**蔬菜类商品的自动定价与补货策略设计**

摘要

在生鲜商超中，蔬菜保鲜期短且品相随时间变差，需每日补货。本文分析蔬菜销售量的分布规律及相互关系，探讨销售总量与成本加成定价的关系，建立数学模型，并给出未来一周的补货和定价策略，制定单品补货计划以最大化收益。同时，提出需采集的相关数据及其对决策的帮助，提高商超收益。

针对问题一,首先对数据进行标准化处理，然后对数据进行可视化处理，以此来分析各蔬菜品类，单品之间的销售规律。结果得到不同品类和不同单品的蔬菜的销量都存在明显季节性差异，通过斯皮尔曼相关系数分析对不同单品之间的销量进行分析，发现每个蔬菜品类之间都具有显著相关性，且发现除了茄类，其余每个品类之间都存在显著正相关，茄类与其他品类之间的存在轻微负相关。通过K-means聚类分析对单品数据进行聚类分析，将单品类型分为持续性畅销型数据，季节性畅销型数据，持续性低迷型数据，间断性畅销型数据，得到有231个数据属于持续性畅销型数据，12个季节性畅销型数据，2个间断性畅销型数据，1个持续低迷型数据。由此，便可得出品类之间，单品之间的销量关系。

针对问题二，首先对数据进行预处理，排除错误和无关数据，将数据进行分组聚合处理。由此可以分析蔬菜单品销售在品类之间的关系，销量和成本加成定价的关系，蔬菜单品销售、销量和时间之间的关系。蔬菜单品销售在品类之间是被包含关系，通过价格弹性模型的分析，得出销售价格和加成比例成线性关系，加成比例和需求量有明显关系。通过线性回归方程和时间序列预测模型可预测出各蔬菜品类未来一周(2023 年 7 月 1-7 日)的日补货总量和定价策略，使得商超收益最大。

针对问题三,对数据进行预处理，排除单品订购量小于2.5千克的单品，然后根据所获利润进行排序，挑选出前30个利润率最高的单品，然后利用灰色预测模型，通过累加求和，灰微分方程，白化微分方程，得到相应预测函数，然后预测7月1日的进货方案与商品定价。

针对问题四，对数据进行预处理，然后通过大数据分析和相关文献查找，发现有很多因素可以进行考虑，例如客户购买行为，季节性需求变化，天气因素，竞争对手，物流和运输数据等，这些因素可以很好的对蔬菜商品的补货和定价决策进行评估。

**关键词：斯皮尔曼相关系数分析，灰色预测，时间预测模型**

# 问题重述

## 问题背景

在生鲜的商场和超市中，一般蔬菜类的商品保鲜时间都比较短，并且新鲜度和品相随销售时间的增长而变差， 大部分品种如果当天没有售出，隔日就无法再卖。因此，商超通常会根据各商品的历史销售和需求情况每天进行补缺和上货。

## 问题提出

由于商超销售的蔬菜品种众多、产地不一定相同，而蔬菜的进货交易时间通常在凌晨的3:00- 4:00，为此商家须在不太清楚知道具体单品和进货价格的情况下，需要做出当天各蔬菜品类的补货决策。从需求侧来看，蔬菜类商品的销售量与时间往往存在一定的关联关系；从供给侧来看，蔬菜的供应品种在4月至10月较为丰富，商超销售空间的限制使得合理的销售组合变得极为重要。通过建立数学模型，解决下列问题：

1. 请分析蔬菜各品类及单品销售量的分布规律及相互关系。
2. 请分析各蔬菜品类的销售总量与成本加成定价的关系，并给出各蔬菜品类未来一周(2023年7月1-7日)的日补货总量和定价策略，使得商超收益最大。
3. 因蔬菜类商品的销售空间有限，商超希望进一步制定单品的补货计划，要求可出售单品总数控制在27-33个，并且各单品订购量满足最小陈列量2.5千克的要求。根据2023年6月24-30日的可售品种，给出7月1日的单品补货量和定价策略，在尽量满足市场对各品类蔬菜商品需求的前提下，使得商超收益最大。
4. 综合考虑制定蔬菜商品的补货和定价决策，给出商超还需要采集哪些相关数据， 这些数据对解决上述问题有何帮助，写出意见和理由。

# 问题分析

## 对问题一的分析

针对问题一，第一小问要分析析蔬菜各品的分布规律及相互关系，首先对数据进行预处理，对附件2、3、4缺失数据，异常数据采取剔除处理，接着针对预处理后的数据，进行探索性数据分析，对附件1每一个品类进行量化，用量化后的数据来分析分布规律和相互关系。第二小问要分析单品销售量的分布规律及相互关系，对附件2各单品销售进行分析，建立探索性数据分析模型，对各单品分布规律和相互关系进行评价分析。

## 对问题二的分析

针对问题二，第一小问要分析各蔬菜品类的销售总量与成本加成定价的关系，由已知数据可知，成本加成定价法可以算出相应的加成比例，然后通过皮尔逊函数模型比较成本加成比例和销售总量之间的相关性，然后用了价格弹性系数模型。第二问要算出各蔬菜品类未来一周(2023 年 7 月 1-7 日)的日补货总量和定价策略， 使得商超收益最大。用时间序列分析和模拟退火模型来分析未来一周(2023 年 7 月 1-7 日)的日补货总量和定价策略最佳方案。

## 对问题三的分析

针对问题三，通过观察6月24日到6月30日的单品陈列量，筛选出最小陈列量大于2.5千克的单品，将筛选后的单品从6月24日到6月30日按照所获利润多少进行排序，利润等于单品的总售价减去单品的总成本，挑选出所获利润最多的27-33种蔬菜单品，通过灰色预测模型和非线性回归模型（因为在问题二中用线性回归模型处理数据效果不太理想）

## 对问题四的分析

经过大数据处理和相关文献的参考，发现还有客户购买行为数据，季节性需求变化数据，天气数据，竞争对手数据，物流和运输数据，实时库存数据，消费者反馈数据进行评估，可以得到更好地制定蔬菜商品的补货和定价决策。

# 模型假设

1：假设各品类间无影响，时间稳定性

2：假设运输时间固定

3：假设商超的销售空间有限

4：假设近期的损耗率数据是准确的，且商超会对品相变差的商品进行合理的打折销售。

5：假设商超能够采集到所有必要的数据，并且这些数据都是准确的。

6：假设可以忽略日常小规模的异常销售波动，例如由于天气变化或临时小规模活动引起的销量变化。

7：假设市场需求是稳定的，并且可以通过历史数据准确预测。

8：假设所有需要补货的蔬菜供应都能按时到货，没有供应短缺的问题。

9：假设成本加成定价策略在未来仍然有效，且能够合理地反映市场需求和竞争情况。

10：假设通过适当的库存管理，商超可以将蔬菜的损耗率控制在历史数据的范围内。

11：每个数据之间不相互影响,及数据独立。

# 符号说明

|  |  |
| --- | --- |
| **符号** | **说明** |
| C𝑖 | 第𝑖 个簇 |
| μ𝑖 | 第 𝑖 个簇的质心 |
| X | 簇 C𝑖 中的一个数据点 |
| ||x−μ𝑖|| | 数据点x 到质心 μ𝑖 的欧氏距离。 |
|  | 斯皮尔曼秩相关系数，衡量两个变量之间的秩相关性。 |
| N | 样本数量，即数据点的数量 |
| R𝑖 | 第𝑖个数据点在变量R 中的排名 |
| S𝑖 | 第𝑖个数据点在变量 S中的排名 |
|  | 变量 R 的平均排名 |
|  | 变量 S的平均排名 |
| ΔQ | 需求量的变化 |
| Q | 原需求量 |
| ΔP | 价格的变化 |
| P | 原价格 |
| X(0) | 原始时间序列数据 |
| X(1) | 一次累加生成序列 |
| P1 | 销售价格 |
| C | 总成本 |
| A | 加成比例 |
| D | 每日销量 |
| P2 | 单品价格 |
| DP | 每日单个单品的销量 |
| S | 损耗率 |
| C1 | 每日每个单品的成本价格 |
| J | 每日进货量 |
| (1)(k) | 时间序列在时刻k的预测值 |
| P0 | 初始价格 |
| Q0 | 初始需求量或供给量 |
| E | 价格弹性 |
| a,b | 灰色微分方程的参数，分别表示发展系数和灰作用量 |
| Δt | 时间步长 |

这里只列出论文各部分通用符号



，个别模型单独使用的符号在首次引用时会进行 说明。

# 模型的建立与求解

## 问题一的模型建立与求解

### 问题一的探索性数据分析模型建立

探索性数据分析（Exploratory Data Analysis，EDA）模型是一种用来初步分析和总结数据集主要特征的方法。EDA 的目标是通过各种统计和图形技术来了解数据的结构、发现模式、检测异常，并生成假设。EDA 的主要目的是为后续的数据分析、建模和决策提供深入的理解和支持。通过 EDA，可以更好地了解数据的特点，选择K-means模型对单品之间的关系进行分析和斯皮尔曼相关性分析模型对品类之间的关系进行进一步分析。

本问题以蔬菜各品类及单品销售量为基础，从而对附录1、2、3、4进行数据分析。

### 对品类中单品销量分布的分析与求解

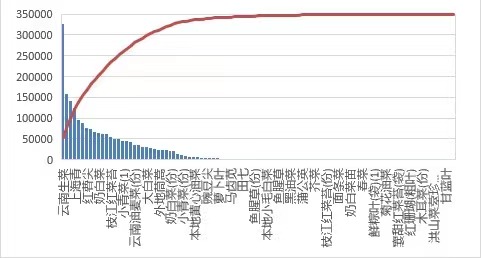
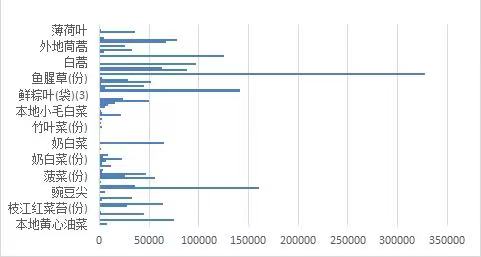
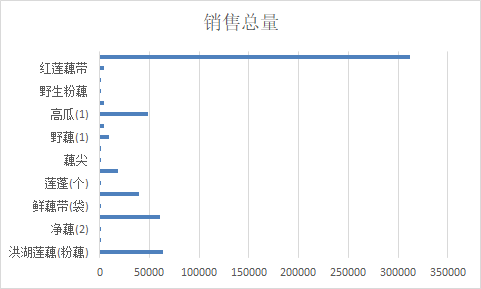
 首先进行数据预处理，用Excel画出销售总量分布图。

图1 花叶类销售总量分布图 （1） 图2 花叶类销售总量分布图（2）

图1中蓝色柱状线代表每个花叶类单品的销售量，红色曲线代表积累频率，表示销售量的累计值。由图1可知，从2020年7月1日到2023年6月30日，花叶类产品的销售表现呈现出显著的长尾分布趋势。云南生菜的销售量最高，代表了一些家庭的常用蔬菜的畅销，销售量的总频率在萝卜叶后趋于稳定，说明萝卜叶后的蔬菜不是家庭常用蔬菜，销售量几乎为零。

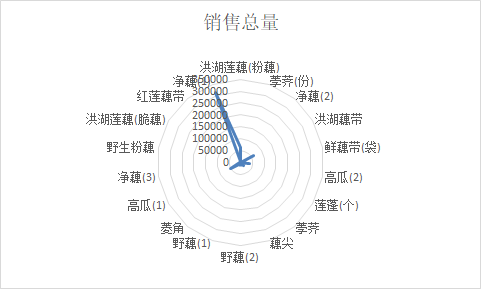
图3 水生根茎类销售统计图 图4 水生根茎类销售雷达图

图3中蓝色柱状图代表每个水生根茎类菜品从2020年7月1日到2023年6月30日的销售总量，其中红莲藕带销售总量最高，占了大多数水生根茎类的销售量，最受欢迎，洪湖莲藕和鲜藕带在第二梯队，销售量远低与红莲藕带，可见，红莲藕带是最受欢迎的。

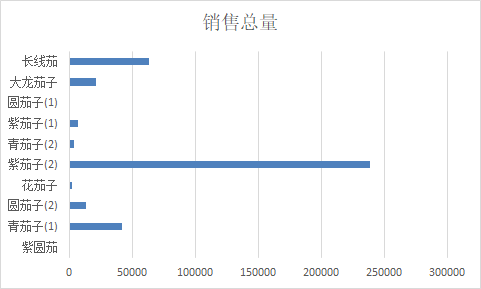
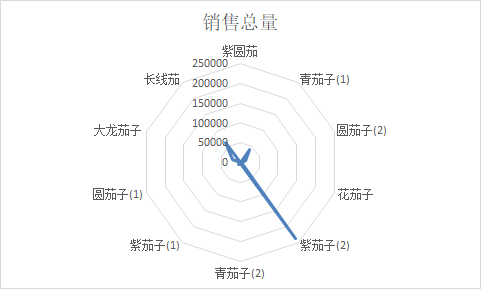


图5 茄类销售量统计图 图6 茄类销售量雷达图

图5是中蓝色柱状图代表每个茄类菜品从2020年7月1日到2023年6月30日的销售总量，其中紫茄子（2）的销售总量最高，远高于其他茄类的销售总量，最受欢迎，圆茄子（1）和紫茄子的销售总量几乎为零，表明在生活中最不常买。

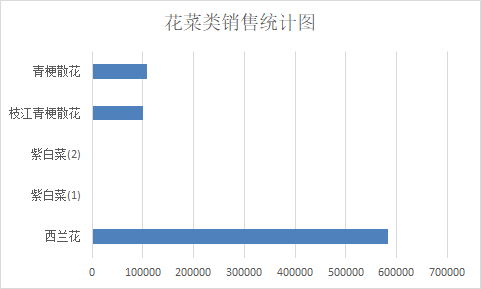
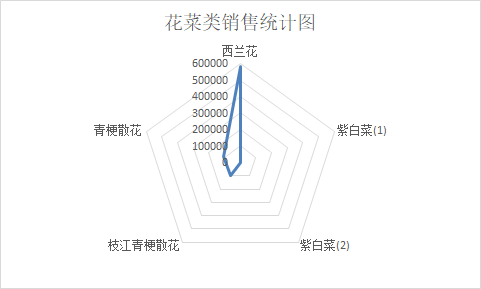
 图7 花菜类销售统计图 图8 花菜类销售雷达图

图7是中蓝色柱状图代表每个花菜类菜品从2020年7月1日到2023年6月30日的销售总量，西兰花的销售量显著高于其他花菜类蔬菜，处于一个持续高销量的状态。西兰花在所有花菜类蔬菜中占据市场领导者地位，销售量遥遥领先于其他品类，表现出持续高销量的趋势。此外，其他花菜类蔬菜（如青梗菜花、枝江青梗菜花和紫白菜）的销售量相对较低。

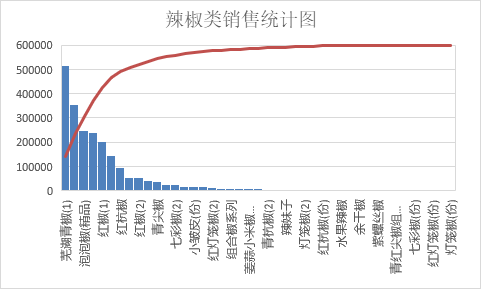
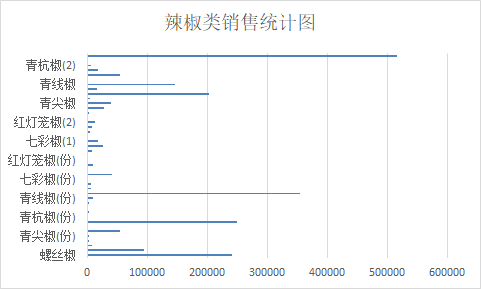


图9 辣椒类销售统计图（1） 图10 辣椒类销售统计图（2）

图9中蓝色柱状线代表每个辣椒类单品的销售量，红色曲线代表积累频率，表示销售量的累计值。由图9可知，从2020年7月1日到2023年6月30日,辣椒类产品的销售表现呈现出显著的长尾分布趋势。芜湖青椒的销售总量最高，其他辣椒类菜品销售量依次降低，芜湖青椒代表了一些家中常用菜的畅销，销售量累计值在青抗椒（2）后趋于稳定，说明青抗椒（2）后的蔬菜购买较少，销售量低，不是家中常用蔬菜。

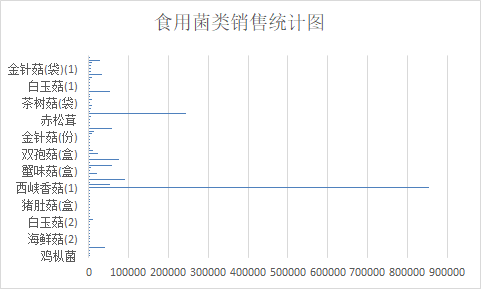
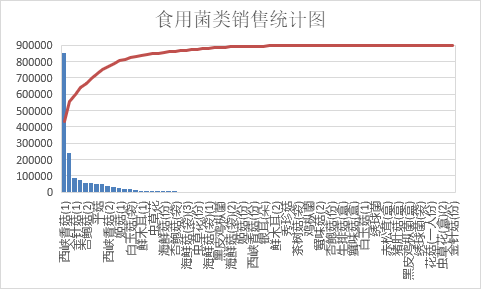
 图11 食用菌类销售统计图 （1） 图12 食用菌类销售统计图 （2）

图11中蓝色柱状线代表每个食用菌类单品的销售量，红色曲线代表积累频率，表示销售量的累计值。由图11可知，从2020年7月1日到2023年6月30。食用菌类产品的销售表现呈现出显著的长尾分布趋势。少数几种菌类（如香菇、平菇等）表现出极高的销售量，形成了市场的主要份额，属于畅销品类。而大多数其他菌类产品的销售量则相对较低，分散在市场的长尾部分。

### 对每一种品类随时间变化的销量分析与求解

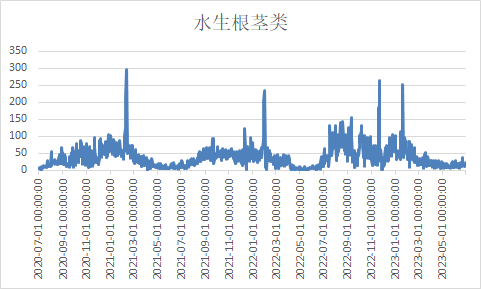


图13 水生根茎类

高峰期分析：图13可以明显看到在某些特定时间点（如2021年初、2022年中、2023年初）出现了销售量的高峰。需要进一步调查这些时间点是否有特殊活动或促销活动。 季节性模式：图13显示每年都有几个高峰，说明可能存在季节性需求。例如，夏季和春节前后可能是销售的高峰期。 周期性变化：观察图表可以看出，每年周期性地出现销售高峰和低谷，表明销售量有较强的周期性特征。整体趋势：尽管存在波动，但总体销售量在2021年到2023年期间呈现一定的稳定性，销售量没有明显的长期上升或下降趋势。

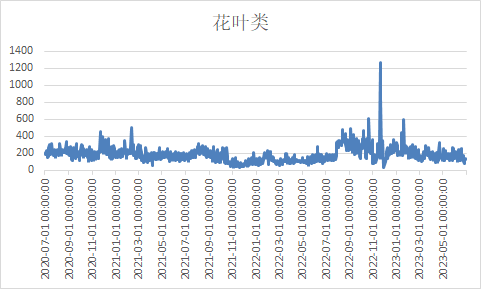


图14 花叶类

高峰期分析：图14可以明显看到在某些特定时间点（如2022年末）出现了销售量的高峰。季节性模式：图14显示没有特定高峰，说明不存在季节性需求。周期性变化：观察图表可以看出，每年周期性地没有出现销售高峰和低谷，表明销售量没有较强的周期性特征。整体趋势：尽管存在波动，但总体销售量在2020年到2022年底期间呈现一定的稳定性，销售量没有明显的长期上升或下降趋势。

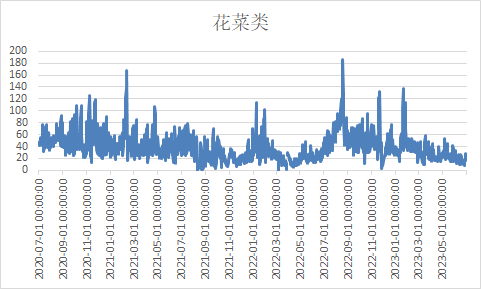


图15 花菜类

高峰期分析：图15可以明显看到在某些特定时间点（如2021年初、2022年初、2022年末、2023年初）出现了销售量的高峰。季节性模式：图15显示每年都有几个高峰，说明可能存在季节性需求。例如，夏季和春节前后可能是销售的高峰期。 周期性变化：观察图表可以看出，每年周期性地出现销售高峰和低谷，表明销售量有较强的周期性特征。整体趋势：存在波动，总体销售量在2021年到2023年期间呈现一定的波动性，销售量没有明显的长期上升或下降趋势。

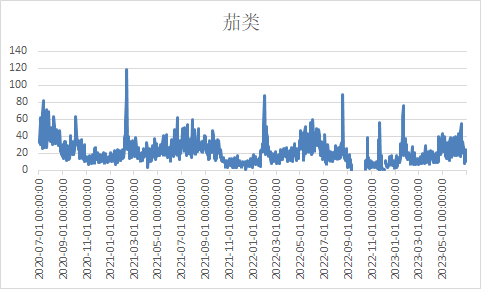


图16 茄类

高峰期分析：图16可以明显看到在某些特定时间点（如2021年初、2022年中、2023年初）出现了销售量的高峰。季节性模式：图16显示每年都有几个高峰，说明可能存在季节性需求。例如，夏季和春节前后可能是销售的高峰期。 周期性变化：观察图表可以看出，每年周期性地出现销售高峰和低谷，表明销售量有较强的周期性特征。整体趋势：尽管存在波动，但总体销售量在2021年到2023年期间呈现一定的稳定性，销售量没有明显的长期上升或下降趋势。

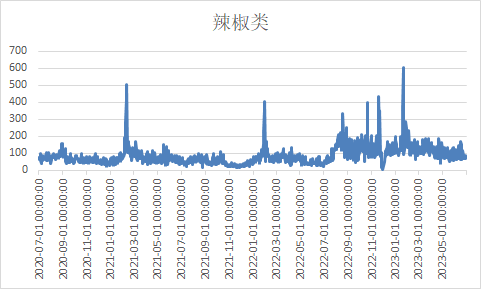


图17 辣椒类

高峰期分析：图17可以明显看到在某些特定时间点（如2021年初、2022年初、2023年初）出现了销售量的高峰。季节性模式：图17显示每年都有几个高峰，说明可能存在季节性需求。例如，夏季和春节前后可能是销售的高峰期。 周期性变化：观察图表可以看出，每年周期性地出现销售高峰和低谷，表明销售量有较强的周期性特征。整体趋势：尽管存在波动，但总体销售量在2021年到2023年期间呈现一定的稳定性，销售量没有明显的长期上升或下降趋势。

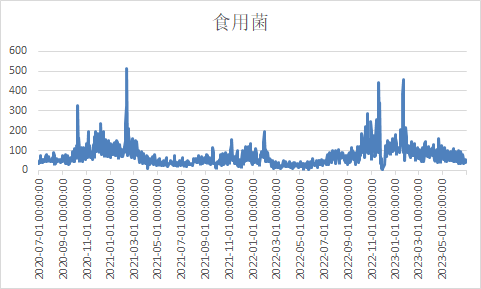


图18 食用菌类

高峰期分析：图18可以明显看到在某些特定时间点（如2020年末、2021年初、2022年末、2023年初）出现了销售量的高峰。季节性模式：图18显示每年都有几个高峰，说明可能存在季节性需求。例如，冬季和春节前后可能是销售的高峰期。 周期性变化：观察图表可以看出，每年周期性地出现销售高峰和低谷，表明销售量有较强的周期性特征。整体趋势：尽管存在波动，但总体销售量在2020年到2023年期间呈现一定的稳定性，销售量没有明显的长期上升或下降趋势。

### 对每一种单品随时间变化的销量分析与求解

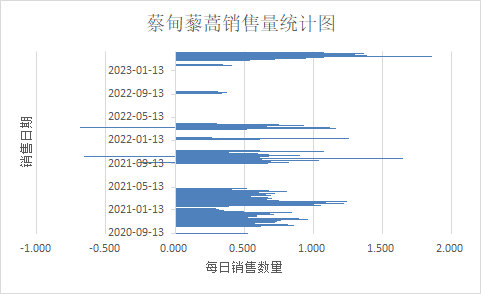
通过观察单品销售量的分布，分析了2020年7月1日——2023年6月30日，我们以70%为全年销售量为基准衡量各个单品销量。

图19 菜甸藜蒿销售量统计图

图19是菜甸藜蒿每日销售量和销售日期的关系

由图19可知，菜甸藜蒿单品销售量呈现极不稳定的趋势，在2021-05-13到2021-09-13及2022-05-13到2022-09-13以及到2023-01-13这几个月内，菜甸藜蒿是时令性蔬菜，其销售量呈现季节性变化，单品销量仅分布在一年中的某几天，整体综合销量较低，销售持续性较差，没有达到70%的标准，故排除。

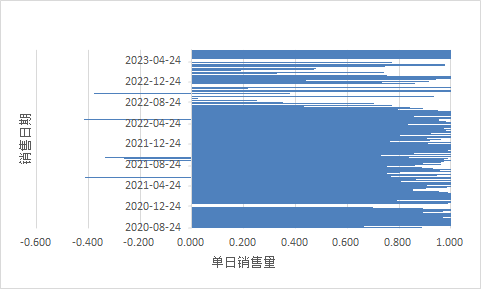


图20 奶白菜销售量统计图

图20是奶白菜每日销售量和销售日期的关系

由图20可知，奶白菜的销售量2020-08-24到2023-04-24有持续高销量，奶白菜是全年可供的蔬菜，由于其销售量最高，最能体现整体销量，因此其变化销量与总销量极为相似。经统计，共有12个单品满足全年销售量达到70%及以上的蔬菜单品，包括：云南生菜，上海青，云南油麦菜，奶白菜，黄白菜（2），西兰花，净藕（1），紫茄子（2），青茄子（1），螺丝椒，芜湖青椒（1），西峡香菇（1）。

### 品类之间的关系与规律

通过数据预处理阶段，可以得到每一种品类从2020.6.30到2023.7.1之间的销售总量，因为数据量庞大且不符合正态分布的相关规律，所以使用斯皮尔曼相关性分析的模型，在该模型中，通过判断两个变量之间的显著性（p值）来判断两个品类之间是否存在显著相关性，如果sig<=0.05,说明两者之间存在显著相关性，反之则不存在显著相关性，并且相关系数的取值范围为【-1，1】，绝对值越大说明越相关。另外，再考虑该问题的过程中，由于数据量过于庞大，所以忽略该问题中存在的极个别的成本过低，损耗率过大，销售量为负等极个别特殊情况。

**模型原理：**

****

 （1）

斯皮尔曼相关性分析（Spearman Rank Correlation）是一种非参数统计方法，用于衡量两个变量之间的相关性。与皮尔逊相关系数不同，斯皮尔曼相关系数不要求数据服从正态分布或线性关系。它适用于衡量单调关系，即当一个变量增加时，另一个变量是持续增加还是减少，不论其具体形式如何。

**模型应用**：

表1 品类间斯皮尔曼相关性分析表

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | 食用菌 | 水生根茎类 | 花叶类 | 花菜类 | 茄类 | 辣椒类 |
| 斯皮尔曼Rho | 食用菌 | 相关系数 | 1.000 | .605\*\* | .596\*\* | .461\*\* | -.090\*\* | .535\*\* |
|  |  | 显著性（双尾） |  | <.001 | <.001 | <.001 | .003 | <.001 |
|  | 水生根茎类 | 相关系数 | .605\*\* | 1.000 | .439\*\* | .395\*\* | -.193\*\* | .334\*\* |
|  |  | 显著性（双尾） | <.001 |  | <.001 | <.001 | <.001 | <.001 |
|  | 花叶类 | 相关系数 | .596\*\* | .439\*\* | 1.000 | .633\*\* | .326\*\* | .595\*\* |
|  |  | 显著性（双尾） | <.001 | <.001 |  | <.001 | <.001 | <.001 |
|  | 花菜类 | 相关系数 | .461\*\* | .395\*\* | .633\*\* | 1.000 | .235\*\* | .430\*\* |
|  |  | 显著性（双尾） | <.001 | <.001 | <.001 |  | <.001 | <.001 |
|  | 茄类 | 相关系数 | -.090\*\* | -.193\*\* | .326\*\* | .235\*\* | 1.000 | .153\*\* |
|  |  | 显著性（双尾） | .003 | <.001 | <.001 | <.001 |  | <.001 |
|  | 辣椒类 | 相关系数 | .535\*\* | .334\*\* | .595\*\* | .430\*\* | .153\*\* | 1.000 |
|  |  | 显著性（双尾） | <.001 | <.001 | <.001 | <.001 | <.001 |  |

\*\*在0.01级别（双尾），相关性显著

**品类之间**的相关性分析：

由表1可知，食用菌类、水生根茎类、花叶类、花菜类和辣椒类与茄类的相关性最差，大部分与茄类的相关性为负相关，说明这几类的销售总量会轻微降低茄类的销售总量，这五类与茄类之间可能存在轻微的相互竞争关系。反之，水生根茎类与食用菌类相关性最强，花叶类与花菜类相关性最强，茄类与花叶类相关性最强，辣椒类与花叶类相关性最强，每两个类之间的销售与其他品类的蔬菜存在相互促进的关系，说明每个类与其它品类之间可能会存在捆绑销售或者交叉销售的销售手段，以此来促进商超蔬菜的销售总量。

### 单品之间的销量与规律

通过数据预处理阶段，可以得到每一种单品在2020.6.30到2023.7.1之间的销售总量，由此便可对数据进行标准化处理，经过标准化处理后，得到246个有效数据，由此便可进行K-means聚类分析，将蔬菜单品之间分为**季节性蔬菜，时令性蔬菜，持续供应型蔬菜，间断供应型蔬菜**，将具有相关性的蔬菜分为一类，由此便可得到单品之间的相关性。另外，在数据预处理阶段，经过数据清洗，已经排除异常数据与缺失数据所带来的影响，由此便可用相关模型进行分析。

**模型原理：**

 （2）

K-means聚类算法是一种基于质心的分割方法，用于将数据集划分为K个簇，使得簇内数据点到簇中心的距离平方和最小。

**模型应用：**

**表2 每个聚类中的个案数**

|  |  |  |
| --- | --- | --- |
|  | 每个聚类中的个案数 |  |
| **聚类** | **1** | **12.000** |
|  | **2** | **1.000** |
|  | **3** | **2.000** |
|  | **4** | **231.000** |
| **有效** |  | **246.000** |
| **缺失** |  | **.000** |

通过K-means聚类分析对单品之间的数量（246种）进行处理和统计。发现共有231种蔬菜属于持续供应型蔬菜，也就是全年超过70%以上的时间都可以进行采购的蔬菜类型。有12种蔬菜属于季节性供应型蔬菜，也就是在某些特殊的时令，该种类型的蔬菜会出现大幅度变化与波动，这可能跟该类型蔬菜在特定时节的生产数量或者人们的需求有关，有2种蔬菜，属于间断供应型蔬菜，经调查发现，该种类型属于蔬菜套装类型，价格相对普通蔬菜来说可能相对昂贵，经推测，可能是在节假日等各别时间内，该类型的蔬菜销量会有销量激增现象的出现。但全年大部分时间，销量显著低于其它蔬菜类型。还有1种蔬菜，全年销量都平稳性偏低，可能原因是该蔬菜类型并不受到大众喜爱，只针对个别人群。

表3 ANOVA数据统计

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **聚类** |  | **误差** |  |  |  |
|  | 均方 | 自由度 | 均方 | 自由度 | F | 显著性 |
| Zscore  （数量） | 73.781 | 3 | 0.098 | 242 | 754.702 | <0.001 |

通过ANOVA数据分析，发现显著性<0.001,说明蔬菜单品之间存在显著性关系

详细过程见附录

## 问题二的模型建立与求解

### 销售总量与成本加成定价的关系的模型建立

**P1=C+（C×A）** （3）

由于销售总价已知，总成本已知，可以算出加成比例，从而可以分析出销售总量与成本加成定价的关系。根据附录2的信息，

P1=∑（D×P2） （4）

C=∑（DP/（1-S）×C1 ） （5）

由此便可以算出在某个时间段内，该超市的水果的总的加成比例，由此，便可通过皮尔逊系数分析出不同单品对加成比例的相关影响。

### 销售总量与成本加成定价的关系的求解

时间序列预测模型（ARIMA）：

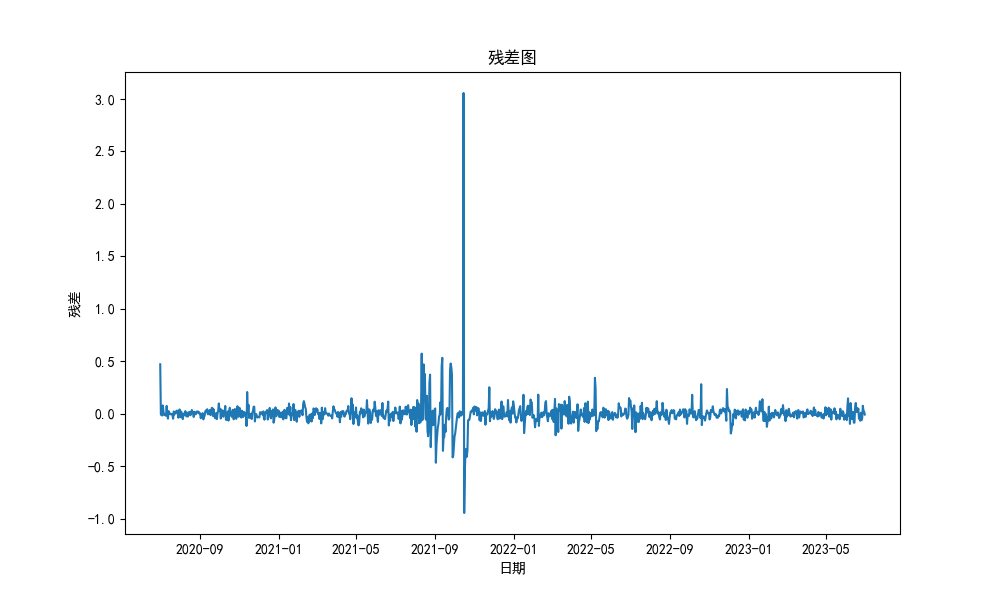


图21 花菜类残差图

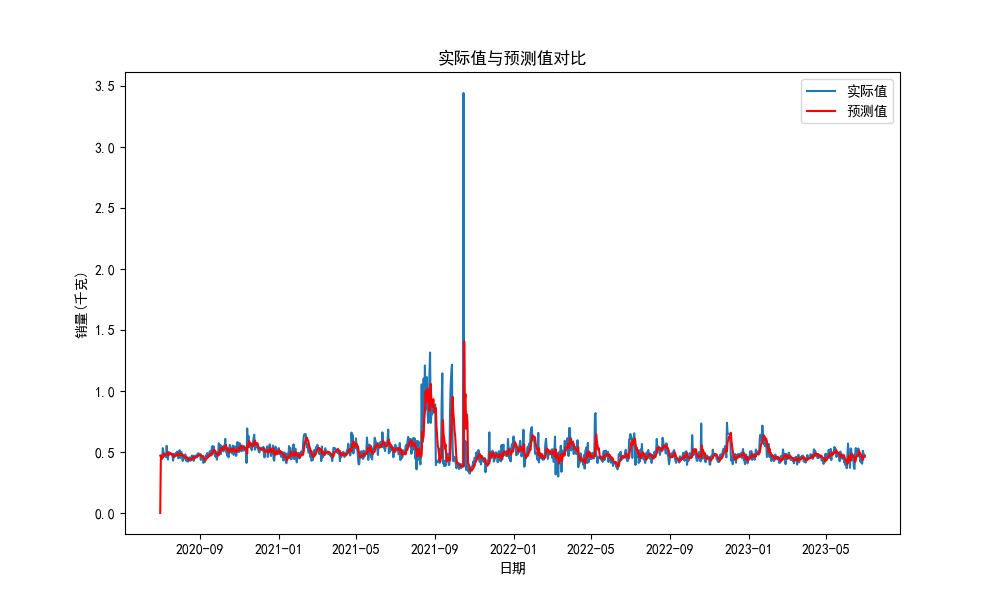


图22 菜花类实际值与预测值对比

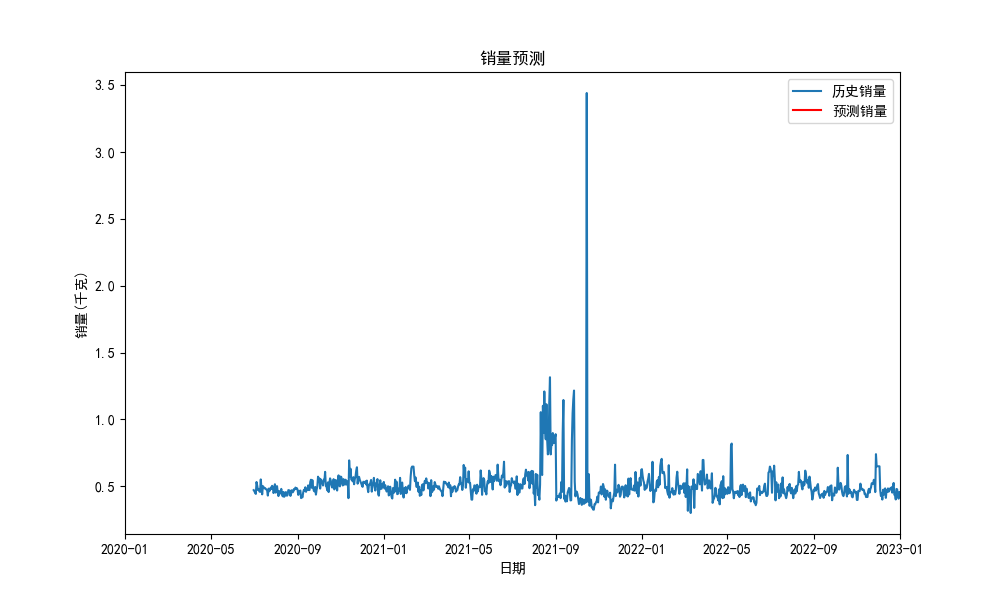


图23 花菜类销量预测

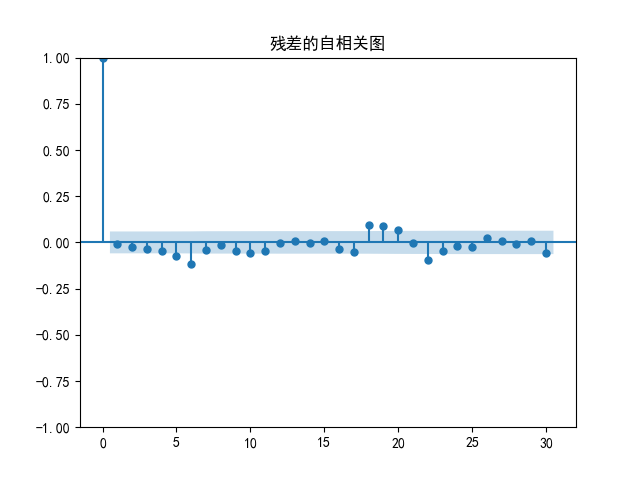


图24 菜花类残差的自相关图

残差在-0.5~0.5范围内，说明模拟效果好；预测量与实际值趋势相吻合，说明模拟效果好

通过数据预处理阶段，将同一种商品的所有销售信息（包括进价、销量、销售单价、销售日期）汇集在一个表格中，由此便可对数据进行标准化处理，表格大致如下：

表4 单品销售信息

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **单品编码** | **单品名称** | **分类编码** | **分类名称** | **销售日期** | **销售单价(元/千克)** | **销量(千克)** | **批发价格(元/千克)** |
| 1.029E+14 | 西兰花 | 1011010201 | 花菜类 | 2020-07-01 00:00:00 | 14 | 34.554 | 9.23 |
| 1.029E+14 | 西兰花 | 1011010201 | 花菜类 | 2020-07-02 00:00:00 | 14 | 30.109 | 8.9 |
| 1.029E+14 | 西兰花 | 1011010201 | 花菜类 | 2020-07-03 00:00:00 | 14 | 26.46 | 8.97 |

其余表格见附录

经过标准化处理，我们排除了一些无关数据（销量为负），由此可进行价格弹性模型测试，价格弹性（Price Elasticity of Demand, PED）定义为需求量的百分比变化相对于价格百分比变化的比率。将蔬菜按品类计算进价，销售价格，销售量，并通过线性回归预测需求量（ΔQ），价格变化量（ΔP）。带入价格弹性计算公式，得出每个单品的价格弹性系数，并将属于该品类的单品的价格弹性系数相加求平均，得到每个品类的价格弹性系数

**模型原理**：

 （6）

表5 品类的价格弹性系数

|  |  |
| --- | --- |
| 品类名称 | 价格弹性系数 |
| 水生根茎类 | 2.6020398419978026 |
| 花叶类 | 2.349510662125202 |
| 花菜类 | 3.951932817678636 |
| 茄类 | 2.803663155572233 |
| 辣椒类 | 2.646429700689547 |
| 食用菌类 | 1.8651139789138345 |

完全无弹性需求（PED = 0）：需求量对价格变化没有反应。

无弹性需求（0 < PED < 1）：需求量对价格变化反应不大。

单位弹性需求（PED = 1）：需求量对价格变化的百分比反应与价格变化的百分比相同。

弹性需求（PED > 1）：需求量对价格变化反应较大。

完全弹性需求（PED = ∞）：价格微小变化会导致需求量极大变化。

由表5得出，品类价格弹性系数在1.80-4.0之间，说明品类的需求量对价格反应需求量比较大，根据成本加成定价公式可得出，由于销售价格和加成比例成线性关系，所以加成比例和需求量有明显关系。

### 各蔬菜品类未来一周(2023 年 7 月 1-7 日)的日补货总量和定价策略， 使得商超收益最大的模型建立

最大模型的建立优化算法：BFGS优化算法（Broyden–Fletcher–Goldfarb–Shanno algorithm）是一种拟牛顿法，用于解决无约束优化问题。该算法通过构建和更新Hessian矩阵的近似值来迭代地找到目标函数的极小值。

计算当前点的梯度

 （7）

检查收敛条件

 如果满足条件，则停止迭代 (8)

计算搜索方向

 (9)

进行线搜索，确定步长 αk使得

最小 （10）

更新位置

 （11）

计算新的梯度

 （12）

计算差值

 （13）

步长差值

 （14）

更新Hessian矩阵的逆矩阵 Hk+1：

 （15）

我们选择这个公式是因为数据的数值范围变化波动较大，受无关变量影响较多，使用该方法梯度下降较慢，可以最大程度寻找全局最优解，避免了无关变量的影响。

首先运用价格弹性模型的预测方案

**模型原理：**

 （16）

 （17）

 （18）

通过这三个公式可得出最佳定价策略和补货策略

### 各蔬菜品类未来一周(2023 年 7 月 1-7 日)的日补货总量和定价策略， 使得商超收益最大的问题求解

### 表6 水生根茎类一的日补货总量和定价策略

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **品类** | **预测销量(千克)** | **初始价格(元)** | **优化价格(元)** | **最大收益(元)** | **天数** |
| 水生根茎类 | 0.385502108 | 12 | 15 | 9.379484398 | 1 |
| 水生根茎类 | 0.387265569 | 12 | 15 | 9.379484398 | 2 |
| 水生根茎类 | 0.371563218 | 12 | 15 | 9.379484398 | 3 |
| 水生根茎类 | 0.388497031 | 12 | 15 | 9.379484398 | 4 |
| 水生根茎类 | 0.384341465 | 12 | 15 | 9.379484398 | 5 |
| 水生根茎类 | 0.379197728 | 12 | 15 | 9.379484398 | 6 |
| 水生根茎类 | 0.383485566 | 12 | 15 | 9.379484398 | 7 |

剩余品类补货策略和定价策略见附录

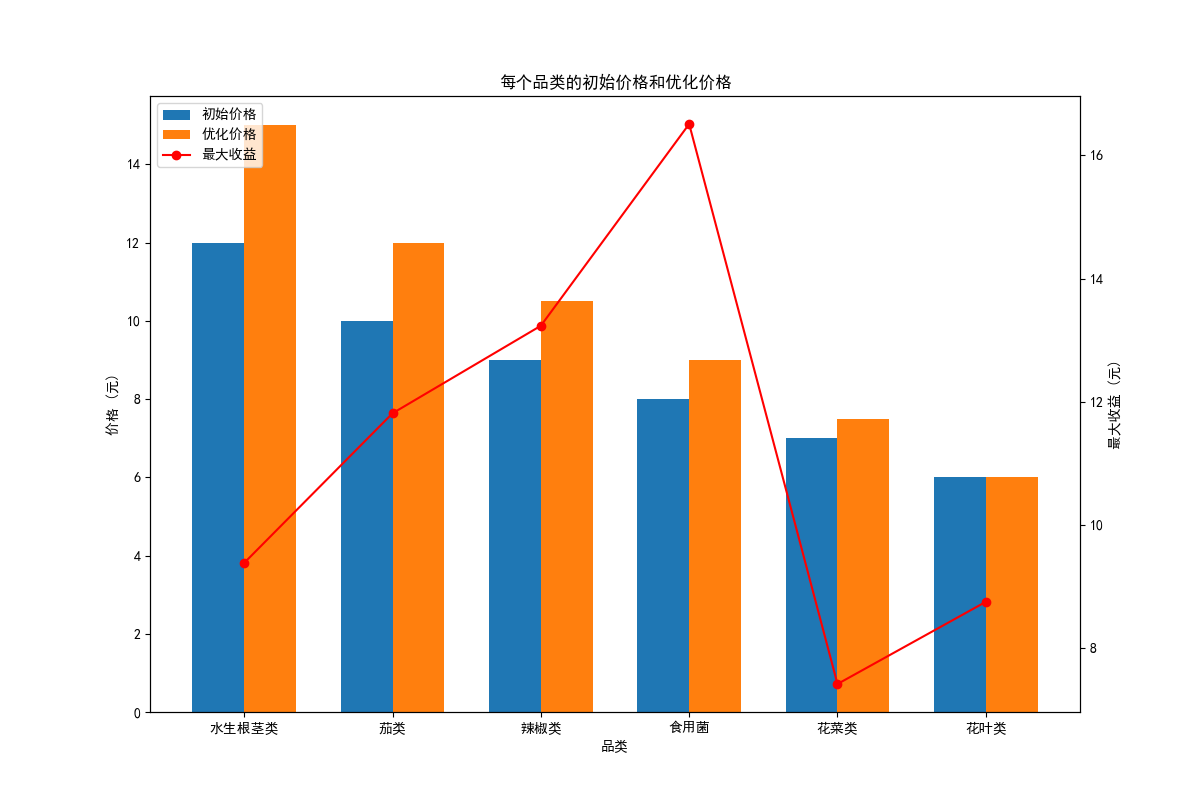


图21 每个品类的初始价格和优化价格

通过对每个品种每个单品分别求解其价格弹性系数，可以得出成本定价与弹性系数之间的关系，然后通过价格弹性系数得出价格与销量的线性回归方程和弹性系数模型预测。

## 问题三的模型建立与求解

### 问题三的模型建立

首先筛选出各单品在6.24到6.30期间订购量满足最小陈列量 2.5 千克的单品，然后对筛选后的数据按照在此期间所获利润的多少进行排序

L+=G（D×P2-C2×J） （19）

J=D/(1-S) （20）

排序完成后筛选出该周利润最多的前30个蔬菜单品：

**上海青、青茄子（1）、奶白菜、木耳菜、螺丝椒、菱角、净藕（1）、金针菇（盒）、洪湖藕带、红薯尖、红莲藕带、海鲜菇（包）、高瓜（1）、菜心、菠菜、菠菜（份）、白玉菇、紫茄子（2）、竹叶菜、长线茄、云南油麦菜、云南油麦菜（份）、云南生菜（份）、小青菜（1）、鲜木耳（份）、西兰花、芜湖青椒（1）、外地茼蒿、娃娃菜。**

因为要预测的是各单品未来一周每天的销售价格，销售总量，以及成本，所以对样本数据进行灰色预测模型处理，首先，分别对每个单品进行原始数据的累加处理，生成新序列

 （21）

然后对累加数列进行平滑处理，生成新的累序列，

 （22）

其中权重，通常均值序列为0.5。

然后建立新的灰微分方程

 （23）

使用GM（1，1）模型得到相应的白化微分方程

 （24）

将灰微分方程移项，得到

 （25）

之后由最小二乘法得到相关矩阵

 （26）

确定参数

 （27）

由此得到a , b，带入白化微分方程得到通解

 (28)

算出预测函数

 (29)

### 问题三模型的求解

使用灰色预测GM（1，1）模型与模拟退火模型优化算法。设定目标函数 f(x)，其中 x 是解的候选解。目标是找到使 f(x) 最小（或最大）的解，用Matlab对每个单品中的销售价格，销量进行单预测扩展，然后使用退火模型，结合预测值，求出全局最优解，通过修正模型的边界条件，退货速度等参数，得到相对较好的模拟结果(代码见附录)，以此来预测未来一周的销量和销售价格以及商超获得的最大利润。

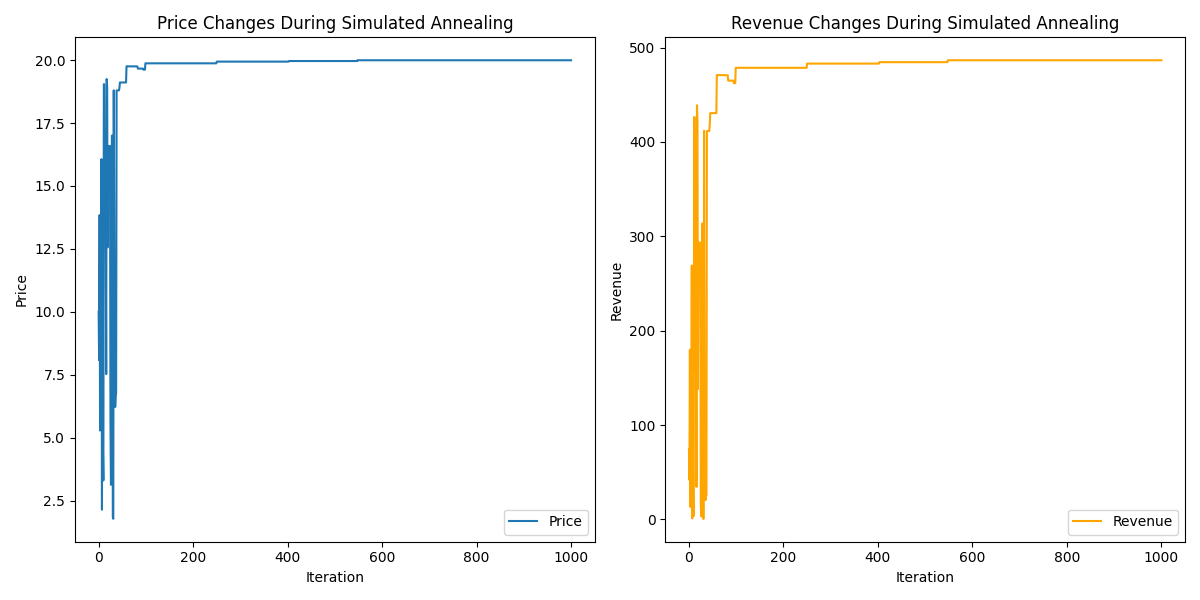
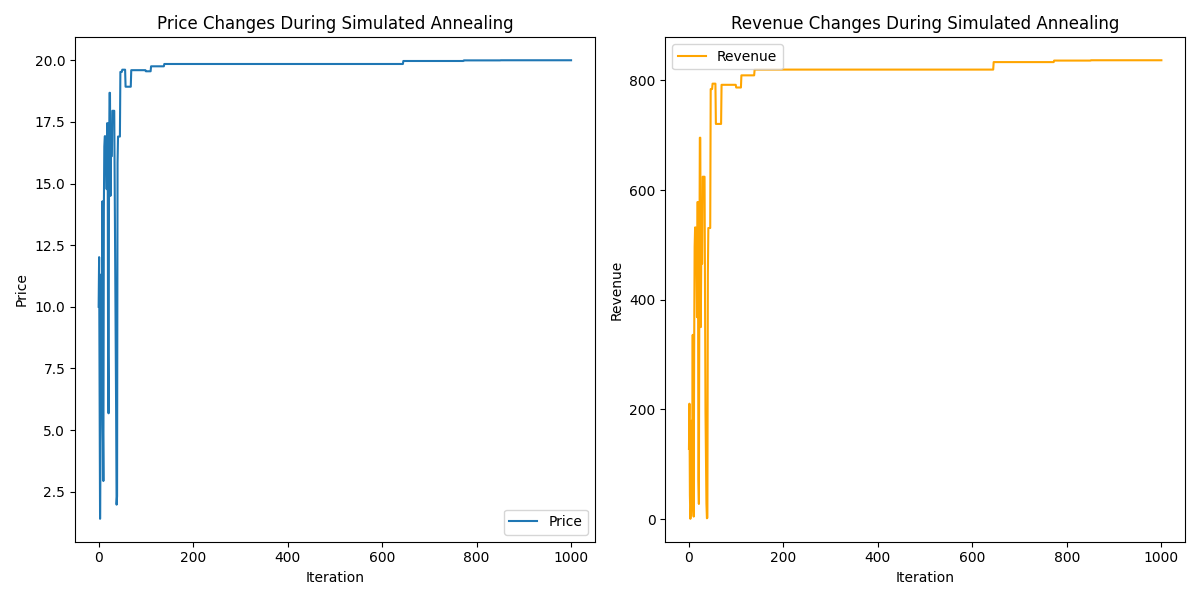
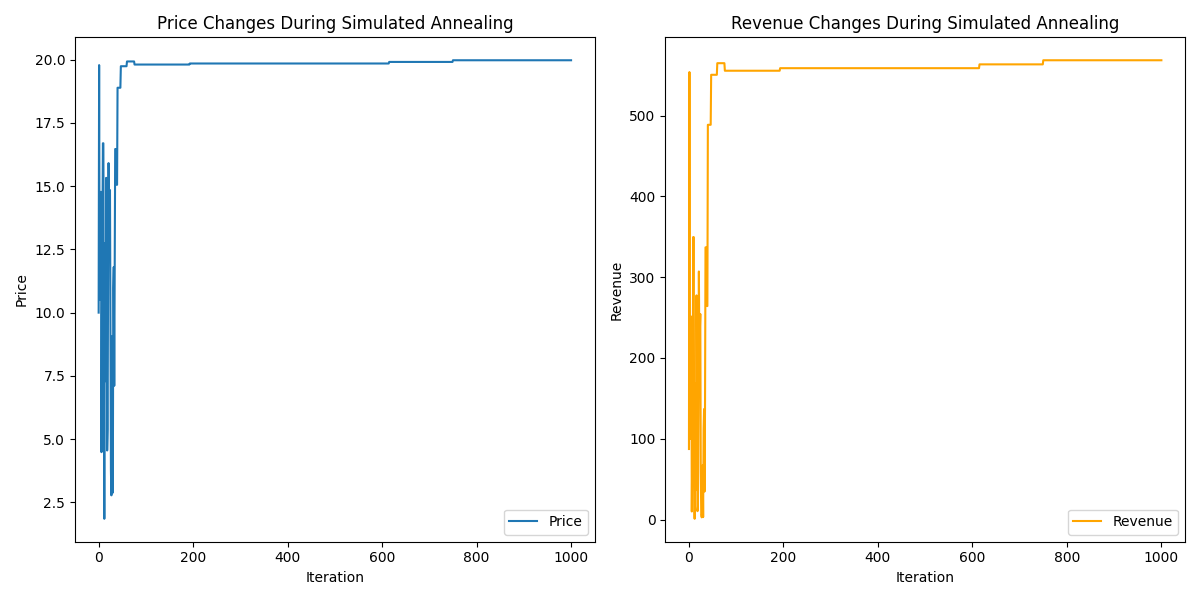
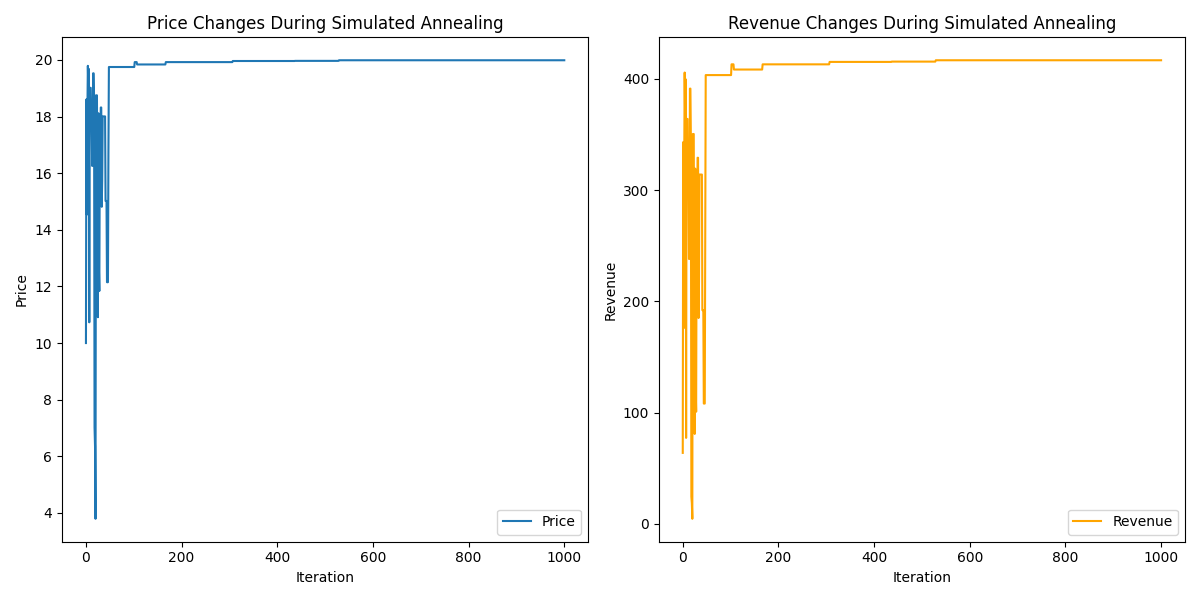
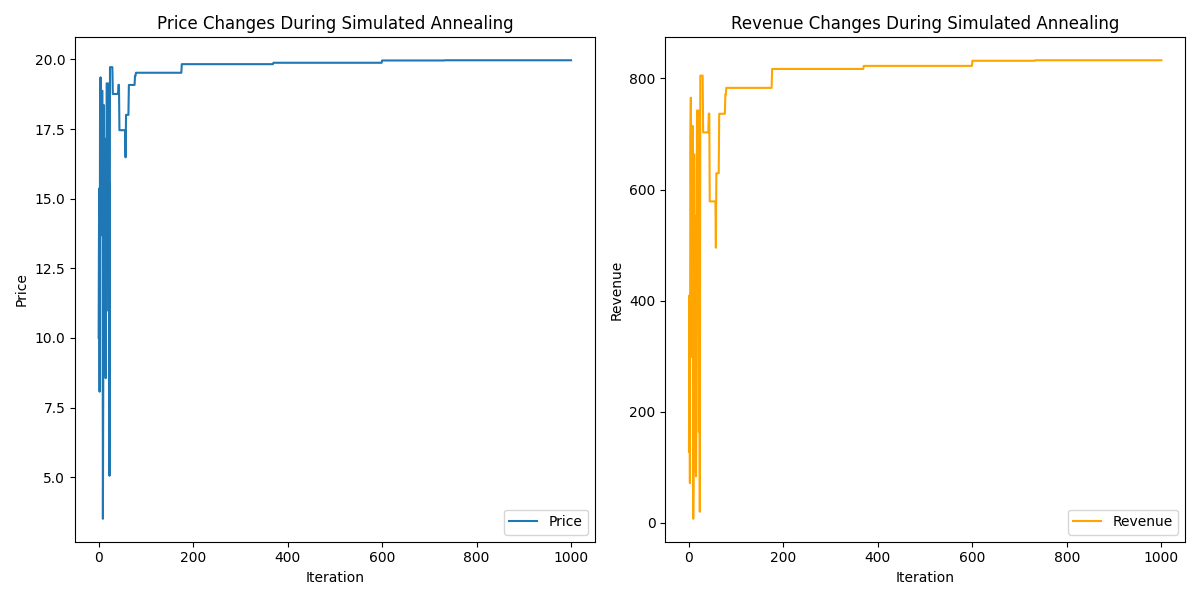
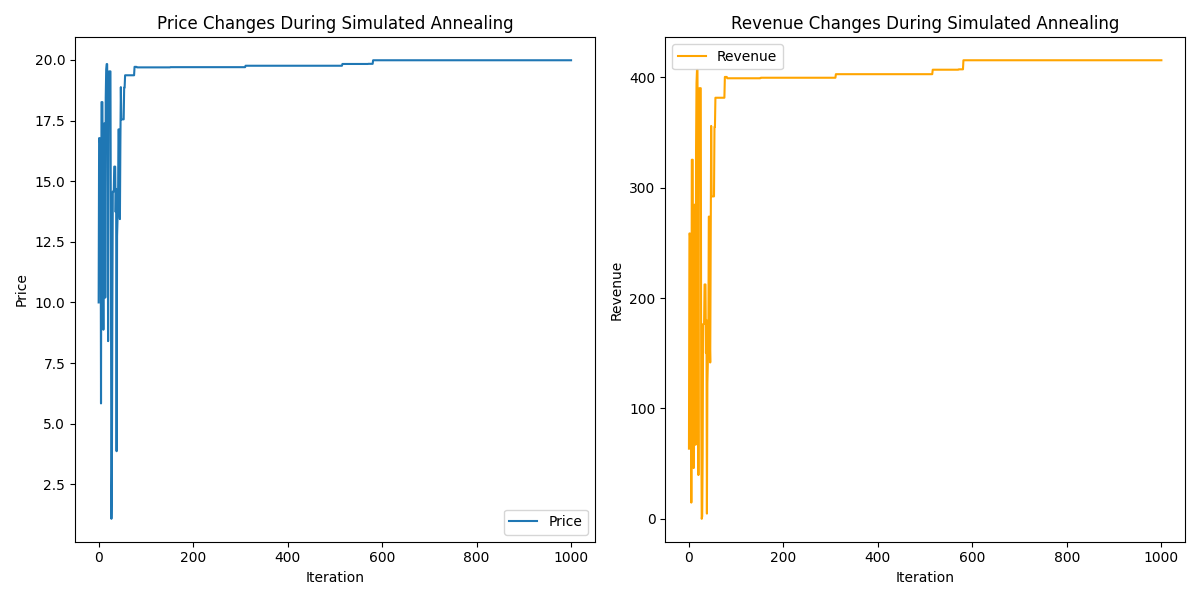
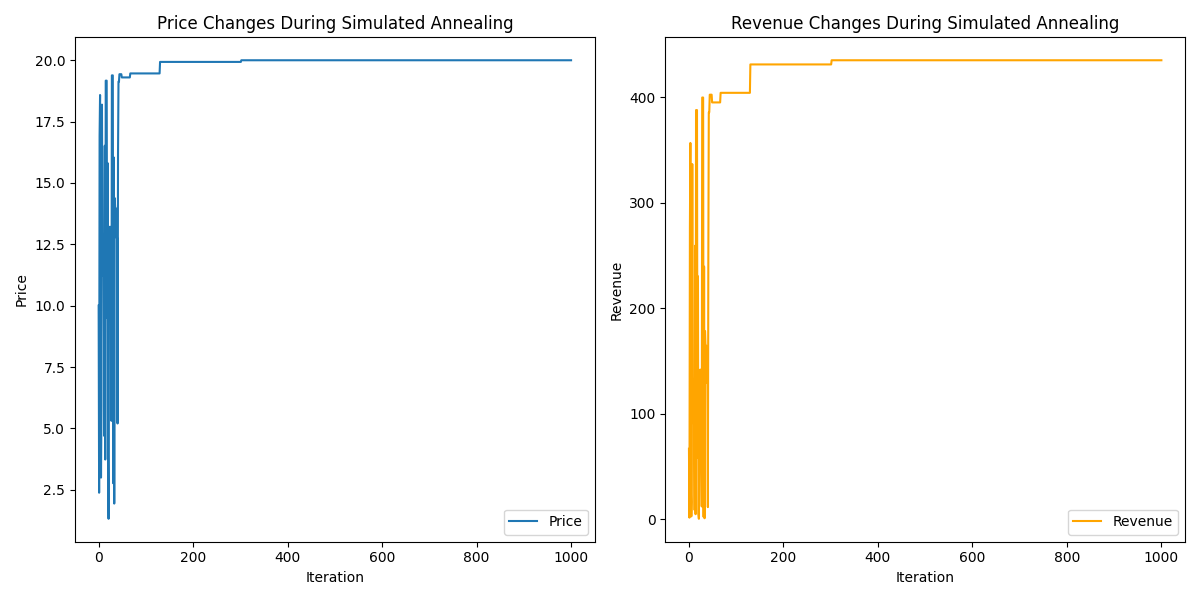
由于数据量过于庞大，这里仅展示十个未来预测结果，其余请见附录。

表8 7月1日商品补货，定价表。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **海鲜菇(包)** | **金针菇(盒)** | **洪湖藕带** | **红椒(2)** | **枝江青梗散花** |
| **最佳定价** | 19.99664023 | 19.99995714 | 19.9995054 | 19.99924344 | 19.9990693 |
| **补货量** | 8.69 | 10.4 | 3.82 | 9.48 | 3.21 |
| **最佳利润** | 868.09932 | 4112.222653 | 558.5480556 | 1702.139123 | 2022.180138 |

该表格显示不同单品在7月1日的补货量和定价策略。

下图针对于海鲜菇（包装）的多次预测结果生成的模拟退火模型图像。



*通过图像的拟合系数，选取最好的模拟结果，以及其对应的最大商超收益和价格参数。*

## 问题四的模型建立与求解

要更好地制定蔬菜商品的补货和定价决策，商超需要采集以下相关数据，这些数据对于解决上述问题至关重要：

1. 客户购买行为数据

数据内容：

客户购买频次、时间段和偏好

客户对不同蔬菜品种的购买量

客户对折扣、促销的敏感度

帮助：

预测需求：了解客户的购买习惯和偏好，可以帮助更准确地预测每天的需求。

定价策略：客户对价格的敏感度可以指导如何进行成本加成定价以及折扣策略的制定。

2. 季节性需求变化数据

数据内容：

不同季节各蔬菜品种的销售量变化

节假日、特殊活动期间的销售变化

帮助：

补货策略：季节性变化数据有助于制定季节性补货计划，避免过量库存或短缺。

定价策略：在需求高峰期适当调整价格，优化收益。

3. 天气数据

数据内容：

每日的气温、降雨量、天气类型（晴天、阴天、雨天等）

帮助：

需求预测：天气对蔬菜需求有显著影响（如夏季高温对绿叶蔬菜需求增加）。

供应链管理：天气状况可以影响蔬菜的供应和运输计划。

4. 竞争对手数据

数据内容：

竞争对手的价格、促销活动

竞争对手的品类和库存情况

帮助：

定价策略：了解市场竞争态势，调整价格以保持竞争力。

补货策略：分析竞争对手的补货策略，优化自身的补货决策。

5. 物流和运输数据

数据内容：

运输时间、费用

运输过程中的损耗率

各供应商的准时交货率

帮助：

补货决策：考虑运输时间和损耗率，优化补货时间和数量。

成本控制：优化物流和运输策略，减少运输成本和损耗。

6. 实时库存数据

数据内容：

每个蔬菜品类的实时库存量

各蔬菜品类的保质期信息

帮助：

补货策略：实时监控库存情况，及时补货，避免库存不足或过量。

损耗管理：根据保质期信息，优先销售即将过期的蔬菜，减少损耗。

7. 消费者反馈数据

数据内容：

消费者对蔬菜质量、价格、服务的评价和反馈

消费者对不同品种蔬菜的需求建议

帮助：

质量控制：根据消费者反馈，提高蔬菜的质量和服务水平。

产品优化：根据消费者需求建议，调整蔬菜品种组合，满足市场需求。

# 模型的评价、改进与推广

## 模型的评价

1. 线性回归模拟各蔬菜品类销售量与成本加成定价的关系，可直观显示各蔬菜品类销售量的变化可引起一定量的成本加成定价的变化。
2. 模拟退火算法的全局优化能力强：能够跳出局部最优，找到全局最优解。简单易实现**：**算法相对简单，容易编程实现。适用范围广：可用于各种优化问题，包括离散和连续问题。
3. K-means模型简单易用，高效性，可伸缩性，可解释性，适应性强。
4. 灰色预测模型数据要求低，处理非线性系统，适用范围广，计算简单，动态预测。
5. 价格弹性模型的优点有定量分析，决策支持，市场理解，预测能力，灵活性。
6. 时间预测序列模有型捕捉趋势和季节性，自动更新，多变量支持，时间依赖性，灵活性，预测区。

## 模型的改进

1. 模拟退火算法计算时间长，参数选择困难。

## 模型的推广

模拟退火算法可以**自动调参，**使用机器学习等方法自动调节模拟退火算法的参数，例如初始温度、冷却速度等，以提高算法性能。

# 参考文献

[1] 顾思弘. 考虑新鲜度变化的H零售商生鲜产品动态定价研究[D].上海:东华大学,2023.

[2] 邹耀斌,齐慧康,孙水发.皮尔逊相关性最大化导向的自动阈值分割方法[J].电子测量技术,2023,46(17):109-117.

[3] 梁晶晶,魏乾.一种船舶历史活动热力图实时生成策略[J/OL].指挥控制与仿真,1-7[2023-09-10].

[4] 黄杏丹. 基于时间序列与深度学习模型的股票价格指数预测研究[D].济南:山东工商学院,2023.

[5] 王晓瑞,刘元,王彤.基于分数整合自回归移动平均模型的山西省手足口病预测研究[J].疾病监测,2023,38(7):865-

871.

[6] 史今驰. 背包问题的实用求解算法研究[D].济南:山东大学,2005

# 附录

问题一单品之间的分析：

最终聚类中心

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | **聚类** |  |  |
|  | 1 | 2 | 3 | 4 |
| Zscore（数量） | 2.21181 | 9.23621 | 5.81378 | -.20522 |

最终聚类中心之间的距离

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 聚类 | 1 | 2 | 3 | 4 |
| 1 |  | 7.024 | 3.602 | 2.417 |
| 2 | 7.024 |  | 3.422 | 9.441 |
| 3 | 3.602 | 3.422 |  | 6.019 |
| 4 | 2.417 | 9.441 | 6.019 |  |

初始聚类中心

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | **初始聚类中心** |  |  |
|  | 1 | 2 | 3 | 4 |
| Zscore（数量） | 2.39278 | 9.23621 | 6.19077 | -.35751 |

描述统计

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **N** | **最小值** | **最大值** | **平均值** | **标准差** |
| 数量 | 246 | 3.500000000 | 854450.1000 | 31844.46179 | 89063.09239 |
| 有效个案数（成例） | 246 |  |  |  |  |

补货策略和定价策略：

食用菌补货策略和定价策略

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **天数** | **最大收益** | **最佳订货量** | **最佳价格** |
| **0** | 1 | 63.76041371 | 7.871615232 | 4.703279002 |
| **1** | 2 | 56.61997646 | 9.10392727 | 3.764104914 |
| **2** | 3 | 92.40240515 | 10.6167684 | 4.534805989 |
| **3** | 4 | 66.74117571 | 7.868301152 | 5.283577119 |
| **4** | 5 | 64.31741418 | 8.777378703 | 4.959911044 |
| **5** | 6 | 66.011421 | 11.12891445 | 4.802181899 |
| **6** | 7 | 62.12555568 | 10.08534514 | 3.954868904 |

水生根茎类补货策略和定价策略

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **天数** | **最大收益** | **最佳订货量** | **最佳价格** |
| **0** | 1 | 58.39691971 | 17.30973337 | 5.044844134 |
| **1** | 2 | 52.75581198 | 13.15815083 | 5.137059702 |
| **2** | 3 | 77.04200774 | 13.13934419 | 5.44344732 |
| **3** | 4 | 114.3457746 | 23.37884678 | 6.141951799 |
| **4** | 5 | 66.29569521 | 19.22715316 | 5.249750912 |
| **5** | 6 | 96.68470445 | 15.61425267 | 5.940900058 |
| **6** | 7 | 97.06461811 | 24.83873944 | 5.048058734 |

花叶类

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **天数** | **最大收益** | **最佳订货量** | **最佳价格** |
| **0** | 1 | 57.64018276 | 15.94751043 | 3.794415886 |
| **1** | 2 | 57.39093241 | 15.16871725 | 4.358653447 |
| **2** | 3 | 55.03987524 | 15.84677347 | 4.302717516 |
| **3** | 4 | 72.384485 | 19.37602381 | 3.838585421 |
| **4** | 5 | 74.87616722 | 21.47991106 | 3.986606362 |
| **5** | 6 | 64.73190328 | 18.60178098 | 4.305691083 |
| **6** | 7 | 68.55526458 | 17.97377725 | 4.090994428 |

花菜类

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **天数** | **最大收益** | **最佳订货量** | **最佳价格** |
| **0** | 1 | 355.9806708 | 49.14329341 | 4.977098188 |
| **1** | 2 | 141.8056474 | 33.52340075 | 3.601952481 |
| **2** | 3 | 366.3184761 | 79.49308973 | 4.993419856 |
| **3** | 4 | 405.7950939 | 78.29108673 | 6.109436882 |
| **4** | 5 | 132.9393036 | 26.49297472 | 4.933486672 |
| **5** | 6 | 182.4109283 | 41.18154734 | 4.024513845 |
| **6** | 7 | 354.2534079 | 46.42500404 | 6.987969537 |

问题三灰色预测模型的代码：

clc

clear

X0 = [];%原始数据

n = length(X0);

%累加数据

x1=zeros(1,n);

x1(1)=X0(1);

for i = 2:n

x1(i)= X0(i)+x1(i-1);

end

%进行光滑,权值取0.4

af=0.5;

z1=zeros(1,n);

z1(1)=0;

for i = 2:n

z1(i)=x1(i)\*af+ (1-af)\*x1(i-1);

end

%构造方程矩阵,X为Z1(K)与1组成的2列矩阵,Y为一介列矩阵,B为a,b的解矩阵.

%由最小二乘法可以知道,最佳解为B= (X(XT)^(-1)(XT)Y

Y=zeros(n-1,1);X=zeros(n-1,2);

for i= 2:n

Y(i-1)=X0(i);

X(i-1)=-z1(i);

X(i-1,2)=1;

end

%求解参数矩阵,inv为求逆运算

B = inv(X'\*X)\*X'\*Y;

a = B(1);

b = B(2);

%求解n+1年的式子为x0(n+1) = (x0(1)-b/a)\*exp(-a\*n)\*(1-e^a);

pred\_n\_1 = (X0(1)-b/a)\*exp(-a\*n)\*(1-exp(a));

x = 0:0.1:4;

y3=(X0(1)-b/a).\*exp(-a.\*x).\*(1-exp(a));

plot(x,y3,"\*")

hold on

n0 = 0:6;

plot(n0,X0,'ro')

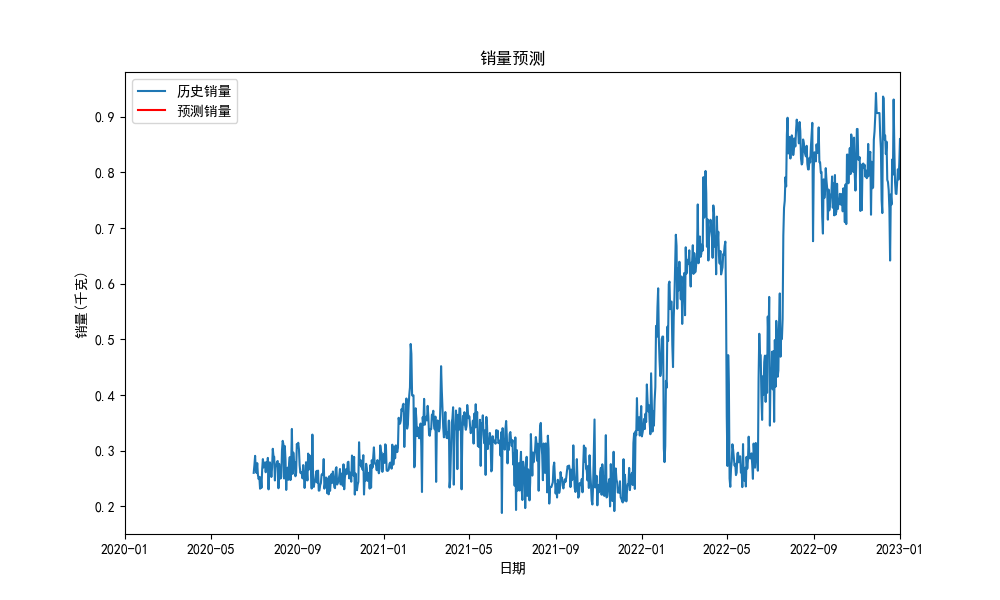
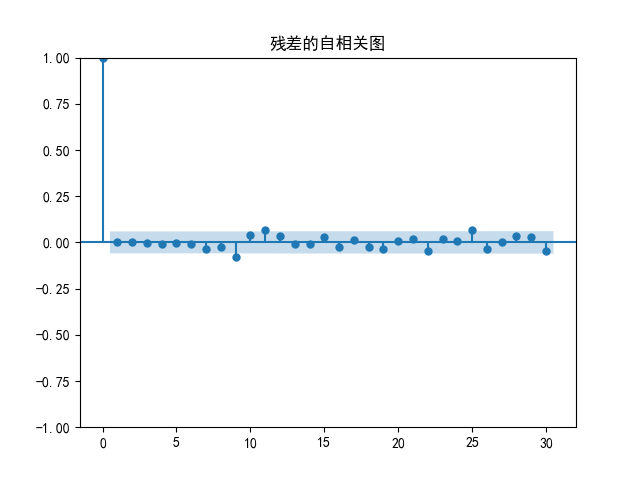
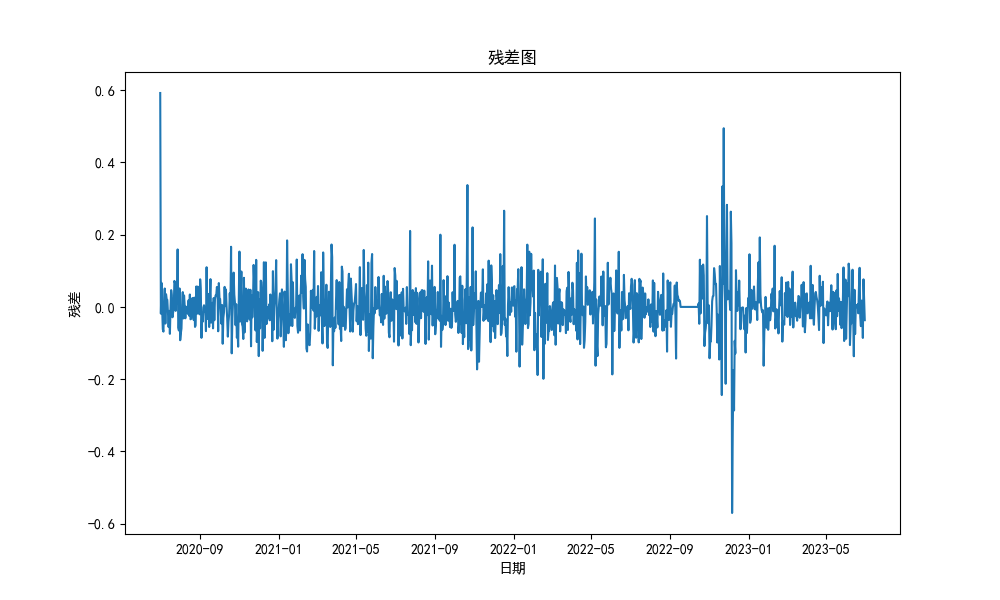
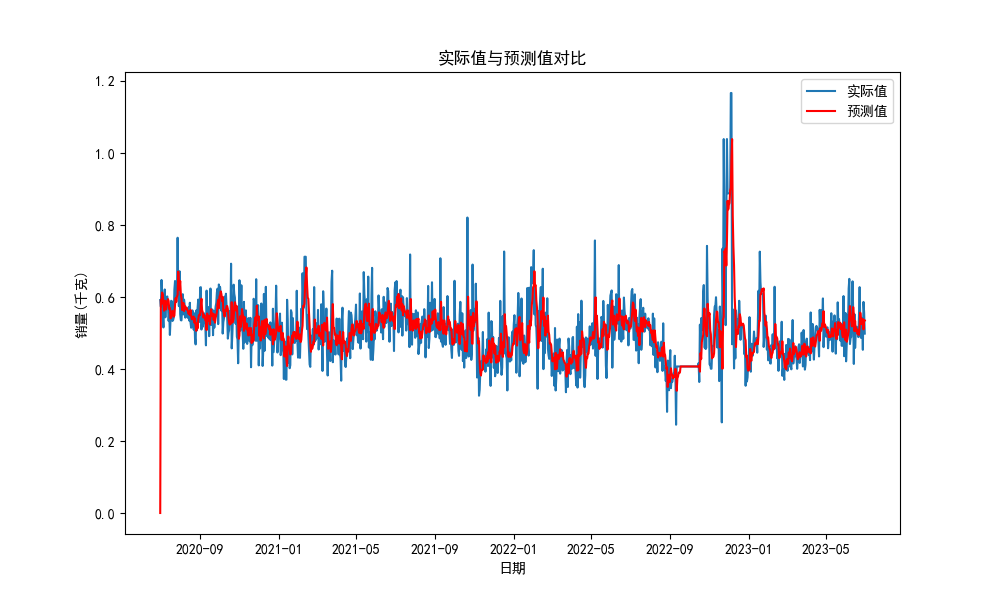
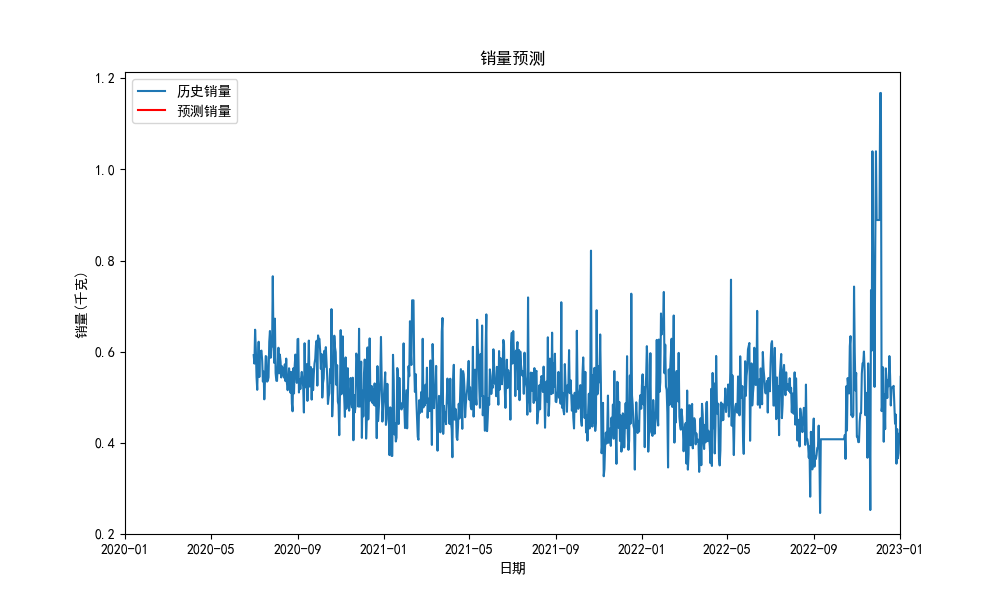
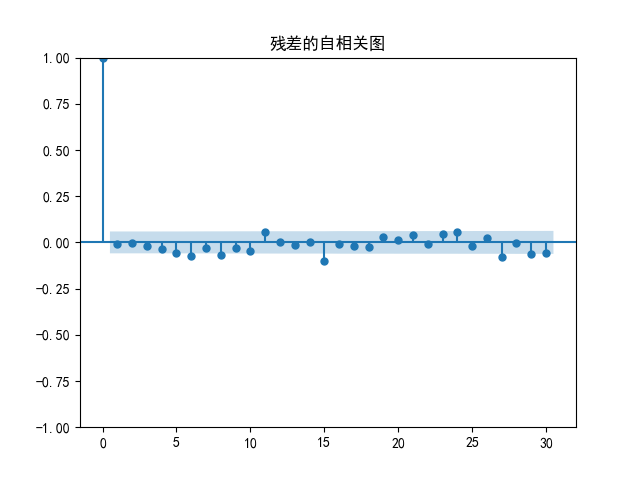
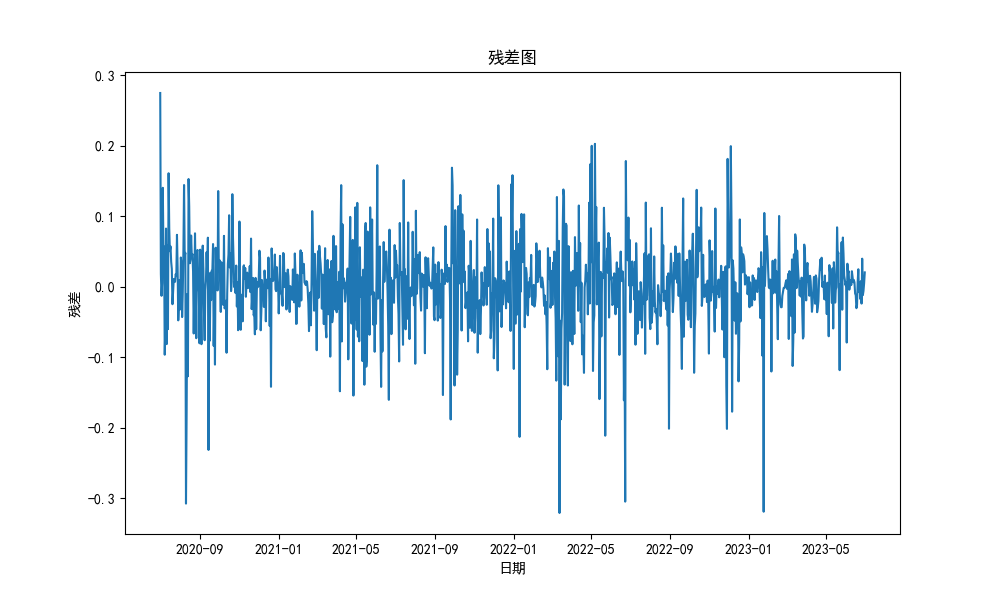
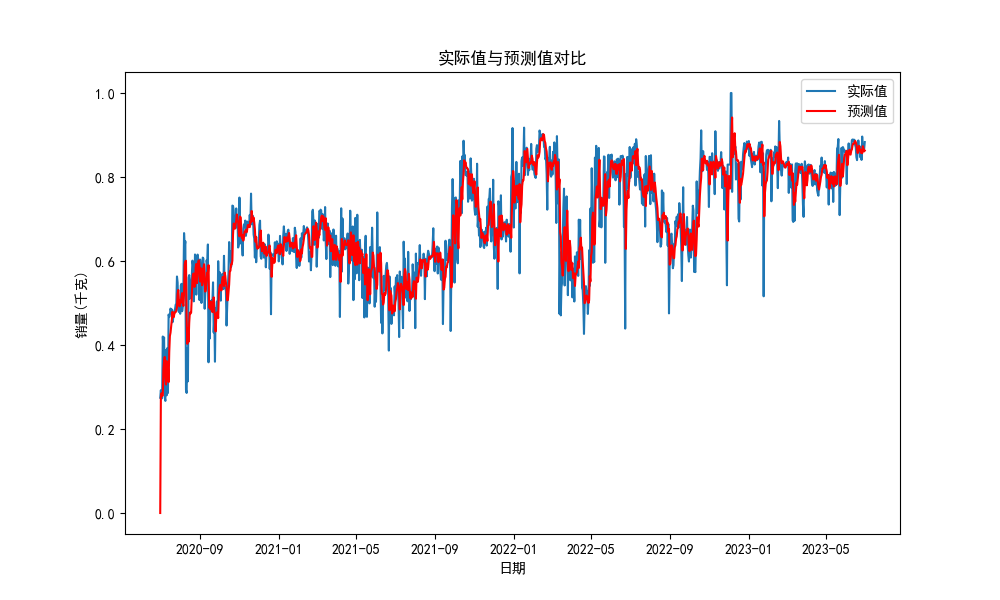
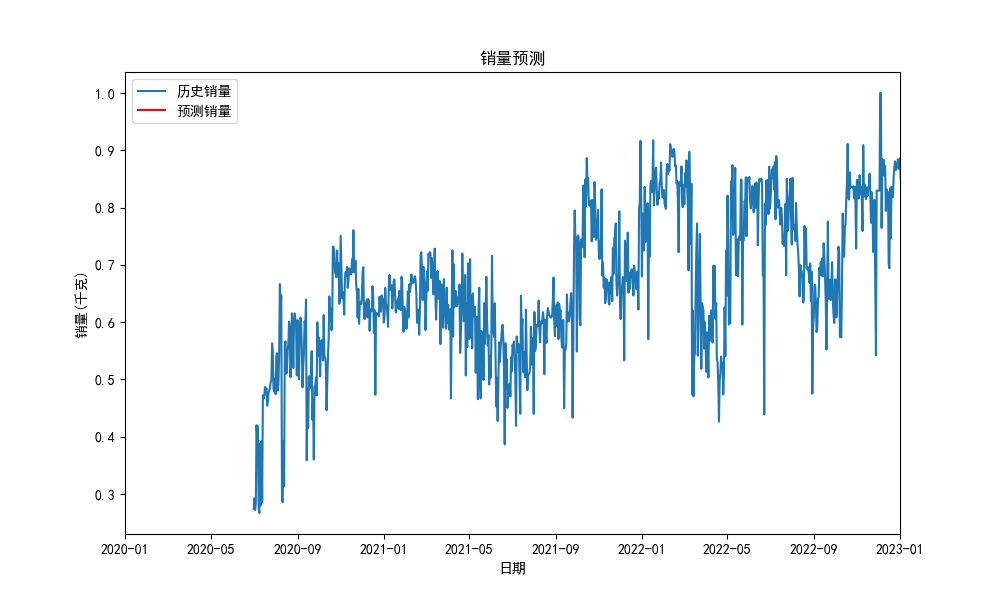
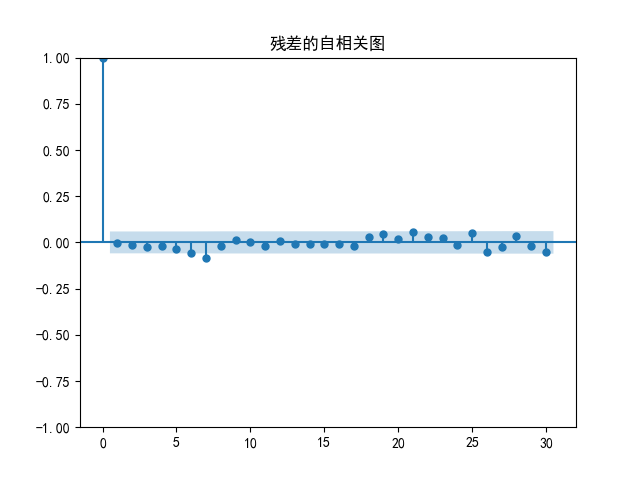
时间序列模型代码：

import time  
  
import pandas as pd  
from statsmodels.tsa.arima.model import ARIMA  
import matplotlib.pyplot as plt  
from matplotlib.font\_manager import FontProperties  
from statsmodels.graphics.tsaplots import plot\_acf, plot\_pacf  
from constant import \*  
# 设置字体为 SimHei  
plt.rcParams['font.sans-serif'] = ['SimHei'] # 用于正常显示中文标签  
plt.rcParams['axes.unicode\_minus'] = False # 用于正常显示负号  
# group2.to\_excel(new\_path)  
t3 = pd.read\_excel(path1)  
ls = set(t3['分类名称'].values)  
print(ls)  
def function():  
 for item in ls:  
 ins\_path = rf'图片数据\{item}'  
 ins\_path2 = rf'数据\{item}预测结果.xlsx'  
 print(item)  
 # 读取数据  
 file\_path = f'整合各单品{item}.xlsx' # 修改为实际的文件路径  
 # file\_path = f'整合各单品花叶类.xlsx'  
 sales\_data = pd.read\_excel(file\_path)  
  
 # 转换日期列为 datetime 类型  
 sales\_data['销售日期'] = pd.to\_datetime(sales\_data['销售日期'])  
  
 # 按日期排序  
 sales\_data = sales\_data.sort\_values('销售日期')  
  
 # 设置日期列为索引，并添加频率信息  
 sales\_data.set\_index('销售日期', inplace=True)  
 sales\_data = sales\_data.asfreq('D') # 设置为每日频率，根据实际情况修改  
  
 # 选择销量数据进行建模  
 sales\_volume = sales\_data['销量(千克)'].fillna(method='ffill') # 如果有缺失值，用前向填充  
  
 # 创建并训练 ARIMA 模型  
 model = ARIMA(sales\_volume, order=(5, 1, 0))  
 model\_fit = model.fit()  
  
 # 残差分析  
 residuals = model\_fit.resid  
  
 plt.figure(figsize=(10, 6))  
 plt.plot(residuals)  
 plt.title('残差图')  
 plt.xlabel('日期')  
 plt.ylabel('残差')  
 plt.savefig(ins\_path+'残差图.png')  
 plt.show()  
  
 plt.figure(figsize=(10, 6))  
 plot\_acf(residuals, lags=30)  
 plt.title('残差的自相关图')  
 plt.savefig(ins\_path+'自相关图.png')  
 plt.show()  
  
 # 模型评估指标  
 aic = model\_fit.aic  
 bic = model\_fit.bic  
 print(f"AIC: {aic}")  
 print(f"BIC: {bic}")  
  
 # 实际值与预测值对比  
 predictions = model\_fit.predict(start=0, end=len(sales\_volume)-1)  
 plt.figure(figsize=(10, 6))  
 plt.plot(sales\_volume, label='实际值')  
 plt.plot(predictions, label='预测值', color='red')  
 plt.title('实际值与预测值对比')  
 plt.xlabel('日期')  
 plt.ylabel('销量(千克)')  
 plt.legend()  
 plt.savefig(ins\_path+'实际与预测')  
 plt.show()  
  
 # 进行未来7天的预测  
 forecast = model\_fit.forecast(steps=7)  
 forecast.to\_csv(ins\_path2)  
 # 可视化预测结果  
 plt.figure(figsize=(10, 6))  
 plt.plot(sales\_volume, label='历史销量')  
 plt.plot(forecast, label='预测销量', color='red')  
 plt.title('销量预测')  
 plt.xlabel('日期')  
 plt.ylabel('销量(千克)')  
 plt.legend()  
 plt.xlim(pd.Timestamp('2020-01-01'), pd.Timestamp('2023-01-01'))  
 plt.savefig(ins\_path+'预测图')  
 forecast.to\_excel(ins\_path2)  
 plt.show()  
 print(forecast)  
 time.sleep(3)  
 plt.close()  
function()

预处理数据代码：

from constant import\*  
import pandas as pd  
import os  
import tqdm  
def merge():  
 new\_path = os.path.join(base\_path,r'问题2新研究\附件2优化数据.xlsx')  
 new\_path1 = os.path.join(base\_path,r'问题2新研究\附件3优化数据.xlsx')  
 t6 = pd.read\_excel(path3)  
 group = t6.groupby(['销售日期','单品编码']).agg({'批发价格(元/千克)':'mean'})  
 group.to\_excel(new\_path1)  
 table1 = pd.read\_excel(path2)  
 grouped\_table = table1.groupby(["销售日期", "单品编码"]).agg(  
 {'销量(千克)': 'mean', '销售单价(元/千克)': 'mean'}).reset\_index()  
  
 grouped\_table.to\_excel(new\_path, index=False)  
 table= pd.read\_excel(new\_path)  
  
 t2 = pd.read\_excel(path3)  
 group2 = t2.groupby(['单品编码','销售日期']).agg({'批发价格(元/千克)':'mean'}).reset\_index()  
 group2.to\_excel(new\_path)  
 t3 = pd.read\_excel(path1)  
 ls = set(t3['分类名称'].values)  
 group4 = t3.groupby('分类名称')  
 for category, group in group4:  
 category\_str = str(category).strip("(),'\"")  
 # 创建导出路径  
 export\_path = os.path.join(base\_path, f'问题2新研究\分类\_{category\_str}.xlsx')  
 # 保存分组数据到Excel文件  
 group.to\_excel(export\_path, index=False)  
 for item in tqdm.tqdm(ls):  
 print(item)  
 # print(item)  
 path6 = f'分类\_{item}.xlsx'  
 # 读取第一个文件  
 df1 = pd.read\_excel(path6)  
  
 # 读取第二个文件  
 df2 = pd.read\_excel(new\_path)  
  
 # 合并两个表格  
 merged\_df = pd.merge(df1, df2, on='单品编码', how='inner')  
  
 # 选择需要的列  
 result\_df = merged\_df[  
 ['单品编码', '单品名称', '分类编码', '分类名称', '销售日期', '销量(千克)', '销售单价(元/千克)']]  
  
 # 保存结果到新的Excel文件  
 # print(result\_df)  
 output\_path = os.path.join(base\_path, f'合并{item}.xlsx')  
 result\_df.to\_excel(output\_path, index=False)  
 for item in tqdm.tqdm(ls):  
 path7 = rf'合并{item}.xlsx'  
 t8 = pd.read\_excel(path7)  
 new\_path7 = f'整合各单品{item}.xlsx'  
 agg = t8.groupby('销售日期').agg({'销量(千克)':'mean','销售单价(元/千克)':'mean'})  
 agg.to\_excel(new\_path7)  
 path7 = '价格弹性系数.xlsx'  
 t7 = pd.read\_excel(path7)  
 agg\_t7 = t7.groupby('分类名称').agg({'价格弹性系数':'mean'}).reset\_index()  
 t9 = pd.read\_excel('附件4.xlsx')  
 merge\_t9 = pd.merge(agg\_t7,t9,left\_on ='分类名称',right\_on='小分类名称')  
 merge\_t9.to\_excel('弹性系数和损耗率.xlsx')  
 print(merge\_t9)  
 # print(t8.head(),'\n',t9)  
 # merge\_78 = pd.merge(t8,t9,on='单品编码')  
 # merge\_78.to\_excel('损耗率与类.xlsx')  
merge()  
# a = pd.read\_excel('附件4.xlsx')  
# print(a)

其它类的自相关图，预测数据与真实数据图。残差图，相关性图



剩余单品种类的模拟退火算法模型图像

