

随机微分博弈模型中的库存管理问题： 马尔可夫链近似和最优策略

欧君恒，卢相刚

广东工业大学数学与统计学院，广东 广州

收稿日期：2024年3月28日；录用日期：2024年4月23日；发布日期：2024年4月30日

摘要

本文研究了在随机参考价格影响下，两个垄断厂商竞争下的生产和定价策略。设定的库存管理系统包括随机参考价格和随机需求。在随机微分博弈模型的框架下研究了库存管理问题，我们给出该背景下支付函数的定义。为了得到最优生产和定价，我们采用动态规划原理的方法，博弈的上下值满足一个耦合的非线性积分微分Hamilton-Jacobi-Isaacs (HJI)方程组。本文还证明了该对策问题鞍点的存在性，由于很难得到封闭形式的解，我们采用马尔可夫链近似来近似值函数和最优控制，并给出了收敛性分析。最后，我们进行了数值实验，并且根据实验结果，提出了相应的管理建议。

关键词

库存控制，随机需求与参考价格，马尔可夫链近似，随机微分对策

Inventory Management Problems in Stochastic Differential Game Models: Markov Chain Approximation and Optimal Policies

Junheng Ou, Xianggang Lu

School of Mathematics and Statistics, Guangdong University of Technology, Guangzhou Guangdong

Received: Mar. 28th, 2024; accepted: Apr. 23rd, 2024; published: Apr. 30th, 2024

Abstract

This paper investigates the production and pricing strategies employed by two monopolies in a

competitive environment, considering the influence of random reference prices. The inventory management system incorporates both random reference prices and random demand. The problem of inventory management is examined within the framework of a stochastic differential game model, with the payment function defined accordingly. To obtain optimal production and pricing decisions, dynamic programming principles are utilized, and coupled nonlinear integral differential Hamilton-Jacobi-Isaacs (HJI) equations that govern the upper and lower values of the game are established. Furthermore, the existence of saddle points for this game problem is proven. Due to challenges in obtaining closed-form solutions, Markov chain approximation is employed to approximate functions and optimal controls while providing convergence analysis. Finally, numerical experiments are conducted to validate our findings, leading to corresponding managerial recommendations.

Keywords

Inventory Control, Stochastic Demand and Reference Price, Markov Chain Approximation, Stochastic Differential Game

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. 引言

库存管理在供应链中起着至关重要的作用, 直接塑造了公司的运营效率和财务业绩。随着市场竞争的日益激烈, 有效地管理和优化库存已成为企业获得竞争优势不可或缺的因素。自从 Whitin 发表了关于库存问题的稳定需求曲线模型以来, 众多学者一直致力于探索如何更有效地降低公司的管理成本, 以实现利润最大化[1]。Bertsimas 和 De Boer 研究了一个周期性的多产品定价和库存控制问题, 考虑了生产和定价收入的影响作为影响因素, 进一步深入研究了库存控制的利润模型, 得出了更符合实际的结果[2]。Chen 等人将这个问题扩展到动态定价和动态生产环境中的库存控制, 考虑了可变条件对最大化预期长期利润的影响[3]。值得注意的是, 在上述研究中, 参考价格效应作为影响消费者购买意愿的关键因素并未得到足够的重视, 这在研究动态定价和生产问题时显然是不合理的。

Greenleaf 强调了参考价格效应对公司盈利能力的重要影响[4]。这种现象表现为, 如果销售价格低于参考价格, 买家会感觉利润并增加购买欲望; 相反, 如果销售价格高于参考价格, 买家会感觉亏损, 导致购买欲望下降。Hu 等人从长期角度关注动态销售定价的整体收入, 考虑消费者囤积行为, 但没有探讨库存方面[5]。Chen 等人进一步考虑了库存的影响, 系统地参考价格效应引入动态库存控制问题[6]。通过一种新的转移方法, 得到了一个凹形的长期期望收益函数, 解释了最优基本库存水平和目标参考价格对当前参考价格的依赖性。Chen 等人假设参考价格是一个固定值; 然而, 将参考价格的可变性视为随机和客户感知的波动似乎更现实[6]。因此, 在此基础上, 我们引入了随机参考价格的概念。

参考价格受顾客的主观感受影响, 受商品平均价格、折扣、品牌效应等因素影响, 具有明显的记忆性。Chen 等人认识到参考价格的随机性和无限连续时间演化下的最优定价, 推导出明确表达的最优定价策略[7]。受此启发, Cao 和 Duan 首次将随机参考价格引入库存问题, 分析了基于不确定需求和参考价格的随机最优库存控制[8]。最优定价与库存水平负相关, 与参考价格水平正相关。制造商利用参考价格效应, 根据客户记忆参数确定价格, 调整库存水平以控制运营成本, 同时确保正常出货水平被视为最佳控制。

在上述研究中, 研究人员经常假设制造业处于垄断地位; 然而, 在现实中, 涉及两家或两家以上公

公司的竞争是普遍存在的。寻求更低的经营风险和占领更多的市场份额已成为企业的重点。公司必须考虑降低其运营风险, 同时最大限度地减少竞争对手风险的影响。Pang 和 Fukushima 提出了一个多领导者和追随者的竞争市场模型, 其中不存在合作, 建立了风险变化下的奖惩机制[9]。Elliott 和 Siu 考虑了在零和博弈背景下进行风险预测的必要性, 通过有限马尔可夫链预测市场变化下的消费者选择变化[10]。Adida 和 Perakis 将上述研究应用于更现实的情景, 设想两个公司在市场上竞争, 寻求获得对方的市场份额以获得垄断优势[11]。一家公司的决策影响另一家公司的市场决策, 遵循纳什均衡。随着时间的推移, 两家公司都寻求通过动态定价和库存控制来提高竞争力, 从而实现利润最大化。尽管研究人员试图尽可能真实地模拟该模型, 但他们没有考虑参考价格对消费者购买意愿的重大影响。考虑到参考价格和消费者购买意愿之间的密切相关性, 理解公司产品的市场竞争力应该包括考虑消费者的看法。为了获得更真实的结果, 将随机参考价格的影响纳入博弈模型是必要的。

这个模型福尔斯一个追逐 - 逃避博弈, 是一种定性微分博弈。Berkovitz 和 Drescher 使用变分方法显式求解了战争模型中的追逃博弈的最优解问题, 提供了一种使用微分方程求解最优控制的方法[12]。Fleming 专注于有限连续时间微分对策的收敛性, 进一步描述了在预定初始条件下的对策解[13]。Elliott 和 Kalton 提出用微分方程组来求解零和动态对策, 并分析了微分对策中值的存在性[14]。Jin 等人引入了两家保险公司之间利润再投资的博弈场景, 使用状态转移模型来捕捉环境不确定性的变化[15]。它们之间的竞争, 可以用微分方程来分析。考虑到 Hamilton-Jacobi-Isaacs (HJI) 方程组很难得到封闭解, 利用马尔可夫链近似观察盈余变化, 预测动态规划中鞍点的存在性。

马尔可夫链方法在博弈模型中的应用, 显著地强调了近似结果在决策预测中的关键作用。采用马尔可夫链近似有助于处理模型中的非均匀可分析分段函数所带来的挑战。最初, Pan 和 Basar 研究了状态跳跃下的不完美测量方案, 提出了一种当系统函数呈现分段特征时使用离散时间马尔可夫链进行近似的方法[16]。Yin 等人继续了 Pan 和 Basar 的方法, 通过研究奇异摄动马尔可夫链来解决离散时间中大规模复杂系统的控制优化问题, 显著降低了系统的复杂性[17]。

受 Jin 等人的启发, 他们使用马尔可夫链近似来构造离散时间控制的马尔可夫链以获得最优解, 我们将马尔可夫链近似应用于解决企业中库存控制和产品运营的博弈问题[15]。建立了考虑参考价格影响的定价与库存控制模型。在该模型中, 企业的初始品牌价值和利润收入不相等, 市场需求不确定, 消费者对两家企业产品的记忆参数不同, 初始参考价格也不同。企业可以在连续的时间内随时调整自己的经营策略, 寻求挤压竞争对手的市场份额, 直到它们之间的利润差距扩大或缩小到一定值。这表明一家公司无法再挑战另一家公司的市场地位, 该模式停止。

本文的主要贡献如下:

(1) 本文首次利用马尔可夫链近似方法求解了两个公司在参考价格作用下的联合生产定价模型的随机最优控制问题。

(2) 进行了数值实验, 我们对得到的结果给出分析和管理的见解。

其余的工作组织如下: 在第 2 节中, 我们提出了两个制造商公司竞争下的随机微分对策的一般形式和假设。在第 3 节中, 我们讨论了马尔可夫链近似方法的数值算法。用近似马尔可夫链很好地逼近了博弈的上、下界, 并给出了动态规划方程。在第 4 节中, 我们讨论了近似格式的收敛性。证明了鞍点的存在性。最后, 在第 5 节中, 我们进行了数值实验, 并给出了我们的结果。在此基础上, 对数值模拟结果进行了合理的分析和讨论, 并给予相应的管理意见。

2. 模型

我们使用一个随机微分方程来表示公司 1 的参考价格:

$$dR^{(1)}(t) = \beta(p_1(t) - R^{(1)}(t))dt + \epsilon^{(1)}\sqrt{R^{(1)}(t)}dY^{(1)}(t) \quad (1)$$

平方根扩散项 $\epsilon^{(1)}\sqrt{R^{(1)}(t)}$ 表示随机参考价格的波动性。这里, $\epsilon^{(1)}$ 为参考价格的波动性参数。参考价格的内存参数或调整速度用 β 表示。制造公司 1 的累积消费者需求 $D^{(1)}(t)$ 是一个随机过程, 由:

$$dD^{(1)}(t) = (a - bp_1(t) + \gamma(R^{(1)}(t) - p_1(t)))dt - \sigma dW(t), t \geq 0 \quad (2)$$

潜在需求的波动性参数用 $\sigma^{(1)}$ 表示, 维纳过程 $w^{(1)}(t)$ 仍然与 $Y^{(1)}(t)$ 保持不相关。库存的连续时间动态受生产率和需求率之间的相互作用的控制, 遵循一个随机微分方程:

$$dX^{(1)}(t) = u_1(t)dt - dD^{(1)}(t) \quad (3)$$

考虑到公式(2), 上述库存动态变为:

$$dX^{(1)}(t) = (u_1(t) - (a - bp_1(t) + \gamma(R^{(1)}(t) - p_1(t))))dt + \sigma dW(t), X^{(1)}(0) = x^{(1)}, t \geq 0. \quad (4)$$

假设连续时间公司盈利过程包括价格 $p_1(t)$ 、需求率 $D_1(t)$ 、生产成本 $C(u_1(t))$ 和库存成本 $H(X^{(1)}(t))$, 它满足随机微分方程:

$$dZ^{(1)}(t) = [p_1(t)dD^{(1)}(t) - (C(u_1(t)) + H(X^{(1)}(t)))]dt \quad (5)$$

考虑到(2), 制造公司 1 的剩余过程可以写为:

$$dZ^{(1)}(t) = [p_1(t)(a - bp_1(t) + \gamma(R^{(1)}(t) - p_1(t))) + p_1(t)(a - bp_1(t) + \gamma(R^{(1)}(t) - p_1(t))) - (C(u_1(t)) + H(X^{(1)}(t)))]dt - p_1(t)\sigma_1 dW^{(1)}(t), Z^{(1)}(0) = z^{(1)}, t \geq 0$$

同样, 与制造公司 1 竞争的制造公司 2 也出现了随后的盈余

$$dZ^{(2)}(t) = [p_2(t)(a - bp_2(t) + \gamma(R^{(2)}(t) - p_2(t))) + p_2(t)(a - bp_2(t) + \gamma(R^{(2)}(t) - p_2(t))) - (C(u_2(t)) + H(X^{(2)}(t)))]dt - p_2(t)\sigma_2 dW^{(2)}(t), Z^{(2)}(0) = z^{(2)}, t \geq 0.$$

在本研究中, 我们将描述两家制造企业之间的竞争。这两家公司的表现是通过其盈余 $Z^{(1)} - Z^{(2)}$ 的差值来衡量的。在不失一般性的前提下, 我们假设 $Z^{(1)} > Z^{(2)}$ 。盈余越大的公司努力扩大盈余差距, 而盈余越小的公司则努力减少盈余差距, 使差距最小。因此, 这两家公司之间的竞争形成了一个双人游戏, 每个参与者都可以根据其生产和定价策略来响应其竞争对手的策略。设 $\bar{Z}(t) = Z^{(1)} - Z^{(2)}$ 。因此, 两个盈余 $\bar{Z}(t)$ 的差异由以下动态控制:

$$d\bar{Z}(t) = [p_1(t)(a - bp_1(t) + \gamma(R^{(1)}(t) - p_1(t))) - (C(u_1(t)) + H(X^{(1)}(t))) - p_2(t)(a - bp_2(t) + \gamma(R^{(2)}(t) - p_2(t))) - p_1(t)\sigma_1 dW^{(1)}(t) + p_2(t)\sigma_2 dW^{(2)}(t), \bar{Z}(0) = \bar{z}, t \geq 0. \quad (6)$$

其中 $\bar{z} = z^{(1)} - z^{(2)}$ 。为了更好地研究公司 1 和公司 2 的需求过程之间的关系, 我们假设公司 1 和公司 2 的累积需求过程之间存在相关性。设 α 为 $w^{(1)}(\cdot)$ 与 $w^{(2)}(\cdot)$ 、 $\alpha \in (-1, 1)$ 之间的相关系数。通过这种方式, 我们可以将 $w^{(2)}(\cdot)$ 重写如下:

$$dW^{(2)}(t) = \nu dW^{(1)}(t) + \sqrt{1 - \nu^2} d\tilde{W}^{(2)}(t) \quad (7)$$

$\tilde{W}^{(2)}(\cdot)$ 是一个标准的布朗运动, 并且独立于 $w^{(1)}(\cdot)$ 。我们允许将盈余以价格投资于金融市场上的资产 $M(t)$, μ 为资产的收益率, σ_3 为相应的波动率, $w^{(3)}(t)$ 为标准布朗运动。并且我们假设它满足以下形式:

$$\frac{dM(t)}{M(t)} = \mu dt + \sigma_3 dW^{(3)}(t) \quad (8)$$

因此, 结合(7)和(8), 我们可以将等式(6)改写如下:

$$\begin{aligned} dZ(t) = & [\mu Z(t) + p_1(a - bp_1(t) + \gamma(R^{(1)}(t) - p_1(t))) \\ & - C(u_1(t) - H(X^{(1)}(t))) - p_2(t)(a - bp_2(t) \\ & \gamma(R^{(2)}(t) - p_2(t)) + C(u_2(t)) + H(X^{(2)}(t))] dt \\ & + (p_2(t)\sigma_2\nu - p_1(t)\sigma_1)dW^{(1)}(t) + p_2(t)\sigma_2\sqrt{1-\nu^2}d\tilde{W}^{(2)}(t) \\ & + \sigma_3 Z(t)dW^{(3)}(t) \end{aligned} \quad (9)$$

为了简化过程, 我们将在下面用 6 维的形式来表示它:

$$\begin{aligned} dL(t) &= f(L(t), \theta(t))dt + \sigma(L(t))d\bar{W} \\ L(t) &= (Z(t), X^{(1)}(t), X^{(2)}(t), R^{(1)}(t), R^{(2)}(t))^T, \quad f = (f_1, f_2, f_3, f_4, f_5)^T, \\ \sigma &= \begin{pmatrix} p_2(t)\sigma_2\nu - p_1(t)\sigma_1 & p_2(t)\sigma_2\sqrt{1-\nu^2} & 0 & \sigma_3 Z(t) & 0 & 0 \\ \sigma_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \epsilon_1\sqrt{R^{(1)}(t)} & 0 \\ 0 & 0 & 0 & 0 & 0 & \epsilon_2\sqrt{R^{(2)}(t)} \end{pmatrix} \end{aligned}$$

$W(t) = (W^{(1)}(t), \tilde{W}^{(2)}(t), W^{(3)}(t), Y^{(1)}(t), Y^{(2)}(t))$, 和 f_1, f_2, f_3, f_4, f_5 具有以下形式

$$\begin{aligned} f_1 &= \mu Z(t) + p_1(t)(a - bp_1(t) + \gamma(R^{(1)}(t) - p_1(t))) \\ & \quad - (C(u_1(t)) + H(X^{(1)}(t))) - p_2(t)(a - bp_2(t) \\ & \quad + \gamma(R^{(2)}(t) - p_2(t))) + (C(u_2(t)) + H(X^{(2)}(t))) \\ f_2 &= u_1(t) - (a - bp_1(t) + \gamma(R^{(1)}(t) - p_1(t))) \\ f_3 &= u_2(t) - (a - bp_2(t) + \gamma(R^{(2)}(t) - p_2(t))) \\ f_4 &= \beta(p_1(t) - R^{(1)}(t)) \\ f_5 &= \beta(p_2(t) - R^{(2)}(t)) \end{aligned}$$

$L(0) = l = (z, x^{(1)}, x^{(2)}, r^{(1)}, r^{(2)})$, 对于所有 $t < \tau$, $\tau = \inf\{t \geq 0: Z(t) \notin (a, b)\}$ 表示退出游戏的时间, a, b 是满足一个 $a < b$ 的常数。我们假设, 如果盈余之差的绝对值太大, 游戏就会停止。无论是积极的还是消极的方向, 巨大的差异意味着一家公司主导市场并赢得游戏。

我们根据两个公司在到达下势垒之前的盈余差的可能性, 建立了值函数。在这个游戏中, 公司 1 努力最大化这个概率, 而公司 2 则努力最小化相同的概率。随后, 我们确定了值函数。用 $\rho > 0$ 表示折扣系数。设 Γ_{11} 和 Γ_{21} 分别是所有生产策略的集合, 假设它们是紧集。 Γ_{12} 和 Γ_{22} 分别是定价策略, 它们也被假定为紧集。让 $\theta = (\theta_1, \theta_2)$, $\theta_k = (u_k, p_k), k = 1, 2$, $\theta_1 \in \Gamma_1 = \Gamma_{11} \times \Gamma_{12}, \theta_2 \in \Gamma_2 = \Gamma_{21} \times \Gamma_{22}$, 以及 $\theta \in \Gamma = \Gamma_{11} \times \Gamma_{12} \times \Gamma_{21} \times \Gamma_{22}$ 。

假设当剩余差异的大小变得过于重要时, 游戏就会停止。无论是在积极的还是消极的领域, 巨大的差距意味着一家公司的市场主导地位的游戏和胜利。我们将收益函数建立为两个参与者在达到 5 个下阈值之前达到上阈值的盈余差的可能性。因此, 在这个游戏中, 公司 1 的目标是最大化概率, 而公司 2 则寻求最小化相同的概率。用 $\rho > 0$ 表示折扣系数。让 $\theta = (\theta_1, \theta_2)$ 和 $\theta_i = (u_i, p_i), i=1,2, \theta \in \Gamma = \Gamma_1 \times \Gamma_2$ 。对于任何可接受的控制 θ , 支付函数是

$$\begin{aligned} \mathcal{J}(l, \theta) &= \mathbb{P}\{Z(\tau) = b \mid L(0) = l\} \\ &= E_z \left[\int_0^\tau e^{-\rho t} K(L(t), \theta(t)) dt \right] \end{aligned} \tag{10}$$

如果 u_1, p_1, u_2, p_2 满足以下条件, 则控制 $\theta = (\theta_1, \theta_2)$ 被认为是可接受的。

- (1) $u_1(t), p_1(t), u_2(t), p_2(t)$ 对于任何 $t \geq 0$ 都为非负值;
- (2) 对于任何 $t \leq \tau, Z(t) \in (a, b)$;
- (3) $u_1(t), p_1(t), u_2(t), p_2(t)$ 是对至少包含 $\sigma\{W(s), 0 \leq s \leq t\}$ 的 \mathcal{F}_t 是可适的;
- (4) $\mathcal{J}(l, \theta) < \infty$ 对于任何 $l \in G$ 和容许对 $\theta = (u_1, p_1, u_2, p_2)$, 其中 \mathcal{J} 是在(10)中定义的函数。

假设 Γ 是可能的保留级别 $\theta(t)$ 的集合。在本文中, 我们假设 Γ 是一个给定的紧集, 并且函数 $K(t)$ 相对于 (θ_1, θ_2) 是凹凸的。

对于 $k=1,2$, 设 $\mathcal{B}(\Gamma_k \times [0, \infty))$ 为 $\Gamma_k \times [0, \infty)$ 的 Borel 子集的 σ 代数。对于每个 $t \geq 0$, 有一个可接受的松弛控制 $m_k(\cdot, \cdot)$ 是对 $\Gamma_k \times [0, \infty)$ 的一个度量, 这样的 $m_k(\Gamma_k \times [0, t]) = t$ 。在给定的概率空间下, 如果 $m_k(\cdot)$ 是 $W(\cdot)$ 或 $(m(\cdot), W(\cdot))$ 的, 并且 $m_k(\cdot, \cdot, \omega)$ 是所有的 $n_k(A \times B \times [0, t])$ 的确定性松弛控制。有一个导数 $m_{t,k}(\cdot, \cdot)$, 使得 $m_{t,k}(\cdot, \cdot)$ 对于所有的 $A \times B \in \mathcal{B}(\Gamma_k)$ 是 \mathcal{F}_t 可适的。给定 $\theta_k(\cdot)$ 的放松控制 $m_k(\cdot, \cdot)$, 其中 $k=1,2$ 。我们定义导数 $m_{t,k}(\cdot, \cdot)$ 使得

$$m_1(A, B) = \int_{\Gamma_{11} \times \Gamma_{12} \times [0, \infty)} I_{\{(\theta_1, \cdot) \in \mathcal{B}(\Gamma_1 \times [0, \infty))\}} m_{t,1}(d\phi) dt$$

对于所有有界的 Borel 集 $A \times B \in \mathcal{B}(\Gamma_1 \times [0, \infty))$ 和对于每个 $t, m_{t,1}(\cdot, \cdot)$ 是对满足 $m_{t,1}(\Gamma_1) = 1$ 在 $\mathcal{B}(\Gamma_1)$ 的测度。例如, 我们可以用任何方便的方式定义 $t=0$ 的 $m_{t,1}(\cdot, \cdot)$, 并定义为 $t > 0$ 的左导数,

$$m_{t,1}(A, B) = \lim_{\delta \rightarrow 0} \frac{n_1(A \times B \times [t - \delta, t])}{\delta}, A \times B \in \mathcal{B}(\Gamma_1)$$

要注意的是 $m_k(d\phi dt) = m_{t,1}(d\phi) dt$, 自然可以定义 $m_1(\cdot, \cdot)$ 的松弛控制表示 $\theta_1(\cdot, \cdot)$,

$$m_{t,1}^h(A) = I_{\{\theta_1(\cdot) \in A \times B\}}, \forall A \times B \in \mathcal{B}(\Gamma_1)$$

类似地, 当 $k=1,2$ 时, 我们有

$$m_2(A, B) = \int_{\Gamma_2 \times [0, \infty)} I_{\{(\theta_2, \cdot) \in \mathcal{B}(\Gamma_2 \times [0, \infty))\}} m_{t,2}(d\phi) dt$$

用导数 $m_t(\cdot, \cdot, \cdot) = (m_{t,1}(\cdot, \cdot) \times m_{t,2}(\cdot, \cdot))$ 来定义松弛控制 $m(\cdot, \cdot, \cdot) = (m_1(\cdot, \cdot) \times m_2(\cdot, \cdot))$ 。因此 $m(\cdot, \cdot, \cdot)$ 是 $(\Gamma_{11} \times \Gamma_{12} \times \Gamma_{21} \times \Gamma_{22}) \times [0, \infty)$ 的 Borel 集上的一个测度。基于[18]的方法和假设, 我们开始定义上值、下值和鞍点。设 Y_k 是 θ_k 的可接受的普通控制。对于 $\Delta > 0$, 用 $Y_k(\Delta)$ 表示分段常数控制 $\theta_k = (u_{1,k}(\cdot), p_{1,k}(\cdot))$ 上的区间 $[[n_1\Delta, (n_1+1)\Delta], [n_2\Delta, (n_2+1)\Delta]]$, $n_1 = 0, 1, 2, \dots, n_2 = 0, 1, 2, \dots$, 其中 $\theta_k(n_1\Delta, n_2\Delta)$ 是 $\mathcal{F}_{n\Delta}$ 测度, 并且 $Y_k(\Delta, \Delta) \subset Y_k$ 。我们设 A_i 为 Γ_1 的 Borel 子集, $\mathcal{R}_1(\Delta, \Delta) \subset Y_1(\Delta, \Delta)$ 表示由 $\mathcal{R}_1(\Delta, \Delta) \subset Y_1(\Delta, \Delta), i = 0, 1, 2, \dots, j = 0, 1, 2, \dots$ 所表示的分段常数控制集的条件概率类型:

$$P\{\theta_1(in_1, jn_2) \in A_i \mid W(s_1, s_2), \theta_2(s_1, s_2), s_1 < in_1, s < jn_2; \theta_1(n_1\Delta, n_2\Delta), n_1 < i, n_2 < j\} \tag{11}$$

在此 $F_{1,ij}(A_i, \cdot, \cdot)$ 是 $A_i \in \mathcal{B}(\Gamma_1)$ 的可测函数。如果控制规则 $\theta_1(\theta_2)$ 由(11)给出, 则我们表示 $\theta_1(\theta_2)$ 来强调

θ_1 对 θ_2 的依赖性。类似地, 我们定义了 $\mathcal{R}_2(\Delta, \Delta)$ 和相关的控制规则 $\theta_2(\theta_1)$ 。为了继续, 我们定义上值和下值。上面的值被定义为

$$V^+(l) = \lim_{\Delta \rightarrow 0} \inf_{\theta_1 \in \mathcal{R}_1(\Delta, \Delta)} \sup_{\theta_2 \in \Gamma_2} \mathcal{J}(l, \theta_1(\theta_2), \theta_2)$$

同样的, 我们可以定义下值为

$$V^-(l) = \lim_{\Delta \rightarrow 0} \inf_{\theta_2 \in \mathcal{R}_2(\Delta, \Delta)} \sup_{\theta_1 \in \Gamma_1} \mathcal{J}(l, \theta_2(\theta_1), \theta_1)$$

如果下值和上值相等, 那么我们就说存在一个鞍点。

$$V(l) = V^+(l) = V^-(l) \tag{12}$$

对于任意的 $\theta \in \Gamma$ 和 $V(\cdot, \cdot, \cdot, \cdot) \in C^5(\mathbb{R})$, 定义一个操作符 \mathcal{L}^θ

$$\begin{aligned} \mathcal{L}^\theta V = & V_z [\mu z + p_1 (a - bp_1 + \gamma (R^{(1)}(t) - p_1)) - (C(u_1) + H(X^{(1)}))] \\ & - p_2(t) (a - bp_2 + \gamma (R^{(2)}(t) - p_2)) + (C(u_2) + H(X^{(2)})) \\ & + V_{x^{(1)}} [u_1 - (a - bp_1 + \gamma (R^{(1)}(t) - p_1))] \\ & + V_{x^{(2)}} [u_2 - (a - bp_2 + \gamma (R^{(2)}(t) - p_2))] \\ & + V_{r^{(1)}} [\beta (p_1 - r^{(1)})] + V_{r^{(2)}} [\beta (p_2 - r^{(2)})] \\ & - \frac{1}{2} \sigma_1^2 p_1 V_{x^{(1)}z} - \frac{1}{2} \sigma_1^2 p_1 V_{zx^{(2)}} + \frac{1}{2} \sigma_2^2 p_2 V_{x^{(2)}z} + \frac{1}{2} \sigma_2^2 p_2 V_{zx^{(2)}} \\ & + \frac{1}{2} \sigma_1^2 V_{x^{(1)}x^{(1)}} + \frac{1}{2} \sigma_2^2 V_{x^{(2)}x^{(2)}} + \frac{1}{2} \epsilon_1^2 r^{(1)} V_{r^{(1)}r^{(1)}} + \frac{1}{2} \epsilon_2^2 r^{(2)} V_{r^{(2)}r^{(2)}} \\ & + \frac{1}{2} (\sigma_1^2 p_1^2 + \sigma_2^2 p_2^2 + \sigma_3^2 z^2) V_{zz} \end{aligned}$$

其中 $V_i = \partial V / \partial i$ 以及 $V_{ii} = \partial^2 V / \partial i^2$, $i = z, x^{(1)}, x^{(2)}, r^{(1)}, r^{(2)}$ 。另一方面, 鞍点 $V(l)$, $l = (z, x^{(1)}, x^{(2)}, r^{(1)}, r^{(2)})$ 满足 HJI 方程

$$\begin{aligned} & \inf_{\theta_1 \in \Gamma_1} \sup_{\theta_2 \in \Gamma_2} (\mathcal{L}^\theta (V(x)) - \rho V(x) + K(l, \theta)) \\ & = \sup_{\theta_2 \in \Gamma_2} \inf_{\theta_1 \in \Gamma_1} (\mathcal{L}^\theta (V(x)) - \rho V(x) + K(l, \theta)) \\ & = 0, l \in (a, b) \\ & V(x) = \tilde{h}(l), l = a, b. \end{aligned} \tag{13}$$

3. 数值算法

现在, 我们寻找一个对 $L(t)$ 的有限状态的马尔可夫链近似, 并且这个马尔可夫链的形式有利于离散化的模拟计算。利用弱收敛理论证明了我们构造的马尔可夫链收敛于 $L(t)$ 。设 $h > 0$ 为离散步长, 边界点 a, b 为 h 的整数倍, e_i 为第 i 个坐标方向上的单位向量, $i = 1, 2, 3, 4, 5$ 。近似受控马尔可夫链具有五维晶格的状态空间, n_i 为整数, $i = 1, 2, 3, 4, 5$ 。

$$L_h^5 = \left\{ l = h \sum_{i=1}^5 n_i e_i \right\}$$

设 G_h 表示 G 上的有限差分网格, 其中 $G = [a, b]$ 和 $G^0 = (a, b)$ 。也就是说, 如果 $l \in L_h^5$, 那么它存在非负

整数 n_1, n_2, n_3, n_4, n_5 使得 $l = \sum_i e_i h n_i$, 定义 $G_h = I_h^S \cap G$ 。现在, 对于一些函数 $\Delta t^h(y)$, 它是 $h \rightarrow 0$, 我们假设马尔可夫链满足以下要求:

$$\begin{aligned} E_l [\xi_{n+1}^h - \xi_n^h | \xi_n^h = y] &= f(y) \Delta t^h(y) + o(\Delta t^h(y)) \\ \text{cov}_l [\xi_{n+1}^h - \xi_n^h | \xi_n^h = y] &= 2a(y) \Delta t^h(y) + o(\Delta t^h(y)) \\ P_l \{ |\xi_{n+1}^h - \xi_n^h| \geq \epsilon | \xi_n^h = y \} &= o(\Delta t^h(y)) \\ 2a(y) &= \sigma(y) \sigma'(y) \end{aligned}$$

设 $\theta^h := \theta_n^h, n > 0$ 为控制动作的顺序。如果 $\theta_n^h = (\theta_{1,n}^h, \theta_{2,n}^h)$ 是 $\sigma\{\xi_0^h, \dots, \xi_n^h, \theta_0^h, \dots, \theta_{n-1}^h\}$ 可适的, 那么序列 θ^h 可以被认为是允许的。由于我们考虑的是连续时间情况, 控制离散时间马尔可夫链改写为

$$\xi^h = \xi_k^h, \theta^h = \theta_k^h,$$

对于 $t \in [t_n^h, t_{n+1}^h)$, 其中 $t_0^h := 0, t_n^h = \sum_{k=0}^{n-1} \Delta t^h(\xi_k^h, \theta_k^h)$, 让 $\eta_h := \inf\{n: \xi_n^h \in \partial G\}$ 。那么 ξ^h 从 G 的第一次是 $\tau^h = \tau_{\eta_h}^h$ 。让 $\xi_0^h = l \in G_h$ 和 $\xi_0^h = l \in G_h$ 是容许控制, 将受控马尔可夫链的收益函数定义为

$$\mathcal{J}^h(l, \theta^h) = \mathbb{P}_{z^h} \{ Z^h(\tau^h) = b | Z^h(0) = z^h \} = E_l \sum_{k=1}^{\eta_h-1} e^{-\rho_k^h} [K(\xi_k^h, \theta_k^h) \Delta t_k^h]$$

我们用 $\theta_n^h = (\theta_{1,n}^h, \theta_{2,n}^h) \subset \Gamma$ 表示随机变量, 即链在 n 时刻的规则控制作用。设 $\theta^h := \{\theta_n^h, n \geq 0\}$ 为控制动作的顺序。此外, 设 $\Gamma_{1,k}^h$ 表示公司 k 为首发的控制集合, 它由一系列可测函数 $F_n^h(\cdot)$ 决定, 使得

$$\theta_{k,n}^h = F_n^h(\xi_i^h, i \leq n)$$

$\Gamma_{2,k}^h$ 表示播放 k 随后进行的普通控制的集合, 该策略由一个可测函数 $\tilde{F}_n^h(\cdot)$ 的序列定义

$$\theta_{k,n}^h = \tilde{F}_n^h(\xi_i^h, i \leq n; \theta_i^h, i \leq n; \theta_{j,n}^h, j \neq k)$$

将 \mathcal{O}_t^h 定义为由 $\{\xi^h(s), \theta^h(s), s < t\}$ 生成的最小 σ 代数。此外, $\Gamma_k^h = \Gamma_{1,k}^h \times \Gamma_{2,k}^h$ 等价于关于 \mathcal{O}_t^h 的所有分段常数容许控制的集合。现在, 我们可以用 $\mathcal{J}^h(\cdot)$ 和 Γ^h 有来定义上下值

$$V^{h,+}(l) = \inf_{\theta_1^h \in \Gamma_{1,1}^h} \sup_{\theta_2^h \in \Gamma_{2,2}^h} \mathcal{J}^h(l, \theta_1^h, \theta_2^h)$$

而下值也可以被定义为

$$V^{h,-}(l) = \sup_{\theta_1^h \in \Gamma_{1,2}^h} \inf_{\theta_2^h \in \Gamma_{2,1}^h} \mathcal{J}^h(l, \theta_1^h, \theta_2^h)$$

如果存在一个鞍点, $V^h(l) = V^{h,+}(l) = V^{h,-}(l)$ 。可以很容易地验证了 $V^h(l)$ 满足以下动态规划方程

$$\begin{aligned} V^{h,+}(l) &= \inf_{\theta_1 \in \Gamma_1} \left\{ \sup_{\theta_2 \in \Gamma_2} \sum_y e^{-\rho^h} p^\theta((x, y) | \theta) V^h(y) \right\} \\ V^{h,-}(l) &= \sup_{\theta_2 \in \Gamma_2} \left\{ \inf_{\theta_1 \in \Gamma_1} \sum_y e^{-\rho^h} p^\theta((x, y) | \theta) V^h(y) \right\} \end{aligned} \tag{14}$$

让

$$\begin{aligned} f_1(l) &= \mu z + p_1(a - bp_1 + \gamma(r^{(1)} - p_1)) - (C(u_1) + H(x^{(1)})) \\ &\quad - p_2(a - bp_2 + \gamma(r^{(2)} - p_2)) - (C(u_2) + H(x^{(2)})) \\ f_2(r^{(1)}) &= u_1 - (a - bp_1 + \gamma(r^{(1)} - p_1)) \end{aligned}$$

$$\begin{aligned}
 f_3(r^{(2)}) &= u_2 - (a - bp_2 + \gamma(r^{(2)} - p_2)) \\
 f_4(r^{(1)}) &= \beta(p_1 - r^{(1)}), f_5(r^{(2)}) = \beta(p_1 - r^{(2)})
 \end{aligned} \tag{15}$$

以及 $f_i^+ = \max(f_i, 0)$, $f_i^- = \max(-f_i, 0)$ 。其中, 我们称 f_i^+ 和 f_i^- 分别为 f_i 的正负部分, $i=1,2,3,4,5$ 。对 V 的一阶导数和二阶导数使用以下近似值:

$$\begin{aligned}
 \Delta^2 V_{l_i} &= \frac{V(l+he_i) + V(l-he_i) - 2V(l)}{h^2} \\
 \Delta^2 V_{l_i} &= \frac{V(l+he_i) + V(l-he_i) - 2V(l)}{h^2} \\
 \Delta^2_+ V_{l_i l_j} &= \frac{2V(l) + V(l+he_i + he_j) + V(l-he_i - he_j)}{2h^2} \\
 &\quad - \frac{V(l+he_i) + V(l-he_i) + V(l+he_j) + V(l-he_j)}{2h^2} \\
 \Delta^2_- V_{l_i l_j} &= -\frac{2V(l) + V(l+he_i - he_j) + V(l-he_i + he_j)}{2h^2} \\
 &\quad + \frac{V(l+he_i) + V(l-he_i) + V(l+he_j) + V(l-he_j)}{2h^2}
 \end{aligned} \tag{16}$$

接下来, 我们使用上面给出的有限差分来近似 $V(l)$, $i=1,2,3,4,5$ 。

$$\begin{aligned}
 V(l) &\rightarrow V^h(l), \\
 V_{l_i}(l) &\rightarrow \Delta^+ V_{l_i}^h(l), f_i > 0, \\
 V_{l_i}(l) &\rightarrow \Delta^- V_{l_i}^h(l), f_i < 0, \\
 V_{l_i l_j}(l) &\rightarrow \Delta^2_+ V_{l_i l_j}^h(l), a_{ij}(l) \geq 0, i \neq j \\
 V_{l_i l_j}(l) &\rightarrow \Delta^2_- V_{l_i l_j}^h(l), a_{ij}(l) \leq 0, i \neq j
 \end{aligned}$$

通过将(15)和(16)代入(12), 并重新排列这些项, 我们得到了动态规划的等价形式为

$$\begin{aligned}
 &\rho V^h(l) \\
 &= f_1^+(l) \frac{V^h(l+he_1) - V^h(l)}{h} - f_1^-(l) \frac{V^h(l) - V^h(l-he_1)}{h} \\
 &\quad + f_2^+(l) \frac{V^h(l+he_2) - V^h(l)}{h} - f_2^-(l) \frac{V^h(l) - V^h(l-he_2)}{h} \\
 &\quad + f_3^+(l) \frac{V^h(l+he_3) - V^h(l)}{h} - f_3^-(l) \frac{V^h(l) - V^h(l-he_3)}{h} \\
 &\quad + f_4^+(l) \frac{V^h(l+he_4) - V^h(l)}{h} - f_4^-(l) \frac{V^h(l) - V^h(l-he_4)}{h} \\
 &\quad + f_5^+(l) \frac{V^h(l+he_5) - V^h(l)}{h} - f_5^-(l) \frac{V^h(l) - V^h(l-he_5)}{h} \\
 &\quad + \sigma_1^2 p_1 \left[\frac{2V^h(l) + V^h(l+he_1 - he_2) + V^h(l-he_1 + he_2)}{2h^2} \right] \\
 &\quad - \sigma_1^2 p_1 \left[\frac{V^h(l+he_1) + V^h(l-he_1) + V^h(l-he_2) + V^h(l+he_2)}{2h^2} \right]
 \end{aligned}$$

$$\begin{aligned}
 & +\sigma_2^2 p_2 \left[\frac{2V^h(l) + V^h(l + he_1 + he_3) + V^h(l - he_1 - he_3)}{2h^2} \right] \\
 & -\sigma_2^2 p_2 \left[\frac{V^h(l + he_1) + V^h(l - he_1) + V^h(l + he_3) + V^h(l - he_3)}{2h^2} \right] \\
 & + \frac{1}{2} (\sigma_1^2 p_1^2 + \sigma_2^2 p_2^2 + \sigma_3^2 z^2) \frac{V^h(l + he_1) + V^h(x - he_1) - V^h(l)}{h^2} \\
 & + \frac{1}{2} \sigma_1^2 \frac{V^h(l + he_3) + V^h(l - he_2) - 2V(l)}{h^2} \\
 & + \frac{1}{2} \sigma_2^2 \frac{V^h(l + he_3) + V^h(l - he_3) - 2V^h(l)}{h^2} \\
 & + \frac{1}{2} \epsilon_1^2 r^{(1)} \frac{V^h(l + he_4) + V^h(l - he_4) - 2V^h(l)}{h^2} \\
 & + \frac{1}{2} \epsilon_2^2 r^{(2)} \frac{V^h(l + he_5) + V^h(l - he_5) - 2V^h(l)}{h^2} \\
 & + K(x^{(1)}, x^{(2)}, r^{(1)}, r^{(2)}, \theta)
 \end{aligned} \tag{17}$$

简化(17), 并将结果与(14)进行比较, 我们得到了一步转移概率的具体表达式如下:

$$\begin{aligned}
 p^\theta(l, l \pm he_1) &= \frac{(\sigma_1^2 p_1^2 + \sigma_2^2 p_2^2 + \sigma_3^2 z^2) / 2 - (\sigma_1^2 p_1 + \sigma_2^2 p_2) + hf_1^\pm(l)}{Q_h} \\
 p^\theta(l, l \pm he_2) &= \frac{\sigma_1^2 / 2 - \sigma_1^2 p_1 + hf_2^\pm(r^{(1)})}{Q_h} \\
 p^\theta(l, l \pm he_3) &= \frac{\sigma_2^2 / 2 - \sigma_2^2 p_2 + hf_3^\pm(r^{(2)})}{Q_h} \\
 p^\theta(l, l \pm he_4) &= \frac{(\epsilon_1^2 r^{(1)}) / 2 + hf_4^\pm(r^{(1)})}{Q_h} \\
 p^\theta(l, l \pm he_5) &= \frac{(\epsilon_2^2 r^{(2)}) / 2 + hf_5^\pm(r^{(2)})}{Q_h} \\
 p^\theta(l, l + he_1 - he_2) &= p^\theta(l, l - he_1 + he_2) = -\frac{(\sigma_1^2 p_1) / 2}{Q_h} \\
 p^\theta(l, l + he_1 + he_3) &= p^\theta(l, l - he_1 - he_2) = -\frac{(\sigma_2^2 p_2) / 2}{Q_h} \\
 p^\theta(l, y) &= 0, \text{allother, } l, y \in G^h
 \end{aligned} \tag{18}$$

$$Q_h = (\sigma_1^2 p_1^2 + \sigma_2^2 p_2^2 + \sigma_3^2 z^2 + \sigma_1^2 + \sigma_2^2 + \epsilon_1^2 r^{(1)} + \epsilon_2^2 r^{(2)}) - (\sigma_1^2 p_1 + \sigma_2^2 p_2) + h \sum_{i=1}^4 |f_i(l)| + h^2$$

为了满足非负性, 我们需要做出一些假设

$$\begin{aligned}
 & (\sigma_2 p_2 \varrho - \sigma_1 p_1)^2 + \sigma_2^2 p_2^2 (1 - \varrho^2) + \sigma_3^2 Z^2 - |\sigma_1 \sigma_2 p_2 - \sigma_1^2 p_1| \geq 0 \\
 & \sigma_1^2 - |\sigma_1 \sigma_2 p_2 - \sigma_1^2 p_1| \geq 0
 \end{aligned} \tag{19}$$

基于之前的零售价格假设, 零售价格 p 有一个上界 \bar{p} 和一个下界 \underline{p} 。然后, 利用零售价格 p 的上界和下界进行简单的尺度计算, 我们可以得到满足公式(19)的条件如下

$$\begin{aligned} \sqrt{\frac{\sigma_1}{2\sigma_2(1-\varrho)}} \leq p_1 \leq \frac{\sigma_1}{\sigma_1 + \sigma_2 \rho} \\ \sqrt{\frac{\sigma_1}{2\sigma_2(1-\varrho)}} \leq p_2 \leq \frac{\sigma_1}{\sigma_1 + \sigma_2 \rho} \end{aligned} \quad (20)$$

此外, 动态规划方程组是对(14)的修正。对于 $l \in G^h$:

$$V^h(l) = \sup_{\theta_2 \in \Gamma_2} \left\{ \inf_{\theta_1 \in \Gamma_1} \sum_y e^{-\rho h} p^\theta(x, y | \theta) V^h(y) + K(l, \theta) \right\} \quad (21)$$

4. 收敛性分析

在本节中, 我们给出了连续时间马尔可夫链 $\{\xi_n^h, n=1, 2, \dots, n\}$ 的收敛性分析, 使马尔可夫链可行地逼近连续时间过程 $L(t)$, 包括我们定义的控制 $\theta^h(\cdot)$ 。

设 $\tau_n = t_n^h$ 是 $\xi^h(\cdot)$ 的第一个退出时间 G_n , \mathcal{D}_n^h 是 $\{\xi_k^h, \theta_k^h, k \leq n\}$ 的最小 σ -代数。当然, τ_n 是一个 \mathcal{D}_n^h 停止时间。在连续时间插值的情况下, 我们可以改写(10)为

$$\mathcal{J}^h(l, \theta^h) = E_l \left[\int_0^{\tau_n} e^{-\rho t} K(\xi^h(t), \theta^h(t)) dt \right] \quad (22)$$

定义 $\xi_0^h = l$ 并使用 E_n^h 表示基于时间 n 的信息的期望, 即基于 \mathcal{D}_n^h 的期望。结合局部一致性, 我们有

$$\begin{aligned} \xi_n^h &= l + \sum_{k=0}^{n-1} \Delta \xi_k^h \\ &= l + \sum_{k=0}^{n-1} E_k^h \Delta \xi_k^h + \sum_{k=0}^{n-1} (\Delta \xi_k^h - E_k^h \Delta \xi_k^h) \\ &= l + \sum_{k=0}^{n-1} f(\xi_k^h, \theta^h) \Delta t^h(\xi_k^h) + M_n^h \end{aligned} \quad (23)$$

定义

$$M_n^h = \sum_{k=0}^{n-1} (\Delta \xi_k^h - E_k^h \Delta \xi_k^h)$$

我们试图表示类似于(2.10)中的扩散项的 $M^h(t)$ 。定义 $W^h(\cdot)$ 为

$$W^h(t) = \sum_{k=0}^{n-1} \frac{\Delta \xi_k^h - E_k^h \Delta \xi_k^h}{\sigma(\xi_k^h)} = \int_0^t \sigma^{-1}(\xi^h(s)) dM^h(s) \quad (24)$$

结合(23)和(24), 我们将(23)重写为

$$\xi^h(t) = l + \int_0^t f(\xi^h(s), \theta^h) ds + \int_0^t \sigma(\xi^h(s)) dW^h(s) + \epsilon^h(t) \quad (25)$$

其中 $\epsilon^h(t)$ 是一个可以忽略的误差, 满足 $\limsup_{h \rightarrow \infty} E|\epsilon^h(t)| \rightarrow 0$ 对于任何 $0 < T < \infty$ 。

我们也可以将(24)改写为

$$L(t) = l + \int_0^t f(L(t), \theta) ds + \int_0^t \sigma(L(t)) dW(s) \quad (26)$$

设 \mathcal{F}_t^h 是一个过滤, 它表示度量的最小 σ -代数

$$\{\xi^h(s), m_s^h(\cdot), W^h(s), N^h(s), s \leq t\}$$

使用 Θ^h 表示关于 $W^h(\cdot)$ 的允许松弛控制的集合 $m^h(\cdot)$, 这样 $m_t^h(\cdot)$ 在给定 \mathcal{F}_t^h 的区间内是一个固定的概率测度。那么 $\Theta^h = \Theta_1^h \times \Theta_2^h$ 是一个包含 Γ^h 的大控制空间。用上面给出的松弛控制的符号, 我们可以将(25)写为

$$\xi^h(t) = l + \int_0^t \int_{\Theta_1^h \times \Theta_2^h} f(\xi^h(s), n) \times m_s^h(dn_1 \times dn_2) ds + \int_0^t \sigma(\xi^h(s)) dW^h(s) + \epsilon^h(t) \tag{27}$$

引理 4.1 假设让 $\{L_n(\cdot)\}$ 在 $C^m[c, d]$ 上有一个连续的路径, 并且和 $\{L_n(\cdot)\}$ 是紧的, 当且仅当 $\forall \eta > 0, \exists N_h < \infty$, 使得 $P\{L_n(c) \geq N_h\} \leq \eta$, 对于所有 n , 当且仅当存在 $k, a > 0, b > 0$, 使得 $E|L_n(t) - L_n(s)|^2 \leq k|t - s|^{1+b}$ 。

证明: 该证明与[17]类似, 在此省略。

定理 4.2 让近似链 $(\xi_n^h, n < \infty)$ 由转移概率定义, $n^h(\cdot)$ 是 $\{\theta_n^h, n < \infty\}$, ξ^h 是连续时间插值, $\{\tilde{\tau}_h\}$ 是 \mathcal{F}_t^h 停止时间的序列, 则 $\{\xi^h(\cdot), n^h(\cdot), W^h(\cdot), \tilde{\tau}_h\}$ 是紧的。

证明: 使用一点紧化的证明, $\tilde{\tau} \in [0, \infty)$ 。让 $T < \infty$, \tilde{v}_h 是一个 \mathcal{F}_t 停时。对于 $\delta > 0$, 我们有

$$E_{\tilde{v}_h}^{\theta^h} (W^h(\tilde{v}_h + \delta) - W^h(\tilde{v}_h))^2 = \delta + \tilde{\epsilon}_h$$

其中 $\tilde{\epsilon}_h \rightarrow 0$ 一致收敛于 \tilde{v}_h , 取 $\limsup \rightarrow 0$ 和 $\lim \delta \rightarrow 0$, 其紧密性为 $\{W^h(\cdot)\}$ 。一个类似的论点产生了 $M^h(\cdot)$ 的紧密性。序列 $\{m^h(\cdot)\}$ 是紧的, 因为它的范围空间是紧致的。这些结果以及 $c(\cdot)$ 和 $\theta(\cdot)$ 的有界性暗示了 $\{\xi^h(\cdot)\}$ 的紧密性, 则 $\{\xi^h(\cdot), n^h(\cdot), W^h(\cdot), \tilde{\tau}_h\}$ 是紧的。

定理 4.3 让 $\{\xi(\cdot), \theta^h(\cdot), W(\cdot), \tilde{\tau}\}$ 是一个弱收敛序列, \mathcal{F}_t 是由 $\{L(s), \theta(s), W(s), s \leq t\}$ 生成的 σ 代数, 那么 $W(\cdot)$ 分别是一个标准的 \mathcal{F}_t -维纳过程, $\tilde{\tau}$ 是一个 \mathcal{F}_t 停时, $\theta(\cdot)$ 是一个允许的控制, 让索赔时间用 v_n 表示, 则(27)是成立的。

证明: 该证明与[15]类似, 在此省略。

定理 4.4 $V^+(l), V^-(l), V^{h,+}(l), V^{h,-}(l)$ 在上一节中分别定义的值函数。然后我们有

$$\lim_{h \rightarrow 0} V^{h,+}(l) = V^+(l) \tag{28}$$

$$\lim_{h \rightarrow 0} V^{h,-}(l) = V^-(l) \tag{29}$$

证明: 该证明请参考[15], 在此省略。

引理 4.5 设 G_1 和 G_2 为紧致空间。假设一个连续函数 $K(l, y) \in G_1 \times G_2$ 满足凹凸条件, 即对于所有 $y \in G_2$, $f(\cdot, y)$ 是凸的, 而对于所有 $l \in G_1$, $K(l, \cdot)$ 是凹的。然后

$$\inf_{l \in G_1} \sup_{y \in G_2} K(l, y) = \sup_{y \in G_2} \inf_{l \in G_1} K(l, y) \tag{30}$$

定理 4.6 对于 $l \in S_h$, 一个在(18)中定义的马尔可夫链, 如果满足(20), 则存在一个鞍点

$$V^{h,+}(l) = V^{h,-}(l) \tag{31}$$

证明: 该证明请参考[15], 在此省略。

定理 4.7 假设满足定理 4.6 中的条件, 则鞍点存在为

$$V^+(l) = V^-(l) \tag{32}$$

证明: 因为对于在(27)中定义的近似马尔可夫链, 我们可以实现

$$\lim_{h \rightarrow 0} V^{h,-}(l) = V^-(l)$$

$$\lim_{h \rightarrow 0} V^{h,+}(l) = V^+(l) \tag{33}$$

通过在定理 4.6 中使用类似的方法。根据定理 4.7, 我们得到存在一个鞍点

$$V^+(l) = V^-(l) \tag{34}$$

5. 数值结果

在本节中, 我们使用给定的数值格式和证明的收敛结果对模型进行了数值实验。数值实验中涉及的参数值见[8] [15] [19]和[20]。

数值模拟示例

示例 5.1: 让我们假设总生产成本为 $C_1(u_1) = -5u_1 + 5u_1^2, C_2(u_2) = -3u_2 + 3u_2^2$ 。我们假设总库存持有成本和总积压成本为 $H_1(x_1) = -2x_1 + x_1^2, H_2(x_2) = -2x_2 + 0.5x_2^2$ 。我们设置了参考价格 $\beta = 2$ 的调整速度, 参考价格的波动性参数 $\epsilon_1 = 0.1, \epsilon_2 = 0.2$ 。平均市场潜力为 $a = 7$, 价格敏感性为 $b = 0.3$, 资产的回报率为 $\mu = 0.4$, 参考价格效应的大小为 $\gamma = 0.5$ 。同时, 我们假设波动率参数 $\sigma_1 = 0.25, \sigma_2 = 0.45, \sigma_3 = 0.75$ 。最后, 我们假设贴现因子 $\rho = 0.1$ 。

为了简单起见, 我们考虑离散事件有两个状态的情况。也就是说, 连续时间马尔可夫链有两个给定索赔额分布的状态。此外, 我们假设索赔额分布是相同的, 在每个政权。利用数值迭代方法, 数值求解了最优控制问题。我们将支付函数作为两个参与者之间的盈余差在到达下障碍之前到达上障碍的概率。因此, 在博弈中, 公司 1 想要最大化概率, 而公司 2 想要最小化相同的概率。

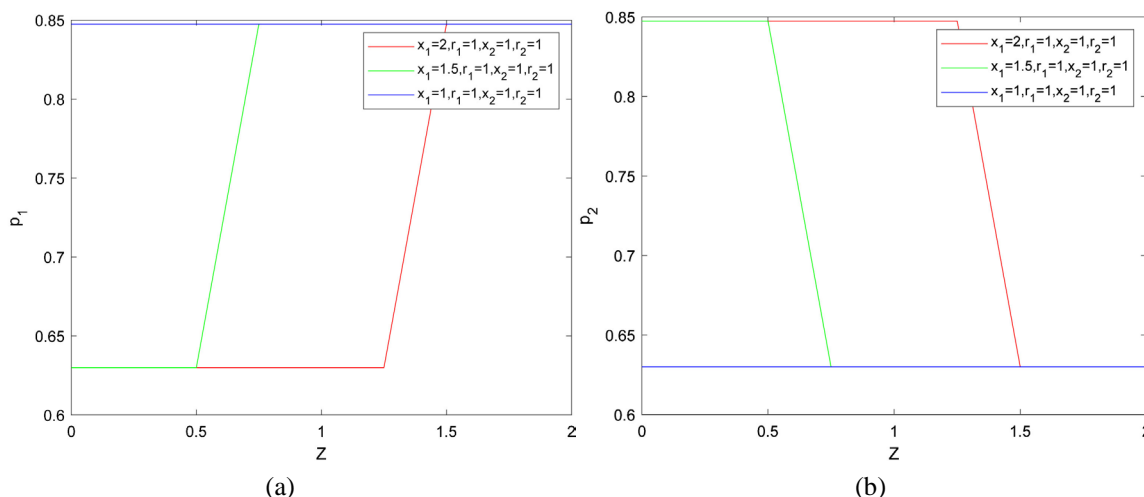


Figure 1. Optimal retail price under some fixed parameters

图 1. 某些固定参数下的最优零售价格

从图 1(a)-(b), 我们知道当其他参数保持不变, 当公司 1 的库存水平与公司 2 一样, 两家公司的定价策略是公司 1 设置价格高于公司 2, 定价差异是最大之间的三个曲线。可能, 公司 1 和公司 2 有相同的库存水平, 也就是说, 相同的库存过剩压力。这类似于我们生活中大品牌公司和小品牌公司之间的商业竞争。同时, 在公司 1 的库存水平大于公司 2 的情况下, 当公司 2 与公司 1 的利润差迅速达到下限时, 公司 1 的定价低于公司 2。这是因为当两家公司的竞争相等时, 公司 1 的库存压力很大, 只能以利润小、营业额快的形式保持竞争力; 当公司 2 与公司 1 具有相同的库存水平时, 无论达到什么竞争水平, 公司 2 的价格总是低于公司 1, 以增加赢得竞争的可能性。

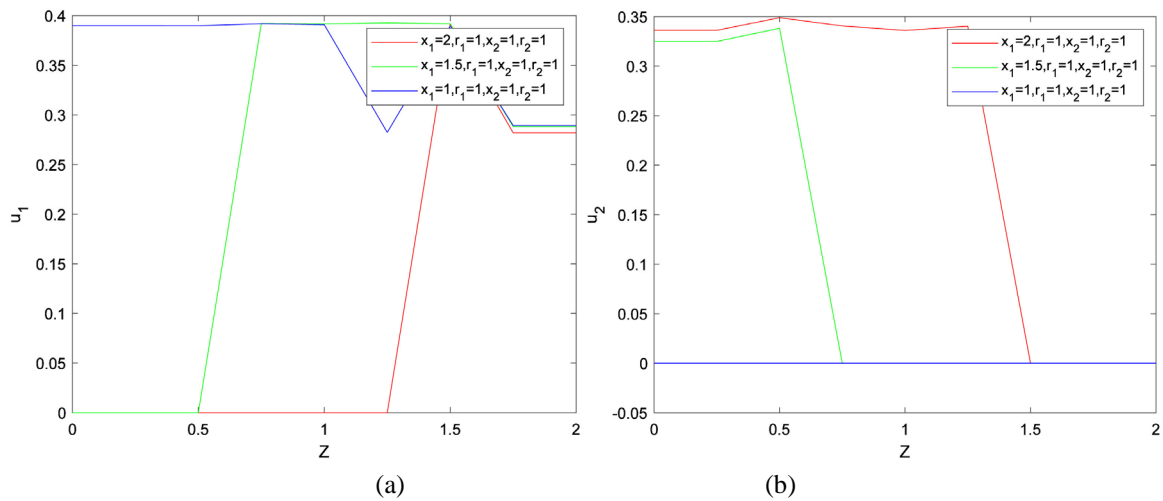


Figure 2. Optimal production rate under some fixed parameters
图 2. 某些固定参数下的最优生产速率

从图 2(a)-(b)可以看出, 当公司 1 的库存水平超过公司 2, 两家公司之间的利润差迅速接近下界时, 公司 1 的生产率下降。这是由于在平等竞争的情况下, 公司 1 的高库存水平导致了货物的积累和成本的增加。因此, 公司 1 采取了降低生产速度、减少库存的战略, 通过降低成本来提高盈利能力。当公司 1 的库存水平超过公司 2, 两家公司的利润差迅速接近上限时, 最优策略表明公司 1 的生产力提高, 公司 2 的生产力下降, 公司 1 的价格上升, 公司 2 的价格下降。这意味着当公司 1 在高库存水平的条件下能够获得巨大的利润时, 最优策略意味着利润小, 快速周转对公司 2 不再可行, 最好的利润策略是减少生产和定价, 通过降低成本提高竞争胜率。在现实的环境中, 一个类似于公司 2 的策略往往是失败的, 这就足以基本上结束公司 1 的胜利。

数值例子提供了动态策略和库存水平分析, 为决策者在减少错误决策造成的重大损失的风险方面提供了实用的指导。虽然在扩散模型中包含跳跃过程在捕捉市场转变方面更为现实, 但它也给开发有效的数值算法带来了更大的复杂性和计算挑战。此外, 在一个随机博弈框架内, 公司可以根据对手的行动, 灵活地调整其最优策略, 以最大化他们成功的机会。

6. 结论

本文研究的是基于随机参考价格的库存管理问题, 在随机博弈的框架下, 每个公司都可以根据自己的对手调整最优再保险策略。根据问题的背景条件给出支付函数和马尔可夫链逼近方法来解决随机博弈的最优控制问题。这两个保险公司之间的竞争看作是随机微分博弈。两个公司给出两公司的支付函数, 各个公司都设法找到一个最佳策略, 其中一个公司以最大化支付函数, 而另一个公司试图最小化相同的支付函数。利用动态规划原理, 我们证明了该背景的博弈模型的上限值和下限值满足一个耦合的非线性积分微分 HJI 方程组。而因为 HJI 方程组非常难以求得封闭解, 所以我们使用一个可行的数值方法去求解, 即使用马尔可夫链近似方法来获得数值解。同时我们还证明了该对策问题鞍点的存在性, 证明了逼近序列收敛于跳扩散过程, 逼近序列分别收敛于对策上值和下值。最后的数值例子提供了动态策略和库存水平分析, 为决策者在减少错误决策造成的重大损失的风险方面提供了实用的指导。

另外, 本文存在以下不足, 只考虑了两家公司间的竞争关系, 不适用于多公司博弈的情况。同时没有考虑跳跃的情形。若扩散模型中包含跳跃过程则在捕捉市场转变方面更为现实, 但它也给开发有效的数值算法带来了更大的复杂性和计算挑战。

参考文献

- [1] Whittin, T.M. (1955) Inventory Control and Price Theory. *Management Science*, **2**, 61-68. <https://doi.org/10.1287/mnsc.2.1.61>
- [2] Bertsimas, D. and De Boer, S. (2005) Special Issue Papers: Dynamic Pricing and Inventory Control for Multiple Products. *Journal of Revenue and Pricing Management*, **3**, 303-319. <https://doi.org/10.1057/palgrave.rpm.5170117>
- [3] Chen, L., Chen, Y. and Pang, Z. (2010) Dynamic Pricing and Inventory Control in a Make-to-Stock Queue with Information on the Production Status. *IEEE Transactions on Automation Science and Engineering*, **8**, 361-373. <https://doi.org/10.1109/TASE.2010.2071865>
- [4] Greenleaf, E.A. (1995) The Impact of Reference Price Effects on the Profitability of Price Promotions. *Marketing Science*, **14**, 82-104. <https://doi.org/10.1287/mksc.14.1.82>
- [5] Hu, Z., Chen, X. and Hu, P. (2016) Technical Note—Dynamic Pricing with Gain-Seeking Reference Price Effects. *Operations Research*, **64**, 150-157. <https://doi.org/10.1287/opre.2015.1445>
- [6] Chen, X., Hu, P., Shum, S., et al. (2016) Dynamic Stochastic Inventory Management with Reference Price Effects. *Operations Research*, **64**, 1529-1536. <https://doi.org/10.1287/opre.2016.1524>
- [7] Chen, X., Hu, Z.Y. and Zhang, Y.H. (2019) Dynamic Pricing with Stochastic Reference Price Effect. *Journal of the Operations Research Society of China*, **7**, 107-125. <https://doi.org/10.1007/s40305-018-0227-1>
- [8] Cao, Y. and Duan, Y. (2020) Joint Production and Pricing Inventory System under Stochastic Reference Price Effect. *Computers & Industrial Engineering*, **143**, Article ID: 106411. <https://doi.org/10.1016/j.cie.2020.106411>
- [9] Pang, J.S. and Fukushima, M. (2005) Quasi-Variational Inequalities, Generalized Nash Equilibria, and Multi-Leader-Follower Games. *Computational Management Science*, **2**, 21-56. <https://doi.org/10.1007/s10287-004-0010-0>
- [10] Elliott, R.J. and Siu, T.K. (2010) On Risk Minimizing Portfolios under a Markovian Regime-Switching Black-Scholes Economy. *Annals of Operations Research*, **176**, 271-291. <https://doi.org/10.1007/s10479-008-0448-5>
- [11] Adida, E. and Perakis, G. (2010) Dynamic Pricing and Inventory Control: Uncertainty and Competition. *Operations Research*, **58**, 289-302. <https://doi.org/10.1287/opre.1090.0718>
- [12] Berkovitz, L.D. and Drescher, M. (1959) A Game-Theory Analysis of Tactical Air War. *Operations Research*, **7**, 599-620. <https://doi.org/10.1287/opre.7.5.599>
- [13] Fleming, W.H. (1961) The Convergence Problem for Differential Games. *Journal of Mathematical Analysis and Applications*, **3**, 102-116. [https://doi.org/10.1016/0022-247X\(61\)90009-9](https://doi.org/10.1016/0022-247X(61)90009-9)
- [14] Elliott, R.J. and Kalton, N.J. (1972) The Existence of Value in Differential Games of Pursuit and Evasion. *Journal of Differential Equations*, **12**, 504-523. [https://doi.org/10.1016/0022-0396\(72\)90022-8](https://doi.org/10.1016/0022-0396(72)90022-8)
- [15] Jin, Z., Yin, G. and Wu, F. (2013) Optimal Reinsurance Strategies in Regime-Switching Jump Diffusion Models: Stochastic Differential Game Formulation and Numerical Methods. *Insurance: Mathematics and Economics*, **53**, 733-746. <https://doi.org/10.1016/j.insmatheco.2013.09.015>
- [16] Pan, Z. and Başar, T. (1995) H_∞ -Control of Markovian Jump Systems and Solutions to Associated Piecewise-Deterministic Differential Games. In: Jan Olsder, G., Ed., *New Trends in Dynamic Games and Applications*, Birkhäuser, Boston, 61-94. https://doi.org/10.1007/978-1-4612-4274-1_4
- [17] Yin, G., Zhang, Q. and Badowski, G. (2003) Discrete-Time Singularly Perturbed Markov Chains: Aggregation, Occupation Measures, and Switching Diffusion Limit. *Advances in Applied Probability*, **35**, 449-476. <https://doi.org/10.1239/aap/1051201656>
- [18] Kushner, H.J. (2002) Numerical Approximations for Stochastic Differential Games. *SIAM Journal on Control and Optimization*, **41**, 457-486. <https://doi.org/10.1137/S0363012901389457>
- [19] Herbon, A. and Kogan, K. (2014) Time-Dependent and Independent Control Rules for Coordinated Production and Pricing under Demand Uncertainty and Finite Planning Horizons. *Annals of Operations Research*, **223**, 195-216. <https://doi.org/10.1007/s10479-014-1616-4>
- [20] Maimon, O., Khmelnitsky, E. and Kogan, K. (1998) Optimal Flow Control in Manufacturing Systems: Production Planning and Scheduling. Springer Science & Business Media, Berlin. <https://doi.org/10.1007/978-1-4757-2834-7>