微分方程模型实验报告

江鹏飞　PB20030896

2023 年 5 月 15 日

摘要

本次实验我们基于国家统计局公布的截至到 2021 年的人口数据（包括人口总数，年龄结构，性别比等数据），建立微分方程模型对我国未来 10 到 20 年的人口进行预测。其中模型可分为两大类，第一类直接基于人口总数进行预测，包括灰色预测模型（GM），logistic 及改进的 logistic 模型，以及基于感知机的模型；第二类则考虑进了年龄结构，性别比，生育率，死亡率等因素，即 Leslie 模型。我们对上述模型预测所得结果可视化以进行分析和比较，并提出若干改进建议。

一、前言

我国是一个人口大国，人口增长受多种因素影响，如何根据人口数据做出准确的预测具有重要的指导意义。一个合适的人口模型可以对我国近期的人口做出合理预测，从而对一些决策起到指导性作用。本次实验中我们提出的第一类模型，直接基于人口总数进行预测，是比较经典的预测方法，在此基础上，我们对这些经典模型进行改进并比较分析；第二类模型则考虑了出生率，性别比等多种因素，相对更具有普适性。

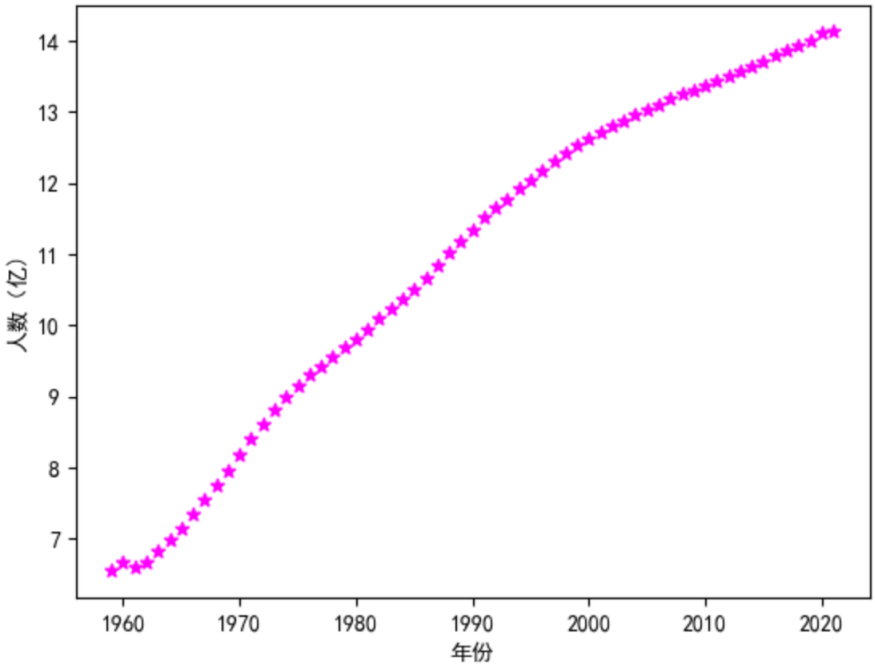


图 1: 人口数据总览

二、问题分析

使用微分方程模型求解人口预测问题，即抽象为如下问题：设 $N(t)$ 表示 t 时刻 (年份) 的总人口数量, $r(t)$ 表示 t 时刻的人口自然增长率, 由此可以通过定义直接列出方程

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = r(t)N(t)$$

给定一个初值

$$N\left(t_0\right) = N_0$$

是 t_0 年份的人口数, 就可以预测后续年份的人口数。关于 $r(t)$ 的具体表示, 根据模型的不同，对于其的假设也不同。而若考虑进年龄结构，性别比等因素，有如下的分析：

设 $F(r,t)$ 表示 t 时刻年龄不超过 r 岁的人口数, $p(r,t)$ 表示 t 时刻的年龄密度函数, 即

$$p(r,t) = \frac{\partial F(r,t)}{\partial r}$$

取最大年龄 m, 有总人口数 $N(t) = F(m,t)$ 。引入 $d(r,t)$ 表示 t 时刻 r 年龄的死亡率, 可以列出偏微分方程

$$p_t + p_r = -dp$$

它有初值条件

$$p(r,0) = p_0(r)$$

是 $t = 0$ 时给定的年龄分布; 以及边值条件限制

$$p(0,t) = p_1(t), \quad p(m,t) = 0,$$

其中 $p_1(t)$ 表示 t 时刻单位时间的出生人数, 第二个式子是最大年龄的限制。

三、符号说明

序号	符号	含义
1	t	表示年份（选定初始年份的 $t = 0$)
2	$a_i(i = 0, 1, 2, \cdots, 90)$	第 i 年龄组的生育率
3	$b_i(i = 0, 1, 2, \cdots, 90)$	第 i 年龄组的生存率
4	$x_i^{(k)}, i = 1, 2, \cdots, n$	时刻 t_k 该第 i 个年龄组中人口数目
5	$X^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)}\right)^T$	2001 年各年龄组的分布向量
6	L	Leslie 矩阵
7	r	人口增长率
8	x_m	自然环境条件下所能容纳的最大人口数量
9	R^2	可决系数的平方
10	$n_i(t), i = 1, 2, \cdots m$	在时间段 t 第 i 年龄组的人口总数
11	$b_i(i = 0, 1, 2, \cdots, 90)$	第 i 年龄组的生育率
12	$d_i(i = 0, 1, 2, \cdots, 90)$	第 i 年龄组的死亡率
13	$s_i(i = 0, 1, 2, \cdots, 90)$	第 i 年龄组的存活率

四、数学模型-微分方程模型

4.1 灰色预测模型

所谓灰色预测模型，实质就是用无规律的原始数据生成新的数据，而新生成的数据可以表现原始数据的内在规律。生成的方式主要有累加生成，累减生成，均值生成和级比生成。灰色预测模型的建立过程如下：设原始序列为

$$\mathrm{X}^{(0)} = \left\{\mathrm{x}^{(0)}(1), \mathrm{x}^{(0)}(2), \ldots, \mathrm{X}^{(0)}(\mathrm{n})\right\}$$

则 $\mathrm{X}^{(0)}$ 的 1-AGO 序列为：

$$\mathrm{X}^{(1)} = \left\{\mathrm{x}^{(1)}(1), \mathrm{x}^{(1)}(2), \ldots, \mathrm{x}^{(1)}(\mathrm{n})\right\}$$

其中 $x^{(1)}(k)=\sum_{i=1}^{k}x^{(0)}(i),\mathrm{k}=1,2,\ldots,\mathrm{n}$ 。$X^{(1)}$ 的紧邻均值生成序列 $Z^{(1)}=\left\{z^{(1)}(1),z^{(1)}(2),\ldots,Z^{(1)}(n)\right\}$, 其中

$$z^{(1)}(k)=\left[x^{(1)}(k)+x^{(1)}(k-1)\right]/2,$$
$$k=2,3,\cdots,n$$

称一阶线性微分方程

$$\frac{\mathrm{d}x^{(1)}}{\mathrm{d}t}+ax^{(1)}=b$$

为灰色微分方程 (即灰色 GM(1,1) 模型)

$$x^{(0)}(k)+az^{(1)}(k)=b$$

的白化方程。用上述微分方程的解

$$\widehat{x}^{(1)}(t)=\left(x^{(0)}(1)-\frac{b}{a}\right)\mathrm{e}^{-a(t-1)}+\frac{b}{a}$$

在 $t=k(k=1,2,\ldots,n)$ 处的值来逼近或描述 $x^{(1)}(k)$ 。微分方程中的系数 a 与常数项 b 是用下述方法确定:

$$\widehat{\boldsymbol{a}}=\begin{pmatrix}a\\b\end{pmatrix}=\left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{Y}$$

其中

$$\boldsymbol{Y}=\begin{pmatrix}x^{(0)}(2)\\x^{(0)}(3)\\\vdots\\x^{(0)}(n)\end{pmatrix},\quad\boldsymbol{B}=\begin{pmatrix}-z^{(1)}(2)&1\\-z^{(1)}(3)&1\\\vdots&\vdots\\-z^{(1)}(n)&1\end{pmatrix}$$

则还原值为**:**

$$\widehat{x^{(0)}}(k+1)=\widehat{x^{(1)}}(k+1)-\widehat{x^{(1)}}(k)=\left(1-e^{a}\right)\left(x^{(0)}(1)-\frac{b}{a}\right)\cdot e^{-ak}$$

其中 $k=1,\ldots,n+m-1$ (m 为预测数据的个数) 。

4.2 Logistic 模型及其改进

4.2.1 Logistic 模型

传统的马尔萨斯人口增长模型将人口增长率 r 定为常数, 从而人口数量 x 呈指数增长。但现实中, 人口增长受到自然资源和环境资源等的限制, 存在一个上限, 即最大人口容量 x_m 。因此, 人口增长率 r 应与人口数量 x 有关, 以此可建立模型:

$$\begin{cases}\frac{dx}{dt}=r(x)x\\x(0)=x_0\end{cases}$$

其中, 受到人口容量的影响, $r(x)$ 应该为减函数, 假设为:

$$r(x)=r_0-\alpha x$$

r_0 为固有增长率。当 x 达到最大人口容量 x_m 时, $r\left(x_m\right)=0$, 则得到 $\alpha=\frac{r_0}{x_m}$, 故上述模型整理为:

$$\begin{cases}\frac{dx}{dt}=\left(r_0-\frac{r_0x}{x_m}\right)x\\x(0)=x_0\end{cases}$$

解微分方程, 得:

$$x(t) = \frac{x_m}{1 + \left(\frac{x_m}{x_0} - 1\right) e^{-r_0 t}}$$

4.2.2 Logistic 模型改进

原 Logstic 模型仅假设 r 是人口 x 的函数，实际上，人口自然增长率还受到时间 t 的影响。假设 r 是时间 t 和人口 $x(t)$ 的函数, 即有 $r(t,x) = r(t)\phi(x)$, 从而 $x(t)$ 满足微分方程:

$$\begin{cases} \dfrac{dx}{dt} = r(t)\phi(x)x, \quad x(0) = x_0 \end{cases}$$

解方程组得

$$\int \frac{dx}{x\phi(x)} = \int r(t)dt, \quad x(0) = x_0$$

假设 $\phi(x)$ 为 **x** 的一次函数：
假设 $\phi(x) = 1 - \alpha x (\alpha > 0$ 且为常数), 代入上式, 得

$$x(t) = \frac{x_0 e^{\int_0^t r(s)ds}}{1 - \alpha x_0 + \alpha x_0 e^t r(s)ds}$$

记 x_m 为自然资源和环境条件限制所能容许的最大人口容量总数, 应有 $r(x_m) = 0$, 从而 $\alpha = \frac{1}{x_m}$, 代入得

$$x(t) = \frac{x_m}{1 + \left(\frac{x_m}{x_0} - 1\right) e^0 r(s)ds}$$

假设 $r(t)$ 为 t 的二次函数假设 $r(t) = a + 2bt + 3t^3$ ，代入，得

$$x(t) = \frac{x_m}{1 + \left(\frac{x_m}{x_0} - 1\right) e^{-(at + bt^2 + ct^3)}}$$

利用 1969 年—-2012 年的数据对参数 a, b, c, x_m 进行拟合, 可以得到预测人口数目。

4.3 感知机模型

常规的感知机模型一般用于分类问题，目标是寻找一个超平面将数据分为两部分，其算法的大致流程如下：
输入: 训练数据集 $T = \{(x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)\}$, 其中 $x_i \in \mathcal{X} = \mathbf{R}^n, y_i \in \mathcal{Y} = -1, +1\}, i = 1, 2, \cdots, N$; 学习率 $\eta (0 < \eta < 1)$;
输出: w, b; 感知机模型 $f(x) = \text{sign}(w \cdot x + b)$.
(1) 选取初值 w_0, b_0
(2) 在训练集中选取数据 (x_i, y_i)
(3) 如果 $y_i (w \cdot x_i + b) \leqslant 0$

$$w \leftarrow w + \eta y_i x_i$$

$$b \leftarrow b + \eta y_i$$

(4) 转至 (2), 直至训练集中没有误分类点.
感知机的损失函数是误分类点到超平面 S 的总距离

$$d = \frac{1}{\|w\|} \left| w^T x + b \right|$$

对于误分类的点：

$$y_i\left(w_i^T x + b\right) \leq 0$$

假设误分类点的集合为 M ，所有误分类点到超平面 S 的距离：

$$d = -\frac{1}{\|w\|}\sum_{x_i \in M} y_i\left(w^T x_i + b\right)$$

所以感知机的损失函数为:

$$L(w,b) = -\frac{1}{\|w\|}\sum_{x_i \in M} y_i\left(w^T x_i + b\right)$$

我们的问题就是要找到最优的 w ， b ，使得损失函数最小。

$$\min_{w,b} L(w,b) = -\frac{1}{\|w\|}\sum_{x_i \in M} y_i\left(w^T x_i + b\right)$$

而对于预测人口这类回归问题，只需要把算法中最后一步输出值和损失函数修改即可：
将最后的 $sign$ 函数修改为：

$$f(x) = x$$

那么，最后的输出值就是一个实数而不是 1 或-1 中的一个值了，这样就达到了回归的目的。对于损失函数选取，通常情况下，回归问题使用的损失函数为：

$$e = \frac{1}{2}(y - \hat{y})^2$$

y 表示训练样本里面的标记，也就是实际值；而 $\hat{y}$ 表示模型计算的出来的预测值。在 n 个样本的数据集中，可以将总误差 E 记为：

$$E = \frac{1}{2}\sum_{i=1}^{n}(y^i - \hat{y}^i)^2$$

我们的目的，是训练模型，求取到合适的 w，使上述损失函数取得最小值。

4.4 Leslie 模型

4.4.1 构建单性别的 Leslie 矩阵

我们先考察女性单性别模型, 然后再扩展到双性别模型。假设有 18 个年龄分组, 即 $0 \sim 85$ 岁及其以上, 年龄间隔为 5 岁。初始人口 $P^{(0)}$

$$\mathbf{P}^{(0)} = ({}_5P_0, {}_5P_5, \cdots, {}_5P_{80}, {}_5P_{85+})'$$

有 18 个元素, 分别对应 $0 \sim 85$ 岁的 17 个 5 岁年龄组, 加上 85 岁及其以上的 1 个开放年龄组, 共计 18 个年龄组。上标的 (0) 表示的是预测起点。向量前 17 个元素 ${}_5\mathrm{P}_x$, 是指确切年龄 $(x, x+5)$ 的人口数, 最后一个元素是 85 岁及其以上年龄组的人口数。下述变量也有两个意义类似的下标。存活矩阵 S 是一个 18×18 的方阵:

$$\mathbf{S} = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ {}_5s_0 & & & & & \\ & {}_5s_5 & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & {}_5s_{80} & s_{85+} \end{pmatrix}$$

各元素是相应年龄组的存活率。显然, $\mathrm{SP}^{(0)}$ 就是 5 年后该初始人口仍然存活着的人口列向量。得到的列向量第一个元素是 0 , 最后一个元素是由两项相加的。除了与死亡相关的 S 矩阵, 我们还需要另一个出生矩阵 B :

$$\mathbf{B}=\begin{pmatrix} b_0 & \cdots & b_{85+} \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

它也是一个 18×18 的矩阵, 除了第一行 $\left\{b_0, b_5, b_{10}, \cdots, b_{80}, b_{85_+}\right\}$, 其余元素均为 0 , 第一行中也只有第 4 到第 10 个元素为非 0 , 即它们对应的是 $15\sim 49$ 岁育龄妇女组成的 7 个 5 岁年龄组。这些非 0 项在后面将再讨论。Leslie 矩阵 $M=S+B$.M 的形式如下所示:

$$\mathbf{M}=\begin{pmatrix} b_0 & \cdots & \cdots & \cdots & \ldots & b_{85+} \\ {}_5s_0 & & & & & \\ & {}_5s_5 & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & {}_5s_{80} & s_{85+} \end{pmatrix}$$

5 年期的人口预测即可表示为:

$$\mathrm{P}^{(5)}=\mathrm{MP}^{(0)}$$

我们把 M 也附上一个上标, 则上式可以改写为通式:

$$\mathrm{P}^{(\mathrm{n}+5)}=\mathrm{M}^{(\mathrm{n}+5)}\mathrm{P}^{(\mathrm{n})}$$

这里, $\mathrm{P}^{(\mathrm{n}+5)}$ 为 $\mathrm{n}+5$ 年后的人口数, $\mathrm{M}^{(\mathrm{n}+5)}$ 是 5 年期间的 Leslie 转移矩阵, P^{n} 为第 n 年的人口数。

4.4.2 扩展为双性别模型

在单性别模型的基础上, 将此矩阵扩展为双性别的。首先将这个扩大后的 36×36 矩阵写出来, 然后再给出相应的解释:
同时, 将 $\mathrm{P}^{(0)}$ 改写成双性别的, 其中星号代表的是男性人口的相应变量:

$$\mathbf{P}^{(0)}=\left({}_5P_0, {}_5P_5, \cdots, P_{85+}, {}_5P_0^*, {}_5P_5^*, \cdots, P_{85+}^*\right)'$$

很显然, 大矩阵 M 是由 4 个 18×18 的子矩阵构成的。左上角的子矩阵即是女性单性别的 Leslie 矩阵。左下角子矩阵第一行各元素与左上角子矩阵第一行各元素是相对应的, 将在后面予以详细讨论。那么, 在已经改写为双性别模型的公式 $P^{(n+5)}=M^{(n+5)}P^{(n)}$ 里, 大矩阵 M 的第 19 行与 $P^{(n)}$ 相乘, 得到的是的第 19 个元素, 也就是 5 年后男性人口中 $0\sim 4$ 岁年龄组的人, 其它年龄以此类推。最后，右上角是零矩阵，右下角的子矩阵是男性单性别的存活矩阵。

4.4.3 矩阵各元素

对于存活矩阵。对一个 5 年期内的 5 岁年龄组, 可以得到年龄组 $(x, x+5)$ 的 5 个死亡率预测值, 记为 $q_{x1}, \cdots, q_{x5}$, 用 $\exp\left(-\sum_{i=1}^{5} q_{xi}\right)$ 作为 5 年期存活率的预测 (Keyfitz & Caswell, 2005), 其中 x 是指

$$\mathbf{M}=\left(\begin{array}{cccccc|cccccc}
b_0 & \cdots & \cdots & \cdots & \cdots & b_{85+} & 0 & 0 & \cdots & \cdots & \cdots & 0 \\
{}_5s_0 & & & & & & 0 & 0 & \cdots & \cdots & \cdots & 0 \\
 & {}_5s_5 & & & & & 0 & 0 & \cdots & \cdots & \cdots & 0 \\
 & & \ddots & & & & \vdots & \vdots & & & & \vdots \\
 & & & \ddots & & & \vdots & \vdots & & & & \vdots \\
 & & & & {}_5s_{80} & s_{85+} & \vdots & \vdots & & & & \vdots \\
\hline
b_0^* & \cdots & \cdots & \cdots & \cdots & b_{85+}^* & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & {}_5s_0^* & & & & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & {}_5s_5^* & & & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & & \ddots & & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & & & \ddots & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & & & & {}_5s_{80}^* & s_{85+}^*
\end{array}\right)$$

年龄组。我们在这里其实假设了年龄组的所有人口都集中在该年龄组的左端点, 这是针对年鉴数据所作的一个近似。这样就得到了存活矩阵中 ${}_5\,\mathrm{s}_0,\cdots,{}_5\,\mathrm{s}_{80}$ 和 s_{85+} 的指数式表达, 也类似地得到男性人口相应的带星变量。

对于出生矩阵，与存活矩阵相似, 也在类似的假设上用一个指数式来描述生育率。得到的是 $(x,x+5)$ 年龄组的 5 个生育率的预测, 记为 $f_{x1},\cdots,f_{x5}$, 用 $\exp\left(\sum_{i=1}^{5}f_{xi}\right)-1$ 作为 5 年期生育率的预测。特别要注意, 统计数据中给出的年龄别生育率是不分出生人口的性别的, 换言之, 这里得到的 $\exp\left(\sum_{i=1}^{5}\mathrm{f_{xi}}\right)-1$ 实际上是双性别的加和。我们将通过出生性别比 s 把两者分离开来。另外, 此处感兴趣的是 $0\sim4$ 岁年龄组的存活人口数, 所以还需要乘以一个 $0\sim4$ 岁年龄组的存活率 ${}_5\,\mathrm{s}_0$ 和 s_0^* 。因此, 实际上是通过下面的公式得到出生矩阵中的非 0 元素:

$$\begin{cases} b_x=\frac{100}{100+s}\left(\exp\left(\sum_{i=1}^{5}f_{xi}\right)-1\right)\cdot{}_5s_0 \\ b_x^*=\frac{s}{100+s}\left(\exp\left(\sum_{i=1}^{5}f_{xi}\right)-1\right)\cdot{}_5s_0^* \end{cases}$$

其, $x=15,20,\cdots,45$ 。

五、实验结果

5.4.4 GM 模型

我们首先展示灰色模型对于 2010-2021 年的人口预测情况：

年份	实际人数	预测人数	误差	相对误差
2010	134091	134850.7	−759.3	0.00566603
2011	134735	135543.1	−808.9	0.00599825
2012	135404	136199.6	−709.2	0.00587579
2013	136072	136819.8	−747.8	0.00549587
2014	136782	137403.6	−213.1	0.00454482
2015	137462	137950.8	−107.4	0.00355638
2016	138271	138461.3	−104.7	0.00137642
2017	139008	138934.8	−144.8	0.00052619
2018	139538	139371.3	255.8	0.00119434
2019	140005	139770.6	162.5	0.00167371
2020	141212	140132.7	1077.1	0.00764281
2021	141260	140457.4	803.6	0.00568116

其对于此后十年的人口预测情况如下：

年份	2022	2023	2024	2025	2026
预测人数	140744.8	140994.7	141207.1	141382.0	141519.4
年份	2027	2028	2029	2030	2031
预测人数	141619.4	141681.9	141707.0	141694.7	141645.2

5.4.5 Logistic 模型及其改进

该部分的实验结果如下：可以看到 logistoc 模型在改进前后对于已知数据的拟合的准确率都是较高

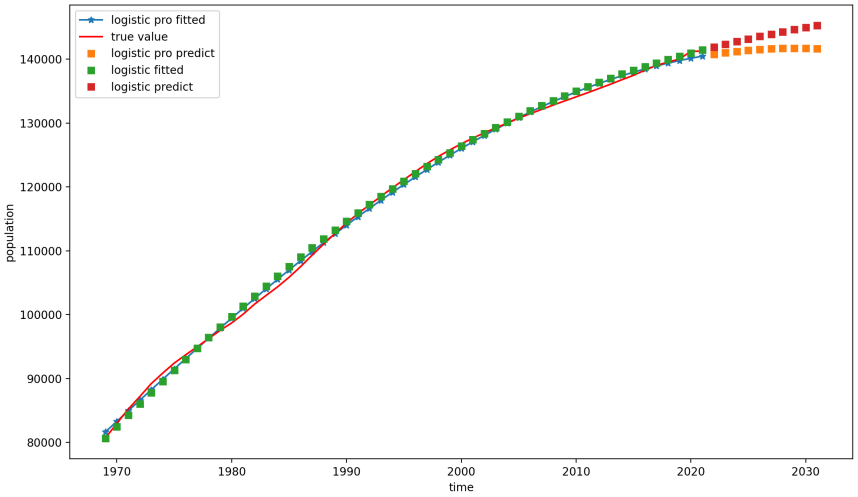


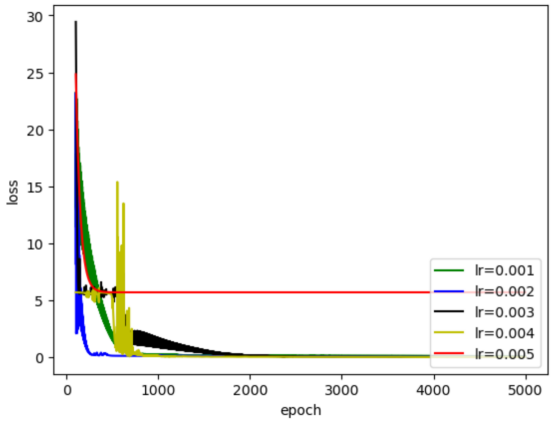
图 2: logistic 模型预测结果对比

的，但在未来人口的预测上原始 logistic 模型的预测明显偏高，因其是经典的人口增长模型，只是考虑到环境的最大承受能力，这种考量对于处理封闭体系的自然环境，例如某区域的动物种群，是合理的。但

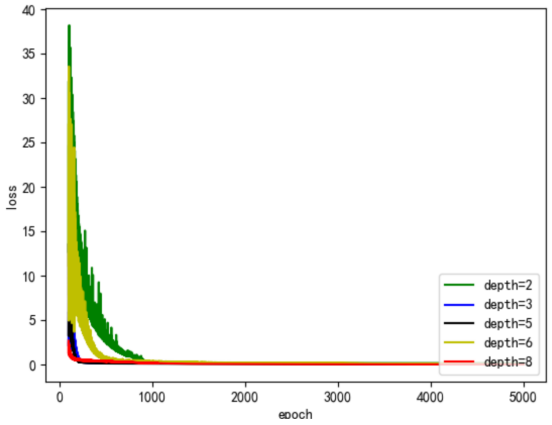
是，考虑人口模型时，因受到生育率、老龄化程度影响，自然环境的最大承受能力的影响反而稍弱。相较来说，经改进后，预测值更加合理，因人口增长率 $r(t,x)$ 不会无休止地增加，也不会一直减小。

5.5 感知机模型

感知机模型对于人口的预测主要受学习率，网络深度的影响，我们绘制调参过程中的损失曲线如下：



(a) lr-loss 曲线



(b) depth-loss 曲线

最优参数设置如下：

```
LR = 0.003 # 学习率
HIDDEN_SIZE = 32 # 隐藏层网络宽度
HIDDEN_LAYERS = 8 # 隐藏层深度
```

其对于未来 10 年的预测结果如下：

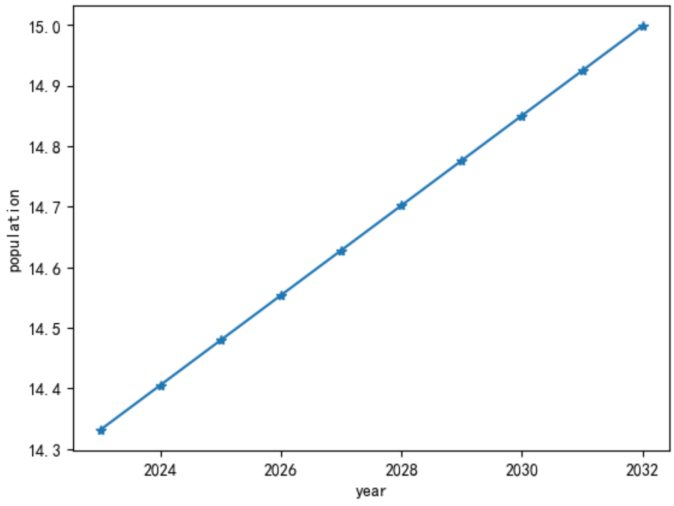


图 3: 感知机模型预测结果

5.6 Leslie 模型

调整 β 的不同取值，我们可以得到在不同 β 值下的预测曲线：

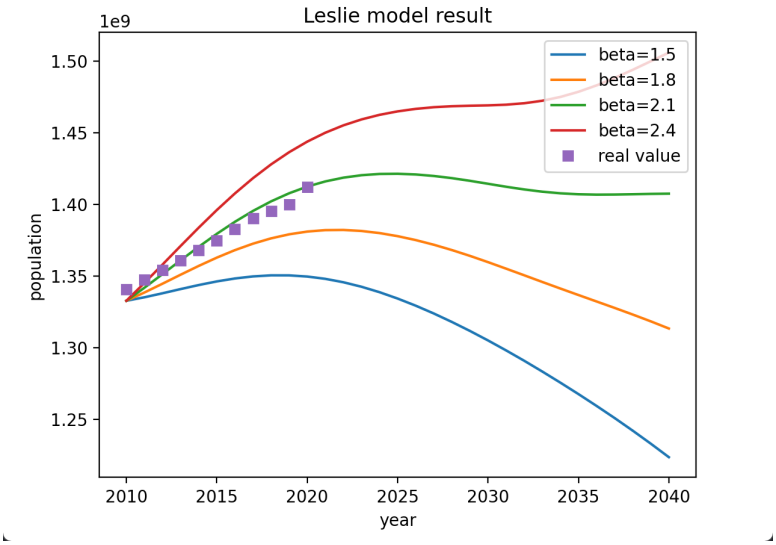


图 4: 不同 β 值下的预测曲线

不同的 β 值，体现不同生育率对于人口增减的影响。由上述结果可知参数 β 值为 2.1 时，对于数据拟合较好，在 β 过大的时候，人口将进一步逐年上升；在 β 过小的时候，较高的老龄化程度和较低的生育率共同作用使得人口出现下降的趋势。

上述参数 β 的取值均为定值，考虑新政策影响时，2015 年国家统计局通过 1% 抽样调查的方式也给出了女性的总和生育率，约为 1.55，又根据中华人民共和国卫计委的目标，在一段时间内将中国的总和生育率提升至 1.7 1.8 左右，从而我们得到这一模型下 $\beta(t)$ 所需要满足的条件：

（1）从政策实行到人民响应有一个过程，因而总和生育率应随时间递增

（2）2015 年时总和生育率为 1.55，一段时间后应接近目标值 1.8

并在这里假设一开始几年大家积极响应国家政策，综合生养率增加较快（即总和生育率的一阶导数递减）。由以上条件和假设，选择其中一个较为适合的反比例函数:

$$\beta(t) = 1.8 - \frac{0.25}{x - x_0}$$

通过使用改进后的 $\beta(t)$ 重新运行程序，得到的人口随时间变化的图像如下：

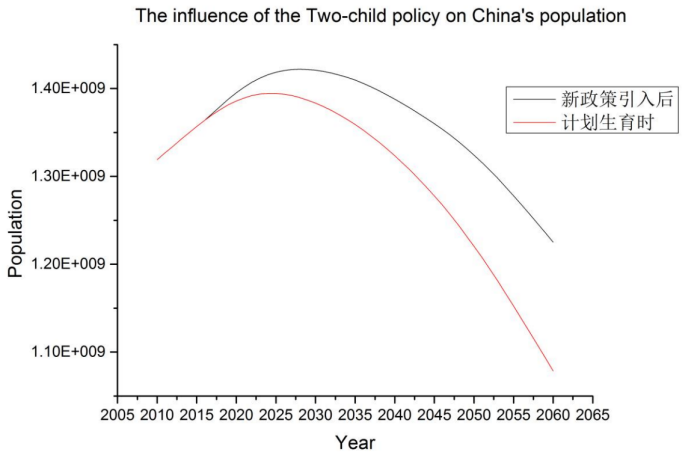


图 5: 改进后的人口预测图

六、结论

本次实验主要建立了两类人口预测模型，GM 模型，Logistic 模型，或感知机模型，这些传统人口预测模型一般对于短期预测具有较好的表现，但其不足之处也比较明显，如考虑影响人口的因素过少，建模过于理想；而相较于传统模型，Leslie 模型则综合考虑进了多种因素，对于长期预测也具有较好的结果。即便如此，我们的模型还有需要改进的地方。例如，死亡率 d 等因素被视作常数，这样考虑显然简化了问题，一种可能的改进方法是，采用随机方法，估计出生育率和死亡率等参量，拟合出带有随机项的平稳时间序列，进而对未来的各变量进行估计。

七、参考文献

[1] 付艳茹，基于 MATLAB 的人口预测研究，华东师范大学硕士论文，2010 年 11 月
[2] 宋佩锋，人口预测方法比较研究，武汉科技大学硕士论文，2013 年 5 月
[3] 任强; 侯大道，人口预测的随机方法: 基于 Leslie 矩阵和 ARMA 模型，35(2)，28-42，2011