مجله محاسبات نرم، جلد ۱۱، شماره ۱ (بهار و تابستان ۱۴۰۱)، صفحه ۴۸ـ۵۹
https://doi.org/10.22052/scj.2022.242848.0

مجله محاسبات نرم
SOFT COMPUTING JOURNAL
تارنمای مجله: scj.kashanu.ac.ir

University of Kashan

تعیین نقطه تعادل در بازی‌های پویای مارکفی گسسته دونفره با احتمالات انتقال صرفا تحت تاثیر استراتژی‌های رقیب ✧

رامین صادقیان[1]*، دانشیار
[1] گروه مهندسی صنایع، دانشگاه پیام نور، تهران، ایران.

اطلاعات مقاله

تاریخچه مقاله:
دریافت ۳۱ فروردین ماه ۱۴۰۰
پذیرش ۰۵ آبان ماه ۱۴۰۱

کلمات کلیدی:
بازی پویای مارکفی
بازی چندمرحله‌ای
زنجیره مارکف
ماتریس احتمال انتقال

چکیده

این مقاله به نوع خاصی از بازی‌ها با عنوان بازی‌های پویای مارکفی می‌پردازد. در این نوع بازی‌ها استراتژی هر بازیکن یک حالت از زنجیره مارکف در نظر گرفته می‌شود. در هر مرحله بازیکنان ممکن است با توجه به شرایط و موقعیت خود، همان استراتژی یا استراتژی دیگری را با احتمال مشخصی انتخاب نمایند. البته این انتخاب‌ها بستگی به استراتژی بازیکنان دیگر هم دارد. در این مقاله یک بازی دونفره گسسته مارکفی با احتمالات انتقال از قبل تعیین شده و مستقل که فقط تحت تاثیر استراتژی‌های بازیکن رقیب است، در نظر گرفته شده و نحوه تعیین نقاط تعادل در حالت بازی‌های پویا به صورت مارکفی در قالب یک مثال عددی مورد ارزیابی و تحلیل قرار گرفته است.

۱. مقدمه

در طی زمان، رویکرد تئوری بازی در حوزه‌های متفاوت پیشرفت‌ها و تغییرات زیادی داشته است. این رویکردها بر اساس متغیرهای گوناگون دسته‌بندی می‌شوند. از جمله این متغیرها می‌توان به دسته‌بندی بازی‌ها بر حسب تعداد بازیکنان، گسسته یا پیوسته بودن بازی، زمان ارائه تصمیم‌گیری شامل فرم‌های نرمال، پویا و ائتلاف، با همکاری یا بدون همکاری بودن بازی، قطعی یا غیرقطعی بودن استراتژی‌ها و انتخاب‌ها، یک مرحله‌ای یا چند مرحله‌ای بودن و از این دست اشاره کرد [۱]،[۲]. دقت داشته باشید که برخی از این دسته‌بندی‌ها مانند بازی‌های مخلوط[1] و بازی‌های چند مرحله‌ای[2] که در دنیای واقعی بیشتر اتفاق می‌افتند، در پژوهش‌های علمی کمتر مورد توجه و تجزیه و تحلیل قرار گرفته‌اند [۳]. در بین آنها بازی‌های مارکفی نوع خاصی از بازی‌ها هستند که بایستی بیشتر مورد ارزیابی قرار گیرند. اگرچه این ارزیابی پیش از این در مقالاتی نیز انجام شده (به عنوان مثال مراجع [۴] و [۵] را ببینید)، اما تحلیل نقاط تعادل با استفاده از زنجیره مارکف به ندرت مورد توجه قرار گرفته و به صورت الگوریتمی ارائه شده است. چنین تحلیلی در قیاس با بازی‌های معمولی از چند جهت دارای مزیت است از جمله چند مرحله‌ای بودن بازی‌ها که بازی‌ها تا زمانی که بازیکنان به تعادل پایدار نرسند، ادامه می‌یابد. دقت

[1] Mixed Strategies Games
[2] Multi Stage

✧ نوع مقاله: پژوهشی
* نویسنده مسئول
پست(های) الکترونیک: akrami.ab@uoz.ac.ir (صادقیان)

نحوه ارجاع به مقاله: صادقیان، رامین، «تعیین نقطه تعادل در بازی‌های پویای مارکفی گسسته دونفره با احتمالات انتقال صرفا تحت تاثیر استراتژی‌های رقیب»، مجله محاسبات نرم، جلد ۱۱، شماره ۱، ص۴۸ـ۵۹، بهار و تابستان ۱۴۰۱.

کنید که در بازی‌های سنتی بازی یا فقط یکبار انجام می‌شود یا در صورت تکرار در چند مرحله قبل از رسیدن به نقطه تعادل پایدار متوقف می‌شود. در اکثر بازی‌هایی که به عنوان مثال در زنجیره تامین کار می‌شود و حتی از نوع پویا و استکلبرگ گرفته می‌شود به طور معمول تا یک مرحله یا حداکثر چند مرحله ادامه می‌یابد و تا رسیدن به نقطه تعادل یا همان احتمال انتقال نهایی پیش نمی‌رود، ولی با اعمال به صورت مارکفی این امکان برای تعیین احتمال انتقال نهایی وجود دارد.

مزیت دیگر قطعی نبودن لزومی استراتژی‌های قابل انتخاب برای هر یک از بازیکنان است، جایی که آنها می‌توانند ترکیبی از استراتژی‌ها را انتخاب کنند. در ضمن این دسته از بازی‌ها لزوما قطعی نیستند و می‌توانند دارای عدم قطعیت هم در نوع ترکیب استراتژی‌ها و هم در احتمال رویداد آنها باشند. به عبارت دیگر، در این نوع بازی‌ها اگر پیش از این در نظر بازیکنان استراتژی خاصی مدنظر بود که اقدام کنند با اقدام بازیکن رقیب این مجموعه استراتژی‌ها و احتمال انتخاب آنها به طور کلی ممکن است تغییر یابد. به عنوان مثال در یک رقابت بین تولیدکننده و خرده‌فروش، استراتژی‌های مرسوم مشخص است و خرده‌فروش می‌داند به ازای حرکت بعدی رقیب که یا تغییر کیفیت محصول است یا تغییر قیمت محصول یا تغییر زمان تحویل محصول و یا ... چه اقداماتی و با چه احتمال‌هایی انجام دهد، با این همه به ناگهان تولیدکننده اقدام جدیدی مانند ارائه محصول جدید و ارزان قیمت را ارائه می‌دهد. در بازی‌های سنتی اگر از قبل استراتژی برای مقابله با این عملکرد در ماتریس‌ها دیده نشده باشد، می‌بایست از بین استراتژی‌های موجود اقدامی انجام گیرد ولی در بازی‌های مارکفی امکان اینکه استراتژی جدیدی اضافه شود و یا احتمال انتخاب آنها بر اساس نیاز تغییر یابد، وجود دارد. در بازی شطرنج نیز به طور مشابه به ازای حرکت‌های پیش‌بینی نشده رقیب می‌توان استراتژی و احتمال‌های مختلفی را بررسی کرد. اگر در بازی‌های سنتی همیشه این مجموعه استراتژی‌ها و احتمال رخداد آنها ثابت بود در بازی‌های مارکفی در هر حرکت قابل تغییر و اصلاح متناسب با حرکت رقیب است. مزیت دیگری امکان دخل و تصرف بر نوع عوامل اثرگذار بر احتمال‌های انتقال شامل عوامل مرتبط با هر بازیکن یا بازیکنان رقیب و حتی اثرات مرتبط با گذر زمان اشاره کرد.

با توجه به آنچه تا به حال بیان شد، می‌توان هدف اصلی این مقاله را معرفی بازی‌های مارکفی با شرایط مذکور و نیز نحوه تعیین و تحلیل نقاط تعادل این گونه بازی‌ها دانست. لذا در این مقاله، بازی‌های پویا و نحوه تعیین نقاط تعادل برای آنها بررسی شده است. دقت داشته باشید که بازی‌ها در این مقاله به صورت دونفره و گسسته در نظر گرفته شده‌اند و احتمال‌های انتقال فقط تحت تاثیر استراتژی‌های انتخاب شده توسط بازیکن رقیب تعریف می‌شوند.

در ادامه این مقاله، ابتدا تعاریف مورد نیاز از جمله بازی‌های مارکفی طرح شده و سپس الگوریتم و مدل پیشنهادی بیان می‌شوند. آنگاه در قالب یک مثال عددی با سه روش پیشنهادی مدل حل و تشریح خواهد شد.

۲. تعاریف

۲,۱. بازی‌های تکراری

در بازی‌های استراتژی، تصمیم‌گیری‌های متوالی سازنده ساختار بازی هستند، یعنی تصمیم‌گیرندگان به طور متوالی حرکات خود را انجام می‌دهند. در ضمن پس از انتخاب طرح اقدام توسط هر تصمیم‌گیرنده، وقوع هیچ رویدادی آن را تغییر نمی‌دهد. به عبارت دیگر، طرح اقدام تنها یکبار برای همیشه انتخاب می‌شود. در مقابل در بازی‌های تعمیم‌یافته[1]، ساختار توالی تصمیم‌گیری‌ها روشن و مشخص است و امکان تغییر تصمیم در موقعیت‌های مختلف را وابسته به وقوع رویدادهای گوناگون برای تصمیم‌گیرنده فراهم می‌کند. در بازی‌های تعمیم‌یافته با اطلاعات کامل[2]، بازیکنان، مجموعه ترجیح‌های آنها، ترتیب حرکت بازیکنان و اقدام‌های آنها باید معین باشد. در بسیاری از تعاملات استراتژیک، تعاملات با همان یک نفر تکرار می‌شود که سبب ایجاد فرصتی برای کسب مزیت بر دیگر بازیکنان می‌شود. از نتایج یک بازی استراتژیک، نتیجه یک بازی تکراری دنباله‌ای

[1] Extensive Games
[2] Perfect Information

است [۶].

در بازی‌های تکراری[1] به بازی پایه یک بازی جزء یا بازی مرحله‌ای گفته می‌شود که در چند مرحله بازی می‌شود. اگر این بازی پایه به تعداد بینهایت بار رخ دهد، ایجاد یک ابربازی می‌کند. در عوض چنانچه در طی زمان ساختار بازی تغییر کند به آن بازی تفاضلی (دیفرانسیلی)[2] گفته می‌شود. در واقع ساختار بازی‌های تکراری مستقل از زمان و بازی‌های دیفرانسیلی وابسته به زمان است. با این همه این بدان معنا نیست که استراتژی بازی‌های تکراری مستقل از زمان هستند. به عبارت دقیق‌تر، یک بازی می‌تواند بینهایت بار تکرار شود ولی استراتژی‌های آن وابسته به کارهای قبلی باشد. به طور معمول تحلیل ابربازی‌ها[3] ساده‌تر از بازی‌های دیفرانسیلی است زیرا در تحلیل بازی‌های دیفرانسیلی به برنامه‌ریزی پویا نیاز است [۷].

در صورت عدم وجود دنباله معینی برای یک مساله، بازی باید به شکل حرکت همزمان انجام شود در حالی که وجود چنین دنباله ای منجر به حرکات دنباله‌ای می‌شود. لازم به ذکر است که در اکثر بازی‌ها، دنباله اقدامات روی استراتژی‌ها و در نتیجه روی نقطه تعادل اثر می‌گذارد و اغلب بازیکنی که حرکت اول را انجام می‌دهد دارای مزیت بیشتری است [۷].

۲,۲. زنجیره مارکف و ماتریس احتمالات انتقال

زنجیره‌های مارکف[4] مرسوم‌ترین نوع فرآیندهای احتمالی گسسته هستند و به صورت زیر تعریف می‌شوند.

فرض کنید $\{X_t\}, \forall t \in T$ یک متغیر تصادفی با مقادیری در فضای قابل شمارش باشد، آن یک زنجیره مارکف است اگر شرط زیر (خاصیت مارکفی) تامین شود:

- برای هر $t \in T$ و $Pr(X_0 = i_0, \dots, x_{t-1} = i_{t-1}) > 0$ داشته باشیم:

$$\begin{aligned}\Pr(X_t = i_t | X_0 = i_0, \dots, X_{t-1} = i_{t-1}) \\ = \Pr(X_t = i_t | X_{t-1} = i_{t-1})\end{aligned} \tag{۱}$$

اگر X_t وضعیت سیستم در دوره زمانی t باشد، احتمال رسیدن به وضعیت قطعی در دوره t فقط به وضعیت سیستم در دوره $t-1$ بستگی دارد و این که چگونه به وضعیت X_{t-1} رسیده است اهمیتی ندارد [۸].

احتمالات شرطی $\Pr(X_t = i_t | X_{t-1} = i_{t-1})$ احتمالات انتقال نامیده می‌شود. با استفاده از خاصیت مارکفی، رابطه بین احتمالات توام و شرطی به دست می‌آید که به معادله چاپمن–کلموگروف[5] معروف است و در تعیین توزیع‌های محدود بسیار مفید است [۸].

$$\begin{aligned}\Pr(X_0 = i_0, \dots, X_t = i_t) \\ = \Pr(X_0 = i_0) \\ \times \Pr(X_1 = i_1 | X_0 = i_0) \times \dots \\ \times \Pr(X_t = i_t | X_{t-1} = i_{t-1})\end{aligned} \tag{۲}$$

زنجیره مارکف $\{X_t\}, \forall t \in T$ همگن[6] است اگر و تنها اگر احتمالات رابطه (۱) به t وابسته نباشند و در غیر این صورت ناهمگن است. در حال حاضر ما فقط حالت همگن را در نظر می‌گیریم.

$$\Pr(X_t = j | X_{t-1} = i) = p_{ij} \tag{۳}$$

و ماتریس P به صورت زیر تعریف می‌شود:

$$P = [p_{ij}] \tag{۴}$$

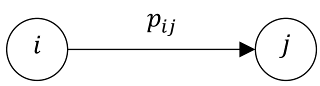
درایه‌های ماتریس P به صورت زیر تعریف می‌شوند:

$$\begin{aligned}&(i)\ p_{ij} > 0 \qquad \forall i, j \in I \\ &(ii) \sum_{j \in I} p_{ij} = 1 \qquad \forall i \in I\end{aligned} \tag{۵}$$

اگر ماتریس P دو شرط فوق را داشته باشد، به آن ماتریس مارکف یا یک ماتریس احتمالات انتقال[7] گویند.

به ازای هر ماتریس انتقال می‌توان یک گراف انتقال تخصیص داد که راس‌ها همان حالات هستند. همانطور که در شکل (۱) قابل مشاهده است، کمانی بین دو راس i و j وجود خواهد داشت اگر و تنها اگر $p_{ij} > 0$ باشد.

[1] Repeated Games
[2] Differential Games
[3] Super-Games
[4] Markov Chain
[5] Chapman-Kolmogorov Equation
[6] Homogeneous
[7] Transition Probabilities Matrix



شکل (۱): گراف انتقال

حال می‌توان احتمال انتقال $p_{ij}^{(n)}$ را به صورت زیر تعریف کرد:

$$p_{ij}^{(n)} = \Pr(X_{v+n} = j \,|X_v = i) \tag{۶}$$

از خاصیت مارکفی رابطه (۱) واضح است که شرط مربوطه با توجه به X_{v+1} به صورت زیر تبدیل می‌شود:

$$p_{ij}^{(2)} = \sum_k p_{ik} p_{kj} \tag{۷}$$

با استفاده از نماد ماتریسی داریم:

$$P^{(2)} = \left[p_{ij}^{(2)}\right] \tag{۸}$$

و با توجه به رابطه (۷) داریم:

$$P^{(2)} = P^2 \tag{۹}$$

و در حالت کلی:

$$P^{(n)} = \left[p_{ij}^{(n)}\right] \tag{۱۰}$$

و به ازای هر $n \geq 1$ داریم:

$$P^{(n)} = P^n \tag{۱۱}$$

رابطه (۱۱) نتیجه می‌دهد که ماتریس احتمال انتقال در n گام با توان n ام ماتریس P برابر است.

برای توزیع‌های حاشیه‌ای مرتبط با X_n به ازای هر $i \epsilon I$ و $n \geq 0$ داریم:

$$p_i(n) = \Pr(X_n = i) \tag{۱۲}$$

که این احتمالات می‌توانند به صورت زیر محاسبه شوند:

$$p_i(n) = \sum_j p_j p_{ji}^{(n)} \qquad i \in I \tag{۱۳}$$

اگر بنویسیم:

$$P^{(0)} = I \text{ یا } p_{ji}^{(0)} = \delta_{ji} \tag{۱۴}$$

آنگاه رابطه (۱۳) به ازای هر $n \geq 0$ صحیح است. در ضمن اگر

$$P(n) = \big(p_1(n), \dots, p_m(n)\big) \tag{۱۵}$$

آنگاه رابطه (۱۳) می‌تواند با استفاده از نماد ماتریسی به صورت زیر بیان شود [۹]:

$$P(n) = p \times P^n \tag{۱۶}$$

۳. بازی‌های مارکفی

در هر مرحله از یک بازی چند مرحله‌ای، هر بازیکن می‌تواند یک استراتژی را با احتمال معینی انتخاب کند. به طور معمول این انتخاب به انتخاب قبلی صورت گرفته توسط بازیکن رقیب وابستگی زیادی دارد. چنین بازی، یک بازی مارکفی[۱] نامیده می‌شود [۳]، [۱۰]– [۱۳]. در حالی که می‌توان بازی‌ها را به صورت دو نفره[۲] یا چند نفره[۳] [۱۴]– [۱۷]، گسسته یا پیوسته [۱۸]– [۲۰]، یک یا چند مرحله‌ای [۳]، [۲۱]– [۲۳] دسته‌بندی کرد، آنها را می‌توان به شکل مارکفی نیز در نظر گرفت [۴]–[۵]، [۲۴]–[۲۸]. در این نوع از بازی‌ها، هر بازیکن ابتدا یک استراتژی را با احتمال معینی انتخاب کرده و سپس با توجه به استراتژی رقیب، در انتخاب‌های بعدی استراتژی خود را تغییر می‌دهد. به عبارت دیگر هر بازیکن به شکل پویا و بر حسب شرایط تصمیم‌گیری می‌نماید. بیشتر بازی‌ها در دنیای به همین صورت انجام می‌شوند و بازیکنان انتخاب‌های خود را با شرایط تطبیق می‌دهند. به عنوان نمونه در یک بازی شطرنج حرکات هر بازیکن می‌تواند بر اساس حرکات رقیب تغییر نماید. از دید بازیکنان استراتژی‌ها دارای احتمالات مشخصی هستند حتی اگر نتوانند آن را به طور شفاف مشخص نمایند. ما در [۲۶]، صرفا بازی‌های ایستا و فرم نرمال را بررسی کرده‌ایم، در حالی که در این مقاله به بازی‌های فرم پویا که تکامل یافته آن است، پرداخته می‌شود.

۴. مدل پیشنهادی

در مدل پیشنهادی، سه سناریو مشابه مقاله [۲۶] استفاده شده است. دقت داشته باشید که هدف اصلی این مقاله نحوه عملکرد هر یک از بازی‌های مارکفی بیان شده در هر سناریو و در نهایت

[1] Markovian Games
[2] Two Player Games
[3] Multi-Player Games

تعیین نقطه تعادل برای این بازی‌ها است. بیان این نکته ضروری است که احتمال انتقال در این سناریوها صرفا تحت تاثیر استراتژی‌های رقیب است و این امر به عنوان یکی از مفروضات مهم این مقاله در نظر گرفته شده است. به عبارت دقیق‌تر، اگر بازیکن A استراتژی A_1 را انتخاب کند، بازیکن B با احتمال r_{11} استراتژی B_1 را انتخاب می‌کند. مفروضات مدل پیشنهادی همانند مرجع [۲۶] بوده و به صورت زیر دسته‌بندی می‌شود:

- بازی به صورت پویا است.
- بازی دو نفره و گسسته است.
- احتمالات انتقال صرفا تحت تاثیر استراتژی‌های رقیب بوده و در واقع مقادیر تابع مطلوبیت نیز روی مقادیر آنها اثرگذار است.
- استراتژی تمام بازیکنان در تمام مراحل ثابت است.
- احتمالات انتقال در تمام مراحل ثابت است.

با تبدیل یک بازی پویا به یک زنجیره مارکف که در ادامه این مقاله توضیح داده خواهد شد، برای مشخص کردن احتمالات پایدار یا نهایی [۸]- [۹] (که در حقیقت احتمالات در بینهایت هستند)، از ماتریس مربعی احتمال انتقال T و دستگاه معادلات ذیل استفاده می‌شود [۲۶]:

$$\begin{cases} q.T = q \\ \sum_i q_i = 1 \end{cases} \Rightarrow \begin{cases} q.T - q.I = O \\ \sum_i q_i = 1 \end{cases} \Rightarrow \begin{cases} q.(T - I) = O \\ \sum_i q_i = 1 \end{cases} \qquad (۱۷)$$

در این دستگاه T یک ماتریس مربعی $m \times m$ و q یک ماتریس برداری $1 \times m$ و معادل عبارت ذیل است که در آن احتمالات در بینهایت به یک مقدار ثابت میل خواهند کرد [۲۶]:

$$q = \lim_{n\to\infty} T^n \qquad (۱۸)$$

معادله اول از دستگاه معادله رابطه (۱۷) یعنی $q.(T - I) = O$ دارای m معادله است، ولی چون این معادله‌ها همگن هستند لذا $m - 1$ معادله از بین آنها می‌تواند برای حل دستگاه انتخاب شود و همراه با معادله $\sum_i q_i = 1$ تشکیل یک دستگاه کامل معادلات جهت حل و تعیین مقادیر q را دهد [۸]- [۹].

۴،۱. بازی‌های پویای دونفره (A بازیکن اول است)

اگر یک بازی گسسته دونفره به صورت ماتریس زیر وجود داشته باشد:

$$\begin{array}{c|c} & \begin{array}{cccc} B_1 & B_2 & \cdots & B_n \end{array} \\ \begin{array}{c} A_1 \\ A_2 \\ \vdots \\ A_m \end{array} & \begin{bmatrix} (f_{11}^A, f_{11}^B) & (f_{12}^A, f_{12}^B) & \cdots & (f_{1n}^A, f_{1n}^B) \\ (f_{21}^A, f_{21}^B) & (f_{22}^A, f_{22}^B) & \cdots & (f_{2n}^A, f_{2n}^B) \\ \vdots & \vdots & \ddots & \vdots \\ (f_{m1}^A, f_{m1}^B) & (f_{m2}^A, f_{m2}^B) & \cdots & (f_{mn}^A, f_{mn}^B) \end{bmatrix} \end{array} \qquad (۱۹)$$

که $A_1, A_2, \dots, A_m$ استراتژی‌های بازیکن A و $B_1, B_2, \dots, B_m$ استراتژی‌های بازیکن B هستند و (f_{ij}^A, f_{ij}^B) مطلوبیت بازیکنان A و B به ازای استراتژی های i و j است، در این صورت فرض می‌شود بازی پویا بوده و بازیکن A شروع کننده بازی است و بازی به صورت تکراری در چندین مرحله تا رسیدن به نقطع تعادل پایدار برگزار می‌شود. از این‌رو در مرحله اول بازی، بازیکن A با احتمالات p_i یکی از استراتژی‌های $A_1, A_2, \dots, A_m$ را انتخاب می‌کند. در تمامی بازی‌های مختلط و زنجیره‌های مارکف، این احتمالات انتقال تماما بر اساس تجربه و خبرگی بازیکنان تعیین می‌شوند. البته ممکن است تحت شرایطی استدلال‌های منطقی هم پشت این موضوع باشد ولی اکثر مواقع ممکن است تحت معیارهای ذهنی بازیکنان و تصمیم‌گیرندگان تعیین و اعمال شوند.

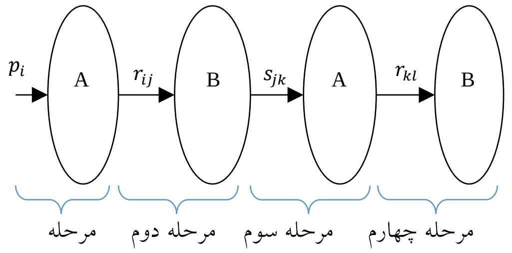
سپس در مرحله دوم، بازیکن B با احتمال r_{ij} استراتژی B_j را انتخاب می‌کند مشروط بر آن که بازیکن A استراتژی A_i را انتخاب کرده باشد. در مرحله سوم بازیکن A با احتمال s_{jk} استراتژی A_k را انتخاب می‌کند مشروط بر آن که بازیکن B استراتژی B_j را انتخاب کرده باشد. یکی از مفروضات این مقاله این است که از مرحله چهارم به بعد، مراحل دوم به بعد تکرار خواهند شد. به این معنی که احتمالات در مراحل دوم و چهارم و ششم و ... با هم برابر و احتمالات در مراحل سوم و پنجم و هفتم و ... نیز با هم برابرند. شکل (۲) نمای کلی زنجیره مارکف حاصل از بازی مارکفی فوق الذکر را نشان می‌دهد و شکل (۳) زنجیره مارکف حاصل از بازی مارکفی را با جزئیات احتمالات موجود نمایش می‌دهد.

ماتریس‌های احتمال انتقال حاصل از بازی مارکفی به صورت ذیل خواهند بود:

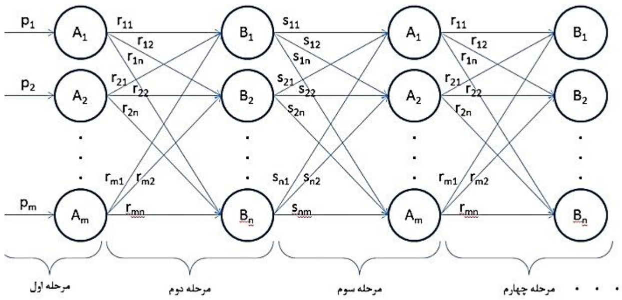
$$P_{1\times m} = [p_1 \quad p_2 \quad \cdots \quad p_m]$$

$$R_{m\times n} = \begin{array}{c} \\ A_1 \\ A_2 \\ \vdots \\ A_m \end{array} \begin{array}{c} \begin{matrix} B_1 & B_2 & \cdots & B_n \end{matrix} \\ \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix} \end{array} \qquad (۲۰)$$

$$R_{m\times n} = \begin{array}{c} \\ B_1 \\ B_2 \\ \vdots \\ B_n \end{array} \begin{array}{c} \begin{matrix} A_1 & A_2 & \cdots & A_m \end{matrix} \\ \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1m} \\ s_{21} & s_{22} & \cdots & s_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{nm} \end{bmatrix} \end{array}$$



شکل (۲): نمای کلی زنجیره مارکف حاصل از بازی مارکفی پویا وقتی A شروع کننده بازی است.



شکل (۳): زنجیره مارکف حاصل از بازی مارکفی پویا همراه با جزئیات وقتی A شروع کننده بازی است.

دو تفاوت عمده در ماتریس‌های احتمال انتقال فوق نسبت به ماتریس‌های احتمال انتقال مرسوم وجود دارد: اول این که ماتریس‌های R و S لزوما مربعی نیستند و دوم این که در ابتدای شروع حرکت بازیکن A احتمالاتی مانند P نیز اثر گذار است که باید مد نظر قرار گیرد، در حالی که در زنجیره‌های مارکف مرسوم وجود ندارند. از این‌رو نحوه برخورد با ماتریس‌های احتمال انتقال فوق الذکر باید اندکی متفاوت از روش‌های مرسوم در زنجیره‌های مارکف باشد. برای این منظور سه روش محاسبه جهت تعیین احتمالات پایدار و نهایی پیشنهاد می‌گردد که جزئیات آنها در ادامه آمده است.

۴،۱،۱. روش اول (استفاده از زنجیره مارکف پنهان)

در این روش در حقیقت احتمال نهایی به صورت رابطه ذیل تعیین می‌شود:

$$\Pi = P_{1\times m} \times R_{m\times n} \times S_{n\times m} \times R_{m\times n} \times S_{n\times m} \times \dots \qquad (۲۱)$$

برای این منظور ابتدا حاصل‌ضرب $R_{m\times n} \times S_{n\times m}$ به صورت یک ماتریس $m \times m$ با نام $RS_{m\times m}$ تعیین می‌شود، سپس با در نظر گرفتن $T = RS$ و با استفاده از دستگاه معادلات رابطه (۱۷) مقدار q محاسبه شده و با استفاده از عملگر زیر که در حقیقت ضرب درونی (درایه به درایه) همراه با نرمالیزه کردن است، مقدار احتمالات نهایی برای بازیکن A تعیین می‌شود:

$$\Pi = P\otimes \lim_{n\to\infty}(R\times S)^n = P\otimes \lim_{n\to\infty}(RS_{m\times m})^n = P\otimes Q \qquad (۲۲)$$

$$P\otimes Q = [p_1 \quad \cdots \quad p_m]\otimes[q_1 \quad \cdots \quad q_m] = \left[\frac{p_1\times q_1}{\sum_{i=1}^{m} p_i\times q_i} \quad \cdots \quad \frac{p_m\times q_m}{\sum_{i=1}^{m} p_i\times q_i}\right] \qquad (۲۳)$$

پس از تعیین احتمالات نهایی برای بازیکن A، این بازیکن آن استراتژی را انتخاب خواهد کرد که بیشترین مقدار را در احتمالات نهایی کسب کرده است. سپس با توجه به ماتریس بازی و اینکه بازی پویا است، استراتژی بازیکن B بر اساس استراتژی انتخابی بازیکن A و در نهایت نقطه تعادل تعیین خواهد شد.

با توجه به اینکه در این روش با جایگزینی $T = RS$ در حقیقت ماتریس‌های احتمال انتقال S علیرغم اینکه روی احتمالات نهایی اثرگذار هستند ولی ابعاد آنها در دستگاه معادلات جهت تعیین احتمالات نهایی مد نظر قرار نمی‌گیرند و در واقع حالات مربوط به بازیکن B در محاسبات پنهان شده‌اند. از این‌رو این زنجیره مارکف، زنجیره مارکف پنهان نامیده می‌شود. به هر حال، با توجه به اینکه در این روش حالات نهایی B جهت تعیین نقطه تعادل نادیده گرفته می‌شود و برای تعیین نقطه تعادل، از

اولین ماتریس بازی استفاده می‌شود نه از حالات حدی، می‌توان این مورد را یکی از نقاط ضعف این روش به حساب آورد. لذا روش‌های دوم و سوم، که در ادامه بیان خواهند شد، می‌توانند جایگزین‌های دیگری برای این روش باشند.

۴،۱،۲. روش دوم (تجمیع ماتریس‌های احتمالات انتقال با یکبار نرمالیزه کردن)

در این روش کل حالات مربوط به استراتژی‌های بازیکنان A و B با هم تجمیع شده و به صورت یک زنجیره مارکف واحد در نظر گرفته می‌شود. در این روش ابتدا ماتریس احتمال انتقال T به صورت زیر شکل می‌گیرد که $(m+n)\times(m+n)$ است:

$$T_{(m+n)\times(m+n)} = \begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{bmatrix} O & R \\ S & O \end{bmatrix} \end{array}$$

$$= \begin{array}{cc} & \begin{array}{cccccc} A_1 & \cdots & A_m & B_1 & \cdots & B_n \end{array} \\ \begin{array}{c} A_1 \\ \vdots \\ A_m \\ B_1 \\ \vdots \\ B_n \end{array} & \begin{bmatrix} 0 & \cdots & 0 & r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{m1} & \cdots & r_{mn} \\ s_{11} & \cdots & s_{1m} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ s_{n1} & \cdots & s_{nm} & 0 & \cdots & 0 \end{bmatrix} \end{array} \quad (۲۴)$$

حال با استفاده از رابطه (۱۷) مقدار Q که میزان احتمالات حدی را به ازای هر دو بازیکن A و B نشان می‌دهد، تعیین می‌شود. توجه شود که در این روش با استفاده از رابطه (۱۷) فقط $m+n-1$ معادله از معادلات همگن استفاده شده و یک معادله نیز جهت نرمالیزه کردن و جمع احتمالات برابر با ۱ استفاده می‌شود. پس از تعیین مقدار Q، ابتدا Q با استفاده از رابطه (۲۵) به تفکیک بازیکنان نرمالیزه می‌شود. سپس می‌بایست عملگر معرفی شده در رابطه (۲۳) این بار به صورت زیر فقط برای احتمالات مربوط به بازیکن A با ضرب آن عملگر بین دو ماتریس احتمالات P و NQ استفاده شوند، یعنی داریم:

$$\begin{aligned} Q &= [q_1 \quad \cdots \quad q_m \quad q_{m+1} \quad \cdots \quad q_{m+n}] \\ NQ &= \left[\frac{q_1}{\sum_{i=1}^{m} q_i} \quad \cdots \quad \frac{q_m}{\sum_{i=1}^{m} q_i} \quad \frac{q_{m+1}}{\sum_{i=m+1}^{m+n} q_i} \quad \cdots \quad \frac{q_{m+n}}{\sum_{i=m+1}^{m+n} q_i}\right] \end{aligned} \quad (۲۵)$$

$$\begin{aligned} &P \otimes NQ \\ &= [p_1 \quad \cdots \quad p_m] \otimes [Nq_1 \quad \cdots \quad Nq_m \quad Nq_{m+1} \quad \cdots \quad Nq_{m+n}] \\ &= \left[\frac{p_1 \times Nq_1}{\sum_{i=1}^{m} p_i \times Nq_i} \quad \cdots \quad \frac{p_m \times Nq_m}{\sum_{i=1}^{m} p_i \times Nq_i} \quad Nq_{m+1} \quad \cdots \quad Nq_{m+n}\right] \end{aligned} \quad (۲۶)$$

با استفاده از مقادیر به دست آمده، هر بازیکن بیشترین احتمال مربوط به استراتژی‌ها را به عنوان استراتژی مورد نظر انتخاب خواهد کرد. از آنجا که در این روش جمع احتمالات برابر با ۱ برای تمام بازیکنان استفاده شده و پس از محاسبات، نرمالیزه کردن مجدد با استفاده از رابطه (۲۵) مورد نیاز است، لذا روش سوم که اصلاح شده این روش است می‌تواند استفاده شود.

۴،۱،۳. روش سوم (تجمیع ماتریس‌های احتمالات انتقال با دوبار نرمالیزه کردن)

در این روش نیز مانند روش دوم کل حالات مربوط به استراتژی‌های بازیکنان A و B با هم تجمیع شده و به صورت یک زنجیره مارکف واحد در نظر گرفته می‌شوند. در این روش نیز ابتدا ماتریس احتمال انتقال T به صورت رابطه (۸) ایجاد شده و سپس با استفاده از رابطه‌ای شبیه رابطه (۱۷) به صورت رابطه زیر، مقدار Q که میزان احتمالات حدی را به ازای هر دو بازیکن A و B نشان می‌دهد، تعیین می‌شود. توجه شود که تفاوت این روش با روش دوم در این است که در این روش فقط $m+n-2$ معادله از معادلات همگن استفاده شده و دو معادله دیگر نیز جهت نرمالیزه کردن و جمع احتمالات برابر با ۱ به ازای هر دو بازیکن A و B استفاده می‌شود.

$$\begin{cases} q.(T-I) = O \\ \sum_{i\in A} q_i = 1 \\ \sum_{i\in B} q_i = 1 \end{cases} \quad (۲۷)$$

پس از تعیین مقدار Q می‌بایست عملگری مشابه عملگر رابطه (۲۶) به صورت زیر فقط برای احتمالات مربوط به بازیکن A با ضرب آن عملگر بین دو ماتریس احتمالات P و Q استفاده شود، یعنی داریم:

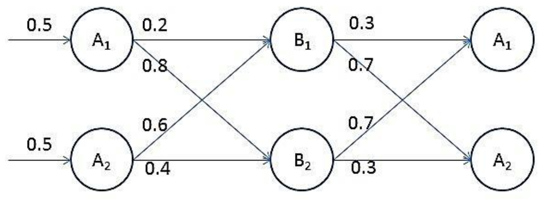
$$\begin{aligned} &P \otimes Q \\ &= [p_1 \quad \cdots \quad p_m] \otimes [q_1 \quad \cdots \quad q_m \quad q_{m+1} \quad \cdots \quad q_{m+n}] \\ &= \left[\frac{p_1 \times q_1}{\sum_{i=1}^{m} p_i \times q_i} \quad \cdots \quad \frac{p_m \times q_m}{\sum_{i=1}^{m} p_i \times q_i} \quad q_{m+1} \quad \cdots \quad q_{m+n}\right] \end{aligned} \quad (۲۸)$$

با توجه به مقادیر حاصله، هر بازیکن بیشترین احتمال مربوط به استراتژی‌ها را به عنوان استراتژی مورد نظر انتخاب خواهد کرد.

۵. ارائه مثال عددی

به عنوان مثال اول، فرض کنید یک ماتریس بازی دونفره گسسته به صورت زیر و نیز زنجیره مارکف با احتمالات انتقال به صورت شکل (۴) موجود است.

$$\begin{array}{ccc} & B_1 & B_2 \\ A_1 & \multicolumn{2}{c}{\multirow{2}{*}{$\begin{bmatrix}(2,0) & (3,2)\\(3,4) & (2,2)\end{bmatrix}$}} \\ A_2 & & \end{array} \quad (۲۹)$$



شکل (۴): زنجیره مارکف مربوط به بازی پویای مثال اول

ماتریس‌های احتمالات انتقال این مثال عبارتند از:

$$P_{1\times2} = [0.5 \quad 0.5]$$
$$R_{2\times2} = \begin{bmatrix}0.2 & 0.8\\0.6 & 0.4\end{bmatrix} \quad (۳۰)$$
$$s_{2\times2} = \begin{bmatrix}0.3 & 0.7\\0.7 & 0.3\end{bmatrix}$$

اگر این بازی به صورت پویا برگزار گردد و بازیکن A نفر اول در تصمیم‌گیری باشد، محاسبات مربوط به هر سه روش به قرار ذیل خواهد بود.

روش اول: در این روش ابتدا $R\times S$ را محاسبه کرده و سپس احتمالات حدی را طبق روابط (۲۲) و (۲۳) بر اساس آن تعیین می‌کنیم. لذا داریم:

$$RS = \begin{bmatrix}0.2 & 0.8\\0.6 & 0.4\end{bmatrix}\begin{bmatrix}0.3 & 0.7\\0.7 & 0.3\end{bmatrix} = \begin{bmatrix}0.62 & 0.38\\0.46 & 0.54\end{bmatrix}$$

$$\begin{cases} q.(RS-I) = O \\ \sum_i q_i = 1 \end{cases} \Rightarrow \begin{cases} -0.38q_1 + 0.46q_2 = 0 \\ q_1 + q_2 = 1 \end{cases} \Rightarrow \begin{cases} q_1 = 0.548 \\ q_2 = 0.452 \end{cases} \quad (۳۱)$$

$$\Pi = [0.5 \quad 0.5]\otimes[0.548 \quad 0.452] = [0.548 \quad 0.452]$$

با توجه به نتیجه تعیین شده بازیکن A استراتژی A_1 را انتخاب می‌کند، از این‌رو با توجه به ماتریس بازی رابطه (۲۹)، بازیکن B در این بازی پویا استراتژی B_2 را انتخاب خواهد کرد و بنابراین نقطه (A_1,B_2) نقطه تعادل موجود است. این نقطه تعادل در حالت حدی و با توجه به احتمالات انتقال تعیین شده است، در حالی که اگر بازی پویای رابطه (۲۹) یک مرحله‌ای و بدون احتمالات انتقال در نظر گرفته شود، نقاط تعادل آن نقاط (A_1,B_2) و (A_2,B_1) می‌باشند.

روش دوم: در این روش از ماتریس احتمالات انتقال تجمیعی و یک معادله نرمالیزه استفاده می‌شود. بر اساس روابط (۲۴) تا (۲۶) داریم:

$$RS = \begin{bmatrix}0 & 0 & 0.2 & 0.8\\0 & 0 & 0.6 & 0.4\\0.3 & 0.7 & 0 & 0\\0.7 & 0.3 & 0 & 0\end{bmatrix}$$

$$\begin{cases} q.(RS-I) = O \\ \sum_i q_i = 1 \end{cases}$$
$$\Rightarrow \begin{cases} -q_1 + 0.3q_3 + 0.7q_4 = 0 \\ -q_2 + 0.7q_3 + 0.3q_4 = 0 \\ 0.2q_1 + 0.6q_2 - q_3 = 0 \\ q_1 + q_2 + q_3 + q_4 = 1 \end{cases} \quad (۳۲)$$
$$\Rightarrow \begin{cases} q_1 = 0.274 \\ q_2 = 0.226 \\ q_3 = 0.310 \\ q_4 = 0.190 \end{cases}$$
$$\Rightarrow \begin{cases} Nq_1 = 0.548 \\ Nq_2 = 0.452 \\ Nq_3 = 0.620 \\ Nq_4 = 0.380 \end{cases}$$

$$\Pi = [0.5 \quad 0.5]\otimes[0.548 \quad 0.452 \quad 0.620 \quad 0.380] = [0.548 \quad 0.452 \quad 0.620 \quad 0.380]$$

با استفاده از مقادیر تعیین شده، نقطه (A_1,B_1) نقطه تعادل محسوب می‌شود. تفاوت این نقطه با نقطه تعیین شده در روش اول به این دلیل است که در روش قبل تعیین استراتژی مورد علاقه B صرفا از روی ماتریس بازی و نه به صورت حدی تعیین شد. البته چون بازی به صورت حدی مطرح است، نقطه تعیین شده توسط این روش می‌تواند بیشتر قابل اتکا باشد.

روش سوم: در این روش از ماتریس احتمالات انتقال تجمیعی و دو معادله نرمالیزه استفاده می‌شود. لذا بر اساس روابط (۲۷) و (۲۸) داریم:

$$RS=\begin{bmatrix}0&0&0.2&0.8\\0&0&0.6&0.4\\0.3&0.7&0&0\\0.7&0.3&0&0\end{bmatrix}$$

$$\begin{cases}q.(RS-I)=O\\ \sum_{i\in A} q_i=1\\ \sum_{i\in B} q_i=1\end{cases}\Rightarrow\begin{cases}-q_1+0.3q_3+0.7q_4=0\\ 0.2q_1+0.6q_2-q_3=0\\ q_1+q_2=1\\ q_3+q_4=1\end{cases} \quad (۳۳)$$

$$\Rightarrow\begin{cases}q_1=0.548\\ q_2=0.452\\ q_3=0.619\\ q_4=0.381\end{cases}$$

$$\Pi=[0.548\quad 0.452\quad 0.619\quad 0.381]$$

همانگونه که ملاحظه می‌شود جواب حاصل از این روش با روش دوم با دقت بسیار بالایی تطابق دارد و در این روش نیز نقطه (A_1,B_1) نقطه تعادل محسوب می‌شود.

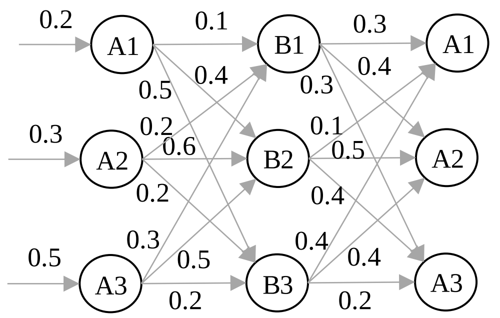
به عنوان مثال دوم، شکل (۵) و ماتریس‌های مربوط به آن را در نظر بگیرید.

$$\begin{array}{c|c} & \begin{array}{ccc}B_1&B_2&B_3\end{array}\\ \begin{array}{c}A_1\\A_2\\A_3\end{array} & \begin{bmatrix}(2,0)&(3,2)&(1,4)\\(3,4)&(2,2)&(1,0)\\(4,4)&(1,2)&(2,3)\end{bmatrix}\end{array} \quad (۳۴)$$

در حالت ایستا

$$\begin{array}{c|c} & \begin{array}{ccc}B_1&B_2&B_3\end{array}\\ \begin{array}{c}A_1\\A_2\\A_3\end{array} & \begin{bmatrix}(2,0)&(3^*,2)&(1,4^*)\\(3,4^*)&(2,2)&(1,0)\\(4^*,4^*)&(1,2)&(2^*,3)\end{bmatrix}\end{array} \quad (۳۵)$$

استراتژی (A_3,B_1) نقطه تعادل محسوب می‌شود.



شکل (۵): زنجیره مارکف مربوط به بازی پویای مثال دوم

ماتریس‌های مربوط به احتمالات انتقال این مثال در رابطه (۳۶) بیان شده‌اند.

$$P_{1\times 3}=[0.2\quad 0.3\quad 0.5]$$

$$R_{3\times 3}=\begin{bmatrix}0.1&0.4&0.5\\0.2&0.6&0.2\\0.3&0.5&0.2\end{bmatrix} \quad (۳۶)$$

$$s_{3\times 3}=\begin{bmatrix}0.3&0.4&0.3\\0.1&0.5&0.4\\0.4&0.4&0.2\end{bmatrix}$$

اگر بازی پویا و بازیکن A نفر اول باشد داریم:

روش اول:

$$RS=\begin{bmatrix}0.27&0.44&0.29\\0.2&0.46&0.34\\0.22&0.45&0.33\end{bmatrix}$$

$$\begin{cases}q.(RS-I)=O\\ \sum_i q_i=1\end{cases}$$

$$\Rightarrow\begin{cases}-0.73q_1+0.2q_2+0.22q_3=0\\ 0.44q_1-0.54q_2+0.45q_3=0\\ 0.29q_1+0.34q_2-0.67q_3=0\\ q_1+q_2+q_3=1\end{cases} \quad (۳۷)$$

$$\Rightarrow\begin{cases}q_1=0.222\\ q_2=0.452\\ q_3=0.326\end{cases}$$

$$\Pi=[0.129\quad 0.396\quad 0.475]$$

پس بازیکن A استراتژی A_3 و طبق ماتریس بازی، بازیکن B استراتژی B_1 را انتخاب کرده و نقطه تعادل (A_3,B_1) می‌شود.

روش دوم:

$$RS=\begin{bmatrix}0&0&0&0.1&0.4&0.5\\0&0&0&0.2&0.6&0.2\\0&0&0&0.3&0.5&0.2\\0.3&0.4&0.3&0&0&0\\0.1&0.5&0.4&0&0&0\\0.4&0.4&0.2&0&0&0\end{bmatrix}$$

$$\begin{cases}q.(RS-I)=O\\ \sum_i q_i=1\end{cases}$$

$$\Rightarrow\begin{cases}-q_1+0.3q_4+0.1q_5+0.4q_6=0\\ -q_2+0.4q_4+0.5q_5+0.4q_6=0\\ -q_3+0.3q_4+0.4q_5+0.2q_6=0\\ 0.1q_1+0.2q_2+0.3q_3-q_4=0\\ 0.4q_1+0.6q_2+0.5q_3-q_5=0\\ 0.5q_1+0.2q_2+0.2q_3-q_6=0\\ q_1+q_2+q_3+q_4+q_5+q_6=1\end{cases} \quad (۳۸)$$

$$\Rightarrow\begin{cases}q_1=0.111\\ q_2=0.226\\ q_3=0.163\\ q_4=0.105\\ q_5=0.262\\ q_6=0.133\end{cases}$$

$$\Pi=[0.130\quad 0.395\quad 0.475\quad 0.210\quad 0.524\quad 0.266]$$

لذا بازیکن A استراتژی A_3 و بازیکن B استراتژی B_2 را انتخاب کرده و نقطه تعادل در شرایط حدی (A_3, B_2) خواهد شد.

روش سوم: ماتریس RS مساوی روش دوم است، لذا از بیان آن خودداری می‌کنیم. در ادامه داریم:

$$\begin{cases} q.(RS-I)=0 \\ \sum_{i\in A} q_i = 1 \\ \sum_{i\in B} q_i = 1 \end{cases}$$

$$\Rightarrow \begin{cases} -q_1+0.3q_4+0.1q_5+0.4q_6=0 \\ -q_2+0.4q_4+0.5q_5+0.4q_6=0 \\ -q_3+0.3q_4+0.4q_5+0.2q_6=0 \\ 0.1q_1+0.2q_2+0.3q_3-q_4=0 \\ 0.4q_1+0.6q_2+0.5q_3-q_5=0 \\ 0.5q_1+0.2q_2+0.2q_3-q_6=0 \\ q_1+q_2+q_3=1 \\ q_4+q_5+q_6=1 \end{cases} \quad (۳۹)$$

$$\Rightarrow \begin{cases} q_1=0.222 \\ q_2=0.452 \\ q_3=0.326 \\ q_4=0.210 \\ q_5=0.532 \\ q_6=0.267 \end{cases}$$

$$\Pi = [0.130 \quad 0.395 \quad 0.475 \quad 0.210 \quad 0.523 \quad 0.267]$$

بنابراین بازیکن A استراتژی A_3 و بازیکن B استراتژی B_2 را انتخاب خواهد کرد و در نتیجه نقطه تعادل در شرایط حدی (A_3, B_2) خواهد شد.

۶. نتیجه‌گیری و کارهای آتی

در انجام یک بازی گسسته به شکل مرحله‌ای، می‌توان با در نظر گرفتن احتمالات انتقال و همچنین نگاشت استراتژی‌های بازی به حالت، یک زنجیره مارکف ایجاد کرد و با تعیین احتمالات حدی این زنجیره، نقاط تعادل حدی را به دست آورد. در این مقاله بازی‌های پویا مورد بررسی قرار گرفت و سه روش برای تعیین نقاط تعادل در این بازی‌ها معرفی گردید. به عنوان مزیت روش پیشنهادی یعنی استفاده از زنجیره مارکوف برای تعیین نقطه تعادل حدی، می‌توان به زمان و محاسبات کمتر اشاره کرد. البته یکی از معایب مهم این روش، لزوم اثرپذیری احتمالات انتقال از انتخاب استراتژی در هر مرحله توسط رقیب و همچنین اثرپذیری احتمالات از توابع مطلوبیت هر بازی است. این امر برای دستیابی به بهترین نتیجه در تعیین نقاط تعادل حدی لازم است که چنانچه احتمالات انتقال به صورت تابعی از مقادیر ذکر شده در نظر گرفته شود، این مشکل قابل رفع است.

به عنوان تحقیقات آتی بازی‌های گسسته چند نفره، پیوسته دو نفره و چند نفره و نیز زمانی که احتمالات انتقال تحت تاثیر استراتژی‌های رقیب نباشد و یا تحت تاثیری سایر موارد باشد، نیز قابل بحث و بررسی خواهد بود.

تعارض منافع: نویسندگان اعلام می‌کنند که هیچ تعارض منافعی ندارند.

مراجع

[1] M.J. Asgharpour, Group Decision Making and Game Thory: an Approach on Operations Research, Samt Press, Tehran University, Tehran, Iran, 2003, [In Persian].

[2] J.B. Krawczyk and V. Petkov, “Multistage Games,” In: Handbook of Dynamic Game Theory, T. Basar and G. Zaccour, Springer, Cham, 2018, doi: 10.1007/978-3-319-44374-4_3.

[3] B.R. Myerson, “Multistage Games with Communication”, Econometrica, vol. 54, no. 2, pp. 323-358, 1986, doi: 10.2307/1913154.

[4] L.M. Littman, “Value-function reinforcement learning

in Markov games", Cogn. Syst. Res., vol. 2, no. 1, pp. 55-66, 2001, doi: 10.1016/S1389-0417(01)00015-8.

[5] P. Vrancx, "Decentralised Reinforcement Learning in Markov Games", Ph.D. dissertation, Brussel University, Brussels, Belgium, 2010.

[6] J.M. Osborne, An introduction to Game Theory, Oxford University Press, Oxford, UK, 2000.

[7] M. Finus, Game theory and international environmental cooperation, Edward Elgar Press, Massachusetts, USA, 2001.

[8] K. Mollering, Inventory Rationing: A New Modeling Approach Using Markov Chain Theory, Springer Press, Koln, Germany, 2019.

[9] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter, Markov Chain Monte Carlo in Practice, Chapman & Hall Press, London, UK, 1996.

[10] S. Balaji, E.G. Julie, Y.H. Robinson, R. Kumar, P.H. Thong, and L.H. Son, "Design of a security-aware routing schemein Mobile Ad-hoc Network using repeated game model," Comput. Stand. Interfaces, vol. 66, 2019, doi: 10.1016/j.csi.2019.103358.

[11] Y. Jie, X. Tang, K.-K. R. Choo, S. Su, M. Li, and C. Guo, "Online task scheduling for edge computing based on repeated Stackelberg game," J. Parallel Distributed Comput., vol. 122, pp. 159-172, 2018, doi: 10.1016/j.jpdc.2018.07.019.

[12] N.T. Cason and V.-L. Mui, "Individual versus group choices of repeated game strategies: A strategy method approach," Games Econ. Behav., vol. 114, pp. 128-145, 2019, doi: 10.1016/j.geb.2019.01.003.

[13] G. Ashkenazi-Golan and E. Lehrer, "Blackwell's comparison of experiments and discounted repeated games," Games Econ. Behav., vol. 117, pp. 163-194, 2019, doi: 10.1016/j.geb.2019.06.003.

[14] E. Ianovski and L. Ong, "The complexity of decision problems about equilibria in two-player Boolean games," Artif. Intell., vol. 261, pp. 1-15, 2018, doi: 10.1016/j.artint.2018.04.006.

[15] F. Gensbittel and C. Rainer, "A Two-Player Zero-sum Game Where Only One Player Observes a Brownian Motion," Dyn. Games Appl., vol. 8, no. 2, pp. 280-314, 2018, doi: 10.1007/s13235-017-0219-5}.

[16] V.O. Baskov, "Equilibrium payoffs in repeated two-player zero-sum games of finite automata," Int. J. Game Theory, vol. 48, no. 2, pp. 423-431, 2019, doi: 10.1007/s00182-018-0634-x.

[17] Z. Wang, Q. Wei, and D. Liu, "Event-triggered adaptive dynamic programming for discrete-time multi-player games," Inf. Sci., vol. 506, pp. 457-470, 2020, doi: 10.1016/j.ins.2019.05.071.

[18] R. Song and L. Zhu, "Stable value iteration for two-player zero-sum game of discrete-time nonlinear systems based on adaptive dynamic programming," Neurocomputing, vol. 340, pp. 180-195, 2019, doi: 10.1016/j.neucom.2019.03.002.

[19] P.M. Abraham and A.A. Kulkarni, "An Approach Based on Generalized Nash Games and Shared Constraints for Discrete Time Dynamic Games," Dyn. Games Appl., vol. 8, no. 4, pp. 641–670, 2018, doi: 10.1007/s13235-017-0231-9.

[20] A.Yadollahi, J. Salimi-Sartaghti, and S. Goli-Bidgoli, "Modeling the security of virtual machines in the cloud using iterative game theory," Soft Comput. J., vol. 10, no. 1, pp. 2-15, 2021, doi: 10.22052/scj.2021.242842.0 [In Persian].

[21] B.A. Eisenbruch, L.R. Grillot, D. Maestripieri, and R.J. Roney, "Evidence of partner choice heuristics in a one-shot bargaining game," Evol. Hum. Behav., vol. 37, no. 6, pp. 429-439, 2016, doi: 10.1016/j.evolhumbehav.2016.04.002.

[22] R.A. Laird, "Sequential interactions – in which one player plays first and another responds – promote cooperation in evolutionary-dynamical simulations of single-shot Prisoner's Dilemma and Snowdrift games," J. Theor. Biol., vol. 452, pp. 69-80, 2018, doi: 10.1016/j.jtbi.2018.05.007.

[23] G. Charness, L. Rigotti, and A. Rustichini, "Social surplus determines cooperation rates in the one-shot Prisoner's Dilemma," Games Econ. Behav., vol. 100, pp. 113-124, 2016, doi: 10.1016/j.geb.2016.08.010.

[24] J.-L. Tan, C. Lei, H. Zhang, and Y.-q. Cheng, "Optimal strategy selection approach to moving target defense based on Markov robust game," Comput. Secur., vol. 85, pp. 63-76, 2019, doi: 10.1016/j.cose.2019.04.013.

[25] S.E. Albarran and J.B. Clempner, "A Stackelberg security Markov game based on partial information for strategic decision making against unexpected attacks," Eng. Appl. Artif. Intell., vol. 81, pp. 408-419, 2019, doi: 10.1016/j.engappai.2019.03.010.

[26] R. Sadeghian, "Determining the Equilibrium Solution in Two-Player Static Discrete Markovian Games," Mod. Res. Decis. Mak., vol. 5, no. 4, pp. 85-99, 2021 [In Persian].

[27] G. Wu, G. Tan, J. Deng, and D. Jiang, "Distributed reinforcement learning algorithm of operator service

slice competition prediction based on zero-sum markov game,'' Neurocomputing, vol. 439, pp. 212-222, 2021, doi: 10.1016/j.neucom.2021.01.061.

[28] Y. Zhao, L. Huang, C. Smidts, and Q. Zhu, "Finite-horizon semi-Markov game for time-sensitive attack response and probabilistic risk assessment in nuclear power plants,'' Reliab. Eng. Syst. Saf., vol. 201, p. 106878, 2020, doi: 10.1016/j.ress.2020.106878.