Matlab 计算机程序实现加权马尔可夫链预测太阳黑子数

陈楚，马英钧，陈岩，赵东方
（华中师范大学数学与统计学学院，武汉 430079）

摘要：首先根据太阳黑子的周期约为 11 年，利用最优分割算法将 1900-2010 年太阳黑子数划分为 6 个等级，以 12 年作为步长，建立加权马尔可夫链预测模型。其次，经过模型检验，得到该模型在预测未来 11 年太阳黑子数时具有较高的精确度。最后，利用该模型预测 2011-2021 年这 11 年的太阳黑子数。Matlab 程序实现加权马尔可夫链预测过程，得到 2011 与 2012 年的太阳黑子数都处于偏高值段。2013-2014 这 2 年降为中值段。而 2015 年为低值段。2016-2021 年太阳黑子数都将处于最低值段。从而得知这 11 年太阳黑子数的等级经过一个周期数，表明加权马尔可夫链在预测太阳黑子数中应用的可行性。

关键词：加权马尔可夫链；太阳黑子数；Matlab 程序
中图分类号：O211

The prediction of weighted Markov chain of sunspot numbers by Matlab program .

Chen Chu, Ma Yingjun, Chen Yan, Zhao Dongfang
(School of Mathematics and Statistics,Huazhong Normal University, WuHan 430079)

Abstract: Firstly,as the sunspot cycle is 11 years, the number of sunspot from1900 to 2010 are divided into 6 levels by the optimal partition method to set up a weighted Markov chain prediction model with 12 years as one step.Secondly,after being tested, the model is affirmed with a high accuracy to predict the sunspot numbers in the next 11 yeaes,Finally, the model is used to predict sunspot numbers from 2011 to 2021 .The process is achieved by Matlab program .The result displays that the sunspot numbers of 2011 and 2012 are in a higher value section. In next 2 years ,it reduces to a median section,drops down to the lowest section in 2015 and remains the lowest section from 2016 to 2021.Thus the grade of the 11-year sunspot numbers has been through one cycle.The result shows it is feasible to forecast the sunspot numbers by the weighted Markov chain.

Keywords: weighted Markov chain; sunspot numbers; Matlab program

0 引言

太阳黑子是人们最熟悉的一种太阳表面活动。太阳黑子数反映了太阳活动强弱的变化。而这些变化对地球的影响很大，诸如大气运动、地磁变化、气候异常等这些都和太阳黑子数有关。沃尔夫（R.Wolfer）根据在过去的 288 年(1700 年-1987 年)间每年太阳黑子出现的数量和大小的观测数据推算出 11 年的周期规律。那么结合太阳黑子数的周期性，考虑太阳黑子数的数学模型极为重要。因为由此模型，我们不仅可以能够更好地研究太阳黑子的变化规律，还可以通过预测未来黑子数的数目，进一步为气象、地磁等方面预报提供重要的预报因子。

研究太阳黑子数的方法有许多，吴令云[2]等人通过建立太阳黑子的多项式趋势自回归条件异方差模型来预测太阳黑子数。基于太阳黑子数的产生及观测存在随机性，我们从概率统计角度建立加权马尔可夫链预测模型。

作者简介：陈楚，（1989-），女，研究生，数序建模。
通信联系人：赵东方，男，教授，数学建模。E-mail: 710030245@qq.com

1 加权马尔可夫链预测模型

加权马尔可夫链预测模型最近提出的一种预测模型。夏乐天[3]利用加权马尔可夫链预测和分析未来农作物的丰欠年景。王艳[4]研究了加权马尔可夫链预测武汉降雨量情况。我们考虑太阳黑子数这一列相依的随机变量序列，其各阶的自相关系数刻画了各种步长序列的相关关系及其强弱程度。因此，我们可以考虑分别依据前面若干年的太阳黑子数的数据，根据其相关关系，通过加权求和，利用已知的信息，进行太阳黑子数的加权马尔可夫链预测。

实现加权马尔可夫链预测的步骤[5]：

(1)利用最优分割算法（参考文献[4]）对太阳黑子数序列进行状态划分

（2）对太阳黑子数进行马氏性行检验

对离散状态的太阳黑子数的马氏性检验，我们可利用 χ^2 统计量进行检验。

用 f_{ij} 表示指标序列 $X_1, X_2, ..., X_n$ 中从状态 i 进过一步转移到状态 j 的频数，以所有的 f_{ij} 为元素的矩阵就是转移频数矩阵，将转移频数矩阵各列之和分别除以各行各列的总和就可以得到边缘频率，并把其作为边缘概率的估计，记为 $\hat{P}_j$，即 $\hat{P}_j = \dfrac{\sum_{i=1}^{m} f_{ij}}{\sum_{i=1}^{m}\sum_{j=1}^{m} f_{ij}}$，则统计量 $\chi^2 = 2\sum_{i=1}^{m}\sum_{j=1}^{m} f_{ij}\left|\ln\dfrac{P_{ij}}{P_j}\right|$ 服从自由度为 $(m-1)^2$ 的 χ^2 分布，其中 $\hat{P}_j$ 为 $m \times m$ 的转移概率矩阵。给点显著性水平 α，查表可得分位点 $\chi_\alpha{}^2(m-1)^2$ 值，通过计算 χ^2，若 $\chi^2 > \chi_\alpha{}^2(m-1)^2$，则可以认为太阳黑子数序列通过马氏性检验，否则太阳黑子数不可作为马尔可夫链进行处理。

（3）计算各阶权值

设 $X_t (t = 1,2,...,n)$ 表示太阳黑子数序列，计算序列各阶自相关系数的公式为 $r_k = \dfrac{\sum_{t=1}^{n-k}(X_t - \overline{X})(X_{t+k} - \overline{X})}{\sum_{t-1}^{n}(X_t - \overline{X})^2}$，其中 r_k 表示第 k 阶（滞后期为 k 年）的自相关系数，n 为太阳黑子数序列长度。归一化为各阶自相关系数 $\overline{w_k} = \dfrac{|r_k|}{\sum_{k=1}^{m}|r_k|}$，并将它们作为各阶步长的马尔可夫链的权重。

（4）构造状态转移概率矩阵

设太阳黑子数序列从状态 E_i 经过 m 步转移到状态 E_j 的次数记为 $m_{ij}{}^{(m)}$，则称以 $m_{ij}{}^{(m)}$ 为元素构成的 m 步状态转移频数矩阵，$P_{ij}{}^{(m)}$ 为由元素构成的矩阵为 m 步状态转移概率矩阵，其计算公式为 $P_{ij}{}^{(m)} = \dfrac{m_{ij}{}^{(m)}}{M_i}$，其中 M_i 为状态 E_i 出现的总次数。

（5）加权预测

第一步：以前面若干个时段各自太阳黑子数为初始状态，结合其相应的状态转移概率矩阵，就可以预测出现在阶段太阳黑子数的状态概率 $P_i^{(k)}$。（i 为状态，k 为步长）

第二步:将同一个状态的各个预测概率加权，并作为太阳黑子数处于该状态的预测概率，即 $P_i=\sum_{i=1}^{n} w_k P_i^{(k)}$，根据隶属度最大原则，取 $\max\{P_i, i\in I\}$ 所对应的状态 i 为我们所预测的状态。

2 模型拟合检验

由于太阳黑子数的周期约为 11 年,因此利用步长为 12 年的加权马尔可夫链进行预测可以充分的利用一个周期内太阳黑子的完整信息。而预测 12 年的太阳黑子数，也可以利用太阳黑子的周期为 11 年的性质对我们的预测结果进行检验。为了验证步长为 12 的加权马尔可夫链对太阳黑子的预测精确度，本文利用 1900-1989 年的太阳黑子数预测 1999-2010 年的太阳黑子数。并通过 Matlab 编程来实现。

2.1 等级划分

考虑到太阳黑子数的周期为 11 年，在这一周期内太阳黑子数将会达到一个波峰值，在波峰两边的太阳黑子数几乎对称。于是我们利用最优分割算法将 1900-2010 年的太阳黑子数划分为 6 个等级，见图 1。由图 1 可以看出，当分类数为 6 的时候，损失函数 L[P(111,6)]小于 1，此时，损失的信息较少，分级标准见表 1。

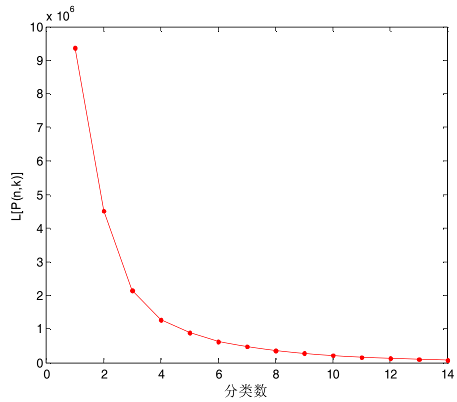

图 1 最优分割点

表 1 太阳黑子数分级标准

状态	分级标准	级别
1	x<=258.8	最低值段
2	258.8<x<=582.5	低值段
3	582.5<x<=1007.1	中值段
4	1007.1<x<=1434.4	高值段
5	1434.4<x<1905	偏高值段
6	X>1905	最高值

2.2 1999-2010 年太阳黑子预测及结果分析

下面我们将通过 1900-1998 年太阳黑子数，利用步长为 12 的加权马尔科夫链预测 1999-2010 年太阳黑子数的数据，预测的一般步骤如下：

步骤一：预测 1999 年的太阳黑子数

（1） 卡方检验。求得 1900-1998 年太阳黑子数序列的 χ^2 值为 132.9869，给定显著性水平 $\alpha = 0.05$，查表可得分位点 $\chi_\alpha{}^2(25) = 50.998$。132.9869>50.998 因此 1900-1998 年的太阳黑子数序列符合马尔科夫性，可利用加权马尔可夫链预测 1999 年的太阳黑子。

（2） 利用步长为 12 的加权马尔科夫链预测 1999 年的降雨量。通过 matlab 程序根据 1900-1998 年太阳黑子数预测 1999 年的转移矩阵如表 2 所示。

表 2 状态转移矩阵

初始年	状态	滞时（年）	权重	转移概率					
				1	2	3	4	5	6
1998	3	1	0.1481	0	9/19	8/19	2/19	0	0
1997	1	2	0.0753	8/27	10/27	4/27	4/27	1/27	0
1996	1	3	0.0089	3/26	5/26	3/13	7/26	2/13	1/26
1995	1	4	0.0740	1/25	3/25	8/25	1/5	6/25	2/25
1994	2	5	0.1031	2/23	6/23	4/23	7/23	3/23	1/23
1993	3	6	0.0893	4/9	2/9	0	1/6	1/9	1/18
1992	4	7	0.0395	7/13	3/13	0	2/13	1/13	0
1991	5	8	0.0293	1/5	1/5	1/10	1/5	1/5	1/10
1990	5	9	0.0970	1/9	0	1/9	2/9	1/3	2/9
1989	5	10	0.1345	0	1/8	0	1/8	1/2	1/4
1988	4	11	0.1242	0	0	1/3	1/12	1/2	1/12
1987	2	12	0.0768	9/22	5/22	1/11	3/22	1/11	1/22
Pi（加权和）				0.1443	0.2046	0.1793	0.1613	0.2228	0.0877

从表格中可以看出，当 $i=5$ 时，$P_i = 0.2228$ 为最大值。这说明 1999 年的太阳黑子数的状态等级为 5，取第 5 等级太阳黑子的平均数 1766.7 做为 1999 年的太阳黑子数。

步骤二：预测 2000-2010 年的太阳黑子数。

将 1999 年太阳黑子的预测等级和 1900-1998 年的太阳黑子数的实际等级作为原始数据，利用第一步的方法预测 2000 年的太阳黑子数。依次进行 2001-2010 年的太阳黑子的预测。其中，每一次预测的 χ^2 值分别为 133.2773、 136.8706、140.3784、143.7990、142.4256 、143.9289 、 145.0989 、 146.5755 、 148.0374 、 149.4842 、 150.9153, 都 大 于 $\chi_\alpha{}^2(25) = 50.998$。 因此，我们可以利用步长为 12 的加权马尔科夫链来预测 2000-2010 的太阳黑子数。

步骤三：结果分析。

利用 matlab 程序求得 2000-2010 年的太阳黑子的预测等级和预测值，将其与实际值对照如表 3。

表 3 1999-2010 实际与预测的太阳黑子数

时间	实际太阳黑子等级	实际太阳黑子数	预测太阳黑子等级	预测太阳黑子数
1999	4	1118.1	5	1766.7
2000	4	1434.4	5	1766.7
2001	4	1331.1	5	1766.7
2002	4	1249.1	5	1766.7
2003	3	762.8	3	799.26
2004	2	485.3	2	441.84
2005	2	357.4	1	135.47
2006	1	182.2	1	135.47
2007	1	90	1	135.47
2008	1	34.4	1	135.47
2009	1	37.2	1	135.47
2010	1	198	5	1766.7

从表 3，我们可以得到利用加权马尔可夫链预测 1999-2010 年这 12 年的太阳黑子数的过程中，1999-2002 以及 2005 年的太阳黑子数等级与实际相差一个等级，误差相对较小。而第 12 年预测得到的太阳黑子等级为 5，而实际太阳黑子的等级为 1，误差较大。因此，利用加权马尔可夫链预测未来 11 年的太阳黑子数具有较高的精确度。

3 加权马尔可夫链预测 2011-2021 年太阳黑子数

由前面的预测检验可知，选用步长为 12 年的加权马尔可夫链来预测未来 11 年的太阳黑子数具有较高的精确度。现运用该模型预测 2011-2021 这 11 年的太阳黑子数，预测结果如表 4 所示。

表 4 2011-2021 太阳黑子数预测等级与预测值

预测时间	太阳黑子等级	太阳黑子数
2011	5	1766.7
2012	5	1766.7
2013	3	797.52
2014	3	797.52
2015	2	440.27
2016	1	131.37
2017	1	131.37
2018	1	131.37
2019	1	131.37
2020	1	131.37
2021	1	131.37

从表 4 容易看出，2011-2021 年太阳黑子数的等级几乎经过一个周期的变化，与太阳黑子数的周期为 11 年相符合，这也说明步长为 12 的加权马尔可夫链预测太阳黑子数的预测精度较高。

为了进一步了解今年（2013 年）太阳黑子的预测情况，现给出预测 2013 年的太阳黑子状态转移矩阵，如表 5 所示。

表 5 预测 2013 年太阳黑子状态转移矩阵

初始年	状态	滞时（年）	权重	转移概率					
				1	2	3	4	5	6
2012（预测）	5	1	0.1449	0	0	1/12	1/4	7/12	1/12
2011（预测）	5	2	0.0681	0	0	5/11	2/11	3/11	1/11
2010	1	3	0.0126	5/32	5/32	3/16	9/32	3/16	1/32
2009	1	4	0.0750	2/31	3/31	8/31	8/31	8/31	2/31
2008	1	5	0.1040	0	1/30	1/3	3/10	3/10	1/30
2007	1	6	0.0922	0	3/29	11/29	10/29	5/29	0
2006	1	7	0.0459	1/28	1/4	13/28	1/4	0	0
2005	2	8	0.0208	8/25	2/25	1/5	7/25	3/25	0
2004	2	9	0.0905	5/12	1/24	5/12	1/8	0	0
2003	3	10	0.1326	1/20	1/20	1/2	7/20	1/20	0
2002	4	11	0.1287	1/17	0	5/17	2/17	8/17	1/17
2001	4	12	0.0846	0	1/8	1/4	1/8	7/16	1/16
Pi（加权和）				0.0670	0.0563	0.3229	0.2353	0.2787	0.0398

由表 5 可知， $i=3, Pi=0.3229$ 为最大值，这说明 2013 年太阳黑子数状态为 3（中值段），即太阳黑子数落在区间 [582.5，1007.1]。取 1900-2012 处于该状态的均值 797.52 作为 2013 年太阳黑子数。这对研究今年（2013 年）太阳黑子数与地球气候变化间的关系具有一定的实际意义。

4 结论

目前，预测太阳黑子数预测方法很多，但是当预测期较长时，往往精度较低。本文根据太阳黑子数周期约为 11 年的规律，利用最优分割算法将太阳黑子划分为 6 个等级，以 12 年为步长建立加权马尔可夫链预测太阳黑子数模型。其主要特点有：

（1） 本文利用加权马尔可夫链对太阳黑子数的等级进行预测，而预测的结果为相应等级太阳黑子数的均值。因此实现了非点值的状态预测，同时预测的范围也扩大了，其精确度也提高了。

（2） 利用 12 年为步长的加权马尔可夫链来预测未来的太阳黑子数，能更充分地、更合理地利用前期太阳黑子数的相关信息。另外，预测期可以达到 11 年且精度较高，较好的实现了太阳黑子长期的预测。

（3） 本文整个预测过程利用 Matlab 程序实现，预测过程更为客观准确，增加了结论的信服度。

[参考文献] (References)

[1] 1900-2010 年的太阳黑子数数据来源于人大经济论坛 http://bbs.pinggu.org/thread-1102135-1-1.html

[2] 吴令云,赵远东.太阳黑子的多项式趋势-自回归-条件异方差模型[J].应用基础与工程科学学报.2003(11): 241-245

[3] 夏乐天，彭志行，沈永梅.加权马尔可夫链在农作物年景预测中的应用[J].数学的实践与认识.2005(35):30-35

[4] 王艳.最优分割法的计算机程序实现与武汉市洪涝灾害预测[D].武汉：华中师范大学,2007

[5] 王艳,毛明志,范晶,赵东方.最优分割算法确定的加权马尔可夫链在降雨量预测中的应用[J].理论新探.2009(11):17-18

[6] 赵东方.数学模型与计算[M].北京：科学出版社，2007.

[7] 赵东方.数学实验与数学建模[M].武汉：华中师范大学出版社,2003

[8] 郭汉伟,乐贵明.太阳黑子群周期的小波分析[J].自然科学进展.2004(14):597-599

[9] 徐克红,程鹏飞.太阳黑子数时间序列的奇异谱分析和小波分析[J].测绘科学.2007(32):35-38