

应用马尔可夫链模型预测零售业库存需求的精准度研究

吴宇霞, 宋鑫

上海理工大学管理学院, 上海

收稿日期: 2023年2月17日; 录用日期: 2023年3月17日; 发布日期: 2023年3月28日

摘要

随着零售业的发展, 库存管理成为零售业运营中非常重要的一环, 一定程度上影响到企业资金链, 进而影响运营效率。因此精准预测库存需求量对于实现零售新物流改革十分重要。本文提出利用马尔可夫链模型来预测库存数据, 通过选取M超市14个品类在2021年12个月的库存量数据进行分析研究。由于各品类量纲不同, 首先将库存数据进行Z-Score标准化, 结合历史数据采用马尔可夫过程计算每月库存状态转移概率, 并对下月库存状态进行预测, 将预测数据与当月实际数据进行对比。结果表明: 应用马尔可夫链模型预测的库存数据准确度高达85.71%, 该模型在一定程度上可以帮助管理者更好地进行库存管理, 减少企业存货成本管理中的缺货成本, 库存成本等主要问题, 提高物流平台的运行质量。

关键词

马尔可夫链, 模型预测, 库存需求

Research on the Accuracy of Applying Markov Chain Model to Forecast Retail Industry Inventory Demand

Yuxia Wu, Xin Song

Business School, University of Shanghai for Science and Technology, Shanghai

Received: Feb. 17th, 2023; accepted: Mar. 17th, 2023; published: Mar. 28th, 2023

Abstract

With the development of the retail industry, inventory management has become a very important

part of the operation of the retail industry, affecting the enterprise's capital chain to a certain extent, and thus affecting the operation efficiency. Therefore, accurate prediction of inventory demand is very important for realizing the new retail logistics reform. This paper proposes to use the Markov chain model to predict the inventory data, and analyzes the inventory data of 14 categories of M Supermarket in the 12 months of 2021. Due to the different dimensions of each category, the inventory data is first standardized by Z-Score, and the Markov process is used to calculate the transfer probability of monthly inventory status in combination with historical data, and the inventory status of the next month is predicted. The predicted data is compared with the actual data of the current month. The results show that the accuracy of the inventory data predicted by the Markov chain model is up to 85.71%. The model can help managers to better manage inventory, reduce the main problems such as shortage cost and inventory cost in enterprise inventory cost management, and improve the operation quality of the logistics platform.

Keywords

Markov Chain, Model Prediction, Inventory Demand

Copyright © 2023 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. 引言

随着消费者的个性化需求和对新品类消费的增加,若运营管理者不能对库存量做精准预测,不仅会增加企业的经济成本,也会提高缺货成本或者影响消费者的时效体验,进而影响销售收入。因此,如何对门店与配送中心的品类和库存进行科学地配置是新零售企业面临的难题之一[1]。目前国内外学者已经对库存预测模型开展了大量研究,李敬泉和满秀芳(2014)运用报童模型研究了三种库存模式下的零售商和供应商的最优库存策略,结果显示零售商分散化管理库存策略和供应商共享库存策略有利于零售商和供应商分别发挥各自优势,实现风险共担,利益共享,提高供应链整体利润水平[2]。罗晓萌等(2014)基于时间序列法对目前系统中历史销售数据进行分析以预测未来需求,同时考虑价格对需求的影响,通过验证发现预测误差一般能控制在10%以下[3]。杨静雅和孙林夫(2015)提出支持向量回归(SVR)的预测方法,采用量子粒子群算法(QPSO)对SVR参数进行优化选择,设计基于QPSO-SVR的售后配件库存需求预测流程。实验和比较结果表明,QPSO-SVR预测方法有效可行,其预测精度和泛化能力较强[4]。

由于零售业库存量在一般情况下是由上期库存以及当期的销售量所决定,不受之前状态所影响,即具备马尔可夫性质[5]。本文第一部分对马尔可夫链及其预测模型进行介绍,第二部分通过选取M超市2021年共14个品类每月的库存量资料,通过运用马尔可夫链模型对其库存量进行预测。最后预测结果表明该模型为企业库存决策有一定帮助,为零售业库存预测提供新思路。

2. 马尔可夫链预测法

2.1. 马尔可夫链

马尔可夫链是指状态空间中从某一状态转换到另一个状态的随机过程,该过程要求具备“马尔可夫性质”,即“无记忆性”,下一状态的概率分布只由当期状态决定,不受时间序列中前期状态所影响[6]。设随机序列 $\{X(n), n=0,1,2,\dots\}$ 具有离散状态空间 E 。若对于任意 m 个非负整数 n_1, n_2, \dots, n_m

($0 \leq n_1 < n_2 < \dots < n_m$)和任意自然数 k , 以及任意 $i_1, i_2, \dots, i_m, j \in E$, 满足 $P\{X(n_m+k) = j | X(n_1) = i_1, X(n_2) = i_2, \dots, X(n_m) = i_m\} = P\{X(n_m+k) = j | X(n_m) = i_m\}$, 在上式中, 如果 n_m 表示现在时刻, n_1, n_2, \dots, n_{m-1} 表示过去时刻, n_m+k 表示将来时刻, 则表明该随机过程在将来 n_m+k 时刻处于状态 j 的概率仅依赖于现在 n_m 时刻的状态 i_m , 而与 $m-1$ 各时刻所处的状态无关, 可以将该随机序列称之为马尔可夫链[7]。

2.2. 转移矩阵

记 $P_{ij}(n, n+k) = P\{X(n+k) = j | X(n) = i\}$ 表示已知 n 时刻处于状态 i , 经 k 个单位时间后过程处于状态 j 的概率, 简记 $P_{ij}(k)$ 。当 $k=1$ 时, $P_{ij}(1)$ 称为一步转移概率, 此时 $P_{ij} = P_{ij}(1) = P\{X(n+1) = j | X(n) = i\}$,

一步转移概率可写成矩阵形式, 对有限状态空间 $E = \{1, 2, \dots, N\}$, 矩阵 $P = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \vdots & \vdots & & \vdots \\ P_{n1} & P_{n2} & \dots & P_{nn} \end{bmatrix}$ 一步转移

概率矩阵, 具有性质: ① $0 \leq P_{ij} \leq 1, i, j = 1, 2, \dots$; ② $\sum_{j \in E} P_{ij} = 1, i = 1, 2, \dots$ [8]。

由 C-K 方程可知, 马尔可夫链的转移概率之间有下列关系: $p_{ij}(n) = p_{ij}(k+l) = \sum_r p_{ir}(k) p_{rj}(l), i, j = 1, 2, \dots$, 一般地, 有 $P(n) = [P(1)]^n$ [9]。

2.3. 马尔可夫预测模型

马尔可夫预测模型是指通过在过去的时间序列中各状态之间的转移概率来建立初始转移概率矩阵。若想得到 t_n 时刻处于各状态的概率, 只需知道 t_0 时刻的状态, 通过 n 次状态转移, 利用 C-K 方程求出。

在零售业库存预测模型中, 通过 2021 年的库存数据, 建立库存量状态的转移概率矩阵, 即可建立马尔可夫模型。

3. M 超市库存需求精准度测试

下面以 M 超市 14 个品类在 2021 年 12 个月的库存量数据为例建立初始转移概率矩阵, 据此预测 2021 年库存量(见表 1)。

Table 1. M Supermarket's inventory in 2021
表 1. M 超市 2021 年库存量

品类	一月	二月	三月	四月	五月	六月	七月	八月	九月	十月	十一月	十二月
个护美容	16,547	15,570	14,673	14,320	13,994	12,332	12,236	12,585	13,306	12,644	12,663	13,740
家庭二	7240	5973	6808	6843	6815	6420	6235	6146	6929	7081	7425	7952
家庭一	25,258	21,481	20,836	20,875	21,048	21,053	19,967	22,133	21,721	21,073	20,747	21,120
酒窖	3451	3041	3028	3350	2979	3144	3077	3060	3233	2911	2769	3115
美食	544	671.8	376.3	346.7	336.46	345.2	388.96	554.44	499.92	439.6	314.6	248.8
面包水吧	2893.7	1969.1	1414.7	1142.4	1198.6	1434.4	1525.3	2249.6	1511.3	45	729.6	74
母婴	6910	6451	6175	5722	5418	4993	4732	5152	4928	4787	5031	5085

Continued

其他	29,576	33,977	31,910	32,274	31,992	30,539	32,101	48,218	48,579	45,078	45,778	17,894
速食	15526	13375	13070	13063	12571	12714	11919	11718	11738	11249	10346	12467
西式美食	290.97	210.65	234	231.9	200.81	205.97	184.98	163.85	188.91	133.46	88.808	71.3
鲜食	16,791	14,982	14,170	13,955	10,946	11,292	10,229	12,775	18,604	18,549	19,166	18,954
休闲	34,125	28,067	30,614	30,001	27,596	27,554	26,422	29,744	29,162	27,286	28,464	29,838
饮料酒水	28,075	24,823	27,273	29,218	26,622	28,383	30,409	31,604	27,348	22,387	18,137	22,120
营养时尚	528	1344	1117	1039	899	785	1195	1054	987	936	829	732

3.1. 库存量数据标准化

由于在 M 超市 14 各品类中, 大多品类的量纲不同, 使得数据不具有可比性, 故应先将库存量数据进行标准化。数据标准化可消除变量间的量纲关系, 本文采用的是 Z-Score 标准化[10]。由上表数据可知, 均值为 \bar{X}_i , 标准差为 s_i , 每月库存量数据为 X_{ij} , 则经标准化后的数据 $Z_{ij} = (X_{ij} - \bar{X}_i) / s_i$, 见表 2。

Table 2. Inventory data after standardization
表 2. 经标准化后库存量数据

标准化数据	一月	二月	三月	四月	五月	六月	七月	八月	九月	十月	十一月	十二月
个护美容	2.07	1.35	0.70	0.44	0.20	-1.01	-1.08	-0.83	-0.30	-0.78	-0.77	0.02
家庭二	0.73	-1.49	-0.02	0.04	-0.01	-0.71	-1.03	-1.19	0.19	0.45	1.06	1.98
家庭一	2.90	0.03	-0.46	-0.43	-0.30	-0.30	-1.12	0.53	0.21	-0.28	-0.53	-0.25
酒窖	1.92	-0.30	-0.37	1.37	-0.64	0.26	-0.11	-0.20	0.74	-1.00	-1.77	0.10
美食	0.99	2.04	-0.38	-0.62	-0.70	-0.63	-0.27	1.08	0.63	0.14	-0.88	-1.42
面包水吧	1.88	0.75	0.08	-0.25	-0.18	0.10	0.21	1.10	0.20	-1.59	-0.75	-1.55
母婴	2.05	1.41	1.02	0.38	-0.04	-0.64	-1.01	-0.42	-0.73	-0.93	-0.59	-0.51
其他	-0.65	-0.18	-0.40	-0.36	-0.39	-0.55	-0.38	1.35	1.39	1.01	1.09	-1.91
速食	2.36	0.69	0.46	0.45	0.07	0.18	-0.43	-0.59	-0.57	-0.95	-1.65	-0.01
西式美食	1.72	0.43	0.81	0.77	0.27	0.36	0.02	-0.32	0.08	-0.81	-1.53	-1.81
鲜食	0.53	-0.02	-0.26	-0.32	-1.23	-1.13	-1.45	-0.68	1.07	1.06	1.24	1.18
休闲	2.47	-0.49	0.75	0.45	-0.72	-0.74	-1.30	0.33	0.04	-0.87	-0.30	0.37
饮料酒水	0.44	-0.40	0.23	0.74	0.07	0.52	1.04	1.35	0.25	-1.03	-2.13	-1.10
营养时尚	-1.93	1.77	0.74	0.39	-0.25	-0.77	1.10	0.46	0.15	-0.08	-0.57	-1.01

3.2. 库存量状态划分

通过观察经标准化后库存数据, 本文将库存量按照下面的标准划分为 5 种状态。① 库存量极少 S1 (-1.2 以下); ② 库存量偏少 S2 (-1.2 至 -0.4); ③ 库存量正常 S3 (-0.4 至 0.4) ④ 库存量偏多 S4 (0.4 至

1.2); ⑤ 库存量极多 S5 (1.2 以上)。库存量具备“无记忆性”的特点, 故可以将其看作是一个以 $\{S1, S2, S3, S4, S5\}$ 为状态空间的马尔可夫链, 据此建立马尔可夫链模型。

3.3. 建立转移概率矩阵

根据各月份的状态数量变化情况, M_{ij} 表示由状态 i 转移到状态 j 的个数, M_i 表示处于状态 i 的个数, 通过计算 $P_{ij} = \frac{M_{ij}}{M_i}$ ($i=1, 2, \dots, n$) [11], 得出一步概率矩阵 P_1 。

$$P_1 = \begin{bmatrix} 0.10 & 0.40 & 0.40 & 0.00 & 0.00 \\ 0.17 & 0.48 & 0.26 & 0.10 & 0.00 \\ 0.04 & 0.33 & 0.48 & 0.12 & 0.04 \\ 0.06 & 0.06 & 0.39 & 0.36 & 0.12 \\ 0.00 & 0.12 & 0.24 & 0.47 & 0.18 \end{bmatrix}$$

通过第一步状态矩阵计算第二步状态矩阵 P_2 。

$$P_2 = (P_1)^2 = \begin{bmatrix} 0.09 & 0.38 & 0.36 & 0.14 & 0.04 \\ 0.12 & 0.39 & 0.36 & 0.12 & 0.04 \\ 0.09 & 0.34 & 0.39 & 0.15 & 0.04 \\ 0.05 & 0.22 & 0.40 & 0.24 & 0.09 \\ 0.06 & 0.19 & 0.37 & 0.29 & 0.10 \end{bmatrix}$$

基于 C-K 方程, 经过矩阵变换后可得到 P_3, P_4, P_5 。

3.4. 库存量预测

以母婴品类为例, 由表 2 数据对应库存量状态划分标准可知, 2021 年 1~5 月的库存状态依次为 S5, S5, S4, S3, S3, 分别对应步长 5~1, 根据各阶步长转移概率矩阵中对应状态的概率, 将各状态下的概率求和, 预测 6 月库存量, 结果见表 3。

Table 3. Prediction and calculation of inventory status of maternal and infant categories

表 3. 母婴品类库存量状态预测计算

初始月	状态	步长	S1	S2	S3	S4	S5
5 月	S3	1	0.04	0.33	0.48	0.12	0.04
4 月	S3	2	0.09	0.34	0.39	0.15	0.04
3 月	S4	3	0.07	0.28	0.38	0.20	0.07
2 月	S5	4	0.08	0.31	0.39	0.19	0.06
1 月	S5	5	0.09	0.32	0.38	0.17	0.06
Σ (概率和)			0.37	1.58	2.02	0.83	0.27

由表数据可知, 在马尔可夫链模型下, 6 月库存量在 S1 状态下的概率为 $0.37/(0.37 + 1.58 + 2.02 + 0.83 + 0.27) = 7.30\%$, 可以算出最大概率在 S3 状态水平为 39.84%。以此类推, 结合 2 月至 6 月的实际库存数据来预测 7 月库存量, 处于 S2 状态下的概率最大, 为 35.48%。重复上述过程, 得出 8 月至 12 月的库存量最大概率均在 S2 状态水平下, 与实际情况相比, 两者符合的概率计算结果为 85.71%, 结果见图 1。

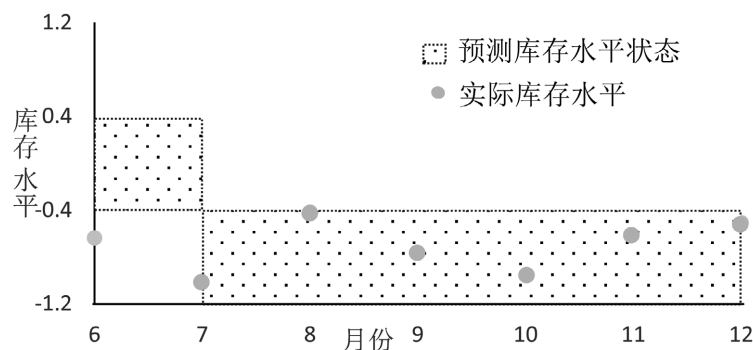


Figure 1. Comparison between actual inventory and predicted value of maternal and infant categories

图 1. 母婴品类实际库存量与预测值对比

本文在模型中初始转移概率矩阵的建立过程中, 不仅考虑了其他品类的库存变化, 还有一年内不同月份的库存变化。将数据进行标准化, 削弱了门店销售状况以及季节变化对产品销售带来的影响。使得该模型在实际应用中具有一定的参考价值。

4. 结论

随着消费者日益增长的个性化需求与缺货容忍度逐步降低, 对于库存需求预测的精准度要求越来越高。本文基于零售业库存量所具有的马尔可夫性质, 以 M 超市整体库存量进行分析研究, 利用企业已有的历史库存数据建立转移概率矩阵, 对各品类的未来库存量进行预测。本研究结果表明, 应用马尔可夫链模型预测的库存数据准确度高达 85.71%, 本模型在一定程度上解决了 M 超市库存管理上的不足, 目前已被 M 超市采纳, 且在试用阶段, 该模型对未来库存数据的预测具有较高的准确性及可靠性。但由于历史数据量的局限性, 在后续模型设计中, 选取更多的数据作为支撑, 预测结果将更具说服力。

基金项目

本文受到教育部青年基金项目(20YJC790118)的资助。

参考文献

- [1] 周永务, 李斐. 新零售运营管理面临的问题与挑战[J]. 系统管理学报, 2022, 31(6): 1041-1055.
- [2] 李敬泉, 满秀芳. 零售业多渠道库存策略研究[J]. 商业研究, 2014(11): 28-33.
- [3] 罗晓萌, 李建斌, 胡鹏. 基于时间序列预测的电子商务库存优化策略[J]. 系统工程, 2014, 32(6): 91-98.
- [4] 杨静雅, 孙林夫. 基于 QPSO-SVR 的售后配件库存需求预测[J]. 计算机工程与设计, 2015, 36(9): 2539-2543+2571.
- [5] Tian, X., Wang, H.Q. and Erjiang, E. (2021) Forecasting Intermittent Demand for Inventory Management by Retailers: A New Approach. *Journal of Retailing and Consumer Services*, 62, Article ID: 102662. <https://doi.org/10.1016/j.jretconser.2021.102662>
- [6] 余维, 荣欣鹏, 刘炜, 田钊. 基于马尔可夫链的生成式区块链隐蔽通信模型[J]. 通信学报, 2022, 43(10): 121-132.
- [7] 苏玉泽, 孟相如, 康巧燕, 韩晓阳. 基于半马尔可夫过程的虚拟网络生存性模型[J]. 科学技术与工程, 2019, 19(24): 260-267.
- [8] 陈梁胜, 邱郡, 赵攸乐. 基于马尔可夫链模型的机械式停车设备安全评估[J]. 起重运输机械, 2022(18): 23-26.
- [9] 陈雯, 吕王勇, 李思奇, 代娟, 邓桢. 两种马尔可夫链状态转移概率矩阵的估计与比较[J]. 重庆理工大学学报(自然科学), 2021, 35(8): 217-223.
- [10] 李栋, 王珊, 任晓菲. 中国数字经济规模预测模型构建[J]. 统计与决策, 2022, 38(10): 5-9.

<https://doi.org/10.13546/j.cnki.tjyj.2022.10.001>

- [11] 翟永, 刘津, 陈杰, 邢绪超, 杜娟, 李恒, 朱杰. 基于马尔可夫链的无桩共享单车车辆投放规模分析[J]. 北京交通大学学报, 2019, 43(5): 27-36.