

10.7690/bgzdh.2015.05.007

用于飞行动作评估的飞参数数据预处理方法

周胜明, 赵育良, 张玉叶, 王淑娟
(海军航空工程学院青岛校区, 山东 青岛 266041)

摘要:为了预处理得到正确反映飞机飞行状态的飞行参数记录数据,介绍一种有效的用于飞行动作评估的飞参数预处理方法。利用主成分分析相关性分析法提取与飞行动作最相关的几个参数数据,将各飞行参数的相关性融入多维自回归数据的预测估计过程,并对某型飞机的部分飞行参数数据进行实验分析,利用多维时间序列建模对其中5个参数的缺失数据进行一步预报得到估计值,与真实值进行比较得到的相对误差值都比较小(平均值0.03),该结果表明:该方法能得到正确反映和重现飞机飞行动作的时间序列数据,可推广到任意一种多元时间序列来表示目标状态的情况。

关键词:飞行动作; 飞行参数数据; 主成分分析; 多维时间序列建模

中图分类号: TP206 文献标志码: A

The Preprocess Method of the Flight Parameter Data Applied to the Evaluation of Acrobatic Maneuver

Zhou Shengming, Zhao Yuliang, Zhang Yuye, Wang Shujuan
(Qingdao Branch, Navy Aeronautical & Astronautical University, Qingdao 266041, China)

Abstract: In order to get the record data of flight parameters correctly reflect the aircraft flight conditions by the means of preprocessing pretreatment, introduce an effective pretreatment method of flight parameter data for acrobatic maneuver evaluation. The several parameters which has higher relevance with acrobatic maneuver was extracted their data using the principal component analysis. And the correlation of the parameters was imported into the multidimensional auto-regression prediction data estimation process. We used a part of a certain aircraft flight parameters data for experimental analysis. In the experiment, the estimated data of 5 parameters' missing data by one step forecasting of multivariate time series modeling were compared with the true data, and the relative errors were all small that the average value was of 0.03. The results show that this method can get the time series data which can reflect or reproduce the acrobatic maneuver correctly. And the method can be extended to any kind of multivariate time series to represent the status of objectives.

Keywords: acrobatic maneuver; flight parameters data; principal component analysis; multivariate time series modeling

0 引言

机载飞参系统在飞机飞行过程中记录各种飞行参数信息,这为评估飞行员操作飞机的水平提供了重要依据。目前评估飞行员对各种飞行动作的完成情况,主要通过对某一个或2个飞行参数值的变化范围进行检测,如果变化范围没有超出指定飞行动作规定的参数值要求,那么操作合格。然而实际操作中,因为飞机的飞行动作取决于若干参数,比如爬升动作,它取决于飞行高度、速度、俯仰角、倾斜角和俯仰角速度等参数;而巡航动作,取决于飞行高度、速度、横滚角和倾斜角等参数。所以要对飞行员操作的飞行动作进行正确评估,需要综合考虑多个参数的信息。而飞参数数据中包含十几个与飞行动作有关的参数,若将所有参数用于每个动作的飞行评估,复杂度太高。考虑到不同的参数与不同飞行动作的相关度不同,笔者采用相关性分析法提取与飞行动作最相关的参数,用于飞行动作评估。

由于飞参记录仪是复杂的电子系统,在其工作采集、记录数据时,会因为噪音干扰、电子干扰等因素,产生一些数据丢失、数据失真和记录的非正常开始或结束等异常情况。而这些异常的飞参数数据,将严重影响飞行员操作水平的评估,所以在评估之前,要保证飞参数数据的完整、真实可靠,也就是要对失真数据及缺失数据进行估计替换^[1-4]。用于飞行动作评估的参数之间有一定的相关性,如果将这些参数的相关性加入到对失真及缺失数据的估计过程,则估计结果将更加准确。因此,笔者利用主成分分析法提取出与飞行动作相关性最大的几项;然后对这几项参数数据进行多维自回归建模,结合参数相关度权重定阶,得到飞参数数据模型表达式;进而对其中的失真数据及缺失数据进行估计预测。

1 主成分分析法选择参数

飞参数数据具有参数维数高、数据量大的特点,

收稿日期: 2014-12-21; 修回日期: 2015-02-06

作者简介: 周胜明(1962—),男,山东人,硕士,教授,从事飞行器仿真相关研究。

而用于飞行动作评估时仅需要与飞行动作相关性比较大的参数。主成分分析法^[5]是将多个变量通过线性变换以选出较少个数重要变量的一种多元统计分析方法，可以有效解决这个问题。

实现的基本思想：将多个具有一定相关性的参数，重新组合成一组新的互无关的综合指标来代替原来的参数。设要分析的数据集涉及 n 个参数，分别用 x_1, x_2, \dots, x_n 表示，这 n 个参数构成 n 维随机向量 $X=(x_1, x_2, \dots, x_n)'$ 。对 X 进行线性变换，可以形成新的综合变量，用 Y 表示：

$$\begin{aligned} Y_1 &= u_{11}x_1 + u_{21}x_2 + \dots + u_{n1}x_n \\ Y_2 &= u_{12}x_2 + u_{22}x_2 + \dots + u_{n2}x_n \\ &\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ Y_n &= u_{1n}x_1 + u_{2n}x_2 + \dots + u_{nn}x_n \end{aligned}$$

其中： u_{in} 表示参量 x_n 上的线性变换系数，也称作主成分在参量上的载荷。由于可以任意地对原始变量进行上述线性变换，由不同的线性变换得到的综合变量 Y 的统计特性也不尽相同；并且为了取得较好的结果，总是希望 $Y_i = u_i X$ 的方差尽可能大且各 Y_i 之间互相独立。因此这里将线性变换约束在下面的原则之下：

- 1) $u_i^T u_i = 1$ ，即 $u_{11}^2 + u_{21}^2 + \dots + u_{n1}^2 = 1$ ，($i = 1, 2, \dots, n$)；
- 2) 在所有的线性组合中选取 Y_1 应该是方差最大的，称为第一主成分；为了有效反映原来的信息， Y_1 已有的信息就不再出现在 Y_2 中，也就是要求 $Cov(Y_1, Y_2) = 0$ ， Y_i 与 Y_j 相互无关 ($i \neq j; i, j = 1, 2, \dots, n$)，称 Y_2 是第二主成分，依此类推，可以构造出 n 个主成分。

最后，通过计算主成分的累积贡献率来确定主成分的个数，该方法需要进一步从其中找到反映飞行动作信息的主成分 Y_d ，从该主成分的线性变换表达式 $Y_d = u_{1d}x_1 + u_{2d}x_2 + \dots + u_{nd}x_n$ 中，比较各参量上载荷 u_{id} 的大小，选择与飞行动作相关性最大的 m 个参数，从而降低数据处理的维数。

2 飞参数据的多维 AR 模型建模

通过主成分分析法降维得到了与飞行动作高度相关的参数数据，要用于飞行动作评估，首先要对这些数据中存在的失真值或者缺失值进行估计替换预处理。因为飞行参数是时间序列数据^[6-7]，具有连续性和平稳性，所以可以用时序建模分析方法完成数据的建模预测。虽然可能只有某项参数数据缺失，

但是如果仅对单项参数的时间序列数据建模，估计得到的数值准确度不高。因为各参数之间都有一定的相关性，如果综合多个相关参数值进行建模分析，能够提高估计的准确度。常用时序模型有自回归(AR)模型、滑动平均(MA)模型以及自回归滑动(ARMA)模型。后两类模型都可用高阶 AR 模型^[8]来逼近，其逼近程度取决于自回归模型的阶次。由于 AR 模型的参数估计为线性回归过程，其计算简单、速度快，实际物理系统也往往是全极点系统，因此本方法采用多维 AR 模型对飞参数据建模来估计数据。

用 $X_t = (x_{1t}, x_{2t}, \dots, x_{mt})^T$ 来表示某架次 m 个飞行参数的时间序列，则 X_t 的 m 维 AR 模型一般形式为

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \dots + a_p X_{t-p} + \varepsilon_t \quad (1)$$

其中： p 是模型的阶数； a_1, a_2, \dots, a_m 是一个 $m \times m$ 阶待估参数矩阵； $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{mt})^T$ 是 m 维白噪声向量，也称 ε_t 为模型 (1) 的残差，它的元素均值为 0，方差为一个固定值。

1) 数据的平稳化及正态化处理。

首先分析飞参序列数据的平稳性，对数据进行平稳化处理并进行正态化检验。对数据进行一次差分以后得到平稳序列 $\{Y_t\}$ ， $\{Y_t\} = \{X_t - X_{t-1}\}$ 。经过差分处理后的时间序列的自相关系数有显著下降趋势，说明经过差分变换的序列已符合 ARMA 建模的条件。然后对 $\{Y_t\}$ 进行标准化正态处理： $X_t = (Y_t - u_y)/\sigma_y$ 。式中， u_y 为序列 $\{X_t\}$ 的均值， σ_y 为 $\{X_t\}$ 的均方差。进而对 $\{X_t\}$ 进行多维 AR 建模。

2) 多维 AR 模型的参数估计。

m 维的 AR(p) 模型可以表示为：

$$\begin{cases} X_{p+1}(i) = \Phi_{p+1} A(i) + \varepsilon_{p+1}(i) \\ \vdots \\ X_N(i) = \Phi_N A(i) + \varepsilon_N(i) \end{cases}, (i = 1, 2, \dots, m)$$

式中： $\Phi_i = [X_{t-1}^T, X_{t-2}^T, \dots, X_{t-p}^T]^T$ ； $A(i) = [a_1(i), a_2(i), \dots, a_p(i)]^T$ 。设 N 为最终的观测时刻，令

$$\mathbf{L}_N(i) = \begin{bmatrix} x_{p+1}(i) \\ \vdots \\ x_N(i) \end{bmatrix}, \mathbf{B}_N = \begin{bmatrix} \Phi_{p+1} \\ \vdots \\ \Phi_N \end{bmatrix}, \mathbf{E}_N = \begin{bmatrix} \varepsilon_{p+1}(i) \\ \vdots \\ \varepsilon_N(i) \end{bmatrix}$$

则式 (1) 可以转化为

$$\mathbf{L}_N(i) = \mathbf{B}_N \mathbf{A}(i) + \mathbf{E}_N(i), (i = 1, 2, \dots, m)$$

当残差平方和达到极小时，即

$$J = \sum_{t=p+1}^N \varepsilon_t^T \varepsilon_t = \sum_{t=p+1}^N \left(x_t - \sum_{i=1}^p a_i x_{t-i} \right)^T \left(x_t - \sum_{i=1}^p a_i x_{t-i} \right) = \min$$

必然有

$$J(i) = \sum_{t=p+1}^N |E_N(i)|^T |E_N(i)| = \min, (i=1, 2, \dots, m).$$

这样基于时刻 N 的所有观测值的 $A(i)$ 的最小二乘估计值为

$$\hat{A}_N(i) = (B_N^T B_N)^{-1} B_N^T L_N(i), (i=1, 2, \dots, m).$$

3) AIC 准则定阶。

模型 (1) 的阶数 p 由最小 (akaike's information criterion, AIC) 准则确定。设模型的一步预报残差 $\varepsilon_t = X_t - \hat{X}_t$, $\delta_E^2(p)$ 是拟合残差方差, 认为它是模型阶数 p 的函数, 定义 AIC 准则函数为

$$AIC(p) = \lg \delta_E^2(p) + 2p / N, p \in (0, 1, \dots, n)$$

因为 ε_t 是一个多维向量, 对它计算 $\delta_E^2(p)$, 可以结合参数相关度权重定阶。由主成分分析得到的 m 个参数, 根据在主成分线性表达式中的载荷, 取绝对值并归一化计算权重, m 个参数的权重分别为 $|C|_1, |C|_2, \dots, |C|_m$, 那么

$$\delta_E^2(p) = \frac{\delta_1^2}{|C|_1} + \frac{\delta_2^2}{|C|_2} + \dots + \frac{\delta_m^2}{|C|_m}$$

式中: δ_i^2 表示 m 维参数时间序列

$X_t = (x_{1t}, x_{2t}, \dots, x_{mt})^T$ 中 x_{it} 的拟合残差方差。

对准则函数 $AIC(p)$, 真实模型的阶数 p_0 应

$COEFF =$

x'_1 : 高度'	-0.490 3	-0.084 4	0.177 1	-0.298 7	0.158 7	-0.180 7	0.173 0	0.737 6
x'_2 : 速度'	-0.451 6	-0.183 2	0.294 3	-0.137 1	0.233 1	0.505 1	-0.534 9	-0.248 3
x'_3 : 航向'	-0.040 1	-0.069 8	-0.776 1	-0.559 7	-0.158 8	0.145 1	-0.174 4	0.035 7
x'_4 : 倾斜角'	0.494 8	0.135 3	0.001 0	0.157 9	0.103 6	0.591 5	-0.164 6	0.569 4
x'_5 : 倾斜角'	-0.415 5	0.326 1	-0.009 5	0.164 7	-0.579 0	0.459 2	0.383 6	-0.022 8
x'_6 : 盘旋角速度'	-0.055 9	-0.639 9	-0.062 6	0.396 8	-0.516 4	-0.120 4	-0.310 7	0.219 9
x'_7 : 俯仰角速度'	-0.318 2	-0.067 8	-0.517 5	0.574 1	0.518 6	0.072 3	0.150 3	0.008 3
x'_8 : 横滚角速度'	0.179 4	-0.644 8	0.089 2	-0.199 1	0.117 1	0.332 7	0.600 1	-0.141 0
Explained	= 45.791 0	24.721 6	16.853 7	6.402 5	4.133 4	1.698 1	0.313 8	0.086 0

表 1 参数相关关系矩阵

参数变量	气压高度	指示空速	真航向	俯仰角	倾斜角	盘旋角速度	俯仰角速度	横滚角速度
气压高度	1.000 0	0.929 2	-0.028 5	-0.942 2	0.624 3	0.107 1	0.397 5	-0.162 9
指示空速	0.929 2	1.000 0	-0.176 9	-0.828 5	0.535 7	0.227 2	0.348 2	0.010 3
真航向	-0.028 5	-0.176 9	1.000 0	-0.130 6	0.016 5	0.074 5	0.406 6	0.024 2
俯仰角	-0.942 2	-0.828 5	-0.130 6	1.000 0	-0.637 3	-0.265 7	-0.526 2	0.164 4
倾斜角	0.624 3	0.535 7	0.016 5	-0.637 3	1.000 0	-0.205 0	0.402 3	-0.702 7
盘旋角速度	0.107 1	0.227 2	0.074 5	-0.265 7	-0.205 0	1.000 0	0.220 4	0.701 0
俯仰角速度	0.397 5	0.348 2	0.406 6	-0.526 2	0.402 3	0.220 4	1.000 0	-0.217 8
横滚角速度	-0.162 9	0.010 3	0.024 2	0.164 4	-0.702 7	0.701 0	-0.217 8	1.000 0

这样得到第一主成分表达式

$$Y_1 = -0.4903x_1 - 0.4516x_2 - 0.0401x_3 + 0.4948x_4 - 0.4155x_5 - 0.0559x_6 - 0.3182x_7 + 0.1794x_8$$

该主成分是飞行动作的综合成分, 比较它在每个标准化变量上的载荷大小可以发现, 该动作与高度、速度、俯仰角、倾斜角和俯仰角速度相关性最大, 由此提取这5个参数数据进行动作的评估。同时对5个参数的载荷值, 取绝对值并归一化, 计算得到权重向量:

$$(|C_1|, |C_2|, |C_3|, |C_4|, |C_5|) = (0.226, 0.208, 0.228, 0.191, 0.147)$$

2) 多维AR模型建模。

对600条包含5个参数的飞参数数据中的前500条进行多维AR建模。首先对数据进行一次差分及标准化正态处理以后得到的平稳时间序列, 然后利用最小二乘法计算得到系数矩阵。利用AIC准则结合参数相关度权重定阶, 得到AIC函数曲线图(如图1), 从图中判断阶数取3时AIC函数值最小。从而得到五维AR(3)模型的3阶系数矩阵如表2所示。

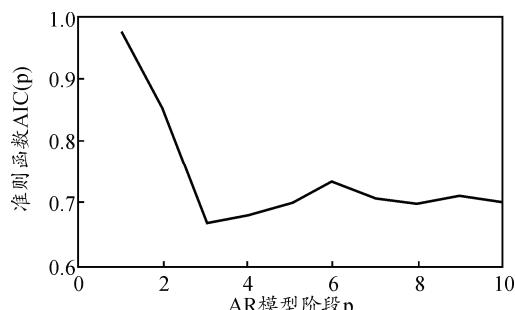


图1 AIC 函数曲线

表2 模型参数估计

5维AR(3)模型3阶系数矩阵	1	2	3	4	5
A(1)	0.2782	0.3333	-7.4444	1.3264	-57.3333
	0.0803	-0.6042	5.8889	-0.1840	-5.3333
	0.0250	-0.1433	0.8344	0.0406	4.2733
	0.5046	0.5169	-24.7903	-1.2447	0.5777
	0.0007	0.0384	-0.4390	-0.0021	0.7902
A(2)	-0.8264	0.3438	0.7639	-0.7361	38.2222
	-0.0972	0.4688	-0.5278	-0.1528	-12.4444
	-0.0330	0.0478	-0.8401	-0.0864	-3.1022
	0.6814	-0.5304	5.9629	0.2738	-17.2396
	0.0137	-0.0165	0.0658	-0.0165	0.7024
A(3)	0.1094	-0.7708	2.1111	1.5432	-58.6667
	0.1719	0.7292	0.2778	0.2867	25.3333
	0.0048	-0.0311	0.0433	0.0366	3.4667
	-0.0342	0.2819	-4.7482	-0.7031	33.7899
	0.0014	0.0439	0	0.0070	-0.1756

计算模型的残差 $\varepsilon_t = X_t - \hat{X}_t$ 是否满足白噪声序列的要求, 通过计算发现 $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{5t})^\top$ 中各子向量都符合均值为零的正态分布, 所以AR(3)参数模型是合适的。由 $Y_t = \sigma_y X_t + \mu_y$, 逆变换得到序

列 $\{X_t\}$, 这样可以进行未知数据的估计。

3) 数据的预测估计验证。

利用所得模型进行500条以上飞参数数据的估计预测, 将预测值与真实数据比较。表3是利用模型分别对5个参数的缺失数据进行一步预报得到的估计值与真实值的相对误差比较, 可以看到相对误差值都比较小, 说明该方法可以用于缺失或失真数据的估计预处理。

表3 数据预测估计结果

参数	真实数据	估计值	误差值	相对误差
飞行速度	6 859	6 865	6	0.0008
高度	600	598	2	0.003
俯仰角	6.59	6.57	0.02	0.003
倾斜角	-2.02	-2.16	0.48	0.069
俯仰角速度	0.4309	0.3979	0.033	0.076

4 结论

为了预处理得到正确反映飞机飞行状态的飞行参数记录数据, 笔者首先利用相关性分析法提取与飞行动作最相关的几个参数数据, 然后为了估计其中的缺失数据, 将各飞行参数的相关性融入多维自回归数据的预测估计过程, 从实验的预测结果看, 该方法是可行的。由于飞参数数据是一种典型的多元时间序列数据, 所以虽然该方法针对用于飞行动作评估的飞行参数数据进行预处理, 但是可以推广到任意一种多元时间序列来表示目标状态的情况。

参考文献:

- [1] 魏燕明, 端木京顺, 高海龙. KPLS-SVM在缺失飞参数估计中的应用[J]. 火力指挥与控制, 2009, 34(11): 186-189.
- [2] 倪世宏, 史忠科, 谢川, 等. 飞行事故调查时缺失飞行参数的综合估计方法[J]. 计算机工程与应用, 2004, 40(32): 206-208.
- [3] 谢川, 倪世宏, 张忠麟. 一种缺失飞行参数预处理的新方法[J]. 计算机仿真, 2005, 22(4): 27-31.
- [4] 吴建刚, 陈志伟, 李曙林, 等. 飞参记录数据计算机处理的有关问题研究[J]. 计算机仿真, 2007, 24(2): 18-21.
- [5] 李丹丹, 路辉, 郎荣玲. 基于主成分分析和支持向量机的飞参阶段划分研究[J]. 现代电子技术, 2010, 33(8): 134-137.
- [6] Brockwell P J, Davis R A. Time Series: Theory and Methods[M]. New York: Springer-Verlag, 1991: 61-62.
- [7] Zeigler Bernard P, Praehofer Herbert, Tag Gon Kim. Theory of modeling and simulation: integrating discrete event and continuous complex dynamic systems[M]. London: Academic Press, 2000: 137-139.
- [8] 李慧, 吴国富, 齐全跃. 舰船运动姿态的多维自回归建模[J]. 系统仿真学报, 2003, 15(5): 617-619.