显著减小。







Fig. 8 Probability density curves of posterior distribution

采用经典的最小二乘法反演坝体、坝基变形模量,与本文提出的反演方法进行对比,在同一坝段,同样观测数据情况下,两种反演方法的结果比较如表2 所示。结果表明,两种算法都能较好地给出参数的 反演结果,最小二乘法尽管误差平方和更小,但缺少 对不确定因素的考虑,只给出了一个确定值的结果;

表 2 两种反演方法结果						
Tab. 2 Comparison between two inversion methods						
计算方法	坝体弹性模量/GPa	坝基变形模量/GPa	误差平方和/mm ²			
MCMC	33.49	29.53	1.71			
最小二乘法	35.60	28.53	1.48			

而 MCMC 算法充分考虑了不确定因素的影响,反演结果能更好地反映坝体、坝基变形模量的分布规律。

2.4 坝体坝基变形模量敏感性分析

位移的不确定性决定了坝体弹性模量及坝基变形模量反演结果的不确定性。通过模拟不同的 σ_H进行 坝体弹性模量及坝基变形模量的敏感性分析。正垂线 PL11-1(高程 342.00 m)与 PL11-3(高程 222.75 m)测 点实测样本标准差分别为 σ_H= 0.12 mm 和 σ_H= 0.24 mm。图 9 为不同的观测样本标准差(随机误差)与坝体 弹性模量、坝基变形模量标准差的关系,易发现随着观测样本随机性(波动)的增大,反演的坝体、坝基变形 模量变异系数相应增大。





Fig. 9 Relationships between standard deviations of observation samples and variation coefficients of elastic modulus of dam body and deformation modulus of dam foundation

3 结 语

针对混凝土坝材料力学参数反演中存在大量不确定性的问题,提出了高混凝土重力坝坝体弹性模量与 坝基变形模量的 MCMC 随机反演方法,给出了无似然函数的马尔可夫链蒙特卡罗法在工程中的实际应用, 并研究了坝体变形不确定性对反演结果的影响。研究结果表明:(1)该随机反演方法可以得出所需反演参 数(坝体弹性模量、坝基变形模量)的分布;(2)反演参数的后验分布变异性较先验分布有明显降低;(3)在考 虑变形观测值离散性变化情况下,后验分布变异性与观测值离散性呈正相关关系。

MCMC 随机反演方法计算简便、高效,具有很强的工程实用性,为混凝土坝参数随机反演提供了重要 计算方法。但有时为了得到较为精确的参数尾部分布,需要的抽样数较大,这就导致有限元计算的总耗时 巨大,因此需要结合有限元原理对该算法进行优化。

参考文献:

- [1] 黄宏伟, 孙钧. 基于Bayesian广义参数反分析[J]. 岩石力学与工程学报, 1994, 13(3): 219-229. (HUANG Hongwei, SUN Jun. Generalized parameters back analysis method based on Bayesian theory[J]. Chinese Journal of Rock Mechanics and Engineering, 1994, 13(3): 219-229. (in Chinese))
- [2] 顾冲时, 吴中如. 大坝与坝基安全监控理论和方法及其应用[M]. 南京: 河海大学出版社, 2006. (GU Chongshi, WU Zhongru. Safety monitoring of dams and dam foundations: theories & methods and their application[M]. Nanjing: Hehai University Press, 2006. (in Chinese))
- [3] 杨杰, 胡德秀, 吴中如. 基于最大熵原理的贝叶斯不确定性反分析方法[J]. 浙江大学学报 (工学版), 2006, 40(5): 810-815, 835. (YANG Jie, HU Dexiu, WU Zhongru. Bayesian uncertainty inverse analysis method based on pome[J]. Journal of Zhejiang University (Engineering Science), 2006, 40(5): 810-815, 835. (in Chinese))
- [4] 苏怀智, 雷鹏, 顾冲时, 等. 混凝土坝材料参数区间反演分析方法[J]. 河海大学学报(自然科学版), 2008, 36(5): 654-658.
 (SU Huaizhi, LEI Peng, GU Chongshi, et al. Interval back analysis for mechanical parameters of a concrete dam[J]. Journal of Hohai University (Natural Sciences), 2008, 36(5): 654-658. (in Chinese))
- [5] 杜永峰, 李万润, 李慧. 基于测量数据不确定性的结构参数识别[J]. 振动、测试与诊断, 2012, 32(4): 629-633. (DU Yongfeng, LI Wanrun, LI Hui. Structural parameters identification based on uncertainty of measurement data[J]. Journal of Vibration, Measurement & Diagnosis, 2012, 32(4): 629-633. (in Chinese))
- [6] 雷鹏, 苏怀智, 张贵金. 基于RNN模型的坝体和岩基区间参数反演方法研究[J]. 岩土力学, 2011, 32(2): 547-552. (LEI Peng, SU Huaizhi, ZHANG Guijin. Study of interval parameters back analysis of dam body and rock foundation based on RNN model[J]. Rock and Soil Mechanics, 2011, 32(2): 547-552. (in Chinese))
- [7] 方圣恩, 张秋虎, 林友勤, 等. 不确定性参数识别的区间响应面模型修正方法[J]. 振动工程学报, 2015, 28(1): 73-81.
 (FANG Sheng'en, ZHANG Qiuhu, LIN Youqin, et al. Uncertain parameter identification using interval response surface model updating[J]. Journal of Vibration Engineering, 2015, 28(1): 73-81. (in Chinese))
- [8] 左自波,张璐璐,程演,等. 基于MCMC法的非饱和土渗流参数随机反分析[J]. 岩土力学, 2013, 34(8): 2393-2400. (ZUO Zibo, ZHANG Lulu, CHENG Yan, et al. Probabilistic back analysis of unsaturated soil seepage parameters based on Markov chain Monte Carlo method[J]. Rock and Soil Mechanics, 2013, 34(8): 2393-2400. (in Chinese))
- [9] 王丽萍, 王渤权, 李传刚, 等. 基于贝叶斯统计与MCMC思想的水库随机优化调度研究[J]. 水利学报, 2016, 47(9): 1143-1152. (WANG Liping, WANG Boquan, LI Chuangang, et al. Reservoir stochastic optimization scheduling research based on Bayesian statistics and MCMC[J]. Journal of Hydraulic Engineering, 2016, 47(9): 1143-1152. (in Chinese))
- [10] MA C F, LI X, NOTARNICOLA C, et al. Uncertainty quantification of soil moisture estimations based on a Bayesian probabilistic inversion[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(6): 3194-3207.
- [11] MARJORAM P, MOLITOR J, PLAGNOL V, et al. Markov Chain Monte Carlo without likelihoods[J]. Proceedings of the National Academy of Sciences of the United States of America, 2003, 100(26): 15324-15328.
- [12] WEGMANN D, LEUENBERGER C, EXCOFFIER L. Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood[J]. Genetics, 2009, 182(4): 1207-1218.
- [13] BROOKS S P, GELMAN A. General methods for monitoring convergence of iterative simulations[J]. Journal of Computational and Graphical Statistics, 1998, 7(4): 434-455.

Stochastic inversion of deformation moduli of concrete dam body and foundation based on Markov chain Monte Carlo method

CHENG Jing¹, LI Peicong¹, LI Tongchun¹, YUAN Ping²

(1. College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China; 2. Zhongye Changtian International Engineering Co., Ltd., Changsha 410007, China)

Abstract: In view of the large number of uncertainties in the inversion of mechanical parameters of concrete dam materials, the Markov chain Monte Carlo (MCMC) stochastic inversion method for elastic modulus of the dam body and deformation modulus of the dam foundation of the high concrete gravity dam is proposed. The deformation moduli of the dam body and dam foundation are considered as random variables. Based on the Bayesian theory, the Markov chain Monte Carlo without likelihoods method is used to sample the posterior distribution of the random parameters. The stationary Markov chain is used to obtain the random samples of the posterior distribution of parameters, and then the corresponding expected values and standard deviation are obtained. Taking the Longtan high concrete gravity dam as an example, based on the two-dimensional plane model for typical dam section, the deformation moduli of the dam body and foundation are inversed by the proposed method, and the distribution of the required inversion parameters (elastic modulus of the dam body and deformation modulus of the dam foundation is obtained. The relationships between the statistical characteristics of the deformation modulus distribution of the dam body and foundation and the fluctuation of the observed values are analyzed. It is concluded that the variability of the posterior distribution is positively correlated with the discreteness of the observed values.

Key words: gravity dam; deformation modulus; Markov chain Monte Carlo method; stochastic inversion