

离散时间随机 LQ Stackelberg 博弈的线性最优闭环策略

李洪丹, 张焕水

(山东科技大学 电气与自动化工程学院, 山东 青岛 266590)

摘要: 离散时间随机线性二次(LQ)Stackelberg 博弈的线性最优闭环策略的复杂性在于经典的最优控制方法已不再适用, 即便是对于 LQ Stackelberg 博弈, 其解也可能呈现非线性特性。在这种情况下, 假设领导者和跟随者的闭环策略是历史和当前新息的线性组合, 这一假设符合闭环信息模式的定义, 因此这是合理的。通过运用矩阵极大值原理, 得到了最优的线性闭环 Stackelberg 策略。本研究首次揭示最优的线性闭环解采用了 Kalman 滤波反馈的形式, 对于未来研究 Stackelberg 策略中的信息非对称控制具有重要意义。

关键词: 线性最优闭环策略; 随机 LQ Stackelberg 博弈; 新息组合; Kalman 滤波反馈; 矩阵极大值原理

中图分类号: TP273.1

文献标志码: A

Linear optimal closed-loop strategies for discrete-time stochastic LQ Stackelberg games

LI Hongdan, ZHANG Huanshui

(College of Electrical and Automation Engineering, Shandong University of Science and Technology, Qingdao 266590, China)

Abstract: The complexity of linear optimal closed-loop strategies for discrete-time stochastic linear quadratic (LQ) Stackelberg games lies in the fact that traditional optimal control methods are no longer applicable because even in the case of LQ Stackelberg games, the solution may exhibit nonlinear characteristics. In such cases, the assumption that the closed-loop strategies for both the leader and the follower are linear combinations of historical and current innovations aligns with the definition of a closed-loop information pattern, making it a reasonable and widely applicable choice. By applying the matrix maximum principle, the optimal linear closed-loop Stackelberg strategies are obtained. This study, for the first time, reveals that the optimal linear closed-loop solution takes the form of Kalman filtering feedback, which is a finding of significant importance for future research in handling information asymmetry in Stackelberg strategies.

Key words: linear closed-loop strategies; stochastic LQ Stackelberg games; combinations of innovations; Kalman filtering feedback; matrix maximum principle

Stackelberg 解的概念最初是由 Stackelberg^[1]提出, 由 Chen 等^[2]引入到控制和博弈领域。Stackelberg 博弈的特点是信息的非同步, 即其中一方(领导者)可以宣布并强制执行其策略给另一方(跟随者), 这是与其他博弈的主要区别, 由此导致 Stackelberg 博弈的研究相对复杂。

在 Stackelberg 博弈中, 考虑玩家的信息集至关重要。与文献[3]中的定义类似, 信息模式可以分为开环、反馈、闭环三种。开环 Stackelberg 博弈意味着领导者的可用信息仅限于初始值, 即 x_0 。由于信息结构

收稿日期: 2023-10-21

基金项目: 国家自然科学基金原创探索计划项目(62450004); 山东省自然科学基金重大基础研究项目(ZR2021ZD14); 山东省自然科学基金项目(ZR2025MS982); 山东省高等学校青年创新团队发展计划项目(2022KJ222); 青岛西海岸新区高层次人才团队项目(RCTDJC-2019-05)

作者简介: 李洪丹(1990—), 女, 山东济宁人, 副教授, 博士, 主要从事随机最优控制的研究。

张焕水(1963—), 男, 山东莱芜人, 教授, 博士, 主要从事随机系统最优滤波与控制、时滞系统随机控制的研究, 本文通信作者, E-mail: hszhang@sdu.edu.cn

相对简单,因此早期已经有了很多关于开环 Stackelberg 策略的研究,例如开环 Stackelberg 博弈的差分 Riccati 方程的渐近行为^[4]、存在性和唯一性的充要条件^[5]。需要注意的是,尤其在动态博弈中,仅仅了解初始值是不够的。Cruz 等^[6-7]提出了反馈 Stackelberg 策略的概念,其中每个玩家的可用信息集在每个阶段都会发生变化,即当前信息,这是一种满足贝尔曼最优性原理的自适应策略。因此,一些经典的最优方法可以用来解决反馈 Stackelberg 博弈。例如,Junger^[8]通过块矩阵方法研究了一种带有反馈信息结构的离散 Stackelberg 策略,Li 等^[9]在考虑跟随者稳定性情况下获得了领导者的反馈激励策略。

事实上,在大多数文献中,反馈通常被视为一种特殊的闭环结构,即无记忆闭环策略。对于有记忆的闭环信息结构,如文献[3]所定义的,即使对于线性二次(linear quadratic, LQ)博弈,也很少有结果。其原因正如 Basar^[10-11]所述,经典的最优控制理论,特别是动态规划,无法得到闭环 Stackelberg 策略应满足的合理的必要条件。这是一个非经典的最优控制问题,因为领导者的一阶偏导数存在于跟随者优化问题的必要条件中,使得传统的最优性方法失效。

为了避免这一困难,文献[11-14]进行了一些尝试。总的来说,到目前为止有两种方法。一种方法是将领导者的最优问题转化为一个团队最优问题。例如,假设领导者是线性且具有一步记忆的,而跟随者则以线性反馈形式实现,Basar 等^[11-12]得到了离散时间博弈和微分博弈的闭环 Stackelberg 策略,但这种方法缺点是得到的领导者策略对跟随者可能的非最优控制选择敏感,由此导致非最优全局解,并会降低领导者策略的可信度。另一种方法是,Medanic^[13]假设领导者策略具有特定的参数结构,在这种情况下,跟随者的响应可以清晰确定,领导者面临的问题可根据跟随者响应集施加的约束来优化这些参数。

需要注意的是,上述提到的方法都是间接的,已有文献也没有直接解决闭环 Stackelberg 博弈的方法。正如 Tolwinski^[15]指出,即使对于 LQ 博弈,其最优闭环 Stackelberg 策略也是非线性的。从这个角度看,要得出一个解析明确的 Stackelberg 闭环解相对复杂。因此,一些学者研究了线性形式的最优闭环 Stackelberg 策略。与已有文献不同的是,每个玩家根据状态的观测信息作决策,即不完全信息。这种问题最早由 Castanon 等^[16]提出,并得到了嵌套信息模式下,两名玩家线性二次高斯(linear quadratic Gaussian, LQG)非零和离散时间动态博弈问题的反馈 Stackelberg 解。在非嵌套信息模式下,Basar^[17]研究了一类具有非高斯随机变量的静态 Stackelberg 博弈问题,Li 等^[18]得到了确定情形下一步记忆的闭环 Stackelberg 策略。

基于上述分析,先前研究中对闭环 Stackelberg 策略的关注主要集中在无记忆闭环方面,即仅依赖当前时刻信息并采用反馈策略。本研究专注于有记忆闭环策略,即考虑基于全部历史信息 and 当前时刻信息的策略设计,这为问题的解决带来了显著挑战。此外,现有文献大多侧重于状态信息完全已知的情形,这明显与实际不符。因此,本研究聚焦在不完全状态信息的情境下,即基于测量信息的闭环策略,这也为本研究增添了相当的难度。为解决该问题,考虑到两个事实:一是在观测变量下的条件数学期望等同于在新息序列下的条件数学期望,二是新息序列具有正交性,进一步假设跟随者和领导者的策略都是新息的线性组合。从上述分析来看,该假设在逻辑上是合理的。在此背景下,本研究目标是找到相应的参数矩阵,以最小化两个成本函数。为了获得线性闭环控制器,运用矩阵极大值原理来获取控制器的增益矩阵。值得注意的是,本研究首次证明了带有不完全信息的线性闭环 Stackelberg 控制实际上等同于卡尔曼滤波反馈,这为未来研究闭环 Stackelberg 策略提供了重要的理论基础和思路。

1 问题描述

考虑下列随机线性系统:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k^f + \bar{\mathbf{B}}\mathbf{u}_k^l + \mathbf{w}_k, \quad (1)$$

$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k. \quad (2)$$

式中: $\mathbf{x}_k \in \mathbf{R}^n$, $\mathbf{u}_k^f \in \mathbf{R}^{m_1}$, $\mathbf{u}_k^l \in \mathbf{R}^{m_2}$ 分别表示状态、跟随者及领导者; $\mathbf{A}$ 、 $\mathbf{B}$ 、 $\bar{\mathbf{B}}$ 、 $\mathbf{H}$ 分别为具有适合维数的矩阵; $\mathbf{w}_k$ 、 $\mathbf{v}_k$ 表示均值为零,方差分别为 $\mathbf{R}_w$ 、 $\mathbf{R}_v$ 的独立高斯白噪声;初始状态 $\mathbf{x}_0$ 为高斯随机变量, $\mathbf{E}[\mathbf{x}_0] = \bar{\mathbf{x}}_0$, $\mathbf{E}[(\mathbf{x}_0 - \bar{\mathbf{x}}_0)(\mathbf{x}_0 - \bar{\mathbf{x}}_0)^T] = \mathbf{R}_0$, 并且与 $\mathbf{w}_k$ 、 $\mathbf{v}_k$ 相互独立; $\mathbf{y}_k$ 表示观测,观测信息集为 $\mathbf{Y}_k = \{\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_k\}$, $\mathbf{y}_k$

的估计值表示为 $\hat{\mathbf{y}}_{k|k-1} = \mathbf{E}[\mathbf{y}_k | \mathbf{Y}_{k-1}]$ 。

为方便讨论,引入新息序列 $\{\boldsymbol{\varepsilon}_k\}_{k \geq 0}$, $\boldsymbol{\varepsilon}_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}$, 并且定义一个新的信息集 $\mathbf{S}_k = \{\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_k\}$ 。新息序列的性质^[18]有 $\hat{\mathbf{y}}_{k|k-1} = \mathbf{E}[\mathbf{y}_k | \mathbf{Y}_{k-1}] = \mathbf{E}[\mathbf{y}_k | \mathbf{S}_{k-1}]$ 以及正交性

$$\mathbf{E}[\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_j^T] = \begin{cases} 0, & i \neq j; \\ \mathbf{H} \boldsymbol{\Sigma}_{i|i-1} \mathbf{H}^T + \mathbf{R}_v, & i = j. \end{cases} \quad (3)$$

式中: $\boldsymbol{\Sigma}_{i|i-1} = \mathbf{E}[(\mathbf{x}_i - \hat{\mathbf{x}}_{i|i-1})(\mathbf{x}_i - \hat{\mathbf{x}}_{i|i-1})^T]$, $\hat{\mathbf{x}}_{0|-1} = \bar{\mathbf{x}}_0$ 为误差协方差矩阵, 且满足下列迭代关系:

$$\boldsymbol{\Sigma}_{i|i-1} = \mathbf{A} \boldsymbol{\Sigma}_{i-1} \mathbf{A}^T + \mathbf{R}_w, \quad (4)$$

$$\boldsymbol{\Sigma}_i = [\mathbf{I} - \boldsymbol{\Sigma}_{i|i-1} \mathbf{H}^T (\mathbf{H} \boldsymbol{\Sigma}_{i|i-1} \mathbf{H}^T + \mathbf{R}_v)^{-1} \mathbf{H}] \boldsymbol{\Sigma}_{i|i-1}, \quad (5)$$

$$\boldsymbol{\Sigma}_{0|-1} = \mathbf{R}_0. \quad (6)$$

定义下列性能指标:

$$J^f = \mathbf{E} \left\{ \sum_{k=0}^N [\mathbf{x}_k^T \mathbf{Q} \mathbf{x}_k + (\mathbf{u}_k^f)^T \mathbf{R}_1 \mathbf{u}_k^f + (\mathbf{u}_k^l)^T \mathbf{R}_2 \mathbf{u}_k^l] + \mathbf{x}_{N+1}^T \mathbf{P}_{N+1} \mathbf{x}_{N+1} \right\}, \quad (7)$$

$$J^l = \mathbf{E} \left\{ \sum_{k=0}^N [\mathbf{x}_k^T \bar{\mathbf{Q}} \mathbf{x}_k + (\mathbf{u}_k^f)^T \bar{\mathbf{R}}_1 \mathbf{u}_k^f + (\mathbf{u}_k^l)^T \bar{\mathbf{R}}_2 \mathbf{u}_k^l] + \mathbf{x}_{N+1}^T \bar{\mathbf{P}}_{N+1} \mathbf{x}_{N+1} \right\}. \quad (8)$$

式中: 状态加权矩阵 $\mathbf{Q}$ 、 $\bar{\mathbf{Q}}$ 、 $\mathbf{P}_{N+1}$ 以及 $\bar{\mathbf{P}}_{N+1}$ 是半正定矩阵; 控制加权矩阵 $\mathbf{R}_i$ 、 $\bar{\mathbf{R}}_i$ ($i=1, 2$) 是正定矩阵。

根据信息结构^[12]的定义: ①若可用信息是 $\{\boldsymbol{\varepsilon}_0\}$, 称为开环不完全信息结构; ②若可用信息是 $\{\boldsymbol{\varepsilon}_k\}$, 称为反馈不完全信息结构; ③若可用信息是 $\{\boldsymbol{\varepsilon}_0, \boldsymbol{\varepsilon}_1, \dots, \boldsymbol{\varepsilon}_k\}$, 称为闭环不完全信息结构。

因为经典的最优控制方法失效, 即使对于 LQ 情形, 求解闭环 Stackelberg 策略仍然很困难。到目前为止, 还没有直接的方法来解决闭环 Stackelberg 策略。然而, 根据闭环不完全信息结构的定义, 线性闭环 Stackelberg 策略可以给出如下表示:

$$\mathbf{u}_k^f = \sum_{i=0}^k \mathbf{K}_k^i \boldsymbol{\varepsilon}_i + \mathbf{G}_k \hat{\mathbf{x}}_{0|-1}, \quad (9)$$

$$\mathbf{u}_k^l = \sum_{i=0}^k \bar{\mathbf{K}}_k^i \boldsymbol{\varepsilon}_i + \bar{\mathbf{G}}_k \hat{\mathbf{x}}_{0|-1}. \quad (10)$$

式中: $\mathbf{K}_k^i$ 、 $\mathbf{G}_k$ 、 $\bar{\mathbf{K}}_k^i$ 、 $\bar{\mathbf{G}}_k$ 为具有合适维数的待确定矩阵。

根据上述准备工作, 提出本研究要讨论的主要问题。

问题 1 寻找线性闭环控制器 $\mathbf{u}_k^f$ 和 $\mathbf{u}_k^l$, 如式(9)和式(10)中所定义, 以最小化性能指标(式(7)和式(8))。

显然, 这与团队博弈存在显著差异, 因为团队博弈通常要求所有参与者共同优化某一性能指标。此外, 在纳什博弈中, 每个参与者可以独立优化其自身的性能指标, 而无需考虑其他方的策略。对于 Stackelberg 博弈, 玩家在不平等的基础上被对待, 这在数学上带来了很大的困难。

2 线性最优闭环 Stackelberg 解

定义下列 Riccati 型差分方程:

$$\mathbf{P}_k = \mathbf{Q} + \mathbf{A}^T \mathbf{P}_{k+1} \mathbf{A} - \mathbf{M}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k, \quad (11)$$

$$\bar{\mathbf{P}}_k = \bar{\mathbf{Q}} + \mathbf{A}^T \bar{\mathbf{P}}_{k+1} \mathbf{A} + (\bar{\boldsymbol{\Gamma}}_k \mathbf{M}_k - \bar{\mathbf{M}}_k)^T \bar{\boldsymbol{\Gamma}}_k^{-1} (\bar{\boldsymbol{\Gamma}}_k \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k - \bar{\mathbf{M}}_k) - \bar{\mathbf{M}}_k^T \bar{\boldsymbol{\gamma}}_k^{-1} \bar{\mathbf{M}}_k - \bar{\mathbf{M}}_k^T \bar{\boldsymbol{\Gamma}}_k^{-1} \bar{\mathbf{M}}_k. \quad (12)$$

式中:

$$\mathbf{M}_k = \mathbf{B}^T \mathbf{P}_{k+1} \mathbf{A}, \quad (13)$$

$$\boldsymbol{\Gamma}_k = \mathbf{R}_1 + \mathbf{B}^T \mathbf{P}_{k+1} \mathbf{B}, \quad (14)$$

$$\bar{\boldsymbol{\Gamma}}_k = \bar{\mathbf{R}}_1 + \mathbf{B}^T \bar{\mathbf{P}}_{k+1} \mathbf{B}, \quad (15)$$

$$\bar{\mathbf{M}}_k = \mathbf{B}^T \bar{\mathbf{P}}_{k+1} \mathbf{A}, \quad (16)$$

$$\bar{\mathbf{M}}_k = (\mathbf{M}_k^B)^T \boldsymbol{\Gamma}_k^{-1} \bar{\mathbf{R}}_1 \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k + \mathbf{B}_k^T \bar{\mathbf{P}}_{k+1} \mathbf{A}_k, \quad (17)$$

$$\bar{\boldsymbol{\gamma}}_k = \bar{\mathbf{R}}_2 + (\mathbf{M}_k^B)^T \boldsymbol{\Gamma}_k^{-1} \bar{\mathbf{R}}_1 \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k^B + \mathbf{B}_k^T \bar{\mathbf{P}}_{k+1} \mathbf{B}_k, \quad (18)$$

以及

$$\mathbf{M}_k^B = \mathbf{B}^T \mathbf{P}_{k+1} \bar{\mathbf{B}}, \quad (19)$$

$$\mathbf{B}_k = \bar{\mathbf{B}} - \mathbf{B} \Gamma_k^{-1} \mathbf{M}_k^B, \quad (20)$$

$$\mathbf{A}_k = \mathbf{A} - \mathbf{B} \Gamma_k^{-1} \mathbf{M}_k. \quad (21)$$

为了确保上述迭代过程能够继续进行,提出以下假设。

假设 1 $\mathbf{\Gamma}_k$ 、 $\bar{\mathbf{\Gamma}}_k$ 以及 $\bar{\mathbf{Y}}_k$ 是可逆矩阵。

假设 2 下列正则条件成立,即

$$(\mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T] \text{Pinv} \langle \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T] \rangle - \mathbf{I}) \hat{\mathbf{x}}_{0|-1} = 0. \quad (22)$$

现在给出本研究的主要结果。

定理 1 如果假设 1 和假设 2 成立,使得性能指标(式(7)和式(8))最小的增益矩阵可以表示为:

$$\bar{\mathbf{K}}_k^i = -\bar{\gamma}_k^{-1} \bar{\mathbf{M}}_k \mathbf{E}[\mathbf{x}_k \boldsymbol{\varepsilon}_i^T] (\mathbf{H} \boldsymbol{\Sigma}_{i|i-1} \mathbf{H}^T + \mathbf{R}_v)^{-1}, \quad (23)$$

$$\bar{\mathbf{G}}_k = -\bar{\gamma}_k^{-1} \bar{\mathbf{M}}_k \mathbf{E}[\mathbf{x}_k \hat{\mathbf{x}}_{0|-1}^T] \text{Pinv} \langle \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T] \rangle, \quad (24)$$

$$\mathbf{K}_k^i = -\mathbf{\Gamma}_k^{-1} (\mathbf{M}_k \mathbf{E}[\mathbf{x}_k \boldsymbol{\varepsilon}_i^T] \mathbf{E}[\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^T]^{-1} + \mathbf{M}_k^B \bar{\mathbf{K}}_k^i), \quad (25)$$

$$\mathbf{G}_k = -\mathbf{\Gamma}_k^{-1} (\mathbf{M}_k \mathbf{E}[\mathbf{x}_k \hat{\mathbf{x}}_{0|-1}^T] + \mathbf{M}_k^B \bar{\mathbf{G}}_k \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T]) \text{Pinv} \langle \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T] \rangle. \quad (26)$$

而且,问题 1 的解以 Kalman 滤波反馈的形式给出:

$$\mathbf{u}_k^f = -\mathbf{\Gamma}_k^{-1} \mathbf{M}_k \hat{\mathbf{x}}_k, \quad (27)$$

$$\mathbf{u}_k^l = -\bar{\gamma}_k^{-1} \bar{\mathbf{M}}_k \hat{\mathbf{x}}_k. \quad (28)$$

式中:

$$\mathbf{M}_k = \mathbf{M}_k - \mathbf{M}_k^B \bar{\gamma}_k^{-1} \bar{\mathbf{M}}_k, \quad (29)$$

$$\hat{\mathbf{x}}_k = \mathbf{A} \hat{\mathbf{x}}_{k-1} + \mathbf{B} \mathbf{u}_{k-1}^f + \bar{\mathbf{B}} \mathbf{u}_{k-1}^l + \boldsymbol{\Sigma}_{k|k-1} \mathbf{H}^T (\mathbf{H} \boldsymbol{\Sigma}_{k|k-1} \mathbf{H}^T + \mathbf{R}_v)^{-1} \boldsymbol{\varepsilon}_k. \quad (30)$$

证明:证明概括为以下两步。

第一步,优化跟随者性能指标 $J^f(\cdot)$ 。

令 $\mathbf{X}_k = \mathbf{E}[\mathbf{x}_k \mathbf{x}_k^T]$ 。根据式(1)、式(2)可得:

$$\begin{aligned} \mathbf{X}_{k+1} = & \mathbf{A} \mathbf{X}_k \mathbf{A}^T + \sum_{i=0}^k [2 \mathbf{A} \mathbf{E}[\mathbf{x}_k \boldsymbol{\varepsilon}_i^T] (\mathbf{B} \mathbf{K}_k^i + \bar{\mathbf{B}} \bar{\mathbf{K}}_k^i)^T + (\mathbf{B} \mathbf{K}_k^i + \bar{\mathbf{B}} \bar{\mathbf{K}}_k^i) (\mathbf{H} \boldsymbol{\Sigma}_{i|i-1} \mathbf{H}^T + \mathbf{R}_v) (\mathbf{B} \mathbf{K}_k^i + \bar{\mathbf{B}} \bar{\mathbf{K}}_k^i)^T] + \\ & \mathbf{R}_w + \mathbf{A} \mathbf{E}[\mathbf{x}_k \hat{\mathbf{x}}_{0|-1}^T] \mathbf{G}_k^T \mathbf{B}^T + \mathbf{B} \mathbf{G}_k \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \mathbf{x}_k^T] \mathbf{A}^T + \mathbf{B} \mathbf{G}_k \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T] \mathbf{G}_k^T \mathbf{B}^T + \mathbf{A} \mathbf{E}[\mathbf{x}_k \hat{\mathbf{x}}_{0|-1}^T] \bar{\mathbf{G}}_k^T \bar{\mathbf{B}}^T + \\ & \bar{\mathbf{B}} \bar{\mathbf{G}}_k \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \mathbf{x}_k^T] \mathbf{A}^T + \bar{\mathbf{B}} \bar{\mathbf{G}}_k \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T] \bar{\mathbf{G}}_k^T \bar{\mathbf{B}}^T + \mathbf{B} \mathbf{G}_k \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T] \bar{\mathbf{G}}_k^T \bar{\mathbf{B}}^T + \bar{\mathbf{B}} \bar{\mathbf{G}}_k \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T] \bar{\mathbf{G}}_k^T \bar{\mathbf{B}}^T. \end{aligned} \quad (31)$$

因此,跟随者的 Hamiltonian 函数可以表示为:

$$\begin{aligned} H_k^f = & \text{Tr}(\mathbf{Q} \mathbf{X}_k + (\mathbf{G}_k^T \mathbf{R}_1 \mathbf{G}_k + \bar{\mathbf{G}}_k^T \mathbf{R}_2 \bar{\mathbf{G}}_k) \hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T + \mathbf{P}_{k+1} \mathbf{A} \mathbf{X}_k \mathbf{A}^T + \mathbf{P}_{k+1} \mathbf{R}_w + \\ & \sum_{i=0}^k [(\mathbf{K}_k^i)^T \mathbf{R}_1 \mathbf{K}_k^i + (\bar{\mathbf{K}}_k^i)^T \mathbf{R}_2 \bar{\mathbf{K}}_k^i] (\mathbf{H} \boldsymbol{\Sigma}_{i|i-1} \mathbf{H}^T + \mathbf{R}_v) + 2 \mathbf{P}_{k+1} \mathbf{A} \mathbf{E}[\mathbf{x}_k \boldsymbol{\varepsilon}_i^T] (\mathbf{B} \mathbf{K}_k^i + \bar{\mathbf{B}} \bar{\mathbf{K}}_k^i)^T + \\ & \mathbf{P}_{k+1} (\mathbf{B} \mathbf{K}_k^i + \bar{\mathbf{B}} \bar{\mathbf{K}}_k^i) (\mathbf{H} \boldsymbol{\Sigma}_{i|i-1} \mathbf{H}^T + \mathbf{R}_v) (\mathbf{B} \mathbf{K}_k^i + \bar{\mathbf{B}} \bar{\mathbf{K}}_k^i)^T + 2 \mathbf{P}_{k+1} \mathbf{A} \mathbf{E}[\mathbf{x}_k \hat{\mathbf{x}}_{0|-1}^T] \mathbf{G}_k^T \mathbf{B}^T + \\ & \mathbf{P}_{k+1} \mathbf{B} \mathbf{G}_k \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T] \mathbf{G}_k^T \mathbf{B}^T + 2 \mathbf{P}_{k+1} \mathbf{A} \mathbf{E}[\mathbf{x}_k \hat{\mathbf{x}}_{0|-1}^T] \bar{\mathbf{G}}_k^T \bar{\mathbf{B}}^T + \\ & 2 \mathbf{P}_{k+1} \bar{\mathbf{B}} \bar{\mathbf{G}}_k \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T] \bar{\mathbf{G}}_k^T \bar{\mathbf{B}}^T + \mathbf{P}_{k+1} \bar{\mathbf{B}} \bar{\mathbf{G}}_k \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T] \bar{\mathbf{G}}_k^T \bar{\mathbf{B}}^T + \mathbf{P}_{N+1} \mathbf{X}_{N+1} \}。 \end{aligned} \quad (32)$$

根据矩阵极大值原理,可以得到最优解存在的必要条件:

$$0 = \frac{\partial H_k^f}{\partial \mathbf{K}_k^i} = \mathbf{R}_1 \mathbf{K}_k^i (\mathbf{H} \boldsymbol{\Sigma}_{i|i-1} \mathbf{H}^T + \mathbf{R}_v) + \mathbf{B}^T \mathbf{P}_{k+1} \mathbf{A} \mathbf{E}[\mathbf{x}_k \boldsymbol{\varepsilon}_i^T] + \mathbf{B}^T \mathbf{P}_{k+1} (\mathbf{B} \mathbf{K}_k^i + \bar{\mathbf{B}} \bar{\mathbf{K}}_k^i) (\mathbf{H} \boldsymbol{\Sigma}_{i|i-1} \mathbf{H}^T + \mathbf{R}_v). \quad (33)$$

$$0 = \frac{\partial H_k^f}{\partial \bar{\mathbf{K}}_k^i} = \mathbf{R}_2 \bar{\mathbf{K}}_k^i (\mathbf{H} \boldsymbol{\Sigma}_{i|i-1} \mathbf{H}^T + \mathbf{R}_v) + \bar{\mathbf{B}}^T \mathbf{P}_{k+1} \mathbf{A} \mathbf{E}[\mathbf{x}_k \hat{\mathbf{x}}_{0|-1}^T] + \bar{\mathbf{B}}^T \mathbf{P}_{k+1} \mathbf{B} \mathbf{G}_k \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T] + \bar{\mathbf{B}}^T \mathbf{P}_{k+1} \bar{\mathbf{B}} \bar{\mathbf{G}}_k \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T]. \quad (34)$$

由假设 1、假设 2 可以得到式(25)、式(26)所示的 $\mathbf{K}_k^i$ 和 $\mathbf{G}_k$ 的表达形式。进一步,基于正则条件式(22),以及 Kalman 滤波

$$\hat{\mathbf{x}}_k = \mathbf{E}[\mathbf{x}_k] + \sum_{i=0}^k \mathbf{E}[\mathbf{x}_k \boldsymbol{\varepsilon}_i^T] \mathbf{E}[\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^T]^{-1} \boldsymbol{\varepsilon}_i, \quad (35)$$

可以推出

$$\mathbf{u}_k^f = -\boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k \hat{\mathbf{x}}_k - \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k^B \mathbf{u}_k^l. \quad (36)$$

将式(36)代入式(32),可得

$$\begin{aligned} H_k^f = & \text{Tr}\{(\mathbf{Q} + \mathbf{A}^T \mathbf{P}_{k+1} \mathbf{A}) \mathbf{E}[\tilde{\mathbf{x}}_k \tilde{\mathbf{x}}_k^T] + \sum_{i=0}^k (\bar{\mathbf{K}}_k^i)^T [\mathbf{R}_2 + \bar{\mathbf{B}}^T \mathbf{P}_{k+1} \bar{\mathbf{B}} - (\mathbf{M}_k^B)^T \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k^B] \bar{\mathbf{K}}_k^i \mathbf{E}[\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^T] + \\ & \sum_{i=0}^k \mathbf{A}^T \mathbf{P}_{k+1} \mathbf{B}_k \bar{\mathbf{K}}_k^i \mathbf{E}[\boldsymbol{\varepsilon}_i \mathbf{x}_k^T] + \sum_{i=0}^k (\bar{\mathbf{K}}_k^i)^T \mathbf{B}_k \mathbf{P}_{k+1} \mathbf{A} \mathbf{E}[\mathbf{x}_k \boldsymbol{\varepsilon}_i^T] + \\ & [\mathbf{Q} + \mathbf{A}^T \mathbf{P}_{k+1} \mathbf{A} - \mathbf{M}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k] \hat{\mathbf{X}}_k + \mathbf{P}_{k+1} \mathbf{R}_w - \mathbf{P}_k \hat{\mathbf{X}}_k - \mathbf{P}_k \mathbf{E}[\tilde{\mathbf{x}}_k \tilde{\mathbf{x}}_k^T]\}. \end{aligned} \quad (37)$$

因此,可以得到

$$\mathbf{P}_k = \frac{\partial H_k^f}{\partial \hat{\mathbf{X}}_k} = \mathbf{Q} + \mathbf{A}^T \mathbf{P}_{k+1} \mathbf{A} - \mathbf{M}_k^T \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k \mathbf{P}_k. \quad (38)$$

在此基础上,进一步考虑了领导者的优化问题。由式(8)、式(36)可以得到如下领导者的 Hamiltonian 函数:

$$\begin{aligned} H_k^l = & \text{Tr}\{(\bar{\mathbf{Q}} + \mathbf{A}^T \bar{\mathbf{P}}_{k+1} \mathbf{A}) \mathbf{X}_k - 2\mathbf{A}^T \bar{\mathbf{P}}_{k+1} \mathbf{B} \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k \mathbf{E}[\mathbf{x}_k \hat{\mathbf{x}}_k^T] + \\ & 2\mathbf{A}^T \bar{\mathbf{P}}_{k+1} (\bar{\mathbf{B}} - \mathbf{B} \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k^B) \mathbf{E}[\mathbf{x}_k (\mathbf{u}_k^l)^T] + \mathbf{M}_k^T \boldsymbol{\Gamma}_k^{-1} (\bar{\mathbf{R}}_1 + \mathbf{B}^T \bar{\mathbf{P}}_{k+1} \mathbf{B}) \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k \mathbf{E}[\hat{\mathbf{x}}_k \hat{\mathbf{x}}_k^T] + \\ & 2\mathbf{M}_k^T \boldsymbol{\Gamma}_k^{-1} [(\bar{\mathbf{R}}_1 + \mathbf{B}^T \bar{\mathbf{P}}_{k+1} \mathbf{B}) \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k - \mathbf{B}^T \bar{\mathbf{P}}_{k+1} \bar{\mathbf{B}}] \mathbf{E}[\hat{\mathbf{x}}_k (\mathbf{u}_k^l)^T] + [(\mathbf{M}_k^l)^T \boldsymbol{\Gamma}_k^{-1} \bar{\mathbf{R}}_1 \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k^B + \\ & (\bar{\mathbf{B}} - \mathbf{B})^T \bar{\mathbf{P}}_{k+1} (\bar{\mathbf{B}} - \mathbf{B} \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k^B) + \bar{\mathbf{R}}_2] \mathbf{E}[\mathbf{u}_k^l (\mathbf{u}_k^l)^T]\}. \end{aligned} \quad (39)$$

根据矩阵极大值原理,得到下列关系式:

$$\begin{aligned} 0 = \frac{\partial H_k^l}{\partial \bar{\mathbf{K}}_k^i} = & [\bar{\mathbf{R}}_2 + (\mathbf{M}_k^B)^T \boldsymbol{\Gamma}_k^{-1} \bar{\mathbf{R}}_1 \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k^B + (\bar{\mathbf{B}} - \mathbf{B} \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k^B)^T \bar{\mathbf{P}}_{k+1} (\bar{\mathbf{B}} - \mathbf{B} \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k^B)] \bar{\mathbf{K}}_k^i (\mathbf{H} \boldsymbol{\Sigma}_{i|i-1} \mathbf{H}^T + \mathbf{R}_v) + \\ & (\mathbf{M}_k^B)^T \boldsymbol{\Gamma}_k^{-1} \bar{\mathbf{R}}_1 \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k \mathbf{E}[\mathbf{x}_k \boldsymbol{\varepsilon}_i^T] + (\bar{\mathbf{B}} - \mathbf{B} \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k^B)^T \bar{\mathbf{P}}_{k+1} \mathbf{A} \mathbf{E}[\mathbf{x}_k \boldsymbol{\varepsilon}_i^T] - \\ & (\bar{\mathbf{B}} - \mathbf{B} \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k)^T \bar{\mathbf{P}}_{k+1} \mathbf{B} \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k \mathbf{E}[\mathbf{x}_k \boldsymbol{\varepsilon}_i^T], \end{aligned} \quad (40)$$

$$\begin{aligned} 0 = \frac{\partial H_k^l}{\partial \bar{\mathbf{G}}_k} = & [(\mathbf{M}_k^B)^T \boldsymbol{\Gamma}_k^{-1} (\bar{\mathbf{R}}_1 + \mathbf{B}^T \bar{\mathbf{P}}_{k+1} \mathbf{B}) \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k^B + \bar{\mathbf{R}}_2 + \bar{\mathbf{B}}^T \bar{\mathbf{P}}_{k+1} \bar{\mathbf{B}} - (\mathbf{M}_k^B)^T \boldsymbol{\Gamma}_k^{-1} \mathbf{B}^T \bar{\mathbf{P}}_{k+1} \bar{\mathbf{B}} - \\ & \bar{\mathbf{B}}^T \bar{\mathbf{P}}_{k+1} \mathbf{B} \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k^B] \bar{\mathbf{G}}_k \mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T] + [(\mathbf{M}_k^B)^T \boldsymbol{\Gamma}_k^{-1} (\bar{\mathbf{R}}_1 + \mathbf{B}^T \bar{\mathbf{P}}_{k+1} \mathbf{B}) \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k - \\ & (\mathbf{M}_k^B)^T \boldsymbol{\Gamma}_k^{-1} \mathbf{B}^T \bar{\mathbf{P}}_{k+1} \mathbf{A} + \bar{\mathbf{B}}^T \bar{\mathbf{P}}_{k+1} \mathbf{A} - \bar{\mathbf{B}}^T \bar{\mathbf{P}}_{k+1} \mathbf{B} \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k] \mathbf{E}[\mathbf{x}_k \hat{\mathbf{x}}_{0|-1}^T]. \end{aligned} \quad (41)$$

因此,可以得到式(23)、式(24)。进一步,由式(35),可以得到

$$\mathbf{u}_k^l = -\bar{\boldsymbol{\gamma}}_k^{-1} \bar{\mathbf{M}}_k \left\{ \sum_{i=0}^k \mathbf{E}[\mathbf{x}_k \boldsymbol{\varepsilon}_i^T] \mathbf{E}[\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^T]^{-1} \boldsymbol{\varepsilon}_i + \mathbf{E}[\mathbf{x}_k \hat{\mathbf{x}}_{0|-1}^T] \text{Pinv}[\mathbf{E}[\hat{\mathbf{x}}_{0|-1} \hat{\mathbf{x}}_{0|-1}^T]] \hat{\mathbf{x}}_{0|-1} \right\} = -\bar{\boldsymbol{\gamma}}_k^{-1} \bar{\mathbf{M}}_k \hat{\mathbf{x}}_k. \quad (42)$$

类似地,由式(36)可以推出

$$\mathbf{u}_k^f = -\boldsymbol{\Gamma}_k^{-1} [\mathbf{M}_k - \mathbf{M}_k^B \bar{\boldsymbol{\gamma}}_k^{-1} \bar{\mathbf{M}}_k] \hat{\mathbf{x}}_k = -\boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k \hat{\mathbf{x}}_k. \quad (43)$$

式中: $\mathbf{M}_k \triangleq \mathbf{M}_k - \mathbf{M}_k^B \bar{\boldsymbol{\gamma}}_k^{-1} \bar{\mathbf{M}}_k$ 。

将式(42)和式(43)代入式(39),可以得到 Riccati 型方程:

$$\begin{aligned} \bar{\mathbf{P}}_k = \frac{\partial H_k^l}{\partial \hat{\mathbf{X}}_k} = & \bar{\mathbf{Q}} + \mathbf{A}^T \bar{\mathbf{P}}_{k+1} \mathbf{A} + \mathbf{M}_k^T \boldsymbol{\Gamma}_k^{-1} \bar{\mathbf{R}}_1 \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k + \bar{\mathbf{M}}_k^T \bar{\boldsymbol{\gamma}}_k^{-1} \bar{\mathbf{R}}_2 \bar{\boldsymbol{\gamma}}_k^{-1} \bar{\mathbf{M}}_k - \mathbf{A}^T \bar{\mathbf{P}}_{k+1} (\mathbf{B} \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k + \bar{\mathbf{B}} \bar{\boldsymbol{\gamma}}_k^{-1} \bar{\mathbf{M}}_k) - \\ & (\mathbf{B} \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k + \bar{\mathbf{B}} \bar{\boldsymbol{\gamma}}_k^{-1} \bar{\mathbf{M}}_k)^T \bar{\mathbf{P}}_{k+1} \mathbf{A} + (\mathbf{B} \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k + \bar{\mathbf{B}} \bar{\boldsymbol{\gamma}}_k^{-1} \bar{\mathbf{M}}_k)^T \bar{\mathbf{P}}_{k+1} (\mathbf{B} \boldsymbol{\Gamma}_k^{-1} \mathbf{M}_k + \bar{\mathbf{B}} \bar{\boldsymbol{\gamma}}_k^{-1} \bar{\mathbf{M}}_k). \end{aligned} \quad (44)$$

进一步化简可以得到式(22)。

经典的最优控制方法,如动态规划、庞特里亚金最大值原理,不能直接解决闭环 Stackelberg 策略的最优

问题,这仍然是一个具有挑战性的问题。通过采用矩阵极大值原理,上述定理提供了一般的线性闭环解。

由上述结果可知,以新息的线性组合形式表示的线性闭环解实际上等于卡尔曼滤波反馈的形式。然而,需要强调的是,增益矩阵 $-\bar{\gamma}_k^{-1}\bar{M}_k$ 和 $-\Gamma_k^{-1}M_k$ 与给定的 Stackelberg 策略下得到的增益矩阵 $K_k\hat{x}_k$ 和 $\bar{K}_k\hat{x}_k$ 是不同的,将在下面的定理中给出详细的解释。

3 反馈 Stackelberg 解

假设

$$u_k^f = K_k \hat{x}_k, u_k^l = \bar{K}_k \hat{x}_k. \quad (45)$$

式中: $\hat{x}_k$ 表示 Kalman 滤波,如(30)所示; $K_k, \bar{K}_k$ 为待确定增益矩阵。

基于式(45),系统(式(1))可以写成

$$x_{k+1} = Ax_k + (BK_k + \bar{B}\bar{K}_k)\hat{x}_k + w_k. \quad (46)$$

那么, X_{k+1} 可以重新表示为:

$$X_{k+1} = AX_k A^T + AX_k \hat{X}_k (BK_k + \bar{B}\bar{K}_k)^T + (BK_k + \bar{B}\bar{K}_k) \hat{X}_k A^T + (BK_k + \bar{B}\bar{K}_k) \hat{X}_k (BK_k + \bar{B}\bar{K}_k)^T + R_w. \quad (47)$$

式中, $\hat{X}_k = E[x_k x_k^T]$ 。反馈 Stackelberg 的问题可描述为:

问题 2 寻找形如式(45)所示的 Kalman 滤波反馈控制器 u_k^f 和 u_k^l ,使得性能指标(式(7)和式(8))最小。应用矩阵极大值原理,问题 2 的解可以由定理 2 给出。

定理 2 使得性能指标(式(7)和式(8))最小化的增益矩阵 $K_k, \bar{K}_k$ 的表达式分别为:

$$K_k = -(\Gamma_k^l)^{-1}M_k^l, \quad (48)$$

$$\bar{K}_k = -(\bar{\gamma}_k^l)^{-1}\bar{M}_k^l. \quad (49)$$

式中:

$$\Gamma_k^l = R_1 + B^T P_{k+1}^l B, \quad (50)$$

$$M_k^l = M_k^l - M_k^{Bl} (\bar{\gamma}_k^l)^{-1} \bar{M}_k^l, \quad (51)$$

$$\bar{\gamma}_k^l = \bar{R}_2 + (M_k^{Bl})^T (\Gamma_k^l)^{-1} \bar{R}_1 (\Gamma_k^l)^{-1} M_k^{Bl} + (B_k^l)^T \bar{P}_{k+1}^l B_k^l, \quad (52)$$

$$\bar{M}_k^l = (M_k^{Bl})^T (\Gamma_k^l)^{-1} \bar{R}_1 (\Gamma_k^l)^{-1} M_k^l + (B_k^l)^T \bar{P}_{k+1}^l A_k^l, \quad (53)$$

$$P_k^l = Q + K_k^T R_1 K_k + \bar{K}_k^T R_2 \bar{K}_k + (A + BK_k + \bar{B}\bar{K}_k)^T P_{k+1}^l (A + BK_k + \bar{B}\bar{K}_k), \quad (54)$$

$$\bar{P}_k^l = \bar{Q} + \bar{K}_k^T \bar{R}_1 K_k + \bar{K}_k^T \bar{R}_2 \bar{K}_k + (A + BK_k + \bar{B}\bar{K}_k)^T \bar{P}_{k+1}^l (A + BK_k + \bar{B}\bar{K}_k), \quad (55)$$

$$M_k^l = B^T P_{k+1}^l A, \quad (56)$$

$$M_k^{Bl} = B^T P_{k+1}^l \bar{B}, \quad (57)$$

$$B_k^l = \bar{B} - B (\Gamma_k^l)^{-1} M_k^{Bl}, \quad (58)$$

$$A_k^l = A - B (\Gamma_k^l)^{-1} M_k^l. \quad (59)$$

证明:证明的大纲与定理 1 中的类似。首先,优化跟随者的性能指标。根据式(7)和式(47),定义跟随者的 Hamiltonian 函数:

$$\begin{aligned} \tilde{H}_k^f = & \text{Tr}\{QX_k + (K_k^T R_1 K_k + \bar{K}_k^T R_2 \bar{K}_k) \hat{X}_k + P_{k+1}^l [AX_k A^T + AX_k \hat{X}_k (BK_k + \bar{B}\bar{K}_k)^T + \\ & (BK_k + \bar{B}\bar{K}_k) \hat{X}_k A^T + (BK_k + \bar{B}\bar{K}_k) \hat{X}_k (BK_k + \bar{B}\bar{K}_k)^T + R_w]\}. \end{aligned} \quad (60)$$

应用矩阵极大值原理,得到

$$0 = \frac{\partial \tilde{H}_k^f}{\partial X_k} = R_1 K_k \hat{X}_k + B^T P_{k+1}^l A \hat{X}_k + B^T P_{k+1}^l (BK_k + \bar{B}\bar{K}_k) \hat{X}_k, \quad (61)$$

$$\begin{aligned} P_k^l = \frac{\partial \tilde{H}_k^f}{\partial \hat{X}_k} = & Q + (K_k^T R_1 K_k + \bar{K}_k^T R_2 \bar{K}_k) + A^T P_{k+1}^l A + A^T P_{k+1}^l (BK_k + \bar{B}\bar{K}_k)^T + \\ & (BK_k + \bar{B}\bar{K}_k)^T P_{k+1}^l A + (BK_k + \bar{B}\bar{K}_k)^T P_{k+1}^l (BK_k + \bar{B}\bar{K}_k). \end{aligned} \quad (62)$$

因此,可以得到 K_k 和 P_k^l 。

$$K_k = -(R_1 + B^T P_{k+1}^l B)^{-1} (B^T P_{k+1}^l A + B^T P_{k+1}^l \bar{B} \bar{K}_k) \triangleq -(\Gamma_k^l)^{-1} (M_k^l + M_k^{Bl} \bar{K}_k), \quad (63)$$

$$P_k^l = Q + (K_k^T R_1 K_k + \bar{K}_k^T R_2 \bar{K}_k) + (A + BK_k + \bar{B} \bar{K}_k)^T P_{k+1}^l (A + BK_k + \bar{B} \bar{K}_k). \quad (64)$$

接下来,优化领导者的性能指标。为此,由式(63),式(47)可以重新表达为:

$$\begin{aligned} X_{k+1} = & AX_k A^T + A \hat{X}_k [-B(\Gamma_k^l)^{-1} M_k^l + B_k^l \bar{K}_k]^T + [-B(\Gamma_k^l)^{-1} M_k^l + B_k^l \bar{K}_k] \hat{X}_k A^T + \\ & [-B(\Gamma_k^l)^{-1} M_k^l + B_k^l \bar{K}_k] \hat{X}_k [-B(\Gamma_k^l)^{-1} M_k^l + B_k^l \bar{K}_k]^T + R_w. \end{aligned} \quad (65)$$

因此,领导者的 Hamiltonian 函数可以定义为:

$$\begin{aligned} \tilde{H}_k = & \text{Tr}\{\bar{Q} X_k + (M_k^l + M_k^{Bl} \bar{K}_k)^T (\Gamma_k^l)^{-1} \bar{R}_1 (\Gamma_k^l)^{-1} (M_k^l + M_k^{Bl} \bar{K}_k) \hat{X}_k + \bar{K}_k^T \bar{R}_2 \bar{K}_k \hat{X}_k + \\ & \bar{P}_{k+1}^l [A X_k A^T + A \hat{X}_k (-B(\Gamma_k^l)^{-1} M_k^l + B_k^l \bar{K}_k)^T + (-B(\Gamma_k^l)^{-1} M_k^l + B_k^l \bar{K}_k) \hat{X}_k A^T + \\ & (-B(\Gamma_k^l)^{-1} M_k^l + B_k^l \bar{K}_k) \hat{X}_k (-B(\Gamma_k^l)^{-1} M_k^l + B_k^l \bar{K}_k)^T + R_w]\}. \end{aligned} \quad (66)$$

基于矩阵极大值原理,有

$$\begin{aligned} 0 = \frac{\partial \tilde{H}_k}{\partial \bar{K}_k} = & \text{Tr}\{[\bar{R}_2 + (M_k^{Bl})^T (\Gamma_k^l)^{-1} \bar{R}_1 (\Gamma_k^l)^{-1} M_k^{Bl} + (B_k^l)^T \bar{P}_{k+1}^l B_k^l] \bar{K}_k + \\ & (M_k^{Bl})^T (\Gamma_k^l)^{-1} \bar{R}_1 (\Gamma_k^l)^{-1} M_k^l + (B_k^l)^T \bar{P}_{k+1}^l A - (B_k^l)^T \bar{P}_{k+1}^l B (\Gamma_k^l)^{-1} M_k^l] \hat{X}_k\}, \end{aligned} \quad (67)$$

$$\begin{aligned} \bar{P}_k = \frac{\partial \tilde{H}_k}{\partial \hat{X}_k} = & \bar{Q} + (M_k^l + M_k^{Bl} \bar{K}_k)^T (\Gamma_k^l)^{-1} \bar{R}_1 (\Gamma_k^l)^{-1} (M_k^l + M_k^{Bl} \bar{K}_k) + \bar{K}_k^T \bar{R}_2 \bar{K}_k + \\ & (A - B(\Gamma_k^l)^{-1} M_k^l + B_k^l \bar{K}_k)^T \bar{P}_{k+1}^l (A - B(\Gamma_k^l)^{-1} M_k^l + B_k^l \bar{K}_k). \end{aligned} \quad (68)$$

因此,可以得到 $\bar{K}_k = -(\bar{\gamma}_k^l)^{-1} \bar{M}_k^l$ 。

将定理 1 的结果与定理 2 的结果进行比较,两个控制器都以 Kalman 滤波反馈的形式存在,不同之处在于增益矩阵。实际上,新息的线性组合是线性闭环 Stackelberg 解的最一般形式。通过比较,问题 2 是通过来自领导者和跟随者更具体的信息情况下解决的。基于这一观点,优化对这两种情况的影响是不同的。

4 例子

在本节中,通过举例来说明定理 1 的优化效果。考虑下列参数情形下的 Stackelberg 博弈控制问题:

$A=1, B=0.5, \bar{B}=1, H=1, R_1=\bar{R}_1=1, R_2=\bar{R}_2=2, Q=1, \bar{Q}=2, R_0=1, N=20, P_{N+1}=\bar{P}_{N+1}=1$, 并且随机变量满足 $x_0 \sim N(0.5, 1), w_k \sim N(0, 0.5), v_k \sim N(0, 0.1)$ 。

首先,由方程式(14)、式(15)及式(18)容易得出假设 1 成立,并且由 $\hat{x}_{0|-1} = \bar{x}_0 = 0.5$ 可知假设 2 显然成立。应用 MATLAB 可以得到 Riccati 方程式(11)、式(12)的解如图 1 所示。进一步,根据定理 1,可以计算出跟随者及领导者的最优性能指标为 $J^f = 25.2119, J^l = 40.4430$ 。

5 结论

本研究讨论了离散时间随机 LQ Stackelberg 博弈的闭环解。根据闭环信息结构的定义,使用传统的最优控制方法直接解决该问题非常困难。此外,由于领导者最优必要条件中存在二阶偏导数,最优解很可能是非线性的。因此,自然地考虑了线性最优闭环解。假设闭环 Stackelberg 策略是历史和当前新息的线性组

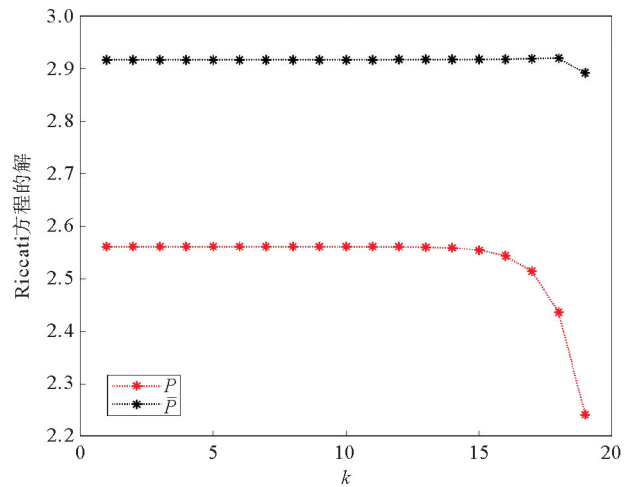


图 1 Riccati 方程的解

Fig. 1 The solution to the Riccati equation

合,利用矩阵极大值原理,给出了最优策略的表达式,其中涉及耦合的 Riccati 方程。研究发现,线性最优闭环策略实际上等于 Kalman 滤波反馈。此外,还给出了两名玩家的最优策略都假定为 Kalman 滤波反馈的情况。由于可用信息不同,优化效果也不同。基于这一思路,未来可以进一步探讨非对称信息的 Stackelberg 博弈,即当参与博弈的双方所持有的信息并不完全相同时,经典的 Stackelberg 博弈理论可能不再适用,此时博弈的结果将不再是基于精确的预期,而是基于双方的推测,这也将是一个非常值得探讨的问题。

参考文献:

- [1] STACKELBERG H. Marktjorm und gleichgewicht[M]. Vienna:Springer,1934.
- [2] CHEN C,CRUZ J. Stackelberg solution for two-person games with biased information patterns[J]. IEEE Transactions on Automatic Control,1972,17(6):791-798.
- [3] BASAR T,OLSDER J. Dynamic noncooperative game theory[M]. Philadelphia:Society for Industrial and Applied Mathematics,1999.
- [4] FREILING G,JANK G,ABOU-KANDIL H. Discrete-time Riccati equations in open-loop Nash and Stackelberg games[J]. European Journal of Control,1999,5(1):56-66.
- [5] XU J J,ZHANG H S,CHAI T. Necessary and sufficient condition for two-player Stackelberg strategy[J]. IEEE Transactions on Automatic Control,2015,60(5):1356-1361.
- [6] CRUZ J,SIMAAN M. On the Stackelberg strategy in nonzero-sum games[J]. Journal of Optimization Theory and Applications,1973,11(5):533-555.
- [7] CRUZ J,SIMAAN M. Additional aspects of the Stackelberg strategy in nonzero-sum games[J]. Journal of Optimization Theory and Applications,1973,11(6):613-626.
- [8] JUNGERS M. Matrix block formulation of closed-loop memoryless Stackelberg strategy for discrete-time games[C]//2008 47th IEEE Conference on Decision and Control,2008:4454-4459.
- [9] LI M,CRUZ J,SIMAAN M A. An approach to discrete-time incentive feedback Stackelberg games[J]. IEEE Transactions on Systems,Man,and Cybernetics-Part A:Systems and Humans,2002,32(4):472-481.
- [10] BASAR T,SELBUZ H. A new approach for the derivation of closed-loop Stackelberg strategies[C/OL]//Proceedings of the IEEE Conference on Decision and Control. San Diego,1979. DOI:10.1109/CDC.1978.268106.
- [11] BASAR T,SELBUZ H. Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems [J]. IEEE Transactions on Automatic Control,1979,AC-24(2):166-179.
- [12] BASAR T,OLSDER J. Team-optimal closed-loop Stackelberg strategies in hierarchical control problems[J]. Automatica,1980,16:409-414.
- [13] MEDANIC J. Closed-loop Stackelberg strategies in linear-quadratic problems[J]. IEEE Transactions on Automatic Control,1978,23(4):632-637.
- [14] PAPA VASSILOPOULOS G,CRUZ J. Nonclassical control problems and Stackelberg games[J]. IEEE Transactions on Automatic Control,1979,24(2):155-166.
- [15] TOLWINSKI B. Closed-loop Stackelberg solution to multistage linear quadratic games[J]. Journal of Optimization Theory and Application,1981,34(4):485-501.
- [16] CASTANON D,ATHANS M. On the stochastic dynamic Stackelberg strategies[J]. Automatica,1976,12(2):177-183.
- [17] AOKI M. New trends in dynamic systems theory and economics[M]. London:Academic Press,1979.
- [18] ANDERSON B,MOORE J. Optimal filtering[M]. New York:Dover Publications,2005.

(责任编辑:齐敏华)