

文章编号: 1000-8152(2005)06-0855-06

基于滚动优化的对偶控制策略

钱富才¹, 宋 俐², 陈小可¹

(1. 西安理工大学 自动化与信息工程学院, 陕西 西安 710048; 2. 西安理工大学 机械与精密仪器工程学院, 陕西 西安 710048)

摘要: 考虑具有未知参数的随机系统的最优控制问题. 提出了一种新的基于滚动优化的对偶控制算法. 在动态规划泛函方程中, 用 Kalman 滤波对系统的状态进行估计; 用线性化方法对阶段损失函数中的后验概率进行近似, 然后, 用滚动优化策略对控制与学习之间的耦合关系进行解耦. 从而获得了原不可解泛函方程的解析递推表达式和一个易于实施的控制律的解析解. 用一个例子说明了控制律的性能, 仿真结果表明: 该控制律具有良好的对偶性质, 在学习和控制之间实现了较好的平衡.

关键词: 随机系统; 动态规划; 滚动优化; 对偶控制

中图分类号: TP275 **文献标识码:** A

Dual control strategy based on receding horizon

QIAN Fu-cai¹, SONG Li², CHEN Xiao-ke¹

(1. School of Automation and Information Engineering, Xi'an University of Technology, Xi'an Shaanxi 710048, China;

2. School of Mechanical and Instrumental Engineering, Xi'an University of Technology, Xi'an Shaanxi 710048, China)

Abstract: The control of a stochastic system with unknown parameter was considered. A novel dual control algorithm based on receding horizon was proposed. In dynamic programming functional equation, the Kalman filter is used to estimate the state of such system, and the linearization method is applied to approximate the nonlinear cost-to-go function with respect to posterior probabilities, the receding horizon strategy is then proposed to decouple the coupling relation between control and learning. Thus, the explicit recursive expression of the original unsolvable functional equation and a control law with easy implementation are obtained. The performance of the control law derived in this paper is also illustrated with an example. The simulation results show that the control law with dual property can achieve preferable compromise between learning and control.

Key words: stochastic systems; dynamic programming; receding horizon; dual control

1 引言 (Introduction)

1960 年, 前苏联学者 Feldbaum 在他的系列论文中阐述了对偶控制 (dual control) 的思想^[1], 2000 年被 IEEE Control Systems Society 列为 20 世纪对控制领域最具影响力的 25 篇论文之一, 而这一难题至今尚未得到解决.

对随机系统施行最优控制, 其对偶性质体现在: 一方面, 控制信号迫使系统的输出趋向期望的目标 (称为对系统的调节作用或控制作用); 另一方面, 控制信号的作用还要有助于减小系统参数的不确定性 (称为对参数不确定性的学习作用或估计作用). 这两种作用在控制律的实现中是矛盾的, 前者往往要求控制信号的变化趋向平缓, 而后者要求维持一定幅度的激励, 因此对偶控制策略就是在控制和学习之间进行最佳折衷.

对偶控制的基本性质与特点已经得到了广泛的研究, 文献[2~5]给出了较为全面的综述. 动态规划是求解随机线性离散系统最优控制问题的有力工具. 然而, 对于具有未知参数的 LQG 问题, 除了少数几个极其简单的特例外, 在绝大多数情况下, 最优控制的解析解无法得到. 因此, 寻求具有对偶特点的次优控制策略成为一种必然途径^[2,3].

文献[6]研究了仅在输出方程中含有参数不确定性的对偶控制问题, 采用最小方差方法得到了在统计意义下的最优对偶控制. 对于更一般的情况, 即在状态方程和输出方程中均包含参数不确定性, 文献[7]运用嵌入原理和方差最小化方法获得了具有主动学习特点的次优对偶控制, 然而, 该控制律是开环控制, 这是工程应用中所不期望的. 对于具有参数不确定性的最小方差控制模型, 文献[8]通过把一个

N 级控制问题用 N 个单步控制问题来近似, 每一步的性能指标中包含控制效果度量与未知参数辨识度量, 尽管这种策略所获得的控制有对偶特点, 但是明显地控制具有近似行为, 因为单步控制只能保证下一步的性能指标, 不可能兼顾整个控制过程.

事实上, 求解对偶控制问题有两个难点: 一个是未来时刻的后验概率与当前及以前时刻施加于系统的控制有关, 同时确定当前时刻的控制又需要知道未来的后验概率, 这样当前控制与未来后验概率之间形成一个环, 如何打破这个环是必须解决的难点之一; 另一个难点是, Bellman 动态规划方程中的最优阶段损失函数不像 LQG 那样, 在每一阶段没有一个闭合形式, 因此, 该方程根本无法递推求解, 即使放弃解析解, 转而考虑数值解, 维数灾难不可避免. 本文受文献[9]的启发, 用滚动优化的思想克服了求解对偶控制问题中的两个难点, 充分利用了控制过程中所获得的信息, 在一定程度上既保留了优化的多步性, 又得到了形式较为简单的解析解, 同时保持了控制律具有对偶特性, 仿真结果表明了算法的有效性.

2 问题的描述(Problem description)

考虑具有参数不确定性的离散线性随机系统的最优控制问题(P):

$$\begin{aligned} \min E \{ & x(N)^T Q x(N) + \sum_{k=0}^{N-1} [x(k)^T Q x(k) + \\ & u(k)^T R u(k)] \} \\ \text{s.t. } & x(k+1) = A(\theta)x(k) + B(\theta)u(k) + v(k), \\ & k = 0, 1, \dots, N-1, \quad (1) \end{aligned}$$

$$y(k) = C(\theta)x(k) + w(k), \quad k = 1, 2, \dots, N. \quad (2)$$

其中: $x(k)$ 是 n 维状态向量, $u(k)$ 是 r 维控制向量, $y(k)$ 是 m 维输出向量, $v(k)$, $w(k)$ 和 $x(0)$ 为相互独立的高斯白噪声, 其分布分别为

$$v(k) \sim N(0, \sigma_v),$$

$$w(k) \sim N(0, \sigma_w),$$

$$x(0) \sim N(0, P(0)).$$

$A(\theta)$, $B(\theta)$, $C(\theta)$ 为适当维数的矩阵, 它们与参数 θ 有关, 并且为 θ 的线性函数. 假定 θ 在整个控制过程中是未知常量, 其取值于有限集合 $\Omega_\theta = \{\theta_1, \theta_2, \dots, \theta_s\}$, θ 的验前概率假定为

$$q_i(0) = P(\theta = \theta_i), \quad i = 1, 2, \dots, s.$$

性能指标中的 Q , R 为适当维数的半正定和正定对称阵.

假定当前时刻为 k , k 时刻控制器能够利用的实时信息为 $I^k = \{y(1), y(2), \dots, y(k), u(0),$

$u(1), \dots, u(k-1)\}$.

对偶控制问题就是求具有如下形式的反馈控制律

$$u(k) = f_k(I^k),$$

使问题(P)的性能指标达到最小.

如果在问题(P)中, 不含有未知参数 θ , 则系统中的不确定性仅有来自外界的随机干扰, 即 $\{v(k)\}, \{w(k)\}$. 此类不确定性无法减少只能通过 Kalman 滤波进行最优估计, 这就是经典的 LQG 最优控制问题. 在当前的问题(P)中, 未知参数 θ 模型化了另外一种不确定性, 如系统运行中参数的变化、元件的故障及一些未知的动态行为. 研究表明, 通过设计具有学习作用的控制器, 这种不确定性能够减少, 甚至完全消除, 参看本文给出的仿真实例.

3 最优对偶控制(Optimal dual control)

当 $\theta = \theta_i$ 时, 为书写方便, 令 $A_i = A(\theta_i)$, $B_i = B(\theta_i)$, $C_i = C(\theta_i)$. 定义 $\hat{x}_i(k|k)$ 是 k 时刻第 i 个模型的状态估计, 即

$$\hat{x}_i(k|k) = E\{x(k) | \theta = \theta_i, I^k\}.$$

上式可以用如下的 Kalman 滤波回归得到

$$\hat{x}_i(k|k) = A_i \hat{x}_i(k-1|k-1) + B_i u_i(k-1) + K_i(k) v_i(k), \quad (3)$$

$$K_i(k) = P_i(k|k-1) C_i^T [C_i P_i(k|k-1) + \sigma_w]^{-1}, \quad (4)$$

$$P_i(k|k-1) = A_i P_i(k-1|k-1) A_i^T + \sigma_v, \quad (5)$$

$$P_i(k|k) = [I - K_i(k) C_i] P_i(k|k-1). \quad (6)$$

其中: $v_i(k) = y(k) - C_i \hat{x}_i(k|k)$ 是新息序列, $y(k)$ 为系统的实际输出. 定义 k 时刻第 i 个模型的后验概率为

$$q_i(k) = P(\theta = \theta_i | I^k).$$

根据文献[6], 用 Bayes 公式, 基于当前的观测 $y(k)$, $q_i(k)$, 能够递推计算

$$q_i(k) = L_i(k) q_i(k-1) / \sum_{j=1}^s q_j(k-1) L_j(k).$$

其中初值 $q_i(0)$ 是如前所述的已知的验前概率, 且

$$L_i(k) = \left| P_{y_i}(k|k-1, \theta_i) \right|^{-1/2} \times \exp\left[-\frac{1}{2} v_i(k)^T P_{y_i}^{-1}(k|k-1, \theta_i) v_i(k)\right],$$

$$P_{y_i}(k|k-1, \theta_i) = C_i P_i(k|k-1, \theta_i) C_i^T + \sigma_w.$$

定义

$$J(k) = x(k)^T Q x(k) + u(k)^T R u(k),$$

$$k = 0, 1, \dots, N-1,$$

$$J(N) = x(N)^T Q x(N).$$

根据期望算子的光滑性,使问题(P)中性能指标达到最小的闭环最优控制是下面序列最小化问题的最优解:

$$\begin{aligned} \min J = & \min_{u(0), \dots, u(N-1)} E\{x(N)^T Qx(N) + \\ & \sum_{k=0}^{N-1} [x(k)^T Qx(k) + u(k)^T Ru(k)] | I^0\} = \\ & \min_{u(0)} E\{J(0) + \min_{u(1)} E\{J(1) + \dots + \\ & \min_{u(N-1)} E\{J(N) | I^{N-1}\} \dots I^1\} | I^0\}. \end{aligned} \quad (7)$$

为求解问题(7),定义阶段损失函数 $V^*(k)$ 为 k 时刻到终端时刻 N 的最优性能指标,即

$$V^*(k) = \min_{u(k), \dots, u(N-1)} E\left\{ \sum_{j=k}^N J(j) | I^k \right\},$$

那么根据式(7)

$$V^*(k) =$$

$$\begin{aligned} \min_{u(k)} E\{J(k) + \min_{u(k+1), \dots, u(N-1)} E\left\{ \sum_{j=k+1}^N J(j) | I^{k+1} \right\} | I^k\} = \\ \min_{u(k)} E\{J(k) + V^*(k+1) | I^k\}. \end{aligned} \quad (8)$$

边界条件为 $V^*(N) = 0$.

这就是求解问题(P)的动态规划方程,理论上有了方程(8),通过对其进行求解,就能获得所需要的对偶控制律.然而,问题的复杂性与困难从方程(8)才刚刚开始.事实上,时刻 k 以后的所有后验概率是未知,而 $V^*(k+1)$ 是未来后验概率的非线性函数,再加上期望算子的复合,导致了根本无法对方程(8)进行求解;与此同时,未来的后验概率依赖于时刻 k 及其之前的所有控制,而确定时刻 k 时的控制又要知道未来的后验概率,这是一个耦合.由上述分析可以看出,处理这类问题的主要思想在于设法打开这一耦合圈,下面将更为具体的分析这些困难.用动态规划原理,在时刻 $N-1$,

$$\begin{aligned} V(N-1) = & E\{J(N) + J(N-1) + V^*(N) | I^{N-1}\} = \\ & E\{x(N)^T Qx(N) + x(N-1)^T Qx(N-1) + \\ & u(N-1)^T Ru(N-1) | I^{N-1}\} = \\ & E\{[A(\theta)x(N-1) + B(\theta)u(N-1) + v(N-1)]^T \times \\ & Q[A(\theta)x(N-1) + B(\theta)u(N-1) + v(N-1)] + \\ & x(N-1)^T Qx(N-1) + u(N-1)^T Ru(N-1) | I^{N-1}\} = \\ & \sum_{i=1}^5 q_i(N-1) \{ \hat{x}_i(N-1 | N-1) \}^T \times \\ & [A_i^T Q A_i + Q] \hat{x}_i(N-1 | N-1) + \\ & \text{tr}[(A_i^T Q A_i + Q P_i(N-1 | N-1))] + \\ & 2u(N-1)^T B_i^T Q A_i \hat{x}_i(N-1 | N-1) + \end{aligned}$$

$$u(N-1)^T [B_i^T Q B_i + R] u(N-1) + \text{tr}[Q \sigma_v]. \quad (9)$$

其中 $\text{tr}(\cdot)$ 表示矩阵的迹.

求解原问题在 $N-1$ 时刻的最优控制,就是通过最小化式(9),得到相应的 $u^*(N-1)$.可以看出,函数 $V(N-1)$ 是 $u(N-1)$ 的二次凸函数,因此其最小解是唯一的.得到 $u^*(N-1)$ 后,将其代入 $V(N-1)$,相应的最小值记为 $V^*(N-1)$.然后用动态规划导出的递推泛函方程(8)对 $N-2$ 时刻的最优控制进行求解,这时泛函方程(8)有如下形式:

$$V^*(N-2) = \min_{u(N-2)} E\{J(N-2) + V^*(N-1) | I^{N-2}\}. \quad (10)$$

如果 $V^*(N-1)$ 具有 LQG 问题中的闭合形式,继续对如上得到的各阶段的泛函方程反向递推求解,直至时刻 0,就可以求出各时刻的最优控制.然而,对偶控制问题远远没有这么简单.这一求解过程出现两个困难:其一,随着控制的进行,信息序列 $\{I^k\}$ 的维数不断增加,出现所谓的“维数灾”,这一困难用 Kalman 滤波(3)~(6),可通过状态估计及其估计方差两个变量来总结系统在 k 时刻以前的信息,这样,系统在任何时刻的实时信息可用这两个变量来代替,从而解决了维数灾问题;其二,由式(9)可以看出,显然在 $N-1$ 时刻,最优控制 $u^*(N-1)$ 为后验概率和状态估计的函数,这样 $V^*(N-1)$ 为后验概率 $q_i(N-1)$ 的非线性函数,使得 $V^*(N-1)$ 没有数学上的闭合形式,导致了无法对泛函方程(10)去求 $N-2$ 时刻的最优控制 $u^*(N-2)$,相应的 $V^*(N-2)$ 也无法求出.至此不能用动态规划原理去求解 $N-1$ 时刻以前的所有最优控制 $u^*(N-2), u^*(N-3), \dots, u^*(0)$,这表明求解 LQG 问题的有力工具—动态规划,对于问题(P)已经无能为力.用动态规划求解经典的 LQG 问题有两个突出特点:一个是能够得到闭环最优控制的解析解,另一个是控制律的结构具有分离性.就计算的复杂性和控制律的反馈结构而言,对随机系统施行最优控制,这两点是人们所期望的.为使动态规划能够用于求解问题(P),本文引入对偶控制的思想,使求解问题得到简化.这样所求的控制同时具有控制与辨识特点.获得该控制的关键在于用滚动优化策略打开了当前控制与未来后验概率间的耦合.就这点而言,本文提出的算法被称为基于滚动优化的对偶控制算法.

4 基于滚动优化的对偶控制(Receding horizon dual control)

对于 LQG 问题,可通过分离性原理和确定性等

价原理进行求解. 首先, 依据分离性原理, 将控制与状态估计分离开来, 通过求解 Bellman 方程得到最优控制增益, 运用 Kalman 滤波方法得到状态的最优估计. 再按照确定性等价原理, 将控制增益和状态估计合成为控制律. 然而, 对于存在不确定参数的线性二次型高斯问题, 分离性原理不成立. 鉴于问题的困难存在于对控制增益的求解而不在于状态估计, 采用如下的滚动优化方法对问题进行求解:

$$\begin{aligned} & \text{针对于不确定参数 } \theta_i, i = 1, 2, \dots, s, \text{ 定义} \\ & \bar{A} = \text{diag}(A_1, A_2, \dots, A_s), \\ & \bar{B} = [B_1^T, B_2^T, \dots, B_s^T]^T, \\ & \bar{W}(k, Q) = \\ & \text{diag}(q_1(k-1)Q, q_2(k-1)Q, \dots, q_s(k-1)Q), \\ & \bar{P}(k-1) = \\ & \text{diag}(P_1(k-1|k-1), P_2(k-1|k-1), \dots, \\ & P_s(k-1|k-1)). \end{aligned}$$

这样式(9)可以重新写为等价形式

$$\begin{aligned} V(N-1) = & \hat{X}^T(N-1)[\bar{A}^T \bar{W}(N, Q) \bar{A} + \\ & \bar{W}(N, Q)] \hat{X}(N-1) + \text{tr}[\bar{A}^T \bar{W}(N, Q) \bar{A} \bar{P}(N-1)] + \\ & 2u^T(N-1) \bar{B}^T \bar{W}(N, Q) \bar{A} \hat{X}(N-1) + \\ & u^T(N-1)[\bar{B}^T \bar{W}(N, Q) \bar{B} + R]u(N-1) + \\ & \text{tr}[Q\sigma_v]. \end{aligned}$$

关于 $u(N-1)$ 最小化上式, 则有

$$u^*(N-1) = -\Gamma(N-1)\hat{X}(N-1). \quad (11)$$

其中

$$\begin{aligned} \Gamma(N-1) &= D(N-1)^{-1} \bar{B}^T \bar{W}(N, Q) \bar{A}, \\ D(N-1) &= \bar{B}^T \bar{W}(N, Q) \bar{B} + R, \\ \hat{X}(N-1) &= \\ & [\hat{x}_1(N-1|N-1)^T, \\ & \hat{x}_2(N-1|N-1)^T, \dots, \hat{x}_s(N-1|N-1)^T]^T. \end{aligned}$$

把最优控制(11)代回 $V(N-1)$, 有

$$\begin{aligned} V^*(N-1) &= \\ & \hat{X}^T(N-1)S(N-1)\hat{X}(N-1) + T(N-1). \end{aligned} \quad (12)$$

其中

$$\begin{aligned} S(N-1) &= \bar{A}^T \bar{W}(N, Q) \bar{A} + \bar{W}(N, \bar{Q}) - \\ & \Gamma^T(N-1)D(N-1)\Gamma(N-1), \quad (13) \\ T(N-1) &= \text{tr}[\bar{A}^T \bar{W}(N, Q) \bar{A} \bar{P}(N-1)] + \text{tr}[Q\sigma_v]. \end{aligned}$$

由于 $\Gamma(N-1)$ 是 $\bar{W}(N, Q)$ 的非线性函数, 而 $\bar{W}(N, Q)$ 又是后验概率 $q_i(N-1)$ 的函数, 根据

式(13), $S(N-1)$ 为后验概率 $q_i(N-1)$ 的非线性函数. 随着控制的进行, $S(N-1)$ 的复杂程度也不断增加, 这就是求解最优对偶控制的本质困难, 用不同的方法克服这一困难, 可以产生不同的次优控制律, 然而由于目前的系统模型中存在不确定性参数, 因此期望所获得的次优控制具有学习特点, 为达到这一目的本文采用下面的方法.

由于在 $N-1$ 时刻的前一时刻 $N-2$ 的后验概率 $q_i(N-2)$ ($i = 1, 2, \dots, s$) 是已知的, 因此, 在 $q_i(N-2)$ 处, 对 $S(N-1)$ 进行线性近似, 即把 $S(N-1)$ 在 $q_i(N-2)$ 处一阶展开, 忽略高次项, 这样就可以得到 $S(N-1)$ 的近似表达式

$$\begin{aligned} \tilde{S}(N-1) &= \\ & q_1(N-1)U_1(q(N-2)) + \\ & q_2(N-1)U_2(q(N-2)) + \dots + \\ & q_s(N-1)U_s(q(N-2)). \end{aligned}$$

其中: $U_i(q(N-2))$ 与 $q_i(N-1)$ 无关, 仅与 $q_i(N-2)$ 有关, 而 $q_i(N-2)$ 在 $N-1$ 时刻是已知. 这样 $\tilde{S}(N-1)$ 关于每个 $q_i(N-1)$ 是线性的.

动态规划方程(8), 在 $N-2$ 时刻的形式为

$$V^*(N-2) = \min_{u(N-2)} E\{J(N-2) + V^*(N-1) | I^{N-2}\}. \quad (14)$$

在上面的式(14)中, 用 $\tilde{S}(N-1)$ 代替 $V^*(N-1)$ 中的 $S(N-1)$ 有

$$\begin{aligned} \tilde{V}(N-2) &= \\ & E\{x(N-2)^T Q x(N-2) + u(N-2)^T R u(N-2) + \\ & T(N-1) + \hat{X}^T(N-1)\tilde{S}(N-1)\hat{X}(N-1) | I^{N-2}\}. \end{aligned}$$

由于 $\tilde{S}(N-1)$ 为 $q_i(N-1)$ 的线性函数, 因此上式中的第四项可以用期望算子的光滑性显式求出, 用与 $N-1$ 阶段相同的方法, 对 $\tilde{V}(N-2)$ 关于 $u(N-2)$ 求极小, 可以推导出 $N-2$ 时刻的控制为

$$u^*(N-2) = -\Gamma(N-2)\hat{X}(N-2).$$

其中

$$\begin{aligned} \Gamma(N-2) &= D(N-2)^{-1} \bar{B}^T \bar{W}(N-1, \tilde{S}(N-1)) \bar{A}, \\ D(N-2) &= \bar{B}^T \bar{W}(N-1, \tilde{S}(N-1)) \bar{B} + R. \end{aligned}$$

把 $u^*(N-2)$ 代入对应的 $\tilde{V}(N-2)$, 可以得到时刻 $N-2$ 的最优性能指标的近似解析表达式, 其形式仍然是二次型, 即

$$\begin{aligned} \tilde{V}^*(N-2) &= \\ & \hat{X}^T(N-2)S(N-2)\hat{X}(N-2) + T(N-2). \end{aligned}$$

再用与 $N-2$ 时刻类似方法, 在 $q_i(N-3)$ 处, 对 $S(N-2)$ 进行线性化, 得到

$$\tilde{S}(N-2) = \sum_{i=1}^4 q_i(N-2) U_i(q(N-3)),$$

这样可以得到 $N-3$ 时刻的控制。

如上所述,继续类似的推导,能够得到一个简单而具有解析表达形式的次优控制

$$u^*(k) = -\Gamma(k)\hat{X}(k), \quad k = 0, 1, \dots, N-1.$$

上述控制策略与文献[9]主要区别在于,在对 $S(k)$ 近似时,针对的是由滚动优化得到的后验概率,而不是初始的先验概率。显然,随着控制的不断进行,后验概率必然逐渐接近真实参数的概率。如果在对 $S(k)$ 进行近似时,仅针对初始先验概率,而不考虑当前时刻的后验概率,就未能充分利用所得信息。因此,在文献[9]的基础上,提出“滚动”优化的方法是很自然的想法。即将权值表达式 $S(k)$ 关于当前时刻的后验概率 $q(k)$ 进行线性近似,得到控制量后,施加于系统,观测输出,再次更新后验概率,再重新对 $S(k)$ 进行线性近似得到 $\tilde{S}(k)$, 从而算出新的控制量,直到整个控制过程结束。

滚动优化算法较针对先验概率的近似方法而言,由于在整个过程中都需要反复运用动态规划的方法,对权值表达式 $S(k)$ 进行近似,增大了运算量。然而,由于对信息的充分利用,获得了较好的控制效果。这种以牺牲简洁性换取的控制效果,是对偶思想的具体体现。从上述推导过程可以看出:本文提出的滚动对偶控制算法,一方面,把未来未知的后验概率固定在当前已知的后验概率上,打破了控制与后验概率间的耦合,同时保证了所得的控制具有学习特点,但是显然,这种控制是开环控制,因为它忽略了控制器未来的学习作用;另一方面,在每个时刻 k ,对 $V^*(k)$ 中的 $S(k)$,在前一时刻的后验概率 $q_i(k-1)$ 处进行线性化,保证了动态规划的顺利运用。但是值得指出的是:滚动对偶控制律的产生基于上述的近似。因此,所得到的控制尽管有对偶特点,但必定是次优的,至于最优对偶控制,目前还是一个挑战性课题。

5 仿真分析(Simulation analysis)

考虑如下包含参数不确定性的二阶系统

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + w(k), \\ y(k) &= Cx(k) + v(k). \end{aligned}$$

其中: $A = \begin{bmatrix} 1 & 0.1 \\ 0.1 & 1 - 0.2\theta \end{bmatrix}$, $B = [0.05, 0.1]^T$, $C = [0.3, 0.7]$, $Q = 0.4$, $R = 0.3$ 。不确定参数 $\theta \in \Omega_\theta = \{\theta_1, \theta_2\}$, 性能指标为

$$J = E\{x(50)^T Qx(50) + \sum_{k=0}^{49} [x(k)^T Qx(k) + u(k)Ru(k)]\}.$$

初始信息 I^0 为 $q(0) = [0.5, 0.5]^T$, 初始状态 $x(0)$ 的均值为 $[0, -6]^T$, 方差为二阶单位阵 I_2 。给定 $\theta_1 = 0.1, \theta_2 = 0.9, \theta_1$ 对应的模型为真实模型。采用不同的控制策略对系统施行控制。最优控制为参数已知时,即 $\theta = \theta_1$,由动态规划得到的控制律;非真实模型的控制是参数 θ_2 对应的动态规划的最优解;基于先验概率近似的对偶控制是采用针对初始先验概率 $q(0)$ 求解次优控制量,这是一种固定近似方法;基于滚动优化的对偶控制如前所述。图1~3分别表示了100次 Monte Carlo 仿真的性能指标和单次运行控制量和单次运行各模型的后验概率。图中: k 为运算步数, J 为性能指标, u 为控制量, p 为后验概率。

图1中,对偶控制算法的性能指标明显优于非对偶控制,滚动优化算法的性能指标在前期较基于先验概率近似的对偶控制有较小的性能指标。这说明该算法在该阶段有更好的辨识与控制的权衡。随着控制的进行,辨识也逐渐结束,滚动优化控制的优势逐渐失去。滚动优化控制与基于先验概率近似的对偶控制趋于重合。

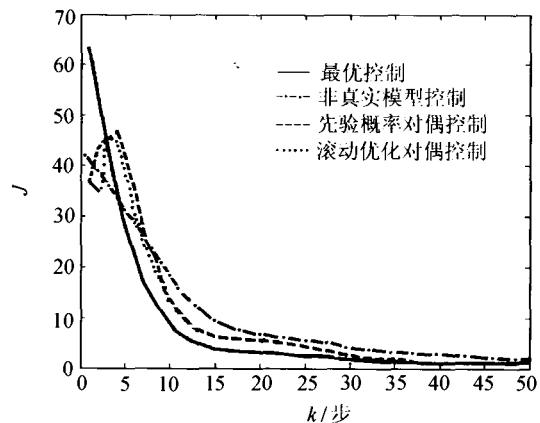
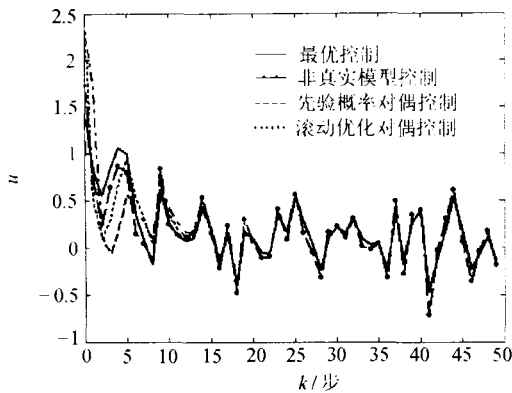
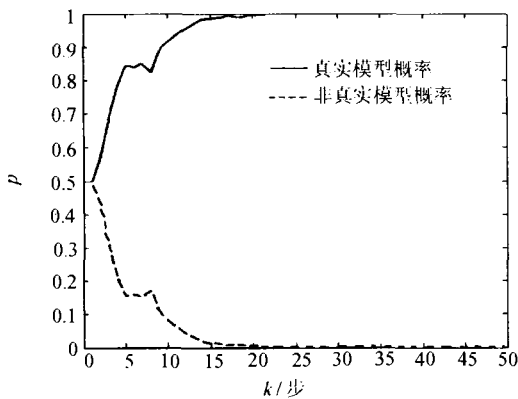


图1 Monte Carlo 仿真
Fig. 1 Monte Carlo simulations

参照图2和图3,分析控制与探测的过程。可以看出,在对偶控制对系统的辨识结束(正确模型后验概率趋于1,非正确模型后验概率趋于0)以前,即大约在20步前,4种控制的控制量区别较大,但滚动优化的控制更接近最优控制,因而确保了较小的性能指标。在辨识结束后,由于探测出正确的参数,对偶控制和最优控制保持一致。非对偶控制无法消除不确定性,故控制量与最优控制存在偏差。

图2 控制律 u Fig. 2 Control law u 图3 后验概率 p Fig. 3 Posterior probabilities p

对上述系统,改变不确定参数的取值,分别进行100次 Monte Carlo 仿真,表1给出了不同不确定参数下各种控制算法的性能指标.由表1知,不确定性的加大(θ_1, θ_2 差值的加大),引起性能指标的增加.非对偶控制的性能指标明显较大,其原因在于没有考虑不确定性的因素.基于滚动优化的对偶控制较固定近似对偶控制的性能指标小,是充分运用了计算得到的后验概率中所包含的信息的结果.

表1 不同不确定参数下 Monte Carlo 仿真性能指标比较
Table 1 Performance index of Monte Carlo runs comparison for different uncertain parameters

θ_1, θ_2	最优控制	非对偶控制	固定近似	滚动优化
0.1, 0.9	408.4479	517.5988	492.2280	488.6375
0.3, 0.7	283.7501	369.5799	348.8661	319.6473
0.4, 0.6	200.4503	272.5656	252.4170	222.1949

6 结论(Conclusion)

本文分析了求解最优对偶控制问题的难点,对于有参数不确定性的 LQG 控制问题,引入 Kalman 滤波解决信息集的维数增长问题,通过对阶段损失函数中权值矩阵关于前一时刻的后验概率进行线性

化近似,打破了当前控制与未来未知后验概率的耦合环,克服了 Bellman 方程中遇到的困难,从而得到了简单并且具有解析形式的阶段损失函数的表达式.为了更好的保持控制算法的对偶特性,充分利用得到的信息,运用了滚动优化的策略求解该问题.仿真表明,本文提出的次优控制方法,较好的解决了控制与探测的矛盾与冲突,消除了系统中的参数不确定性,呈现出良好对偶的特性.

参考文献(References):

- [1] FELDBAUM A A. Dual control theory I-II [J]. *Automatic Remote Control*, 1960, 21(4): 1033 - 1039.
- [2] 梁军. 对偶自适应控制[J]. *控制理论与应用*, 1997, 14(3): 297 - 305.
(LIANG Jun. Dual adaptive control [J]. *Control Theory & Applications*, 1997, 14(3): 297 - 305.)
- [3] FILATOR N M, UNBEHAUEN H. Survey of adaptive dual control methods [J]. *IEE Proc - Control Theory and Applications*, 2000, 147(1): 118 - 128.
- [4] LI Duan, QIAN Fucui, FU Peilin. Research on dual control [J]. *Acta Automatica Sinica*, 2005, 31(1): 32 - 42.
- [5] LI Duan, QIAN Fucui, FU Peilin. Variance minimization in stochastic systems [M] // YAO Daivede, ZHANG Hanqin, ZHOU Xinyu. *Stochastic Models and Optimization*. New York: Springer-Verlag, 2002.
- [6] LI Duan, QIAN Fucui, FU Peilin. Variance minimization approach for a class of dual control problems [J]. *IEEE Trans on Automatic Control*, 2002, 47(12): 2010 - 2020.
- [7] FU Peilin, LI Duan, QIAN Fucui. Active dual control for Linear-Quadratic Gaussian system with unknown parameters [C] // *Proc of the 15th IFAC World Congress*. Barcelona, Spain: Pergamon, 2002.
- [8] 钱富才, 李云霞, 刘丁. 基于两级算法的对偶控制[J], 2004, 21(1): 89 - 93.
(QIAN Fucui, LI Yunxia, LIU Ding. Dual control based on two-level algorithm [J]. *Control Theory & Applications*, 2004, 21(1): 89 - 93.)
- [9] STEIN G, SARIDIS G N. A parameter adaptive control technical [J]. *Automatica*, 1971, 5(2): 731 - 739.

作者简介:

钱富才 (1963—),男,教授,1998年获西安交通大学系统工程研究所博士学位,1999年在香港中文大学系统工程与工程管理系从事博士后研究工作,已在 *IEEE Trans on Automatic Control* 等刊物上发表论文 40 多篇,研究领域为大系统理论、非线性系统、最优控制、随机系统的辨识与控制、自适应控制, E-mail: fcqian@xaut.edu.cn;

宋俐 (1958—),女,教授,1999年获西安交通大学固体力学博士学位,研究领域为机械工程、土木工程、结构优化设计, E-mail: songli@xaut.edu.cn;

陈小可 (1978—),男,2004年获得西安理工大学硕士学位,研究方向为随机系统控制, E-mail: xkchen@sohu.com.