

文章编号:1000-0887(2000)04-0402-07

动力模型参数识别中的 Bayes 方法^{*}

李 书¹, 卓家寿², 任青文²

(1. 北京航空航天大学 飞机设计所,北京 100083; 2. 河海大学 土木学院,南京 210098)

(叶庆凯推荐)

摘要: 将统计分析中的 Bayes 方法应用到参数识别问题中,提出了利用测量频率的 Bayes 估计识别动力学模型的方法,该方法是基于广义逆特征值问题的解。考虑到试验数据的随机性,测量频率用正态分布来描述。此外,将工程师关于测量频率的置信度进行量化,并且与识别过程相结合。数值算例证明了这一方法的有效性。

关键词: 参数识别; 动力学模型; Bayes 估计; 特征值反问题; 先验分布; 后验分布
中图分类号: O327;O241.6 **文献标识码:** A

引 言

物理现象数学模型中的常见问题是用观测的数据来对模型参数进行估计。近年来,参数识别方法在试验和理论分析之间架起了桥梁。在实施动力学模型参数识别之前,需要确保观测到的模态和频率是准确的。但是,在振动试验中,很难避免测量误差,很多情况都会影响测量的结果。由振动试验来确定结构的自然频率和振型可能包括多个误差源,如仪器设备本身的误差、噪声过大、故障、不正确的截断位置以及不适当的方法等。可见,任何观测过程都应考虑测量时的不确定性。因此,如何处理试验数据的随机特性引起了广泛的关注^[1,2],参数识别问题已有了许多的研究成果,大多数的文献仅考虑了试验数据的随机性而忽视了工程师的经验。为了克服这个缺点,我们采用著名的 Bayesian 方法得到试验的谱数据(即自然频率)的 Bayes 估计,用 Bayes 估计值来进行动力模型的参数识别。我们的目的是降低不确定因素对参数识别结果的影响。

Bayesian 方法是将样本信息和其它可以利用的确切的先验信息结合起来的一种统计估计方法。换句话说,一位工程师能否完成某件事的概率,是由他自己的经验和知识决定的,本文运用 Bayesian 观点来获得参数识别的数值解。在众多的 Bayes 估计研究中,一般的做法先是假定先验分布函数^[3],然后用 Bayes 规则来获得后验概率密度函数的描述,通常假定估计值的条件概率密度函数是高斯分布,主要由于两方面原因:一是有许多特殊的数学性质;二是高斯分布与很多物理随机现象的概率规律很相似。

本文方法将动力学试验与 Bayesian 理论相结合,把试验值的精度和对原模型值的置信度

* 收稿日期: 1998-12-01; 修订日期: 1999-12-18

作者简介: 李书(1965~),男,江苏镇江人,副教授,博士;卓家寿(1938~),男,福建福州人,教授,博士生导师;任青文(1942~),男,浙江宁波人,教授,博士生导师,科研处长。

都考虑到参数识别过程中。Bayes 方法给我们的研究提供了一个有用的工具。

1 Bayes 理论

本文是用连续的正态分布来近似离散分布。首先确定随机变量的统计特征。设 X_1, X_2, \dots, X_n 是独立相同分布(i.i.d.)的随机变量,其概率分布函数为正态分布,形如 $N(\mu, \delta)$, 其中 μ 和 δ 为分布参数,同时也分别为均值和标准方差。分布参数也被认为是具有同样分布函数的随机变量,并且作为先验分布 $f(\theta)$ 。为了简单,考虑未知函数为一实数 $\theta (-\infty < \theta < \infty)$ 。然后观测数据的均值 \bar{X} 能够通过 Bayes 理论与 θ 的先验分布结合得到一个修正的后验分布 $f(\theta | \bar{X})$ 。

假定测量数据的样本均值 \bar{X} 是一个随机变量具有正态分布 $N(\mu, \delta)$, 则其密度函数为

$$f(\bar{X}_n | \theta) = \frac{1}{\sqrt{2\pi}\delta} \exp\left[-\frac{1}{2}\left(\frac{\bar{X} - \theta}{\delta}\right)^2\right]. \quad (1)$$

假设 θ 的先验正态分布 $N(\lambda, \sigma)$, 其表达式为

$$f(\theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{\theta - \lambda}{\sigma}\right)^2}, \quad (2)$$

则, \bar{X} 和 θ 的联合密度函数为

$$f(\bar{X}, \theta) = f(\bar{X} | \theta) \cdot f(\theta) = \frac{1}{2\pi\delta\sigma} \exp\left[\left(-\frac{1}{2}\right)\left\{\left(\frac{\bar{X} - \theta}{\delta}\right)^2 + \left(\frac{\theta - \lambda}{\sigma}\right)^2\right\}\right]. \quad (3)$$

通过对 μ 的加、减,很容易得到

$$(\bar{X} - \theta)^2 = (\mu - \theta)^2. \quad (4)$$

将(4)代入(3), (3)式变为

$$f(\bar{X}, \theta) = \exp\left[\left(-\frac{1}{2}\right)\left\{\left(\theta - \frac{\sigma^2\mu + \delta^2\lambda}{\sigma^2 + \delta^2}\right) / \sqrt{\frac{\sigma^2\delta^2}{\sigma^2 + \delta^2}}\right\}^2\right]. \quad (5)$$

Bayes 的预测分布(predictive distribution)是观测随机变量的边缘分布,为

$$g(\bar{X}) = \int_{-\infty}^{\infty} f(\bar{X}, \theta) d\theta = \sqrt{2\pi} \sqrt{\frac{\sigma^2\delta^2}{\sigma^2 + \delta^2}}. \quad (6)$$

θ 在给出 \bar{X} 时的后验概率分布为

$$f(\theta | \bar{X}) = \frac{f(\bar{X}, \theta)}{g(\bar{X})} = \frac{1}{\sqrt{2\pi} \sqrt{\frac{\sigma^2\delta^2}{\sigma^2 + \delta^2}}} \exp\left[\left(-\frac{1}{2}\right) \times \left\{\left(\theta - \frac{\sigma^2\mu + \delta^2\lambda}{\sigma^2 + \delta^2}\right) / \sqrt{\frac{\sigma^2\delta^2}{\sigma^2 + \delta^2}}\right\}^2\right],$$

其中均值为 $(\sigma^2\mu + \delta^2\lambda)/(\sigma^2 + \delta^2)$, 标准方差为 $\sqrt{\sigma^2\delta^2/(\sigma^2 + \delta^2)}$ 。因此可以表明 θ 的 Bayes 估计是后验的期望^[3], 即

$$\hat{\theta}(\bar{X}) = \frac{\sigma^2\mu + \delta^2\lambda}{\sigma^2 + \delta^2}. \quad (7)$$

在前面的表达式中,假设先验分布和后验分布都是连续的, θ 的后验分布描述了基本随机变量 \bar{X} 的分布函数中参数的统计不确定性。 \bar{X} 的基本分布反映了随机变量固有的不确定性, 并且和参数 θ 的不确定性结合起来, 得到 \bar{X} 修正的概率分布。方程(7)表明先验估计被修正, 这意味着修正的估计依赖于先验的信息和观测。

2 参数识别模型的描述

2.1 广义逆特征值问题

为求解参数识别问题,将频率和振型的物理变化与结构参数变化直接联系起来,归结为一个广义逆特征值问题. 结构动力特征方程为

$$K\Phi = M\Phi\Omega, \quad (8)$$

其中 K 为结构刚度矩阵, M 为质量矩阵, $\Phi = (\phi_1, \phi_2, \dots, \phi_m)$ 为振型矩阵, $\Omega = \text{diag}(\omega_1^2, \omega_2^2, \dots, \omega_m^2)$ 是频率矩阵, 为对角矩阵.

$$\Phi^T M \Phi = I, \quad \Phi^T K \Phi = \Omega. \quad (9)$$

对大型结构而言,要测量所有的频率和振型是不现实的,因此

$$\Phi_r^T M \Phi_r = I_r, \quad \Phi_r^T K \Phi_r = \Omega_r, \quad (10)$$

其中 $\Phi_r = (\phi_1, \phi_2, \dots, \phi_r)$ 和 $\Omega_r = \text{diag}(\omega_1^2, \omega_2^2, \dots, \omega_r^2)$ 分别是测量的振型和频率矩阵,这些数据被直接用来识别结构参数,这个识别过程需要建立每个单元的刚度矩阵 K_i 和未知材料的关系. 考虑一结构有 n 个单元和 m 个位移自由度,总体刚度矩阵是单个刚度矩阵的和,对质量矩阵同样处理. 令 $p = (p_1, p_2, \dots, p_n)$ 是设计参数的集合,例如代表单元的截面积. 最后组装的刚度矩阵和质量矩阵可写为如下形式

$$K(p) = \sum_{i=1}^n K_i(p), \quad M(p) = \sum_{i=1}^n M_i(p). \quad (11)$$

这里的问题就是确定参数,使得相应的特征值方程具有给定的特征值. 在 Friedland 等人^[4]的文献里系统完整地含参数逆特征值问题作了研究,基于文献[4]的理论和方法,我们将动力模型参数识别问题视为一类广义逆特征值问题,通过假定特征值是参数的函数,利用频率的理论值和测量值应当相等,建立起一组非线性方程组,用 Newton-Raphson 方法解这个方程组. 因此,参数识别问题本质上等价于解下面非线性方程组

$$\omega_i(p) - \omega_i^* = 0 \quad i = 1, \dots, r. \quad (12)$$

如果我们设

$$g_i(p) = \omega_i(p) - \omega_i^* \quad i = 1, \dots, r, \quad (13)$$

其中 $g(p) = (g_1(p), g_2(p), \dots, g_r(p))^T$, ω_i^* 是测量的频率,设 $p = (p_1, p_2, \dots, p_n)^T$ 是含 n 个参数的矢量, p^* 是问题的解. 为了应用 Newton 方法解方程(12),我们需要计算 $g'(p^{(k)})$, 其中 $g'(p) = (g'_1(p), g'_2(p), \dots, g'_r(p))^T$.

对(13)关于 p 求导,其中令 $\lambda_i(p) = \omega_i^2(p)$,按照文献[5],有

$$\frac{\partial \lambda_i(p)}{\partial p_j} = \phi_i^T \left(\frac{\partial K(p)}{\partial p_j} - \lambda_i \frac{\partial M(p)}{\partial p_j} \right) \phi_i. \quad (14)$$

我们用 Newton-Raphson 方法,通过解下面线性方程组,

$$g(p^{(k)}) \Delta p^{(k)} = g(p^{(k)}), \quad (15)$$

并且方程(15)是线性方程组,其中

$$\mathbf{g}'(\mathbf{p}) = \begin{bmatrix} \frac{\partial g_1}{\partial p_1} & \frac{\partial g_1}{\partial p_2} & \cdots & \frac{\partial g_1}{\partial p_n} \\ \frac{\partial g_2}{\partial p_1} & \frac{\partial g_2}{\partial p_2} & \cdots & \frac{\partial g_2}{\partial p_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial g_n}{\partial p_1} & \frac{\partial g_n}{\partial p_2} & \cdots & \frac{\partial g_n}{\partial p_n} \end{bmatrix}.$$

假定 p^k 是第 k 步逼近 p^* , 令 $\lambda(p^k)$ 是第 k 步迭代 $\mathbf{p} = p^k$ 时的特征值, 那么第 $k+1$ 步的逼近结果由下式给出

$$p^{(k+1)} = p^{(k)} - \Delta p^{(k)}, \quad (16)$$

最终, 系数将被确定. 解方程(15)时, 一般情况 $r \neq n$, 这时我们用 Moore-Penrose 广义逆来求解. 然后下一步对解 p^* 的逼近解 p^{k+1} 可以从解方程(16)中获得.

文献[6]中的 Kantorovitch 收敛定理指出, 初值选取的好坏是牛顿迭代法成功的关键. 通常要使选取的初值满足 Kantorovitch 收敛定理的假设是不现实的. 实际上的初值选取是在有一定经验的猜测基础上进行的, 并且在计算过程中自动调整. 在参数识别中, 初值向量的元素代表结构的参数, 其取值应当是正数. 另外, 如果能利用一些模型的信息, 对初值选取比较有利. 本文算例中初值取不同正数, 收敛效果都很好.

2.2 重特征值导数

无论在文献[7]中, 还是我们的计算经验都表明将两个很接近的特征值当作重特征值处理效果较好. 当重特征值出现时, 特征值一般在 p^* 点不可微, 而且特征向量 ϕ_i 不唯一, 一般不能在 p^* 点定义为参数 \mathbf{p} 的连续函数. 假定 $\lambda_i = \omega_i^2$ 表示方程(8) 在 $\mathbf{p} = p^*$ 点的 t 个重特征值, 相应的特征向量为 $\phi_i, \phi_{i+1}, \dots, \phi_{i+t-1}$, 令

$$\Phi_i = (\phi_i, \phi_{i+1}, \dots, \phi_{i+t-1}) \in R^{n \times t},$$

设

$$D_j(\lambda_i) = \left(\frac{\partial \mathbf{K}(\mathbf{p})}{\partial p_j} \right)_{\mathbf{p}=p^*} - \lambda_i \left(\frac{\partial \mathbf{M}(\mathbf{p})}{\partial p_j} \right)_{\mathbf{p}=p^*}, \quad (17)$$

根据孙继广的工作[7],

$$d_{p_j}(\lambda_i) = \rho(\Phi_i^T D_j(\lambda_i) \Phi_i) \quad (18)$$

定义为重特征值关于 p_i 的导数, 其中 $\rho(\cdot)$ 为谱半径.

2.3 基本步骤

广义逆特征值问题的计算过程总结如下:

- 1) 选择初始逼近解 p^0 , 令迭代指标 $k = 0$;
- 2) 利用(11)在 $\mathbf{p} = p^k$ 点形成刚度矩阵 \mathbf{K} 和质量矩阵 \mathbf{M} ;
- 3) 计算 $\mathbf{K}\Phi = \mathbf{M}\Phi\Omega$ 的最小 r 阶特征值 $\lambda_i(p^k) = \omega_i^2(p^k)$, $i = 1, \dots, r$, 和关于质量矩阵 \mathbf{M} 正交的特征向量 $\phi_i(p^k)$, $i = 1, \dots, r$;
- 4) 检查收敛性, 如果 $\|\omega(p^k) - \omega^*\|$ 足够小, 则停止计算;
- 5) 利用(14)或(18)计算 $\mathbf{g}'(p^k)$;
- 6) 用(16)计算逼近解;

7) 令 $k = k + 1$ 并重复 2) 以后的步骤。

3 数值算例

数值算例是一个简单的大跨度桥, 图 1 是该桥的模型, 分成 4 个单元, 桥长 $L = 400$ m, 横截面积 $A = 10$ m², 关于 y 和 z 轴面积矩分别为 $I_y = I_z = 1.0 \times 10^4$ m⁴。

假定我们已知桥的前 5 阶频率的先验正态分布, 在试验中又测得桥的前 5 阶频率, 样本的均值和方差很容易获得, 利用(7)求后验分布。参数的分布列入表 1, 其中第 1 和第 2 阶频率的分布分别见图 2 和图 3。

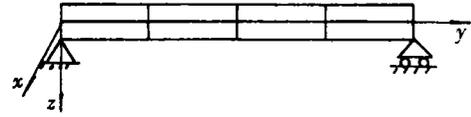


图 1 桥的模型

然后我们用后验分布中提供的均值来识别桥梁模型的参数。本文讨论的问题是识别弹性模量 E 和质量密度 ρ , 测量数据为频率。根据图 1 模型, 式(11)可写为如下形式, 其中矢量 p 包含 E 和 ρ 作为其元素

表 1 不同频率分布下的频率均值与方差 Hz

先验分布		边缘分布		后验分布	
均值 λ	方差 σ	均值 μ	方差 δ	均值 $\frac{\sigma^2 \mu + \delta^2 \lambda}{\sigma^2 + \delta^2}$	方差 $\sqrt{\frac{\sigma^2 \delta^2}{\sigma^2 + \delta^2}}$
1.92	0.15	2.11	0.04	2.09	0.037
12.85	0.13	13.01	0.17	13.00	0.031
12.94	0.24	13.11	0.07	13.09	0.072
36.31	0.28	36.63	0.04	36.62	0.038
36.39	0.24	36.67	0.04	36.64	0.049

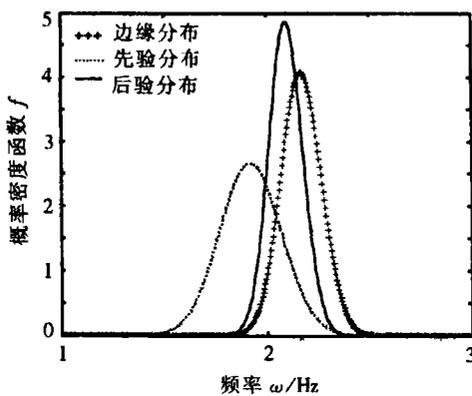


图 2 1 阶频率分布

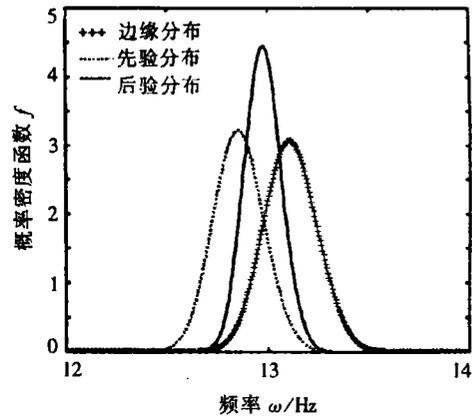


图 3 2 阶频率分布

$$K(p) = \sum_{i=1}^n p_i K_i, \quad M(p) = \sum_{i=1}^n p_i M_i. \quad (19)$$

图 2 表明后验概率密度分布曲线比先验概率密度分布曲线和预测概率密度分布曲线更“尖”,这意味着它包含更多的信息^[8],因此和 Bayes 思想是一致的。基于频率后验信息分布识别得到的结果见表 2。

表 2 不同频率分布下识别的参数比较

先验分布		边缘分布		后验分布	
模量 $E/N/m^2$	密度 $\rho/kg/m^3$	模量 $E/N/m^2$	密度 $\rho/kg/m^3$	模量 $E/N/m^2$	密度 $\rho/kg/m^3$
$6.899\ 0 \times 10^{11}$	2 029.1	$7.016\ 5 \times 10^{11}$	1 995.3	$7.012\ 8 \times 10^{11}$	1 996.4

该算例说明了 Bayesian 原理实际上就是在分析中,对已有的信息(不论是先验得到的还是观测得到的)进行加权平均。

4 结 论

本文研究了结构动力模型的参数识别问题。测量误差具有随机性,为探讨随机性对参数识别的影响,基于 Bayes 理论提出了改善识别效果的方法。算例表明本文方法是有效的,数学概念也比较简单易于接受,因此适合于实际问题,该方法在结构分析和试验之间架设一座桥梁。

算例同时也证明了统计识别方法为模型参数识别提供了一个有效的工具,结果清楚地表明我们的方法优于没有同时考虑试验数据随机特性和工程师经验的方法,另外,本文方法可以用于某些有工程意义的问题。如利用振动试验来确定结构的有限元模型。

[参 考 文 献]

- [1] Benjamin J R, Cornell C A. *Probability Statistics Decision for Civil Engineers*[M]. New York: McGraw-Hill, 1970.
- [2] Collins J D, Hart G C, Hasselman T K, et al. Statistical identification of structures[J]. *AIAA J*, 1974, 12(2): 185 ~ 190.
- [3] Zacks S. *Parametric Statistical Inference*[M]. Oxford: Pergamon Press Ltd, 1981.
- [4] Friedland S, Nocedal J, Overton M L. The formulation and analysis numerical methods for inverse eigenvalue problems[J]. *SIAM J Number Anal*, 1987, 24(3): 634 ~ 667.
- [5] Ojalvo I U. Efficient computation of mode shape derivatives for large dynamic systems[J]. *AIAA J*, 1987, 25(10): 1384 ~ 1390.
- [6] Dennis J E. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*[M]. Englewood Cliffs NJ: Prentice-Hall, 1983.
- [7] Sun Jiguang. A note on local behavior of multiple eigenvalue[J]. *SIAM J Matrix Anal Appl*, 1989, 10(4): 533 ~ 541.
- [8] 黄克中,毛善培. 随机方法与模糊数学应用[M]. 上海:同济大学出版社,1987.

Parameter Identification of Dynamic Models Using a Bayes Approach

Li Shu¹, Zhuo Jiashou², Ren Qingwen²

(1. *Institute of Aircraft Design, Beijing University of Aeronautics & Astronautics, Beijing 100083, P R China;*

2. *Institute of Civil Engineering, Hehai University, Nanjing 210098, P R China*)

Abstract: The Bayesian method of statistical analysis has been applied to the parameter identification problem. A method is presented to identify parameters of dynamic models with the Bayes estimators of measurement frequencies. This is based on the solution of an inverse generalized eigenvalue problem. The stochastic nature of test data is considered and a normal distribution is used for the measurement frequencies. An additional feature is that the engineer's confidence in the measurement frequencies is quantified and incorporated into the identification procedure. A numerical example demonstrates the efficiency of the method.

Key words: parameter identification; dynamic models; Bayes estimators; inverse eigenvalue problem; prior distribution; posterior distribution