

装甲车辆润滑油监测数据的多元回归分析

张玲 李卫东¹ 吕玺² 董玉才

(装甲兵工程学院非线性研究所, 科研部研究所¹, 北京 100072; 西安军代局驻 845 厂军代室², 西安 710302)

摘要 利用多元回归分析方法对装甲车辆变速箱润滑油监测数据进行了分析, 说明了油样中铁谱磨粒浓度和油液污染度之间的线性相关性, 并建立了两者的线性模型, 为预测油液污染度提供了一种方法。

关键词 油液污染度 铁谱数据 相关系数 多元回归分析

中图法分类号 TJ810.321; **文献标志码** A

污染度和铁系金属含量经常用来判别被测油样的质量以及车辆传动系统磨损程度。大量油液检测试验结果表明: 重型车辆传动系统润滑油中, 铁系金属磨粒的含量和油样的污染度存在着很强的关联性。因此可以通过建立模型来分析这 2 个指标间的定量关系, 以达到通过其中一个指标判断另一个指标发展趋势的目的。

本文通过回归分析的统计方法, 分析变量之间的统计关系, 考察变量之间的数量变化规律, 并通过建立回归方程的形式描述和反映这种关系。

1 多元回归原理

在回归分析之前, 要研究变量之间线性关系的紧密程度。Pearson 简单相关系数可以用来度量定距型变量间的线性相关关系, 其计算公式为

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}},$$

其中, n 为样本容量, x_i 和 y_i 为两变量对应的样本值。Pearson 简单相关系数的检验统计量为 t 统

计量,

$$t = \frac{r \sqrt{n-2}}{\sqrt{1-r^2}}。$$

其中 t 统计量服从自由度为 $n-2$ 的 t 分布。一般认为, 当相关系数的绝对值大于 0.8 时, 两个变量之间具有较强的线性关系。

若变量之间具有较强的线性关系, 则可以进行线性回归分析。多元线性回归是研究一个随机变量与多个变量之间相关关系的方法, 其数学模型为

$$\begin{cases} y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_m x_m + \varepsilon \\ \varepsilon \sim N(0, \sigma^2) \end{cases}$$

现在的任务是根据 n 个样本值 $(x_{i1}, x_{i2}, \cdots, x_{im}; y_i)$, $i = 1, 2, \cdots, n$ 求未知参数 $b_0, b_1, b_2, \cdots, b_m$ 的最小二乘估计值 $\hat{b}_0, \hat{b}_1, \hat{b}_2, \cdots, \hat{b}_m$, 从而得到 m 元线性回归方程

$$\hat{Y} = \hat{b}_0 + \hat{b}_1 x_1 + \hat{b}_2 x_2 + \cdots + \hat{b}_m x_m。$$

确定模型后, 还必须对回归方程进行各种统计检验, 包括模型的显著性检验、回归系数的显著性检验、残差分析等。常用的 F 检验法的统计量为

$$F = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \cdot \frac{(n-2)}{1}。$$

其中 $\bar{y}$ 表示被解释变量 y 的样本均值, 统计量 F 服从自由度为 $(1, n-2)$ 的 F 分布。

2 实验数据的获取

实验油样来自于某中型坦克变速箱,由同一车辆变速箱中获取实验数据^[1],包括该油液中大磨粒的浓度 D_L ;直径为(1~2) μm 的小磨粒的浓度 D_S ;坦克变速箱中润滑油的污染度。具体数据如表 1 所示。

表 1 油样污染度和直读铁谱数据

油样号	污染度/($\text{mg} \cdot \text{L}^{-1}$)	直读铁谱数据/(%)	
		D_L	D_S
1	25.03	88.9	64.21
2	24.97	64.21	46.01
3	28.19	95.01	70.61
4	21.82	58.16	32.20
5	35.44	131.8	103.5
6	35.16	110.51	80.14

3 油液污染度与铁谱数据回归分析

3.1 污染度与铁谱数据的相关分析

利用 SPSS 软件^[2]将实验数据进行相关分析,计算结果如表 2。

表 2 相关系数

		污染度	D_L	D_S
Pearson 相关系数	污染度	1.000	0.928	0.915
	D_L	0.928	1.000	0.992
	D_S	0.915	0.992	1.000
显著性检验 的概率值	污染度	.	0.004	0.005
	D_L	0.004	.	0.000
	D_S	0.005	0.000	.
N (个数)	污染度	6	6	6
	D_L	6	6	6
	D_S	6	6	6

结果表明,污染度与 D_L 、 D_S 的 Pearson 相关系数分别为 0.928 和 0.915,单尾概率 p 值分别为 0.004和 0.005,在显著性水平为 0.005 的情形下,可以认为污染度与 D_L 、 D_S 之间具有较强的线性相关性。

3.2 线性回归分析

在 SPSS 中进行多元线性回归分析,得到表 3,即多元线性回归的系数列表。

表 3 回归系数表

模型	非标准化参数		标准化参数	
	回归系数	参数的标准误差	标准化回归系数	
1	常数项系数	9.702	7.573	
	D_L	0.266	0.356	1.301
	D_S	-0.085	0.393	-0.376

因变量:污染度

由此可建立多元线性回归方程为:污染度 = $9.702 + 0.266D_L - 0.085D_S$ 。

3.3 显著性检验

表 4 是多元回归的方差分析表,从表中可以看出,模型的 F 统计量的观察值为 9.469,概率 p 值为 0.051,可以认为污染度与铁谱数据间的模型具有显著性。

表 4 方差分析表

模型		平方和	D_f 值	均方	F 统计量的观察值	概率 P 值
1	回归的均方	139.600	2	69.800	9.469	0.051(a)
	剩余的均方	22.114	3	7.371		
	总和	161.714	5			

(a) 预测因子:常量, D_S 、 D_L , (b) 因变量:污染度

表 5 是回归模型的残差统计量,从表 5 中可以看出标准化残差的绝对值最大为 1.065,没有超过默认值 3,没有奇异值。

表 5 残差统计量

	最小值	最大值	均数	标准差	N /个数
预测值	22.453 4	36.010 9	28.435 0	5.283 93	6
标准化预测值	-1.132	1.434	0.000	1.000	6
预测值的标准误差	1.123 65	2.404 52	1.850 51	0.560 02	6
调整后的预测值	16.278 1	36.914 1	27.588 4	6.804 67	6
残差	-2.892 3	2.835 9	0.000 0	2.103 07	6
标准差	-1.065	1.045	0.000	0.775	6

因变量:污染度

4 结 论

文献[1]中用灰色关联度分析变量之间的相异关联程度,实质上就是比较数据列之间曲线几何形状的接近程度,但这一分析方法应用于多元回归分析中的变量筛选却存在明显的缺陷^[3]。本文利用多元回归中的相关性分析方法,计算变量之间的线性相关系数,得出污染度与铁谱数据间存在明显的

线性关系,从而建立多元线性回归模型,并且通过模型可以由铁谱数据预测污染度。

参 考 文 献

1 张 雨,刘义乐. 装甲车辆润滑油监测数据关联分析. 装甲兵工程学院学报,2009;23(1):42—44
2 张红兵,贾来喜. SPSS 宝典. 北京:电子工业出版社,2007
3 张恒喜,郭基联. 小样本多元数据分析方法及其应用. 西安:西北工业大学出版社,2002

Multi Factor Line Regression of Lubrication Oil Monitor in Armored Vehicles

ZHANG Ling,LI Wei-dong¹,LÜ Xi²,DONG Yu-cai

(Institute of Nonlinear Science, Armored Force Engineering, Beijing 100072, P. R. China;
Institute of Department of Science Research Armored Force Engineering¹, Beijing 100072;
Section of PLA Representation in No. 845 Factory²,Xi'an 710302, P. R. China)

[Abstract] Some test data of lubrication oil in gear-box through multi factor line regression is analyzed and discussed line correlation between pollution concentration and ferrograph grind concentration. At last the multi factor regression equation is obtained. The result provides a new method to forecast pollution concentration.

[Key words] oil pollution concentration ferrograph data correlation multi factor line regression

(上接第 7125 页)

Development of Virtual Instruments Measurement Uncertainty
Evaluation Based on Monte Carlo Method

YUAN Min, LIU Gui-xiong*

(School of Mechanical & Automotive Engineering, South China University of Technology, Guangzhou 510640, P. R. China)

[Abstract] Improvement of measurement uncertainty evaluation of virtual instruments based on Monte Carlo method is proposed in tow aspects. According to the mathematical modules, roundoff error module is built based on interval analysis method, in order to evaluate the measurement uncertainty of virtual instruments caused by roundoff errors. The uncertainty evaluation is effected by the quality of simulation, in order to reduced the impaction, an simulation quality testing and controlling module is built in the uncertainty simulation module. An example according to pressure measuring system was showed, and the result prove the improvement.

[Key words] virtual instruments measurement uncertainty Monte Carlo method roundoff errors