

噪声未知的LQG控制问题研究

钱富才, 朱少平, 刘 丁

(西安理工大学 自动化与信息工程学院, 陕西 西安 710048)

摘要: 对具有未知方差的线性二次高斯(LQG)问题, 本文提出了具有学习特点的对偶自适应控制方法. 控制器一方面能够控制系统朝着期望的目标运行, 另一方面又能对未知参数进行学习, 仿真结果表明了控制方法的有效性.

关键词: 鲁棒控制; LQG控制; 对偶控制; Kalman滤波

中图分类号: TP273 **文献标识码:** A

On LQG problems with unknown noises

QIAN Fu-cai, ZHU Shao-ping, LIU Ding

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

Abstract: For the control problem of linear-quadratic-Gaussian(LQG) with unknown variances of noises, a dual adaptive control approach with learning features is proposed. The controller, on the one hand, drives the system toward a desired state; on the other hand, it performs the learning process to reduce the uncertainties of unknown parameters. Simulation results validate the proposed approach.

Key words: robust control; LQG control; dual control; Kalman filter

1 引言(Introduction)

LQG问题以其工程上的实用性和数学上的易处理性, 受到了控制理论界和工程应用界的普遍重视, 迄今为止, 航空、航天、工业系统和社会经济系统中的大量问题都可以在LQG框架下获得令人满意的控制效果. 然而, 与确定性系统不同, 对于运行在随机环境下的系统, 尽管系统是线性的、性能指标为二次型, 但要设计出最优控制器, 必须知道作用于状态方程和观测方程中随机噪声的确切统计特性, 它们必须是验前已知. 只有这样才能设计出理想的最优控制器, 理想之处在于控制器与滤波器具有分离性, 即控制律等于控制增益与系统当前时刻最佳状态估计之积. 控制增益与状态估计无关, 可以用动态规划离线求出; 状态估计可以用Kalman滤波在线获得, 两者之间可以独立计算, 在线合成. 但是, 实际遇到的系统由于随机干扰的非平稳性、系统未建模部分的非线性等原因, 往往不能验前确切地获得随机干扰的统计特性. 这些非理想因素的影响, 导致在对随机过程的统计描述时, 不可避免的存在一定程度的未知性和不确定性.

噪声统计特性未知的滤波与控制问题已经获得

了一些研究结果^[1~3], 传统的方法是在对策论的框架下, 讨论了当噪声的二阶矩未知, 但在某个凸集内变动时, 对于其对应的最坏二次型性能指标, 施行最优控制, 获得了所谓的最小-最大控制, 给出了刻化鞍点的充分必要条件. 最小-最大意义下的最优控制对应的最坏过程噪声和测量噪声, 不同于最优估计器对应的最坏噪声, 只有当噪声的未知统计特性仅出现在状态方程, 或者仅存在于测量方程时, 最优控制对应的模型与最优估计器对应的模型才是一致的, 这无疑增加了问题的求解难度, 限制了这种方法的应用范围. 另一种方法就是最近几十年里发展起来的LMI方法, 首先把控制问题转化为一个松弛的优化问题, 再用LMI技术求解优化问题, 获得控制增益^[4~6]. 这样设计出的控制器是针对最坏情况下的噪声, 而系统在运行过程中的大多数时间范围内可能“坏”不到这种程度, 因此, 这种方法设计出的控制器有明显的保守性.

本文假定影响系统噪声的结构特性是已知的, 为高斯噪声, 噪声的方差未知, 但其真值是在某个有限集合内. 对这类问题, 在对偶控制框架下, 设计出了具有学习作用的控制律. 控制器一方面能使性能指标最小, 另一方面又具有学习作用, 如在仿真例子中

收稿日期: 2009-05-02; 收修改稿日期: 2009-09-19.

基金项目: 国家自然科学基金资助项目(60874033); 陕西省自然科学基金资助项目(SJ08F27); 高等学校博士学科点专项科研基金资助项目(20060700007).

所揭示的那样,最终能够学习出噪声二阶矩的真值.基于本文方法设计出的控制器与传统的最小-最大和现代的LMI方法的最大区别,就是降低、甚至完全消除了控制器的保守性.

2 问题的描述(Problem statement)

考虑如下随机系统:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + w(k), \\ k &= 0, 1, \dots, N-1, \end{aligned} \quad (1)$$

$$z(k) = Cx(k) + v(k), k = 1, 2, \dots, N, \quad (2)$$

其中: $x(k) \in \mathbb{R}^n$ 为状态向量, $z(k) \in \mathbb{R}^m$ 为量测向量, $u(k) \in \mathbb{R}^p$ 为控制向量. 矩阵 A , B 和 C 具有适当维数, 且已知. 初始状态向量 $x(0)$ 是均值为 $\bar{x}_0$, 方差为 P_0 的高斯白噪声, $w(k)$ 和 $v(k)$ 为不相关的过程噪声和测量噪声, 它们与随机的初始状态 $x(0)$ 也不相关. $w(k)$ 和 $v(k)$ 是均值为 0, 方差分别为 W 和 V 的高斯白噪声, 即:

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

$$v(k) \sim N(0, V).$$

其中: W 和 V 是噪声的方差矩阵, 假设是未知的, 但是是常数, 分别取值于有限空间 Ω_w 和 Ω_v , 这里:

$$\Omega_w = \{W_1, W_2, \dots, W_s\},$$

$$\Omega_v = \{V_1, V_2, \dots, V_s\}.$$

系统(1)和(2)中, 噪声对 $(w(k), v(k))$ 的方差可能的取值集合为

$$\Omega = \{(W_1, V_1), (W_2, V_2), \dots, (W_s, V_s)\}.$$

度量系统的性能指标为

$$J = x^T(N)Qx(N) + \sum_{k=0}^{N-1} [x^T(k)Qx(k) + u^T(k)Ru(k)], \quad (3)$$

其中: Q 为有适当维数的半正定矩阵, R 为有适当维数的正定矩阵. 如果矩阵 A , B , C , Q 和 R 是时变的, 本文的方法同样适用, 为书写简单起见, 考虑时不变情形.

对于随机系统(1), 在 k 时刻, 控制器能够知道系统的信息为除了先验信息外, 还有 k 时刻以前, 控制器对系统施加的控制 $\{u(0), u(1), \dots, u(k-1)\}$ 及直至 k 时刻得到系统的测量信息 $\{z(1), z(2), \dots, z(k)\}$, 因此定义实时信息集为

$$I^k = \{u(0), \dots, u(k-1), z(1), \dots, z(k)\}.$$

本文要解决的问题是, 求具有 $u(k) = f_k(I^k)$ 形式的控制律, 使性能指标 $E\{J\}$ 最小, 即:

$$(P) \min E\{J\},$$

$$\text{s.t. } x(k+1) = Ax(k) + Bu(k) + w(k),$$

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

$$z(k) = Cx(k) + v(k), k = 1, 2, \dots, N.$$

控制律约束成 $u(k) = f_k(I^k)$ 这种形式, 是工程实际的要求, 因为在当前时刻控制器能够知道的系统信息是 I^k , 因此, 这样的控制律在物理上才是可实现的, 是允许的. 与随机系统不同, 对于确定性系统而言, 系统当前的状态完全概括了系统的历史信息, 这也是确定性的控制律一般约束为状态反馈形式的原因, 粗略地说, 在随机系统中实时信息集 I^k 相当与确定性系统中的状态, 它能够确定系统的行为, 而随机系统的状态是以概率的形式转移或演化的.

当干扰于系统的过程噪声与测量噪声的方差 W 和 V 已知时, 上述控制问题就是经典的LQG问题, 已经有成熟的求解方法; 当 W 和 V 未知时, 且其在某个凸集内变化, 文献[4]是在对策论框架下, 获得了最小最大控制律的典型方法. 本文的问题是作用于系统的外界干扰的强度是未知的, 但对这样的两个未知量并非不知道任何信息, 假定知道其真值在有限集 Ω_w 和 Ω_v 内. 这是和LQG与文献[4]完全不同的情况. 另外, 本文解决的问题(P)与LQG不同, 后者只有外部噪声的不确定性, 而前者除了外部噪声的不确定性外, 噪声的方差还是未知的常数, 这类不确定性称为参数不确定性. 外部噪声的不确定性是客观存在, 不能减少, 但控制器可以用Kalman滤波技术对系统的真实状态进行估计, 而系统中的参数不确定性, 通过控制器不断的从系统中获得信息, 可以学习出未知参数, 因此这类不确定性可以减少, 甚至完全消除.

能使性能指标 $E\{J\}$ 达到最优, 同时又能兼顾学习出未知参数 W 和 V 的真值的控制器称其为对偶控制. 作者在文献[7~12]中对其进行了深入研究, 获得了一些极有意义的结果.

3 对偶控制策略(Dual control strategy)

对于给定的方差对 (W, V) 则有一个LQG问题与之对应, 解这个LQG问题, 可以获得此种噪声级别下的最优控制. 根据 Ω 的定义, Ω 内有 s 个不同元素对, 因此, 方差对 (W, V) 有 s 个. 这意味着需要求解 s 个LQG问题. 为了后面讨论方便起见, 这里约定, 用参数 θ 表示 (W, V) 对, $\theta = \theta_j$ 表示 (W, V) 取值为 (W_j, V_j) , 此时对应的系统模型被称为第 j 个模型.

定义 1 假设 k 时刻实时信息为 I^k , 第 j 个系统的状态 $x_j(k)$ 基于 I^k 的估计为

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

$\hat{x}_j(k|k)$ 可用Kalman滤波方程递推得到:

$$\begin{aligned}\hat{x}_j(k|k) &= A\hat{x}_j(k-1|k-1) + Bu_j(k-1) + \\ &F_j(k)v_j(k),\end{aligned}\quad (4)$$

其中:

$$F_j(k) = P_j(k|k)C^T[CP_j(k|k-1)C^T + V_j]^{-1}, \quad (5)$$

$$P_j(k|k-1) = AP_j(k-1|k-1)A^T + W_j, \quad (6)$$

$$P_j(k|k) = [I - F_j(k)C]P_j(k|k-1), \quad (7)$$

$$v_j(k) = z(k) - C\hat{x}_j(k|k). \quad (8)$$

边界条件为 $\hat{x}_j(0|0) = \hat{x}(0)$ 和 $P_j(0|0) = P_0$.

定理 1 对于第 j 个模型, 使性能指标 $E\{J\}$ 达到最小的控制律为

$$u_j^*(k) = -\Gamma(k)\hat{x}_j(k|k), \quad (9)$$

其中对于 $k = N-1, N-2, \dots, 0$,

$$\Gamma(k) = D^{-1}(k)B^T S(k+1)A, \quad (10)$$

$$D(k) = B^T S(k+1)B + R, \quad (11)$$

$$S(k) = A^T S(k+1)A + Q - \Gamma^T(k)D(k)\Gamma(k), \quad (12)$$

边界条件为 $S(N) = Q$.

证 对于给定的 $j, j \in \{1, 2, \dots, s\}$, 定义:

$$\begin{aligned}J_j^*(k) &= \min_{u(k), u(k+1), \dots, u(N-1)} E\{x^T(N)Qx(N) + \\ &\sum_{j=k}^{N-1} [x^T(j)Qx(j) + u^T(j)Ru(j)] | I^k, \theta = \theta_j\}.\end{aligned}$$

根据实时信息集 I^k 的定义, 显然, $I^k \subset I^{k+1}$, 因此, 期望算子如下的光滑性成立:

$$E\{(\cdot)|I^k\} = E\{E\{(\cdot)|I^{k+1}|I^k\}.$$

根据随机动态规划的最优性原理和上述期望算子的光滑性, 有

$$\begin{aligned}J_j^*(k) &= \min_{u(k)} E\{x^T(k)Qx(k) + u^T(k)Ru(k) + \\ &J_j^*(k+1) | I^k, \theta = \theta_j\}.\end{aligned}\quad (13)$$

边界条件为

$$J_j^*(N) = E\{x^T(N)Qx(N) | I^N, \theta = \theta_j\}. \quad (14)$$

到此为止, 获得了第 j 个模型对应的Bellman动态规划方程, 如果每个阶段的控制都符合方程(13), 用这些控制量就可以对第 j 个模型表示的系统施行最优控制.

下面用归纳法解方程(13). 在阶段 $N-1$, 有

$$\begin{aligned}J_j(N-1) &= E\{x_j^T(N-1)Qx_j(N-1) + \\ &u_j^T(N-1)Ru_j(N-1) + J_j^*(N) | I^{N-1}\}.\end{aligned}\quad (15)$$

把边界条件(14)代入上式(15), 经过简单的计算整理, 有

$$\begin{aligned}J_j(N-1) &= \hat{x}_j^T(N-1|N-1)[A^TQA + Q] \times \hat{x}_j(N-1| \\ &N-1) + 2u_j^T(N-1)B^TQA\hat{x}_j(N-1) + \\ &u_j^T(N-1)[B^TQB + R]u_j(N-1) + \\ &\text{tr}(A^TQA + Q)P_j(N-1|N-1) + \text{tr}QW_j.\end{aligned}\quad (16)$$

泛函 $J_j(N-1)$ 关于 $u_j(N-1)$ 最小化, 就可以获得 $N-1$ 阶段的最优控制

$$u_j^*(N-1) = -\Gamma(N-1)\hat{x}_j(N-1|N-1),$$

其中 $\Gamma(N-1)$ 与式(10)一致. 把 $N-1$ 阶段的最优控制 $u_j^*(N-1)$ 代入性能指标(16)中, 则

$$\begin{aligned}J_j^*(N-1) &= \hat{x}_j^T(N-1|N-1)S(N-1) \times \\ &\hat{x}_j(N-1|N-1) + s_j(N-1),\end{aligned}$$

其中 $S(N-1)$ 与式(12)一致, $s(N-1)$ 与控制 and 状态无关, 且

$$s_j(N-1) = \text{tr}[Q(AP_j(N-1|N-1)A^T + W_j)].$$

假定在 $k+1$ 阶段

$$\begin{aligned}J_j^*(k+1) &= \hat{x}_j^T(k+1|k+1)S(k+1) \times \\ &\hat{x}_j(k+1|k+1) + s_j(k+1).\end{aligned}\quad (17)$$

根据Bellman动态规划方程(13), 则

$$\begin{aligned}J_j^*(k) &= \min_{u(k)} E\{x_j^T(k)Qx_j(k) + u_j^T(k)Ru_j(k) + \\ &J_j^*(k+1) | I^k\},\end{aligned}\quad (18)$$

把Kalman滤波方程(4)代入方程(17), 再把其结果代入方程(18), 经过简单整理可得

$$\begin{aligned}J_j(k) &= \hat{x}_j^T(k|k)[A^TQA + Q]\hat{x}_j(k|k) + \\ &2u_j^T(k)B^TQA\hat{x}_j(k) + \\ &u_j^T(k)[B^TQB + R]u_j(k) + s_j(k+1) + \\ &\text{tr}(A^TQA + Q)P_j(k|k) + \text{tr}QW_j.\end{aligned}$$

泛函 $J_j(k)$ 关于 $u_j(k)$ 最小化, 就可以获得 k 阶段的最优控制

$$u_j^*(k) = -\Gamma(k)\hat{x}_j(k|k),$$

其中 $\Gamma(k)$ 与式(10)一致. 把 k 阶段的最优控制 $u_j^*(k)$

代入性能指标 $J_j(k)$ 中,则

$$J_j^*(k) = \hat{x}_j^T(k|k)S(k)\hat{x}_j(k|k) + s_j(k), \quad (19)$$

其中 $S(k)$ 与式(12)一致, $s(k)$ 与控制 and 状态无关,

$$s_j(k) = s_j(k+1) + \text{tr}[Q(AP_j(k|k)A^T + W_j)].$$

证毕.

由定理1可以看出, 第 j 个模型对应的最优控制 $u_j^*(k)$ 由两部分组成, 一部分为控制增益 $\Gamma(k)$, 另一部分为状态估计 $\hat{x}_j(k|k)$, 前者由式(10)~(12)确定, 后者由式(4)~(8)确定, 明显两者之间具有分离性. 特别是式(10)~(12)与状态无关, 只与系统矩阵和性能指标中的权矩阵有关, 因此, $\Gamma(k)$ 可以离线计算; $\hat{x}_j(k|k)$ 中包含新息项, 它需要在线计算. 这个特点为工程应用带来了极大的方便, 首先离线计算出控制增益 $\Gamma(k)$, 然后, 在线测量, 用测量结果对状态进行估计, 获得 $\hat{x}_j(k|k)$, 两者之积就是当前的最优控制, 减少了在线计算负担.

由于噪声参数 θ 有 s 个, 根据定理1每一个参数 θ 对应一个最优控制, 这样 s 个参数 θ 对应 s 个最优控制 $u_1^*(k), u_2^*(k), \dots, u_s^*(k)$, 因此, 从系统中获得一次测量就有 s 个最优控制与之对应, 到底哪一个最优控制施加于实际系统, 是一个必须解决的问题, 下面的定理回答了这个问题.

定理 2 假设在时刻 k , 第 j 个模型的后验概率为 $q_j(k)$, 则施加于实际系统的控制为

$$u^*(k) = \sum_{j=1}^s q_j(k)u_j^*(k). \quad (20)$$

证 定义 $q_j(k)$ 是第 j 个模型在时间 k 的后验概率, 则

$$q_j(k) = P(\theta = \theta_j | I^k), \quad k = 0, 1, \dots, N-1. \quad (21)$$

用文献[13]的方法, 后验概率 $q_j(k)$ ($j = 1, 2, \dots, s$) 可以如下递推获得

$$q_j(k) = \frac{L_j(k)}{\sum_{j=1}^s L_j(k)q_j(k-1)} q_j(k-1), \quad (22)$$

其中:

$$\begin{aligned} L_j(k) &= |P_z(k|k-1)|^{-\frac{1}{2}} \times \\ &\quad \exp[-\frac{1}{2}v_j^T(k)P_z^{-1}(k|k-1)v_j(k)], \\ P_z(k+1|k+1) &= CP_j(k+1|k)C^T + V_j, \\ v_j(k) &= z(k) - C\hat{x}_j(k|k). \end{aligned}$$

方程(22)的边界条件为:

$$q_1(0) = \frac{1}{s}, q_2(0) = \frac{1}{s}, \dots, q_s(0) = \frac{1}{s},$$

上述边界条件表明, 在初始时刻, s 个参数模型取真值的可能性是相同的, 没有偏好任何一个模型, 因此这种假定在概率意义下是最坏的假定, 即便如此, 从后面的仿真例子可以看出, 控制器最终能够学习出参数的真值.

如果系统的输出具有遍历性, 假设参数的真值为 $\theta = \theta_i$, 文献[14]证明了, 当 $k \rightarrow +\infty$, 对于 $j = i$, $q_j(k) \rightarrow 1$; 当 $k \rightarrow +\infty$, 对于 $j \neq i$, $q_j(k) \rightarrow 0$. 根据后验概率的上述性质, 施加于系统的最优控制 $u^*(k) = \sum_{j=1}^s q_j(k)u_j^*(k)$ 在 k 趋向无穷时, 趋于真值 θ_i 对应的最优控制 $u_i^*(k)$. 证毕.

从后验概率的收敛性可以看出, 只有在 k 趋向无穷时, 真值的后验概率趋向于1, 然而, 问题(P)为有限时间, 因此, 当 k 足够大时, 控制器才有学习性质. 每个控制律 $\{u_j^*(k)\}$ 能使性能指标 $E\{J\}$ 最小, 把这些控制律以后验概率为权系数, 加权 $\sum_{j=1}^s q_j(k)u_j^*(k)$ 施加于实际系统, 而权系数具有学习特点, 这样用本文方法设计的控制器具有学习与控制的双重作用.

值得注意的是, 噪声参数的个数与取值确定是一个非常重要的问题. 目前, 可以用如下两种方法来处理:

i) 噪声参数的个数可验前确定. 如空中飞行器, 在飞行过程中一般要跨几个区域, 每个区域有不同的噪声级别, 这些噪声级别就可以确定噪声参数的个数与取值.

ii) 用表示参数来确定. 就一个凸多面体而言, 对于其中的任何一点都可以用顶点(极点)的凸组合来表示, 这些顶点可以看成凸多面体的表示参数; 对于描述不确定性的参数 θ , 如果在其变化空间内, 可以找到这组表示参数, 显然, 表示参数集合就可以解决噪声参数的个数与取值问题. 一个简单情况, 就是参数 θ 是一维的、未知的, 但是是常数, 其取值于某个区间 $[a, b]$. 用一组分点 $a = a_1 < a_2 < \dots < a_s = b$, 把区间 $[a, b]$ 分成若干个互不相交的子区间, 即 $[a, b] = \bigcup_{i=1}^{s-1} [a_i, a_{i+1}]$, 取每个小区间的中点 c_i 为表示参数, 则 $\{c_1, c_2, \dots, c_{s-1}\}$ 就是表示参数集合. 用本文的方法, 一定能够学习到某个 c_j , 它与 θ 的真值最为接近, 这种情况下, c_j 对应的区间 $[a_j, a_{j+1}]$ 包含了 θ 的真值, 该区间的长度与分点的个数有关, 分点越多, 每个子区间的长度越短, 因此, 取适当的分点个数, 一定能使包含真值的最小区间的区间长度小于某个预先给定的正数 ε , 只是 ε 越小, 噪声参数的个数越多, 计算的代价越大, 这里的 ε 可视为验前给定的允许学习误差.

4 仿真分析(Simulation analysis)

综上所述, 前面导出的控制方法可通过下面的算法来实施.

Step 1 用递推方程(5)~(8)离线求出所有的滤波增益 $F_j(k)$;

Step 2 用递推方程(10)~(12)离线求出所有的控制增益 $\Gamma(k)$;

Step 3 令 $k = 0$;

Step 4 测量系统输出 $z(k)$, 用方程(4)对第 j 个模型的状态进行估计, 获得 $\hat{x}_j(k|k)$;

Step 5 计算第 j 个模型的最优控制 $u_j^*(k) = -\Gamma(k)\hat{x}_j(k|k)$;

Step 6 用递推关系(22)计算出 k 阶段 s 个后验概率 $q_1(k), q_2(k), \dots, q_s(k)$;

Step 7 把 $u^*(k) = \sum_{j=1}^s q_j(k)u_j^*(k)$ 用于实际系统;

Step 8 如果 k 等于 $N - 1$, 结束; 否则, 置 $k := k + 1$, 返回Step 4.

下面通过一个实例, 来说明用本文方法设计出的控制器的控制与学习特点.

考虑随机系统:

$$x(k+1) = 0.8x(k) + 0.5u(k) + w(k),$$

其中: $w(k) \sim N(0, W)$, $W \in \Omega_w = \{1, 0.2\}$, 初始状态, $x(0) \sim N(\bar{x}_0, P_0)$, $\bar{x}_0 = 2$, $P_0 = 1$.

测量方程为

$$z(k) = x(k) + v(k),$$

其中: $v(k) \sim N(0, V)$, $V \in \Omega_v = \{0.1, 0.8\}$. $\Omega = \{(1, 0.1), (0.2, 0.8)\}$, (W, V) 的真值对为 $(1, 0.1)$.

系统的性能指标为

$$J = E\{x^2(20) + \sum_{k=0}^{19} [x^2(k) + u^2(k)]\}.$$

先验概率 $q_1(0) = 0.5$, $q_2(0) = 0.5$.

首先, 根据定理1求出 (W, V) 的真值对 $(1, 0.1)$ 的最优控制和性能指标, 然后, 用本节的算法, 设计出对偶控制器, 通过两者的比较, 揭示本节算法的控制与学习特点. 值得指出的是, 最优控制是方差已知时的系统控制, 因此, 它的性能指标值是对偶控制对应的性能指标值的下界.

在以上条件下, 对例子中的动态系统, 用对偶控制和最优控制进行1000次Monte Carlo仿真, 比较结果如表1所示.

从表1可以看出, 本文给出的对偶控制的性能指

标在统计意义下与最优控制的性能指标非常接近, 而且两者之间的标准方差和最大方差也几乎相等, 这些表明, 对偶控制具有最优性和可靠性.

表1 对偶控制与最优控制的比较

Table 1 Comparison of dual control algorithm with optimal control algorithm

性能指标	最优控制	对偶控制
平均方差	3.4730	3.4925
最大方差	39.6382	41.9046
标准方差	4.1119	4.2037

图1和图2是一次仿真对应的后验概率、对偶控制和最优控制的演化情况, 它们说明了对偶控制的学习特点.

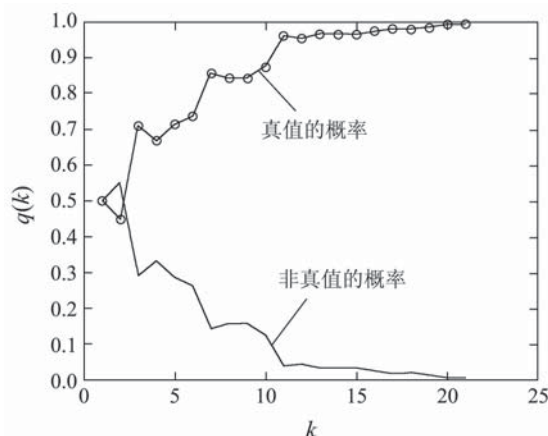


图1 两组参数的后验概率

Fig. 1 A-posteriori probabilities of two sets of parameters

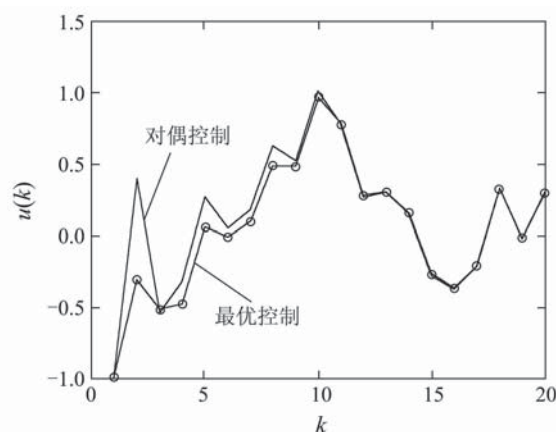


图2 最优控制与对偶控制

Fig. 2 Optimal control and dual control

观测图1, 参数的真值在不足20个单位时间内已经被完全获得, 说明对偶控制消除了系统中参数未知的不确定性. 与最大-最小控制器相比较, 对偶控

制对本例来说,减少了保守性;观测图2,除了最初的一段时间外,后继时间内对偶控制能够密切跟踪参数已知的最优控制,再次表明该方法设计出的控制器没有保守性. 初始阶段对偶控制的能量大于最优控制的能量,这是因为对于未知的参数,对偶控制器需要学习,因此,控制器除了追求控制目标外,还必须付出控制以外的能量以对未知参数进行学习,当参数学习成功后,对偶控制和最优控制几乎完全一致,一般来说,对偶控制器很难精确的学习出未知参数的真值,即便误差很小,但总会存在误差,这会导致对偶控制和最优控制在图形上难以完全重合. 对于当前仿真的例子,当控制时间超过20个单位时间后,对偶控制几乎和最优控制相同,特别在稳态时,系统可以认为是以最优的方式在运行.

5 结论(Conclusion)

本文研究了具有未知噪声的随机线性离散系统在二次性能指标下的最优控制问题,设计出了具有学习与调节作用的对偶控制器,仿真结果也表明了控制器的最优性、可靠性和学习特点. 用本文的方法设计出的控制器容易实现,由于控制器中的控制增益和滤波增益可以离线计算,在线调用,因此计算量小. 同时,与最小-最大控制、LMI导出的控制律机理比较来看,本文的算法为消除控制器的保守性提供了新方法.

参考文献(References):

- [1] CAREW B, BELAXGER P R. Identification of optimum filter steady-state gain for systems with unknown noise covariance[J]. *IEEE Transactions on Automatic Control*, 1973, 18(6): 582 – 587.
- [2] LOOZE D P, POOR H V, VADTOLA K, et al. Minimax control of linear stochastic systems with noise uncertainty[J]. *IEEE Transactions on Automatic Control*, 1983, 28(9): 882 – 888.
- [3] VERDU S, POOR H V. Minimax linear observers and regulators for stochastic systems with uncertain second-order statistics[J]. *IEEE Transactions on Automatic Control*, 1984, 29(6): 499 – 511.
- [4] 褚健, 俞立, 苏宏业. 鲁棒控制理论及应用[M]. 浙江: 浙江大学出版社, 2000.
- [5] 梅生伟, 申铁龙, 刘康志. 现代控制理论与应用[M]. 北京: 清华大学出版社, 2004.
- [6] 俞立. 鲁棒控制-线性矩阵不等式处理方法[M]. 北京: 清华大学出版社, 2002.
- [7] LI D, QIAN F C, FU P L. Variance minimization approach for a class of dual control problems[J]. *IEEE Transactions on Automatic Control*, 2002, 47(12): 2010 – 2020.
- [8] FU P L, LI D, QIAN F C. Active dual control for linear-quadratic Gaussian system with unknown parameters[C] // *Proceedings of the 15th IFAC World Congress*. Barcelona, Spain: Pergamon Press, 2002, 7: 21 – 26.
- [9] LI D, QIAN F C, FU P L. Variance minimization in stochastic systems[M] // ZHOU X Y (Ed). *Stochastic Modeling and Control*. New York: Springer, 2002.
- [10] LI D, QIAN F C, FU P L. Optimal nominal dual control for discrete-time LQG problem with unknown parameters[J]. *Automatica*, 2008, 44(1): 119 – 127.
- [11] LI D, QIAN F C, FU P L. Research on dual control[J]. *Acta Automatica Sinica*, 2005, 31(1): 32 – 42.
- [12] LI D, FU P L, QIAN F C. Guaranteed performance for discrete-time LQG problem performance[C] // *The 16th IFAC World Congress*. Prague, Czech Republic: Pergamon Press, 2005, 7: 4 – 8.
- [13] STEIN G, SARIDIS G N. A parameter adaptive control technique[J]. *Automatica*, 1969, 5(6): 731 – 739.
- [14] SARIDIS G N, DAO T K. A learning approach to the parameter adaptive self-organization control problem[J]. *Automatica*, 1972, 8(5): 589 – 597.

作者简介:

钱富才 (1963—), 男, 教授, 博士生导师, 已在IEEE Transactions on Automatic Control, Automatica等期刊发表论文60多篇, 主要研究方向为随机控制、系统辨识、非线性控制、最优控制和大规模系统, E-mail: fqcqian@xaut.edu.cn;

朱少平 (1963—), 男, 副教授, 目前研究方向为非线性控制, E-mail: zhushaoping@126.com;

刘丁 (1957—), 男, 教授, 博士生导师, 获多项国家和省部级科技奖, 发表论文200多篇, 主要研究方向为智能控制、复杂系统建模、计算机控制和控制理论与应用, E-mail: liud@xaut.edu.cn.