研究生教材 研究生教材

最优停止理论及其应用

金治明 著

国家自然科学基金资助课题

最优停止理论及其应用

金治明 著

国防科技大学出版社

内 容 简 介

本书系统地叙述了随机序列、随机过程、马氏过程以及多指标随机过程的最优停止理论，并对若干应用模型，如统计中的序贯方法、股票与投资、战争问题中的最优决策作了介绍。

本书可作为概率论专业的研究生教材，也可供大学高年级学生及应用研究者参考。

最优停止理论及其应用
金治明 著
责任编辑 何 晋
责任校对 曹 红

国防科技大学出版社出版发行
（地址：长沙市砚正街47号 邮编：410073）
新华书店总店科技发行所经销
湖 南 大 学 印 刷 厂 印 装

*

开本：850×1168 1/32 印张：14 字数：351千
1995年8月第1版 1996年1月第2次印刷 印数：1201—2200册

ISBN 7—81024—332—2
O·42 定价：17.00元

序

最优停止理论是概率论中一个具有很强应用背景的领域，它的产生可溯源很久，但 60 年代以来的发展是主要的。Chou Yuan-shih，H. Robbins，D. Sigmund 合著的“Great Expectation：The Theory of optimal stopping”与前苏联学者 A. N. Shirayayev 的“Optimal stopping rules”是该领域的两本专著。前者主要讲离散时间的最优停止，后者主要讨论马氏过程的最优停止。十几年来，我们一直以这两本著作为主要的材料为硕士研究生开设了这一方向的课程，同时开展了最优停止理论及其应用的研究工作。本书包括了我们的一部分研究成果，也概括了前两本书的主要内容。

第一章讨论有限情形，特别是各类秘书问题。这节也可供大学本科生阅读，可以作为指导学士论文的参考。

第二章一般理论，主要讨论离散情形。与文献[1]相比，我们增加了最大、最小最优停时，最优停时的唯一性，Rasche 方法，约束最优停时与多目标最优停止等内容。

第三、五章是关于马尔可夫序列与马尔可夫过程的最优停止。

第四章讨论连续时间的最优停止，特别是停时类的紧性与 BC 拓扑。

第六章讨论多指标最优停止，研究了两指标最优停点的存在性与构造。

最优停止理论这一领域的丰富内容决非一本书所能包容的，如果它能对有兴趣的研究者有所帮助，那就达到了我们的愿望。

本书的编写与我们的讨论班及硕士点的工作是紧密相联的，所以我要感谢参加讨论班与硕士点工作的同志，特别要感谢罗建书、易东云两位同志为第六章提供了初稿，感谢国防科技大学研究

生院及系统工程与数学系为本书的出版给予的支持。

本书的出版得到了国家自然科学基金的资助，特此鸣谢！

由于水平与学识所限，错误及不妥之处，敬请读者批评指正。

作　者

1993.6月于长沙

符号说明表

$\mathscr{R}$——实数集

$\mathscr{R}_+$——非负实数集

$\mathscr{N}$——自然数集

$\mathscr{N}_+$——非负整数集

$\mathscr{R}_+^2=\mathscr{R}_+\times\mathscr{R}_+$

$\mathscr{R}^2=\mathscr{R}\times\mathscr{R}$

$\mathscr{N}_+^2=\mathscr{N}_+\times\mathscr{N}_+$

$\overline{\mathscr{R}},\overline{\mathscr{R}}_+,\overline{\mathscr{N}}_+$——分别为 $\mathscr{R},\mathscr{R}_+,\mathscr{N}_+$ 与 $\{\infty\}$ 之并

$\overline{\mathscr{R}}_+^2,\overline{\mathscr{N}}_+^2$——分别为 $\mathscr{R}_+^2,\mathscr{N}_+^2$ 与 $\{\infty,\infty\}$ 之并

$\mathscr{B}$ 或 $\mathscr{B}(\mathscr{R})$——$\mathscr{R}$ 上 Borel 集

$(\Omega,\mathscr{F},P)$——完备的概率空间

$(\mathscr{F}_t)_{t\in\mathscr{R}_+}$ 或 $(\mathscr{F}_n)_{n\in\mathscr{N}}$——$\mathscr{F}$ 的子 σ 代数族，满足通常条件

E——对 $\mathrm{d}P$ 的数学期望

$\int_A X_n$——表示 $\int_A X_n\mathrm{d}P$

$\mathbf{E}$——对 $\mathrm{d}P\times\mathrm{d}\lambda$ 的数学期望

$E(\cdot\,|\mathscr{F}_n)$ 或 $E_{\mathscr{F}_n}(\cdot)$——对 $\mathscr{F}_n$ 的条件期望

$L^p(\mathscr{F})$ 或 $L^p(\Omega)$——Ω 上 P 方可积的随机变量全体

x^+,x^-——x 的正部，x 的负部

a,s——几乎处处

$\xrightarrow{P}$ 依概率收敛

$\xrightarrow{L^p}$ P 方平均收敛

$C([0,\infty])$——$[0,\infty]$ 上连续函数全体

$C_B([0,\infty])$——$[0,\infty]$ 上有界连续函数全体

$\mathscr{B}_B([0,\infty])$——$[0,\infty]$ 上有界 Borel 可测函数全体

$C(B\times B)$——$B\times B$ 上有界连续函数全体

supp——函数的支撑

sup——上确界

inf——下确界

esssup——本质上确界

[ ]——取整运算

$\mathscr{T}$ $(\overline{\mathscr{T}})$——$(\mathscr{F}_t)$或$(\mathscr{F}_n)$停止规则(停时)全体

$\mathscr{T}_t(\overline{\mathscr{T}}_t)$——$\{\tau\in\mathscr{T}(\overline{\mathscr{T}}):\tau\geqslant t\}$

C $(\overline{C})$——如所论报酬为(Z_t),C,$\overline{C}=\{\tau\in\mathscr{T}(\overline{\mathscr{T}}):EX_\tau^-<\infty\}$

$\overline{C}(g)$——如所论马氏过程为(X_t),报酬函数为 g,则 $\overline{C}(g)=\{\tau\in\mathscr{T}:E_xg(X_\tau)^-<\infty\}$

$C(g)=C\bigcap\overline{C}(g)$

$C_t=C\bigcap T_t$

$\overline{C}_t=\overline{C}\bigcap\mathscr{T}_t$

$\mathscr{T}_B$——有界停时全体

$\vee$——取大运算

$\wedge$——取小运算

$\uparrow$——单调不减

$\gamma_t,\gamma_n,\overline{\gamma_t},\overline{\gamma_n},\gamma_n(x),\overline{\gamma_n}(x)$——Snell 包或 Snell 包函数

Γ——随机化停时全体

Σ——a,s 有限停点全体

$\Sigma_t=\{\sigma\in\Sigma:\sigma\geqslant t\}$

$\mathscr{Q}$——策略集

$\mathscr{Q}_z$——从 z 出发的策略全体

$\mathscr{Q}_z^*$——$\{T\in\mathscr{Q}_z:T=((V_t),\tau),\tau<\infty\quad a.s.\}$

$\mathscr{Q}_z^N$——有限步停止的策略

Θ——全体随机化策略

$\mathscr{S}$——全体随机化可选增道路

$D(z)$——z之后继

目　　录

引　论

我们常常面临各种各样的决策，决策理论的主要任务是告诉决策人，采取什么样的策略将会得到最大的利益。有许多决策都可以简单地归结为是继续（做某件事）或是停止？所谓最优停止理论就是研究何时停止最为有利的数学理论。当然一件事的停止往往是另一件事情的开始。比如战争的停止就是和平的开始，对甲工程投资的停止就可能是对某乙工程投资的开始。在这个意义上，最优停止理论也就是最优开始理论。

赌博是生活中的丑恶现象，但我们却常常以此为例，因为它确是一个很直观的随机现象的模型，许多经济活动也都带有"机会"的色彩。

一个赌徒在进行了若干次赌博之后（他总是有输有赢），很关心如何决定他的策略，是继续赌下去有利还是停止赌博为好？

一个统计试验工作者常常要问：虽然继续试验可以获得更多的信息，但每做一次试验总要耗费人力与物力，因此在做了若干次试验之后是继续做试验还是停止呢？

甲、乙两个军事集团进行了若干战斗之后（此时可能各有胜负），从某方利益而言，自然要提出是停战还是继续打下去的问题。

一个投资者在向甲企业投资了若干年之后，观察企业的效益与整个经济形势自然也会提出是停止投资（即转变投资对象）还是继续投资更为有利的问题。

如此等等，我们可以举出更多的实例。本书并不是直接来解决

上述种种现实的问题，但它们的数学抽象却为人们指出了解决这些问题的途径。

让我们从几个简单的例子开始，说明什么是最优停止理论。

例 1（加倍或输光） 重复掷一均匀的硬币。每掷一次我们都可以停止或再掷一次，每一步所做的决断取决于迄今为止所掷出的结果。假定我们考虑的是有限次总要结束的游戏。若掷 n 次后停止，假设所得的报酬是

$$X_n = \frac{2n}{(n+1)}\prod_{i=1}^{n}(y_i+1) \tag{1}$$

其中，$y_i=1$ 表示第 i 次掷出正面，否则 $y_i=-1$. 于是当前 n 次都掷出正面的报酬是 $2^{n+1}\cdot n/(n+1)$，否则（即只要在前 n 次中有一次出现反面），$X_n=0$. 一般说来，停止在第 n 步的报酬 $X_n=f(y_1,\cdots,y_n)$，其中 f 是某一函数。

一个停时 t 是一个随机变量。在例 1 的情形 $t<\infty$. 一般地，t 取值于 $\{1,2,3,\cdots,+\infty\}$，并且 $\{t=n\}\in\sigma(y_1,y_2,\cdots,y_n)$. 这表明何时停止取决于迄今为止的观察结果（或称为信息）。称

$$\bar{V} = \sup EX_t$$

为报酬序列 $\{X_n\}$ 的值，这里上确界是对一切使得 EX_t 有意义的停时 t 来取的。而 E 表示数学期望，对于随机现象来说，$X_t(\omega)$ 最大是没有实际意义的，我们只能讨论在概率平均的意义上（即数学期望）的最大值问题，并且总假定 $\forall\ n, EX_n$ 有限。

记 $\bar{C}=\{t$ 为停时：$EX_t^-<+\infty\}$，容易证明

$$\bar{V} = \sup_{t\in\bar{C}} EX_t$$

事实上，$\bar{V}\geqslant\sup\limits_{t\in\bar{C}} EX_t$ 是显然的，设 t 是一个停时，且 EX_t 存在。如果 $EX_t^-=+\infty$，则 $EX_t=-\infty$，而 $EX_1>-\infty$. 于是 $\sup\limits_{t\in\bar{C}} EX_t\geqslant EX_1>EX_t$. 因此，$\bar{V}=\sup\limits_{t\in\bar{C}} EX_t$.

如果存在 $\sigma\in\bar{C}$，使得 $EX_\sigma=\bar{V}$，则称 σ 为最优停时。

记 $C=\overline{C}\cap\{t$ 为停时，且 $t<\infty \quad a.s.\}$，称 C 中的停时为停止规则（简称为规则），并记

$$V=\sup_{t\in C}EX_t$$

如果存在 $\sigma\in C$，使 $EX_\sigma=V$，则称 σ 为最优停止规则（简称为最优规则）。

现在继续讨论例 1，令 $C^k=\{t\in C;t\leqslant k\}$，则 $C^k\uparrow C$，显然

$$V\geqslant\lim_{k\to\infty}\sup_{t\in C^k}EX_t$$

又因为 $t<\infty \quad a.s, \lim\limits_{k\to\infty}X_{t\wedge k}=X_t$. 由 Fatou 引理易知

$$EX_t=E\lim_{k\to\infty}X_{t\wedge k}\leqslant\varliminf_{k\to\infty}EX_{t\wedge k}\leqslant\lim_{k\to\infty}\sup_{t\in C^k}EX_t$$

所以，$V=\lim\limits_{k\to\infty}\sup\limits_{t\in C^k}EX_t$. 我们还可证明：$\forall\ t\in C^k, EX_t\leqslant EX_k$. 事实上，在 $\mathscr{F}\triangleq\sigma(y_1,\cdots,y_n)$ 的原子 $A=\{\omega;y_1(\omega)=\cdots=y_n(\omega)=1\}$ 上，$X_n=n\cdot 2^{n+1}/(n+1)$，因此当 $w\in A$ 时

$$E(X_{n+1}|\mathscr{F}_n)(\omega)=\frac{1}{2}\cdot(n+1)\cdot 2^{n+2}/(n+2)>X_n(\omega) \qquad (2)$$

而在 $\{y_j(\omega)=-1\}$ 上，$X_j=X_{j+1}=\cdots=X_{n+1}=0$，于是总有 $E(X_{n+1}|\mathscr{F}_n)\geqslant X_n$. 这样对一切 $t\in c^k$

$$EX_t=\sum_{i=1}^{k}\int_{[t=i]}X_i\leqslant\sum_{i=1}^{k}\int_{[t=i]}X_{i+1}=EX_{t+1}\cdots\leqslant EX_k$$

（$\int_{[t=i]}X_i$ 表示 $\int_{[t=i]}X_i\mathrm{d}p$，省掉 $\mathrm{d}p$ 的记号，余同。）从而

$$\begin{aligned}V&=\sup_{t\in C}EX_t\\&=\lim_{k\to\infty}\sup_{t\in C^k}EX_t\\&=\lim_{k\to\infty}EX_k\\&=\lim_{k\to\infty}\frac{1}{2^k}\cdot\frac{k2^{k+1}}{k+1}+(1-\frac{1}{2^k})\cdot 0\\&=2\end{aligned}$$

但是本例并不存在最优规则。事实上，$\forall\ t\in C$，由于 $E(X_{n+1}|\mathscr{F}_n)\geqslant$

X_n，不难知道

$$EX_t = \sum_{k=1}^{\infty}\int_{[t=k]} X_k = \sum_{k=1}^{\infty}\int_{[t=k,y_1=\cdots=y_k=1]} X_k + \sum_{k=1}^{\infty}\int_{[t=k,\exists\ i\leqslant k,\text{使}y_i=-1]} X_k$$

$$< \sum_{k=1}^{\infty}\int_{[t=k,y_1=\cdots=y_k=1]} X_{k+1} + \sum_{k=1}^{\infty}\int_{[t=k,\exists\ i\leqslant k\text{使}y_i=-1]} X_{k+1}$$

$$= EX_{t+1}$$

所以任何 $t\in C$ 都不是最优的。

顺便指出，(2)式表明在任何时刻 n，如一直只有正面出现，那么再掷一次一定会获得更大的期望报酬。所以在全是正面时停止总是“愚蠢”的。但是如果我们不在某一步“愚蠢”地停下来，也即在一直出现正面时一直赌下去，那就必然会出现一次反面，而最后的报酬永远是零。所以每一步似乎是“聪明”的策略将导致最坏的结局。这是一个富有启发性的例子。由例 1 可见：

(1) 最优规则可能不存在，何时存在呢？如果存在，怎么表达呢？

(2) 既然每一步的“聪明”策略都可能是拙劣的，但一列好的策略的极限形式也可能是拙劣的，那么在什么条件下不会发生这种情形呢？

显然，最优规则是否存在与所给定问题的概率结构有关，也与报酬函数的给定有关。

例 2(报酬等于平均数)　诸 y_n 同例 1，设报酬序列为

$$X_n = \frac{1}{n}(y_1 + y_2 + \cdots + y_n) \tag{3}$$

先考虑特殊的停止规则

$$t=\begin{cases}1, y_1=1\\ n, y_1=-1\text{ 且 } n \text{ 是第一个整数，使 } y_1+\cdots+y_n=0\end{cases} \tag{4}$$

为了证明 t 是停止规则，即 $P(t<\infty)=1$，我们有

命题 1　设诸 y_n 如例 1，$S_n=\sum_{i=1}^{n} y_i$，$S_0=1$，则

$$P(\limsup_{n\to\infty} S_n = +\infty) = P(\liminf_{n\to\infty} S_n = -\infty) = 1$$

$$P(S_n \text{取到每个整数值无穷次}) = 1$$

证明 记 $q_j = P(\sup_{n\geq 0} S_n \geq j)$，对 $j\geq 1$

$$q_j = P\left(y_1 = 1, \sup_{n\geq 1}\sum_{i=2}^{n} y_i \geq j-1\right)$$

$$+ P\left(y_1 = -1, \sup_{n\geq 1}\sum_{i=1}^{n} y_i \geq j+1\right)$$

$$= \frac{1}{2}P(\sup_{n\geq 0} S_n \geq j-1) + \frac{1}{2}P(\sup_{n\geq 0} S_n \geq j+1)$$

$$= \frac{1}{2}(q_{j-1} + q_{j+1})$$

所以 $q_j = Cj + q_0$，因为 $q_j \leqslant 1$，所以 $C=0$. 于是 $q_j = q_0 = 1$，即 $P(\sup_{n\geq 1} S_n \geqslant j) = 1$，对任何 $j\geqslant 1$ 成立。因而

$$P(\limsup_{n\to\infty} S_n = +\infty)$$

$$= P(\operatorname{Sup}_{n\geq 1} S_n = +\infty)$$

$$= P(\bigcap_{j=1}^{\infty}\{\sup_{n\geq 1} S_n \geq j\}) = 1$$

由对称性

$$P(\liminf_{n\to\infty} S_n = -\infty) = 1$$

从而 $P(S_n$ 取到每个整数值无穷次$)=1$.

由命题 1，对几乎所有的 ω，总存在(S_n)的子列趋于$\pm\infty$. 因此不论 $y_1=1$ 或-1，总存在有限的 n，使 $S_n=0$. 于是 $P(t<\infty)=1$.

对(4)式定义的 t，$EX_t = \frac{1}{2}\cdot\frac{1}{1}+\frac{1}{2}\cdot 0 = \frac{1}{2}$. 从而 $V\geqslant\frac{1}{2}$，但 $X_n\leqslant 1$，故 $V\leqslant 1$. 但是我们却找不到一个使 $EX_t > 0.9$ 的例子，也难以证明 $V\leqslant 0.9$，更不用说去找到 V 的精确值和决定最优规则了。对这个问题的研究已有一系列论文，读者可参考文献[5,6]。

如果限于在$\{1,2,\cdots,N\}$中考虑最优规则，情形就不同了。因

为对于有限向量$(y_1,y_2,\cdots,y_N)$，$y_i=\pm1$，Ω 可以划分为 2^N 个原子。在每个原子上 t 至多取 N 个不同的值，所以一切可能的停止规则至多有 $N\cdot2^N$ 种不同的取值，从理论上来说最优规则一定可以找到。但其实这是办不到的。比如取 $N=100$，则 $N\cdot2^N\approx10^{32}$. 如用亿次机进行计算，即使把算每个 EX_t 算做一次运算，也得花 3000 亿年！虽然如此，但先限于 N 的思想是可取的。如果限于 N，求出最优规则 σ_N 及 V^N，再令 $N\to\infty$，考虑其极限。因为

$$V^N=\sup_{t\in c^N}EX_t$$

是 N 的单调非降函数，所以 $V'=\lim\limits_{N\to\infty}V^N$ 是存在的，但是 $V=V'$ 吗？即使 $V=V'$，那么最优规则 σ 是否存在，且 $\sigma=\lim\limits_{n\to\infty}\sigma_N$ 吗？如果 $V\neq V'$，能否找到有效的方法把 V 介于两个彼此接近的可计算的上下界之间呢？

下面的例子说明了 $V\neq V'$.

例 3 诸 y_n 如例 1，设

$$X_n=\min(1,y_1+\cdots+y_n)-\frac{n}{n+1},\quad n=1,3,\cdots$$

考虑特殊的停止规则

$$t=\inf\{n\geqslant1,:y_1+\cdots+y_n=1\}\tag{5}$$

由命题 1 可证 $P(t<\infty)=1$. 且

$$EX_t=1-E\left(\frac{t}{t+1}\right)>0$$

因此 $V>0$. 为了计算 V 的值，我们先求 $P(t=n)$。设想有一质点在 $(-\infty,0]$ 上作随机游动，其中 0 为吸收壁，质点每次移动 1 个单位，且向左或向右移动的概率都是 $\frac{1}{2}$. 于是一个处在 -1 处的质点在第 n 步被 0 点吸收的概率恰为 $P(t=n)$。记

$$u_{z,n}=P(\text{处于 } z \text{ 处的质点在第 } n \text{ 步被吸收的概率})$$

则当 $z\leqslant-1$，$n\geqslant1$ 时，有