

文章编号: 1000-8152(2009)07-0709-06

## 含扩散项不可靠生产系统最优生产控制的数值求解

宋春跃, 李 平

(浙江大学 工业控制技术国家重点实验室 工业控制研究所, 浙江 杭州 310027)

**摘要:** 针对含扩散项不可靠随机生产系统最优生产控制的优化命题, 采用数值解方法来求解该优化命题最优控制所满足的模态耦合的非线性偏微分HJB方程. 首先构造Markov链来近似生产系统状态演化, 并基于局部一致性原理, 把求解连续时间随机控制问题转化为求解离散时间的Markov决策过程问题, 然后采用数值迭代和策略迭代算法来实现最优控制数值求解过程. 文末仿真结果验证了该方法的正确性和有效性.

**关键词:** 不可靠生产系统; 生产控制; 数值解; Markov决策过程

**中图分类号:** TP11      **文献标识码:** A

## Numerical solution for optimal production control of unreliable production systems with diffusion terms

SONG Chun-yue, LI Ping

(State Key Lab of Industrial Control Technology, Institute of Industrial Process Control, Zhejiang University, Hangzhou Zhejiang 310027, China)

**Abstract:** The optimal production control for unreliable stochastic production system always involves in solving a mode-coupled nonlinear partial differential equation, i.e., HJB(Hamilton-Jacobi-Bellman) equation, which is the necessary and sufficient condition of optimal control. Numerical method for stochastic control problems in continuous time is adopted to solve the optimal production control problem involving diffusion terms by constructing Markov chains to approximate the evolution of the system states, and then, the associated HJB equation is transformed into a discrete time Markov decision process(MDP) under local consistence. Based on the MDP, an algorithm including numerical iteration and policy iteration is then proposed. Finally, some numerical examples of production system are presented to illustrate the usefulness of the numerical method.

**Key words:** unreliable production systems; production control; numerical method; Markov decision process

### 1 引言(Introduction)

柔性制造系统(FMS)是由若干可进行多种加工工序的工作站(workstation)及传送装置组成. 由于FMS具有很好的柔性, 故FMS制造系统较传统的制造系统具有很大的优势, 以至越来越多的生产企业采用FMS工作方式. 然而, FMS高昂的成本要求对FMS的生产和维护进行有效的管理和调度成为必然, 以应付诸如设备故障、需求波动、销售退货及库存损耗等不确定事件的发生, 来满足市场的需求和降低生产成本<sup>[1]</sup>. 由于系统规模大及结构复杂, 使得对FMS的管理和控制具有递阶结构, 即上层为管理层, 下层为生产控制(production control)层. 管理层是离线实现的, 而生产控制层则是在线实现, 动态调度, 以优化产品的生产速度和切换次序.

对生产控制的研究由来已久, 现在世界上形成

了以MIT的“制造和生产力研究实验室”(laboratory for manufacturing and productivity, <http://web.mit.edu/lmp/>)和Boston大学的“制造系统生产控制实验室”(production control of manufacturing systems laboratory, <http://www.bu.edu/pcms/>)为主的生产控制研究中心, 在对生产控制策略性质及值函数性质的研究上, 得出很多有意义的结论: 例如生产控制的开关结构和值函数的凸性以及最优生产控制策略所要满足的条件等<sup>[2~5]</sup>, 为生产控制的工程应用奠定了理论基础. 对生产控制的研究可分为两种途径, 一种是把工作站作为服务中心来对待, 把待加工零件作为服务对象, 用排队论来研究此类系统; 另一类是基于对大批量产品生产的近似, 以流率化模型来研究生产控制<sup>[8]</sup>. 不论那种研究方法, 最优状态反馈控制求取都面临求解HJB方程的难题.

本文采用后一种生产控制建模方式,针对含扩散项的不可靠随机生产系统最优生产控制的优化命题,为求解含扩散项不可靠生产系统最优生产控制中模态耦合的非线性偏微分HJB方程,利用数值解方法,通过构造Markov链来近似生产系统状态的演化,在满足局部一致性意义下,把求解原始HJB方程转化为求解离散时间的Markov决策过程问题,并由此设计了数值迭代和策略迭代算法来实现数值求解过程.仿真结果分析表明了该方法的正确性和有效性.

## 2 问题描述(Description of the problem)

本文所研究的不可靠生产系统具有文献[1]中对象相同的结构,即由 $L$ 台工作站组成,每台工作站由 $L_j$ 个相同的设备组成, $j = 1, \dots, L$ .系统可生产 $N$ 类不同的产品类型 $P_i$ , $i = 1, \dots, N$ .令 $x(t) = [x_1(t), \dots, x_N(t)]^T \in X \subseteq \mathbb{R}^N$ 为产品盈余(surplus),即当 $x_i \geq 0$ 时表示产品 $P_i$ 的库存状态(inventory),反之为该产品的缺货状态(shortage). $x(t)$ 的动态演化可由如下随机微分方程描述:

$$dx(t) = (u(t) - z(t))dt + \sigma dw(t), x(0). \quad (1)$$

其中: $u(t) = [u_1(t), \dots, u_N(t)]^T$ 为系统对产品的生产率向量,是控制变量; $z(t) = [z_1(t), \dots, z_N(t)]^T$ 为市场或下游设备对产品的需求向量,有 $z_i(t) \geq 0$ ;  $\sigma$ 为 $N \times P$ 矩阵, $w(t) = [w_1(t), \dots, w_P(t)]^T$ 为标准Wiener过程; $\sigma dw(t)$ 为方程(1)的扩散项,其物理意义可解释为需求的波动、库存的损耗及销售的退货等.

由于设备的不可靠性,在生产过程中每台设备都存在随机故障的发生和修复.令 $s_j(t)$ 为 $t$ 时刻第 $j$ 台工作站的状态,则系统工作站状态为 $s(t) = [s_1(t), \dots, s_L(t)]^T$ .

$s(t)$ 的动态演化与设备老化程度有关,令 $a_j(t) \geq 0$ 为工作站 $j$ 在 $t$ 时刻的老化程度, $a_j(t)$ 可由生产率的累计量及维护频率来刻画.为简化问题说明,假设工作站老化程度仅是维护频率 $m_j$ 的函数, $a_j(t) = g_j(m_j(t))$ .维护频率 $m_j(t)$ 的倒数 $1/m_j(t)$ 表示对工作站 $j$ 的维护间隔时间.若 $m_j(t) = 0$ ,则表示不用对该工作站进行维护;由于企业人力物力的限制及每台工作站的重要性不同,有 $0 \leq m_j(t) \leq m_j^{\max}$ . $s_j(t)$ 的动态演化可由如下转移概率刻画:

$$P\{s_j(t+dt) = l+1 \mid s_j(t) = l, a_j(t)\} = \begin{cases} (L_j-l)p_{j_l j_{l+1}}^j(a_j(t))dt + o(dt), & 0 \leq l < L_j, \\ 0, & \text{其他情况.} \end{cases} \quad (2a)$$

$$P\{s_j(t+dt) = l-1 \mid s_j(t) = l, a_j(t)\} =$$

$$\begin{cases} (L_j-l)p_{j_l j_{l-1}}^j(a_j(t))dt + o(dt), & 0 < l \leq L_j, \\ 0, & \text{其他情况.} \end{cases} \quad (2b)$$

令 $m(t)$ 为系统的维护频率,即 $m(t) = [m_1(t), \dots, m_L(t)]^T$ ;  $a(t) = [a_1(t), \dots, a_L(t)]^T$ ,  $g(\cdot) = [g_1(\cdot), \dots, g_L(\cdot)]^T$ ,则有

$$\dot{a}(t) = g(m(t))dt, a(0). \quad (3)$$

假设备老化程度 $a$ 为非减的,故有 $g \geq 0$ .若 $\alpha, \beta$ 为 $s(t)$ 的两个取值,定义 $P\{s(t+dt) = \alpha \mid s(t) = \beta, a(t)\} = \lambda_{\alpha\beta}(a(t))dt + o(dt)$ , $\alpha \neq \beta$ .

以上各式中 $dt > 0$ 为一个非常小数, $\lambda_{\alpha\beta}$ 为系统状态 $\alpha$ 跳跃到状态 $\beta$ 的跳跃率,有 $\sum_{\beta} \lambda_{\alpha\beta}(\cdot) = 0$ .当 $a$ 为恒值时,即 $\dot{a} = 0$ 时,表示该工作站故障的发生和修复与设备的维护无关, $s(t)$ 就退化为平稳过程.经过以上定义后, $s(t)$ 可由其生成矩阵 $Q = (\lambda_{\alpha\beta}(a))$ 完全描述.由于设备不可靠性的存在及设备自身生产能力的约束,使得生产率控制向量 $u(t)$ 存在以下约束:

$$0 \leq u_i(t) \leq s(t)^T r_i. \quad (4)$$

其中: $r_i = [r_{i1}, \dots, r_{iL}]^T$ , $r_{ij}$ 为工作站 $j$ 对产品 $P_i$ 的最大生产率.对所描述系统,最优生产控制问题可表述为:已知初始盈余 $x(0)$ 、工作站的初始状态 $s(0)$ 及初始老化状态 $a(0)$ ,寻求最优的容许生产控制策略 $u(t)$ 和设备维护频率 $m(t)$ ,在满足方程(1)(3)和相应控制变量约束的情况下,使如下长期折扣费用目标泛函最小:

$$J_{(x(0), s(0), a(0))}(u(\cdot), m(\cdot)) = \min_{u(\cdot), m(\cdot)} E\left(\int_0^{\infty} e^{-\rho v} l(x(v))dv \mid x(0), s(0), a(0)\right). \quad (5)$$

其中: $0 < \rho < 1$ 为折扣因子, $l(x)$ 为产品成本运行函数.定义目标函数(5)的值函数为 $V(x(t), s(t), m(t))$ ,其表达式如下:

$$V(x(t), s(t), a(t)) = \min_{u(\cdot), m(\cdot)} E\left(\int_t^{\infty} e^{-\rho(v-t)} l(x(v))dv \mid x(t), s(t), a(t)\right). \quad (6)$$

生产率 $u(t)$ 和维护频率 $m(t)$ 为最优的充要条件由以下引理给出:

**引理 1** 对于由式(1)(3)描述的不可靠生产系统,若使目标泛函(5)最小,则对于每个状态 $x(t) = x$ , $s(t) = \alpha$ 及 $a(t) = a$ ,容许生产控制策略 $u(t)$ 和设备维护频率 $m(t)$ 为最优的充分必要条件是其所由

式(6)所定义的值函数满足如下HJB方程:

$$\begin{aligned} & \min_{u(t), m(t)} \{V_x(x, \alpha, a)(u(t) - z(t)) + \\ & V_a(x, \alpha, a)g(m(t)) + \frac{1}{2}\text{tr}[V_{xx}(x, \alpha, a)\sigma\sigma^T] + \\ & \sum_{\beta} \lambda_{\alpha\beta}(a)V(x, \beta, a) - \rho V(x, \alpha, a)\} + l(x) = 0. \end{aligned} \quad (7)$$

证 把 $a(t)$ 及设备维护频率 $m(t)$ 分别作为辅助状态变量和辅助控制率引入文献[1]的Theorem 6.2中, 该引理可直接得到. 证毕.

显然有 $J_{(x(0), S(0), a(0))}(u(\cdot), m(\cdot)) = V(x(0), s(0), a(0))$ . 若 $\sigma$ 有如下形式,  $\sigma = \text{diag}\{\sigma_1, \dots, \sigma_N\}$ , 其中 $\sigma_i > 0, i = 1, \dots, N$ . 则式(7)可改写为

$$\begin{aligned} & \min_{u(t), m(t)} \left\{ \sum_{i=1}^N \frac{\partial}{\partial x_i} V(x, \alpha, a)(u_i(t) - z_i(t)) + \right. \\ & \sum_{i=1}^L \frac{\partial}{\partial a_i} V(x, \alpha, a)g_i(m_i(t)) + \\ & \left. \frac{1}{2} \sum_{i=1}^N \frac{\partial^2 V(x, \alpha, a)}{\partial x_i^2} \sigma_i^2 + \sum_{\beta} \lambda_{\alpha\beta}(a)V(x, \beta, a) - \right. \\ & \left. \rho V(x, \alpha, a) \right\} + l(x) = 0. \end{aligned} \quad (8)$$

### 3 数值求解(Numerical method)

式(8)为一个模态耦合的非线性偏微分方程, 求解这样的方程实非易事, 故往往求助于数值求解. Kushner 在弱存在性(weak existence)和弱唯一性(weak uniqueness)概念下创立了求解连续时间随机控制问题的数值解方法<sup>[6]</sup>. 据作者所知, 把该方法首先应用于求解不可靠生产系统生产控制问题的是文献[4], 用于处理生产率不耦合的分段确定性模型; 文献[3]则用该数值解方法来处理生产率耦合的分段确定性模型. 当系统状态演化方程中包含扩散项时, 还没有看到有相关文献进行数值求解. 本文仍采用Kushner数值解方法, 通过构造Markov链来近似生产系统状态的演化. 在分析了该方法用于求解生产控制优化命题的局部特性后, 基于满足局部一致性(local consistent)的前提下给出求解涉及扩散项的生产优化控制问题的数值解方法及算法. 不失一般性, 这里取 $\sigma = \text{diag}\{\sigma_1, \dots, \sigma_N\}$ .

#### 3.1 值函数近似(Approximation of value function)

Kushner数值解方法是构造Markov链, 在满足局部一致性条件下来近似连续随机过程, 然后利用该Markov链的值函数来逼近原始问题的值函数. 令

$$\begin{aligned} R(h) &= \sum_{i=1}^N \sigma_i^2 + h \left( \sum_{i=1}^N (u_i + z_i) + \sum_{i=1}^L g_i(m_i) \right) + \\ & (\rho - \lambda_{\alpha\alpha}(a))h^2, \quad \Delta t(h) = \frac{h^2}{R(h)}. \end{aligned}$$

当 $h \rightarrow 0$ 时,  $\Delta t(h) \rightarrow 0$ . 构造Markov链向量

$$\begin{aligned} \chi^k &= [\chi_1^k, \dots, \chi_N^k]^T, \quad \xi^k = [\xi_1^k, \dots, \xi_L^k]^T, \\ k &= 1, 2, \dots \end{aligned}$$

使其转移概率有如下形式:

$$P_{\chi_i^k}^+ = P(\chi_i^{k+1} | \chi_i^k, u_i^k) = \frac{\sigma_i^2 + 2hu_i^k}{2R(h)}, \quad \Delta\chi_i^k > 0, \quad (9a)$$

$$P_{\chi_i^k}^- = P(\chi_i^{k+1} | \chi_i^k, u_i^k) = \frac{\sigma_i^2 + 2hz_i^k}{2R(h)}, \quad \Delta\chi_i^k < 0, \quad (9b)$$

$$P_{\chi_i^k} = P(\chi_i^{k+1} | \chi_i^k, u_i^k) = \frac{h^2\rho}{R(h)}, \quad \Delta\chi_i^k = 0, \quad (9c)$$

$$P_{\xi_j^k} = P(\xi_j^{k+1} | \xi_j^k, m_j^k) = \frac{hg_j(m_j^k)}{R(h)}, \quad (9d)$$

$$P^\beta = \frac{h^2\lambda_{\alpha\beta}}{R(h)}. \quad (9e)$$

式中:  $\Delta\chi_i^k = \chi_i^{k+1} - \chi_i^k$ ,  $h$ 可解释为变量的有限差分.  $u^k$ 及 $m^k$ 分别为 $k$ 时刻所施加的控制. 令 $V^h(x, \alpha, a)$ 为Markov链 $\chi, \xi$ 的值函数, 其表达式如下:

$$\begin{aligned} V^h(x, \alpha, a) &= \\ & \min_{u(\cdot), m(\cdot)} E \left( \sum_{k=0}^{\infty} e^{-\rho t_k} l(\chi_k) \Delta t(h) \mid x, \alpha, a \right). \end{aligned} \quad (10)$$

其中 $t_k = \sum_{i=0}^{k-1} \Delta t(h)$ , 则有

$$\begin{aligned} V^h(x, \alpha, a) &= \\ & \min_{u_k, m_k} \left\{ \sum_{i=1}^N (P_{\chi_i^k}^+ V^h(x + e_i h, \alpha, a) + \right. \\ & P_{\chi_i^k}^- V^h(x - e_i h, \alpha, a)) + \\ & \sum_{j=1}^L P_{\xi_j^k}^+ V^h(x, \alpha, a + e_j h) + \\ & \left. \sum_{\beta \neq \alpha} P^\beta V(x, \beta, a) \right\} + \Delta t(h)l(x). \end{aligned} \quad (11)$$

#### 3.2 局部一致性(Local consistent)

对于系统(1)而言, 在非常小数 $dt > 0$ 时, 系统状态 $x(t)$ 有如下局部特性:

$$E[x(t+dt) - x(t) \mid x(t)] = (u(t) - z(t))dt, \quad (12a)$$

$$\text{Cov}[x(t+dt) - x(t) \mid x(t)] = \sigma\sigma^T dt. \quad (12b)$$

$E(\cdot)$ 和 $\text{Cov}(\cdot)$ 分别表示期望和方差. 所构造的Markov链 $\chi$ 向量在转移概率(9)情况下, 其局部特性如下:

$$E[\chi^{k+1} - \chi^k \mid \chi^k] =$$

$$\sum_{i=1}^N (P_{\chi_i^k}^+ - P_{\chi_i^k}^-)h = (u^k - z^k)\Delta t(h), \quad (13a)$$

$$\text{Cov}[\chi^{k+1} - \chi^k | \chi^k] = \sigma\sigma^T \Delta t(h) + o(\Delta t(h)). \quad (13b)$$

其中 $z^k$ 为需求 $z(t)$ 在 $k$ 时刻的值. 对比式(12)与式(13), 不难看出, 当 $\Delta t(\cdot) \rightarrow 0$ 时所构造的Markov链 $\chi$ 在局部特性上与原系统 $x(t)$ 是一致的, 即符合Kushner方法所要求的局部一致性条件. 同样可以验证 $a(t)$ 与所构造Markov链 $\xi$ 的局部一致性也是符合的. 式(11)为离散时间的Markov决策过程的一般形式, 当 $h \rightarrow 0$ 时, 由方程(11)所得到的 $V^h(x, \alpha, a)$ 趋向于由HJB方程(7)所决定的 $V(x, \alpha, a)$ , 相应的最优控制策略 $u^k$ ,  $m^k$ 序列可看作是原始优化问题最优控制策略解的近似,  $k = 1, 2, \dots$ , 这种近似有如下引理确保:

**引理 2** 当 $h \rightarrow 0$ 时, 由Markov决策过程式(11)所求得的价值函数 $V^h(\cdot)$ 为所研究不可靠生产系统价值函数 $V(\cdot)$ 的极限, 即 $\lim_{h \rightarrow 0} V^h(\cdot) = V(\cdot)$ .

**证** 在局部一致性条件满足下, 引理2可由文献[6]中Chap. 10中Theorem 5.2得到.

证毕.

引理2表明, 具有Markov特性的连续时间优化控制命题可转化为相同状态空间下的离散时间Markov决策过程问题. 由式(11)求得的最优控制 $u^k$ ,  $m^k$ 即为原始问题最优解 $u, m$ 的数值解.

### 3.3 算法设计(Algorithm)

根据前文所述, 可设计如下包含值迭代和策略迭代的算法来实现生产优化控制策略的数值求解过程. 需要说明的是, 用离散时间Markov决策过程(11)进行数值求解时, 应先网格化所考虑状态空间 $X$ , 定义网格化后的状态空间为 $G$ , 则算法描述如下:

**数值求解算法框架:**

**Input:** 网格化状态空间 $G$ ; 对系统参数进行赋值;

**Initialization:**  $\forall x \in G, \alpha, a$ , 令 $V^h(\cdot) = 0$ ;

**Loop 1:** 值迭代. 直到设定标准满足;

**Loop 2:** 策略迭代. 给定策略, 由式(11)求解相应的价值函数 $V^h(\cdot)$ 及最优控制策略 $u^k(\cdot), m^k(\cdot)$ ;

**End**

**End**

**Output:** 最优控制策略  $u^k(\cdot), m^k(\cdot)$  及价值函数  $V^h(\cdot)$ .

其他适合离散时间MDP问题的优化算法都可以用来具体化本算法, 这里不再详细讨论.

## 4 仿真分析(Simulation and analysis)

为验证方法的正确性和有效性, 如下是两个仿真实例. 为简化具体数值求解过程, 不妨令 $\dot{a}(t) = 0$ .

**例 1** 不可靠生产系统是只有由一台设备组成的工作站, 只生产一种产品. 设备有故障和正常工作两种状态, 以0表示故障状态, 1表示正常状态. 故障状态下, 设备不能进行任何生产. 并假设对产品的需求 $z$ 为恒值. 采用如下形式的成本运行函数表达式:

$$l(x(t)) = c^+ x^+(t) + c^- x^-(t).$$

其中:  $x^+(t) = \max\{0, x\}$ ,  $x^-(t) = \max\{0, -x\}$ . 系统其他参数如下:  $z = 0.45$ ,  $\rho = 0.9$ ,  $\lambda_{01} = 0.9$ ,  $\lambda_{10} = 0.14$ ,  $c^+ = 1.0$ ,  $c^- = 6.0$ ,  $r = 1.0$ . 若不考虑扩散项, 即令 $\sigma = 0$ . 这是最简单情况下的不可靠设备生产控制的优化命题, 文献[5]给出了该情况下的解析解, 即最优反馈控制策略为 $u(t) = r$ 当 $x(t) < x^*$ 时; 当 $x(t) > x^*$ 时 $u(t) = 0$ ; 当 $x(t) = x^*$ 时 $u(t) = z(t)$ ; 其中

$$x^* = \max\left\{0, \frac{1}{\lambda_-} \ln\left[\frac{c^+}{c^+ + c^-} (1 + \frac{\rho z}{\lambda_{10} z - (\rho + \lambda_{01} + z \lambda_-)(r - z)})\right]\right\},$$

$\lambda_-$ 为矩阵  $\begin{bmatrix} \frac{\rho + \lambda_{10}}{r - z} & \frac{\lambda_{10}}{z - r} \\ \frac{\lambda_{01}}{z} & -\frac{\lambda_{01} + \rho}{z} \end{bmatrix}$  的唯一负特征根.

采用本文数值求解仿真结果如图1(a)所示. 图1(a)中实线表示设备状态为0时的值函数曲线; 点划线为设备状态为1时的值函数曲线, 虚线为最优控制策略 $u$ . 采用文献[5]公式计算得 $x^* = 0$ , 数值求解结果与理论相符. 当系统参数 $\lambda_{01} = 0.6$ ,  $\lambda_{10} = 0.3$ 时, 意味着设备故障时间增加而正常生产时间减少, 直觉告诉笔者此时的最优库存也应增大, 利用公式计算 $x^* = 0.106$ . 这种情况下数值求解的仿真如图1(b)所示(图1(c)是图1(b)的局部放大图). 图中点划线所对应的是当仍取 $x^* = 0$ 时的值函数 $V(\cdot)$ 曲线及控制策略; 虚线是取 $x^* = 0.5$ 时的值函数 $V(\cdot)$ 曲线及控制策略; 而实线是取 $x^* = 0.1$ 时的值函数 $V(\cdot)$ 曲线及控制策略. 这两个仿真结果验证了本文所述方法结果的正确性.

以上数值求解都是系统不存在扩散项时的结果. 当扩散项存在时, 即 $\sigma \neq 0$ 时, 目前仍没有解析解存在. 图1(d)是系统存在扩散项而其他参数不变时的数值解结果, 仍取 $\lambda_{01} = 0.6$ ,  $\lambda_{10} = 0.3$ . 直觉上此时系统存在除设备随机故障外的随机扰动, 最优库存应该比没有扩散项时大. 从图1可以看出, 当 $\sigma = 0.05$ 时,  $x^* = 0.13$ , 的确大于 $\sigma = 0$ 时的最优库存; 然而, 当 $\sigma$ 继续增大时, 这种直觉就被打破了. 当 $\sigma$ 取0.1及

以上值时, 仿真结果显示此时的最优库存小于不存在扩散项时的最优库存. 且随着扩散项参数的增大, 不同设备状态下的值函数曲线也在彼此靠拢, 向  $V(\cdot) = c^+x, V(\cdot) = c^-x$  靠近, 如图 1(d) 所示. 经过对对象的分析, 由于系统中含有标准 Wiener 过程扩散项, 这种现象也是合理的和必然的, 是系统本身所决定的. 即当存在扩散项时, 系统的设备故障因素被“噪声”给弱化了, 使得故障因素的影响比重变小, 从而值函数曲线彼此靠近; 当扩散项不断增大, 该项在系统的演化进程中所扮演的角色越来越重, 以至淹没了系统状态  $x(t)$  受  $s(t)$  所驱动演化的部分, 系统逐渐退化为不涉及随机事件的随机控制系统; 而 Wiener 过程零均值的特性及目标泛函为期望的表达式, 使得值函数在每个状态都趋向于  $c^+x$  或  $c^-x$ .

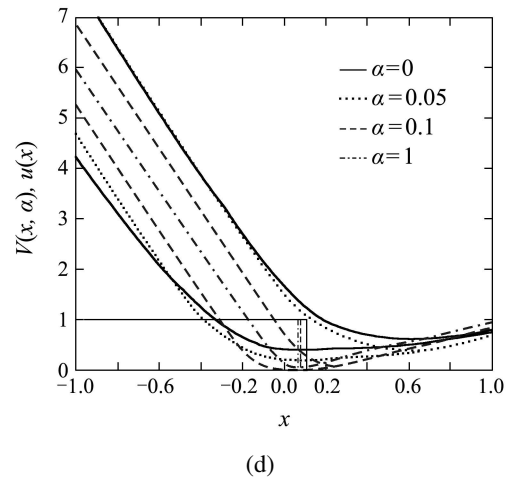
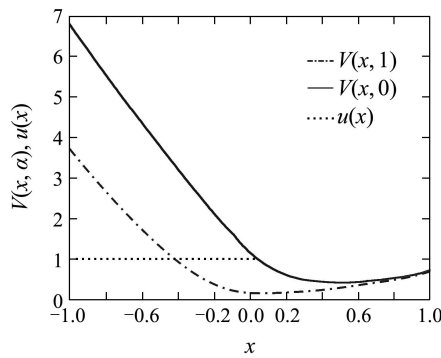
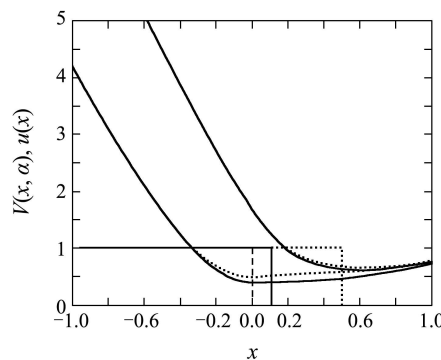


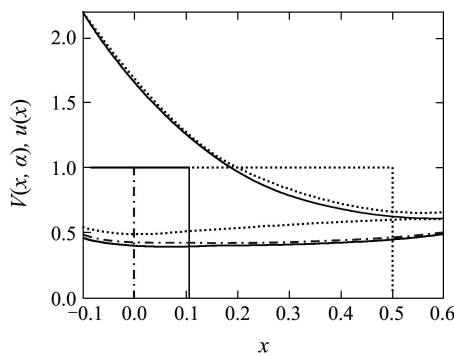
图 1 例 1 值函数  $V$  及控制策略  $u$  的数值解  
Fig. 1 Numerical solution to  $V, u$  with respect to  $x$  for example 1, respectively



(a)



(b)



(c)

**例 2** 不可靠生产系统生产 2 种产品情况, 设备状态同例 1 相同. 并假设对每种产品的需求  $z_i$  为恒值,  $i = 1, 2$ . 产品盈余运行函数表达式为  $l(x(t)) = \sum_{i=1}^2 (c_i^+ x_i^+(t) + c_i^- x_i^-(t))$ . 系统其他参数如下:  $z_1 = 0.4, z_2 = 0.6, \rho = 0.9, \lambda_{01} = 0.6, \lambda_{10} = 0.3, c_1^+ = c_2^+ = 1.0, c_1^- = 3.0, c_2^- = 6.0, r_1 = r_2 = 1.0, \sigma_1 = \sigma_2 = 0.02$ .

若系统由两台或多台设备组成, 产品生产率之间不耦合时, 不涉及产品的切换序列, 可直接由第 3 节所述算法进行仿真, 是例 1 在高维空间的直接推广. 若系统仍只有由一台设备组成的工作站, 设备在已知时刻只能生产一种产品, 在一台设备上进行不同产品的加工, 生产率之间存在耦合, 涉及产品的切换次序, 则需要对生产率增加一个约束如下:

$$\begin{cases} 0 \leq u_i(t) \leq s(t)r_i, & i = 1, 2, \\ u_j(t) = 0, & j \neq i. \end{cases}$$

按照上式约束, 在进行算法仿真时需对  $R(\cdot)$  进行相应的调整, 即

$$R(h) = \sum_{i=1}^N \sigma_i^2 + h \left( \sum_{i=1}^2 z_i(t) + u_i(t) I_{\{P_i\}} \right) + (\rho - \lambda_{\alpha\alpha}) h^2,$$

其中  $I_{\{\cdot\}}$  为指示函数, 表示当系统生产  $P_i$  时  $I_{\{\cdot\}} = 1$ , 否则为 0. 图 2 是生产率耦合情况下的仿真.  $V(x, 1)$  和  $V(x, 0)$  图分别是值函数在设备状态为正常和故障时产品盈余所对应的值函数曲面.  $u_1(x)$  和  $u_2(x)$  图分别代表了当设备状态为正常时最优生产控制策略与产品盈余的关系.

仿真图表明, 在生产能力耦合时, 系统的状态空间被分割为几个子空间, 分别对应相应的最优

控制策略,也就是说这类不可靠生产系统,在最优目标泛函为无限折扣费用情况下,最优控制策略具有开关结构(例1也有相同结论),与文献理论结果相一致<sup>[1]</sup>.图中还可以看出,最优控制策略的开关流形曲线并不是一个常值.值函数仿真曲线(面)显示出 $V(\cdot)$ 为凸函数,这也是对文献研究值函数凸特性的一个证实.数值仿真结果表明该数值求解方法正确有效,并可以扩展到更复杂情况.但从第3节的算法描述和Markov链的构造中也可以看出,当系统的状态维数增加时,计算所消耗的时间也急剧增长,算法的时间复杂性为 $\|\frac{X}{h_X}\|^N \|\frac{U}{h_U}\|^N$ ,其中 $|X|^N$ 为算法网格化时所考虑的状态空间大小, $|U|^N$ 为容许控制策略的空间大小,而 $h_X$ 和 $h_U$ 分别表示状态和策略空间上的迭代步长.为减少计算复杂性,一方面可以改进算法,采用效率更高的求解MDP问题的相关算法;另一方面可根据已有的生产控制文献对生产控制策略属性和结构的研究结果,用启发式方法来减少策略的搜索空间.基于该数值求解方法还可仿真分析系统参数对最优控制策略的影响,本文不再讨论.

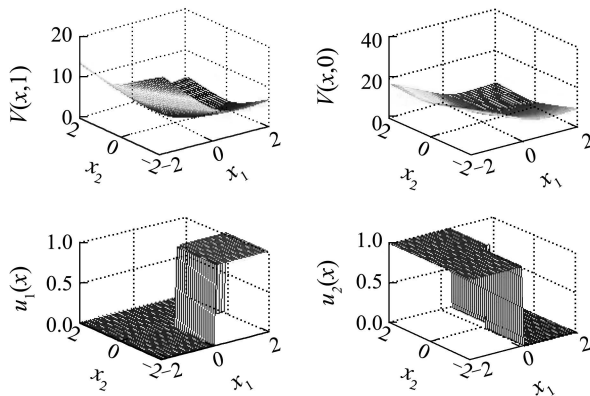


图2 例2值函数 $V$ 及控制策略 $u$ 的数值解

Fig. 2 Numerical solution to  $V$ ,  $u$  with respect to  $x$  for example 2, respectively

## 5 结论(Conclusion)

针对具有随机故障发生并涉及扩散项的不可靠生产系统,通过构造Markov链来近似系统的状态演化,并基于求解原始命题转化后的MDP问题,给出了求解模态耦合的非线性偏微分方程的数值解方法.对生产系统的数值仿真分析,验证了该方法的正确性和有效性.需要说明的是,不可靠随机生产系统可看作随机混杂系统,而随机混杂系统的研究大门才刚刚开启<sup>[7~10]</sup>.本文这些结论和方法为求解连续时

间的具有和不可靠生产系统相似特性的随机混杂系统优化控制命题提供了一种解决方案.遗憾的是,正如第4节中所说,由于该数值解方法计算复杂性是与状态空间和控制策略空间有关,当系统的状态空间为高维和控制策略空间也为高维时,所需计算时间急剧增长.开发高效率的算法是一个很有实际意义的工作,这也将是研究组下一步的工作重点.

## 参考文献(References):

- [1] MARINAL K G, ARISTOTLE A, STEVEN I M. Optimal control of switching diffusions with application to flexible manufacturing systems[J]. *SIAM Journal on Control and Optimization*, 1993, 31(5): 1183 – 1204.
- [2] YAN H, YIN G, ZHANG Q. *Stochastic Processes, Optimization, and Control Theory: Applications in Financial Engineering, Queueing Networks, and Manufacturing Systems*[M]. Chap. 4, 8. Berlin; NY: Springer, 2006.
- [3] YAN H, ZHANG Q. A numerical method in optimal production and setup scheduling of stochastic manufacturing systems[J]. *IEEE Transactions on Automatic Control*, 1997, 42(10): 1452 – 1455.
- [4] BOUKAS E K, HAURIE A. Manufacturing flow control and preventive maintenance: A stochastic control approach[J]. *IEEE Transactions on Automatic Control*, 1990, 35(9): 1024 – 1031.
- [5] AKELLA R, KUMAR P R. Optimal control of production rate in a failure-prone manufacturing systems[J]. *IEEE Transactions on Automatic Control*, 1986, 31(2): 116 – 126.
- [6] KUSHNER H J, DUPUIS P G. *Numerical Methods for Stochastic Control Problems in Continuous Time*[M]. 2nd ed. Berlin; NY: Springer, 2002.
- [7] BLOM H A P, LYGEROS J. *Stochastic Hybrid Systems: Theory and Safety Critical Applications*[M] // *Lecture Notes in Computer Science(LNCS)*, Chap. 1. Berlin: Springer-Verlag, 2006.
- [8] CASSANDRAS C G, LYGEROS J. *Stochastic Hybrid Systems*[M]. Chap 1,2,6,7. Boca Raton: Taylor & Francis, 2007.
- [9] 李俊领,解学军.间接型的混杂模型参考自适应控制[J].控制理论与应用,2007,24(6):1029–1033.  
(LI Junling, XIE Xuejun. Indirect hybrid model reference adaptive control[J]. *Control Theory & Applications*, 2007, 24(6): 1029 – 1033.)
- [10] HE P, JAGANNATHAN S. Reinforcement-based neural network controller for nonlinear discrete-time systems with input constraints[J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 2007, 37(2): 425 – 437.

## 作者简介:

宋春跃 (1972—),男,博士,副教授,目前研究方向为混杂系统优化控制、生产调度及控制、信号处理等, E-mail: cysong@iipc.zju.edu.cn;

李平 (1954—),男,教授,博士生导师,目前研究方向为企业综合自动化、流程工业CIPS、工业过程中的先进控制理论及应用等, E-mail: pli@iipc.zju.edu.cn.