

# 基于均值迭代法的稳态数据处理方法

马昌\*, 李密, 薛文鹏

中国飞行试验研究院 发动机所, 陕西 西安, 710089

**摘要:** 参数不确定度的分析及合成是动力装置性能试飞中最为关键的环节之一, 而参数名义值的获取以及剔除异常数据是不确定度分析的基础。本文介绍了一种基于均值迭代的稳态数据处理算法, 详细分析了该方法的工作原理、名义值求取步骤。算例计算表明, 本方法对数据样本的概率分布无要求, 适用性强, 避免了一些传统算法的复杂计算, 而且可以实时地评估稳定段数据的抖动程度, 适合于参数样本分布未知的航空发动机关键截面参数的数据预处理。

**关键词:** 均值迭代; 数据关联; 稳态飞行数据

中图分类号: V233.7 文献标识码: A 文章编号: 1007-5453 (2016) 07-0048-05

在动力装置性能试飞科目中, 需要得到发动机稳定状态的关键参数, 但飞行中的稳态数据采集, 会因各种不确定的原因出现一些异常值<sup>[1]</sup>, 而异常值产生的原因很多, 如瞬时或者间断性的测量系统故障或者是电磁干扰、机械冲击等<sup>[2,3]</sup>。所以数据预处理的一个重要组成部分就是剔除异常值, 目前, 在工程中广泛应用的异常值剔除方法基本为莱茵达准则、格拉布斯准则、狄克逊准则、肖维勒准则、AEDC法<sup>[4]</sup>。但这些方法大都需要预判测量数据的统计特性<sup>[5]</sup>, 如样本分布规律等, 且对样本点数量有一定要求, 如当测量数据样本为均匀分布时, 莱茵达准则和 AEDC 法就会失效<sup>[6]</sup>。所以, 在飞行试验中的测量参数的统计特性事先无法得知的情况下, 这些方法的使用就受到了局限<sup>[7,8]</sup>。因此, 本文提出了基于均值迭代的稳态数据处理算法, 不需要预估测量参数样本的概率分布, 避免了一些传统算法的复杂计算过程, 适用于飞行试验中的数据预处理, 以某型发动机飞行试验数据为例进行分析, 并与其他计算方法进行对比和验证, 表明本文介绍的数据预处理方法有效、可行, 满足发动机数据分析的精度需求。

## 1 计算方法

本文提出的均值迭代法不需要预估测量参数样本的概率分布, 对数据样本点的个数也没有限制, 能有效地处理稳态数据, 提高数据的真实性。在介绍均值迭代法之前, 首先

介绍一下在工程中广泛应用的莱茵达准则和 AEDC 法, 以便在实际应用中对几种方法的处理结果进行对比<sup>[9]</sup>。

### (1) 莱茵达准则

莱茵达准则通常又叫三西格玛准则, 其表达式如下:

$$|x_i - \bar{x}| > 3\sigma \quad (1)$$

式中:  $x_i$  为第  $i$  个测量数据 (应用莱茵达准则样本数应当大于 10),  $\bar{x}$  为该参数的平均值,  $\sigma$  为该参数的标准差。即当某个数据  $x_i$  满足式 (1) 时, 便认为该数据为异常值, 应当剔除<sup>[10]</sup>。

### (2) AEDC 法

AEDC 法是阿诺德工程研发中心在莱茵达准则的基础上发展的一种适用于各种样本的剔除异常值的方法<sup>[11]</sup>, 与莱茵达准则不同的是, 其表达式如下:

$$|x_i - \bar{x}| > C\sigma \quad (2)$$

式中:  $C$  是基于工程判断得到理想数据的期望区间的系数, 其公式如下:

$$C = \frac{-1.6819236 + 1.6386898N - 0.00721312N^2}{1.0 + 0.59286772N - 0.00355709N^2} \quad (3)$$

式中:  $N$  为样本数。如果  $N$  小于 65, 通过该式 (3) 计算  $C$ , 如果样本数大于等于 65, 则  $C$  等于 3, 此时, 与莱茵达准则相同<sup>[12]</sup>。

收稿日期: 2016-03-09; 录用日期: 2016-05-26

\* 通讯作者. Tel.: 13572438326 E-mail: machang1980@163.com

引用格式: MA Chang, LI Mi, XUE Wenpeng. Method of steady state data processing based on mean iteration [J]. Aeronautical Science & Technology, 2016, 27 (07): 48-52. 马昌, 李密, 薛文鹏. 基于均值迭代法的稳态数据处理方法 [J]. 航空科学技术, 2016, 27 (07): 48-52.

这2种数据预处理方法的前提是假设测量参数样本满足正态分布,所以当发动机截面参数等关键参数的分布为其他分布,此方法就可能存在较大误差<sup>[13]</sup>。

(3) 均值迭代法

均值迭代法采用平均值迭代更新测量数据中的最大和最小值,从而使异常值在每一次的迭代过程中对整个稳态数据的影响降低,直到收敛于某一精度。只要迭代次数足够,就可以将数据中的异常值“剔出”,如果无法收敛,就可以根据迭代次数和迭代曲线对测量参数进行评估,判断此稳定段数据是否存在不可接受的抖动。流程如图1所示,具体算法步骤如下:

- (a) 计算测量参数样本的算术平均值或者几何平均值  $X_{mean1}$ ,同时找出测量参数样本的最大值  $X_{max}$ ,令  $X_{mean1}$  替换  $X_{max}$ 。
- (b) 计算更新后的测量参数样本的算术平均值  $X_{mean2}$ ,同时找出测量参数样本的最小值  $X_{min}$ ,令  $X_{mean2}$  替换  $X_{min}$ 。
- (c) 再次计算测量参数样本的平均值  $X_{mean3}$ ,并求出测量参数样本的方差  $e$ ,若方差  $e$  不满足要求,则进行迭代转到 (a) 步,否则迭代终止。

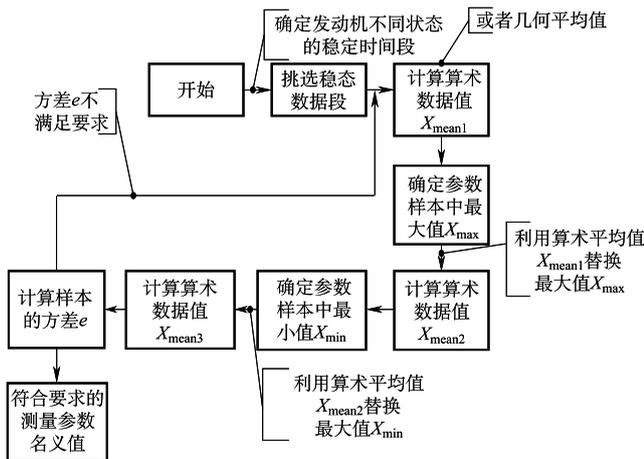


图1 均值迭代法流程简图  
Fig.1 The diagram of mean iteration method

2 算例及分析

以某型发动机地面台架试验数据为例,编写C语言程序<sup>[14]</sup>,使用均值迭代法对不同的稳态数据段进行处理。如图2~图7为发动机地面台架试验数据中台架推力、发动机物理转速、燃油流量、涡轮后总温、涡轮后总压等参数的迭代收敛曲线,台架推力、涡轮后总压、压气机可调叶片角度都在较少的迭代次数下收敛(收敛条件为  $1 \times 10^{-6}$ ),如图2、图4、图6所示。而涡轮后总温、相对物理转速、

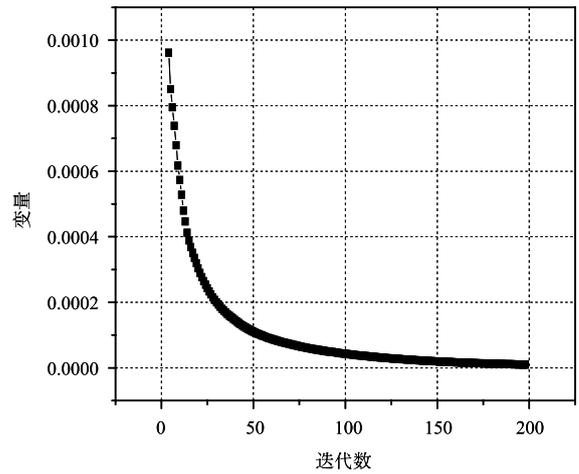


图2 迭代收敛曲线(台架推力)  
Fig.2 Convergence curve (test bed)

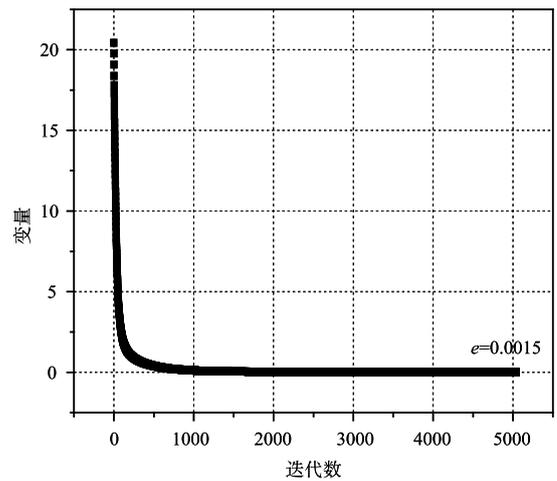


图3 迭代未收敛曲线(涡轮后总温)  
Fig.3 Non convergence curve (temperature after turbine)

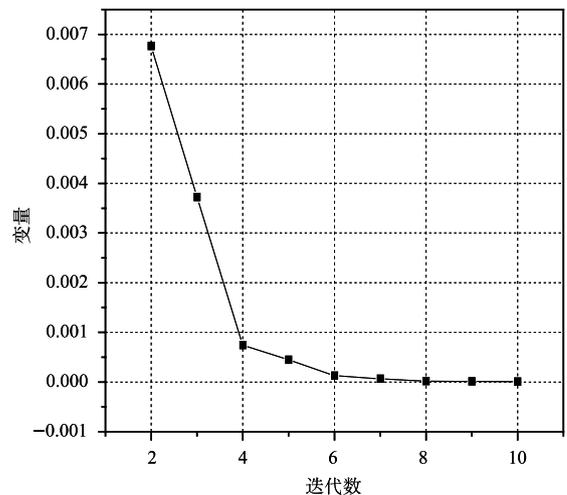


图4 迭代收敛曲线(涡轮后总压)  
Fig.4 Convergence curve (total pressure after turbine)

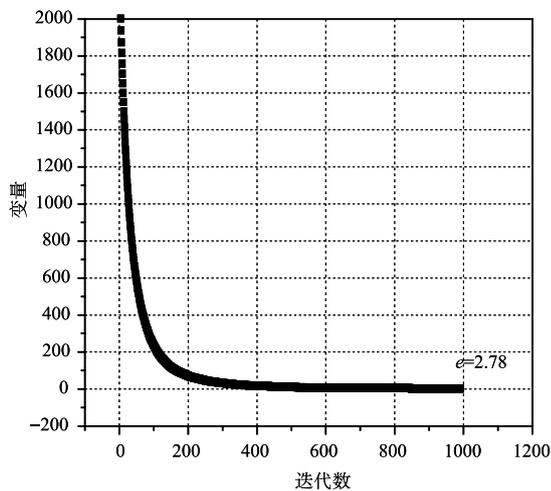


图5 迭代未收敛曲线(相对物理转速)

Fig.5 Non convergence curve (relative speed)

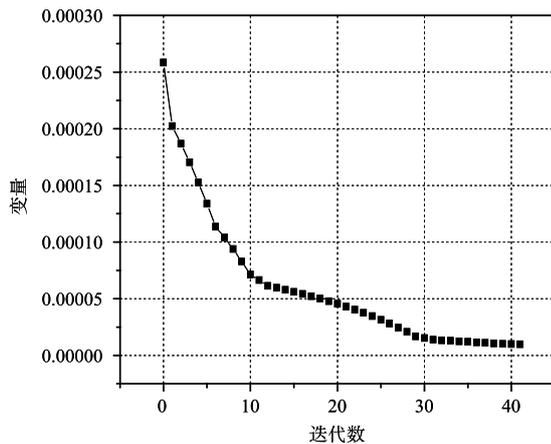


图6 迭代收敛曲线(压气机可调导叶角度)

Fig.6 Convergence curve (IGV)

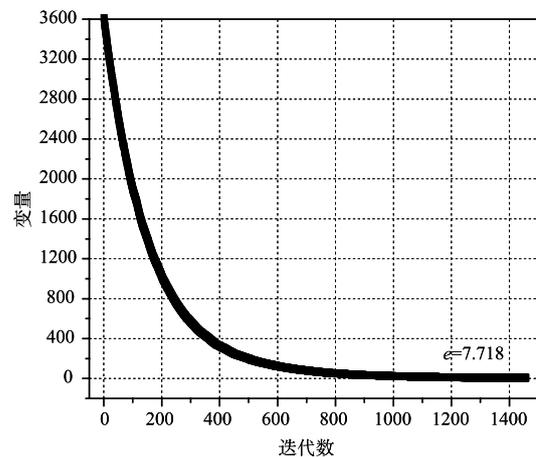


图7 迭代未收敛曲线(燃油流量)

Fig.7 Non convergence curve (fuel flow)

燃油流量,最终迭代不收敛,如图3、图5、图7所示,即可判断此测量参数的稳定段存在“抖动”,若使用AEDC法进行剔点,将会剔除真实的试验数据,会降低数据处理的精度。

虽然燃油流量等测量参数的样本方差未收敛于规定值,但整个收敛曲线随迭代次数的增加而趋于固定值,即可看出均值迭代法并未改变参数样本的分布情况。

由于AEDC法是经过NASA大量试验验证的有效方法,根据发动机测量参数的数据预处理的经验,相对物理转速、台架实测推力、可调导叶角度的样本分布满足正态分布,满足AEDC法的使用条件<sup>[15]</sup>。所以,分别采用均值迭代法和AEDC法对其进行预处理并进行对比,结果如表1所示。可以看出,2种方法的预处理结果接近。可见,本文提出的方法满足工程上与处理阶段对异常值剔除的要求。

表1 两种方法对稳态数据段的预处理结果比较

Table 1 Comparison of two methods for the pretreatment of steady state data

可调导叶角度			相对物理转速			台架推力		
均值迭代法	莱茵达准则	相对误差 /%	均值迭代法	莱茵达准则	相对误差 /%	均值迭代法	莱茵达准则	相对误差 /%
0.99703	0.99521	0.18327	0.72913	0.72858	0.07455	0.20043	0.19888	0.78037
0.68344	0.68292	0.07562	0.80372	0.80346	0.03258	0.31705	0.31529	0.56012
0.4726	0.47189	0.14929	0.85462	0.85403	0.06899	0.4432	0.43879	1.00491
0.26236	0.26221	0.05692	0.90328	0.90315	0.01366	0.61537	0.61134	0.65978
0.09512	0.0941	1.07854	0.94287	0.94257	0.03195	0.78464	0.77958	0.64958
0.05981	0.05931	0.84046	0.99994	0.99989	0.00479	0.95505	0.94808	0.73507

注: 稳态段测量参数的名义值均为相对值

### 3 结论

航空发动机中一些关键参数接近于正态分布,即使用如AEDC法的传统稳态数据预处理方法完全可以满足精度要求,但当测量参数样本不遵循正态分布而为均匀分布时,

AEDC方法就不能使用。本文提出的均值迭代法不需要对数据的统计特性进行预判,人为假设较少,避免了因为样本分布不满足预处理方法要求而产生的误差。均值迭代法存在以下优点:

(1) 对测量参数样本分布无要求,即使用于各种分布,通用性强。

(2) 对样本数无要求,如莱茵达准则和 AEDC 方法均要求样本数大于 65,而均值迭代法可以使用于任何数量的样本个数。

(3) 可以根据收敛曲线评估稳态数据段的稳定程度。

(4) 避免了传统算法的复杂计算过程,计算速度快,可用于实时数据处理。

AST

### 参考文献

- [1] Abernethy R B, Thompson J W. Handbook, uncertainty in gas turbine measurements[Z].1973.
- [2] 叶川, 伍川辉, 张嘉怡. 计量测试中异常数据剔除方法比较[J]. 计量与测试技术, 2007, 34 (7): 26-27.  
YE Chuan, WU Chuanhui, ZHANG Jiayi. Comparision about how to get rid of abnormal data in metrology & measurement[J]. Metrology & Measurement Technique, 2007, 34 (7) : 26-27. (in Chinese)
- [3] Negrao J, Fanton R , Fuente R , et al. In-flight thrust determination[J]. SAE Technical Paper, 1998.
- [4] Angiulli F Basta S, Pizzuti C.Distance-based detection and prediction of outliers[J]. IEEE Transactions on Knowledge and Data Engineering , 2006, 18 (02) : 145-160.
- [5] Jiang S Y, Li Q H, Li KL , et al. A new approach formining local outlier [C]// Proceedings of the second international conference on machine learning and cybernetics, 2003.
- [6] 孙传友,孙晓斌,汉泽西,等. 测控系统原理与设计[M]. 北京: 北京航空航天大学出版社, 2002.  
SUN Chuanyou, SUN Xiaobin, HAN Zexi, et al.Principle and design of measurement and control system[M].Beijing: Beihang University Press, 2002. (in Chinese)
- [7] 盛骤,谢式千,潘承毅. 概率论与数理统计[M]. 北京: 高等教育出版社, 1989.  
SHENG Zhou, XIE Shiqian, PAN Chengyi. Probability theory and mathematical statistics[M]. Beijing: Higher Education Press, 1989. (in Chinese)
- [8] 耿素军,余剑. 智能测量系统中粗大误差的处理[J]. 电气电子教学学报, 2005, 27 (3) : 37-39.  
GENG Sujun, YU Jian. Grosserror processing in intelligent measuring system[J]. Journal of Electrical & Electronic Education, 2005, 27 (3) : 37-39. (in Chinese)
- [9] 余剑. 高精度智能测量系统中粗大误差的处理技术[J]. 测试技术学报, 2003, 17 (3) : 258-261.  
YU Jian. Gross error processing technology in intelligence measuring system with high precision[J].Journal of Test and Measurement Technology, 2003, 17 (3) : 258-261. (in Chinese)
- [10] 吴天鹏. 对粗大误差判别的理论探讨[J]. 黄石高等专科学校学报, 1995 (2) : 62-66.  
WU Tianpeng. A theoretical study on the discrimination of gross errors [J]. Journal of Huangshi Polytechnic College, 1995 (2) : 62-66. (in Chinese)
- [11] 肖艳军, 李建勋. 抗野值多速率模型及交互式状态估计[J]. 上海交通大学学报, 2005, 39 (9) : 1106-1112.  
XIAO Yanjun , LI Jianxun. An outlier rejecting multirate model for interacting state estimation[J].Journal of Shanghai Jiaotong University, 2005, 39 (9) : 1106-1112. (in Chinese)
- [12] 王正明,易东云. 测量数据建模与参数估计[M]. 长沙: 国防科技大学出版社, 1996.  
WANG Zhengming, YI Dongyun. Measurement data modeling and parameter estimation[M]. Changsha: National University of Defense Technology Press, 1996. (in Chinese)
- [13] 张文博,李凯,朱尤攀,等. 光电稳定跟踪平台中微机电陀螺滤波方法研究[J]. 红外技术, 2006, 28 (5) : 249 - 252.  
ZHANG Wenbo, LI Kai, ZHU Youpan , et al. Research on the denoising methods of MEMS gyro used in servo system with optic sensors[J]. Infrared Technology, 2006, 28 (5) : 249 - 252. (in Chinese)
- [14] 沈一鹰, 冉启文, 刘永坦. 改进的格拉布斯准则在信号检测门限估值中的应用[J]. 哈尔滨工业大学学报, 1999, 31 (3) : 111-113.  
SHEN Yiying, RAN Qiwen, LIU Yongtan. Application of improved gurbbs' criterion to estimation of signal detection threshold[J]. Journal of Harbin Institute of Technology, 1999 , 31 (3) : 111-113. (in Chinese)
- [15] 谭浩强. C 程序设计[M]. 第 2 版. 北京: 清华大学出版社, 2000.  
TAN Haoqiang .C program design [M]. 2nd ed. Beijing: Tsinghua University Publication, 2000. (in Chinese)

**作者简介**

马昌(1980—)男,硕士,工程师。主要研究方向:发动机整机试验技术研究。

Tel: 13572438326

E-mail: machang1980@163.com

李密(1986—)男,硕士,工程师。主要研究方向:发动机推力测量计算技术。

薛文鹏(1987—)男,硕士,工程师。主要研究方向:航空发动机建模与控制系统仿真。

## Method of Steady State Data Processing Based on Mean Iteration

MA Chang\* , LI Mi, XUE Wenpeng

*China Flight Test Establishment, Xi'an 710089, China*

**Abstract:** The analysis and synthesis of parameter uncertainty is one of the most important procedures in the performance test of power plant, for which the acquisition of the nominal value of parameters and the elimination of abnormal data are basis. In this paper, a method of steady state data processing based on mean iteration was introduced include a detailed analysis of the working principle of the method and the steps of obtaining the nominal value. Calculation shows that, the method has strong applicability because it not only has no requirement for the probability distribution of data sample, but also avoids the complex calculation in some traditional algorithms. Furthermore, data jitter in stable period could be evaluated real-time so that it could be more suitable for critical aeroengine section parameter data preprocessing in which the parameters distribution is unknown.

**Key Words:** mean iteration; data correlation; steady state flight test data;

---

Received: 2016-03-09; Accepted: 2016-05-26

\*Corresponding author. Tel. : 13572438326 E-mail: machang1980@163.com