中国科学: 数学 2025 年 第 55 卷 第 ? 期: 1 ∼ 22
SCIENTIA SINICA Mathematica
综 述

随机线性二次最优控制的一些研究进展

孙景瑞*, 袁吕宁, 张佳琪

南方科技大学理学院数学系, 深圳 518055
E-mail: sunjr@sustech.edu.cn, yuanln@sustech.edu.cn, zhangjq3@sustech.edu.cn

收稿日期: 2024-06-27; 接受日期: 2024-12-25; 网络出版日期: 2025-01-22; * 通信作者
国家自然科学基金 (批准号: 12322118 和 12271242) 和深圳市自然科学基金 (批准号: JCYJ20220530112814032) 资助项目

摘要 本文综合了作者及其合作者们的最新相关工作, 概述了随机线性二次最优控制问题的一些研究进展. 文章主题由密切相关的三部分组成, 分别是有限时区上的随机线性二次最优控制, 无限时区上的随机线性二次最优控制, 以及随机线性二次最优控制的 turnpike 性质.

关键词 随机最优控制 线性二次 Riccati 方程 turnpike 性质

MSC (2020) 主题分类 49N10, 49N35, 93E15, 93E20

1 引言

为方便起见, 首先引入一些记号. 记 $\mathbb{S}^n$ 为 $n\times n$ 阶实对称矩阵组成的集合, $\mathbb{S}^n_+$ 为 $\mathbb{S}^n$ 中的正定矩阵组成的集合, I_n 为 n 阶单位矩阵, $M^\top$ 为矩阵 M 的转置, $M^\dagger$ 为 M 的广义逆, $\mathscr{R}(M)$ 为矩阵 (或算子) M 的像空间. 对于矩阵 $M,N\in\mathbb{S}^n$, 若 $M-N$ 为半正定矩阵 (对应地, 正定矩阵), 记为 $M\geqslant N$ (对应地, $M>N$). 设 $(\Omega,\mathcal{F},\mathbb{P})$ 为一个完备概率空间, 其上可定义一维标准 Brown 运动 $W=\{W(t);t\geqslant 0\}$. 令 $\mathbb{F}=\{\mathcal{F}_t\}_{t\geqslant 0}$ 为 W 生成的 σ 域流的通常扩张 (usual augmentation). 我们将关于 $\mathcal{F}_t$ 可测的随机变量 ξ 记作 $\xi\in\mathcal{F}_t$, 将关于域流 $\mathbb{F}$ 循序可测的随机过程 X 记作 $X\in\mathbb{F}$. 设 T 为一个正实数, $\mathbb{H}$ 为一个 Euclid 空间 (如 $\mathbb{R}^n$, $\mathbb{R}^{n\times m}$ 和 $\mathbb{S}^n$ 等), 其上赋予 Frobenius 内积 $\langle\cdot\,,\cdot\rangle$, 该内积诱导的范数记为 $|\cdot|$. 令

$$
\begin{aligned}
&C([0,T];\mathbb{H})\triangleq\{\varphi:[0,T]\to\mathbb{H}\mid \varphi(\cdot)\text{ 为连续映射}\},\\
&L^2(0,T;\mathbb{H})\triangleq\left\{\varphi:[0,T]\to\mathbb{H}\,\middle|\,\varphi(\cdot)\text{ Lebesgue 可测, 且 }\int_0^T|\varphi(t)|^2dt<\infty\right\},\\
&L^2_{\mathbb{F}}(0,T;\mathbb{H})\triangleq\left\{\varphi:[0,T]\times\Omega\to\mathbb{H}\,\middle|\,\varphi\in\mathbb{F}\text{ 且 }\mathbb{E}\int_0^T|\varphi(t)|^2dt<\infty\right\}.
\end{aligned}
$$

类似地, 可以定义 $C([0,\infty);\mathbb{H})$, $L^2(0,\infty;\mathbb{H})$ 和 $L^2_{\mathbb{F}}(0,\infty;\mathbb{H})$.

英文引用格式: Sun J R, Yuan L N, Zhang J Q. Some recent developments in stochastic linear-quadratic optimal control (in Chinese). Sci Sin Math, 2025, 55: 1–22, doi: 10.1360/SSM-2024-0217

考虑受控的线性随机微分方程 (stochastic differential equation, SDE)

$$\begin{cases} dX(t) = [A(t)X(t) + B(t)u(t) + b(t)]dt + [C(t)X(t) + D(t)u(t) + \sigma(t)]dW(t), \quad t \geqslant 0, \\ X(0) = x, \end{cases} \tag{1.1}$$

以及二次型代价泛函

$$\begin{aligned} J_T(x; u(\cdot)) \triangleq \mathbb{E}\bigg\{ \int_0^T & [\langle Q(t)X(t), X(t)\rangle + 2\langle S(t)X(t), u(t)\rangle + \langle R(t)u(t), u(t)\rangle \\ & + 2\langle q(t), X(t)\rangle + 2\langle r(t), u(t)\rangle]dt + \langle GX(T), X(T)\rangle + 2\langle g, X(T)\rangle \bigg\}, \end{aligned} \tag{1.2}$$

其中, $X(\cdot)$ 是取值于 $\mathbb{R}^n$ 的状态过程, $u(\cdot)$ 是取值于 $\mathbb{R}^m$ 的控制过程. 在 (1.2) 中, 等号右侧的积分项称为期望运行成本, 剩余两项的和称为期望终端成本.

关于 (1.1) 和 (1.2) 中的系数, 我们作如下假设:

(H1) $A, C : [0, \infty) \to \mathbb{R}^{n\times n}$, $B, D : [0, \infty) \to \mathbb{R}^{n\times m}$, $b, \sigma, q : [0, \infty) \to \mathbb{R}^n$, $Q : [0, \infty) \to \mathbb{S}^n$, $S : [0, \infty) \to \mathbb{R}^{m\times n}$, $R : [0, \infty) \to \mathbb{S}^m$, $r : [0, \infty) \to \mathbb{R}^m$ 均为确定性的、Lebesgue 可测的有界函数, $G \in \mathbb{S}^n$ 和 $g \in \mathbb{R}^n$ 为常值函数.

令 $\mathscr{U}[0, T] \triangleq L^2_{\mathbb{F}}(0, T; \mathbb{R}^m)$. 显然, 在假设 (H1) 下, 对任意给定的初始状态 $x \in \mathbb{R}^n$ 以及任意的控制过程 $u(\cdot) \in \mathscr{U}[0, T]$, 状态方程 (1.1) 在 $[0, T]$ 上存在唯一强解 $X(\cdot) \equiv X(\cdot, x, u(\cdot))$, 并且代价泛函 $J_T(x; u(\cdot))$ 中的积分项均是良定的. 从而, 可以将有限时区上的随机线性二次 (linear-quadratic, LQ) 最优控制问题表述如下.

问题 (SLQ)$_T$ 对于任意给定的初始状态 $x \in \mathbb{R}^n$, 寻找控制过程 $\bar{u}(\cdot) \in \mathscr{U}[0, T]$, 使得

$$J_T(x; \bar{u}(\cdot)) = \inf_{u(\cdot)\in\mathscr{U}[0,T]} J_T(x; u(\cdot)) \equiv V_T(x). \tag{1.3}$$

称满足 (1.3) 的 $\bar{u}(\cdot) \in \mathscr{U}[0, T]$ 为对应于初始状态 x 的开环最优控制, 称对应的状态过程 $\bar{X}(\cdot) \equiv X(\cdot\,; x, \bar{u}(\cdot))$ 与二元组 $(\bar{X}(\cdot), \bar{u}(\cdot))$ 为最优状态过程与最优对, 称 $V_T(\cdot)$ 为问题 (SLQ)$_T$ 的值函数.

当 $b(\cdot)$ 和 $\sigma(\cdot)$ 恒为 0 时, 称控制系统 (1.1) 为齐次的, 并记相应的状态过程为 $X^0(\cdot) \equiv X^0(\cdot\,; x, u(\cdot))$, 它满足如下方程:

$$\begin{cases} dX^0(t) = [A(t)X^0(t) + B(t)u(t)]dt + [C(t)X^0(t) + D(t)u(t)]dW(t), \quad t \geqslant 0, \\ X^0(0) = x. \end{cases} \tag{1.4}$$

此外, 当 $q(\cdot)$, $r(\cdot)$ 和 g 也恒为 0 时, 记对应的代价泛函为 $J^0_T(x; u(\cdot))$, 它由下式给出:

$$\begin{aligned} J^0_T(x; u(\cdot)) = \mathbb{E}\bigg\{ \int_0^T & [\langle Q(t)X^0(t), X^0(t)\rangle + 2\langle S(t)X^0(t), u(t)\rangle + \langle R(t)u(t), u(t)\rangle]dt \\ & + \langle GX^0(T), X^0(T)\rangle \bigg\}. \end{aligned} \tag{1.5}$$

当状态方程满足 (1.4) 且代价泛函满足 (1.5) 时, 记相应的最优控制问题为问题 (SLQ)^0_T, 并称之为有限时区上的齐次随机 LQ 最优控制问题. 此时的值函数记为 $V^0_T(\cdot)$.

当 $T = \infty$ 时, 考虑如下形式的随机控制系统:

$$\begin{cases} dX(t) = [AX(t) + Bu(t) + b(t)]dt + [CX(t) + Du(t) + \sigma(t)]dW(t), \quad t \geqslant 0, \\ X(0) = x, \end{cases} \tag{1.6}$$

以及如下形式的代价泛函:

$$J_\infty(x;u(\cdot)) \triangleq \mathbb{E}\int_0^\infty [\langle QX(t),X(t)\rangle + 2\langle SX(t),u(t)\rangle + \langle Ru(t),u(t)\rangle + 2\langle q(t),X(t)\rangle + 2\langle r(t),u(t)\rangle]dt, \tag{1.7}$$

其中, 系数 A, B, C, D, Q, S 和 R 均为常值矩阵, 状态方程中的非齐次项 $b(\cdot)$ 和 $\sigma(\cdot)$ 以及代价泛函中的一阶系数 $q(\cdot)$ 和 $r(\cdot)$ 满足以下条件:

(H2) $b,\sigma,q:[0,\infty)\to\mathbb{R}^n$, $r:[0,\infty)\to\mathbb{R}^m$ 均为确定性的 Lebesgue 可测函数, 且满足

$$\int_0^\infty [|b(t)|^2 + |\sigma(t)|^2 + |q(t)|^2 + |r(t)|^2]dt < \infty.$$

在无限时区下, 为确保代价泛函 (1.7) 中的被积过程可积, 定义关于初始状态 x 的容许控制集为

$$\mathscr{U}_{ad}(x) \triangleq \left\{ u(\cdot)\in L^2_{\mathbb{F}}(0,\infty;\mathbb{R}^m) \,\middle|\, \mathbb{E}\int_0^\infty |X(t;x,u(\cdot))|^2 dt < \infty \right\}. \tag{1.8}$$

从而, 无限时区 $[0,\infty)$ 上的随机 LQ 最优控制问题可表述如下:

问题 (SLQ)$_\infty$ 对于任意给定的初始状态 $x\in\mathbb{R}^n$, 寻找容许控制 $\bar u(\cdot)\in\mathscr{U}_{ad}(x)$, 使得

$$J_\infty(x;\bar u(\cdot)) = \inf_{u(\cdot)\in\mathscr{U}_{ad}(x)} J_\infty(x;u(\cdot)) \equiv V_\infty(x).$$

与有限时区的情况类似, 当 $b(\cdot)$, $\sigma(\cdot)$, $q(\cdot)$ 和 $r(\cdot)$ 均恒等于 0 时, 记相应的最优控制问题为问题 (SLQ)$^0_\infty$, 并称之为无限时区 $[0,\infty)$ 上的齐次随机 LQ 最优控制问题. 此时的代价泛函和值函数分别记为 $J^0_\infty(x;u(\cdot))$ 和 $V^0_\infty(\cdot)$.

关于 LQ 最优控制问题的研究, 可追溯到 1960 年左右由 Bellman 等 [5], Kalman [18] 和 Letov [21] 等所进行的关于常微分方程 (ordinary differential equation, ODE) 系统的开创性工作. 关于 ODE 的 LQ 理论的标准结果可以参见文献 [1,20,42,44] 等的先驱性工作. 同时也可参阅最优控制理论的经典著作 (参见文献 [45]).

随着确定性最优控制研究的不断发展, 借助现代概率论的工具, 随机 LQ 问题也在 20 世纪 60 年代得到了开创性研究, 这得益于 Kushner [19] 和 Wonham [43] 的先驱性工作. 关于经典的随机 LQ 问题的研究, 可以参见文献 [7,12] 等著名研究. 1998 年, Chen 等 [10] 发现, 在随机 LQ 问题中, 代价泛函中的加权矩阵在某些条件下可以是不确定的. 特别地, 代价泛函 (1.2) 中的权重矩阵 $R(\cdot)$ 的半正定性不再是随机 LQ 问题中最优控制存在的必要条件. 此外, Yong 和 Zhou [45] 在其专著中详细介绍了 21 世纪之前关于随机 LQ 问题的研究进展.

自 2010 年初以来, 本文作者及其合作者们开始从不同角度研究关于 SDE 的随机 LQ 最优控制问题, 引入了有限性、开环和闭环可解性等概念, 并探讨了这些概念之间的关系以及相应的刻画准则. 特别地, 凭借微分/代数 Riccati 方程的解, 开环最优控制通常具有闭环表示. 这方面的具体研究可参见文献 [28,29,31–34,36] 等工作.

本文探讨的另一个主题是随机 LQ 最优控制的 turnpike 性质 (有学者将其译为 "大道性质"). Turnpike 性质的命名源自美国的收费公路系统. 在美国, 收费公路通常被称为 turnpike. 这些公路是为了长途旅行和运输而设计的高速公路, 通常是直线或者近乎直线的. 在最优控制理论中, turnpike 现象的命名正是因为它与这种收费公路的性质相似. 具体而言, turnpike 现象描述了在一个有限但足够长的控制时间段内, 系统的最优控制策略会使系统的状态轨迹在大部分时间接近某个稳定状态, 而只有

在起始和终止的时间段内会有显著的偏离. 这与收费公路的直线性质类似, 即在大部分行程中车辆会保持直线行驶, 只有在入口和出口处才会进行较大的转向或调整. 这种性质在经济学、工程学和其他应用领域的长期规划中具有重要意义, 因为它表明在多数时间内系统的控制策略会趋向于一个相对稳定的状态, 利用这个稳定状态作为近似最优解, 从而简化长时间控制问题的分析和设计.

关于 turnpike 性质的研究可追溯到 1945 年 von Neumann [41] 在无限时间尺度下针对确定性最优增长问题开展的工作. 1958 年, Dorfman 等 [13] 首次提出了 "turnpike" 这一术语. 该概念在文献 [9] 中得到了详细阐述. 几十年来, turnpike 理论吸引了众多学者的关注, 因为它通常能够直观地描述最优解的长期稳定特性, 无需进行复杂的解析计算, 从而显著简化了求解这类最优控制问题的数值方法. 在离散时间和连续时间的确定性系统背景下, 学者们对有限维和无限维问题进行了广泛的研究, 相关文献包括但不限于文献 [8, 11, 14–16, 22, 27, 30, 37, 38, 46–48]. 这些研究不仅扩展了理论的深度和广度, 还为实际应用提供了重要的指导和方法.

关于确定性 LQ 最优控制问题, Porretta 和 Zuazua [26] 以及 Trélat 和 Zuazua [40] 证明了在状态方程能控和代价泛函能观的条件下, 最优对具有指数 turnpike 性质. 具体而言, 针对以下形式的 ODE 控制系统:

$$\begin{cases} \dot{X}(t) = AX(t) + Bu(t) + b, \quad t \geqslant 0, \\ X(0) = x, \end{cases}$$

以及以下形式的代价泛函:

$$\begin{aligned} J_T(x; u(\cdot)) \triangleq \int_0^T & [\langle QX(t), X(t)\rangle + 2\langle SX(t), u(t)\rangle + \langle R(t)u(t), u(t)\rangle \\ & + 2\langle q, X(t)\rangle + 2\langle r, u(t)\rangle]dt, \end{aligned}$$

文献 [26, 40] 证明了存在与控制时长 T 无关的常数 $K, \lambda > 0$, 使得确定性 LQ 问题的最优对 $(\bar{X}_T(\cdot), \bar{u}_T(\cdot))$ 满足不等式

$$|\bar{X}_T(t) - x^*| + |\bar{u}_T(t) - u^*| \leqslant K[\mathrm{e}^{-\lambda t} + \mathrm{e}^{-\lambda(T-t)}], \quad \forall\, t \in [0, T], \tag{1.9}$$

其中, (x^*, u^*) 是以下静态优化问题 (有限维的条件极值问题) 的解:

$$\begin{aligned} &\min\ F_0(x, u) \triangleq \langle Qx, x\rangle + 2\langle Sx, u\rangle + \langle Ru, u\rangle + 2\langle q, x\rangle + 2\langle r, u\rangle \\ &\ \text{s.t.}\ \ Ax + Bu + b = 0. \end{aligned} \tag{1.10}$$

这一指数 turnpike 性质 (1.9) 意味着, 对任一常数 $\delta \in (0, \frac{1}{2})$, 都有以下估计:

$$|\bar{X}_T(t) - x^*| + |\bar{u}_T(t) - u^*| \leqslant 2K\mathrm{e}^{-\lambda\delta T}, \quad \forall\, t \in [\delta T, (1-\delta)T].$$

由于 T 较大时, 即便 δ 很小, 不等式右边的这一项 $2K\mathrm{e}^{-\lambda\delta T}$ 也非常接近于 0, 所以在控制区间的大部分时间里, 最优对 $(\bar{X}_T(\cdot), \bar{u}_T(\cdot))$ 近似等于静态优化问题 (1.10) 的解 (x^*, u^*).

本文将介绍在随机情形下相应的 LQ 最优控制问题的 turnpike 结果. 为叙述简便起见, 本文只讨论常系数的情形, 即状态方程具有如下形式:

$$\begin{cases} dX(t) = [AX(t) + Bu(t) + b]dt + [CX(t) + Du(t) + \sigma]dW(t), \quad t \geqslant 0, \\ X(0) = x, \end{cases} \tag{1.11}$$

代价泛函具有如下形式:

$$J_T(x;u(\cdot)) \triangleq \mathbb{E}\int_0^{\infty}[\langle QX(t),X(t)\rangle + 2\langle SX(t),u(t)\rangle + \langle R(t)u(t),u(t)\rangle \\ + 2\langle q,X(t)\rangle + 2\langle r,u(t)\rangle]dt, \tag{1.12}$$

其中, 系数 A, B, C, D, Q, R 和 S 均为常数矩阵, b, σ, q 和 r 均为常数向量. 令 $(\bar{X}_T(\cdot),\bar{u}_T(\cdot))$ 为问题 $(\mathrm{SLQ})_T$ 关于初始状态 x 的最优对, 并分别令 $\mu_T(t;x)$ 和 $\nu_T(t;x)$ 为 $\bar{X}_T(t)$ 和 $\bar{u}_T(t)$ 的概率分布. 本文将阐述分布意义下的指数 turnpike 性质, 即存在唯一的与初始状态 x 和控制时长 T 无关的概率分布 μ^* 和 ν^* (分别定义在 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 上) 以及与 x 和 T 无关的常数 $K,\lambda>0$, 使得

$$d(\mu_T(t;x),\mu^*) + d(\nu_T(t;x),\nu^*) \leqslant K(|x|^2+1)[\mathrm{e}^{-\lambda t}+\mathrm{e}^{-\lambda(T-t)}], \quad \forall\, t\in[0,T],$$

其中, $d(\cdot\,,\cdot)$ 为概率分布的 Wasserstein 度量.

本文将聚焦于随机 LQ 最优控制问题. 考虑到该领域的发展庞大且内容广泛, 本文并未意图全面覆盖所有相关的研究进展, 而是重点总结了作者及其合作者们的近期工作成果.

本文的余下内容将按如下方式展开. 第 2 节从开环和闭环两个角度介绍有限时区上的随机 LQ 最优控制问题的可解性. 第 3 节深入探讨随机 LQ 最优控制问题在无穷控制区间上的开环和闭环可解性. 第 4 节进一步讨论随机 LQ 最优控制问题的 turnpike 性质.

2　有限时区上的随机线性二次最优控制问题

本节讨论有限时区上的随机 LQ 最优控制问题. 我们将介绍该问题的开环和闭环可解性, 讨论它们的关系, 并分别给出相应的刻画准则.

首先, 给出开环可解性的定义.

定义 2.1　若存在 $\bar{u}(\cdot)\in\mathscr{U}[0,T]$ 使等式 (1.3) 成立, 则称问题 $(\mathrm{SLQ})_T$ 在初始状态 $x\in\mathbb{R}^n$ 下开环可解. 此时, 称 $\bar{u}(\cdot)$ 为对应于初始状态 x 的开环最优控制, 称 $\bar{X}(\cdot)\equiv X(\cdot\,;x,\bar{u}(\cdot))$ 和 $(\bar{X}(\cdot),\bar{u}(\cdot))$ 为相应的开环最优状态过程和开环最优对. 若对任意的初始状态 $x\in\mathbb{R}^n$, 问题 $(\mathrm{SLQ})_T$ 均在 x 下开环可解, 则称问题 $(\mathrm{SLQ})_T$ 是开环可解的.

由于状态方程的漂移项和扩散项是状态—控制组合 $(X(\cdot),u(\cdot))$ 的仿射函数, 并且代价泛函由 $(X(\cdot),u(\cdot))$ 的二次型和线性项组成, 故在假设 (H1) 下, 代价泛函具有如下泛函表示:

$$J_T(x;u(\cdot)) = \langle \mathcal{M}u(\cdot),u(\cdot)\rangle + 2\langle \mathcal{L}(x),u(\cdot)\rangle + \mathcal{N}(x), \tag{2.1}$$

其中, $\mathcal{M}:\mathscr{U}[0,T]\to\mathscr{U}[0,T]$ 为有界自伴算子, $\mathcal{L}(x)\in\mathscr{U}[0,T]$, $\mathcal{N}(x)$ 为依赖 x 的标量. 在 (2.1) 中, 我们用之前的记号 $\langle\cdot\,,\cdot\rangle$ 表示 Hilbert 空间 $\mathscr{U}[0,T]$ 中的内积. 下文用 $\mathcal{I}$ 表示从 $\mathscr{U}[0,T]$ 映射到其自身的恒等算子.

利用上述泛函表示, 我们对问题 $(\mathrm{SLQ})_T$ 的开环可解性有如下刻画, 具体的证明过程可参见文献 [25, 36].

命题 2.1　假设 (H1) 成立. 问题 $(\mathrm{SLQ})_T$ 在初始状态 $x\in\mathbb{R}^n$ 下开环可解当且仅当

$$\mathcal{M}\geqslant 0, \quad \mathcal{L}(x)\in\mathscr{R}(\mathcal{M}). \tag{2.2}$$

换言之, 问题 $(\mathrm{SLQ})_T$ 是开环可解的当且仅当 $\mathcal{M}$ 是一个非负算子, 并且 $\mathcal{L}(x)$ 在 $\mathcal{M}$ 的值域中. 当 (2.2) 成立时, $\bar{u}(\cdot) \in \mathscr{U}[0,T]$ 是对应于初始状态 x 的开环最优控制当且仅当

$$\mathcal{M}\bar{u}(\cdot) + \mathcal{L}(x) = 0.$$

另外, 若存在常数 $\delta > 0$ 使得

$$\mathcal{M} - \delta\mathcal{I} \geqslant 0, \tag{2.3}$$

则对任一初始状态 $x \in \mathbb{R}^n$, 均存在唯一的开环最优控制 $\bar{u}(\cdot)$, 它由下式给出:

$$\bar{u}(\cdot) = -\mathcal{M}^{-1}\mathcal{L}(x).$$

条件 (2.3) 等价于齐次随机 LQ 最优控制问题的代价泛函 $J_T^0(0; u(\cdot))$ 在初始状态 x 为 0 时关于控制 $u(\cdot)$ 的一致凸性, 即对于某个常数 $\delta > 0$, $J_T^0(0; u(\cdot))$ 满足

$$J_T^0(0; u(\cdot)) \geqslant \delta\, \mathbb{E}\int_0^T |u(t)|^2 dt, \quad \forall\, u(\cdot) \in \mathscr{U}[0,T]. \tag{2.4}$$

将命题 2.1 转化为随机微分方程的语言, 便得到问题 $(\mathrm{SLQ})_T$ 的开环可解性的如下刻画.

定理 2.1 (参见文献 [36, 定理 2.3.2]) 假设 (H1) 成立. 问题 $(\mathrm{SLQ})_T$ 在初始状态 $x \in \mathbb{R}^n$ 下开环可解当且仅当以下两个条件成立:

(i) $u(\cdot) \mapsto J_T^0(0; u(\cdot))$ 为凸映射, 即

$$J_T^0(0; u(\cdot)) \geqslant 0, \quad \forall\, u(\cdot) \in \mathscr{U}[0,T];$$

(ii) 最优性系统

$$\begin{cases} d\bar{X}(t) = [A(t)\bar{X}(t) + B(t)\bar{u}(t) + b(t)]dt + [C(t)\bar{X}(t) + D(t)\bar{u}(t) + \sigma(t)]dW(t), \\ d\bar{Y}(t) = -[A(t)^\top\bar{Y}(t) + C(t)^\top\bar{Z}(t) + Q(t)\bar{X}(t) + S(t)^\top\bar{u}(t) + q(t)]dt + \bar{Z}(t)dW(t), \\ \bar{X}(0) = x, \quad \bar{Y}(T) = G\bar{X}(T) + g, \\ B(t)^\top\bar{Y}(t) + D(t)^\top\bar{Z}(t) + S(t)\bar{X}(t) + R(t)\bar{u}(t) + r(t) = 0 \end{cases} \tag{2.5}$$

存在一个适定解 $(\bar{X}(\cdot), \bar{u}(\cdot), \bar{Y}(\cdot), \bar{Z}(\cdot))$.

在这种情况下, $(\bar{X}(\cdot), \bar{u}(\cdot))$ 为一个开环最优对.

系统 (2.5) 中的第二个方程为倒向随机微分方程 (backward stochastic differential equation, BSDE), 其适应解是关于 $\mathbb{F}$ 循序可测的二元组随机过程 $(\bar{Y}(\cdot), \bar{Z}(\cdot))$. 该二元组被称作 $(\bar{X}(\cdot), \bar{u}(\cdot))$ 的伴随过程. 系统 (2.5) 是一个正倒向随机微分方程 (forward-backward stochastic differential equation, FBSDE), 通过最后一个等式实现耦合. 系统 (2.5) 最后的等式也被称作平稳条件.

定理 2.1(ii) 实际上是 Pontryagin 最大值原理在随机 LQ 最优控制问题中的表现形式, 它是最优控制满足的一个必要条件. 在 LQ 框架下, 当凸性条件 (i) 成立时, 它同样也是一个充分条件. 定理 2.1 虽然刻画了问题 $(\mathrm{SLQ})_T$ 的开环可解性, 但并没有说明如何构造开环最优控制.

接下来, 给出一个有趣的例子.

例 2.1　考虑受控的 SDE

$$\begin{cases} dX(t)=u(t)dW(t), \quad t\in[0,T], \\ X(0)=x, \end{cases}$$

以及代价泛函

$$J_{T}(x;u(\cdot))\triangleq\mathbb{E}\bigg[2|X(T)|^2-\int_0^T|u(t)|^2dt\bigg].$$

显然,

$$\mathbb{E}|X(t)|^2=\mathbb{E}\bigg[|x|^2+\int_0^t|u(s)|^2ds\bigg], \quad t\in[0,T].$$

所以,

$$J_{T}(x;u(\cdot))=\mathbb{E}\bigg[\int_0^T|u(t)|^2dt+2|x|^2\bigg].$$

因此, 对应的随机 LQ 问题有唯一的开环最优控制 $\bar{u}(\cdot)=0$. 然而, 在这个例子中, $R(\cdot)=-1<0$. 这意味着与确定性 LQ 问题不同, 在随机情形, $R(\cdot)\geqslant 0$ 并不是开环最优控制存在的必要条件.

考虑开环最优控制的构造问题. 假设对所有的 $t\geqslant 0$, $R(t)^{-1}$ 均存在, 则开环最优控制可由下式给出:

$$\bar{u}(t)=-R(t)^{-1}[B(t)^{\top}\bar{Y}(t)+D(t)^{\top}\bar{Z}(t)+S(t)\bar{X}(t)+r(t)], \quad t\in[0,T].$$

这意味着, 为了确定控制 $\bar{u}(\cdot)$ 在时刻 t 的值 $\bar{u}(t)$, 除了开环最优过程 $\bar{X}(\cdot)$ 的值 $\bar{X}(t)$ 以外, 还需要知道系统 (2.5) 中 BSDE 的适应解 $(\bar{Y}(\cdot),\bar{Z}(\cdot))$ 的值 $(\bar{Y}(t),\bar{Z}(t))$. 而 $(\bar{Y}(t),\bar{Z}(t))$ 由 $\bar{X}(\cdot)$ 的未来值 $\bar{X}(T)$ 决定. 因此, 上述结论虽有其数学意义, 但缺乏实际应用价值.

为了构造更为 “实用” 的最优控制, 令 $\boldsymbol{\Theta}[0,T]=L^2(0,T;\mathbb{R}^{m\times n})$, 并引入以下定义.

定义 2.2　假设问题 $(\mathrm{SLQ})_T$ 开环可解. 若存在独立于初始状态 x 的二元组 $(\Theta(\cdot),v(\cdot))\in\boldsymbol{\Theta}[0,T]\times\mathscr{U}[0,T]$, 使得对应于初始状态 x 的开环最优控制 $\bar{u}(\cdot\,;x)$ 具有如下形式:

$$\bar{u}(t;x)=\Theta(t)\bar{X}(t)+v(t), \quad t\in[0,T],$$

其中, $\bar{X}(\cdot)$ 是闭环系统

$$\begin{cases} d\bar{X}(t)=\{[A(t)+B(t)\Theta(t)]\bar{X}(t)+B(t)v(t)+b(t)\}dt \\ \qquad\qquad +\{[C(t)+D(t)\Theta(t)]\bar{X}(t)+D(t)v(t)+\sigma(t)\}dW(t), \quad t\in[0,T], \\ \bar{X}(0)=x \end{cases}$$

的解, 则称开环最优控制 $\bar{u}(\cdot\,;x)$ 具有一个闭环表示 (也称作状态反馈表示).

为了从开环最优控制的刻画 (定理 2.1) 中探求闭环表示的可能性, 我们尝试对最优性系统 (2.5) 作如下假设:

$$\bar{Y}(t)=P(t)\bar{X}(t)+\eta(t), \quad t\in[0,T], \tag{2.6}$$

其中, $P(\cdot)$ 为某个取值于 $\mathbb{S}^n$ 的可微确定性函数, $\eta(\cdot)$ 为以下具有待定系数 $\gamma(\cdot)$ 的 ODE 的解:

$$d\eta(t)=\gamma(t)dt, \quad \eta(T)=g.$$

同时, 我们希望 $\bar{Z}(\cdot)$ 也能类似地利用 $\bar{X}(\cdot)$ 的非预期形式表示, 即在不使用 $\bar{X}(\cdot)$ 的未来值的情况下确定 $\bar{Z}(\cdot)$ 在时刻 t 的值 $\bar{Z}(t)$, 从而原则上能够实际确定 $\bar{u}(t)$. 这种方法在文献 [6] 中被称为不变嵌入, 也曾被 Ma 等 [23] 及 Ma 和 Yong [24] 用于解耦一般的 FBSDE. 若 (2.6) 成立, 则由 Itô 公式可得 (为书写简便, 以下省略时间变量 t)

$$\begin{aligned} &-(A^\top \bar{Y} + C^\top \bar{Z} + Q\bar{X} + S^\top \bar{u} + q)dt + \bar{Z}dW \\ &\quad = d\bar{Y} \\ &\quad = [\dot{P}\bar{X} + P(A\bar{X} + B\bar{u} + b) + \gamma]dt + P(C\bar{X} + D\bar{u} + \sigma)dW. \end{aligned} \tag{2.7}$$

比较上述扩散项可得

$$\bar{Z} = P(C\bar{X} + D\bar{u} + \sigma). \tag{2.8}$$

将 (2.6) 和 (2.8) 代入系统 (2.5) 中的平稳条件, 可得

$$\begin{aligned} 0 &= B^\top \bar{Y} + D^\top \bar{Z} + S\bar{X} + R\bar{u} + r \\ &= B^\top (P\bar{X} + \eta) + D^\top P(C\bar{X} + D\bar{u} + \sigma) + S\bar{X} + R\bar{u} + r \\ &= (B^\top P + D^\top PC + S)\bar{X} + (R + D^\top PD)\bar{u} + B^\top \eta + D^\top P\sigma + r. \end{aligned}$$

假定对于某个常数 $\delta > 0$, 下式成立:

$$R(t) + D(t)^\top P(t)D(t) \geqslant \delta I_m, \quad \text{a.e. } t \in [0, T]. \tag{2.9}$$

从而 $R + D^\top PD$ 可逆, 并有

$$\bar{u} = -(R + D^\top PD)^{-1}[(B^\top P + D^\top PC + S)\bar{X} + B^\top \eta + D^\top P\sigma + r]. \tag{2.10}$$

接下来, 通过比较等式 (2.7) 等号两边的漂移项, 并将 (2.6) 和 (2.8) 代入可得

$$\begin{aligned} 0 &= A^\top \bar{Y} + C^\top \bar{Z} + Q\bar{X} + S^\top \bar{u} + q + \dot{P}\bar{X} + P(A\bar{X} + B\bar{u} + b) + \gamma \\ &= (\dot{P} + PA + A^\top P + C^\top PC + Q)\bar{X} + (PB + C^\top PD + S^\top)\bar{u} \\ &\quad + A^\top \eta + C^\top P\sigma + Pb + q + \gamma. \end{aligned}$$

将 (2.10) 代入到上式中, 立得

$$\begin{aligned} 0 = &[\dot{P} + PA + A^\top P + C^\top PC + Q - (PB + C^\top PD + S^\top)(R + D^\top PD)^{-1}(B^\top P + D^\top PC + S)]\bar{X} \\ &+ (A + B\Theta)^\top \eta + (C + D\Theta)^\top P\sigma + \Theta^\top r + Pb + q + \gamma, \end{aligned}$$

其中,

$$\Theta \triangleq -(R + D^\top PD)^{-1}(B^\top P + D^\top PC + S).$$

因此, $P(\cdot)$ 应满足微分 Riccati 方程

$$\begin{cases} \dot{P} + PA + A^\top P + C^\top PC + Q \\ \quad -(PB + C^\top PD + S^\top)(R + D^\top PD)^{-1}(B^\top P + D^\top PC + S) = 0, \quad t \in [0, T], \\ P(T) = G, \end{cases} \tag{2.11}$$

$\eta(\cdot)$ 应满足 ODE

$$\begin{cases}\dot{\eta}(t)+(A+B\Theta)^\top\eta+(C+D\Theta)^\top P\sigma+\Theta^\top r+Pb+q=0, \quad t\in[0,T],\\ \eta(T)=g.\end{cases} \tag{2.12}$$

综合上述分析, 不难证明如下结论:

命题 2.2 设 (H1) 成立. 若问题 $(\mathrm{SLQ})_T$ 开环可解, 并且微分 Riccati 方程 (2.11) 有满足 (2.9) 的解 $P(\cdot)$ (从而 ODE (2.12) 唯一可解), 则开环最优控制 $\bar{u}(\cdot)$ 具有闭环表示 (2.10).

上述结论提供了一个重新思考问题 $(\mathrm{SLQ})_T$ 的角度, 下面就此展开讨论. 考虑如下的 SDE:

$$\begin{cases}dX(t)=\{[A(t)+B(t)\Theta(t)]X(t)+B(t)v(t)+b(t)\}dt\\ \qquad\qquad +\{[C(t)+D(t)\Theta(t)]X(t)+D(t)v(t)+\sigma(t)\}dW(t),\\ X(0)=x,\end{cases}$$

其中, $(\Theta(\cdot),v(\cdot))\in\boldsymbol{\Theta}[0,T]\times\mathscr{U}[0,T]$ 被称为闭环策略. 称上述系统为闭环策略 $(\Theta(\cdot),v(\cdot))$ 下的闭环系统. 在 (H1) 成立的条件下, 该闭环系统存在唯一解 $X_x^{\Theta,v}(\cdot)\equiv X(\cdot\,;x,\Theta(\cdot),v(\cdot))$. 称

$$u(\cdot)\triangleq\Theta(\cdot)X_x^{\Theta,v}(\cdot)+v(\cdot)$$

为闭环策略 $(\Theta(\cdot),v(\cdot))$ 在初始状态 x 下的输出控制. 对应于这一输出控制的代价泛函可展开成如下形式:

$$\begin{aligned}J_T(x;\Theta(\cdot)X_x^{\Theta,v}(\cdot)+v(\cdot))&\triangleq J_T(x;u(\cdot))\\ &=\mathbb{E}\bigg\{\int_0^T[\langle Q(t)X_x^{\Theta,v}(t),X_x^{\Theta,v}(t)\rangle+2\langle S(t)X_x^{\Theta,v}(t),\Theta(t)X_x^{\Theta,v}(t)+v(t)\rangle\\ &\quad+\langle R(t)[\Theta(t)X_x^{\Theta,v}(t)+v(t)],\Theta(t)X_x^{\Theta,v}(t)+v(t)\rangle+2\langle q(t),X_x^{\Theta,v}(t)\rangle\\ &\quad+2\langle r(t),\Theta(t)X_x^{\Theta,v}(t)+v(t)\rangle]dt+\langle GX_x^{\Theta,v}(T),X_x^{\Theta,v}(T)\rangle+2\langle g,X_x^{\Theta,v}(T)\rangle\bigg\}.\end{aligned}$$

定义 2.3 若闭环策略 $(\bar{\Theta}(\cdot),\bar{v}(\cdot))\in\boldsymbol{\Theta}[0,T]\times\mathscr{U}[0,T]$ 满足

$$\begin{gathered}J_T(x;\bar{\Theta}(\cdot)X_x^{\bar{\Theta},\bar{v}}(\cdot)+\bar{v}(\cdot))\leqslant J_T(x;\Theta(\cdot)X_x^{\Theta,v}(\cdot)+v(\cdot)),\\ \forall\,(\Theta(\cdot),v(\cdot))\in\boldsymbol{\Theta}[0,T]\times\mathscr{U}[0,T],\quad\forall\,x\in\mathbb{R}^n,\end{gathered}$$

则称其为问题 $(\mathrm{SLQ})_T$ 的一个闭环最优策略. 此时, 称问题 $(\mathrm{SLQ})_T$ 闭环可解. 若 $(\bar{\Theta}(\cdot),\bar{v}(\cdot))$ 存在且唯一, 则称问题 $(\mathrm{SLQ})_T$ 唯一闭环可解.

注 2.1 闭环最优策略 $(\bar{\Theta}(\cdot),\bar{v}(\cdot))$ 与初始状态 x 无关.

注 2.2 Başar 在其早期研究 (如 20 世纪 70 年代的工作 [2,4] 及 1998 年专著 [3]) 中定义了开环解、闭环解和反馈解, 这些定义适用于广泛的控制模型. 尽管本文中的闭环解定义与 Başar 的广义定义在形式上有所不同, 但这主要源于模型的差异. 更多细节, 可以参考 Başar 的相关文献.

关于闭环可解性, 有以下等价命题.

命题 2.3 (参见文献 [36, 命题 2.1.5]) 设 (H1) 成立, 并令 $(\bar{\Theta}(\cdot),\bar{v}(\cdot))\in\boldsymbol{\Theta}[0,T]\times\mathscr{U}[0,T]$, 则下列命题等价.

(i) $(\bar{\Theta}(\cdot),\bar{v}(\cdot))$ 是问题 $(\mathrm{SLQ})_T$ 的一个闭环最优策略.

(ii) 对任意的 $x\in\mathbb{R}^n$ 和 $v(\cdot)\in\mathscr{U}[0,T]$, 下式均成立:

$$J_T(x;\bar{\Theta}(\cdot)X_x^{\bar{\Theta},\bar{v}}(\cdot)+\bar{v}(\cdot))\leqslant J_T(x;\bar{\Theta}(\cdot)X_x^{\bar{\Theta},v}(\cdot)+v(\cdot)).$$

(iii) 对任意的 $x\in\mathbb{R}^n$ 和 $u(\cdot)\in\mathscr{U}[0,T]$, 下式均成立:

$$J_T(x;\bar{\Theta}(\cdot)X_x^{\bar{\Theta},\bar{v}}(\cdot)+\bar{v}(\cdot))\leqslant J_T(x;u(\cdot)).$$

根据上述命题不难看出, 问题 $(\mathrm{SLQ})_T$ 的闭环最优策略 $(\bar{\Theta}(\cdot),\bar{v}(\cdot))$ 具有以下性质.

(1) 考虑下述随机 LQ 最优控制问题. 状态方程为

$$\begin{cases}dX(t)=[A^{\bar{\Theta}}(t)X(t)+B(t)v(t)+b(t)]dt+[C^{\bar{\Theta}}(t)X(t)+D(t)v(t)+\sigma(t)]dW(t),\\X(0)=x,\end{cases}\tag{2.13}$$

代价泛函为

$$\begin{aligned}J_T^{\bar{\Theta}}(x;v(\cdot))\triangleq\mathbb{E}\bigg\{\int_0^T&[\langle Q^{\bar{\Theta}}(t)X(t),X(t)\rangle+2\langle S^{\bar{\Theta}}(t)X(t),v(t)\rangle+\langle R(t)v(t),v(t)\rangle\\&+2\langle q^{\bar{\Theta}}(t),X(t)\rangle+2\langle r(t),v(t)\rangle]dt+\langle GX(T),X(T)\rangle+2\langle g,X(T)\rangle\bigg\},\end{aligned}\tag{2.14}$$

其中,

$$\begin{cases}A^{\bar{\Theta}}(t)\triangleq A(t)+B(t)\bar{\Theta}(t),\quad C^{\bar{\Theta}}(t)\triangleq C(t)+D(t)\bar{\Theta}(t),\\Q^{\bar{\Theta}}(t)\triangleq Q(t)+\bar{\Theta}(t)^\top S(t)+S(t)^\top\bar{\Theta}(t)+\bar{\Theta}(t)^\top R(t)\bar{\Theta}(t),\\S^{\bar{\Theta}}(t)\triangleq S(t)+R(t)\bar{\Theta}(t),\quad q^{\bar{\Theta}}(t)\triangleq q(t)+\bar{\Theta}(t)r(t).\end{cases}$$

如果 $(\bar{\Theta}(\cdot),\bar{v}(\cdot))$ 是问题 $(\mathrm{SLQ})_T$ 的闭环最优策略, 则由命题 2.3(ii) 可知, $\bar{v}(\cdot)$ 是上述随机 LQ 最优控制问题的开环最优控制.

(2) 命题 2.3(iii) 表明, 输出控制 $\bar{u}(\cdot)\triangleq\bar{\Theta}(\cdot)X_x^{\bar{\Theta},\bar{v}}(\cdot)+\bar{v}(\cdot)$ 是问题 $(\mathrm{SLQ})_T$ 在初始状态 x 下的开环最优控制.

因此, 有如下推论.

推论 2.1 设 (H1) 成立. 若问题 $(\mathrm{SLQ})_T$ 闭环可解, 则它也是开环可解的.

上述结论的逆命题在通常情况下并不成立. 下面给出一个例子.

例 2.2 考虑一维受控 SDE

$$\begin{cases}dX(t)=u_1(t)dt+u_2(t)dW(t),\quad t\in[0,T],\\X(0)=x,\end{cases}$$

以及代价泛函

$$J_T(x;u_1(\cdot),u_2(\cdot))\triangleq\mathbb{E}\bigg[\int_0^T|u_2(t)|^2dt+|X(T)|^2\bigg].$$

此处, $(u_1(\cdot),u_2(\cdot))^\top$ 为取值在 $\mathbb{R}^2$ 上的控制过程. 这一随机 LQ 问题是开环可解的. 事实上, 对任意的 $x\in\mathbb{R}$ 和任意的 $0<\delta<T$, 若令

$$\bar{u}_1^\delta(t)\triangleq-\frac{x}{\delta}\mathbf{1}_{[0,\delta]}(t),\quad\bar{u}_2(t)\triangleq0,\quad t\in[0,T],$$

则 $X(T)=0$, 且

$$J_T(x;\bar u_1^\delta(\cdot),\bar u_2(\cdot))=0=\inf_{(u_1(\cdot),u_2(\cdot))\in\mathscr{U}[0,T]}J_T(x;u_1(\cdot),u_2(\cdot)).$$

因此, 这样的控制是开环最优控制. 但该问题并非闭环可解. 若不然, 设

$$\bar\Theta(\cdot)\triangleq\begin{pmatrix}\bar\Theta_1(\cdot)\\ \bar\Theta_2(\cdot)\end{pmatrix},\quad \bar v(\cdot)\triangleq\begin{pmatrix}\bar v_1(\cdot)\\ \bar v_2(\cdot)\end{pmatrix}$$

为一个闭环最优策略. 则闭环系统满足 (令 $\bar X(\cdot)\triangleq X_x^{\bar\Theta,\bar v_1,\bar v_2}(\cdot)$)

$$\begin{cases}d\bar X(t)=[\bar\Theta_1(t)\bar X(t)+\bar v_1(t)]dt+[\bar\Theta_2(t)\bar X(t)+\bar v_2(t)]dW(t),\\ \bar X(0)=x.\end{cases}$$

由于输出控制

$$\bar u_1(t)\triangleq\bar\Theta_1(t)\bar X(t)+\bar v_1(t),\quad \bar u_2(t)\triangleq\bar\Theta_2(t)\bar X(t)+\bar v_2(t)$$

为开环最优控制, 故必有

$$\bar\Theta_2(t)\bar X(t)+\bar v_2(t)=0,\quad \bar X(T)=0.$$

继而,

$$d\bar X(t)=[\bar\Theta_1(t)\bar X(t)+\bar v_1(t)]dt,\quad \bar X(0)=x,\quad \bar X(T)=0.$$

因此, 根据闭环最优策略的定义, 下式对所有的 $x\in\mathbb{R}$ 成立:

$$0=\bar X(T)=x\mathrm{e}^{\int_0^T\bar\Theta(\tau)d\tau}+\int_0^T\mathrm{e}^{\int_s^T\bar\Theta(\tau)d\tau}\bar v_1(s)ds.$$

这显然是不可能的.

我们的下一个目标是刻画闭环最优策略. 为此, 先假设 $(\bar\Theta(\cdot),\bar v(\cdot))$ 是一个闭环最优策略. 由命题 2.3(ii) 可知, $\bar v(\cdot)$ 是随机 LQ 问题 (2.13)–(2.14) 的开环最优控制. 再由定理 2.1(ii) 可知 (记 $\bar X(\cdot)\triangleq X_x^{\bar\Theta,\bar v}(\cdot)$)

$$\begin{cases}d\bar X(t)=[A^{\bar\Theta}(t)\bar X(t)+B(t)\bar v(t)+b(t)]dt+[C^{\bar\Theta}(t)\bar X(t)+D(t)\bar v(t)+\sigma(t)]dW(t),\\ d\bar Y(t)=-[A^{\bar\Theta}(t)^\top\bar Y(t)+C^{\bar\Theta}(t)^\top\bar Z(t)+Q^{\bar\Theta}(t)\bar X(t)+S^{\bar\Theta}(t)^\top\bar v(t)+q^{\bar\Theta}(t)]dt+\bar Z(t)dW(t),\\ \bar X(0)=x,\quad \bar Y(T)=G\bar X(T)+g,\\ B(t)^\top\bar Y(t)+D(t)^\top\bar Z(t)+S^{\bar\Theta}(t)\bar X(t)+R(t)\bar v(t)+r(t)=0.\end{cases}$$

为了强调对初始状态 x 的依赖性, 记 $(X^x(\cdot),Y^x(\cdot),Z^x(\cdot))$ 为上述系统的适应解. 令

$$\hat X(\cdot)\triangleq X^x(\cdot)-X^0(\cdot),\quad \hat Y(\cdot)\triangleq Y^x(\cdot)-Y^0(\cdot),\quad \hat Z(\cdot)\triangleq Z^x(\cdot)-Z^0(\cdot).$$

于是,

$$\begin{cases}d\hat X(t)=A^{\bar\Theta}(t)\hat X(t)dt+C^{\bar\Theta}(t)\hat X(t)dW(t),\\ d\hat Y(t)=-[A^{\bar\Theta}(t)^\top\hat Y(t)+C^{\bar\Theta}(t)^\top\hat Z(t)+Q^{\bar\Theta}(t)\hat X(t)]dt+\hat Z(t)dW(t),\\ \hat X(0)=x,\quad \hat Y(T)=G\hat X(T),\\ B(t)^\top\hat Y(t)+D(t)^\top\hat Z(t)+S^{\bar\Theta}(t)\hat X(t)=0.\end{cases}$$

令 x 的取值跑遍整个 $\mathbb{R}^n$, 则下列矩阵值系统具有适应解 $(\mathbb{X}(\cdot),\mathbb{Y}(\cdot),\mathbb{Z}(\cdot))$:

$$\begin{cases} d\mathbb{X}(t)=A^{\bar\Theta}(t)\mathbb{X}(t)dt+C^{\bar\Theta}(t)\mathbb{X}(t)dW(t),\\ d\mathbb{Y}(t)=-[A^{\bar\Theta}(t)^\top\mathbb{Y}(t)+C^{\bar\Theta}(t)^\top\mathbb{Z}(t)+Q^{\bar\Theta}(t)\mathbb{X}(t)]dt+\mathbb{Z}(t)dW(t),\\ \mathbb{X}(0)=I_n,\quad \mathbb{Y}(T)=G\mathbb{X}(T),\\ B(t)^\top\mathbb{Y}(t)+D(t)^\top\mathbb{Z}(t)+S^{\bar\Theta}(t)\mathbb{X}(t)=0. \end{cases}$$

显然, 对任意的 $t\in[0,T]$, $\mathbb{X}(t)$ 都是可逆的, 并且 $\mathbb{X}(\cdot)^{-1}$ 满足如下 SDE:

$$\begin{cases} d[\mathbb{X}(t)^{-1}]=-\mathbb{X}(t)^{-1}[A^{\bar\Theta}(t)-C^{\bar\Theta}(t)^2]dt-\mathbb{X}(t)^{-1}C^{\bar\Theta}(t)dW(t),\\ \mathbb{X}(0)^{-1}=I_n. \end{cases}$$

定义

$$P(t)\triangleq\mathbb{Y}(t)\mathbb{X}(t)^{-1},\quad \varGamma(t)\triangleq\mathbb{Z}(t)\mathbb{X}(t)^{-1},\quad t\in[0,T].$$

根据 Itô 公式, 有 (为书写简便, 以下省略时间变量 t)

$$\begin{aligned} dP&=\{-[(A^{\bar\Theta})^\top\mathbb{Y}+(C^{\bar\Theta})^\top\mathbb{Z}+Q^{\bar\Theta}\mathbb{X}]\mathbb{X}^{-1}-\mathbb{Y}\mathbb{X}^{-1}[A^{\bar\Theta}-(C^{\bar\Theta})^2]-\mathbb{Z}\mathbb{X}^{-1}C^{\bar\Theta}\}dt\\ &\quad+(\mathbb{Z}\mathbb{X}^{-1}-\mathbb{Y}\mathbb{X}^{-1}C^{\bar\Theta})dW\\ &=-\{(A^{\bar\Theta})^\top P+(C^{\bar\Theta})^\top\varGamma+Q^{\bar\Theta}+P[A^{\bar\Theta}-(C^{\bar\Theta})^2]+\varGamma C^{\bar\Theta}\}dt+(\varGamma-PC^{\bar\Theta})dW. \end{aligned}$$

令 $\varLambda\triangleq\varGamma-PC^{\bar\Theta}$, 则上式转化为

$$dP=-[PA^{\bar\Theta}+(A^{\bar\Theta})^\top P+(C^{\bar\Theta})^\top PC^{\bar\Theta}+\varLambda C^{\bar\Theta}+(C^{\bar\Theta})^\top\varLambda+Q^{\bar\Theta}]dt+\varLambda dW.$$

上式以及终端条件 $P(T)=G$ 构成了一个确定性系数且终端项非随机的 BSDE. 由适应解的唯一性, 必有 $\varLambda(\cdot)=0$. 因此, $P(\cdot)$ 实际上是以下 ODE 的解:

$$\begin{cases} \dot P+PA^{\bar\Theta}+(A^{\bar\Theta})^\top P+(C^{\bar\Theta})^\top PC^{\bar\Theta}+Q^{\bar\Theta}=0,\\ P(T)=G. \end{cases}\tag{2.15}$$

另外, 平稳条件等价于

$$0=B^\top P+D^\top\varGamma+S^{\bar\Theta}=B^\top P+D^\top PC+S+(R+D^\top PD)\bar\Theta.$$

上式意味着

$$\mathscr{R}(B^\top P+D^\top PC+S)\subseteq\mathscr{R}(R+D^\top PD),$$

以及对于某个 $\varPi(\cdot)\in\boldsymbol{\Theta}[0,T]$, 有

$$\bar\Theta=-(R+D^\top PD)^\dagger(B^\top P+D^\top PC+S)+[I_m-(R+D^\top PD)^\dagger(R+D^\top PD)]\varPi.$$

利用这一表达式, 可知

$$PA^{\bar\Theta}+(A^{\bar\Theta})^\top P+(C^{\bar\Theta})^\top PC^{\bar\Theta}+Q^{\bar\Theta}$$

$$
\begin{aligned}
&= PA + A^\top P + C^\top PC + Q + (PB + C^\top PD + S^\top)\bar{\Theta} + \bar{\Theta}^\top(B^\top P + D^\top PC + S)\\
&\quad + \bar{\Theta}^\top(R + D^\top PD)\bar{\Theta}\\
&= PA + A^\top P + C^\top PC + Q - (PB + C^\top PD + S^\top)(R + D^\top PD)^\dagger(B^\top P + D^\top PC + S).
\end{aligned}
$$

从而, (2.15) 转化为以下微分 Riccati 方程 (该方程推广了 (2.11)):

$$
\begin{cases}
\dot{P}(t) + P(t)A(t) + A(t)^\top P(t) + C(t)^\top P(t)C(t) + Q(t)\\
\qquad - [P(t)B(t) + C(t)^\top P(t)D(t) + S(t)^\top][R(t) + D(t)^\top P(t)D(t)]^\dagger\\
\qquad \times [B(t)^\top P(t) + D(t)^\top P(t)C(t) + S(t)] = 0,\\
P(T) = G.
\end{cases}
\tag{2.16}
$$

定义 2.4　如果函数 $P(\cdot) \in C([0,T];\mathbb{S}^n)$ 满足 (2.16) 以及下列条件:

$$
\begin{cases}
R(t) + D(t)^\top P(t)D(t) \geqslant 0, \quad \text{a.e. } t \in [0,T],\\
\mathscr{R}(B(t)^\top P(t) + D(t)^\top P(t)C(t) + S(t)) \subseteq \mathscr{R}(R(t) + D(t)^\top P(t)D(t)), \quad \text{a.e. } t \in [0,T],\\
[R(\cdot) + D(\cdot)^\top P(\cdot)D(\cdot)]^\dagger[B(\cdot)^\top P(\cdot) + D(\cdot)^\top P(\cdot)C(\cdot) + S(\cdot)] \in \boldsymbol{\Theta}[0,T],
\end{cases}
\tag{2.17}
$$

则称其为微分 Riccati 方程 (2.16) 在 $[0,T]$ 上的正则解.

借助上述定义, 介绍一些有关闭环最优策略的重要结论.

定理 2.2 (参见文献 [36, 定理 2.4.3])　设 (H1) 成立. 问题 $(\text{SLQ})_T$ 闭环可解当且仅当以下条件成立:

(i) 微分 Riccati 方程 (2.16) 存在正则解 $P(\cdot) \in C([0,T];\mathbb{S}^n)$;

(ii) ODE

$$
\begin{cases}
\dot{\eta}(t) + [A(t) + B(t)\Theta(t)]^\top \eta(t) + [C(t) + D(t)\Theta(t)]^\top P(t)\sigma(t)\\
\qquad \eta(t) + \Theta(t)^\top r(t) + P(t)b(t) + q(t) = 0,\\
\eta(T) = g
\end{cases}
\tag{2.18}
$$

的解 $\eta(\cdot)$ 满足

$$
\begin{aligned}
&\kappa(t) \triangleq B(t)^\top \eta(t) + D(t)^\top P(t)\sigma(t) + r(t) \in \mathscr{R}(R(t) + D(t)^\top P(t)D(t)),\\
&\nu(\cdot) \triangleq -[R(\cdot) + D(\cdot)^\top P(\cdot)D(\cdot)]^\dagger \kappa(\cdot) \in \mathscr{U}[0,T],
\end{aligned}
$$

其中, 在 (2.18) 中,

$$
\Theta(t) \triangleq -[R(t) + D(t)^\top P(t)D(t)]^\dagger[B(t)^\top P(t) + D(t)^\top P(t)C(t) + S(t)].
$$

进一步地, 当问题 $(\text{SLQ})_T$ 闭环可解时, 下式给出了所有的闭环最优策略 $(\bar{\Theta}(\cdot), \bar{v}(\cdot))$:

$$
\begin{aligned}
&\bar{\Theta}(t) = \Theta(t) + \{I_m - [R(t) + D(t)^\top P(t)D(t)]^\dagger[R(t) + D(t)^\top P(t)D(t)]\}\Pi(t),\\
&\bar{v}(t) = \nu(t) + \{I_m - [R(t) + D(t)^\top P(t)D(t)]^\dagger[R(t) + D(t)^\top P(t)D(t)]\}\pi(t),
\end{aligned}
$$

其中, $(\Pi(\cdot),\pi(\cdot))\in\boldsymbol{\Theta}[0,T]\times\mathscr{U}[0,T]$ 是任意的. 并且, 值函数具有如下表示:

$$V_T(x)=\langle P(0)x,x\rangle+2\langle\eta(0),x\rangle+\int_0^T[\langle P(t)\sigma(t),\sigma(t)\rangle+2\langle\eta(t),b(t)\rangle\\ -\langle[R(t)+D(t)^\top P(t)D(t)]^\dagger\kappa(t),\kappa(t)\rangle]dt.$$

在上述定理中, $\eta(\cdot)$ 满足的 ODE 与前文中的 (2.12) 相同, 可以通过定义

$$\eta(t)\triangleq\bar{Y}(t)-P(t)\bar{X}(t)$$

以及对 $\eta(\cdot)$ 应用 Itô 公式推导得到. 这大体上可以证明定理 2.2 的必要性. 对 $\langle P(\cdot)X(\cdot),X(\cdot)\rangle$ 应用 Itô 公式和配方法可证得定理 2.2 的充分性. 同时, 在证明中可看到, (2.17) 中的第一个条件保证了映射 $u(\cdot)\mapsto J_T^0(0;u(\cdot))$ 的凸性.

定理 2.2 虽然指出了问题 $(\mathrm{SLQ})_T$ 的闭环可解性与微分 Riccati 方程 (2.16) 的正则可解性之间的等价关系, 却没有解决 "何时闭环可解" 这一问题. 下面的结果给出了后者的回答.

定理 2.3 (参见文献 [36, 定理 2.5.6]) 设 (H1) 成立. 以下命题等价:

(i) $u(\cdot)\mapsto J_T^0(0;u(\cdot))$ 为一致凸映射, 换言之, 存在常数 $\delta>0$ 使得不等式 (2.4) 成立;

(ii) 微分 Riccati 方程 (2.16) 在 $[0,T]$ 上存在正则解 $P(\cdot)$, 并且使得不等式 (2.9) 对某个常数 $\delta>0$ 成立.

这里又一次涉及了映射 $u(\cdot)\mapsto J_T^0(0;u(\cdot))$ 的一致凸性. 下面的命题给出了一个使得一致凸性成立的简单条件.

命题 2.4 (参见文献 [36, 命题 2.5.1]) 设对于某个常数 $\delta>0$, 标准条件

$$G\geqslant 0,\quad R(t)\geqslant\delta I_m,\quad Q(t)-S(t)^\top R(t)^{-1}S(t)\geqslant 0,\quad \text{a.e. } t\in[0,T] \tag{2.19}$$

成立, 则映射 $u(\cdot)\mapsto J_T^0(0;u(\cdot))$ 具有一致凸性.

由此, 结合本节之前的讨论, 当条件 (2.19) 成立时, 问题 $(\mathrm{SLQ})_T$ 是闭环可解的, 从而也是开环可解的.

3 无穷时区上的随机线性二次最优控制问题

本节讨论无限时区上的随机 LQ 最优控制问题. 考虑状态系统 (1.6) 和代价泛函 (1.7). 此时的系数 A, B, C, D, Q, S 和 R 均为常值矩阵, 状态方程中的非齐次项 $b(\cdot)$ 和 $\sigma(\cdot)$ 以及代价泛函中的一阶系数 $q(\cdot)$ 和 $r(\cdot)$ 满足假设 (H2).

在假设 (H2) 下, 对任意给定的初始状态 $x\in\mathbb{R}^n$ 和控制 $u(\cdot)\in\mathscr{U}[0,\infty)\triangleq L_{\mathbb{F}}^2(0,\infty;\mathbb{R}^m)$, 状态方程 (1.6) 存在唯一解 $X(\cdot)\equiv X(\cdot\,;x,u(\cdot))$. 该解局部平方可积, 但在 $[0,\infty)$ 上不一定全局平方可积. 因此, $\mathscr{U}[0,\infty)$ 中的控制并非全部是容许的, 即 $u(\cdot)\in\mathscr{U}[0,\infty)$ 不一定能保证 (1.7) 中的被积过程在 $[0,\infty)$ 上可积. 因此, 用 (1.8) 作为问题 $(\mathrm{SLQ})_\infty$ 关于初始状态 x 的容许控制集的定义, 即

$$\mathscr{U}_{ad}(x)\triangleq\left\{u(\cdot)\in\mathscr{U}[0,\infty)\,\middle|\,\mathbb{E}\int_0^\infty|X(t;x,u(\cdot))|^2dt<\infty\right\}.$$

容许控制集的这一定义虽然保证了代价泛函 (1.7) 中的积分有意义, 但 $\mathscr{U}_{ad}(x)$ 的结构并不清晰, 我们甚至无法从上述定义中确定 $\mathscr{U}_{ad}(x)$ 是否非空.

为了进一步了解 $\mathscr{U}_{ad}(x)$, 首先考虑无控制的齐次 SDE (即, 令 (1.6) 中的 $u(\cdot)=b(\cdot)=\sigma(\cdot)=0$):

$$dX(t)=AX(t)dt+CX(t)dW(t),\quad t\geqslant 0.$$

为方便起见, 记上述系统为 $[A,C]$.

定义 3.1　若对任意的初始状态 $x\in\mathbb{R}^n$, 系统 $[A,C]$ 的解 $X(\cdot)\equiv X(\cdot\,;x)$ 都满足

$$\mathbb{E}\int_0^\infty |X(t)|^2dt<\infty,$$

则称该系统为 L^2 稳定的.

下面的结论刻画了系统 $[A,C]$ 的 L^2 稳定性, 其证明可参见文献 [36].

命题 3.1 (参见文献 [36, 定理 3.2.3])　系统 $[A,C]$ 是 L^2 稳定的当且仅当存在 $P\in\mathbb{S}^n_+$ 使得以下 Lyapunov 不等式成立:

$$PA+A^\top P+C^\top PC<0.$$

此时, 对任意的 $\Lambda\in\mathbb{S}^n$, Lyapunov 方程

$$PA+A^\top P+C^\top PC+\Lambda=0$$

存在唯一解 $P\in\mathbb{S}^n$, 该解由 $P=\mathbb{E}\int_0^\infty\Phi(t)^\top\Lambda\Phi(t)dt$ 给出, 其中, $\Phi(\cdot)$ 是下列矩阵值 SDE 的解:

$$\begin{cases}d\Phi(t)=A\Phi(t)dt+C\Phi(t)dW(t),\quad t\geqslant 0,\\ \Phi(0)=I_n.\end{cases}$$

接下来, 进一步考虑非齐次 SDE

$$\begin{cases}dX(t)=[AX(t)+f(t)]dt+[CX(t)+g(t)]dW(t),\quad t\in[0,\infty),\\ X(0)=x,\end{cases}\tag{3.1}$$

以及 BSDE

$$dY(t)=-[A^\top Y(t)+C^\top Z(t)+h(t)]dt+Z(t)dW(t),\quad t\in[0,\infty).\tag{3.2}$$

我们有如下结论.

命题 3.2 (参见文献 [36, 命题 3.2.4])　设系统 $[A,C]$ 是 L^2 稳定的, 则对任意的 $f(\cdot),g(\cdot),h(\cdot)\in L^2_{\mathbb{F}}(0,\infty;\mathbb{R}^n)$, SDE (3.1) 存在唯一解 $X(\cdot)\in L^2_{\mathbb{F}}(0,\infty;\mathbb{R}^n)$, 该解满足

$$\mathbb{E}\int_0^\infty|X(t)|^2dt\leqslant K\bigg\{|x|^2+\mathbb{E}\int_0^\infty[|f(t)|^2+|g(t)|^2]dt\bigg\}.$$

同时, BSDE (3.2) 也存在唯一的 L^2 稳定适应解 $(Y(\cdot),Z(\cdot))$, 该解满足

$$\mathbb{E}\bigg[\sup_{0\leqslant t<\infty}|Y(t)|^2+\int_0^\infty|Z(t)|^2dt\bigg]\leqslant K\mathbb{E}\int_0^\infty|h(t)|^2dt,$$

其中 K 是与 x, $f(\cdot)$, $g(\cdot)$ 和 $h(\cdot)$ 无关的正常数.

注 3.1 BSDE (3.2) 的 L^2 稳定适应解是满足以下条件的一对随机过程 $(Y(\cdot),Z(\cdot))$:

(i) $Y(\cdot),Z(\cdot)\in L^2_{\mathbb{F}}(0,\infty;\mathbb{R}^n)$;

(ii) $(Y(\cdot),Z(\cdot))$ 满足 (3.2) 的积分形式

$$Y(t)=Y(0)-\int_0^t[A^\top Y(s)+C^\top Z(s)+h(s)]ds+\int_0^t Z(s)dW(s),\quad t\geqslant 0.$$

现在考虑受控的齐次 SDE

$$dX(t)=[AX(t)+Bu(t)]dt+[CX(t)+Du(t)]dW(t),\quad t\geqslant 0.$$

为简便起见, 记上述系统为 $[A,C;B,D]$.

定义 3.2 若存在矩阵 $\Theta\in\mathbb{R}^{m\times n}$ 使得无控制系统 $[A+B\Theta,C+D\Theta]$ 是 L^2 稳定的, 则称系统 $[A,C;B,D]$ 是 L^2 能稳的. 此时, 称 Θ 为系统 $[A,C;B,D]$ 的一个稳定器, 并记 $\mathscr{S}[A,C;B,D]$ 为系统 $[A,C;B,D]$ 的所有稳定器构成的集合.

下面的结论给出了容许控制集 $\mathscr{U}_{ad}(x)$ 非空的充分必要条件.

命题 3.3 (参见文献 [36, 定理 3.3.5]) 设 (H2) 成立, 则以下命题等价:

(i) 对任意的 $x\in\mathbb{R}^n$, 有 $\mathscr{U}_{ad}(x)\neq\emptyset$;

(ii) $\mathscr{S}[A,C;B,D]\neq\emptyset$;

(iii) 代数 Riccati 方程 (algebraic Riccati equation, ARE)

$$PA+A^\top P+C^\top PC+I_n-(PB+C^\top PD)(I_m+D^\top PD)^{-1}(B^\top P+D^\top PC)=0$$

存在解 $P\in\mathbb{S}^n_+$.

当 (iii) 成立时,

$$-(I_m+D^\top PD)^{-1}(B^\top P+D^\top PC)\in\mathscr{S}[A,C;B,D],$$

且对任意的 $\Theta\in\mathscr{S}[A,C;B,D]$, 容许控制集具有如下结构:

$$\mathscr{U}_{ad}(x)=\{\Theta X_x^{\Theta,v}(\cdot)+v(\cdot)\mid v(\cdot)\in\mathscr{U}[0,\infty)\},\quad\forall\, x\in\mathbb{R}^n.\tag{3.3}$$

上式中, $X_x^{\Theta,v}(\cdot)$ 为下列闭环系统的解:

$$\begin{cases}dX(t)=[(A+B\Theta)X(t)+Bv(t)+b(t)]dt+[(C+D\Theta)X(t)+Dv(t)+\sigma(t)]dW(t),\\ X(0)=x.\end{cases}$$

从上述命题中可以看出, 当 $\mathscr{U}_{ad}(x)=\emptyset$ 时, 问题 $(\text{SLQ})_\infty$ 无意义. 因此, 我们作出如下假设:

(H3) 系统 $[A,C;B,D]$ 是 L^2 能稳的, 换言之, $\mathscr{S}[A,C;B,D]\neq\emptyset$.

另外, 由 (3.3) 可知, 对任意的 $\Theta_1,\Theta_2\in\mathscr{S}[A,C;B,D]$, 都有

$$\{\Theta_1X_x^{\Theta_1,v}(\cdot)+v(\cdot)\mid v(\cdot)\in\mathscr{U}[0,\infty)\}=\{\Theta_2X_x^{\Theta_2,v}(\cdot)+v(\cdot)\mid v(\cdot)\in\mathscr{U}[0,\infty)\},\quad\forall\, x\in\mathbb{R}^n.$$

因此, 针对容许控制的构造, 只需任取一个方便的稳定器 $\Theta\in\mathscr{S}[A,C;B,D]$ 即可.

现在, 可以效仿问题 $(\text{SLQ})_T$, 给出如下定义.

定义 3.3 (i) 若存在 $\bar{u}(\cdot)\in\mathscr{U}_{ad}(x)$ 使得

$$J_\infty(x;\bar{u}(\cdot))=V_\infty(x),$$

则称问题 $(\mathrm{SLQ})_\infty$ 在初始状态 $x\in\mathbb{R}^n$ 下开环可解. 此时, 称 $\bar{u}(\cdot)$ 为对应于 x 的开环最优控制, 称 $\bar{X}(\cdot)\equiv X(\cdot\,;x,\bar{u}(\cdot))$ 和 $(\bar{X}(\cdot),\bar{u}(\cdot))$ 分别为相应的开环最优状态过程和开环最优对. 若对任意的 $x\in\mathbb{R}^n$, 开环最优控制均存在, 则问题 $(\mathrm{SLQ})_\infty$ 是开环可解的.

(ii) 称 $\mathscr{S}[A,C;B,D]\times\mathscr{U}[0,\infty)$ 中的元素为问题 $(\mathrm{SLQ})_\infty$ 的闭环策略. 若存在闭环策略 $(\bar{\Theta},\bar{v}(\cdot))$ 使得

$$J_\infty(x;\bar{\Theta}X_x^{\bar{\Theta},\bar{v}}(\cdot)+\bar{v}(\cdot))\leqslant J_\infty(x;\Theta X_x^{\Theta,v}(\cdot)+v(\cdot)),\quad \forall\,(x,\Theta,v(\cdot))\in\mathbb{R}^n\times\mathscr{S}[A,C;B,D]\times\mathscr{U}[0,\infty),$$

则称问题 $(\mathrm{SLQ})_\infty$ 闭环可解, 称 $(\bar{\Theta},\bar{v}(\cdot))$ 为问题 $(\mathrm{SLQ})_\infty$ 的一个闭环最优策略.

接下来, 介绍本节的重要结论. 由于接下来的定理证明与定理 2.1 类似, 故仅给出结论, 省略具体证明.

定理 3.1 设 (H2)–(H3) 成立. 设 $\Theta\in\mathscr{S}[A,C;B,D]$, $\bar{v}(\cdot)\in\mathscr{U}[0,\infty)$. 若

$$v(\cdot)\mapsto J_\infty^0(0;\Theta X_0^{\Theta,v}(\cdot)+v(\cdot))$$

为凸映射, 则容许控制

$$\bar{u}(\cdot)\triangleq\Theta\bar{X}_x^{\Theta,\bar{v}}(\cdot)+\bar{v}(\cdot)$$

为初始状态 x 下的开环最优控制当且仅当 BSDE

$$d\bar{Y}(t)=-[A^\top\bar{Y}(t)+C^\top\bar{Z}(t)+Q\bar{X}(t)+S^\top\bar{u}(t)+q(t)]dt+\bar{Z}(t)dW(t)$$

存在 L^2 稳定适应解 $(\bar{Y}(\cdot),\bar{Z}(\cdot))$, 且使得下述平稳条件成立:

$$B^\top\bar{Y}(\cdot)+D^\top\bar{Z}(\cdot)+S\bar{X}(\cdot)+R\bar{u}(\cdot)+r(\cdot)=0.$$

在介绍下一个结论之前, 需要先引入以下定义.

定义 3.4 若矩阵 $P\in\mathbb{S}^n$ 满足 ARE

$$\begin{cases}PA+A^\top P+C^\top PC+Q\\ \quad-(PB+C^\top PD+S^\top)(R+D^\top PD)^\dagger(B^\top P+D^\top PC+S)=0,\\ \mathscr{R}(B^\top P+D^\top PC+S^\top)\subseteq\mathscr{R}(R+D^\top PD),\\ R+D^\top PD\geqslant 0,\end{cases}\tag{3.4}$$

且存在 $\Pi\in\mathbb{R}^{m\times n}$ 使得

$$-(R+D^\top PD)^\dagger(B^\top P+D^\top PC+S)+[I_m-(R+D^\top PD)^\dagger(R+D^\top PD)]\Pi\in\mathscr{S}[A,C;B,D],$$

则称 P 为 ARE (3.4) 的一个稳定化解.

定理 3.2 (参见文献 [33, 定理 4.5]) 设 (H2) 和 (H3) 成立, 则下列命题等价:

(i) 问题 $(\mathrm{SLQ})_\infty$ 开环可解;

(ii) 问题 $(\mathrm{SLQ})_\infty$ 闭环可解;

(iii) ARE (3.4) 存在稳定化解 $P \in \mathbb{S}^n$, 且存在可微的 $\eta(\cdot) \in L^2(0,T;\mathbb{R}^n)$ 使得

$$\begin{aligned}
&\dot{\eta}(t) + [A - B(R + D^\top PD)^\dagger(B^\top P + D^\top PC + S)]^\top \eta(t) \\
&\quad + [C - D^\top(R + D^\top PD)^\dagger(B^\top P + D^\top PC + S)]^\top P\sigma(t) \\
&\quad - (B^\top P + D^\top PC + S)^\top(R + D^\top PD)^\dagger r(t) + Pb(t) + q(t) = 0; \\
&\theta(t) \triangleq B^\top \eta(t) + D^\top P\sigma(t) + r(t) \in \mathscr{R}(R + D^\top PD), \quad \text{a.s. } t \in [0,\infty).
\end{aligned}$$

在上述情形下, 下面的式子给出了所有的闭环最优策略 $(\bar{\Theta}, \bar{v}(\cdot))$:

$$\begin{cases}
\bar{\Theta} = -(R + D^\top PD)^\dagger(B^\top P + D^\top PC + S) + [I_m - (R + D^\top PD)^\dagger(R + D^\top PD)]\Pi, \\
\bar{v}(t) = -(R + D^\top PD)^\dagger\theta(t) + [I_m - (R + D^\top PD)^\dagger(R + D^\top PD)]\pi(t),
\end{cases}$$

其中, $\Pi \in \mathbb{R}^{m\times n}$ 使得 $\bar{\Theta} \in \mathscr{S}[A,C;B,D]$, $\pi(\cdot) \in \mathscr{U}[0,\infty)$ 是任意的. 此外, 对任一给定初始状态 x, 每个开环最优控制 $\bar{u}(\cdot)$ 均可由一个闭环最优策略 $(\bar{\Theta}, \bar{v}(\cdot))$ 闭环表示:

$$\bar{u}(t) = \bar{\Theta} X_x^{\bar{\Theta},\bar{v}}(t) + \bar{v}(t), \quad t \geqslant 0.$$

并且, 值函数具有如下表达式:

$$\begin{aligned}
V_\infty(x) = \langle Px, x\rangle + 2\langle \eta(0), x\rangle + \int_0^\infty [\langle P\sigma(t), \sigma(t)\rangle + 2\langle \eta(t), b(t)\rangle \\
- \langle (R + D^\top PD)^\dagger\theta(t), \theta(t)\rangle] dt.
\end{aligned}$$

与有限时区上的随机 LQ 最优控制问题类似, 上述结论只给出了开/闭环可解性与 ARE 之间的等价关系, 并没有解决问题 $(\text{SLQ})_\infty$ 何时开/闭环可解这一问题. 下面的条件保证了问题 $(\text{SLQ})_\infty$ 的可解性.

(H4) 对某个 $\Theta \in \mathscr{S}[A,C;B,D]$, $v(\cdot) \mapsto J_\infty^0(0;\Theta X_0^{\Theta,v}(\cdot) + v(\cdot))$ 为一致凸映射. 换言之, 存在常数 $\delta > 0$, 使得

$$J_\infty^0(0;\Theta X_0^{\Theta,v}(\cdot) + v(\cdot)) \geqslant \delta\, \mathbb{E}\int_0^\infty |v(t)|^2 dt, \quad \forall\, v(\cdot) \in \mathscr{U}[0,\infty).$$

命题 3.4 设 (H2)–(H4) 成立, 则问题 $(\text{SLQ})_\infty$ 开环可解.

4 随机线性二次最优控制的 turnpike 性质

在前面两节的基础上, 本节进一步讨论问题 $(\text{SLQ})_T$ 的 turnpike 性质. 为了简洁呈现核心思想, 考虑常值系数的状态方程 (1.11) 和代价泛函 (1.12). 本节假设 (H3) 和下述条件成立:

(H5) 代价泛函 (1.12) 中的加权矩阵满足

$$R > 0, \quad Q - S^\top R^{-1} S > 0.$$

此时, 相应的微分 Riccati 方程 (2.16) 可写成

$$\begin{cases}
\dot{P}(t) + P(t)A + A^\top P(t) + C^\top P(t)C + Q \\
\quad -[P(t)B + C^\top P(t)D + S^\top][R + D^\top P(t)D]^{-1}[B^\top P(t) + D^\top P(t)C + S] = 0, \quad t \in [0,T], \\
P(T) = 0.
\end{cases} \tag{4.1}$$

由命题 2.4、定理 2.3 和 2.2 可知, 上述方程存在唯一解 $P(\cdot)\in C([0,T];\mathbb{S}^n)$, 并使得下式成立:

$$R+D^\top P(t)D>0,\quad \forall\, t\in[0,T].$$

进而, 问题 $(\mathrm{SLQ})_T$ 唯一闭环可解 (从而唯一开环可解). 注意, 方程 (4.1) 的解依赖区间长度 T, 为了强调这一点, 将 (4.1) 的解记为 $P_T(\cdot)$.

相应的 ARE (3.4) 可写成

$$\begin{aligned}&PA+A^\top P+C^\top PC+Q\\&\quad-(PB+C^\top PD+S^\top)(R+D^\top PD)^{-1}(B^\top P+D^\top PC+S)=0.\end{aligned}\tag{4.2}$$

类似于有限时区的情形, 可以证明 (参见文献 [17,35]), 在假设 (H3) 和 (H5) 下, 方程 (4.2) 存在唯一解 $P\in\mathbb{S}^n$, 并使得

$$R+D^\top PD>0,\quad \varTheta\triangleq-(R+D^\top PD)^{-1}(B^\top P+D^\top PC+S)\in\mathscr{S}[A,C;B,D].$$

通过对比 (4.1) 和 (4.2), 可以猜测 P 和 $P_T(\cdot)$ 之间或许存在某些联系. 下面的结论阐明了这一点, 并且对 turnpike 性质的建立起着关键的作用.

命题 4.1 (参见文献 [30, 定理 4.1])　设 (H3) 和 (H5) 成立. 存在与 T 无关的常数 $K,\lambda>0$, 使得

$$|P_T(t)-P|\leqslant K\mathrm{e}^{-\lambda(T-t)},\quad \forall\, t\in[0,T].$$

接下来, 介绍与随机 LQ 最优控制相对应的静态优化问题. 与确定性情形类似, 约束空间为

$$\mathscr{V}\triangleq\{(x,u)\in\mathbb{R}^n\times\mathbb{R}^m\mid Ax+Bu+b=0\}.$$

但此时的目标函数与确定性情形具有显著差异. 随机情形的目标函数由下式给出:

$$\begin{aligned}F(x,u)&=\langle Qx,x\rangle+\langle Ru,u\rangle+2\langle Sx,u\rangle+2\langle q,x\rangle+2\langle r,u\rangle\\&\quad+\langle P(Cx+Du+\sigma),Cx+Du+\sigma\rangle.\end{aligned}$$

也就是说, 此时的静态优化问题如下.

问题 (O)　寻找二元组 $(x^*,u^*)\in\mathscr{V}$ 使得

$$F(x^*,u^*)=\min_{(x,u)\in\mathscr{V}}F(x,u)\equiv V.$$

利用 Lagrange 乘子法, 容易得到以下结果.

命题 4.2 (参见文献 [38, 命题 3.1])　设 (H3) 和 (H5) 成立. 问题 (O) 存在唯一解, 并且二元组 (x^*,u^*) 是问题 (O) 的解当且仅当存在 $\lambda^*\in\mathbb{R}^n$ 使得

$$\begin{cases}Ax^*+Bu^*+b=0,\\A^\top\lambda^*+Qx^*+C^\top P(Cx^*+Du^*+\sigma)+S^\top u^*+q=0,\\B^\top\lambda^*+Ru^*+D^\top P(Cx^*+Du^*+\sigma)+Sx^*+r=0.\end{cases}\tag{4.3}$$

令 (x^*, u^*) 是问题 (O) 的解, $\lambda^* \in \mathbb{R}^n$ 是 (4.3) 中的向量. 记

$$\sigma^* \triangleq Cx^* + Du^* + \sigma, \quad \Theta \triangleq -(R + D^\top PD)^{-1}(B^\top P + D^\top PC + S).$$

考虑以下 SDE:

$$dX^*(t) = (A + B\Theta)X^*(t)dt + [(C + D\Theta)X^*(t) + \sigma^*]dW(t), \quad t \in [0, \infty). \tag{4.4}$$

由于 $[A + B\Theta, C + D\Theta]$ 是 L^2 稳定的, 所以存在一个具有有限二阶矩的唯一初始分布 ρ^*, 使得 (4.4) 的解 $X^*(\cdot)$ 具有平稳分布. 换言之, 对于任意的 Borel 集合 $\Gamma \in \mathbb{R}^n$, 下式均成立:

$$\mathbb{P}[X^*(t) \in \Gamma] = \rho^*(\Gamma), \quad \forall\, t \geqslant 0,$$

具体证明详见文献 [37]. 定义

$$\boldsymbol{X}^*(t) \triangleq X^*(t) + x^*, \quad \boldsymbol{u}^*(t) \triangleq \Theta X^*(t) + u^*.$$

显然, $\boldsymbol{X}^*(\cdot)$ 和 $\boldsymbol{u}^*(\cdot)$ 也具有平稳分布. 分别记 $\boldsymbol{X}^*(\cdot)$ 和 $\boldsymbol{u}^*(\cdot)$ 的平稳分布为 μ^* 和 ν^*.

有了上述准备工作, 现在给出本节的主要结果, 即问题 $(\mathrm{SLQ})_T$ 的指数 turnpike 性质.

定理 4.1 (参见文献 [38, 定理 3.2]) 设 (H3) 和 (H5) 成立. 令 $(\bar{X}_T(\cdot), \bar{u}_T(\cdot))$ 为问题 $(\mathrm{SLQ})_T$ 关于初始状态 x 的最优对, 则存在与 T 和 x 无关的常数 $K, \lambda > 0$, 使得

$$\mathbb{E}|\bar{X}_T(t) - \boldsymbol{X}^*(t)|^2 + \mathbb{E}|\bar{u}_T(t) - \boldsymbol{u}^*(t)|^2 \leqslant K(|x|^2 + 1)[\mathrm{e}^{-\lambda t} + \mathrm{e}^{-\lambda(T-t)}], \quad \forall\, t \in [0, T].$$

上述定理有一些直接推论. 下面的第一个推论建立了问题 $(\mathrm{SLQ})_T$ 的积分和均方 turnpike 性质, 第二个推论指出了问题 $(\mathrm{SLQ})_T$ 的值函数在时间平均的意义下收敛到问题 (O) 的最优值.

推论 4.1 (参见文献 [38, 推论 3.3]) 设 (H3) 和 (H5) 成立, 则

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathbb{E}[|\bar{X}_T(t) - \boldsymbol{X}^*(t)|^2 + |\bar{u}_T(t) - \boldsymbol{u}^*(t)|^2]dt = 0,$$
$$\lim_{T\to\infty} \frac{1}{T}\int_0^T [|\mathbb{E}[\bar{X}_T(t)] - x^*|^2 + |\mathbb{E}[\bar{u}_T(t)] - u^*|^2]dt = 0.$$

推论 4.2 (参见文献 [38, 推论 3.4]) 设 (H3) 和 (H5) 成立, 则

$$\lim_{T\to\infty} \frac{1}{T} V_T(x) = V, \quad \forall\, x \in \mathbb{R}^n.$$

最后, 利用定理 4.1 和 Wasserstein 度量的定义, 可以得到分布意义下的指数 turnpike 性质.

定理 4.2 (参见文献 [39, 定理 5.3]) 设 (H3) 和 (H5) 成立. 分别令 $\mu_T(t;x)$ 和 $\nu_T(t;x)$ 为 $\bar{X}_T(t)$ 和 $\bar{u}_T(t)$ 的概率分布, 则存在与 T 和 x 无关的常数 $K, \lambda > 0$, 使得

$$d(\mu_T(t;x), \mu^*) + d(\nu_T(t;x), \nu^*) \leqslant K(|x|^2 + 1)[\mathrm{e}^{-\lambda t} + \mathrm{e}^{-\lambda(T-t)}], \quad \forall\, t \in [0, T],$$

其中, $d(\cdot\,,\cdot)$ 为概率分布的 Wasserstein 度量.

参考文献

1 Anderson B D O, Moore J B. Linear Optimal Control. Englewood Cliffs: Prentice-Hall, 1971

2 Başar T. On the uniqueness of the Nash solution in linear-quadratic differential games. Internat J Game Theory, 1976, 5: 65–90
3 Başar T, Olsder G J. Dynamic Noncooperative Game Theory. Philadelphia: SIAM, 1998
4 Başar T, Selbuz H. Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems. IEEE Trans Automat Control, 1979, 24: 166–179
5 Bellman R, Glicksberg I, Gross O A. Some Aspects of the Mathematical Theory of Control Processes. Santa Monica: Rand Corporation, 1958
6 Bellman R, Kalaba R, Wing G M. Invariant imbedding and the reduction of two-point boundary value problems to initial value problems. Proc Natl Acad Sci USA, 1960, 46: 1646–1649
7 Bensoussan A. Lecture on Stochastic Control. Lecture Notes in Mathematics, vol. 972. Berlin: Springer-Verlag, 1981
8 Breiten T, Pfeiffer L. On the turnpike property and the receding-horizon method for linear-quadratic optimal control problems. SIAM J Control Optim, 2020, 58: 1077–1102
9 Carlson D A, Haurie A B, Leizarowitz A. Infinite Horizon Optimal Control—Deterministic and Stochastic Systems, 2nd ed. Berlin: Springer-Verlag, 1991
10 Chen S, Li X, Zhou X Y. Stochastic linear quadratic regulators with indefinite control weight costs. SIAM J Control Optim, 1998, 36: 1685–1702
11 Damm T, Grüne L, Stieler M, et al. An exponential turnpike theorem for dissipative discrete time optimal control problems. SIAM J Control Optim, 2014, 52: 1935–1957
12 Davis M H A. Linear Estimation and Stochastic Control. London: Chapman and Hall, 1977
13 Dorfman R, Samuelson P A, Solow R M. Linear Programming and Economics Analysis. New York: McGraw-Hill, 1958
14 Faulwasser T, Grüne L. Turnpike properties in optimal control: An overview of discrete-time and continuous-time results. Handb Numer Anal, 2022, 23: 367–400
15 Grüne L, Guglielmi R. Turnpike properties and strict dissipativity for discrete time linear quadratic optimal control problems. SIAM J Control Optim, 2018, 56: 1282–1302
16 Grüne L, Guglielmi R. On the relation between turnpike properties and dissipativity for continuous time linear quadratic optimal control problems. Math Control Relat Fields, 2021, 11: 169–188
17 Huang J, Li X, Yong J. A linear-quadratic optimal control problem for mean-field stochastic differential equations in infinite horizon. Math Control Relat Fields, 2015, 5: 97–139
18 Kalman R E. Contributions to the theory of optimal control. Bol Soc Mat Mexicana, 1960, 5: 102–119
19 Kushner H J. Optimal stochastic control. IRE Trans Auto Control, 1962, 7: 120–122
20 Lee E B, Markus L. Foundations of Optimal Control Theory. New York: John Wiley & Sons, 1967
21 Letov A M. The analytical design of control systems. Automat Remote Control, 1961, 22: 363–372
22 Lou H, Wang W. Turnpike properties of optimal relaxed control problems. ESAIM Control Optim Calc Var, 2019, 25: 74
23 Ma J, Protter P, Yong J. Solving forward-backward stochastic differential equations explicitly—A four-step scheme. Probab Theory Related Fields, 1994, 98: 339–359
24 Ma J, Yong J. Forward-Backward Stochastic Differential Equations and Their Applications. Lecture Notes in Mathematics, vol. 1702. Berlin: Springer-Verlag, 1999
25 Mou L, Yong J. Two-person zero-sum linear quadratic stochastic differential games by a Hilbert space method. J Ind Manag Optim, 2006, 2: 95–117
26 Porretta A, Zuazua E. Long time versus steady state optimal control. SIAM J Control Optim, 2013, 51: 4242–4273
27 Sakamoto N, Zuazua E. The turnpike property in nonlinear optimal control—A geometric approach. Automatica, 2021, 134: 109939
28 Sun J. Mean-field stochastic linear quadratic optimal control problems: Open-loop solvabilities. ESAIM Control Optim Calc Var, 2017, 23: 1099–1127
29 Sun J, Li X, Yong J. Open-loop and closed-loop solvabilities for stochastic linear quadratic optimal control problems. SIAM J Control Optim, 2016, 54: 2274–2308
30 Sun J, Wang H, Yong J. Turnpike properties for stochastic linear-quadratic optimal control problems. Chin Ann Math Ser B, 2022, 43: 999–1022
31 Sun J, Xiong J, Yong J. Indefinite stochastic linear-quadratic optimal control problems with random coefficients: Closed-loop representation of open-loop optimal controls. Ann Appl Probab, 2021, 31: 460–499
32 Sun J, Yong J. Linear quadratic stochastic differential games: Open-loop and closed-loop saddle points. SIAM J Control Optim, 2014, 52: 4082–4121
33 Sun J, Yong J. Stochastic linear quadratic optimal control problems in infinite horizon. Appl Math Optim, 2018, 78: 145–183

34 Sun J, Yong J. Linear-quadratic stochastic two-person nonzero-sum differential games: Open-loop and closed-loop Nash equilibria. Stochastic Process Appl, 2019, 129: 381–418
35 Sun J, Yong J. Stochastic Linear-Quadratic Optimal Control Theory: Differential Games and Mean-Field Problems. Cham: Springer, 2020
36 Sun J, Yong J. Stochastic Linear-Quadratic Optimal Control Theory: Open-Loop and Closed-Loop Solutions. Cham: Springer, 2020
37 Sun J, Yong J. Turnpike properties for stochastic linear-quadratic optimal control problems with periodic coefficients. J Differential Equations, 2024, 400: 189–229
38 Sun J, Yong J. Turnpike properties for mean-field linear-quadratic optimal control problems. SIAM J Control Optim, 2024, 62: 752–775
39 Sun J, Yong J. Long-time behavior of zero-sum linear-quadratic stochastic differential games. arXiv:2406.02089, 2024
40 Trélat E, Zuazua E. The turnpike property in finite-dimensional nonlinear optimal control. J Differential Equations, 2015, 258: 81–114
41 von Neumann J. A model of general economic equilibrium. Rev Econ Stud, 1945, 13: 1–9
42 Willems J C. Least squares stationary optimal control and the algebraic Riccati equation. IEEE Trans Automat Control, 1971, 16: 621–634
43 Wonham W M. On a matrix Riccati equation of stochastic control. SIAM J Control, 1968, 6: 681–697
44 Wonham W M. Linear Multivariable Control: A Geometric Approach, 2nd ed. New York: Springer-Verlag, 1979
45 Yong J, Zhou X Y. Stochastic Controls: Hamiltonian Systems and HJB Equations. New York: Springer-Verlag, 1999
46 Zaslavski A J. Turnpike Properties in the Calculus of Variations and Optimal Control. New York: Springer, 2006
47 Zaslavski A J. Turnpike Theory of Continuous-time Linear Optimal Control Problems. Cham: Springer, 2015
48 Zuazua E. Large time control and turnpike properties for wave equations. Annu Rev Control, 2017, 44: 199–210

Some recent developments in stochastic linear-quadratic optimal control

Jingrui Sun, Lvning Yuan & Jiaqi Zhang

Abstract In this paper, we present an overview of recent developments in stochastic linear-quadratic optimal control, integrating the latest relevant works of the authors and their collaborators. The paper is structured into three closely related sections: stochastic linear-quadratic optimal control in finite time horizons, its counterpart in infinite time horizons, and an examination of the turnpike property within stochastic linear-quadratic optimal control.

Keywords stochastic optimal control, linear-quadratic, Riccati equation, turnpike property

MSC(2020) 49N10, 49N35, 93E15, 93E20

doi: 10.1360/SSM-2024-0217