第11卷第6期　　空 军 工 程 大 学 学 报(自然科学版)　　Vol. 11　No. 6
2010年12月　　JOURNAL OF AIR FORCE ENGINEERING UNIVERSITY(NATURAL SCIENCE EDITION)　　Dec. 2010

飞机生存力评价指标之间相关性的解决策略

王怀威，　李曙林，　陈　宁
（空军工程大学　工程学院，陕西　西安　710038）

摘　要:针对以往构建的生存力评价指标体系中评价指标的合理性缺乏考证的问题,提出了评价指标之间潜在相关性的处理方法。以文献中的生存力数据为例,首先采用模糊聚类的方法,在不同置信水平下把原有的10个评价指标重新分类,依据定义的有效性指标选择出有效性最高的分类结果为4类;对新的4个综合评价指标进行因子分析,验证了模糊聚类的结果,并得到了新指标与原指标的数值关系。结果表明,相较于原指标体系,新指标体系简单有效,聚类方法合理,聚类效果良好。该方法可以用于生存力评价指标体系的构建。

关键词:飞机生存力;评价指标;相关性;模糊聚类;因子分析

DOI:10. 3969/j. issn. 1009 -3516. 2010. 06. 002

中图分类号: V271. 4　　**文献标识码**:A　　**文章编号**:1009 -3516(2010)06 -0007 -05

生存力是各类军用飞机设计的重要指标之一,对提高飞机综合作战效能有突出作用。飞机生存力评价研究主要包括评价指标体系的建立和综合评价方法的探索。目前,研究重点多集中在评价方法[1-3],由于生存力设计工作没有具体型号的实际应用,对评价指标的选择往往较随意,其合理性缺乏考证。同时由于指标之间不可避免地存在潜在关联,将会显著增大评价的复杂性和重复性,经过赋权后又必定夸大对指标中的关联部分,影响评价结果的有效性和可靠性。因此,有必要采取合适的方法解决评价指标之间的相关性,构建科学合理的评价指标体系,真实地评价飞机生存力。本文在考查飞机生存力指标数据间相似关系的基础上,对指标变量进行聚类分析和因子分析,解决指标间相关性的同时有效降低评价指标维数,并依据应用算例,对生存力评价指标体系的建立给出建议。

1　模糊聚类分析

模糊聚类分析是非监督模式识别的重要分支,它把一个没有类别标记的样本集按某种准则划分成若干个子集(类),使相似的样本尽可能归为一类[4],已经在模式识别[5]、工程力学[6-7]、机械故障诊断[8]等领域得到了广泛应用。

1.1　初始样本数据

设有 n 架飞机,每架飞机各有 m 个需要考虑的评价指标参数,即 m 维聚类特性指标,x'_{ij}表示第 i 架飞机的第 j 项指标的数据值($i=1,2,\cdots,n;j=1,2,\cdots,m$),这样 n 架飞机的所有特性指标构成的初始矩阵记作:

$$\boldsymbol{X}^{*}=\begin{bmatrix} x'_{11} & \cdots & x'_{1m} \\ \cdots & & \cdots \\ x'_{n1} & \cdots & x'_{nm} \end{bmatrix} \tag{1}$$

1.2　数据归一化

* **收稿日期**:2010 -05 -07
基金项目:国防预研基金资助项目(9140A25010406JB3205)
作者简介:王怀威(1981 -),男,安徽淮南人,博士生,主要从事飞机生存力研究;E -mail:awei121_0@yahoo. cn
李曙林(1959 -),男,河北威县人,教授,博士生导师,主要从事飞机结构强度、可靠性及生存力研究.

为了平衡各项模糊聚类指标的作用,消除差异,需要进行无量纲归一化处理,得到新的归一化数据阵 $\boldsymbol{X}$。对于效益型数据和对于成本型数据,分别有:

$$x_{ij}=\frac{x'_{ij}-\min(x'_{ij})}{\max(x'_j)-\min(x'_j)} \tag{2}$$

$$x_{ij}=\frac{\max(x'_{ij})-x_{ij}}{\max(x'_j)-\min(x'_j)} \tag{3}$$

式中 $\max(x'_j)$ 和 $\min(x'_j)$ 分别是初始矩阵中第 j 项指标的最大值和最小值。经过归一化处理后,$0\leqslant x_{ij}\leqslant 1$。

1.3 构造模糊相似关系

$\boldsymbol{X}$ 的每一列是指标参数,指标间的相似程度用相似系数 r_{ij} 来表示,并构成 $m\times m$ 维关系矩阵,即 $\boldsymbol{R}=[r_{ij}]_{m\times m}$。用标准化的数据标定计算各个样本间的相似系数的方法很多,本文采用相关系数法:

$$r_{ij}=\frac{\mathrm{cov}(x_i,x_j)}{\sqrt{D(x_i)D(x_j)}} \tag{4}$$

式中:$\mathrm{cov}(x_i,x_j)$ 为指标 x_i 和指标 x_j 的协方差;$D(x_i)$ 和 $D(x_j)$ 分别是 x_i 和 x_j 的方差。

1.4 确定模糊等价矩阵

上节中构造的 $\boldsymbol{R}$ 一般不具有模糊等价关系,需要按照传递闭包原理进行改造。令:

$$\boldsymbol{R}^2=\boldsymbol{R}\circ\boldsymbol{R}=[r_{ij}^2]_{m\times m} \tag{5}$$

式中 $r_{ij}^2=\bigvee_{k=1}^{m}(r_{ik}\wedge r_{jk})$,$\wedge$ 为取小运算,$\vee$ 为取大运算。再令 $\boldsymbol{R}^4=\boldsymbol{R}^2\circ\boldsymbol{R}^2$,以此类推,计算到 $\boldsymbol{R}^{2p}=\boldsymbol{R}^{2p-1}$ 为止,其中,$p=\mathrm{ceil}(\ln m)$,ceil 为取不小于 $\ln m$ 的最小整数。则 $\boldsymbol{R}^*=\boldsymbol{R}^{2p}=\boldsymbol{R}^{2p-1}$ 便是 $\boldsymbol{R}$ 的传递闭包,也是本文需要的模糊等价矩阵。

1.5 聚类

指标集之间的模糊等价关系 $\boldsymbol{R}^*$ 确定后,给定不同的置信水平值 λ,就可以得到模糊等价矩阵的 λ 水平截距矩阵 $\boldsymbol{R}_\lambda^*$,即决定了在 λ 水平下的一个模糊分类。给定不同的 λ 值,就可以得到不同的分类,随着 λ 的减小,分类由精到粗,分类数由多到少,逐层归并,形成一个动态的聚类图。

1.6 聚类有效性检验

不同聚类数目的聚类质量需要有一个准确的度量,本文用样本离开相应聚类中心间距的平均值与聚类中心最小间距的比值来衡量聚类的有效性[9]。假设聚类后的聚类数目为 q,每个样本 x_l 被分到一个聚类块 c_i 中,每个 c_i 包含的样本数为 m_i,对应的聚类中心为 $v_1,v_2,\cdots,v_q$。c_i 中各个样本到聚类中心的距离为:

$$\delta_i=\frac{1}{m_i}\sum_{x_l\in c_i}\|x_l-v_i\|^2,\quad i=1,2,\cdots,q \tag{6}$$

对于所有的样本,其离开相应聚类中心的距离为:

$$\delta=\frac{1}{m}\sum_{i=1}^{q}m_i\delta_i \tag{7}$$

而各个聚类中心之间的间距的最小值为:

$$L=\min_{i,j}\|v_i-v_j\|^2,\quad i,j=1,2,\cdots,q;i\neq j \tag{8}$$

则聚类有效性表示为 $S(q)=\delta/L$。对于一个紧致的、良性划分的聚类应具有尽可能大的聚类中心间距,而各类内部样本与其中心间距尽可能小,所以 $S(q)$ 的值越小越好。

2 因子分析

用因子分析法可以得到多元数据的一组公因子,这些公因子是多元数据原指标中潜在的、不能直接观测的随机变量,具有一些特殊的性质,可以成为新的指标。本文应用因子分析法对模糊聚类的结果进行检验和验证,并确定新指标与原指标之间对应的数值关系。

2.1 因子分析的前提条件

因子分析要求原变量之间应具有较强的相关关系。常用的相关性检验方法有相关系数检验、巴特利特

球形度检验和 KMO 检验[10]。通常认为相关系数小于 0.3 和 KMO 值小于 0.5 时不适合做因子分析。

2.2　因子分析算法

因子分析模型描述如下:

$$\begin{bmatrix} x_1 \\ \cdots \\ x_p \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & & \cdots \\ a_{p1} & \cdots & a_{pm} \end{bmatrix} \begin{bmatrix} F_1 \\ \cdots \\ F_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \cdots \\ \varepsilon_p \end{bmatrix} \tag{9}$$

简记为 $\boldsymbol{X}_{p\times 1} = \boldsymbol{A}_{p\times m}\boldsymbol{F}_{m\times 1} + \varepsilon_{p\times 1}, m < p$。式中:$x_1, x_2, \cdots, x_p$ 为原有变量,是均值为 0、方差为 1 的标准化向量;$F_1, F_2, \cdots, F_m$ 称为因子变量;$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_p$ 是特殊因子变量;$\boldsymbol{A}$ 是因子载荷矩阵,a_{ij}表示第 j 个因子变量对第 i 个原有变量的解释程度,在各个因子变量不相关的情况下,因子载荷 a_{ij}就是第 i 个原有变量和第 j 个因子变量的相关系数,a_{ij}绝对值越大,因子变量和原有变量的相关度就越高[11]。

由于因子变量能够反映原有变量的相关关系,往往需要将因子变量表示为原有变量的线性组合,即 $\boldsymbol{F}_{m\times 1} = \beta_{m\times p}\boldsymbol{X}_{p\times 1}$,这样,原有的 p 个相关变量就被简化为 m 个不相关的因子,在综合评价中,这些因子就成为新的评价指标,赋权后就可以对目标进行评价。

3　应用算例

在文献[1]中,给出生存力评价指标体系为{雷达散射截面积,涡轮前温度,电子对抗能力,机动性,致命性部件比例,致命性部件余度概率,致命性部件防护/遮挡概率,致命性结构部件的平均安全系数,可达性,平均抢修时间},分别以 $x_1 - x_{10}$表示这 10 个指标,对应指标值见表 1。

表 1　飞机生存力评价指标值

Tab. 1　Evaluation index values of aircraft survivability

机型	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
A	0.1	1 923	1	31.2	0.31	0.28	0.41	1.6	0.64	0.8
B	11.3	1 672	0.75	25.5	0.42	0.23	0.35	1.4	0.55	1.5
C	4.9	1 672	0.5	24.5	0.45	0.21	0.37	1.8	0.6	1.2
D	5.8	1 528	0.75	25.65	0.39	0.2	0.31	2	0.42	2
E	10.8	1 747	0.75	27.3	0.42	0.24	0.37	2	0.45	1.5
F	9.1	1 650	0.75	26.3	0.47	0.23	0.29	1.5	0.42	1.5
G	3.6	1 298	0.25	17.8	0.52	0.14	0.19	1.5	0.25	3
H	3.8	1 370	0	17.55	0.55	0.16	0.2	1.5	0.23	2.5

根据前述的模糊聚类分析方法,首先对表 1 数据进行归一化处理,获得归一化矩阵 $\boldsymbol{X}$,再求各指标间的相关系数矩阵 $\boldsymbol{R}$。

$$\boldsymbol{X} = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0.3333 & 1 & 1 \\ 0 & 0.4016 & 0.75 & 0.5824 & 0.5417 & 0.6429 & 0.7273 & 0 & 0.7805 & 0.6818 \\ 0.5714 & 0.4016 & 0.5 & 0.5092 & 0.4167 & 0.5 & 0.8182 & 0.6667 & 0.9024 & 0.8182 \\ 0.4911 & 0.632 & 0.75 & 0.5934 & 0.6667 & 0.4286 & 0.5455 & 1 & 0.4634 & 0.4545 \\ 0.0446 & 0.2816 & 0.75 & 0.7143 & 0.5417 & 0.7143 & 0.8182 & 1 & 0.5366 & 0.6818 \\ 0.1964 & 0.4368 & 0.75 & 0.641 & 0.3333 & 0.6429 & 0.4545 & 0.1667 & 0.4634 & 0.6818 \\ 0.6875 & 1 & 0.25 & 0.0183 & 0.125 & 0 & 0 & 0.1667 & 0.0488 & 0 \\ 0.6696 & 0.8848 & 0 & 0 & 0 & 0.1429 & 0.0455 & 0.1667 & 0 & 0.2273 \end{bmatrix}$$

$$
\boldsymbol{R}=\begin{bmatrix}
1 & 0.088\,6 & -0.209\,6 & -0.125\,3 & 0.107\,3 & -0.144\,3 & -0.140\,2 & -0.095\,2 & -0.031 & -0.091\,4\\
0.088\,6 & 1 & -0.856\,2 & -0.951\,1 & -0.843\,1 & -0.985\,7 & -0.948\,9 & -0.219\,8 & -0.895\,8 & -0.967\,4\\
-0.209\,6 & -0.856\,2 & 1 & 0.962\,6 & 0.908\,6 & 0.895\,5 & 0.836\,6 & 0.281\,5 & 0.783\,3 & 0.789\,9\\
-0.125\,3 & -0.951\,1 & 0.962\,6 & 1 & 0.918 & 0.968\,1 & 0.916\,4 & 0.331\,6 & 0.847\,2 & 0.904\,6\\
0.107\,3 & -0.843\,1 & 0.908\,6 & 0.918 & 1 & 0.850\,6 & 0.862\,2 & 0.354\,7 & 0.816\,1 & 0.770\,7\\
-0.144\,3 & -0.985\,7 & 0.895\,5 & 0.968\,1 & 0.850\,6 & 1 & 0.91 & 0.178\,5 & 0.847\,3 & 0.941\,8\\
-0.140\,2 & -0.948\,9 & 0.836\,6 & 0.916\,4 & 0.862\,2 & 0.91 & 1 & 0.388 & 0.949\,8 & 0.940\,8\\
-0.095\,2 & -0.219\,8 & 0.281\,5 & 0.331\,6 & 0.354\,7 & 0.178\,5 & 0.388 & 1 & 0.180\,9 & 0.196\,7\\
-0.031 & -0.895\,8 & 0.783\,3 & 0.847\,2 & 0.816\,1 & 0.847\,3 & 0.949\,8 & 0.180\,9 & 1 & 0.933\\
-0.091\,4 & -0.967\,4 & 0.789\,9 & 0.904\,6 & 0.770\,7 & 0.941\,8 & 0.940\,8 & 0.196\,7 & 0.933 & 1
\end{bmatrix}
$$

$\boldsymbol{R}$ 中元素小于 0.3 的不足 30%,指标间相关性较大,有必要进行模糊聚类分析。求解模糊等价矩阵 $\boldsymbol{R}^{*}=\boldsymbol{R}^{8}$,即为:

$$
\boldsymbol{R}^{8}=\begin{bmatrix}
1 & 0.088\,6 & 0.107\,3 & 0.107\,3 & 0.107\,3 & 0.107\,3 & 0.107\,3 & 0.107\,3 & 0.107\,3 & 0.107\,3\\
0.088\,6 & 1 & 0.088\,6 & 0.088\,6 & 0.088\,6 & 0.088\,6 & 0.088\,6 & 0.088\,6 & 0.088\,6 & 0.088\,6\\
0.107\,3 & 0.088\,6 & 1 & 0.962\,6 & 0.918 & 0.962\,6 & 0.940\,8 & 0.388 & 0.940\,8 & 0.941\,8\\
0.107\,3 & 0.088\,6 & 0.962\,6 & 1 & 0.918 & 0.968\,1 & 0.940\,8 & 0.388 & 0.940\,8 & 0.941\,8\\
0.107\,3 & 0.088\,6 & 0.918 & 0.918 & 1 & 0.918 & 0.918 & 0.388 & 0.918 & 0.918\\
0.107\,3 & 0.088\,6 & 0.962\,6 & 0.968\,1 & 0.918 & 1 & 0.940\,8 & 0.388 & 0.940\,8 & 0.941\,8\\
0.107\,3 & 0.088\,6 & 0.940\,8 & 0.940\,8 & 0.918 & 0.940\,8 & 1 & 0.388 & 0.949\,8 & 0.940\,8\\
0.107\,3 & 0.088\,6 & 0.388 & 0.388 & 0.388 & 0.388 & 0.388 & 1 & 0.388 & 0.388\\
0.107\,3 & 0.088\,6 & 0.940\,8 & 0.940\,8 & 0.918 & 0.940\,8 & 0.949\,8 & 0.388 & 1 & 0.940\,8\\
0.107\,3 & 0.088\,6 & 0.941\,8 & 0.941\,8 & 0.918 & 0.941\,8 & 0.940\,8 & 0.388 & 0.940\,8 & 1
\end{bmatrix}
$$

分别选取不同的 λ 水平,获得聚类结果见表 2。可见,最好的聚类结果应是 4 组,其不仅有效性最好,而且 λ 置信水平范围最广。对应的聚类情况是 $c_1=\{x_1\}$,$c_2=\{x_2\}$,$c_3=\{x_3,x_4,x_5,x_6,x_7,x_9,x_{10}\}$ 和 $c_4=\{x_8\}$。对于 c_3,可以进行因子分析,考查验证聚类后的性能变化并获得其中 7 个指标的线性组合方式。

表 2 不同 λ 水平取值范围的聚类结果

Tab.2 Outcome of fuzzy cluster analysis with different λ

聚类结果	λ 值域							
	[0.088 6,0.107 3)	[0.107 3,0.388)	[0.388,0.918)	[0.918,0.940 8)	[0.940 8,0.941 8)	[0.941 8,0.949 8)	[0.949 8,0.962 6)	[0.962 6,0.968 1)
聚类数目	2	3	4	5	6	7	8	9
δ	0.060 2	0.033 7	0.009 7	0.010 5	0.008 2	0.006 3	0.003 8	0.001 3
L	1.785 6	1.378 2	1.378 2	0.171 3	0.163 9	0.126 9	0.101 6	0.100 5
$S(q)$	0.033 7	0.024 5	0.007 0	0.061 3	0.050 0	0.049 6	0.037 4	0.012 9

观察 c_3 中 7 个指标之间的相关系数,最小的为 0.771,绝大部分在 0.85 以上,相关度非常高;巴特利特球形度检验拒绝相关系数阵为单位阵的零假设,另外计算 KMO 值为 0.542,适合因子分析。按主成分分析法,获得一个主因子,该因子可以解释 7 个原指标造成的总方差的 90.055%。因子与原指标间的关系见表 3。由因子载荷可以看出,该因子对 7 个指标的解释程度都在 92% 以上。

表 3 因子与原指标间的关系

Tab.3 The relation between new indices and original indices

指标	因子载荷	因子得分系数
电子对抗能力	0.930	0.147
机动性	0.981	0.156
致命部件比例	0.922	0.146
致命性部件余度概率	0.966	0.153
致命性部件防护/遮挡概率	0.966	0.153
可达性	0.930	0.147
平均抢修时间	0.946	0.150

一般地,如果对 m 个指标进行因子分析后获得了 n 个主因子,可以解释总方差的 80% 以上,则表明这 m 个指标应该归为 n 类;同时,因子载荷越高,说明因子对原指标的解释程度越高,因子分析的结果越合理[10]。因此,本例中因子分析的结果既表明了因子分析的有效性,又验证了 7 个指标聚为一类的合理性,模糊聚类的效果良好。综合表 3 和聚类情况 c_1、c_2、c_3、c_4,得到聚类后新的指标因子为:$F_1=x_1$,$F_2=x_2$,$F_3=0.147x_3+$

$0.156x_4 + 0.146x_5 + 0.153x_6 + 0.153x_7 + 0.147x_9 + 0.150x_{10}$，$F_4 = x_8$。

以这 4 个指标因子作为新的指标，可以采取 2 种方式对飞机生存力进行综合评价：一是采用原有 10 个指标的数据，但只对 $F_1 - F_4$ 赋权，可以显著减轻评价工作量；二是考虑对原有的 10 个指标予以取舍，减少评价指标数量，提高评价效率，如电子对抗能力、机动性、致命性部件比例、致命性部件余度概率、致命性部件防护/遮挡概率、可达性和平均抢修时间这 7 个指标之间关联程度较高，选择其中有代表性的 1 个指标即可达到与原来 7 个指标接近的效果。

4　结束语

本文通过考查现有生存力评价指标体系中各个指标间的相关系数，发现其部分指标间存在较强的相关性。通过模糊聚类和因子分析，既降低了评价指标维数，又使得新指标间的相关性得到了有效的削减。相较于原来的评价指标体系，新指标体系变得更为简单和有效。

由于指标之间潜在的相关性是通过对指标数据的统计分析观察到的，因此建议在制定评价指标体系时，先初步确定评价指标，再获取指标数据进行分析，通过有效地聚类或取舍，消除待定指标间的相关性，最终给出科学合理的评价指标体系。

参考文献：

[1]　童中翔，李寿安，叶广强，等. 基于协商定权的飞机生存力评价体系研究[J]. 兵工学报，2007，28(4)：508－512.
TONG Zhongxiang, LI Shouan, YE Guangqiang, et al. Evaluation System of Aircraft Survivability Based on Bargaining Weight [J]. Acta Armamentarii, 2007, 28(4): 508－512. (in Chinese)

[2]　李寿安，张恒喜，李曙林，等. 飞机生存力评估与综合权衡方法研究[J]. 航空学报，2005，26(1)：23－26.
LI Shouan, ZHANG Hengxi, LI Shulin, et al. Research on Aircraft Survivability Evaluation and Synthetic Tradeoff Methods [J]. Acta Aeronautica et Astronautica Sinica, 2005, 26(1): 23－26. (in Chinese)

[3]　李寿安，王礼沅，张恒喜，等. 基于灰色关联投影法的飞机生存力设计方案评估[J]. 电光与控制，2006，13(1)：33－35.
LI Shouan, WANG Liyuan, ZHANG Hengxi, et al. Aircraft Survivability Design Scheme Evaluation Based on Grey Correlation Projection [J]. Electronics Optics & Control, 2006, 13(1): 33－35. (in Chinese)

[4]　潘寒尽，张多林，方冬进，等. 模糊聚类与信息熵综合评价法在防空 C^3I 系统中的应用[J]. 空军工程大学学报：自然科学版，2006，7(5)：19－21.
PAN Hanjin, ZHANG Duolin, FANG Dongjin, et al. The Application of Synthetical Evaluation Based on Fuzzy Clustering and Information Entropy to the Effective Evaluation of Antiaircraft System [J]. Journal of Air Force Engineering University: Natural Science Edition, 2006, 7(5): 19－21. (in Chinese)

[5]　Liang G S, Chou T Y, Han T C. Cluster Analysis Based on Fuzzy Equivalence Relation [J]. European Journal of Operational Research, 2005, 166(1): 160－171.

[6]　李枝军，李爱群，韩晓林，等. 基于最大熵谱和模糊聚类分析的斜拉桥拉索索力测试与评估[J]. 工程力学，2006，26(11)：88－94.
LI Zhijun, LI Aiqun, HAN Xiaolin, et al. Measurement and Estimation of the Cable Tension Based on Maximum Entropy Spectral and Fuzzy Clustering[J]. Engineering Mechanics, 2006, 26(11): 88－94. (in Chinese)

[7]　张鹏，张在明，杨宇友，等. 模糊聚类在地层分析中的应用[J]. 岩土力学，2009，30(8)：2348－2352.
ZHANG Peng, ZHANG Zaiming, YANG Yuyou, et al. Application of Fuzzy Cluster to Geotechnical Stratum Analysis[J]. Rock and Soil Mechanics, 2009, 30(8): 2348－2352. (in Chinese)

[8]　张国钢，李宇，汤翔，等. 模糊聚类分析用于断路器状态评估因素分类[J]. 高电压技术，2008，34(2)：350－354.
ZHANG Guogang, LI Yu, TANG Xiang, et al. Clustering Analysis in Condition Assessment Factor Classification of Circuit Breaker[J]. High Voltage Engineering, 2008, 34(2): 350－354. (in Chinese)

[9]　吴今培，孙德山. 现代数据分析[M]. 北京：机械工业出版社，2006.
WU Jinpei, SUN Deshan. Modern Data Analysis[M]. Beijing: China Machine Press, 2006. (in Chinese)

（下转第 26 页）