应用概率统计　第 38 卷
第 4 期　2022 年 8 月

Chinese Journal of Applied Probability and Statistics
Aug., 2022, Vol. 38, No. 4, pp. 531-545
doi: 10.3969/j.issn.1001-4268.2022.04.004

具有共有寿命和取消订货的非抢占优先权排队库存系统 *

罗煦香*　　刘再明

(佛山科学技术学院数学与大数据学院, 佛山, 528000)　　(中南大学数学与统计学院, 长沙, 410083)

摘　要: 本文研究了具有取消订货和共有寿命的非抢占优先权排队库存系统, 其中顾客到达服从泊松过程, 服务时间服从指数分布. 我们构建了一个水平相依的拟生灭过程 (LDQBD 过程), 并利用 Neuts-Rao 截断法得到了系统的平稳条件和稳态概率向量, 同时给出了一些性能指标和期望成本函数. 通过数值模拟, 我们得到了最优库存容量和最小成本. 最后, 我们通过对系统参数的敏感性分析, 给管理者提供了一些有益的建议.

关键词: 非抢占优先权; 排队库存系统; LDQBD; 共有寿命; 取消订货

中图分类号: O226

英文引用格式: LUO X X, Liu Z M. Analysis of a non-preemptive priority queueing-inventory system with common life time and cancellation [J]. Chinese J Appl Probab Statist, 2022, 38(4): 531–545. (in Chinese)

§1. 引　　言

在日常生活中, 人们经常会预定产品以备将来使用, 或者由于某些原因取消已订购产品. 例如飞机票、火车票等订票系统. 在这类需求库存系统中, 座位表示库存. 当没有库存时, 服务台无法进行工作, 此时将影响服务台的正常运营; 而当飞机/火车驶离, 座位票失效, 即该类库存产品具有共有寿命, 一旦到达共有寿命, 系统中的库存将立即失效. 此外, 部分顾客由于行程更改等原因而取消订票, 此时会导致库存量的增加. 因此, 在现实生活中, 企业管理者需要综合考虑上述各种因素并制定相应的管理机制, 通过控制订货率或订货水平来调控整个系统的运营, 以使在相应的费用结构下实现经济效益最大化.

Krishnamoorthy 等[1] 首次将预订、取消和共有寿命引入到排队库存系统中. 在该模型中, 顾客到达服从柏松分布, 服务时间服从指数分布, 库存中的物品具有随机的共有寿命并且到达共有寿命后, 缓冲区中的顾客数会被清除. 之后, 他们又将上述模型拓展到 $GI/M/1$ 排队库存系统[2], 考虑了两种顾客从重试空间转移到缓冲区策略下的单服务器系统模型, 得到了几个系统性能指标, 并讨论了两种模型的成本函数. Shajin 和

*粤港澳智能微纳光电技术联合实验室 (批准号: 2020B1212030010) 和国家自然科学基金项目 (批准号: 12071487) 资助.

*通讯作者, E-mail: luoxuxiang@fosu.edu.cn.

本文 2020 年 8 月 26 日收到, 2021 年 1 月 8 日收到修改稿.

Krishnamoorthy[3] 提出了带有不耐烦顾客和超额预订的排队库存系统, 其中顾客按照 Markovian 到达过程到达, 服务时间具有位相型分布, 商品的使用寿命呈指数分布且允许预订和取消. Shajin 等[4] 考虑了 K 个不同时间单位前的库存情况, 研究了具有提前预订、取消、共有寿命和超额预订的单服务排队库存系统, 并对系统的性能特性进行了计算.

在上述研究的排队库存模型中, 通常都假设到达系统的顾客服从先到先服务的排队规则. 然而在实际生活中, 往往会出现优先处理的情况. 例如 VIP 客户可以优先订票、加急快递享有优先处理的权利、银行的 VIP 客户可以优先得到服务、医院里重症患者优先安排救治等. 尤其是人们对于差异化和个性化服务需求的不断增加, 优先权问题也越来越普遍, 许多研究人员在这方面开展了大量的研究工作. Miller[5] 研究了带抢占和非抢占优先 $M/M/1$ 排队模型的稳态概率分布问题. Takine 和 Hasegawa[6] 分析了 $MAP/G/1$ 抢占优先排队系统, 并推导出系统中工作量的 Laplace-Stieljes 变换. Zhu 等[7] 考虑了带优先权的 $M/G/1$ 可修重试排队系统, 给出了系统中的优先队列、轨道和系统中客户数量的稳态分布以及平均等待时间. 胡根生等[8] 研究了带有两个优先权的 N 策略 $M/G/1$ 排队系统在通信网络中的应用, 通过补充变量法得到了对长分布母函数以及通信网缓冲器中的队列长度. Shalev-Oren 等[9] 讨论了优先权在制造系统方面的应用, 并对制造系统的各种优先级分配策略进行了分析. de Souza 等[10] 分析了带优先权的排队系统在紧急医疗服务方面的应用, 分析了具有多个优先级系统和一个等待客户的队列, 评估了每一类紧急呼叫的平均响应时间.

本文研究了具有取消订货和共有寿命的非抢占优先权排队库存系统, 其中顾客到达是一个泊松过程, 服务时间和共有寿命都服从指数分布. 我们构建了一个水平相依的拟生灭过程 (LDQBD 过程), 并利用 Neuts-Rao 截断法[11] 得到了系统的平稳条件和稳态概率向量, 同时给出了一些性能指标和期望成本函数. 通过数值模拟, 我们得到了最优库存容量和最小成本. 最后, 我们分析了最优库存容量和系统性能指标关于系统参数的变化, 并基于变化结果, 给管理者提供了建议.

§2. 模型描述

本文考虑一个具有共有寿命和取消订货的非抢占优先权排队库存系统, 模型基本假设如下:

1. 系统由服务台、无限等待空间以及最大容量为 S 的仓库构成.
2. 考虑两类顾客: 高优先级顾客和低优先级顾客, 两类顾客都是按照 Poisson 过程独立的到达, 到达率分别为 λ_1 和 λ_2.
3. 系统只有一个服务台, 具有无限等待空间. 当库存非 0 时, 顾客依次接受服务, 服务完成时每个顾客带走单位 1 的库存产品, 且服务时间分别服从参数为 μ_1 和 μ_2 的指数

分布.

4. 高优先权的顾客采取非抢占优先的规则获得服务. 非抢占优先的排队规是一个高优先权的顾客到达系统时, 若发现服务员正在服务, 则不论正在被服务的顾客优先级是否比他低, 高优先级的顾客都要等到服务员服务完当前顾客才能按规则为他服务. 在同一等级类型的顾客中, 采取先到先服务规则.

5. 新顾客到达时, 若发现服务台空闲且库存非 0, 顾客立即进入服务台接受服务; 若发现服务台忙, 顾客进入等待空间等待服务. 当服务完的顾客离开系统后, 若系统至少存在一个顾客且库存非 0 时, 立即开始下一个服务; 若系统准备服务一个顾客且库存为 0 时, 该服务在下一次补给到达时立即开始.

6. 库存中的产品具有共有寿命, 共有寿命服从参数为 γ 的指数分布. 当到达共有寿命时, 库存中的产品失效且被及时清除, 等待空间中的高优先权顾客立即离开系统, 低优先权顾客留在系统等待下一个周期的服务.

7. 定义一个周期: 从库存立即补给到最大库存量 S 时开始, 到库存到达共有寿命时终止.

8. 系统共有两种库存补给策略. 一是瞬时库存补给策略, 即: 到达共有寿命且开始下一个周期时, 库存立即得到补给, 补给量为最大库存容量 S. 二是取消订货补给策略, 即: 顾客在产品到达共有寿命前取消订货, 库存得到增加. 顾客取消订货时间服从参数为 $(S-i)\beta$ 的指数分布, i 为当前库存水平. 注意, 通过取消订货补给, 库存水平不会超过最大库存量 S.

9. 顾客的到达时间间隔、服务时间、共有寿命及取消订货时间相互独立.

系统的模型示意图如图 1 所示: 在 t 时刻, 用 $N_1(t)$ 表示系统中高优先权的顾客数, $N_1(t) \in$

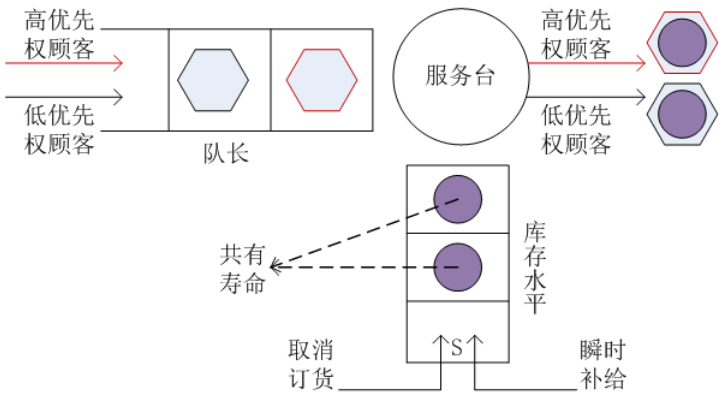


图 1　非抢占优先权排队库存系统

$\{0,1,2,\cdots\}$; 用 $N_2(t)$ 表示系统中低优先权的顾客数, $N_2(t) \in \{0,1,2,\cdots\}$; 用 $I(t)$ 表示当

前库存水平, $I(t) \in \{0,1,2,\cdots,S\}$; 用 $J(t)$ 表示正在接受服务的顾客类别,

$$J(t)=\begin{cases}0, & \text{没有顾客正在接受服务;}\\ 1, & \text{高优先权的顾客正在接受服务;}\\ 2, & \text{低优先权的顾客正在接受服务.}\end{cases}$$

则 $\{N_1(t),N_2(t),J(t),I(t),t\geqslant 0\}$ 是一个四维的 Markov 过程, 其状态空间 Ω 为

$$\begin{aligned}\Omega=&\{(n_1,n_2,0,0):n_1\geqslant 0,\ n_2\geqslant 0\}\\ &\cup\{(0,0,0,i):0<i\leqslant S\}\\ &\cup\{(n_1,0,1,i):n_1>0,\ 0<i\leqslant S\}\\ &\cup\{(0,n_2,2,i):n_2>0,\ 0<i\leqslant S\}\\ &\cup\{(n_1,n_2,j,i):n_1>0,\ n_2>0,\ j\in\{1,2\},\ 0<i\leqslant S\}\\ &\cup\{(-,n_2,0,S):n_2\geqslant 0\},\end{aligned}$$

其中 $(-,n_2,0,S)$ 表示吸收态, 即到达共有寿命时, 系统顾客数为 n_2, 库存立即补给到最大库存量 S.

由于两类顾客的队列长度都趋近于无限, 通过标准矩阵分析法, 该系统是一个双无限拟生灭过程 (QBD 过程), 我们很难用数学方法有效地处理双无限 QBD 过程. 为了解决这个问题, 我们采用由 Lian 和 Liu[12]及 Lian 等[13] 提出的一种特殊的状态排序法, 使得系统过程成为一个水平相依的拟生灭过程. 这种排序法有利于我们计算联合稳态分布.

我们按照以下顺序对 $\{N_1(t),N_2(t),J(t),I(t),t\geqslant 0\}$ 的状态排序:

1. 将两类顾客的队列长度之和 n_1+n_2 作为水平.
2. 具有相同水平的所有状态都在同一行中.
3. 行按水平升序排列.
4. 一行中的状态按 n_2, j 和 i 的顺序升序排列.

则, 我们有:

水平 0: $(0,0,0,0),(0,0,0,1),\cdots,(0,0,0,S);\ (0,S)$.

水平 1: $(1,0,0,0),(1,0,1,1),\cdots,(1,0,1,S);$
$(0,1,0,0),(0,1,2,1),\cdots,(0,1,2,S);\ (1,S)$.

水平 2: $(2,0,0,0),(2,0,1,1),\cdots,(2,0,1,S);$
$(1,1,0,0),(1,1,1,1),\cdots,(1,1,1,S);\ (1,1,2,1),(1,1,2,2),\cdots,(1,1,2,S);$
$(0,2,0,0),(0,2,2,1),\cdots,(0,2,2,S);\ (2,S)$.

水平 3: $(3,0,0,0),(3,0,1,1),\cdots,(3,0,1,S)$;
$(2,1,0,0),(2,1,1,1),\cdots,(2,1,1,S)$; $(2,1,2,1),(2,1,2,2),\cdots,(2,1,2,S)$;
$(1,2,0,0),(1,2,1,1),\cdots,(1,2,1,S)$; $(1,2,2,1),(1,2,2,2),\cdots,(1,2,2,S)$;
$(0,3,0,0),(0,3,2,1),\cdots,(0,3,2,S)$; $(3,S)$.

$\cdots$　　$\cdots$

对任意水平 $k\in\{0,1,2,\cdots\}$, 唯一可能的状态转移为 $k\to k$, $k\to k+1$ 和 $k\to k-1$. 按水平 k 排序, 相应的生成元矩阵 $\boldsymbol{Q}$ 可表示为

$$\boldsymbol{Q}=\begin{pmatrix}\boldsymbol{B}_0 & \boldsymbol{A}_0 & & & \\ \boldsymbol{C}_1 & \boldsymbol{B}_1 & \boldsymbol{A}_1 & & \\ & \boldsymbol{C}_2 & \boldsymbol{B}_2 & \boldsymbol{A}_2 & \\ & & \ddots & \ddots & \ddots\end{pmatrix}.$$

$\boldsymbol{A}_0$, $\boldsymbol{B}_0$ 和 $\boldsymbol{C}_1$ 分别是 $(S+2)\times(2S+3)$, $(S+2)\times(S+2)$ 和 $(2S+3)\times(S+2)$ 维矩阵, 具体如下:

$$\boldsymbol{A}_0=\begin{pmatrix}\lambda_1 & & \lambda_2 & & 0\\ & \ddots & & \ddots & \vdots\\ & & \lambda_1 & & \lambda_2 & 0\\ & & \lambda_1 & & \lambda_2 & 0\end{pmatrix},$$

$$\boldsymbol{B}_0=\begin{pmatrix}-\lambda_1-\lambda_2-\gamma-S\beta & S\beta & & & & & \gamma\\ & \ddots & \ddots & & & & \vdots\\ & & -\lambda_1-\lambda_2-\gamma-\beta & \beta & & \gamma\\ & & & -\lambda_1-\lambda_2-\gamma & \gamma\\ & & & & -\lambda_1-\lambda_2\end{pmatrix},$$

$$\boldsymbol{C}_1=\begin{pmatrix}0 & & & & 0\\ \mu_1 & 0 & & & 0\\ & \ddots & \ddots & & \vdots\\ & & \mu_1 & 0 & 0\\ 0 & & & & 0\\ \mu_2 & 0 & & & 0\\ & \ddots & \ddots & & \vdots\\ & & \mu_2 & 0 & 0\\ & & & & 0\end{pmatrix}.$$

易知, 水平为 k $(k>0)$ 时有 $2kS+k+2$ 个状态, 则 $\boldsymbol{A}_k$, $\boldsymbol{B}_k$ 和 $\boldsymbol{C}_{k+1}$ 分别是 $(2kS+k+2)\times(2kS+2S+k+3)$, $(2kS+k+2)\times(2kS+k+2)$ 和 $(2kS+2S+k+3)\times(2kS+k+2)$ 维矩阵, 具体如下:

$$\boldsymbol{A}_k=\begin{pmatrix} \boldsymbol{a}_0^1 & \boldsymbol{a}_0^2 & & & & \boldsymbol{0}\\ & \ddots & \ddots & & & \vdots\\ & & & \boldsymbol{a}_k^1 & \boldsymbol{a}_k^2 & \boldsymbol{0}\\ & & & \widetilde{\boldsymbol{\lambda}}_1 & \widetilde{\boldsymbol{\lambda}}_2 & 0 \end{pmatrix},$$

$$\boldsymbol{B}_k=\begin{pmatrix} \boldsymbol{b}_0 & & & \gamma\boldsymbol{e}\\ & \ddots & & \vdots\\ & & \boldsymbol{b}_k & \gamma\boldsymbol{e}\\ & & & -\lambda_1-\lambda_2 \end{pmatrix},$$

$$\boldsymbol{C}_k=\begin{pmatrix} \boldsymbol{c}_0^1 & & & & \boldsymbol{0}\\ \boldsymbol{c}_1^2 & \boldsymbol{c}_1^1 & & & \boldsymbol{0}\\ & \ddots & \ddots & & \vdots\\ & & \boldsymbol{c}_{k-1}^2 & \boldsymbol{c}_{k-1}^1 & \boldsymbol{0}\\ & & & \boldsymbol{c}_k^2 & \boldsymbol{0}\\ & & & & 0 \end{pmatrix}, \qquad k>1,$$

其中, $\boldsymbol{a}_0^1$, $\boldsymbol{a}_k^2$, $\boldsymbol{b}_0$, $\boldsymbol{b}_k$, $\boldsymbol{c}_0^1$ 和 $\boldsymbol{c}_k^2$ 是 $(S+1)\times(S+1)$ 维矩阵, $\boldsymbol{a}_k^1$ 和 $\boldsymbol{a}_0^2$ 是 $(S+1)\times(2S+1)$ 维矩阵, $\boldsymbol{c}_1^2$ 和 $\boldsymbol{c}_{k-1}^1$ 是 $(2S+1)\times(S+1)$ 维矩阵, $\boldsymbol{a}_m^1$ $(0<m<k)$, $\boldsymbol{a}_m^2$ $(0<m<k)$, $\boldsymbol{b}_m$ $(0<m<k)$, $\boldsymbol{c}_m^1$ $(0<m<k-1)$, $\boldsymbol{c}_m^2$ $(1<m<k)$ 是 $(2S+1)\times(2S+1)$ 维矩阵, $\boldsymbol{e}$ 为适当维数的所有分量取值为 1 的列向量, $\boldsymbol{I}$ 是适当维数的单位矩阵, $\widetilde{\boldsymbol{\lambda}}_1=(0,\cdots,0,\lambda_1,0,\cdots,0)$, λ_1 位于第 $2kS+k+1$ 列, $\widetilde{\boldsymbol{\lambda}}_2=(0,\cdots,0,\lambda_2,0)$, λ_2 位于第 $2kS+2S+k+2$ 列, 且 $\boldsymbol{a}_m^1=\lambda_1\boldsymbol{I}$ $(0\leqslant m<k)$, $\boldsymbol{a}_m^2=\lambda_2\boldsymbol{I}$ $(1\leqslant m\leqslant k)$,

$$\boldsymbol{a}_k^1=\begin{pmatrix} \lambda_1 & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \lambda_1\boldsymbol{I} \end{pmatrix}, \qquad \boldsymbol{a}_0^2=\begin{pmatrix} \lambda_2 & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \lambda_2\boldsymbol{I} & \boldsymbol{0} \end{pmatrix},$$

$$\boldsymbol{b}_0=\begin{pmatrix} b_{00}^0 & S\beta & & \\ & \ddots & \ddots & \\ & & b_{S-1,S-1}^0 & \beta\\ & & & b_{S,S}^0 \end{pmatrix}, \qquad \boldsymbol{b}_k=\begin{pmatrix} b_{00}^k & S\beta & & \\ & \ddots & \ddots & \\ & & b_{S-1,S-1}^k & \beta\\ & & & b_{S,S}^k \end{pmatrix},$$

$$\boldsymbol{b}_m=\begin{pmatrix} b_{00}^0 & S\beta & & & & & & & \\ & \ddots & \ddots & & & & & & \\ & & b_{S-1,S-1}^0 & \beta & & & & & \\ & & & b_{S,S}^0 & & & & & \\ & & & & b_{11}^k & (S-1)\beta & & & \\ & & & & & \ddots & \ddots & & \\ & & & & & & b_{S-1,S-1}^k & \beta & \\ & & & & & & & b_{S,S}^k & \end{pmatrix},$$

$$\boldsymbol{c}_0^1=\begin{pmatrix} 0 & & & \\ \mu_1 & 0 & & \\ & \ddots & \ddots & \\ & & \mu_1 & 0 \end{pmatrix},\qquad \boldsymbol{c}_m^1=\begin{pmatrix} 0 & & & & \\ \mu_1 & 0 & & & \\ & \ddots & \ddots & & \\ & & \mu_1 & 0 & \\ & & & & \boldsymbol{0} \end{pmatrix},\qquad 1\leqslant m\leqslant k-2,$$

$$\boldsymbol{c}_{k-1}^1=\begin{pmatrix} 0 & & & \\ \mu_1 & 0 & & \\ & \ddots & \ddots & \\ & & \mu_1 & 0 \\ & & & \boldsymbol{0} \end{pmatrix},\qquad \boldsymbol{c}_1^2=\begin{pmatrix} \boldsymbol{0} & & & \boldsymbol{0} \\ \mu_2 & 0 & & \\ & \ddots & \ddots & \\ & & \mu_2 & 0 \end{pmatrix},$$

$$\boldsymbol{c}_m^2=\begin{pmatrix} 0 & & & \boldsymbol{0} \\ \mu_2 & 0 & & \\ & \ddots & \ddots & \\ & & \mu_2 & 0 \end{pmatrix},\qquad 2\leqslant m\leqslant k-1,\qquad \boldsymbol{c}_k^2=\begin{pmatrix} 0 & & \\ \mu_2 & 0 & \\ & \ddots & \ddots \\ & & \mu_2 & 0 \end{pmatrix},$$

$b_{ii}^0=-\lambda_1-\lambda_2-\mu_1 I_{i>0}-(S-i)\beta-\gamma\ (0\leqslant i<S)$, $b_{S,S}^0=-\lambda_1-\lambda_2-\mu_1-\gamma$, $b_{ii}^k=-\lambda_1-\lambda_2-\mu_2 I_{i>0}-(S-i)\beta-\gamma\ (0\leqslant i<S)$, $b_{S,S}^k=-\lambda_1-\lambda_2-\mu_2-\gamma$.

§3. 稳态分析

令

$$P(n_1,n_2,j,i)=\lim_{t\to\infty}\mathrm{P}\{N_1(t)=n_1,N_2(t)=n_2,J(t)=j,I(t)=i\}$$

表示在状态 $(n_1,n_2,j,i)\in\Omega$ 下的稳态概率, $\boldsymbol{\pi}$ 表示稳态概率向量, 当系统平稳时, 有

$$\boldsymbol{\pi Q}=\boldsymbol{0},\qquad \boldsymbol{\pi e}=1.$$

对于水平相依的生成元 $\boldsymbol{Q}$, 我们将其联合稳态分布 $\boldsymbol{\pi}$ 分割成 $\boldsymbol{\pi}=(\boldsymbol{\pi}_1,\boldsymbol{\pi}_2,\cdots)$, 其中 $\boldsymbol{\pi}_k$ 表示水平为 k 时的稳态概率, 为

$$\boldsymbol{\pi}_k=(P(n_1,n_2,j,i):n_1+n_2=k,j\in\{0,1,2\},i\in\{0,1,\cdots,S\})\cup(P(-,k,0,S)).$$

1) 稳态条件

采用 Neuts-Rao 截断法, 我们先确定一个合适的正整数 N, 对于 $n\geqslant N$, $\boldsymbol{A}=\boldsymbol{A}_n$, $\boldsymbol{B}=\boldsymbol{B}_n$, $\boldsymbol{C}=\boldsymbol{C}_n$, 我们得到下列生成元:

$$\boldsymbol{Q}'=\begin{pmatrix}
\boldsymbol{B}_0 & \boldsymbol{A}_0 & & & & & & \\
\boldsymbol{C}_1 & \boldsymbol{B}_1 & \boldsymbol{A}_1 & & & & & \\
 & \boldsymbol{C}_2 & \boldsymbol{B}_2 & \boldsymbol{A}_2 & & & & \\
 & & \ddots & \ddots & \ddots & & & \\
 & & & \boldsymbol{C}_{N-1} & \boldsymbol{B}_{N-1} & \boldsymbol{A}_{N-1} & & \\
 & & & & \boldsymbol{C} & \boldsymbol{B} & \boldsymbol{A} & \\
 & & & & & \ddots & \ddots & \ddots
\end{pmatrix}. \tag{1}$$

令 $\boldsymbol{\pi}'$ 为 $\boldsymbol{H}$ 的平稳概率向量

$$\boldsymbol{H}=\begin{pmatrix}
\boldsymbol{d}_0 & \boldsymbol{a}_0^2 & & & & \gamma\boldsymbol{e} \\
\boldsymbol{c}_1^2 & \boldsymbol{d}_1 & \boldsymbol{a}_1^2 & & & \gamma\boldsymbol{e} \\
 & \ddots & \ddots & \ddots & & \vdots \\
 & & \boldsymbol{c}_N^2 & \boldsymbol{d}_N & \boldsymbol{a}_N^2 & \gamma\boldsymbol{e} \\
 & & & \widetilde{\boldsymbol{\lambda}}_1 & \widetilde{\boldsymbol{\lambda}}_2 & -\lambda_1-\lambda_2
\end{pmatrix},$$

其中, $\boldsymbol{d}_m=\boldsymbol{a}_m^1+\boldsymbol{c}_m^1+\boldsymbol{b}_m$ $(0\leqslant m\leqslant N-1)$, $\boldsymbol{d}_N=\boldsymbol{a}_N^1+\boldsymbol{b}_N$. 则有

$$\boldsymbol{\pi}'\boldsymbol{H}=\boldsymbol{0},\qquad \boldsymbol{\pi}'\boldsymbol{e}=1.$$

$\boldsymbol{\pi}'$ 的组成元素由下列方程获得:

$$\boldsymbol{\pi}'_k=\boldsymbol{\pi}'_{k+1}\boldsymbol{\Phi}_k,\qquad 0\leqslant k\leqslant N, \tag{2}$$

其中

$$\boldsymbol{\Phi}_k=\begin{cases}
-\boldsymbol{c}_1^2\boldsymbol{d}_0^{-1}, & k=0;\\
-\boldsymbol{c}_{k+1}^2(\boldsymbol{\Phi}_{k-1}\boldsymbol{a}_{k-1}^2+\boldsymbol{d}_k)^{-1}, & 1\leqslant k\leqslant N-1;\\
-\widetilde{\boldsymbol{\lambda}}_1(\boldsymbol{\Phi}_{N-1}\boldsymbol{a}_{N-1}^2+\boldsymbol{d}_N)^{-1}, & k=N.
\end{cases}$$

通过方程

$$(\boldsymbol{\pi}_0' + \boldsymbol{\pi}_1' + \cdots + \boldsymbol{\pi}_N')\gamma\boldsymbol{e} - \boldsymbol{\pi}_{N+1}'(\lambda_1 + \lambda_2) = 0$$

和规范化条件 $\boldsymbol{\pi}'\boldsymbol{e} = 1$, 我们得到

$$\boldsymbol{\pi}_{N+1}' = \frac{\gamma}{\lambda_1 + \lambda_2 + \gamma}.$$

则当前排队–库存系统的平稳条件由下列定理给出.

定理 1　当前排队库存系统是平稳的当且仅当

$$\boldsymbol{\pi}_{N+1}'\Big(\sum_{i=0}^{N}\prod_{k=i}^{N}\boldsymbol{\Phi}_k'\boldsymbol{a}_i^1 + \sum_{i=0}^{N-1}\prod_{k=i}^{N}\boldsymbol{\Phi}_k'\boldsymbol{a}_i^2 + \widetilde{\boldsymbol{\lambda}}_1\Big)\boldsymbol{e} < \boldsymbol{\pi}_{N+1}'\Big(\sum_{i=0}^{N-1}\prod_{k=i}^{N}\boldsymbol{\Phi}_k'\boldsymbol{c}_i^1 + \sum_{i=1}^{N}\prod_{k=i}^{N}\boldsymbol{\Phi}_k'\boldsymbol{c}_i^2\Big)\boldsymbol{e}. \quad (3)$$

证明:　具有形如式 (1) 的生成元的排队库存系统是平稳的当且仅当 [14]

$$\boldsymbol{\pi}'\boldsymbol{A}\boldsymbol{e} < \boldsymbol{\pi}'\boldsymbol{C}\boldsymbol{e}. \quad (4)$$

由 $\boldsymbol{A}$ 和 $\boldsymbol{C}$ 的组成元素, 我们得到

$$\boldsymbol{\pi}'\boldsymbol{A}\boldsymbol{e} = \Big(\sum_{i=0}^{N}\boldsymbol{\pi}_i'\boldsymbol{a}_i^1 + \sum_{i=0}^{N-1}\boldsymbol{\pi}_i'\boldsymbol{a}_i^2 + \boldsymbol{\pi}_{N+1}'\widetilde{\boldsymbol{\lambda}}_1\Big)\boldsymbol{e},$$
$$\boldsymbol{\pi}'\boldsymbol{C}\boldsymbol{e} = \Big(\sum_{i=0}^{N-1}\boldsymbol{\pi}_i'\boldsymbol{c}_i^1 + \sum_{i=1}^{N}\boldsymbol{\pi}_i'\boldsymbol{c}_i^2\Big)\boldsymbol{e}.$$

将方程 (2) 代入上述方程中, 利用平稳条件 (4), 我们得到了式 (3).　　□

2) 稳态概率向量

假设平稳条件 (3) 满足, 我们得到 $n \geqslant N$ 时的稳态概率向量 $\boldsymbol{\pi}_n$,

$$\boldsymbol{\pi}_{n+N-1} = \boldsymbol{\pi}_{N-1}\boldsymbol{R}^n, \qquad n \geqslant 1,$$

其中 $\boldsymbol{R}$ 是下列矩阵方程的最小非负解:

$$\boldsymbol{A} + \boldsymbol{R}\boldsymbol{B} + \boldsymbol{R}^2\boldsymbol{C} = 0.$$

求解下列方程组我们可以得到平稳概率 $\boldsymbol{\pi}_k$, $0 \leqslant k \leqslant N-1$,

$$\begin{aligned}
&\boldsymbol{\pi}_0\boldsymbol{B}_0 + \boldsymbol{\pi}_1\boldsymbol{C}_0 = \boldsymbol{0},\\
&\boldsymbol{\pi}_0\boldsymbol{A}_0 + \boldsymbol{\pi}_1\boldsymbol{B}_1 + \boldsymbol{\pi}_2\boldsymbol{C}_2 = \boldsymbol{0},\\
&\boldsymbol{\pi}_{k-1}\boldsymbol{A}_{k-1} + \boldsymbol{\pi}_k\boldsymbol{B}_k + \boldsymbol{\pi}_{k+1}\boldsymbol{C}_{k+1} = \boldsymbol{0}, \qquad 2 \leqslant k \leqslant N-2,\\
&\boldsymbol{\pi}_{N-2}\boldsymbol{A}_{N-2} + \boldsymbol{\pi}_{N-1}(\boldsymbol{B}_{N-1} + \boldsymbol{R}\boldsymbol{C}) = \boldsymbol{0}.
\end{aligned}$$

规范化条件为

$$\sum_{k=0}^{N-2}\boldsymbol{\pi}_k\boldsymbol{e}+\boldsymbol{\pi}_{N-1}(\boldsymbol{I}-\boldsymbol{R})^{-1}\boldsymbol{e}=1,$$

其中

$$\boldsymbol{\pi}_k=\boldsymbol{\pi}_0\prod_{i=1}^{k}\boldsymbol{\Psi}_i,\qquad 1\leqslant k\leqslant N-1,$$

$$\boldsymbol{\Psi}_i=\begin{cases}-\boldsymbol{A}_0(\boldsymbol{B}_1+\boldsymbol{\Psi}_2\boldsymbol{C}_2)^{-1}, & i=1;\\ -\boldsymbol{A}_{i-1}(\boldsymbol{B}_i+\boldsymbol{\Psi}_{i+1}\boldsymbol{C}_{i+1})^{-1}, & 2\leqslant i\leqslant N-2;\\ -\boldsymbol{A}_{N-2}(\boldsymbol{B}_{N-1}+\boldsymbol{R}\boldsymbol{C})^{-1}, & i=N-1.\end{cases}$$

§4. 稳态性能指标

在这一节中, 我们给出了一些稳态性能指标.

1. 期望等待成本 E_{W}

$$E_{\mathrm{W}}=b_1\sum_{n_1=1}^{\infty}n_1\sum_{n_2,j,i}\boldsymbol{\pi}(n_1,n_2,j,i)+b_2\sum_{n_2=1}^{\infty}n_2\sum_{n_1,j,i}\boldsymbol{\pi}(n_1,n_2,j,i),$$

其中, b_1 和 b_2 分别为高优先权顾客和低优先权顾客的单位等待成本.

2. 库存期望水平 E_{I}

$$E_{\mathrm{I}}=\sum_{i=1}^{S}i\sum_{n_1,n_2,j}\boldsymbol{\pi}(n_1,n_2,j,i).$$

3. 期望取消订货量 E_{CR}

$$E_{\mathrm{CR}}=\frac{1}{\gamma}\sum_{i=0}^{S-1}\frac{(S-i)\beta}{\gamma+(S-i)\beta}\sum_{n_1,n_2,j}\boldsymbol{\pi}(n_1,n_2,j,i).$$

4. 库存期望失效水平 E_{LI} (即产品到达共有寿命时库存失效的数量, 比如航班起飞时飞机内的空座位数)

$$E_{\mathrm{LI}}=\sum_{i=1}^{S}i\Big[\sum_{n_1>0}\sum_{n_2\geqslant 0}p_1^i\boldsymbol{\pi}(n_1,n_2,1,i)+\sum_{n_2>0}\sum_{n_1\geqslant 0}p_2^i\boldsymbol{\pi}(n_1,n_2,2,i)\Big],$$

其中,

$$p_1^i=\frac{\gamma}{\lambda_1+\lambda_2+\mu_1+(S-i)\beta+\gamma},\qquad p_2^i=\frac{\gamma}{\lambda_1+\lambda_2+\mu_2+(S-i)\beta+\gamma}.$$

5. 期望总成本函数 $C(S)$

$$C(S)=E_{\mathrm{W}}+hE_{\mathrm{I}}+c_1E_{\mathrm{CR}}+c_2E_{\mathrm{LI}},$$

其中, h 为单位保管成本, c_1 为单位取消订货损失成本, c_2 为单位失效库存损失成本.

§5. 数值模拟

1) 最优库存策略

由于成本函数的复杂性, 我们很难理论分析成本函数与最大库存量 S 的关系. 通过数值模拟, 我们得出成本函数关于最大库存量 S 是凸的. 因此, 我们能够得到最优库存容量 S^*, 使得 $C(S^*) = \min_S C(S)$.

考虑如下满足平稳条件 (3) 的系统参数:

$$\begin{aligned} &\lambda_1 = 0.5, \quad \lambda_2 = 0.1, \quad \mu_1 = 6, \quad \mu_2 = 2, \\ &\gamma = 2, \quad \beta = 0.1, \quad b_1 = 5, \quad b_2 = 2, \\ &h = 0.1, \quad c_1 = 0.5, \quad c_2 = 3, \quad N = 25. \end{aligned}$$

成本函数关于最大库存量 S 的变化由图 2 给出. 从图 2 中可以看出, 最优库存容量 $S^* = 4$,

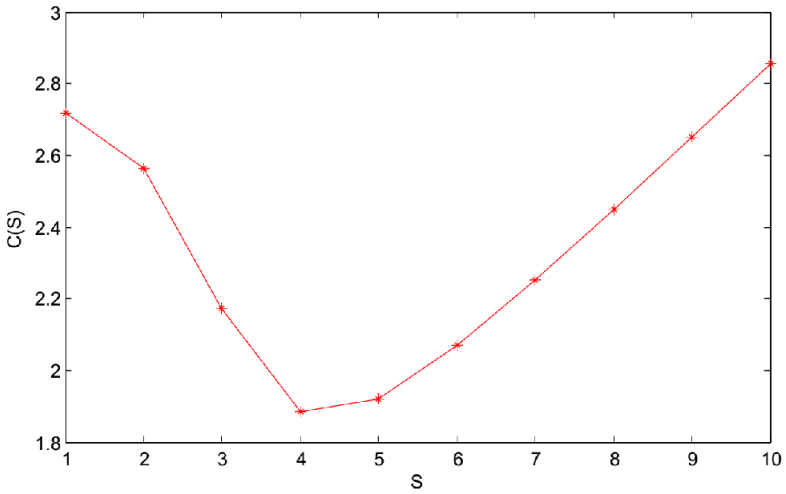


图 2　期望总成本随最大库存量 S 的变化

相应的最小成本 $C(S^*) = 1.8833$, 误差为 $\|\boldsymbol{\pi}_L^L - \boldsymbol{\pi}_{L-1}^L\|_1 = 6.9265 \times 10^{-4}$, 其中 $\|x\|_1 := |x_1| + |x_2| + \cdots + |x_d|$, $x \in \mathbb{R}^d$, L 为截断水平, $L = 25$. 更多的数值实验包括最优库存容量 S^* 在下节给出, 结果进一步表明了成本函数关于最大库存量 S 是凸的.

2) 敏感性分析

在本节中, 我们分析上述性能指标和最大库存量关于系统参数 λ_1, λ_2, μ_1, μ_2, b_1, b_2, γ 和 β 的敏感性.

表 1 到表 5 分别给出了在最优库存容量 S^* 条件下的系统性能指标关于系统参数 λ_i $(i=1,2)$, μ_i, b_i, γ 和 β 的变化, 其他未变化参数与上节一致.

表 1　系统性能指标关于 λ_1 和 λ_2 的变化

(λ_1,λ_2)	S^*	$E_{\rm W}$	$E_{\rm I}$	$E_{\rm CR}$	$E_{\rm LI}$	$C(S^*)$
$(0.3,0.1)$	4	0.617	0.392	0.003	0.148	1.102
$(0.4,0.1)$	4	0.890	0.471	0.004	0.172	1.456
$(0.5,0.1)$	4	1.240	0.551	0.005	0.195	1.883
$(0.6,0.15)$	5	1.735	0.901	0.007	0.316	2.777
$(0.6,0.2)$	4	2.229	0.735	0.007	0.258	3.080
$(0.6,0.25)$	4	2.402	0.778	0.008	0.274	3.305

表 2　系统性能指标关于 μ_1 和 μ_2 的变化

(μ_1,μ_2)	S^*	$E_{\rm W}$	$E_{\rm I}$	$E_{\rm CR}$	$E_{\rm LI}$	$C(S^*)$
$(7,1.5)$	4	1.617	0.602	0.006	0.238	2.394
$(8,1.5)$	4	1.378	0.583	0.006	0.231	2.134
$(9,1.5)$	4	1.178	0.564	0.005	0.224	1.910
$(10,2)$	4	0.572	0.428	0.004	0.148	1.059
$(10,3)$	3	0.458	0.237	0.003	0.064	0.676
$(10,4)$	3	0.396	0.200	0.002	0.046	0.555

表 3　系统性能指标关于 b_1 和 b_2 的变化

(b_1,b_2)	S^*	$E_{\rm W}$	$E_{\rm I}$	$E_{\rm CR}$	$E_{\rm LI}$	$C(S^*)$
$(3,1)$	4	0.661	0.551	0.005	0.195	1.304
$(4,1)$	4	0.742	0.551	0.005	0.195	1.385
$(5,1)$	4	0.823	0.551	0.005	0.195	1.466
$(6,1.5)$	4	1.113	0.551	0.005	0.195	1.756
$(6,2)$	4	1.322	0.551	0.005	0.195	1.965
$(6,2.5)$	5	1.338	0.711	0.005	0.253	2.172

根据计算结果, 注意到最优库存容量 S^* 和系统性能指标具有以下变化并从管理者的角度给出一些建议:

1. 最优库存容量 S^* 关于 λ_2, μ_2 是非增的, 关于 μ_1, b_2 是非降的. 最优库存容量 S^* 对

表 4　系统性能指标关于 γ 的变化

γ	S^*	E_{W}	E_{I}	E_{CR}	E_{LI}	$C(S^*)$
1.2	4	0.808	0.496	0.014	0.122	1.230
1.4	4	0.840	0.513	0.011	0.142	1.321
1.6	4	0.914	0.531	0.008	0.161	1.455
1.8	4	1.048	0.544	0.007	0.180	1.644
2	4	1.240	0.551	0.005	0.195	1.883
2.2	4	1.437	0.548	0.004	0.206	2.111

表 5　系统性能指标关于 β 的变化

β	S^*	E_{W}	E_{I}	E_{CR}	E_{LI}	$C(S^*)$
0.1	4	1.240	0.551	0.005	0.195	1.883
0.2	4	1.185	0.553	0.009	0.193	1.825
0.3	4	1.145	0.554	0.012	0.191	1.780
0.4	4	1.115	0.559	0.015	0.190	1.749
0.5	4	1.093	0.561	0.017	0.189	1.724
0.6	4	1.075	0.564	0.018	0.188	1.705

低优先权顾客具有更强的敏感性, 管理者需根据低优先权顾客的到达率和服务率适当调整最大库存量.

2. 期望等待成本 E_{W} 关于 λ_i, b_i, γ 递增, 关于 μ_i, β 递减. 为了减少等待成本, 管理者需提高服务率.
3. 期望库存水平 E_{I} 关于 λ_1, b_2, β 非减, 关于 λ_2, μ_i 非增.
4. 期望取消订货量 E_{CR} 关于 λ_i, β 是递增的, 关于 μ_i, γ 是递减的.
5. 库存期望失效水平 E_{LI} 关于 μ_i, β 是非增的, 关于 λ_1, b_i, γ 是非降的.
6. 期望总成本 $C(S^*)$ 关于 λ_i, b_i, γ 递增, 关于 μ_i, β 递减. 为了减少期望总成本, 管理者需降低顾客等待成本, 增加共有寿命并提高服务率.

§6. 结论与展望

本文通过构建一个水平相依的拟生灭过程, 研究了具有取消订货和共有寿命的非抢占优先权排队库存系统, 得到了系统的平稳条件和稳态概率向量, 同时给出了一些性能指标

和期望成本函数. 通过数值模拟, 我们得到了最优库存容量和最小成本. 最后, 我们通过对系统参数分析, 给管理者提供了建议.

当前工作可以往两个方向拓展, 第一个是将排队系统中的一些特殊模型结合起来, 比如休假排队模型和带反馈排队模型; 另一个是考虑库存系统中的其他补给策略, 比如 (r,Q) 订货策略等.

参 考 文 献

[1] KRISHNAMOORTHY A, SHAJIN D, LAKSHMY B. On a queueing-inventory with reservation, cancellation, common life time and retrial [J]. *Ann Oper Res*, 2016, **247(1)**: 365–389.

[2] KRISHNAMOORTHY A, SHAJIN D, LAKSHMY B. $GI/M/1$ type queueing-inventory systems with postponed work, reservation, cancellation and common life time [J]. *Indian J Pure Appl Math*, 2016, **47(2)**: 357–388.

[3] SHAJIN D, KRISHNAMOORTHY A. On a queueing-inventory system with impatient customers, advanced reservation, cancellation, overbooking and common life time [J]. *Oper Res*, 2021, **21(2)**: 1229-1253.

[4] SHAJIN D, KRISHNAMOORTHY A, DUDIN A N, et al. On a queueing-inventory system with advanced reservation and cancellation for the next K time frames ahead: the case of overbooking [J]. *Queueing Syst*, 2020, **94(1-2)**: 3–37.

[5] MILLER D R. Computation of steady-state probabilities for $M/M/1$ priority queues [J]. *Oper Res*, 1981, **29(5)**: 945–958.

[6] TAKINE T, HASEGAWA T. The workload in the $MAP/G/1$ queue with state-dependent services: its application to a queue with preemptive resume priority [J]. *Comm Statist Stochastic Models*, 1994, **10(1)**: 183–204.

[7] ZHU Y J, ZHOU Z H, FENG Y G. $M/G/1$ retrial queue system with priority and repair [J]. *Acta Automat Sinica*, 2008, **34(2)**: 195–201.

[8] 胡根生, 朱翼隽, 陈洋, 等. 优先权的 N 策略 $M/G/1$ 排队在通信网中的应用 [J]. 江苏大学学报 (自然科学版), 2003, **24(4)**: 82–86.

[9] SHALEV-OREN S, SEIDMANN A, SCHWEITZER P J. Analysis of flexible manufacturing systems with priority scheduling: PMVA [J]. *Ann Oper Res*, 1985, **3(3)**: 113–139.

[10] DE SOUZA R M, MORABITO R, CHIYOSHI F Y, et al. Incorporating priorities for waiting customers in the hypercube queuing model with application to an emergency medical service system in Brazil [J]. *European J Oper Res*, 2015, **242(1)**: 274–285.

[11] NEUTS M F, RAO B M. Numerical investigation of a multiserver retrial model [J]. *Queueing Syst*, 1990, **7(2)**: 169–189.

[12] LIAN Z T, LIU L M. A tandem network with MAP inputs [J]. *Oper Res Lett*, 2008, **36(2)**: 189–195.

[13] LIAN Z, LIU X, LIU L. Discriminatory processor sharing queues and the DREB method [J]. *Stoch Models*, 2008, **24(1)**: 19–40.

[14] NEUTS M F. Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach [M]. Baltimore: Johns Hopkins University Press, 1981.

Analysis of a Non-preemptive Priority Queueing-Inventory System with Common Life Time and Cancellation

LUO Xuxiang

(*School of Mathematics and Big Data, Foshan University, Foshan, 528000, China*)

LIU Zaiming

(*School of Mathematical and Statistics, Central South University, Changsha, 410083, China*)

Abstract: A queueing-inventory system with cancellation, common life time and non-preemptive priority is considered, in which customer arrives according to a Poisson process and service time follows an exponentially distribution. By formulating the system process into a level-dependent quasi-birth-and-death process (LDQBD), the stability condition and steady-state probability vectors under Neuts-Rao truncation method are obtained. Some performance measures and expected cost function are also given. Then, the optimal maximum inventory level and minimum cost are achieved through numerical simulations. Finally, the sensitivities analysis on major parameters are performed to provide more managerial insights.

Keywords: non-preemptive priority; queueing-inventory system; LDQBD; common life time; cancellation

2020 Mathematics Subject Classification: 90B22