

基于动态规划模型的蔬菜商品定价与补货问题

王欣宇^{1*}, 王子杰², 李宇鹏^{1#}

¹延安大学数学与计算机科学学院, 陕西 延安

²西安科技大学地质与环境学院, 陕西 西安

收稿日期: 2024年7月1日; 录用日期: 2024年7月26日; 发布日期: 2024年8月1日

摘要

本文研究了如何对蔬菜类商品进行自动定价与补货。通过分析近三年各蔬菜品类及各单品的销售信息, 剔除掉不合理数据, 整理出最优数据, 建立动态规划模型, 结合添加附加条件预测出未来特定时间内的最优补货量和最优定价策略。针对问题一, 本文首先对数据进行预处理, 通过箱线图查找异常值并剔除不合理数据, 将销量进行量纲归一化处理, 分析相关性。利用MATLAB数据可视化工具箱绘制销量随季度变化的图表和品类相关性矩阵图, 揭示各品类销售总量变化规律和相关程度。对于单品销售量, 利用归一化数据绘制归一化表并通过Excel数据透视表进行分析。结果显示花叶类商品销量稳定高, 茄类销量较低; 水生根茎类和花菜类相关性高, 辣椒类与食用菌类相关性较高。在单品销售中, 芜湖青椒与西兰花销量接近, 金针菇与云南生菜销量最高。针对问题二, 第一小问利用公式计算各蔬菜单品的成本加成价格, 并通过MATLAB数据拟合分析品类销量与成本加成定价的关系。随着成本加成定价的降低, 销售总量分布更分散且较低; 成本加成定价提高后, 销售总量更聚集, 与拟合直线重叠较多。第二小问使用MATLAB创建动态规划的状态空间和价值函数矩阵, 建立模型并通过内部嵌套循环程序分析, 利用disp函数得出各蔬菜品类每天的最优补货量和定价, 以获得商超的最大收益值, 当销量超过前一天库存时, 最优补货量设为0。针对问题三, 首先对总表进行数据预处理, 得到六天的蔬菜单品总销量、销售价格和成本加成定价数据矩阵; 其次建立一个线性规划模型, 通过定义目标函数和约束条件, 得到蔬菜商超的收益最大值、当日蔬菜单品的最优补货量和最优定价。针对问题四, 为了更好地帮助商超解决前三个问题, 可以首先提出蔬菜需求量预测数据, 并利用基于集成学习stacking方法的算法来预测未来顾客对蔬菜购买量的需求, 以帮助商超调整蔬菜商品的定价。其次, 统计供应商信息, 考虑进货商可靠性等因素, 以选择最优供应商, 提高进货效率和蔬菜品质。然后, 统计蔬菜商品市场价格竞争情况, 帮助商超分析同行的定价策略。最后, 考虑顾客满意度问题, 统计顾客反馈和需求数据以及退货原因, 不断改进以吸引更多忠实客户。

关键词

动态规划模型, 预测模型, 量纲归一化, 相关系数分析法, 拟合分析

*第一作者。

#通讯作者。

Vegetable Product Pricing and Replenishment Problem Based on Dynamic Programming Model

Xinyu Wang^{1*}, Zijie Wang², Yupeng Li^{1#}

¹School of Mathematics and Computer Science, Yan'an University, Yan'an Shaanxi

²School of Geology and Environment, Xi'an University of Science and Technology, Xi'an Shaanxi

Received: Jul. 1st, 2024; accepted: Jul. 26th, 2024; published: Aug. 1st, 2024

Abstract

This study explores how to automatically price and replenish vegetable products. By analyzing sales data of various vegetable categories and individual products over the past three years, filtering out unreasonable data, organizing optimal data, establishing a dynamic programming model, and adding additional conditions to predict the optimal replenishment quantity and pricing strategy for specific future time periods. For Problem 1, this paper first preprocesses the data, identifies outliers using box plots, normalizes sales quantities, and analyzes correlations. Using MATLAB's data visualization toolbox, charts showing sales variation by quarter and category correlation matrices are generated to reveal the trends and correlations in total sales volumes by category. For individual product sales, normalized tables are created and analyzed using Excel pivot tables. Results show stable high sales for leafy vegetables, lower sales for tomato products; high correlation between aquatic root vegetables and cauliflower, and between chili peppers and edible fungi. Among individual product sales, sales are similar for Wuhu green peppers and broccoli, while enoki mushrooms and Yunnan lettuce have the highest sales. For Problem 2, in the first sub-question, the cost-plus pricing for each vegetable product is calculated using a formula, and the relationship between category sales and cost-plus pricing is analyzed through MATLAB data fitting. As cost-plus pricing decreases, total sales distribution becomes more dispersed and lower; with increased cost-plus pricing, total sales become more concentrated and overlap with the fitted line. In the second sub-question, a dynamic programming model is created using MATLAB to establish the state space and value function matrix. An internal nested loop program is used to analyze and determine the optimal replenishment quantity and pricing for each vegetable category daily, aiming to maximize the supermarket's revenue, with the optimal replenishment quantity set to zero when sales exceed the previous day's inventory. For Problem 3, initial data preprocessing is performed on the overall table to obtain matrices of total sales, selling prices, and cost-plus pricing data for vegetable products over six days. A linear programming model is then established to define the objective function and constraints, resulting in the maximum profit for the vegetable supermarket, along with the optimal replenishment quantity and pricing for the vegetable products on that day. For Problem 4, to further assist the supermarket in addressing the previous three problems, it is suggested to first propose vegetable demand forecasting data. Utilizing an algorithm based on ensemble learning stacking methods to predict future customer demand for vegetable purchases can aid the supermarket in adjusting vegetable product pricing. Secondly, gathering supplier information and using the Analytic Hierarchy Process to select vegetable suppliers, considering factors such as supplier reliability to choose the best supplier, thereby enhancing procurement efficiency and vegetable quality. Additionally, analyzing market price competition of vegetable products can help the supermarket understand competitor pricing strategies. Finally, focusing on customer satisfaction, collecting feedback, demand data, and return reasons can help continuously improve to attract more loyal customers.

Keywords

Dynamic Programming Model, Forecasting Model, Normalization, Correlation Coefficient Analysis, Fitting Analysis

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 问题重述

1.1. 问题背景

在蔬菜商品超市中,销售量和种类繁多,提前准备蔬菜商品至关重要。商家需考虑销售商品的季节性变化和保质期限制。根据历史销售数据预测未来销售情况,选择适当的商品类型进行批发。销售量和批发价格之间存在一定的关联性,因此根据历史销售情况确定未来商品的定价和补货量至关重要。

1.2. 问题重述

根据问题背景和附件数据,需要解决以下 4 个问题:

在蔬菜商超中,不同蔬菜品类之间的销售量具有一定的规律性、相关性和分布特点,不同蔬菜单品之间的销售量也有上述此关系。

为了使蔬菜商超获得最大收益,分析其各蔬菜品类的销总售量与成本加成定价之间的关联性,为该商超提供未来一周各蔬菜品类的最优补货量和最优定价提供优化方法。

基于问题二,以可售单品至多 33 个、至少 27 个,各单品订购量不小于 2.5 千克为限制条件,制定该蔬菜商超各可单品未来某日的最优补货量和最优定价,使蔬菜商超获得最大收益。

通过对前面三个问题的分析计算以及数据的整理,思考更多的可以帮助解决前三个问题的数据,从而更好地帮助商超制定蔬菜商品的补货和定价决策。

2. 问题分析

2.1. 问题一的分析

针对问题一,根据题目把其分为两小问,第一小问要分析不同蔬菜品类之间销售量的规律性、相关性和分布特点,首先合并附件 1 和附件 2,对其中数据进行预处理,剔除异常值,处理缺失的数据,整理商品的销售流水明细,因为蔬菜是一类季节性的农产品,季节与销售量之间或存在一定的关系,所以计算出每个季度各品类蔬菜的销售总量,最后使用 MATLAB 数据可视化工具箱,将各蔬菜品类进行比较和可视化展示,绘制出直方图、折线图和品类相关性矩阵图,来反映各品类销售总量随季度的变化规律和其两两之间的相关程度。对于第二小问要求根据销售数量分析不同蔬菜单品之间的分布规律和相互关系,观察到各蔬菜销售量的单位不同,要利用量纲归一化方法进行不同单位的数据统一转换,首先利用 excel 的数据透视表整理出各蔬菜单品四个季度的销售总量,其次进行销售量归一化处理[1],利用 excel 可视化图表绘制各单品销售量直方图,观察它的分布规律。

2.2. 问题二的分析

针对问题二,把其分为两小问,第一小问分析超市蔬菜品类销量和成本加成定价的关系。首先将附

件 1 至附件 4 的数据进行合并, 利用“成本加成定价公式”, 通过将“销量”、“销售单价”、“批发价格”和“单品损耗率”带入公式, 得出每件蔬菜单品的成本加成定价。最后通过 matlab 数据拟合得到两者之间的二维关系及其与时间之间的三维关系。对于第二小问要分析 2020~2023 年间的销售量数据, 筛选出三年里 7 月 1 日~7 月 7 日的的数据, 将其作为基础数据, 使用 excel 透视表对该数据进行分析、分类、汇总、取平均, 得到各蔬菜品类的销售量、销售单价和成本加成定价矩阵。再用 matlab 创建动态规划的状态空间和价值函数的矩阵, 建立动态规划模型, 用于计算在商超中每天的最优补货量和最优定价, 将先前得到的分类数据带入模型, 最终实现最大收益模型的建立与解决。

2.3. 问题三的分析

针对问题三, 需要建立一个线性规划模型, 其约束条件为 7 月 1 日可以出售的单品总数和各蔬菜单品的订购量, 目标函数为蔬菜商超的收益最大值、当日蔬菜单品的最优补货量和最优定价。首先需要对总表进行数据预处理, 得到六天的单品总销量、销售价格和成本加成定价; 其次对其透视表进行行分类提取和处理, 形成 44×7 型“销售单价”、“成本加成定价”、“销售量归一化”三个矩阵; 再使用 matlab 程序, 通过定义目标函数和约束条件, 调用线性规划程序, 得到最优解, 并根据最优解进行进一步的分析 and 决策。

2.4. 问题四的分析

针对问题四, 需要查找其他有效数据来更好的帮助解决前三个问题, 从而帮助商超制定蔬菜商品的补货和定价决策, 通过对问题的整体思考, 首先, 提出增加蔬菜需求量预测数据, 利用基于集成学习 stacking 方法的预测和算法来预测未来顾客对蔬菜购买量的需求[2]。最后, 考虑顾客满意程度问题, 统计顾客的反馈和需求数据以及退货原因, 并具体分析这五类数据的意义。

3. 模型假设

- 1) 不考虑保质期和品相随时间变差的问题;
- 2) 排除生活中的小概率事件, 如台风、地震等;
- 3) 仅考虑问题中的核心因素如售价、销量等, 不考虑次要因素的影响如每次蔬菜的质量是否统一;
- 4) 为了简化模型, 假设一周内不会因为突发灾害而不售卖。

4. 符号说明

见表 1 符号假设表, 符号 D 为销售单价, 即为每种蔬菜类的单价; 符号 J 为批发价格即为本文中各种商品的批发价格; 符号 K 为单品损耗率; 符号 C 为单品销量。

Table 1. Symbol hypothesis table

表 1. 符号假设表

符号	说明	单位
D	销售单价	元
J	批发价格	元
K	单品损耗率	%
C	单品销量	kg

5. 问题一模型的建立与求解

5.1. 数据预处理

数据分析与处理

数据准备：汇总并合并各蔬菜商品的销售流水明细及其相对应的批发价格数据。

数据处理：对数据进行清选、处理和筛查。通过对原始数据的可视化分析，对异常数据、缺失数据进行处理[3]，如删除批发价高于单品售价的数据，删除突变数据等不合理数据，又如通过取出一类菜品中的日期列与销售单价绘制直方图来观察突变值和异常值，去除异常值、处理缺失数据、计算最终的销售额。

5.2. 针对问题一的第一小问

5.2.1. 分布规律分析

分布规律分析法是一种通过分析数据的分布，用来得出数据背后规律趋势的方法。

该小问利用 excel 中数据透视表功能计算出每个季度各蔬菜品类的销售总量。

花叶类商品在四个季度中的销售量都居于高位，而茄类在四季中销售量都较少，可以看出顾客在选择菜品时更倾向于去购买花叶类蔬菜。辣椒类与食用菌类在四个季度中销售量的变化趋势相近。

5.2.2. 相关性分析

相关性分析法[4]是一种统计方法，用于研究两个数据或多个数据变量之间的关系。

本文通过对数据的分析得出蔬菜商品各品类之间的相关性矩阵图来衡量数据之间的相关程度，从而得出各蔬菜品类之间的相关性。

可以在数据中得到水生根茎类蔬菜和花菜类蔬菜的相关性偏高，为 0.8，辣椒类蔬菜与食用菌类蔬菜相关性相对较高，其相关系数接近 0.7，而茄类和食用菌类、花叶类和食用菌类的相关性相反，两者都呈负线性相关关系，分为-0.6 和-0.4。

可以看出顾客更倾向于选择水生根茎类和花菜类一起购买，辣椒类与食用菌类一起购买。

5.3. 针对问题一的第二小问

5.3.1. 量纲归一化

本题中观察到蔬菜销量数据的单位不同，需要先对单品销量进行量纲归一化，把不同单位的数据转化为统一的尺度。

归一化公式：

$$y = \frac{\text{new}(x) - \min(x)}{\max(x) - \min(x)}$$

其中：

$\text{new}(x)$ ——当前销量值 x ；

$\max(x)$ ——最大销量值 x ；

$\min(x)$ ——最小销量值 x 。

利用 excel 中蔬菜单品的销量数据。对各单品销量进行量纲归一化处理。去除接近于 0 的数据。

5.3.2. 单品销量随四季度变化规律

首先利用 excel 中的数据透视表功能整理出各蔬菜单品四个季度的销售总量，接着使用 excel 制作功

能绘制直方。

从蔬菜单品的总销量，以及各个季度的销量在总销量中的占比部分可以看到，西兰花总销量最高，其次为云南生菜、金针菇、芜湖青椒、紫茄子、西岐花菇等。蔬菜单品销售量趋势为第一季度和第三季度销售数量增加，第二季度和第四季度销售数量减少，顾客在七月到九月购买蔬菜的需求量最多。

6. 问题二模型的建立与求解

6.1. 针对问题二的第一小问

6.1.1. 计算成本加成定价

首先处理汇总表，得到各个蔬菜单品的三年的销售量、销售单价、批发价格、单品损耗率，根据成本加成定价公式，得出各个蔬菜单品的成本加成定价。

成本加成定价法公式

$$\begin{aligned} \text{价格} &= \text{单位成本} \times (1 + \text{加成率}) \\ \text{单位成本} &= (\text{固定成本总额} + \text{变动成本总额}) / \text{销售量} \\ &= \text{单位固定成本} + \text{单位变动成本} \\ \text{加成率} &= ((\text{售价} - \text{进价}) / \text{进价}) \times 100\% \end{aligned}$$

成本加成定价 = ((批发价格 × 销量 + 单品损耗率/100 × 销量)/销量) × (1 + (销售单价 - 批发价格)/批发价格)。

6.1.2. 数据拟合

通过数据处理得到总的销量在三年中的随时间变化的趋势。再通过 matlab 曲线拟合工具，使用 polyfit 函数得到各蔬菜品类销量和成本加成定价之间的二维关系；通过 polyfit 函数得到的拟合系数 p 绘制实际数据的散点图，得到蔬菜品类销量、成本加成定价与时间三者之间的三维关系。

单品销售总量与加成成本定价之间的关系为：当成本加成定价较低时，销售总量分布较为分散，销售总量也较低；当成本加成定价上涨时，单品的销售总量有所提高；当成本加成定价再次提高时，单品销售总量的分布较为聚集，与最终的拟合直线有较多重叠。

6.2. 针对问题二的第二小问

6.2.1. 数据预处理

根据 3 年中特定时间段(7 月 1 日~7 月 7 日)的各类蔬菜品种的销售单价、销售量、成本加成定价，将已有的附件里的数据，通过以单品编号为主键的方式，做成一张新的筛选后的汇总表。通过对其进行数据透视表操作，形成以品类为行向量、日期为列向量的 6 × 7 型销售总量、售价和成本加成定价矩阵。

6.2.2. 建立、求解、分析动态规划模型

1) 建立模型

使用动态规划分析统计方法建立销售总量与定价的数学模型，预测未来七日的销售量，并考虑损耗率和库存情况，确定每个品类的日补货总量。

编辑模型内部嵌套循环程序，计算最大收益，遍历各品类和七天的补货量、定价和收益，得出每日最优补货计划和定价策略。

将数据输入动态规划模型，得出各品类每日补货总量和定价策略，以及商店的最大收益。

2) 模型分析

本段代码使用动态规划算法计算商超每天的最优补货量和定价，以实现最大收益。步骤如下：

定义 6 个蔬菜品类和预测天数。

创建销售价格和批发成本价格矩阵，定义销售总量矩阵。

建立动态规划的状态空间和价值函数矩阵。

输入数据矩阵到模型中求解。

3) 模型求解

定义状态空间和价值函数矩阵；

从倒数第二天开始，逐天计算最大收益；

计算当前天的补货总量，计算当前天的收益，比较更新最大收益和补货计划；

计算当前天的补货总量，使用 \max 函数限制不小于 0 的最小值；

外层循环结束后，得到最大收益和补货计划。

6.2.3. 结果分析

该蔬菜商超最大收益为 414.8298 元；

根据 2020 年至 2022 年每年 7 月 1~7 日六类蔬菜销售价格和成本价格的数据，分析每个蔬菜品类的销售趋势和周期性，再将数据导入 MATLAB，预测未来七日计划每日最优补货量和最优定价结果见如下表 2 和表 3 所示，表格中花菜类 7 月 2 日计划补货量为 0，原因是销售总量超过了前一天的库存量。

Table 2. Optimal daily replenishment (kg)

表 2. 每日最优补货量(kg)

种类	日期	7月1日	7月2日	7月3日	7月4日	7月5日	7月6日	7月7日
	花菜类		826.38	0	31.529	0	425.108	13.848
花叶类		796.101	0	15.33	0	384.801	72.452	398.2
辣椒类		812.136	0	19.761	0	439.888	33.485	103.37
茄类		811.491	0	818.224	58.487	446.401	0	51.014
食用菌		800.601	0	52.88	0	442.828	5.075	91.228
水生根茎类		813.248	0	73.481	0	57.051	0	53.727

Table 3. Optimal daily pricing (Yuan)

表 3. 每日最优定价(元)

种类	日期	7月1日	7月2日	7月3日	7月4日	7月5日	7月6日	7月7日
	花菜类		11.5498	11.1011	10.8797	11.615	11.157	10.8425
花叶类		7.5969	7.221	7.3189	7.5185	7.1104	7.3136	7.2112
辣椒类		11.6403	13.0662	12.0352	12.0722	11.5757	12.9185	14.3391
茄类		5.2755	5.6754	5.7528	5.7061	6.5969	6.1051	5.937
食用菌		13.0341	14.0056	14.6016	14.6867	13.8262	15.4592	15.0341
水生根茎类		16.7865	17.7358	13.1062	14.7839	15.0309	13.89	14.331

7. 问题三模型的建立与求解

7.1. 数据预处理

对剔除异常数据后的总表进行数据筛选, 选出 2023 年 6 月 24~30 日的蔬菜单品数据, 对其进行数据透视表操作, 最后得到 44 种蔬菜单品每日的销量总和、平均销售价格、平均成本加成定价。

对上述数据进行分类提取和处理, 形成“销售单价”、“成本加成定价”、“销售量归一化”三张表格, 其都为 44×7 型矩阵。

7.2. 模型建立

使用线性规划来解决此问题:

通过分析 44 种蔬菜单品的数量、天数、销售总量、成本加成定价和销售单价等, 构造一个目标函数系数矩阵, 其中每个元素表示每种蔬菜在每天的销售收益或成本加成价;

分别构造了两个不等式约束矩阵和不等式约束右侧向量, 限制每种蔬菜品类在每天的补货总量必须在 27~33 范围内, 同时限制它们在每天的陈列量必须满足最小要求 2.5 kg;

调用 matlab 函数求解线性规划问题, 得到最优解, 并对其进行处理, 确定每种蔬菜单品在 7 月 1 日的最优补货量和最优定价, 使蔬菜商超收益最大。

该模型使用线性规划方法来优化每种蔬菜单品的日补货量和定价, 可以最大化蔬菜商超的收益。通过定义目标函数和约束条件, 然后调用线性规划程序, 可以得到最优解, 并根据最优解进行进一步的分析 and 决策。

8. 问题四模型的建立与求解

通过对前面三个问题的分析计算以及数据的整理, 我们认为以下数据可以更好的制定蔬菜商品的补货和定价决策。

8.1. 蔬菜需求量预测数据

通过使用基于集成学习 stacking 方法的预测和算法可以很有效的预测未来顾客对蔬菜购买量的需求。

首先使用我们已有的销售数据, 包括蔬菜的商品名、销售数量和销售金额。其次整理天气数据, 包括对气温、风力大小。天气数据可以根据 Python 爬虫得到相同时间跨度的历史天气数据[3]这有利于商超更加准确地预测各个蔬菜品类的未来需求量。蔬菜需求预测数据可以有效的帮助商超制定合理的定价策略。通过了解未来的蔬菜需求量, 商超可以根据得到的供需关系来调整蔬菜商品的定价。如果预测的需求量较低, 商超可以适度降低蔬菜商品的销售单价, 从而减少库存风险, 制定更加合理的补货策略。

8.2. 供货商信息

选择最适合的蔬菜供应商可以建立更加科学高效率的供应链。供应商的诚信度、提供菜品能力、交货的时间、准时率都会影响到菜品的售价及销售量。

商超通过了解供应商的诚信的、近阶段的历史交货记录以及客户的评价来评估供应商的可靠性。若有可靠性较高的供应商, 商超就可以对其选择, 甚至可以与其长期建立合作关系, 更加提高了今后进货效率。而若有些供应商的诚信度较低, 商超就可以不考虑与其合作, 去寻找更适合的供应商, 但如若必须与其合作, 商超也可以通过数据采取措施来减少损耗并降低风险。

商超也可以根据各个供应商给出的成本价和进购批发价进行分析, 通过应用层次分析法[4]来考虑进

购商可靠性、提供菜品能力、交货的时间和准时率等各项因素来选择最优的供货商。

基于模糊层次法对蔬菜供应商的选择[2]就可以很好的帮助我们实现挑选品质更高的供应商。

8.3. 价格竞争情况

通过统计蔬菜超市竞争对手的价格竞争情况,商超可以更好的了解市场上同类蔬菜商品的定价情况。它们可以根据这些数据,结合自身的产品特点和品质优势来确定或调整定价策略。

商超还可以利用竞争对手的价格竞争情况预测顾客对蔬菜的需求量和销售趋势,并及时调整自己的最优补货计划。例如,竞争对手的定价较低,商超也可以适度下调自己的蔬菜销售价格,以吸引更多的顾客并提高销售量。

8.4. 顾客反馈和需求数据

商超可以通过顾客调查问卷法、顾客反馈收集法、顾客访谈和分析销售数据等方法,整理出顾客对不同蔬菜品类的喜爱程度和需求量从而调整补货和定价策略,为顾客提供更符合其需求的商品来提升顾客的满意程度,积攒更多的回头客。

参考文献

- [1] 张鹤. 基于集成学习的蔬菜销售预测[D]: [硕士学位论文]. 昆明: 云南师范大学, 2021.
- [2] 平瑛, 吴昊. 基于模糊层次分析法的蔬菜供应商选择研究[J]. 广东农业科学, 2013, 40(16): 217-220.
- [3] 刘全伟. 基于网络爬虫的排行榜系统设计与实现[D]: [硕士学位论文]. 北京: 北京邮电大学, 2018.
- [4] 张建林, 钱江. 浅析供应商选择决策——应用层次分析法的问题[J]. 技术经济与管理研究, 2002(1): 76-77.