文章编号: 1001-0920(2004)07-0817-03

基于两级参数算法的对偶控制

李云霞, 康 波
(电子科技大学 自动化学院, 四川 成都 610054)

摘 要: 针对具有未知参数的随机系统, 从随机次优的角度出发, 将原不可解的动态规划问题转化为优化一个效用函数, 该效用函数考虑了输出调节要求和参数学习要求及两者之间的折衷, 充分利用了对偶控制较常规自适应控制的优越性. 提出利用两级参数算法来最小化效用函数, 从而获得控制信号. 仿真结果表明, 这种控制器具有良好的对偶性质, 能得到较好的学习和控制效果.

关键词: 随机系统; 最小方差控制; 对偶控制
中图分类号: TP275 **文献标识码**: A

Dual control based on two-level parametric algorithm

LI Yun-xia, KANG Bo
(School of Automation, University of Electronic Science and Technology of China, Chengdu 610054, China. Correspondent: LI Yun-xia, E-mail: diwana@163.com)

Abstract: A dual control method is proposed for stochastic systems with unknown parameters. The utility function to be minimized indicates the desired degree of compromise between the regulate effect of the output and the learning effect of the unknown parameters, taking advantage of the superior performance of dual control compared to conventional adaptive control. The two-level parametric algorithm is introduced to derive the control law. Simulation results show the dual property of the controller that can achieve good estimation results as well as good control performance.

Key words: stochastic system; minimum variance control; dual control

1 引 言

常规自适应控制用于解决具有参数不确定性随机控制问题的基本依据是确定性等价原理, 即首先对系统的未知参数进行估计, 然后在设计控制器时将参数的估计值作为真值. 这种确定性等价原理的缺点是在设计控制器时没有考虑到参数辨识的不精确性. 因此, 将控制目标与辨识目标相结合作为一个混合问题考虑, 是对确定性等价原理的重大突破, 这就是对偶控制的本质. 对偶控制思想是1960年由Feldbaum[1]提出的. 随机最优控制策略的对偶性质体现在: 一方面, 控制信号对系统输出具有调节作用; 另一方面, 控制信号还要对未知参数进行主动学习. 这两种作用在控制律的实现中是矛盾的, 对偶控制策略旨在获取调节和学习的最佳折衷.

对偶控制中控制与学习之间存在偶合关系, 导致动态规划的回归方程无法处理而得不到解析解, 因此既保持对偶性质又能加以实现的次优随机控制策略成为对偶控制的研究方向之一. 1974年, Alster[2]等人将控制目标由极小化N步超前输出方差简化为极小化一步超前输出方差, 导出了一步超前控制器; 1982年, Milito[3]等人在控制指标泛函的基础上, 以一个学习因子引入新息序列的方差, 从而实现了控制作用和估计作用的折衷; 1994年, Maitelli[4]等人提出了两步超前控制器, 同时极小化

收稿日期: 2003-07-14; **修回日期**: 2003-10-09.
作者简介: 李云霞(1977—), 女, 四川武胜人, 博士生, 从事随机控制、最优控制的研究; 康波(1968—), 男, 重庆人, 副教授, 从事混沌研究.

一步超前和两步超前输出方差; 1999 年,Maitelli 又提出了 M 步超前控制器[5], 虽然控制效果随 M 的增加稍有改善, 但却增加了巨大的计算量; 陈翰馥、郭雷教授在参数辨识与稳定性方面做出了杰出的贡献[6]; 梁军教授就对偶控制理论作了详尽的综述[7]. 本文从随机次优的角度出发研究了对偶控制问题, 将原来不可解的动态规划问题转化为优化一个可同时表达输出调节要求和参数学习要求的效用函数. 为求解这个优化问题, 本文采用了两级参数算法, 该算法收敛速度快, 减少了求取控制信号的计算量.

2 问题的提出

考虑如下具有未知参数的随机系统:

$$\begin{aligned} y(k) = {} & a_1(k)y(k-1) + a_2(k)y(k-2) + \cdots + \\ & a_m(k)y(k-m) + b_1(k)u(k-1) + \cdots + \\ & b_n(k)u(k-n) + e(k), \\ & k = 0, 1, \cdots, N-1. \end{aligned} \tag{1}$$

其中: $a_1, a_2, \cdots, a_m, b_1, \cdots, b_n$ 是未知的参数; $u(k)$ 和 $y(k)$ 分别是控制序列和输出序列; $e(k)$ 是均值为 0、方差为 σ^2 的高斯白噪声, 并与未知参数相互独立. 定义如下向量:

$$\begin{aligned} & x(k) = \\ & [a_1(k), a_2(k), \cdots, a_m(k), b_1(k), b_2(k), \cdots, b_n(k)]^{\mathrm{T}}, \\ & \Phi(k) = [y(k-1), y(k-2), \cdots, y(k-m), \\ & \quad u(k-1), u(k-2), \cdots, u(k-n)]^{\mathrm{T}}, \end{aligned} \tag{2}$$

那么式(1) 可变换为紧凑格式的形式, 即

$$y(k) = \Phi^{\mathrm{T}}(k)x(k) + e(k). \tag{3}$$

其初值分布: $x(0) \sim N(\hat{x}(0), P(0))$, 假定 $\hat{x}(0)$ 和 $P(0)$ 均已知或已给出. 初始条件 $I(0) = \{y(0), \cdots, y(-m+1), u(-1), \cdots, u(-n+1)\}$ 已知. 控制随机系统的输出 $y(k)$ 使之跟踪期望目标 $y_r(k)$, 利用 $y(k)$ 与 $y_r(k)$ 的最小方差函数的条件均值作为代价函数

$$J = E\left\{\sum_{k=0}^{N-1}[y(k+1) - y_r(k+1)]^2 \,\Big|\, I(k)\right\} \tag{4}$$

其中: $I(k) = \{y(k), \cdots, y(1), u(k-1), \cdots, u(0), I(0)\}$ 为当前信息序列, 也是在时刻 k 所能得到的所有信息. 设计控制器的目的是选择合适的输入 $u(k) = f_k(I_k)$, 使式(4) 取得最小值. $x(k+1)$ 在条件 $I(k)$ 下的均值 $\hat{x}(k+1)$ 及估计方差 $P(k+1)$ 可由 Kalman 滤波公式给出.

3 基于两级参数算法的对偶控制

由于目标函数式(4) 无法求得解析解, 本文从随机次优的角度出发, 提出了基于两级参数算法的对偶控制(DCTP) 的效用函数, 即在时刻 k 对系统(1) 施加控制量 $u(k)$, $k = 0, 1, \cdots, N-1$, 使得

$$J(k) = J_1(k) - \lambda\frac{J_2(k)}{J_3(k)} \tag{5}$$

的损失函数值最小. 式中: $J_1(k)$ 为一步超前输出方差中的可控部分, 表达了对输出的调节要求; $J_2(k)$ 和 $J_3(k)$ 为新息中的可控部分, $\dfrac{J_2(k)}{J_3(k)}$ 表达了对参数的学习要求; λ 为学习因子, 表达了对控制要求和学习要求的折衷, $0 \le \lambda \le 1$; $J_1(k)$, $J_2(k)$ 和 $J_3(k)$ 的表达式分别为:

$$\begin{aligned} & J_1(k) = \\ & E\{[y(k+1) - y_r(k+1)]^2 \,|\, I(k)\} - \sigma^2 = \\ & [\Phi^{\mathrm{T}}(k+1)\hat{x}(k+1) - y_r(k+1)]^2 + \\ & \Phi^{\mathrm{T}}(k+1)P(k+1)\Phi(k+1), \end{aligned} \tag{6}$$

$$\begin{aligned} & J_2(k) = E\{v^2(k+1) \,|\, I(k)\} - \sigma^2 = \\ & \Phi^{\mathrm{T}}(k+1)P(k+1)\Phi(k+1), \end{aligned} \tag{7}$$

$$\begin{aligned} & J_3(k) = E\{v^2(k+2) \,|\, I(k)\} - \sigma^2 = \\ & \Phi^{\mathrm{T}}(k+2)P(k+2)\Phi(k+2). \end{aligned} \tag{8}$$

根据 $P(k+1)$, $P(k+2)$, $\Phi(k+1)$ 和 $\Phi(k+2)$ 的定义, 不难得出: $J_1(k)$, $J_2(k)$ 和 $J_3(k)$ 都是关于 $u(k)$ 和 $u(k+1)$ 的二次函数, 且 $J_3(k)$ 始终大于 0, 直接优化效用函数(5) 比较困难. 本文采用两级参数算法解决该问题, 最小化式(5) 的数学表达式

$$\min_{u(k), u(k+1)}\left\{J_1(k) - \lambda\frac{J_2(k)}{J_3(k)}\right\} \tag{9}$$

针对问题(9), 构造参数辅助问题

$$J(\alpha): \min_{u(k), u(k+1)}\{J_1(k)J_3(k) - \lambda J_2(k) - \alpha J_3(k)\}. \tag{10}$$

定理 1 $u^*(k)$, $u^*(k+1)$ 为问题(9) 的解, 当且仅当 $u^*(k)$, $u^*(k+1)$ 为问题的 $J(\alpha^*)$ 解, 而 α^* 满足

$$\begin{aligned} & \alpha^* = J_1^*(k) - \lambda\frac{J_2^*(k)}{J_3^*(k)}, \\ & J_i^*(k) = J_i(k)\,\big|_{u^*(k), u^*(k+1)}, \\ & \quad i = 1, 2, 3. \end{aligned} \tag{11}$$

证明 1) 必要性. 利用反证法, 设 $u^*(k)$, $u^*(k+1)$ 为问题(9) 的解, 但不是问题 $J(\alpha^*)$ 的解. 那么, 存在问题 $J(\alpha^*)$ 的一个可行解 $\hat{u}(k)$, $\hat{u}(k+1)$, 使得

$$\begin{aligned} & \hat{J}_1(k)\hat{J}_3(k) - \lambda\hat{J}_2(k) - \alpha^*\hat{J}_3(k) < \\ & J_1^*(k)J_3^*(k) - \lambda J_2^*(k) - \alpha^* J_3^*(k). \end{aligned} \tag{12}$$

其中 $\hat{J}_i(k) = J_i(k)\,\big|_{\hat{u}(k), \hat{u}(k+1)}$, $i = 1, 2, 3$. 将式(11)

代入式(12),得

$$\hat{J}_1(k)\hat{J}_3(k) - \lambda\hat{J}_2(k) - \alpha^*\hat{J}_3(k) < 0. \qquad (13)$$

因为$J_3(k)$ 始终大于 0,将不等式(13) 两边同除以 $J_3(k)$,则有

$$\alpha^* > \hat{J}_1(k) - \lambda\frac{\hat{J}_2(k)}{\hat{J}_3(k)}. \qquad (14)$$

将式(11) 代入式(14),有

$$J_1^*(k) - \lambda\frac{J_2^*(k)}{J_3^*(k)} > \hat{J}_1(k) - \lambda\frac{\hat{J}_2(k)}{\hat{J}_3(k)},$$

这与 $u^*(k), u^*(k+1)$ 为问题(9) 的解的最优性相矛盾,因此定理得证. □

2) 充分性: 仍利用反证法,方法同上,略.

该定理的意义在于,在参数满足条件(11) 时,问题 (9) 与问题(10) 的解是一致的. 因此,若求解问题(9),只需求出对应最优参数 α 的问题(10) 的解. 两级参数算法思路为: 上级首先给定参数 α 的值,然后利用梯度法对问题(10) 进行求解,其结果返回上级,上级更新参数 α,再送给下级,如此循环,直到参数 α 满足最优性条件.

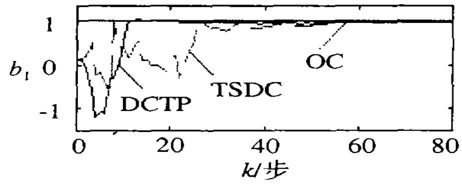
4　仿真分析

考虑式(1) 描述的系统,其中: $b_1(k) = 1, b_2(k) = 0.9, a(k) = -1, e(k) \sim N(0,1)$. 初始条件设为: $I(0) = [u(-1), y(0)] = [-0.5, -0.5], \hat{x}(1) = [0.1, 0.1, 0.1]^T, P(1) = 10I_3$. 以最优控制器(OC,即参数已知条件下的控制器) 作基准,比较本文提出的基于两级参数算法的对偶控制(DCTP) 和文献[4] 提出的两步次优对偶控制(TSDC) 的性能. 代价函数均定义为

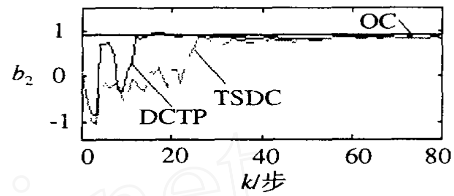
$$J = \frac{1}{M}\left\{\sum_{j=1}^{M}\frac{1}{N}\sum_{k=0}^{N-1}[y(k+1) - y_r(k+1)]^2 \middle| I(k)\right\}$$

运算步数$N = 80$,Monte Carlo 运行次数$M = 100$.

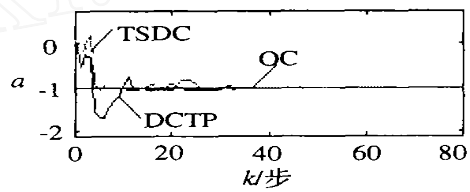
图 1 为参数学习曲线. 该图显示了参数学习结果,可以看出,DCTP 和 TSDC 分别经过大约 15 步和 30 步运算后,其参数的估计值逐渐逼近真实值,表明了DCTP 具有良好的学习性能. 由图 2 可以看出,在大约 10 步前,参数估计值不确定的情况下,对偶控制器运用了谨慎控制,控制信号幅值比较小; 10 步后,参数估计的不确定性减小了,对偶控制器与最优控制器的控制曲线基本吻合. 由图 3 的代价函数的Monte Carlo 曲线反映出,DCTP 与最优控制器的代价函数曲线更接近,表明在均值意义下,DCTP 的控制性能较 TSDC 好. 另外,为了比较两种算法的计算量,选择其运行时间作为比较对象. 仿真结果表明,DCTP和TSDC的总运行时间(100次) 分别为 20 s 和 25 s,说明前者的计算量较小.



(a)　参数 b_1 的学习曲线



(b)　参数 b_2 的学习曲线



(c)　参数 a 学习曲线

图 1　参数学习曲线

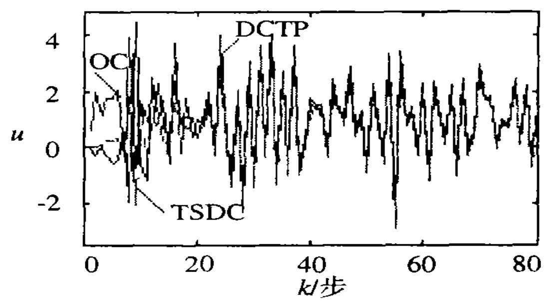


图 2　控制信号曲线

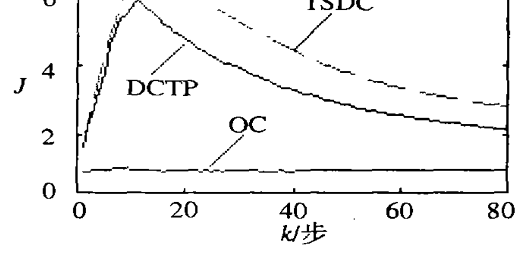


图 3　性能指标曲线

(下转第 823 页)

$$\max_{\{\pi_t\}_{0\leq t\leq T}} E[E(\ln(X_T)|F_0)].$$

最后, 由 μ 的正态性, 定理得证 □

另一方面, 若投资者只获得部分信息, 则由定理 1, 他得到的最大期望效用为

$$J_P = \ln(X_0) + \left[r + \frac{\gamma_0 + (m_0 - r)^2}{2\sigma^2}\right]T - \frac{1}{2}\ln\left(\frac{\sigma^2 + \gamma_0 T}{\sigma^2}\right). \tag{27}$$

比较式(26) 与(27) 得到如下信息价值测算公式:

定理 2(内部交易者获得的) 信息价值为

$$V \equiv J_F - J_P = \ln\left(1 + \frac{\gamma_0}{\sigma^2}T\right). \tag{28}$$

6　结　　论

1) 信息价值公式(28) 的计算很简单, 易于投资者操作. 例如, 由观测到的风险资产历史价格序列, 借助随机过程统计学[6] 及金融计量学[8] 理论, 导出参数 γ_0, σ的估计值; 然后根据目标期 T 的大小, 由式(28) 计算出信息价值V. 2) 信息价值只与 γ_0, T 和 σ有关, 无风险资产平均收益率 μ 的不确定性 γ_0 越大, 信息价值越大; 目标期T 越长, 信息价值越大; 无风险资产价格波动性 σ越大, 信息价值越小. 所有这些, 与直观是相吻合的, 但直观无法给出它们之间的精确数量关系. 3) 文中得到的信息价值式(28) 对投资者决策具有参考意义, 如: 是否值得去寻求内部信息决定于信息价值与信息成本的差额; 此外, 结合该公式与基于消费的资产定价模型(CCAPM), 有可能得到更深刻的资产定价理论. 4) 若投资者的效用函数不是对数型, 或风险资产的变动不满足式(1) 和(2), 则式(28) 不成立, 相关结果的推算是一个有待解决的问题, 根据本文分析, 不可能得到解析解.

参考文献(References):

[1] Merton R C. *Continuous-Time Finance*[M]. Cambridge: Blackwell, 1990.

[2] Lakner P. Utility maximization with partial information[J]. *Stochastic Process Appl*, 1995, 56: 247-273.

[3] Lakner P. Optimal trading strategy for an investor: The case of partial information[J]. *Stochastic Process Appl*, 1998, 76: 77-97.

[4] 杨昭军, 李致中, 邹捷中. 部分信息下的最优投资消费策略显式解[J]. 应用概率统计, 2001, 16(4): 390-398. (Yang Z J, Li Z Z, Zou J Z. Explicit Sloution for the optimal strategy for the investment-consumption with partial information[J]. *Chinese J of Applied Probability and Statistics*, 2001, 16(4): 390-398.)

[5] Yang Z J, Ma C Q. Optimal trading strategy with partial information: The simplified and generalized models[J]. *Int J of Heoretical & Applied Finance*, 2001, 4(5): 759-772.

[6] Liptser R S, Shiryayev A N. *Statistics of Random Process*[M]. New York: Springer-Verlag, 1977.

[7] Föllmer H, Schied A. *Stochastic Finance: An Introduction in Discrete Time*[M]. Berlin: Walter de Gruyter, 2002. 60-73.

[8] Campbell J, Lo A, Mackinlay A C. *The Econometrics of Financial Markets*[M]. New Jersey: Princeton University Press, 1997.

(上接第819 页)

5　结　　论

本文对具有未知恒定参数的随机系统的学习与控制问题进行了研究, 提出了基于两级参数算法的对偶控制方法. 原问题采用目标函数(4) 是不可解的, 而采用效用函数(5) 代替原目标函数, 再利用两级参数算法求解控制信号, 问题则变得非常容易. 尽管所得的解仍是次优的, 但它具有对偶特性, 对未知参数的学习能力和对输出的调节作用较强. 与 TSDC 的比较结果表明, 本文控制策略可得到更好的参数辨识结果与控制性能. 本文提出的对偶控制方法只适用于未知参数恒定的情况, 对于未知参数变化的随机控制问题是继续研究的方向.

参考文献(References):

[1] Feldbaum A A. Dual control theory I-IV[J]. *Automat Remote Contr*, 1960, 21(4): 874-1033.

[2] Alster J, Belanger P R. A technical for the dual adaptive control[J]. *Automatica*, 1974, 10(3): 627-634.

[3] Milito R, Padilla C S, Padilla R A. An innovations approach to dual control[J]. *IEEE Trans on Automat Control*, 1982, 27(1): 132-137.

[4] Maitelli A L, Yoneyama T. Two-stage suboptimal dual controller for systems with stochastic parameters using optimal predictors[J]. *IEE Proc D Control Theory Application*, 1994, 141(4): 253-260.

[5] Maitelli A L, Yoneyama T. A multistage suboptimal dual controller using optimal predictors[J]. *IEEE Trans on Automat Control*, 1999, 44(5): 1002-1008.

[6] Chen H F, Guo L. A robust stochastic adaptive controller[J]. *IEEE Trans on Automat Contr*, 1988, 33(11): 1035-1043.

[7] Liang J. Dual adaptive control[J]. *Control Theory and Applications*, 1997, 14(3): 297-305.