

考虑非线性环境因素影响的 结构损伤预警方法研究

郑泓, 段忠东

(哈尔滨工业大学(深圳)土木与环境工程学院, 广东 深圳 518055)

摘要: 在长期监测土木工程结构损伤的过程中, 环境因素(如温度)变化引起的结构特性改变往往会被误认为损伤, 从而导致损伤预警结果的可信度大大降低。针对此类问题, 本文将高斯混合模型(Gaussian Mixture Model, GMM)的聚类算法引入到基于协整(Cointegration, CI)理论的损伤预警方法中, 发展非线性环境因素影响下的结构损伤预警方法。该方法以识别的结构模态频率作为协整变量, 根据分段线性化的思想, 利用GMM对不同环境条件下的频率样本进行聚类; 对相同簇群的频率样本建立线性协整方程; 以协整残差作为损伤指标, 通过X-bar控制图实现损伤预警。将所提方法应用于瑞士Z24桥的现场测试数据, 结果表明, 该方法能够有效消除非线性环境温度的影响, 减少损伤误报。

关键词: 结构健康监测; 损伤预警; 环境影响; 高斯混合模型; 协整

中图分类号: TU311.3; TU312+.3 **文献标志码:** A **文章编号:** 1004-4523(2021)06-1101-11

DOI: 10.16385/j.cnki.issn.1004-4523.2021.06.001

引言

土木工程结构在服役期间容易受到环境腐蚀、车辆荷载、自然灾害等的共同作用而出现功能退化, 严重时甚至会导致结构失效而引发灾难性事故。为保证人员生命和财产安全, 有必要对在役结构的状态进行监测^[1-2]。近几十年来, 基于振动的结构损伤预警方法已成为结构健康监测(Structural Health Monitoring, SHM)的重要手段^[3-5]。该方法认为损伤会改变结构动力参数(如模态频率等), 通过比较损伤前后这些参数的变化可以识别损伤。但是, 环境因素(如温度、车载等)的改变也会引起结构动力参数的变化^[6-9], 这种变化往往会掩盖结构的真实损伤。如果不消除损伤预警过程中环境因素的影响, SHM的可信度将大打折扣。为此, 考虑环境影响的损伤预警方法开始受到学者的关注。

国内外学者对这一问题开展了大量研究。Sohn和Farrar^[10]将消除环境因素影响的过程称为“数据标准化(Data Normalization)”。根据是否需要测量环境信息, 该过程分为基于模型的方法和非模型的方法。前者需要借助采集的环境信息建立结构动力参数与环境变量之间相互关联的数学模型, 以量化环境因素对结构动力参数的影响。Peeters和

De Roeck^[6]采用ARX模型建立了Z24桥的环境温度与模态频率的关系; 文献[11-12]先后采用支持向量回归模型(Support Vector Regression, SVR)和人工神经网络(Artificial Neural Network, ANN)量化温度对香港汀九桥模态频率的影响。Yang等^[13]采用线性回归方法建立润扬长江大桥的模态频率和温度的回归关系。基于模型的方法为理解环境因素如何影响结构物理特性的机理奠定了基础, 但实际工程中某些环境变量很难测量, 准确建立环境因素和结构动力参数之间的关系难度很大。非模型的方法无需利用环境信息, 而是将环境变量当作隐藏变量, 通过机器学习的手段从大量样本中找出环境因素影响的内在规律^[14]。基于主成分分析(Principle Component Analysis, PCA)的方法^[15]假定环境因素改变是结构动力参数出现变化的主要原因, 采用PCA可以将环境因素以主成分的形式分离出来。林友新等^[16]认为结构模态频率受环境源和损伤源控制, 通过盲源分离(Blind Source Separation, BSS)将二者分开后, 可以根据损伤源信号是否发生突变识别损伤。刁延松和任红^[17]对结构响应数据建立自回归模型, 利用因子分析(Factor Analysis)去除温度对自回归系数的影响。Santos等^[18-19]采用改进的高斯混合模型学习结构模态频率在不同环境条件下的概率分布规律, 利用聚类后的簇群表示不同环境因素的影

响。邱雷等^[20]利用 GMM 建立导波监测特征参数在环境因素影响下的概率模型,采用 KL 距离衡量基准 GMM 和待测 GMM 的迁移趋势,从而实现损伤预警。

近几年,基于协整理论的损伤预警方法得到了快速发展。Cross 等^[21-22]最早对结构健康监测中的非平稳数据进行协整研究,发现模态频率存在的协整关系不会因为环境条件的改变而发生变化,但损伤的出现会破坏这种均衡关系,通过协整残差可以对损伤进行定量。刁延松等^[23]提出基于自回归模型系数和协整的结构损伤预警方法,并通过海洋平台振动台模型验证该方法的有效性。Huang 等^[24]将结构模态频率之间的协整系数作为卡尔曼滤波的状态向量,通过观察协整关系的改变实现在线损伤预警。李秀娟等^[25]将协整理论引入到压电阻抗法中,对温度变化影响下的阻抗谱峰值频率进行协整,利用协整残差的突变识别损伤。基于协整理论的方法从非平稳时间序列分析入手,利用协整关系消除环境因素引起的数据非平稳性,相比于传统复杂的数据分析算法具有原理简单、易于实现的特点。然而,实际工程中的协整关系往往具有一定的非线性^[21],采用传统的线性协整理论往往无法描述真实的协整关系,因此需要发展基于非线性协整的损伤预警方法^[26-27]。

本文结合高斯混合模型的聚类算法和协整理论进行非线性环境因素影响下的损伤预警研究。该方法以结构模态频率作为协整变量,利用高斯混合模型对不同环境条件下的频率样本进行概率分布拟合,然后依据高斯后验概率将这些样本分为不同的簇群;根据分段线性化思想,如果簇群的数量(聚类数)取值合理,那么相同簇群里的样本近似满足线性协整关系,采用 Johansen 检验可求出各个簇群的协整方程;对于待测样本,同样采用高斯后验概率确定其簇群归属,之后代入相应的协整方程求出协整残差,通过 X-bar 控制图^[28]实现结构损伤预警。

1 高斯混合聚类

高斯混合模型是高斯模型的简单扩展^[29]。假设随机向量 $\boldsymbol{x} \in \mathbb{R}^{l \times 1}$ 服从多元高斯分布,其概率密度函数可由均值向量 $\boldsymbol{\mu} \in \mathbb{R}^{l \times 1}$ 和协方差矩阵 $\boldsymbol{\Sigma} \in \mathbb{R}^{l \times l}$ 决定,具体表达式为

$$p(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{l}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})} \quad (1)$$

由此可以定义 GMM 的概率密度函数为

$$p(\boldsymbol{x}) = \sum_{i=1}^m \alpha_i p(\boldsymbol{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \quad (2)$$

式中 m 代表高斯分量的个数; α_i 代表第 i 个高斯分量的混合系数,满足 $\sum_{i=1}^m \alpha_i = 1$; $p(\boldsymbol{x}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ 代表第 i 个高斯分量的概率密度函数,其均值向量和协方差矩阵分别为 $\boldsymbol{\mu}_i$ 和 $\boldsymbol{\Sigma}_i$ 。

GMM 聚类算法^[29]采用多个高斯分布的线性组合作为数据分布的概率密度函数,通过概率模型对应的后验概率确定簇群的划分。如果样本集 $\boldsymbol{X} = [\boldsymbol{x}_1 \ \boldsymbol{x}_2 \ \cdots \ \boldsymbol{x}_s]$ 由式(2)所表示的 GMM 生成,引入新的随机变量 $z_j \in \{1, 2, \dots, m\}$ 代表生成样本 \boldsymbol{x}_j 的高斯分量,显然其先验概率 $p(z_j = i) = \alpha_i$ 。根据贝叶斯定理, z_j 的后验概率可以表示为

$$p(z_j = i | \boldsymbol{x}_j) = \frac{p(z_j = i) p(\boldsymbol{x}_j | z_j = i)}{p(\boldsymbol{x}_j)} = \frac{\alpha_i p(\boldsymbol{x}_j | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)}{\sum_{k=1}^m \alpha_k p(\boldsymbol{x}_j | \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)} \quad (3)$$

采用期望最大(Expectation Maximization, EM)算法求出 GMM 的未知参数后,可根据式(3)计算出后验概率最大的高斯分量作为 \boldsymbol{x}_j 的归属。相比于传统的 K 均值聚类, GMM 聚类产生的簇群形状可以是任意的椭圆,实际应用范围更广。

2 协整理论

协整最早用于解决非平稳时间序列建模时引起的虚假回归问题。其核心思想是利用多个非平稳变量之间存在的长期均衡(协整)关系,通过线性组合的方式将非平稳性时间序列转化为平稳时间序列。

如果非平稳时间序列 $y_t (t = 1, 2, \dots)$ 经过 d 次差分之后刚好成为平稳时间序列,则称 y_t 具有 d 阶单整性,记为 $y_t \sim I(d)$ 。显然,平稳时间序列的单整性为 $I(0)$ 。假定随机向量 $\boldsymbol{y}_t = \{y_{1t} \ y_{2t} \ \cdots \ y_{lt}\}^T$ 的每个分量都具有 1 阶单整性^[30],而且存在 l 维向量 $\boldsymbol{\beta} = \{\beta_1 \ \beta_2 \ \cdots \ \beta_l\}^T$ 使得 \boldsymbol{y}_t 的线性组合变为平稳时间序列,即 $\boldsymbol{\beta}^T \boldsymbol{y}_t \sim I(0)$,那么 \boldsymbol{y}_t 存在协整关系,其协整方程表示为

$$\beta_1 y_{1t} + \beta_2 y_{2t} + \cdots + \beta_l y_{lt} = \varepsilon_t \quad (4)$$

式中 $\boldsymbol{\beta}$ 为协整向量; ε_t 为协整残差。

目前,协整性检验主要分为 EG(Engle-Granger)两步法和 Johanson 法。EG 检验^[31]是以残差为基础的检验方法,首先采用最小二乘法对变量进行回归建模,再通过回归残差的平稳性检验确定协整关

系是否成立。以随机向量 y_t 为例,选择 y_{1t} 作为因变量建立如下回归方程

$$y_{1t} = \hat{\beta}_2 y_{2t} + \cdots + \hat{\beta}_l y_{lt} + e_t \quad (5)$$

式中 $\hat{\beta}_2, \dots, \hat{\beta}_l$ 为回归系数, e_t 为回归残差。如果 e_t 能通过 ADF (Augmented Dickey-Fuller) 平稳性检验^[31], 则 y_t 存在协整关系。回归分析需要指定自变量和因变量, 所以 EG 法每次只能检验一个协整关系, 其缺点是不同因变量对应的协整向量不一定满足独立性。通常在协整变量较少的时候采用 EG 法。

Johanson 检验^[31]是以向量误差修正 (Vector Error Correction, VEC) 模型为基础的多变量协整性检验方法。VEC 模型不仅包含变量间的长期协整关系, 而且考虑了变量短期波动对协整方程的影响。因此, Johanson 法在协整向量的估计精度上要高于采用简单回归分析的 EG 法。仍以随机向量 y_t 建立如下 VEC 模型

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^p \Psi_i \Delta y_{t-i} + u_t = AB^T y_{t-1} + \sum_{i=1}^p \Psi_i \Delta y_{t-i} + u_t \quad (6)$$

式中 $\Pi \in \mathbb{R}^{l \times l}$ 代表长期均衡矩阵, 与变量协整关系相关; $\Psi_i \in \mathbb{R}^{l \times l}$ 代表短期动态矩阵, 与变量瞬时波动特征相关; $\Delta y_t \in \mathbb{R}^{l \times 1}$ 为 y_t 的差分项; p 为滞后阶次; u_t 为 l 维白噪声向量。矩阵 Π 由修正矩阵 $A \in \mathbb{R}^{l \times r}$ (协整残差对 Δy_t 的修正速度) 和协整向量矩阵 $B \in \mathbb{R}^{l \times r}$ (每一列代表一个协整向量) 组成, 其中 r 代表矩阵 Π 的秩 ($r=0$ 说明 y_t 不存在协整关系)。因为 $y_t \sim I(1)$, 则 $\Delta y_t, \Delta y_{t-i} \sim I(0)$, 式(6)中的所有变量都具有平稳性, 采用极大似然法估计协整向量矩阵 B , 然后选择第一列协整向量 (对应特征值最大、平稳性最强^[31]) 建立协整方程。

3 结构损伤预警方法

结构各阶模态频率受环境因素影响存在协整关系^[30], 损伤的出现将破坏这种关系而导致协整残差出现突变, 通过比较损伤前后的协整残差可以实现损伤预警。但是, 环境因素影响往往存在一定的非线性, 协整变量间难以保证较好的线性协整关系, 导致基于线性协整理论的损伤预警方法精度不高。为解决上述问题, 本文根据分段线性化思想, 利用 GMM 聚类将非线性协整关系转化为多个线性协整关系, 弥补协整理论无法处理非线性环境因素影响的问题。

3.1 几何解释

为说明 GMM 聚类结合线性协整理论具有处理非线性环境因素影响的能力, 以模态频率 f_1 和 f_2 组成的二维样本为例。如图 1 所示, 由于环境因素的影响, f_1 和 f_2 存在明显的双线性关系^[32], 线性协整理论无法准确描述二者的非线性协整关系。本文采用 GMM 聚类将这些频率样本分割成簇群 1 和簇群 2。根据分段线性化思想, 假定聚类后每个簇群都不存在非线性特征, 采用 Johanson 法建立两个线性协整方程近似原来的非线性协整关系。

对于损伤样本点 A_0 , 计算其后验概率可确定其簇群归属: 如果 A_0 属于簇群 1 的概率较大, 那么利用协整方程 1 求出协整残差 (线段 $A_0 A_1$ 的长度) 作为损伤指标; 反之, 如果 A_0 属于簇群 2 的概率较大, 那么根据协整方程 2 求得协整残差 (线段 $A_0 A_2$ 的长度) 作为损伤指标。

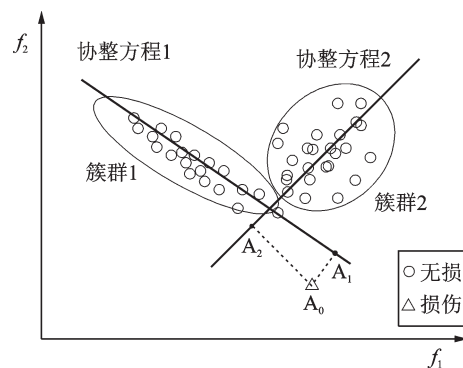


图1 非线性环境因素影响下模态频率 f_1 和 f_2 的散点分布图
Fig. 1 Scatter plot of modal frequencies f_1 and f_2 under non-linear environmental effects

3.2 损伤预警流程

土木工程结构可能产生的损伤千差万别, 而且在服役期间的大部分时段都处于正常状态, 缺少或很难建立起完整的损伤样本库, 无法通过样本匹配的方式确定未知样本是否属于某类损伤样本。最常用的方法是将损伤样本看作离群点, 采用 X-bar 控制图^[28]实现损伤预警。针对正态分布的样本总体, X-bar 控制图根据小概率事件原理确定警戒线, 对异常事件进行实时监测和报警。

假定结构无损伤 (参考) 样本的协整残差满足正态分布, 根据 99% 置信度确定控制上限 (Upper Control Limit, UCL) 和控制下限 (Lower Control Limit, LCL) 分别为

$$\begin{aligned} UCL &= CL + 2.58\sigma, \\ LCL &= CL - 2.58\sigma \end{aligned} \quad (7)$$

式中 CL 和 σ 分别代表协整残差的均值和标准差。当待测样本的协整残差落在上下控制限范围之外,

则说明结构出现损伤。

因为 GMM 聚类得到的不同簇群可能存在不同的离散程度(即不同簇群的样本标准差 σ 存在差异),如图 1 中簇群 2 的样本离散性大于簇群 1,导致根据式(7)确定的控制上下限也会出现不同:采用簇群 2 对应的 X-bar 控制图进行损伤预警,由于其控制上下限之间的区域较宽,某些归属于簇群 1 的损伤样本可能会被误认为无损样本;反之,采用簇群 1 对应的 X-bar 控制图会导致某些归属于簇群 2 的健康样本被误判为损伤样本。因此,需要消除不同簇群的样本离散性对 X-bar 控制图的影响,最简单的方法是对归属于簇群 i 的协整残差样本 $\epsilon_i^{(i)}$ 进行标准化处理,处理后的协整残差 $\bar{\epsilon}_i^{(i)}$ 可表示为

$$\bar{\epsilon}_i^{(i)} = \frac{\epsilon_i^{(i)}}{\sigma^{(i)}} \quad (8)$$

式中 $\sigma^{(i)}$ 为簇群 i 所有参考样本的协整残差标准差。理论上,标准化后所有参考样本的协整残差都满足标准正态分布 $N(0, 1)$,可采用统一的 X-bar 控制图对不同簇群的待测样本进行监测。

综上,基于 GMM 聚类和协整理论(简称为 GMM-CI)的损伤预警方法分为以下 4 步。

Step 1: 模态参数识别。采集无损伤结构在不同环境条件下的振动响应数据,并采用运行模态分析技术^[33](如随机子空间法、频域分解法等)识别结构模态频率作为协整变量。

Step 2: GMM 聚类。将 Step 1 得到的频率样本作为参考样本,利用 GMM 聚类对其进行聚类分析。其中,聚类数 m ,可根据贝叶斯信息准则^[34](Bayesian Information Criterion, BIC)确定,其表达式为

$$BIC(m) = -2\ln L(m) + n(m)\ln s \quad (9)$$

式中 $L(m)$ 为似然函数; $n(m)$ 为待估计参数个数; s 为参考样本数量。随着聚类数 k 的增加, BIC 值会先减少后增大,选择 BIC 最小值对应的 m 值作为最佳聚类数。

Step 3: 估计协整方程。对聚类后相同簇群的参考样本进行 Johansen 检验,选择最大特征值对应的协整向量建立如式(4)所示的协整方程,同时获得该簇群的协整残差样本。

Step 4: 建立 X-bar 控制图。计算每个簇群协整残差样本的标准差 $\sigma^{(i)}$,并按照式(8)对样本进行标准化处理,之后根据式(7)确定 X-bar 控制图的控制上下限。

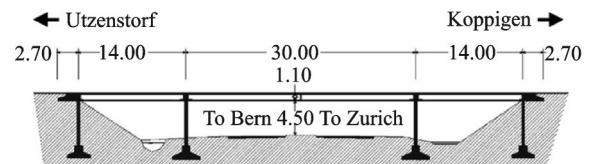
Step 5: 损伤预警。对于结构未知状态的振动响应,首先按照 Step 1 得到其频率样本,然后通过式(3)计算 GMM 后验概率确定簇群归属,之后利用对应的协整方程求出协整残差,最后根据式(8)对协整残差进行标准化处理,如果标准化后的协整残差没

有落在 X-bar 控制图上下限的范围内,则说明结构出现损伤。

4 试验数据验证

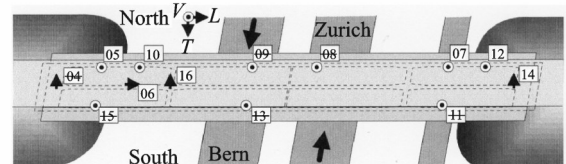
为验证上述方法的有效性,将其应用于瑞士 Z24 桥的现场试验数据。该桥为预应力混凝土箱梁桥,主跨 30 m,两边跨均为 14 m,如图 2(a)所示。为研究环境因素以及损伤对结构动力参数的影响,从 1997 年 11 月 11 日到 1998 年 9 月 11 日, Z24 桥安装的健康监测系统采集了结构加速度响应数据和包括环境温湿度、风速、风向等在内的环境数据。布置的 16 个加速度传感器在监测过程中只有 8 个(编号 3, 5, 6, 7, 10, 12, 14, 16)保持正常工作,如图 2(b)和 (c)所示。结构渐进破坏试验持续了一个月(1998 年 8 月 9 日到 1998 年 9 月 9 日),设置的损伤^[32]依次是桥墩沉降、混凝土脱落、墩台滑坡、混凝土铰接失效、锚头失效和钢筋破裂。

Peeters 和 De Roeck^[6]采用随机子空间法识别 Z24 桥的前 4 阶模态频率,其变化如图 3 所示。因为环境因素的影响,即使结构没有出现损伤(样本点 1 到 3470),其模态频率仍出现不同程度的波动,而且



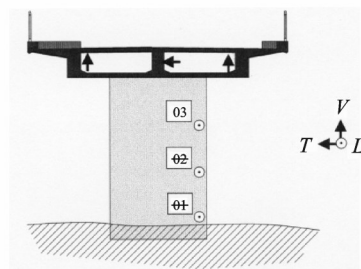
(a) 侧视图(单位: m)

(a) Lateral view (Unit: m)



(b) 桥面处加速度传感器布点

(b) Layout of accelerometers on the bridge deck



(c) 桥墩处加速度传感器布点

(c) Layout of accelerometers on the pier

图 2 瑞士 Z24 桥及其加速度传感器分布^[32]

Fig. 2 Switzerland Z24 bridge and the layout of acceleration sensors^[32]

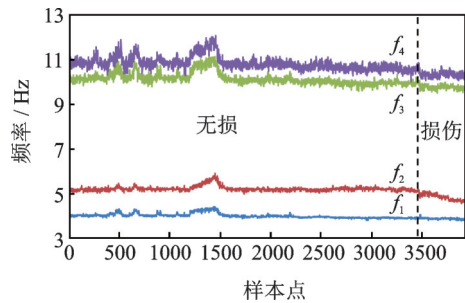


图 3 瑞士 Z24 桥前 4 阶模态频率的变化趋势
Fig. 3 Variations of the first four modal frequencies of the Z24 Bridge

高阶模态频率的波动性明显大于低阶模态。渐进破坏试验(样本点 3471 到 3932)导致结构各阶模态频率均呈现下降的趋势。

Z24 桥的模态频率随着桥梁路面沥青层温度的升高而降低,同时存在明显的双线性关系(如图 4 所示),其原因^[32]可解释为:当温度大于 0 °C 时,路面沥青的弹性模量改变对结构刚度的影响起主导作用,温度越高弹性模量越低,导致结构模态频率随着温度升高而降低;当温度小于 0 °C 时,结冰引起的结构边界条件改变起主导作用,低温增强了边界约束,同时也增大结构刚度,出现结构模态频率随着温度降低而增大的现象。从协整理论的角度看,以模态频率作为协整变量至少存在两种均衡关系,是一种非线性协整关系。

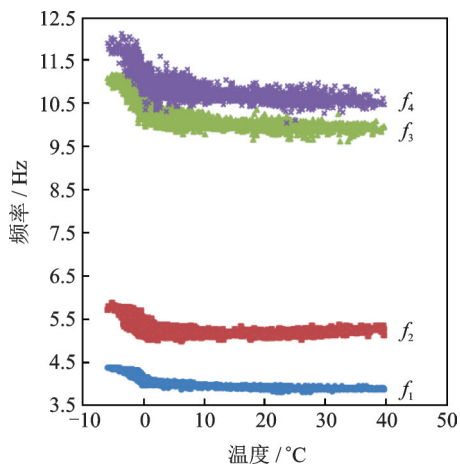


图 4 结构模态频率和温度的双线性关系
Fig. 4 Bi-linear relationship between structural modal frequency and temperature

4.1 损伤预警结果

以 f_1 和 f_2 作为协整变量,将图 3 中编号 1 到 3000 的样本作为参考(训练)样本,编号 3001 到 3932 的样本作为待测样本(其中编号 3001 到 3470 的样本是无损伤样本,编号 3471 到 3932 的样本是损伤样本),采用本文所提基于高斯混合聚类和协整理论的损伤预

警方法预报结构损伤。

首先采用 BIC 准则确定高斯混合聚类的最佳聚类数。如图 5 所示,聚类数为 3 时的 BIC 值达到最小。因此,采用包含 3 个高斯分量的 GMM 对参考样本进行聚类,其结果如图 6 所示。根据各簇群样本对应的环境温度数据(见图 7)可知,簇群 1(对应的环境温度大于 0 °C)和簇群 3(对应的环境温度小于 0 °C)代表上述两种由温度主导的均衡机制;簇群 2(既有大于 0 °C 的样本,也有小于 0 °C 的样本)可能是其他环境因素(风速或车载等)主导的均衡机制。

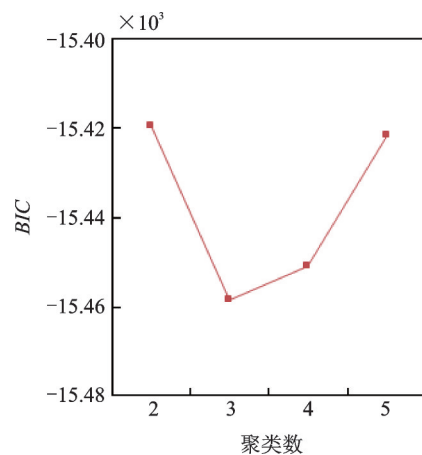


图 5 BIC 准则确定 GMM 聚类数
Fig. 5 BIC criterion to determine the cluster number

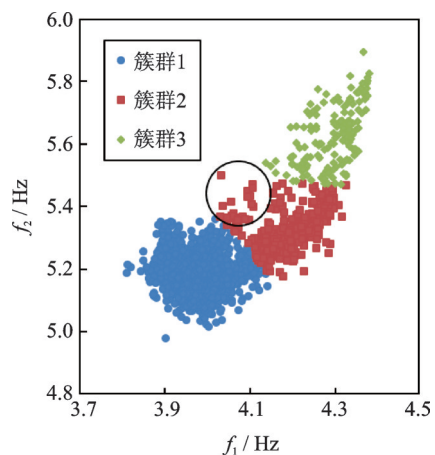


图 6 参考样本 GMM 聚类结果
Fig. 6 GMM clustering results of reference samples

对各簇群的参考样本进行 ADF 检验,设定置信水平为 5% 置信水平,则 $p < 5%$ (或者 t -统计量小于 -1.941) 代表变量是平稳时间序列。从表 1 可出,模态频率 f_1 和 f_2 是非平稳时间序列,但其一阶差分是平稳时间序列,即模态频率 f_1 和 f_2 都具有一阶单整性。

确定模态频率 f_1 和 f_2 都满足一阶单整条件后,采用 Johanson 检验计算各簇群对应的协整方程,其表达式为

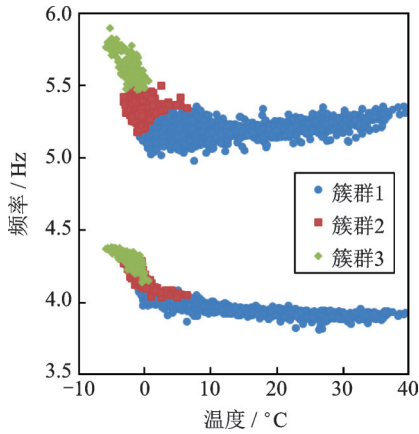


图7 各簇群模态频率与温度对应关系

Fig. 7 Relationship between modal frequencies and temperature in each cluster

$$\epsilon_i = \begin{cases} -4.73f_1 - 21.76f_2 + 131.57, & \text{if 簇群1} \\ 9.44f_1 - 15.05f_2 + 40.61, & \text{if 簇群2} \\ -20.02f_1 + 10.56f_2 + 26.50, & \text{if 簇群3} \end{cases} \quad (10)$$

利用上式可得不同簇群参考样本的协整残差,之后根据式(8)对协整残差进行标准化处理,最后确定 X-bar 控制图的控制上下限。

对于待测样本,根据式(3)计算其属于不同簇群的概率,选择最大概率值对应的簇群作为最终归属,之后将样本代入式(10)中对应簇群的协整方程,求得协整残差并按照式(8)对其进行标准化处理,最后通过 X-bar 控制图进行损伤预警,结果如图8所示。从图中可以看出:参考样本的协整残差不存在明显的变化趋势(样本点1500附近的异常突变与图6中画圈部分样本的模态识别偏差有关),说明采用本文所提方法能够有效消除非线性环境因素的影响;待测损伤样本的协整残差随损伤的累积有变大的趋势,说明协整残差能体现损伤的相对严重程度。需要指出的是,待测样本中有4.89%的无损伤样本被误判为损伤样本,2.38%的损伤样本被误判为无损伤样本。为进一步分析造成上述误差的原因,图9给出了归属于簇群1的参考样本以及所有待测样本的分布情况(待测样本的聚类结果均为簇群1)。从

图中可以看出,出现误判的待测样本主要位于警戒线2附近:对于待测无损伤样本,理论上99%置信度确定的 X-bar 控制图会有1%的样本均匀分布在两条警戒线之外,而实际位于警戒线2以下的样本有19个,位于警戒线1以上的样本只有4个,其原因最有可能是局部样本的频率识别误差过大,导致位于警戒线2以下的误判样本远超过位于警戒线1以上的样本;至于损伤样本,由于渐进破坏试验初期损伤引起的协整残差变化不明显,再加上测量噪声引起的样本不确定性,导致个别损伤样本被误认为无损伤样本。

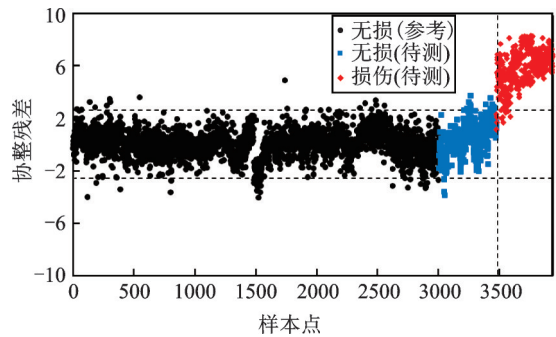


图8 基于GMM-CI方法的损伤预警结果(以 f_1 和 f_2 作为协整变量)

Fig. 8 Damage alert results based on GMM-CI method (using f_1 and f_2 as cointegration variables)

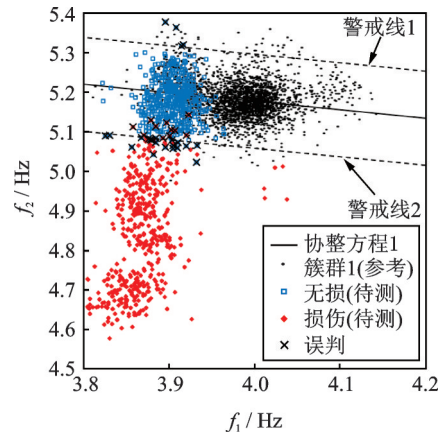


图9 参考样本与待测样本的散点分布

Fig. 9 Scatter distribution of reference sample and samples to be tested

表1 模态频率 f_1 和 f_2 及其一阶差分的 ADF 检验结果(5%置信水平)

Tab. 1 ADF test for modal frequencies f_1 , f_2 and their first differences (5% confidence level)

变量	簇群1		簇群2		簇群3	
	p值/t-统计量	是否平稳	p值/t-统计量	是否平稳	p值/t-统计量	是否平稳
f_1	0.57/-0.23	否	0.65/-0.01	否	0.52/-0.36	否
f_2	0.57/-0.21	否	0.60/-0.14	否	0.66/0.04	否
Δf_1	0.00/-76.60	是	0.00/-18.11	是	0.00/-12.40	是
Δf_2	0.00/-87.36	是	0.00/-24.81	是	0.00/-15.01	是

4.2 不同协整方法比较结果

为比较不同协整方法的结果,图10给出了传统基于线性协整理论的损伤预警结果。由于线性协整模型无法完全消除非线性环境因素的影响,损伤样本的误判率高达41.77%,在识别精度上远低于本文所提的基于GMM聚类相结合的协整理论的损伤预警方法。

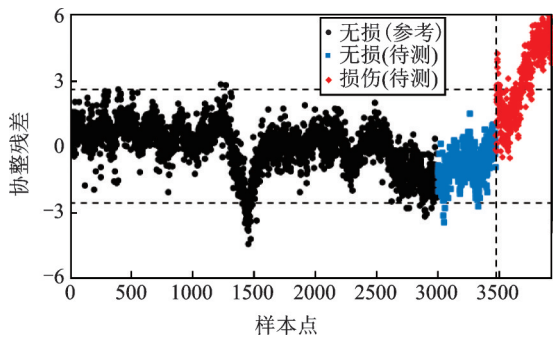


图10 基于线性协整理论的损伤预警结果

Fig. 10 Damage alert using linear co-integration theory

图11为Shi等^[31]提出的基于区制转移的非线性协整方法的损伤预警结果。该方法利用环境温度信息识别出模态频率之间存在两种协整关系(以 $0.98\text{ }^{\circ}\text{C}$ 为界),通过待测样本对应的温度信息划分其协整关系归属,之后根据对应的协整方程求得协整残差作为损伤指标。由于考虑了温度信息,基于区制转移的非线性协整方法的误判率(无损样本为2.77%,损伤样本为1.51%)均低于本文所提方法。但是,考虑环境因素的监测难度以及成本控制(比如边界条件、车载等),本文所提方法无需借助环境信息,更易于实际应用。

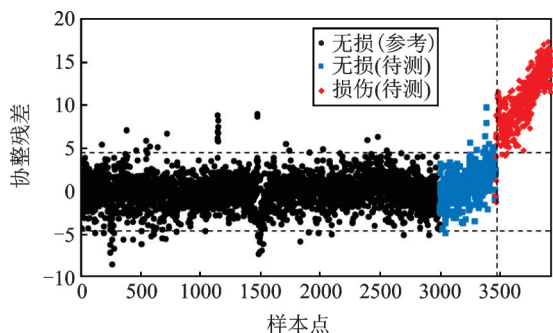


图11 基于区制转移的非线性协整方法得到的协整残差(利用 f_1 - f_4 和温度数据)

Fig. 11 Cointegration residuals based on regime-switching cointegration approach (using f_1 - f_4 and temperature data)

4.3 聚类个数对损伤预警结果的影响

BIC准则作为理想的模型选择方法,在实际应用中易受样本分布和测量噪声的影响,很有可能产生错误的结果,有必要研究不同聚类数对损伤预警

精度的影响,并提出合理的聚类数建议。如图12所示,当聚类数取2时,待测损伤样本的误判率较大,出现欠拟合现象;当聚类数取3,4和5时,损伤预警的误判率均处于5%附近。显然,欠拟合会大大降低本文所提方法的损伤预警能力,必须设法避免聚类数取值过小导致损伤预警失效。

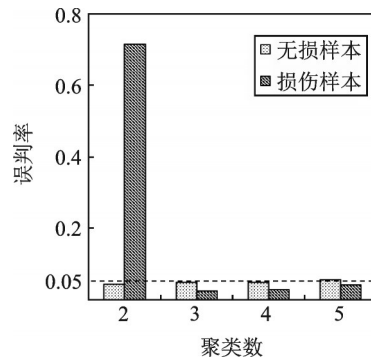


图12 不同聚类数对损伤预警精度的影响

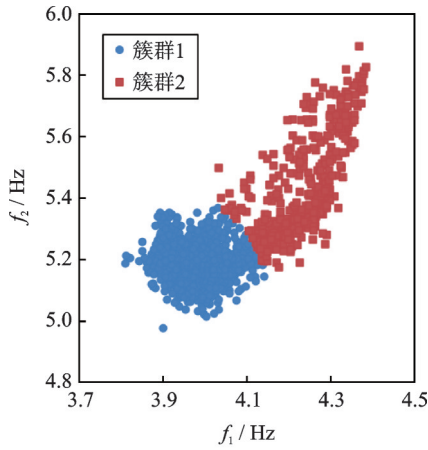
Fig. 12 Influences of different cluster numbers on damage alert accuracy

考虑GMM聚类的目的是将非线性协整关系转化为多个线性协整关系的组合,因此聚类数的选择理论上应保证聚类后各个簇群都不存在非线性特征。图13(a)中,簇群2由于聚类数不足仍具有非线性特征,采用线性协整理论进行建模必然会有误差。根据分段线性化思想,簇群数越多(分段数越多)越能近似非线性特征,比如图13(b)簇群1,4和5相比于图13(a)簇群1有更多细节上的线性特征,因此聚类数取4或5仍能较为准确地预警损伤。但是,聚类数取值太大不仅会造成过拟合现象,而且严重影响GMM聚类的计算效率和精度。所以本文建议在BIC准则的基础上,借助交叉验证^[35]方法产生多个备选聚类数,选择较大者作为最终聚类数。

4.4 多协整变量的损伤预警

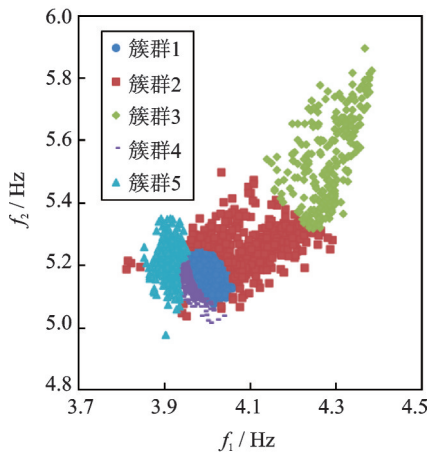
高维数据一般存在多种协整关系,不同的协整关系对损伤预警精度的影响很大。实际应用中可以采用主成分分析^[36](Principle Component Analysis, PCA)将高维数据进行降维处理,通常以第一主成分和第二主成分作为新的协整变量,利用二维数据协整关系的唯一性特点避开高维数据协整关系的选择难题。除此之外,降维不仅能够简化GMM聚类的复杂度,而且二维空间便于人工判断BIC准则确定的聚类数以及GMM聚类结果是否合理。

以模态频率 f_1 , f_2 和 f_3 为例,同样选择图3前3000个数据作为参考样本,采用PCA对数据进行降维处理,将得到的第一主成分 P_1 和第二主成分 P_2 作为新的协整变量,然后采用GMM聚类对降维后的参考样本进行聚类,其结果如图14所示。相比于图



(a) 聚类数取2

(a) Cluster number is 2



(b) 聚类数取5

(b) Cluster number is 5

图13 不同聚类数对GMM聚类结果的影响

Fig. 13 GMM clustering results with different cluster numbers

6,其簇群分布显得更加紧凑,但依然能够区分出3种协整关系。对于待测样本,同样需要经过PCA降维,再确定其簇群归属。

表2为各聚类参考样本的ADF检验结果。从表中可以看出,主成分 P_1 和 P_2 是非平稳时间序列,但其一阶差分是平稳时间序列,说明PCA分析得到的 P_1 和 P_2 都具有一阶单整性。

确定主成分 P_1 和 P_2 都满足一阶单整条件后,采用Johanson检验计算出各簇群对应的协整方程,其表达式为

表2 主成分 P_1 和 P_2 及其一阶差分的ADF检验结果(5%置信水平)

Tab. 2 ADF test for principle components P_1 , P_2 and their first differences (5% confidence level)

变量	簇群1		簇群2		簇群3	
	ρ 值/ t -统计量	是否平稳	ρ 值/ t -统计量	是否平稳	ρ 值/ t -统计量	是否平稳
P_1	0.56/-0.24	否	0.62/-0.08	否	0.61/-0.11	否
P_2	0.36/-0.80	否	0.52/-0.36	否	0.62/-0.08	否
ΔP_1	0.00/-84.42	是	0.00/-22.10	是	0.00/-15.76	是
ΔP_2	0.00/-86.18	是	0.00/-24.96	是	0.00/-16.68	是

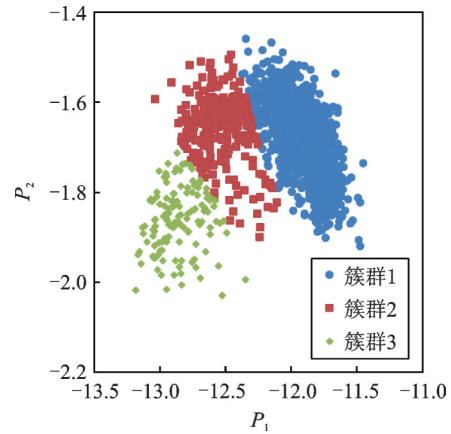


图14 PCA降维后参考样本的GMM聚类结果

Fig. 14 GMM clustering results of reference samples after PCA

$$\epsilon_t = \begin{cases} 4.18P_1 - 11.55P_2 + 32.20, & \text{if 簇群1} \\ -1.55P_1 + 12.15P_2 + 0.78, & \text{if 簇群2} \\ 7.37P_1 + 20.08P_2 + 121.45, & \text{if 簇群3} \end{cases} \quad (11)$$

之后的损伤预警参考4.1节。

图15和16分别给出了未采用PCA和采用PCA后本文所提方法的损伤预警结果。前者以最小特征值对应的协整向量建立各簇群的协整方程,无损伤样本和损伤样本的误判率分别为3.40%和2.60%;后者利用PCA降维技术避开协整向量的选择问题,无损伤样本和损伤样本的误判率分别为2.98%和3.03%。二者误判率相差不大,说明采用PCA对多维协整变量进行降维处理对损伤预警精

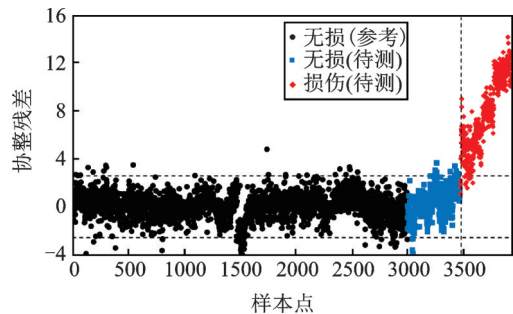


图15 基于GMM-CI方法的损伤预警结果(最小特征值对应的协整向量)

Fig. 15 Damage alert results based on GMM-CI method (using the cointegration vector corresponding to the smallest eigenvalue)

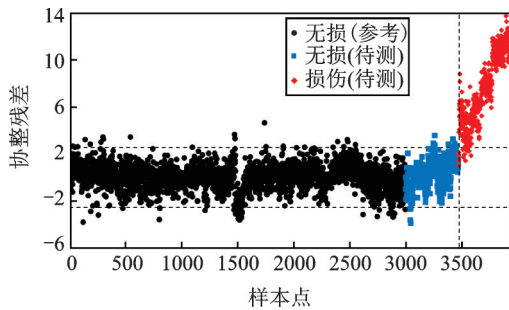


图16 基于PCA-GMM-CI方法的损伤预警结果

Fig. 16 Damage alert results based on PCA-GMM-CI method

度的影响不大。

最后,图17给出了基于线性协整方法的损伤预警结果。由于未采用GMM聚类对非线性协整关系进行线性化处理,其损伤样本的误判率仍有42.64%。需要说明的是,无论是基于GMM-CI的方法,还是基于线性协整的方法,都存在协整向量的选择问题。如果选择最大特征值对应的协整方程进行损伤预警,两种方法的损伤样本误判率均高达90%以上。其原因可解释为:虽然最大特征值对应的协整方程平稳性最强,但也说明这种协整状态很难被损伤打破(或者说对损伤不敏感),比如基于线性协整的方法中对应于最大特征值的协整方程为

$$\epsilon_t = 27.00f_1 + 0.73f_2 - 11.34f_3 + 3.23 \quad (12)$$

根据图3可知, f_2 对损伤最为敏感,但式(12)中其权重只有0.73,导致损伤引起的协整残差 ϵ_t 变化过小而被噪声覆盖。

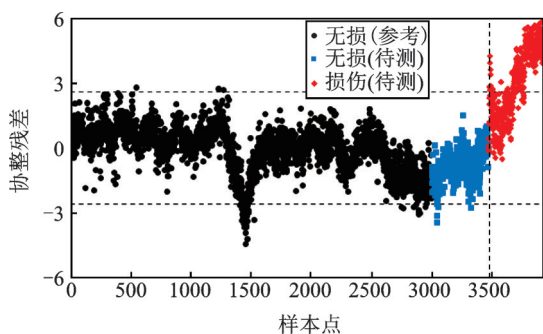


图17 基于线性协整方法的损伤预警结果(最小特征值对应的协整向量)

Fig. 17 Damage alert results based on linear cointegration method (using the cointegration vector corresponding to the smallest eigenvalue)

5 结论

为解决传统线性协整理论无法消除非线性环境因素影响的问题,本文将高斯混合聚类引入到基于协整的损伤预警方法。该方法以结构模态频率的协

整关系为基础,采用高斯混合聚类对存在非线性协整关系的协整变量进行分段线性化处理,聚类后相同簇群的样本便可采用线性协整理论进行建模。对于待测样本,先通过计算GMM后验概率确定其簇群归属,再根据对应的协整方程求得协整残差,最后利用X-bar控制图监测协整残差是否出现异常进行损失预警。

通过Z24桥的长期监测数据和破坏试验对上述方法进行验证,得到以下结论:(1)采用高斯混合聚类将协整变量间的非线性关系转化为多个线性关系的集合,从而将基于协整理论的损伤预警方法扩展到非线性协整理论;(2)利用主成分分析对协整变量进行降维处理可以在保证损伤预警精度的前提下,避开高维数据协整关系的选择问题,同时简化高斯混合模型的复杂度;(3)本文所提方法不需要测量环境信息,能够有效消除实际工程中非线性环境因素对损伤预警的影响,降低损伤预警的误判率。

GMM聚类采用EM算法估计模型参数,聚类结果易受EM算法的初值影响,导致损伤预警结果产生显著变化,如何提高GMM聚类的精度和稳定性是未来重点研究方向。

致谢 感谢比利时鲁汶大学的Guido De Roeck教授提供Z24桥的模态频率和环境温度数据。

参考文献:

- [1] Sohn H, Farrar C R, Hemez F M, et al. A review of structural health monitoring literature:1996-2001[R].LA-13976-MS, Los Alamos National Laboratory, 2004: 1-7.
- [2] Farrar C R, Worden K. An introduction to structural health monitoring[J]. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2006, 365(1851): 303-315.
- [3] Yan Y J, Cheng L, Wu Z Y, et al. Development in vibration-based structural damage detection technique[J]. Mechanical Systems & Signal Processing, 2007, 21(5): 2198-2211.
- [4] Fan W, Qiao P. Vibration-based damage identification methods: A review and comparative study[J]. Structural Health Monitoring, 2011, 10(1): 83-111.
- [5] Amezcua-Sanchez J P, Adeli H. Signal processing techniques for vibration-based health monitoring of smart structures[J]. Archives of Computational Methods in Engineering, 2016, 23(1): 1-15.
- [6] Peeters B, De Roeck G D. One-year monitoring of the Z24-Bridge: Environmental effects versus damage events [J]. Earthquake Engineering & Structural Dy-

- namics. 2001, 30(2): 149-171.
- [7] Sohn H. Effects of environmental and operational variability on structural health monitoring[J]. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2007, 365(1851): 539-560.
- [8] 杨 鸥, 刘 洋, 李 惠, 等. 时变环境与损伤耦合下桥梁结构频率及阻尼比的统计分析[J]. 计算力学学报, 2010, 27(3): 78-84.
Yang Ou, Liu Yang, Li Hui, et al. Cable bridge modal parameter statistical analysis under the time varying environment coupled with damage[J]. Chinese Journal of Computational Mechanics, 2010, 27(3): 78-84.
- [9] Xia Y, Chen B, Weng S, et al. Temperature effect on vibration properties of civil structures: A literature review and case studies [J]. Journal of Civil Structural Health Monitoring, 2012, 2(1): 29-46.
- [10] Sohn H, Farrar C R. Damage diagnosis using time series analysis of vibration signals [J]. Smart Material Structures, 2001, 10(3): 446-451.
- [11] Ni Y Q, Hua X G, Fan K Q, et al. Correlating modal properties with temperature using long-term monitoring data and support vector machine technique [J]. Engineering Structures, 2005, 27(12): 1762-1773.
- [12] Zhou H F, Ni Y Q, Ko J M. Constructing input to neural networks for modeling temperature-caused modal variability: Mean temperatures, effective temperatures, and principal components of temperatures[J]. Engineering Structures, 2010, 32(6): 1747-1759.
- [13] Deng Y, Ding Y L, Li A Q. Structural condition assessment of long-span suspension bridges using long-term monitoring data [J]. Earthquake Engineering and Engineering Vibration, 2010, 1(9): 123-131.
- [14] Santos A, Figueiredo E, Silva M F M, et al. Machine learning algorithms for damage detection: Kernel-based approaches [J]. Journal of Sound & Vibration, 2016, 363: 584-599.
- [15] Soo Lon Wah W, Chen Y T, Roberts G W, et al. Separating damage from environmental effects affecting civil structures for near real-time damage detection [J]. Structural Health Monitoring, 2018, 17(4): 850-868.
- [16] 林友新, 周 翠, 李宏男. 基于盲源分离的损伤识别方法[J]. 地震工程与工程振动, 2013, 33(6): 158-163.
Lin Youxin, Zhou Cui, Li Hongnan. Damage identification method based on blind source identification [J]. Earthquake Engineering and Engineering Dynamics, 2013, 33(6): 158-163.
- [17] 刁延松, 任 红. 基于 AR 模型和因子分析的结构损伤预警研究[J]. 振动与冲击, 2014, 33(18): 115-119.
Diao Yansong, Ren Hong. Study on the structural damage alarm based on AR model and factor analysis [J]. Journal of Vibration and Shock, 2014, 33(18): 115-119.
- [18] Santos A, Santos R, Silva M, et al. A global expectation-maximization approach based on memetic algorithm for vibration-based structural damage detection [J]. IEEE Transactions on Instrumentation & Measurement, 2017, 66(4): 661-670.
- [19] Santos A, Figueiredo E, Silva M, et al. Genetic-based EM algorithm to improve the robustness of Gaussian mixture models for damage detection in bridges [J]. Structural Control and Health Monitoring, 2017, 24(3): e1886.
- [20] 邱 雷, 房 芳, 袁慎芳, 等. 导波强化裂变聚合概率模型损伤监测方法[J]. 振动、测试与诊断, 2018, 38(3): 438-445.
Qiu Lei, Fang Fang, Yuan Shenfang, et al. Guided wave and enhanced split merge probability model based on damage evaluation method [J]. Journal of Vibration, Measurement & Diagnosis, 2018, 38(3): 438-445.
- [21] Cross E J, Worden K. Approaches to nonlinear cointegration with a view towards applications in SHM [J]. Journal of Physics Conference, 2011, 305: 12069.
- [22] Cross E J, Manson G, Worden K, et al. Features for damage detection with insensitivity to environmental and operational variations [J]. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2012, 468(2148): 4098-4122.
- [23] 刁延松, 曹亚东, 孙玉婷. 环境变化下基于 AR 模型系数和协整的海洋平台结构损伤识别[J]. 工程力学, 2017, 34(2): 179-188.
Diao Yansong, Cao Yadong, Sun Yuting. Structural damage identification based on ar model coefficients and cointegration for offshore platform under environmental variations [J]. Engineering Mechanics, 2017, 34(2): 179-188.
- [24] Huang J Z, Li D S, Li H N, et al. Damage identification of a large cable-stayed bridge with novel cointegrated Kalman filter method under changing environments [J]. Structural Control & Health Monitoring, 2018(4): e2152.
- [25] 李秀娟, 屈文忠, 肖 黎. 压电阻抗损伤识别温度影响的协整消除方法研究[J]. 振动与冲击, 2019, 38(18): 128-134.
Li Xiujuan, Qu Wenzhong, Xiao Li. Cointegration approach for the temperature effect compensation in EMI damage detection [J]. Journal of Vibration and Shock, 2019, 38(18): 128-134.
- [26] Zolna K, Dao P B, Staszewski W J, et al. Towards homoscedastic nonlinear cointegration for structural health monitoring [J]. Mechanical Systems & Signal Processing, 2016, (75): 94-108.
- [27] 李东升, 黄杰忠, 李宏男. 环境变化下基于核典型相关

- 分析与协整的损伤识别方法[J]. 中国公路学报, 2019, 32(11): 71-82.
- Li Dongsheng, Huang Jiezhong, Li Hongnan. Structural Damage identification based on kernel canonical correlation analysis and cointegration under changing environments [J]. China Journal of Highway and Transport, 2019, 32(11): 71-82.
- [28] 王毓芳,肖诗唐. 统计过程控制的策划与实施[M]. 北京: 中国经济出版社, 2005: 136-155.
- Wang Yufang, Xiao Shitang. Statistical Process Control in Planning and Implementation [M]. Beijing: China Economic Press, 2005: 136-155.
- [29] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016: 197-224.
- Zhou Zhihua. Machine Learning [M]. Beijing: Tsinghua University Press, 2016: 197-224.
- [30] Cross E J, Worden K. Cointegration and why it works for SHM [J]. Journal of Physics Conference Series, 2012, 382: 12-46.
- [31] Shi H, Worden K, Cross E J. A regime-switching cointegration approach for removing environmental and operational variations in structural health monitoring [J]. Mechanical Systems & Signal Processing, 2018, 103: 381-397.
- [32] Reynders E, De Roeck G. Vibration-based damage identification: The Z24 bridge benchmark [J]. Encyclopedia of Earthquake Engineering, 2014, 482: 1-8.
- [33] Magalhães F, Cunha Á. Explaining operational modal analysis with data from an arch bridge [J]. Mechanical Systems & Signal Processing, 2011, 25 (5): 1431-1450.
- [34] Schwarz G. Estimating the dimension of a model [J]. The Annals of Statistics, 1978, (6): 461-464.
- [35] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016: 24-28.
- Zhou Zhihua. Machine Learning [M]. Beijing: Tsinghua University Press, 2016: 24-28.
- [36] Atasever U H, Gunen M A, Besdok E. A new unsupervised change detection approach based on PCA based blocking and GMM clustering for detecting flood damage [J]. Fresenius Environmental Bulletin, 2018, 27 (3): 1688-1694.

Structural damage alert with consideration of the nonlinear environmental effects

ZHENG Hong, DUAN Zhong-dong

(School of Civil and Environmental Engineering, Harbin Institute of Technology, Shenzhen 518055, China)

Abstract: Damage accumulation always happens to civil engineering structures during their operations, degrades the structural performance, and may eventually lead to catastrophic event. Therefore, it is necessary to monitor the state of structures and alert dangerous behavior. However, during the long-period monitoring of structures, the changes of structural characteristics resulted from the variations of environmental factors (such as temperature) often mask the real structural damages, and lead to false warning. In this paper, a damage warning method based on Gaussian Mixture Model (GMM) clustering and cointegration (CI) theory is developed to remove the nonlinear environmental effects. This method uses the identified structural modal frequencies as cointegration variables. Assuming piecewise linear effect of environmental factors on structural frequencies, GMM is applied to cluster the frequency samples extracted from different environmental conditions. Then, the linear cointegration relationship between structural frequencies is established using Johansen test from the frequency samples in each cluster group. Finally, the damage is identified through the X-bar control chart using the cointegration residuals as damage indicators. The proposed method is applied to the in-situ test results of the Z24 Bridge in Switzerland, and the results show that the proposed method can effectively eliminate the nonlinear effect of temperature on the structural frequencies, and reduces the chance of false alert of structural damages.

Key words: structural health monitoring; damage alert; environmental variation; Gaussian mixture model; cointegration

作者简介: 郑泓(1986-),男,博士研究生。电话:15889601972;E-mail:zhenghong319@126.com

通讯作者: 段忠东(1968-),男,教授。E-mail:duanzd@hit.edu.cn