



南京航空航天大学

本科毕业设计（论文）

题 目 机械磨损界限值制定方法研究

学生姓名	卢冠宇
学 号	071951919
学 院	通用航空与飞行学院
专 业	交通运输（民航维修工程）
班 级	0720902
毕设地点	南京航空航天大学
指导教师	陈果 教授

二〇二四年七月

南京航空航天大学

本科毕业设计（论文）诚信承诺书

本人郑重声明：所呈交的毕业设计（论文）是本人在导师的指导下独立进行研究所取得的成果。尽我所知，除了文中特别加以标注和致谢的内容外，本设计（论文）不包含任何其他个人或集体已经发表或撰写的成果作品。对本设计（论文）所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

作者签名：_____

日期： 20 年 月 日

南京航空航天大学

毕业设计（论文）使用授权书

本人完全了解南京航空航天大学有关收集、保留和使用本人所送交的毕业设计（论文）的规定，即：本科生在校攻读学位期间毕业设计（论文）工作的知识产权单位属南京航空航天大学。学校有权保留并向国家有关部门或机构送交毕业设计（论文）的复印件和电子版，允许论文被查阅和借阅，可以公布论文的全部或部分内容，可以采用影印、缩印或扫描等复制手段保存、汇编论文。保密的论文在解密后适用本声明。

论文涉密情况：

不保密

保密，保密期（起讫日期：_____）

作者签名：_____

导师签名：_____

日期： 20 年 月 日

日期： 20 年 月 日

摘 要

航空发动机附件机匣是动力传输的“桥梁”，具有转速高、工作环境恶劣、可靠性要求高等特点，极易发生各种磨损故障。因此采用油液分析技术对航空发动机附件机匣的磨损情况进行监控与故障诊断就有十分重要的意义。本文聚焦于磨损界限值制定方法研究，主要工作内容如下：

(1) 介绍了传动界限值制定方法，包括：直方图法、线性回归法、增长率法，针对航空发动机附件机匣油液数据不一定服从正态分布的情况，研究了基于概率密度函数估计的界限值制定方法，包括：最大熵法、Parzen 窗法以及 K 近邻法；

(2) 通过生成正态分布、F 分布、威布尔分布的仿真数据，验证了基于概率密度函数估计界限值方法的有效性及应用范围；

(3) 对真实航空发动机附加机匣磨损数据进行基于概率密度函数估计，阐述上述方法在实际工程状态监测中的应用效果。

关键词：航空发动机附件机匣、磨损监测、界限值制定、概率密度函数

ABSTRACT

Aero engine accessory casing is the "bridge" of power transmission, which has the characteristics of high speed, bad working environment and high reliability, and is prone to various wear failures. Therefore, it is very important to monitor and diagnose the wear of aero-engine accessory casing by using oil analysis technology. This paper focuses on the research on the formulation method of wear limit value, and the main contents are as follows:

(1) The formulation methods of transmission limit value are introduced, including: histogram method, linear regression method and growth rate method. In view of the situation that the oil data of aero-engine accessories casing may not follow the normal distribution, the formulation methods of limit value based on probability density function estimation are studied, including maximum entropy method, Parzen window method and K-nearest neighbor method.

(2) By generating the simulation data of normal distribution, F-distribution and Weibull distribution, the validity and application range of the method based on probability density function are verified;

(3) The wear data of the real aeroengine additional casing is estimated based on probability density function, and the application effect of the above method in the actual engineering condition monitoring is expounded.

Keywords: Aero engine accessory casing, Wear monitoring, Limit value formulation, Probability density function

目 录

第一章 绪论.....	1
1.1 项目研究的背景及意义.....	1
1.2 国内外研究现状.....	2
1.3 问题提出.....	3
1.4 论文主要工作及章节安排.....	3
第二章 航空发动机附件机匣磨损元素界限值制定方法.....	4
2.1 增长率法.....	4
2.2 线性回归法.....	5
2.3 直方图法.....	6
2.4 最大熵法.....	9
2.5 Parzen 窗法.....	10
2.6 K 近邻法.....	12
2.7 注意事项.....	13
第三章 基于仿真信号的界限值制定方法对比验证.....	14
3.1 服从正态分布数据集.....	14
3.2 服从 F 分布数据集.....	15
3.3 服从威布尔分布数据集.....	17
第四章 基于附件机匣磨损元素的界限值制定方法对比验证.....	20
4.1 真实航空发动机附件机匣滑油光谱数据集.....	20
4.1 Cu 元素界限值制定.....	21
4.2 Ag 元素界限值制定.....	22
4.3 Ni 元素界限值制定.....	24
4.4 Mg 元素界限值制定.....	25
第五章 总结与展望.....	28
5.1 研究总结.....	28
5.2 研究展望.....	28
参考文献.....	29
致谢.....	30

第一章 绪论

1.1 项目研究的背景及意义

航空发动机的附件机匣在飞机传动系统中扮演着至关重要的角色，对于发动机的正常运行起着不可或缺的作用。所以，可靠高效地监控故障就变得尤为重要。这不仅有助于提升我国航空发动机的整体性能水平，还可为未来技术发展提供重要指导^[1-3]。

当前，航空发动机传动系统被广泛认为是制约发动机技术进步的关键因素之一，同时也是国内外军用航空发动机故障率较高的系统之一。传动系统目前的技术水平受到结构和技术限制的影响，无法提供冗余的设计，因此经常会在出现故障时引发严重后果。据统计数据显示，自 2005 年至 2013 年，我国某型新型飞机列装后，传动系统接连出现故障。这些故障不仅导致了多次空中停机等紧急情况，还引发了数起二级重大飞行事故和飞机迫降事件。这一系列问题直接导致了数十亿元的经济损失，严重威胁了空军的正常运行和国家财产的安全。图 1.1 (a) 为某发动机运行中振动、温度超标，停机后检查出附件齿轮箱滚动轴承外圈疲劳剥落，图 1.1 (b) 为某型发动机在工作过程中，因垫圈与固定螺钉脱落，后被游星齿轮碾碎，造成的齿轮磨损。



因此，对航空发动机附件机匣进行运行机理和故障诊断研究具有重要工程意义^[4-5]，可实现：

(1) 增强传动系统的安全性。通过监测和诊断传动系统关键部件，能够及时察觉早期故障和故障迹象，以便及时处理，提升飞机的安全性。

(2) 应用健康管理系统可根据监测数据判断传动系统的运行状态，减少维护检查次数，延长定期检测周期。

(3) 提高传动系统的可靠性。通过健康监测和故障诊断，减少对无故障部件的检查，延长检查周期，从而降低维护成本。

1.2 国内外研究现状

在航空发动机领域，界限值是指一个跨越状态的关键数值，通常被称为阈值或阈值^[6]。飞机发动机的状态评估和故障诊断可以通过比较实际的磨损数据和这些阈值来进行。因此，实际评估与诊断的效果直接受到界限值的设定质量的影响^[7]。

本文深入探讨了如何科学制定航空发动机附件机匣的滑油光谱磨损数据的阈值。研究分析了国内外在制定阈值的方法，特别是自 20 世纪 40 年代美国开始使用光谱分析技术以来的发展历程。美国海军和空军在光谱分析技术的应用方面取得了显著成效^[8]，特别是在 SOAP 评估标准和 JOAP 机构的建立上。

考虑到传统方法的局限性，本文提出了采用概率密度函数、支持向量机（SVM）、GM(0, N)模型等高级统计技术来估计光谱数据的分布，并据此制定更加科学合理的界限值。这些高级模型可以更精确地拟合实际数据，反映元素间的相关关系，并通过实际数据验证其有效性。此外，利用遗传算法和 BP 神经网络建立的动态界限值模型，可以更科学地监测发动机的状态，有效预防因超出阈值而导致的设备故障或停机检查，从而提高飞机的安全性和运行效率^[10]。这种对传统阈值制定方法的改进和高级技术的应用，不仅增强了光谱分析技术在航空发动机维护中的应用，也为其他行业提供了参考，展示了数据分析在设备维护和故障诊断中的重要价值。

闫辉等^[11]考虑到机械设备的运行环境、工作环境以及数据污染的影响，提出了改进的油液监测界限值传统计算方法。他们通过引入设备本身信息，采用故障率，使得界限值制定更合理，且评估效果优于传统方法。

为了突破传统光谱数据界限值制定方法的约束，新的研究已经开展，旨在通过概率密度函数更准确地估计光谱数据的分布。李爱等人^[12]采用支持向量机（SVM）来估计光谱数据的概率密度函数，成功实现了更科学的界限值设定。孙金哲等^[13]基于 GM(0, N)模型，开发了一种新的油液分析界限值制定方法，这种方法不仅能够准确拟合数据，还能揭示元素间的相互关系，并已通过实际数据得到验证。

其他方法还包括基于比例模型的动态界限值模型和基于神经网络的动态界限值模型等。

对于提高界限值的精度和动态程度，有研究者提出了更为科学的方法。比如，朱焕勤^[14]等对积分方式变化的动态调整问题进行了研究，通过分析实验条件和建立工作曲线，建立了 MOA 检测不同油品的关系式。周平等^[15]以稳健回归模型为基础，提出了提高抗异常点干扰能力的方法，从而更加真实地监控设备的运行状态。曹轶乾^[16]运用遗传算法与 BP 神经网络，

创立了一种动态界限值模型，以更科学地监测发动机的运行状态。万耀青^[17]提议，在设定界限值时应留有一定的保守余地，并通过加密采样精确评估设备状态，以减少因超出界限值而导致的机器停机检查。

1.3 问题提出

传统界限值设定方法通常基于假设数据服从正态分布，通过统计数据的均值和方差来确定警告和异常阈值。然而，航空发动机附件机匣的磨损数据往往分布不均，不一定遵循正态分布，若仍采用正态分布计算界限值，易形成数值偏差，造成机械运行健康状态的错报、误报，因此这种传统方法面临局限，这一挑战促使研究者探索新的界限值设定技术。

1.4 论文主要工作及章节安排

本论文围绕机械磨损界限值制定方法研究展开，主要工作内容如下：

第一章为绪论，介绍了本文的研究背景和意义，综述了机械磨损界限值的制定方法的研究现状，分析了航空发动机附件制定界限值所面临的问题，并对本文的主要研究内容进行了介绍。

第二章界限值制定方法介绍，阐述了传动界限值的制定方法，包括：直方图法、线性回归法、增长率法，针对航空发动机附件机匣油液数据不一定服从正态分布的情况，本文对使用概率密度函数来估计界限值，有许多种方法，例如：K 近邻法、Parzen 窗法、最大熵法；

第三章是通过生成正态分布、F 分布、威布尔分布的仿真数据，验证了基于概率密度函数估计界限值方法的有效性及其适应范围；

第四章对真实航空发动机附加机匣磨损数据进行基于概率密度函数估计，阐述上述方法在实际工程状态监测中的应用效果。

第五章是总结与展望。总结本文主要研究工作和创新点，并阐述后续需深入开展的工作。

第二章 航空发动机附件机匣磨损元素界限值制定方法

磨损元素的界限值对于预测航空发动机的磨损故障至关重要，这能有效消除潜在故障风险。在精确界定飞机发动机磨损元素阈值的研究中，本章详细介绍了两种主要的方法：传统的基于正态分布的假设方法和一种先进的基于概率密度函数的方法。在第一种方法中，我们假设润滑油光谱磨损数据遵循正态分布，并通过计算这些数据的均值和方差来确定相应的警告和异常阈值。具体地，我们将警告阈值设定为均值加上两倍标准差（ σ ），异常阈值设定为均值加上三倍标准差（ σ ）。这种方法简单直观，易于实施，但它依赖于数据的正态分布假设，可能不适用于所有类型的数据集。第二种方法采用概率密度函数来更精确地估计数据的分布特性。此方法首先收集大量的润滑油光谱磨损数据样本，对这些数据样本进行详细的概率密度函数分析，以确定数据的实际分布情况。然后，根据这个分布，我们计算出使累积样本数达到整个样本总数的 97.72% 对应的数值作为警告阈值，99.87% 对应的数值作为异常阈值。这种基于概率密度的方法不仅提供了一种不依赖于正态分布假设的界定阈值手段，而且能更准确地反映实际数据的统计特性。

2.1 增长率法

为了进一步提高磨损元素阈值设定的科学性和实用性，研究还引入了增长率法（也称为控制图法）。这种方法通过分析样本数据的变化趋势和波动情况来确定阈值。它基于统计学中的小概率事件理论，在本理论中有两个限值，分别为 σ_1 和 σ_2 ，通过计算得到特定的增长率，从而设定警告和异常的界限。这种方法特别适用于那些数据变化具有明显趋势性和周期性的情况，如航空发动机附件机匣的磨损监测。

$$\alpha_1 = \bar{y} + 2\sigma, \quad \alpha_2 = \bar{y} + 3\sigma \quad (2.1)$$

式 (2.1) 中， σ 为样本的方差， \bar{y} 为样本的均值，即

$$\begin{aligned} \bar{y} &= \frac{1}{n} \sum_{i=1}^n y_i \\ \sigma &= \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 \end{aligned} \quad (2.2)$$

其中， n 是样本容量， y_i 为某元素的浓度或浓度增长率（梯度）。

增长率法流程图如图 2.1 所示：

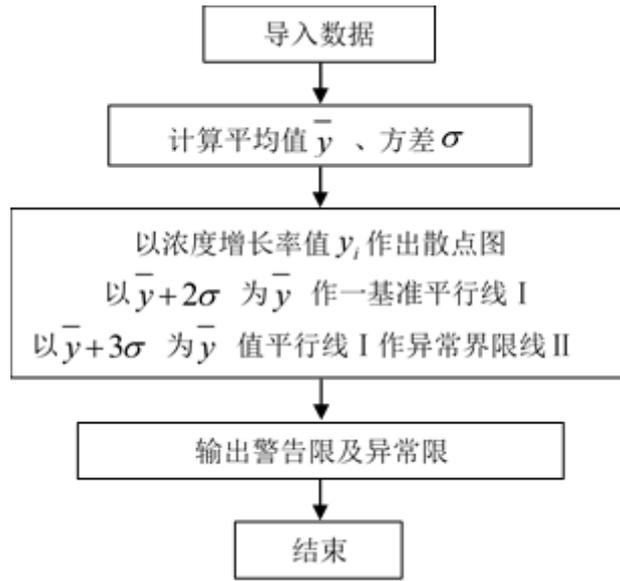


图 2.1 增长率法流程图

2.2 线性回归法

在研究航空发动机附件机匣的磨损量时，发现磨损量与耐磨寿命之间通常呈线性关系，这表明随着运行时间的增长，磨损量以可预测的方式递增。因此，本研究通过构建一元线性回归数学模型来拟合磨损元素的变化趋势，模型形式为 $y=ax+b$ 。为求解模型参数 a 和 b ，采用最小二乘法，并通过公式计算回归直线的精度 S 。

$$S = \sqrt{\frac{1}{N-2} \left[\sum_{i=1}^N (y_i - \bar{y})^2 - \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{\sum_{i=1}^N (x_i - \bar{x})^2} \right]} \quad (2.3)$$

根据统计数据，约有 95.4% 的点位于 $y = a + bx \pm 2S$ 两直线之间，而约有 99.7% 的点位于 $y = a + bx \pm 3S$ 两条直线之间。

以上探讨基于假设曲线具有形式 $y = ax + b$ 的前提，进一步地，为验证模型的适用性，本研究对回归直线的显著性进行了相关系数检验。通过统计分析，95.4% 的数据点落在预测的两条直线之间，显示出较高的预测准确度。此外，本研究还运用假设检验方法评估了磨损量与耐磨寿命的线性关系是否成立，以确保模型的科学性和实用性。通过这些方法，本研究为航空发动机的维护提供了理论依据和实践指南。本论文中的假设检验如下：

$$H_0 : b = 0; H_1 : b \neq 0 \quad (2.4)$$

众所周知，在概率中对事件的显著性检验是通过相关系数 γ 来检测的，相关系数的公式为：

$$\gamma = \frac{Cov(x, y)}{\sqrt{D(x)}\sqrt{D(y)}} = \frac{\sum_{i=0}^N x_i y_i - \bar{x} \sum_{i=0}^N y_i}{\sqrt{(\sum_{i=0}^N x_i^2 - \bar{x} \sum_{i=0}^N x_i)(\sum_{i=0}^N y_i^2 - \bar{y} \sum_{i=0}^N y_i)}} \quad (2.5)$$

相关系数用来衡量变量 x 和 y 之间的线性关系程度。绝对值较大的相关系数表示 x 和 y 之间的线性关联性较强，而较小的绝对值则意味着关联性较弱。为了精确评估变量之间的线性关系，我们采用了定量的分析方法，即通过比较计算得到的相关系数与临界值。如果相关系数超过了临界值，我们可以认定变量 x 和 y 之间存在显著的线性相关性；若低于该值，则认为两者之间无显著线性关系。具体的分析步骤包括：首先，计算样本数据的相关系数 γ ，以及通过线性回归确定的参数 a 和 b ，并评估回归的精度 S 。然后，根据样本数量 n （自由度 $-2n-2$ ）和设定的显著性水平 α ，查找相关系数的临界值。最终，将计算得到的 γ 与这一临界值进行比较，以此判定变量间是否具有线性相关性。此方法不仅增强了研究的严密性，还为实际应用提供了重要依据。

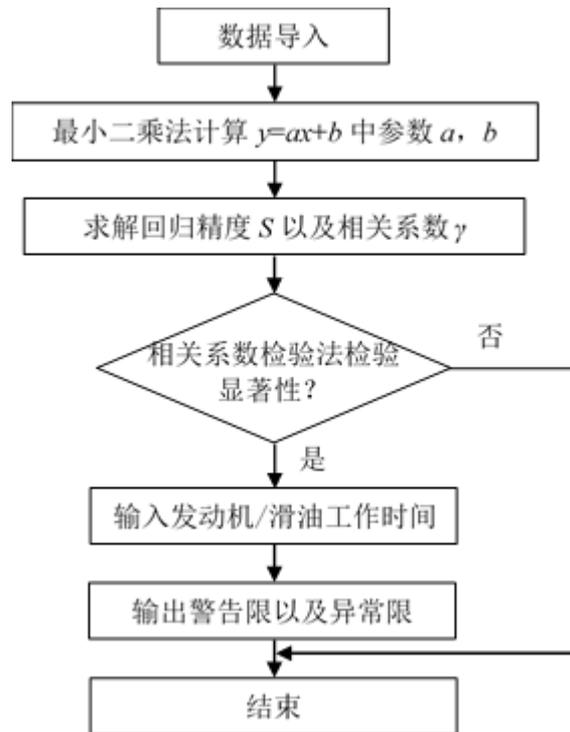


图 2.2 线性回归法流程图

为了精确评估航空发动机磨损元素的阈值，本文介绍了一种基于线性回归的方法，并提供了流程图（图 2.2）。

2.3 直方图法

在此方法中，置信度 α 和样本容量 n 是关键参数。特定时刻 t 的元素界限值通过计算 $a*t+b$ 加上 $2S$ （警告限）和 $3S$ （异常限）来确定。本章还详细阐述了直方图法，该方法通过分析样

本数据的直方图来识别航空发动机磨损状态的变化。这一方法将样本数据的累积频数与总样本数的特定百分比（97.72%和 99.87%）相比较，以设定警告限和异常限。

为了精确分析样本数据，并制定合理的界限值，首先需要采用直方图法来确定样本的分组数量和每个组的区间长度。接下来，对每个组内的样本数据进行频数、均值和方差的统计分析。尽管计算样本均值的常规方法是对所有数据进行累加后求平均，如公式（2.2）所示，这种方法在处理大规模数据集时可能会因异常值等因素出现误差。因此，本研究采用了一种多级均值算法，这种间接算法可以有效提高数据处理的精度和可靠性。此外，通过设置多个计算阶段，其中数据点数 $N=2 \times m$ ，并在每个阶段应用该计算公式，可以进一步优化数据分析的结果。

$$X = X_m(0) = \frac{1}{2}(X_{m-1}(0) + X_{m-1}(1)) \quad (2.6)$$

采用 Kendall 和 Stuart 给出的总体样本量 N 和分组数目 K 应满足的最优关系式，进而确定组区间的个数 K ：

$$K = 1.87(N-1)^{0.4} \quad (2.7)$$

为了验证样本数据是否呈正态分布，本研究采用了 χ^2 拟合优度检验法，这种方法可以准确评估数据的分布特性。首先，设定假设认为样本数据的区间频数符合正态分布。然后，根据该假设，计算 χ^2 统计量，此统计量基于正态分布概率密度函数和具体的样本区间频数得出。之后，将得到的 χ^2 值与卡方检验的临界值进行比较。如果 χ^2 值低于临界值，则确认样本频数遵循正态分布，否则，认为数据不符合正态分布，难以确定准确的界限值。

$$\chi^2 = \sum_{i=1}^K \frac{(f_i - np_i)^2}{np_i} \quad (2.8)$$

此外，对于区间的数量 K 和置信度进行调整，确保每个区间的频数适宜，并在适当的置信度下制定界限值。如图 2.3 所示，直方图法流程清晰展现了整个检验过程。

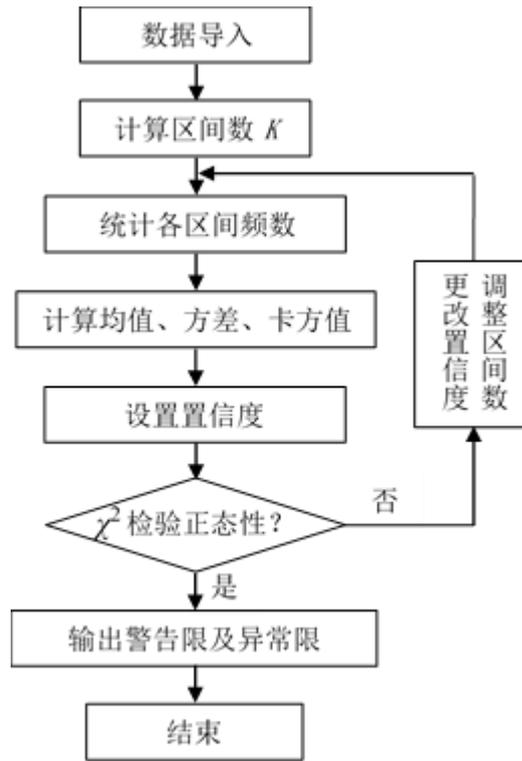


图 2.3 直方图法算法流程

在航空发动机磨损数据分析中，虽然常用方法基于正态分布的假设，实际上磨损数据的分布往往更加复杂且未知。因此，本文提出需要通过收集大量的磨损数据来精确估计其概率密度函数，以揭示数据的实际分布特性。通过这一分布估计，可以更准确地确定磨损元素的阈值。磨损元素界限值计算方法为：

如果想要估计概率密度函数

$$F(x) = P(X \leq x) = \int_{-\infty}^x p(t) dt \quad (2.7)$$

就必须要求线性算子方程

$$\int_{-\infty}^{\infty} \theta(x-t)p(t) dt = F(x) \quad (2.8)$$

的解，其中

$$\theta(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (2.9)$$

同时，方程的解还必须满足下列条件：

$$p(x) \geq 0, \int_{-\infty}^{\infty} p(x) dx = 1 \quad (2.10)$$

在式（2.8）中，分布函数 $F(x)$ 的表达式未知，但是给出了一组样本 x_1, \dots, x_i 。由概率论可

知该组样本是独立同分布的。

现在利用样本 x_1, \dots, x_l ，来构造经验分布函数

$$F_l(x) = \frac{1}{l} \sum_{i=1}^l \theta(x - x_i), \quad i = 1, 2, \dots, l \quad (2.11)$$

对于概率密度函数的估计，本研究分析了两种主要的方法：参数估计和非参数估计。参数估计依赖于预设的分布假设（如正态或二项分布），而非参数估计方法则无需此类假设，直接从数据学习其分布，如最大熵法、Parzen 窗法和 K 近邻法等，提供了更大的灵活性和广泛的应用可能。这些方法的选择和应用，依据具体的数据特性和分析需求来确定。

2.4 最大熵法

最大熵法是基于信息论原理，在已知数据约束条件下，通过优化熵值来推导概率分布的一种方法。对于一个具备特定概率分布 $p = \{p_1, p_2, \dots, p_n\}$ 的离散随机变量 $X = \{x_1, x_2, \dots, x_n\}$ ，此方法在满足数据正规化和若干矩约束的前提下，寻求一种概率分布，使得其熵值最大化。

$$\sum_{i=1}^n p_i = 1 \quad (2.16)$$

$$\sum_{i=1}^n g_{ri} p_i = a_r, \quad (r = 1, 2, \dots, m) \quad (2.17)$$

这里， g_{ri} 、 a_r 为已知的定值。Shannon 熵定义：

$$S(p) = - \sum_{i=1}^n p_i \ln p_i \quad (2.18)$$

这种方法以 Shannon 熵为基础，旨在在已知的约束条件下，推导出一种最无偏见且统一的概率分布，从而保证了分布的客观性和广泛的适应性。通过使用拉格朗日乘子法可以得到：

$$p_i = \exp(-\lambda_0 - \lambda_1 g_{1i} - \lambda_2 g_{2i} - \dots - \lambda_m g_{mi}), \quad (i = 1, 2, \dots, n) \quad (2.19)$$

$$\lambda_0 = \ln \left[\sum_{i=1}^n \exp(-\sum_{j=1}^m \lambda_j g_{ji}) \right] \quad (2.20)$$

$$a_r = \frac{\sum_{i=1}^n g_{ri} \exp(-\sum_{j=1}^m \lambda_j g_{ji})}{\sum_{i=1}^n \exp(-\sum_{j=1}^m \lambda_j g_{ji})}, \quad (r = 1, 2, \dots, m) \quad (2.21)$$

在公式 2.17 当中，公式通过整理形成了 m 个非线性方程组，这些方程组的解一般是通过 Newton 迭代法来进行计算的，计算公式还可以写成如下的表达式：

$$f(\lambda)_r = a_r \sum_{i=1}^n \exp(-\sum_{j=1}^m \lambda_j g_{ji}) - \sum_{i=1}^n g_{ri} \exp(-\sum_{j=1}^m \lambda_j g_{ji}), \quad \lambda = \lambda_1, \lambda_2, \dots, \lambda_m \quad (2.22)$$

以非线性方程组

$$F(\lambda) = 0 \quad (2.23)$$

来表示式（2.20），根据 Newton 法有

$$\lambda_{k+1} = \lambda_k - [F'(\lambda)]^{-1} F(\lambda_k) \quad (2.24)$$

在上述公式中， F 的雅可比矩阵为 $F'(\lambda)$ ，根据上述公式进行迭代就能计算出 λ 。再将其代入到公式中，就能得到概率密度函数，具体的计算过程如下图：

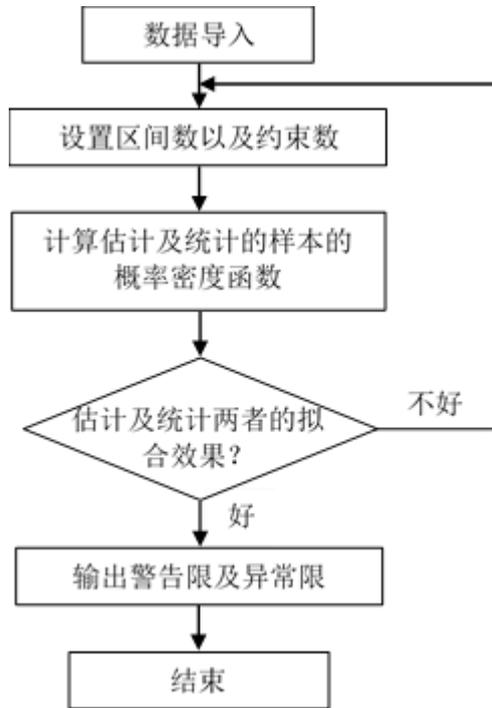


图 2.4 最大熵法算法流程图

2.5 Parzen 窗法

Parzen 窗估计法，一种根据已有样本推断整体概率密度分布的非参数技术，主要通过计算特定区域内样本点密度的平均值来实现。此方法在非参数密度估计中因其坚实的理论基础和优异性能而广泛应用。

具体操作中，设 x 为 d -维空间中的点，以 x 为中心构造一个边长为 h_N 的超立方体 V_N ，计算其体积并统计落入该立方体内的样本点数 k_N 。通过这种方式，可以精确地估算点 x 处的概率密度 $P(x)$ ，从而有效地进行数据分析和概率预测。

$$\varphi(u) = \begin{cases} 1, & \text{当 } |u_j| \leq \frac{1}{2}, j=1,2,\dots,d \\ 0, & \text{其他} \end{cases} \quad (2.12)$$

并使 $\varphi(\mu)$ 满足式 (2.10) 条件, 则落入超立方体内的样本数为

$$k_N = \sum_{i=1}^N \varphi\left(\frac{x-x_i}{h_N}\right) \quad (2.13)$$

将式(2.11)代入 $P_N(x)$ 的基本公式:

$$\hat{P}_N(x) = \frac{k_N/N}{V_N} \quad (2.14)$$

得到概率密度函数的估计值为:

$$\hat{P}_N(x) = \frac{1}{N} \sum_{i=1}^N \frac{1}{V_N} \varphi\left(\frac{x-x_i}{h_N}\right) \quad (2.15)$$

公式 (2.28) 是 Parzen 窗法估计中的关键表达式, 它描述了如何通过平均的方式对概率密度函数 $p(x)$ 进行估计。此方法考虑了变量 x 与样本点 x_i 之间的某种函数关系。窗函数 $\phi(\mu)$, 常用的有超立方体窗函数 (方窗) 等, 其核心功能是作为内插工具, 通过与样本的距离来决定样本在估计中的权重。此外, 窗宽 h_N 作为一个决定性参数, 在样本数量 N 有限的情况下对估计效果影响显著。整个估计流程在图 2.5 中有详细展示。

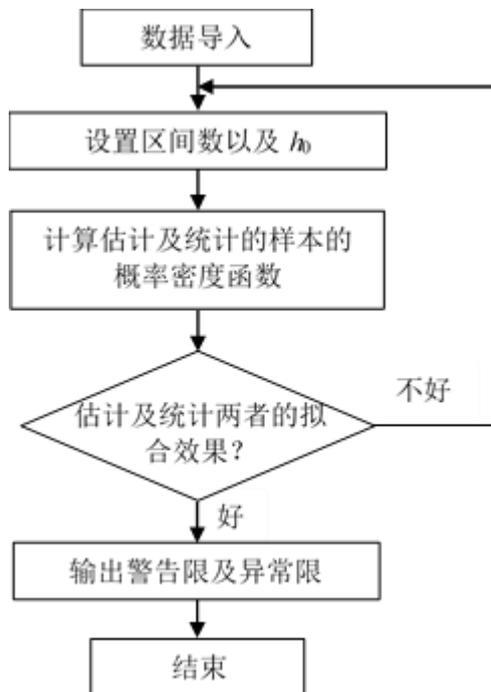


图 2.5 Parzen 窗法算法流程图

2.6 K 近邻法

k_N 近邻法核心在于通过分析特定点 x 周围的局部样本密度来估计 x 的概率 $p(x)$ 。这种方法不单纯依赖样本数量 N 的简单统计，而是通过确定一个包含 k_N 个样本的体积，这个体积以 x 为中心扩展直至囊括足够的近邻样本。体积大小反映了周围样本的密度：密度高的区域体积较小，密度低的区域则体积较大。该体积增长直到达到高密度区域后将不再扩展。此方法通过公式 (2.14) 实施，确保精确计算 x 的概率密度，这对理解样本分布情况至关重要。

$$(1) \lim_{N \rightarrow \infty} V_N = 0$$

$$(2) \lim_{N \rightarrow \infty} k_N = \infty$$

$$(3) \lim_{N \rightarrow \infty} k_N = 0$$

当样本数量 N 无限增大时， k_N 也将逐渐增加，这有利于利用公式 (2.14) 更精确地估算在体积 V_N 内点的概率分布。然而，为避免 k_N 增速过快而导致 V_N 无限缩小，需对 k_N 的增长速度加以适当限制。一种常见的做法是设 k_N 为 N 的函数，如 $k_N = k_1 \sqrt{N}$ ，其中 k_1 是任意大于零的常数。在样本数量有限的情况下， k_1 的选择会影响估计精度。估计过程将朝着真实总体分布 $p(x)$ 收敛。该方法的流程详见图 2.6。

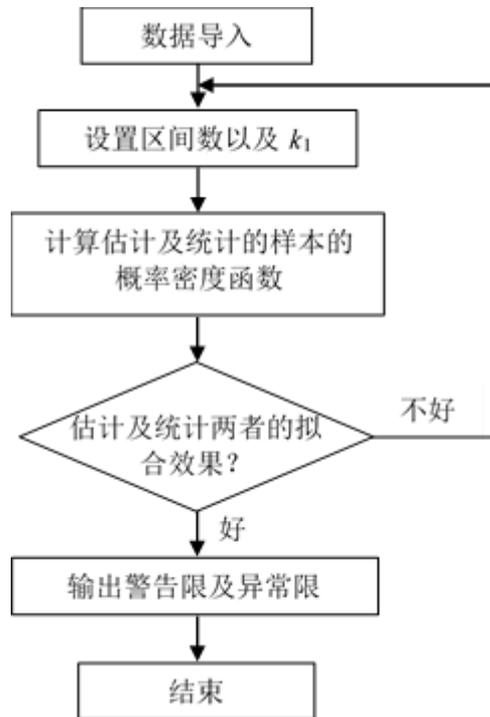


图 2.6 K 近邻法算法流程图

在进行 k_N 近邻估计时，我们面临的主要挑战包括处理大规模数据所需的庞大计算和存储资源，尤其在处理多维数据时这一问题尤为突出。例如，处理一维数据通常需要数百个样本

以达到可靠的分析结果，而处理二维数据时，所需样本数可能上升至数千，随着数据维度的增加，所需样本数以指数级增长，对计算资源的需求相应增加。

2.7 注意事项

（1）浓度限值与浓度梯度限值相辅相成，共同作用于故障的早期识别。例如，某些情况下，尽管磨损元素的浓度未超过预设的限值，其梯度的快速上升可能暗示潜在的异常磨损风险，这要求同时监测这两项指标。

（2）限值应根据设备的使用情况和寿命期进行动态调整。设备不同阶段的油液磨损元素含量会有所不同，更换的零部件和不同的摩擦副磨损状态也可能需要调整现有的限值。例如，美国三军便定期每两年更新一次设备的限值标准，以适应设备的实际使用状况。

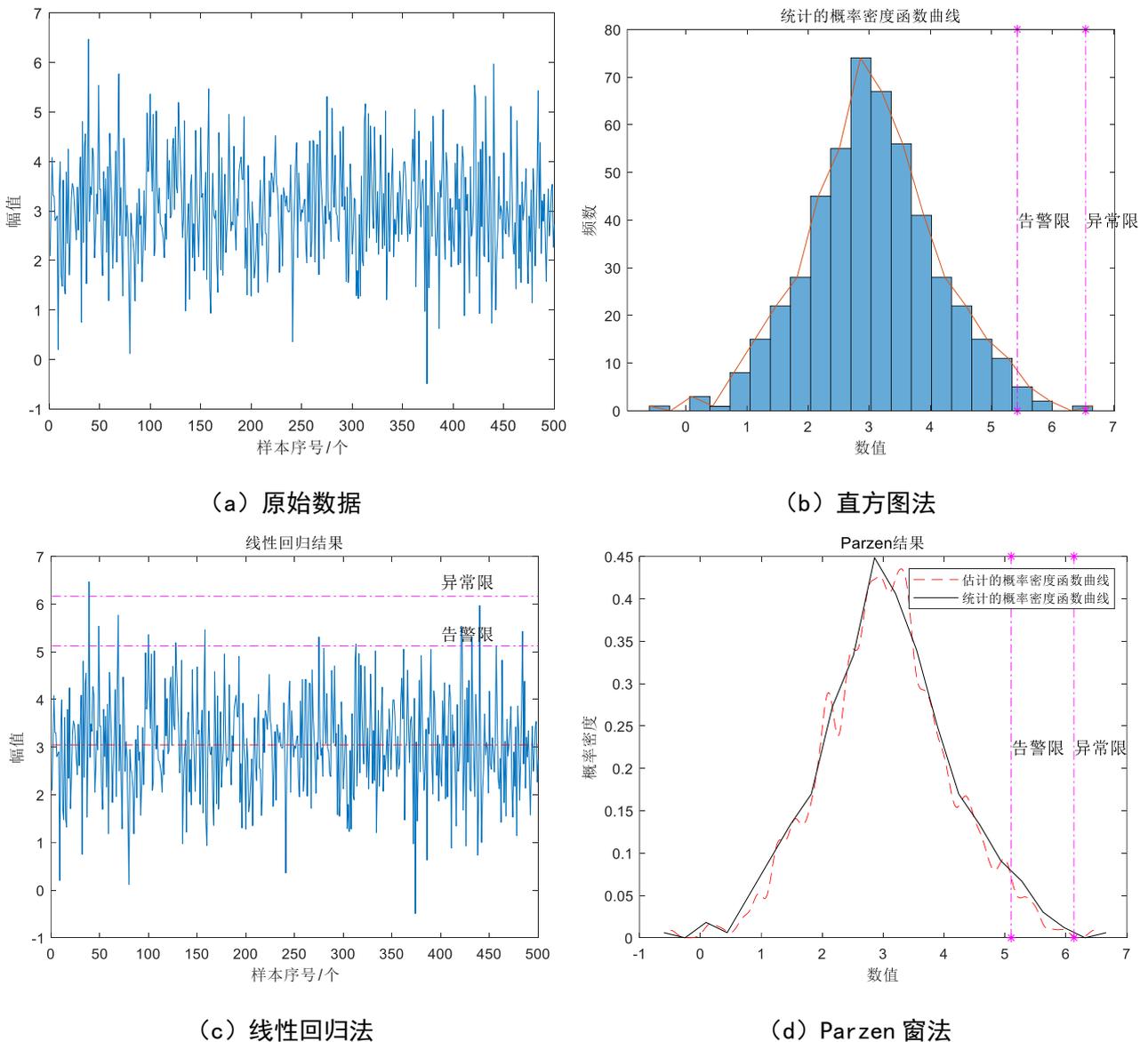
（3）限值的确定往往具有一定的模糊性，通常国内外的标准限值之间的边界差异极小，例如常常只有 $1\mu\text{g}/\text{mL}$ 的差距。这种设定基于确保机器的安全运行，同时留有足够的余地以应对无法立即停机的情况，允许在发现异常指标时有时间进行更详细的取样和分析，从而减少因突然停机导致的损失。此外，磨损进程的逐渐性也要求在限值设置时必须预留足够的操作空间。

第三章 基于仿真信号的界限值制定方法对比验证

在本文的前文中，我们探讨了六种界限值设定的算法及其步骤。为了深入理解这些算法在实际应用中的效果，本研究使用 MATLAB 生成了仿真数据来进行验证。所生成的数据集涵盖了正态分布、F 分布和威布尔分布三种概率分布类型。通过这种方式，可以有效地评估各算法对不同类型数据的处理能力和准确性，进而为实际应用中的数据分析和问题解决提供科学依据。这不仅帮助验证理论的实用性，也展示了各算法在处理复杂数据时的潜力和局限。

3.1 服从正态分布数据集

对于遵循正态分布的数据样本，本研究应用了三种传统的界限值设定方法以及三种基于概率密度函数的估计方法进行分析。每种方法的计算结果以及效果均在图 3.1 中展示。



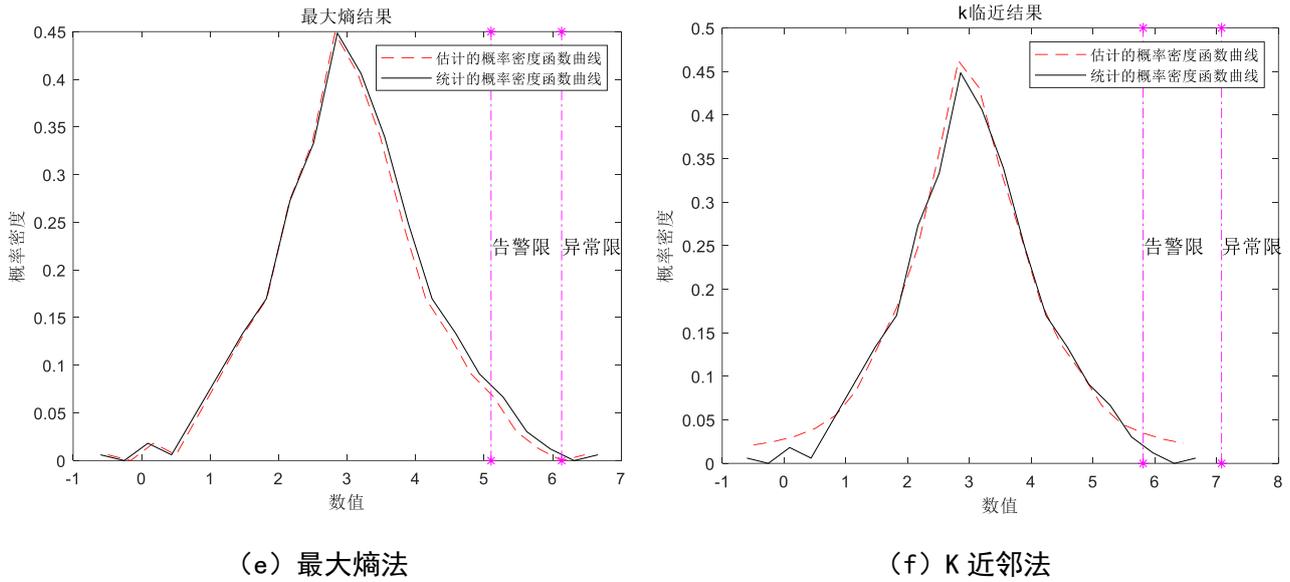


图 3.1 正态分布数据集估计的概率密度函数曲线

图 3.1(a) 显示了正态分布数据的原始分布图，图 3.1(b) 展示了相对应的频数分布图，其频数分布与理论上的正态分布钟形曲线非常接近。图 3.1(c) 则描绘了利用线性回归方法制定界限值的效果。由图可知，数据并未呈现明显的上升趋势。拟合线、警告限及异常限等曲线基本保持水平，这意味着所制定的界限值在时间上并不随之变化；图 3.1 (d)、(e)、(f) 为概率密度估计的界限值制定方法，可以看出各种方法估计出的概率密度曲线与统计结果吻合良好，因此由此制定出的界限值比较科学可靠。

此外，这些方法确定的警告限和异常限详见表 3.1，本文还将就这些方法在界定正态分布数据的界限值的效果进行详细比较和讨论。这种综合分析有助于评估各方法的有效性，并指导实际应用中的选择。由表 3.1 可知，各方法的界限值制定结果有稍许不同，但是数值大小基本上相差不大，其中 K 近邻法计算得到的界限值稍大于其他 5 种方法，但仍可以认定是在合理误差范围之内；故可以认为面对服从正态分布的数据，这 6 种界限值的制定方法都适用。

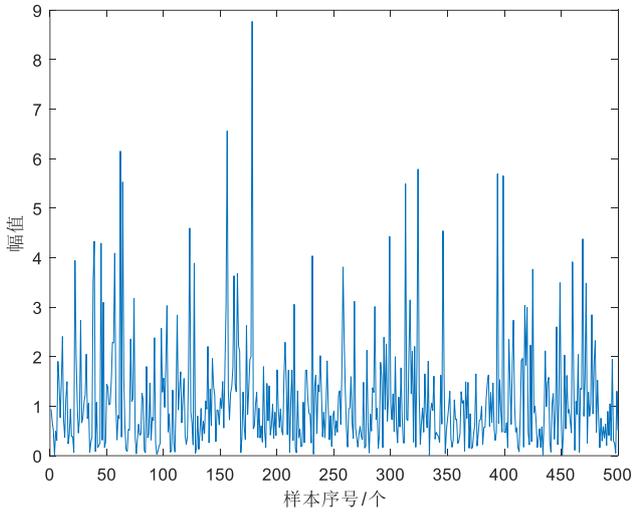
表 3.1 服从正态分布数据集计算结果

方法	直方图法	回归法	增长率法	Parzen窗法	最大熵法	K近邻法
参数值				$h=1$	区间10, 约束5	$k=2$
警告限	5.4263	5.1250	5.1273	5.0461	5.0949	5.8111
异常限	6.5450	6.1634	6.1669	6.0656	6.1360	7.0797

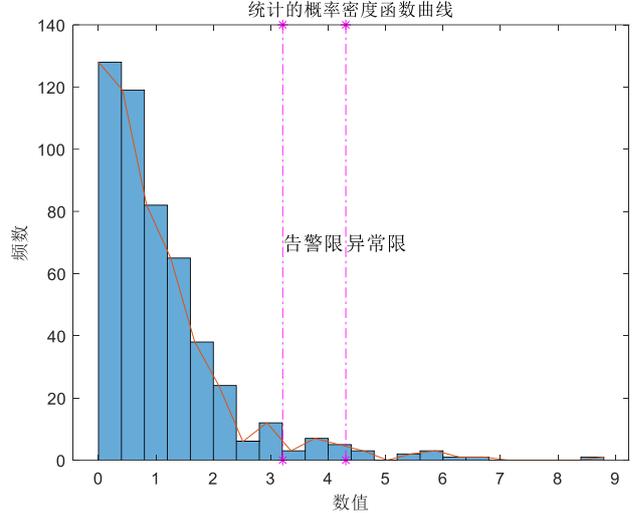
3.2 服从 F 分布数据集

在分析实际航空发动机附件机匣中的滑油磨损元素数据时，发现其分布特征显示在低浓度区间的样本数据量大，频数高，而在高浓度区间的样本数据量较小，频数低，这与 F 分布的特点相似。因此，本研究采用符合 F 分布的数据进行更深入的分析。

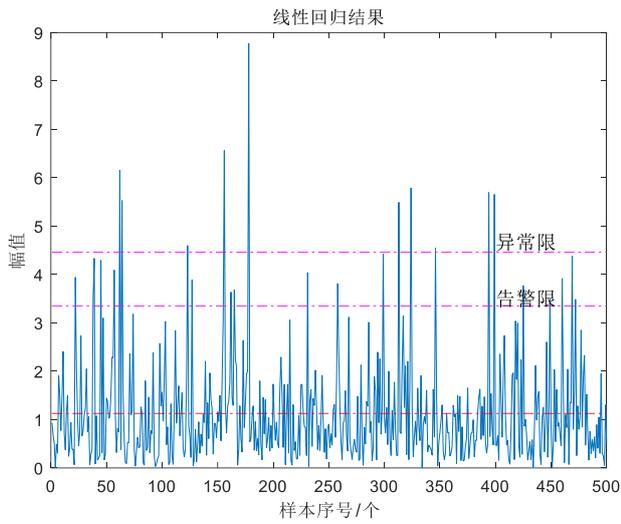
本文运用了多种统计方法，包括直方图法、Parzen 窗法、最大熵法和 K 近邻法来设定界限值，相应的图形曲线展示在图 3.2 中，而从这些方法得出的警告限和异常界限值则详细列在表 3.2 中。这些分析工具帮助我们更精确地定义和理解数据的统计界限，为未来的数据处理和决策提供科学依据。



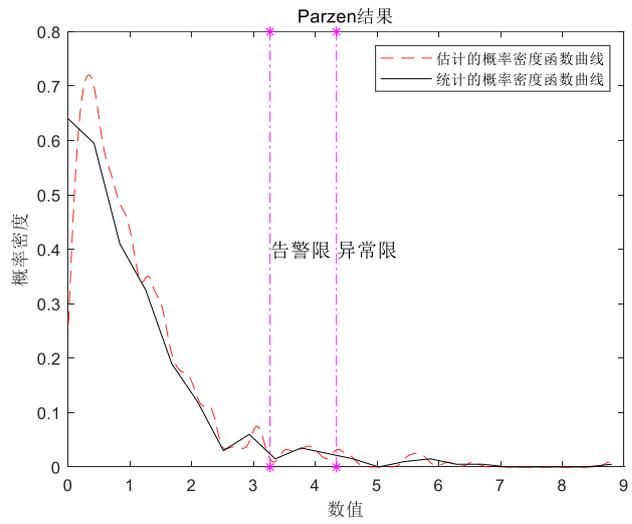
(a) 原始数据



(b) 直方图法



(c) 线性回归法



(d) Parzen 窗法

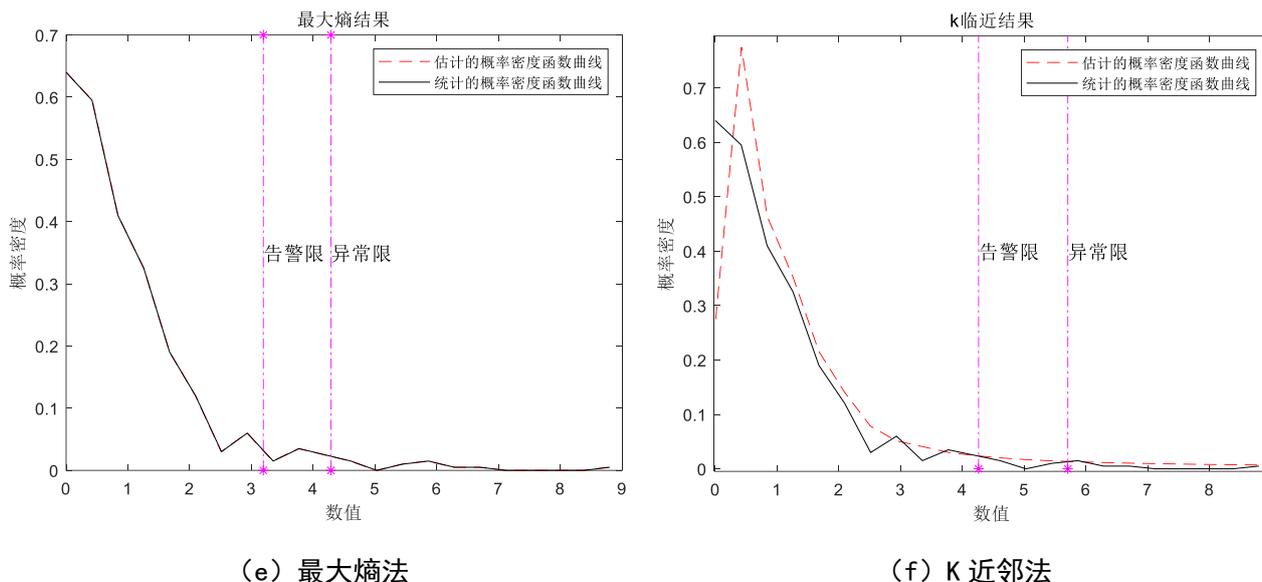


图 3.2 F 分布数据集估计的概率密度函数曲线

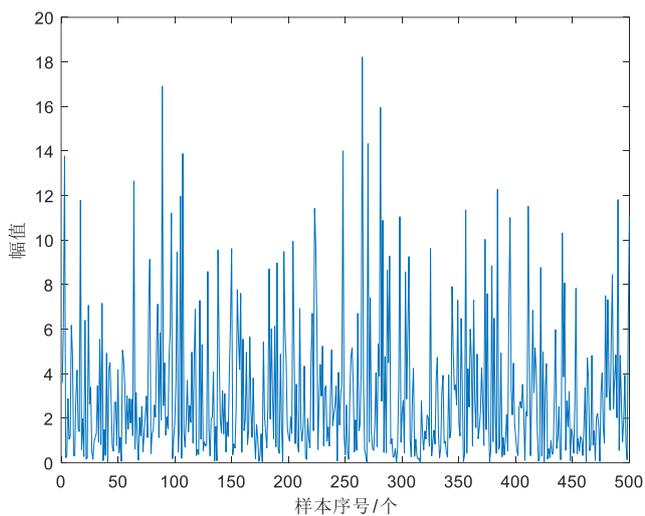
从图 3.2(e)可以看出，最大熵法得到的频数图与 F 分布的数据特征吻合，尤其是在低浓度区间样本频数较高，而高浓度区间则显著降低。此外，图 3.2(b)、(c)、(d)展示的数据分布情况也明显呈现出与 F 分布相似的趋势，表明这些统计方法能有效反映出数据的分布特性。这种分布特征的一致性说明了这些方法在处理符合 F 分布的数据集时的适用性和准确性。并且统计的概率密度函数曲线与估计的概率密度函数曲线拟合效果很好，由此制定出的界限值较为科学可靠，界限值结果汇总见表 3.2。

表 3.2 服从 F 分布数据集计算结果

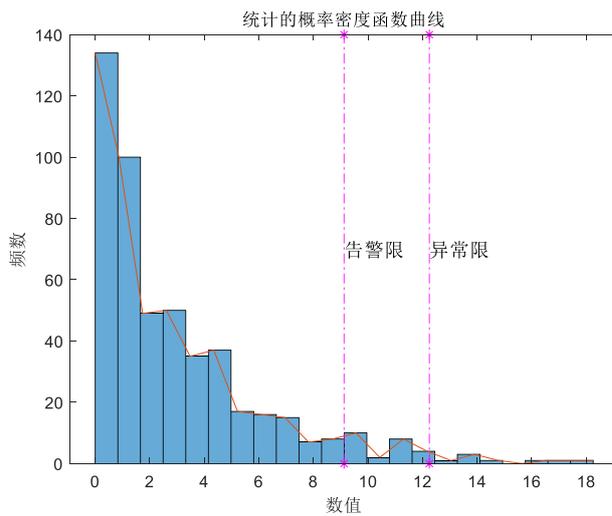
方法	直方图法	回归法	增长率法	Parzen窗法	最大熵法	K近邻法
参数值				$h=1$	区间10, 约束5	$k=2$
警告限	3.2072	3.3499	3.3533	3.2691	3.1927	4.2618
异常限	4.3062	4.4610	4.4662	4.3441	4.2859	5.7067

3.3 服从威布尔分布数据集

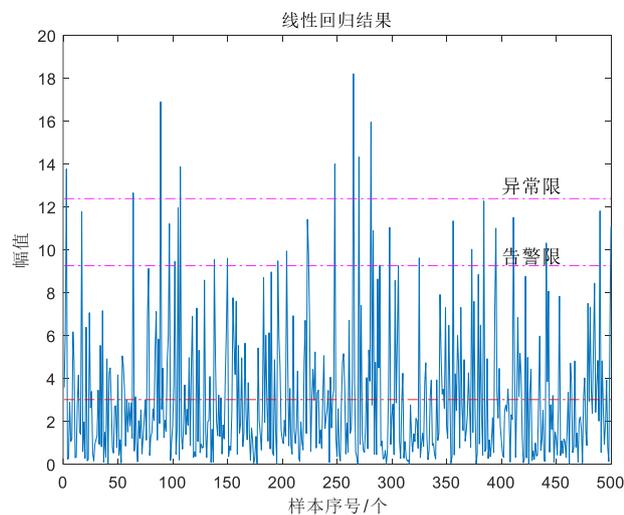
威布尔分布广泛应用于产品可靠性研究，尤其适合分析磨损失效数据。鉴于航空发动机附件机匣在运行过程中的磨损失效特性与威布尔分布的相似性，本研究进一步探讨了使用威布尔分布进行界限值设定的有效性。为此，我们比较了不同方法在处理符合威布尔分布的样本数据时的效果，并通过图 3.3 展示了这些方法的拟合结果，以便全面评估各种方法在设定界限值时的适用性和准确性。



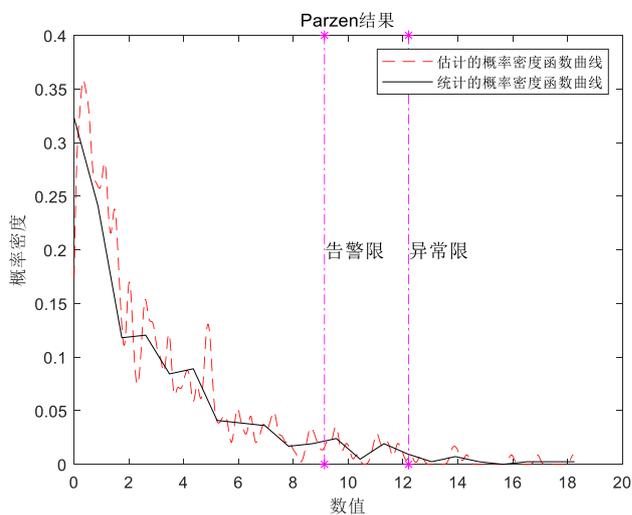
(a) 原始数据(a) 原始数据



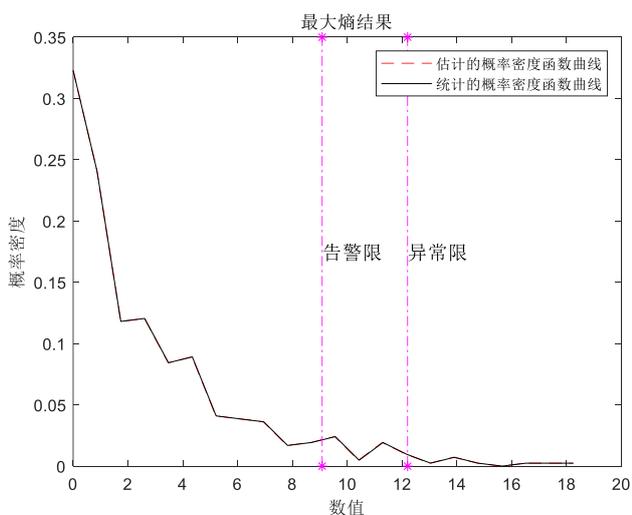
(b) 直方图法



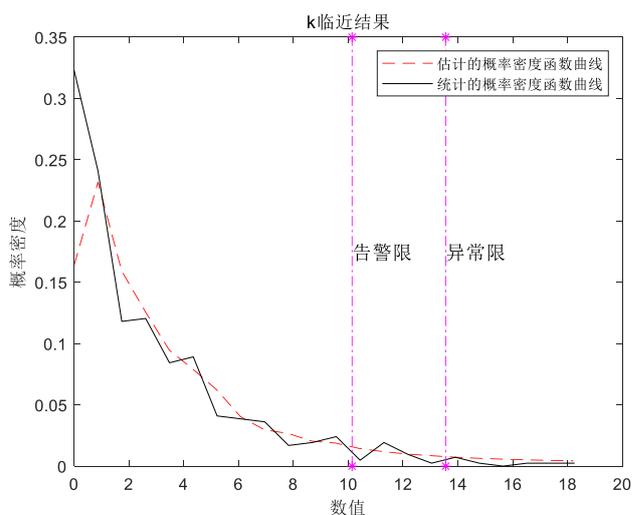
(c) 线性回归法



(d) Parzen 窗法



(e) 最大熵法



(f) K 近邻法

图 3.3 威布尔分布数据集估计的概率密度函数曲线

由图 3.3 (b)、(e) 可知，统计出的频数图与最大熵法得到的概率密度函数十分吻合。而图 3.3 (d) 中，Parzen 窗得到概率密度函数出现了较大的波动，图 3.3 (f) 中，K 近邻法得到的概率密度函数无法完全反应数据的波形性，但两者概率密度函数曲线变化趋势与统计的概率密度曲线一致，在低浓度区间，样本数据量很大，频数高；而在高浓度区间，数据量较少，频数低，能够反映威布尔分布的实际分布情况。

为了更加直白的对各个方法得到的界限值准确度的判断，论文也对服从威布尔分布的结果进行了相关的统计，记录的结果见下表：

表 3.3 服从威布尔分布数据集计算结果

方法	直方图法	回归法	增长率法	Parzen窗法	最大熵法	K近邻法
参数值				$h=1$	区间10, 约束5	$k=2$
警告限	9.1231	9.2654	9.0512	9.4548	9.0864	10.1414
异常限	12.2513	12.3896	12.0683	12.5901	12.2032	13.5563

根据表 3.3 的结果，各方法确定的界限值存在一定差异，尽管这些差异相对于基于 F 分布的界限值来说较小。这表明，仿真生成的服从威布尔分布的数据相比服从 F 分布的数据更接近于正态分布。此外，从实际曲线图中也能观察到数据分布与正态分布曲线更为接近。因此，这些差异说明了各方法在界限值确定上的差异。

当样本数据符合正态分布时，数据的界限值往往通国六种界限值的方法来确定，但是，结果不是符合正态分布的，那么假设的这种方法就不再适用于实际应用。此外，当样本数据的分布与正态分布相似时，界限值的制定结果差异较小；相反，如果分布差异较大，则制定的界限值也会存在显著差异。

第四章 基于附件机匣磨损元素的界限值制定方法对比验证

在本文上一章节中使用 3 种不同分布的仿真数据对各种算法的效果进行了对比验证。下面，将使用真实航空发动机附件机匣 Cu、Ag、Ni、Mg 四种磨损元素进行界限值制定，充分对比验证各种方法的差异，并对实际试车中磨损元素界限值制定提供可行思路与途径。

4.1 真实航空发动机附件机匣滑油光谱数据集

本章将采用真实航空发动机附件机匣滑油光谱数据集来验证上述方法，该数据集包含 Cu、Ag、Ni、Mg 四种元素随试车进程的变化量，每种元素数量为 26 个，如表 4.1 所示。

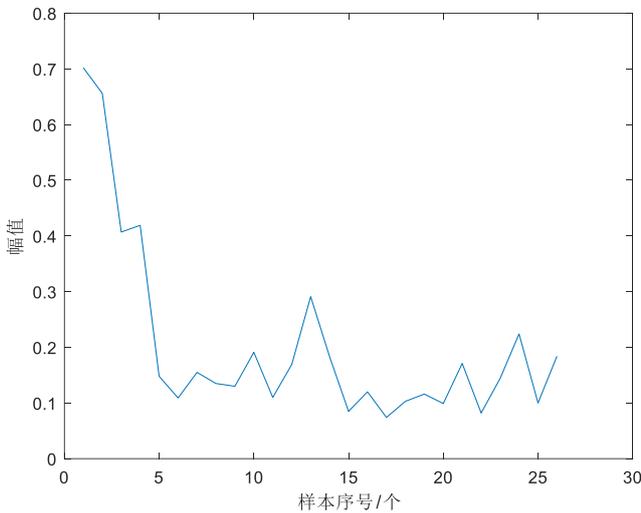
表 4.1 真实航空发动机附件机匣滑油光谱数据集

Cu/ppm	Ag/ppm	Ni/ppm	Mg/ppm
0.702	0.035	0.251	0.453
0.656	0.040	0.015	0.571
0.407	0.035	0.112	0.247
0.419	0.020	0.028	0.379
0.148	0.023	0.001	0.181
0.109	0.008	0.145	0.335
0.155	0.021	0.064	0.334
0.135	0.069	0.153	0.260
0.130	0.048	0.195	0.269
0.191	0.054	0.177	0.412
0.110	0.048	0.107	0.274
0.169	0.050	0.220	0.349
0.291	0.044	0.172	0.430
0.183	0.049	0.203	0.342
0.085	0.047	0.195	0.529
0.120	0.048	0.136	0.251
0.074	0.033	0.146	0.394
0.103	0.038	0.129	0.393
0.116	0.044	0.086	1.206
0.099	0.034	0.133	1.286
0.171	0.055	0.098	1.407
0.082	0.035	0.152	0.689
0.144	0.045	0.087	1.155

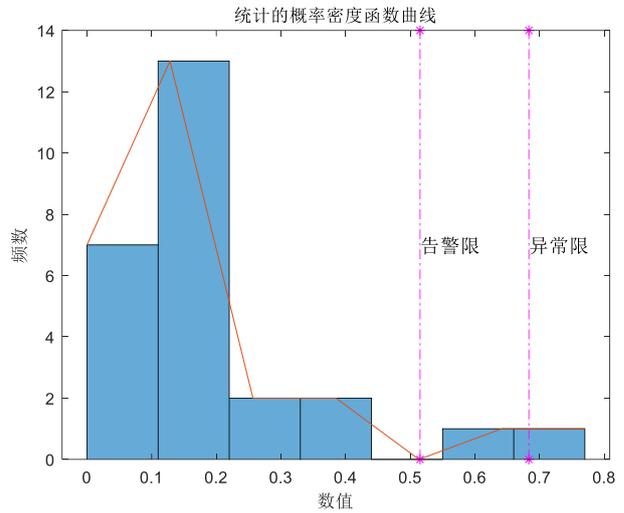
0.224	0.060	0.102	0.940
0.100	0.038	0.103	0.510
0.184	0.041	0.089	0.914

4.1 Cu 元素界限值制定

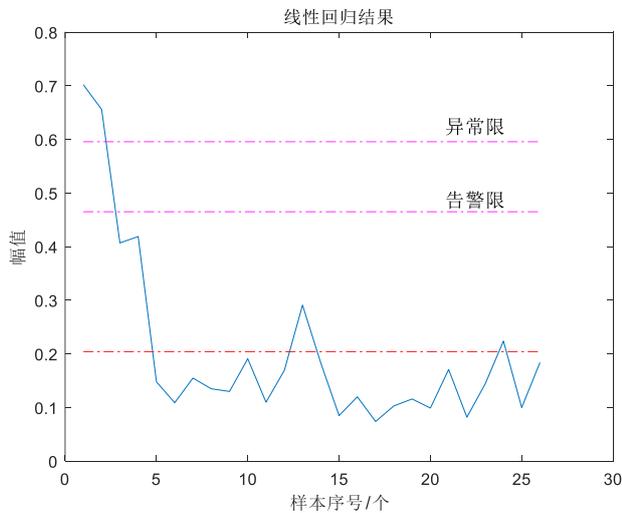
对真实航空发动机附件机匣运行过程中的滑油光谱 Cu 元素数据进行分析，利用直方图法、Parzen 窗法、最大熵法和 K 近邻法等方法进行界限值制定时的图形曲线如图 4.1 所示，界限值结果如表 4.2 所示。由图 4.1 所示，最大熵法的概率密度函数虽与统计得到的概率密度函数在低浓度范围下有一定的偏差，但曲线整体吻合效果要比 Parzen 窗法以及 K 近邻法效果好。但是，由表 4.2 可知，最大熵法异常限偏小，K 近邻法的警告限偏大；同时，最大熵法警告限、异常限与 Parzen 窗法的警告限接近。



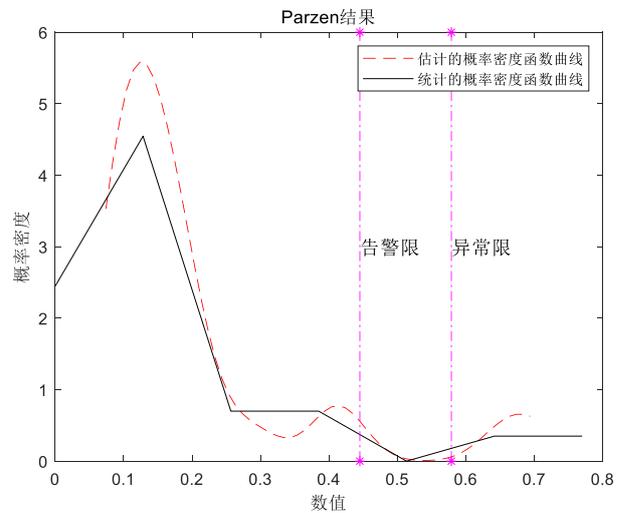
(a) 原始数据



(b) 直方图法



(c) 线性回归法



(d) Parzen 窗法

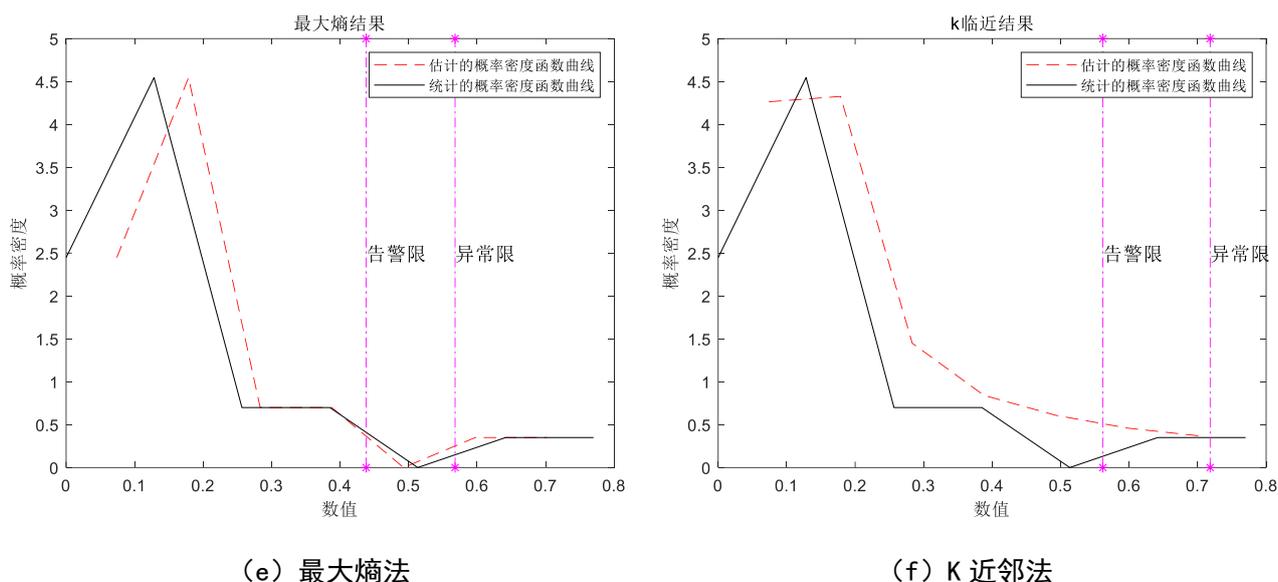


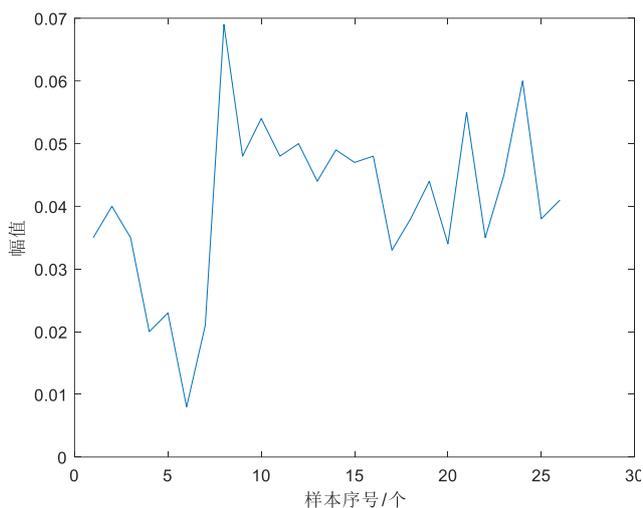
图 4.1 Cu 元素估计的概率密度函数曲线

表 4.2 某型航空发动机附件机匣滑油光谱 Cu 元素磨损数据界限值

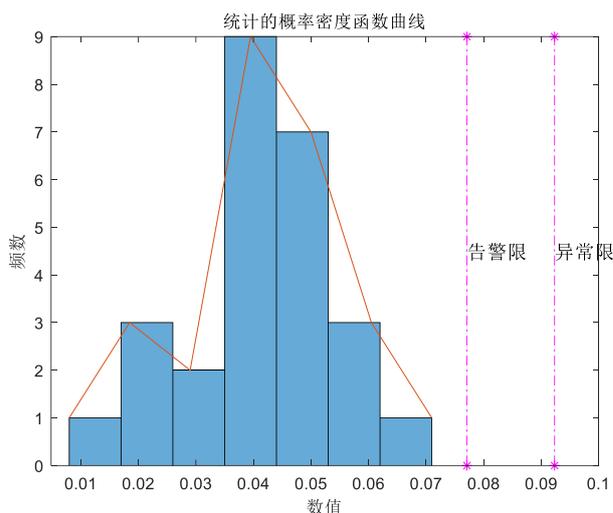
方法	直方图法	回归法	增长率法	Parzen窗法	最大熵法	K近邻法
警告限	0.5151	0.4652	0.5342	0.4451	0.4380	0.5616
异常限	0.6834	0.5958	0.6993	0.5789	0.5679	0.7183

4.2 Ag 元素界限值制定

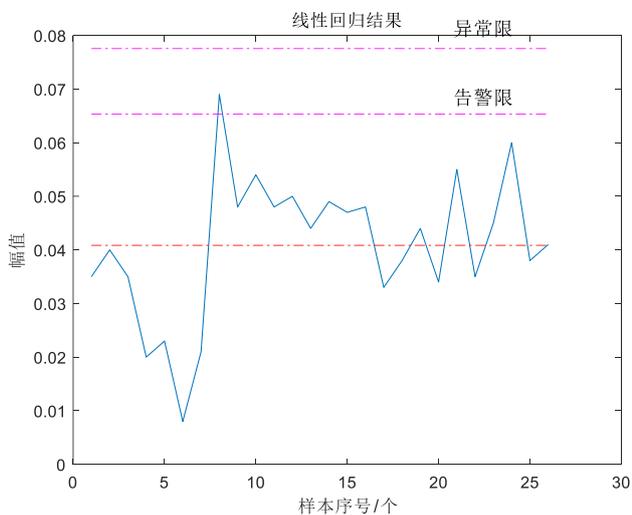
对真实航空发动机附件机匣运行过程中的滑油光谱 Ag 元素数据进行分析，在使用直方图法、Parzen 窗法、最大熵法以及 K 近邻法等多种方法设定界限值时，图 4.2 展示了这些方法的图形曲线对比，而表 4.3 则详细列出了各方法设定的界限值结果。从图 4.2 的分析可以看出，最大熵法在拟合概率密度函数方面相比于 Parzen 窗法和 K 近邻法展现出更优的吻合效果，这表明最大熵法在处理本数据集时可能提供更准确的结果预测。这种比较有助于评估各种方法在实际应用中的效果，为选择最适合的方法提供科学依据。这能够反映 Ag 元素数据的整齐变化趋势，而 Parzen 法和 K 近邻法无法准确的反映出数据的分布情况。此外，由表 4.3 可知，K 近邻法计算得到的警告限和异常限偏大；同时，最大熵法计算得到的警告限与异常限要略大于 Parzen 窗法得到的计算结果。



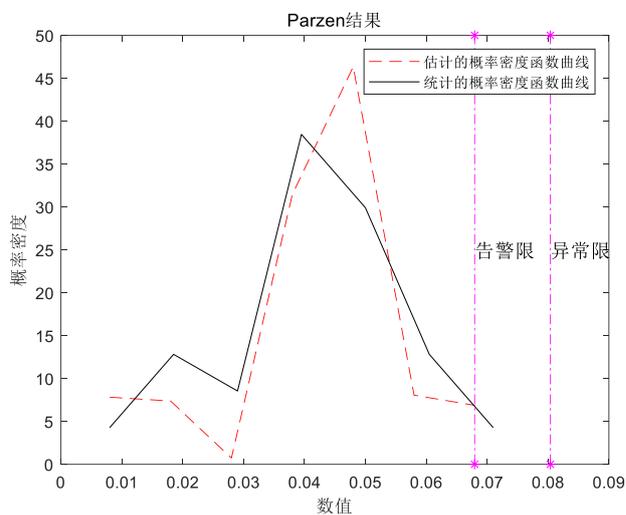
(a) 原始数据



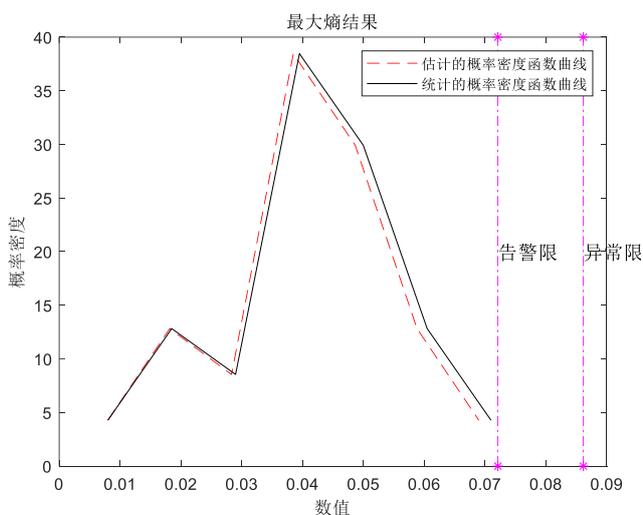
(b) 直方图法



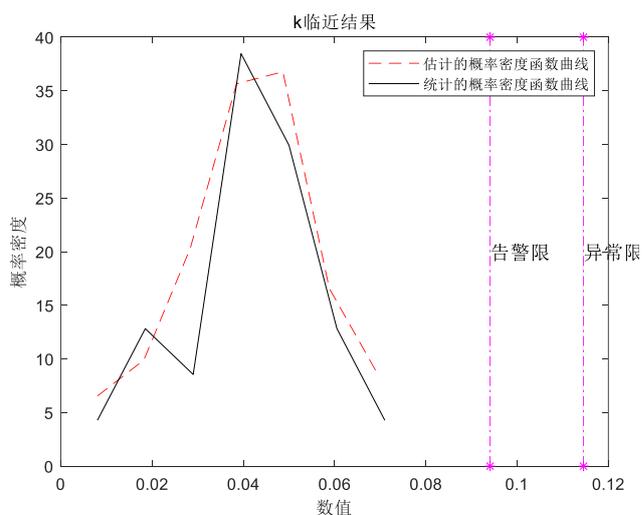
(c) 线性回归法



(d) Parzen 窗法



(e) 最大熵法



(f) K 近邻法

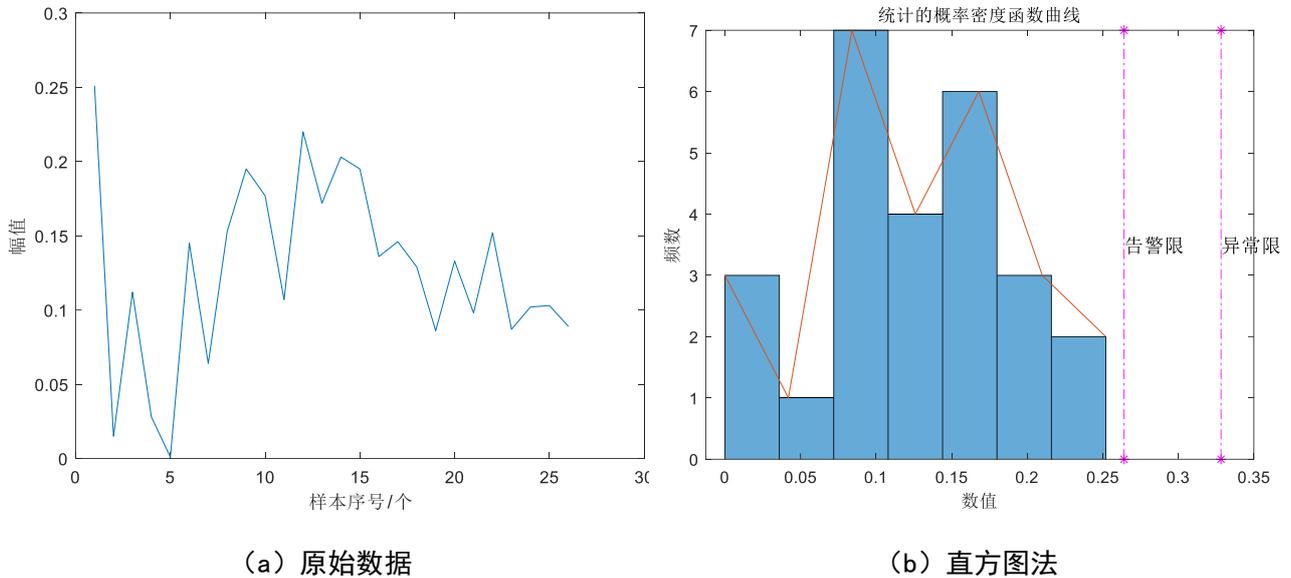
图 4.2 Ag 元素估计的概率密度函数曲线

表 4.3 某型航空发动机附件机匣滑油光谱 Ag 元素磨损数据界限值

方法	直方图法	回归法	增长率法	Parzen窗法	最大熵法	K近邻法
警告限	0.0771	0.0653	0.0673	0.0679	0.0721	0.0941
异常限	0.0923	0.0775	0.0805	0.0804	0.0862	0.1146

4.3 Ni 元素界限值制定

对真实航空发动机附件机匣运行过程中的滑油光谱 Ni 元素数据进行分析, 在使用直方图法、Parzen 窗法、最大熵法和 K 近邻法等多种统计方法确定界限值时, 各方法的效果通过图 4.3 中的图形曲线进行了展示, 同时具体的界限值结果列于表 4.4。根据图 4.3 的显示, 最大熵法在概率密度函数的拟合上相较于 Parzen 窗法和 K 近邻法展现了更佳的效果, 这指示了其在精确匹配实际数据分布方面的优越性。这些发现有助于指导实际应用中针对不同方法的选择和使用。可更加准确的反映出 Ni 元素数据的分布情况, 而 Parzen 窗法和 K 近邻法无法准确反映出数据的分布的波动和差异。此外, 由表 4.4 可知, 最大熵法与 K 近邻法结果相近, Parzen 窗法得到的结果要小于最大熵法与 K 近邻法。



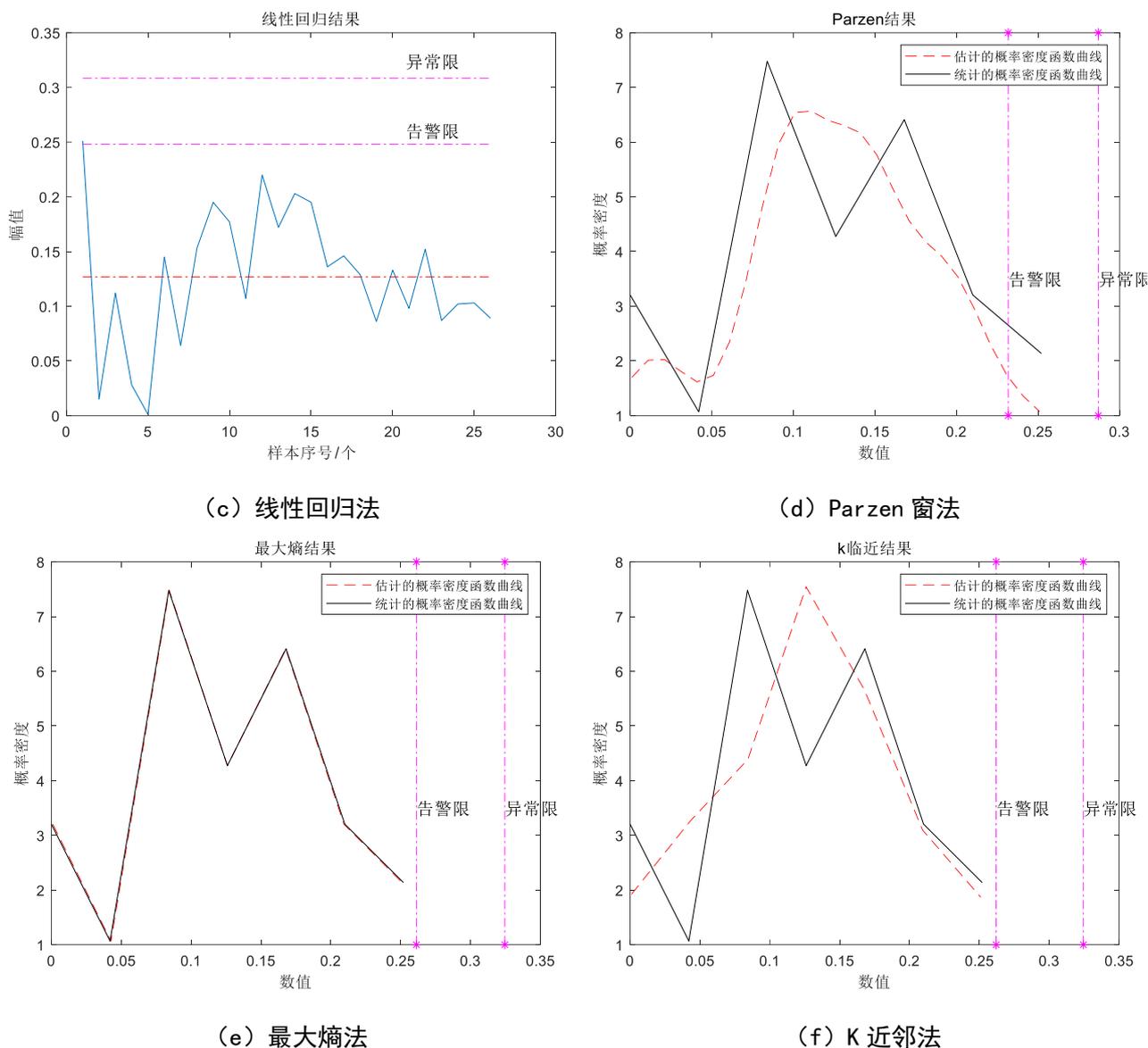


图 4.3 Ni 元素估计的概率密度函数曲线

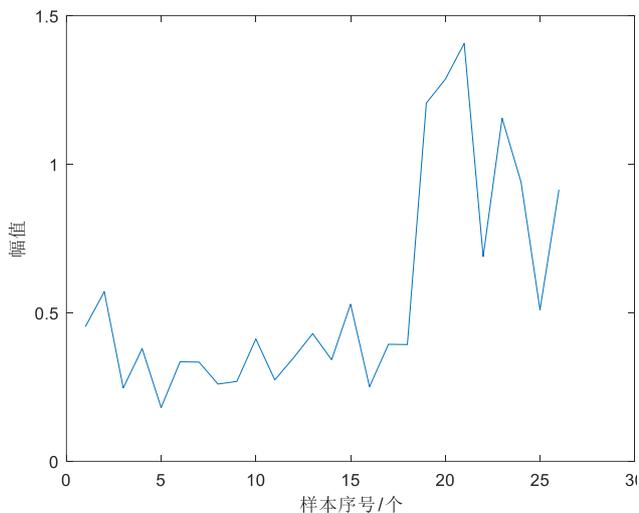
表 4.4 某型航空发动机附件机匣滑油光谱 Ag 元素磨损数据界限值

方法	直方图法	回归法	增长率法	Parzen窗法	最大熵法	K近邻法
警告限	0.2641	0.2479	0.2503	0.2318	0.2612	0.2620
异常限	0.3283	0.3085	0.3121	0.2869	0.3246	0.3244

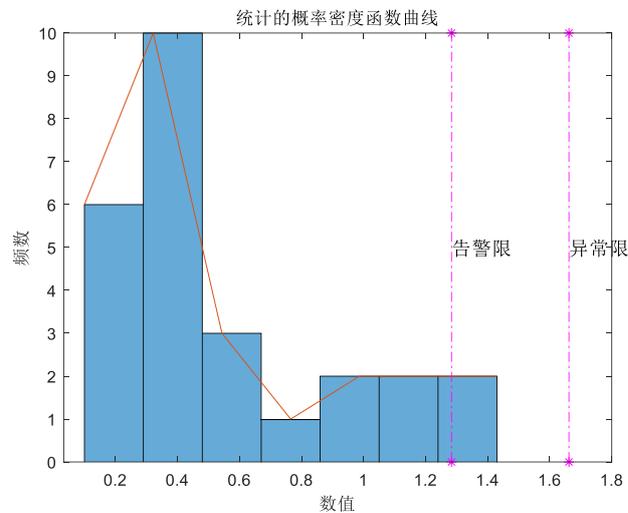
4.4 Mg 元素界限值制定

对真实航空发动机附件机匣运行过程中的滑油光谱 Mg 元素数据进行分析，利用直方图法、Parzen 窗法、最大熵法和 K 近邻法等方法进行界限值制定时的图形曲线如图 4.4 所示，界限值结果如表 4.5 所示。由图 4.4 所示，最大熵法的概率密度函数曲线吻合效果相比 Parzen

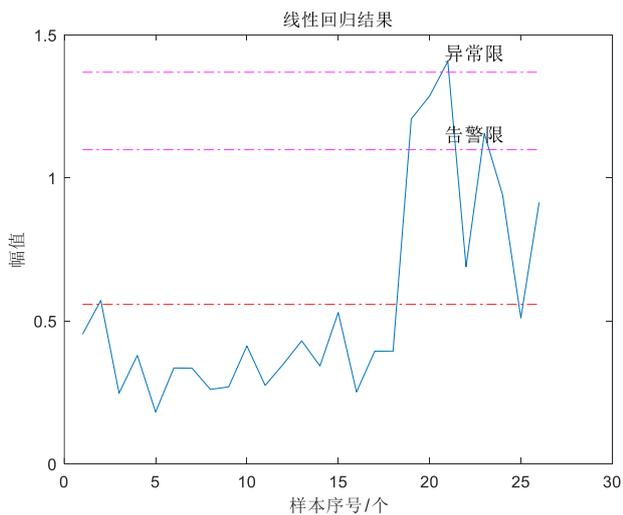
窗法以及 K 近邻法效果好。此外，由表 4.5 可知，最大熵法计算结果大于 Pazren 窗法，小于 K 近邻法。



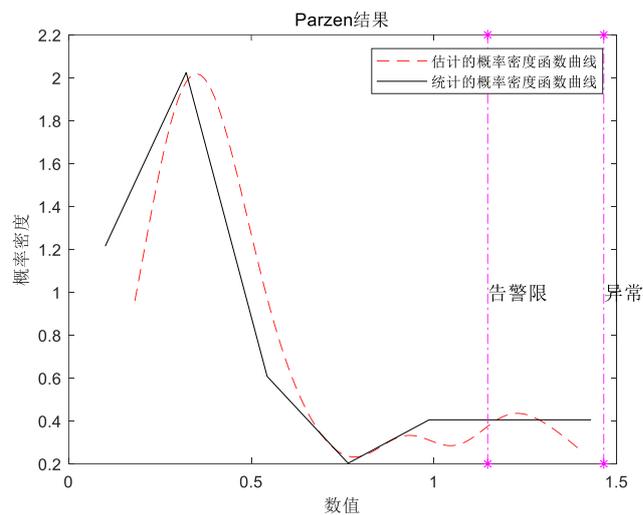
(a) 原始数据



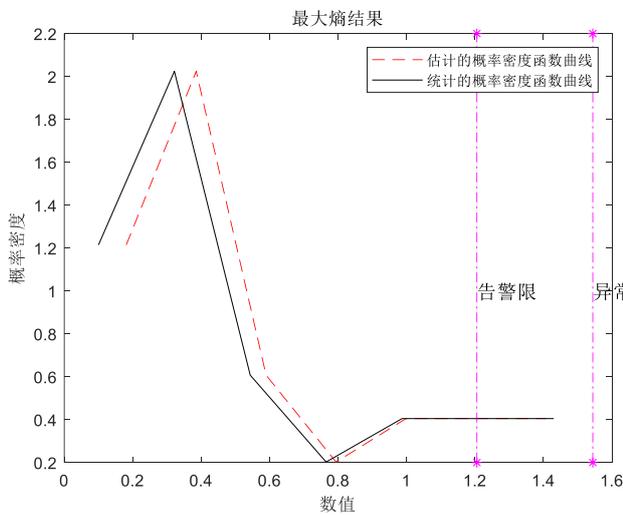
(b) 直方图法



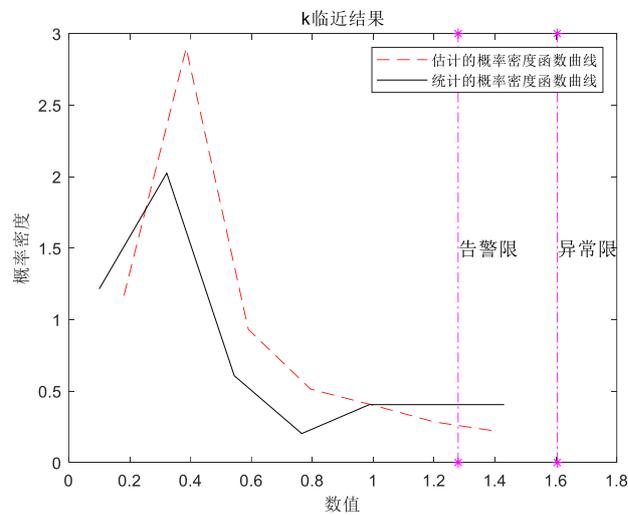
(c) 线性回归法



(d) Parzen 窗法



最大熵结果



k 临近结果

(e) 最大熵法

(f) K 近邻法

图 4.4 Mg 元素估计的概率密度函数曲线

表 4.5 某型航空发动机附件机匣滑油光谱 Mg 元素磨损数据界限值

方法	直方图法	回归法	增长率法	Parzen窗法	最大熵法	K近邻法
警告限	1.2834	1.0988	1.2771	1.1471	1.2054	1.2788
异常限	1.6619	1.3691	1.6366	1.4650	1.5440	1.6041

第五章 总结与展望

5.1 研究总结

本研究探讨了航空发动机附件机匣的滑油光谱数据阈值设定问题，并提出了几种新的阈值定义方法。传统的阈值设定常基于假设数据遵循正态分布，但实际观察表明，这一假设并非总是成立。为了克服此限制，本文采用了基于概率密度函数的方法，通过分析大量数据的统计特性和概率密度，来更精确地确定阈值，从而提高阈值设定的科学性和准确性。此方法通过细致的数据分析，旨在实现更合理的阈值设定，以适应复杂的实际应用需求。

5.2 研究展望

本文仅是对各种磨损元素进行单独的界限值制定，如何将多种元素进行融合，从而进行界限值的融合制定，是之后的研究方向。

参考文献

- [1] 陈果, 李爱. 航空器检测与诊断技术导论[M]. 北京: 航空工业出版社, 2012.
- [2] 汪瑾, 陈果, 王洪伟, 等. 基于后验概率 SVM 的航空发动机滚动轴承状态评估[J]. 轴承, 2015(5):55-59.
- [3] 杨士亮, 郝敬团, 杨宏伟, 等. 飞机发动机油液故障诊断技术的现状与发展趋[J]. 当代化工, 2014(11):2421.2422.
- [4] 陈果, 左洪福, 杨新. 基于神经网络的多种油样分析技术融合诊断[J]. 摩擦学学报, 2003, 23(5):431.434.
- [5] 宋兰琪, 汤道宇, 陈立波, 等. 航空发动机滑油光谱专家系统知识库建立[J]. 航空学报, 2000, 21(5):453.456.
- [6] 万耀青, 郑长松, 马彪. 原子发射光谱仪作油液分析故障诊断的界限值问题[J]. 机械强度, 2006, 28(4):485.488.
- [7] 曹轶乾, 李克武, 曹成玉, 等. 发动机油液光谱分析界限值动态调整问题研究[J]. 测控技术, 2013, 32(10):155.158.
- [8] 张培林. 自行火炮油液光谱分析研究[D]. 南京理工大学, 2003.
- [9] 吴汪洋, 张巍. 基于光谱分析趋势值的机械设备油液监测方法研究[J]. 液压与气动, 2009(9):77.79.
- [10] 宋兰琪, 毛美娟. 航空发动机滑油光谱故障诊断专家系统[J]. 航空发动机, 1999(1):45.49.
- [11] 闫辉, 贺石中, 毛宁, 等. 油液监测界限值传统计算方法的改进[J]. 润滑与密封, 2005(4):81.83.
- [12] 李爱, 陈果. 基于 SVM 的航空发动机油样光谱诊断界限值制定[J]. 航空动力学报, 2011, 26(4):771.778.
- [13] 孙金哲, 陈军, 沈兴国, 等. 基于 GM(0,N)模型的油液光谱分析界限值制定[J]. 润滑与密封, 2012(8):103.106.
- [14] 朱焕勤, 张子阳, 费逸伟. 基于积分方式变化的润滑油 MOA 检测界限值动态调整问题研究[J]. 润滑与密封, 2008, 33(7):84.86.
- [15] 周平, 刘东风, 石新发, 等. 基于稳健回归的油液光谱分析界限值制定[J]. 润滑与密封, 2010, 35(5):85.88.
- [16] 万耀青, 郑长松, 马彪. 原子发射光谱仪作油液分析故障诊断的界限值问题[J]. 机械强度, 2006, 28(4):485.488.
- [17] 张玉峰. 智能信息系统[M]. 武汉: 武汉大学出版社, 2008.
- [18] 李龙康. 基于智能网格分类的天然气短期负荷预测模型的研究与实现[D]. 上海师范大学, 2015.
- [19] 高雪. K-means 聚类算法在面板数据分析中的改进及实证研究[D]. 太原理工大学, 2015.
- [20] 毛宇星, 陈彤兵, 施伯乐. 一种高效的多层和概率关联规则挖掘方法[J]. 软件学报, 2011, 22(12):2965.2980.
- [21] 邹翔, 张巍, 刘洋, 等. 分布式序列模式发现算法的研究[J]. 软件学报, 2005, 16(7):1262.1269.
- [22] 张鹏, 张建业, 杜军, 等. 基于神经网络的时间序列相似模式发现方法[J]. 模式识别与人工智能, 2008, 21(3):401.405.
- [23] 邓超, 熊范纶. 数据库中混沌模式的发现和预测[J]. 软件学报, 1999, 10(5):469.474.
- [24] 侯春燕, 李云春, 李巍. 基于分布式环境一种动态依赖关系发现的方法[J]. 微计算机信息, 2006, 22(30):195.197.
- [25] 高强, 李良敏, 孟庆丰, 等. EMD 趋势分析方法及其应用研究[J]. 振动与冲击, 2007, 26(8):98.100.
- [26] Iverson D. Data Mining Applications for Space Mission Operations System Health Monitoring[C]// 2008.
- [27] Martinezheras J, Donati A. Data Mining at the Service of Space Operations[C]// SpaceOps 2006 Conference. 2006.
- [28] Chiba K, Obayashi S. Knowledge Discovery for Flyback.Booster Aerodynamic Wing Using Data Mining[J]. Journal of Spacecraft & Rockets, 2015, 45(5):975.987.

致谢

感谢党和国家！