

中国社会消费品零售总额的时间序列分析

陈鹏蕾¹, 连雨欣², 王 美¹, 武业文^{1*}

¹南京信息工程大学数学与统计学院, 江苏 南京

²吉林建筑大学经济与管理学院, 吉林 长春

收稿日期: 2023年2月28日; 录用日期: 2023年4月13日; 发布日期: 2023年4月21日

摘 要

本文基于2000年1月~2022年12月的社会消费品零售总额数据, 通过时序图和确定性因素分解初步了解该序列具有明显的季节性和趋势性。ARIMA模型为时间序列数据分析中最常用的模型, 在对模型添加季节因素后可以很好地对具有该性质的数据进行建模, Holt-Winters三参数指数平滑模型中的指数包含了季节项和趋势项, 也可以很好地分析和预测该数据, 因此本论文采用ARIMA乘积季节模型和Holt-Winters三参数指数平滑模型对社会消费品零售总额数据进行建模, 并根据建立的模型预测未来五个月(2022年8月~2022年12月)的社会消费品零售总额, 探究不同模型在处理该类数据时的优劣。从结果来看, Holt-Winters三参数指数平滑模型优于ARIMA模型, 我国短时间内社会消费品零售总额仍然将保持增长状态。

关键词

ARIMA模型, Holt-Winters三参数指数平滑模, 趋势预测

Time Series Analysis of Total Retail Sales of Consumer Goods in China

Penglei Chen¹, Yuxin Lian², Mei Wang¹, Yewen Wu^{1*}

¹School of Mathematics and Statistics, Nanjing University of Information Science & Technology, Nanjing Jiangsu

²School of Economics and Management, Jilin Jianzhu University, Changchun Jilin

Received: Feb. 28th, 2023; accepted: Apr. 13th, 2023; published: Apr. 21st, 2023

Abstract

Based on the total retail sales data of social consumer goods from January 2000 to December 2022,

*通讯作者。

文章引用: 陈鹏蕾, 连雨欣, 王美, 武业文. 中国社会消费品零售总额的时间序列分析[J]. 运筹与模糊学, 2023, 13(2): 919-932. DOI: 10.12677/orf.2023.132095

this paper preliminarily understands that the sequence has obvious seasonality and trend through time series diagram and deterministic factor decomposition. ARIMA model is the most used model in the analysis of time series data. After adding seasonal factors to the model, data with this nature can be well modeled. The index in the Holt-Winters three-parameter exponential smoothing model contains seasonal and trend items, which can also analyze and predict the data well, therefore, this paper uses ARIMA product seasonal model and Holt-Winters three-parameter exponential smoothing model to model the total retail sales of social consumer goods, and predicts the total retail sales of social consumer goods in the next five months (August 2022~December 2022) according to the established model, and explores the advantages and disadvantages of different models in processing such data. From the results, the Holt-Winters three-parameter exponential smoothing model is better than ARIMA model, and the total retail sales of consumer goods in China will still maintain growth in a short time.

Keywords

ARIMA Model, Holt-Winters Three-Parameter Exponential Smoothing Model, Trend Prediction

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

社会消费品零售总额是指批发零售业、贸易业、餐饮业、制造业和其他行业向城乡居民和社会集团以及农民提供的消费品零售额的总和。通过该指标可以反映居民和社会集团能否通过多种商品流通渠道获取到足够的生活消费品来满足他们的生活需求，因此是深入研究人民生活、社会消费品购买力、货币流通等问题的重要指标。随着消费环境的逐步改善，人们的消费能力不断增强，从而直接推动社会消费品零售总额发展，面对复杂多变的国内外形势，特别是应对国际金融危机以及新冠肺炎疫情带来的冲击，国家出台了一系列扩大内需、促进消费等政策措施，消费品市场的稳定发展对我国缓冲金融危机起到了明显的积极作用，消费需求已经成为经济增长的重要组成部分。

中国消费品零售业的发展即将进入新阶段，会更深入参与到国际竞争中。零售总额数据会受诸多因素所影响，只有可靠准确的数据系统才能让政府更好地对经济进行宏观决策。因此，预测我国消费品零售总额具有积极意义。

本文基于我国 2000 年 1 月~2022 年 12 月的月社会消费品零售总额数据，利用时间序列分析方法对其进行分析并预测。时间序列分析是根据有限的时间序列数据记录，探究其包含的随着时间变化的信息，建立能够比较准确地反映其动态依存关系的数学模型，并借以对系统的未来行为进行预报。通过建立 ARIMA 模型和 Holt-Winters 三参数指数平滑模型，既得到我国社会消费品零售总额的增长趋势和季节趋势，反映出我国的经济走向以及经济发展水平和繁荣程度；也得到了未来一年内社会消费品零售总额的月数据，可以为政府的宏观决策提供依据。

2. 课题研究概述

2.1. 研究背景

改革开放以来，中国经济保持着持续平稳较快增长的良好态势，居民收入稳步增长。伴随着消费环境的逐步改善，人们的消费能力不断增强，直接带动社会消费品零售总额的发展[1]。

社会消费品零售总额是指批发、零售、住宿、餐饮等行业直接向城乡居民和社会团体销售的消费品零售总额。其中，消费品零售是指向城乡居民销售日常消费的生活用品。社会团体消费品零售是指销售给党政机关、社会团体、军队、学校、企业、事业单位、居民委员会、村民委员会等，主要与非生产性、非经营性消费共用的商品。简单来说，就是涉及城乡居民衣食住行基本消费的销售总额。我们可以看到一段时期内人们物质文化生活水平的提高，反映了社会商品购买力的实现程度和零售市场的规模[2]。通过对社会消费品零售总额的预测，可以看出中国未来的经济发展水平。

2.2. 研究目的及意义

2.2.1. 目的

社会消费品零售总额这个指标充分反映了特定时期内人们的生活水平提升程度、社会零售商品的支付实现情况、社会生产、零售市场的规模概况、货币流通以及物价发展变动趋势。在各类与消费相关的统计数据中，该指标最直观地表现了国内消费需求，研究了国内零售市场的变动态势、宏观经济的景气程度[3]，因此对社会消费品零售总额的发展趋势进行预测，就可以知道未来我国经济发展的大致水平，并可以为未来时间内的宏观调控提供可靠的科学理论支持。

2.2.2. 意义

随着人民生活水平的提高、双重征税的减少、社会良性循环的建立以及法人实体税负减轻，国家从2016年开始全面开展“转业增收”工作，以实现税收征管的合理化和规范化。在这种情况下，全面的“商业改革”是否会对人类消费水平和国民经济水平产生影响？社会消费品零售总额通常是衡量人们消费水平的重要指标之一，也是国民经济发展的指标之一。因此，在复杂的“转型与增长”条件下，研究社会消费品零售总额的发展趋势对我国经济发展具有十分重要的意义。

3. 时间序列方法简介

时间序列分析是基于预测对象的历史和当前时间数据，研究其基本趋势和规律，建立拟合度高的数学模型，并根据模型外推预测对象的定量预测方法，其预测的根本依据是“惯性原理”，因此一般运用该方法预测时包含以下两个假定前提：一是承认预测对象所具备的惯性。利用过去的统计数据，就能延续预测出事物未来几年的数据。二是设定预测对象的趋势性仅与时间这个变量相关。任何事物的发展都会受到偶然因素的影响，因此假定主要考虑的只有一个影响因素 - 时间。其内容包括：其一，收集并整理需要预测数据的历史值；其二，检验所采集的数值，依照次序组成数列；其三，观察时间数列，找到研究指标随时间变化而变化的趋势性；其四，拟合适当的模型来预测该统计指标以后年份的数据值[4]。

3.1. X11 模型

X11 模型可以对序列进行因素分解，分别监测季节波动和趋势效应；其是在基于中心移动平均、Henderson 加权移动平均和 Musgrave 非对称移动平均这三大类移动平均方法，使用多次移动平均反复迭代进行因素分解，其计算流程如下。

迭代第一阶段：

第 1 步：进行 $M_{2 \times 12}$ 复合移动平均，剔除周期效应，得到趋势效应初始估计值。

$$T_t^{(1)} = M_{2 \times 12}(x_t) \quad (1)$$

第 2 步：从原序列 $\{x_t\}$ 中剔除趋势效应，得到季节 - 不规则成分，不妨记作 $\{y_t^{(1)}\}$ 。

$$y_t^{(1)} = S_t^{(1)} \times I_t^{(1)} = \frac{x_t}{T_t^{(1)}} \quad (2)$$

第 3 步: 计算 $\{y_t^{(1)}\}$ 序列的季节指数。

$$S_t^{(1)} = \frac{y_t^{(1)}}{\bar{y}_t^{(1)}} = \frac{M_{3 \times 3}(y_t^{(1)})}{M_{2 \times 12}(y_t^{(1)})} \quad (3)$$

第 4 步: 从原序列 $\{x_t\}$ 中剔除季节效应, 得到趋势 - 不规则成分, 不妨记作 $\{x_t^{(2)}\}$ 。

$$x_t^{(2)} = T_t \times I_t = \frac{x_t}{S_t^{(1)}} \quad (4)$$

迭代第二阶段:

第 5 步: 用 13 期 Henderson 加权移动平均, 并使用 Musgrave 非对称移动平均填补 Henderson 加权移动平均不能获得的最后估计值, 得出趋势效应估计值。

$$T_t^{(2)} = H_{13}(x_t^{(2)}) \quad (5)$$

第 6 步: 从序列 $\{x_t^{(2)}\}$ 中剔除趋势效应, 得到季节 - 不规则成分, 不妨记作 $\{y_t^{(2)}\}$ 。

$$y_t^{(2)} = S_t^{(2)} \times I_t^{(2)} = \frac{x_t^{(2)}}{T_t^{(2)}} \quad (6)$$

第 7 步: 计算 $\{y_t^{(2)}\}$ 序列的季节指数 $S_t^{(2)}$ 。

$$S_t^{(2)} = \frac{y_t^{(2)}}{\bar{y}_t^{(2)}} = \frac{M_{3 \times 3}(y_t^{(2)})}{M_{2 \times 12}(y_t^{(2)})} \quad (7)$$

第 8 步: 从序列 $\{x_t^{(2)}\}$ 中剔除季节效应, 得到季节调整后序列。不妨记作 $\{x_t^{(3)}\}$ 。

$$x_t^{(3)} = T_t \times I_t = \frac{x_t^{(2)}}{S_t^{(2)}} \quad (8)$$

迭代第三阶段:

第 9 步: 根据 $\{x_t^{(3)}\}$ 波动性的大小, 程序自动选择适当期数的 Henderson 加权移动平均, 并使用 Musgrave 非对称移动平均填补 Henderson 加权移动平均不能获得的估计值, 计算最终趋势效应。

$$T_t^{(3)} = H_{2k+1}(x_t^{(3)}) \quad (9)$$

第 10 步: 从 $\{x_t^{(3)}\}$ 中剔除趋势效应, 得到随机波动。

$$I_t^{(3)} = \frac{x_t^{(3)}}{T_t^{(3)}} \quad (10)$$

通过上面三个迭代阶段, 得到的是最终的因素分解结果:

$$x_t = S_t^{(2)} \times T_t^{(3)} \times I_t^{(3)} \quad (11)$$

3.2. ARMA 模型

把具有如下结构的模型称为自回归移动平均模型, 简记为 ARMA(p, q):

$$\begin{cases} x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \\ \phi_p \neq 0, \theta_q \neq 0 \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ E(x_s \varepsilon_t) = 0, \forall s < t \end{cases} \quad (12)$$

当 $\phi_0 = 0$ ，该模型被称为中心化 ARMA(p, q)模型，若缺省默认条件，中心化 ARMA(p, q)模型可以简写为：

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (13)$$

ARMA(p, q)模型可以通过自相关系数和偏自相关系数来描述其统计性质，通过这两个系数，我们可以总结出如下规律：当自相关系数拖尾而偏自相关系数 p 阶截尾，选择 AR(p)模型；当自相关系数 q 阶截尾而偏自相关系数拖尾，选择 MA(q)模型；当自相关系数和偏自相关系数都是拖尾时，选择 ARMA(p, q)模型，具体见表 1。

Table 1. Characteristics of autocorrelation coefficient and partial autocorrelation coefficient of model

表 1. 模型自相关系数和偏自相关系数特征

模型	自相关系数	偏自相关系数
AR(p)	拖尾	p 阶截尾
MA(q)	q 阶截尾	拖尾
ARMA(p, q)	拖尾	拖尾

由于样本存在随机性，所以其相关系数不会呈现出理论上的完美的截尾情况，本应截尾的相关系数仍会呈现出小幅度振荡的情况。平稳时间序列由于通常都具有短期相性，随着延迟阶数的增大，相关系数都会衰减至零值附近，呈现出小值波动的现象。

一个正态分布的随机变量在任意方向上超出两倍标准差的概率约为 0.05。因此，可以通过自相关和偏自相关估计值序列的直方图来大致判断在 0.05 的显著性水平下模型的自相关系数和偏自相关系数不为零的个数，进而大致判断序列应选择的具体模型形式。至于相对最优模型的选择，一般利用 AIC 准则和 BIC 准则评判拟合模型的相对优劣，即使得上述两个 AIC 和 BIC 函数值达到最小的模型为相对最优模型。

3.3. ARIMA 乘积季节模型

ARIMA 模型主要是针对非平稳数据的建模过程，而 ARIMA 季节乘积模型是对 ARIMA 模型进行了分解。

在现实情况中，我们所遇到的时间序列并不是像理论一样，通过简单的周期步长差分以及低阶差分就可以将季节效应与趋势信息提取出来，有时候他们之间存在着复杂的交互影响。这时就需要采用 ARIMA 季节乘积模型来进行拟合[5]。季节乘积模型 ARIMA 的结构如下：

$$\nabla^d \nabla_s^D x_t = \frac{\Theta(B)\Theta_s(B)}{\Phi(B)\Phi_s(B)} \varepsilon_t \quad (14)$$

3.4. Holt-Winters 三参数指数平滑

假设要进行指数平滑的序列为 $\{x_t\}$ ， $\{x_t\}$ 序列既含有趋势因素又含有季节因素，记 a_t 为该序列的水平部分， b_t 为该序列的趋势部分， s_t 为该序列的季节因子(假设一个季节周期长度为 π)，这个季节因子可以随着每年的具体情况波动，构造 Holt-Winters 三参数指数平滑模型。

如果季节和趋势是加法模型，则 Holt-Winters 三参数指数平滑模型构造如下：

$$\begin{aligned} a_t &= \alpha(x_t - s_{t-\pi}) + (1-\alpha)(a_t + b_{t-1}) \\ b_t &= \beta(a_t - a_{t-\pi}) + (1-\beta)b_{t-1} \\ s_t &= \gamma(x_t - a_t) + (1-\gamma)s_{t-\pi} \end{aligned} \quad (15)$$

如果季节和趋势是乘法模型，则 Holt-Winters 三参数指数平滑模型构造如下：

$$\begin{aligned} a_t &= \alpha(x_t/s_{t-\pi}) + (1-\alpha)(a_t + b_{t-1}) \\ b_t &= \beta(a_t - a_{t-\pi}) + (1-\beta)b_{t-1} \\ s_t &= \gamma(x_t/a_t) + (1-\gamma)s_{t-\pi} \end{aligned} \quad (16)$$

4. 实证分析

4.1. 确定性因素分解

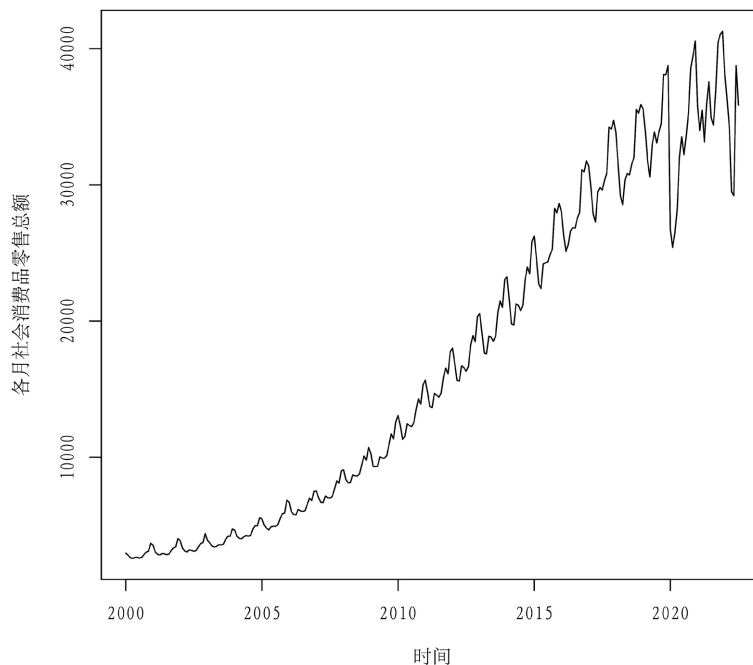


Figure 1. Timing chart of total retail sales of social consumer goods

图 1. 社会消费品零售总额时序图

通过图 1 的时序图中可以对社会消费品零售总额有一个整体的把握，总体来看社会总消费品零售总额有逐年增长趋势，带有明显的趋势性，因此可以初步判断该序列是非平稳的，同时，我们还发现社会消费品零售总额呈现上升 - 回落 - 上升的周期性，其中 2020 年的一个大幅度下降是受新冠肺炎疫情的影响。如此动态上升，则表明该序列具有明显的季节性；接下来对数据进行确定性因素分解。

对该序列进行确定性因素分解，得到图 2。其中 seasonal 行表示季节性的波动，能够很明显地观察到季节性；trend 行表示趋势变化，可以发现一个很明显的向上增长的趋势；random 行表示随机波动的部分，围绕在 1 附近上下波动。综上所述，该序列具有明显的季节性和趋势性。

将季节指数绘制成图 3，可以清晰地观看到月度变迁对社会消费品零售总额的影响。在上图中可以看出，季节指数呈现出先下降后上升的趋势。季节指数的大小表明了月社会零售销售总额的高低，1 月、12 月的季节指数明显高于其他月份，这说明 1 月和 12 月的社会零售销售总额明显高于其他月份。由于 12 月和 1 月此阶段正值春节期间，人们置办年货消费大量的商品，促使整个社会消费品零售总额增长，因此出现一个波峰；7 月和 8 月往往人们的消费心理并不是那么活跃，主要受到传统消费的影响，没有明确的消费目标；8 月份之后又有了明显的上升，此期间节日较多，如国庆节、“双 11”、圣诞节、元

且等，在节日期间人们往往购买大量的商品，“双11”对于消费的带动更是如此。并且在过节期间商家销售力度会加大，更加迎合人们的消费心理[6]。

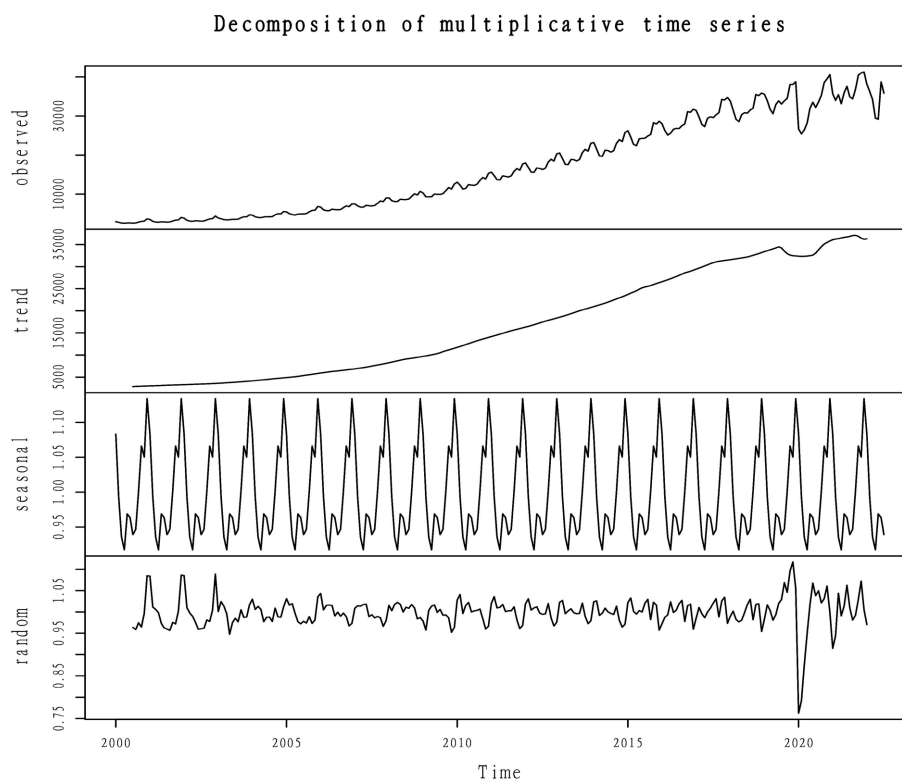


Figure 2. Sequence breakdown diagram
图 2. 序列分解图

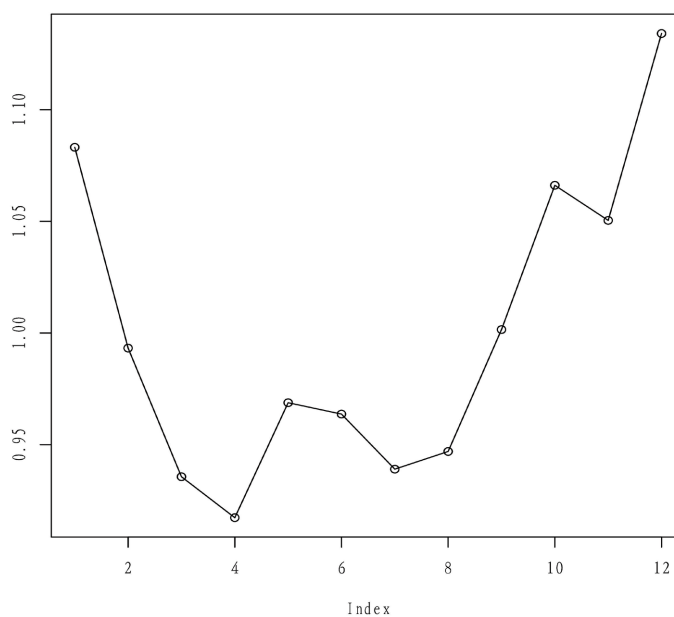


Figure 3. Seasonal index chart of total retail sales of consumer goods in China
图 3. 中国社会消费品零售总额序列季节指数图

4.2. ARIMA 模型建立及预测

接下来基于 2000 年 1 月~2022 年 7 月的社会消费品零售总额数据建立时间序列模型，并基于该模型对未来 5 个月(2022 年 8 月~2022 年 12 月)社会消费品零售总额数据进行预测，同时对比预测值和真实值来检验模型预测效果。

4.2.1. 平稳性检验

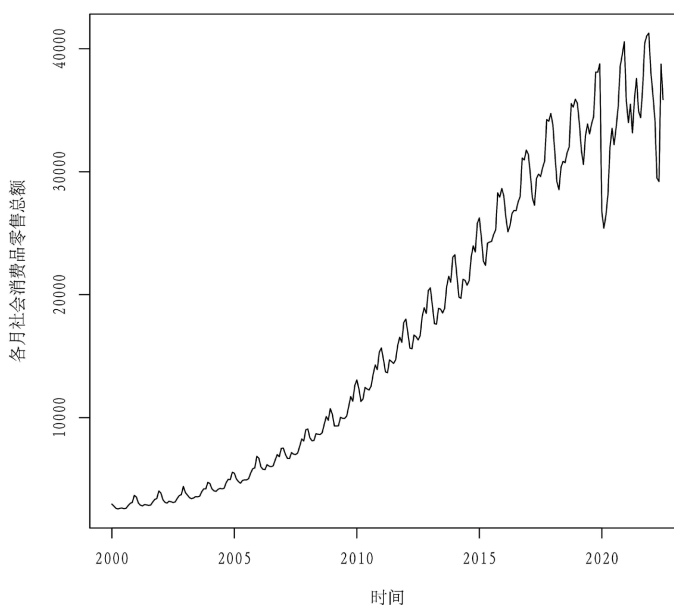


Figure 4. Timing chart of total retail sales of consumer goods from January 2000 to July 2022

图 4. 2000 年 1 月~2022 年 7 月社会消费品零售总额时序图

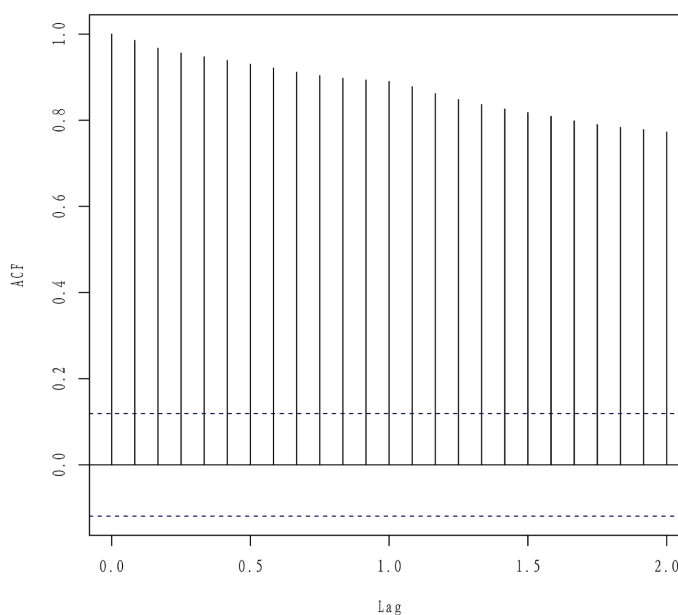


Figure 5. Autocorrelation coefficient graph of original sequence

图 5. 原序列的自相关系数图

从时序图(图 4)可以看出, 序列存在明显的趋势性, 由自相关系数图(图 5)可以发现该序列的自相关系数递减的过程十分缓慢, 在很长的延迟时期内, 自相关系数始终为正, 且始终在 2 倍的标准差之外, 没有很强的短期相关性。因此, 综合时序图和自相关系数可以得出该序列是非平稳序列, 通过差分处理将该序列转换为平稳序列。

4.2.2. 差分处理

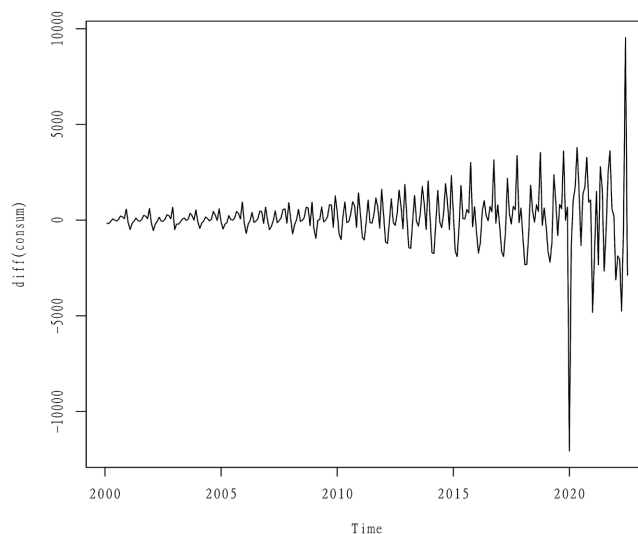


Figure 6. Sequence diagram of first-order differential sequence
图 6. 一阶差分序列时序图

从上述时序图(图 4)可以看出社会总消费品零售总额序列蕴含一个近似线性的递增趋势, 因此对该序列进行 1 阶差分运算如图 6, 从图中可以看出该序列 1 阶差分后的线性递增信息被提取出来, 但还存在稳定的季节波动和随机波动, 因此再进行 12 步的周期差分, 提取季节波动信息。

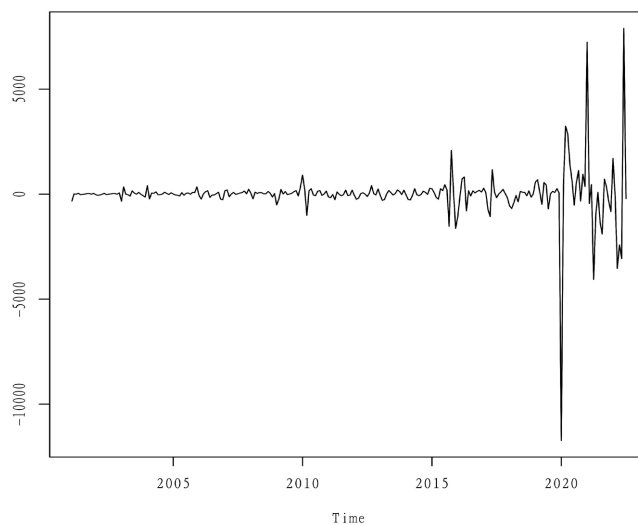


Figure 7. Sequence diagram after first-order difference plus 12-step difference
图 7. 一阶差分加 12 步差分后的时序图

从图 7 的时序图中可以看出差分后的数据始终围绕在 0 附近上下波动，说明差分后的序列没有了明显趋势或周期，只是在 2020 年往后的波动幅度会稍稍变大，但可以视为平稳序列；也说明了周期差分可以非常好的提取周期信息，1 阶差分加 12 步差分能够提取出原序列中蕴含的长期趋势效应和季节效应等确定性因素。

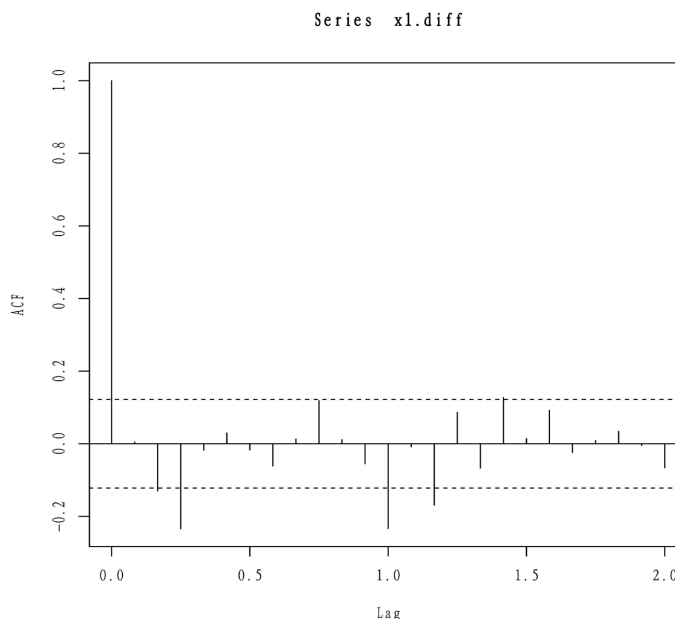


Figure 8. Autocorrelation coefficient diagram after difference
图 8. 差分后自相关系数图

对差分后的序列绘制自相关系数图(图 8)，该序列的自相关系数大多数比较小，且始终控制在 2 倍的标准差范围之内，可以认为该序列自始至终都在零轴附近波动，这是随机性非常强的平稳时间序列通常具有的自相关图特征，因此差分后的序列通过平稳性检验。

4.2.3. 白噪声检验

Table 2. White noise test results
表 2. 白噪声检验结果

延迟阶数	X-squared	P 值
6	18.887	0.004359
12	38.412	0.0001315

对此数据进行白噪声检验，显著性水平为 0.05。根据表 2 的检验结果，得到前 6 期和前 12 期的 P 值分别为 0.004359、0.0001315，远远小于显著性水平，因此拒绝原假设，认为该序列不属于纯随机波动，差分后序列通过白噪声检验。

4.2.4. 模型定阶

从时序图中可以看出，该序列显示出随着趋势的递增，每个季节的振幅也在增大，说明季节效应受到趋势的影响，这时通常会选择建立乘积季节模型。

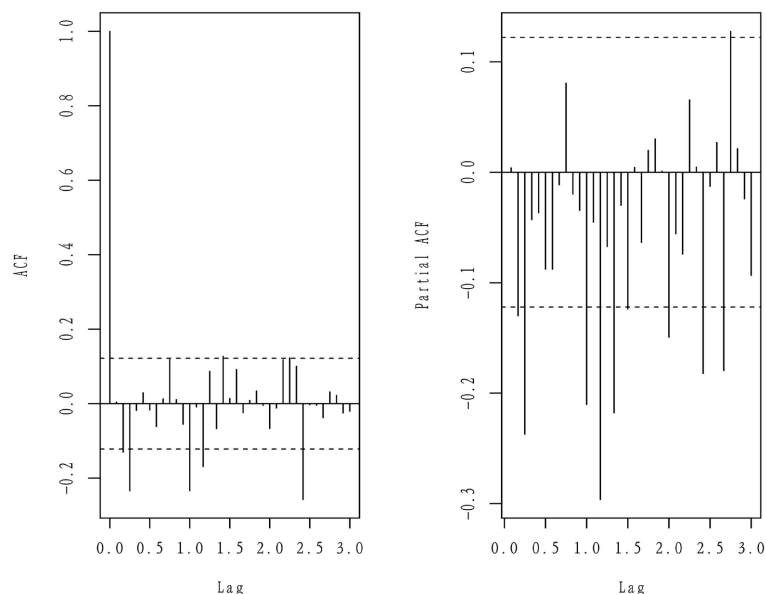


Figure 9. Autocorrelation coefficient diagram and partial autocorrelation coefficient diagram
图 9. 自相关系数图和偏自相关系数图

首先确定短期相关模型的阶数，从偏自相关图中可以看出，偏自相关系数呈现拖尾的现象，其衰减到零的过程中反复出现大幅度的波动。从自相关图中，自相关系数在 2、3 阶都超过了 2 倍标准差范围，且 3 阶超过 2 倍标准差的现象十分明显，而后绝大部分的自相关系数落在 2 倍的标准差范围内，而且由非零自相关系数衰减至零附近的过程比较突然，因此可以认为自相关系数 3 阶截尾，因此用 ARMA(0, 3) 来提取差分后序列短期自相关信息。

在考虑季节自相关特征时，考察延迟 12 阶、24 阶等以 12 为周期长度单位的自相关系数和偏自相关系数的特征(其在横轴上表现的刻度分别为 1、2 阶延迟系数)。可以看出图 9 左边的自相关图延迟 12 阶、13 阶和 24 阶的自相关系数显著非零，且 24 阶自相关系数最为明显超过了 2 倍标准差范围，其余绝大部分自相关系数落入 2 倍标准差范围内。由非零自相关系数衰减至零附近的过程比较突然，因此可以视为自相关系数截尾，截尾的阶数为 2；图 9 右边的偏自相关图显示延迟 12 阶和 24 阶的偏自相关系数都显著非零，且偏自相关系数始终在零值附近保持较大幅度的波动。综上所述，可以认为季节自相关特征是自相关系数 2 阶截尾，偏自相关系数拖尾，这时以 24 步为周期的 ARMA(0, 1) 24 模型提取差分后序列的季节自相关信息。

4.2.5. 拟合乘积季节模型——ARIMA(0, c(1, 24), 3) × (0, 1, 1)

Table 3. Product seasonal model results
表 3. 乘积季节模型结果

模型系数	ma1	ma2	ma3	sma1
均值	-0.1890	-0.4463	-0.1096	-0.3961
方差	0.0644	0.0686	0.0651	0.0654

根据自相关系数图和偏自相关系数图建立 ARIMA(0, c(1, 24), 3) × (0, 1, 1)模型，表 3 为模型拟合的结果，根据模型拟合输出结果，我们得到的拟合模型为：

$$\nabla\nabla^{24}x_t = (1 - 0.189B - 0.4463B^2 - 0.1096B^3)(1 - 0.3961B^{24})\varepsilon_t$$

为了检验模型是否可以用来预测,对模型的残差进行白噪声检验,同时对模型系数进行显著性检验,显著性水平为 0.05。

Table 4. Model residual white noise test results

表 4. 模型残差白噪声检验结果

延迟阶数	X-squared	P 值
6	1.3888	0.9665
12	16.227	0.1811

表 4 是对乘积季节模型的残差进行白噪声检验,得知检验结果 P 值为 0.9665 和 0.1811,均大于 0.05,因此接受原假设,认为残差为白噪声。

以下将对模型的系数进行显著性检验,共有 4 个系数,它们的 P 值分别为 0.001829433, 2.19988e-10, 0.04677773, 2.638146e-09,所有系数的 P 值明显小于 0.05,均通过了系数的显著性检验,因此该乘积季节模型可行。

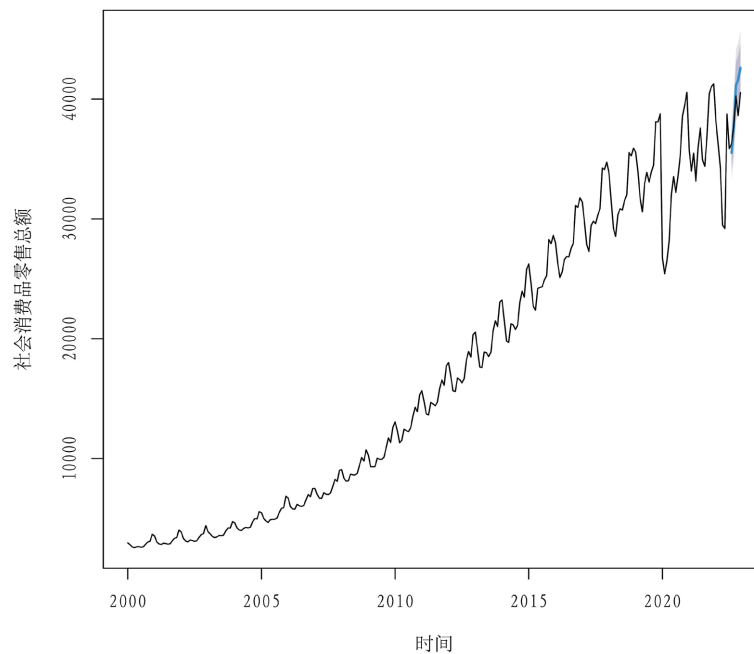


Figure 10. Forecast trend chart from November 2021 to March 2022

图 10. 2021.11~2022.03 月的预测趋势图

Table 5. Model prediction results

表 5. 模型预测结果

年 - 月	预测值	80%置信区间下限	80%置信区间上限	95%置信区间下限	95%置信区间上限
2022-8	35508.91	34055.81	36962.01	33286.59	37731.24
2022-9	37630.90	35759.97	39501.83	34769.56	40492.24
2022-10	41160.07	39215.53	43104.62	38186.15	44134.00
2022-11	41617.11	39637.53	43596.70	38589.60	44644.63
2022-12	42605.99	40591.98	44620.00	39525.82	45686.15

图 10 为乘积季节模型的预测结果图，表 5 给出了模型预测的预测值和置信区间。从图中可以看出，除了 11 和 12 月份的预测结果值与实际有较明显偏差，其余月份的预测效果都很好。通过将预测值与真实值作对比，得到平均绝对误差为 1363.652，误差值相对较小。

4.3. Holt-Winters 三参数指数平滑预测

在对序列进行确定性因素分解时发现，序列受到明显的长期趋势，季节效应和随机波动这三个因素的影响，由于 Holt-Winters 三参数模型指数含有水平项、趋势项、季节项，是用于预测含趋势项、季节项、随即项的时间序列，因此对该序列应用 Holt-Winters 三参数指数模型并进行预测。

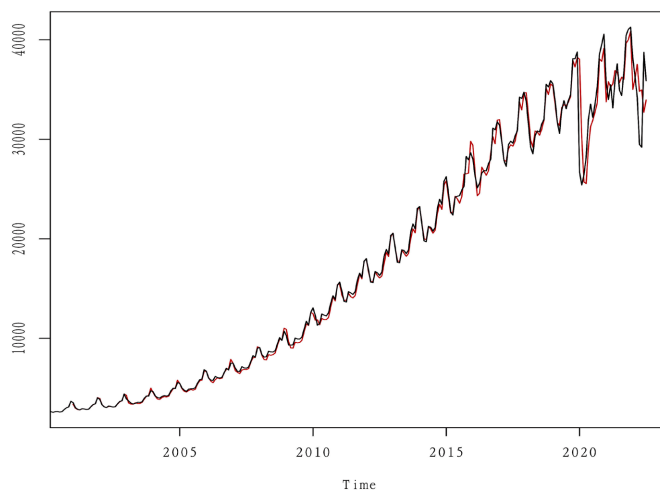


Figure 11. Holt-Winters three-parameter exponential smoothing fitting effect diagram

图 11. Holt-Winters 三参数指数平滑拟合效果图

首先对此序列进行绘制 Holt-Winters 三参数指数平滑拟合效果图，如图 11，黑线为原始数据的时序图，红线为拟合后的时序图，可以看出红线与黑线的趋势大致重合，拟合效果较为理想。

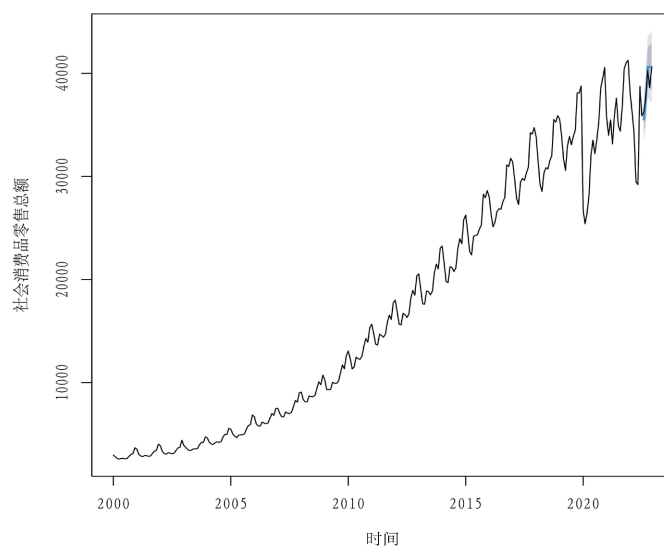


Figure 12. Forecast trend chart from 2022.8 to 2022.12

图 12. 2022.8~2022.12 的预测趋势图

根据 Holt-Winters 三参数指数平滑预测五个月社会消费品零售总额得到结果如图 12, 可以发现与 ARIMA 模型预测趋势图大致相符, 且该模型在测试集中的平均绝对误差为 702.77, 该误差值与乘积季节模型的误差值相比减少了接近一半, 因此 Holt-Winters 三参数指数平滑更适合用于该数据的拟合预测, 更能对数据中的趋势性和周期性做出准确的判断识别。

4.4. 误差分析

通过将实际结果与预测结果相对比, 2022 年 9 月的预测是最准确的, 所有月份的真实值均在预测值的 95%置信区间内, 真实值总体上低于预测值, 这是因为, 根据往期趋势来看, 用于预测的月份在周期内属于上升阶段, 而模型放大了其上升趋势, 导致总体上预测值高于真实值, 且在真实值中, 11 月的数据是下降的, 该趋势没有被模型捕捉到。

5. 总结

本文基于 2000 年 1 月~2022 年 8 月的社会消费品零售总额, 采用乘积季节模型对其建模, 建立 $ARIMA(0, c(1, 24), c(2, 3)) \times (0, 1, 1)$ 模型, 该模型口径为 $\nabla \nabla^{24} x_t = (1 - 0.189B - 0.4463B^2 - 0.1096B^3)(1 - 0.3961B^{24})\varepsilon_t$, 并根据此模型预测未来五个月的社会消费品零售总额。除此之外还采用了 Holt-Winters 三参数指数平滑模型来进行建模预测, 得到的结果优于乘积季节模型, 说明在分析此类数据时, Holt-Winters 三参数指数平滑更能够提取出序列中的趋势性和周期性。

现社会消费品零售总额一直呈增长状态, 且增长速度加快, 由社会消费品零售总额的增长趋势可以反映出我国的经济走向以及经济发展水平和繁荣程度, 说明我国经济发展水平在稳步增长。但在此增长过程中还存在着小幅度的波动, 这要求我国要不断完善相关制度及政策, 来稳定其他因素对经济的影响, 保证经济的稳步增长。

参考文献

- [1] 罗中德, 赖美艳. 中国社会消费品零售总额的预测分析[J]. 统计与决策, 2013(2): 143-145.
- [2] 张懿滢. 关于我国社会消费品零售总额的统计分析[J]. 现代商业, 2015(24): 127-128.
- [3] 单筱婷. 厦门构建区域性消费中心的路径与政策研究[D]: [硕士学位论文]. 厦门: 厦门大学, 2014.
- [4] 马爱华, 杨右栋, 李莉. 基于时间序列分析法对社会消费品零售总额的预测——以甘肃省张掖市为例[J]. 经济研究导刊, 2019(12): 61-63.
- [5] 王镜乔. 社会消费品零售总额的预测——基于季节 ARIMA 模型[J]. 时代金融, 2018(24): 220-221.
- [6] 徐金红, 刘亚倩. 基于季节模型的中国社会消费品零售总额分析[J]. 河北企业, 2019(1): 62-63.