

doi: 10.3969/j.issn.1006-1576.2010.04.015

一种新的基于灰色贝叶斯的装备库存预测模型

刘文辉¹, 张玉林¹, 党丽琼²

(1. 西安通信学院 装备教研室, 陕西 西安 710106; 2. 陕西航天职工大学 计算机教研室, 陕西 西安 710105)

摘要: 针对现代复杂战场环境下装备保障的随机性影响因素, 提出用灰色预测解决装备保障中备件库存的预测问题。考虑到传统统计学在处理小样本数据时存在的不足, 运用贝叶斯理论来对预测结果进行评估, 通过推导伽玛分布的卡方变化, 给出了简单的概率区间公式。最后通过实例进行了分析, 证明了该方法优于统计学估计, 特别是在历史数据较少的情况下, 预测效果明显, 对装备保障有现实的指导意义。

关键词: 灰色系统; MGM (1, 1) 模型; 贝叶斯概率区间

中图分类号: N941.5 **文献标识码:** A

A New Prediction Model of Equipment Stock Based on Gray-Bayesian

LIU Wen-hui¹, ZHANG Yu-lin¹, DANG Li-qiong²

(1. Staff Room of Equipment, Xi'an Communication University, Xi'an 710106, China;

2. Staff Room of Computer, Shanxi Spaceflight Employee University, Xi'an 710105, China)

Abstract: Aiming at the random factors to equipment support at complex battlefield environment nowadays, introduce to use gray theory to solve the prediction of spare parts inventory during equipment support. Taking into account in dealing with small samples of traditional statistical data shortcomings, the use of Bayesian theory to predict the results of the evaluation, through the derivation of inverted gamma distribution of chi-square change, give a simple formula for the probability interval. Give a simple formula of probability intervals. Finally, an example is analyzed to prove that this method is better than statistical estimates, especially in the case of less historical data to predict effects obviously, which has practical guiding significance to equipment support.

Keywords: Gray system; MGM (1, 1) model; Bayesian probability intervals

0 引言

在明确任务的前提下, 准确而灵活地确定备件数量, 并合理地权衡费用成了装备综合保障领域的一个研究热点。目前, 国内外针对这方面的研究主要包括可用度模型、遗传算法模型、灰色系统模型和仿真模型等^[1-3]。

现实情况表明, 备件需求受多种因素的影响, 备件消耗过程是一个随机过程, 消耗量是个随机变量, 且在一定的范围内呈现出随机的波动性。灰色系统对随机性因素的处理具有很大的优越性, 灰色系统把任何随机过程都看成是在一定幅值范围、一定时区内变化的灰色量, 通过生成变换将无规律的随机性序列变成可以处理的有规律的灰色序列, 对灰色序列的处理在另一层面反映了随机序列的变化, 这就是灰色系统的本质所在, 称为灰色过程^[4]。基于此, 提出用灰色理论来解决备件的库存预测问题, 同时把贝叶斯区间估计引入到预测的评估中, 以解决灰色预测在区间估计方面的不足。

1 灰色预测

自邓聚龙教授于 20 世纪 80 年代提出用于控制和预测的灰色理论和方法以来, 灰色预测逐渐成为灰色系统理论的重要组成部分, 其中应用较为广泛的是文献[5]提出的 GM (1,1) 预测模型。

1.1 MGM (1, 1) 模型

灰色预测模型选择新陈代谢 GM (1, 1) 模型, 为了叙述方便, 记新陈代谢 GM (1, 1) 为 MGM (1, 1), M 代表新陈代谢。MGM (1, 1) 与 GM (1, 1) 的计算方法在本质上是相同的, 它的特点在于通过不断地加入新元素, 去除旧元素达到数据序列更新的目的, 理论上说, MGM (1, 1) 的应用在一定程度上可改善 GM (1, 1) 的预测效果。

设有原始数列: $X^{(0)}(k) = \{x^{(0)}(1), x^{(0)}(2) \cdots x^{(0)}(n)\}$, 子序列: $X_j^{(0)}(k) = \{x^{(0)}(j), x^{(0)}(j+1) \cdots x^{(0)}(j+m-1)\}$, 一次累加: $X_j^{(1)}(k) = \{x^{(1)}(j), x^{(1)}(j+1) \cdots x^{(1)}(j+m-1)\}$, 其中原始数据长度为 n , m 为子序列长度, j 为子序列代数。

1.2 模型求解

第 1 步: 建立第 j 代子序列灰微分方程

收稿日期: 2009-11-06; 修回日期: 2009-12-22

作者简介: 刘文辉 (1983-), 男, 陕西人, 西安通信学院在读硕士研究生, 从事军事装备技术保障研究。

$X_j^{(0)}(k) + aZ_j^{(1)}(k) = b$, 式中, $X_j^{(0)}(k)$ 为灰导数, 符合白因灰果律, $Z_j^{(1)}(k)$ 为 $X_j^{(0)}(k)$ 的白化背景值, $Z_j^{(1)}(k) = \lambda X_j^{(1)}(k) + (1-\lambda)X_j^{(1)}(k+1)$, λ 取 0.5.

第2步: 参数求解

令: $y_{jN} = [X^{(0)}(j+1), X^{(0)}(j+2) \cdots X^{(0)}(j+m-1)]^T$

$$B_j = \begin{bmatrix} -Z^{(1)}(2) & 1 \\ -Z^{(1)}(3) & 1 \\ \dots & \dots \\ -Z^{(1)}(m) & 1 \end{bmatrix} \quad \hat{a}_j = \begin{bmatrix} a \\ b \end{bmatrix}$$

则有: $y_{jN} = B_j \hat{a}_j$

在最小二乘准则下有:

$$\hat{a}_j = (B_j^T B_j)^{-1} B_j^T y_{jN}$$

第3步: 求得预测递推公式

$$\begin{cases} \hat{X}_j^{(0)}(j) = X_j^{(1)}(j) \\ \hat{X}_j^{(0)}(k+1) = (X_j^{(0)}(1) - \frac{b}{a})(1 - e^{ak})e^{ak} \quad k=2,3,\dots,m \end{cases}$$

第4步: 子序列 j 加 1, 进入下一轮计算.

对于上述灰色预测结果, 可通过经典统计学样本矩分析求解出预测结果的概率分布区间, 而在实际情况下, 当要分析的样本很少时, 采用经典统计学方法得到的评估结构精度不够高, 这时可以采用贝叶斯方法建立贝叶斯可信概率区间, 且贝叶斯区间估计明显优于经典统计学的频率法^[6,7].

2 贝叶斯评估模型

2.1 贝叶斯原理

贝叶斯学派的最基本观点是: 任何一个未知量 θ 都可以看作一个随机变量, 应用一个概率分布去描述 θ 的未知状况^[8], 这个概率分布即为先验分布. 贝叶斯分析实质上是未知参数先验分布的后验分布过程, 与传统的经典统计学相比, 其主要差别在于它在样本、总体信息的基础上, 充分利用了随机变量的先验信息.

2.2 确定待估参数

在一元线性回归模型 $Y = \beta X + \alpha$ 中, 通常假定随机误差向量 α 服从正态分布. 而 GM(1, 1) 预测中的参数求解实质上是一种线性回归的最小二乘估计, 因此, 可以假设预测残差 $\varepsilon_i = X^{(0)}(i) - \hat{X}^{(0)}(i)$ 满

足正态分布. 且 ε_i 服从均值为 0 的正态分布 $N(0, \sigma^2)$, 即 σ 表现为序列预测过程中的随机性干扰因子, 下面所要做的工作就是在正态分布均值已知情况下对未知参数 σ^2 作贝叶斯区间估计.

2.3 贝叶斯区间求解

由于误差序列 $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ 服从均值为 0, 方差为 σ^2 的正态分布, 所以可得未知参数 σ^2 的似然函数: $L(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n | \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\varepsilon_i^2 / 2\sigma^2)$. 取 σ^2 的共轭密度函数倒伽玛分布 $IGa(a, \lambda)$ 作为其先验密度函数^[9].

倒伽玛分布 $IGa(a, \lambda)$ 表达式为:

$$P(x) = \frac{\lambda^a}{\Gamma(a)} x^{-(a+1)} \exp(-\lambda/x), a > 0, \lambda > 0$$

由此可得: σ^2 的先验分布 $\pi(\sigma^2) \propto IGa(a, \lambda) = \frac{\lambda^a}{\Gamma(a)} \sigma^{-2(a+1)} \exp(-\lambda/\sigma^2)$, 则 σ^2 的后验密度函数

$$h(\sigma^2 | \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) \propto \pi(\sigma^2) L(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n | \sigma^2) = \frac{\lambda^a}{\Gamma(a)} \sigma^{-2(a+1)} \exp(-\lambda/\sigma^2) \cdot \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp(-\sum \varepsilon_i^2 / 2\sigma^2)$$

令 $\bar{X}^2 = \frac{1}{n} \sum_{i=1}^n \varepsilon_i^2$, 上式可简化为:

$$\frac{\lambda^a}{\Gamma(a)} \sigma^{-2(a+n/2+1)} \exp(-(\frac{\lambda + n\bar{X}^2}{\sigma^2})) \cdot \exp(-n/2) \left(\frac{1}{\sqrt{2\pi}}\right)^n$$

取其分布密度核函数: $h(\sigma^2 | \varepsilon) \propto \pi(\sigma^2) L(\varepsilon | \sigma^2) = \frac{\lambda^a}{\Gamma(a)} \sigma^{-2(a+n/2+1)} \exp(-(\frac{\lambda + n\bar{X}^2}{\sigma^2})) \sim IGa(a + n/2, \lambda + n\bar{X}^2/2)$, 即其分布密度核函数服从参数为: $a + n/2, \lambda + n\bar{X}^2/2$ 的倒伽玛分布.

为了简化计算, 对倒伽玛分布作如下变换^[10]: 设 $\xi \sim IGa(a, \lambda)$, 令 $y = k\lambda f(\xi)$, 则:

$$F(y) = P(k\lambda f(\xi) < y) = P(\xi < f^{-1}(y) / k\lambda) = \int_0^{f^{-1}(y)/k\lambda} [\lambda^a / \Gamma(a)] x^{-(a+1)} \exp(-\lambda/x) dx$$

$$F'(y) = f(y) = [\lambda^a / \Gamma(a)] (f^{-1}(y) / k\lambda)^{-(a+1)} \cdot \exp(-\lambda / (f^{-1}(y) / k\lambda)) (f^{-1}(y) / k\lambda)'$$

当 $k=2$, $f(\xi) = 1/\xi$ 时,

$$f(y) = [(1/2)^a / \Gamma(a)] y^{a-1} \exp(-y/2) \quad (1)$$

$$\sim Ga(a, 1/2) \sim \chi^2(2a)$$

即伽玛分布经过参数变换可用卡方分布表达出来,而卡方分布的区间求解可通过查表容易得到。

2.4 区间表示

由式 (1) 得: $2(\lambda + n\bar{X}^2/2)/\sigma \sim \chi^2(2(a+n/2))$, 取置信水平为 α , $0 < \alpha < 1$, $|2(a+n/2)|$ 代表取整数。查找卡方分布上下限 $\chi_{1-\alpha/2}^2 |2(a+n/2)|$, $\chi_{\alpha/2}^2 |2(a+n/2)|$ 得:

$$P(\chi_{1-\alpha/2}^2 |2(a+n/2)| \leq 2(\lambda + n\bar{X}^2/2)/\sigma^2 \leq \chi_{\alpha/2}^2 |2(a+n/2)|) = 1 - \alpha$$

从而得 σ 的 $1-\alpha$ 后验置信概率的贝叶斯区间估计为:

$$\left[\sqrt{\frac{2\lambda + n\bar{X}^2}{\chi_{\alpha/2}^2 |2(a+n/2)|}}, \sqrt{\frac{2\lambda + n\bar{X}^2}{\chi_{1-\alpha/2}^2 |2(a+n/2)|}} \right] \quad (2)$$

由贝叶斯区间估计可知, σ 在区间上限和下限的带状分布中满足: $P(\sigma_{\text{下限}} < \sigma < \sigma_{\text{上限}}) = 1 - \alpha$ 。

为了简化描述, 取 $(p_x - \sigma_{\text{上限}}, p_x + \sigma_{\text{上限}})$ 为可

信水平为 $1-\alpha$ 的灰色贝叶斯概率预测区间, 其中 p_x 为预测值, $\sigma_{\text{上限}}$ 为预测最大误差因子, 则该概率区间为基于灰色贝叶斯的装备库存概率预测区间。

3 算例分析

设有以下某备件 1995-2005 年之间的历史需求信息: (113, 116, 123, 128, 134, 139, 142, 145, 146, 149, 153), 现要求根据现有历史信息, 预测备件未来需求。

下面将给出灰色新陈代谢模型子序列长度 $m=4$ 的详细预测结果, 并在贝叶斯区间估计的基础上给出备件预测概率分布区间。

在 Matlab 中编程解得 MGM (1, 1) 模型的预测结果 (113, 116, 122, 129, 136, 141, 146, 147, 148, 148, 151, 157), 其中最后一项 157 为 2006 年的预测结果。预测误差序列为: (0, 0, -1, 1, 2, 2, 4, 2, 2, -1, -2)。由式 (2) 可得预测误差序列未知参数 σ 的可信水平为 $1-\alpha=0.95$ 的后验贝叶斯区间如表 1。

表 1 贝叶斯概率区间

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
上限	2.368 1	2.440 9	1.745 0	2.236 5	2.341 9	3.525 8	3.372 0	3.350 7	3.069 6	3.043 6
下限	0.196 9	0.270 7	0.363 7	0.569 1	0.685 2	1.145 4	1.188 9	1.262 4	1.222 2	1.269 9

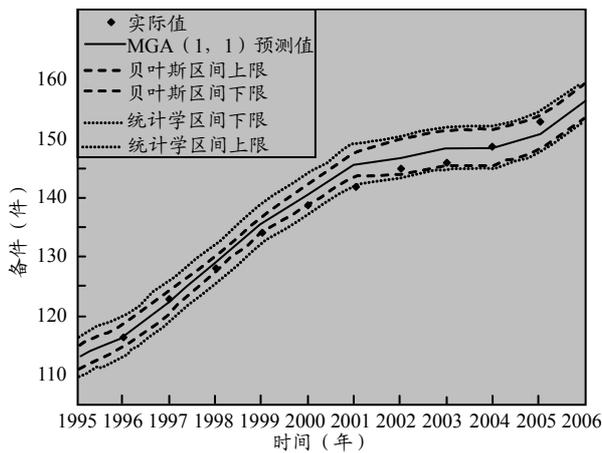


图 1 备件概率区间分布 (m=4)

由表 1 可以看出, 贝叶斯概率区间估计较好的反映了原始数据变化的趋势, 当 2000 年数据变化趋势突然改变时, 2002 年的区间上限达到了最大值 3.525 8, 但很快能够在 2003 以后降低下来, 且区间长度基本上稳定在均值 1.8 左右。

下面给出预测子序列长度 $m=4$ 、5 时的概率预

测区间分布图, 为了对比分析, 图中加入了同样置信水平下统计学概率区间。如图 1、图 2。注: 贝叶斯的可信水平与统计学的置信水平在形式上相同, 但两者概念不同, 具体区别文献[9]中有详细说明。

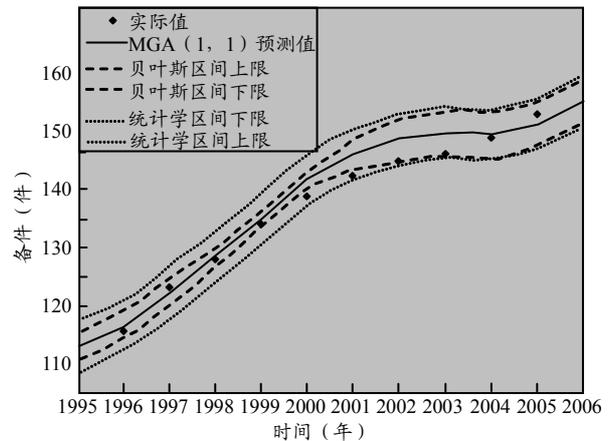


图 2 备件概率区间分布 (m=5)

从图 1、图 2 中对历史数据的拟合结果可以看出, 预测曲线上下两侧的包络线在很大程度上覆盖

了真实值, 即贝叶斯概率区间估计反映了置信水平在预测结果上的表现。通过设置合适的子序列长度可以达到很高的预测精度, 如本例中, 当预测子序列长度 m 取 4 时, 除了在曲线最大拐点处预测区间失效外, 其余各点的区间估计达到了预期的目标, 且预测最大概率区间明显优于传统经典统计学估计; 同时, 也可见, 当数据长度增大时, 两者区间估计接近, 这是因大容量样本空间下, 统计学估计在极限上趋于贝叶斯估计, 这也证明了贝叶斯估计在小样本估计时的优势。因此, 提出的灰色贝叶斯预测概率区间对装备保障决策支持意义很明显, 特别是在备件历史数据匮乏的情况下, 按照灰色理论有 4 个以上的数据即可取得理想的预测分析结果。

4 小结

此简要介绍了现代战争条件下装备保障的复杂性, 提出了用灰色预测解决装备保障中备件库存的预测问题, 并结合贝叶斯理论, 推导了共轭先验分布为倒加码分布的卡方分布变化, 以简单的形式给出了贝叶斯概率区间估计, 最后得到了备件预测的概率分布区间, 通过实例分析, 证明了贝叶斯区间估计对灰色预测的扩展意义明显。

参考文献:

(上接第 47 页)

表 2 实验数据

实验序号	1	2	3	4	5	6	7	8	9	10
配送方案总迭代次数	1 028	1 034	993	1 047	1 028	1 015	1 045	1 071	1 028	948
平均每个编组方案最优解所需迭代次数	34	34	33	35	34	34	35	36	34	32
满意解计算总耗时 (s)	6.40	6.54	6.22	6.30	6.34	6.35	6.28	6.28	6.32	6.20

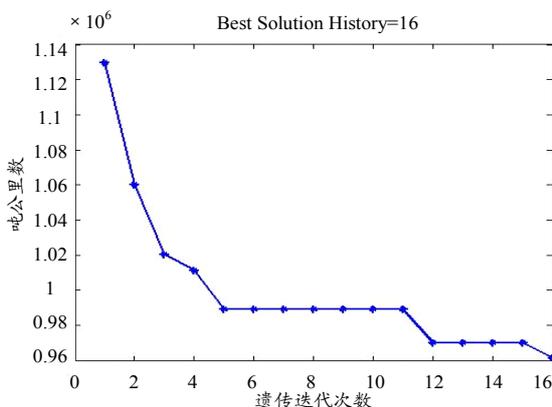


图 1 改进遗传算法寻优过程

4 结束语

该算法将贪婪扫描法与改进的基于种群扩展的多变异遗传算法进行有机融合, 有效地解决了车辆编组和配送次序问题。实验验证了该算法的简捷性、

[1] 杨军, 于丹. 一种基于可用度的装备贮存模型及其应用[J]. 系统工程理论与实践, 2004(7): 111-113.
 [2] 郭继周, 宋贵宝, 彭绍雄. 装备使用保障费用灰色建模分析[J]. 系统工程与电子技术, 2004, 26(1): 64-66.
 [3] 刘勇, 武昌, 孙鹏, 等. 战损条件下装备备件供应保障仿真研究[J]. 系统仿真学报, 2009, 21(5): 1470-1471.
 [4] Xin-ping Xiao, Shu-hua Mao. A new modeling technique of GM(1,1) prediction model[J]. Proceeding of the scnd international conference machine learning and cybernetics, 2003(11): 2-5.
 [5] 邓聚龙. 灰色预测与决策[M]. 武汉: 华中理工大学出版社, 1988: 38-40.
 [6] 杨明, 叶鹰, 李萍. 在频率学派的区间估计中引入贝叶斯先验[J]. 统计与决策(理论版), 2007(9): 36-37.
 [7] Weiping Zhang, Laisheng Wei, Yaning Yang. The superiority of empirical Bayes estimator of parameters in linear model[J]. Statistics & Probability Letters, 2005, 72: 43-50.
 [8] David G. Whiting, H. Dennis Tolley, Gilbert W. Fellingham. An empirical Bayes procedure for adaptive forecasting of shrimp yield[J]. Aquaculture, 2000, 182: 215-228.
 [9] 茆诗松. 贝叶斯统计[M]. 北京: 中国统计出版社, 1999: 20-44.
 [10] 陶靖轩, 李秀兰. 贝叶斯区间估计[J]. 中国计量学院学报, 2002, 13(2): 103-106.
 [11] 张福光, 单岳春, 洪亮, 等. 装备维护管理信息系统的设计与实现[J]. 四川兵工学报, 2009(9): 13-15.

实用性和有效性, 为今后军事物流配送路径的决策提供了参考。

参考文献:

[1] 王宗喜, 徐东. 军事物流学[M]. 北京: 清华大学出版社, 2007.
 [2] 石玉峰. 战时不确定性运输路径优化[D]. 成都: 西南交通大学博士学位论文, 2006.
 [3] 李军. 车辆调度问题的分派启发式算法[J]. 系统工程理论与实践, 1999(1): 27-33.
 [4] 张涛, 张杰, 王梦光. 不确定车辆数的车辆路径问题模型和混合算法[J]. 系统工程理论方法应用, 2002, 11(2): 201-205.
 [5] Christofides N, Mingozzi A and Toth P, The Vehicle Routing Problem. Combinatorial Optimization[M]. New York: 1979.
 [6] 谢金星, 刑文训. 网络优化[M]. 北京: 清华大学出版社, 2000.