

基于马氏距离的航空装备故障预测研究

辛 龙, 周越文, 孔庆春, 杨 召

(空军工程大学 自动测试系统实验室, 西安 710038)

摘要: 故障状态类型的判别是航空装备故障预测系统的核心环节, 它直接影响到故障预测的准确性; 针对航空装备的故障状态类型判别问题, 提出一种基于马氏距离的故障预测方法; 首先介绍了马氏距离, 其次建立了状态数据库矩阵及状态判别模型, 并给出了基于马氏距离的故障预测流程; 最后将该方法用于某型飞机火控系统的故障预测中, 使得在线和离线的平均故障预测准确度分别达到 98.48% 和 97.77%, 表明马氏距离在航空装备的故障预测中有较好的应用和推广价值。

关键词: 马氏距离; 航空装备; 故障预测

Study on Aerial Equipment Failure Forecast Based on Mahalanobis Distance

Xin Long, Zhou Yuewen, Kong Qingchun, Yang Zhao

(Automatic Test System Lab., Air Force Engineering University, Xi'an 710038, China)

Abstract: Distinguishing failure states is a core technology in aerial equipment failure forecast system, and it influences the accuracy directly. To solve the problem of failure distinction on aerial equipments, a method of failure forecast based on Mahalanobis distance is brought out. Firstly, Mahalanobis distance was introduced. Then, the state database and distinguishing model were founded, and the failure forecast process was given. In the end, the method was applied to the failure forecast of a certain airplane fire control system. The result shows that the average precision of on-line forecast is 98.48%, while the one of off-line forecast is 97.77%. It can be apparently indicated that Mahalanobis distance has a high value in application and it deserves to be spread widely, especially in the area of aerial equipment failure forecast.

Keywords: Mahalanobis distance; aerial equipments; failure forecast

0 引言

航空装备故障预测^[1]是指基于特征预测信息, 对航空装备部件或系统进行准确的剩余寿命预测, 并对可能发生的故障作出预先诊断的方法。它是 PHM (Prognostics and Health Management) 技术在航空领域的重要应用^[2], 在提高装备的可利用率和可维护性、延长装备的寿命等方面具有很高的军事应用价值。航空装备故障预测系统的基本过程包括信息提取与处理、状态类型判别以及故障预测与诊断决策。其中, 状态类型的判别是故障预测的核心环节, 它直接影响预测的准确度, 常用的判别方法有时域分析法、模式识别法及统计分析法等。

马氏距离是统计学中两组样本集数据之间的协方差, 它代表了两个样本集之间的相似度^[3-5]。由于马氏距离可以消除量纲的影响, 并且综合考虑了各种特性之间的联系, 在状态类型的判别方面具有特征分析全面、准确度高等优势, 适合应用于多故障模式的故障预测研究中。

1 马氏距离

假设两个变量 X 和 Y 的取值分别为 x_1, x_2, \dots, x_n 和 y_1, y_2, \dots, y_n , 那么 X 和 Y 之间的样本协方差被定义为:

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

$$\text{其中: } \bar{X} = \sum_{i=1}^n X_i, \bar{Y} = \sum_{i=1}^n Y_i。$$

对于 P 个变量, 可以建立 $p \times p$ 维的协方差矩阵, 同时为了消除协方差受 X 和 Y 取值范围的约束, 通过标准化的方法可以得到 X 和 Y 间的样本相关系数:

$$\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\left(\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2 \right)^{\frac{1}{2}}}$$

由此可以得到两个 p 维测量值 x_i 和 y_j 之间的马氏距离, 定义如下:

$$d_{MH}(i, j) = ((x_i - y_j)' \sum^{-1} (x_i - y_j))^{\frac{1}{2}} \quad (1)$$

其中: $(x_i - y_j)'$ 表示矩阵 $(x_i - y_j)$ 的转置, \sum^{-1} 为相关系数矩阵。 $d_{MH}(i, j)$ 值越小, 表示测量值 x_i 与 y_j 之间的相似度越高。

2 基于马氏距离的故障预测模型

2.1 建立装备状态数据库矩阵

假设复杂装备系统的历史数据类有 n 组, 称之为样品, 并且每个样品测得指标个数为 p 个, 则原始数据可用一个属性矩阵表示:

$$\mathbf{X}_{TF} = \begin{matrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{matrix} \begin{bmatrix} x_{11} & x_{12} & \cdots \\ x_{21} & x_{22} & \cdots \\ \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots \end{bmatrix} \quad (2)$$

同时, 装备系统的历史检测数据中还存在有 m 组正常运行情况下的状态数据, 其属性矩阵为:

收稿日期:2014-02-10; 修回日期:2014-04-06。

作者简介:辛 龙(1988-),男,内蒙古包头人,工学硕士研究生,主要从事机载导弹控制与检测方向的研究。

周越文(1970-),男,陕西西安人,教授,研究生导师,主要从事航空装备检测自动化及智能化方向的研究。

$$\mathbf{Y}_T = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{bmatrix} \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & \cdots & y_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mp} \end{bmatrix} (j = 1, 2, \cdots, p) \quad (3)$$

在工程应用中, 影响系统状态的各指标的量纲并不是统一的, 为了消除量纲不统一给数据处理带来的不利影响, 还需要对上述数据进行标准化处理。首先, 利用式 (2) 中的正常状态数据求得各个指标测量值的均值:

$$\bar{y}_j = \frac{1}{m} \sum_{i=1}^m y_{ij} (j = 1, 2, \cdots, p) \quad (4)$$

对于故障状态下的数据, 就可以利用各指标测量值偏离正常状态的程度来刻画各故障的具体情况, 即:

$$x'_{ij} = \frac{x_{ij} - \bar{y}_j}{y_j} (i = 1, 2, \cdots, n; j = 1, 2, \cdots, p) \quad (5)$$

由于式 (5) 消除了各指标的量纲, 实现了量纲的归一化, 因此, 原始数据的状态矩阵式 (2) 可以表示成式 (6) 的形式:

$$\mathbf{X}'_{TF} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \begin{bmatrix} x'_{11} & x'_{12} & \cdots & x'_{1p} \\ x'_{21} & x'_{22} & \cdots & x'_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ x'_{n1} & x'_{n2} & \cdots & x'_{np} \end{bmatrix} \quad (6)$$

通过以上处理, 定义第 i 个样品和第 j 个样品间的相关系数为:

$$r_{ij} = \frac{\sum_{k=1}^p (x'_{ik} - \bar{x}'_i)(x'_{jk} - \bar{x}'_j)}{\sqrt{\sum_{k=1}^p (x'_{ik} - \bar{x}'_i)^2 \cdot \sum_{k=1}^p (x'_{jk} - \bar{x}'_j)^2}} \quad (7)$$

其中: $\bar{x}'_i = \frac{1}{p} \sum_{k=1}^p x'_{ik}$, $\bar{x}'_j = \frac{1}{p} \sum_{k=1}^p x'_{jk}$, $i = 1, 2, \cdots, n; j = 1, 2, \cdots, p$ 。

所以, 由式 (7) 得各样品之间的相关系数矩阵为:

$$\mathbf{R} = (r_{ij})_{n \times n} = \begin{bmatrix} 1 & & & \\ r_{21} & 1 & & \\ \vdots & \vdots & \cdots & \\ r_{n1} & r_{n2} & \cdots & 1 \end{bmatrix} \quad (8)$$

由于式 (8) 的元素满足 $-1 \leq r_{ij} \leq 1$, 因此可以根据相关系数矩阵 \mathbf{R} 对 n 组观测值进行分类: r_{ij} 越接近 1, 则两种状态越相似。再根据专家经验和历史故障数据分析确定出阈值 δ 后, 对装备的运行状态进行分类, 得到 m 类装备的状态数据矩阵^[6-7]:

$$\mathbf{X}''_{TF} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_m \end{bmatrix} \begin{bmatrix} x_1 & x'_2 & \cdots & x'_p \\ x''_{11} & x''_{12} & \cdots & x''_{1p} \\ x''_{21} & x''_{22} & \cdots & x''_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ x''_{m1} & x''_{m2} & \cdots & x''_{mp} \end{bmatrix} \quad (9)$$

2.2 建立故障规则知识库

通过对历史故障数据库的分析, 可以确定出复杂装备的大部分故障模式, 如表 1 所示。表中, 故障状态下的数据均由故障数据的平均值经归一化处理得到。因此, 我们可以分别获得系统在正常状态下和故障状态下的属性矩阵, 可以看作是一个案例库或故障规则知识库:

$$\mathbf{X}_{TF} = \begin{bmatrix} \bar{X}_0 \\ \bar{X}_1 \\ \vdots \\ \bar{X}_n \end{bmatrix} \begin{bmatrix} 0 & 0 & \cdots & 0 \\ x_{11} & x_{12} & \cdots & x_{1p} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \quad (10)$$

其中, \bar{X}_0 中的元素 0 表示系统正常运行情况下测量值偏离正常状态的程度为 0。

2.3 状态判别模型的建立与算法的实现

基于故障规则知识库进行故障预测时, 对于任意一次的在线检测或历史记录数据, 若它与第 i 类故障模式的重心距离最近, 则认为它最有可能发生第 i 类故障, 需要先修正这类故障, 如果没有故障就去修正与检测值距离次近的第 j 类故障, 以此类推。显然, 要预测可能发生的故障并给予及时维护, 关键在于判别检测值 $\hat{X} = (\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_n)'$ 最有可能属于的故障类别。为此, 根据马氏距离的形式, 定义复杂装备系统的状态趋势预测值与故障模式重心 $\bar{X}_i = (\bar{x}_{i1}, \bar{x}_{i2}, \cdots, \bar{x}_{in})'$ 之间的马氏距离为:

$$d_{MH}(\hat{X}, \bar{X}_i) = ((\hat{X} - \bar{X}_i)' \sum^{-1} (\hat{X} - \bar{X}_i))^{\frac{1}{2}} \quad (11)$$

在实际工程应用时, 每一个系统检测值通过归一化处理, 分别计算与复杂系统故障规则库中 n 类故障模式重心的马氏距离, 然后进行大小排序, 给出系统运行的分析报告。系统维护时先检测复杂系统是否存在 $d_{MH}(\hat{X}, \bar{X}_i)$ 最小的故障类, 再依次从小到大检测, 如果存在可能的故障就进行维护处理。

在复杂装备的故障诊断和预测过程中, 在线故障预测与离线故障预测的处理方式略有不同。在线故障预测仅需测量一次实时数据, 处理起来较简单, 预测流程如图 1 所示。离线故障预测由于要处理系统运行与设备检测时记录的大量数据, 处理起来相对复杂, 预测流程如图 2 所示。

在线故障预测算法的实现流程如下所述。

第 1 步: 利用统计学相关知识、专用设备检测诊断和专家分析, 对系统历史数据及检测数据进行处理得到初始故障规则知识库 $\{\bar{X}_0, \bar{X}_1, \bar{X}_2, \cdots, \bar{X}_n\}$, 以矩阵的形式预存到计算机程序中;

第 2 步: 利用式 (5) 对系统在线检测值进行归一化处理, 消除系统状态量量纲不统一的影响;

第 3 步: 根据式 (11) 计算在线检测值和各类故障重心的马氏距离, 如果 $d_{MH}(\hat{X}, \bar{X}_0) = 0$, 说明系统运行完全正常; 如果

表 1 故障信息表

故障现象	故障模式	故障状态数据	解决方案
故障现象 1	故障模式 1	对应故障状态下的数据 $\bar{X}_1: \bar{x}_{11}, \bar{x}_{12}, \cdots, \bar{x}_{1p}$	对应解决方案 1
	故障模式 2	对应故障状态下的数据 $\bar{X}_2: \bar{x}_{21}, \bar{x}_{22}, \cdots, \bar{x}_{2p}$	对应解决方案 2
故障现象 2	故障模式 3	对应故障状态下的数据 $\bar{X}_3: \bar{x}_{31}, \bar{x}_{32}, \cdots, \bar{x}_{3p}$	对应解决方案 3
	故障模式 4	对应故障状态下的数据 $\bar{X}_4: \bar{x}_{41}, \bar{x}_{42}, \cdots, \bar{x}_{4p}$	对应解决方案 4
...

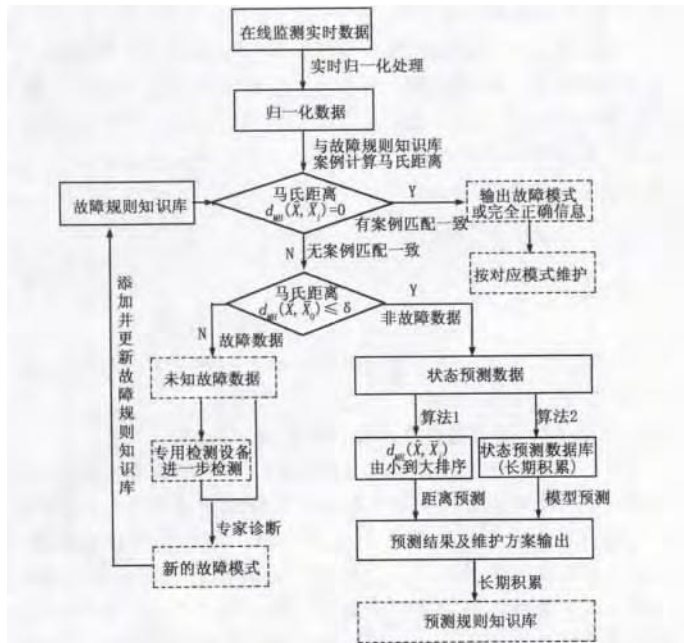


图1 在线故障预测流程图

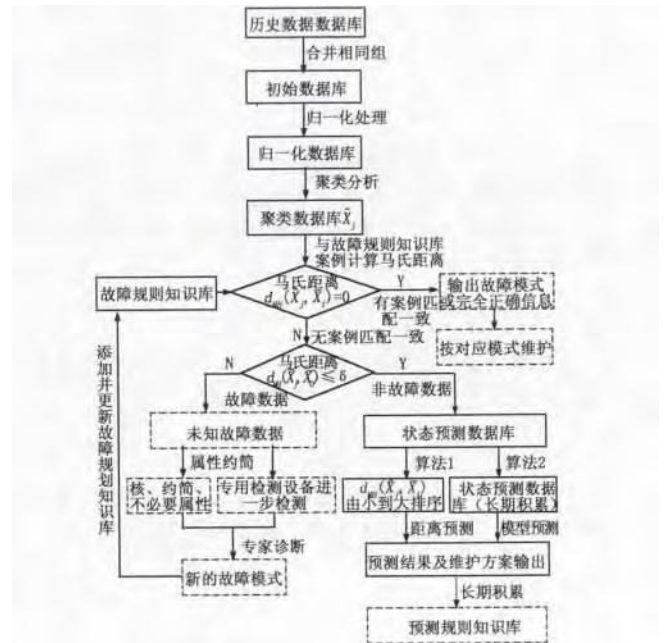


图2 离线故障预测流程图

$d_{MH}(\hat{X}, \bar{X}_i) = 0 (i = 1, 2, \dots, n)$, 说明系统存在第 i 类故障, 应按对应模式对系统进行排查维护; 如果马氏距离以上两条均不满足, 进入下一步;

第 4 步: 如果 $d_{MH}(\hat{X}, \bar{X}_0) \neq 0$, 且 $d_{MH}(\hat{X}, \bar{X}_0) \leq \delta$, 则认为系统虽然运行正常, 但此时系统性能已发生异常变化, 系统可能在未来的某一时刻发生故障, 进入第 5 步; 如果 $d_{MH}(\hat{X}, \bar{X}_0) > \delta$, 说明系统存在未知模式故障, 需要利用专用检测设备进行详细的检测确认, 得到新的故障模式并充实故障规则知识库;

第 5 步: ① 对 $d_{MH}(\hat{X}, \bar{X}_i)$ (其中 $i = 1, 2, \dots, n$) 进行从小到大排序, 给出分析报告 (包括预测结果、维护方案等), 先检测系统是否存在马氏距离最小的故障模式, 再从小到大依次检测维护; ② 把归一化处理后的在线检测值添加到长期积累的状态预测数据库中, 利用预测模型对系统的状态进行预测, 给出预测结果和维护方案。

离线故障预测算法的实现流程如下所述。

第 1 步: 对历史数据库中的所有检测值 X_i 进行初步处理, 合并相同组得到初始数据库 (2);

第 2 步: 利用式 (5) 对合并后的数据进行归一化处理, 得到归一化数据库 (6);

第 3 步: 利用式 (7) 计算归一化数据库 (6) 各类状态之间的相关系数, 并进行聚类处理得到 m 类装备状态的数据库 (9);

第 4 步: 根据公式 (11) 分别计算聚类数据库中各类检测值 \hat{X}_j 和故障规则知识库中各类故障重心的马氏距离, 如果 $d_{MH}(\hat{X}_j, \bar{X}_0) = 0$, 说明系统运行完全正常; 如果 $d_{MH}(\hat{X}_j, \bar{X}_i) = 0 (i = 1, 2, \dots, n)$, 说明系统存在第 i 类故障, 应按对应模式对系统进行排查维护; 如果马氏距离以上两条均不满足, 进入下一步;

第 5 步: 如果 $d_{MH}(\hat{X}_j, \bar{X}_0) \neq 0$, 且 $d_{MH}(\hat{X}_j, \bar{X}_0) \leq \delta$, 则认为系统虽然正常运行, 但此时系统性能已发生异常变化, 系统可能在未来的某一时刻发生故障, 进入下一步; 如果

$d_{MH}(\hat{X}_j, \bar{X}_0) > \delta$, 说明系统存在未知模式故障, 需要根据状态数据确认故障原因, 或利用专用检测设备进行详细的检测确认, 从而得到新的故障模式并充实故障规则知识库;

第 6 步 ① 对 $d_{MH}(\hat{X}_j, \bar{X}_i)$ (其中 $i = 1, 2, \dots, n$) 进行从小到大排序, 给出分析报告 (包括预测结果、维护方案等), 先检测系统是否存在马氏距离最小的故障模式, 再从小到大依次检测维护; ② 把归一化处理后的在线检测值添加到长期积累的状态预测数据库中, 利用预测模型对系统的状态进行预测, 给出预测结果和维护方案。

3 某型飞机火控系统故障预测应用

以某型飞机火控系统为研究对象, 在真实工作环境的情况下, 经人工故障设置 (故障注入) 后, 采用基于马氏距离的故障预测方法进行 500 次故障预测实验。注入的故障模式有 9 种: X_1 —天线故障、 X_2 —高频接收机故障、 X_3 —低频接收机故障、 X_4 —主控振荡器故障、 X_5 —发射机故障、 X_6 —同步器故障、 X_7 —变换部件故障、 X_8 —电源故障、 X_9 —计算机故障。

3.1 在线故障预测

将系统运行时测得的实时数据利用式 (5) 归一化处理后, 再根据式 (11) 计算在线检测值和各类故障重心的马氏距离 $d_{MH}(\hat{X}_j, \bar{X}_i)$, 给出预测结果; 实时观测系统的运行状态, 并记录下各故障的故障模式;

3.2 离线故障预测

对历史数据库中的所有检测值经合并相同组、归一化处理及聚类处理后, 得到矩阵形式的装备状态的数据库 (9), 再根据公式 (11) 计算数据库矩阵 (9) 中各类检测值和故障重心的马氏距离 $d_{MH}(\hat{X}_j, \bar{X}_i)$, 给出预测结果。

最后, 整合实验数据, 将每一故障模式发生的预测次数与实测次数进行比较, 并分别计算出故障预测准确度, 实验结果如表 2 所示。

(下转第 2058 页)

