

# 带Markov跳的离散时间随机控制系统的最大值原理

蔺香运<sup>1†</sup>, 王鑫瑞<sup>1</sup>, 张维海<sup>2</sup>

(1. 山东科技大学 数学与系统科学学院, 山东 青岛 266590; 2. 山东科技大学 电气与自动化工程学院, 山东 青岛 266590)

**摘要:** 本文研究一类同时含有Markov跳过程和乘性噪声的离散时间非线性随机系统的最优控制问题, 给出并证明了相应的最大值原理. 首先, 利用条件期望的平滑性, 通过引入具有适应解的倒向随机差分方程, 给出了带有线性差分方程约束的线性泛函的表示形式, 并利用Riesz定理证明其唯一性. 其次, 对带Markov跳的非线性随机控制系统, 利用针状变分法, 对状态方程进行一阶变分, 获得其变分所满足的线性差分方程. 然后, 在引入Hamilton函数的基础上, 通过一对由倒向随机差分方程刻画的伴随方程, 给出并证明了带有Markov跳的离散时间非线性随机最优控制问题的最大值原理, 并给出该最优控制问题的一个充分条件和相应的Hamilton-Jacobi-Bellman方程. 最后, 通过一个实际例子说明了所提理论的实用性和可行性.

**关键词:** 最大值原理; 最优控制; Markov跳; 倒向随机差分方程; Hamilton-Jacobi-Bellman方程

**引用格式:** 蔺香运, 王鑫瑞, 张维海. 带Markov跳的离散时间随机控制系统的最大值原理. 控制理论与应用, 2024, 41(5): 895 – 904

DOI: 10.7641/CTA.2022.10807

## A maximum principle for optimal control of discrete-time stochastic systems with Markov jump

LIN Xiang-yun<sup>1†</sup>, WANG Xin-rui<sup>1</sup>, ZHANG Wei-hai<sup>2</sup>

(1. College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao Shandong 266590, China;

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

**Abstract:** The maximum principle (MP) of the discrete-time nonlinear stochastic optimal control problem is proved, in which the control systems are driven by both Markov jumps and multiplicative noise. Firstly, based on the adapted solutions of the backward stochastic difference equation, the linear functional with the constraint of a linear difference equation is represented. The Riesz theorem is used to prove the uniqueness of such representation. Secondly, the spike variation method is extended to the nonlinear stochastic difference equation with Markov jumps. The variation equation of such state equation is obtained. Thirdly, by introducing a Hamiltonian function, a necessary condition of the discrete-time nonlinear stochastic optimal control system with Markov jump is obtained. It is proved that the adjoint equation of the maximum principle of the system is a pair of backward stochastic difference equations. Moreover, a sufficient condition is also given and the corresponding Hamilton-Jacobi-Bellman equation is derived. Finally, a practical example is given to illustrate the practicability and feasibility of the proposed theory.

**Key words:** maximum principle; optimal control; Markov jump; backward stochastic difference equations; Hamilton-Jacobi-Bellman equations

**Citation:** LIN Xiangyun, WANG Xinrui, ZHANG Weihai. A maximum principle for optimal control of discrete-time stochastic systems with Markov jump. *Control Theory & Applications*, 2024, 41(5): 895 – 904

## 1 引言

最大值原理是Pontryagin等<sup>[1]</sup>在20世纪50年代提出的解决最优控制问题的重要方法, 基本思想就是最优控制问题的最优状态轨道可以用一个具有边值问题的Hamilton系统来刻画. 对于连续时间的随机最优

控制系统, 相应的Hamilton系统可以用倒向随机微分方程或正倒向随机微分方程来描述<sup>[2]</sup>, 因此倒向随机微分方程理论在随机最优控制系统的最大值原理研究中具有重要的地位和作用. Pontryagin最大值原理主要针对连续时间的确定性系统<sup>[3]</sup>. 20世纪60年代中

收稿日期: 2021-08-27; 录用日期: 2022-03-10.

<sup>†</sup>通信作者. E-mail: lxy9393@sina.com.

本文责任编辑: 裴海龙.

国家自然科学基金项目(62273212, 61973198), 山东省泰山学者项目研究基金项目, 山东省自然科学基金项目(ZR2020MF062)资助.

Supported by the National Natural Science Foundation of China (62273212, 61973198), the Research Fund for the Taishan Scholar Project of Shandong Province of China and the National Natural Science Foundation of Shandong Province of China (ZR2020MF062).

期, Kushner<sup>[4-6]</sup>首先考虑将最大值原理扩展到随机的情况. 随后几十年间, Bismut<sup>[7-8]</sup>, Bensoussan<sup>[9]</sup>, Haussmann<sup>[10]</sup>等进一步研究了随机系统的最优控制问题, 为早期随机最优控制问题的最大值原理研究奠定了基础. 20世纪90年代初, 针对正向随机最优控制问题, 彭实戈<sup>[11]</sup>通过二阶变分技术和二阶伴随方程, 创造性地给出了一般最大值原理. 对于正倒向随机最优控制问题, 彭实戈<sup>[12]</sup>在1993年首次研究了控制域为凸的情形, 得到其最大值原理. 在此基础上, 吴臻分别在1998年和2013年给出了状态完全耦合且控制域为凸情形下的最大值原理<sup>[13]</sup>和在控制域非凸且倒向生成元依赖 $z$ 的情形下的一般最大值原理<sup>[14]</sup>, 王光臣等<sup>[15]</sup>给出了局部信息下平均场型最优控制的随机最大值原理, 这些研究成果极大地丰富和完善了随机最优控制理论体系<sup>[16-22]</sup>. 随着连续时间随机最优控制理论的发展, 离散时间随机最优控制问题的研究也逐渐被学者关注, 如文献[23]给出了具有乘性噪声的离散时间随机最优控制系统的最大值原理, 文献[24]研究了一类具有马尔可夫跳变参数的平均场随机线性二次最优控制问题, 文献[25]给出了由分数噪声和白噪声共同驱动的离散时间随机控制系统的最大值原理等.

在实际的工程应用中存在许多由于突变因素而引起系统状态和参数跳变的现象, 这种现象可以用带Markov跳的系统模型来描述<sup>[26]</sup>, 因此带Markov跳的随机模型被广泛地应用于金融、生物和工程实际等诸多领域<sup>[27-32]</sup>. 由于带Markov跳的含乘性噪声的离散时间随机最优控制系统结构复杂, 在其最大值原理研究中, 不仅要考虑其伴随倒向随机差分方程的结构形式, 还要处理Markov链与乘性噪声之间的关系, 这些问题都会影响到控制器的结构. 为了解决这些问题, 本文借鉴文献[24]中控制器的结构, 将文献[23]所研究的离散时间最优控制的最大值原理推广到带Markov跳情形. 问题具体叙述如下.

目标泛函为

$$J(u) = \mathbb{E} \sum_{k=0}^{N-1} l(r_k, x_k, u_k) + \mathbb{E}h(r_N, x_N). \quad (1)$$

状态约束方程为

$$\begin{cases} x_{k+1} = g(r_k, x_k, u_k) + \sigma(r_k, x_k, u_k)\omega_k, \\ x_0 \in \mathbb{R}^n, r_0 \in \mathcal{S}, \\ k = 1, 2, \dots, N-1, \end{cases} \quad (2)$$

其中:  $x_k$ 是状态变量,  $u_k$ 是容许控制变量,  $\omega_k$ 是相互独立的 $d$ 维随机变量,  $r_k$ 是齐次Markov链.

本文依然在针状变分法的基础上进行研究最优控制问题(1)–(2)的最大值原理. 该结论需要依赖具有适应解的带Markov跳的倒向随机差分方程. 同时, 通过引入最优值函数以Hamilton-Jacobi-Bellman方程的形

式给出了求解该最优控制问题的充分条件.

本文使用了以下符号:  $\mathbb{R}^n$ : 所有实 $n$ 维向量的集合;  $\mathbb{R}^{n \times d}$ :  $n \times d$ 维实矩阵的集合;  $A^T, x^T$ : 矩阵 $A$ 或向量 $x$ 的转置;  $n_u$ : 向量 $u$ 的维数;  $\langle \cdot, \cdot \rangle$ : 两个向量或矩阵的内积, 即 $\langle x, y \rangle = \text{tr}(x^T y)$ 且相应的范数为 $|x| = \sqrt{\langle x, x \rangle}$ ;  $A \otimes B$ : 两个矩阵 $A$ 和 $B$ 的克罗内克积;  $\delta_{km}$ : 克罗内克 $\delta$ 函数, 即 $k = m$ 时,  $\delta_{km} = 1, k \neq m$ 时,  $\delta_{km} = 0$ ;  $I_d$ :  $d \times d$ 维单位矩阵;  $0_d$ :  $d \times d$ 维零矩阵;  $\mathbb{E}[X]$ : 随机变量 $X$ 的数学期望;  $\mathbb{E}[X|\mathcal{F}_k]$ : 随机变量 $X$ 关于 $\sigma$ 域 $\mathcal{F}_k$ 的条件期望.

## 2 预备知识

在给定的完备概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上, 设 $\{\omega_k\}$ 是随机变量序列,  $\{r_k\}$ 是齐次Markov链, 它们满足以下条件( $i, j, k = 0, 1, \dots, N$ ):

1)  $\omega_0, \omega_1, \dots, \omega_N$ 是相互独立的 $d$ 维随机变量序列, 即 $\omega_k = [\omega_k^1 \ \dots \ \omega_k^d]^T \in \mathbb{R}^d$ 且 $\omega_k^1, \dots, \omega_k^d$ 也是相互独立的 $\mathbb{R}^d$ 值随机变量, 并满足如下条件:

$$\mathbb{E}\omega_k = 0_{d \times 1}, \mathbb{E}[\omega_i \omega_j^T] = \delta_{ij} I_d. \quad (3)$$

2) 齐次马尔可夫链 $r_k$ 的状态空间为

$$\mathcal{S} = \{1, 2, \dots, M\},$$

其一步转移概率为

$$q_{ij} = P(r_{k+1} = j | r_k = i), \quad i, j \in \mathcal{S}, \quad q_{ij} \geq 0, \quad (4)$$

且

$$\sum_{j \in \mathcal{S}} q_{ij} = 1, \quad \forall i \in \mathcal{S}. \quad (5)$$

3)  $\{\omega_k\}$ 与 $\{r_k\}$ 是相互独立的随机过程.

令 $\mathcal{F}_k^\omega$ 是由 $\omega_0, \omega_1, \dots, \omega_{k-1}$ 生成的 $\sigma$ 域, 即

$$\mathcal{F}_k^\omega = \sigma\{\omega_0, \omega_1, \dots, \omega_{k-1}\}.$$

$\mathcal{F}_k$ 是由 $\omega_0, \omega_1, \dots, \omega_{k-1}$ 和 $r_0, r_1, \dots, r_{k-1}$ 生成的 $\sigma$ 域, 即

$$\mathcal{F}_k = \sigma\{\omega_0, \omega_1, \dots, \omega_{k-1}\} \vee \sigma\{r_0, r_1, \dots, r_{k-1}\} = \mathcal{F}_k^\omega \vee \sigma\{r_0, r_1, \dots, r_{k-1}\},$$

并且 $\mathcal{F}_0 = \{\emptyset, \Omega\}$ ( $\emptyset$ 是空集). 设 $F = \{\mathcal{F}_k\}_{k=0}^N$ . 如果对每个 $k = 0, 1, \dots, N$ , 随机过程 $y_k$ 是 $\mathcal{F}_k$ 可测的随机变量, 则随机序列 $y = \{y_k\}_{k=0}^N$ 称为 $F$ -适应的<sup>[33]</sup>. 记 $L^2(\Omega, \mathcal{F}_k; \mathbb{R}^n)$ 为 $\mathcal{F}_k$ 可测随机变量全体且其中随机变量 $X$ 满足 $\mathbb{E}|X|^2 < \infty$ . 记

$$\mathcal{H}_N = \{L^2(\Omega, \mathcal{F}_k; \mathbb{R}^n) \times L^2(\Omega, \mathcal{F}_k; \mathbb{R}^{n \times d})\}_{k=0}^{N-1},$$

对任意 $\alpha, \beta \in \mathcal{H}_N$ , 定义内积如下:

$$\langle \alpha, \beta \rangle_{\mathcal{H}_N} = \mathbb{E} \sum_{k=0}^{N-1} [\langle \alpha_k^{(1)}, \beta_k^{(1)} \rangle + \langle \alpha_k^{(2)}, \beta_k^{(2)} \rangle],$$

其中:

$$\alpha = \{(\alpha_k^{(1)}, \alpha_k^{(2)})\}_{k=0}^{N-1}, \quad \beta = \{(\beta_k^{(1)}, \beta_k^{(2)})\}_{k=0}^{N-1},$$

则对上述内积, 线性空间 $\mathcal{H}_N$ 是Hilbert空间<sup>[34]</sup>.

**引理 1**(Riesz定理) 如果 $f$ 是Hilbert空间 $\mathcal{H}$ 上的连续线性泛函, 即

$$f : \mathcal{H} \rightarrow \mathbb{R},$$

那么存在唯一的 $y_f \in \mathcal{H}$ 满足

$$f(x) = \langle y_f, x \rangle_{\mathcal{H}}, \forall x \in \mathcal{H},$$

其中 $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ 为Hilbert空间上的内积.

下面将上述Riesz表示定理推广到具有线性差分方程约束的情形. 对任意给定的 $F$ 适应的随机过程 $\phi = \{\phi_k\}_{k=0}^{N-1}$ 和 $\psi = \{\psi_k\}_{k=0}^{N-1}$ , 考虑如下线性差分方程:

$$\begin{cases} z_{k+1} = g_k(r_k)z_k + \phi_k + \sum_{s=1}^d (\sigma_k^s(r_k)z_k + \psi_k^s)\omega_k^s, \\ z_0 = 0, k = 0, 1, \dots, N-1. \end{cases} \quad (6)$$

其中:  $g_k(r_k), \sigma_k^s(r_k) \in L^2(\Omega, \mathcal{F}_k; \mathbb{R}^{n \times n})$ 是在 $\mathbb{R}^{n \times n}$ 空间中取值的有界随机变量, 适应随机过程

$$\phi_k \in L^2(\Omega, \mathcal{F}_k; \mathbb{R}^n),$$

$$\psi_k = (\psi_k^1, \psi_k^2, \dots, \psi_k^d) \in L^2(\Omega, \mathcal{F}_k; \mathbb{R}^{n \times d}).$$

对给定的 $\tilde{l}_k(r_k) \in L^2(\Omega, \mathcal{F}_k; \mathbb{R}^n)$ , 考虑以下线性泛函:

$$I(\phi, \psi) = \mathbb{E} \sum_{k=0}^N \langle \tilde{l}_k(r_k), z_k \rangle, \quad (7)$$

其中 $z_k$ 是方程(6)的解, 即

$$z(\phi, \psi) = \{z_k\}_{k=0}^N,$$

则 $I(\cdot, \cdot)$ 是Hilbert空间 $\mathcal{H}_N$ 上的连续线性泛函<sup>[11]</sup>.

下面讨论 $I(\cdot, \cdot)$ 的表示问题, 即寻找

$$\alpha_I = (p, Q) \in \mathcal{H}_N,$$

使得对任意 $\alpha = (\phi, \psi) \in \mathcal{H}_N$ 有

$$I(\phi, \psi) = \langle \alpha_I, \alpha \rangle, \quad (8)$$

因此, 需要构造如下带Markov跳的耦合倒向随机差分方程:

$$\begin{cases} p_k(i) = \mathbb{E} \left\{ \sum_{j \in S} q_{ij} [\tilde{l}_{k+1}(j) + g_{k+1}^T(j)p_{k+1}(j) + \sum_{s=1}^d \sigma_{k+1}^s(j)^T Q_{k+1}^s(j)] | \mathcal{F}_k^\omega \right\}, \\ Q_k(i) = \mathbb{E} \left\{ \sum_{j \in S} q_{ij} [\tilde{l}_{k+1}(j) + g_{k+1}^T(j)p_{k+1}(j) + \sum_{s=1}^d \sigma_{k+1}^s(j)^T Q_{k+1}^s(j)] \omega_k^T | \mathcal{F}_k^\omega \right\}, \\ p_{N-1}(i) = \mathbb{E} [\sum_{j \in S} q_{ij} \tilde{l}_N(j) | \mathcal{F}_{N-1}^\omega], \\ Q_{N-1}(i) = \mathbb{E} [\sum_{j \in S} q_{ij} \tilde{l}_N(j) \omega_{N-1}^T | \mathcal{F}_{N-1}^\omega], \\ k = 0, 1, \dots, N-2, \end{cases} \quad (9)$$

其中 $Q_k^s$ 为 $Q_k = (Q_k^1, \dots, Q_k^d)$ 的第 $s$ 列, 从方程构造上能够看出倒向随机差分方程(9)在 $\mathcal{H}_N$ 中有唯一的适应解 $(p, Q) = \{(p_k(r_k), Q_k(r_k))\}_{k=0}^{N-1}$ .

**引理 2** 设 $(p, Q)$ 是倒向随机差分方程(9)的解, 那么由式(7)定义的泛函 $I(\cdot, \cdot)$ 可以唯一地表示为

$$I(\phi, \psi) = \mathbb{E} \left\{ \sum_{k=0}^{N-1} [\langle p_k(r_k), \phi_k \rangle + \langle Q_k(r_k), \psi_k \rangle] \right\}. \quad (10)$$

**证** 令 $A_k(r_k) = g_k(r_k) + \sum_{s=1}^d \sigma_k^s(r_k)\omega_k^s$ , 则差分方程(6)的解为

$$z_k = \sum_{t=0}^{k-1} A_{t+1}^k (\phi_t + \sum_{s=1}^d \psi_t^s \omega_t^s),$$

其中:  $k = 1, \dots, N-1, A_t^k, 1 \leq t \leq k \leq N$ 的定义如下:

当 $t = k$ 时,  $A_t^k = I_n$ ;

当 $1 \leq t < k$ 时,

$$A_t^k = A_{k-1}(r_{k-1})A_{k-2}(r_{k-2}) \cdots A_t(r_t),$$

为了简洁, 证明中 $\tilde{l}_k(r_k)$ 简记为 $\tilde{l}_k$ , 将 $z_k$ 代入式(7), 并利用条件期望的平滑性得

$$I(\phi, \psi) = \mathbb{E} \left\{ \sum_{k=0}^{N-1} [\mathbb{E} [\sum_{t=k+1}^N (A_{t+1}^k)^T \tilde{l}_t | \mathcal{F}_k], \phi_k] + \langle \mathbb{E} [\sum_{t=k+1}^N (A_{t+1}^k)^T \tilde{l}_t \omega_k^T | \mathcal{F}_k], \psi_k \rangle \right\}.$$

令

$$p_k = \mathbb{E} [\sum_{t=k+1}^N (A_{t+1}^k)^T \tilde{l}_t | \mathcal{F}_k], \quad (11)$$

$$Q_k = \mathbb{E} [\sum_{t=k+1}^N (A_{t+1}^k)^T \tilde{l}_t \omega_k^T | \mathcal{F}_k], \quad (12)$$

利用数学归纳法可证 $(p_k, Q_k)$ 满足方程(9). 显然 $(p, Q) = \{(p_k(i), Q_k(i))\}_{k=0}^{N-1}$ 是 $F$ 适应的且满足式(10). 再由Riesz定理可知, 满足式(10)的过程 $(p, Q)$ 是唯一的. 证毕.

### 3 带Markov跳的离散时间随机系统的最大值原理

设最优控制问题(1)-(2)的容许控制集为

$$\mathcal{U} = \{\mathcal{U}_k\}_{k=0}^{N-1},$$

其中:  $\mathcal{U}_k = L^2(\Omega, \mathcal{F}_k; \mathbb{R}^{n_u}), \{r_k\}_{k=0}^N$ 是齐次Markov链且满足式(4),  $\{\omega_k = [\omega_k^1 \ \dots \ \omega_k^d]^T\}_{k=1}^N$ 是相互独立的随机变量且满足式(3). 同时做如下假设.

**假设 1**  $\phi$ 关于 $x$ 和 $u$ 是二阶连续可微的,  $\phi_x, \phi_{xx}, \phi_u, \phi_{uu}, l_{uu}, l_{xx}, h_{xx}$ 有界, 其中 $\phi = g, \sigma \times \omega_k$ 有四阶矩, 即 $\mathbb{E}|\omega_k|^4 < \infty$ .

设 $u^* = \{u_k^*\}_{k=0}^{N-1}$ 是最优控制问题(1)-(2)的一个最优控制, 相应的最优轨道为 $x^* = \{x_k^*\}_{k=0}^N$ . 对于固

定的 $m, 0 \leq m \leq N - 1$ , 记

$$\bar{u}_k = u_k^* + \delta_{km}\varepsilon\Delta u, \Delta u \in \mathcal{U}_m,$$

其中 $\Delta u$ 是取值于 $\mathcal{U}_m \subset L^2(\Omega, \mathcal{F}_m; \mathbb{R}^{n_u})$ 的随机变量, 满足

$$\sup_{\omega \in \Omega} |\Delta u(\omega)| < \infty, \quad (13)$$

其中 $\varepsilon > 0$ , 它的大小表明了 $\bar{u}_k$ 与 $u_k^*$ 之间的差别, 且当 $\varepsilon \rightarrow 0^+$ 时,  $\bar{u}_k \rightarrow u_k^*$ , 也就是 $\bar{u}_k$ 的极限是最优控制 $u_k^*$ . 令 $\bar{x} = \{\bar{x}_k\}_{k=0}^N$ 是系统(2)在初值为 $\bar{x}_0 = x_0$ 控制为 $\bar{u}$ 时对应的解.

**引理 3** 在假设1的条件下, 下列不等式成立:

$$\sup_{0 \leq k \leq N} \mathbb{E}[|\bar{x}_k - x_k^*|^2] \leq K_1 \varepsilon^2 \mathbb{E}|\Delta u|^2, \quad (14)$$

$$\sup_{0 \leq k \leq N} \mathbb{E}[|\bar{x}_k - x_k^*|^4] \leq K_1 \varepsilon^4 \mathbb{E}|\Delta u|^4. \quad (15)$$

**证** 记 $\hat{x}_k = \bar{x}_k - x_k^*$ , 则 $\hat{x}_0 = 0$ . 下面对 $k$ 分3种情况进行讨论:

- 1) 当 $0 \leq k \leq m$ 时, 有 $\hat{x}_k = 0$ , 显然结论成立.
- 2) 当 $k = m + 1$ 时, 利用Taylor公式, 得

$$\hat{x}_{m+1} = g_u(r_m, x_m, u_m^* + \theta_m \Delta u) \Delta u + \sum_{s=1}^d [\sigma_u^s(r_m, x_m, u_m^* + \theta_m \Delta u) \Delta u] \omega_m^s,$$

其中 $0 < \theta_m < 1$ , 由假设1中 $g_u, \sigma_u^s$ 的有界性可得得不等式(14)–(15)成立.

3) 当 $k > m + 1$ 时, 与情形2)类似, 可证存在 $C_3 > 0, C_4 > 0$ 使得

$$\mathbb{E}[|\hat{x}_{k+1}|^2] \leq C_3 \mathbb{E}[|\hat{x}_k|^2], \mathbb{E}[|\hat{x}_{k+1}|^4] \leq C_4 \mathbb{E}[|\hat{x}_k|^4],$$

利用上式的递推关系, 令 $K_1 = \max_{m+1 \leq k \leq N} C_1 C_3^{k-m}$ , 则可得出不等式(14)成立. 类似地, 不等式(15)可证.

设 $y = \{y_k\}_{k=0}^N$ 是以下线性差分方程的解:

$$\begin{cases} y_{k+1} = g_x(r_k, x_k, u_k) y_k + \sum_{s=1}^d \sigma_x^s(r_k, x_k, u_k) y_k \omega_k^s + \delta_{km} g_u(r_m, x_m, u_m) \varepsilon \Delta u + \sum_{s=1}^d \delta_{km} \sigma_u^s(r_m, x_m, u_m) \varepsilon \Delta u \omega_k^s, \\ y_0 = 0. \end{cases} \quad (16)$$

下面引理4表明方程(16)中的 $y$ 为方程(2)中 $x$ 的一阶变分, 因此, 方程(16)称为方程(2)的一阶变分方程.

**证** 毕.

**引理 4** 在假设1的条件下, 存在 $M > 0$ 使得下列不等式成立:

$$\sup_{0 \leq k \leq N} [|\bar{x}_k - x_k^* - y_k|^2] \leq M \varepsilon^4 \mathbb{E}[|\Delta u|^4]. \quad (17)$$

**证** 记 $\hat{x}_k = \bar{x}_k - x_k^*, \hat{x}_0 = 0$ , 由Taylor公式, 得

$$\begin{aligned} \hat{x}_{k+1} &= g_x(r_k, x_k^*, u_k^*) \hat{x}_k + \Phi_k^1 + \delta_{km} [g_u(r_m, x_m, u_m^*) \varepsilon \Delta u + \Phi_m^2] + \sum_{s=1}^d [\sigma_x^s(r_k, x_k^*, u_k^*) \hat{x}_k + \Psi_k^{1,s} + \delta_{km} (\sigma_u^s(r_m, x_m, u_m^*) \varepsilon \Delta u + \Psi_m^{2,s})] \omega_k^s, \end{aligned}$$

当 $k \geq m + 1$ 时,

$$\begin{aligned} \Phi_k^1 &= \frac{1}{2} (\hat{x}_k^T \otimes I_n) g_{xx}(r_k, x_k^* + \theta_k \hat{x}_k, u_k^*) (\hat{x}_k \otimes 1_n), \\ \Psi_k^{1,s} &= \frac{1}{2} (\hat{x}_k^T \otimes I_n) \sigma_{xx}^s(r_k, x_k^* + \theta_k \hat{x}_k, u_k^*) (\hat{x}_k \otimes 1_n), \end{aligned}$$

当 $k \leq m$ 时,  $\Phi_k^l = 0$ . 此外,

$$\begin{aligned} \Phi_m^2 &= \frac{1}{2} \varepsilon^2 (\Delta u^T \otimes I_n) g_{uu}(r_m, x_m, u_m^* + \theta_m \varepsilon \Delta u) \times (\Delta u \otimes 1_n), \\ \Psi_m^{2,s} &= \frac{1}{2} \varepsilon^2 (\Delta u^T \otimes I_n) \sigma_{uu}^s(r_m, x_m, u_m^* + \theta_m \varepsilon \Delta u) \times (\Delta u \otimes 1_n), \end{aligned}$$

其中符号 $\varphi_x, 1_n$ 按如下方式定义:

$$\varphi_x = \begin{bmatrix} \varphi_{x_1}^1(i, x, u) & \varphi_{x_2}^1(i, x, u) & \cdots & \varphi_{x_n}^1(i, x, u) \\ \varphi_{x_1}^2(i, x, u) & \varphi_{x_2}^2(i, x, u) & \cdots & \varphi_{x_n}^2(i, x, u) \\ \vdots & \vdots & & \vdots \\ \varphi_{x_1}^n(i, x, u) & \varphi_{x_2}^n(i, x, u) & \cdots & \varphi_{x_n}^n(i, x, u) \end{bmatrix},$$

$\varphi = [\varphi^1 \cdots \varphi^n]^T, \varphi = g, \sigma^s, s = 1, \dots, d, \varphi_u$ 也有类似的含义,  $1_n = [1 \ 1 \ \cdots \ 1]^T \in \mathbb{R}^n$ .

相应的符号 $\varphi_{xx}, \varphi_{uu}$ 按如下方式定义:

$$\begin{aligned} \varphi_{xx}(i, x_k + \theta_k \hat{x}_k, u_k) &= \begin{bmatrix} \varphi_{xx}^1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \varphi_{xx}^n \end{bmatrix}, \\ \varphi_{uu}(i, x_k, u_k + \theta_k \hat{u}_m) &= \begin{bmatrix} \varphi_{uu}^1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & \varphi_{uu}^n \end{bmatrix}, \end{aligned}$$

其中所有的 $\theta, \theta_k^s, \tilde{\theta}_k^s$ 满足 $0 < \theta < 1$ , 且 $\hat{u}_m = \varepsilon \Delta u$ .

由 $\varphi_{xx}$ 的有界性及引理3, 有

$$\begin{aligned} \mathbb{E}[|\Phi_k^1|^2] &\leq K_1 \mathbb{E}[|\hat{x}_k|^4] \leq K_2 \varepsilon^4 \mathbb{E}[|\Delta u|^4], \\ \mathbb{E}[|\Phi_m^2|^2] &\leq K_2 \varepsilon^4 \mathbb{E}[|\Delta u|^4], \\ \mathbb{E}[|\Psi_k^{1,s}|^2] &\leq K_2 \varepsilon^4 \mathbb{E}[|\Delta u|^4], \\ \mathbb{E}[|\Psi_m^{2,s}|^2] &\leq K_2 \varepsilon^4 \mathbb{E}[|\Delta u|^4], \\ k > m, s &= 1, \dots, d. \end{aligned}$$

记 $Z_k = \bar{x}_k - x_k^* - y_k, k = 0, 1, \dots, N - 1$ , 下面对 $k$ 的不同取值情况分别讨论:

- 1) 当 $k \leq m$ 时, 显然有 $Z_k = 0$ .
- 2) 当 $k = m + 1$ , 有

$$Z_{m+1} = \bar{x}_{m+1} - x_{m+1}^* - y_{m+1} =$$

$$\Phi_m^2 + \sum_{s=1}^d \Psi_m^{2,s} \omega_k^s, \mathbb{E}[|Z_{m+1}|^2],$$

再利用假设1有

$$\mathbb{E}[|Z_{m+1}|^2] \leq K_3 \varepsilon^4 \mathbb{E}[|\Delta u|^4].$$

3) 当  $k > m + 1$  时,

$$\begin{aligned} Z_{k+1} &= \sum_{s=1}^d [\sigma_x^s(r_k, x_k^*, u_k^*) Z_k + \Psi_k^{1,s}] \omega_k^s + \\ &g_x(r_k, x_k^*, u_k^*) Z_k + \Phi_k^1, \\ \mathbb{E}[|Z_{k+1}|^2] &= \sum_{s=1}^d \mathbb{E}[|\sigma_x^s(r_k, x_k^*, u_k^*) Z_k + \Psi_k^{1,s}|^2] + \\ &\mathbb{E}[|g_x(r_k, x_k^*, u_k^*) Z_k + \Phi_k^1|^2] \leq \\ &K_3 \mathbb{E}[|Z_k|^2] + K_3 \varepsilon^4 \mathbb{E}[|\Delta u|^4], \end{aligned}$$

通过迭代, 可以证明不等式(17). 证毕.

### 3.1 最优控制存在的必要条件

由于方程(16)所确定的过程  $y = \{y_k\}_{0 \leq k \leq N}$  是方程(2)的解  $x = \{x_k\}_{0 \leq k \leq N}$  的一阶变分, 因此值函数  $J(u)$  的变分是

$$\begin{aligned} \mathcal{L}J(u) &= \mathbb{E} \sum_{k=0}^{N-1} [\langle l_x(r_k, x_k, u_k), y_k \rangle] + \\ &\mathbb{E} \sum_{k=0}^{N-1} [\delta_{km} \langle l_u(r_m, x_m, u_m), \varepsilon \Delta u \rangle] + \\ &\mathbb{E} \langle h_x(r_N, x_N), y_N \rangle. \end{aligned}$$

定义如下Hamilton函数:

$$\begin{aligned} H(i, x, u, p, Q) &= \\ &\langle p, g(i, x, u) \rangle + \sum_{s=1}^d \langle Q^s, \sigma^s(i, x, u) \rangle + l(i, x, u), \end{aligned} \tag{18}$$

其中:

$$\begin{aligned} p &\in \mathbb{R}^n, Q = (Q^1, \dots, Q^d) \in \mathbb{R}^{n \times d}, \\ x &\in \mathbb{R}^n, u \in \mathbb{R}^{n_u}, i \in \mathcal{S}. \end{aligned}$$

构造如下离散时间倒向随机差分方程:

$$\begin{cases} p_k(i) = \mathbb{E} \left\{ \sum_{j \in \mathcal{S}} q_{ij} [l_x(j, x_{k+1}^*, u_{k+1}^*) + \right. \\ \quad \left. g_x^T(j, x_{k+1}^*, u_{k+1}^*) p_{k+1}(j) + \right. \\ \quad \left. \sum_{s=1}^d \sigma_x^s [j \ x_{k+1}^* \ u_{k+1}^*]^T \times \right. \\ \quad \left. Q_{k+1}^s(j) \right] | \mathcal{F}_k^\omega \}, \\ Q_k(i) = \mathbb{E} \left\{ \sum_{j \in \mathcal{S}} q_{ij} [l_x(j, x_{k+1}^*, u_{k+1}^*) + \right. \\ \quad \left. g_x^T(j, x_{k+1}^*, u_{k+1}^*) p_{k+1}(j) + \right. \\ \quad \left. \sum_{s=1}^d \sigma_x^s [j \ x_{k+1}^* \ u_{k+1}^*]^T \times \right. \\ \quad \left. Q_{k+1}^s(j) \right] \omega_k | \mathcal{F}_k^\omega \}, \\ p_{N-1}(i) = \mathbb{E} \left[ \sum_{j \in \mathcal{S}} q_{ij} h_x(j, x_N) | \mathcal{F}_{N-1}^\omega \right], \\ Q_{N-1}(i) = \mathbb{E} \left[ \sum_{j \in \mathcal{S}} q_{ij} h_x(j, x_N) \omega_{N-1} | \mathcal{F}_{N-1}^\omega \right], \\ j \in \mathcal{S}, k = 0, 1, \dots, N-2. \end{cases} \tag{19}$$

注 1 由式(18)可得

$$\begin{aligned} \frac{\partial H}{\partial x}(i, x, u, p, Q) &= l_x(i, x, u) + g_x^T(i, x, u) p + \\ &\sum_{s=1}^d \sigma_x^s [i \ x \ u]^T Q^s, \end{aligned} \tag{20}$$

因此倒向随机差分方程(19)还可以表示为如下形式:

$$\begin{cases} p_k(i) = \mathbb{E} \left\{ \sum_{j \in \mathcal{S}} q_{ij} \frac{\partial H}{\partial x}(j, x_{k+1}^*, u_{k+1}^*) p_{k+1}(j), \right. \\ \quad \left. Q_{k+1}(j) \right] | \mathcal{F}_k^\omega \}, \\ Q_k(i) = \mathbb{E} \left\{ \sum_{j \in \mathcal{S}} q_{ij} \frac{\partial H}{\partial x}(j, x_{k+1}^*, u_{k+1}^*) p_{k+1}(j), \right. \\ \quad \left. Q_{k+1}(j) \right] \omega_k | \mathcal{F}_k^\omega \}, \\ p_{N-1}(i) = \mathbb{E} \left[ \sum_{j \in \mathcal{S}} q_{ij} h_x(j, x_N) | \mathcal{F}_{N-1}^\omega \right], \\ Q_{N-1}(i) = \mathbb{E} \left[ \sum_{j \in \mathcal{S}} q_{ij} h_x(j, x_N) \omega_{N-1} | \mathcal{F}_{N-1}^\omega \right], \\ j \in \mathcal{S}, k = 0, 1, \dots, N-2. \end{cases} \tag{21}$$

在上述工作的基础上, 下面给出并证明最优控制问题(1)–(2)的最大值原理.

**定理 1** 如果最优控制问题(1)–(2)中的系数  $g, \sigma, h$  满足假设1的条件, 设  $\{u_k^*\}_{k=0}^{N-1}$  和  $\{x_k^*\}_{k=0}^N$  分别为它的最优解和最优轨道,  $(p, Q)$  是倒向随机差分方程(19)的解, 那么, 对任意的  $i \in \mathcal{S}$ , 有

$$H_u(i, x_m^*, u_m^*, p_m(i), Q_m(i)) = 0 \quad \text{a.s.} \tag{22}$$

证 为了简洁, 把  $\phi_u(i, x_m, u_m)$  简记为  $\phi_u$ , 其中  $\phi = g, \sigma, l$ . 由引理2, 有

$$\begin{aligned} \mathcal{L}J(u) &= \mathbb{E} \left\{ \sum_{k=0}^{N-1} [\langle p_k, \delta_{km} g_u \varepsilon \Delta u \rangle + \right. \\ &\quad \left. \langle Q_k, \delta_{km} \sigma_u \varepsilon \Delta u \rangle + \delta_{km} \langle l_u, \varepsilon \Delta u \rangle] \right\}, \end{aligned}$$

所以

$$\mathcal{L}J(u) = \mathbb{E}[\langle H_u(i, x_m, u_m, p_m(i), Q_m(i)), \varepsilon \Delta u \rangle],$$

由于  $u^*$  是最优控制, 有

$$\mathcal{L}J(u^*) = 0,$$

即对任意  $\Delta u \in \mathcal{U}$ , 有

$$\mathbb{E}[\langle H_u(i, x_m^*, u_m^*, p_m(i), Q_m(i)), \Delta u \rangle] = 0, \tag{23}$$

对于每个整数  $K > 0$ , 令

$$\begin{aligned} \Delta u &= H_u(i, x_m^*, u_m^*, p_m(i), Q_m(i)) \times \\ &\quad \mathbf{1}_{\{|H_u(i, x_m^*, u_m^*, p_m(i), Q_m(i))| < K\}} \in \mathcal{U}. \end{aligned}$$

由式(23), 有

$$\begin{aligned} \mathbb{E}[|H_u(i, x_m^*, u_m^*, p_m(i), Q_m(i))|^2 \times \\ \mathbf{1}_{\{|H_u(i, x_m^*, u_m^*, p_m(i), Q_m(i))| < K\}}] &= 0, \end{aligned}$$

从而可得

$$\begin{aligned} |H_u(i, x_m^*, u_m^*, p_m(i), Q_m(i))|^2 \times \\ \mathbf{1}_{\{|H_u(i, x_m^*, u_m^*, p_m(i), Q_m(i))| < K\}} &= 0 \quad \text{a.s.} \end{aligned}$$

令  $K \rightarrow \infty$ , 可以得到式(22).

由于式(22)是最优控制问题(1)–(2)的必要条件, 因此定理1给出了最优控制问题(1)–(2)的最大值原理. 证毕.

### 3.2 最优控制问题的充分条件

下面讨论最优控制问题(1)–(2)的最优控制存在的充分条件. 这需要引入下列辅助函数:

$$V_k : \mathcal{S} \times \mathbb{R}^n \longrightarrow \mathbb{R}^+,$$

为了简洁, 引入符号

$$\Delta_u V_k(i, x) = \mathbb{E}[\sum_{j \in \mathcal{S}} q_{ij} V_{k+1}(j, g(i, x, u) + \sigma(i, x, u)\omega_k)] - V_k(i, x),$$

其中:  $x \in \mathbb{R}^n, u \in \mathcal{U} \subset \mathbb{R}^{n_u}$ .

**定理 2** 若存在函数序列  $V_k(i, x)$  和容许控制  $u^* = \{u_k^*\} \in \mathcal{U}, k = 0, 1, \dots, N$ , 满足

$$\Delta_{u^*} V_k(i, x) + l(i, x, u^*) \equiv 0, \forall i \in \mathcal{S}, \quad (24)$$

$$\Delta_u V_k(i, x) + l(i, x, u) \geq 0, \forall u \in \mathbb{R}^{n_u}, \quad (25)$$

且

$$V_N(i, x) = h(i, x), i \in \mathcal{S}, x \in \mathbb{R}^n, \quad (26)$$

则  $u^* = \{u_k^*\}_{k=0}^N$  是最优控制问题(1)–(2)的最优控制, 即

$$J(u^*) = \min_{u \in \mathcal{U}} J(u), \quad (27)$$

且对给定的  $r_0 \in \mathcal{S}$ , 有

$$J(u^*) = V_0(r_0, x_0). \quad (28)$$

**证** 对每个容许控制  $u = \{u_k\}$ , 设  $x = \{x_k\}$  为相应的状态轨道, 则

$$\begin{aligned} & \mathbb{E}[V_{k+1}(r_{k+1}, x_{k+1}) | \mathcal{F}_k] - V_k(r_k, x_k) = \\ & \mathbb{E}[V_{k+1}(r_{k+1}, g(r_k, x_k, u_k) + \sigma(r_k, x_k, u_k)\omega_k) - \\ & V_k(r_k, x_k) + l(r_k, x_k, u_k) - l(r_k, x_k, u_k) | \mathcal{F}_k], \end{aligned}$$

由于  $l(r_k, x_k, u_k)$  为  $\mathcal{F}_k$  可测, 因此利用条件期望的平滑性, 可得, 在  $r_k = i$  的条件下, 有

$$\begin{aligned} & \mathbb{E}[V_{k+1}(r_{k+1}, x_{k+1}) | \mathcal{F}_k] - V_k(r_k, x_k) = \\ & [\Delta_u V_k(i, x) + l(i, x, u)]_{x=x_k, u=u_k} - \\ & l(r_k, x_k, u_k), \end{aligned} \quad (29)$$

在式(29)中令  $u = \{u_k^*\}, x = \{x_k^*\}$ . 由式(24)得

$$l(r_k, x_k^*, u_k^*) = V_k(r_k, x_k^*) - \mathbb{E}[V_{k+1}(r_{k+1}, x_{k+1}^*) | \mathcal{F}_k],$$

两边取期望求和, 再利用条件  $V_N(i, x) = h(i, x)$ , 得

$$\mathbb{E}[\sum_{k=0}^{N-1} l(i, x_k^*, u_k^*)] + \mathbb{E}[h(i, x_N^*)] = V_0(r_0, x_0),$$

即

$$J(u^*) = V_0(r_0, x_0). \quad (30)$$

下证  $u^*$  为最优控制. 设  $u = \{u_k\}$  为容许控制,  $x_k$  为相应的状态, 由式(29), 并利用条件(25), 有

$$l(i, x_k, u_k) \geq V_k(i, x_k) - \mathbb{E}[\sum_{j \in \mathcal{S}} q_{ij} V_{k+1}(j, x_{k+1}) | \mathcal{F}_k^\omega],$$

两边取数学期望并求和得

$$J(u) \geq V_0(r_0, x_0),$$

由式(30)得

$$J(u) \geq J(u^*),$$

因此  $u^* = \{u_k^*\}$  是最优控制问题(1)–(2)的最优控制.

证毕.

**注 2** 将定理2中的条件(24)–(25)整合, 可得到  $u^*$  是如下Hamilton-Jacobi-Bellman方程的解:

$$\min_{u \in \mathcal{U}} [\Delta_u V_k(i, x) + l(i, x, u)] = 0, \forall i \in \mathcal{S}. \quad (31)$$

**注 3** 定义Hamilton函数如下:

$$\tilde{H}_k(i, x, u) = \Delta_u V_k(i, x) + l(i, x, u), \quad (32)$$

则Hamilton-Jacobi-Bellman方程(31)可表示为

$$\min_{u \in \mathcal{U}} \tilde{H}_k(i, x, u) = 0, \forall i \in \mathcal{S}. \quad (33)$$

定理2给出了最优控制问题的充分条件, 但求解Hamilton-Jacobi-Bellman方程(31)或方程(33)存在困难. 在  $\tilde{H}_k$  关于  $u$  具有二阶连续偏导数的条件下, 可以用如下的定理3来获取最优控制.

**定理 3** 设  $\tilde{H}_k$  关于  $u \in \mathcal{U} \subset \mathbb{R}^{n_u}$  二阶连续可微, 如果存在函数  $\zeta_k(i, x)$  满足以下条件:

- 1)  $\tilde{H}_k(i, x, \zeta_k(i, x)) \equiv 0, \forall i \in \mathcal{S}, x \in \mathbb{R}^n;$
- 2)  $\frac{\partial \tilde{H}_k}{\partial u}(i, x, u)|_{u=\zeta_k(i, x)} = 0, \forall i \in \mathcal{S}, x \in \mathbb{R}^n;$
- 3)  $\frac{\partial^2 \tilde{H}_k}{\partial u^2}(i, x, u) \geq 0, \forall i \in \mathcal{S}, x \in \mathbb{R}^n, u \in \mathbb{R}^{n_u},$

那么具有  $\{u_k^* = \zeta_k(i, x)\}_{k=1, \dots, N-1}$  结构的控制器是最优控制问题(1)–(2)的最优控制.

**证** 对任意  $x \in \mathbb{R}^n$ , 在  $u_k = \zeta_k(i, x)$  处, 对  $\tilde{H}_k$  进行泰勒展开, 有

$$\begin{aligned} \tilde{H}_k(i, x, u) &= \tilde{H}_k(i, x, \zeta_k(i, x)) + \langle \frac{\partial \tilde{H}_k}{\partial u} |_{u=\zeta_k(i, x)}, \\ & u - \zeta_k(i, x) \rangle + \frac{1}{2} [u - \zeta_k(i, x)]^T \times \\ & \frac{\partial^2 \tilde{H}_k}{\partial u^2} |_{u=\zeta_k(i, x) + \theta(u - \zeta_k(i, x))} \times \\ & [u - \zeta_k(i, x)], \end{aligned}$$

其中  $0 < \theta < 1$ . 由条件3), 有

$$\begin{aligned} \tilde{H}_k(i, x, u) &\geq \tilde{H}_k(i, x, \zeta_k(i, x)) + \\ & \langle \frac{\partial \tilde{H}_k}{\partial u} |_{u=\zeta_k(i, x)}, u - \zeta_k(i, x) \rangle, \end{aligned}$$

再由条件1)和2), 有

$$\tilde{H}_k(i, x, u) \geq 0, \forall i \in \mathcal{S}, \quad (34)$$

由定理2可得  $u_k^* = \zeta_k(i, x)$  是最优控制.

**注 4** 下面讨论定理1与定理3之间的关系. 为了方便, 假设  $V_k(i, x)$  与  $i$  无关, 则

$$\Delta_u V_k(i, x) = E[V_{k+1}(g(i, x, u) + \sigma(i, x, u)\omega_k)] - V_k(x),$$

相应地, 式(32)中的Hamilton函数即为

$$\tilde{H}_k(i, x, u) = E[V_{k+1}(g(i, x, u) + \sigma(i, x, u)\omega_k)] - V_k(x) + l(i, x, u),$$

首先, 对定理3中的Hamilton函数  $\tilde{H}_k$  求关于  $x$  的一阶偏导数, 则有

$$\frac{\partial \tilde{H}_k}{\partial x} = E[l_x(i, x, u) + g_x^T(i, x, u) \frac{\partial V_{k+1}}{\partial x} + \sum_{s=1}^d \sigma_x^s [i \ x \ u]^T \frac{\partial V_{k+1}}{\partial x} \omega_k^s] - \frac{\partial V_k}{\partial x}(i, x),$$

然后, 由定理3中条件1)知, 对任意  $x \in \mathbb{R}^n$ , 恒有

$$\tilde{H}_k(i, x, \zeta_k(i, x)) = 0,$$

若设  $\zeta(i, x)$  具有连续导数, 则上式两边可对  $x$  求导, 则

$$\frac{\partial \tilde{H}_k}{\partial x} + \sum_{s=1}^d \zeta_x^s [i \ x]^T \frac{\partial \tilde{H}_k}{\partial u^s}(i, x, \zeta_k(i, x)) = 0,$$

注意到  $u^* = \zeta(i, x)$ , 再由定理3中条件2)可知

$$\frac{\partial \tilde{H}_k}{\partial x}(i, x, u^*) = 0,$$

即

$$\frac{\partial V_k}{\partial x} = E[l_x(i, x, u^*) + g_x^T(i, x, u^*) \frac{\partial V_{k+1}}{\partial x} + \sum_{s=1}^d \sigma_x^s [i \ x \ u^*]^T \frac{\partial V_{k+1}}{\partial x} \omega_k^s],$$

在  $u, x$  分别取最优控制  $u_{k+1}^*$  和最优轨道  $x_{k+1}^*$  时, 令

$$\begin{aligned} p_{k+1}(i) &= \frac{\partial V_{k+1}}{\partial x}(i, x_{k+1}^*), \\ Q_{k+1}^s(i) &= \frac{\partial V_{k+1}}{\partial x}(i, x_{k+1}^*) \omega_k^s, \\ Q_{k+1}(i) &= (Q_{k+1}^1(i), \dots, Q_{k+1}^d(i)), \end{aligned}$$

通过上式与倒向随机差分方程(19)比较可知,  $p_k, Q_k$  满足离散时间倒向随机差分方程(19).

证毕.

#### 4 例子

**例 1** 对如下带Markov跳的线性随机控制系统:

$$\begin{cases} x_{k+1} = A(r_k)x_k + B(r_k)u_k + C(r_k)x_k\omega_k, \\ x_0 \in \mathbb{R}^n, k = 0, 1, \dots, N-1, \end{cases} \quad (35)$$

考虑如下线性二次问题:

$$\begin{aligned} J(u) = E\{ & \sum_{k=0}^{N-1} [x_k^T \tilde{Q}(r_k)x_k + 2x_k^T \tilde{S}(r_k)u_k + \\ & u_k^T \tilde{R}(r_k)u_k] \} + E[x_N^T \tilde{Q}(r_N)x_N], \end{aligned} \quad (36)$$

其中:  $A, C, \tilde{Q} \in \mathbb{R}^{n \times n}; B, S \in \mathbb{R}^{n \times n_u}; R \in \mathbb{R}^{n_u \times n_u}$

为矩阵; 随机噪声  $\omega_k$  是相互独立的1维随机变量且  $E[\omega_k] = 0, E[\omega_k^2] = 1; r_k$  是状态空间为  $\mathcal{S}$  的齐次Markov链, 转移概率满足式(4)–(5),  $\{r_k\}$  与  $\{\omega_k\}$  相互独立.

将文献[24]定理3.1中的平均场部分去掉, 可以得到最优控制问题(35)–(36)的状态反馈最优控制为

$$u_k^* = -\Pi^{-1}(r_k, \mathcal{P}_{k+1})\Xi(r_k, \mathcal{P}_{k+1})x_k^*, \quad (37)$$

其中  $\{\mathcal{P}_k\} \geq 0$  是以下Riccati方程的解:

$$\begin{cases} P_k(i) = \Gamma(i, \mathcal{P}_{k+1}) - \Xi[i \ \mathcal{P}_{k+1}]^T \times \\ \quad \Pi^{-1}(i, \mathcal{P}_{k+1})\Xi(i, \mathcal{P}_{k+1}), \\ P_N(i) = \tilde{Q}_N(i), i \in \mathcal{S}, k = 0, 1, \dots, N-1, \end{cases} \quad (38)$$

其中: 符号  $\Pi, \Xi$  和  $\Gamma$  分别为

$$\begin{aligned} \Pi(i, \mathcal{P}_{k+1}) &= \sum_{j \in \mathcal{S}} q_{ij} [B_k^T(i) \mathcal{P}_{k+1}(j) B_k(i) + \tilde{R}(i)], \\ \Xi(i, \mathcal{P}_{k+1}) &= \sum_{j \in \mathcal{S}} q_{ij} [B_k^T(i) \mathcal{P}_{k+1} A_k(i) + \tilde{S}^T(i)], \\ \Gamma(i, \mathcal{P}_{k+1}) &= \sum_{j \in \mathcal{S}} q_{ij} [A(i)^T \mathcal{P}_{k+1}(j) A(i) + \\ & \quad C(i)^T \mathcal{P}_{k+1} C(i) + \tilde{Q}(i)], \end{aligned}$$

取Hamilton函数为

$$\begin{aligned} H(i, x, u, p, Q) = & x^T Q(i)x + x^T S(i)x + \\ & u^T R(i)u + \langle p, A(i)x + \\ & B(i)u \rangle + \langle Q, C(i)x \rangle, \end{aligned} \quad (39)$$

设  $x_k^*$  是式(37)确定的最优控制  $u^*$  对应的状态轨道, 由定理1知最优控制问题(35)–(36)对应的伴随倒向随机差分方程为

$$\begin{cases} p_k(i) = E\{ \sum_{j \in \mathcal{S}} q_{ij} [2\tilde{Q}(i)x_{k+1}^* + 2\tilde{S}(i)u_{k+1}^* + \\ \quad A^T(i)p_{k+1}(j) + \\ \quad C^T(i)p_{k+1}(j)Q_{k+1} | \mathcal{F}_k^\omega] \}, \\ Q_k(i) = E\{ \sum_{j \in \mathcal{S}} q_{ij} [2\tilde{Q}(i)x_{k+1}^* + 2\tilde{S}(i)u_{k+1}^* + \\ \quad A^T(i)p_{k+1}(j) + C^T(i)p_{k+1}(j) \times \\ \quad Q_{k+1}] \omega_k | \mathcal{F}_k^\omega \}, \\ p_{N-1}(i) = 2E\{ \sum_{j \in \mathcal{S}} q_{ij} \tilde{Q}(j)x_N | \mathcal{F}_{N-1}^\omega \}, \\ Q_{N-1}(i) = 2E\{ \sum_{j \in \mathcal{S}} q_{ij} \tilde{Q}(j)x_N \omega_{N-1} | \mathcal{F}_{N-1}^\omega \}, \\ i \in \mathcal{S}, k = 0, 1, \dots, N-2, \end{cases}$$

由定理1可知, 对任意  $m = 0, 1, \dots, N-1$  有

$$H_u(i, x_m^*, u_m^*, p_m(i), Q_m(i)) = 0 \quad \text{a.s.}$$

**例 2** 考虑图1所示的风光储联合发电的微电网优化管理系统<sup>[35]</sup>.

该系统包括风力发电  $P_{\text{wind}}$ 、太阳能发电  $P_{\text{solar}}$ 、输送电网  $P_{\text{grid}}$  和储能设备  $P_{\text{sto}}$  4个模块. 下面对该系统进行建模, 设  $x_t$  是该系统在时刻  $t$  储存设备的能量

状态,  $u_t$ 是能量储存策略, 马尔科夫链 $r_t$ 代表着风能和太阳能由于阴天下雨、光照强弱、昼夜交替、风力强度大小等自然变化而引起的发电模式之间的随机切换. 这里设马尔可夫链 $r_t$ 的影响表现在风光发电的总电能 $P = P_{\text{wind}} + P_{\text{solar}}$ 与预测电量 $\tilde{P}$ 之间的差别, 即预测误差

$$\Delta P = P - \tilde{P},$$

设 $\Delta P$ 满足如下差分方程:

$$\Delta P(t+1) = a(r_t)\Delta P(t) + \sigma(r_t)\Delta P(t)\omega_t, \quad (40)$$

其中:  $r_t = 1$ 表示风力发电;  $r_t = 2$ 为光电;  $\omega_t$ 表示随机因素对 $\Delta P$ 的影响,  $\omega_t$ 是相互独立的随机变量且  $E[\omega_t] = 0, E[\omega_t^2] = 1$ .

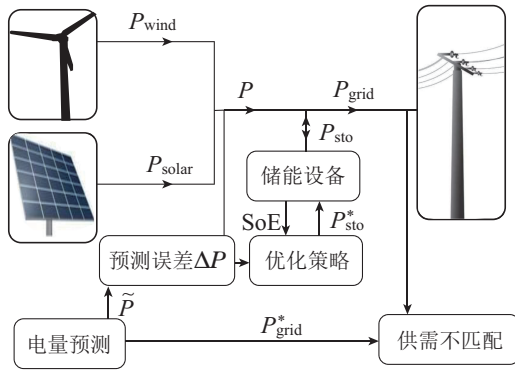


图1 风光储联合发电的微电网优化管理系统

Fig. 1 Microgrid optimization management system for wind-solar-energy-storage power plant

SoE( $t$ )是 $t$ 时刻电能储存设备(蓄电池)的能量状态,  $P_{\text{sto}}(t)$ 是 $t$ 时刻的风光发电设备对蓄电池的输入或输出电量, 若 $P_{\text{sto}}(t) > 0$ 表示发电量供应电网需求 $P_{\text{grid}}$ 的电量充足有结余并将多余电量存入蓄电池(充电过程), 若 $P_{\text{sto}}(t) < 0$ 表示发电量供应电网需求 $P_{\text{grid}}$ 的电量不足, 需要用蓄电池向电网补充供电(放电), 以满足电网需求, 若 $P_{\text{sto}}(t) = 0$ 表示发电量恰好满足电网需求, 即 $P_{\text{sto}}(t)$ 对应着电池充放电过程的调控策略. 因此, 电池在 $t$ 时刻的电能SoE( $t$ )满足如下方程:

$$\text{SoE}(t+1) = \text{SoE}(t) + P_{\text{sto}}(t). \quad (41)$$

优化目标是使得多出的预测电量 $\Delta P$ 能够尽可能地储存起来, 即预测误差 $\Delta P$ 与电量储存 $P_{\text{sto}}$ 之间的差别尽可能地小, 因此引入如下的优化目标:

$$\min_{P_{\text{sto}}} E \sum_{t=0}^T |\Delta P(t) - P_{\text{sto}}(t)|^2, \quad (42)$$

约束条件为式(40)–(41).

记

$$x = \begin{bmatrix} \text{SoE} \\ \Delta P \end{bmatrix}, u = P_{\text{sto}}, A(r_t) = \begin{bmatrix} a(r_t) & 0 \\ 0 & 1 \end{bmatrix},$$

$$B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, C(r_t) = \begin{bmatrix} \sigma(r_t) & 0 \\ 0 & 0 \end{bmatrix},$$

$$Q = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, S = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, R = 1,$$

则上述优化问题的目标函数为

$$J(u) = E \left[ \sum_{t=0}^{T-1} (x_t^T Q x_t + 2x_t^T S u_t + R u_t^2) \right] + E[x_T^T Q x_T], \quad (43)$$

其中 $u(T) = 0$ .

状态约束条件为

$$\begin{cases} x_{t+1} = A(r_t)x_t + B u_t + C(r_t)x_t \omega_t, \\ x_0 \in \mathbb{R}^2, r_t \in \mathcal{S} = \{1, 2\}, t = 0, 1, \dots, T-1, \end{cases} \quad (44)$$

根据例1的结论可知, 优化问题(43)–(44)的最优控制、Hamilton函数和伴随倒向随机差分方程分别由式(37)–(40)确定. 在本例中, 设

$$a(1) = 0.8, a(2) = 1.2, \sigma(1) = 0.02, \sigma(2) = 0.1,$$

Markov链的一步转移概率矩阵为

$$\pi = \begin{bmatrix} 0.6 & 0.4 \\ 0.7 & 0.3 \end{bmatrix},$$

$\omega_k$ 为高斯噪声, 即 $\omega_k$ 是相互独立的服从标准正态分布的随机变量, 相应的状态轨道与控制策略如图2所示.

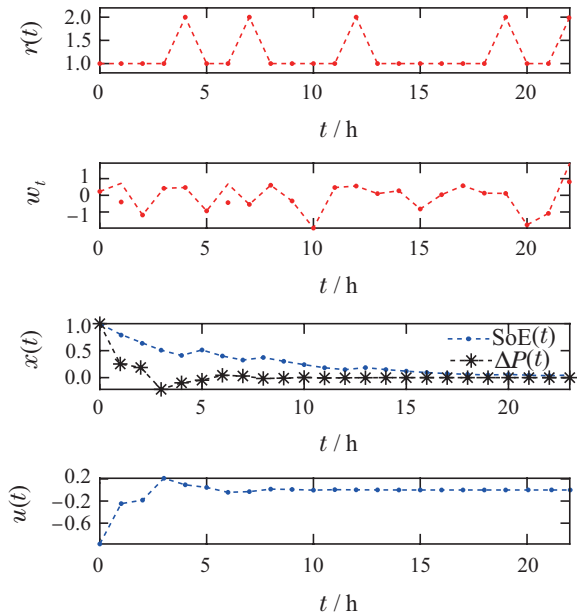


图2 Markov链、高斯噪声、状态轨道和控制轨道曲线

Fig. 2 Trajectories of Markov chain, Gaussian noise, system state and controller.

由图2可以看出, 在最优控制策略 $u_t$ 的作用下, 虽然系统(44)受到随机噪声 $\omega_t$ 和Markov跳过程 $r_t$ 的影



响, 预测误差 $\Delta P$ 和储能设备的状态SoE仍趋于平衡的, 从而达到控制目标.

## 5 结论

本文把离散时间随机最优控制问题中的针状变分法推广到一类同时含有Markov链和乘性噪声的离散时间非线性随机系统的最优控制问题, 并获得相应的最大值原理和动态规划原理. 利用针状变分法, 对状态方程进行一阶变分, 获得相应的线性差分方程. 通过由倒向随机差分方程刻画的伴随方程的适应解, 结合条件期望的平滑性, 给出并证明了同时含有Markov跳和乘性噪声的离散时间非线性随机最优控制系统的最大值原理, 并通过引入Hamilton函数获得了该最优控制问题的必要条件. 同时, 在引入辅助函数的基础上, 结合差分方程解的马氏性, 通过Hamilton-Jacobi-Bellman方程, 给出并证明了该最优控制问题的一个充分条件, 获得相应的动态规划原理. 最后, 通过一个具体例子说明所研究问题的实际意义和可行性.

## 参考文献:

- [1] BOLTJANSKII V G, GAMKRELIDZE R V, PONTRYAGIN L S. On the theory of optimal processes. *Doklady Akademii Nauk SSSR*, 1956, 110(1): 7 – 10.
- [2] YONG J, ZHOU X Y. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. New York: Springer-Verlag, 1999.
- [3] PONTRYAGIN L S, BOLTJANSKI V G, GAMKRELIDZE R V. *The Mathematical Theory of Optimal Control Processes*. New York: John Wiley, 1962.
- [4] KUSHNER H J, SCHWEPPE F C. A maximum principle for stochastic control systems. *Journal of Mathematical Analysis and Applications*, 1964, 8(2): 287 – 302.
- [5] KUSHNER H J. On the stochastic maximum principle: Fixed time of control. *Journal of Mathematical Analysis and Applications*, 1965, 11: 78 – 92.
- [6] KUSHNER H J. On the stochastic maximum principle with average constraints. *Journal of Mathematical Analysis and Applications*, 1965, 12(1): 13 – 26.
- [7] BISMUT J M. On optimal control of linear stochastic equations with a linear-quadratic criterion. *SIAM Journal on Control and Optimization*, 1977, 15(1): 1 – 4.
- [8] BISMUT J M. An introductory approach to duality in optimal stochastic control. *SIAM Review*, 1978, 20(1): 62 – 78.
- [9] BENSOUSSAN A. Stochastic maximum principle for distributed parameter systems. *Journal of the Franklin Institute*, 1983, 315(5/6): 387 – 406.
- [10] HAUSSMANN U G. Generalized solutions of the Hamilton-Jacobi equation of stochastic control. *SIAM Journal on Control and Optimization*, 1994, 32(3): 728 – 743.
- [11] PENG S G. A general stochastic maximum principle for optimal control problems. *SIAM Journal on Control and Optimization*, 1990, 28(4): 966 – 979.
- [12] PENG S G. Backward stochastic differential equations and applications to optimal control. *Applied Mathematics and Optimization*, 1993, 27(2): 125 – 144.
- [13] WU Z. Maximum principle for optimal control problem of fully coupled forward-backward stochastic systems. *Journal of Mathematics and System Science*, 1998, 11(3): 249 – 259.
- [14] WU Z. A general maximum principle for optimal control of forward-backward stochastic systems. *Automatica*, 2013, 49(5): 1473 – 1480.
- [15] WANG G C, ZHANG C H, ZHANG W H. Stochastic maximum principle for mean-field type optimal control under partial information. *IEEE Transactions on Automatic Control*, 2014, 59(2): 522 – 528.
- [16] CHANG D J, WU Z. Stochastic maximum principle for non-zero sum differential games of FBSDEs with impulse controls and its application to finance. *Journal of Industrial and Management Optimization*, 2015, 11(1): 27 – 40.
- [17] WANG G C, WU Z, XIONG J. A linear-quadratic optimal control problem of forward-backward stochastic differential equations with partial information. *IEEE Transactions on Automatic Control*, 2015, 60(11): 2904 – 2916.
- [18] HUANG H, WANG X R, LIU M J. A maximum principle for fully coupled forward-backward stochastic control system driven by Lévy process with terminal state constraints. *Journal of Systems Science and Complexity*, 2018, 31(4): 859 – 874.
- [19] MOON J. The risk-sensitive maximum principle for controlled forward-backward stochastic differential equations. *Automatica*, 2020, 120(3): 109069.
- [20] ZHANG T L, DENG F Q, SUN Y, et al. Fault estimation and fault-tolerant control for linear discrete time-varying stochastic systems. *Science China Information Sciences*, 2021, 64(10): 200201.
- [21] JIANG X S, ZHAO D Y. Event-triggered fault detection for nonlinear discrete-time switched stochastic systems: A convex function method. *Science China Information Sciences*, 2021, 64(10): 200204.
- [22] HUANG Z, WANG Y, WANG X R. A mean-field optimal control for fully coupled forward-backward stochastic control systems with Lévy processes. *Journal of Systems Science & Complexity*, 2022, 35(1): 205 – 220.
- [23] LIN X Y, ZHANG W H. A maximum principle for optimal control of discrete-time stochastic systems with multiplicative noise. *IEEE Transactions on Automatic Control*, 2015, 60(4): 1121 – 1126.
- [24] NI Y H, LI X, ZHANG J F. Mean-field stochastic linear-quadratic optimal control with Markov jump parameters. *Systems & Control Letters*, 2016, 93: 69 – 76.
- [25] HAN Y C, LI Z. Maximum principle of discrete stochastic control system driven by both fractional noise and white noise. *Discrete Dynamics in Nature and Society*, 2020, 2020: 1959050.
- [26] ZHOU Shaowei, CHEN Bing, LIU Hongxia. Finite-time stability for nonhomogeneous Markov jump systems with multiplicative noises. *Control and Decision*, 2018, 33(3): 565 – 570. (周绍伟, 陈兵, 刘洪霞. 带乘性噪声的非齐次Markov跳跃系统有限时间稳定性. *控制与决策*, 2018, 33(3): 565 – 570.)
- [27] STUTZER M. Persistence of averages in financial Markov switching models: A large deviations approach. *Physica A: Statistical Mechanics and Its Applications*, 2020, 553: 124237.
- [28] MALALLAH A, PEREZ H, DATTA-GUPTA A, et al. Multiscale data integration with Markov random fields and Markov chain Monte

- Carlo: A field application in the Middle East. *SPE Reservoir Evaluation and Engineering*, 2004, 7(6): 416 – 426.
- [29] COWAN R. Expected pattern frequency in Markov chain realizations conditional upon transition counts: Application of whittle's combinatoric formula to problems in DNA sequences. *Theory of Probability and Its Applications*, 2006, 37(2): 371 – 374.
- [30] YAKASIRI M, AVREL J, SHARMA S, et al. A stochastic approach for the state-wise forecast of wind speed using discrete-time Markov chain. *Region 10 Conference on Technology, Knowledge, and Society*. Kochi, India: IEEE, 2019: 575 – 580.
- [31] PENG S, BOUKAS E K. Control for Markovian jumping discrete-time systems with different forms of uncertainties. *Proceedings of the 1998 American Control Conference*. Philadelphia: IEEE, 1998: 728 – 732.
- [32] ZHANG X, ELLIOTT R, SIU T. A stochastic maximum principle for a Markov regime-switching jump-diffusion model and its application to finance. *SIAM Journal on Control and Optimization*, 2012, 50(2): 964 – 990.
- [33] WANG Jiagang. *Foundations of Modern Probability*. Shanghai: Fudan University Press, 2005.  
(汪嘉冈. 现代概率论基础. 上海: 复旦大学出版社, 2005.)
- [34] CHENG Qixiang. *Fundamentals of Real Variable Functions and Functional Analysis*. Beijing: Higher Education Press, 1983.  
(程其襄. 实变函数与泛函分析基础. 北京: 高等教育出版社, 1983.)
- [35] LATIMIER R L G, BOUEDEC E L, MONBET V. Markov switching autoregressive modeling of wind power forecast errors. *Electric Power Systems Research*, 2020, 189: 106641.

#### 作者简介:

**蔺香运** 副教授, 博士, 硕士生导师, 主要研究方向为随机最优控制和随机 $H_\infty$ 控制, E-mail: lxy9393@sina.com;

**王鑫瑞** 硕士研究生, 主要研究方向为随机最优控制和随机 $H_\infty$ 控制, E-mail: Wangxinrui0206@163.com;

**张维海** 教授, 博士, 博士生导师, 主要研究方向为随机控制理论及其应用, E-mail: w.hzhang@163.com.