

Behavior Analysis of Sensor Fault Detection Using PCA Approach

HUANG Xiao-bin, NIU Zheng, NIU Yu-guang, LIU Ji-zhen

(Department of Automatics, North China Electric Power University, Baoding Hebei 071003, China)

Abstract: Principle component analysis is an effective multivariate statistical method for process monitoring and fault diagnosis. In this paper, a systematical analysis of detection behaviors of PCA for various sensor faults is presented. The characters of two basic detection statistics, Hotelling T^2 and Q are analyzed under different sensor fault types. Then fault detectability for every individual sensor is given. The validity of analytical results is demonstrated by a real-world application to a boiler system.

Key words: principle component analysis; sensor; fault detection

利用主元方法进行传感器故障检测的行为分析^①

黄孝彬, 牛 征, 牛玉广, 刘吉臻

(华北电力大学自动化系, 河北 保定 071003)

摘要: 主元分析方法(PCA)是基于多元统计分析的过程监测和故障诊断手段。在假设过程只存在传感器故障的情况下, 系统地分析了 PCA 方法在传感器典型故障下的检测行为。首先导出了 Hotelling T^2 和 Q 两个检测统计量在传感器不同故障下的变化关系和规律, 然后从理论上给出了每个传感器故障的可检测性条件。最后通过火电厂锅炉过程中传感器故障检测实例验证了所得到的结论。

关键词: 主元分析; 传感器; 故障检测

中图分类号: TP273

文献标识码: A

文章编号: 1004-1699(2003)04-0419-05

工业过程中存在着大量相关的过程变量, 这些测量值的相关性为传感器的故障诊断提供了宝贵的解析冗余。基于主元分析(PCA)的过程监测和故障诊断方法是利用过程变量间的相关关系, 建立正常工况下的主元模型, 通过检验新的数据样本相对于主元模型的背离程度, 从而发现异常和故障。通常的检测方法是判断 Hotelling T^2 和 Q 统计量是否超出它们的控制限。引起上述两个统计量超限的原因主要是过程变量的均值、方差或变量间的相关关系矩阵(或协方差矩阵)的变化。由于 PCA 方法完全基于过程的测量数据, 而各测量值实际上是真实变

量的值和传感器本身故障等附加信号的合成, 因此通过 PCA 方法实际上很难区分过程工况变动、过程故障和传感器故障。PCA 分析方法在传感器的故障检测与诊断方面已有一些研究和应用的报道^[1~3], 然而其检测的行为还有待深入的研究。

本文在假设过程处于稳态的运行状况的情况下, 细致地分析了 PCA 在传感器发生各种典型故障(恒偏差, 增益变化, 卡死)的情况下的检测行为。从理论上给出了在不同的传感器故障下, T^2 和 Q 统计量的变化规律, 并讨论了传感器故障的可检测性条件。

① 收稿日期: 2003-07-01

作者简介: 黄孝彬(1977—), 男, 博士, 从事控制系统故障检测与诊断, 大机组智能优化控制, bin_huangcn@yahoo.com.cn.

©1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

1 基于 PCA 的故障检测模型

PCA 方法通过主元分析提取过程变量间的相关关系,进行故障的检测和诊断。取过程正常运行条件下的一段数据样本集 $X \in R^{n \times p}$, 其中 n 是样本的个数, p 是传感器测量变量的个数。将 X 进行标准化, 得到 $\bar{X} = [X - 1_n u^T] D_\sigma^{-1/2}$, 其中 1_n 是所有元素都为 1 的 n 维列向量, $u = [u_1, \dots, u_p]^T$, $D_\sigma = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2)$ 分别为均值向量和方差矩阵。变量的相关系数矩阵为: $R = \frac{\bar{X}^T \bar{X}}{(n-1)}$ 。对标准化数据集 \bar{X} 作主元分析, 首先需求得 R 的 p 个特征值 λ_i 和对应的单位化正交特征向量 p_i 其中 $i = 1, \dots, p$, $\lambda_1 \geq \lambda_2 \dots \geq \lambda_p \geq 0$, 测量数据矩阵可分解为:

$$\bar{X} = TP^T + TP^T = TP^T + E \quad (1)$$

其中 $P = [p_1, p_2, \dots, p_a] \in R^{p \times a}$, $P = [p_{a+1}, p_{a+2}, \dots, p_p] \in R^{p \times (p-a)}$ 分别为主元负荷矩阵和残差负荷矩阵, $T \in R^{n \times a}$, $T \in R^{n \times (n-a)}$ 分别为主元得分矩阵和残差得分矩阵, a 为主元个数, E 为残差矩阵, $[P \ P]$ 是单位正交阵, $[T \ T]$ 为正交阵。这样过程数据向量通过主元分解后被投影到两个正交的互补空间: 主元子空间和残差子空间。当变量间具有很强的相关性时, 主元空间利用维数很少的几个变量就可以描述过程变量的大部分变化信息, 而残差子空间表示数据中的噪声成分。利用建立好的主元统计模型, 新的采集变量 x (标准化后为 \bar{x}) 可以表达为:

$$\bar{x} = PP^T \bar{x} + e = Pt + e \quad (2)$$

其中, $t = P^T \bar{x}$ 称为主元得分向量, e 为误差向量。故障的检测可以在两个子空间中进行, 具体就是在主元子空间中利用 Hotelling T^2 统计量进行统计检验; 在残差子空间中利用 Q 统计量进行统计检测^[4]。

T^2 统计量定义为

$$T^2 = t^T D_\lambda^{-1} t = \bar{x}^T P D_\lambda^{-1} P^T \bar{x} \leq T_\alpha^2 \quad (3)$$

$D_\lambda = \text{diag}(\lambda_1, \dots, \lambda_a)$ 为前 a 个主元的方差矩阵, T_α^2 是显著性水平为 α 的 T^2 统计量控制限。

Q 统计量定义为

$$Q = e^T e = (PP^T \bar{x})^T (PP^T \bar{x}) = \bar{x}^T PP^T \bar{x} \leq Q_\alpha^2 \quad (4)$$

Q_α^2 是显著性水平为 α 的 Q 统计量控制限, 控制限的具体算法见文献[5]。由于 T^2 反映了主元统计模型

内部的变动情况, 而 Q 反映了新样本在主元模型下的误差, 通过判断两个统计量的值是否超过其控制限, 可以检测被控过程是否出现了故障。

2 传感器故障下的 PCA 检测行为分析

传感器存在的典型故障有: 传感器恒偏差故障, 传感器增益故障, 以及传感器卡死故障等^[6]。当传感器发生故障时, 其测量信号中将携带故障信号, 相应的测量值的统计特性也将发生明显的变化。假定同时只有一个传感器发生故障, 则第 i 个传感器存在故障时测量向量可以表示为:

$$x' = x + E f_i(t) \quad (5)$$

其中, E_i 为单位阵的第 i 列, 表示故障的位置; $f_i(t)$ 为一标量函数, 表示故障类型。令 $\theta_i = E f_i(t)$ 为故障信号向量, 通常可认为, 故障信号与正常情况下的测量变量 x 相互独立。此时:

$$E(x') = E(x) + E(\theta_i) = u + \Delta u = u' \quad (6)$$

$$\text{var}(x') = \text{var}(x) + \text{var}(\theta_i) = D_\sigma + \Delta D_\sigma = D'_\sigma \quad (7)$$

其中, $\Delta u = [0, \dots, \Delta u_i, \dots, 0]^T$, $\Delta D_\sigma = \text{diag}(0, \dots, \Delta \sigma_i^2, \dots, 0)$, 首先给出传感器在正常情况下 Q 和 T^2 统计量的均值计算公式^[7]:

$$E(Q) = E((P^T \bar{x})^T (P^T \bar{x})) = \text{tr}\{P^T R P\} \quad (8)$$

$$E(T^2) = \text{tr}\{D_\lambda^{-1/2} P^T R P D_\lambda^{-1/2}\} \quad (9)$$

其中 $R = E(\bar{x} \bar{x}^T) = D_\sigma^{-1/2} E[(x - u)(x - u)^T] D_\sigma^{-1/2}$ 为变量的相关系数矩阵。对于新采集的样本数据 x' , 可导出 Q 和 T^2 统计量的均值分别为

$$E(Q') = \text{tr}\{P^T (I + \Delta D) R' (I + \Delta D) P\} + \text{tr}\{P^T D_\sigma^{-1/2} (\Delta u) (\Delta u)^T D_\sigma^{-1/2} P\} \quad (10)$$

$$E(T'^2) = \text{tr}\{D_\lambda^{-1/2} P^T (I + \Delta D) R' (I + \Delta D) P D_\lambda^{-1/2}\} + \text{tr}\{D_\lambda^{-1/2} P^T D_\sigma^{-1/2} (\Delta u) (\Delta u)^T D_\sigma^{-1/2} P D_\lambda^{-1/2}\} \quad (11)$$

其中

$$R' = (D'_\sigma)^{-1/2} E[(x' - u')(x' - u')^T] (D'_\sigma)^{-1/2},$$

$$D_\sigma^{-1/2} (D'_\sigma)^{1/2} = I + \Delta D.$$

2.1 传感器恒偏差故障情形

第 i 个传感器发生恒偏差故障时, 传感器测量值表示为 $x'_i = x_i + b_i$, 其中 b_i 为偏差故障的幅度, $\Delta u_i = b_i$ 为常数。根据式(6)(7), 此时第 i 个测量变量的均值发生了变化, 方差没有变化, 与其他的变量间的相关系数没有发生变化, 故有 $R' = R$, 由式(8)~(11), 可以证明统计量期望的变化为

$$\begin{aligned}\Delta E(Q) &= E(Q') - E(Q) \\ &= \text{tr}\{P^T D_\sigma^{-1/2} (\Delta u) (\Delta u)^T D_\sigma^{-1/2} P\} \\ &= \frac{(\Delta u_i)^2}{\sigma_i^2} \sum_{m=a+1}^p p_{mi}^2\end{aligned}\quad (12)$$

$$\begin{aligned}\Delta E(T^2) &= E(T'^2) - E(T^2) \\ &= \text{tr}\{D_\lambda^{-1/2} P^T D_\sigma^{-1/2} (\Delta u) (\Delta u)^T D_\sigma^{-1/2} P D_\lambda^{-1/2}\} \\ &= \frac{(\Delta u_i)^2}{\sigma_i^2} \sum_{m=1}^a \frac{p_{mi}^2}{\lambda_m}\end{aligned}\quad (13)$$

其中: p_{mi} 为第 m 个特征向量 p_m 的第 i 个元素, 定义 $S_Q = \sum_{m=a+1}^p p_{mi}^2$ 和 $S_T = \sum_{m=1}^a \frac{p_{mi}^2}{\lambda_m}$ 分别为 Q 统计量和 T^2 统计量的敏感系数, 由式(12)(13)可知, 第 i 个传感器恒偏差故障时 Q 和 T^2 统计量都将增大, 期望值增大的幅度和恒偏差的大小有关, 还与各自的敏感系数有关。

2.2 传感器增益故障情形

第 i 个传感器发生增益变化故障时, 传感器测量值表示为 $x'_i = \beta x_i$, 其中 $\beta > 0$ ($\beta \neq 1$) 为故障后传感器的增益系数, $E(x'_i) = \beta E(x_i) = u_i + (\beta - 1)u_i$, $\text{var}(x'_i) = \beta^2 \text{var}(x_i)$ 。此时第 i 个变量和其他的变量间的相关系数不变, 说明相关系数矩阵不变 $R' = R$, 而 $\Delta D = \text{diag}(0, 0, \dots, \alpha, \dots, 0)$, $\alpha = \beta - 1$ 。于是可导出统计量的均值变化为

$$\begin{aligned}\Delta E(Q) &= \text{tr}\{P^T [(\Delta D)R + R(\Delta D) + (\Delta D)R(\Delta D)] P\} \\ &\quad + \text{tr}\{P^T D_\sigma^{-1/2} (\Delta u) (\Delta u)^T D_\sigma^{-1/2} P\} \\ &= 2\alpha \sum_{m=a+1}^p \lambda_m p_{mi}^2 + \alpha^2 \sum_{m=a+1}^p p_{mi}^2 + \\ &\quad \frac{(\alpha \Delta u_i)^2}{\sigma_i^2} \sum_{m=a+1}^p p_{mi}^2\end{aligned}\quad (14)$$

$$\begin{aligned}\Delta E(T^2) &= 2\alpha \sum_{m=1}^a p_{mi}^2 + \alpha^2 \sum_{m=1}^a \frac{p_{mi}^2}{\lambda_m} + \\ &\quad \frac{(\alpha \Delta u_i)^2}{\sigma_i^2} \sum_{m=1}^a \frac{p_{mi}^2}{\lambda_m}\end{aligned}\quad (15)$$

当传感器增益变大时 ($\beta > 1$), Q 和 T^2 统计量都将增大; 当传感器的增益减小时 ($\beta < 1$), 根据式(14), Q 统计量增大的条件是:

$$2 \sum_{m=a+1}^p \lambda_m p_{mi}^2 + \alpha(1 + \Delta u_i^2 / \sigma_i^2) \sum_{m=a+1}^p p_{mi}^2 < 0 \quad (16)$$

即

$$\beta < 1 - \frac{2 \sum_{m=a+1}^p \lambda_m p_{mi}^2}{(1 + \Delta u_i^2 / \sigma_i^2) \sum_{m=a+1}^p p_{mi}^2} = \beta_Q \quad (17)$$

类似地, T^2 统计量增大的条件是:

$$\beta < 1 - \frac{2 \sum_{m=1}^a p_{mi}^2}{(1 + \Delta u_i^2 / \sigma_i^2) \sum_{m=1}^a \frac{p_{mi}^2}{\lambda_m}} = \beta_T \quad (18)$$

称 β_Q, β_T 为增益故障临界值, 当满足式(17)、(18)时传感器增益减小故障才会被检测出来。

2.3 传感器卡死故障情形

第 i 个传感器发生卡死故障对应于传感器的失效, 即无论被测信号如何变化, 其输出均固定在一个值不变。此时传感器的测量值可表示为 $x'_i = c$, $E(x'_i) = u_i + (c - u_i)$, $\text{var}(x'_i) = 0$, 第 i 个变量和其他的变量间的相关系数将变为 0, 即相关系数矩阵的第 i 行和第 i 列的非对角线元素将变为 0, 而 $\Delta D = \text{diag}(0, 0, \dots, -1, \dots, 0)$ 。可以计算, 此时检测统计量的变化正好是 2.2 中 $\beta = 0$ 的情况。因此, 根据 2.2 的分析结果, 统计量的期望变化的幅值可按式(14)(15)进行计算。

3 传感器故障的可检测性条件

下面分析传感器故障的可检测性条件, 即传感器发生多大的幅值的故障时可以被检测出来。对式(5)表示的传感器输出进行标准化,

$$\bar{x}' = D_\sigma^{-1/2} (x + E_i f_i(t) - u) = \bar{x} + D_\sigma^{-1/2} E_i f_i(t) \quad (19)$$

检测统计量为

$$Q' = \|P^T \bar{x}'\|^2 = \|P^T \bar{x} + P^T D_\sigma^{-1/2} E_i f_i(t)\|^2 \quad (20)$$

$$\begin{aligned}T'^2 &= \|D_\lambda^{-1/2} P^T \bar{x}'\|^2 \\ &= \|D_\lambda^{-1/2} P^T \bar{x} + D_\lambda^{-1/2} P^T D_\sigma^{-1/2} E_i f_i(t)\|^2\end{aligned}\quad (21)$$

应用向量的三角不等式得:

$$\|P^T \bar{x} + P^T D_\sigma^{-1/2} E_i f_i(t)\| \geq \|P^T \bar{x}\| - \|P^T D_\sigma^{-1/2} E_i f_i(t)\|$$

则 Q 统计量能检测出第 i 个传感器故障的幅值充分条件为: $\|P^T D_\sigma^{-1/2} E_i f_i(t)\| > 2Q_\alpha$, 即

$$\|f_i(t)\|_Q \geq \frac{2\sigma_i Q_\alpha}{\left(\sum_{k=a+1}^p p_{ki}^2\right)^{1/2}} \quad (22)$$

同理用 T^2 统计量检测出第 i 个传感器故障的幅值充分条件为

$$\|f_i(t)\|_T \geq \left[\frac{2\sigma_i T_a}{\sum_{k=1}^a \frac{p_{ki}^2}{\lambda_k}} \right]^{1/2} \quad (23)$$

定义两个故障检测幅值的比值为

$$r_f = \frac{\|f_i(t)\|_T}{\|f_i(t)\|_Q} = \frac{T_a}{Q_a} \left[\frac{S_Q}{S_T} \right]^{1/2} \quad (24)$$

称 r_f 为相对检测能力指标, 该指标表明两个统计量具有不同检测能力, 例如, r_f 大于 1 时, 则对于相同的故障, Q 统计量检测出故障时, T^2 统计量却一定不能检测出来, 即它对故障的敏感程度较低, 检测的能力较差。

4 实例研究

以火电厂锅炉生产过程实例说明 PCA 在传感器故障检测中的行为。图 1 为锅炉水位控制系统的流程图。汽包水位是反映锅炉汽水系统物质平衡的标志, 也是锅炉安全运行的重要参数。通过控制两台汽动给水泵的转速和一台电动给水泵的出口调节阀可调节锅炉的给水流量, 从而达到控制水位的目的。锅炉产生的蒸汽经过过热后得到具有一定温度和压力的蒸汽进入汽轮机做功。在正常负荷下, 只有两台汽动给水泵运行, 因此本文取 R_1 (汽泵 A 转速 r/min), R_2 (汽泵 B 转速 r/min), P_1 (汽包压力 MPa), F_1 (给水流量 t/h), P_2 (调节级压力 MPa), P_3 (主汽压力 MPa) 等 6 个传感器测量变量进行检测。

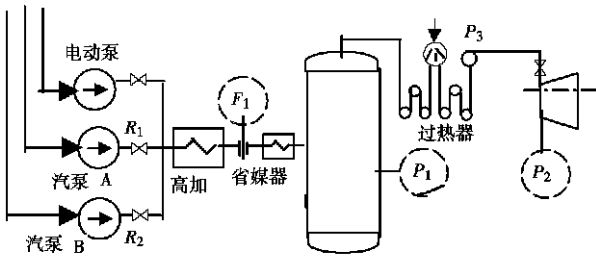


图 1 锅炉给水系统生产流程图

试验数据来源于某 300 MW 机组稳定负荷状况下的 1 000 组锅炉运行数据, 前 800 组数据用于建立主元检测模型, 后 200 组数据用于模拟传感器的不同故障。由于前 2 个主元的方差累计贡献率为 89.35%, 故取主元个数 $a=2$, 相应的 T^2 和 Q 统计量的控制限分别为 $T_a^2=9.26$ ($\alpha=99\%$) 和 $Q_a^2=4.22$ ($\alpha=99\%$)。

4.1 检测行为的分析结果

主元模型建立后, 根据前述分析可定量地计算

出传感器故障检测的特性参数, 以用于检测行为的分析。检测参数由表 1 给出, 其中 $S_Q, \beta_Q, \|\bar{f}_i(t)\|_Q$ 分别为 Q 统计量对应的敏感系数、临界增益以及故障检测幅值; $S_T, \beta_T, \|\bar{f}_i(t)\|_T$ 分别为 T^2 统计量对应的敏感系数、临界增益以及故障检测幅值; r_f 为相对检测能力指标。

由表 1 的计算结果可知, 当传感器发生增益故障时, 若临界增益 β_Q 和 β_T 都趋近于 1, 则可认为无论增益如何变化(增大或减小), 两个检测统计量均会增大。当用某个统计量检测故障时, 通过故障检测幅值可以预先估计出传感器发生多大幅值的故障就一定能够被检测出来。另外, 利用相对检测能力指标可以比较 T^2 和 Q 统计量的故障检测能力, r_f 大于 1 说明 Q 统计量相对于 T^2 统计量更灵敏, 在故障幅值比较小的情况下就能检测出故障。

4.2 实验结果

首先给出传感器发生偏差故障时, T^2 和 Q 统计量具有不同的检测能力的比较。以给水流量传感器 F_1 的故障为例说明 PCA 的检测的结果。传感器的故障模拟用采集数据中的后 200 组数据, 设采样时刻为 20 时 F_1 发生幅值为 66 t/h 的恒偏差, 到 100 个采样时刻故障幅值增大到 300 t/h。图 2 给出了上述故障情况下的检测结果。采样时刻 20 到 100 的过程中, 由于 P_1 的故障幅值在 Q 统计量检测幅值附近, 因此 Q 统计量必然很快超限而检测出传感器故障; 根据表 1 知, T^2 的检测能力较差, 因此只有当故障幅值发展到较大程度时才能检测出故障。

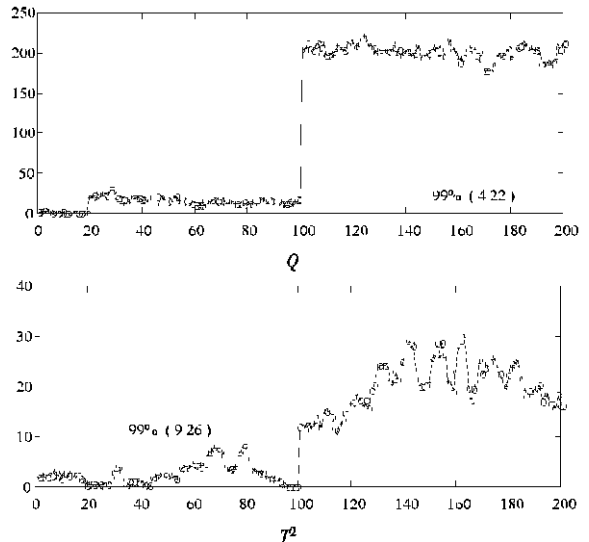


图 2 检测统计量在传感器恒偏差故障时的表现

下面验证传感器增益故障下临界增益值对于检

测行为的影响。由表 1 知, 除给水流量传感器 $F1$ 外, 其他传感器的故障临界增益都近似为 1; 对于传感器 $F1$ 可以判断, 当 $F1$ 的增益减小为 0.9 以下才有可能被 T^2 统计量检测出来, 减小为 0.95 以下才有可能被 Q 统计量检测出来。图 3 中给出了 $F1$ 在增益减小过程中的检测结果, 在采样时间为 50 到

100 之间 $F1$ 增益变为 0.95, 此时两个统计量都检测不出这一故障; 而采样时间为 100 后, $F1$ 的增益变为 0.9, 此时 Q 统计量检测出故障, 而 T^2 统计量仍不能检测出此故障。传感器卡死故障情况与增益故障类似, 此处不再赘述。

表 1 传感器故障检测的特性参数

传感器	u_i	σ_i	S_Q	S_T	β_Q	β_T	$\ \tilde{f}_i(t)\ _Q$	$\ \tilde{f}_i(t)\ _T$	r_f
R1	5 209.265	16.786	0.688	0.139	0.984	0.964	83.156	273.892	3.29
R2	5 165.320	16.189	0.517	0.205	0.982	0.964	92.518	217.818	2.35
P1	18 512	0.047	0.987	0.004	0.990	0.966	0.192	4.374	22.82
F1	972.046	14.660	0.809	0.067	0.952	0.892	66.981	343.545	5.13
P2	12 186	0.028	0.248	0.326	0.986	0.971	0.228	0.295	1.29
P3	16 700	0.069	0.750	0.128	0.981	0.960	0.324	1.159	3.58

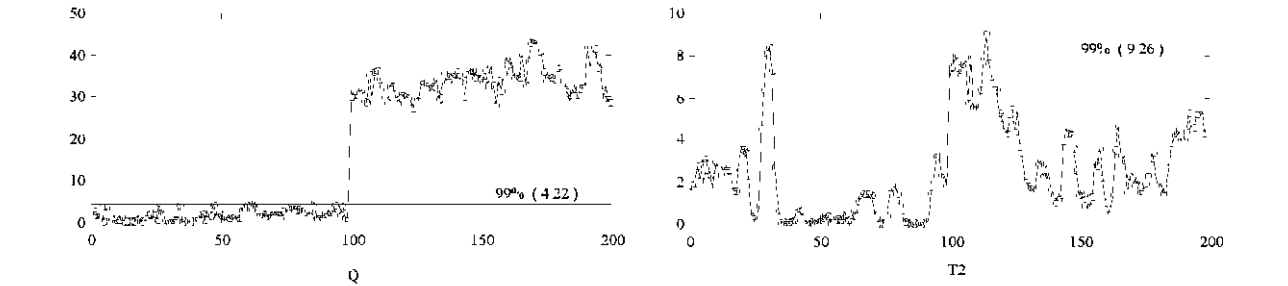


图 3 检测统计量在传感器增益故障时的表现

5 结 论

本文在假设只存在传感器故障的条件下, 系统分析了 PCA 检测方法不同传感器故障情况下检测行为, 给出了不同故障情况下检测统计量 T^2 和 Q 的变化规律以及两者检测能力的比较。同时检测特性参数的计算只依赖于 PCA 模型的参数, 不需要先验的知识, 因此利用本文的分析方法可以在故障检测之前就对 PCA 的检测行为定量地进行分析, 从而更深入的把握 PCA 的检测结果, 这对于进一步实现传感器故障诊断将起到积极作用。

参考文献:

[1] Dunia R, Qin S J. Joint diagnosis of process and sensor faults using principle component analysis[J]. Control Engineering Practice, 1998, 6: 457—469.
[2] Qin S J, Li W H. Detection and identification of faulty sensors with maximum sensitivity[C]. In: Proceedings of Ameri-

can Control Conference, California, 1999, 613—617.
[3] 陆宁云, 杨英华, 王福利 (Lu N Y, Yang Y H, Wang F L). 基于迭代主成分分析的过程监测方法的研究与实现 (Research and implementation of process monitoring based on iterative PCA)[J]. 控制与决策 (Control and Decision), 2002, 17(3): 215—218.
[4] Chiang L. H., Russell E. L., and Braatz R. D. Fault detection and diagnosis in industrial systems[M]. London: Springer, 2000.
[5] Jackson J. E. A User's guide to principle components[M]. Wiley, New York, 1991.
[6] 周东华, 叶银忠 (Zhou D H, Ye Y Z). 现代故障诊断与容错控制 (Modern fault diagnosis and fault-tolerant control)[M]. 北京: 清华大学出版社 (Tsinghua Univeristy Press), 2000.
[7] 王海清, 宋执环, 王慧 (Wang H Q, Song Z H, Wang H). PCA 过程监测方法的故障检测行为分析 (Fault detection behavior analysis of PCA-based performance monitoring approach)[J]. 化工学报 (J. of Chemical Industry and Engineering, China), 2002, 53(3): 297—301.