

电商仓储缺货损失与库存成本协同优化

——基于共轭梯度法的帕累托前沿求解

贺 灿

贵州大学数学与统计学院, 贵州 贵阳

收稿日期: 2026年8月4日; 录用日期: 2026年8月18日; 发布日期: 2026年10月15日

摘 要

在供应链快速发展与电商需求高波动背景下, 库存缺货损失与运营成本的权衡成为电商仓储管理的核心难题。本文针对随机需求与固定补货提前期场景, 构建以最小化年期望缺货损失、最小化总库存运营成本为双目标的库存优化模型, 以安全库存系数与订货批量为决策变量, 通过对数变换将原问题转化为无约束光滑向量优化问题, 并采用多目标共轭梯度法结合强Wolfe线搜索求解帕累托前沿。基于Kaggle物流仓储数据集进行参数标定与数值实验, 得到35个非支配帕累托最优解, 结果表明缺货损失与库存成本随服务水平提升呈边际递增替代关系, 且模型在参数 $\pm 20\%$ 扰动下具有良好稳定性。依据帕累托前沿将决策方案划分为高服务保障、均衡稳健、低成本运营三类区间, 提出商品差异化履约、精益补货、大促弹性调度及缺货成本优先标定等实操策略, 可为电商企业从经验式补货转向数据驱动、动态优化的精细化库存管理提供量化方法与决策支撑, 助力企业在库存成本、资金周转与客户服务体验之间实现科学均衡。

关键词

电商库存, 双目标优化, 共轭梯度法, 帕累托前沿, 安全库存

Collaborative Optimization of Stockout Loss and Inventory Cost in E-Commerce Warehouses

—Pareto Frontier Solution Based on Conjugate Gradient Method

Can He

College of Mathematics and Statistics, Guizhou University, Guiyang Guizhou

Received: August 4, 2026; accepted: August 18, 2026; published: October 15, 2026

Abstract

Against the background of rapid supply chain development and highly volatile e-commerce demand, the trade-off between stockout loss and operating cost has become a core issue in e-commerce warehouse management. This paper addresses the scenario of stochastic demand and fixed replenishment lead time, and constructs a bi-objective inventory optimization model aiming to minimize the annual expected stockout loss and the total inventory operating cost. Taking the safety stock coefficient and order quantity as decision variables, the original problem is transformed into an unconstrained smooth vector optimization problem through logarithmic transformation. A multi-objective conjugate gradient method combined with strong Wolfe line search is adopted to solve the Pareto frontier. Based on the Kaggle logistics warehouse dataset, parameter calibration and numerical experiments are carried out, and 35 non-dominated Pareto optimal solutions are obtained. The results show that the stockout loss and inventory cost exhibit a marginally increasing substitution relationship as the service level improves, and the model remains robust under $\pm 20\%$ parameter perturbations. According to the Pareto frontier, the decision schemes are divided into three categories: high service guarantee, balanced robustness, and low-cost operation. Practical strategies such as differentiated commodity fulfillment, lean replenishment, elastic inventory scheduling during promotion periods, and prior calibration of stockout cost are proposed. This study can provide quantitative methods and decision support for e-commerce enterprises to transform from empirical replenishment to data-driven and dynamically optimized refined inventory management, helping them achieve a scientific balance among inventory cost, capital turnover and customer service experience.

Keywords

E-Commerce Inventory, Bi-Objective Optimization, Conjugate Gradient Method, Pareto Frontier, Safety Stock

Copyright © 2026 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

在供应链高速发展背景下,库存管理直接影响企业运营效率、资金周转与客户满意度。实际经营中,市场需求存在随机波动,采购与运输环节具有固定提前期,供需双重不确定性易引发库存短缺或积压,导致缺货赔付、客户流失与仓储资金过度占用。电商生鲜前置仓更面临品类繁多、库容有限、履约时效严苛等约束,缺货不仅造成订单损失,更损害长期活跃度与品牌声誉;头部平台实践亦表明,智能补货优化能显著提升运营效率与消费者体验。

传统库存优化研究多以总成本最小化为单一目标,将缺货损失以惩罚项形式纳入成本函数,难以单独呈现缺货风险与服务水平的实际状态及其动态权衡关系。部分多目标研究采用线性加权法将多目标转化为单目标求解,但权重赋值依赖主观判断,不同决策者的偏好差异可导致结果显著偏离,策略的可复现性与一致性难以保证,决策参考价值有限。Qi 等[1]的研究亦指出,传统的“先预测-再优化”两阶段方法在面对大规模、高波动的电商需求时存在结构性效率损失,端到端的数据驱动框架正成为新的探索方向。

多目标优化通过帕累托最优解集刻画目标间的天然权衡,无需预设权重,可输出一组彼此非支配的均衡方案,适配不同风险偏好与经营战略的决策需求。共轭梯度算法具有存储量小、收敛速度快、迭代稳定性强的优势,Fliege 和 Svaiter [2]建立了多目标最速下降的理论基础, Lucambio Pérez 和 Prudente [3]

进一步将共轭梯度法推广至向量值优化问题，为其在多目标连续优化中的应用提供了理论支撑。

本文将共轭梯度算法拓展至双目标优化场景，引入强 Wolfe 线搜索准则保证迭代的下降性与计算精度，基于真实物流数据集完成参数标定，构建以最小化期望缺货损失和库存持有与订购成本为双目标的优化模型。通过多起点多目标共轭梯度法求解帕累托前沿，分析收敛点分布与前沿形态变化规律，并通过敏感性分析检验关键参数对结论的影响，从而为随机需求环境下的电商库存优化决策提供科学参考。

2. 数学建模

2.1. 问题描述

本文关注的核心问题是在随机需求与固定补货提前期并存的电商仓库环境下，如何在 (R, Q) 连续盘点策略框架下合理设定安全库存系数与订货批量，以平衡缺货带来的客户流失与库存持有带来的运营成本压力。随着电商促销常态化、用户需求波动日益剧烈，这一矛盾日趋尖锐。为此，本研究构建了一个面向电商仓库的双目标优化模型：一个目标是最小化年期望缺货损失，反映订单履约能力与客户服务水平；另一个目标是最小化总库存运营成本，由周期库存持有成本、安全库存持有成本与订货固定成本三部分构成。

2.2. 模型符号、参数定义

为清晰阐述研究问题并构建数学模型，本节统一界定模型所用各类符号与参数。表 1 汇总了模型中主要的符号及其含义。

Table 1. Nomenclature
表 1. 模型符号说明

符号	含义
μ	日均需求量
σ	日需求标准差
L	采购提前期
σ_L	提前期需求标准差
D_y	年度总需求量
h	单位产品年度库存持有成本
K_f	单次订货固定成本
c_s	单位产品缺货损失成本
z	安全库存系数
Q	单次订货批量

2.3. 目标函数构建

本文构建双目标优化模型，设决策变量集合为 S ，目标一为最小化年期望缺货损失 $f_1(x)$ ，目标二为最小化总库存运营成本 $f_2(x)$ 。

2.3.1. 期望缺货损失目标

引入标准正态损失函数[4]描述随机需求下的缺货程度：

$$L(z) = \phi(z) - z(1 - \phi(z))$$

(1)

其中， $\phi(x)$ 为标准正态分布概率密度函数， $\varphi(x)$ 为标准正态分布累积分布函数，表示服务水平。

假设提前期内总需求 $X \sim N(\mu L, L\sigma^2)$ ，则每次补货周期的期望缺货量为 $\sigma\sqrt{L} \cdot L(z)$ ，年期望订货次数约为 $\frac{D_y}{Q}$ ，故年期望缺货损失为：

$$f_1(z, Q) = c_s \cdot \frac{D_y}{Q} \cdot \sigma\sqrt{L} \cdot L(z) \quad (2)$$

2.3.2. 总库存运营成本目标

在连续盘点 (R, Q) 策略[5]下，库存量随时间呈现“锯齿形”变化：当库存降低到 ROP ，触发一次订货，批量为 Q ；在提前期 L 内，安全库存 $SS = z \cdot \sigma_L$ 是用来缓存需求和供应的不确定性的；周期库存是指两次到货之间，库存从 $Q + SS$ 降到 SS 的平均值，即 $Q/2$ 。其中标准的 ROP 公式为：

$$ROP = L\mu + z \cdot \sigma\sqrt{L} \quad (3)$$

1) 年均持有成本。

平均库存 $S_a = Q/2 + z \cdot \sigma_L$ 。乘以单位年持有成本 h ，得到年均持有成本 m_y 为：

$$m_y = h \cdot \left(\frac{Q}{2} + z \cdot \sigma_L \right) \quad (4)$$

2) 年订货固定成本。

年期望订货次数年期望订货次数约为 $\frac{D_y}{Q}$ 。乘以每次订货的固定成本 K_f ，得到年订货固定成本 m_f 为：

$$m_f = \frac{D_y}{Q} \cdot K_f \quad (5)$$

3) 总库存运营成本。

将年均持有成本与年订货固定成本相加，即得第二个优化目标：

$$f_2(z, Q) = h \cdot \left(\frac{Q}{2} + z \cdot \sigma_L \right) + \frac{D_y}{Q} \cdot K_f \quad (6)$$

2.4. 优化模型

双目标优化问题可表述为：

$$\min_{z \in R, Q > 0} F(z, Q) = (f_1(z, Q), f_2(z, Q))^T \quad (7)$$

为消除 $Q > 0$ 的边界约束，引入对数变换 $u = \ln Q$ ，则 $Q = e^u$ ，且 $u \in R$ 。令 $x = (z, u)^T$ ，得到完全无约束的光滑向量优化问题：

$$\min_{x \in R^2} \tilde{F}(x) = (f_1(z, e^u), f_2(z, e^u))^T \quad (8)$$

通过对数变换，原问题的任意满足 $Q > 0$ 的可行解与 R^2 中的点一一对应，且不改变原问题的帕累托最优解集。

3. 算法设计

在多目标优化中，常见求解方法包括加权和法、 ε -约束法、NSGA-II 等算法。加权和法结构简单，但需要预先设定权重，难以完整刻画帕累托前沿的形态； ε -约束法可生成均匀解，但需要多次求解单目标子

问题, 计算效率较低; NSGA-II 算法无需梯度信息, 可处理非凸、非光滑问题, 但迭代次数多、收敛速度较慢。相比之下, 共轭梯度法具有存储量小、收敛速度快、迭代稳定性强等优势, 特别适合本文构建的无约束、光滑、低维双目标库存优化模型。该模型经过对数变换后具有连续可微的梯度结构, 满足共轭梯度法的使用条件; 同时, 共轭梯度法仅需存储当前与前一步的梯度和方向信息, 计算量小、收敛更快, 能够在较少迭代步数内稳定收敛到帕累托驻点。综合计算效率、对问题特性的适配性以及求解质量与稳定性考虑, 本文以共轭梯度法求解上节所建立的无约束光滑向量优化问题。具体的算法流程如下:

Step1: 梯度计算。

目标 $f_1(x)$ 的梯度:

$$\begin{cases} \frac{\partial f_1}{\partial z} = c_s \cdot \frac{D_y}{Q} \cdot \sigma_L \cdot (\varphi(z) - 1) \\ \frac{\partial f_1}{\partial u} = -c_s \cdot \frac{D_y}{Q} \cdot \sigma_L \cdot L(z) \end{cases} \quad (9)$$

目标 $f_2(x)$ 的梯度:

$$\begin{cases} \frac{\partial f_2}{\partial z} = h \cdot \sigma_L \\ \frac{\partial f_2}{\partial u} = \frac{h}{2} \cdot Q - \frac{D_y}{Q} \cdot K_f \end{cases} \quad (10)$$

Step2: 种群初始化。

在策略空间内随机生成 $N = 250$ 个初始点。安全库存系数 z 的取值范围为 $[-1.5, 3.5]$, 对应的服务水平覆盖率约 7.01% 至 99.97%; 订货批量 Q 的取值范围为 $[10, 300]$ 件, 涵盖了从极小批量到接近数据集隐含补货量上限的合理区间。经过对数变换得到初始搜索点 $x_i^{(0)} = (z_i^{(0)}, \ln Q_i^{(0)})^T$, $i = 1, \dots, N$ 。随机种子固定为 42, 确保了实验结果的可复现性。

Step3: 个体评价。

对当前搜索点 $x^{(k)}$, 计算目标向量 $\tilde{F}(x^{(k)}) = (f_1, f_2)^T$ 及梯度矩阵 $J^{(k)} = [\nabla f_1^{(k)}, \nabla f_2^{(k)}] \in R^{2 \times 2}$ 。

Step4: 构造下降方向。

定义如下的 $H(x, d)$ 函数:

$$H(x, d) = \sup \left\{ \langle J^{(k)} d, p \rangle \mid p \in X \right\} \quad (11)$$

其中 X 为一个基集。

方向按以下规则更新:

$$d^{(k)} = \begin{cases} v(x), & k = 0 \\ v(x) + \beta^{(k)} d^{(k-1)}, & k \geq 1 \end{cases} \quad (12)$$

其中方向 $v(x)$ 定义为:

$$v(x) = \arg \min \left\{ H(x, d) + \frac{\|d\|^2}{2} \mid d \in R^n \right\} \quad (13)$$

参数 $\beta^{(k)}$ 定义为:

$$\beta^{(k)} = \max \left\{ \frac{-H(x^{(k)}, v(x^{(k)})) + H(x^{(k-1)}, v(x^{(k)}))}{-H(x^{(k-1)}, v(x^{(k-1)}))}, 0 \right\} \quad (14)$$

Step5: 非精确线搜索。

沿方向 $d^{(k)}$ 执行满足 Wolfe 条件的非精确线搜索, 要求两个目标函数同时满足充分下降与曲率条件, 步长 $\alpha_k > 0$ 。若搜索失败或步长过小, 终止该路径。

Step6: 更新与裁剪。

接受步长 α_k 后, 更新 $x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$ 。同时为维持数值稳定性, 将 u 分量裁剪至 $[\ln 1, \ln 10^6]$, 即限制 $Q \in [1, 10^6]$ 件, 防止指数映射溢出。

Step7: 终止判定。

- 1) 梯度范数充分小(小于 10^{-6}), 表明达到帕累托驻点;
- 2) 两梯度近乎反向, 表明两目标在该点强烈冲突, 无法继续共同改进;
- 3) 线搜索失效: 步长过小或方向非下降。

Step8: 帕累托前沿提取。

所有起点收敛后, 对所得解集进行非支配排序, 采用“支配计数”方法: A 支配 B 当且仅当 $f_1(A) \leq f_1(B)$ 且 $f_2(A) \leq f_2(B)$ 且至少有一个不等式严格成立。剔除所有受支配解, 剩余非支配解按 f_1 升序排序, 即构成无约束光滑向量优化问题的近似帕累托前沿, 对应原问题的帕累托最优安全库存策略集。

4. 数值检验

4.1. 数据来源

本研究采用 Kaggle 公开的 Logistics Warehouse Dataset 数据集, 该数据集包含 3204 条仓储物品记录, 涵盖库存水平、需求特征、成本结构、履约表现等字段, 属于典型的电商仓储运营数据。

4.2. 参数提取

模型所需参数全部从该数据集中计算提取或由数据集相似数据作近似计算提取, 具体如表 2 所示。

Table 2. Model parameter calibration

表 2. 模型参数标定

参数	取值
μ	25.436 件
σ	5.260 件
L	6 天
σ_L	12.885
D_y	9284 件
h	381.787 元/(件·年)
K_f	27.771 元/次
c_s	15.494 元/件

缺货成本采用每单毛利乘以失约率进行近似计算, 具体计算公式如下:

$$c_s = (p - h_c) \cdot (1 - r) \quad (15)$$

其中 p 为单价, h_c 为单位处理成本, r 为履约率。

订购固定成本由单位处理成本聚合估算。所有参数均来自同一数据集, 确保了内部一致性。

4.3. 实验结果与分析

全部实验在 Python 2026 环境下完成。250 个初始点全部成功收敛，经非支配排序共获得 35 个帕累托最优解，表 3 列出了 17 组具有代表性的帕累托最优解的详细结果。

Table 3. Pareto optimal solution
表 3. 帕累托最优解

方案	z	Q (件)	缺货损失 f_1 (元/年)	库存成本 f_2 (元/年)	服务水平
P1	3.350	297.116	0.660	74062.160	99.96%
P2	3.318	292.574	0.750	73054.054	99.95%
P8	3.333	32.543	6.380	30531.842	99.96%
P9	3.244	37.255	7.810	29992.220	99.94%
P10	3.041	39.906	15.330	29039.460	99.88%
P11	2.887	39.326	26.720	28263.363	99.81%
P15	2.015	61.457	245.880	25839.633	97.81%
P16	2.141	36.473	293.220	24563.811	98.39%
P17	2.040	38.181	369.650	24078.180	97.93%
P18	1.726	26.692	1191.930	23244.542	95.78%
P21	1.197	30.115	3476.080	20197.088	88.43%
P25	0.587	44.268	7211.630	17162.571	72.14%
P26	0.205	51.468	10970.760	15844.451	58.13%
P30	-0.720	59.153	26883.550	12109.293	23.58%
P33	-1.127	55.367	39909.650	9680.840	12.98%
P34	-1.182	43.661	52653.010	8424.432	11.86%
P35	-1.475	26.328	105996.160	7564.635	7.01%

按缺货损失 f_1 升序排列，构成光滑、连续且分布均匀的帕累托前沿，如图 1 所示。前沿整体呈外凸下降形态，边际替代率逐步递减，表明每降低一单位缺货损失所需付出的库存成本增量随着服务水平提高而加速上升。

根据表 3 和图 1 可知，帕累托前沿包含多组非支配最优方案，覆盖电商不同经营决策场景，可为平台及商家提供多元化、灵活的决策参考：

1) 高服务水平方案：位于帕累托前沿左侧区域，对应目标 f_1 取极小值、目标 f_2 取较大值的非支配解，典型代表为 P1 至 P9。在此区域内，服务水平均维持在 99.9% 以上，缺货损失被压制在 7.810 元/年以内的极低水平，但库存成本高达 29992.220~74062.160 元/年。例如，极端方案 P1 ($z = 3.350$, $Q = 297.116$) 将缺货损失控制在 0.660 元/年，而其库存持有与订购成本却为所有方案之最(74062.160 元/年)。此类方案以高昂的库存投资换取近乎绝对的货源可得性，特别适配注重消费者购物体验、严控缺货断货风险的电商经营主体，常见于生鲜电商、医药电商、即时零售等刚需品类赛道。

2) 低成本运营方案：位于帕累托前沿右侧区域，以目标 f_2 取较小值、 f_1 取较大值为特征，适用于成本导向型电商商家与中小电商企业。自 P17 起，服务水平跌破 98%，库存成本继续下探，而缺货损失进入快速增长通道。P17 至 P25 方案可为经营者提供清晰的分级降本图谱：例如 P17 的库存成本为 24078.180 元/年，P21 进一步降至 20197.088 元/年，P25 则压低至 17162.571 元/年。在此过程中，缺货损失相应从

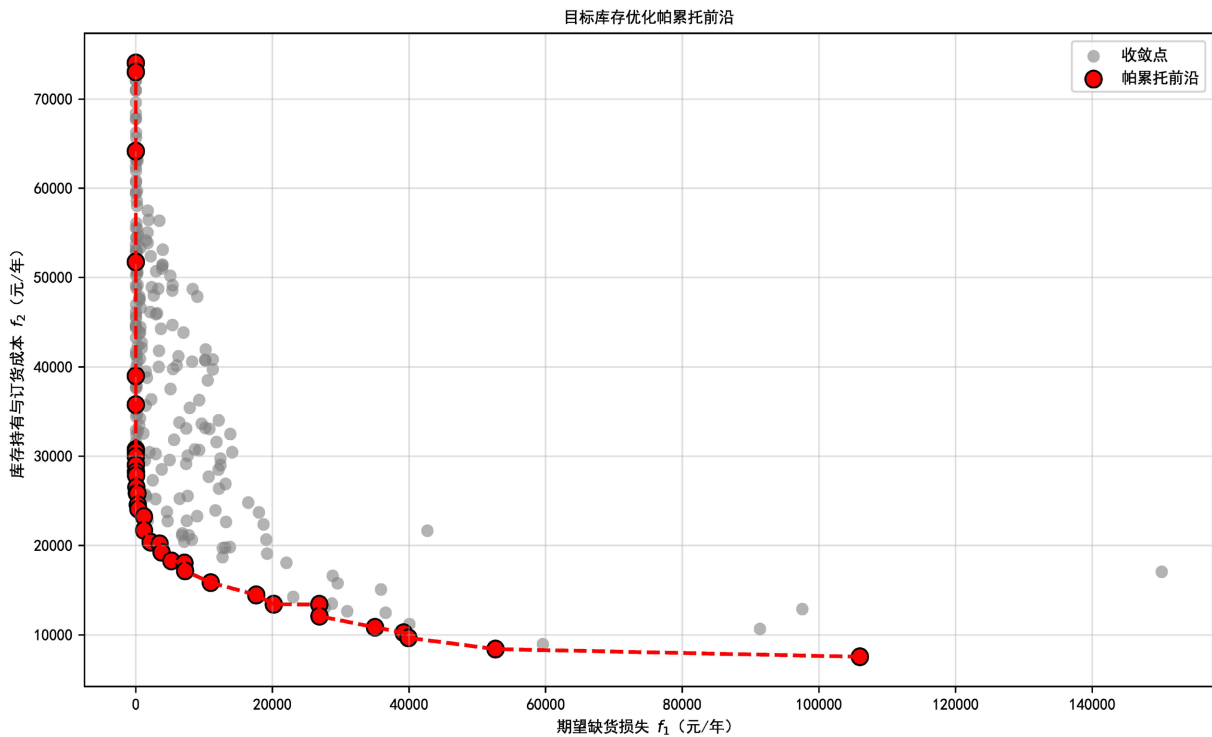


Figure 1. Pareto frontier plot

图 1. 帕累托前沿图

369.650 元/年升至数千元量级。此类商家通过合理容忍适度缺货，压缩仓储库存、降低订货频次与履约成本，实现降本增效，适用于市场扩张期、价格战阶段及长尾、非急需品类，以可控缺货代价换取成本弹性与现金流空间。需警惕的是，前沿最右侧的极端方案(P25 之后)会使缺货损失飙升、远超库存成本，易引发客户负面体验的链式风险，除清仓去库或边缘业务场景外，通常不建议采纳。

3) 均衡决策方案：随着对库存成本约束的增强，前沿进入中段区域，缺货损失与库存成本均被控制在合理区间内，形成一批高性价比的折衷方案，以 P10 至 P16 最为典型。该段服务水平保持在 97.81%~99.88%，缺货损失绝对值仍处于百元量级，但库存成本已较左侧大幅回落至 24563.811~29039.460 元/年。尤其 P10 的库存成本与左侧 P9 几乎持平，且缺货损失仅增至 15.330 元/年，几乎实现了服务不降级而成本微降的优良特性。P15 和 P16 则进一步在不显著侵蚀客户交付质量的前提下，将库存成本压缩到较低水平。与前段相比，中段边际替代率相对平缓，缺货损失与库存成本的权衡成本较低，是综合型电商平台、全品类商家的主流优选策略，能够在保障整体交付质量的同时合理控制仓储与履约成本，兼顾经营稳健性与客户满意度。

4.4. 敏感性分析

考虑到缺货成本 c_s 与订购固定成本 K_f 由数据集中的变量估算得到，其取值存在一定的不确定性。为检验上述参数估计偏差对帕累托前沿及结论稳健性的影响，本文对 c_s 和 K_f 分别进行 $\pm 20\%$ 的独立扰动，生成四组参数组合，每组参数下均采用与原始实验相同的初始点集，利用多目标共轭梯度法重新求解，得到对应的帕累托前沿，五组前沿叠加对比结果如图 2 所示。

由图 2 可知， c_s 的扰动引起前沿整体左右平移，但前沿的外凸形态和曲率几乎保持不变。 $c_s^{+20\%}$ 前沿位于原始前沿右上方， $c_s^{-20\%}$ 前沿位于左下方，三条曲线近似平行。这是因为 c_s 仅通过放大或缩小 f_1 的数值影响前沿位置，而不改变两目标函数的梯度结构； K_f 的扰动对前沿的影响幅度相对于 c_s 来说较小，这

是因为在低服务、高缺货损失区域，订购批量 Q 相对较小，订购成本在 f_2 中的占比上升， K_f 变动的影响才得以体现。以上结果表明，本文所构建的模型在参数 $\pm 20\%$ 的扰动范围内均未发生实质性改变，模型结论对变量估计偏差具有可靠的稳定性。

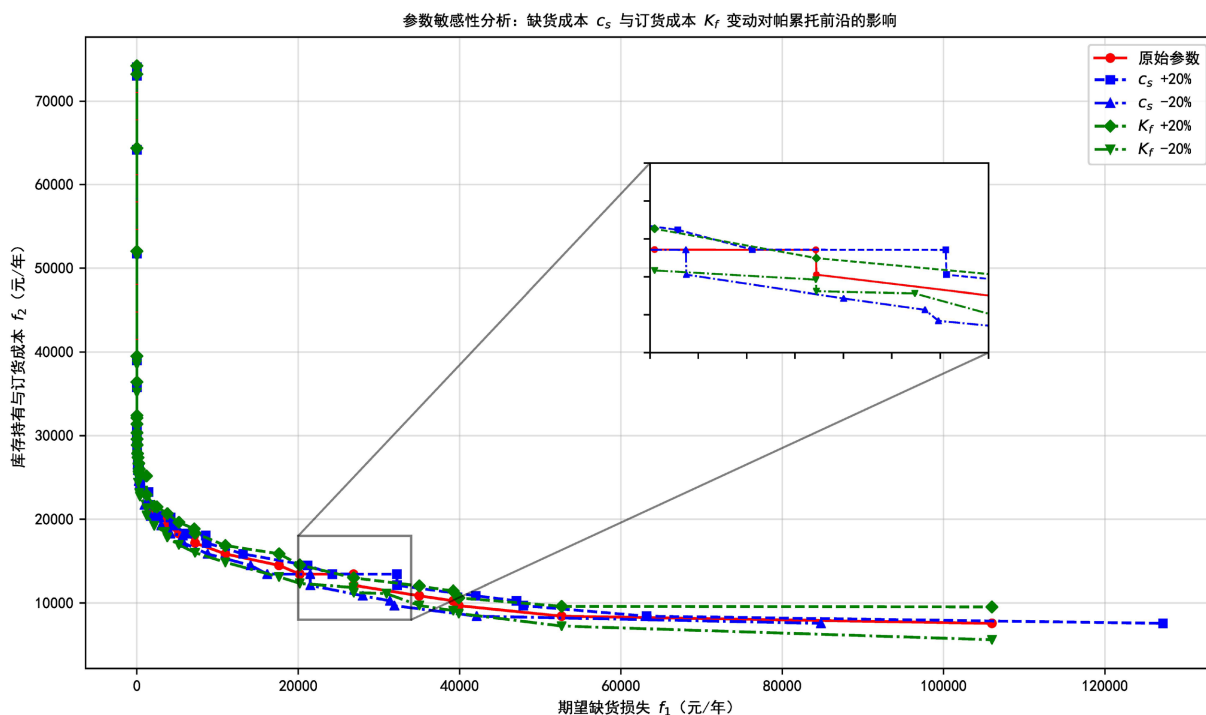


Figure 2. Comparison of pareto fronts

图 2. 帕累托前沿对比图

5. 结论与展望

5.1. 研究结论

本文面向电商仓储库存管理中的缺货损失与库存成本权衡问题，构建了以安全库存系数 z 和订货批量 Q 为联合决策变量的双目标优化模型。模型以 Kaggle 公开的 Logistics Warehouse Dataset 为数据基础，全部 3204 条真实仓储记录统一提供了需求、成本、补货周期等参数，从根本上保证了数据来源的一致性与研究的可复现性。通过对数变换将 $Q > 0$ 的边界约束消除后，采用多目标无约束共轭梯度法直接求解帕累托前沿，避免了传统加权求和法对先验权重的依赖。

数值实验获得 35 个分布均匀的帕累托最优解，缺货损失 f_1 横跨 0.660 元至 105996.160 元，服务水平从 99.96% 覆盖至 7.01%，库存成本相应地在 7564.635 元至 74062.160 元之间有序变化。前沿呈现典型的外凸下降形态，清晰揭示了每降低一单位缺货损失所需付出的库存成本增量随服务水平提高而加速上升的客观规律。敏感性分析表明，缺货成本和持有成本在 $\pm 20\%$ 范围内的独立扰动未改变前沿的基本形态与策略分区，模型具有可靠的稳定性。结合以上数值实验结论，从电商运营的实际场景出发，提出以下库存管理策略：

1) 差异化履约管理。依据帕累托前沿的三区结构，将高毛利核心爆品、医药健康类刚需品及即时零售高频品类定位于高服务水平区，以充裕现货保障“下单即到”的确定性承诺，巩固用户信任与复购黏性；常规日销品置于中段均衡区，以可控的适度缺货释放库存资金，投向新品引入与营销推广；临期品、换季清仓款及长尾商品推至低成本区，主动接受较高缺货率以压缩持有成本，通过快速出清与预售集单对冲风险[6]。

2) 精益补货。利用订货批量与安全库存的非单调协同关系,以“多频次小批量”快反补货替代“少频次大批量”传统模式[7],在不增加缺货风险的前提下降低平均库存,释放仓储容量与流动资金,为供应链协同提供量化依据。

3) 弹性库存调度。以帕累托前沿为标尺:大促前,核心商品沿前沿向左上滑动,临时加高安全库存并放大订货批量,以短期成本换取全程零缺货口碑;大促后沿前沿向右下回收,通过多批次小额补货消化高位库存,平稳回归日常均衡区。由此将库存管理升级为随需求节奏同步波动的弹性治理模式。

4) 数据化决策。缺货成本是影响前沿定位的最敏感参数,应优先依托下单行为追踪、缺货流失率统计及客户终身价值折现精细化测算[8];标定后即可在前沿上直接读取最优策略点。持有成本等不敏感参数以行业均值与仓库合同公开数据估算即可,降低模型落地门槛。

5.2. 研究展望

本文基于正态需求与固定提前期的标准假设,构建了双目标库存优化模型,为需求平稳、补货周期相对稳定的电商仓储场景提供量化决策依据。为提升模型的适用性,未来可从以下三方面深化研究:

1) 向非正态需求延伸[9]。电商长尾商品、促销款及低频周转品常表现出需求间断、波动剧烈且分布右偏的现象,若仍采用正态假设,会低估尾部缺货风险,导致安全库存设置偏低、实际缺货损失高于预期,进而改变低服务水平区间的帕累托前沿形态与边际替代率。未来可采用负二项分布、泊松分布或经验分布函数更准确地拟合真实需求,从而提高缺货成本与库存水平的测算精度。

2) 引入随机提前期。固定提前期假设忽略了补货延迟的不确定性:提前期随机化会使提前期内需求方差显著增大,同等服务水平下所需安全库存上升,库存成本与缺货损失的权衡曲线更陡峭,帕累托前沿整体上移。将其纳入模型,能更真实地反映多重不确定性共同作用下的库存决策规律。

3) 扩展至多仓多产品联合补货,并形成动态管理闭环。将双目标框架扩展至多仓、多产品的联合补货决策,结合在线学习机制实现“沿帕累托前沿滑动”的动态库存管理闭环[10];若能基于电商平台真实交易数据对缺货成本进行实证标定并嵌入该框架,即可完成从理论优化到企业级落地的完整验证闭环。

参考文献

- [1] Qi, M., Shi, Y., Qi, Y., Ma, C., Yuan, R., Wu, D., *et al.* (2022) A Practical End-to-End Inventory Management Model with Deep Learning. *Management Science*, **69**, 759-773. <https://doi.org/10.1287/mnsc.2022.4564>
- [2] Fliege, J. and Svaiter, B.F. (2000) Steepest Descent Methods for Multicriteria Optimization. *Mathematical Methods of Operations Research (ZOR)*, **51**, 479-494. <https://doi.org/10.1007/s001860000043>
- [3] Lucambio Pérez, L.R. and Prudente, L.F. (2018) Nonlinear Conjugate Gradient Methods for Vector Optimization. *SIAM Journal on Optimization*, **28**, 2690-2720. <https://doi.org/10.1137/17m1126588>
- [4] Zipkin, P.H. (2000) Foundations of Inventory Management. McGraw-Hill.
- [5] Nahmias, S. and Cheng, Y. (2008) Production and Operations Analysis. 6th Edition, McGraw-Hill.
- [6] 李昀洲, 陈璐. 考虑需求关联性的多供应商电商库存分配研究[J]. 上海理工大学学报, 2023, 45(2): 180-188.
- [7] Silva, P.M., Gonçalves, J.N.C., Martins, T.M., Marques, L.C., Oliveira, M., Reis, M.I., *et al.* (2022) A Hybrid Bi-Objective Optimization Approach for Joint Determination of Safety Stock and Safety Time Buffers in Multi-Item Single-Stage Industrial Supply Chains. *Computers & Industrial Engineering*, **168**, Article 108095. <https://doi.org/10.1016/j.cie.2022.108095>
- [8] Yan, X., Chao, X. and Lu, Y. (2024) Optimal Control Policies for Dynamic Inventory Systems with Service Level Dependent Demand. *European Journal of Operational Research*, **314**, 935-949. <https://doi.org/10.1016/j.ejor.2023.11.005>
- [9] 柯蓉. 非正态需求分布函数条件下的多阶段库存管理: 数理统计学的应用[M]. 上海: 上海浦江教育出版社, 2013.
- [10] Tang, J., Duenyas, I., Shi, C. and Yang, N. (2025) Multiproduct Inventory Systems with Upgrading: Replenishment, Allocation, and Online Learning. *Manufacturing & Service Operations Management*, **28**, 537-557. <https://doi.org/10.1287/msom.2024.0974>