

参赛密码 _____
(由组委会填写)

第十一届华为杯全国研究生数学建模竞赛

学 校 青岛科技大学

参赛队号 10426019

1.王玉江

队员姓名 2.陈桂兵

3.严春梅

参赛密码 _____
(由组委会填写)



第十一届华为杯全国研究生数学建模竞赛

题 目 人体营养健康角度的中国果蔬发展战略研究

摘 要：

水果和蔬菜是重要的农产品，主要为人体提供矿物质、维生素、膳食纤维。因此，预测我国果蔬的消费与生产趋势，科学地规划与调整我国果蔬的中长期的种植模式，具有重要的战略意义。

针对问题一，首先选取需要研究的主要水果、蔬菜品种，以其总计含量分别达到各自总产量的 90% 为约束条件，采用多目标规划（产量最多、营养含量最高、种类最少）筛选出主要的水果、蔬菜品种（水果 10 种、蔬菜 8 种，结果见表 4、表 5）。随后，查询 2002 至 2010 年主要品种的产量，通过两种方法（损耗率、进出口差量）计算人均消费量，用灰色预测 GM(1,1) 和曲线拟合两种模型预测未来人均消费量，并分别检验拟合效果、进行误差分析，发现基于损耗率估计消费量灰色预测的效果较理想。所以以苹果为例，综合考虑损耗率和进出口量利用灰色预测模型估计其 2010 至 2020 年的人均消费量（见表 9、表 12），拟合出曲线图（见图 2、图 6），分析其发展趋势。

针对问题二，首先依据 10 种营养成分在主要果蔬产品中含量的相似性，运用 Spss 软件对 10 种营养成分进行 Q 型聚类降维得 6 大类营养成分（结果见表 12），每类营养成分都筛选出一个营养成分作为代表分别是膳食纤维、维生素 A、维生素 C、维生素 E、钙、锌。采用正态分布中间型模型计算每种营养成分年均实际摄入量关于标准摄入量的隶属度，隶属度越接近 1 表明实际摄入量越趋近于标准量，以此评价 2014 年的营养年摄入水平。用维生素 A、C、E 隶属度的乘积量化维生素这一大类营养元素趋于健康的程度，同样对钙锌采取相同的措施作为矿物质的量化值，关于时间作图可知 2014 至 2019 年中国居民的营养健康状况趋于好转，2019 至 2020 年趋于恶化。

针对问题三，考虑到不同地区不同季节的蔬菜水果价格不同，以北方地区为例进行研究。模型中，首先充分考虑了水果和蔬菜之间的相互替代关系，根据其营养成分含量的相似性对蔬菜水果分别进行聚类分析，共分为七类。其中每类蔬菜水果的营养成分含量相似，以每类中价格较低的蔬菜或水果来代替此类果蔬。其次，以七种果蔬为样本，

建立以购买成本最低为目标，以六种营养成分分别满足人体营养需求为约束条件的目标优化模型。通过模型求解，得出合理的人均年度消费量（见表 20）。

针对问题四，首先，在综合考虑多方面因素的基础上，以满足获得人体必需的营养元素摄入量，进出口量、土地面积、种植规模变动等因素进行约束，建立以居民购买成本最低和种植者收益最大的双目标优化模型。对模型的求解，采用线性加权法转化成单目标规划问题求解，借助 *Matlab* 软件求解，得到六类主要蔬菜水果的人均年消费量的最优解（见表 23），并且进行了模型的灵敏度分析。最后，通过比较分析给出到 2020 年我国果蔬类产品的调整战略。

最后，结合对前述问题的分析结果，给出了我国果蔬产业发展的政策建议。

关键词： 多目标规划 灰色预测 聚类分析 隶属度 灵敏度分析

1、问题重述

据相关资料显示，人体需要的营养素主要有蛋白质、脂肪、维生素、矿物质、糖和水。其中维生素对于维持人体新陈代谢的生理功能是不可或缺的，多达 30 余种，分为脂溶性维生素（如维生素 A、D、E、K 等）和水溶性维生素（如维生素 B1、B2、B6、B12、C 等）。矿物质无机盐等亦是构成人体的重要成分，约占人体体重的 5%，主要有钙、钾、硫等以及微量元素铁、锌等。另外适量地补充膳食纤维对促进良好的消化和排泄固体废物有着举足轻重的作用。

水果和蔬菜是重要的农产品，主要为人提供矿物质、维生素、膳食纤维。近年来，中国水果和蔬菜种植面积和产量迅速增长，水果和蔬菜品种也日益丰富，中国居民生活水平不断提高，人们对人体营养均衡的意识也有所增强。然而多数中国居民喜食、饱食、偏食、忽视人体健康所需的营养均衡的传统饮食习惯尚未根本扭转，这就使得我国的果蔬消费（品种和数量）在满足居民身体健康所需均衡营养的意义下，近乎盲目无序，进而影响到果蔬生产。

因此，预测我国果蔬的消费与生产趋势，科学地规划与调整我国果蔬的中长期的种植模式，具有重要的战略意义。为此请你们协助完成以下几项任务：

第一，科学决策的基础是比较准确地掌握情况。但我国蔬菜和水果品种繁多，无论是中国官方公布的数据还是世界粮农组织（FAO）、美国农业部（USDA）等发布的数据均不完整，缺失较为普遍，而且品种、口径不一。我们既不可能也没有必要了解全部数据，对这样的宏观问题，恰当的方法是选取主要的水果和蔬菜品种进行研究。因此，要求主要的水果、蔬菜品种不仅总计产量应分别超过它们各自总产量的 90%，而且这部分品种所蕴含营养素无论在成分上还是在含量上都满足研究的需要。请你们运用数学手段从附件表格中筛选出主要的水果和蔬菜品种，并尝试用多种方法建立数学模型对其消费量进行估计，研究其发展趋势。

第二、摸清我国居民矿物质、维生素、膳食纤维等营养素摄入现状。请结合为保障人体健康所需要的各种营养成分的范围（见附件和参考文献）和前面预测的人均消费结果，评价中国居民目前矿物质、维生素、膳食纤维等营养的年摄入水平是否合理。按照水果和蔬菜近期的消费趋势，至 2020 年，中国居民的人体营养健康状况是趋于好转还是恶化？请给出支持你们结论的充分依据。

第三，不同的蔬菜、水果尽管各种营养素含量各不相同，但营养素的种类大致相近，存在着食用功能的相似性。所以，水果与水果之间、蔬菜与蔬菜之间、水果与蔬菜之间从营养学角度在一定程度上可以相互替代、相互补充。由于每种蔬菜、水果所含有的维

生素、矿物质、膳食纤维成分、含量不尽相同，价格也有差异，因而在保证营养均衡满足健康需要条件下，如何选择消费产品是个普遍的问题。请你们为当今中国居民（可以分区域分季节）提供主要的水果和蔬菜产品的按年度合理人均消费量，使人们能够以较低的购买成本（假定各品种价格按照原有趋势合理变动）满足自身的营养健康需要。

第四，为实现人体营养均衡满足健康需要，国家可能需要对水果和蔬菜各品种的生产规模做出战略性调整。一方面国家要考虑到居民人体的营养均衡，并使营养摄入量尽量在合理范围内；另一方面也要顾及居民的购买成本，使其购买成本尽可能的低；同时还要使种植者能够尽量获得较大收益；而且，作为国家宏观战略，还要考虑进出口贸易、土地面积等其他因素。请你们基于上述考虑，建立数学模型重新计算中国居民主要的水果和蔬菜产品的按年度合理人均消费量，并给出到 2020 年我国水果和蔬菜产品生产的调整战略。

第五，结合前面的研究结论，给相关部门提供 1000 字左右的政策建议。

2、模型假设

- 1、水果、蔬菜从产生到人消费损耗只发生在田间地头到大市场、市场到零售市场、零售市场到餐桌。
- 2、居民能够吸收所消费果蔬的全部营养物质。
- 3、膳食纤维、维生素和矿物质摄入量超出需求量是不合理的。
- 4、相对人体健康状况，膳食纤维、维生素和矿物质的重要性是等同的。
- 5、水果和蔬菜只向人体提供膳食纤维、维生素和矿物质。

3、通用符号说明

序号	符号	符号说明
1	m_{ij}	第 i 种水果第 j 种营养成分含量
2	n_{ij}	第 i 种蔬菜第 j 种营养成分含量
3	M_{ij}	第 i 种水果第 j 种营养成分含量无量纲化数据
4	N_{ij}	第 i 种蔬菜第 j 种营养成分含量无量纲化数据
5	S_i	第 i 种水果或蔬菜的年总产量
6	B_i	第 i 种水果或蔬菜的营养成分无量纲化总量
7	A	水果或蔬菜总产量

8	f_1	田间地头到大市场损耗率
9	f_2	大市场到零售市场损耗率
10	f_3	零售市场到餐桌的损耗率

4、问题一模型的建立与求解

4.1 问题分析

问题一要求选取主要的水果和蔬菜品种，并采用多种方法建立模型从其产量对消费量进行估计，研究其发展趋势。

选取主要水果和蔬菜品种的要求有两个：一是选取的水果、蔬菜品种不仅总计产量应分别超过它们各自总产量的 90%；二是选取的水果（蔬菜）品种含有水果（蔬菜）蕴含的所有营养素并且含量尽可能多。因此，选取过程建立多目标规划模型，以主要的水果品种选取为例，目标有三个：首先是选取的水果总产量最多，其次是水果含有的总计营养成分最多，最后是选取的水果种类最少；约束条件为选取的水果和蔬菜种类总计产量分别超过他们各自产量的 90%。

对选取的主要水果和蔬菜的产量拟采用两种方法对消费量进行估计。第一种是基于损耗率考虑，国内产量减去田间地头到大市场损耗、大市场到零售市场损耗、零售市场到餐桌的损耗得到消费量；第二种是基于进出口量考虑，产量加进口量减去出口量得到消费量。第一种方法的求解过程采用灰色预测模型和曲线拟合模型，对两种预测模型进行检验，确定一种更合理的模型；第二种方法的求解选用第一种方法中的合理预测模型估计蔬菜水果主要品种的消费量。对两种方法预测消费量的过程进行分析，选择预测更理想的方法估计主要果蔬品种消费量，并对其发展趋势进行研究。

4.2 主要果蔬品种筛选的多目标规划模型

4.2.1 模型准备

为了计算方便，对 15 种水果进行编号，结果见表 1。

表 1 15 种水果对应的编号

水果	苹果	梨	柑橘	香蕉	菠萝	荔枝	龙眼	猕猴桃
编号	1	2	3	4	5	6	7	8
水果	桃子	葡萄	大枣	柿子	西瓜	甜瓜	草莓	
编号	9	10	11	12	13	14	15	

对 14 种营养成分进行编号，结果见表 2。

表 2 14 种营养成分对应的编号

营养成分	维 A	B1	B2	B6	C	E	钙
编号	1	2	3	4	5	6	7
营养成分	铁	钾	钠	铜	镁	锌	硒
编号	8	9	10	11	12	13	14

第 i 种水果第 j 种营养成分含量无量纲化公式：

$$M_{ij} = \frac{m_{ij}}{\max(m_{ij})} \quad (i = 1, 2, \dots, 15, j = 1, 2, \dots, 14) \quad (1)$$

15 种水果的 14 种营养成分的无量纲化数据见附录 1。

第 i 种水果的营养成分无量纲化后的总量：

$$B_i = \sum_{j=1}^{14} M_{ij} \quad (i = 1, 2, \dots, 15) \quad (2)$$

统计 15 种水果营养成分无量纲化总量计算结果，见表 3。

表 3 15 种水果 B_i 数据表

B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8
3.33	3.93	5.28	6.16	3.11	2.50	4.29	6.04
B_9	B_{10}	B_{11}	B_{12}	B_{13}	B_{14}	B_{15}	
2.20	2.55	5.85	2.87	2.75	4.90	3.57	

统计 13 种蔬菜营养成分无量纲化总量计算结果，见表 4。

表 4 13 种蔬菜 B_i 数据表

B_1	B_2	B_3	B_4	B_5	B_6	B_7
3.83	3.45	5.85	2.87	2.75	4.90	4.29
B_8	B_9	B_{10}	B_{11}	B_{12}	B_{13}	
3.20	2.12	5.65	3.87	6.78	2.25	

4.2.2 模型建立

第一步：约束条件^[4]的确定

由问题分析，选取主要水果品种的约束条件：选取的水果总计产量超过总产量的 90%。

设 i 为水果编号，对编号的选取引入 0-1^[2]变量 $x_i (i = 1, 2, \dots, 15)$ 。

令

$$x_i = \begin{cases} 1, & \text{当编号为 } i \text{ 的水果被选中,} \\ 0, & \text{当编号为 } i \text{ 的水果没被选中. } \end{cases} \quad i = 1, 2, \dots, 15$$

所以，约束条件为：

$$s.t \begin{cases} x_i = 0 \text{ or } 1 \\ z_1 \geq 0.9A \end{cases} \quad (z_1 : \text{选取水果品种的总计产量}) \quad (5)$$

第二步：建立目标函数

选取主要的水果品种，需要建立三个目标：

$$\begin{cases} z_1: \text{主要水果品种的总计产量最大} \\ z_2: \text{主要水果品种所含营养成分含量最大} \\ z_3: \text{选取的主要水果品种数最少} \end{cases}$$

且这三个目标均要求恰好达到目标值，即即正、负偏差变量都要尽可能地小，这时

$$\begin{cases} \min z_1 = f(d_1^+ + d_1^-) \\ \min z_2 = f(d_2^+ + d_2^-) \\ \min z_3 = f(d_3^+ + d_3^-) \end{cases}$$

由问题分析得，这三个目标的优先权是 $z_1 > z_2 > z_3$ 。因此，给 z_1 赋于优先因子 P_1 ， z_2 赋于优先因子 P_2 ， z_3 赋于优先因子 P_3 ，并规定 $P_1 \gg P_2 \gg P_3$ 。

所以目标函数：

$$\begin{aligned} \min & P_1(d_1^- + d_1^+) + P_2(d_2^- + d_2^+) + P_3(d_3^- + d_3^+) \\ & \begin{cases} \max z_1 = \sum_{i=1}^n S_i x_i + d_1^- - d_1^+ \\ \max z_2 = \sum_{i=1}^n B_i x_i + d_2^- - d_2^+ \\ \min z_3 = \sum_{i=1}^n x_i + d_3^- - d_3^+ \\ d_k^-, d_k^+ \geq 0, \quad k=1,2,3 \end{cases} \quad i=1,2,\dots,15 \end{aligned} \quad (6)$$

S_i 表示第 i 种水果的产量； B_i 表示第 i 种水果的营养成分含量； A 表示水果总产量。

4.2.3 模型求解

1、选取主要水果品种求解

由 Matlab 程序求解得编号为 i 的水果 x_i 的 0-1 取值,见表 5

表 5 x_i 的 0-1 取值表

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	1	1	1	1	0	0	0
x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	
0	1	1	1	1	1	0	

由表 5 知，符合目标条件选取 10 种水果为苹果，梨，柑橘，香蕉，桃，葡萄，红枣，柿子，西瓜，甜瓜。

2、选取主要蔬菜品种求解

将蔬菜相关数据带入，令 $i=1,2,\dots,13$ ， S_i 表示第 i 种蔬菜的产量； B_i 表示第 i 种蔬菜的营养成分含量； A 表示蔬菜总产量。

由 Matlab 程序求解得编号为 i 的蔬菜 x_i 的 0-1 取值，见表 6。

表 6 x_i 的 0-1 取值

x_1	x_2	x_3	x_4	x_5	x_6	x_7
1	1	0	1	1	0	0
x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	
0	1	1	0	1	1	

由表 5 知，符合目标条件选取 8 种蔬菜为萝卜，胡萝卜大蒜，大葱，茄子，黄瓜，大白菜，芹菜。

4.3 消费量估计模型的建立与求解

4.3.1 基于损耗率消费量估计模型

4.3.1.1 基于损耗率计算人均消费量的公式

由假设水果、蔬菜的损耗只发生在从田间地头到大市场、从市场到零售市场、从零售市场到餐桌的过程中，基于损耗率计算人均消费量：

$$x(i)=\frac{S_i \times (1-f_1)(1-f_2)(1-f_3)}{N} \quad (7)$$

其中， S_i ：第 i 种水果或蔬菜的年总产量， f_1 ：田间地头到大市场损耗率， f_2 ：大市场到零售市场损耗率， f_3 ：零售市场到餐桌的损耗率， N ：全国总人口数

4.3.1.2 利用灰色预测^[2]估计消费量

1、模型建立

第一步：数据的检验与处理

首先，为了保证灰色预测建模的可行性，需要对主要水果、蔬菜品种已知 n 年的人均消费量做必要的检验处理。设 n 年的参考数据为数列

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)) \quad (8)$$

计算数列的级比：

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, k = 2, 3, \dots, n \quad (9)$$

计算容覆盖：
$$\Theta = (e^{\frac{2}{n+1}}, e^{\frac{2}{n+2}}) \quad (10)$$

如果所有的级比 $\lambda(k)$ 都落在可容覆盖内，则数列 $x^{(0)}$ 可以作为模型 $GM(1,1)$ 的数据进行灰色预测。否则，需要对数列 $x^{(0)}$ 作必要的变换处理，使其落入可容覆盖内。

第二步：建立灰色预测模型

$GM(1,1)$ ^[2]即表示模型是人均消费量关于时间的 1 阶方程，仅含时间 1 个自变量。

下面对其建立过程做详细阐述：

设 $x^{(0)}$ 为 n 年人均消费量的数列 $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ ， $x^{(0)}$ 的一次累加生成列为 $x^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n))$ ，其中 $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), (k=1, 2, \dots, n)$

则定义 $x^{(1)}$ 的灰导数为： $dx^{(1)}(k) = x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k-1)$

令 $z^{(1)}$ 为数列 $x^{(1)}$ 的紧邻均值数列，即： $z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1), k=2, 3, \dots, n$

则灰微分方程模型为： $dx^{(1)}(k) + az^{(1)}(k) = b$

$$\text{即： } x^{(0)}(k) + az^{(1)}(k) = b$$

其中 $x^{(0)}(k)$ 称为灰导数， a 称为发展系数， $z^{(1)}(k)$ 称为白化背景值， b 称为灰作用量。

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix},$$

令 $Y = ((x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n))^T, u = (a, b)^T$

称 Y 为数据向量， B 为数据矩阵， u 为参数向量，则灰色微分方程模型可表述为矩阵方程 $Y = Bu$ 。

由最小二乘法可以求得

$$\hat{u} = (\hat{a}, \hat{b})^T = (B^T B)^{-1} B^T Y$$

最后可得预测值

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} + \frac{b}{a}, k=0, 1, \dots, n-1, \dots$$

而且 $\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k), k=1, 2, \dots, n-1, \dots$

第三步：检验预测值^[1]

(1) 残差检验：令残差为 $\varepsilon(k)$, 计算

$$\varepsilon(k) = \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)}, k = 1, 2, \dots, n$$

这里 $\hat{x}^{(0)}(1) = x^{(0)}(1)$ ，如果 $\varepsilon(k) < 0.2$ 则可认为达到一般要求；如果 $\varepsilon(k) < 0.1$ ，则认为达到较高要求。

(2) 级比偏差值检验：首先由参考数据 $x^{(0)}(k-1), x^{(0)}(k)$ 计算出级比 $\lambda(k)$ ，再用发展系数 a 求出相应的级比偏差

$$\rho(k) = 1 - \left(\frac{1 - 0.5a}{1 + 0.5a} \right) \lambda(k)$$

如果 $\rho(k) < 0.2$ ，则可认为达到一般要求；如果 $\rho(k) < 0.1$ ，则认为达到较高的要求。

2、模型求解

以水果为例，对建立的灰色预测模型进行求解。

第一步：对 10 个主要水果品种的人均消费量进行级比检验，由公式 (7) 得 10 种水果 2002 年至 2010 年共 9 年的人均消费量，见附录

将 $n=9$ 带入公式 (8) 和公式 (9)，由 Matlab 程序计算得 10 种水果的级比，见表 7。

表 7 10 种水果的级比数值表

苹果	0.907	0.887	0.981	0.917	0.930	0.928	0.937	0.947
梨	0.946	0.916	0.935	0.940	0.925	0.947	0.944	0.942
柑橘	0.887	0.895	0.935	0.885	0.865	0.878	0.919	0.947
香蕉	0.937	0.970	0.924	0.940	0.880	0.989	0.882	0.918
桃	0.847	0.873	0.915	0.923	0.903	0.944	0.944	0.954
葡萄	0.911	0.850	0.804	0.811	1.002	0.829	0.851	0.945
红枣	0.965	0.894	0.910	0.937	0.897	0.944	0.951	0.950
柿子	1.073	1.004	0.955	0.963	0.992	0.982	0.964	0.944
西瓜	1.044	1.017	0.909	0.924	0.914	0.861	0.976	0.985
甜瓜	0.837	0.909	0.945	1.039	0.996	0.930	0.901	0.941

由表 7 可知，10 种水果的最大级比是 1.073，最小级比是 0.837。将 $n=9$ 带入公式 (10)，得 15 种水果的可容覆盖 $\Theta = (0.819, 1.221)$ 。10 种水果的级比都落在可容覆盖内，所以 10 种水果 2002 年—2010 年的数据 $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(9))$ 均可以作为模型 $GM(1,1)$ 的数据进行灰色预测。

第二步：GM(1,1) 预测值求解

以苹果的人均消费量数据为例详细分析求解过程

(1) 对苹果原始数据 $x^{(0)}$ 作一次累加

$$x^{(1)} = (14.349, 30.161, 47.989, 66.161, 85.986, 107.293, 130.254, 154.77, 180.665)$$

(2) 构造数据矩阵 B 及数据向量 Y

$$B = \begin{bmatrix} -z^{(1)}(x^{(1)}(1) + x^{(1)}(2)) & 1 \\ -z^{(1)}(x^{(1)}(2) + x^{(1)}(3)) & 1 \\ \vdots & \vdots \\ -z^{(1)}(x^{(1)}(8) + x^{(1)}(9)) & 1 \end{bmatrix}, Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(9) \end{bmatrix}$$

(3) 计算 \hat{u}

$$\hat{u} = (\hat{a}, \hat{b})^T = (B^T B)^{-1} B^T Y = \begin{bmatrix} 0.069 \\ -13.939 \end{bmatrix}, \text{ 于是得到 } a = 0.069, b = -13.939$$

(4) 求解模型方程为 $\frac{dx^{(1)}}{dt} + 0.06x^{(1)} = -13.939$

$$\text{解得 } \hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} + \frac{b}{a} = 216.361e^{-0.069k} - 202.012, \quad k = 0, 1, \dots, 8$$

(5) 求取数列预测值 $\hat{x}^{(1)}(k+1)$ 及模型还原值 $\hat{x}^{(0)}(k+1)$

令 $k = 1, 2, \dots, 8$ ，由上面的时间响应函数可算得 $\hat{x}^{(0)}(k+1) = x^{(1)}(k+1) - x^{(1)}(k)$

模型还原数列：

$$\begin{aligned} \hat{x}^{(0)} &= (\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(9)) \\ &= (13.642, 15.342, 16.433, 17.602, 18.855, 20.196, 21.633, 23.172, 24.820) \end{aligned}$$

3、模型检验

由 Matlab 程序运行得到苹果消费量灰色预测模型的各项指标检验值，结果见表 8。

表 8 苹果消费量灰色预测模型检验值表

年份	苹果原始值	预测值	相对误差	残差
2002	14.349	14.349	0.00%	0
2003	15.812	15.742	-0.44%	-0.07
2004	17.828	17.733	-0.53%	-0.095
2005	18.172	18.182	0.06%	0.01
2006	19.825	19.855	0.15%	0.03

2007	21.307	21.296	-0.05%	-0.011
2008	22.961	22.933	-0.12%	-0.028
2009	24.516	24.472	-0.18%	-0.044
2010	25.895	25.82	-0.29%	-0.075

由 Matlab 程序作图得

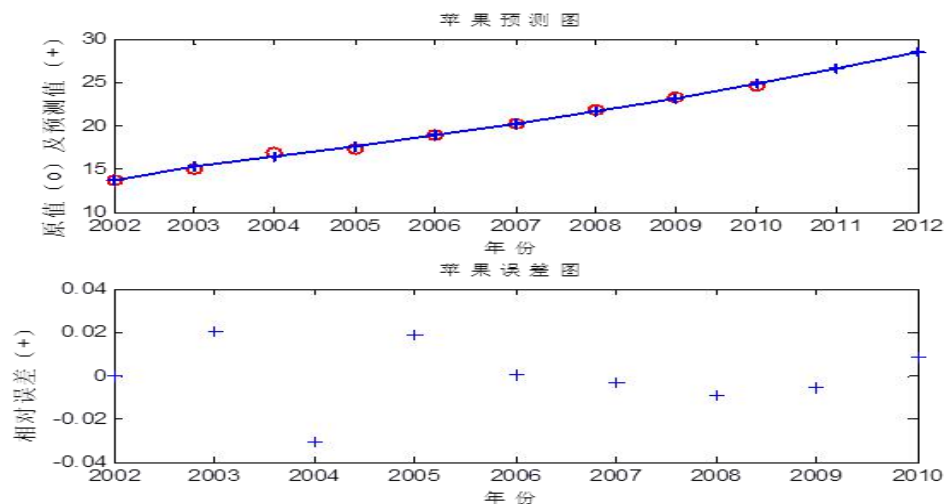


图 1 苹果消费量预测图和误差图

由图 1 预测图及表 8 中残差检验可知，该模型的残差 $|\varepsilon_1(k)|_{\max} = 0.095 < 0.1$ ，级比偏差 $\rho(k) < 0.1$ ，精度达到较高要求，可以对苹果人均消费量进行合理预测。

由 $\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a}\right)e^{-ak} + \frac{b}{a} = 216.361e^{-0.069k} - 202.012$, $k = 9, 10, \dots, 8$ 预 测

2011—2020 年的苹果人均消费量，见表 9。

表 9 2011 至 2020 年的苹果消费量估计值（单位/千克）

2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
26.586	28.477	30.503	32.673	34.998	37.488	40.154	43.011	46.071	49.349

4、研究消费量发展趋势

由表 9 中基于损耗率的 2011 至 2020 年的苹果消费量估计值，用 Matlab 软件作图得

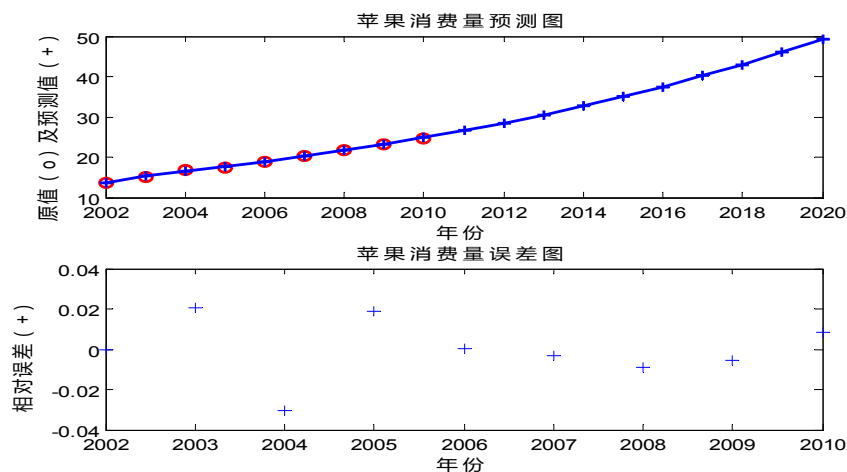


图 2 2010 至 2020 年苹果消费量估计图

由图 2，可以看出苹果消费量与时间呈正相关关系，从 2010 到 2020 年增长速度整体较快，说中国居民对苹果的需求不断增加。其中从 2010 到 2016 年曲线斜率逐渐增加，从 2016 到 2020 年曲线斜率基本保持不变，中国居民对苹果的需求持续递增。

4.3.1.3 利用曲线拟合^[2]估计消费量

由 Matlab 程序对苹果 2002 至 2010 年的人均消费量进行曲线拟合，结果见图 2。

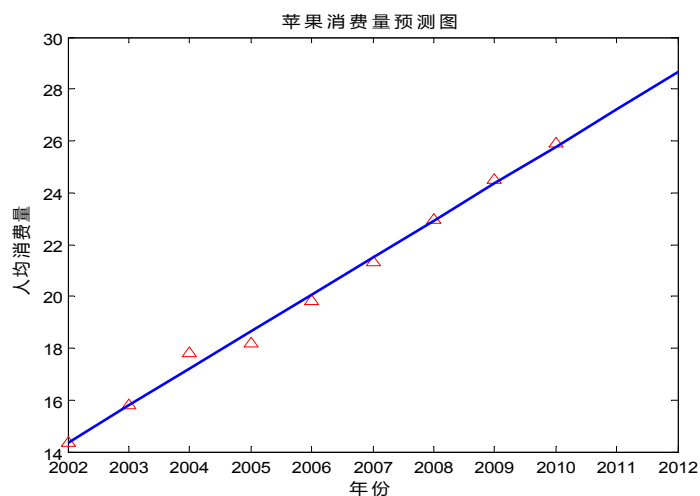


图 3 苹果消费量拟合图

由图 3 可以看出还原数据与实际数据有较大差距，实际数据大部分在拟合曲线附近波动，有较少的点在曲线上，拟合准确性较差。

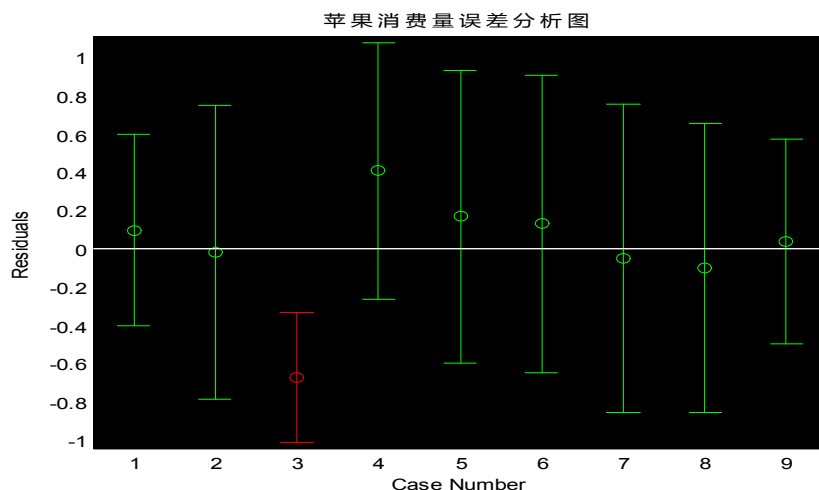


图 4 苹果误差分析图

由图 4 可知，经曲线拟合，所有数据均含有误差，其中 9 个数据中 4 个数据的误差较大，1 个数据的误差极大属于异常点。所以，用曲线拟合估计消费量存在较大误差。

用回归分析对曲线拟合结果进行检验，得到 $\hat{\beta}_0 = -2.1813$ ， $\hat{\beta}_1 = 140.6194$ ， $\hat{\beta}_0$ 的置信区间是 $[11.1276, 6.7649]$ ， $\hat{\beta}_1$ 的置信区间是 $[0.3230, 2.1369]$ ； $R^2 = 0.9944$ ， $F = 534.0707$ ， $p = 0.1139$ 。

对检验数据进行分析： $R^2 = 0.9944$ 较大，但是 $p = 0.1139$ ，取 $\alpha = 0.05$ 时模型不能通过检验。说明曲线拟合针对未来消费量的预测效果较差，此方法不适用估计水果和蔬菜的消费量。

4.3.2 基于进出口量的消费量估计模型

4.3.2.1 基于进出口量计算人均消费量的公式

由问题分析，水果或蔬菜的国内产量减去出口量再加进口量即为中国居民消费总量。基于进出口量计算人均消费量：

$$x(i) = \frac{S_i - P_i + Q_i}{N} \quad (11)$$

其中， P_i ：第 i 种水果或蔬菜的年出口量， Q_i ：第 i 种水果或蔬菜的年进口量

4.3.2.1 利用灰色预测估计消费量

1、模型求解

以苹果为例分析求解过程。

由公式（8）得到苹果的人均消费量数据，见表 10

表 10 2002 年至 2010 年苹果消费量（单位/千克）

2002	2003	2004	2005	2006	2007	2008	2009	2010
14.035	15.312	17.174	17.540	18.983	20.649	22.011	23.553	24.977

由以损耗率为标准的灰色预测求解过程

(1) 计算 \hat{u}

$$\hat{u} = (\hat{a}, \hat{b})^T = (B^T B)^{-1} B^T Y = \begin{bmatrix} 0.068 \\ -14.020 \end{bmatrix}, \text{ 于是得到 } a = 0.068, b = -14.020$$

(2) 求解模型方程为 $\frac{dx^{(1)}}{dt} + 0.06x^{(1)} = -14.020$

解得

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} + \frac{b}{a} = 220.208e^{-0.068k} - 206.173, \quad k = 0, 1, \dots, 8$$

(3) 求取数列预测值 $\hat{x}^{(1)}(k+1)$ 及模型还原值 $\hat{x}^{(0)}(k+1)$

令 $k = 1, 2, \dots, 8$ ，由上面的时间响应函数可算得 $\hat{x}^{(0)}(k+1) = x^{(1)}(k+1) - x^{(1)}(k)$

模型还原数列：

$$\begin{aligned} \hat{x}^{(0)} &= (\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(9)) \\ &= (14.035, 15.568, 16.669, 17.847, 19.109, 20.460, 21.907, 23.455, 25.114) \end{aligned}$$

2、模型检验

由 Matlab 软件运行得到苹果灰色预测模型的各项指标检验值，结果见表 11。

表 11 苹果消费量灰色预测模型检验值表

年份	苹果原始值	预测值	相对误差	残差
2002	14.035	14.035	0.00%	0
2003	15.312	15.368	0.37%	0.056
2004	17.174	17.17	-0.02%	-0.004
2005	17.54	17.547	0.04%	0.007
2006	18.983	19.009	0.14%	0.026
2007	20.649	20.646	-0.01%	-0.003
2008	22.011	22.007	-0.02%	-0.004
2009	23.553	23.455	-0.42%	-0.098
2010	24.977	25.114	0.55%	0.137

由 Matlab 程序作图得

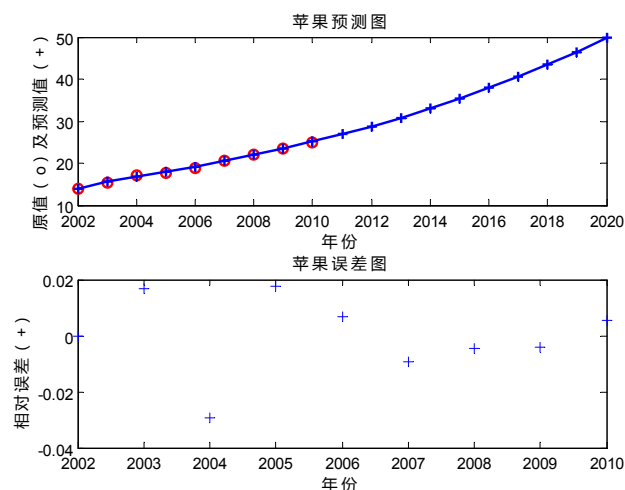


图 5 苹果消费量预测图和误差图

由图 5 预测图及表 11 中残差检验可知，该模型的残差 $|\varepsilon_2(k)|_{\max} = 0.098 < 0.1$ ，级比偏差 $\rho(k) < 0.1$ ，精度达到较高要求，可以对苹果人均消费量进行合理预测。

$$\text{由 } \hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-ak} + \frac{b}{a} = 220.208e^{-0.068k} - 206.173, \quad k = 0, 1, \dots, 8 \text{ 预测 2011}$$

至 2020 年的苹果消费量，见表 12。

表 12 2011 至 2020 年的苹果消费量预测值（单位/千克）

2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
26.889	28.790	30.826	33.005	35.338	37.837	40.512	43.376	46.442	49.726

3、研究消费量发展趋势

由表 12 中基于进出口量的 2011 至 2020 年的苹果消费量估计值，用 Matlab 软件作图得

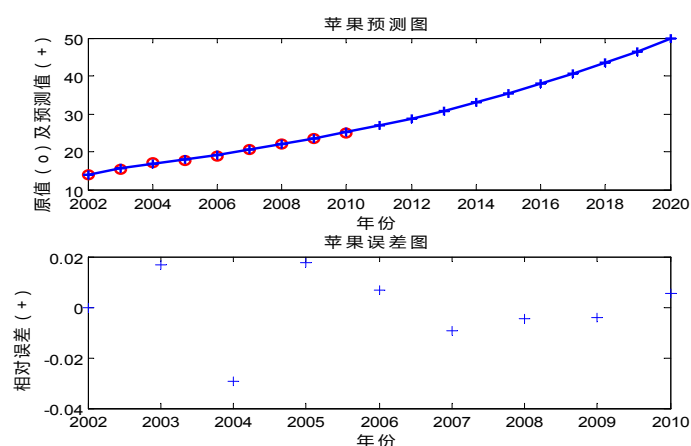


图 6 基于进出口量 2010 至 2020 年苹果消费量估计图

由图 6，可以看出苹果消费量与时间呈正相关关系，从 2010 到 2020 年增长速度整体较快，说中国居民对苹果的需求增加。其中从 2010 到 2016 年曲线斜率逐渐增加，从 2016

到 2020 年曲线斜率基本保持不变，中国居民对苹果的需求持续递增。

5、问题二模型的建立与求解

5.1 问题分析

本题要求评价中国居民目前矿物质、维生素、膳食纤维的年摄入水平是否合理。根据问题一筛选出的 10 种水果和 8 种蔬菜各自含有的 10 种营养成分数据，拟采用 Spss 软件对 10 种营养成分含量进行聚类，每一类筛选出一种作为代表重点分析。由问题一中预测模型求出 2014—2020 年 18 种物质的人均消费量，对 18 种物质各类营养成分分别求和得人均年摄入量。将其与保障人体健康所需要的各类营养成分期望值比较，建立隶属度分析模型判断各营养成分的人均年摄入水平是否合理，对各类营养成分设立合理临界值，当隶属度大于该临界值时即认为摄入量是合理的。

2014—2020 年 18 种物质各类营养成分总摄入量与标准值的隶属度用 Excel 软件作图，可分析某类营养成分的走势。因为水果和蔬菜提供的膳食纤维、矿物质、维生素都是人体不可缺少的营养物质，可以把各类营养成分隶属度求均值作为营养总摄入量的隶属度。把营养总摄入量的隶属度对时间作图，根据曲线走势可判断居民营养健康状况是否趋于好转。

5.2 营养成分聚类分析模型建立与求解

聚类分析方法是基于数值分类法的思想建立起来的，又称系统聚类法，本文基于 10 种营养成分进行 Q 型聚类。

1、数据处理

由于 10 种营养成分除膳食纤维（单位为克）其余量纲均为毫克，18 种水果和蔬菜 $A_i^{(1)} (i=1,2,\dots,18)$ 的各营养成分含量数据矩阵 $x_{ij}^{(1)} (j=1,2,\dots,10)$ 可直接使用。

2、计算样本的相似度

确定营养成分样本的相似度量时，采用相关系数来衡量两个营养成分样本的相似性。记样本 $\hat{x}_j^{(1)}$ 的取值 $(x_{1j}, x_{2j}, \dots, x_{18j})^T \in R^n (j=1,2,\dots,10)$ ，则两个样本的相关系数

$$r_{jk}^{(1)} = \frac{\sum_{i=1}^{12} (x_{ij}^{(1)} - \bar{x}_j^{(1)})(x_{ik}^{(1)} - \bar{x}_k^{(1)})}{\left[\sum_{i=1}^{12} (x_{ij}^{(1)} - \bar{x}_j^{(1)})^2 \sum_{i=1}^{12} (x_{ik}^{(1)} - \bar{x}_k^{(1)})^2 \right]^{1/2}} \quad (12)$$

相似性度量中 $|r_{jk}^{(1)}|$ 越接近于 1， $\hat{x}_j^{(1)}$ 与 $\hat{x}_k^{(1)}$ 越相关； $|r_{jk}^{(1)}|$ 越接近 0， $\hat{x}_j^{(1)}$ 与 $\hat{x}_k^{(1)}$ 的相关性越弱。

3、由 Spss 软件运行得 10 种营养成分聚类结果

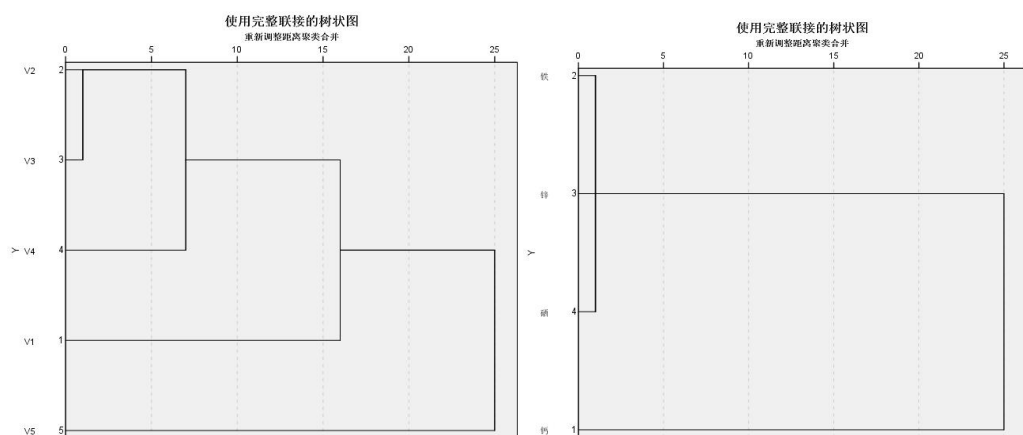


图 7 聚类分析结果图

由图 7 对聚类结果分析：

10 种营养成分划分成 6 类的结果见表 12

表 12 营养成分分类表

第 1 类	膳食纤维
第 2 类	维 生 素 A
第 3 类	维生素 C、维生素 B1、维生素 B2
第 4 类	维生素 E
第 5 类	钙
第 6 类	铁 锌 硒

5.3 营养摄入量模糊隶属度模型建立与求解

5.3.1 模型准备

由问题一中模型预测得到 2014 至 2020 年 6 类营养成分人均摄入量数据，见表 13。

表 13 2014 至 2020 年 6 类营养成分人均摄入量

年份	2014	2015	2016	2017	2018	2019	2020
膳食纤维	6023.87	6312.33	6618.83	6944.72	7291.44	7660.57	8053.82
维生素 A	260.91	277.13	294.70	313.75	334.43	356.90	381.34
维生素 C	11114.63	11652.76	12231.32	12854.09	13525.28	14249.52	15031.95
维生素 E	5916.90	6223.84	6552.09	6903.40	7279.69	7683.09	8115.87
钙	129814.0	135266.0	141097.8	147343.80	154042.3	161235.4	168969.4
锌	1424.01	1488.52	1557.09	1630.04	1707.72	1790.48	1878.74

查询资料得保障健康的人均营养摄入量标准值数据，见表 14。

表 14 保障健康的人均营养摄入量标准值数据表

营养成分	膳食纤维	维生素 A	维生素 C	维生素 E	钙	锌
健康标准值	6570.00	273.25	19163.00	8468.00	164250.00	1971.00

5.3.2 模型建立

模糊数学的基本思想是隶属度的思想。应用模糊数学方法建立数学模型的关键是建立符合实际的隶属函数。

定义论域 X_m ($m=1,2,\dots,6$) (18 种物质第 m 类营养成分 2014 至 2020 年含量数列) 中的任意元素 x_{mn} ($n=1,2,\dots,7$) 到闭区间 $[0,1]$ 上的任意映射都确定论域 X_m 上的一个模糊集合 A_m , μ_{mA} 叫做 A_m 的隶属函数, 即:

$$\begin{cases} \mu_{mA} : X_m \rightarrow [0,1] \\ x_{mn} \rightarrow \mu_{mA}(x_{mn}) \end{cases} \quad (13)$$

$\mu_{mA}(x_{mn})$ 叫做 x_{mn} 对模糊集合 A_m 的隶属度, 记为: $A_m = \{(x_{mn}, \mu_{mA}(x_{mn})) | x_{mn} \in X_m\}$, 使 $\mu_{mA}(x_{mn}) = 0.5$ 的点 x_0 成为模糊集 A_m 的过渡点, 此点最具有模糊性。所以 18 种物质第 m 类营养物质总含量的模糊集合 A_m 完全由每种物质的隶属函数 μ_{mA} 来刻画, 当 $\mu_{mA}(x_{mn}) = \{0,1\}$ 时, A_m 退化为一个普通集。

确定隶属函数是模糊集合的关键, 本文采用可靠性领域广泛应用的待定系数法, 根据论域、边界、分布形式和过渡形式等确定隶属函数, 根据实际情况选择一种可以表现评价对象模糊集特性的隶属函数作为分布, 并根据人体营养所需膳食宝塔等确定分布参数。当取较低的隶属度时, 采用正态分布能够屏蔽更多的隶属度低的信息, 获取更多隶属度高、有价值的评价信息, 提高结果的可信任度。最后确定隶属度函数为:

$$\mu_{mA}(x_m) = \left\{ \exp \left(\frac{x_{mn} - a_m}{\sigma_m} \right)^2 \right\} \quad (14)$$

$$\text{方差 } \sigma_m = \sqrt{\sum_{n=1}^7 \frac{(x_{mn} - \bar{x}_m)^2}{7}} \quad (15)$$

其中, μ_{mA} 是参量 x_{mn} 的隶属度

a_m 为第 m 种营养成分含量分布期望值, 即人体所需营养成分期望值。

设 $A_m \in F(U)$ ($m=1,2,\dots,6$), 对 $u_0 \in U$, 若存在 m_0 , 使

$$A_{m_0}(u_0) = \max \{A_1(u_0), A_2(u_0), \dots, A_6(u_0)\}$$

则认为 u_0 相对地隶属于 A_m 。

5.3.3 模型求解

1、隶属度函数求解

由表 13 中的 6 类营养成分 2014 至 2020 年的人均摄入量计算得到其各自的隶属度函数

$$\text{膳食纤维: } \mu_A(x_i) = \left\{ \exp\left(\frac{x_i - 6570}{3125}\right)^2 \right\}$$

$$\text{维生素A: } \mu_A(x_i) = \left\{ \exp\left(\frac{x_i - 273.75}{323}\right)^2 \right\}$$

$$\text{维生素C: } \mu_A(x_i) = \left\{ \exp\left(\frac{x_i - 19163}{11935}\right)^2 \right\}$$

$$\text{维生素E: } \mu_A(x_i) = \left\{ \exp\left(\frac{x_i - 8468}{4934}\right)^2 \right\}$$

$$\text{钙: } \mu_A(x_i) = \left\{ \exp\left(\frac{x_i - 164250}{81214}\right)^2 \right\}$$

$$\text{锌: } \mu_A(x_i) = \left\{ \exp\left(\frac{x_i - 1971}{984}\right)^2 \right\}$$

2、隶属度计算

由 Matlab 软件运行得出 2014 至 2020 年 6 类营养物质的隶属度，见表 15。

表 15 6 类营养物质的隶属度 μ_{m_A} 值表

年份	2014	2015	2016	2017	2018	2019	2020
膳食纤维	0.943	0.979	0.998	0.996	0.968	0.912	0.828
维生素 A	0.998	1.000	0.995	0.982	0.960	0.926	0.879
维生素 C	0.638	0.676	0.716	0.759	0.802	0.846	0.888
维生素 E	0.771	0.818	0.863	0.907	0.945	0.976	0.995
钙	0.831	0.877	0.920	0.956	0.984	0.999	0.997
锌	0.741	0.792	0.843	0.890	0.933	0.968	0.992

由上表中的隶属度趋近于 1 的程度可判断 6 类营养物质摄入量趋近于人体健康标准值的程度， μ_{m_A} 越趋近于 1，表明该类营养物质人均摄入量越合理。

以膳食纤维为例，由隶属度曲线方程画出其分布图，见图 8，其他五种营养物质的隶属度分布图见附录 2。

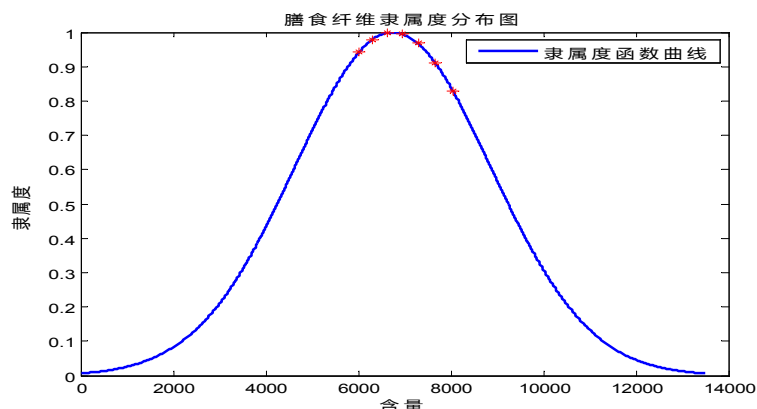


图 8 膳食纤维隶属度曲线方程拟合图

3、人均摄入量合理化分析

因为 6 类营养成分人体所需含量的不同，合理临界值也不同，根据正态分布^[3]，计算得每类营养成分人均摄入量隶属度的合理临界值 p_m 。

表 16 合理临界值 p_m 表

营养成分	膳食纤维	维生素 A	维生素 C	维生素 E	钙	锌
合理临界值	0.833	0.933	0.667	0.827	0.933	0.889

以 2014 年中国居民 6 类营养成分摄入量为例分析其年摄入水平是否合理。

由 6 类营养物质的 μ_{m_i} 值表可知，因为 $0.943 > 0.833$ ，所以 2014 年中国居民的膳食纤维摄入量是合理的； $0.998 > 0.933$ ，维生素 A 的摄入量也是合理的； $0.638 < 0.667$ ，维生素 C 类摄入量不合理且偏低； $0.771 < 0.827$ ，维生素 E 类摄入量不合理且偏低； $0.832 < 0.933$ ，钙类摄入量不合理且偏低； $0.741 < 0.889$ ，锌类摄入量不合理且偏低。

可以看出中国居民营养摄入量仅膳食纤维和维生素 A 是合理的，其他大量来源于水果和蔬菜的营养成分均偏低，这与目前中国居民喜食肉类，尤其是猪肉，食蔬菜水果较少息息相关。

4、2014 至 2020 年中国居民营养健康状况趋势分析

设 2014 至 2020 年对应的编号分别为 1 到 7，以膳食纤维为例，对膳食纤维人均年摄入量 7 年的隶属度描点连线得图 9，维生素和矿物质隶属度连线图见附录 3。

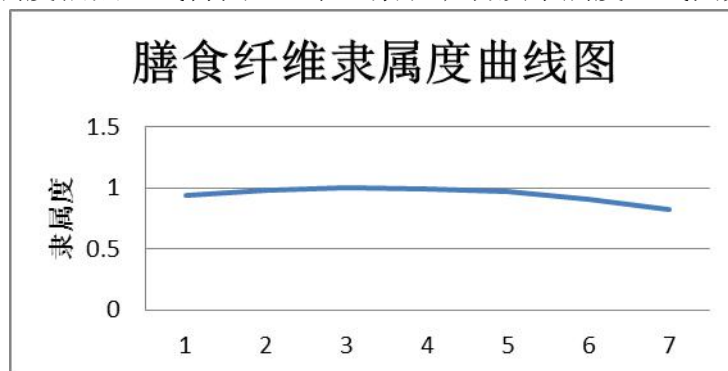


图 9 膳食纤维隶属度曲线图

发现中国居民的膳食纤维摄入量对合理值的隶属度先上升在降低，在 2016 年、2017 年达到预期理想值，即膳食纤维摄入量最佳。

因为膳食纤维、矿物质和维生素是人体不可或缺的重要成分，所以计算人体营养摄入总量隶属度时分别对膳食纤维、维生素A、维生素C、维生素E、钙、锌的隶属度赋予相同的权值。即

$$\text{人体营养年均摄入总量隶属度} A = \sum_{i=1}^6 \frac{\mu_{m,A}}{6} (m=1,2,\dots,6)$$

表 17 2014 至 2020 年隶属度 A 值表

年份	2014	2015	2016	2017	2018	2019	2020
隶属度	0.820	0.857	0.889	0.915	0.932	0.938	0.930

用 Excel 软件作图得

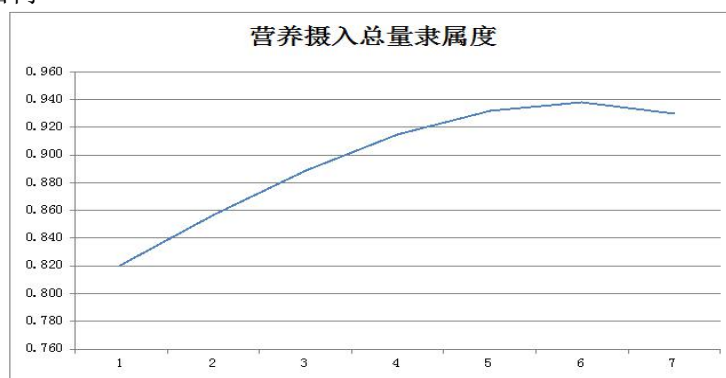


图 10 营养摄入总量隶属度

可以看出，营养摄入总量隶属度在 2014 到 2019 年呈上升趋势，从 0.820 增加到 0.938，说明这段时间内中国居民的人体营养健康状况是趋于好转且较接近于合理的标准值。2019 年到 2020 年，营养摄入总量隶属度下降，从 0.938 减少至 0.930，说明从 2019 年开始，中国居民的人体营养状况由好转趋于恶化。

6. 问题三模型建立与求解

6.1 问题分析

不同的蔬菜、水果尽管各种营养素含量各不相同，但营养素的种类大致相近，存在着食用功能的相似性。所以，水果与水果之间、蔬菜与蔬菜之间、水果与蔬菜之间从营养学角度在一定程度上可以相互替代、相互补充。本文拟采用聚类分析法根据各种蔬菜水果在营养成分含量的相似性进行分类，以达到功能上相似的水果蔬菜互相替代的目的。考虑到不同地区在不同季节的价格和供求量都不尽相同，我们对不同地区在四个季度上分别进行优化。按照蔬菜水果的价格和产地以及供求关系的不同拟将我国分为三个主要地区：北方地区、南方地区和偏远地区。通过查阅中国统计年鉴等资料得到各种蔬菜水果在四个季度的价格，根据价格和人居需求量可以得到蔬菜水果的购买成本。

果蔬类品种很多，营养元素也有很多种，不宜对所有果蔬品种和所有的营养成分进行逐一分析，拟用问题二的十八种果蔬类品种和六种营养元素进行分析。在此基础上，拟采用目标规划数学模型。模型中，以购买成本最低为目标函数、以满足人体营养成分需要量为约束条件。在确定人体健康所需的营养成分标准时，我们根据不同年龄段居民对营养成分的不同需求按照人口比例得到人均合理营养成分的范围。在满足各个营养元素满足机体需要的基础上建立目标函数，使居民在果蔬类上的投入费用最少，给出各个蔬菜品种的消费量的最优解。

6.2 模型准备

1、营养成分单位的转化，将每种果蔬产品的每 100g 所含六种营养成分换算成每公斤所含营养成分的量，求解表达式如下，

$$c_{ij} = \frac{c'_{ij}}{100} \times 1000,$$

其中 c'_{ij} 表示 100 克第 j 种蔬菜或者水果所含有的第 i 种营养成分的含量； c_{ij} 表示每千克第 j 种蔬菜或者水果所含有的第 i 种营养成分的含量。

2、六种营养元素的标准摄入范围的计算。通过查询中国统计年鉴等数据库，我们得到不同年龄段的居民人口比例，再根据附件 4 中我国不同年龄段居民膳食营养素参考日摄入量计算得到我国人均六种营养物质的标准摄入量范围。计算公式如下，

$$C_k = p_1 C_1 + p_2 C_2 + p_3 C_3$$

其中， p_1, p_2, p_3 分别表示青少年、中壮年和老年人的人口比例； C_1, C_2, C_3 分别表示青少年、中壮年和老年人对营养成分的标准摄入量范围； C_k 表示第 k 种营养物质人均需求量范围。

3、统计不同种果蔬产品不同季度的价格（见附录 4）。

6.3 水果蔬菜种类的聚类分析模型

6.3.1 模型建立

不同的蔬菜、水果尽管各种营养素含量各不相同，但营养素的种类大致相近，存在着食用功能的相似性。所以，水果与水果之间、蔬菜与蔬菜之间、水果与蔬菜之间从营养学角度在一定程度上可以相互替代、相互补充。本文通过聚类分析的方法将十八种果蔬类产品六种营养元素进行汇总，建立一个 18×6 的矩阵：

$$\begin{pmatrix} C_{1,1} & \dots & C_{1,6} \\ \dots & \dots & \dots \\ C_{18,1} & \dots & C_{18,6} \end{pmatrix}$$

根据每一行的数据的相关性，采用最短距离法进行聚类，即将不同果蔬品种根据其所含的六种营养元素的含量进行聚类，找出可替代的果蔬品种，达到相互替代相互补充的目的。为了均衡居民摄入水果和蔬菜的量，分别对蔬菜类产品和水果类产品进行聚类，即分别对数列的前八行和后十行进行聚类：

$$f_1, f_2, \dots, f_n \quad n < 18$$

聚类分析方法是基于数值分类法的思想建立起来的，又称系统聚类法。采用聚类分析法对 18 种蔬菜水果根据营养成分进行聚类，称为 R 型聚类。聚类分析作为一种定量方法，将从数据分析的角度，给出一个更准确、细致的分类工具。在于可以指出由粗到细的多种分类情况，典型的系统聚类结果可由一个聚类图展示出来。系统归类过程与计

算类和类之间的距离有关，采用不同的距离定义，有可能得出不同的聚类结果。利用最短距离法对 18 中蔬菜水果做聚类分析,从而将果蔬进行分类。

步骤如下： $\Omega = \{w_1, w_2, \dots, w_n\}$

(1)计算 n 个本点两两之间的距离 $\{I_{ij}\}$ 记为矩阵，记为矩阵 $L = (I_{ij})_{n \times n}$ ；

(2)首先构造 n 个类，每一个类中只包含一个样本点，每一类平台高度均为零；

(3)合并距离最近的两类为新类，并且以这两类间的距离值为聚类图中的平台高度；

(4)计算新类与当前各类的距离，若类个数已等于 1，转步骤否则，转步骤 3；

(5)画聚类图；

(6)决定类的个数和类；

首先，对蔬菜水果进行聚类分析可得，10 种水果和 8 种蔬菜分别对应 $\Omega = \{w_1, w_2, \dots, w_{10}\}$, $\Omega = \{w_1, w_2, \dots, w_8\}$, 蔬菜、水果的营养物质含量为 $(G_b^{(i)}, I_{b2}^{(i)})$ 和 $(G_h^{(i)}, I_{h2}^{(i)})$ ，使用绝对值距离来测量点与点之间的距离，使用最短距离法来测量蔬菜水果之间的距离，即：

$$l(w_i, w_j) = \sum_{k=1}^2 |G_b^{(ik)} - I_{b2}^{(ik)}|$$

$$L(G_p, G_q) = \min \{l(w_i, w_j)\}$$

用距离的长短将 18 种蔬菜水果进行分类。

根据季节不同各种果蔬产品的价格不同，分别从 f_i 中找出价格相对较低的果蔬品种，作为居民对该类产品营养摄入的主要来源，建立一个行的矩阵 $x_i (i=1,2,3,\dots,n)$ 代表在满足人体所需的营养成分时第 f_i 类果蔬产品的居民人均年消耗量建立 $n \times 6$ 的矩阵 $c_{n,j} (j=1,2,3,4,5,6)$ 代表每一类果蔬品种的膳食纤维、维生素 A、维生素 C、维生素 E、钙、锌六种营养元素的含量。建立一维矩阵 Q_i ，表示每类果蔬产品的价格。

6.3.2 模型求解

样本间的相似性采用欧氏距离度量，水果蔬菜类间距离的计算采用类平均法，我们根据其营养成分含量将 18 种水果蔬菜分别进行聚类，共分为 7 类，聚类分析树状图见图 11，以及聚集成员分布表，见表 18。

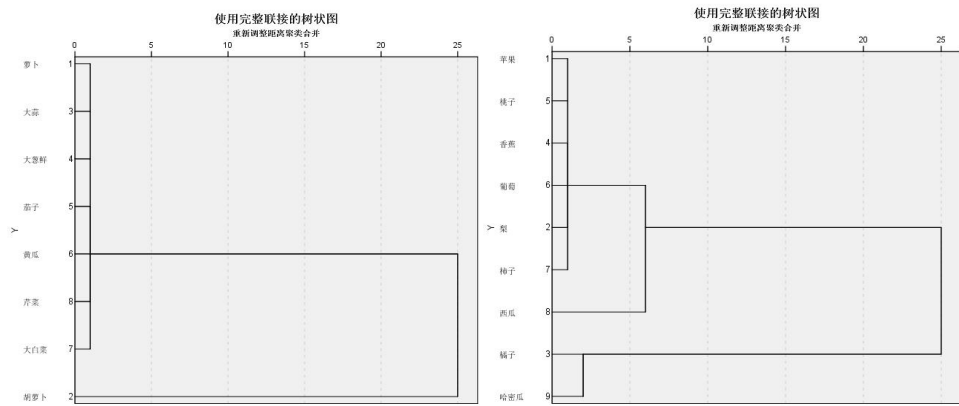


图 11 聚类分析树状图

表 18 群集成员分布表

群集成员					
苹果	1	1	案例	4 群集	3 群集
梨	1	1	萝卜	1	1
橘子	2	2	胡萝卜	2	2
香蕉	1	1	大蒜	1	1
桃子	1	1	大葱	1	1
葡萄	1	1	茄子	1	1
柿子	1	1	黄瓜	1	1
西瓜	3	3	大白菜	3	3
甜瓜	4	2	芹菜	4	1

最后得到水果蔬菜分类汇总：

表 19 水果蔬菜聚类表

水果聚 类	苹果、梨、香蕉、桃子、葡萄、柿子	橘子	西瓜	哈密瓜
蔬菜聚 类	萝卜、大蒜、大葱、茄子、黄瓜、芹菜	胡萝卜	大白菜	

经过市场调研发现，不同季节的同一类果蔬中价格也不尽相同，汇总得到不同季节需要考虑的水果品种。根据同类水果蔬菜中的价格不同，选用价格相对较低的水果蔬菜来代表整类物品。最终筛选出来的蔬菜水果有萝卜、胡萝卜、大白菜、柑橘、香蕉、西瓜和甜瓜七种。

6.4 人均消费量的目标优化模型

6.4.1 模型建立

问题三要求给出今中国居民提供主要的水果和蔬菜产品的按年度合理人均消费量，

使人们能够以较低的购买成本满足自身的营养健康需要。在该模型中，各种蔬菜水果的营养成分大致固定，显然，这是一个优化问题。

一般线性规划问题的标准型为

$$\min \quad z = \sum_{j=1}^n c_j x_j \quad (16)$$

$$\text{s.t.} \quad \sum_{j=1}^n a_{ij} x_j \leq b_i \quad i = 1, 2, \dots, m \quad (17)$$

满足约束条件（17）的解 $x = (x_1, x_2, x_3, \dots, x_n)$ ，称之为现行规划问题的可行解，而使目标函数（16）达到最小值的可行解叫最优解。所有可行解构成的集合称为问题的可行域，记为 R 。目标规划就是在可行域内在满足约束条件的前提下求得目标函数的最优解。为使居民营养元素摄入量达到最佳的状态，即要求每种营养元素都要达到标准范围内。建立以购买成本最低为目标函数，以满足人体自身的营养健康需要为约束条件的目标优化模型。考虑到不同地区在不同季节的价格和蔬菜水果的供应有所不同，本文按照蔬菜水果的分布情况以及人民生活消费水平，将中国分为三大地区，北方地区、南方地区和偏远地区。本文在目标规划模型的基础上，对不同地区按照不同季度对整年的合理人均消费量求解最优解。

1、目标函数的确定

由题意可知，目标函数是当今中国居民能够以最低的成本购买蔬菜水果。查阅中国统计年鉴得到不同地区每个季度各种水果蔬菜的单价，居民在蔬菜水果上的购买成本就是每种水果蔬菜的单价与相应购买量的乘积，对所有的蔬菜水果进行累加求和。所以目标函数可表示为：

$$\min z = \sum_{i=1}^n a_i x_i$$

其中 a_i 表示第 i 种蔬菜或者水果的价格； x_i 表示第 i 种蔬菜或者水果的需求量

2、约束条件的确定

达到购买成本的目标的前提是满足自身的营养需要，也就是说要将中国居民从蔬菜水果中人均摄取营养量在标准人体需求量的范围之内。根据第二问中的聚类分析，我们可以得出，人体营养健康的需要可以由膳食纤维、维生素 A 、维生素 C 、维生素 E 和钙锌六中主要营养成分来表示。所以在分析约束条件时，只需要将这六种营养成分达到标准即可。问题二中，已经根据不同年龄段对六中营养成分按照人口比例得到人均标准。所以约束条件可以表述为：

$$s.t. \begin{cases} 1368.75 < \sum_i^n c_{i,1}x_i < 1916.25, \\ 63.875 < \sum_i^n c_{i,2}x_i < 73.225, \\ 3193.75 < \sum_i^n c_{i,3}x_i < 6375, \\ 1752 < \sum_i^n c_{i,4}x_i < 2482, \\ 38325 < \sum_i^n c_{i,5}x_i < 43800, \\ 438 < \sum_i^n c_{i,6}x_i < 547.5, \end{cases}$$

其中 $c_{i,1}, c_{i,2}, c_{i,3}, c_{i,4}, c_{i,5}, c_{i,6}$ 表示第 i 种蔬菜或者水果所含的膳食纤维、维生素 A 、维生素 C 、维生素 E 和钙锌成分含量； x_i 表示第 i 种蔬菜或者水果的需求量。对不同地区，分别根据不同季节的数据，重复建立以上数学模型，求得不同季节应选择的果蔬品种及消费量，最后汇总得到整个年度的果蔬消费品种和消费量。

6.4.2 模型求解

不同地区在不同季节蔬菜水果的价格有所不用，导致相应的消费量有所不同。本文以北方地区为例在四个季度分别进行优化得到在满足人体健康需要前提下购买成本最低的人均消费量。由目标函数和约束条件用 *MATLAB* 软件求解得到蔬菜水果人均消费量，结果见表 20。

表 20 不同季度所需蔬菜水果消费量表（单位/千克）

	第一季度	第二季度	第三季度	第四季度
萝卜	45.5	56.3	26.5	38.9
胡萝卜	59.1	1.0	0	0
大白菜	81.4	42.1	82.5	120.7
柑橘	32.3	79.9	76.8	32.8
目标值（元）	484.91	413.72	457.54	396.58

由表 20 可知，在第一季度，以萝卜、胡萝卜、大白菜和柑橘为主要蔬菜水果摄取营养物质，可得到整个季度最低购买成本为 484.91 元，平均每天 5.39 元。在第二季度，以萝卜、大白菜和柑橘为主要蔬菜水果摄取营养物质，可得到整个季度最低购买成本为 413.72 元，平均每天 4.60 元。在第三季度，以萝卜、大白菜和柑橘为主要蔬菜水果摄取营养物质，可得到整个季度最低购买成本为 457.54 元，平均每天 5.08 元。在第四季度，以萝卜、大白菜和柑橘为主要蔬菜水果摄取营养物质，可得到整个季度最低购买成本为 396.58 元，平均每天 4.41 元。可见在第二三四季度都是以萝卜、大白菜和柑橘为主要水果蔬菜。

6.4.3 目标优化的灵敏度分析

在讨论线性规划问题时，假定系数 a_i, b_i, c_{ij} 都是常数。但实际上这些系数往往是估计值和预测值。如市场条件一变，价格 a_i 值就会变化； c_{ij} 往往是因自然条件的改变而改变； b_i 是决策选择。因此提出这样两个问题：当这些系数有一个或几个发生变化时，已求得的线性规划问题的最优解会发生变化；或者这些系数在什么范围内变化时，线性规划问题的最优解或最优基不变。因此要对决策变量进行灵敏度分析。

灵敏度分析法是研究与分析一个模型的状态或输出变化对系统参数或周围条件变化的敏感程度的方法。在最优化方法中经常利用灵敏度分析来研究原始数据不准确或发生变化时最优解的稳定性。通过灵敏度分析还可以决定哪些参数对系统或模型有较大的影响。

附录

表 21 各个季度蔬菜水果价格表

种类	第一季度	第二季度	第三季度	第四季度
萝卜	1.78	1.41	1.55	1.46
胡萝卜	2.08	2.63	2.44	2.12
大白菜	1.26	1.76	1.63	1.14
柑橘	5.97	6.71	4.67	5.45
香蕉	4.82	9.58	9.92	7.40
西瓜	5.19	3.05	2.86	5.01
甜瓜	5.95	7.23	4.76	6.03

表 22 各种营养物质的含量表

含量 (mg/kg)	膳食纤维	维生素	维生素	维生素	钙	锌
萝卜	10	0.03	21	9.2	360	3
胡萝卜	13	6.88	16	4.1	320	1.4
大白菜	8	0.42	31	7.6	500	3.8
柑橘	4	1.48	28	9.2	350	0.8
香蕉	12	0.1	8	2.4	70	1.8
西瓜	3	0.75	6	1	80	1
甜瓜	2	1.53	12	2	40	1.3

7.问题四模型建立与求解

7.1 问题分析：

果蔬类农业结构的优化和调整涉及多种复杂的因素，为了有效的改善果蔬类农产品的种植结构，重点考虑其中的几个方面的因素。各种果蔬类产品的生产产量可由果蔬类产品的人均消费量得到。要确定各种果蔬类产品的人均消费量，一方面要满足我国居民的人均需要，即能够满足人体各营养元素摄入标准的产量，在此基础上将价格引入作为考虑因素，考虑居民的购买成本，使居民在满足营养摄入的基础上的果蔬类产品的消费成本最少，使各种果蔬类产品的人均消费量进一步优化。另一方面，农业生产在我国的比重很大，是农民的重要收入来源，为此，在考虑消费者的利益的基础上再考虑生产者的利益，使生产者的收益最大，使结果更优化。要满足以上两个方面的目标，每类蔬菜和水果还有自身的限制条件。如消费量受生产量、损耗率、进出口贸易食品加工等因素的影响。生产量受种植面积，其它果蔬类产品种植量的影响等。拟建立多目标规划数学模型进行，求解我国居民主要的蔬菜水果产品的年度合理人均消费量和生产量，分析我国 2020 年水果和蔬菜产品的调整战略。

7.2 模型准备：

- 1.调查汇总我国的主要果蔬类产品的种植面积^[5]。
- 2.调查汇总我国的主要果蔬类产品的进出口量、损耗率、食品加工率^[6]。
- 3.调查汇总我国的主要果蔬类产品的每公斤的种植收益情况^[7]。

7.3 多目标优化模型^[4]的建立

为实现人体营养均衡满足健康需要，国家可能需要对水果和蔬菜各品种的生产规模做出战略性调整。一方面国家要考虑到居民人体的营养均衡，并使营养摄入量尽量在合理范围内；另一方面也要顾及居民的购买成本，使其购买成本尽量低；同时还要使种植者能够尽量获得较大收益。探讨用多目标规划模型研究我国果蔬产业的调整战略，模型中以购买成本最低和种植者收益最大为目标，满足人体营养需求、进出口量、种植面积以及种植规模等约束条件。不考虑季节的变动时选择蔬菜水果品种的变动，只从所有的蔬菜水果中筛选出主要的水果蔬菜六种进行多目标规划。

7.3.1 目标函数的确定：

1、最低购买成本的目标函数确定：

由分析可知，此目标函数是当今中国居民能够以最低的成本购买蔬菜水果。查阅中国统计年鉴得到主要水果蔬菜产品的单价（求得年均价），居民在蔬菜水果上的购买成本就是水果蔬菜的单价与相应购买量的乘积，对所有的蔬菜水果进行累加求和。所以目标函数可表示为：

$$\min z = \sum_{i=1}^n a_i x_i$$

其中 a_i 表示第 i 种蔬菜或者水果的价格； x_i 表示第 i 种蔬菜或者水果的总消费量。

2、种植者最大收益的目标函数确定：

由分析可知，此目标函数是当今中国果蔬类产品生产者能够获得最大的产值。查阅中国年鉴等相关资料得到主要蔬菜水果产品的收购价格，生产者在蔬菜水果上的收益近似的用产值来表示。则果蔬类生产者的收益即为水果蔬菜的收购价格与相应生产量的乘积，并对所有水果蔬菜进行累加求和，得到果蔬类产品的总产值。目标函数可表示为：

$$\max f = \sum_{i=1}^n J_i y_i$$

其中 J_i 表示第 i 种果蔬类产品的收购单价， y_i 表示第 i 种果蔬类产品的生产量。

7.3.2 约束条件的确定

约束条件分为四大类，需要考虑营养元素满足机体需要、进出口量、种植面积以及种植规模的调整等。

1、营养摄入的约束

营养成分含量能达到人体所需要的标准，如问题三的条件约束：

$$B_j < \sum_{i=1}^n c_{ij} \times x_i < B'_j$$

其中 c_{ij} 表示第种果蔬产品的第种营养成分的含量。 x_i 表示第种果蔬类产品的人均消费量， B_j, B'_j 分别表示第种营养成分的人体必需的下限量和上限量。

2、种植面积的约束

此约束是固有种植结构约束，即蔬菜类产品和水果类产品的种植面积应在往年的种植面积呈正负百分之十的范围内波动。约束条件为：

$$\begin{aligned} 0.9C_1 &< P_1 + P_2 + \dots P_n < 1.1C_1 \\ 0.9C_2 &< P'_1 + P'_2 + \dots P'_n < 1.1C_2 \end{aligned}$$

其中 P_i 表示第 i 种蔬菜品种的种植面积， P'_i 表示第 i 种水果的种植面积， C_1, C_2 分别代表蔬菜总种植面积和水果的总种植面积。其中总种植面积是由总产量与单位亩产量之比确定的。

3、种植规模变动的约束

此约束是可控种植结构约束，即不同种蔬菜水果的种植面积具有一定的联系，这是由于种植者会根据不同产品的价格变动相应的增减产品的种植面积。搜索最近几年的主要产品的种植面积，进行平稳性检验。建立一个时间序列，通过该序列的时间路径图来判断是否平稳的。平稳的时间序列的图形上表现出围绕其均值不断波动的过程，而非平稳序列则往往表现出在不同时间段的均值（如持续上升或持续下降）

对于非平稳变量，为避免伪回归现象，对于不同变量进行协整检验，用 OLS 法估计方程 $Y_t = \alpha_0 + \alpha_1 X_t + \mu_t$ 并计算非均衡误差，得到

$$\begin{aligned} \hat{Y} &= \hat{\alpha}_0 + \hat{\alpha}_1 X_t \\ e_t &= Y_t - \hat{Y}_t \end{aligned}$$

再对 e_t 检验单整性。如果 e_t 为稳定序列，则认为变量为协整。

对协整变量组合进行回归分析，从中选出各变量组合的系数显著，残差平稳且相对独立的回归方程，整理后得到刚性约束条件：

$$x_i = \alpha_0 x_1 + \alpha_1 x_2 + \dots \alpha_n x_n$$

根据最近四年六种果蔬产品的种植面积数据（见附录 4），借助 Eviews6.0 软件对六类蔬菜水果种植面积进行回归分析，得到具体的约束方程：

$$st \begin{cases} x_1 = 0.750628x_2 + 0.330210x_3 - 0.542810x_5 + 0.432510x_6 \\ x_2 = 0.894736x_3 + 0.652411x_5 + 0.987325x_6 \\ x_4 = 0.642589x_5 + 0.756423x_6 \end{cases}$$

4、生产量与消费量的约束

生产量和消费量之间有一定的数量关系。将每种产品的进出口量、损耗量及食品加

工消耗的量进行汇总，除进口量取正值，其它的均取负值，得到附加量 F_i 值，则约束条件为

$$x_i = (y_i - F_i) / R$$

其中 R 表示全国总人口。

7.3.3. 模型的建立

综上所述建立模型如下：

目标函数：

$$\begin{aligned} \min z_1 &= \sum_{i=1}^6 a_i x_i \\ \max z_2 &= \sum_{i=1}^6 J_i y_i \end{aligned} \quad (18)$$

$$st. \begin{cases} B_j < \sum_{i=1}^6 c_{ij} \times x_i < B_j, j=1,2,\dots,6, \\ 0.9C_1 < P_1 + P_2 + \dots + P_n < C_1, \\ 0.9C_2 < P_1 + P_2 + \dots + P_n < C_2, \\ x_1 = 0.750628x_2 - 0.330210x_3 + 0.542810x_5, \\ x_2 = 0.894736x_3 + 0.652411x_5 + 0.987325x_6, \\ x_4 = 0.642589x_5 + 0.756423x_6, \\ y_i = R \times x_i - F, \end{cases}$$

7.4 模型求解

在实际问题中两个目标表示不同的物理意义，并且数值也相差很大，因此必须统一每个目标函数的量纲和数量级，再进行线性加权化为单目标函数。优先考虑提高种植者的收入，将第二个目标设置相对较大的权值。建立单目标函数：

$$\max z = -\omega_1 z_1 + \omega_2 z_2$$

其中 ω_1, ω_2 分别代表 z_1, z_2

约束条件不变（式（18））。

运用规划软件 Matlab 对上述单目标线性规划模型求解得到最优解：

表 23 最优化人均消费量表

种类	萝卜	大白菜	胡萝卜	香蕉	苹果	西瓜
人均消费量(千克/年)	42.54	76.39	17.22	8.98	28.58	12.81

7.5 模型灵敏度分析

在讨论线性规划问题时，假定系数 a_i, b_i, c_{ij} 都是常数。但实际上这些系数往往是估计值和预测值。如市场条件一变，价格 a_i 值就会变化； c_{ij} 往往是因自然条件的改变而改变；

b_i 是决策选择。因此提出这样两个问题：当这些系数有一个或几个发生变化时，已求得的线性规划问题的最优解会发生变化；或者这些系数在什么范围内变化时，线性规划问题的最优解或最优基不变。因此要对决策变量进行灵敏度分析。

灵敏度分析法是研究与分析一个模型的状态或输出变化对系统参数或周围条件变化的敏感程度的方法。在最优化方法中经常利用灵敏度分析来研究原始数据不准确或发生变化时最优解的稳定性。通过灵敏度分析还可以决定哪些参数对系统或模型有较大的影响。灵敏度分析方法有以下特点：

(1) 灵敏度分析法研究的是各因素对模型的全局影响（不仅是在某点处，而是在不同位置处）；

(2) 因素的范围可扩展到因素的整个定义域，各因素可同时变化，能够对非线性、非叠加、非单调模型进行研究和分析。

目前最常见的全局灵敏度分析方法是Sobol'法。Sobol'灵敏度分析方法是一种基于方差的蒙特卡罗法。定义一个 k 维的单元体 Ω^k 作为输入因素的空间域，表示为 $\Omega^k = \{x | 0 \leq x_i \leq 1; i = 1, 2, \dots, k\}$ 。

Sobol'方法的中心思想是将函数 $f(x)$ 分解为子项之和

$$f(x_1, x_2, \dots, x_k) = f_0 + \sum_{i=1}^k f_i(x_i) + \sum_{i=1}^k f_{ij}(x_i, x_j) + \dots + f_{1,2,\dots,k}(x_1, x_2, \dots, x_k) \quad (19)$$

上式右端共有 2^k 个子项，且有多种分解方法，其中给予多重积分的分解方法具有一般代表性。该分解方法的特点如下：

(1) f_0 为常数项，各子项对其所包含的任一因素的积分为

$$\int_0^1 f_{i_1, i_2, \dots, i_s}(x_{i_1}, x_{i_2}, \dots, x_{i_s}) dx_{i_j} = 0 (1 \leq j \leq s)$$

(2) 各子项之间正交。即如果：

$$(i_1, i_2, \dots, i_s) \neq (j_1, j_2, \dots, j_l), \text{ 则 } \int_{\Omega^k} f_{i_1, i_2, \dots, i_s} \cdot f_{j_1, j_2, \dots, j_l} dx = 0。$$

(3) 式(19)中分解形式唯一，且各阶子项可有多重积分求得。如：

$$f_0 = \int_{\Omega^k} f(x) dx$$

$$f_i(x_i) = -f_0 + \int_0^1 \cdots \int_0^1 f(x) dx_{-i} (1 \leq i \leq k)$$

根据统计学的知识，模型输出 $f(x)$ 的总方差为 $D = \int_{\Omega^k} f^2(x) dx - f_0^2$ ，总方差与各阶偏方差的关系：总偏方差等于各阶偏方差之和。即

$$D = \sum_{i=1}^k D_i + \sum_{1 \leq i < j \leq k} D_j + \cdots + D_{1,2,\dots,k}$$

将各阶灵敏度系数定义为各阶偏方差与总方差的比值。 s 阶灵敏度 $S_{1,2,\dots,k} = \frac{D_{1,2,\dots,k}}{D} (1 \leq i \leq k)$ 。这里， S_i 成为因素 x_i 的一阶灵敏度系数，表示 x_i 对输出的主要影响； $S_{ij} (i \neq j)$ 为二阶灵敏度系数，表示来那个因素之间的交叉影响。

在 Sobol' 方法中，各积分可由蒙特卡罗法求出。因此 f_0 ， D 及 D_i 可通过蒙塔卡罗估计得出

$$\hat{f}_0 = \frac{1}{n} \sum_{m=1}^n f(x_m),$$

$$\hat{D} = \frac{1}{n} \sum_{m=1}^n f^2(x_m) - \hat{f}_0^2,$$

$$\hat{D}_i = \frac{1}{n} \sum_{m=1}^n f(x_{im}^{(1)} x_{(-i)m}^{(1)}) f(x_{im}^{(1)} x_{(-i)m}^{(2)}) - \hat{f}_0^2,$$

经灵敏度分析，决策变量中除了萝卜和西瓜的人均年消费量对进出口量、种植面积、种植规模等因素非常敏感，最优解会发生微小变化，其它的决策变量的最优解受约束条件的变化而保持基本不变。

7.6 到 2020 年的战略调整

将第一问的 2020 年各种蔬菜水果的消费量作为种植者的生产量，见表 24：

表 24 各种蔬菜水果的人均生产量表

种类	萝卜	大白菜	胡萝卜	香蕉	苹果	西瓜
人均生产量（千克/年）	38.02	88.5	16.84	10.10	49.35	16.38

对表 23 和表 24 进行比较得到人均生产多余人均消耗的果蔬产品有大白菜、香蕉、苹果和西瓜。国家应该宏观调控，相应缩小这三种产品的种植规模或者提高种植质量鼓励出口。其它的产品如萝卜和胡萝卜的生产量小于消费量，因此国家应采取相应的鼓励

措施，扩大这两类的生产规模，或提高生产质量，提高产量。

8、模型评价

问题一在选取主要的蔬菜水果品种时，能充分考虑产量、营养成分含量、品种数，较为全面的满足了研究的需要。另外，在对消费量进行估计时，采用灰色预测和曲线拟合两种建模方法，并通过检验确定出采用灰色预测估计消费量更为科学合理。

问题二中对营养成分含量分析时，采用聚类分析简化了问题，建立的隶属度模型使中国居民营养物质的摄入合理性量化，更加直观地表达出中国居民的人体营养健康状况。然而，在分析过程中，忽略了营养物质的摄入量不等同于吸收量的问题，对人体营养健康状况的分析造成一定的偏差，模型有待优化。

问题三中，综合考虑了地区、季节变化对整个目标函数最优解的影响，考虑合理性更强，并且能够给出不同地区的居民在不同季节的最优方案。同时将营养元素通过聚类分类将其进行分类，只选择其中几种进行分析，使结果更具代表性。

问题四，约束条件考虑的很全面。在第三问的基础上，综合考虑了生产量、进出口量、损耗率和加工率等因素从而约束了人均年消费量，同时还考虑了各种蔬菜水果的固有种植面积和种植规模的变动的的影响，从而约束生产量。建立双目标优化方案，使结果更客观。

9. 问题五：关于我国果蔬业发展政策的建议

我国地跨寒、温、热三带，自然条件极其复杂，植物种类繁多，果树蔬菜资源尤为丰富，其种类和产量位居世界前列，但在果蔬产业的生产与销售过程中还存在一些不足。

近 20 年来，我国蔬菜产业已取得很大进展，但从整体上看，供大于求的现象依然存在，产品质量及单位面积产量不高，产业链的产前、产中、产后各环节比例不协调，贮藏加工落后，经济效益起伏不定等突出问题亟待解决。主要表现在蔬菜产品总体质量不高如形状大小营养成分等难以达到标准，农药残留严重超标，在生产过程中由于不科学的管理产生三废问题，蔬菜产业结构不合理等。

针对我国蔬菜种植业存在的问题，未来应该向以下三个方向发展：一是由传统的蔬菜生产转向绿色无公害蔬菜生产。二是加大发展大棚蔬菜的力度，尽量减少露天蔬菜的栽培，提高蔬菜培养的科技含量，生产反季节蔬菜，解决蔬菜盛产期卖蔬菜难的问题。三是科研部门销售部门生产部门建立繁殖销售生产一体化道路。大力减少农村合作社工程，加强蔬菜的创新和新品种的选育。同时进一步增加蔬菜采后加工业的科技投入，加强国际合作，提高产业核心竞争力，如可通过项目合作、引进技术与管理、人员培训、互访、信息交流等方式了解国外该领域的科研现状、发展方向等。

我国水果产业是种植业中仅次于粮食和蔬菜的第三大产业，在国民经济中占有非常重要的地位。但我国的水果产业中也存在诸多的问题。例如，由于体制的约束，生产与流通相互联系薄弱，对外贸易对国内生产和需求信号反应不敏感。水果产品的出口体系建设落后，出口渠道不顺畅使中国失去了许多水果出口机会，水果产品布局和品种结构不合理，科技含量不高，生产标准不完善，水果加工能力不足，水果消费的产品链过断，水果贸易市场体系不健全等。

针对上述问题，我国水果产业应采取相应的措施。首先实现水果品种和区域化合理配置，有关部门要加大科技、人才、信息、政策等方面的扶持力度，形成集科研开发、技术推广、品种改良、种苗繁育、生产加工的体系，大幅度提高我国水果的生产能力和

质量水平。发展专业合作组织，扶持水果龙头企业。加强果品市场体系建设，维护水果市场秩序，政府应对水果的产销市场秩序进行管制，防治水果价格大幅度波动，积极开拓市场，建立完善的市场体系和信息网络体系，加大水果批发市场的建设力度，通过信息网络及时反馈各地的市场批发价格和主要产地的生产状况，沟通，衔接水果产销，宣传扩大产品的知名度。保护国内市场，积极开拓国际市场促进水果出口。

参考文献

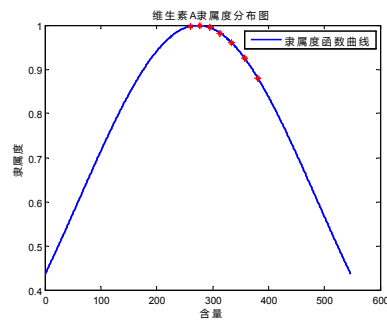
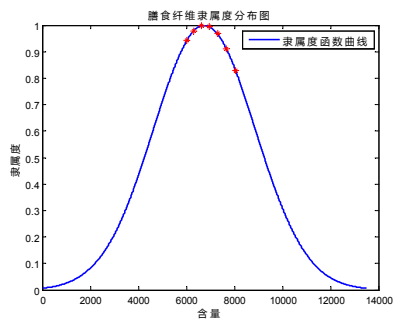
- [1], 姜启源, 谢金星, 叶俊, 数学建模[M], 北京: 高等教育出版社, 325-332
- [2], 卓金武, 魏永生, 秦健, 李必文, MATLAB 在数学建模中的应用[M], 北京: 北京航空航天大学出版社, 5-9, 26-28, 30-35, 2011.4
- [3], 周博, 张惟, 侯纲领, MATLAB 科学计算范例实战速查宝典[M], 北京: 清华大学出版社, 313-315, 2013.7
- [4], 汪晓音, 周保平, 数学建模与数学实验[M], 北京: 科学出版社, 89-91, 2012.8
- [5], 胡小松 肖华志 牛丽影 廖小军, 中国果蔬产业发展状况与预测, Vol.2, No.2: 25-27, 2002.3
- [6], 中华人民共和国农业部种植业管理司, 蔬菜数据库, <http://202.127.42.157/moazzys/shucaixcx.aspx>, 2014年9月21日
- [7], 中华人民共和国国家统计局, 茶叶、水果产量, <http://data.stats.gov.cn/workspace/index?m=hgnd>, 2014年9月21日

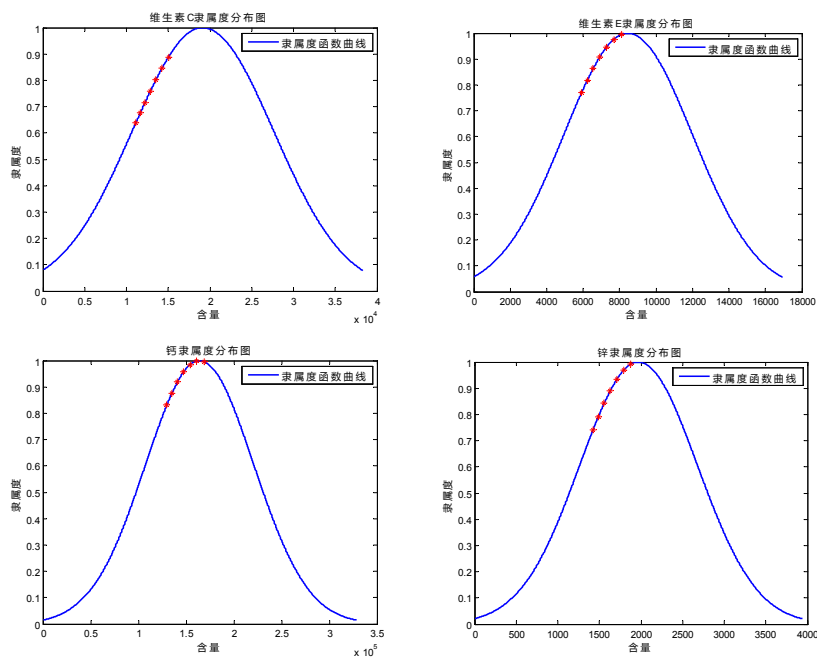
附录

附录 1

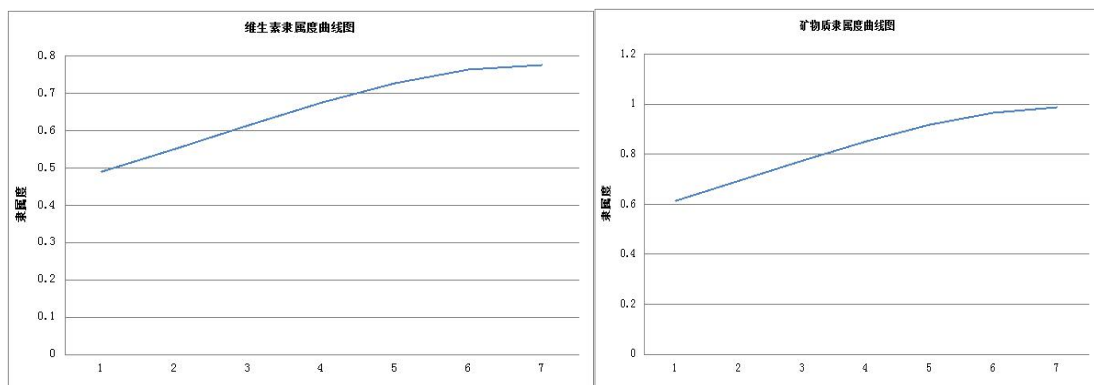
水果	2003	2004	2005	2006	2007	2008	2009	2010
苹果	0.125	0.214286	0.157895	0.01227	1	0.314286	0.045455	0.004237
梨	0.375	0.214286	0.078947	0.006135	1	0.085714	0.318182	0.243644
柑橘类	0.625	0.285714	0.131579	0.050613	0.308219	1	0.090909	0.375
香蕉	0.25	0.285714	1	0.004601	0.342466	0.914286	0.181818	1
菠萝	1	0.142857	0.210526	0.03681	0	0.514286	0.227273	0.311441
荔枝	0.25	0.428571	0.236842	0.055215	0.068493	0.171429	0.227273	0.408898
龙眼	0.125	1	0.526316	0.065951	0	0.171429	0.090909	0.525424
桃	0.125	0.142857	0.315789	1	0.890411	0.914286	0.136364	0.305085
猕猴桃	0.125	0.214286	0.052632	0.013804	0.479452	0.342857	0.227273	0.305085
葡萄	0.625	0.214286	0.105263	0.006135	0.232877	0.314286	0.090909	0.262712
红枣	0.75	0.357143	0.368421	0.455521	0.068493	0.457143	0.318182	0.269068
柿子	0.25	0.142857	0.157895	0.046012	0.767123	0.257143	0.090909	0.319915
西瓜	0.375	0.285714	0.184211	0.015337	0.068493	0.371429	0.090909	0.254237
甜瓜	0.625	0.071429	0.289474	0.053681	0.136986	0.114286	0.136364	0.402542
草莓	0.375	0.214286	0.105263	0.053681	0.273973	0.428571	1	0.360169

附录 2





附录 3



附录 4

六种果蔬的亩产量表

品种	萝卜	大白菜	胡萝卜	香蕉	苹果	西瓜
亩产（千克）	863	986	575	3750	8500	2460

六种果蔬最近四年的种植面积表

品种(千公顷)	萝卜	大白菜	胡萝卜	香蕉	苹果	西瓜
2010	93.51	126.41	86.51	357.33	2139.94	1812.52
2011	96.55	138.93	87.60	386.04	2177.32	1803.17
2012	94.18	142.44	85.38	401.05	2198.53	1836.21
2013	92.62	159.49	84.63	425.27	2221.41	1853.83

六种果蔬的收购价格表

品种	萝卜	大白菜	胡萝卜	香蕉	苹果	西瓜
收购价格 (元/千克)	0.95	2.3	1.2	3.5	3.8	0.8

附录 5 水果灰色预测模型

```

clear
clc
N=xlsread('shuiguorenjun.xlsx');
olk=xlsread('sunhao.xlsx');
&%原始数据 x0(k)，包含 N 个元素;
mmm=10;
x0=N(mmm,:)*olk(mmm,1);
x0=N;
chang=length(x0);
f=x0;    %f 用于保留原值!
    %灰色预测计算模型
x1=[];
n=chang;
for k=1:n
    if k==1
        x1(k)=x0(k);
    else x1(k)=x1(k-1)+x0(k);
    end
end;                %计算 x1()

n=n-1;
B=ones(n,2);
for i=1:n
    B(i,1)=-0.5*(x1(i)+x1(i+1));
end                %参数 B

x0(:,1)=[];
Y=x0';                %参数 Y

A=inv(B'*B)*B'*Y;    %可知 a 和 u
a=A(1);
u=A(2);
cgb=u/a;

%模型标定求 x'                x11 为累加数，加成数；
% n=n+1+m;                %m 为向后预测的年数,需修改!
n=n+1+10;

```



```

x11(1)=f(1);
for j=2:n
    x11(j)=(f(1)-u/a)*exp(-a*(j-1))+u/a;
end
%用灰色模型计算 x'0
x01(1)=f(1);

```

附录 6 蔬菜灰色预测模型

```

clear
clc
N=xlsread('shucairenjun.xlsx');
olk=xlsread('sunhao.xlsx');
%原始数据 x0(k)，包含 N 个元素
mmm=8;
x0=N(mmm,:)*olk(mmm,1);
chang=length(x0);
f=x0;    %f 用于保留原值！
%灰色预测计算模型
x1=[];
n=chang;
for k=1:n
    if k==1
        x1(k)=x0(k);
    else x1(k)=x1(k-1)+x0(k);
    end
end;          %计算 x1()

n=n-1;
B=ones(n,2);
for i=1:n
    B(i,1)=-0.5*(x1(i)+x1(i+1));
end          %参数 B

x0(:,1)=[];
Y=x0';          %参数 Y

A=inv(B'*B)*B'*Y;    %可知 a 和 u
a=A(1);
u=A(2);
cgb=u/a;

%模型标定求 x'          x11 为累加数，加成数；
% n=n+1+m;              %m 为向后预测的年数,需修改!
n=n+1+12;

```

```

x11(1)=f(1);
for j=2:n
    x11(j)=(f(1)-u/a)*exp(-a*(j-1))+u/a;
end
%用灰色模型计算 x'0
x01(1)=f(1);
for k=2:n
    x01(k)=x11(k)-x11(k-1);    %x01 为预测值;
end
%计算残差 e0(i)和相对误差 t0(i)
for j=1:chang
    e0(j)=x01(j)-f(j);
    t0(j)=e0(j)/f(j);
end

%绘制预测值及原值曲线
n=chang;                                year1=1:n;year01=2003:2008;
year2=1:n+12;    % n+m, 预测 m 年
year02=2003:2020;
subplot(2,1,1);
plot(year01,f,'go',year02,x01,'b-+', 'linewidth',2)
xlabel('年份')
ylabel('原值 (o) 及预测值 (+) ')
title('芹菜消费量预测图')
% subplot(2,1,2);plot(year1,e0,'*',year1,t0,'+')
subplot(2,1,2);
plot(year01,t0,'+')

xlabel('年份')
ylabel('相对误差 (+) ')
title('芹菜消费量误差图')
%判断预测精度
disp('完成预测')
a=abs(a);
if a<=0.3
    disp('可进行中长期预测')
elseif a<=0.5
    disp('可用于短期预测,中长期预测慎用')
elseif a<=0.8
    disp('作短期预测应十分谨慎')
    disp('作短期预测应十分谨慎')
else
    disp('不宜预测')
end
end

```

附录 7 拟合回归分析程序

```
clear
clc
[data22]=xlsread('shucairenjun.xlsx');
t=data22(7,:);
i=1:6;
nian=2003:2008;
A=data22(1,i);
figure(1)
plot(nian,A,'r^');
title('黄瓜消费量预测图')
xlabel('年份')
ylabel('人均消费量')
hold on
a=polyfit(nian,A,1);
[a,s]=polyfit(nian,A,1);
A1=A-polyval(a,nian);
AA1=mean(A1);
t1=2003:1:2010;
A2=polyval(a,t1);
AA2=A2+AA1;
figure(1)
plot(t1,AA2,'b','linewidth',2);
title('黄瓜消费量预测图')
hold on
p=max(AA2);
[y,delta]=polyconf(a,nian,s);           %置信区间;
AA3=y;
t0=ones(6,1);
t1=t';
t2=t1.*t1;
T=[t0 t1 t2];
[b,bint,r,rint,stats]=regress(AA3',T,0.05);
figure(2)
b,bint,stats,rcoplot(r,rint)
title('黄瓜消费量误差分析图')
```

附录 8 隶属度程序

```
clear
clc
N=xlsread('chengfenshujun.xlsx');
i=1;
a=6570;
```

```

a1=N(i,:);
j=1:7;
b(j)=(a1(j)-a).^2;
s=sum(b(j))/6;
u=sqrt(s);
u=std(a1);
N=xlsread('chengfenshuju.xlsx');
x=0:2*a;
y=exp(-((x-a)/u).^2);
plot(x,y,'b','linewidth',2)
title('铋隶属度分布图')
xlabel('含量')
ylabel('隶属度')
legend('隶属度函数曲线')
hold on
x1=N(i,:);
lishudu=exp(-((x1-a)/u).^2);
plot(x1,lishudu,'r*')
hold on
lishudu

```

附录 9 目标优化程序

```

clear
clc
N1=xlsread('jiage1.xlsx');
N2=xlsread('chengfen.xlsx');
i=4;
c1=N1(:,i)';
a=[N2(:,1)';N2(:,2)';N2(:,3)';N2(:,4)';N2(:,5)';N2(:,6)';...
    N2(:,1)';N2(:,2)';N2(:,3)';N2(:,4)';N2(:,5)';N2(:,6)'];
b=[-5475/4*0.9,-255.5/4*0.9,-12775/4*0.9,-7008/4*0.9,-153300/4*0.9,-1752/4*0.9,...
    7665/4*0.9,292.9/4*0.9,25500/4*0.9,9928/4*0.9,175200/4*0.9,2190/4*0.9];
[x,y]=linprog(c,-a,b,[],[],zeros(7,1));

```

附录 10 多目标优化程序

```

clear
clc
N1=xlsread('jiage.xlsx');
N2=xlsread('chengfen.xls');
i=1;
c1=N1(:,i)';
c=[c1(1,1);c1(1,2);c1(1,3);c1(1,4);c1(1,5);...
    c1(1,6);c1(1,7);c1(1,8);c1(1,9);c1(1,10);...
    c1(1,11);c1(1,12);c1(1,13);c1(1,14);c1(1,15);c1(1,16)];
a=[N2(:,1)';N2(:,2)';N2(:,3)';N2(:,4)';N2(:,5)';N2(:,6)'];

```

```
b=[5475;255.5;12775;7008;153300;1752]/4;  
[x,y]=linprog(c,-a,-b,[],[],zeros(16,1)); [x,y]=linprog(c,-a,b,[],[],zeros(7,1));
```