教学班：周四下午6.7.8



**2022数值计算方法大作业报告**

|  |  |
| --- | --- |
| 题 目： | 美国人口增长模型 |
| 姓 名： | 吕智 |
| 学 院： | 电气工程学院 |
| 专 业： | 自动化（电气） |
| 学 号： | 3200105004 |

2022年 4 月 22 日

**题目：美国的人口增长模型**

**一、问题描述**

1800-2000年美国每年的人口记录如下表所示：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **时间(年)** | 1800 | 1810 | 1820 | 1830 | 1840 | 1850 | 1860 |
| **人口(百万人)** | 5.3 | 7.2 | 9.6 | 12.9 | 17.1 | 23.2 | 31.4 |
| **时间(年)** | 1870 | 1880 | 1890 | 1900 | 1910 | 1920 | 1930 |
| **人口(百万人)** | 38.6 | 50.2 | 62.9 | 76.0 | 92.0 | 106.5 | 123.2 |
| **时间(年)** | 1940 | 1950 | 1960 | 1970 | 1980 | 1990 | 2000 |
| **人口(百万人)** | 131.7 | 150.7 | 179.3 | 204.0 | 226.5 | 251.4 | 281.4 |

利用以上数据和数值方法建立适合的人口增长模型，观察美国人口数量的变化规律，并对接下来的每隔10年进行5次人口预测，并对已有的2010和2020数据进行对比分析。（数据来源美国人口调查局）2010年人口数据为308.7百万人，2020年人口为329.4百万人。

**二、问题分析**

对于人口增长的问题，其影响因素很多，比如：人口基数，出生率，死亡率，迁入迁出，年龄组成，男女比例，人口的生育率和生育模式，国家医疗发展情况，国家的政治政策等众多的因素。但是如果把这些因素全部考虑进去，根本无从下手。因此，我们在下面提出一些假设来帮助我们更好的解决这些问题。

上述人口数据是以十年为一组离散的数据，我们可以对它先进行散点图绘制和插值图像绘制，来观察模型的大致情况，然后通过指数增长模型（Malthus模型）分析，进一步优化为一个阻滞增长模型（Logistic模型），进而进行之后五十年的数据预测。

**三、问题基本假设**

1.假设所给的数据真实可靠

2.人口数随时间连续变化并可导

3.人口增长仅和出生率和死亡率相关

4.各个年龄段的性别比例大致保持不变

5.记时刻t时的人口为p(t)，单位为百万人，p(t)为一个连续可微的函数，将1800年记为t=0，且此时人口记为P0，将之后的1810，1820直到2000记为1，2……20，单位为10年，则2010，2020即为21，22。

**四、模型建立过程**

**1.利用表中的数据，使用matlab绘制散点图**

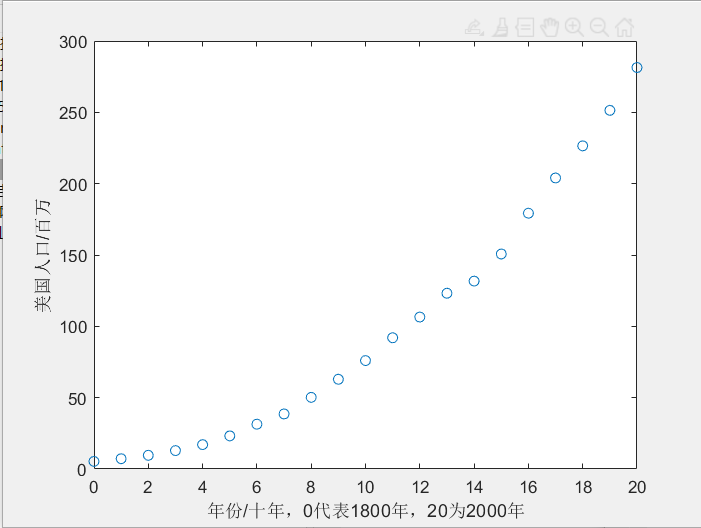


图1：1800-2000年的美国人口的散点数据图

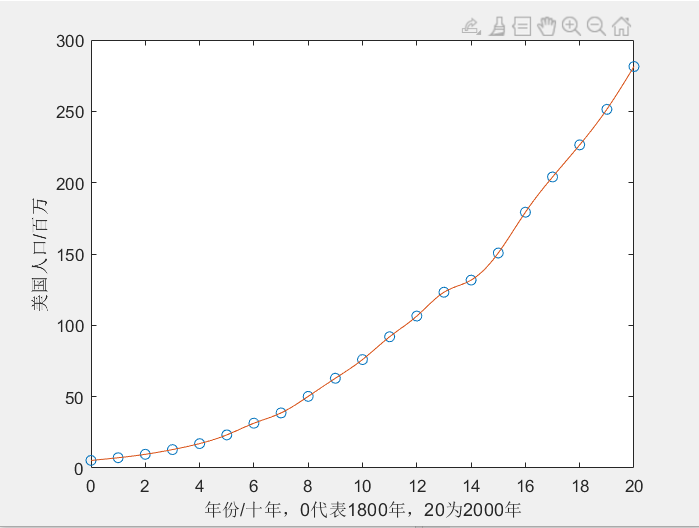
**2.利用三次样条插值方法绘制图像，观察趋势(函数在附录中)**

图2：1800-2000年的美国人口的三次样条插值数据图

观察上述散点图和三次样条插值图可知，美国人口数据的增长趋势基本符合指数增长，即指数增长模型（Malthus模型）。

**3. Malthus模型**

（1）记时刻t时的人口为p(t)，单位为百万人，p(t)为一个连续可微的函数，将1800年记为t=0，且此时人口记为P0，将之后的1810，1820直到2000记为1，2……20，单位为10年。

（2）假设人口增长率为常数r，即单位时间内x(t)的增长量满足下面的微分方程：

**（a）**

解方程，可得，当r>0时，此式将按照指数规律随时间不断增长。

**4.** **阻滞增长模型（Logistic模型）**

（1）由于自然资源，环境条件等因素对人口增长起着阻滞作用，且随着人口的增加，阻滞作用越来越大。

（2）阻滞增长模型在指数增长模型的基本假设之上，引入了自然资源和环境条件能够容纳的最大人口数M，称为人口容量，为增长率。

可得常微分方程如下：

**（b）**

解该常微分方程可得到 **（c）**

**五、模型求解**

**1.马尔萨斯模型求解**

**a.算法设计：** 由已知得，为1800年的人口数据即5.3百万人，但是仍有一个未知数r是未知的，r可以通过实际数据得线性最小二乘法求解，对于原方程直接求解是比较麻烦的，因此再两边取对数，可得,ln已知，利用最小二乘法拟合的方法把r求出来，具体操作如下：

原式可以看成y =ax + b，且b已知的形式，满足最小二乘法的条件为残差和最小，此时满足，代入美国人口数值可求得a = 0.2265，即r为0.2265.

马尔萨斯模型表达式为，图像如下页图3所示，红线表示马尔萨斯人口模型数据，蓝色圆圈代表实际人口数据。（此部分具体程序请看附录部分）

**b. 结果与分析：**从图中可以看到，马尔萨斯模型在前面几个点的吻合度较高，在后面的点距离曲线偏差较大，拟合效果并不好。通过表达式也可以对2010——2050年的数据进行计算，结果如下表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 年份 | 2010 | 2020 | 2030 | 2040 | 2050 |
| 数据 | 616.057 | 772.628 | 968.992 | 1215.3 | 1524.1 |
| 真实数据 | 308.7 | 329.4 |  |  |  |
| 误差 | 308 | 443 |  |  |  |

可以看到，误差是相当大的，这是由于指数函数有着指数爆炸的性质，随着t的增加，函数值增加的非常的快，所以用此模型计算结果并不准确。同时，我们知道人的增长不可能是无限的，所以需要对其加以限制，这就引出了我们的阻滞增长模型。

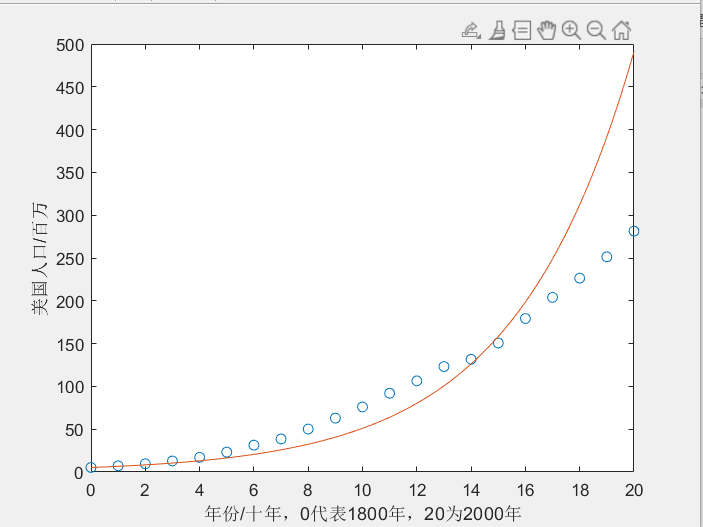


图3：马尔萨斯模型数据图

**2.** **Logistic模型求解**

(1)通过离散模型进行参数拟合

**a.算法设计：**上式为人口的阻滞模型的表达式，其中有两个参数未知，λ和M。我们可以将原方程写乘离散形式，即，其中可以将看成y，P(n)看为x，即写成了线性拟合的表达形式：y=kx+b，k=，b=。以1860为起始时间，对1860——2000年的数据进行拟合可以得到k = -6.887×，b=0.2559。即得到λ = 6.887×，M = 371.9399。将和M代入进常微分方程式中，利用四阶龙格库塔法对每年数据求解，绘制曲线图如下页所示。

**b. 结果与分析：**可以看到logistic曲线图和马尔萨斯模型图相比更加缓慢，数据也更加接近，曲线图更像一个“S”型，而马尔萨斯模型图像更像一个“J”型。

对2010——2050年的数据通过Logistic模型计算结果如下表所示：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 年份 | 2010 | 2020 | 2030 | 2040 | 2050 |
| 数据 | 301.5607 | 315.0201 | 326.2951 | 335.5944 | 343.1662 |
| 真实数据 | 308.7 | 329.4 |  |  |  |
| 误差 | 2.3% | 4.2% |  |  |  |

可以看到，2010年的误差仅为2.3%，2020年的误差为4.2%，也在5%以内，已经较为精确。

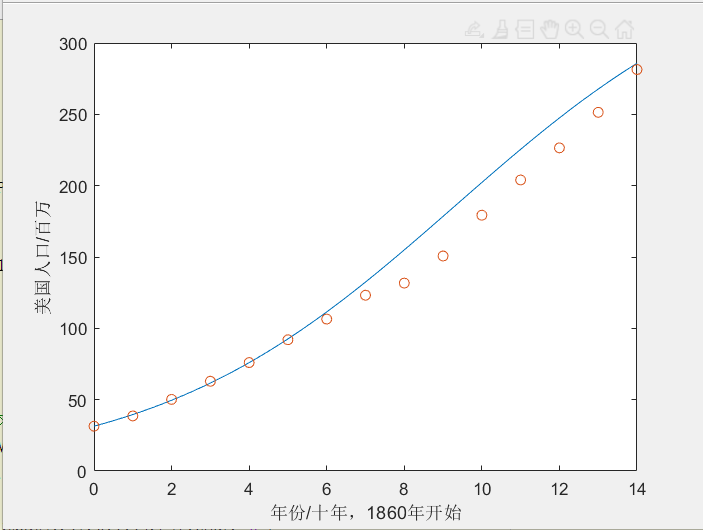


图4：离散阻滞增长模型曲线图

(2)利用数值微分近似导数来参数拟合

**a.算法设计：**原常微分方程可以变换为，可通过数值微分即差商近似微商求得。然后同样可以将看作y，P看为x，进行y=kx+b的最小二乘法线性参数拟合，其中k=-，b = 。以1860——2000年的数据进行拟合，首先先对每一个节点使用数值微分的三点公式求得，然后获得序列，进而线性拟合。可以求得k = -4.517×，b=0.2145。即得到λ =4.517×，M = 474.9105。将和M代入进常微分方程式中，利用四阶龙格库塔法对每年数据求解，绘制曲线图如下页所示。

**b. 结果与分析：**可以看到此图像和真实情况更加吻合，图像效果比离散模型要好。

对2010——2050年的数据通过数值微分Logistic模型计算结果如下表所示：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 年份 | 2010 | 2020 | 2030 | 2040 | 2050 |
| 数据 | 303.3764 | 326.1207 | 347.1195 | 366.1432 | 383.0841 |
| 真实数据 | 308.7 | 329.4 |  |  |  |
| 误差 | 1.5% | 0.99% |  |  |  |

可以看到，2010年的误差仅为1.5%，2020年的误差仅为0.99%，可以看到预测效果非常好，可以认为该模型是相当让人满意的。最终我们预测2030年美国人口将347.1195百万人，2040年美国人口366.1432百万人，2050年美国人口为383.0841百万人。

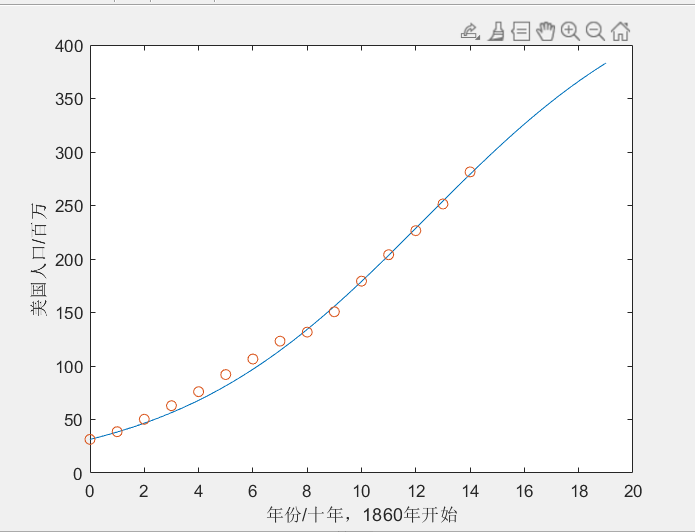


图5：数值微分方法的阻滞增长模型曲线图

**六、模型评价与推广**

Malthus数学模型在短期内具有较好的准确度，简单易行，但是不能准确的预测出人口长期的发展趋势，不具有预测人口长期增长数量的能力。

在人口增长过程中logistic模型预测的数据与题中数据更为接近，能够较好地进行预测，且2010和2020年的数据结果，与真实结果相差不大。

但是logistic模型也并不完善，实际情况下，还会涉及到国家宏观调控，科技进步带来的人口容量提升，生育率改变等多方面的综合影响，所以想要更加准确的预测未来的人口趋势变化，目前的logistic模型还远远不足。

**七、数值计算方法的使用**

**1.** **三次样条插值的使用**

，,i=1,2,3,…,n

式中存在4n个未知数，满足如下条件：

1.端点值相等，i=1,2,3,…,n-1

2.端点的一阶导值相等，i=1,2,3,…,n-1

3.端点的二阶导值相等，i=1,2,3,…,n-1

4.，i=0,2,3,…,n

以上共4n-2个式子，无法满足条件，所以需要边界条件的限制，本例中假设了在起始点和最终点的二阶导数已知且都为0的自然边界条件。

**算法设计：**

利用三弯矩法实现三次样条插值

利用S(x)在节点xi处的二阶导数值Mi=S”(xi)表示Si(x)。用S'(x)在内节点xi上的连续性和边界条件来确定Mi，利用二阶导数连续性和插值条件端点值相等和二阶导数相等

得到此式： i=1,2,3,…,n-1

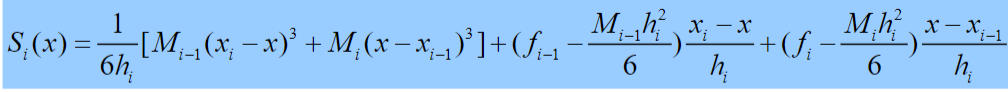
其中：

再根据自然边界条件可得到如下矩阵：

A =

X = [M1 ,…,Mn-1]

B = [g1，g2，…,gn-1]

求解出X，将参数代入到三次样条插值函数公式

绘制图像即可。

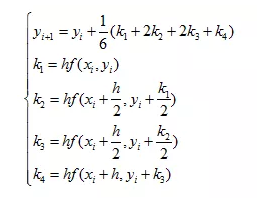
**2.** **最小二乘法曲线拟合**

构造一个能逼近列表数据的近似的数学表达式，使各数据点从总体上最贴近，而不一定要求构造的函数曲线通过所给数据点

**算法设计：**线性拟合方法即

建立此方程组： = 利用高斯消元法或高斯约当法等线性方程组求解方法求解系数a和b

**3.四阶龙格库塔法求解常微分方程**



**4.数值微分**

两点差商近似：

向前差商：

向后差商：

中间差商：

数值微分三点公式

中间差商公式和上面方法相同。

三点向前差商公式：

*三*点向后差商公式：

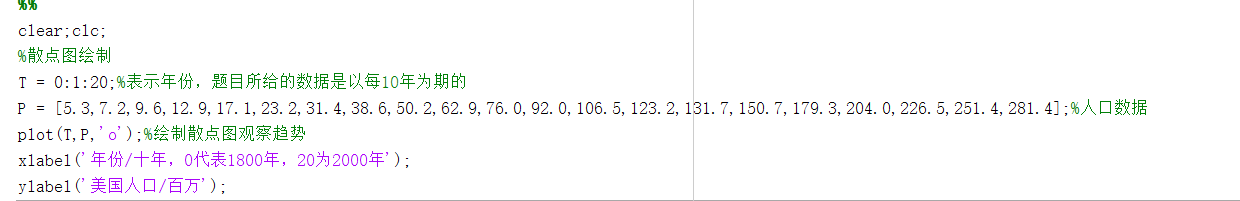
**八、参考资料**

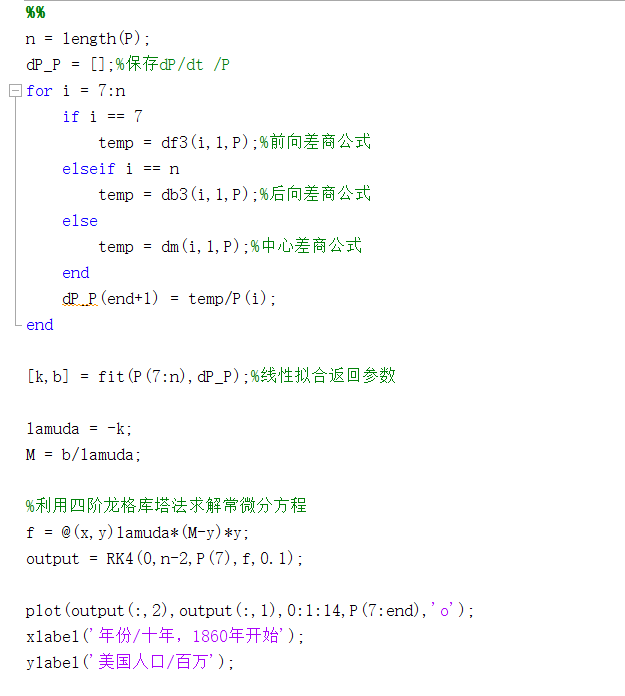
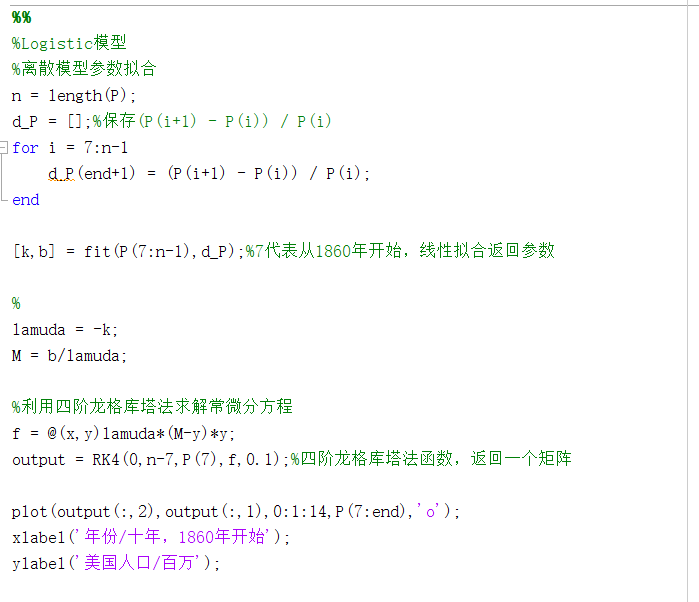
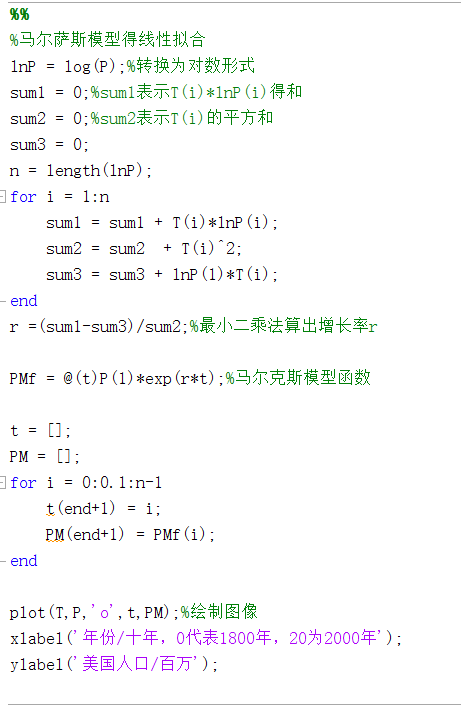
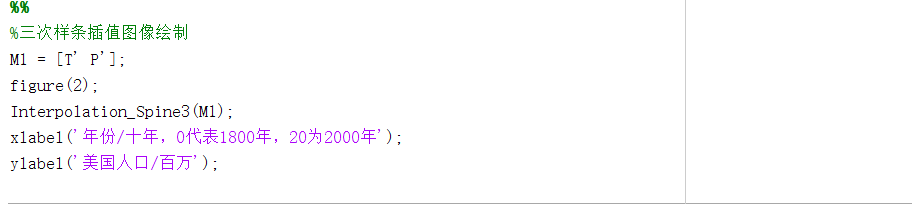
1. <https://zhuanlan.zhihu.com/p/181717510> 人口增长模型的文本介绍

2. [数据建模之人口模型（3）——模型的参数拟合\_哔哩哔哩\_bilibili](https://www.bilibili.com/video/BV1g7411N7KJ/?spm_id_from=333.788.recommend_more_video.-1) 人口增长模型的视频介绍

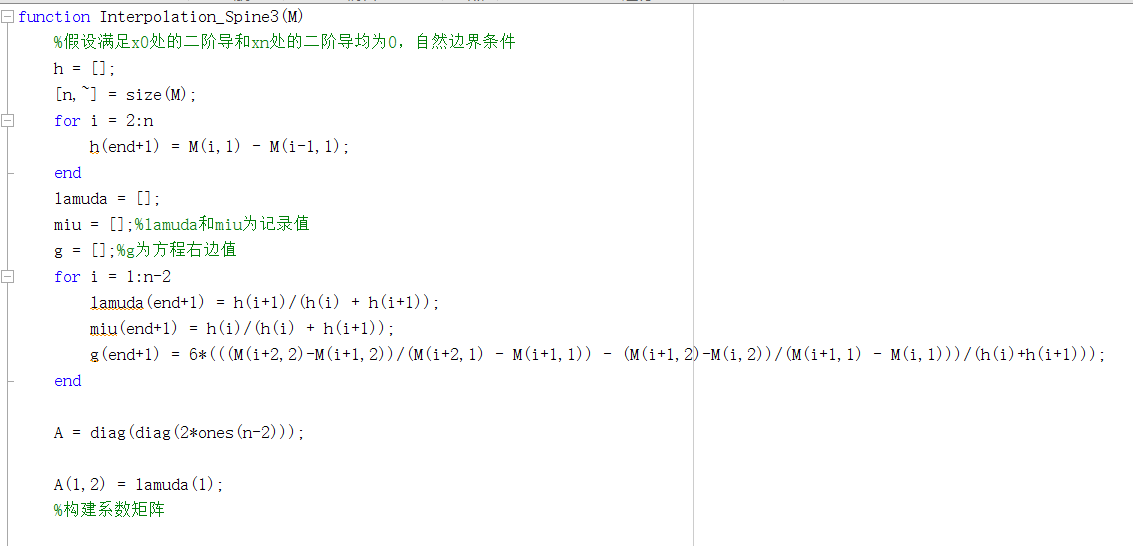
**附录：**

**1. Mian.m 主函数，所有操作在这里进行，分节运行**

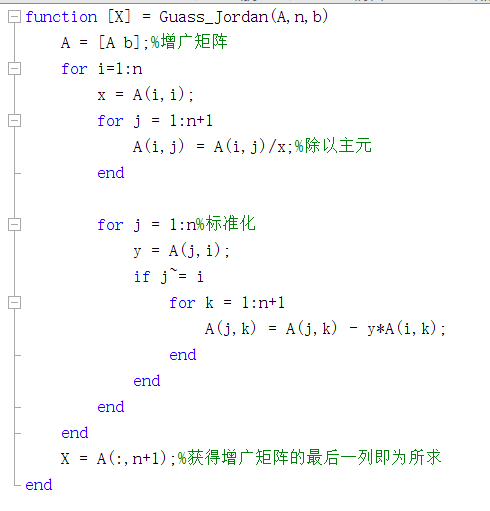




**2.Interpolation\_Spine3.m 三次样条插值函数**

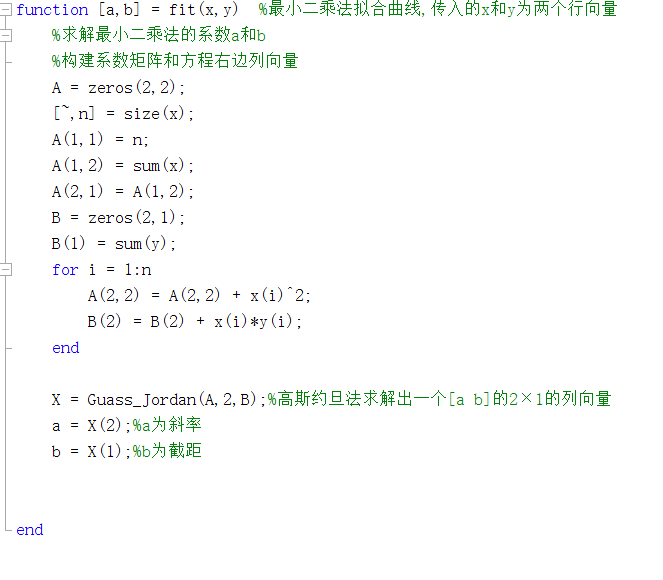


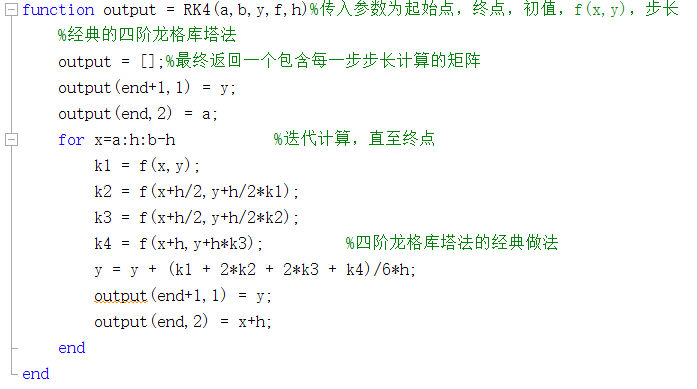
**3.Guass\_Jordan 高斯约当法解线性多项式，在拟合和三次样条插值样条函数里都会用到**



**4.fit.m 进行线性拟合，返回斜率和截距两个参数**

**5.RK4.m 四阶龙格库塔法**





**6.df3.m 和db3.m和dm.m 数值微分三点差商公式 传入参数均为起始点，步长，函数**

