山东大学学报(理学版)2021 年 10 月 第 56 卷 第 10 期
Journal of Shandong University (*Natural Science*), Vol.56, No.10, 2021

E-mail:xblxb@ sdu.edu.cn
http://lxbwk.njournal.sdu.edu.cn
Tel:+86-531-88366917

文章编号:1671-9352(2021)10-0001-10 DOI:10.6040/j.issn.1671-9352.9.2021.012

吴臻,山东大学数学学院教授,教育部“长江学者”特聘教授,国家杰出青年基金获得者,泰山学者攀登计划专家。现任山东大学副校长、泰山学堂常务副院长。目前担任国家自然科学基金数学天元基金学术领导小组成员,山东数学会理事长,国际控制理论权威期刊 *SIAM Journal on Control and Optimization*、国家基金委英文期刊 *Fundamental Research* 数学物理领域编委、SCI 学术期刊 *Statistics & Probability Letters*、国际学术期刊 *Probability, Uncertainty and Quantitative Risk*、*Partial Differential Equations and Applications* 的编委。研究领域涉及控制论、概率论和金融数学等,取得了一系列具有突破性和原创性的科研成果,2019 年荣获中国数学会第十七届陈省身数学奖,2018 年获山东省自然科学奖一等奖。主持国家基金委重点项目、山东省重大基础研究项目等。为国家"万人计划"首批科技创新领军人才入选者和科技部首批国家创新人才推进计划“金融数学”重点领域创新团队负责人,享受国务院政府特殊津贴。

正倒向随机最优控制问题的最大值原理:完全信息和部分信息

吴臻[1*],王光臣[2],李敏[1]
(1.山东大学数学学院, 山东 济南 250100; 2.山东大学控制科学与工程学院, 山东 济南 250061)

摘要:本文是一篇关于正倒向随机最优控制问题研究进展的综述论文。近 30 年来,正倒向随机控制系统的各种理论与应用研究得到了迅猛发展,取得了大量原创性科研成果,吸引了大批国际同行跟进研究。限于论文篇幅和作者意图,本文仅仅聚焦于正倒向随机最优控制问题的最大值原理这一主题,概述其最新研究进展及其在求解线性二次最优控制问题中的简单应用。

关键词:正倒向随机系统;一般最大值原理;最优滤波;倒向分离方法;线性二次控制
中图分类号:O211 **文献标志码**:A
引用格式:吴臻,王光臣,李敏.正倒向随机最优控制问题的最大值原理:完全信息和部分信息[J]. 山东大学学报(理学版),2021,56(10):1-10.

Maximum principle for optimal control of forward-backward stochastic system: full information and partial information

WU Zhen[1*], WANG Guang-chen[2], LI Min[1]
(1. School of Mathematics, Shandong University, Jinan 250100, Shandong, China; 2. School of Control Science and Engineering, Shandong University, Jinan 250061, Shandong, China)

收稿日期:2021-08-24; **网络出版时间**:2021-09-17 10:49:49
网络出版地址:http://kns.cnki.net/kcms/detail/37.1389.N.20210917.0911.002.html
基金项目:国家自然科学基金重点资助项目(11831010); 中港合作项目(61961160732); 国家自然科学基金杰出青年科学基金资助项目(61925306); 山东省自然基金重大项目(ZR2019ZD42, ZR2020ZD24); 山东省泰山学者攀登计划(TSPD20210302)
第一作者简介:吴臻(1971—), 男, 博士, 教授, 博士生导师, 研究方向为随机控制、正倒向随机微分方程、数理金融.
E-mail:wuzhen@ sdu.edu.cn
* 通信作者

Abstract: This paper reviews some progresses on forward-backward stochastic control system. In the past 30 years, various theories and applications of forward-backward stochastic control system have been developed rapidly and a large number of original scientific research results have been obtained, which have attracted many international peers to follow-up research. Limited to the length of this paper and the authors' emphasis, this paper only focuses on maximum principle for optimal control of forward-backward stochastic system, and summarizes its latest research progress and application in solving linear quadratic optimal control problem.
Key words: forward-backward stochastic system; general maximum principle; optimal filtering; backward separation approach; linear-quadratic control

1 引言

线性倒向随机系统产生于正向随机系统最优控制问题的研究过程[1],是变分方程的对偶,它与状态方程通过最大值条件耦合,构成一类正倒向随机系统,其最具代表性的数学结构是

$$\begin{cases}\mathrm{d}x(t)=b(t,\omega,x(t))\mathrm{d}t+\sigma(t,\omega,x(t))\mathrm{d}W_t,\\ -\mathrm{d}y(t)=f(t,\omega,x(t),y(t),z(t))\mathrm{d}t-z(t)\mathrm{d}W_t,\\ x(0)=a,\ y(T)=\Phi(x(T))。\end{cases}$$

该系统具有完美的动态结构和良好的数学性质,系数 b,σ,f 和 Φ 在非常宽泛的条件下即存在唯一解 $(x(\cdot),y(\cdot),z(\cdot))$,其中 $z(\cdot)$ 起着保证 $(x(\cdot),y(\cdot),z(\cdot))$ 具有适应性的重要作用,体现了该系统与正向随机系统在数学结构上的本质不同[2-10]。该系统的最优控制和最优滤波除了在控制理论领域具有重要的科学价值外,还在金融经济领域具有广阔的应用背景。这可从 Wang 和 Wu[11] 列举的例子中窥斑见豹。具体地,令某一资产的内在价值过程满足

$$\begin{cases}\mathrm{d}x(t)=[a+bx(t)+g(\pi(t))]\mathrm{d}t+\delta\mathrm{d}W_t,\\ x(0)=x\in\mathbf{R},\end{cases}$$

这里 a,b 和 δ 是常数,$g(\cdot)$ 是确定性函数,$W_\cdot$ 是标准布朗运动,$\pi(\cdot)$ 是经济因子,可以包括红利率、税率等。假设投资者的消费率为 $c(\cdot)$,终端回报为 $\Phi(x(T))$。根据倒向随机系统理论,投资者关于 $c(\cdot)$ 的递归效用 $y(\cdot)$ 满足

$$\begin{cases}-\mathrm{d}y(t)=f(t,y(t),z(t),c(t))\mathrm{d}t-z(t)\mathrm{d}W_t,\\ y(T)=\Phi(x(T)),\end{cases}$$

其中 $f(\cdot)$ 和 $\Phi(\cdot)$ 是满足某些条件的确定性函数。假定投资者的目标是找到一对 $\mathscr{F}_t$-适应的过程 $(\bar{c}(\cdot),\bar{\pi}(\cdot))$ 使得

$$y^{\bar{c},\bar{\pi}}(0)=\sup_{(c(\cdot),\pi(\cdot))}y(0)。$$

这就构成了一类完全信息下由正倒向随机系统驱动的最优控制问题,简称完全信息下的正倒向随机最优控制问题。

事实上,资产的内在价值过程往往不可直接观测,但可通过其价格过程

$$\mathrm{d}S(t)=S(t)[h(t,x(t),\pi(t))\mathrm{d}t+\sigma\mathrm{d}V_t]$$

被部分地观测到,其中 σ 是常数,$h(\cdot)$ 是确定性函数,$V(\cdot)$ 是与 $W(\cdot)$ 独立的标准布朗运动。令 $Y(\cdot)=\log S(\cdot)$,由伊藤公式可得

$$\mathrm{d}Y(t)=\left[h(t,x(t),\pi(t))-\frac{1}{2}\sigma^2\right]\mathrm{d}t+\sigma\mathrm{d}V_t。$$

如果记 $\mathscr{Y}_t=\sigma\{Y(s):0\leqslant s\leqslant t\}$,那么上述问题就变成了找到一对 $\mathscr{Y}_t$-适应的过程 $(\bar{c}(\cdot),\bar{\pi}(\cdot))$ 使得 $y(0)$ 取上确界,这就是一类部分(可观测)信息下的正倒向随机最优控制问题。

众所周知,最大值原理是求解最优控制的最重要研究工具之一。对于正向随机最优控制问题,Peng[12] 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,也被后世尊称为 Peng 最大值原理。对于正倒向随机最优控制问题,Peng[13] 首次研究了控制域为凸的情形,得到了最大值原理。Xu[14] 研究了 Peng[13] 的简化模型,在正向系统不含控制变量且控制域非凸的假定下得到了最大值原理。Wu[15] 给出了状

态完全耦合且控制域为凸情形下的最大值原理。1999 年，Peng[16] 提出了一个问题：怎样在控制域非凸且倒向生成元依赖于 z 的情形下获得一般最大值原理？如何回答该问题一直是困扰学术界的公开难题。直到 2013 年，Wu[17] 打破常规，提出嵌入方法，获得了一般最大值原理，回答了 Peng 的公开问题。稍后，Hu[18] 用对偶法重解了这个问题。动态规划原理是求解最优控制的另一个重要工具，关于其在正倒向随机控制系统框架下的进展，请参见 Wu 和 Yu[19] 以及 Nie，Shi 和 Wu[20]。

仔细观察不难发现，上述文献基本上都假设决策者能获得控制系统的所有信息，也即决策者可以完全观测到状态方程的噪声函数。事实上，由于受到外部噪声干扰，决策者在大多数情况下只能得到部分信息，因此，研究部分信息下的正倒向随机系统最优控制问题同样具有重要的理论价值和现实意义。与研究完全信息下的最优控制问题不同，这里遇到的一个额外的研究困难是，状态方程通常不具有高斯性，导致传统分离原理[21] 不再适用于将原问题转化为完全信息下的最优控制问题，从而无法得到最优控制，因此，迫切需要改变思维方式，提出新的、有效的转发求解方法。2008 年，受 Li 和 Tang[22]，Tang[23] 启发，Wang 和 Wu[24] 提出了倒向分离方法，得到了一类正倒向随机系统的最优滤波器，解决了一类正向线性二次最优控制问题。倒向分离方法的工作原理是，先通过最大值原理形式推出最优控制，然后借助正倒向随机系统的最优滤波器[25-26] 计算 Hamiltonian 系统的最优滤波，从而可适用于转化求解一大类非高斯系统驱动的最优控制问题，且避免了无穷维空间中繁杂的随机优化运算。例如：对于正倒向随机系统这类典型非高斯系统驱动的最优控制问题，综合运用倒向分离方法和变分法，Wu[27] 在正向扩散系数包含控制变量且控制域为凸的情形下得到了最大值原理；Wang 和 Wu[11] 在正向扩散系数不含控制变量且控制域非凸的情形下得到了最大值原理；Wang，Wu 和 Xiong[28] 在状态和观测具有相关噪声且控制域为凸的情形下得到了最大值原理。对于部分可观测的线性二次正倒向随机最优控制问题，Wang，Wu 和 Xiong[29] 综合运用倒向分离方法和分解技术，求出了解析的最优控制。

本文其他章节的安排如下：第 2 节介绍一类完全信息下的正倒向随机最优控制问题，通过嵌入方法，给出了一般最大值原理。第 3 节概述部分信息下正倒向随机控制问题的最大值原理。第 4 节简述部分可观测正倒向线性二次最优控制问题的求解。第 5 节列出总结与展望。

符号说明。 $\mathbf{R}^n$ 表示 n-维欧几里得空间，$|\cdot|$ 表示矩阵所有分量平方和的平方根。$\langle\cdot,\cdot\rangle$ 表示欧几里得空间中任意两个向量的内积。tr 表示矩阵的迹。上标 $\top$ 表示矩阵或者向量的转置。令 T 是固定常数，$(\Omega,\mathscr{F},(\mathscr{F})_{0\leqslant t\leqslant T},P)$ 是完备概率空间，$(W_\cdot)$ 是定义在该空间上的 $\mathbf{R}$-值标准布朗运动，$\mathscr{F}_t$ 是自然信息流，$\mathscr{F}=\mathscr{F}_T$。如果 $\{\mathscr{F}_t\}_{0\leqslant t\leqslant T}$-适应的随机过程 $g(\cdot):[0,T]\times\Omega\to\mathbf{R}^n$ 满足 $\mathbb{E}\int_0^T|g(t)|^2\mathrm{d}t<\infty$，记 $g(\cdot)\in L^2_{\mathscr{F}}(0,T;\mathbf{R}^n)$。如果 $\mathscr{F}_T$ 可测的随机变量 $\xi:\Omega\to\mathbf{R}^n$ 满足 $\mathbb{E}\,\xi^2<\infty$，记 $\xi\in L^2_{\mathscr{F}_T}(\Omega,\mathbf{R}^n)$。如果 $g(\cdot):[0,T]\to\mathbf{R}$ 是连续函数，记 $g(\cdot)\in C(0,T;\mathbf{R})$。

2　完全信息下正倒向随机系统的最大值原理

考虑受控的正倒向随机系统

$$\begin{cases}\mathrm{d}x(t)=b(t,x(t),v(t))\mathrm{d}t+\sigma(t,x(t),v(t))\mathrm{d}W_t,\\-\mathrm{d}y(t)=f(t,x(t),y(t),z(t),v(t))\mathrm{d}t-z(t)\mathrm{d}W_t,\\x(0)=a,\ y(T)=\Phi(x(T)),\end{cases}\tag{1}$$

其中 $(x(\cdot),y(\cdot),z(\cdot))\in\mathbf{R}^n\times\mathbf{R}^m\times\mathbf{R}^m$ 是(1)的解，$v(\cdot)$ 是控制变量，b,σ,f,Φ 是合适维数的泛函。令 $\mathscr{U}$ 是 $\mathbf{R}^k$ 的一个非空子集，不必凸。控制变量 $v(\cdot)$ 称为容许的，如果 $v(t)$ 是一个取值于 $\mathscr{U}$ 的 $\mathscr{F}_t$-适应的过程且满足 $\sup_{0\leqslant t\leqslant T}\mathbb{E}\,|v(t)|^\iota\leqslant\infty$，$\forall\,\iota=1,2,\cdots$ 所有容许控制变量的全体记为 $\mathscr{U}_{ad}$。$\forall\,v(\cdot)\in\mathscr{U}_{ad}$，引入代价泛函

$$\mathscr{J}(v(\cdot))=\mathbb{E}\left[\int_0^T l(t,x(t),y(t),z(t),v(t))\mathrm{d}t+g(x(T))+\gamma(y(0))\right],\tag{2}$$

其中 l,g 和 γ 是合适维数的泛函，它们满足

假设 1　f,b,σ,Φ,l,g 和 γ 关于 (x,y,z) 二次连续可微。f,b,σ,Φ 关于 (x,y,z) 的一阶和二阶导数有界。$f,b,\sigma,\Phi,l_x,l_y,l_z,g_x$ 和 γ_y 关于 (x,y,z,v) 线性增长，关于 (t,v) 连续。$l_{xx},l_{xy},l_{xz},l_{yz},l_{yy},l_{zz},h_{xx}$ 和 γ_{yy} 有界。

问题（A） 找到一个 $v^*(\cdot)\in\mathscr{U}_{ad}$使得代价泛函(2)最小化。

如果存在这样的 $v^*(\cdot)$,那么称其为最优控制,称相应的$(x^*(\cdot),y^*(\cdot),z^*(\cdot))$为最优状态轨迹。

推导问题(A)的最大值原理的通常做法是,针对倒向随机系统,找到合适的二阶伴随方程,然而这个倒向系统十分复杂,国际同行寻找多年未果。2013 年,受 Kohlmann 和 Zhou[30] 和 Lim 和 Zhou[31] 启发,Wu[17] 打破常规,提出嵌入方法,巧妙规避了难以找到合适伴随方程所导致的技术障碍,建立了一般最大值原理,从而回答了前文提到的 Peng 公开问题。具体求解过程概述如下。

引入辅助

问题（B） 找到一个由 $y_0\in\mathbf{R}^m$, $u(\cdot)\in L^2_{\mathscr{F}}(0,T;\mathbf{R}^m)$, $v(\cdot)\in\mathscr{U}_{ad}$构成的三元组,使其最小化

$$\mathscr{J}(y_0,u(\cdot),v(\cdot))=\mathbb{E}\left[\int_0^T l(t,x(t),y(t),y(t),v(t))\mathrm{d}t+g(x(T))+\gamma(y_0))\right],$$

其中$(x(\cdot),y(\cdot),z(\cdot))$满足正向随机控制系统

$$\begin{cases}\mathrm{d}x(t)=b(t,x(t),v(t))\mathrm{d}t+\sigma(t,x(t),v(t))\mathrm{d}W_t,\\ -\mathrm{d}y(t)=f(t,x(t),y(t),u(t),v(t))\mathrm{d}t-u(t)\mathrm{d}W_t,\\ x(0)=a,\ y(0)=y_0\end{cases}\tag{3}$$

和最优状态约束$\mathbb{E}\,|y(T)-\Phi(x(T))|^2=0$。

显然,问题(A)已被嵌入到问题(B),而问题(B)属于一类带状态约束的正向随机最优控制问题,采用 Peng[12] 介绍的二阶变分技术,即可推出其最优控制$(y_0^*,u^*(\cdot),v^*(\cdot))$满足的一般最大值原理,从而得到问题(A)的最优控制 $v^*(\cdot)$满足的一般最大值原理。故此只需研究问题(B)的一般最大值原理即可。$\forall\, u(\cdot),v(\cdot)\in\mathscr{U}_{ad}$或$L^2_{\mathscr{F}}(0,\mathscr{T};\mathbf{R}^m)$,定义 $\mathscr{U}_{ad}$和$L^2_{\mathscr{F}}(0,T;\mathbf{R}^m)$中的距离为

$$d(u(\cdot),v(\cdot))=\mathbb{E}\,[\,\mathrm{mes}\{t\in[0,T],\ u(t)\neq v(t)\}\,],$$

其中 mes 表示 Lebesgue 测度。易证$(L^2_{\mathscr{F}}(0,T;\mathbf{R}^m),d(\cdot,\cdot))$和$(\mathscr{U}_{ad},d(\cdot,\cdot))$都是完备距离空间。假设$(y_0^*,u^*(\cdot),v^*(\cdot))$是问题(B)的最优控制,$(x^*(\cdot),y^*(\cdot))$是(3)中相应的最优状态轨迹,满足 $y^*(T)=\Phi(x^*(T))$。易证 $\forall\, y_0\in\mathbf{R}^m$, $u(\cdot)\in L^2_{\mathscr{F}}(0,T;\mathbf{R}^m)$,$v(\cdot)\in\mathscr{U}_{ad}$,

$$\mathscr{J}(y_0^*,u^*(\cdot),v^*(\cdot))\leqslant\mathscr{J}(y_0,u(\cdot),v(\cdot))。$$

记$(x^v(\cdot),y^v(\cdot))$是相应于(3)的解。注意$(x^v(\cdot),y^v(\cdot))$不一定满足 $y^v(T)=\Phi(x^v(T))$。定义新的代价泛函

$$\begin{aligned}\mathscr{J}_\rho(y_0,u(\cdot),v(\cdot))=\{&[\mathscr{J}(y_0,u(\cdot),v(\cdot))-\mathscr{J}(y_0^*,u^*(\cdot),v^*(\cdot))+\rho]^2,\\ &+[\mathbb{E}\,|y^v(T)-\Phi(x^v(T))|^2]^2\}^{\frac{1}{2}},\end{aligned}\tag{4}$$

其中 $\rho>0$ 是常数。显然,(3)和(4)构成一类状态不受约束的随机最优控制问题,简记为问题(C)。接下来利用 Ekeland 变分原理,可得到所需的一般最大值原理。具体推导细节从略。

引入 Hamiltonian 函数

$$\begin{aligned}H(t,x,y,u,v,p,q,k,r,\theta)=&\langle p,b(t,x,v)\rangle-\langle q,f(t,x,y,u,v)\rangle\\ &+\langle k,\sigma(t,x,v)\rangle+\langle r,u\rangle+\theta l(t,x,y,u,v)\end{aligned}$$

和 4 个伴随方程

$$\begin{cases}-\mathrm{d}p(t)=[b_x^\top p(t)-f_x^\top q(t)+\sigma_x^\top k(t)+\theta l_x]\mathrm{d}t-k(t)\mathrm{d}W_t,\\ \mathrm{d}q(t)=[f_y^\top q(t)-\theta l_y]\mathrm{d}t+r(t)\mathrm{d}W_t,\\ p(T)=\theta g_x(x^*(T)),q(T)=0,\end{cases}\tag{5}$$

$$\begin{cases}\mathrm{d}P_2(t)=[P_2(t)f_y+f_y^\top P_2(t)-H_{yy}]\mathrm{d}t+Q_2(t)\mathrm{d}W_t,\\ P_2(T)=2\theta_T I_m,\end{cases}\tag{6}$$

$$\begin{cases}\mathrm{d}P_3(t)=[P_3(t)f_y-P_3(t)g_x+2f_x^\top P_2(t)-\sigma_x^\top Q_3(t)-H_{xy}]\mathrm{d}t+Q_3(t)\mathrm{d}W_t,\\ P_3(T)=0,\end{cases}\tag{7}$$

$$\begin{cases}-\mathrm{d}P_1(t)=[b_x^\top P_1(t)+P_1(t)b_x+\sigma_x^\top P_1(t)\sigma_x+\sigma_x^\top Q_1(t)+Q_1(t)\sigma_x-P_3(t)f_x+H_{xx}]\mathrm{d}t-Q_1(t)\mathrm{d}W_t,\\ P_1(T)=\theta g_{xx}+2\theta_T\Phi_x^\top(x^*(T))\Phi_x(x^*(T)),\end{cases}\tag{8}$$

其中 $\phi_x=\phi_x(t,x^*(t),v^*(t))$，$\varphi=\varphi_x(t,x^*(t),y^*(t),z^*(t),v^*(t))$，$\phi=b,\sigma,g_x,\sigma_x$，$\varphi=f,f_x,f_y,l,l_x,l_y$。$L=L(t,x(t),y(t),z^*(t),v^*,p(t),q(t),k(t),r(t),\theta)$，$L=H_{xx},H_{yy},H_{xy}$。

定理 1　在假设 1 下，若 $(v^*(\cdot),x^*(\cdot),y^*(\cdot),z^*(\cdot))$ 是问题(A)的最优解且 $y_0^*=y^*(0)$，则存在两个参数 θ 和 θ_T 满足 $\theta^2+\theta_T^2=1$，$\theta\geqslant 0$，$\theta_T\geqslant 0$，使得 $\forall\, y_0\in\mathbf{R}^m$，$u\in\mathbf{R}^m$ 和 $v\in\mathscr{U}$，有

$$
\begin{aligned}
&H(\tau,x^*(\tau),y^*(\tau),u,v,p(\tau),q(\tau),k(\tau)-P_1(\tau)\sigma(\tau,x^*(\tau),v^*(\tau))\\
&\quad -P_3(\tau)u^*(\tau),r(\tau)-P_2(\tau)u^*(\tau)-P_3^{\top}(\tau)\sigma(\tau,x^*(\tau),v^*(\tau),\theta)\\
&\quad +\frac{1}{2}\mathrm{tr}[\sigma(\tau,x^*(\tau),v^*(\tau))\sigma^{\top}(\tau,x^*(\tau),v^*(\tau))P_1(\tau)]+\frac{1}{2}\mathrm{tr}[u(\tau)u^{\top}(\tau)P_2(\tau)]\\
&\quad +\frac{1}{2}\mathrm{tr}[\sigma^{\top}(\tau,x^*(\tau),v^*(\tau))P_3(\tau)u(\tau)]+\langle\theta\gamma_y(y_0^*),y_0\rangle+\langle q(0),y_0\rangle\\
&\geqslant H(\tau,x^*(\tau),y^*(\tau),u^*(\tau),v^*(\tau),p(\tau),q(\tau),k(\tau)-P_1(\tau)\sigma(\tau,x^*(\tau),v^*(\tau))\\
&\quad -P_3(\tau)u^*(\tau),r(\tau)-P_2(\tau)u^*(\tau)-P_3^{\top}(\tau)\sigma(\tau,x^*(\tau),v^*(\tau),\theta)\\
&\quad +\frac{1}{2}\mathrm{tr}[\sigma(\tau,x^*(\tau),v^*(\tau))\sigma^{\top}(\tau,x^*(\tau),v^*(\tau))P_1(\tau)]+\frac{1}{2}\mathrm{tr}[u^*(\tau)(u^*(\tau))^{\top}P_2(\tau)]\\
&\quad +\frac{1}{2}\mathrm{tr}[\sigma^{\top}(\tau,x^*(\tau),v^*(\tau))P_3(\tau)u^{\top}(\tau)]+\langle\theta\gamma_y(y_0^*),y_0^{\top}\rangle+\langle q(0),y_0^*\rangle,
\end{aligned}
$$

其中 $(p(\cdot),k(\cdot),q(\cdot))$，$(P_1(\cdot),Q_1(\cdot))$，$(P_2(\cdot),Q_2(\cdot))$ 和 $(P_3(\cdot),Q_3(\cdot))$ 是(5)—(8)的解。

特别地，若只考虑正向系统，即 $\Phi=f=y(\cdot)=z(\cdot)=\gamma(\cdot)\equiv 0$ 的情形，令 $\theta_T=0$ 和 $\theta=1$，则 $q(\cdot)=r(\cdot)=P_2(\cdot)=Q_2(\cdot)=P_3(\cdot)=0=Q_3(\cdot)=0$，定理 1 即为 Peng[12] 获得的一般最大值原理。若考虑正倒向系统且正向扩散系数不含控制，即 $\sigma(t,x,v)=\sigma(t,x)$，令 $\theta_T=0$，$\theta=1$，$q(0)=-\gamma_y(y_0^*)$，$r(t)=f_z^{\top}(t,x^*(t),y^*(t),z^*(t),v^*(t))q(t)-l_z(t,x^*(t),y^*(t),z^*(t),v^*(t))$，$0\leqslant t\leqslant T$，定理 1 即为 Xu[14] 的情形。类似，如果控制域是凸的且 σ 包含 v，则定理 1 即为 Peng[13] 的结果。

以上内容主要取自 Wu[17]。关于最大值原理的其他进展，请参见[32-34]。

3　部分信息下正倒向随机系统的最大值原理

今后的两节仅考虑一维控制系统，多维的情况可类似处理。假设 $(W_\cdot,Y(\cdot))$ 是取值于 $\mathbf{R}^2$ 的标准布朗运动。记 $\{\mathscr{F}_t^W\}_{0\leqslant t\leqslant T}$ 和 $\{\mathscr{Y}_t\}_{0\leqslant t\leqslant T}$ 分别为由 $W_\cdot$ 和 $Y(\cdot)$ 生成的信息流，P_W 和 P_Y 分别是 $C(0,T;\mathbf{R})$ 上的概率测度。令 $\Omega=C(0,T;\mathbf{R})\times C(0,T;\mathbf{R})$，$\mathscr{F}_t=\mathscr{F}_t^W\otimes\mathscr{Y}_t$，$P=P_W\times P_Y$。

考虑受控的正倒向随机系统

$$
\begin{cases}
\mathrm{d}x(t)=b(t,x(t),v(t))\mathrm{d}t+\sigma(t,x(t),v(t))\mathrm{d}W_t+\tilde{\sigma}(t,x(t),v(t))\mathrm{d}\widetilde{W}_t^v,\\
-\mathrm{d}y(t)=f(t,x(t),y(t),z(t),\tilde{z}(t),v(t))\mathrm{d}t-z(t)\mathrm{d}W_t-\tilde{z}(t)\mathrm{d}\widetilde{W}_t^v,\\
x(0)=a,\ y(T)=\Phi(x(T)),
\end{cases}
\tag{9}
$$

其中 $v(\cdot)$ 是取值于凸集 $U\subseteq\mathbf{R}$ 的控制变量，$(x(\cdot),y(\cdot),z(\cdot),\tilde{z}(\cdot))\in\mathbf{R}\times\mathbf{R}\times\mathbf{R}\times\mathbf{R}$ 是状态过程。$\widetilde{W}_\cdot^v$ 是依赖于 v 的随机过程，$b,\sigma,\tilde{\sigma},f$ 和 Φ 是合适维数的泛函。假设 $(x(\cdot),y(\cdot),z(\cdot))$ 不能被直接观测到，所能观测到的是与之相关的过程 $Y(\cdot)$，它满足

$$
\begin{cases}
\mathrm{d}Y(t)=h(t,x(t))\mathrm{d}t+\mathrm{d}\widetilde{W}_t^v,\\
Y(0)=0,
\end{cases}
\tag{10}
$$

其中 h 是合适维数的泛函。控制变量 $v(\cdot):[0,T]\times C(0,T;\mathbf{R})\to U$ 称为容许的，如果 $v(t)$ 是 $\mathscr{Y}_t$-适应的过程且满足 $\sup_{0\leqslant t\leqslant T}\mathbb{E}\,|v(t)|^8<\infty$。所有容许控制变量的集合记为 $\mathscr{U}_{ad}$。

假设 2　$b,\sigma,\tilde{\sigma},\Phi,f$ 和 h 关于 (x,v)，$(x,y,z,\tilde{z},v)$ 和 x 连续可微，偏导数 $b_x,b_v,\sigma_x,\sigma_v,\tilde{\sigma}_x,\tilde{\sigma}_v,h_x,f_x,f_y,f_z,f_{\tilde{z}},f_v$ 和 Φ_x 一致有界。$\forall\,(t,x,v)\in[0,T]\times\mathbf{R}\times U$，存在一个常数 K，使得 $|\tilde{\sigma}(t,x,v)|+|h(t,x)|\leqslant K$。

将(10)代入(9)可得

$$
\begin{cases}
dx(t)=[b(t,x(t),v(t))-\tilde{\sigma}(t,x(t),v(t))h(t,x(t))]dt \\
\qquad +\sigma(t,x(t),v(t))dW_t+\tilde{\sigma}(t,x(t),v(t))dY(t), \\
-dy(t)=f(t,x(t),y(t),z(t),\tilde{z}(t),v(t))dt-z(t)dW_t-\tilde{z}(t)d\widetilde{W}_t^v, \\
x(0)=a,\ y(T)=\Phi(x(T))。
\end{cases}
\tag{11}
$$

$\forall v(\cdot)\in\mathscr{U}_{ad}$,(11)在假设2下存在唯一解$(x^v(\cdot),y^v(\cdot),z^v(\cdot),\tilde{z}^v(\cdot))$。

定义 $dP^v=Z^v(t)dP$,其中

$$
\begin{cases}
dZ^v(t)=Z^v(t)h(t,x^v(t))dY(t), \\
Z^v(0)=1。
\end{cases}
\tag{12}
$$

根据 Girsanov 定理及假设2,P^v 是一个新概率,$(W_\cdot,\widetilde{W}_\cdot^v)$是一个定义在$(\Omega,\mathscr{F},\{\mathscr{F}_t\}_{0\leqslant t\leqslant T},P^v)$上的二维标准布朗运动。

引入代价泛函

$$
J(v(\cdot))=\mathbb{E}^v\left[\int_0^T l(t,x^v(t),y^v(t),z^v(t),\tilde{z}^v(t),v(t))dt+g(x^v(T))+\gamma(y^v(0))\right],
\tag{13}
$$

其中 l,g 和 γ 是合适维数的泛函,$\mathbb{E}^v$表示概率空间$(\Omega,\mathscr{F},(\mathscr{F})_{0\leqslant t\leqslant T},P^v)$上的数学期望。

假设 3 函数 l,g 和 γ 分别关于$(x,y,z,\tilde{z})$,x 和 y 连续可微且满足

$$
\mathbb{E}^v\left[\int_0^T l(t,x^v(t),y^v(t),z^v(t),\tilde{z}^v(t),v(t))dt+g(x^v(T))+\gamma(y^v(0))\right]<\infty。
$$

问题(P) 找到一个容许控制 $u(\cdot)\in\mathscr{U}_{ad}$使其代价泛函(13)最小化。

如果这样的$u(\cdot)$存在,称之为问题(P)的最优控制,称相应的(1)的解为$(x(\cdot),y(\cdot),z(\cdot),\tilde{z}(\cdot))$为最优状态轨迹。据此,简记 $\widetilde{W}(\cdot)=\widetilde{W}^u(\cdot)$,$Z(\cdot)=Z^u(\cdot)$。由 Bayes 公式,(13)可改写为

$$
J(v(\cdot))=\mathbb{E}^v\left[\int_0^T Z^v(t)l(t,x^v(t),y^v(t),z^v(t),\tilde{z}^v(t),v(t))dt+Z^v(T)g(x^v(T))+\gamma(y^v(0))\right]。
\tag{14}
$$

如前所述,传统分离原理不适用于处理该问题,然而倒向分离方法仍然有效。Wang,Wu 和 Xiong[28]综合运用倒向分离方法和变分方法,得到了最大值原理,概述如下。

假设 4 (ⅰ)对于使得 $t+\tau\in[0,T]$的任意 t,τ 和 $\mathscr{Y}_t$-可测的有界随机变量 ν,构造控制变量 $v(s)=\nu I_{t,t+\tau}(s)\in U,\ s\in[0,T]$。

(ⅱ)对于任意 $\mathscr{Y}_s$-可测的有界随机过程 $v(s)$,$s\in[0,T]$,存在 $\varepsilon>0$ 使得 $\forall\varepsilon\in(-\varepsilon,\varepsilon)$,有 $u(\cdot)+\varepsilon v(\cdot)\in\mathscr{U}_{ad}$成立。

采用类似于第2节的记号,引入伴随方程

$$
\begin{cases}
-dP(t)=l dt-Q(t)dW_t-\widetilde{Q}(t)d\widetilde{W}_t, \\
P(T)=g(x(T))
\end{cases}
\tag{15}
$$

和

$$
\begin{cases}
dp(t)=[f_y p(t)-l_y]dt+[f_z p(t)-l_z]dW_t \\
\qquad +[(f_{\tilde{z}}-h)p(t)-l_{\tilde{z}}]d\widetilde{W}_t, \\
-dq(t)=\{[b_x-\tilde{\sigma}h_x]q(t)+\sigma_x k(t)+\tilde{\sigma}_x\tilde{k}(t) \\
\qquad +h_x\widetilde{Q}(t)-f_x p(t)+l_x]dt-k(t)dW_t-\tilde{k}(t)d\widetilde{W}_t, \\
p(0)=-\gamma_y(y(0)),\ p(T)=g_x(x(T))-\Phi_x(x(T))p(T)。
\end{cases}
\tag{16}
$$

这里引入(15)的目的是用来对偶处理因$(x(\cdot),y(\cdot),z(\cdot))$部分可观测所导致的(12),此伴随方程在完全信息下随机优化控制问题研究中是不需要的。

定理 2(最大值原理) 在假设2,3和4下,如果$u(\cdot)$是$J(v(\cdot))$的局部最小值,且$\forall v(\cdot)\in\mathscr{U}_{ad}$,函数$g,g_x\in L^2_{\mathscr{F}}(\Omega;\mathbf{R})$,$l,l_x,l_y,l_z,l_{\tilde{z}},l_v\in L^2_{\mathscr{F}}(0,T;\mathbf{R})$,(15)和(16)存在唯一解,那么$u(\cdot)$满足

$$
\mathbb{E}^u[H_v(t,x,y,z,\tilde{z},u,p,q,k,\tilde{k},Q,\widetilde{Q})\mid\mathscr{Y}_t]\geqslant 0,
$$

其中

$$
\begin{aligned}
H(t,x,y,z,\tilde{z},v,p,q,k,\tilde{k},Q,\widetilde{Q})=&\,b(t,x,v)q+\sigma(t,x,v)k+\tilde{\sigma}(t,x,v)\tilde{k}+h(t,x)\widetilde{Q} \\
&-(f(t,x,y,z,\tilde{z},v)-h(t,x)\tilde{z})p+l(t,x,y,z,v)。
\end{aligned}
$$

特别地，若(9)不含有 $\widetilde{W}^v$，则定理 2 即为 Wu[27] 的结果；若 $f=\Phi=0$，则定理 2 退化为正向随机控制系统的情形[23,36-37]。

以上内容主要取自 Wang，Wu 和 Xiong[28]。关于部分可观测正倒向随机最优控制的其他进展可参见文献[35]。

4　部分观测线性二次最优控制问题

假设$(W_\cdot,\overline{W}_\cdot)$是 $\mathbf{R}^2$ 中的标准布朗运动，$\mathscr{F}_t=\sigma\{W_s,\overline{W}_s;\ 0\leqslant s\leqslant t\}$。$\forall\, v(\cdot)\in L^2_{\mathscr{F}}(0,T;\mathbf{R})$，介绍受控的状态方程

$$\begin{cases}\mathrm{d}x^v(t)=[a(t)x^v(t)+b(t)v(t)+\bar{b}(t)]\mathrm{d}t+c(t)\mathrm{d}W_t+\bar{c}(t)\mathrm{d}\overline{W}_t,\\ -\mathrm{d}y^v(t)=[A(t)x^v(t)+B(t)y^v(t)+C(t)z^v(t)+\overline{C}(t)\bar{z}^v(t)\\ \qquad\qquad +D(t)v(t)+\overline{D}(t)]\mathrm{d}t-z^v(t)\mathrm{d}W_t-\bar{z}^v(t)\mathrm{d}\overline{W}_t,\\ x^v(0)=e_0,\ y^v(T)=Fx^v(T)+G,\end{cases}\tag{17}$$

观测方程

$$\begin{cases}\mathrm{d}Y^v(t)=[f(t)x^v(t)+g(t)]\mathrm{d}t+h(t)\mathrm{d}W_t,\\ Y^v(0)=0,\end{cases}\tag{18}$$

和代价泛函

$$\begin{aligned}J(v(\cdot))=\frac{1}{2}\mathbb{E}\int_0^T[&L(t)x^v(t)^2+R(t)v(t)^2+2l(t)x^v(t)+2o(t)y^v(t)\\ &+2r(t)v(t)]\mathrm{d}t+Mx^v(T)^2+2mx^v(T)+Ny^v(0)^2+2ny^v(0),\end{aligned}\tag{19}$$

其中系数满足

假设 5　$a(\cdot),b(\cdot),\bar{b}(\cdot),c(\cdot),\bar{c}(\cdot),f(\cdot),g(\cdot),h(\cdot),1/h(\cdot),A(\cdot),B(\cdot),C(\cdot),\overline{C}(\cdot),D(\cdot)$和$\overline{D}(\cdot)$都是一致有界的确定性函数。$e_0$ 和 F 是常数，$G\in L^2_{\mathscr{F}_T}(\Omega;\mathbf{R})$。

假设 6　$L(\cdot)\geqslant 0$，$R(\cdot)>0$，$1/R(\cdot),l(\cdot),o(\cdot)$和 $r(\cdot)$是一致有界的确定性函数。$M\geqslant 0$，$N\geqslant 0$，m 和 n 是常数。

问题（LQC）　找到一个 $u(\cdot)\in\mathscr{U}_{ad}$最小化代价泛函(19)。

如果这样的 $u(\cdot)$存在，称 $u(\cdot)$是最优控制，相应的$(x^u(\cdot),y^u(\cdot),z^u(\cdot),\bar{z}^u(\cdot))$是最优状态轨迹，$Y^u(\cdot)$是最优观测。简记$(x(\cdot),y(\cdot),z(\cdot),\bar{z}(\cdot))=(x^u(\cdot),y^u(\cdot),z^u(\cdot),\bar{z}^u(\cdot))$，$Y(\cdot)=Y^u(\cdot)$。

问题(LQC)看起来简单，但传统的分离原理和定理 2 都不适用。另一个基本的研究困难是，观测数据 Y^v和控制变量 $v(\cdot)$循环依赖，导致经典的变分法也不能直接应用，迫切需要给出新的求解技术。受 Bensoussan[37] 启发，Wang，Wu 和 Xiong[29] 提出一种分解转化技术，解耦了观测与控制之间的循环依赖，综合运用倒向分离方法和 Riccati 方程，给出了问题的解析结果，概述如下。

引入不受控的状态过程

$$\begin{cases}\mathrm{d}x^0(t)=a(t)x^0(t)\mathrm{d}t+c(t)\mathrm{d}W_t+\bar{c}(t)\mathrm{d}\overline{W}_t,\\ -\mathrm{d}y^0(t)=[A(t)x^0(t)+B(t)y^0(t)+C(t)z^0(t)+\overline{C}(t)\bar{z}^0(t)]\mathrm{d}t\\ \qquad\qquad -z^0(t)\mathrm{d}W_t-\bar{z}^0(t)\mathrm{d}\overline{W}_t,\\ x^0(0)=e_0,\ y^0(T)=Fx^0(T),\end{cases}\tag{20}$$

和观测方程

$$\begin{cases}\mathrm{d}Y^0(t)=f(t)x^0(t)\mathrm{d}t+h(t)\mathrm{d}W_t,\\ Y^0(0)=0。\end{cases}\tag{21}$$

$\forall\, v(\cdot)\in L^2_{\mathscr{F}}(0,T;\mathbf{R})$，引入受控的状态方程和观测方程

$$\begin{cases}\dot{x}^1(t)=a(t)x^1(t)+b(t)v(t)+\bar{b}(t),\\ -\mathrm{d}y^1(t)=[A(t)x^1(t)+B(t)y^1(t)+C(t)z^1(t)+\overline{C}(t)\bar{z}^1(t)\\ \qquad\qquad +D(t)v(t)+\overline{D}(t)]\mathrm{d}t-z^1(t)\mathrm{d}W_t-\bar{z}^1(t)\mathrm{d}\overline{W}_t,\\ x^1(0)=0,\ y^1(T)=Fx^1(T)+G\end{cases}\tag{22}$$

和

$$\begin{cases}\dot{Y}^1(t)=f(t)x^1(t)+g(t),\\ Y^1(0)=0。\end{cases} \tag{23}$$

显然,(20)—(23)存在唯一解。定义

$$x^v(t)=x^0(t)+x^1(t),\ y^v(t)=y^0(t)+y^1(t),$$
$$z^v(t)=z^0(t)+z^1(t),\ \bar{z}^v(t)=\bar{z}^0(t)+\bar{z}^1(t),$$
$$Y^v(t)=Y^0(t)+Y^1(t)。$$

由伊藤公式,易知如上定义的$(x^v(\cdot),y^v(\cdot),z^v(\cdot),\bar{z}^v(\cdot))$和$Y^v(\cdot)$分别是(17)和(18)的唯一解。

设$\mathscr{F}_t^{Y^v}=\sigma\{Y_s^v;\ 0\leqslant s\leqslant t\}$, $\mathscr{F}_t^{Y^0}=\sigma\{Y_s^0;\ 0\leqslant s\leqslant t\}$, $\mathscr{U}_{ad}^0=\{v(\cdot)\mid v(t)$是$\mathscr{F}_t^{Y^0}$-适应的过程满足$\mathbb{E}\sup_{0\leqslant t\leqslant T}v(t)^2<\infty\}$。如果$v(\cdot)\in\mathscr{U}_{ad}^0$且关于$\mathscr{F}_t^{Y^v}$-适应,那么称其为容许的。所有容许控制变量的集合记为$\mathscr{U}_{ad}$。显然,$\mathscr{U}_{ad}$是$\mathscr{U}_{ad}^0$的子集,且$\forall v(\cdot)\in\mathscr{U}_{ad}$, $\mathscr{F}_t^{Y^v}=\mathscr{F}_t^{Y^0}$。

引入记号$\hat{\psi}(t)=\mathbb{E}[\psi(t)\mid\mathscr{F}_t^{Y}]$和 Riccati 方程

$$\begin{cases}\dot{\pi}(t)+2a(t)\pi(t)-\dfrac{1}{R(t)}b(t)^2\pi(t)^2+L(t)=0,\\ \pi(T)=M,\end{cases} \tag{24}$$

$$\begin{cases}\dot{\Sigma}(t)+\left[a(t)+B(t)-\dfrac{1}{R(t)}b(t)^2\pi(t)\right]\Sigma(t)\\ \qquad+\dfrac{1}{R(t)}b(t)D(t)\pi(t)-A(t)=0,\\ \Sigma(T)=-F,\end{cases} \tag{25}$$

$$\begin{cases}\dot{\lambda}(t)+\left[a(t)-\dfrac{1}{R(t)}b(t)^2\pi(t)\right]\lambda(t)-o(t)\Sigma(t)\\ \qquad-\dfrac{1}{R(t)}b(t)r(t)\pi(t)+\bar{b}(t)\pi(t)+l(t)=0,\\ \lambda(T)=m。\end{cases} \tag{26}$$

定理 3 在假设5、6下,若$u(\cdot)$是问题(LQC)的最优控制,则它可以表示为

$$u(t)=\frac{1}{R(t)}[D(t)\hat{p}(t)-b(t)\hat{q}(t)-r(t)],$$

其中$(\hat{p}(\cdot),\hat{q}(\cdot),\hat{k}(\cdot))$和$\hat{K}(\cdot)$分别满足

$$\begin{cases}\mathrm{d}\hat{p}(t)=[B(t)\hat{p}(t)-o(t)]\mathrm{d}t+\left[C(t)\hat{p}(t)+\dfrac{f(t)}{h(t)}[\widehat{x(t)p(t)}-\hat{x}(t)\hat{p}(t)]\right]\mathrm{d}\hat{W}_t,\\ -\mathrm{d}\hat{q}(t)=[a(t)\hat{q}(t)-A(t)\hat{p}(t)+L(t)\hat{x}(t)+l(t)]\mathrm{d}t-\hat{K}(t)\mathrm{d}\hat{W}_t,\\ \hat{p}(0)=-Ny(0)-n,\ \hat{q}(T)=M\hat{x}(T)-F\hat{p}(T)+m,\end{cases}$$

$\hat{K}(t)=\hat{k}(t)+\dfrac{f(t)}{h(t)}[\widehat{x(t)q(t)}-\hat{x}(t)\hat{q}(t)]$,状态的最优滤波$(\hat{x}(\cdot),\hat{y}(\cdot),\hat{z}(\cdot),\hat{\bar{z}}(\cdot))$满足

$$\begin{cases}\mathrm{d}\hat{x}(t)=[a(t)\hat{x}(t)+b(t)u(t)+\bar{b}(t)]\mathrm{d}t+\left[c(t)+\dfrac{P(t)f(t)}{h(t)}\right]\mathrm{d}\hat{W}_t,\\ -\mathrm{d}\hat{y}(t)=[A(t)\hat{x}(t)+B(t)\hat{y}(t)+C(t)\hat{z}(t)+\bar{C}(t)\hat{\bar{z}}(t)\\ \qquad+D(t)u(t)+\bar{D}(t)]\mathrm{d}t-\hat{Z}(t)\mathrm{d}\hat{W}_t,\\ \hat{x}(0)=e_0,\ \hat{y}(T)=F\hat{x}(T)+\hat{G},\end{cases}$$

$\hat{x}(\cdot)$的均方误差估计$P(\cdot)$满足

$$\begin{cases}\dot{P}(t)-2a(t)P(t)+\left[c(t)+\dfrac{P(t)f(t)}{h(t)}\right]^2-[c(t)+\bar{c}(t)]^2=0,\\ P(0)=0,\end{cases}$$

标准布朗运动

$$\hat{W}_t=\int_0^t\frac{1}{h(s)}\{dY(s)-[f(s)]\hat{x}(s)+g(s)\}\mathrm{d}s=\int_0^t\frac{f(s)}{h(s)}[x(s)-\hat{x}(s)]\mathrm{d}s+W_t,$$

$$\hat{Z}(t)=\hat{z}(t)+\frac{f(s)}{h(s)}[\widehat{x(t)y(t)}-\hat{y}(s)\hat{y}(t)]。$$

进一步,最优控制可表示为反馈形式

$$u(t)=\frac{1}{R(t)}\{[D(t)-b(t)\Sigma(t)]\hat{p}(t)-b(t)\pi(t)\hat{x}(t)-b(t)\lambda(t)-r(t)\},$$

其中 $\pi(\cdot)$,$\Sigma(\cdot)$和 $\lambda(\cdot)$分别是(24)—(26)的解。

特别地,若只考虑倒向随机控制系统,也即 $a=b=\bar{b}=c=\bar{c}=f=g=l=r=m=0$, $A=L=M=0$, $h=1$,则定理 3 即为 Huang,Wang 和 Xiong[38]的情形。

以上内容主要取自 Wang,Wu 和 Xiong[29]。关于部分可观测线性二次控制问题的其他进展可参见文献[39-41]。

5　结论

本文简单概述了最近几年正倒向随机系统最大值原理研究方面所取得的部分代表性进展。随着理论研究的逐步深入和金融应用的积极尝试,崭新的挑战性课题如雨后春笋般不断涌现,例如:基于强化学习的随机控制理论、基于高频交易的金融风险度量以及区块链上的最优交易策略设计等,都非常值得我们今后继续深入研究。

参考文献:

[1] BISMUT J M. Linear quadratic optimal stochastic control with random coefficients[J]. SIAM Journal on Control and Optimization, 1976, 14(3):419-444.
[2] PARDOUX E, PENG S G. Adapted solution of a backward stochastic differential equation[J]. Systems & Control Letters, 1990, 14(1):55-61.
[3] ANTONELLI F. Backward-forward stochastic differential equations[J]. The Annals of Applied Probability, 1993, 3(3):777-793.
[4] PARDOUX E, TANG S J. Forward-backward stochastic differential equations and quasilinear parabolic PDEs[J]. Probability Theory and Related Fields, 1999, 114(2):123-150.
[5] MA J, PROTTER P, YONG J M. Solving forward-backward stochastic differential equations explicitly: a four step scheme[J]. Probability Theory and Related Fields, 1994, 98(3):339-359.
[6] HU Y, PENG S G. Solution of forward-backward stochastic differential equations[J]. Probability Theory and Related Fields, 1995, 103(2):273-283.
[7] PENG S G, WU Z. Fully coupled forward-backward stochastic differential equations and applications to optimal control[J]. SIAM Journal on Control and Optimization, 1999, 37(3):825-843.
[8] MA J, YONG J M. Forward-backward stochastic differential equations and their applications[M]//Lecture Notes in Math 1702. New York: Springer-Verlag, 1999.
[9] YONG J M, ZHOU X Y. Stochastic controls: Hamiltonian systems and HJB euqations[M]. New York: Springer-Verlag, 1999.
[10] MA J, WU Z, ZHANG D T, et al. On well-posedness of forward-backward SDEs: a unified approach[J]. The Annals of Applied Probability, 2015, 25(4):2168-2214.
[11] WANG G C, WU Z. The maximum principles for stochastic recursive optimal control problems under partial information[J]. IEEE Transactions on Automatic Control, 2009, 54(6):1230-1242.
[12] PENG S G. A general stochastic maximum principle for optimal control problems[J]. SIAM Journal on Control and Optimization, 1990, 28(4):966-979.
[13] PENG S G. Backward stochastic differential equations and applications to optimal control[J]. Applied Mathematics and Optimization, 1993, 27(2):125-144.
[14] XU W S. Stochastic maximum principle for optimal control problem of forward and backward system[J]. The ANZIAM Journal, 1995, 37(2):172-185.
[15] WU Z. Maximum principle for optimal control problem of fully coupled forward-backward stochastic systems[J]. Journal of

Mathematics and System Science. 1998, 11(3):249-259.
[16] PENG S G. Open problems on backward stochastic differential equations[M]// Control of Distributed Parameter and Stochastic Systems. Boston: Springer, 1999: 265-273.
[17] WU Z. A general maximum principle for optimal control of forward-backward stochastic systems[J]. Automatica, 2013, 49(5):1473-1480.
[18] HU M S. Stochastic global maximum principle for optimization with recursive utilities[J]. Probability, Uncertainty and Quantitative Risk, 2017, 2(1):1-20.
[19] WU Z, YU Z Y. Dynamic programming principle for one kind of stochastic recursive optimal control problem and Hamilton-Jacobi-Bellman equation[J]. SIAM Journal on Control and Optimization, 2008, 47(5):2616-2641.
[20] NIE T Y, SHI J T, WU Z. Connection between MP and DPP for stochastic recursive optimal control problems: viscosity solution framework in the general case[J]. SIAM Journal on Control and Optimization, 2017, 55(5):3258-3294.
[21] WONHAM W M. On the separation theorem of stochastic control[J]. SIAM Journal on Control, 1968, 6(2):312-326.
[22] LI X J, TANG S J. General necessary conditions for partially observed optimal stochastic controls[J]. Journal of Applied Probability, 1995, 32(4):1118-1137.
[23] TANG S J. The maximum principle for partially observed optimal control of stochastic differential equations[J]. SIAM Journal on Control and Optimization, 1998, 36(5):1596-1617.
[24] WANG G C, WU Z. Kalman-Bucy filtering equations of forward and backward stochastic systems and applications to recursive optimal control problems[J]. Journal of Mathematical Analysis and Applications, 2008, 342(2):1280-1296.
[25] WANG G C, ZHANG C H, ZHANG W H. Stochastic maximum principle for mean-field type optimal control under partial information[J]. IEEE Transactions on Automatic Control, 2014, 59(2):522-528.
[26] WANG G C, WU Z, XIONG J. An introduction to optimal control of FBSDE with incomplete information[M]. Cham: Springer, 2018.
[27] WU Z. A maximum principle for partially observed optimal control of forward-backward stochastic control systems[J]. Science China Information Sciences, 2010, 53(11):2205-2214.
[28] WANG G C, WU Z, XIONG J. Maximum principles for forward-backward stochastic control systems with correlated state and observation noises[J]. SIAM Journal on Control and Optimization, 2013, 51(1):491-524.
[29] WANG G C, WU Z, XIONG J. A linear-quadratic optimal control problem of forward-backward stochastic differential equations with partial information[J]. IEEE Transactions on Automatic Control, 2015, 60(11):2904-2916.
[30] KOHLMANN M, ZHOU X Y. Relationship between backward stochastic differential equations and stochastic controls: a linear-quadratic approach[J]. SIAM Journal on Control and Optimization, 2000, 38(5):1392-1407.
[31] LIM A E B, ZHOU X Y. Linear-quadratic control of backward stochastic differential equations[J]. SIAM Journal on Control and Optimization, 2001, 40(2):450-474.
[32] HU M S, JI S L, XUE X L. A global stochastic maximum principle for fully coupled forward-backward stochastic systems [J]. SIAM Journal on Control and Optimization, 2018, 56(6):4309-4335.
[33] MOON J. The risk-sensitive maximum principle for controlled forward-backward stochastic differential equations[J]. Automatica, 2020, 120:109069.
[34] JI S L, LIU H. Maximum principle for stochastic optimal control problem of forward-backward stochastic difference systems [J/OL].(2021-12-29). International Journal of Control. https://arxiv.org/abs/1812.11283.
[35] LI R J, FU F Y. The maximum principle for partially observed optimal control problems of mean-field FBSDEs[J]. International Journal of Control, 2019, 92(10):2463-2472. doi:10.1080/00207179.2018.1441555.
[36] ZHANG S Q, LI X, XIONG J. A stochastic maximum principle for partially observed stochastic control systems with delay [J]. Systems & Control Letters, 2020, 146:1-7. doi:10.1016/j.sysconle.2020.104812.
[37] BENSOUSSAN A. Stochastic control of partially observable systems[M]. Cambridge: Cambridge University Press, 1992.
[38] HUANG J H, WANG G C, XIONG J. A maximum principle for partial information backward stochastic control problems with applications[J]. SIAM Journal on Control and Optimization, 2009, 48(4):2106-2117.
[39] LI N, WANG G C, WU Z. Linear-quadratic optimal control for time-delay stochastic system with recursive utility under full and partial information[J]. Automatica, 2020, 121:109169.
[40] BENSOUSSAN A, FENG X W, HUANG J H. Linear-quadratic-Gaussian mean-field-game with partial observation and common noise[J]. Mathematical Control & Related Fields, 2021, 11(1):23-46.
[41] WANG G C, WANG W C, YAN Z G. Linear quadratic control of backward stochastic differential equation with partial information[J]. Applied Mathematics and Computation, 2021, 403:126164.

(编辑:李晓红)