

带有无界赔付函数的非零和随机对策折扣模型*

杨 洁^{1,2}, 郭先平²

(1. 北京信息科技大学理学院, 北京 100192;
2. 中山大学数学与计算科学学院, 广东 广州 510275)

摘 要: 讨论了赔付函数可能既无上界又无下界的离散时间可数状态非零和随机对策的折扣模型。在零和随机对策中常用的“漂移”和“连续-紧”性条件下, 用 Fan's 不动点定理证明了 Nash 平衡点的存在性。

关键词: 非零和随机对策; 期望折扣赔付准则; Nash 平衡点; 可数状态空间

中图分类号: O211 **文献标识码:** A **文章编号:** 0529-6579 (2008) 05-0023-06

众所周知, 离散时间非零和随机对策期望折扣赔付准则已得到了广泛研究^[1-13]。其中, 在状态和行动空间均有限的情况下, 文献 [2, 10-11, 13] 证明了 Nash 平衡点的存在性。对状态空间可数但行动空间有限的情形, 文献 [6] (对有界费用) 和文献 [12] (对无界费用) 分析了 Nash 平衡点的存在性。当状态空间可数且行动空间为紧集时, 就我们所知只有文献 [3] (对界费用) 和文献 [1] (对无界费用) 证明了 Nash 平衡点的存在性。对于状态空间更一般 (可能不可数) 的情形, 文献 [4, 8] 在赔付函数和转移概率可分且行动空间有限的条件下证明了 Nash 平衡点的存在性, 而文献 [9] 在转移概率独立于初始状态的假设下也获得了 Nash 平衡点的存在性; 在转移概率可分且赔付函数有界的条件下, 文献 [5] 将文献 [4, 8-9] 中的主要结果推广到行动空间为紧集的情形。

本文研究在可数状态空间、紧的行动集、赔付函数可能既无上界又无下界的情况下, 离散时间非零和随机对策的期望折扣赔付准则 Nash 平衡点的存在性。用零和随机对策研究中常用的“漂移”和“连续-紧”性条件, 我们确立了 Nash 平衡点所满足的折扣最优方程, 并进一步证明了 Nash 平衡点的存在性。我们的假设要弱于文献 [1] 的假设 1 (即转移概率是 u -连续的), 而且我们的证明方法也与文献 [1] 中的证明不同。此外, 本文的赔付函数是无界的且转移概率不需要满足可分性条件, 因而文献 [1, 4, 8-9] 中的结果并不能适用于本文的模型。

1 对策模型

若 X 是 Borel 空间 (即完备可分距离空间的 Borel 子集), 则用 $B(X)$ 表示其 Borel σ -代数, $P(X)$ 表示 X 上的概率测度空间, 并赋予弱收敛拓扑。

引进离散时间二人非零和随机对策模型:

$$\{S, (A(i) \subset A, B(i) \subset B, i \in S), q(\cdot | i, a, b), r_k(i, a, b) (k = 1, 2)\}$$

其中, S 是可数状态空间, A 和 B 分别表示局中人甲和乙的行动空间, $A(i)$ 和 $B(i)$ 分别表示系统处于状态 $i \in S$ 时局中人甲和乙的允许行动集, 假设它们是 Borel 空间。显然, 集合

$$K := \{(i, a, b) : i \in S, a \in A(i), b \in B(i)\} \quad (1)$$

是 $S \times A \times B$ 的 Borel 子集。 $q(\cdot | i, a, b)$ 表示关于 $(i, a, b) \in K$ 的转移概率, 是给定 K 在 S 上的随机核。最后, $r_k(i, a, b) (k = 1, 2)$ 表示关于 $(i, a, b) \in K$ 上局中人甲、乙的实值赔付函数。

为介绍本文所关心的期望折扣赔付准则, 我们需要介绍策略类的定义。

定义 1 随机核序列 $\pi^1 := (\pi_t^1, t = 0, 1, \dots)$ 称为局中人甲的随机马氏策略, 若满足

$$\pi_t^1(A(i) | i) = 1, \quad \forall i \in S, t \geq 0$$

我们用 Π_1 表示局中人甲的随机马氏策略类。称随机马氏策略 $\pi^1 := (\pi_t^1, t \geq 0) \in \Pi_1$ 为平稳策略, 如果对任意 $i \in S$, 存在概率测度 $\pi^1(\cdot | i) \in P(A(i))$, 使得

$$\pi_t^1(\cdot | i) = \pi^1(\cdot | i), \quad \forall t \geq 0$$

用 Π_1^s 表示局中人甲的平稳策略类。用 $B(i)$

* 收稿日期: 2008-03-12

基金项目: 国家自然科学基金资助项目 (60574002)

作者简介: 杨洁 (1979 年生), 女, 博士; E-mail: zezhongy@163.com

代替 $A(i)$, 我们类似定义局中人乙的随机马氏策略类 Π_2 和平稳策略类 Π_2^S 。

对任意策略对 $(\pi^1, \pi^2) := ((\pi_i^1, \pi_i^2), t \geq 0) \in \Pi_1 \times \Pi_2$, $i, j \in S$ 和 $t \geq 0$, 我们定义

$$q(j|i, \pi_i^1, \pi_i^2) :=$$

$$\int_B \int_A q(j|i, a, b) \pi_i^1(da|i) \pi_i^2(db|i) \quad (2)$$

$$r_k(i, \pi_i^1, \pi_i^2) = \int_B \int_A r_k(i, a, b) \pi_i^1(da|i) \pi_i^2(db|i) \quad (3)$$

其中 $k=1, 2$ 。特别地, 对 $\pi^k \in \Pi_k^S$, $q(j|i), \pi_i^1, \pi_i^2$ 和 $r_k(i, \pi_i^1, \pi_i^2)$ 分别记为 $q(j|i), \pi^1, \pi^2$ 和 $r_k(i, \pi^1, \pi^2)$ 。

对任意策略对 $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$ 和初始状态 $i \in S$, 由著名的 Tulcea's 定理, 存在唯一的概率空间 $(\Omega, F, P_i^{\pi^1, \pi^2})$ 和定义在 Ω 上的随机过程 $\{(i_t, a_t, b_t), t \geq 0\}$ 使得对任意 $j \in S$ 和 $t \geq 0$, 有

$$P_i^{\pi^1, \pi^2}(i_{t+1} = j | h_t, a_t, b_t) = q(j|i_t, a_t, b_t),$$

$$\forall h_t = (i, a_0, b_0, \dots, i_{t-1}, a_{t-1}, b_{t-1}, i_t) \in H_t \quad (4)$$

这里 $H_0 = S, H_t := K \times H_{t-1} (t \geq 1)$, i_t, a_t 和 b_t 分别表示 $t \geq 0$ 时刻的状态变量和局中人甲和乙的行动变量。关于测度 $P_i^{\pi^1, \pi^2}$ 的期望算子记为 $E_i^{\pi^1, \pi^2}$ 。

下面我们定义期望折扣赔付准则和它的 Nash 平衡点。

对任意策略对 $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$, $i \in S$ 和折扣因子 $\alpha \in (0, 1)$, 局中人甲, 乙的期望折扣准则 $V_\alpha^k(i, \pi^1, \pi^2) (k=1, 2)$ 定义为:

$$V_\alpha^k(i, \pi^1, \pi^2) := \sum_{t=0}^{\infty} \alpha^t E_i^{\pi^1, \pi^2} [r_k(i_t, a_t, b_t)] =$$

$$\sum_{t=0}^{\infty} \alpha^t \sum_{j \in S} q'(j|i, \pi^1, \pi^2) r_k(j, \pi_i^1, \pi_i^2) \quad (5)$$

其中 $q'(j|i, \pi^1, \pi^2) := P_i^{\pi^1, \pi^2}\{i_t = j\}$ 。

定义 2 策略对 $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$ 称为期望折扣准则的 Nash 平衡点, 如果

$$V_\alpha^1(i, \pi^1, \pi^2) \geq V_\alpha^1(i, \pi^1, \pi^{*2}) \quad (6)$$

且

$$V_\alpha^2(i, \pi^{*1}, \pi^2) \geq V_\alpha^2(i, \pi^{*1}, \pi^2) \quad (7)$$

对所有 $i \in S, \pi^1 \in \Pi_1, \pi^2 \in \Pi_2$ 成立。

本文的主要目的是寻找期望折扣赔付模型 Nash 平衡点存在的新条件。

2 主要结论

为保证 Nash 平衡点的存在性, 我们需要下面的假设。由于赔付函数 $r_k(i, a, b) (k=1, 2)$ 可能无界, 为保证 $V_\alpha^k(i, \pi^1, \pi^2) (k=1, 2)$ 的有限性, 我们首先引进下面 (8) 式的“漂移 (期望增长)”条件。

假设 1 (1) 存在常数 $\eta \geq 0, 0 \leq \lambda < 1$ 和 S

上的可测函数 $w \geq 1$ 使得

$$\sum_{j \in S} q(j|i, a, b) w(j) \leq \lambda w(i) + \eta, \quad \forall (i, a, b) \in K \quad (8)$$

(2) 存在常数 $M > 0$ 使得 $|r_k(i, a, b)| \leq Mw(i)$ 对所有 $(i, a, b) \in K$ 和 $k=1, 2$ 成立。

注 1 如果赔付函数 $r_k(i, a, b)$ 一致有界 (即 $\sup_{(i, a, b) \in K, k=1, 2} |r_k(i, a, b)| < \infty$), 假设 1 (2) 显然成立。

为得到 Nash 平衡点的存在性, 我们也需要“连续-紧”性假设^[14]。

假设 2

(1) 对任意 $i \in S$, $A(i)$ 和 $B(i)$ 是紧集。

(2) 对任意 $i \in S$ 和 $k=1, 2$, 赔付函数 $r_k(i, a, b)$ 关于 $(a, b) \in A(i) \times B(i)$ 连续。

(3) 对任意固定 $i \in S$ 和 S 上任意有界可测函数 u 及 $u = w$, $\sum_{j \in S} u(j) q(j|i, a, b)$ 是 $A(i) \times B(i)$ 上的连续函数, 其中 w 是假设 1 中的函数。

下面给出用来证明主要结论 (定理 1) 的一些引理。

对假设 1 中函数 $w \geq 1$, 我们定义 S 上实值函数 u 的加权模上确界范数 $\|u\|_w$ 为

$$\|u\|_w := \sup_{i \in S} [w(i)^{-1} |u(i)|]$$

和 Banach 空间 $L_w^\infty := \{u: \|u\|_w < \infty\}$ 。

对任意 $i, j \in S, \rho \in P(B(i))$ 和平稳策略 $\pi^k = (\pi^k(\cdot|i), i \in S) \in \Pi_k^S, k=1, 2$, 令

$$q(j|i, \pi^1, \rho) := \int_{B(i)} \int_{A(i)} q(j|i, a, b) \pi^1(da|i) \rho(db) \quad (9)$$

$$r_1(i, \pi^1, \rho) := \int_{B(i)} \int_{A(i)} r_1(i, a, b) \pi^1(da|i) \rho(db) \quad (10)$$

取 $v \in P(A(i))$ 代替 $\rho \in P(B(i))$, 并类似定义 $q(j|i, v, \pi^2)$ 和 $r_2(i, v, \pi^2)$ 。

引理 1 如果假设 1 成立, 则

$$(a) E_i^{\pi^1, \pi^2} [w(i_t)] \leq \lambda^t w(i) + \frac{1 - \lambda^t}{1 - \lambda} \eta \quad \text{对所有}$$

$t \geq 0, i \in S$ 和 $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$ 成立。

(b) 对任意 $\alpha \in (0, 1)$ 和 $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$, 存在两个非负常数 M_1 和 M_2 (依赖 α) 使得对任意 $\forall i \in S, k=1, 2$ 成立:

$$|V_\alpha^k(i, \pi^1, \pi^2)| \leq M_1 + M_2 w(i) \quad (11)$$

证明 (a) 和 (b) 的证明见文献 [14] 引

理 1 的证明, 其中 $M_1 = \frac{M\eta}{(1-\lambda)(1-\alpha)}$, $M_2 =$

$$\frac{M}{1-\alpha\lambda}。$$

引理2 在假设1和2下, 下列结论成立:

(a) 对任意策略 $\pi^1 = (\pi_t^1, t \geq 0) \in \Pi_1$ 和 $\pi^2 = (\pi_t^2, t \geq 0) \in \Pi_2$, 如果存在函数 $u \in L_w^\infty$ 对任意 $i \in S, t \geq 0$ 满足

$$u(i) \geq r_k(i, \pi_t^1, \pi_t^2) + \alpha \sum_{j \in S} q(j|i, \pi_t^1, \pi_t^2) u(j) \quad (12)$$

则有 $u(i) \geq V_\alpha^k(i, \pi_t^1, \pi_t^2)$ 对所有 $i \in S$ 和 $k = 1, 2$ 成立。

(b) 对任意策略 $\pi^1 = (\pi_t^1, t \geq 0) \in \Pi_1$ 和 $\pi^2 = (\pi_t^2, t \geq 0) \in \Pi_2$, 如果存在函数 $u \in L_w^\infty$ 对任意 $i \in S, t \geq 0$ 满足

$$u(i) \leq r_k(i, \pi_t^1, \pi_t^2) + \alpha \sum_{j \in S} q(j|i, \pi_t^1, \pi_t^2) u(j)$$

则有 $u(i) \leq V_\alpha^k(i, \pi_t^1, \pi_t^2)$ 对所有 $i \in S$ 和 $k = 1, 2$ 成立。

(c) 对任意平稳策略对 $(\pi^1, \pi^2) \in \Pi_1^S \times \Pi_2^S$ 和 $k = 1, 2$, $V_\alpha^k(i, \pi^1, \pi^2)$ 是下方方程 (13) 式在 L_w^∞ 中的唯一解:

$$u(i) = r_k(i, \pi^1, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi^2) u(j) \quad (13)$$

(d) 对任意平稳策略 $\pi^2 \in \Pi_2^S$, 存在平稳策略 $\pi^{*1} \in \Pi_1^S$ 满足折扣最优方程

$$V_\alpha^1(i, \pi^{*1}, \pi^2) = r_1(i, \pi^{*1}, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^{*1}, \pi^2) V_\alpha^1(j, \pi^{*1}, \pi^2) = \sup_{\pi^1 \in \Pi_1^S} \{r_1(i, \pi^1, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi^2) V_\alpha^1(j, \pi^1, \pi^2)\}, \quad \forall i \in S$$

(e) 对任意平稳策略 $\pi^1 \in \Pi_1^S$, 存在平稳策略 $\pi^{*2} \in \Pi_2^S$ 满足折扣最优方程

$$V_\alpha^2(i, \pi^1, \pi^{*2}) = r_2(i, \pi^1, \pi^{*2}) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi^{*2}) V_\alpha^2(j, \pi^1, \pi^{*2}) = \sup_{\pi^2 \in \Pi_2^S} \{r_2(i, \pi^1, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi^2) V_\alpha^2(j, \pi^1, \pi^2)\} \quad \forall i \in S.$$

(f) 对任意 $i \in S$ 和 $k = 1, 2$, $V_\alpha^k(i, \pi^1, \pi^2)$ 关于 $(\pi^1, \pi^2) \in \Pi_1^S \times \Pi_2^S$ 连续。

证明 (a) 对任意 $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$ 和 $i \in S$, 根据 (12) 式得: 对任意 $i_t \geq 0, t \geq 0, k = 1, 2$, 成立

$$u(i_t) \geq r_k(i_t, \pi_t^1, \pi_t^2) + \alpha \sum_{j \in S} q(j|i_t, \pi_t^1, \pi_t^2) u(j)$$

因此, 由上式及 (2) - (4) 式可得

$$E_i^{\pi^1, \pi^2}[u(i_t)] \geq E_i^{\pi^1, \pi^2}[r_k(i_t, a_t, b_t)] + \alpha E_i^{\pi^1, \pi^2}[u(i_{t+1})] \quad (14)$$

上式两边同乘 α^t , 令 $t = 0, \dots, N$, 将所有式子相加整理有

$$u(i) \geq \sum_{t=0}^N \alpha^t E_i^{\pi^1, \pi^2}[r_k(i_t, a_t, b_t)] + \alpha^{N+1} E_i^{\pi^1, \pi^2}[u(i_{N+1})]$$

在上式中令 $N \rightarrow \infty$, 由引理1(a)和(5)式知引理2(a)成立。

(b) 类似 (a) 的证明可得。

(c) 由引理1(b), 对任意平稳策略 $(\pi^1, \pi^2) \in \Pi_1^S \times \Pi_2^S$, 有 $V_\alpha^k(i, \pi^1, \pi^2) \in L_w^\infty (k = 1, 2)$ 。而且由 (2) - (3) 式和 (5) 式可得

$$r_k(i, \pi^1, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi^2) V_\alpha^k(j, \pi^1, \pi^2) = r_k(i, \pi^1, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi^2) \sum_{l=0}^{\infty} \alpha^l \sum_{l \in S} q'(l|j, \pi^1, \pi^2) r_k(l, \pi^1, \pi^2) =$$

$$r_k(i, \pi^1, \pi^2) + \alpha \sum_{l \in S} \left[\sum_{t=0}^{\infty} \alpha^t \sum_{j \in S} q(j|i, \pi^1, \pi^2) q'(l|j, \pi^1, \pi^2) \right] r_k(l, \pi^1, \pi^2) = r_k(i, \pi^1, \pi^2) + \alpha \sum_{l \in S} \sum_{t=0}^{\infty} \alpha^t q^{t+1}(j|i, \pi^1, \pi^2) r_k(l, \pi^1, \pi^2) =$$

$$r_k(i, \pi^1, \pi^2) + \sum_{t=0}^{\infty} \alpha^{t+1} E_i^{\pi^1, \pi^2}[r_k(i_{t+1}, a_{t+1}, b_{t+1})] =$$

$$r_k(i, \pi^1, \pi^2) + \sum_{t=1}^{\infty} \alpha^t E_i^{\pi^1, \pi^2}[r_k(i_t, a_t, b_t)] =$$

$$V_\alpha^k(i, \pi^1, \pi^2)$$

从而, $V_\alpha^k(i, \pi^1, \pi^2)$ 是 (13) 式的解。再由引理2(a)和(b)知该解是唯一的。

(d) 固定 $\pi^1 \in \Pi_1^S$, 非零和随机对策过程可看作马氏确定过程 (MDP), 由文献 [15] 定理 8.3.6 知引理2(d)成立。

(e) 类似 (d) 的证明可得。

(f) 为证明 (f), 我们只考虑 $k = 1$ 的情况, 对 $k = 2$ 的情况可类似证明。

令 $(\pi_n^1, \pi_n^2) \rightarrow (\pi^1, \pi^2)$, $\limsup_{n \rightarrow \infty} V_\alpha^1(i, \pi_n^1, \pi_n^2) =: v(i) (i \in S)$ 。则存在 $\{(\pi_n^1, \pi_n^2)\}$ 的子列 $\{(\pi_m^1, \pi_m^2)\}$ 对任意 $i \in S$ 满足

$$\lim_{m \rightarrow \infty} V_\alpha^1(i, \pi_m^1, \pi_m^2) = v(i) \quad (15)$$

$$\lim_{m \rightarrow \infty} r_1(i, \pi_m^1, \pi_m^2) = r_1(i, \pi^1, \pi^2)$$

由引理1(b)和(15)式, 可得

$$|v(i)| \leq M_1 + M_2 w(i) \quad \forall i \in S \quad (16)$$

又由引理2(c), 对任意 $i \in S$ 和 $m \geq 1$,

$$V_\alpha^1(i, \pi_m^1, \pi_m^2) = r_1(i, \pi_m^1, \pi_m^2) + \alpha \sum_{j \in S} q(j|i, \pi_m^1, \pi_m^2) V_\alpha^1(j, \pi_m^1, \pi_m^2)$$

因此, 由 (15) - (16) 式, 假设2(1)和文献 [15] 引理 8.3.7, 我们有

$$v(i) = r_1(i, \pi^1, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi^2) v(j)$$

上式结合引理2(c)的“解的唯一性”, 可得

$$\limsup_{n \rightarrow \infty} V_\alpha^1(i, \pi_n^1, \pi_n^2) = v(i) = V_\alpha^1(i, \pi^1, \pi^2) \quad (17)$$

同理可证得

$$\liminf_{n \rightarrow \infty} V_\alpha^1(i, \pi_n^1, \pi_n^2) = v(i) = V_\alpha^1(i, \pi^1, \pi^2) \quad (18)$$

因而, 从 (17) - (18) 式, 我们有

$$\lim_{n \rightarrow \infty} V_{\alpha}^1(i, \pi_n^1, \pi_n^2) = V_{\alpha}^1(i, \pi^1, \pi^2), \forall i \in S$$

故对 $k = 1$ 我们证明了 (f) 成立。

引理 3 在假设 1 和 2 下, 下列结论成立:

(a) 对任意固定 $i \in S$ 和 $k = 1, 2$, 函数 $r_k(i, \pi^1, \pi^2)$ 和 $\sum_{j \in S} q(j|i, \pi^1, \pi^2) V_{\alpha}^k(i, \pi^1, \pi^2)$ 关于 $(\pi^1, \pi^2) \in \Pi_1^S \times \Pi_2^S$ 连续。

(b) 如果实值函数 $f(\pi^1, \pi^2)$ 关于 $(\pi^1, \pi^2) \in \Pi_1^S \times \Pi_2^S$ 连续, 则函数

$$f^*(\pi^1) := \sup_{\pi^2 \in \Pi_2^S} f(\pi^1, \pi^2)$$

关于 $\pi^1 \in \Pi_1^S$ 连续。

证明 (a) 由假设 2 (2) 和 (3), 知对任意 $i, j \in S$, 函数 $r_k(i, \pi^1, \pi^2)$ 和 $q(j|i, \pi^1, \pi^2)$ 连续。因此, 由假设 2 (1), 引理 2 (f) 和文献 [15] 中的引理 8.3.7, 知 (a) 成立。

(b) 因为 S 可数, 由假设 2 (1) 有 $\Pi_2^S = \Pi_{i \in S} P(B(i))$ 是紧集。因而, 对任意 $\pi_1 \in \Pi_1^S, f^*(\pi^1)$ 有定义。同样由假设 2 (1), $\Pi_1^S \times \Pi_2^S$ 也是紧集。因此 f 在紧的距离空间上一致连续, 故 (b) 显然成立。

最后, 我们回忆 Fan's 不动点定理。

引理 4 (Fan's 不动点定理) 设 X 是局部凸拓扑线性空间, Y 是 X 上的紧凸集, $F(Y)$ 是 Y 上非空闭凸子集族, 则对任意上半连续集值映射 $f: Y \rightarrow F(Y)$, 存在一点 $y_0 \in Y$, 使得 $y_0 \in f(y_0)$ 。

证明 见文 [16]。

注 2 紧集 Π_1^S 和 Π_2^S 显然是凸集, 而且由 (2) - (3) 式可得函数

$$\pi^1 \mapsto r_k(i, \pi^1, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi^2) u(j) \quad (19)$$

对于任意固定 $\pi^2 \in \Pi_2^S, u \in L_w^{\infty}, i \in S$ 和 $k = 1, 2$, 是 $\pi^1 \in \Pi_1^S$ 上的凸函数 (如果固定 $\pi^1 \in \Pi_1^S$ 考虑 (19) 式作为 $\pi^2 \in \Pi_2^S$ 的函数凸性亦成立)。

定理 1 在假设 1 和 2 下, 下列结论成立:

(a) 存在策略对 $(\pi^*, \pi^{*2}) \in \Pi_1^S \times \Pi_2^S$ 满足折扣最优方程

$$V_{\alpha}^1(i, \pi^*, \pi^{*2}) = \sup_{\pi^1 \in \Pi_1^S} \{r_1(i, \pi^1, \pi^{*2}) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi^{*2}) V_{\alpha}^1(j, \pi^1, \pi^{*2})\} \quad \forall i \in S \quad (20)$$

$$V_{\alpha}^2(i, \pi^*, \pi^{*2}) = \sup_{\pi^2 \in \Pi_2^S} \{r_2(i, \pi^*, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^*, \pi^2) V_{\alpha}^2(j, \pi^*, \pi^2)\} \quad \forall i \in S \quad (21)$$

(b) 策略对 (π^*, π^{*2}) 是一个 Nash 平衡点。而且 (π^*, π^{*2}) 是一个 Nash 平衡点当且仅当它满足 (20) 和 (21) 式。

证明 对任意固定平稳策略 $\pi^2 \in \Pi_2^S$, 由引理 2 (d) 知存在一个策略 $\pi^{*1} \in \Pi_1^S$ 使得对所有 $i \in S$ 成立

$$V_{\alpha}^1(i, \pi^{*1}, \pi^2) = r_1(i, \pi^{*1}, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^{*1}, \pi^2) V_{\alpha}^1(j, \pi^{*1}, \pi^2) = \sup_{\pi^1 \in \Pi_1^S} \{r_1(i, \pi^1, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi^2) V_{\alpha}^1(j, \pi^1, \pi^2)\}$$

因而, 集合

$$f(\pi^2) := \{\pi^{*1} \in \Pi_1^S : r_1(i, \pi^{*1}, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^{*1}, \pi^2) V_{\alpha}^1(j, \pi^{*1}, \pi^2) = \sup_{\pi^1 \in \Pi_1^S} \{r_1(i, \pi^1, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi^2) V_{\alpha}^1(j, \pi^1, \pi^2)\} \quad \forall i \in S\} \quad (22)$$

非空。对任意给定 $\pi^1 \in \Pi_1^S$, 由引理 2 (e) 有集合

$$f(\pi^1) := \{\pi^{*2} \in \Pi_2^S : r_2(i, \pi^1, \pi^{*2}) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi^{*2}) V_{\alpha}^2(j, \pi^1, \pi^{*2}) = \sup_{\pi^2 \in \Pi_2^S} \{r_2(i, \pi^1, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi^2) V_{\alpha}^2(j, \pi^1, \pi^2)\} \quad \forall i \in S\} \quad (23)$$

非空。从而由引理 3 (b) 知 $f(\pi^1)$ 和 $f(\pi^2)$ 是紧集。而且, 由引理 2 (c) - (e) 和注 2 有 $f(\pi^1)$ 和 $f(\pi^2)$ 是凸集。

我们定义集值映射 $\Phi: \Pi_1^S \times \Pi_2^S \rightarrow 2^{\Pi_1^S \times \Pi_2^S}$, $(\pi^1, \pi^2) \rightarrow f(\pi^2) \times f(\pi^1)$ 。对所有 $n \geq 1$, 令 (π_n^1, π_n^2) 和 (π_n^{*1}, π_n^{*2}) 满足

$$(i) (\pi_n^1, \pi_n^2), (\pi_n^{*1}, \pi_n^{*2}) \in \Pi_1^S \times \Pi_2^S;$$

$$(ii) \lim_{n \rightarrow \infty} (\pi_n^1, \pi_n^2) = (\pi^1, \pi^2)$$

$$\text{和} \lim_{n \rightarrow \infty} (\pi_n^{*1}, \pi_n^{*2}) = (\pi^{*1}, \pi^{*2});$$

$$(iii) (\pi_n^{*1}, \pi_n^{*2}) \in \Phi(\pi_n^1, \pi_n^2)。$$

我们希望证明 $(\pi^{*1}, \pi^{*2}) \in \Phi(\pi^1, \pi^2)$ 。为此对所有 $n \geq 1$, 由 (iii) 可得

$$r_1(i, \pi_n^{*1}, \pi_n^2) + \alpha \sum_{j \in S} q(j|i, \pi_n^{*1}, \pi_n^2) V_{\alpha}^1(j, \pi_n^{*1}, \pi_n^2) = \sup_{\pi^1 \in \Pi_1^S} \{r_1(i, \pi^1, \pi_n^2) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi_n^2) V_{\alpha}^1(j, \pi^1, \pi_n^2)\} \quad (24)$$

令 (24) 式中 $n \rightarrow \infty$, 由 (ii) 和引理 3, 有

$$r_1(i, \pi^{*1}, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^{*1}, \pi^2) V_{\alpha}^1(j, \pi^{*1}, \pi^2) = \sup_{\pi^1 \in \Pi_1^S} \{r_1(i, \pi^1, \pi^2) + \alpha \sum_{j \in S} q(j|i, \pi^1, \pi^2) V_{\alpha}^1(j, \pi^1, \pi^2)\}$$

这就给出了 $\pi^{*1} \in f(\pi^2)$ 。同理可证 $\pi^{*2} \in f(\pi^1)$, 因此有 $(\pi^{*1}, \pi^{*2}) \in f(\pi^2) \times f(\pi^1)$ 。即映射 Φ 是上半连续的。

另一方面, 由引理 4 知存在策略对 $(\pi^*, \pi^{*2}) \in \Pi_1^S \times \Pi_2^S$, 满足 $(\pi^*, \pi^{*2}) \in f(\pi^{*2}) \times f(\pi^*)$, 结合 (22) - (23) 式和引理 2 (c) 我们就有对所有 $i \in S$

$$V_{\alpha}^1(i, \pi^*, \pi^{*2}) = \sup_{\pi^1 \in \Pi_1^S} \{r_1(i, \pi^1, \pi^{*2}) +$$

$$\alpha \sum_{j \in S} q(j|i, \pi^1, \pi^{*2}) V_{\alpha}^1(j, \pi^{*1}, \pi^{*2}) \} \quad (25)$$

和 $V_{\alpha}^2(i, \pi^{*1}, \pi^{*2}) = \sup_{\pi^2 \in \Pi_2^S} \{ r_2(i, \pi^{*1}, \pi^2) +$

$$\alpha \sum_{j \in S} q(j|i, \pi^{*1}, \pi^2) V_{\alpha}^2(j, \pi^{*1}, \pi^{*2}) \} \quad (26)$$

因此, 结论 (a) 成立。

(b) 首先, 固定 $\pi^{*2} \in \Pi_2^S$, 由 (2) - (3) 式, (9) - (10) 式和 (25) 式可得

$$V_{\alpha}^1(i, \pi^{*1}, \pi^{*2}) \geq r_1(i, v, \pi^{*2}) + \alpha \sum_{j \in S} q(j|i, v, \pi^{*2}) V_{\alpha}^1(j, \pi^{*1}, \pi^{*2}) \quad \forall v \in P(A(i)), i \in S \quad (27)$$

因而, 对任意 $\pi^1 = (\pi_t^1, t \geq 0) \in \Pi_1$, 由对所有 $i \in S$ 和 $t \geq 0$, $\pi_t^1(\cdot|i) \in P(A(i))$, 结合 (2) - (3) 式及 (27) 式可得

$$V_{\alpha}^1(i, \pi^{*1}, \pi^{*2}) \geq r_1(i, \pi_t^1, \pi^{*2}) + \alpha \sum_{j \in S} q(j|i, \pi_t^1, \pi^{*2}) V_{\alpha}^1(j, \pi^{*1}, \pi^{*2})$$

对所有 $i \in S$ 成立。上式结合引理 2 (a) 给出

$$V_{\alpha}^1(i, \pi^{*1}, \pi^{*2}) \geq V_{\alpha}^1(i, \pi^1, \pi^{*2})$$

对任意 $i \in S$, $\pi^1 \in \Pi_1$ 成立。

同理由 (2) - (3) 式, (9) - (10) 式和 (26) 式, 可证得

$$V_{\alpha}^2(i, \pi^{*1}, \pi^{*2}) \geq V_{\alpha}^2(i, \pi^{*1}, \pi^2)$$

对任意 $i \in S$ 和 $\pi^2 \in \Pi_2$ 成立。

因而, (π^{*1}, π^{*2}) 是一个 Nash 平衡点。反之, 如果 (π^{*1}, π^{*2}) 是一个 Nash 平衡点, 则由引理 2 (c) 可知它满足 (20) 和 (21) 式。

注 3 满足上述假设 1 和假设 2 但又不满足文 [1 - 12] 中条件的例子是容易给出的, 只因篇幅所限, 我们略去详情。

参考文献:

- [1] ALTMAN E, HORDIJK A, SPIEKSMAN F M. Contraction conditions for average and α -discount optimality in countable state Markov games with unbounded rewards [J]. Math Oper Res, 1997, 22: 588 - 618.
- [2] FINK A M. Equilibrium in a stochastic N-person game [J]. J Sci Hiroshima Univ; Ser A - I Math, 1964, 28: 89 - 93.
- [3] FEDERGRUEN A. On N-person stochastic games with denumerable state space [J]. Adv Appl Prob, 1978, 10 (2): 452 - 471.
- [4] HIMMELBERG C J, PARTHASARATHY T, RAGHAVAN TES, et al. Existence of p-equilibrium and optimal stationary strategies in stochastic games [J]. Proc Amer Math Soc, 1976, 60: 245 - 251.
- [5] NOWAK A S. On a new class of nonzerosum discounted stochastic games having stationary Nash equilibrium points [J]. Int J Game Theory, 2003, 32: 121 - 132.
- [6] PARTHASARATHY T. Discounted, positive, and noncooperative stochastic games [J]. Internat J Game Theory, 1973, 2(1): 25 - 37.
- [7] PARTHASARATHY T, STERN M. Markov games-a survey [C]. Differential Games and Control Theory, Proc 2nd Conf Univ Rhode Island, Kingston RI 1976.
- [8] PARTHASARATHY T. Existence of equilibrium stationary strategies in discounted stochastic games [J]. Shankhya: Ser A, 1982, 44(1): 114 - 127.
- [9] PARTHASARATHY T, SINHA S. Existence of stationary equilibrium strategies in non-zerosum discounted stochastic games with uncountable state space and state-independent transitions [J]. Internat J Game Theory, 1989, 18(2): 189 - 194.
- [10] ROGERS P D. Non-zerosum stochastic games [D]. Berkeley: University of California, 1969.
- [11] SOBEL M J. Noncooperative stochastic games [J]. Ann Math Stat, 1971, 42: 1930 - 1935.
- [12] SENNOTT L I. Nonzero-sum stochastic games with unbounded costs: Discounted and average cost cases [J]. Z Oper Res, 1994, 40(2): 145 - 162.
- [13] TAKAHASHI M. Equilibrium points of stochastic non-cooperative n-person games [J]. J Sci Hiroshima Univ; Ser A-I Math, 1964, 28: 95 - 99.
- [14] GUO X P, YANG J. A new condition and approach for zero-sum stochastic games with average payoffs [J]. Stoc Anal Appl, 2008, 26(3): 537 - 561.
- [15] HERNANDEZ-LERMA O, LASSERRE J B. Further Topics on Discrete-Time Markov Control Processes [M]. New York: Springer-Verlag, 1999.
- [16] FAN K. Minimax theorems [J]. Proc Nat Acad Sci USA, 1953, 39: 42 - 47.

(下转第 36 页)

- two-phase cooling system under periodic temperature boundary conditions [J]. Nuclear power engineering, 2008, 29(2): 96-100.
- [9] 裴念强, 郭开华, 刘杰. 半导体制冷在新型环路热管的应用计算[J]. 低温物理学报, 2007, 129(1): 42-45.
- PEI Nianqiang, LIU Jie, GUO Kaihua. Applied calculation of peltier cooler one new LHP[J]. Chinese Journal of Low Temperature Physics. 2007, 129(1): 42-45.
- [10] 裴念强, 刘杰, 郭开华, 一种空间热控箱及其在运行轨道中的热分析[J]. 工程热物理学报, 2006, 127(S2): 73-77.
- PEI Nianqiang, LIU Jie, GUO Kaihua. A space thermal control box and its thermal analysis in Orbit[J]. Journal of Engineering Thermophysics, 2006, 127(S2): 73-77.
- [11] LIU Jie, PEI Nianqiang, GUO Kaihua, et al. Experimental investigation on startup of a novel two-phase cooling loop [J]. Experimental Thermal and Fluid Science, 2008, 32(4): 939-946.

Experimental Investigation on Cooling Capacity of The MPCL System Under Simulated Space Boundary Conditions

LIU Jie¹, GUO Kai-hua², HE Zhen-hui²

(1. Institute of Environmental and Municipal Engineering, Qingdao Technological University, Qingdao 266033, China;

2. Center of Space TECH., Sun Yat-sen University, Guangzhou 510275, China)

Abstract: The two-phase mechanical pumped cooling loop (MPCL) is a promising thermal control system. Experiments have been carried out to investigate cooling capacity and running characteristics under simulated temperature boundary conditions and variation heat load. It is concluded: when the condensers' heat rejection is higher than heat load on the evaporators, the system average temperature becomes lower and lower, which may lead to working fluid freezing in the condensers. It must decrease the set point evaporative temperature by controlling the accumulator. On the contrary, when the condensers' heat rejection is lower than heat load on the evaporators, the system average temperature becomes higher and higher, which may cause bubble appearing at the pump inlet and pump cavitation corrosion. It is necessary to increase the set point to avoid the pump running in an unstable station. All conclusions proved strong supporting for the system design and space applications

Key words: thermal control; space applications; mechanically pumped; cooling; two-phase

(上接第 27 页)

Nonzero-Sum Stochastic Games with Unbounded Discounted Payoff

YANG Jie^{1,2}, GUO Xian-ping²

(1. School of Science, Beijing Information Science and Technology University, Beijing 100192, China;

2. School of Mathematics and Computational Science, Sun Yat-sen University, Guangzhou 510275, China)

Abstract: Discrete time two-person nonzero-sum stochastic games with the discounted payoff criterion and a countable state space is studied, here the payoff functions might have neither upper nor lower bounds. The existence of Nash equilibria is proved in randomized stationary strategies by using the Fan's fixed-point theorem under the "drift" and the standard "continuous-compact" conditions, which are usually used in zero-sum stochastic games.

Key words: nonzero-sum stochastic game; expected discounted payoff criterion; Nash equilibria; countable state space