

具有学习特点的控制器的设计算法

尚 婷¹ 钱富才^{1,2} 张晓艳¹ 谢 国¹

摘 要 对于普遍存在的具有未知参数的随机最优控制问题, 本文提出了一种具有学习特点的控制器的设计算法. 该算法用 Kalman 滤波估计系统的未知参数, 在滚动优化机制下用动态规划获取控制增益, 为了赋予控制器的学习特点, 在 LQG 控制律中附加使下一时刻估计方差最小的学习控制分量. 仿真结果表明了算法的有效性.

关键词 自适应控制, 不确定系统, LQG 问题, 卡尔曼滤波

引用格式 尚婷, 钱富才, 张晓艳, 谢国. 具有学习特点的控制器的设计算法. 自动化学报, 2017, 43(7): 1202–1207

DOI 10.16383/j.aas.2017.c160493

A Controller Design Algorithm with Learning Property

SHANG Ting¹ QIAN Fu-Cai^{1,2} ZHANG Xiao-Yan¹ XIE Guo¹

Abstract A new controller design algorithm with learning characteristic is proposed for the ubiquitous stochastic optimal control problem with unknown parameters. This algorithm estimates system unknown parameters by Kalman filter and obtains control gains by dynamic programming and continuous rolling optimization mechanism. In order to endow the controller with learning characteristics a learning control component which minimizes next moment estimated variance is attached to the LQG control law. Simulation results show the effectiveness of the algorithm.

Key words Adaptive control, uncertainty systems, LQG problem, Kalman filter

Citation Shang Ting, Qian Fu-Cai, Zhang Xiao-Yan, Xie Guo. A controller design algorithm with learning property. *Acta Automatica Sinica*, 2017, 43(7): 1202–1207

半个多世纪以来, 理论界对于具有不确定性动态系统研究的兴趣与日俱增, 因为在所有的实际控制问题中几乎都存在着不同程度的不确定性, 这些不确定性可能来自于外部扰动、子系统故障、参数波动^[1]. 一些控制对象, 例如飞行器, 其速度、高度、环境噪声由于横跨高山、海洋、陆地以及飞行过程中的一些不太致命的故障, 使得模型中的参数很难确定^[2], 还有一些系统, 即使元器件的参数能够精确知道且系统模型的结构确定不变, 但由于建模简化、近似、工作环境恶劣, 使得模型中的等效参数与实际物理参数不存在确定的对应关系, 可能会以许多不同的形式出现较大的变化或波动^[3]. 因此, 对具有参数不确定性的系统实施最优控制, 关键在于对不确定性的数学描述和采用什么样的方法尽可能地消除

不确定性对控制性能的影响.

解决不确定性问题的途径已经有几种常用的方法. 理论与实践表明, 反馈能使闭环系统在一定程度上对控制过程的扰动变得不灵敏, 因此, 可以通过设计固定参数的控制器使得闭环系统对于过程的微弱扰动具有不敏感性, 圆满解决了扰动较小的不确定性系统的控制问题^[4]. 另一种解决不确定性系统的控制问题的方法就是采用鲁棒控制, 它能够解决有较大范围的不确定性, 但往往有保守性^[5]. 对于具有未知参数的动态系统, 无疑自适应控制要优于鲁棒控制, 因为在 Lyapunov 框架下对未知参数构造出自适应学习律, 只要与控制策略有效结合, 能够使系统全局稳定并能保证学习的收敛性, 克服了鲁棒控制的保守性^[6].

由于具有未知参数随机系统的最优控制问题或决策问题在经济学领域与控制领域普遍存在, 长期以来众多的学者在处理这类不确定性系统方面做出了巨大的贡献. 其中最具影响力的研究之一就是对偶控制^[7–10]. 对偶控制的思想由来已久, 不确定性与方差是两个被频繁提到的概念, 然而, 半个多世纪以来, 如何更好地在控制与辨识这两个相互冲突的目标之间进行折衷却依然悬而未决, 导致虽然用来自系统的信息对未知参数的估计朝着真值方向不断校正, 但所需时间一般较长, 难以保证闭环系统具备

收稿日期 2016-06-27 录用日期 2016-11-03

Manuscript received June 27, 2016; accepted November 3, 2016

国家自然科学基金 (61273127, U1534208, 61533014), 陕西省科技攻关项目 (2016GY-108) 资助

Supported by National Natural Science Foundation of China (61273127, U1534208, 61533014) and Science and Technology Project of Shaanxi Province (2016GY-108)

本文责任编辑 方海涛

Recommended by Associate Editor FANG Hai-Tao

1. 西安理工大学自动化与信息工程学院 西安 710048 2. 西安工业大学陕西省自主系统与智能控制国际联合研究中心 西安 710021

1. School of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048 2. The International Joint Research Center of Autonomous Systems and Intelligent Control, Xi'an Technological University, Xi'an 710021

稳、准、快的性质.

本文模型化了在控制过程中参数的波动现象, 利用 Kalman 滤波对系统中的未知参数进行估计, 在滚动优化机制下利用动态规划求出了最优控制的控制增益, 研究表明对于具有未知参数的 LQG 问题, 强化控制器的学习作用必不可少, 因此, 本文在 LQG 控制律的基础上附加了使下一时刻参数估计方差最小的学习控制分量, 仿真结果表明算法是有效的. 与传统的自适应控制相比, 本文给出的控制器设计方法对未知参数的学习时间不需要太长, 在有限时间内能够快速估计出未知参数的真值, 并能显著改善被控系统的瞬态性能, 使闭环系统具有稳、准、快的品质.

1 问题描述

描述实际随机动态系统常用的数学模型主要有离散时间受控自回归平移平均模型和随机状态空间模型, 两种模型可以相互转化, 取时间变量为离散时间主要是考虑到用控制律去控制实际系统绝大多数都是通过工控机来实现的. 因此, 本文考虑具有如下形式的完全信息不确定离散随机系统

$$x(k+1) = a(k)x(k) + b(k)u(k) + w(k), \\ k = 0, 1, \dots, N-1 \quad (1)$$

式中, $x(k)$ 是状态量, $u(k)$ 是控制量, $a(k)$ 和 $b(k)$ 为系统模型参数, 本文假定它们是未知的; $w(k)$ 为高斯白噪声, 服从 $N(0, \Sigma_w)$ 分布. 给定二次型性能指标

$$J = E \left\{ x^T(N)Q_N x(N) + \sum_{k=0}^{N-1} [x^T(k)Qx(k) + u^T(k)Ru(k)] | I^0 \right\} \quad (2)$$

其中, Q_N, Q 为非负定对称矩阵, R 为正定对称阵, I^0 为初始信息集合, 包含了初始状态和输出、初始未知参数的信息, 即在初始时刻, 初始状态的均值 $\hat{x}(0)$ 和方差 $P_x(0)$, 不确定参数的均值 $\bar{\theta}$ 和方差 $P_{\theta}(0)$, 所有这些信息均可视为已知的.

本文要解决的控制问题是: 寻求控制序列 $\{u^*(k)\}$ ($k = 0, 1, \dots, N-1$), 使动态系统 (1) 关于性能指标 (2) 最小.

该模型形式具有广泛的应用范围, 许多经济学模型、金融工程模型其状态变量是已知, 且大多数模型与该模型类似, 或为在此基础上的变形.

2 控制器设计

由于不确定参数的存在, 使得单纯使用基于“最

优性原理”的动态规划对于求解上述问题显得无能为力. 通过对对偶问题的研究, 许多学者提出了一些次优对偶控制方法^[11-14], 其核心都在于如何更好的将控制与辨识相结合, 并在两者之间实现最佳平衡或者折衷. 研究表明, 要实现更好的控制, 仅获取原问题的一步或两步控制解有其局限性, 设计出的控制器有“近视”性缺点. 为了更好地对原系统进行控制优化和辨识, 关键在于如何充分利用原系统提供的信息, 也就是在利用新息对原系统参数进行估计的同时, 要把估计误差信息充分地运用到对未来控制的求解之中, 实现“滚动”优化, 以便获得满意控制性能. 毫无疑问, 对于本文建立的系统模型, Kalman 滤波是在滚动优化过程中对未知参数进行估计的有力工具.

对于式 (1) 和式 (2) 描述的控制问题, 为能使用 Kalman 滤波方法, 类似 CARMA 模型, 在 k 时刻定义新的状态向量 θ 和回归向量 Φ 分别为

$$\theta(k) = [a(k), b(k)]^T \quad (3)$$

$$\Phi(k) = [x(k), u(k)] \quad (4)$$

则系统模型 (1) 可等价变为

$$x(k+1) = \Phi(k)\theta(k) + w(k) \quad (5)$$

不确定参数 θ 与系统的各个环节相关的连接机理、元器件或者组织管理、统计监测部门密切相关, 即使系统中各元件的物理参数精确已知, 由于各环节之间的耦合, 这些精确的“元器件”参数确定出来的模型参数 θ 也不能描述系统的真实行为, 因此, 需要实时的辨识或者估计模型的等效参数. 为此, 建立参数动态模型

$$\theta(k+1) = \theta(k) + v(k) \quad (6)$$

其中, $v(k)$ 为与过程噪声 $w(k)$ 独立的高斯白噪声, 服从 $N(0, \Sigma_v)$ 分布. 模型 (6) 表明, 系统在运行过程中, 模型参数会在一个标称值附近波动, 波动的大小就是噪声 $v(k)$ 的方差. 一个极端情况就是参数 θ 为未知常数, 即后一时刻的 θ 值与前一时刻的值始终相同, 在模型 (6) 中仅需要取 $\Sigma_v = 0$ 即可. 总之, 这里未知参数 θ 的模型可以根据实际情况, 建立成许多有用的方式.

将模型 (5) 与模型 (6) 联立, 可以得到反映系统状态与参数演化的动态描述

$$\theta(k+1) = \theta(k) + v(k) \quad (7)$$

$$x(k+1) = \Phi(k)\theta(k) + w(k) \quad (8)$$

定义 $\hat{\theta}(k|k)$ 为 k 时刻用状态信息对未知参数的估计, 则通过卡尔曼滤波方程, 可以得到未知参数的条件均值和方差为

$$\hat{\theta}(k) = \hat{\theta}(k-1) + K(k)e(k) \quad (9)$$

$$K(k) = P(k|k-1)\Phi(k)[\Phi^T(k)P(k|k-1) \times \Phi(k) + \Sigma_w]^{-1} \quad (10)$$

$$P(k|k-1) = P(k-1) + \Sigma_v \quad (11)$$

$$P(k) = [I - K(k)\Phi^T(k)]P(k|k-1), \quad k = 0, 1, \dots, N-1 \quad (12)$$

其中, $e(k) = x(k) - \Phi(k)\hat{\theta}(k|k-1)$ 即为新息.

估计优劣可以直接由测量方程来确定, 即,

$$\begin{aligned} \Sigma_x(k+1) &= E\{[x(k+1) - \hat{x}(k+1)] \times \\ &\quad [x(k+1) - \hat{x}(k+1)]^T\} = \\ &= E\{[\Phi(k)\theta(k) + w(k) - \Phi(k)\hat{\theta}(k)] \times \\ &\quad [\Phi(k)\theta(k) + w(k) - \Phi(k)\hat{\theta}(k)]^T\} = \\ &= E\{[\Phi(k)\tilde{\theta}(k) + w(k)] \times \\ &\quad [\Phi(k)\tilde{\theta}(k) + w(k)]^T\} = \\ &= \Phi(k)P(k)\Phi^T(k) + \Sigma_w \end{aligned} \quad (13)$$

根据 Kalman 滤波中估计误差 $P(k)$ 的定义, 对其进行分块

$$P(k) = \begin{pmatrix} P_{aa}(k) & P_{ab}(k) \\ P_{ab}(k) & P_{bb}(k) \end{pmatrix} \quad (14)$$

将上述分块矩阵代入式 (13), 限制状态变量与输入变量为标量, 得到

$$\Sigma_x(k+1) = P_{bb}(k)u^2(k) + 2P_{ab}(k)u(k)x(k) + P_{aa}(k)x^2(k) + \Sigma_w \quad (15)$$

从式 (15) 可以看出, 当前控制与下一步参数估计的误差矩阵 $\Sigma_x(k+1)$ 存在直接关系, 可以选择控制 $u(k)$ 使 $\Sigma_x(k+1)$ 极小. 为此, 令 $\frac{d\Sigma_x(k+1)}{du(k)} = 0$, 可得最优估计控制

$$u_l^*(k) = -\frac{P_{ab}(k)}{P_{bb}(k)}x(k) \quad (16)$$

显然, 上述控制 $u_l^*(k)$ 能使未来一步估计方差最小, 因此, 这样的控制有学习性质. 然而, 本文欲解决的控制问题主要目的是让二次性能指标最小, 可以设想在最优控制律上以某种权系数 α 叠加 $u_l^*(k)$ 就能赋予控制器的学习性质.

对于模型 (1) 中的未知参数 $a(k)$ 和 $b(k)$, 利用上面的 Kalman 滤波可以得到它们的估计值 $\hat{a}(k|k)$ 和 $\hat{b}(k|k)$. 然而, 利用性能指标 (3) 计算系统在当前时刻的最优控制时, 由于性能指标具有二次和的可分性质, 可以根据动态规划从最后一个阶段 N 向当前阶段 k 反向后退, 这必然用到 k 时刻以后未知

参数的值, 为此, 作以下假定: 对于 $t \geq k$, $a(t) = \hat{a}(k|k)$ 和 $b(t) = \hat{b}(k|k)$, 即未来时刻系统模型中的未知参数都用当前时刻的估计值取代. 这样, 依据“确定性等价原理”, 按照一般的动态规划计算方法, 求解贝尔曼泛函方程, 可以得到对应的最优控制策略为

$$u_c^*(k) = -L(k)x(k) \quad (17)$$

$$L(k) = D^{-1}(k)\hat{b}^T S(k+1)\hat{a} \quad (18)$$

$$D(k) = \hat{b}^T S(k+1)\hat{b} + R \quad (19)$$

$$S(k) = \hat{a}^T S(k+1)\hat{a} + Q - L^T(k)D(k)L(k) \quad (20)$$

$$S(N) = Q_N \quad (21)$$

其中, $\hat{a}$ 和 $\hat{b}$ 表示 k 时刻对原系统参数 a 和 b 的估计, 可由 Kalman 滤波 (9) 来确定, 为书写方便, 在参数估计 $\hat{a}$ 和 $\hat{b}$ 中, 省略了时间 k , 这样就完整地实现了参数估计与控制目标的优化结合. 如上得到的最优控制序列是基于“确定性等价原理”和“分离性原理”, LQG 理论表明, 当模型中参数已知时, 分离性原理成立, 当模型参数未知时, 分离性原理不成立. 本文在分离性原理不成立的条件下, 人为地强迫其成立, 因此, 获得的控制为次优控制. 另外, 用动态规划反向后退所得到的最优控制序列为 $\{u^*(k), u^*(k+1), \dots, u^*(N-1)\}$. 我们用第 1 个 $u^*(k)$ 对系统实施控制, 获得新的状态信息 $x(k+1)$, 用新的状态信息通过 Kalman 滤波可以估计出新的参数 $\hat{a}(k+1|k+1)$, $\hat{b}(k+1|k+1)$, 再将未来的参数固定在当前估计值上, 通过动态规划求出新的控制序列, 这是一个“滚动”优化过程.

为了使实际施加于系统的控制既有控制功能又有学习性质, 采用控制

$$u^*(k) = u_c^*(k) + \alpha u_l^*(k) \quad (22)$$

其中, $\alpha \in (0, +\infty)$ 为学习因子, 它的大小反映了控制目标与学习目标间的权衡, α 越大, 对学习要求越高, 反之亦然. 由式 (22) 给出的控制序列 $u^*(k)_{k=0}^{N-1}$, 由于包含第二部分学习项, 因此, 称该控制为学习控制.

学习控制是一个“滚动优化”过程, 在用动态的求解多阶段贝尔曼泛函方程时, 把滤波得到的最新参数运用到控制量的求解当中, 尽可能地利用了探测过程中所得到的最新的信息, 因而具有较好的控制效果.

学习控制具有如下两个性质.

性质 1. 对于本文描述的控制问题, 控制器与滤波器之间的分离性不再成立.

本文描述的控制问题由二次型性能指标 (2)、状态方程 (7) 和测量方程 (8) 组成. 由于模型中含有未

知参数, 因此, 在控制器设计时需对它们用 Kalman 滤波 (9)~(12) 进行估计, 估计精度可用估计方差矩阵 $P(k)$ 的迹或者行列式值或者其他与之相关的指标来度量. 将式 (10) 代入式 (12), 并对式 (12) 的两边取行列式, 则有

$$\begin{aligned}
 |P(k)| &= |[I - K(k)\Phi(k)]P(k|k-1)| = \\
 & \left| I - \frac{P(k|k-1)\Phi^T(k)\Phi(k)}{\Phi(k)P(k|k-1)\Phi^T(k) + \Sigma_w} \right| \times \\
 & |P(k|k-1)| = \\
 & \left| 1 - \frac{\Phi(k)P(k|k-1)\Phi^T(k)}{\Phi(k)P(k|k-1)\Phi^T(k) + \Sigma_w} \right| \times \\
 & |P(k|k-1)| = \\
 & \frac{|P(k|k-1)|\Sigma_w}{\Phi(k)P(k|k-1)\Phi^T(k) + \Sigma_w} = \\
 & \frac{|P(k|k-1)|\Sigma_w}{P_{aa}x^2(k) + 2P_{ab}x(k)u(k) + P_{bb}u^2(k) + \Sigma_w} \quad (23)
 \end{aligned}$$

其中, 第 3 个等号用到了行列式的性质 $|I - AB| = |I - BA|$ (A, B 为同阶方阵), 最后一个等号用到了 $P(k|k-1)$, 与式 (14) 完全相同的分块方式, 为书写方便省去了时间变量.

由式 (23) 可以看出, 未知参数估计的方差矩阵 $P(k)$ 中明显地含有控制 $u(k)$, 因此, 控制 $u(k)$ 不仅对目标有控制作用, 对于未知参数的估计精度也有影响, 表明控制器与滤波器之间不具有分离性质.

显然, 对估计来说, $u(k)$ 越大, $|P(k)|$ 越小, 估计精度越高; 而对二次型控制目标 (2) 来说, $u(k)$ 越大, 目标函数越大, 这与目标函数的极小意义相冲突. 另外, 在动态规划的使用中, 从最后一项开始, 倒退递推至首项, 控制目标与估计目标中皆有 $u(k)$, 且在 $|P(k)|$ 中 $u(k)$ 以二次形式出现于分母, 动态规划每一阶段的 cost-to-go 无法求出解析解, 即使开始阶段有解析解, 在动态规划的倒退过程中每一阶段的 cost-to-go 都有非线性的 $|P(k)|$ 的复合加上期望算子的积分的复合, 使得 cost-to-go 即使是求数值解, 也显得极为困难. 本文的学习算法避免了这一困难.

性质 2. 在学习控制中学习因子 $\alpha \in (0, f]$, f 是一个常数.

学习程度的“好”、“坏”与 Kalman 滤波的方差矩阵有关, 也与“新息” $e(k)$ 有关, 事实上, 可以导出新息 $e(k)$ 平方的均值为

$$\begin{aligned}
 E\{e^2(k)\} &= E\{[x(k+1) - \Phi(k)\hat{\theta}(k|k-1)]^2\} = \\
 E\{[\Phi(k)\theta(k) + w(k) - \Phi(k)\hat{\theta}(k|k-1)]^2\} &=
 \end{aligned}$$

$$\begin{aligned}
 E\{[\Phi(k)\tilde{\theta}(k|k-1) + w(k)]^2\} &= \\
 \Phi(k)P(k|k-1)\Phi^T(k) + \Sigma_w \quad (24)
 \end{aligned}$$

根据式 (23), 可得

$$|P(k)| = \frac{|P(k|k-1)|\Sigma_w}{E\{e^2(k)\}} \quad (25)$$

因此, 学习控制算法在 N 个阶段的控制问题的每一阶段附加了学习控制, 该策略可以近似地模型化为如下优化问题:

$$\min_{u(k)} J_k = \overbrace{Qx^2(k) + Ru^2(k)}^{\text{第 1 项}} - \alpha \overbrace{E\{e^2(k)\}}^{\text{第 2 项}} \quad (26)$$

其中, 第 1 项为控制目标, 第 2 项为学习目标.

将式 (24) 代入式 (26), 并利用矩阵的分块形式 (14), 可得

$$\begin{aligned}
 J_k &= Qx^2(k) + Ru^2(k) - \alpha[P_{aa}x^2(k) + \\
 & 2P_{ab}x(k)u(k) + P_{bb}u^2(k) + \Sigma_w] = \\
 & [R - \alpha P_{bb}]u^2(k) - 2\alpha P_{ab}x(k)u(k) + \\
 & [Q - \alpha P_{aa}]x^2(k) - \alpha\Sigma_w
 \end{aligned}$$

显然, 要获得上式关于 $u(k)$ 的最小值, 二次项前面的系数应该大于零, 即,

$$R - \alpha P_{bb} \geq 0 \quad (27)$$

当控制器在每个阶段获得信息后都要对未知参数进行估计, 累计对系统实施 $N-1$ 步控制, 因此, 式 (27) 共有 $N-1$ 个 P_{bb} , 在有限个 $P_{bb}(k)$ 中, 选最大值, 即可得到使全体不等式成立的 α 的上界 f .

3 滚动学习控制算法

图 1 为滚动学习控制算法框图.

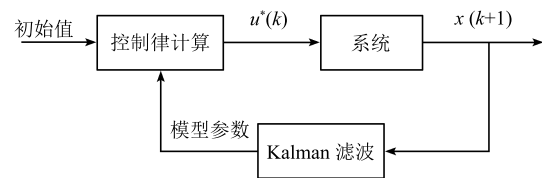


图 1 滚动学习控制算法原理

Fig. 1 Control algorithm principle of rolling learning

用滚动优化求出的控制序列是一种次优序列. 原因为: 1) 对于具有未知参数的 LQG 问题, 强迫分离性成立; 2) 在动态规划实施过程中未来参数固定在当前的估计值上; 3) 计算最优控制时虽然没有忽略参数估计误差, 但学习控制是人为的附加于 LQG 控制.

滚动优化算法工作过程为: 在初始时刻 0, 固定模型中的等效参数 $a(\theta)$ 和 $b(\theta)$ 为其初值 $\hat{a}(0)$ 和 $\hat{b}(0)$, 该参数已知的动态系统与二次性能指标构成一个 LQG 问题, 用动态规划求出最优控制序列 $u^*(0), u^*(1), \dots, u^*(N-1)$, 将第一个控制 $u^*(0)$ 施加于系统, 得到系统的当前状态 $x(1)$, 用该信息进行 Kalman 滤波, 得到新的参数值 $\hat{\theta}(1)$, 即 $\hat{a}(1)$ 与 $\hat{b}(1)$, 再将模型中的等效参数 $a(\theta)$ 和 $b(\theta)$ 固定于 $\hat{a}(1)$ 与 $\hat{b}(1)$. 重复上述过程, 可以对系统实施滚动控制. 算法具体实施步骤如下:

初始化. 置 $k = 0$.

步骤 1. 用 Kalman 滤波式 (9) 和式 (10) 对未知参数 θ 进行估计.

步骤 2. 用式 (17) ~ (21) 求出使性能指标 J 最小的控制 $\{u_c^*(i)\}_{i=k}^{N-1}$.

步骤 3. 用式 (16) 求出学习控制 $u_l^*(k)$.

步骤 4. 用式 (22) 求出施加于实际系统的控制 $u^*(k)$.

步骤 5. 判断 $k = N - 1$ 是否成立, 如果成立, 结束; 否则, 返回步骤 1.

4 仿真分析

考虑如下不确定离散随机系统

$$x(k+1) = a(\theta)x(k) + b(\theta)u(k) + w(k) \quad (28)$$

假定状态的初始条件为 $x(0) = 0.5$, 不确定性参数 a 和 b 为未知常数, 它们的标称值为 $\bar{a} = 0.8$, $\bar{b} = 0.5$, $w(k) \sim N(0, 0.4)$.

系统在运行的过程中, 模型参数在标称值附近波动, 波动的方差 $R_{na} = 0.06$, $R_{nb} = 0.04$, 两者独立变化, 互不相关, 则模型 (7) 中的噪声 $v(k)$ 服从的高斯分布为 $N(0, \Sigma_v)$, $\Sigma_v = \text{diag}\{R_{na}, R_{nb}\}$.

性能指标

$$J = E \left\{ Q_N x^2(N) + \sum_{k=0}^{N-1} [Q x^2(k) + R u^2(k)] \right\} \quad (29)$$

其中, $N = 50$, $Q_N = Q = R = 1$.

如果在系统模型中取参数 a 和 b 为它们的标称值, 即控制过程中完全忽略参数的波动, 由于模型参数已知, 用动态规划可以求出最优控制

$$\bar{u}(k) = -\bar{L}(k)x(k) \quad (30)$$

$$\bar{L}(k) = \bar{D}^{-1}(k)\bar{b}^T\bar{S}(k+1)\bar{a} \quad (31)$$

$$\bar{D}(k) = \bar{b}^T\bar{S}(k+1)\bar{b} + R \quad (32)$$

$$\bar{S}(k) = \bar{a}^T\bar{S}(k+1)\bar{a} + Q - \bar{L}^T(k)\bar{D}(k)\bar{L}(k) \quad (33)$$

$$\bar{S}(N) = Q_N \quad (34)$$

上面导出的控制序列 $\bar{u}(k)_{k=0}^{N-1}$, 由于将本来有波动的模型参数固定于标称值, 因此称其为标称控制, 标称控制是依据已知参数确定的控制律, 它的工作过程不需要滚动. 在学习控制 (22) 中, 如果令学习因子 $\alpha = 0$, 则对应的控制序列 $\{u_c^*(k)\}_{k=0}^{N-1}$ 称为非学习控制, 在滚动学习控制中 $\alpha = 0.58$. 对于该系统用三种控制策略进行控制, 在 100 次 Monte Carlo 仿真后, 性能指标的值如表 1 所示.

表 1 不同控制下的 Monte Carlo 仿真性能指标比较

滚动学习控制	非学习控制	标称控制
150.8072	208.7341	220.9545

表 1 表明, 非学习控制与标称控制性能指标相差不多, 也就是说系统模型中如果存在参数随机波动, 仅采用 Kalman 滤波对参数估计, 控制效果并不显著, 而滚动学习控制的性能指标明显优于其他两个, 显示了学习的重要性.

模型中参数由于受到噪声影响而随机波动, 因此, 在每一时刻 k 参数 a 和 b 都是变化的, 不太容易看出控制器的学习效果. 下面假定模型参数 a 和 b 未知, 但为常数. 图 2 和图 3 给出了控制时段内 Kalman 滤波对参数的估计过程, 图中的实线与虚线分别为参数的估计值和真值, 可以看出估计值在不断的逼近真值.

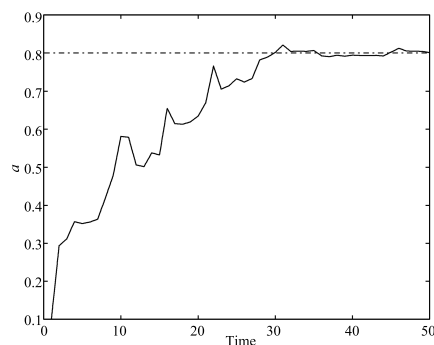


图 2 a 的估计过程

Fig. 2 The estimation process of a

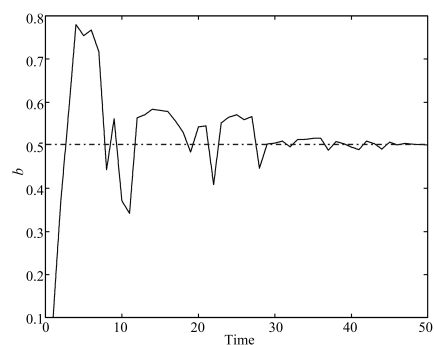


图 3 b 的估计过程

Fig. 3 The estimation process of b

5 结论

本文针对含有不确定参数的随机系统提出了一种滚动优化学习控制方法. 控制器既有控制功能又有学习功能, 但两者是冲突的, 它们之间的权重可由学习因子来实现, 然而, 本文中的学习因子是固定的, 能不能在控制的开始一段时间内取大点, 加强学习力度, 当控制器学习出真值后, 学习因子自动消失, 即最优学习因子的确定是未来需要解决的问题; 本文采用滚动优化, 计算量小, 为控制的实施带来了方便, 但导致了次优策略, 能否得到最优控制, 还有待研究; 本文的研究结果表明, 对于具有未知参数的最优控制问题, 强化控制器的学习必不可少, 虽然控制能量的一部分被分配至学习, 降低了控制性能, 但是, 这部分代价不可避免.

References

- Huang M, Wang X, Wang Z L. Multiple model adaptive control for a class of linear-bounded nonlinear systems. *IEEE Transactions on Automatic Control*, 2015, **60**(1): 271–276
- Ma Y J, Jiang B, Tao G, Cheng Y H. Uncertainty decomposition-based fault-tolerant adaptive control of flexible spacecraft. *IEEE Transactions on Aerospace and Electronic Systems*, 2015, **51**(2): 1053–1068
- De Villiers J P, Laskey K, Jusselme A L, Blasch E, Pavlin G, Costa P. Uncertainty representation, quantification and evaluation for data and information fusion. In: *Proceedings of the 18th International Conference on Information Fusion*. Washington, DC, USA: IEEE, 2015. 50–57
- Bellman R, Kalaba R. *Dynamic Programming and Modern Control Theory*. New York: Academic Press, 1965.
- Apkarian P, Dao M N, Noll D. Parametric robust structured control design. *IEEE Transactions on Automatic Control*, 2015, **60**(7): 1857–1869
- Wen J T, Kreutz-Delgado K, Bayard D S. Lyapunov function-based control laws for revolute robot arms: tracking control, robustness, and adaptive control. *IEEE Transactions on Automatic Control*, 1992, **37**(2): 231–237
- Cao S P, Qian F C, Wang X M. Exact optimal solution for a class of dual control problems. *International Journal of Systems Science*, 2016, **47**(9): 2078–2087
- Qian F C, Huang J R, Liu D, Hu S L. Adaptive dual control of discrete-time LQG problems with unknown-but-bounded parameter. *Asian Journal of Control*, 2015, **17**(3): 942–951
- Bar-Shalom Y, Wall K D. Dual adaptive control and uncertainty effects in macroeconomic systems optimization. *Automatica*, 1980, **16**(2): 147–156
- Filatov N M. *Adaptive Dual Control Theory and Applications*. Berlin Heidelberg: Springer, 2004.
- Li D, Qian F C, Fu P L. Variance minimization approach for a class of dual control problems. In: *Proceedings of the 2002 American Control Conference*. Anchorage, AK, USA: IEEE, 2002. 3759–3764
- Li D, Fu P L, Qian F C. Optimal nominal dual control for discrete-time LQG problem with unknown parameters. In: *Proceedings of the 2003 SICE Annual Conference*. Fukui, Japan: IEEE, 2003. 476–479
- Qian F C, Gao J J, Li D. Complete statistical characterization of discrete-time LQG and cumulant control. *IEEE Transactions on Automatic Control*, 2012, **57**(8): 2110–2115
- Milito R, Padilla C, Padilla R, Cadorin D. An innovations approach to dual control. *IEEE Transactions on Automatic Control*, 1982, **27**(1): 132–137



尚婷 西安理工大学自动化与信息工程学院博士研究生. 主要研究方向为最优控制, 系统辨识, 故障诊断.

E-mail: tshang0722@126.com

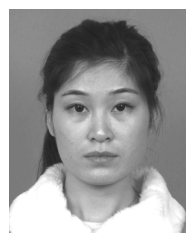
(SHANG Ting Ph.D. candidate at the School of Automation and Information Engineering, Xi'an University of Technology. Her research interest covers optimal control, systems identification, and fault diagnosis.)



钱富才 西安理工大学自动化与信息工程学院教授. 主要研究方向为随机控制, 系统辨识, 非线性控制, 最优控制, 故障诊断和全球定位系统. 本文通信作者.

E-mail: qianfc@xaut.edu.cn

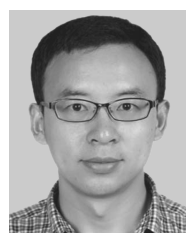
(QIAN Fu-Cai Professor at the School of Automation and Information Engineering, Xi'an University of Technology. His research interest covers stochastic control, systems identification, nonlinear control, optimal control, fault diagnosis and GPS system. Corresponding author of this paper.)



张晓艳 西安理工大学自动化与信息工程学院博士研究生. 主要研究方向为故障诊断, 最优控制, 系统辨识.

E-mail: xyzhang.2016@163.com

(ZHANG Xiao-Yan Ph.D. candidate at the School of Automation and Information Engineering, Xi'an University of Technology. Her research interest covers fault diagnosis, optimal control, and systems identification.)



谢国 西安理工大学副教授. 2013 年获得日本大学工学博士学位. 主要研究方向为轨道交通系统的安全性和可靠性, 最优控制, 随机控制.

E-mail: guoxie@xaut.edu.cn

(XIE Guo Associate professor at Xi'an University of Technology. He received his Ph.D. degree from Nihon University, Tokyo, Japan, in 2013. His research interest covers safety and reliability of railway system, optimal control, and stochastic control.)