

## 基于偏最小二乘回归的飞机维修保障费用预测

郭 风, 张恒喜, 李寿安, 张 琦

(空军工程大学工程学院, 陕西西安 710038)

摘 要: 分析了影响飞机维修保障费用的参数, 提出用偏最小二乘回归方法来预测飞机维修保障费用。该方法对变量进行主成分分析、典型相关分析和多元线性回归, 在处理存在多重线性相关的小样本多元数据方面效果很好。实例证明, 与传统普通多元线性回归方法相比, 偏最小二乘回归在飞机维修保障费用预测中精度更高。

关键词: 维修保障; 费用预测; 偏最小二乘回归

中图分类号: V37 文献标识码: A 文章编号: 1009-3516(2005)03-00010-02

飞机的维修保障费用在其寿命周期费用中居主要地位, 对军用飞机来说, 使用维修费通常占寿命周期费用 65% ~ 80%。准确地估算飞机维修保障费用对寿命周期费用控制, 新型武器装备研制方案决策和现役装备的科学管理有着十分重要的意义<sup>[1-2]</sup>。目前, 维修保障费用数据匮乏、样本容量小、变量间关系复杂已成为飞机维修保障费用估算研究的瓶颈。偏最小二乘回归方法在处理样本容量小、说明性变量多、变量间存在严重多重相关性方面具有独特的优势<sup>[3]</sup>。

### 1 建模步骤

设已知因变量  $y$  和  $k$  个说明性变量  $x_1, x_2, \dots, x_k$ , 样本为  $n$ , 构成数据表  $x[x_1, x_2, \dots, x_k]_{n \times k}$  和  $y = [y]_{n \times 1}$ 。

1) 将  $X$  与  $y$  进行标准化处理, 得到标准化后的自变量矩阵  $E_0$  和因变量矩阵  $F_0$ , 标准化处理的目的是为了公式表达上的方便和减少运算误差。  $x_{ij}^* = \frac{x_{ij} - \bar{x}_j}{s_j}$ ,  $i = 1, \dots, n; j = 1, \dots, k$ 。  $E_0 = (x_{ij}^*)_{n \times k}$ ,  $F_0 = (\frac{y_{ij} - \bar{y}}{s_y})_{n \times 1}$ ,  $i = 1, 2, \dots, n$ 。式中,  $\bar{x}_j$  是  $X_j$  的均值,  $s_j$  是  $X_j$  的标准差;  $\bar{y}$  是  $y$  的均值;  $s_y$  是  $y$  的标准差。

2) 从  $E_0$  中抽取一个成分,  $t_1 = E_0 w_1$ , 其中  $w_1 = E_0' F_0 / \|E_0' F_0\|$ 。实施  $E_0$  和  $F_0$  在  $t_1$  上的回归:  $E_0 = t_1 p_1 + e_1$ ;  $F_0 = t_1 r_1 + F_1$ 。其中,  $p_1, r_1$  是回归系数即  $p_1 = E_0' t_1 / \|t_1\|^2$ ;  $r_1 = F_0' t_1 / \|t_1\|^2$ 。记残差矩阵  $E_1 = E_0 - t_1 p_1$ ;  $F_1 = F_0 - t_1 r_1$ 。检查收敛性, 若  $y$  对  $t_1$  的回归方程已达到满意的精度, 则进行下一步; 否则, 令:  $E_0 = E_1, F_0 = F_1$ , 回到第 2) 步, 对残差矩阵进行新一轮的成分提取和回归分析;

3) 在第  $h$  步 ( $h = 2, \dots, m$ ), 方程满足精度要求(可用交叉有效性确定), 这时得到  $m$  个成分  $t_1, t_2, \dots, t_m$ , 实施  $F_0$  在  $t_1, t_2, \dots, t_m$  上的回归, 得  $\hat{F}_0 = r_1 t_1 + r_2 t_2 + \dots + r_m t_m$ 。由于  $t_1, t_2, \dots, t_m$  均是  $E_0$  的线性组合, 因此,  $\hat{F}_0$  可写成  $E_0$  的线性组合形式, 即  $\hat{F}_0 = r_1 E_0 w_1^* + \dots + r_m E_0 w_m^*$ 。其中,  $w_h^* = \prod_{j=1}^{h-1} (I - w_j p_j')$ ,  $I$  为单位矩阵。最后, 有  $\hat{y}^* = \alpha_1 x_1^* + \dots + \alpha_p x_p^*$ 。  $x_j^*$  的回归系数为  $\alpha_j = \sum_{h=1}^m r_h \omega_{hj}^*$ 。式中,  $\omega_{hj}^*$  是  $w_h^*$  的第  $j$  个分量。

4) 按照标准化的逆过程, 将  $\hat{F}_0(\hat{y}^*)$  的回归方程还原为  $Y$  对  $X$  的回归方程。

从以上建模步骤可以看出, 偏最小二乘回归的建模依据是建立在信息分解与提取的基础之上的。它在

收稿日期: 2004-10-12

基金项目: 军队科研基金资助项目

作者简介: 郭 风(1979-), 男, 河南商丘人, 硕士生, 主要从事飞机效费分析、飞机可靠性、维修性及保障性等研究;  
张恒喜(1937-), 男, 江苏姜堰人, 教授, 博士生导师, 主要从事飞机效费分析、军用飞机型号发展工程学研究。

说明性变量  $x_1, x_2, \dots, x_k$  中逐次提取综合成分  $t_1, t_2, \dots, t_m (m < k)$ , 这相当于对  $x_1, x_2, \dots, x_k$  中的信息进行重新组合和抽取, 从而得到对  $y$  的解释能力最强, 同时又最能概括说明性变量集合  $X$  中信息的综合变量。而与此同时, 对  $y$  没有解释意义的信息被自然地排除掉了。

5) 交叉有效性分析可以确定所应提取的成分个数。记  $y_i$  为原始数据,  $t_1, t_2, \dots, t_m$  是在偏最小二乘回归过程中提取的成分。 $\hat{y}_{hi}$  是使用全部样本点并取  $t_1 \sim t_h$  个成分回归建模后, 第  $i$  个样本点的拟合值。 $\hat{y}_{h(-i)}$  是在建模时删去的样本点  $i$ , 取  $t_1 \sim t_h$  个成分回归建模后, 再用此模型计算的  $y_i$  的拟合值。记  $SS_h = \sum_{i=1}^n (y_i - \hat{y}_{hi})^2$ ;  $PRESS_h = \sum_{i=1}^n (y_i - \hat{y}_{h(-i)})^2$ ;  $Q_h^2 = 1 - PRESS_h / SS_{h-1}$ 。则当  $Q_h^2 \geq (1 - 0.95^2) = 0.0975$  时, 认为  $t_h$  成分的边际贡献是显著的。引进新的主成分  $t_h$  会对模型的预测能力有明显的改善作用<sup>[5]</sup>。

## 2 实例分析

现以某型飞机维修保障费用预测为例。可靠性、维修性、保障性(RMS)是影响飞机维修保障费用的关键因素, 这些参数包括: 任务可靠度(DR)、平均故障间隔时间(MTBF)、平均修复时间(MTTR)、每飞行小时维修工时(DMMH/FH)、更换发动机时间(ERT)、再次出动准备时间(SGR)、故障检测率(FDR)、故障隔离率(FIR)、备件利用率(SCR)。收集的具体数据如表1。

表1 飞机维修保障费用原始数据

机型	A	B	C	D	E	F	G	H	I	K
DR(%)	90.5	85.1	93.0	85.0	84.0	90.0	88.0	90.6	86.0	88.0
MTBF/h	1.75	2.9	5	2.6	2.1	2.5	2.9	2.7	1.5	3.0
MTTR/h	1.78	1.9	3	1.6	1.8	2.0	1.78	1.3	1.0	1.79
DMMH/(FH/h)	11	12	4.5	6.0	7.5	8.0	5.6	10	11.2	5.8
ERT/min	21	30	89	149	300	150	80	640	529	128
SCR/min	15	10	13	18	42	35	15	25	34	16
FDR(%)	95.0	96.0	98.0	96.5	97.0	97.0	95.0	95.0	94.0	97.0
FIR(%)	95.0	96.0	98.0	96.5	96.0	98.5	95.0	90.0	92.0	97.5
SCR(%)	95	96	98	96	95	97	97	95	94	97
C(万元)	36652	50515	40150	58220	95511	76251	51567	18131	15427	49865

为了对建立的模型进行误差分析和预测检验, 选取表中前9个子样进行建模, 以K型机作为检验子样, 并分析误差计算。用偏最小二乘回归方法建立的模型函数为  $C = 27\ 082.79 \times x_1^{-0.02} x_2^{-0.64} x_3^{-0.79} x_4^{-0.26} x_5^{-0.13} x_6^{-0.24} x_7^{-0.013} x_8^{-0.012} x_9^{-0.0015}$

预测结果K型机维修保障费用为42 094万元, 误差为15.6%, 与传统的计算普通多元性回归方法相比精确度较高。见表2, 因此在飞机维修保障费用预测中, 运用偏最小二乘回归进行建模分析是可行的。

表2 预测结果及误差分析

方法	偏最小二乘回归	普通多元线性回归
C(万元)	42094	38620
误差(%)	-15.6	-22.6

## 3 结论

飞机的维修保障费用预测是根据飞机概要设计的情况, 初步估算飞机的维修保障成本。由于收集数据困难、样本容量少、自变量多, 用传统的方法计算起来比较困难。偏最小二乘回归方法将对样本容量小、自变量多、变量间存在严重多重相关性的数据处理产生积极的意义。由以上算例可以看出在较少样本点的情况下, 利用偏最小二乘回归建模进行飞机维修保障费用预测是可行有效的。

### 参考文献:

[1] 刘晓东, 张恒喜. 飞机可靠性与研制费用相关关系研究[J]. 空军工程大学学报(自然科学版), 2000, 1(1): 63-66.