

文章编号: 2097-1974(2025)01-0037-05 DOI: 10.7654/j.issn.2097-1974.20250105

基于k-means聚类熵权评价的飞行器质心调整优化方法

田小川, 郁立勇, 白斌, 陈思, 何文凯
(中国运载火箭技术研究院, 北京, 100076)

摘要: 针对飞行器质心调整流程复杂、耗时长的问题, 运用k-means聚类方法, 对飞行器配重历史数据进行聚类, 基于样本聚类结果, 计算出不同样本下飞行器标准配重, 再通过模拟装配计算增加标准配重后的飞行器质心偏移, 并得出一系列统计数据, 最后采用基于熵权的综合评价方法对比质心调整效果, 选出最优的飞行器标准配重, 进而简化飞行器质心调整流程, 大幅提升飞行器生产效率。

关键词: k-means; 熵权评价模型; 飞行器质心调整; 聚类
中图分类号: V42 **文献标识码:** A

Optimization Method of Aircraft Centroid Adjustment based on k-means Clustering Entropy Weight Evaluation

TIAN Xiaochuan, YU Liyong, BAI Bin, CHEN Si, HE Wenkai
(China Academy of Launch Vehicle Technology, Beijing, 100076)

Abstract: In response to the complex and time-consuming process of adjusting the center of gravity of aircraft, k-means clustering method is used to cluster the historical data of aircraft counterweight. Based on the clustering results of samples, the standard counterweight of aircraft under different samples is calculated. Then, the centroid offset of aircraft with standard counterweight is calculated through simulated assembly, and a series of statistical data are obtained. After that, the comprehensive evaluation method based on entropy weight is used to compare the results of centroid adjustment, and the optimal standard counterweight of aircraft is selected, thus simplifying the adjustment process of aircraft centroid and greatly improving the producing efficiency of aircraft.

Keywords: k-means; entropy weight evaluation model; aircraft centroid adjustment; cluster

0 引言

飞行器质心相对于理论质心的横偏量直接影响飞行器的姿态控制, 当质心横偏量超出规定范围时, 飞行器可能会产生滚转速率与气动俯仰频率的共振“连锁现象”, 进而导致配平攻角急剧增长, 横向载荷超过许用值而破坏飞行器, 还可能导致飞行器某一子午面长时间处于迎风位置, 加剧不对称烧蚀, 最终导致飞行失利。因此, 在飞行器生产过程中, 会对飞行器质心进行测量, 并根据测量结果对飞行器质心进行调整, 该调整流程耗时较长, 过程复杂, 严重影响生产效率。

为简化质心调整流程, 本文基于历史数据, 首先运用k-means数据聚类方法, 对历史数据样本进行分类分组, 通过不同的样本聚类结果, 计算出飞行器理想配重, 将不同样本对应的飞行器配重模拟装配到未

调整质心的飞行器中, 采用基于熵权的综合评价模型评价质心调整效果, 选出最理想的飞行器标准配重, 进而简化飞行器质心调整流程, 将飞行器质心调整工作时间从8天缩短至1天, 优化前的飞行器质心调整流程见图1。

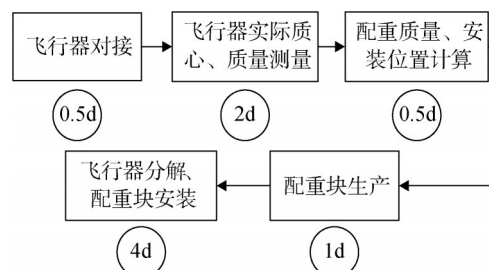


图1 优化前的飞行器质心调整流程
Fig.1 Aircraft centroid adjustment process

1 数据聚类分析

1.1 聚类的基本概念

聚类指将类似的对象分组，划分成多个类的过程，其中对象包括物理对象、抽象对象等的集合^[1]。聚类与数据分类有所不同，它是一种模式识别，且是无监督的。聚类过程是在对象数据库样本类型和类型数量未知的情况下，根据样本内部数据的相似度，对待分析对象数据样本开展分类，将相似程度最大的数据样本分为一类，相似程度较低或者差异较大的样本分为不同类。聚类分析方法有着良好的混合数据类型的处理能力、可伸缩能力和高维数据的处理能力等优点^[2]。

对象的数据库集合可以表示为一个 $n \times k$ 的矩阵如图2所示。行向量表示共有 n 个数据对象，列向量表示数据具有的 k 个属性。聚类过程为对该类型的数据集合中的数据对象，根据属性相似程度进行分类^[3]。

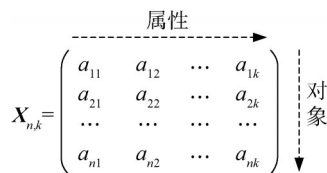


图2 数据矩阵示意

Fig.2 Data matrix

1.2 k-means 聚类算法

k-means 聚类算法也叫 k 均值聚类算法^[4]，流程如图3所示，其算法步骤如下：

a) 对数据进行预处理，主要完成数据的标准化和特征选取工作，之后生成如图2所示的 $n \times k$ 数据矩阵，对数据进行预处理可提高聚类收敛速度，并优化聚类结果；

b) 根据需要选择 k 值，将数据分为 k 组；

c) 根据选择的 k 值，随机选取 k 个聚类中心，聚类中心在数据向量所属区间内选取，否则无法收敛；

d) 选择适合对象数据库样本的距离函数，计算数据矩阵中各向量到聚类中心的距离；

e) 根据各向量到聚类中心的距离，把向量划分到与向量距离最小的聚类中心；

f) 根据所分配的向量，计算并重置聚类中心；

g) 返回步骤d)，循环这个过程，直至聚类中心不再更新；

h) 输出聚类结果。

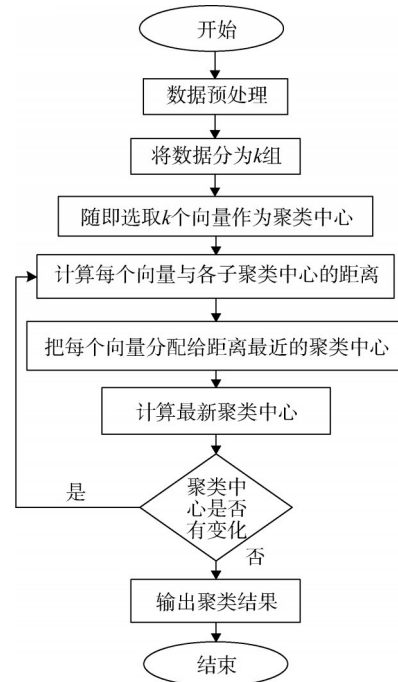


图3 k-means 聚类算法流程

Fig.3 k-means clustering algorithm flow

1.3 相似性度量

数据的相似程度是数据聚类分析的依据，故聚类之前需要进行相似性度量，即对向量之间的相似程度进行定义。相似性度量根据所采用的测度机制的不同，分为密度度量、距离度量、概念度量和连通性度量^[5]。

针对本文研究内容，选择距离度量对样本相似度进行度量，距离越近表示相似程度越高，距离越远表示相似程度越低。目前，聚类分析还没有一个可以适用所有应用场景的距离函数。作为使用最多的距离函数，Minkowski距离计算方便且直观，适用于度量具有连续性特征的向量。对于2个 n 维向量 $\mathbf{a} = (x_{21}, x_{22}, \dots, x_{2n})$ 与 $\mathbf{b} = (x_{11}, x_{12}, \dots, x_{1n})$ ，将两者之间的Minkowski距离定义为

$$\text{曼哈顿距离: } d_{12m} = \sqrt[p]{\sum_{k=1}^n |x_{1k} - x_{2k}|^p}, p = 1;$$

$$\text{欧式距离: } d_{12o} = \sqrt[p]{\sum_{k=1}^n |x_{1k} - x_{2k}|^p}, p = 2;$$

$$\text{切比雪夫距离: } d_{12c} = \sqrt[p]{\sum_{k=1}^n |x_{1k} - x_{2k}|^p}, p \rightarrow \infty.$$

1.4 聚类结果评价

聚类效果好坏一般使用轮廓系数进行评价。它综合考虑分离度和内聚度两种因素，来评价不同算法对聚类结果产生的影响^[6]。

若已经通过某种聚类算法将数据分成了 k 个簇，对于其中一个向量 i 来说， i 向量的轮廓系数为

$$S_i = \frac{b_i - a_i}{\max \{a_i, b_i\}} \tag{1}$$

式中 a_i 为所有它属于的簇中其他所有向量到 i 向量的距离的平均值； b_i 为所有它不属于的簇中所有向量到 i 向量的距离的平均值的最小值。所有向量的轮廓系数的平均值，就是该聚类的轮廓系数。轮廓系数取值范围为 $[-1, 1]$ ，聚类的分离度和内聚度越优，轮廓系数越趋近于 1。

2 飞行器配重样本聚类

对于样本聚类，可以按照样本收集、数据预处理、k-means 聚类、结果评价 4 个步骤进行聚类。

a) 样本收集。收集同一型号飞行器安装配重的全部历史数据，数据包含 2 个属性的内容：安装角度 α (I 象限偏 II 象限) 和配重块质量 M ，2 个属性示意如图 4 所示。本研究共收集了 37 组数据。

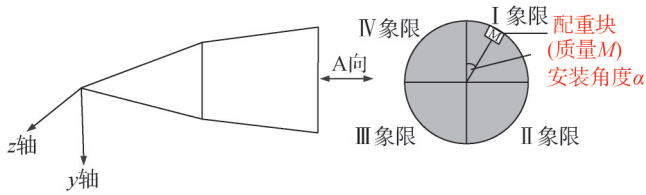


图4 配重块安装示意

Fig.4 Schematic diagram of weight block installation

b) 数据预处理。由于 2 个数据量纲不同，直接进行聚类会导致聚类结果收敛慢且聚类结果偏差较大，因此需对样本进行如下预处理：

$$M_z = M \sin \alpha \tag{2}$$

$$M_y = M \cos \alpha \tag{3}$$

式中 M_z 和 M_y 分别为配重块在飞行器径向沿 z 轴和 y 轴的质量分量，令 M_z 和 M_y 组成新的 37×2 矩阵，对该数据矩阵进行聚类。

c) k-means 聚类。根据数据分布情况，选择 k 值在 $[2, 4]$ 范围内，按照 1.2 节的聚类算法对数据进行聚类。当 $k=2$ 时，聚类结果见图 5，样本聚类为红蓝两类，“o” 形为两个样本的聚类中心；当 $k=3$ 时，聚类结果见图 6，样本聚类为绿、蓝、红 3 类；当 $k=4$ 时，聚类结果见图 7，样本聚类为粉、红、绿、蓝 4 类。

d) 聚类结果评价。按照 1.4 节计算轮廓系数的方

法，对聚类结果进行评价。由于本文聚类样本为二维同量纲向量，故采用 Minkowski 距离中的欧式距离作为相似度度量工具，这样计算简洁且更易直观理解。计算出当 $k=2$ 时，轮廓系数为 0.337；当 $k=3$ 时，轮廓系数为 0.353；当 $k=4$ 时，轮廓系数为 0.367。

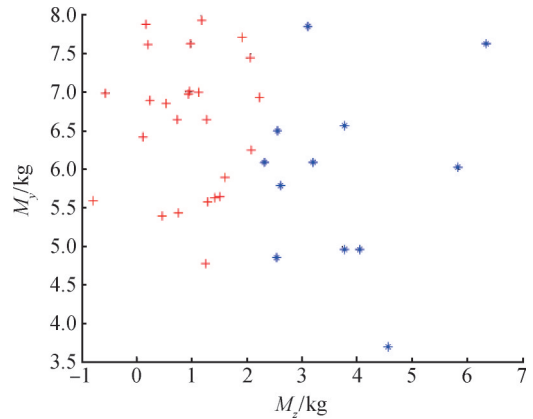


图5 k=2时的聚类结果

Fig.5 Cluster result at k=2

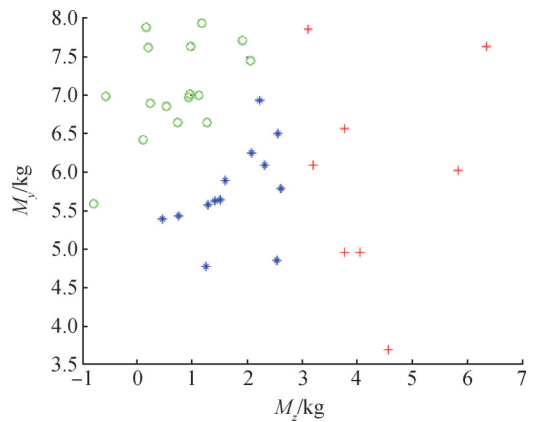


图6 k=3时的聚类结果

Fig.6 Cluster result at k=3

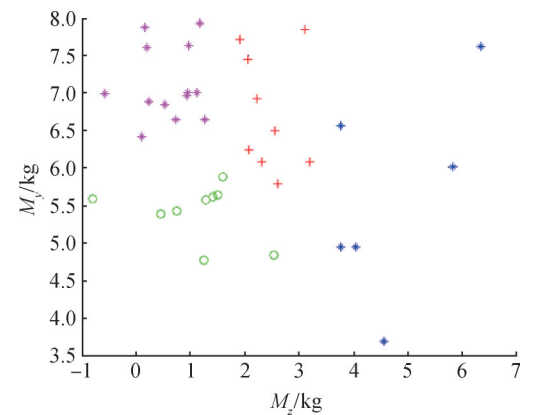


图7 k=4时的聚类结果

Fig.7 Cluster result at k=4

3 熵权法评价

3.1 评价数据

根据聚类结果,共分出9组样本数据,按照以下公式,分别利用每个样本计算出标准配重块质量(见表1)。

$$M_B = (\sum M_T/n) * (\sum L/n)/318 \quad (4)$$

$$\alpha_B = \arctan(L_z/L_y) \quad (5)$$

式中 M_B 为配重块质量; M_T 为飞行器总质量; L 为飞行器质心偏移量; n 为样本数量; α_B 为配重块安装角度; L_z 为飞行器在 z 轴方向的质心偏移量; L_y 为飞行器在 y 轴方向的质心偏移量。

表1 不同样本计算出的标准配重块

Tab.1 Standard weights calculated from different samples

聚类结果	M_B/kg	$\alpha_B/(\circ)$
2-1	6.70	7.05
2-2	7.08	32.18
3-1	7.23	13.98
3-2	6.30	2.68
3-3	7.01	35.98
4-1	7.18	20.08
4-2	7.39	39.96
4-3	5.60	8.54
4-4	7.16	4.59

将上述标准配重块模拟安装到目标飞行器中,并对安装标准配重块后飞行器质心平均偏移量、最大偏移量、偏移量离散度、标准差以及轮廓系数进行统计作为评价样本的指标数据,见表2。

表2 样本评价指标数据

Tab.2 Sample evaluation index data

聚类结果	平均偏移量/cm	最大偏移量/cm	离散度	标准差	轮廓系数
2-1	0.556 4	1.192 2	0.081 1	0.628 6	0.337 3
2-2	1.420 1	2.514 2	0.139 0	1.506 1	
3-1	0.707 1	1.671 7	0.097 4	0.789 6	0.353 4
3-2	0.629 6	1.099 8	0.075 6	0.686 1	
3-3	1.607 4	2.675 8	0.138 5	1.683 1	0.366 5
4-1	0.887 9	1.989 1	0.128 7	1.002 0	
4-2	1.898 9	2.978 9	0.138 5	1.963 5	
4-3	0.654 0	1.163 3	0.090 2	0.730 5	
4-4	0.667 4	1.209 6	0.086 9	0.737 3	

3.2 熵权评价模型

信息熵可以根据指标的重要性计算熵值和权重,用来对信息的有用程度进行评估^[7]。

在 $n \times p$ 的样本数据中,信息熵和权重的计算方法如下^[8]:

x_{ij} 为第 i 个样本第 j 个指标,按下式对其进行标准化:

$$x'_{ij} = \frac{x_{ij} - \min_i(x_{ij})}{\max_i(x_{ij}) - \min_i(x_{ij})} \quad (6)$$

对每个 x'_{ij} 按下式求取比重:

$$p_{ij} = \frac{x'_{ij}}{\sum_{i=1}^n x'_{ij}} \quad (7)$$

计算每个指标的信息熵值 E_j :

$$E_j = \frac{\sum_{i=1}^n p_{ij} \ln p_{ij}}{\ln n} \quad (8)$$

当 $p_{ij} = 0$, $p_{ij} \ln p_{ij} = 0$ 。

计算每个指标的权重 ω_j :

$$\omega_j = \frac{1 - E_j}{\sum_{k=1}^p (1 - E_k)} \quad (9)$$

熵权评价结果 T_i , 样本越优, T_i 值越小, 飞行器质心平均偏移量、最大偏移量、偏移量离散度、标准差均为越小代表样本越优, 而轮廓系数在 $[-1, 1]$ 区间内越大代表样本越优, 故轮廓系数项赋值为负。

$$T_i = \sum_{j=1}^4 \omega_j x_{ij} - \omega_5 x_{i5} \quad (10)$$

3.3 熵权评价结果

按照3.2节对样本进行熵权评价, 计算结果见表3, 可见样本2-1为最优样本, 查表1可知, 使用2-1样本计算标准配重块质量为6.7 kg, 安装位置为I偏II象限7.05°。

表3 熵权评价结果

Tab.3 Entropy weight evaluation results

聚类结果	熵权评价结果
2-1	0.405 2
2-2	1.033 9
3-1	0.561 4
3-2	0.410 0
3-3	1.137 5
4-1	0.706 2
4-2	1.312 8
4-3	0.435 3
4-4	0.448 1

4 试验验证

在真实的飞行器生产过程中, 生产质量为6.7 kg

标准配重块，安装到飞行器I象限偏II象限7.05°，并进行质心测量的结果如表4所示，平均偏移量0.557 2 mm，最大横偏量0.864 mm，满足指标要求。

表4 试验验证结果

Tab.4 Experiment to verify the results

编号	质量/kg	z_c/mm	y_c/mm	L/mm
XXX1	652.64	0.602	-0.238	0.647
XXX2	653.71	0.041	0.014	0.043
XXX5	653.83	0.633	0.348	0.723
XXX6	657.09	0.225	0.835	0.864
XXX7	652.94	0.499	0.099	0.509

5 结束语

本文运用k-means聚类方法，对飞行器配重历史数据进行聚类，并通过熵权评价方法对不同样本进行评价，进而选出最优样本，最后确定飞行器最佳标准配重。通过试验验证，标准配重能够满足飞行器质心调整要求，该方法大幅提升了飞行器生产效率，降低了生产成本。

参 考 文 献

[1] 邵峰晶, 于忠清. 数据挖掘原理与算法[M]. 北京: 中国水利水电出版社, 2003.
SHAO Fengjing, YU Zhongqing. Principles and algorithms of data mining[M]. Beijing: China Water Resources and Hydropower Press, 2003.

[2] HAN J, PEI J, KAMBER M. Data mining: concepts and techniques [M]. Amsterdam: Elsevier, 2011.

[3] 王永红. 新能源汽车数据聚类分析及其应用研究[D]. 苏州: 苏州

大学, 2016.
WANG Yonghong. New energy vehicle data cluster analysis and its application[D]. Suzhou: Soochow University, 2016.

[4] ADNAN R M, PARMAR K S, HEDDAM S, et al. Suspended sediment modeling using a heuristic regression method hybridized with K-means clustering[J]. Sustainability, 2021, 13(9): 1-21.

[5] RICHARD J R, MICHAEL W G. 数据挖掘教程[M]. 翁敬农, 译. 北京: 清华大学出版社, 2003.
RICHARD J R, MICHAEL W G. Data mining tutorials[M]. Translated by WENG Jingnong. Beijing: Tsinghua University Press, 2003.

[6] LACHIHEB O, GOUIDER M S, SAID L B. An improved mapreduce design of K-means with iteration reducing for clustering stock exchange very large Datasets[C]. Beijing: 2015 11th International Conference on Semantics, Knowledge and Grids (SKG), 2015.

[7] 刘阳. 基于加权熵K-means聚类的改进卷积神经网络下喉镜图像分类研究[D]. 长春: 长春工业大学, 2022.
LIU Yang. Research on laryngoscopic image classification based on improved convolutional neural network based on weighted entropy K-means clustering[D]. Changchun: Changchun University of Technology, 2022.

[8] 魏法杰. 基于熵权的复杂装备研制过程技术成熟度评价研究[J]. 中国管理科学, 2012(20): 911-915.
WEI Fajie. Research on technology maturity evaluation of complex equipment development process based on entropy weight[J]. China Journal of Management Science, 2012(20): 911-915.

作 者 简 介

田小川 (1992—), 男, 工程师, 主要研究方向为电气系统设计与型号总体设计。
郁立勇 (1981—), 男, 高级工程师, 主要研究方向为数据评估。
白 斌 (1986—), 男, 高级工程师, 主要研究方向为型号总体设计。
陈 思 (1992—), 男, 工程师, 主要研究方向为型号总体设计。
何文凯 (1992—), 男, 工程师, 主要研究方向为结构总体设计。

(上接第36页)

[13] QIANG Baochen, LE Zhang. Switching antiwindup design on enlarging the domain of attraction for a supercavitating vehicle subject to actuator saturation[J]. Journal of Dynamic Systems, Measurement, and Control, 2017, 139(9): 091010.

[14] Bui Duc Hong Phuc, YouSam-Sang, Natwar Singh Rathore, et al. Dynamical analysis and robust control for dive plane of supercavitating vehicles[J]. Applied Ocean Research, 2019, 84: 259-267.

[15] 韩云涛, 程章龙, 李盼盼, 等. 超空泡航行体LPV鲁棒变增益控制[J]. 华中科技大学学报(自然科学版), 2017, 45(7): 127-132.
HAN Yuntao, CHENG Zhanglong, LI Panpan, et al. Robust variable gain control of supercavitating vehicle LPV[J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2017, 45(7): 127-132.

[16] VANEK B. Control methods for high-speed supercavitating vehicles

[D]. Minneapolis: University of Minnesota, 2008.

[17] SHAO Y, MESBAHI M, PlaningBalas G. Switching and supercavitating flight control[C]. Texas: Proceedings of AIAA Guidance, Navigation and Control Conference and Exhibit, 2003.

[18] 庞爱平, 何朕, 王京华, 等. 超空泡航行体 H_∞ 状态反馈设计[J]. 控制理论与应用, 2018, 35(2): 146-512.
PANG Aiping, HE Zhen, WANG Jinghua, et al. Design of state feedback for supercavitating vehicles with H_∞ [J]. Control Theory and Applications, 2018, 35(2): 146-512.

作 者 简 介

岳玲玮 (1999—), 女, 助理工程师, 主要研究方向为姿控总体设计。
赵长见 (1976—), 男, 研究员, 主要研究方向为飞行器总体设计。
宋志国 (1987—), 男, 研究员, 主要研究方向为飞行控制。