国防科技大学学报
JOURNAL OF NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY
第12卷第3期 1990年9月 Vol.12 No.3

有限马氏链最优停止的线性规划解

谷建湘
（系统工程与应用数学系）

摘　要　离散时间的有限状态马尔可夫链最优停止的值函数存在的一个充分条件是所对应的线性规划有解，且其最优解等于值函数，本文证明这个条件还是必要的。

关键词　随机过程，马尔可夫链，线性规划，停止变量
分类号　O211.62

一个抽象的马尔可夫最优停止问题可以从[1]中找到。本文提到的马尔可夫链最优停止问题的描述如下（详细描述见[2]）：给定一个转移概率矩阵 $P=(p_{ij};\ i,\ j=1,\ 2,\ \cdots)$ 以及状态空间 $S=\{1,\ 2,\ \cdots\}$。在可测空间（Ω，$\mathscr{F}$）上诱导一个概率测度 P_i，$i\in S$，其中 $\Omega=\overset{\infty}{\underset{t=0}{X}}S$，$\mathscr{F}$ 为 Ω 的所有子集组成的 σ 域，有 $P_i\left(iX\overset{\infty}{\underset{t=1}{X}}S\right)=1$．在概率空间（$\Omega$，$\mathscr{F}$，$P_i$）上产生一个离散时间的马尔可夫链 $\{z_t,\ t=0,1,2,\cdots\}$（见[3]），用 z_t 来表示所叙系统在时刻 t 所处的状态。用 $\mathscr{F}_t$ 表示由 z_0，z_1，…，z_t 生成的 σ 域，$t=0$，1，2，…．当系统从初始状态 $z_0=i$ 出发且决策者使用停止变量 σ 时，所得到的总期望支付为

$$J_i(\sigma)=E_i\left\{\sum_{k=0}^{\sigma-1}f(z_k)+g(z_\sigma)I_{(\sigma<+\infty)}\right\} \tag{1}$$

其中 $E_i(\cdot)$ 表示在概率测度 P_i 下的期望；$I_{(\cdot)}$ 表示集$(\cdot)$的示性函数；恒约定 $\sum_{k=0}^{-1}f(k,\cdot)=0$.

记：$\Delta_i^m=\{\sigma;\ \rho$ 为$(\Omega,\mathscr{F},P_i)$中的一个停止变量，且 $P_i(\sigma\leqslant m)=1;\ i\in S;\ m=0,\ 1,\ 2,\ \cdots;\}$

$\Delta_i=\{\sigma;\ \sigma$ 为$(\Omega,\mathscr{F},P_i)$中的一个停止变量，$P_i(\sigma<+\infty)=1$ 且使 $J_i(\sigma)$ 存在有限$\}$

$B(S)=\{u;\ u(\cdot)$ 为 S 上的实函数，且 $\|u\|=\sup_{i\in S}|u(i)|<+\infty\}$。

以后总设 $f\in B(S)$，$g\in B(S)$．在[2]中马尔可夫链最优停止的值函数定义如下：如果对每个 $i\in S$，都存在 $\tau\in\Delta_i$，使得

$$J_i(\tau)=\sup_{\sigma\in\Delta_i}J_i(\sigma)<+\infty$$

1989年3月28日收稿

则称$J_i(\tau)$，$i\in S$为马尔可夫链最优停止的值函数，相应的τ（与i有关）称为最优停止变量。

文[2]证明的一个主要结果如下：

定理 1 设状态空间$S=\{1,2,\cdots,n\}$为有限集，如果线性规划

$$\begin{cases} \min \sum_{i=1}^{n} x_i \\ x_i - \sum_{j=1}^{n} x_j p_{ij} \geqslant f(i),\ i=1,\ 2,\ \cdots,\ n \\ x_i \geqslant g(i),\ i=1,\ 2,\ \cdots,\ n \end{cases} \tag{2}$$

有最优解x_i^0，$i\in S$，则上述马尔可夫链的最优停止问题的值函数存在，且值函数$v(i)=x_i^0$，$i\in S$，相应的最优停止变量$\tau=\inf\{k\geqslant 0;\ x_{z_k}^0=g(z_k)\}$。记$\bar{C}=\{i;\ x_i^0>g(i),\ i\in S\}$，若$\bar{C}=\phi$，则$E_i(\tau)=0$，$i\in S$；若$\bar{C}\neq 0$，则$E_i(\tau)$，$i\in\bar{C}$可以从如下线性方程组解得：

$$x_i - \sum_{j\in\bar{C}} x_j p_{ij} = 1,\ i\in\bar{C}$$

本文将证明此结论反之亦真。从而离散时间、状态有限的马尔可夫链最优停止问题值函数的存在性等价于一个线性规划解的存在性；如果此线性规划有解，相应的停止问题的解的存在性和计算也就解决了。

在$B(S)$上定义一个算子T如下：$\forall u\in B(S)$

$$Tu(i)=\max(g(i), f(i)+\sum_{j\in S} u(j)p_{ij}),\ i\in S$$

对于给定的正整数m，定义

$$v_{m-k}(i)=T^k g(i),\ k=0,1,2,\cdots,m$$

其中$T^k=T(T^{k-1})$，$T^0=$恒等算子，并且定义$v_{m+1}(i)\equiv 0$，显然有

$$Tv_{k+1}=v_k,\ k=0,1,2,\cdots,m-1 \tag{3}$$

首先证明引理。

引理 $$\sup_{\sigma\in\Delta_i^m} E_i\left\{\sum_{k=0}^{\sigma-1} f(z_k)+g(z_\sigma)\right\}=v_0(i),\ i\in S$$

证明

$$\begin{aligned} & E_i\left\{\sum_{k=0}^{m}[v_k(z_k)-E_{z_k}(v_{k+1}(z_{k+1}))]\right\} \\ &=E_i\left\{\sum_{k=0}^{m} v_k(z_k)\right\}-E_i\left\{\sum_{k=0}^{m} E_{z_k}(v_{k+1}(z_{k+1}))\right\} \\ &=v_0(i)+\sum_{k=1}^{m} E_i(v_k(z_k))-\sum_{k=0}^{m} E_i\{E_{z_k}(v_{k+1}(z_{k+1}))\} \\ &=v_0(i)-E_i\{v_{m+1}(z_{m+1})\} \\ &=v_0(i),\ i\in S. \end{aligned} \tag{4}$$

$\forall\sigma\in\Delta_i^m$，

$$\begin{aligned} & E_i\left\{\sum_{k=\sigma}^{m}[v_k(z_k)-E_{z_k}(v_{k+1}(z_{k+1}))]\right\} \\ &=E_i\left\{\sum_{k=\sigma}^{m} v_k(z_k)\right\}-E_i\left\{\sum_{k=\sigma}^{m} E_{z_k}(v_{k+1}(z_{k+1}))\right\} \end{aligned}$$

$$=E_i\left\{\sum_{k=\sigma}^{m}v_k(z_k)\right\}-E_i\left\{E\left[\sum_{k=0}^{m-\sigma}E_{z_{k+\sigma}}(v_{k+\sigma+1}(z_{k+\sigma+1}))\mid\mathscr{F}_\sigma\right]\right\}$$

$$=E_i\left\{\sum_{k=\sigma}^{m}v_k(z_k)\right\}-E_i\left\{\sum_{k=0}^{m-\sigma}E\left[E(v_{k+\sigma+1}(z_{k+\sigma+1})\mid\mathscr{F}_{k+\sigma})\mid\mathscr{F}_\sigma\right]\right\}$$

$$=E_i\left\{\sum_{k=\sigma}^{m}v_k(z_k)\right\}-E_i\left\{\sum_{k=0}^{m-\sigma}E\left[v_{k+\sigma+1}(z_{k+\sigma+1})\mid\mathscr{F}_\sigma\right]\right\}$$

$$=E_i\left\{\sum_{k=\sigma}^{m}v_k(z_k)\right\}-E_i\left\{E\left[\sum_{k=0}^{m-\sigma}v_{k+\sigma+1}(z_{k+\sigma+1})\mid\mathscr{F}_\sigma\right]\right\}$$

$$=E_i\left\{\sum_{k=\sigma}^{m}v_k(z_k)\right\}-E_i\left\{\sum_{k=\sigma}^{m}v_{k+1}(z_{k+1})\right\}$$

$$=E_i(v_\sigma(z_\sigma))-E_i(v_{m+1}(z_{m+1}))$$

$$=E_i(v_\sigma(z_\sigma)),\ i\in S. \tag{5}$$

定义 $\tau=\inf\{k\geqslant 0;\ v_k(z_k)=g(z_k)\}$。

在样本集$(\tau>k)$上，有$v_k(z_k)>g(z_k)$，根据(3)式，在$(\tau>k)$上有

$$v_k(z_k)=\max(g(z_k),\ f(z_k)+\sum_{j\in S}v_{k+1}(j)p_{z_k,j})$$

$$=f(z_k)+\sum_{j\in S}v_{k+1}(j)p_{z_k,j}$$

$$=f(z_k)+E_{z_k}(v_{k+1}(z_{k+1})),\ \text{(a.s成立)}$$

即

$$v_k(z_k)-E_{z_k}(v_{k+1}(z_{k+1}))=f(z_k),\ \text{在}(\tau>k)\text{上a.s成立},$$

从而，由(4)、(5)式，有

$$E_i\left\{\sum_{k=0}^{\tau-1}f(z_k)+g(z_\tau)\right\}=E_i\left\{\sum_{k=0}^{\tau-1}f(z_k)\right\}+E_i\{g(z_\tau)\}$$

$$=E_i\left\{\sum_{k=0}^{\tau-1}[v_k(z_k)-E_{z_k}(v_{k+1}(z_{k+1}))]\right\}+E_i\{v_\tau(z_\tau)\}$$

$$=E_i\left\{\sum_{k=0}^{\tau-1}[v_k(z_k)-E_{z_k}(v_{k+1}(z_{k+1}))]\right\}$$

$$+E_i\left\{\sum_{k=\tau}^{m}[v_k(z_k)-E_{z_k}(v_{k+1}(z_{k+1}))]\right\}$$

$$=E_i\left\{\sum_{k=0}^{m}[v_k(z_k)-E_{z_k}(v_{k+1}(z_{k+1}))]\right\}$$

$$=v_0(i),\ i\in S. \tag{6}$$

另一方面，根据(3)式可知，对每个k有

$$v_k(i)\geqslant f(i)+\sum_{j\in S}v_{k+1}(j)p_{ij},\ i\in S$$

$$v_k(i)\geqslant g(i),\ i\in S$$

所以$\forall\,\sigma\in\Delta_i^m$，

$$E_i\left\{\sum_{k=0}^{\sigma-1}f(z_k)+g(z_\sigma)\right\}$$

$$=E_i\left\{\sum_{k=0}^{\sigma-1}f(z_k)\right\}+E_i\{g(z_\sigma)\}$$

$$
\begin{aligned}
&\leqslant E_i\left\{\sum_{k=0}^{\sigma-1} f(z_k)\right\}+E_i\{v_\sigma(z_\sigma)\}\\
&\leqslant E_i\left\{\sum_{k=0}^{\sigma-1}\left[v_k(z_k)-E_{z_k}(v_{k+1}(z_{k+1}))\right]\right\}\\
&+E_i\left\{\sum_{k=\sigma}^{m}\left[v_k(z_k)-E_{z_k}(v_{k+1}(z_{k+1}))\right]\right\}\\
&=E_i\left\{\sum_{k=0}^{m}\left[v_k(z_k)-E_{z_k}(v_{k+1}(z_{k+1}))\right]\right\}\\
&=v_0(i),\ i\in S
\end{aligned}
\tag{7}
$$

综合(6)，(7)两式，得到

$$
\begin{aligned}
v_0(i)&=E_i\left\{\sum_{k=0}^{\tau-1} f(z_k)+g(z_\tau)\right\}\\
&=\sup_{\sigma\in\Delta_i^m} E_i\left\{\sum_{k=0}^{\sigma-1} f(z_k)+g(z_\sigma)\right\},\ i\in S
\end{aligned}
$$

其中达到上确界是因$\tau\in\Delta_i^m$. （证毕）

空间$B(S)$上的算子T有如下两个简单性质：

性质Ⅰ $\forall u\in B(S)$，$\forall w\in B(S)$且$u\geqslant w$，有$Tu\geqslant Tw$，其中$u\geqslant w$是指$\forall i\in S$，$u(i)\geqslant w(i)$。

性质Ⅱ $\forall u\in B(S)$，$\forall w\in B(S)$，有$\|Tu-Tw\|\leqslant\|u-w\|$

事实上，

$$
\begin{aligned}
Tu(i)&=\max(g(i),\ f(i)+\sum_{j\in S} u(j)p_{ij})\\
&=\max(g(i),\ f(i)+\sum_{j\in S} w(j)p_{ij}+\sum_{j\in S}(u(j)-w(j))p_{ij})
\end{aligned}
\tag{8}
$$

如果$u\geqslant w$，则$\sum_{j\in S}(u(j)-w(j))p_{ij}\geqslant 0$，得到$Tu\geqslant Tw$，性质Ⅰ成立。性质Ⅱ成立是因$u$，$w$任意的，由(8)式得到$Tu\leqslant Tw+\|u-w\|$，从而$\|Tu-Tw\|\leqslant\|u-w\|$。

下面设状态空间S为有限集。设对每个$i\in S$，都有

$$
w_i\triangleq\sup_{\sigma\in\Delta_i} J_i(\sigma)<+\infty
$$

由引理知，对每个正整数m和所有$i\in S$，有

$$
T^m g(i)=v_0(i)=\sup_{\sigma\in\Delta_i^m} J_i(\sigma)\leqslant\sup_{\sigma\in\Delta_i} J_i(\sigma)=w_i
$$

另外显然$Tg\geqslant g$，由算子T的性质Ⅰ递推得到

$$
g\leqslant T^{m-1}g\leqslant T^m g\leqslant w,\ m=1,\ 2,\ \cdots
$$

所以$\{T^m g\}$为不减的有界函数列，从而存在极限，记为$v=\lim_{m\to+\infty} T^m g$，则有

$$
v\geqslant T^{m+1}g=T(T^m g),\quad m=1,2,\cdots \tag{9}
$$

因$\|v-T^m g\|\to 0$ $(m\to+\infty)$，由性质Ⅱ的算子T连续性与(9)式得到

$$
v\geqslant Tv \tag{10}
$$

另外对任何满足$x\geqslant Tx$的x（至少v就是这样一个函数），由性质Ⅰ，$T^{m+1}x\leqslant T^m x\leqslant x$，$m=1,2,\cdots$. 另一方面根据$Tx\geqslant g$得到$T^{m+1}x\geqslant T^m g$，$m=1,2,\cdots$，从而

$$
v=\lim_{m\to\infty} T^m g\leqslant x \tag{11}
$$

(10)和(11)式表明v是所有满足$x\geqslant Tx$的那些元素中的"最小"元，由此得以下定理。

定理 2　设状态空间$S=\{1,2,\cdots,n\}$为有限集，则对每个$i\in S$ $\sup\limits_{\sigma\in\Delta_i}J_i(\sigma)<+\infty$成立的充要条件是线性规划

$$\begin{cases}\min\sum\limits_{i=1}^{n}x_i\\ x_i-\sum\limits_{j=1}^{n}x_jp_{ij}\geqslant f(i),\ i=1,2,\cdots,n\\ x_i\geqslant g(i),\ i=1,2,\cdots,n\end{cases}\tag{12}$$

有最优解，而且最优解$x_i^0=\sup\limits_{\sigma\in\Delta_i}J_i(\sigma)$，$i\in S$.

证明　充分性由定理1即知。必要性是因条件$x\geqslant Tx$等价于线性规划(12)的约束不等式，而v是满足此约束的"最小"元，所以v是线性规划(12)的唯一解。另外因线性规划(12)有解，由文[2]的结论(本文的定理1)可知，最优解$x_i^0=\sup\limits_{\sigma\in\Delta_i}J_i(\sigma)$，$i\in S$.

(证毕)

从定理1和定理2，立即得到：

定理 3　设状态空间$S=\{1,2,\cdots,n\}$为有限集，则所给马尔可夫链最优停止问题的值函数v的存在性等价于线性规划(12)的最优解的存在性。而且值函数v等于线性规划(12)的最优解。

参 考 文 献

[1] Chow Y S, Robbins H, Siegmund D, 何声武，汪振鹏译. 最优停止理论. 上海科技出版社, 1983

[2] 谷建湘. 马尔可夫链最优停止的值函数. 国防科技大学学报, 1990, 12(1):1～11

[3] 王梓坤. 随机过程论. 科学出版社, 1978

A LP Solution to the Problem of the Optimal Stopping Discrete Finite Markov Chains

Gu Jianxiang

(Department of Applied Mathematics and System Engineering)

Abstract

This paper deals with the optimal stopping problem for Markov chains with non—discounted total payment.

Paper [2] proves a sufficient condition:

The value function of the optimal stopping problem for discrete finite Markov chains exists if the corresponding linear programming has a solution. This paper also shows that this condition is also necessary.

Key words: stochastic process, Markov chain, linear programming, stopping variable