

## TRÒ CHƠI NGẪU NHIÊN (stochastic game)

**Trò chơi ngẫu nhiên** là một trò chơi động có tính Markov được chơi theo một chuỗi các giai đoạn. Vào đầu mỗi giai đoạn, trò chơi ở trong một trạng thái nào đó, đấu thủ chọn một hành động và được một phần thưởng phụ thuộc vào trạng thái hiện tại và hành động đã chọn. Tiếp theo, trò chơi chuyển sang trạng thái ngẫu nhiên mới có phân bố phụ thuộc vào trạng thái trước đó và hành động đã được đấu thủ chọn. Thủ tục được lặp lại ở trạng thái mới và tiếp tục lặp lại qua một số hữu hạn hoặc vô hạn các giai đoạn. Tính Markov thể hiện hàm phân phối xác suất chuyển không phụ thuộc vào lịch sử của trò chơi, nghĩa là

$$F(x_k | x_1, s^{(x_1)}; \dots; x_{k-1}, s^{(x_{k-1})}) = F(x_k | x_{k-1}, s^{(x_{k-1})}),$$

trong đó  $F(x_k | x_1, s^{(x_1)}; \dots; x_{k-1}, s^{(x_{k-1})})$  là hàm phân phối xác suất của trò chơi tại thời điểm tương lai  $x_k$  với điều kiện trước đó, ở các thời điểm quá khứ  $(x_1, \dots, x_{k-1})$  trò chơi đã có các trạng thái  $(s^{(x_1)}, \dots, s^{(x_{k-1})})$  tương ứng, còn  $F(x_k | x_{k-1}, s^{(x_{k-1})})$  là hàm phân phối xác suất của trò chơi tại thời điểm  $x_k$  với điều kiện ở thời điểm hiện tại  $x_{k-1}$  trò chơi có trạng thái  $s^{(x_{k-1})}$ .

L. S. Shapley là người đầu tiên đưa ra định nghĩa khái niệm trò chơi ngẫu nhiên vào đầu những năm 1950, khi nghiên cứu trò chơi ngẫu nhiên với tổng bằng 0 của hai người chơi với luật chơi chi trả ngay (được gọi là *trò chơi Shapley*). Trong trò chơi Shapley, tập các trạng thái của trò chơi và các bộ chiến lược cơ bản của các đấu thủ đều là hữu hạn, đồng thời ở bất kỳ bước nào và đối với bất kỳ lựa chọn nào của các đấu thủ, trò chơi đều có thể chấm dứt với một xác suất dương. Với điều kiện này, trò chơi kết thúc với xác suất 1 sau một số lượng hữu hạn các bước, đồng thời kỳ vọng toán của lượng tiền được chi trả của mỗi người chơi là hữu hạn. Trong một trò chơi như vậy, cả hai đấu thủ đều có chiến lược tối ưu dừng, tức là các chiến lược mà ở mọi giai đoạn của trò chơi, đấu thủ lựa chọn chiến lược cơ bản chỉ phụ thuộc vào tình hình tại thời điểm hiện tại. Shapley cũng phát hiện ra một thủ tục để có thể tìm ra chiến lược tối ưu của trò chơi.

Người ta còn nghiên cứu các trò chơi ngẫu nhiên khác với trò chơi Shapley có thể có vô hạn các trạng thái, là các trò chơi ngẫu nhiên với mức chi trả trung bình có giới hạn, nghĩa là trò chơi ngẫu nhiên hai đấu thủ với tổng

bằng 0 mà các chi trả  $h_1$  và  $h_2$  của hai đấu thủ thỏa mãn đẳng thức

$$h_1(P) = h_2(P) = \limsup_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{n} h_i(x_k, s^{x_k}).$$

Dưới điều kiện tính ergodic của xích Markov, sự tồn tại của chiến lược tối ưu dừng của một trò chơi như thế đã được chứng minh. Những kết quả này cũng được tổng quát hoá cho những trường hợp hạn chế số lượng các trạng thái và chiến lược cơ bản, và đối với các dạng luật chi trả khác.

**Trò chơi thống kê.** Bên cạnh khái niệm trò chơi ngẫu nhiên còn có khái niệm *trò chơi thống kê* là trò chơi hai đấu thủ với tổng bằng 0, trong đó người chơi thứ nhất được hiểu là tự nhiên, có chiến lược là một quá trình ngẫu nhiên, còn người chơi thứ hai được coi như một nhà thống kê, có chiến lược là một quy tắc quyết định (hàm quyết định), đồng thời hàm chi trả cho người chơi thứ nhất là hàm tổn thất của nhà thống kê. Chiến lược hỗn hợp của người chơi thứ nhất sẽ là một độ đo xác suất áp lên tập các quá trình ngẫu nhiên nêu trên, chiến lược hỗn hợp của người chơi thứ hai sẽ là một quy tắc quyết định ngẫu nhiên, đồng thời hàm chi trả cho người chơi thứ nhất được định nghĩa bằng kỳ vọng toán của hàm tổn thất của nhà thống kê.

Các trò chơi thống kê phát sinh từ các bài toán thống kê toán học (ví dụ, các bài toán ước lượng tham số, kiểm định giả thuyết, v.v.), trong đó có sử dụng phép kiểm định minimax. Việc mô tả một cách hệ thống về các bài toán thống kê toán học như các trò chơi thống kê được A. Wald giới thiệu lần đầu tiên, ông đã chứng minh định lý minimax cho một lớp rộng các trò chơi thống kê. Định lý này cung cấp một phương pháp giải quyết các vấn đề thống kê toán học, vì trong một số trường hợp có thể dễ dàng hơn và thuận tiện hơn để tìm ra maximin so với minimax, điều này làm cho việc tìm ra một quy tắc ngẫu nhiên hóa tối ưu trở nên dễ dàng hơn.

**Mô hình tổng quát.** Một trò chơi ngẫu nhiên bao gồm: tập hợp hữu hạn đấu thủ  $I$ ; không gian trạng thái  $M$ ; tập hành động  $S^i$  cho từng đấu thủ  $i \in I$ ; xác suất chuyển trạng thái  $\mathbb{P}$  từ  $M \times S$  sang  $M$ , trong đó  $S = \times_{i \in I} S^i$ , với  $\mathbb{P}(A|m, s)$  là xác suất mà trạng thái tiếp theo nằm trong  $A$  với điều kiện trạng thái hiện tại là  $m$  và hành động hiện tại là  $s$ ; và hàm chi trả  $g: M \times S \rightarrow \mathbb{R}^I$ , trong đó hàm véc tơ  $g$  có tọa độ thứ  $i$  là  $g^i$ , cho biết số

tiền thưởng đầu thủ thứ  $i$  sẽ nhận được nếu đang ở trạng thái  $m$  mà chọn hành động  $s$ .

Trò chơi bắt đầu ở trạng thái ban đầu  $m_1$  nào đó. Ở giai đoạn  $t$ , trước tiên các đầu thủ quan sát trạng thái  $m_t$ , rồi cùng một lúc chọn các hành động  $s_t^i$  trong  $S^i$  của mình, sau đó quan sát danh mục hành động  $s_t = (s_t^i : i \in I)$  của tất cả các đầu thủ cùng chơi, tiếp đó cơ chế của trò chơi sẽ tự động chọn ra  $m_{t+1}$  theo xác suất  $\mathbb{P}(\cdot | m_t, s_t)$ . Việc tiến hành trò chơi ngẫu nhiên,  $(m_1, s_1), \dots, (m_t, s_t) \dots$ , sẽ xác định luồng véctơ tiền thưởng được thanh toán  $g_1, g_2, \dots$ , trong đó  $g_t = g(m_t, s_t)$ .

Trò chơi có khấu trừ  $\Gamma_\lambda$  với hệ số khấu trừ  $\lambda \in (0; 1]$  là trò chơi mà người chơi thứ  $i$  được nhận số tiền thưởng bằng  $\lambda \sum_{t=1}^{\infty} (1 - \lambda)^{t-1} g_t^i$ . Trò

chơi  $n$  - giai đoạn là trò chơi mà tiền thưởng cho người chơi thứ  $i$  bằng

$$\bar{g}_n^i := \frac{1}{n} \sum_{t=1}^n g_t^i.$$

Trò chơi ngẫu nhiên có nhiều ứng dụng trong kinh tế, sinh học và nghiên cứu mạng máy tính.

**HỒ ĐĂNG PHÚC**

### Tài liệu tham khảo

1. A. Condon, *The complexity of stochastic games*, Information and Computation 96, 1992, 203–224.
2. J. Filar; K. Vrieze, *Competitive Markov Decision Processes*, Springer-Verlag, New York, 1997.
3. J. F. Mertens; A. Neyman, *Stochastic Games*, International Journal of Game Theory. 10, 1981, No. 2, 53–66.
4. A. Neyman; S. Sorin, *Stochastic Games and Applications*, Kluwer Academic Press, Dordrecht, 2003.
5. L. S. Shapley, *Stochastic games*, Proc. Nat. Acad. Sci., 39, 1953, 1095–1100.
6. N. Vieille. *Stochastic Games: Recent Results*, Handbook of Game Theory. Amsterdam: Elsevier Science 2002, 1833–1850.
7. K. Vrieze, *Zero-sum stochastic games*, Surveys in Game Theory and Related Topics, Tracts 39, 1987, CWI, 103–132.