文章编号: 1000−8152(2010)01−0099−04

具有条件马尔科夫结构的离散随机系统最优控制

方洋旺[1], 王洪强[1,2], 伍友利[1]

(1. 空军工程大学 工程学院, 陕西 西安 710038; 2. 95997部队, 北京 100076)

摘要: 基于Bellman随机非线性动态规划法, 提出了具有条件马尔科夫跳变结构的离散随机系统的最优控制方法, 应用随机变结构系统的性质对最优控制算法进行了简化处理, 并将后验概率密度函数用条件高斯函数来逼近, 针对一类具有条件马尔科夫跳变结构的线性离散随机系统, 给出了其逼近最优控制算法.

关键词: 随机系统; 最优控制; 条件马尔科夫结构; Bellman随机动态规划法

中图分类号: TP271.74 **文献标识码**: A

Optimal control for discrete stochastic system with conditional Markov structure

FANG Yang-wang[1], WANG Hong-qiang[1,2], WU You-li[1]

(1. Engineering Institute, Air Force Engineering University, Xi'an Shaanxi 710038, China; 2. 95997 Army, Beijing 100076, China)

Abstract: Based on Bellman stochastic nonlinear dynamic programming, we develop an optimal control algorithm for the discrete stochastic system with conditional Markov structure. By making use of characters of stochastic variable structure system, we simplify the optimal control algorithm, and approximate the posterior probability density function by a Gaussian function. The optimal control algorithm for a class of linear discrete stochastic systems with conditional Markov structure is derived finally.

Key words: stochastic system; optimal control; conditional Markov structure; Bemllman stochastic nonlinear dynamic programming

1 引言(Introduction)

在工程实际问题中, 存在着大量的动力学系统, 由于随机突变现象引起系统的跳变, 诸如互联子系统的变化、环境条件等的突变、非线性系统经线性化后工作点的迁移、系统元件的故障、参数的改变等等. 人们通过大量的研究发现, 这种随机变化的规律通常遵循Markov过程的变化规律[1]. 具有上述特征的系统, 一般来讲既包含了连续的系统状态, 又包含了跳变的结构状态, 故此类系统又称为随机跳变系统. 目前, 结构随机跳变系统在飞行器控制、通信技术、机器人技术以及其他领域都具有广泛的应用[2~4], 深入研究此类系统的有效的辨识、状态估计与控制方法, 对于有关此类复杂系统问题的解决无疑具有重要的学术意义及应用价值.

根据结构状态和系统状态的关系可将结构随机跳变系统分为两类, 一类为具有马尔科夫结构随机跳变系统, 即系统结构跳变与状态无关, 另一类为具有条件马尔科夫结构随机跳变系统, 即系统结构跳变与状态相关[5]. 目前针对结构随机跳变系统的研究主要集中的第一类系统[6~9], 并且大多假定系统的结构跳变模式已知, 不考虑系统模态的辨识, 而在现实应用过程中由于系统结构跳变的随机性, 系统的模态不一定完全已知. 因此, 此类系统的应用受到严重的限制. 对于第2类系统目前研究较少, 文献[5]研究了具有条件马尔科夫结构随机跳变线性系统最优控制问题, 但算法太抽象, 且没有具体推导过程. 文献[10, 11]也是针对非常一般的模型给出了有关滤波和控制的抽象结论, 并没有针对具有条件马尔科夫结构随机跳变系统滤波问题具体讨论, 其研究结果很难应用.

本文对具有条件马尔科夫结构的离散随机跳变系统最优控制问题进行研究, 利用随机最大值原理, 推导了最优控制算法. 考虑到算法的复杂性和可实

收稿日期: 2008−11−20; 收修改稿日期: 2009−04−20.
基金项目: 国家自然科学基金资助项目(60674031); 国家高科技发展计划"863"计划资助项目(2007AA701405); 空军工程大学工程学院优秀博士论文创新基金资助项目(BC07004).

现性, 将系统状态的后验概率密度函数逼近为已知的函数, 并使用此逼近方法获得随机跳变系统的逼近最优控制算法.

2 条件马尔科夫结构的线性控制系统 (Linear control system with conditional Markov structure)

假设离散时间随机跳变线性系统可以表示为

$$\begin{cases}x_{k+1}=A_k(s_k)x_k+B_k(s_k)u_k+F_k(s_k)\xi_k,\\ z_{k+1}=C_{k+1}(s_{k+1},u_k)x_k+E_{k+1}(s_{k+1},u_k)\zeta_k.\end{cases} \tag{1}$$

$$s_k=\overline{1,m}, k=0,1,\cdots,l.$$

结构状态转移概率为

$$q_{k+1}(s_{k+1}|s_k,x_k,u_k). \tag{2}$$

代价函数:

$$J_1=\lim_{u_{\overline{0,l-1}}}\mathrm{E}[\sum_{k=1}^{l}W_k(x_k,s_k,u_{k-1})]. \tag{3}$$

式中:

$$W_k(\cdot)=x_k^{\mathrm T}H_k(s_k)x_k+u_{k-1}^{\mathrm T}U_k(s_k)u_{k-1}. \tag{4}$$

$H_k(s_k),U_k(s_k)$为关于s_k的确定的已知向量矩阵.

控制目的: 要求寻找最优控制u_k^*, 它依赖于观测向量$z_{\overline{0,k}}$, 使J_1达到最小.

注 在本小节中, u_k不受约束, 且不考虑θ_k.

定理 1 对于离散时间随机跳变线性系统(1), 满足代价函数(3)的最优控制量表示为

$$u_k^*=-\tilde{\varPhi}_{k+1}^{-1}B_k^{\mathrm T}[(\tilde D_{k+1}+\tilde\wedge_{k+1})A_k\hat X_k]. \tag{5}$$

式中:

$$\begin{aligned}\tilde D_k=&(H_k\tilde P_k+\tilde L_k+\tilde N_k\tilde\varOmega_k)\tilde\pi_k+\\&[(H_k\tilde P_k+\tilde L_k+\tilde N_k\tilde\varOmega_k)\tilde\pi_k]^{\mathrm T},\\ \tilde\wedge_k=&\tilde\pi_k^{\mathrm T}\tilde N_k(I-\tilde\varOmega_k)\tilde\pi_k+[\tilde\pi_k^{\mathrm T}\tilde N_k(I-\tilde\varOmega_k)\tilde\pi_k]^{\mathrm T},\\ \tilde\varPhi_k=&2t_r(\tilde P_kU_k)+B_{k-1}^{\mathrm T}(\tilde D_k+\tilde\wedge_k)B_{k-1},\\ \tilde\pi_k=&\tilde P_k^{-1}q_k\hat P_{k-1},\tilde\varOmega_k=\tilde R_kC_k^{\mathrm T}\tilde Q_k^{-1}C_k,\\ \tilde Q_k=&C_k\hat R_kC_k^{\mathrm T}+E_kQ_kE_k^{\mathrm T}.\end{aligned}$$

并且模块参数$\tilde L_k,\tilde N_k,J_k^0$由下列迭代方程给出:

$$\begin{aligned}\tilde L_k=&\frac12A_k^{\mathrm T}\tilde D_{k+1}A_k,\\ \tilde N_k=&\frac12[A_k^{\mathrm T}\tilde\wedge_{k+1}A_k-\tilde K_{k+1}^{\mathrm T}B_k^{\mathrm T}(\tilde D_{k+1}+\tilde\wedge_{k+1})A_k],\\ J_k^0=&J_{k+1}^0+\frac12D_{k+1}F_kG_kF_k^{\mathrm T},\\ \tilde K_k=&\tilde\varPhi^{-1}B_{k-1}^{\mathrm T}(\tilde D_k+\tilde\wedge_k)A_{k-1}.\end{aligned}$$

证 为了求出最优控制算法, 先考虑函数$\tilde W_k(\cdot)$. 根据式(15), 得

$$\begin{aligned}\tilde W_k=&\mathrm E[W_k(x_k,s_k,u_{k-1})]=\\&\mathrm E[x_k^{\mathrm T}H_k(s_k)x_k+u_{k-1}^{\mathrm T}U_k(s_k)U_{k-1}]=\\&\mathrm{tr}[H_k\tilde P_k\tilde\varPhi_k+\tilde P_kU_kU_{k-1}U_{k-1}^{\mathrm T}].\end{aligned} \tag{6}$$

式中$U_k,H_k,\tilde P_k,\tilde\varPhi_k$为模块对角阵. 其元素为$U_k(s_k)$, $H_k(s_k),\tilde P_k(s_k),\tilde\varPhi_k(s_k)$. $\mathrm{tr}(\cdot)$表示求迹运算.

$$\tilde\varPhi_k(s_k)=\mathrm E[x_kx_k^{\mathrm T}|s_k,z_{\overline{0,k-1}},u_{\overline{0,k-1}}].$$

假定J_k^*为

$$J_k^*=\mathrm{tr}(\tilde L_{k-1}\hat\varPhi_{k-1}+\tilde N_{k-1}\hat X_{k-1}\hat X_{k-1}^{\mathrm T})+J_{k-1}^0. \tag{7}$$

式中 $\hat\varPhi_{k-1}$ 为模块对角矩阵, 其元素为 $\hat\varPhi_{k-1}(s_k)$; $\hat X_{k-1}$为模块向量, 其元素为$\hat x_{k-1}(s_{k-1})$, $s_{k-1}=\overline{1,m}$; $\tilde L_{k-1},\tilde N_{k-1}$为待定矩阵, J_{k-1}^0为待定标量函数.

$$\hat\varPhi_{k-1}(s_{k-1})=\mathrm E[x_{k-1}x_{k-1}^{\mathrm T}|s_{k-1},z_{\overline{0,k-1}},u_{\overline{0,k-1}}].$$

由式(7)得

$$J_{k+1}^*=\mathrm{tr}(\tilde L_k\hat\varPhi_k+\tilde N_k\hat X_k\hat X_k^{\mathrm T})+J_k^0. \tag{8}$$

积分函数$\int_{-\infty}^{\infty}J_{k+1}^*(z_k)\bar f(z_k)\mathrm dz_k$可表示为

$$\begin{aligned}&\int_{-\infty}^{\infty}J_{k+1}^*(z_k)\bar f(z_k)\mathrm dz_k=\\&\mathrm{tr}(\tilde L_k\tilde\varPhi_k+\tilde N_k\tilde X_k\tilde X_k^{\mathrm T}+\tilde N_k\tilde\varOmega_k\tilde R_k)+J_k^0.\end{aligned} \tag{9}$$

式中:

$$\tilde\varOmega_k=\tilde R_k\tilde C_k^{\mathrm T}\tilde Q_k^{-1}C_k,\tilde Q_k=C_k\tilde R_kC_k^{\mathrm T}+E_kQ_kE_k^{\mathrm T},$$

$\tilde X_k$为模块向量, 元素为$\tilde x_k(s_k)$; $\tilde R_k,C_k,E_k$为模块对角矩阵, 元素为$\tilde R_k(s_k),C_k(s_k),E_k(s_k)$, $s_k=\overline{1,m}$.

由$\tilde R_k=\tilde\varPhi_k-\tilde x_k\tilde x_k^{\mathrm T}$, 代入式(9)中, 得

$$\begin{aligned}&\int_{-\infty}^{\infty}J_{k+1}^*(z_k)\bar f(z_k)\mathrm dz_k=\\&\mathrm{tr}[(\tilde L_k+\tilde N_k\tilde\varOmega_k)\tilde\varPhi_k+\tilde N_k(I-\tilde\varOmega_k)\tilde X_k\tilde X_k^{\mathrm T}]+J_k^0.\end{aligned} \tag{10}$$

第k步的预测估计可以借助第$k-1$步后验估计得到, 即

$$\begin{aligned}&\mathrm{tr}\,\tilde\varPhi_k=\\&\mathrm{tr}[\tilde\pi_k(A_{k-1}\hat\varPhi_{k-1}A_{k-1}^{\mathrm T}+2A_{k-1}\hat X_{k-1}u_{k-1}^{\mathrm T}B_{k-1}^{\mathrm T}+\\&B_{k-1}u_{k-1}u_{k-1}^{\mathrm T}B_{k-1}^{\mathrm T}+F_{k-1}G_{k-1}F_{k-1}^{\mathrm T})],\end{aligned} \tag{11}$$

$$\begin{aligned}&\mathrm{tr}(\tilde X_k\tilde X_k^{\mathrm T})=\\&\mathrm{tr}[\tilde\pi_k(A_{k-1}\hat X_{k-1}\hat X_{k-1}^{\mathrm T}A_{k-1}^{\mathrm T}+\\&2A_{k-1}\hat X_{k-1}u_{k-1}^{\mathrm T}B_{k-1}^{\mathrm T}+B_{k-1}u_{k-1}u_{k-1}^{\mathrm T}B_{k-1}^{\mathrm T})\tilde\pi_k^{\mathrm T}].\end{aligned} \tag{12}$$

式中: $\tilde\pi_k=\tilde p_k^{-1}q_k\hat p_{k-1}$, B_k为模块向量, 其元素为$B_k(s_k)$, $\hat p_k,A_k,F_k$分别为模块对角阵, 其元素为$\hat p_k(s_k),A_k(s_k),F_k(s_k)$, q_{k+1}为转移概率矩阵, 其元

素为$q_{k+1}(s_{k+1}|s_k,x_k,u_k)$. 利用式(6)(10)(11)(12)得:

$$\begin{aligned}&\tilde{W}_k+\int_{-\infty}^{\infty}J_{k+1}^*(z_k)\bar{f}(z_k)\mathrm{d}z_k=\\&\mathrm{tr}[\frac{1}{2}\tilde{D}_{k-1}A_{k-1}\hat{\varPhi}_{k-1}\hat{A}_{k-1}+\\&(\tilde{D}_k+\tilde{\wedge}_k)(A_{k-1}\hat{X}_{k-1}u_{k-1}^{\mathrm{T}}B_{k-1}^{\mathrm{T}}+\\&\frac{1}{2}B_{k-1}u_{k-1}u_{k-1}^{\mathrm{T}}B_{k-1}^{\mathrm{T}})+\frac{1}{2}\tilde{D}_kF_{k-1}G_{k-1}F_{k-1}^{\mathrm{T}}+\\&\frac{1}{2}\tilde{\wedge}_kA_{k-1}\hat{X}_{k-1}\hat{X}_{k-1}^{\mathrm{T}}A_{k-1}^{\mathrm{T}}+\tilde{P}_kU_ku_{k-1}u_{k-1}^{\mathrm{T}}]+J_k^0.\end{aligned}\tag{13}$$

式中:

$$\begin{aligned}\tilde{D}_k=&(H_k\tilde{P}_k+\tilde{L}_k+\tilde{N}_k\tilde{\varOmega}_k)\tilde{\pi}_k+\\&[(H_k\tilde{P}_k+\tilde{L}_k+\tilde{N}_k\tilde{\varOmega}_k)\tilde{\pi}_k]^{\mathrm{T}},\\\tilde{\wedge}_k=&\tilde{\pi}_k^{\mathrm{T}}\tilde{N}_k(I-\tilde{\varOmega}_k)\tilde{\pi}_k+[\tilde{\pi}_k^{\mathrm{T}}\tilde{N}_k(I-\tilde{\varOmega}_k)\tilde{\pi}_k]^{\mathrm{T}},\end{aligned}$$

I为单位矩阵.

由下列等式可求出最优控制:

$$\begin{aligned}&u_{k-1}^*=\\&\arg\{\frac{\partial}{\partial u_{k-1}}[\tilde{W}_k+\int_{-\infty}^{\infty}J_{k+1}^*(z_k)\bar{f}(z_k)\mathrm{d}z_k]=0\}.\end{aligned}\tag{14}$$

将式(12)代入式(13), 得

$$\begin{aligned}&u_{k-1}^*=\\&-\varPhi_k^{-1}B_{k-1}^{\mathrm{T}}[\tilde{D}_k+\tilde{\wedge}_k]A_{k-1}\hat{X}_{k-1}=\\&-\tilde{K}_k\hat{X}_{k-1}.\end{aligned}\tag{15}$$

式中

$$\tilde{K}_k=\varPhi_k^{-1}B_{k-1}^{\mathrm{T}}[\tilde{D}_k+\tilde{\wedge}_k]A_{k-1}.$$

得到

$$J_k^*=\tilde{W}_k^*+\int_{-\infty}^{\infty}J_{k+1}^{**}(z_k)\bar{f}(z_k)\mathrm{d}z_k.\tag{16}$$

式中$\tilde{W}_k^*, J_{k+1}^{**}$分别为$\tilde{W}_k, J_{k+1}^*$在$u_{k-1}=u_{k-1}^*$时的取值, 将(12)(15)代入(16), 并根据$\hat{\varPhi}_k, \hat{X}_{k-1}\hat{X}_{k-1}^{\mathrm{T}}$前面的系数相等, 得未知矩阵$\tilde{L}_k, \tilde{N}_k, \tilde{g}_k$的逼近迭代方程, 近而得到定理成立.

注意, 上述系数矩阵依赖于状态估计器参数$\tilde{p}_k$, $\hat{p}_k, \tilde{R}_k$, 因此必须通过最优状态估计算法获得. 由于滤波算法推导的复杂性, 滤波结果具体推导过程请参考文献[15].

3　数值算例(Numerical example)

考虑如下的简单2结构随机跳变系统, 状态方程为:

$$\begin{cases}x^{(1)}(k+1)=a_1x^{(1)}(k)+b_1u(k)+f_1\xi(k),\\x^{(2)}(k+1)=a_2x^{(2)}(k)+b_2u(k)+f_2\xi(k).\end{cases}\tag{17}$$

其中参数为

$$\begin{aligned}&a_1=-1, b_1=1, f_1=1,\\&a_2=-2, b_2=1, f_2=1,\end{aligned}$$

$\xi(k)$为强度为1的Gauss白噪声.

观测方程为:

$$z(k)=x(k)+\zeta(k).\tag{18}$$

$\zeta(k)$为强度为1的Gauss白噪声.

仿真中$x(1)=10$, 系统的结构转换强度矩阵

$$q=\begin{bmatrix}0.9&0.1\\0.1&0.9\end{bmatrix}.$$

滤波器初始设置为

$$\begin{aligned}&x^{(1)}(1)=x^{(2)}(1)=0, \hat{p}^{(1)}(1)=0.8,\\&\hat{p}^{(2)}(1)=0.2, R^{(1)}(1)=R^{(2)}(1)=Q.\end{aligned}$$

仿真中代价函数

$$\begin{aligned}J_1=\min_u \mathrm{E}[\sum_{k=1}^{3}&x_k^{\mathrm{T}}H_k(s_k)x_k+\\&u_{k-1}^{\mathrm{T}}U_k(s_k)u_{k-1}].\end{aligned}\tag{19}$$

式中的参数$H_k(s_k)=U_k(s_k)=1$.

首先根据预测算法正向计算各结构的结构概率和协方差如表1所示.

表 1　各结构的结构概率和协方差

Table 1　Structure probability and covariance of each structures

	离散时间		
	1	2	3
$\hat{p}^{(1)}(k)$	0.8	0.74	0.692
$\hat{p}^{(2)}(k)$	0.2	0.26	0.308
$R^{(1)}(k)$	1	2.25	4.4197
$R^{(2)}(k)$	1	9	35.47

控制器参数的终止条件$\tilde{L}_3=\tilde{N}_3=I_{2\times2}$, 通过反向迭代求解控制器相关参数如表2所示. 控制器相关参数确定后可以确定$\tilde{K}(1)=[-0.8550\quad -0.2811]$, 由控制量的表达式确定第1步的控制量为, $u(1)=11.362$, 代入到系统的状态方程参与状态控制. 第2步以后的系统状态、结构估计及协方差等通过观测方程经过滤波算法得到, 得到上述滤波结果后同第1步类似, 首先通过预测算法正向计算各结构以后时间的结构概率和协方差预测值, 在反向迭代求解控制器相关参数, 最后得到控制量的数值参与控制.

表 2 控制器相关参数

Table 2 Prameters of controller

	离散时间		
	1	2	3
$\tilde{L}_k$	$\begin{bmatrix} -28.468 & 176.07 \\ -704.3 & 3617.4 \end{bmatrix}$	$\begin{bmatrix} 5.5005 & 4.2676 \\ 17.07 & 132.4 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$\tilde{N}_k$	$\begin{bmatrix} -58.251 & 527.33 \\ 527.33 & 3267.6 \end{bmatrix}$	$\begin{bmatrix} -6.8133 & -8.4011 \\ -8.4011 & 126.38 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$\tilde{\Lambda}_k$	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 217.41 & 582.23 \\ 582.23 & 1642.8 \end{bmatrix}$	$\begin{bmatrix} -9.0199 & -14.214 \\ -14.214 & -55.855 \end{bmatrix}$
$\tilde{D}_k$	$\begin{bmatrix} 126.32 & 14.873 \\ 14.873 & -638.14 \end{bmatrix}$	$\begin{bmatrix} -56.971 & -352.15 \\ -352.15 & 1808.7 \end{bmatrix}$	$\begin{bmatrix} 11.001 & 8.5351 \\ 8.5351 & 66.2 \end{bmatrix}$

4 结论(Conclusion)

本文利用Bellman随机非线性动态规划法, 将具有马尔科夫结构的离散随机系统的最优控制问题分解为逆向求解Bellman泛函方程和后验概率密度函数, 并通过将后验概率密度逼近条件高斯分布, 获得逼近最优控制算法. 最后给出了一类具有马尔可夫结构的线性离散随机系统的最优控制器. 此算法不需要结构指示器, 易于实现, 具有更广阔的应用前景.

参考文献(References):

[1] COSTA O L V, FRAGOSO M D, MARQUES R P. *Discrete Time Markov Jump Linear Systems*[M]. London: Springer-Verlag, 2005: 52 – 78.

[2] GRAY W S, GONZALEZ O R, DOGAN M. Stability analysis of digital linear flight controllers subject to electromagnetic disturbance[J]. *IEEE Transactions on Aerospace and Electronic systems*, 2000, 36(4): 1204 – 1218.

[3] DO VAL J B R, COSTA E F. Numerical solution for linear-quadratic control problems of Markov jump linear system and weak detectablility concept[J]. *Journal of Optimization Theory and Applications*, 2002, 114(1): 69 – 96.

[4] STOICA A, YAESH I. Jump Markovian-based control of wing deployment for an uncrewed air vehicle[J]. *Journal of Guidance*, 2002, 25(2): 407 – 411.

[5] 吴森堂. 结构随机跳变系统理论及其应用[M]. 北京: 科学出版社, 2007: 131 – 154.

[6] BOUKAS E K, LIU Z K. Robust H_∞-control of discrete-time Markovian jump linear systems with mode-dependent time-delays[J]. *IEEE Transactions on Automatic Control*, 2001, 46(12): 1918 – 1924.

[7] PARK B, KWON W H. Robust one-step receding horizon control of discrete-time Markovian jump uncertain systems[J]. *Automatica*, 2002, 38(7): 1229 – 1235.

[8] MAGDI S M, SHI P. Robust control for Markovian jump linear discrete-time systems with unknown nonlinearities[J]. *IEEE Transactions on Circuits and Systems*, 2002, 49(4): 538 – 542.

[9] SHI P, XIA Y Q, LIU G P, et al. On designing of sliding-mode control for stochastic jump systems[J]. *IEEE Transactions on Automatic Control*, 2006, 51(1): 97 – 105.

[10] BUKHALOV B A. Optimization smoothing in systems with random jump-like structure[J]. *Automatization and Telecontrol*, 1996, 4: 56 – 63.

[11] SOTSKOVA Y L. Representation of Markov processes with jumps backward in time and related matters optimization nonlinear estimation[J]. *Automatization and Telecontrol*, 2006, 9: 23 – 29.

[12] 方洋旺. 随机系统分析与应用[M]. 西安: 西北工业大学出版社, 2006: 45 – 65.

作者简介:

方洋旺 (1966—), 男, 教授, 博士生导师, 研究方向为随机最优控制理论与应用、导弹制导与控制、导弹作战使用、非线性控制等;

王洪强 (1980—), 男, 讲师, 主要研究方向为随机最优控制理论、导弹制导与控制等, E-mail: wanghongqiang2009@163.com;

伍友利 (1979—), 男, 博士研究生, 主要研究方向为导弹制导与控制、导弹作战使用等.