

文章编号: 1000-8055(2002)02-0196-05

发动机飞行任务剖面的主成份聚类法

宋迎东, 高德平

(南京航空航天大学 能源与动力学院, 江苏 南京 210016)

摘要: 本文提出了利用主成份分析对航空发动机飞行任务剖面进行分类的方法, 并对某战斗机发动机 18 个飞行任务剖面进行了聚类分析。选取了飞行高度、飞行马赫数、发动机转速以及发动机重心法向过载等 4 个参数作为分类的原始依据参数。对上述 4 个参数进行主成份分析, 得到 4 个独立的主成份, 其中第一、二主成份的累积贡献率可达 81.1%。因此, 可以根据主平面内各飞行任务剖面的第一、二主成分的分布情况直观地进行定性地分类。最后, 本文利用重心法进行了定量的聚类, 得到了分类的树状图。研究结果表明本文提出的方法是合理可行的。

关键词: 飞行任务剖面; 聚类分析; 主成份; 航空发动机

中图分类号: V 231.91

文献标识码: A

The Principal Component Analysis Method for Engine Flight Mission Profiles Categorization

SONG Ying-dong, GAO De-ping

(Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: In this paper, the principal component analysis method is proposed for engine flight mission profiles categorization. Eighteen flight mission profiles of a fighter engine are categorized. The categorization is based on 4 parameters, namely, flight height, flight Mach number, engine rotational speed and overload on the engine center of gravity. 4 principal components of the mentioned 4 parameters are obtained, the accumulated contribution ratio of the first two principal components arrives at 81.1%. The qualitative categorization can be done from the first and the second principal components distribution range in the principal plane. Finally, a tree-shaped category graph is acquired by quantitative categorization using center of gravity method. The results show that the proposed method is feasible.

Key words: flight mission profile; categorization; principal component; aeroengine

1 引言

如何对发动机飞行任务剖面进行分类是迄今未很好地解决的问题。文献[1]采用剖面参数范围相似的原则进行分类, 是 1 种经验的方法, 在工程应用中较难操作, 随意性较大。原因在于发动机飞行任务剖面分类所依据的原始参数超过 2 个, 是

1 个多维空间内的分类问题。文献[2]提出采用模糊聚类的分析方法, 选取了飞行高度、飞行马赫数和发动机转速等 3 个参数进行分类的方法。然而, 对于航空发动机来说, 除了飞行高度、飞行马赫数以及发动机工作转速影响其应力与寿命之外, 还有由机动飞行引起的外部作用力, 这在国军标^[3]和发动机结构完整性大纲^[4]上有均明确的规定。

收稿日期: 2001-04-02; 修订日期: 2001-07-11

作者简介: 宋迎东(1969-), 男, 安徽太湖人, 南京航空航天大学能源与动力学院副教授, 博士

外部作用力对机匣、主轴以及安装节等零件的寿命具有十分重要的影响,若不充分地考虑可能引起发动机严重的故障。因此,飞行任务剖面的分类不仅应当选取飞行高度、飞行马赫数、发动机转速等 3 个参数,还应该选取反映外部作用力的主要参数——发动机重心法向过载,以这 4 个参数分类将给出更加合理的结果。本文提出飞行任务剖面的主成份分析方法,并由此进行剖面的分类。

2 发动机飞行任务剖面主成份分析原理

飞行剖面的分类就是要将飞行高度、飞行速度、发动机转速以及发动机重心处过载范围相近的剖面进行归类,这是一个四维空间内对飞行数据进行分类的问题,无法直观地进行观察,必须降维处理。主成份分析是一种把多个指标化为少数几个相互独立的综合指标的一种多元统计方法^[5,6],综合指标就是主成份,主成份通常为各单指标的线性组合,主成份之间是相互独立的。根据主成份分析原理,发动机飞行任务剖面主成份分析可采用以下方法。

设有 n_f 种飞行任务剖面,每种飞行任务剖面采样点数为 $m_k(k=1,2,\cdots,n_f)$,则总的采样点数为 n 个 $(n=\sum_{k=1}^{n_f}m_k)$,每一个样本点有 p 个飞行参数,所得到的观测值为 $x_{ij}(i=1,2,\cdots,n;j=1,2,\cdots,p)$,则原始飞行参数数据矩阵为:

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix}$$

飞行任务剖面主成份分析的步骤为:

(1) 原始飞行任务剖面参数的标准化,即:

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma}$$

式中: $\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}$ 为第 $j(j=1,2,\cdots,p)$ 个剖面参数的均值,即剖面参数矩阵 X 第 j 列数据的平均值; $\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}$ 为第 $j(j=1,2,\cdots,p)$ 个剖面参数即剖面参数矩阵 X 第 j 列数据的标准差。

(2) 根据标准化的剖面参数矩阵,计算剖面参数之间的相关系数矩阵。

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ r_{p1} & r_{p2} & \cdots & r_{pp} \end{bmatrix}$$

(3) 计算相关系数矩阵的特征值 $\lambda_j(j=1,2,\cdots,p;\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p)$ 和相应特征向量。

(4) 计算飞行任务剖面的主成份 $y_j(j=1,2,\cdots,p)$ 。计算公式为: $y_j = \sum_{k=1}^p u_{jk}x_{ik} \quad (i=1,2,\cdots,n)$ 其中: u_{jk} 为特征值 λ_j 对应的特征向量。 $y_1, y_2, \cdots, y_p$ 分别称为第一、第二、 $\cdots$ 、第 p 主成份。

(5) 计算主成份的贡献率和累积贡献率。

主成份的贡献率为: $g_k = \lambda_k / \sum_{j=1}^p \lambda_j = \frac{\lambda_k}{p}$
($k=1,2,\cdots,p$)

主成份的累积贡献率为:

$$G_k = \sum_{j=1}^k \lambda_j / \sum_{j=1}^p \lambda_j = \sum_{j=1}^k \lambda_j / p \quad (k=1,2,\cdots,p)$$

根据计算结果,一般选取累积贡献率大于 80% 的 k 个主成份进行对原始飞行任务剖面参数的分析即可以得到较满意的结果^[5,6]。

3 某发动机飞行任务剖面的主成份分析与聚类

3.1 飞行任务剖面的主成份分析

测量了某战斗机发动机 18 个飞行任务中的飞行高度 H 、飞行马赫数 Ma 、发动机转速 N 和发动机中心处的法向过载系数 NY 等 4 个参数,总的采样点数为 38410 个。对原始数据进行标准化以后,计算其相关系数矩阵,见表 1。可见这 4 个飞行参数均具有一定的相关性,其中飞行高度

表 1 飞行任务剖面参数的相关系数矩阵

	H	Ma	N	NY
H	1.00000	0.83376	0.35512	-0.00598
Ma	0.83376	1.00000	0.48786	0.14580
N	0.35512	0.48786	1.00000	0.24494
NY	-0.00598	0.14580	0.24494	1.00000

表 2 特征值、贡献率、累积贡献率

	特征值	贡献率/%	累积贡献率/%
λ_1	2.18103	54.525	54.525
λ_2	1.06298	26.575	81.100
λ_3	0.61096	15.274	96.374
λ_4	0.14503	3.626	100

和飞行马赫数相关性较大, 相关系数为 0.83376。

表 2 的特征值以及累积贡献率表明, 第一、第二主成份的累积贡献率达 81.1%, 因此取第一、第二主成份作为分类的依据, 能够以 80% 以上的精度代表原 4 个剖面参数。

表 3 第一、二特征值对应的特征向量

标准化飞行参数	λ_1	λ_2
飞行高度 H	0.58795	- 0.34403
飞行马赫数 M_a	0.63223	- 0.14928
发动机转速 N	0.47385	0.30583
法向过载 N_Y	0.17337	0.87512

表 3 为第一、第二特征值对应的特征向量, 这样第一、第二主成份为:

$$y_1 = 0.58795H' + 0.63223M_a' + 0.47385N' + 0.17337N_Y' \quad (1)$$

$$y_2 = - 0.34403H' - 0.14928M_a' + 0.30583N' + 0.87512N_Y' \quad (2)$$

可见, 第一主成份主要反映飞行高度、飞行速度和发动机转速的变化, 而第二主成份则反映了发动机法向飞行过载的变化。

3.2 飞行任务剖面的聚类

根据上述分析结果可见, 通过主成份分析将

该发动机的在 4 维空间的飞行任务剖面分类问题转化为二维主平面内的分类问题, 大大地简化了分类的难度, 可以直观地在主平面内看出各剖面之间的差别, 从而进行初步的分类, 如图 1(a)~(d) 分别为第 1, 4, 5, 16 飞行任务剖面在主平面上的分布范围。可见第 1, 4 飞行任务剖面具有大致相同的范围, 可作为一类, 而第 5, 16 飞行任务剖面则与它们具有很大的不同。为了科学地进行分类, 必须在第一、二主平面内进行定量的分析。在多元统计分析中定量聚类方法很多, 如最短距离法、最长距离法、中间距离法、重心法以及离差平方和法等等。每种方法有不同的特点, 通常聚类结果也不完全相同^[6]。重心法可以考虑一类中含多样本点的情况, 因此, 本文选取重心法进行聚类。

将 18 个飞行任务剖面的测量参数代入公式 (1) 和 (2), 计算出每一个飞行任务剖面的第一、第二主成份序列。然后计算出每一个剖面的重心 $(\bar{P}_{i1}, \bar{P}_{i2})$, $(i = 1, 2, \cdots, 18)$ 。主平面内重心之间的距离定义为: $D_{ij} = \sqrt{(\bar{P}_{i1} - \bar{P}_{j1})^2 + (\bar{P}_{i2} - \bar{P}_{j2})^2}$, $(i, j = 1, 2, \cdots, 18)$ 。

表 4 为飞行任务剖面间距离矩阵, 因为是对称矩阵, 所以只列出一半数字。利用重心法逐步分类的步骤为:

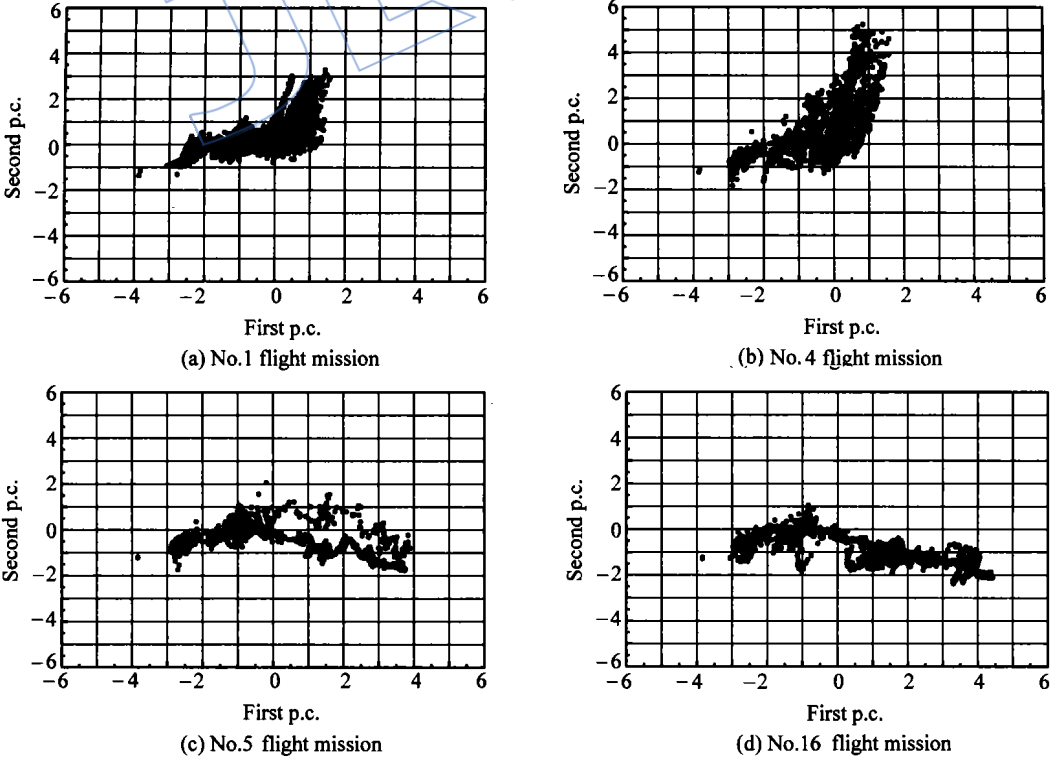


图 1 飞行任务剖面在第一、二主平面内的分布

Fig 1 Flight mission profiles distribution in the first two principal components plane

表 4 飞行剖面间的距离矩阵

剖面号	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0.120	0.560	0.072	1.155	0.655	1.877	0.964	0.278	0.255	0.353	0.312	1.081	0.395	1.099	1.140	3.284	0.480
2		0.679	0.185	1.051	0.555	1.764	0.863	0.304	0.161	0.289	0.365	0.972	0.486	1.180	1.037	3.163	0.391
3			0.494	1.630	1.139	2.381	1.433	0.576	0.775	0.865	0.613	1.574	0.331	0.757	1.611	3.837	0.953
4				1.193	0.692	1.923	0.999	0.247	0.298	0.423	0.338	1.122	0.325	1.031	1.176	3.346	0.511
5					0.501	0.761	0.197	1.055	0.900	1.131	1.393	0.109	1.318	1.747	0.026	2.343	0.684
6						1.242	0.309	0.571	0.400	0.684	0.908	0.435	0.841	1.369	0.485	2.744	0.186
7							0.955	1.809	1.626	1.774	2.073	0.807	2.076	2.491	0.784	1.643	1.428
8								0.858	0.709	0.966	1.212	0.165	1.122	1.572	0.178	2.514	0.489
9									0.275	0.593	0.584	1.002	0.272	0.901	1.035	3.302	0.387
10										0.389	0.525	0.826	0.524	1.175	0.885	3.064	0.230
11											0.322	1.036	0.748	1.452	1.123	3.059	0.575
12												1.306	0.612	1.298	1.381	3.381	0.754
13													1.270	1.736	0.115	2.350	0.621
14														0.706	1.298	3.573	0.659
15															1.722	4.086	1.222
16																2.368	0.667
17																	2.918

表 5 飞行任务剖面间一次并类距离矩阵

类号	2	3	4	6	7	8	9	10	11	12	13	14	15	17	18	19
1	0.120	0.560	0.072	0.655	1.877	0.964	0.278	0.255	0.353	0.312	1.081	0.395	1.099	3.284	0.480	1.140
2		0.679	0.185	0.555	1.764	0.863	0.304	0.161	0.289	0.365	0.972	0.486	1.180	3.163	0.391	1.037
3			0.494	1.139	2.381	1.433	0.576	0.775	0.865	0.613	1.574	0.331	0.757	3.837	0.953	1.611
4				0.692	1.923	0.999	0.247	0.298	0.423	0.338	1.122	0.325	1.031	3.346	0.511	1.176
6					1.242	0.309	0.571	0.400	0.684	0.908	0.435	0.841	1.369	2.744	0.186	0.485
7						0.955	1.809	1.626	1.774	2.073	0.807	2.076	2.491	1.643	1.428	0.761
8							0.858	0.709	0.966	1.212	0.165	1.122	1.572	2.514	0.489	0.178
9								0.275	0.593	0.584	1.002	0.272	0.901	3.302	0.387	1.035
10									0.389	0.525	0.826	0.524	1.175	3.064	0.230	0.885
11										0.322	1.036	0.748	1.452	3.059	0.575	1.123
12											1.306	0.612	1.298	3.381	0.754	1.381
13												1.270	1.736	2.350	0.621	0.109
14													0.706	3.573	0.659	1.298
15														4.086	1.222	1.722
17															2.918	2.343
18																0.667

(1) 首先找出距离最小的 2 个剖面, 根据表 4, 第 5、16 号之间的距离为 0.026, 是最小值。将这两个剖面合成一类, 称为第 19 类。

(2) 计算并类后的距离。只需计算 19 类和其余飞行剖面之间的距离。譬如, 剖面 2 到 19 的距离取剖面 2 到剖面 5 和剖面 16 之间的最小者, 按照表 4 剖面 2 到 19 的距离 $D_{2,19}=1.037$ 。未合并的剖面之间的距离不变, 将计算结果列入表 5 中,

称为一次并类距离矩阵。再在一次并类距离矩阵找出距离最近的 2 个剖面类合并为第 20 类。依次类推, 一直合并到全部剖面完成为止。图 2 为 18 个飞行任务剖面的的树状分类图, 根据该图就可以按照不同的分类需要进行剖面的粗分和细分。

图 3 给出了第 1、4、5 和 16 飞行任务的发动机转速剖面, 它们很大程度上反映了各自任务中发动机的使用载荷情况。可见第 1 和 4 剖面较相

似, 转速变化很剧烈, 从而形成较多的载荷循环,

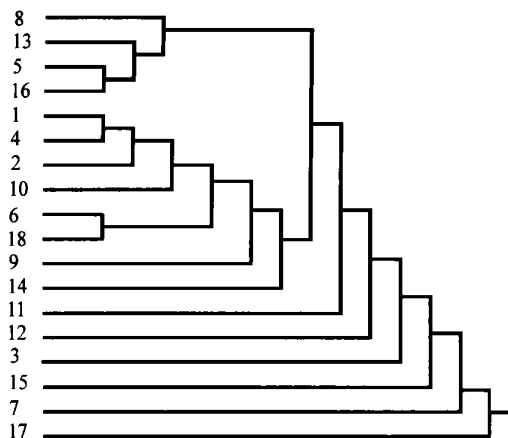


图 2 18 个飞行任务剖面分类的树状图

Fig 2 Tree graph of eighteen flight mission profiles categorization

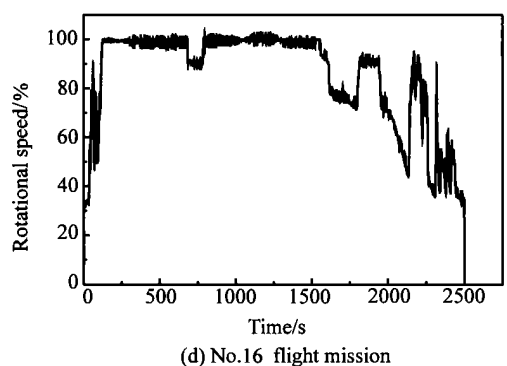
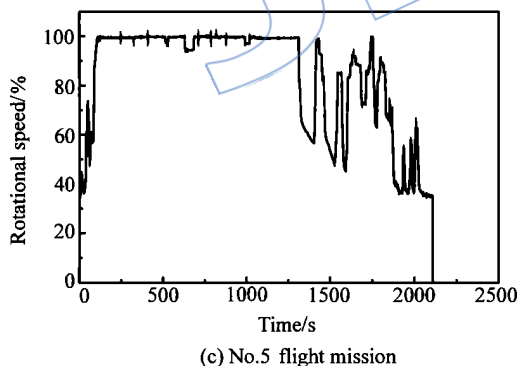
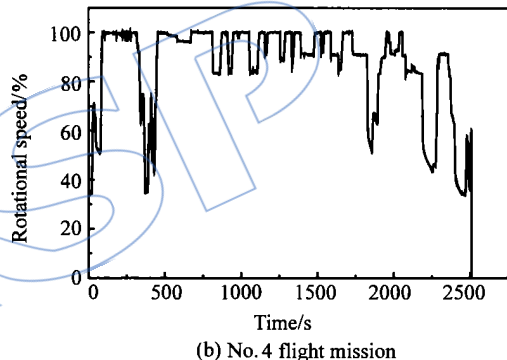
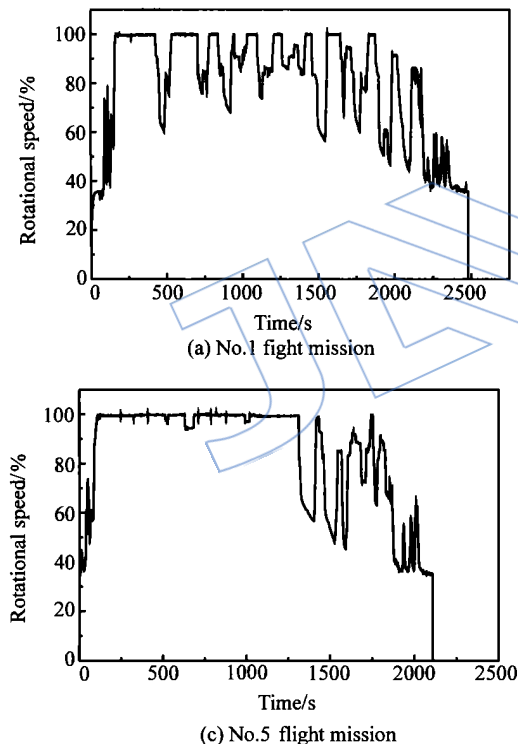


图 3 发动机转速剖面

Fig 3 Engine rotational speed profiles

的方法是合理可行的。

参考文献

- [1] 航空工业部发动机管理局 WP-6 发动机寿命研究报告[R]. 1988
- [2] 程德金 模糊数学在飞行试验数据处理中的应用[J]. 飞行力学, 1994, 12(2):.

与 5 和 16 剖面相差较大, 而 5 与 16 非常接近, 这与图 1 的直观观察结果和图 2 的定量分析结果是一致的, 说明了上述定量分类的合理性。

4 结 论

以飞行高度、飞行马赫数、发动机转速以及发动机重心法向过载等 4 个参数作为某发动机飞行任务分类的原始依据参数, 利用多元统计中的主成份分析方法, 求出其 4 个主成份, 结果表明: (1) 飞行高度和飞行速度具有最大的相关性, 相关系数达 0.83376。(2) 第一、二主成份能以 81.1% 的精度代表 4 个原始参数, 可以在第一、二主平面内进行飞行任务剖面的分类研究。(3) 利用重心法在第一、二主平面内进行逐步聚类, 得到聚类结果的树状图。(4) 采用 4 个剖面参数进行分类更加全面地反映了发动机的飞行载荷状况, 本文提出

- [3] GJB241-87. 航空涡轮喷气与涡轮风扇发动机通用规范[S].
- [4] MIL-STD-1783 发动机结构完整性大纲[S].
- [5] 任若恩, 王惠文著 多元统计数据分析——理论、方法、实例[M]. 北京: 国防工业出版社, 1997.
- [6] 唐守正 多元统计分析方法[M]. 北京: 中国林业出版社, 1986

(责任编辑 王震华)