文章编号: 1000 - 8152(2004)06 - 1032 - 04

半 Markov 过程基于性能势的灵敏度分析和性能优化

李衍杰[1], 殷保群[1], 奚宏生[1], 周亚平[2], 代桂平[1]

(1.中国科学技术大学 自动化系,安徽 合肥 230027; 2.中国科学技术大学 管理科学系,安徽 合肥 230027)

摘要: 基于性能势的方法,研究了一类半 Markov 过程(SMP)的性能灵敏度分析和平均费用下的性能优化问题.将 SMP 转化为与之等价的离散时间 Markov 链(DTMC),利用 DTMC 的性能势,对 SMP 进行灵敏度分析和性能优化,得到了 SMP 基于 DTMC 性能势的灵敏度分析公式和最优性方程.最后给出了一个数值例子以表明该方法的应用.

关键词: 半 Markov 过程; 性能势; 灵敏度分析; 最优性方程

中图分类号: O226　　**文献标识码**: A

Sensitivity analysis and performance optimization of semi-Markov processes based on performance potentials

LI Yan-jie[1], YIN Bao-qun[1], XI Hong-sheng[1], ZHOU Ya-ping[2], DAI Gui-ping[1]

(1. Department of Automation, University of Science and Technology of China, Hefei Anhui 230027, China;
2. Department of Management Science, University of Science and Technology of China, Hefei Anhui 230027, China)

Abstract: Based on performance potential theories, the sensitivity analysis and performance optimization of the semi-Markov process(SMP) under the average cost criterion are studied. By applying the transformation of the SMP to the discrete-time Markov chain(DTMC), the potential of the DTMC is used to obtain the sensitivity formula and optimality equation of SMP. Finally, a numerical example is provided to illustrate the applicability of the method.

Key words: semi-Markov process; performance potential; sensitivity analysis; optimality equation

1 引言(Introduction)

Markov 过程作为一种数学模型已被广泛应用于许多实际系统的建模和分析中,基于 Markov 过程的重要性,曹希仁教授和陈翰馥院士提出了 Markov 性能势理论[1],为 Markov 过程的性能分析提供了一个一致的框架,同时也减弱了对传统摄动分析方法所要求的较强的假设条件.但在许多实际问题中,大部分的系统是非 Markov 型的,如半 Markov 型.与 Markov 过程相比,SMP 具有更广泛的应用范围.文献[2]利用 DTMC 的动态规划方法研究了 SMP 决策问题,给出了平均费用准则下的最优性方程和策略迭代算法.文献[3]在 SMP 中引入了一个与 Markov 过程中无穷小矩阵相似的量,定义了 SMP 的性能势,利用该性能势研究了 SMP 的灵敏度分析和决策问题.本文借鉴文献[2]的思想,从性能势的角度运用 DTMC 的性能势,对一类 SMP 进行灵敏度分析,给出了一种基于单一样本轨道对性能灵敏度的估计算法,同时在平均费用准则下,给出了 SMP 基于 DTMC 性能势的最优性方程.

2 SMP 的稳态性能(Steady-state performance of SMP)

设 $X = \{X_t, t \geqslant 0, E, Q_{ij}(t), f\}$ 为一个 SMP, $E = \{1,2,\cdots,M\}$ 为其状态空间, $Q_{ij}(t)$ 为 SMP 核,则 $Q_{ij}(t) = p_{ij}F_{ij}(t)$,其中 p_{ij} 为其状态转移概率,$F_{ij}(t)$ 为逗留时间分布函数,$f: E \to R$ 为 SMP 的性能函数.设 $Y = \{X_n, n = 1,2,\cdots, E, p_{ij}, \tilde{f}\}$ 是 SMP X 的嵌入 Markov 链,$\tilde{f}: E \to R$ 为嵌入 DTMC 的性能函数.假定 SMP X 是不可约遍历的.记

$$p_{ij}(t) = p(X_t = j \mid X_0 = i),\ Q_i(t) = \sum_{j=1}^{M} Q_{ij}(t),$$

$$\eta_{ij} = \int_0^\infty t\mathrm{d}F_{ij}(t),\ \eta_i = \int_0^\infty t\mathrm{d}Q_i(t).$$

引理 1[4]　对于 SMP X,若对任意 $i \in E$,相继进入状态 i 之间的时间 T_{ii} 具有非格点分布和有限

收稿日期:2003 - 01 - 05; 收修改稿日期:2003 - 12 - 17.
基金项目:国家自然科学基金项目(60274012);安徽省自然科学基金项目(01042308).

均值，则极限 $p_i = \lim_{t\to\infty} p_{ji}(t)$ 存在；将 SMP 看成再生过程，则有

$$p_i = \frac{\mathrm{E}(\text{一个循环中处于 } i \text{ 的时间})}{\mathrm{E}(\text{一个循环的时间})}.$$

这里称 $\mathrm{E}(\cdot)$ 表示数学期望. 称 p_i 为 SMP X 在状态 i 的极限分布，记 $p = (p_1, p_2, \cdots, p_M)$. 定义 $\eta_f = \sum_{i=1}^{M} p_i f(i) = pf$ 为 SMP X 的稳态性能，其中 $f = (f(1), \cdots, f(M))^{\mathrm{T}}$. 若 X 为不可约遍历的，则 Y 存在稳态分布，设稳态分布为 $\pi = (\pi_1, \cdots, \pi_M)$ 有

$$\pi P = \pi, \ \pi e = 1.$$

其中 $e = (1, 1, \cdots, 1)^{\mathrm{T}}$. 定义 Y 的稳态性能为 $\tilde{\eta}_{\tilde{f}} = \sum_{i=1}^{M} \pi_i \tilde{f}(i) = \pi \tilde{f}$，其中 $\tilde{f} = (\tilde{f}(1), \cdots, \tilde{f}(M))^{\mathrm{T}}$.

引理 2[4]　若 SMP X 满足引理 1 的条件，则

$$p_i = \frac{\pi_i \eta_i}{\sum_{j=1}^{M} \pi_j \eta_j}. \tag{1}$$

定理 1　满足引理 1 条件的 SMP X 与其嵌入 DTMC 的稳态性能完全相同，即 $\eta_f = \tilde{\eta}_{\tilde{f}}$，其中 $\tilde{f} = \Lambda f$，

$$\Lambda = \mathrm{diag}\left\{\frac{\eta_1}{\sum_{j=1}^{M} \pi_j \eta_j}, \cdots, \frac{\eta_M}{\sum_{j=1}^{M} \pi_j \eta_j}\right\}.$$

证　由式(1)和 SMP 稳态性能的定义，

$$\eta_f = \sum_{i=1}^{M} p_i f(i) =$$

$$\sum_{i=1}^{M} \frac{\pi_i \eta_i}{\left(\sum_{j=1}^{M} \pi_j \eta_j\right)} f(i) = \sum_{i=1}^{M} \pi_i \tilde{f}(i) = \tilde{\eta}_{\tilde{f}}.$$

证毕.

3　DTMC 的性能势(Performance potentials of DTMC)

在这里简要介绍一下 DTMC 的性能势的有关内容，详见文献[1,5]. 称

$$d_{ij} = \lim_{N\to\infty} \{\mathrm{E}_j[\sum_{n=0}^{N} \tilde{f}(X_n)] - \mathrm{E}_j[\sum_{n=0}^{N} \tilde{f}(X_n)]\}$$

为 DTMC 关于性能函数 $\tilde{f}$ 的实现因子. 矩阵 $D^{\tilde{f}} = [d_{ij}]$ 称为实现矩阵. 这里 E_j 表示对初始状态 $X_0 = j$ 求条件期望. 对任意常数 c 和 $k^* \in E$，定义一个量 x_i，

$$x_i = x_{k^*} + d_{k^* i} = d_{k^* i} + c, \ \forall i \in E, \tag{2}$$

称 x_i 为 DTMC 关于性能函数 $\tilde{f}$ 在状态 i 的性能势，记 $x^{\tilde{f}} = (x_1, \cdots, x_M)^{\mathrm{T}}$ 为性能势向量，由式(2)，显然有 $d_{ij} = x_j - x_i, i, j \in \mathrm{E}$，因此有

$$D^{\tilde{f}} = e(x^{\tilde{f}})^{\mathrm{T}} - x^{\tilde{f}} e^{\mathrm{T}}. \tag{3}$$

引理 3[1]　如果 $\lim_{n\to\infty} \mathrm{E}[|\tilde{f}(X_n)|] = E_\pi[|\tilde{f}|]$，其中 $E_\pi(|\tilde{f}|) = \sum_{i\in E} \pi_i |\tilde{f}(i)|$，则

$$d_{ij} = \mathrm{E}[\sum_{n=0}^{L^j(i)-1} \tilde{f}(X_n) - L^j(i) \tilde{\eta}_{\tilde{f}}]. \tag{4}$$

其中 $L^j(i) = \inf_n \{n \mid X_n = i, X_0 = j\}$.

由此可知实现矩阵可以基于一条样本轨道进行估计.

引理 4[5]　势向量满足如下 Poisson 方程：

$$(I - P) x^{\tilde{f}} = \tilde{f} - \tilde{\eta}_{\tilde{f}} e, \tag{5}$$

且所有性能势向量可表示为

$$x^{\tilde{f}} = (I - P + e\pi)^{-1} \tilde{f} + \beta e. \tag{6}$$

β 为任意常数. 可见，性能势 $x^{\tilde{f}}$ 不是唯一的. 由式(3)得

$$x^{\tilde{f}} = -D^{\tilde{f}} \pi^{\mathrm{T}} \tag{7}$$

是一性能势，所以 $x^{\tilde{f}}$ 也可通过实现矩阵由一条样本轨道估计得到.

4　SMP 的灵敏度分析(Sensitivity analysis of SMP)

假设 SMP X 的状态转移概率 p_{ij} 和逗留时间分布 $F_{ij}(t)$ 为区间 $J \subset \mathbb{R}$ 上变化的参数 θ 的函数，即 $p_{ij} = p_{ij}(\theta), F_{ij}(t, \theta)\ i, j \in E$. 设 SMP 的性能函数 f 也与 θ 有关，同时设 p_{ij}, η_{ij} 和 f 是 θ 的可微函数. 由式(5)微分可得

$$-\frac{\partial P}{\partial \theta} x^{\tilde{f}} + (I - P) \frac{\partial x^{\tilde{f}}}{\partial \theta} =$$

$$\frac{\partial \Lambda}{\partial \theta} f + \Lambda \frac{\partial f}{\partial \theta} - \frac{\partial \tilde{\eta}_{\tilde{f}}}{\partial \theta} e =$$

$$\frac{\partial \Lambda}{\partial \theta} f + \Lambda \frac{\partial f}{\partial \theta} - \frac{\partial \tilde{\eta}_{\tilde{f}}}{\partial \theta} e. \tag{8}$$

由定理 1，式(8)两边左乘 π，得到 SMP 的稳态性能关于参数 θ 的导数公式为

$$\frac{\partial \eta_f}{\partial \theta} = \pi \frac{\partial P}{\partial \theta} x^{\tilde{f}} + \pi \frac{\partial \Lambda}{\partial \theta} f + p \frac{\partial f}{\partial \theta}. \tag{9}$$

对右边第二项，设 $\eta = (\eta_1, \eta_2, \cdots, \eta_M)^{\mathrm{T}}$，则

$$\frac{\partial \Lambda}{\partial \theta} = -\frac{1}{\pi\eta} \frac{\partial \pi\eta}{\partial \theta} \Lambda + \frac{1}{\pi\eta} \mathrm{diag}\left(\frac{\partial \eta_1}{\partial \theta}, \cdots, \frac{\partial \eta_M}{\partial \theta}\right) =$$

$$-\frac{\pi \dfrac{\partial P}{\partial \theta} x^{\eta} + \pi \dfrac{\partial \eta}{\partial \theta}}{\pi\eta} \Lambda + \frac{1}{\pi\eta} \mathrm{diag}\left(\frac{\partial \eta_1}{\partial \theta}, \cdots, \frac{\partial \eta_M}{\partial \theta}\right),$$

其中第二个等式是将 $\pi\eta$ 看作 DTMC Y 的稳态性能

得到的.代入式(9)得到

$$\frac{\partial \eta_f}{\partial \theta} = \pi \frac{\partial P}{\partial \theta} x^{\tilde{f}} - \frac{1}{\pi\eta}[(\pi \frac{\partial P}{\partial \theta} x^{\eta} + \pi \frac{\partial \eta}{\partial \theta})\eta_f - \pi\, \mathrm{diag}(\frac{\partial \eta_1}{\partial \theta}, \cdots, \frac{\partial \eta_M}{\partial \theta}) f] + p \frac{\partial f}{\partial \theta}. \tag{10}$$

下面给出一种基于单一样本轨道,对性能势和性能灵敏度的仿真估计算法.

步骤1 通过一次仿真预先估计 π,由 $\tilde{f} = \Lambda f$ 计算 $\tilde{f}$,并计算 η_i 及其关于 θ 导数.

步骤2 通过第二次仿真估计 $\mathrm{E}[\sum_{n=0}^{L^j(i)-1} \tilde{f}(X_n)]$, $\mathrm{E}[L^j(i)]$, $\eta_{\tilde{f}}$.

步骤3 根据式(4),(7)计算实现矩阵 $D^{\tilde{f}}, D^{\eta}$ 和性能势 $x^{\tilde{f}}, x^{\eta}$.

步骤4 根据式(10)计算稳态性跑导数.

5 SMDP 平均费用准则下的最优性方程(Optimality equation of SMDP under the average cost criterion)

假设处于状态 i 时的行动集为 $D(i)$, $D = \bigcup_{i=1}^{M} D(i)$ 为行动空间,这里仅考虑平稳策略的情况.假设策略 $v: E \to D$ 仅发生在状态转移时刻且仅与状态有关,在该策略空间的任意策略驱动下的SMP仍满足引理1的条件.这时SMP的一切量将与策略有关,用右上角加 v 的对应量表示.设平均费用性能指标为

$$J^v = \lim_{T\to\infty} \frac{1}{T} \mathrm{E}\{\int_0^T f^v(X_t)\mathrm{d}t\}.$$

引理5 设SMP X 满足引理1的条件, T_1 为SMP第一次到达初始状态的时刻,若 $\mathrm{E}\left|\int_0^{T_1} f(X(t))\mathrm{d}t\right| < +\infty$,则有

$$\lim_{T\to\infty} \frac{1}{T} \mathrm{E}\int_0^T f(X_t)\mathrm{d}t = \sum_{j=1}^{M} p_j f(j) = \eta_f.$$

证 根据再生报酬过程

$$\lim_{T\to\infty} \frac{1}{T} \mathrm{E}\int_0^T f(X_t)\mathrm{d}t = \frac{\mathrm{E}\int_0^{T_1} f(X_t)\mathrm{d}t}{(ET_1)},$$

$\mathrm{E}\int_0^{T_1} f(X_t)\mathrm{d}t = \mathrm{E}\sum_{j=1}^{M} f(j)\xi_j = \sum_{j=1}^{M} f(j)\mathrm{E}\xi_j$(ξ_j 是在 $[0, T_1]$ 中处于状态 j 的时间),由引理1知

$$\mathrm{E}\int_0^{T_1} f(X_t)\mathrm{d}t = \sum_{j=1}^{M} f(j) p_j ET_1.$$

证毕.

可知 $J^v = \eta_f^v$.由定理1,对任意策略 v, SMDP X 与 DTMDP Y 的性能优化是完全等价的.由此采用与文献[7]相同的方法可以得到如下定理.

定理2 v^* 是SMDP平均费用性能准则下的最优平稳策略的充分必要条件是下列不等式成立:

$$\Lambda^{v^*} f^{v^*} + P^{v^*}(x^{\tilde{f}})^{v^*} \leqslant \Lambda^v f^v + P^v (x^{\tilde{f}})^{v^*}, \ \forall v \in D$$

或

$$0 = \min_{v\in A}\{\Lambda^v f^v - e\eta^{v^*} + (P^v - I)(x^{\tilde{f}})^{v^*}\}.$$

6 数值例子(Numerical example)

考虑一个具有3个状态的SMP系统,该系统的运行可以由状态转移概率和逗留时间分布来描述, $\theta_1, \theta_2, \theta_3$ 为系统的3个参数.已知系统的转移概率为

$$p_{11}(\theta_1) = 1 - \mathrm{e}^{-\theta_1/2},\ p_{12}(\theta_1) = \frac{7}{8}\mathrm{e}^{-\theta_1/2},$$

$$p_{13}(\theta_1) = \frac{1}{8}\mathrm{e}^{-\theta_1/2},\ p_{21}(\theta_2) = \frac{1-\mathrm{e}^{-\theta_2/4}}{1+\mathrm{e}^{-\theta_2/2}},$$

$$p_{23}(\theta_2) = \frac{1}{3}\mathrm{e}^{-\theta_2/4},\ p_{22}(\theta_2) = 1 - p_{21}(\theta_2) - p_{23}(\theta_2),$$

$$p_{31}(\theta_3) = \frac{1-\mathrm{e}^{-\theta_3/4}}{1+\mathrm{e}^{-\theta_3/4}},\ p_{33}(\theta_3) = 1 - \frac{1-\mathrm{e}^{-\theta_3/3}}{1+\mathrm{e}^{-\theta_3/2}},$$

$$p_{32}(\theta_3) = 1 - p_{31}(\theta_3) - p_{33}(\theta_3).$$

$F_{ij}(t, \theta_j)$ 为 $[0, i\theta_j]$ 上的均匀分布,其性能函数为 $f(i, \theta_i) = \ln[(1+i)] + \sqrt{i}/2\theta_i, i, j = 1,2,3$.在 $\theta_1 = 1, \theta_2 = 1, \theta_3 = 5$ 处,利用第4节给出的仿真算法,在每次仿真次数为1000000次的情况下求得稳态性能关于 $\theta_1, \theta_2, \theta_3$ 导数分别为0.1179, -0.5305,0.0417.而其理论值是

0.1179, -0.5308,0.0421.

通过以上两组数据的比较,可以看到该仿真方法具有较高的仿真精度.

7 结论(Conclusion)

在Markov链性能势的基础上,用转化的方法将SMP的灵敏度分析和性能优化转化为DTMC的灵敏度分析和性能优化,基于DTMC的性能势给出了灵敏度公式和最优性方程,并且利用该势能可由一条样本轨道来估计的特性,给出了一种基于一条样本轨道对性能灵敏度的仿真算法,因而可以应用于半Markov型系统在线性能分析和优化.

参考文献(References):

[1] CAO X R, CHEN H F. Perturbation realization, potentials and sensitivity analysis of Markov Processes [J]. *IEEE Trans on Automatic*

Control, 1997, 42(10): 1382 - 1393.
[2] PUTERMAN M L. *Markov Decision Processes: Discrete Stochastic Dynamic Programming* [M]. New York: Wiley, 1994.
[3] CAO X R. Semi-Markov decision problems and performance sensitivity analysis [J]. *IEEE Trans on Automatic control*, 2003, 48(5): 758 - 769.
[4] ROSS S M. *Stochastic Process* [M]. New York: John Wiley and Sons, 1983.
[5] CAO X R. A unified approach to Markov decision problems and performance sensitivity analysis [J]. *Automatica*, 2000, 36(5): 771 - 774.
[6] 殷保群，周亚平，杨孝先，等. 状态相关闭排队网络的性能指标灵敏度公式[J]. 控制理论与应用，1999, 16(2): 255 - 257.
(YIN Baoqun, ZHOU Yaping, YANG Xiaoxian, et al. Sensitivity formula of performance in close state-dependent queuing networks [J]. *Control Theory & Applications*, 1999, 16(2): 255 - 257.)
[7] 奚宏生，唐昊，殷保群. 连续时间 MCP 在紧致行动集上的最优策略[J]. 自动化学报，2003, 29(2): 206 - 211.)
(XI hongsheng, TANG Hao, Yin Baoqun. Optimal policies for a continuous time MCP with compact action space [J]. *Acta Automatica Sinica*, 2003, 29(2): 206 - 211.)

作者简介：

李衍杰　(1978 —)，男，中国科学技术大学自动化系在读博士生，主要研究方向是离散时间动态系统，E-mail: whylyj@ustc.edu.cn;

殷保群　(1962 —)，男，中国科学技术大学自动化系副教授，博士，主要研究方向为非线性系统展开理论、随机离散事件系统性能分析、优化及在通讯网络中的应用，E-mail: bqyin@ustc.edu.cn;

奚宏生　(1950 —)，男，中国科学技术大学自动化系教授，博士生导师，主要研究方向为鲁棒控制、离散事件动态系统及其应用；

周亚平　(1963 —)，男，中国科学技术大学管理科学系副教授，主要研究方向为经济管理系统、排队网络性能灵敏度仿真估计及优化；

代桂平　(1977 —)，男，中国科学技术大学自动化系在读博士生，主要研究方向是离散时间动态系统.

(上接第 1031 页)

[2] FREUDENBERG J S, LOOZE D P. Half-plane poles and zeros design trade-offs in feedback systems [J]. *IEEE Trans on Automatic Control*, 1985, 30(6): 555 - 565.
[3] GOMEZ G I, GOODWIN G C. Integral constraints on sensitivity vectors for multivariable linear systems [J]. *Automatica*, 1996, 32(4): 499 - 518.
[4] SUNG H K, HARA S. Properties of sensitivity and complementary sensitivity functions in SISO digital control systems [J]. *Int J Control*, 1988, 48(6): 2429 - 2439.
[5] FREUDENBERG J S, LOOZE D P. *Frequency domain properties of scalar and multivariable feedback systems* [M]. Berlin: Springer Verlag, 1988.
[6] MIDDLETON R H, GRAEBE S F. Slow stable open loop poles: To cancel or not to cancel? [R]. Newcastle: University of Newcastle, 1995.
[7] GOODWIN G C, WOODYATT A R, MIDDLETON R H, et al. Fundamental limitations due to $j\omega$-axis zeros in SISO systems [J]. *Automatica*, 1999, 35(5): 857 - 863.
[8] BOYD S, DESOER C A. Subharmonic functions and performance bonds on linear time invariant feedback systems [J]. *IMA J of Math Information Control*, 1995, 2(1): 153 - 170.
[9] CHEN J. Sensitivity integral relations and design trade-offs in linear multivariable feedback systems [C] // *Proc of American Control Conference*. San Franscisco: [s.n.], 1993: 3160 - 3164.
[10] SULE V R, ATHANI V V. Directional sensitivity trade-offs in multivariable feedback systems [J]. *Automatica*, 1991, 27(5): 869 - 872.
[11] HE H L, WANG Z S, LIAO X X. Limitation on the tracking problem due to $j\omega$-axis zeros in SISO feedback control systems [C] // *Proc of Dynamics of Continuous, Discrete and Impulsive Systems - Series B - Applications and Algorithms*. Canada: Watam Press, 2003: 103 - 108.
[12] HE H L, WANG Z S, LIAO X X. Limitation on the tracking problem due to $j\omega$-axis zeros in MIMO feedback control systems [J]. *Control Theory & Applications*, 2003, 20(2): 289 - 292.
(河汉林，王中生，廖晓昕. 多输入输出反馈系统虚轴的零点对跟踪问题的约束[J]. 控制理论与应用，2003, 20(2): 289 - 292.)
[13] VIDYASAGAR M. *Control System Synthesis: a Factorization Approach* [M]. MA: MIT Press, 1985.

作者简介：

何汉林　(1962 —)，男，海军工程大学副教授，1983 年于华中师范大学获学士学位，1989 年于重庆大学获硕士学位，2003 年于华中科技大学获博士学位，主要研究兴趣为反馈控制理论和应用泛函分析，E-mail: hanlinhe62@yahoo.com.cn;

王中生　(1965 —)，男，中原工学院副教授，1990 年于西北大学获硕士学位，2003 年于华中科技大学获博士学位，主要研究兴趣为自适应控制和神经网络；

廖晓昕　(1938 —)，男，华中科技大学控制科学与工程系教授，博士生导师，发表论文 200 余篇，出版专著 3 部，获湖北省自然科学一等奖，教育部科技进步二等奖两项，主要研究兴趣为非线性动力系统的稳定性、人工程神经网络的动力学行为.