باسمه تعالی

پروژه‌ی درس آمار و احتمال مهندسی

مقدمه‌ای بر تئوری صف

استاد درس

دکتر محمّدمهدی مجاهدیان

دانشکده‌ی مهندسی برق
دانشگاه صنعتی شریف

آذر ۱۴۰۳

تیم طراحی:
امیرحسن نقدی، علی صادقیان، محمد محمدیان، ماهان سرافرازی

آخرین مهلت تحویل:
۱۱ دی ۱۴۰۳

فهرست مطالب

۱ آشنایی با تئوری صف

۱.۱ تئوری صف چیست؟

تئوری صف یا نظریه صف‌بندی[۱] شاخه‌ای از ریاضیات کاربردی و علم کامپیوتر است که به مطالعه صف‌ها یا سیستم‌های انتظار[۲] می‌پردازد. این نظریه برای تحلیل و مدل‌سازی سیستم‌هایی استفاده می‌شود که در آن‌ها کارها (که می‌توانند افراد، داده‌ها، درخواست‌ها یا بسته‌های اطلاعاتی باشند) در صف انتظار برای دریافت خدمات از یک یا چند سرور قرار می‌گیرند. از کاربردهای این شاخه می‌توان به شبکه‌های مخابراتی و رایانه‌ای، مدیریت حمل و نقل و ترافیک، تولید و تولیدات صنعتی و مدل سازی سیستم‌های خدماتی اشاره کرد.

۲.۱ اجزای اصلی سیستم‌های صف

۱.۲.۱ ترتیب خدمت‌دهی (Service Order)

نحوه‌ی تعیین اولویت و ترتیب اجرای کارها توسط سرور را به عنوان یکی از اجزاء کیفی سیستم صف در نظر می‌گیریم. در صورتی که ترتیب خدمت‌دهی به طور مشخص بیان نشده باشد، فرض بر این است که سرور از روش «اولین ورود، اولین خروج» (FCFS[۳]) پیروی می‌کند. این بدان معناست که کارها به همان ترتیبی که وارد سیستم می‌شوند، خدمت دریافت می‌کنند.

۲.۲.۱ نرخ ورود متوسط (Average Arrival Rate)

این مقدار که با λ نمایش داده می‌شود، به نرخ متوسط ورود کارها به سرور اشاره دارد. به طور مثال، اگر $\lambda = ۳$ کار بر ثانیه باشد، این بدان معناست که به طور متوسط، در هر ثانیه سه کار وارد سیستم می‌شوند.

۳.۲.۱ میانگین زمان بین ورودها (Mean Interarrival Time)

این مقدار، به میانگین زمان بین دو ورود متوالی کارها اشاره دارد و به صورت معکوس نرخ ورود یعنی $\frac{۱}{\lambda}$ تعریف می‌شود. به عنوان نمونه، اگر $\lambda = ۳$ باشد، میانگین زمان بین ورود کارها برابر با $\frac{۱}{۳}$ ثانیه خواهد بود، یعنی به طور متوسط هر ۰٫۳۳ ثانیه یک کار وارد سیستم می‌شود.

۴.۲.۱ نیاز به خدمت (Service Requirement یا Size)

«سایز» یا «اندازه» یک کار، که معمولاً با متغیر تصادفی S نشان داده می‌شود، به زمانی اشاره دارد که آن کار برای اجرا توسط سرور نیاز دارد، مشروط بر اینکه هیچ کار دیگری در صف نباشد و سرور به صورت اختصاصی آن کار را اجرا کند. به عبارت دیگر، این متغیر نشان‌دهنده‌ی مدت زمانی است که سرور برای اجرای کامل آن کار نیاز دارد. برای مثال، اگر زمان خدمت‌دهی به یک کار ۵ ثانیه باشد، اندازه‌ی آن کار برابر با ۵ ثانیه است.

۵.۲.۱ میانگین زمان خدمت‌دهی (Mean Service Time)

این مقدار، که به صورت $\mathbb{E}[S]$ نمایش داده می‌شود، زمان متوسط مورد نیاز برای خدمت‌دهی به یک کار توسط سرور است و تنها به زمان پردازش آن کار اشاره دارد، بدون در نظر گرفتن مدت زمان انتظار در صف. به عنوان مثال، اگر $\mathbb{E}[S] = \frac{۱}{۴}$ ثانیه باشد، این بدان معناست که به طور متوسط هر کار ۰٫۲۵ ثانیه برای اجرا توسط سرور نیاز دارد.

۶.۲.۱ نرخ خدمت‌دهی متوسط (Average Service Rate)

این مقدار که با μ نمایش داده می‌شود، به نرخ متوسط خدمت‌دهی توسط سرور اشاره دارد. این مقدار معکوس میانگین زمان خدمت‌دهی است و به صورت زیر بیان می‌شود:

$$\mu = \frac{۱}{\mathbb{E}[S]}$$

[۱] Queueing Theory
[۲] queueing systems
[۳] First-Come-First-Served

برای مثال، اگر $\mathbb{E}[S] = \frac{1}{4}$ ثانیه باشد، نرخ خدمت‌دهی متوسط برابر با $\mu = 4$ کار در ثانیه خواهد بود، به این معنا که سرور به طور متوسط می‌تواند ۴ کار را در هر ثانیه پردازش کند.

۷.۲.۱ زمان پاسخ [4] یا زمان اقامت در سیستم [5]

زمان پاسخ، که با T نمایش داده می‌شود، به مدت زمانی که یک کار در سیستم می‌ماند از لحظه ورود تا زمان ترک سیستم، اشاره دارد. این مدت به صورت زیر تعریف می‌شود:

$$T = t_{depart} - t_{arrive}$$

که در آن:

- t_{depart} : زمان خروج کار از سیستم (لحظه‌ای که خدمت‌دهی به ایستان رسیده و کار از سیستم خارج می‌شود).
- t_{arrive} : زمان ورود کار به سیستم.

هدف ما معمولاً محاسبه‌ی میانگین زمان پاسخ ($\mathbb{E}[T]$) ، واریانس زمان پاسخ ($Var(T)$) و بررسی رفتار توزیع انتهایی زمان پاسخ یعنی احتمال اینکه زمان پاسخ بیشتر از مقدار مشخصی $\mathbb{P}[T > t]$ باشد.

۸.۲.۱ زمان انتظار یا تأخیر [6]

زمان انتظار، که با T_Q نمایش داده می‌شود، به مدت زمانی که یک کار در صف منتظر می‌ماند تا خدمت‌دهی به آن آغاز شود، اشاره دارد. به این مدت زمان، «زمان در صف» یا «زمان تلف شده» نیز گفته می‌شود. زمان انتظار شامل زمانی نمی‌شود که کار در حال دریافت خدمت از سرور است؛ بلکه تنها مدت زمانی که کار در صف قرار دارد و در حال انتظار برای شروع خدمت است، مدنظر است.
در سیستم‌هایی که از ترتیب خدمت‌دهی FCFS استفاده می‌شود، زمان انتظار به این صورت تعریف می‌شود: مدت زمانی از ورود کار به سیستم تا لحظه‌ای که برای اولین بار خدمت‌دهی به آن آغاز می‌شود.
رابطه‌ی میانگین زمان پاسخ با زمان انتظار و زمان خدمت‌دهی به صورت زیر است:

$$\mathbb{E}[S] + \mathbb{E}[T_Q] = \mathbb{E}[T]$$

که در آن $\mathbb{E}[S]$ میانگین زمان خدمت‌دهی به هر کار است.

۹.۲.۱ تعداد کارها در سیستم

این مفهوم، تعداد کل کارهایی را که در سیستم قرار دارند شامل می‌شود، که عبارت است از مجموع تعداد کارهایی که در صف انتظار قرار دارند، به علاوه‌ی کاری که در حال حاضر در حال دریافت خدمت است (در صورتی که کار در حال خدمت‌دهی وجود داشته باشد). به عبارت دیگر، N شامل تمامی کارهای در حال انتظار و همچنین کاری که در حال سرویس‌گیری از سرور است، می‌شود.

۱۰.۲.۱ تعداد کارها در صف

این مقدار، که با N_Q نشان داده می‌شود، تنها به تعداد کارهایی که در صف انتظار قرار دارند اشاره دارد و کارهایی که در حال دریافت خدمت از سرور هستند را شامل نمی‌شود. بنابراین، این مفهوم صرفاً تعداد مشتریان یا کارهایی را نشان می‌دهد که هنوز خدمتی دریافت نکرده‌اند و در صف منتظر هستند.

---

[4] Response Time
[5] (Turnaround Time, Sojourn Time)
[6] (Waiting Time or Delay) - T_Q

پرسش تئوری ۱. به طور شهودی بگویید چه اتفاقی می‌افتد اگر $\mu > \lambda$ باشد؟

۳.۱ انواع صف ها

۱.۳.۱ شبکه های باز[۷]

شبکه‌های باز در تئوری صف به سیستم‌هایی گفته می‌شود که دارای ورود و خروج خارجی هستند؛ یعنی کارها از خارج وارد شبکه می‌شوند و پس از دریافت خدمات، از شبکه خارج می‌گردند. این شبکه‌ها معمولاً شامل چندین سرور هستند که هر یک می‌تواند کارها را از سرورهای دیگر دریافت کرده یا به آن‌ها ارسال کند. در شبکه‌های باز، کارها می‌توانند به صورت داخلی بین سرورها حرکت کنند یا از سیستم خارج شوند. سیستم های تک پردازنده،ساده ترین حالت این شبکه ها هستند.

۲.۳.۱ شبکه های بسته[۸]

شبکه‌های بسته در تئوری صف به سیستم‌هایی اشاره دارند که ورود و خروج خارجی ندارند؛ یعنی تعداد کل کارها در سیستم ثابت است و کارها پس از دریافت خدمات در سیستم باقی می‌مانند. شبکه‌های بسته معمولاً به دو دسته اصلی تقسیم می‌شوند:

۱. سیستم‌های تعاملی[۹]:
در این نوع شبکه‌ها، مجموعه‌ای از ترمینال‌ها وجود دارد که نمایانگر کاربران هستند. هر کاربر (ترمینال) یک کار به سیستم مرکزی ارسال می‌کند و سپس منتظر پاسخ می‌ماند. سیستم مرکزی به عنوان یک شبکه از صف‌ها عمل می‌کند و هر کار تا زمانی که پردازش شود در این شبکه باقی می‌ماند. تا زمانی که کار قبلی کاربر به او بازنگردد، کاربر نمی‌تواند کار جدیدی ارسال کند. بنابراین تعداد کل کارهای موجود در سیستم برابر با تعداد کاربران است که به آن بار یا سطح چندبرنامگی[۱۰] می‌گویند.

۲. سیستم‌های دسته‌ای[۱۱]:
این نوع شبکه‌ها شامل یک مجموعه ثابت از کارها هستند که بدون هیچ وقفه یا زمان تفکری بین کارها در سیستم جابه‌جا می‌شوند. کارها به صورت متوالی و به روش از پیش تعیین‌شده در صف‌ها حرکت می‌کنند و پس از پردازش توسط هر سرور، وارد مرحله بعدی می‌شوند.

۳.۳.۱ بهره‌برداری از دستگاه[۱۲]

استفاده از دستگاه، که با ρ_i نمایش داده می‌شود، نشان‌دهنده بخشی از زمان است که دستگاه i در حال کار یا مشغول است. به عبارت دیگر، این مقدار درصد زمانی است که دستگاه مشغول است. در مواقعی که دستگاه مشخصی مورد نظر نباشد، نماد زیرنویس i حذف می‌شود و از ρ به طور کلی استفاده می‌کنیم.
فرض کنیم یک دستگاه i را برای مدت زمان طولانی τ (مدت زمان مشاهده) بررسی کنیم. اگر B نشان‌دهنده مدت زمانی باشد که دستگاه در طول دوره مشاهده مشغول کار بوده آنگاه میزان استفاده از دستگاه به صورت زیر تعریف می‌شود:

$$\rho_i = \frac{B}{\tau}$$

یعنی، نسبت زمان مشغول بودن دستگاه به کل زمان مشاهده.

[۷] Open Networks
[۸] Closed Networks
[۹] Interactive Systems
[۱۰] Multi programming level
[۱۱] Batch Systems
[۱۲] Device Utilization - ρ_i

۴.۳.۱ توان عملیاتی [۱۳]

توان عملیاتی یک دستگاه، که با X_i نمایش داده می‌شود، به نرخ اتمام کارها در دستگاه i اشاره دارد؛ مثلاً تعداد کارهایی که در هر ثانیه پردازش می‌شوند. توان عملیاتی کل سیستم، که با X نمایش داده می‌شود، به نرخ اتمام کارها در کل سیستم اشاره دارد.
اگر C نشان‌دهنده تعداد کل کارهایی باشد که در مدت زمان τ در دستگاه i به اتمام رسیده‌اند، آنگاه توان عملیاتی دستگاه i به صورت زیر تعریف می‌شود:

$$X_i = \frac{C}{\tau}$$

یعنی، تعداد کارهای کامل شده در واحد زمان.

۵.۳.۱ ارتباط بین توان عملیاتی و استفاده از دستگاه

ارتباط میان X_i و ρ_i به این صورت قابل بیان است:

$$X_i = \frac{C}{\tau} = \rho_i \times \frac{C}{B} = \frac{B}{\tau} \times \frac{C}{B}$$

پرسش تئوری ۲. $\frac{C}{B}$ بیانگر چیست؟

---

۲ little's Law

۱.۲ تعریف little's Law

Little's Law از مهم‌ترین نتایج نظریه صف است. بیان این قانون به شرح زیر است
برای هر سیستم باز ارگادیک [۱۴] داریم:

$$\mathbb{E}[N] = \lambda\mathbb{E}[T]$$

پرسش تئوری ۳. چرا این رابطه منطقی به نظر میرسد؟

برای هر سیستم بسته ارگادیک داریم:

$$N = X\mathbb{E}[T]$$

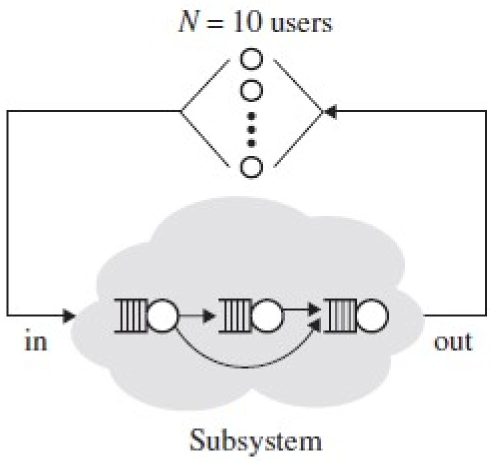
نکته قابل توجه این است که در اینجا چون سیستم بسته است (نه کاری خارج و نه وارد می‌شود) $\mathbb{E}[N] = N$.
نکته دیگر این است که زمان ماندن در سیستم (T) زمان یک لوپ کامل تعریف می‌شود، پس $\mathbb{E}[T] = \mathbb{E}[R]+\mathbb{E}[Z]$ که در آن Rزمان پاسخ سرور و Z مدت زمان لازم برای تفکر کاربر است.

پرسش تئوری ۴. ثابت کنید $\rho = \frac{\lambda}{\mu}$.

پرسش تئوری ۵. یک سیستم صف بسته تعاملی با $N = 10$ داریم. می‌دانیم $\mathbb{E}[Z] = 5$ و $\mathbb{E}[R] = 15$،

---

[۱۳] Device Throughput - X_i
[۱۴] یک فرایند ارگادیک است در صورتی که میانگین آماری آن با میانگین زمانی‌اش برابر باشد. ارگادیک بودن در فصل زنجیره‌ی مارکف تعریف شده



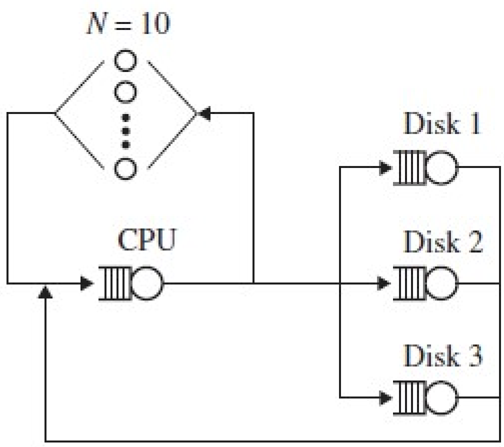
توان عملیاتی سیستم چقدر است؟

برای هر سیستم صف داریم:

$$\bar{N_Q} = X\bar{T_Q}$$

که در آن $\bar{F}$ به معنای میانگین زمانی F است. و N_Q نشان دهنده تعداد کارهای موجود در صف و T_Q نشان دهنده زمان انتظار کارها در صف هستند.

پرسش تئوری ۶. سیستم زیر را در نظر بگیرید،



- توان عملیاتی دیسک سوم $X_3 = 40$ است.
- میانگین زمان پردازش در دیسک سوم $\mathbb{E}[S_3] = 0.0225$ است.
- میانگین تعداد کار ها در دیسک سوم $\mathbb{E}[N_3] = 4$ است.

(آ) نسبت استفاده دیسک سوم (ρ_3) را محاسبه کنید.

(ب) میانگین زمان در صف دیسک سوم $(\mathbb{E}[T_{Q_3}])$ را محاسبه کنید.

(ج) $(\mathbb{E}[N_{Q_3}])$ را محاسبه کنید.

(د) با فرض دانستن $\mathbb{E}[\text{users ready of Number}] = ۷/۵$ ، $\mathbb{E}[\text{Terminals of Number}] = ۱۰$ و $\mathbb{E}[Z] = ۵s$ توان عملیاتی کل سیستم (X) را محاسبه کنید.

(ه) آیا راهی برای محاسبه توان عملیاتی کل با یک معادله هست؟

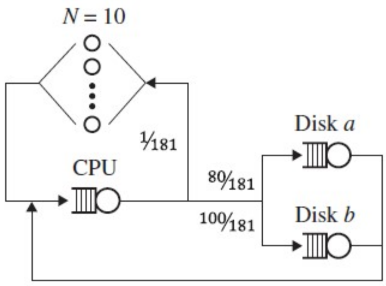
۲.۲ قانون Forced Flow

می‌توان روابط توان عملیاتی اجزای سیستم و توان عملیاتی کل را طبق قانون زیر محاسبه کرد.

$$X_i = X\mathbb{E}[V_i]$$

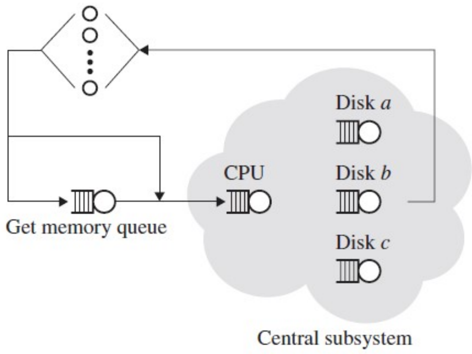
که V_i تعداد دفعات عبور یک کار از زیر سیستم i را بیان می‌کند.

پرسش تئوری ۷. سیستم صفی زیر را در نظر بگیرید.



$\mathbb{E}[V_a]$ ، $\mathbb{E}[V_b]$ و $\mathbb{E}[V_{cpu}]$ را محاسبه کنید.

پرسش تئوری ۸. سیستم نسبتا پیچیده زیر را در نظر بگیرید.



داده های زیر را در مورد این سیستم می‌دانیم.

$N = 23$

$\mathbb{E}[Z] = 21$

$X = 0.45$

$\mathbb{E}[N_{\text{getting memory}}] = 11.65$

$\mathbb{E}[V_{cpu}] = 3$

$\mathbb{E}[S_{cpu}] = 0.21$

با توجه به این داده ها $\mathbb{E}[\text{T central subsystem}]$ را محاسبه کنید.

۳.۲ تقاضای دستگاه

برای هر دستگاه در سیستم صفی متغیری به نام تقاضای دستگاه تعریف می‌شود. برای دستگاه شماره i ام داریم:

$$D_i = \Sigma_{j=۱}^{V_i} S_i^{(j)}$$

این کمیت بیانگر مجموع زمانی است که یک کار در دستگاه i ام پردازش شده است.

پرسش تئوری ۹. ثابت کنید $\mathbb{E}[D_i] = \mathbb{E}[V_i]\mathbb{E}[S_i]$.

اهمیت D_i برای قانونی است که به آن قانون گلوگاه[۱۵] می‌گوییم.

۴.۲ Bottleneck Law

قانون Bottleneck Law به صورت زیر تعریف می‌شود.

$$\rho_i = X\mathbb{E}[D_i]$$

اثبات.

$$\rho_i = X_i\mathbb{E}[S_i] = X\mathbb{E}[V_i]\mathbb{E}[S_i] = X\mathbb{E}[D_i]$$

□

پرسش تئوری ۱۰. Bottleneck Law را ابتدا به صورت شهودی توجیه کنید و سپس با روابط اثبات کنید.

۳ کران های مجانبی برای سیستم‌های بسته

در این بخش قصد داریم پیش‌بینی مناسبی از رفتار سیستم‌های بسته به کمک کران‌های مجانبی داشته باشیم. ابتدا دو کمیت مهم تعریف می‌کنیم.

۱.۳ کران های مجانبی

فرض کنید m تعداد دستگاه‌های سیستم باشد.

$$D = \Sigma_{i=1}^{m}\mathbb{E}[D_i]$$
$$D_{max} = \max_i(\mathbb{E}[D_i])$$

به کمک دو کمیت بالا برای یک سیستم در حالت کلی تعاملی (کافی است قرار دهیم $\mathbb{E}[Z] = 0$ تا سیستم تبدیل به سیستم دسته‌ای شود) کران های زیر برقرار هستند.
برای هر سیستم بسته با N کار داریم:

$$X \leq \min\{\frac{N}{D + \mathbb{E}[Z]}, \frac{1}{D_{max}}\}$$
$$\mathbb{E}[R] \geq \max\{D, N.D_{max} - \mathbb{E}[Z]\}$$

نکته مهم این است که ترم اول هر دو بند کران مجانبی برای N های کوچک است. بالعکس این موضوع برای ترم دوم صدق می‌کند و این ترم کران مجانبی برای N های بزرگ است.

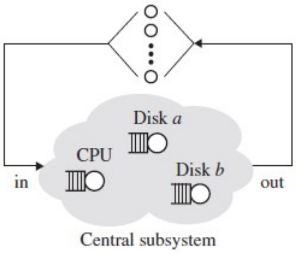
پرسش تئوری ۱۱. با توجه به قوانینی که تا کنون یاد گرفته‌اید این کران ها را ثابت کنید.

$$X \leq \min(\frac{N}{D + \mathbb{E}[Z]}, \frac{1}{D_{max}})$$

$$\mathbb{E}[R] \geq \max(D, N.D_{max} - \mathbb{E}[Z])$$

[۱۵] (Bottleneck Law)

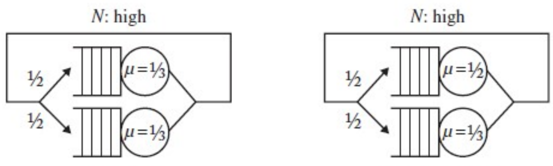
پرسش تئوری ۱۲. در سیستم بسته زیر کمیت های زیر اندازه‌گیری شده‌اند،



- $\mathbb{E}[Z] = 18$
- $\mathbb{E}[D_{CPU}] = 5$
- $\mathbb{E}[D_{disk_a}] = 4$
- $\mathbb{E}[D_{disk_b}] = 3$

به ازای N های مختلف کران های مجانبی را محاسبه کنید و X و $\mathbb{E}[R]$ را به عنوان تابعی از N رسم کنید. (هر دو کران بالا و پایین را رسم کرده و نمودار نهایی را مجانب به آنها رسم کنید). چه نتیجه ای می‌گیرید؟

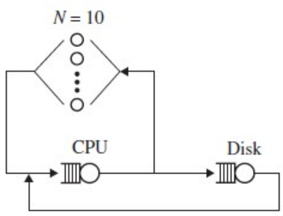
پرسش تئوری ۱۳. سوال سیستم a را در نظر بگیرید. این سیستم برای N های بزرگ کار می‌کند. برای بهبود عملکرد (در اینجا افزایش توان عملیاتی)، سیستم b را جایگزین می‌کنیم.



(a) Original (b) "Improved"

مقادیر D و D_{max} را برای هر دو سیستم به دست آورید. آیا توان عملیاتی سیستم بهبود می‌یابد؟ پاسخ خود را توجیه کنید.

پرسش تئوری ۱۴. سیستم بسته زیر را در نظر بگیرید



می‌دانیم $N = 20$ و $\mathbb{E}[Z] = 5$. یک بار CPU و Disk را به گونه ای انتخاب می‌کنیم که $D_{CPU} = 4.6$ و $D_{Disk} = 4.0$ باشد. بار دیگر به گونه ای سیستم را می‌سازیم که $D_{CPU} = 4.9$ و $D_{Disk} = 1.9$. کدام سیستم توان عملیاتی بهتری دارد؟

پرسش تئوری ۱۵. اندازه گیری های زیر روی یک سیستم انجام شده است.

- $T = 650s$
- $B_{cpu} = 400s$
- $B_{slowdisk} = 100s$
- $B_{fastdisk} = 600s$

- $C = C_{cpu} = 200jobs$
- $C_{slowdisk} = 2000jobs$
- $C_{fastdisk} = 20000jobs$
- $\mathbb{E}[Z] = 15s$
- $N = 20users$

ما چهار ایده برای بهبود عملکرد این سیستم داریم که به شرح زیر هستند.

(آ) CPU را دو برابر سریع تر می‌کنیم.

(ب) دیسک سریع و کند را متعادل می‌کنیم. به این معنی که تقاضای آنها را برابر می‌کنیم.

(ج) یک دیسک سریع دیگر مانند دیسک سریع سیستم می‌خریم تا نیمی از کار دیسک سریع فعلی را به عهده گیرد.

(د) هرسه کار را می‌کنیم. یعنی هم یک دیسک سریع دیگر می‌خریم، هم هر سه را بالانس می‌کنیم و هم CPU را دو برابر سریع می‌کنیم.

تاثیر هر کدام را روی عملکرد سیستم بررسی کنید.

آیا این کران ها برای سیستم های باز هم معتبر هستند؟ پاسخ مثبت است. اما نکته قابل توجه این است که این کران ها دیگر کران های مجانبی محسوب نمی‌شوند. ساده ترین نکته برای نشان دادن این موضوع این است که در سیستم های باز $X = \lambda$ و λ چیزی نیست که ما روی آن کنترلی داشته باشیم. با این حال این کران ها به صورت نامجانبی برقرار هستند. اما همانطور که مشخص است قدرتی که برای توصیف سیستم های بسته دارند را برای سیستم های باز ندارند. به همین جهت به ابزار های دیگری احتیاج داریم که در فصل های بعد بررسی می‌شوند.

۴ فرایند های پواسون

۱.۴ توزیع های نمایی و پواسون

در طول درس با متغیر های تصادفی پواسون و نمایی آشنا شدید. در این بخش ضمن دوره این مباحث، ارتباط بین آن ها را نیز بررسی می‌کنیم.

توزیع نمایی، توزیعی پیوسته است که دارای چگالی احتمال زیر است:

$$f_X(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

پرسش تئوری ۱۶. اگر T متغیری تصادفی با توزیع نمایی با پارامتر λ باشد، مقادیر زیر را برای آن محاسبه کنید.

(آ) $\mathbb{E}[T]$

(ب) $Var(T)$

پرسش تئوری ۱۷. خاصیت بی حافظگی متغیر های تصادفی نمایی را اثبات کنید.

پرسش تئوری ۱۸. اگر $T_1 \sim exp(\lambda_1)$ و $T_2 \sim exp(\lambda_2)$ و مستقل از هم باشند. توزیع $min(T_1, T_2)$ را محاسبه کنید. سپس با استقرا برای n متغیر نمایی مستقل از هم با پارامتر های $T_i \sim exp(\lambda_i)$ توزیع متغیر تصادفی $S = min(T_1, T_2, ..., T_n)$ را بیان کنید.

پرسش تئوری ۱۹. اگر $T_1 \sim exp(\lambda_1)$ و $T_2 \sim exp(\lambda_2)$ و مستقل از هم باشند. ثابت کنید :

$$\mathbb{P}(X_1 < X_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

پرسش تئوری ۲۰. فرض کنید $X_1, X_2, X_3, \ldots, X_n$ متغیرهای تصادفی مستقل و هم‌توزیع نمایی با نرخ λ باشند. تعریف کنید:

$$Z = \max\{X_1, X_2, \ldots, X_n\}.$$

(آ) امید ریاضی $\mathbb{E}[Z]$ را به دست آورید.

(ب) تقریباً شکل $\mathbb{E}[Z]$ به عنوان تابعی از n و λ وقتی n بزرگ باشد، چگونه خواهد بود؟

(ج) توزیع Z را بدست آورید.

پرسش تئوری ۲۱. فرض کنید $X \sim exp(\lambda_X)$ و $Y \sim exp(\lambda_Y)$ باشند، به طوری که $X \perp Y$ (مستقل هستند). همچنین تعریف کنید:

$$Z = \min\{X, Y\}.$$

نشان دهید:

$$\mathbb{P}\{X > t \mid X < Y\} = \mathbb{P}\{Z > t\}.$$

توزیع پواسون، یک توزیع احتمالی گسسته است که احتمال اینکه یک حادثه به تعداد مشخصی در فاصله زمانی یا مکانی ثابتی رخ دهد را شرح می‌دهد؛ به شرط اینکه این حوادث با نرخ میانگین مشخصی و مستقل از زمان آخرین حادثه رخ دهند. میگوییم متغیر تصادفی X توزیع پواسون با میانگین λ دارد اگر داشته باشیم :

$$\mathbb{P}[X = n] = e^{-\lambda}\frac{\lambda^n}{n!} \quad for \ \ n = 0, 1, 2, 3, \ldots$$

پرسش تئوری ۲۲. اگر $X \sim Poisson(\lambda)$ اثبات کنید :

$$\mathbb{E}[X(X-1)...(X-k+1)] = \lambda^k$$

پرسش تئوری ۲۳. اگر X_i ها متغیر هایی مستقل از باشند و داشته باشیم $X_i \sim Poisson(\lambda_i)$ اثبات کنید:

$$X_1 + ... + X_k = Poisson(\lambda_1 + ... + \lambda_n)$$

می‌دانیم در صورتی که در یک توزیع دوجمله‌ای $n \to \infty$ را می‌توان با یک توزیع پواسون تخمین زد. در سوال بعد قصد داریم این موضوع را اثبات کنیم:

پرسش تئوری ۲۴. فرض کنید می‌خواهیم در بازه ای به طول T به صورت کاملا یکنواخت و مستقل از هم n نقطه انتخاب کنیم. ما طول بازه و هم چنین تعداد نقاط را به سمت بی نهایت می‌بریم. ($T \to \infty$ $n \to \infty$) طوری که تعداد نقاط انتخاب شده در واحد زمان ثابت باشد، یعنی نسبت $\frac{n}{T}$ ثابت و برابر λ باشد. متغیر تصادفی X را تعداد نقاط تعریف می‌کنیم که از این n نقطه در بازه ی مشخصی به طول τ قرار گیرند.
می‌دانیم احتمال قرار گرفتن در این بازه برای هر نقطه به صورت $p = \frac{\tau}{T} \to 0$ است اما چون تعداد نقاط نیز به سمت بی‌نهایت می‌رود، به طوری که $\frac{n}{T} = \lambda$ می‌توان توزیع X را به دست آورد. ثابت کنید X دارای توزیع پوآسون با پارامتر $\lambda\tau$ است. $(X \sim Poisson(\lambda\tau))$

حال که با رویکرد بالا به توزیع پواسون رسیدیم، آماده‌ایم تا ارتباط بین توزیع های نمایی و پواسون را بررسی کنیم:

پرسش تئوری ۲۵. در سوال قبل، چون نقاط به صورت تصادفی انتخاب می‌شوند فاصله بین آن ها نیز یک متغیر تصادفی پیوسته است. از خواص مهم نقاط که تعداد آن ها در یک بازه به طول مشخص از توزیع پواسون پیروی می‌کند آن است که فاصله بین نقاط دارای توزیع نمایی است. ثابت کنید فاصله ی بین نقاط مطرح شده در سوال قبل توزیع نمایی دارد.

۲.۴ فرایند های پواسون

فرایند پواسون یکی از فرایندهای تصادفی اساسی در نظریه احتمال و کاربردهای آن در مدل‌سازی رخدادهای نادر است. این فرایند توزیع تعداد وقوع یک حادثه را در بازه‌های زمانی ثابت، که این رخدادها مستقل از هم هستند، مدل می‌کند. فرایند پواسون را می‌توان به صورت یک دنباله از تعداد رخدادها، $\{N(t) : t \geq ۰\}$، تعریف کرد که شرایط زیر را دارا باشد:

- **شرط ابتدایی:** در زمان صفر هیچ رخدادی رخ نداده است، یعنی $N(۰) = ۰$.
- **توزیع پواسون رخدادها:** تعداد رخدادها در بازه‌ای با طول Δt (مثلاً از t تا $t + \Delta t$) دارای توزیع پواسون با پارامتر $\lambda\Delta t$ است:
$$N(t + \Delta t) - N(t) \sim Poisson(\lambda\Delta t).$$
- **افزایش‌های مستقل:** تعداد رخدادها در بازه‌های زمانی جداگانه مستقل از هم هستند. به عبارت دیگر، اگر $۰ \leq t_۰ < t_۱ < \cdots < t_n$ باشد، آنگاه اختلافات $N(t_۱)-N(t_۰), N(t_۲)-N(t_۱), \ldots, N(t_n)-N(t_{n-۱})$ متغیرهای تصادفی مستقل هستند.

با توجه به سوالات قبلی، می‌توان گفت که این خواص توزیع پواسون و نمایی به ما کمک می‌کند فرایند پواسون را بهتر درک کنیم. به طور خاص، تعداد رخدادها در یک بازه زمانی مشخص از توزیع پواسون پیروی می‌کند و فاصله بین رخدادهای پیاپی نیز دارای توزیع نمایی با پارامتر λ است.

تعریف :

فرض کنید $\tau_۱, \tau_۲, \ldots$ متغیرهای تصادفی نمایی مستقل با نرخ λ باشند. تعریف می‌کنیم:

$$T_n = \tau_۱ + \tau_۲ + \cdots + \tau_n \quad \text{برای} n \geq ۱, \quad T_۰ = ۰$$

و همچنین

$$N(s) = \max\{n \mid T_n \leq s\}.$$

در اینجا τ_n را به عنوان زمان بین ورود مشتری‌ها به یک دستگاه خودپرداز (ATM) در نظر می‌گیریم، به طوری که $T_n = \tau_۱ + \cdots + \tau_n$ زمان ورود مشتری nام است. آنگاه $N(s)$ تعداد ورودها تا زمان s را نشان می‌دهد. برای نشان دادن این‌که این یک فرایند پواسون را تشکیل می‌دهد، تک تک شروط را بررسی می‌کنیم.

لم ۱. $N(s)$ دارای توزیع پواسون با میانگین s است.

اثبات. داریم $N(s) = n$ اگر و تنها اگر $T_n \leq s < T_{n+۱}$، یعنی مشتری n_ام قبل از زمان s می‌رسد ولی مشتری $(n+۱)$_ام بعد از زمان s می‌رسد. با توجه به توزیع $T_n = t$ و توجه به این که $T_{n+۱} > s$ به معنی $\tau_{n+۱} > s - t$ است و $\tau_{n+۱}$ مستقل از T_n است، داریم:

$$\mathbb{P}(N(s) = n) = \int_۰^s f_{T_n}(t)\mathbb{P}(\tau_{n+۱} > s - t)\, dt.$$

با جایگذاری چگالی T_n و ساده‌سازی، به دست می‌آوریم:

$$\mathbb{P}(N(s)=n)=\int_{0}^{s} e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}\cdot e^{-\lambda(s-t)}\,dt=\frac{e^{-\lambda s}(\lambda s)^n}{n!},$$

که تابع احتمال توزیع پواسون با میانگین λs است. $\square$

لم ۲. $N(t+s)-N(s)$، برای $t\geq 0$، یک فرایند پواسون با نرخ λ و مستقل از $N(r)$ برای $0\leq r\leq s$ است.

اثبات. فرض کنید تا زمان s چهار ورود در زمان‌های T_1,T_2,T_3,T_4 رخ داده است. زمان انتظار برای ورود پنجم باید شرط $\tau_5>s-T_4$ را داشته باشد. با استفاده از خاصیت بدون حافظه توزیع نمایی، داریم:

$$\mathbb{P}(\tau_5>s-T_4+t\mid \tau_5>s-T_4)=\mathbb{P}(\tau_5>t)=e^{-\lambda t}.$$

بنابراین، توزیع اولین ورود بعد از s نمایی(λ) و مستقل از زمان‌های ورود قبلی T_1,T_2,T_3,T_4 است. از آنجا که $\tau_6,\tau_7,\ldots$ نیز مستقل از T_1,T_2,T_3,T_4 و τ_5 هستند، نتیجه می‌گیریم که $N(t+s)-N(s)$، برای $t\geq 0$، یک فرایند پواسون است. $\square$

لم ۳. $N(t)$ دارای افزایش‌های مستقل است.

اثبات. قسمت قبل نشان دادیم که $N(t_n)-N(t_{n-1})$ مستقل از $N(r)$ برای $r\leq t_{n-1}$ است، و لذا مستقل از $N(t_{n-1})-N(t_{n-2}),\ldots,N(t_1)-N(t_0)$ نیز است. این نتیجه با استقرا اثبات می‌شود. $\square$

پرسش تئوری ۲۶. فرض کنید دو فرایند پواسون مستقل داریم که فرایند اول با نرخ λ_1 و فرایند دوم با نرخ λ_2 رخ می‌دهند. نشان دهید که ترکیب این دو فرایند یک فرایند پواسون با نرخ $(\lambda_1+\lambda_2)$ است.

پرسش تئوری ۲۷. فرض کنید یک فرایند پواسون داریم و می‌دانیم که تا زمان t یک رخداد در این فرایند اتفاق افتاده است. ثابت کنید که این رخداد با احتمال یکسان در هر نقطه‌ای از بازه‌ی $[0,t]$ می‌توانسته رخ داده باشد.

پرسش تئوری ۲۸. فرض کنید T یک متغیر تصادفی نمایی با نرخ λ باشد.

(آ) از تعریف امید شرطی استفاده کنید تا $\mathbb{E}[T\mid T<c]$ را به دست آورید.

(ب) مقدار $\mathbb{E}[T\mid T<c]$ را با استفاده از رابطه زیر تعیین کنید:

$$\mathbb{E}[T]=\mathbb{P}[T<c]\times\mathbb{E}[T\mid T<c]+\mathbb{P}[T>c]\times\mathbb{E}[T\mid T>c].$$

پرسش تئوری ۲۹. چه زمانی دانشجو از جاده عبور می‌کند؟ فرض کنید ترافیک جاده از یک فرایند پواسون با نرخ λ (تعداد ماشین در دقیقه) پیروی می‌کند. یک دانشجو برای عبور از جاده به یک فاصله حداقل به اندازه c دقیقه در ترافیک نیاز دارد. برای محاسبه زمانی که دانشجو باید برای عبور از جاده منتظر بماند، فرض کنید $t_1,t_2,t_3,\ldots$ زمان‌های بین‌ورود ماشین‌ها هستند و تعریف کنید:

$$J=\min\{j\mid t_j>c\}.$$

اگر $T_n=t_1+\cdots+t_n$ باشد، آنگاه دانشجو در زمان T_{J-1} شروع به عبور از جاده می‌کند و در زمان $T_{J-1}+c$ به آن طرف جاده می‌رسد. از تمرین قبلی استفاده کنید تا نشان دهید:

$$\mathbb{E}(T_{J-1}+c)=\frac{e^{\lambda c}-1}{\lambda}.$$

۵ زنجیره مارکف[۱۶]

زنجیره مارکف، یک مدل ریاضی است که برای توصیف سیستم‌های تصادفی با ویژگی عدم وابستگی به گذشته به کار می‌رود. این مدل به ویژه در تحلیل‌هایی که در آن‌ها حالت فعلی سیستم تنها به حالت قبلی آن وابسته است، کاربرد دارد. به عبارت دیگر، پیش‌بینی آینده یک سیستم، تنها نیازمند در نظر گرفتن حالت کنونی و نه تمام وقایع قبلی است. زنجیره مارکف به دو نوع اصلی گسسته و پیوسته تقسیم می‌شود که هر یک از این نوع‌ها کاربردها و ویژگی‌های خاص خود را دارند.

۱.۵ زنجیره مارکف گسسته در زمان

حال به بررسی عمیق زنجیره‌های مارکوف می‌پردازیم و با زنجیره‌های مارکوف زمان گسسته (DTMCs[۱۷]) آغاز می‌کنیم. در یک DTMC، جهان به مراحل زمانی همزمان تقسیم می‌شود. یک رویداد (ورود یا خروج) تنها در انتهای یک مرحله زمانی می‌تواند رخ دهد.
در زنجیره‌های مارکوف زمان پیوسته (CTMCs[۱۸]) رویدادها می‌توانند در هر لحظه‌ای از زمان رخ دهند. این ویژگی، CTMCها را برای مدل‌سازی سیستم‌ها بسیار مناسب می‌سازد.

تعریف :

زنجیره مارکوف زمان گسسته (DTMC) یک فرآیند تصادفی $\{X_n | n = ۰, ۱, ۲, \ldots\}$ است که در آن X_n نمایانگر حالت در مرحله زمانی n بوده و به صورت زیر تعریف می‌شود:
$\forall n \geq ۰$، $\forall i, j$ و $\forall i_۰, \ldots, i_{n-۱}$ داریم:

$$\mathbb{P}\{X_{n+۱} = j \mid X_n = i, X_{n-۱} = i_{n-۱}, \ldots, X_۰ = i_۰\} = \mathbb{P}\{X_{n+۱} = j \mid X_n = i\} = P_{ij}$$

که در آن P_{ij} مستقل از مرحله زمانی و تاریخچه گذشته است.

ویژگی زنجیر مارکفی بیان می‌کند که توزیع شرطی هر حالت آینده $X_{n+۱}$ با توجه به حالات گذشته $X_۰, X_۱, \ldots X_{n-۱}$ و با توجه به حالت فعلی X_n مستقل از حالات گذشته است و فقط به حالت فعلی X_n بستگی دارد.

تعریف :

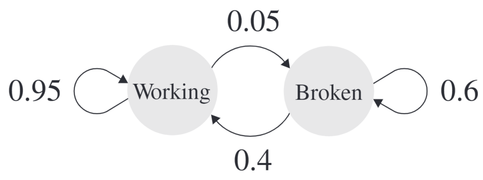
ماتریس احتمال انتقال مرتبط با هر زنجیره مارکوف زمان گسسته (DTMC) یک ماتریس P است که در آن عنصر (i, j)، یعنی P_{ij}، نمایانگر احتمال انتقال به حالت j در انتقال بعدی، با توجه به اینکه حالت کنونی i است، است.

مثال: یک دستگاه یا در حال کار است یا در مرکز تعمیر. اگر امروز کار می‌کند، پس به احتمال ۰٫۹۵ فردا کار می‌کند. اگر امروز در مرکز تعمیرات باشد به احتمال ۰٫۴۰ فردا کار می‌کند.
سوال: (DTMC) را برای مسئله توضیح دهید.
پاسخ: دو حالت **در حال کار** و **خراب** وجود دارد که **خراب** نشان می‌دهد که دستگاه در حال تعمیر است.

$$P = \begin{bmatrix} ۰٫۹۵ & ۰٫۰۵ \\ ۰٫۴ & ۰٫۶ \end{bmatrix}$$

پرسش تئوری ۳۰. حال فرض کنید پس از ۴ روز خرابی دستگاه، دستگاه با دستگاه جدید جایگزین شود. نمودار و ماتریس احتمال انتقال (DTMC) چگونه تغییر می‌کند؟

[۱۶] Markov Chains
[۱۷] Discrete-Time Markov Chains
[۱۸] Continuous-Time Markov Chains



شکل ۱: دیاگرام زنجیره مارکف مسئله خرابی دستگاه

۱.۱.۵ زنجیره مارکف با تعداد حالت محدود

ما این فصل را با تمرکز بر DTMC ها با تعداد محدودی از حالت‌ها آغاز می‌کنیم. بعداً در این فصل، به DTMC هایی با تعداد بی نهایت حالت تعمیم می‌دهیم.

پرسش تئوری ۳۱. فرض کنید $P^n = P \times P \cdots \times P$، n بار ضرب ماتریسی شود. ما از نماد P^n_{ij} برای نشان دادن $(P^n)_{ij}$ استفاده می‌کنیم.
سوال: P^n_{ij} نشان دهنده چیست (منظور از P^n_{ij}، درایه (i,j) از ماتریس P^n می باشد)؟

اکنون به بررسی حدی می‌پردازیم. ورودی (i,j) ام ماتریس توان P_n را برای n بزرگ در نظر بگیرید:

$$\pi_j = \lim_{n\to\infty} P^n_{ij} = (\lim_{n\to\infty} P^n)_{ij}$$

این کمیت نشان دهنده احتمال حدی در حالت j در آینده بسیار دور است، با توجه به اینکه ما در حالت i شروع کردیم.
نشان خواهیم داد که π_j مقدار احتمال حدی حضور در حالت j مستقل از حالت شروع i خواهد بود. برای یک DTMC با M حالت، با حالت های $0, 1, \cdots, M-1$ داریم:

$$\vec{\pi} = (\pi_0, \pi_1, ..., \pi_{M-1}), \qquad \sum_{i=0}^{M-1} \pi_i = 1$$

$\vec{\pi}$ نشان دهنده توزیع حدی برای هر حالت است (اینکه چرا این حد موجود هست را در ادامه خواهیم دید).

پرسش تئوری ۳۲. در مورد اینکه چرا P^n وقتی که n عدد خیلی بزرگی است، دارای سطرهای برابری می‌شود، توضیح دهید (پاسخ این سوال به عبارت **مستقل از حالت شروع** i جواب می‌دهد).

۲.۱.۵ حالت ایستا

تعریف :

یک توزیع احتمال $\vec{\pi} = (\pi_0, \pi_1, ..., \pi_{M-1})$ را توزیع احتمال ایستا[۱۹] گوییم اگر:

$$\vec{\pi} \times P = \vec{\pi} \quad and \quad \sum_{i=0}^{M-1} \pi_i = 1$$

برای درک بهتر:

$$\sum_{i=0}^{M-1} \pi_i P_{ij} = \pi_j \leftrightarrow \forall j \quad and \quad \sum_{i=0}^{M-1} \pi_i = 1$$

[۱۹] Stationary probability distribution

سمت چپ، نشان دهنده احتمال قرار گرفتن در حالت j برای یک انتقال از هم اکنون است، با توجه به اینکه توزیع احتمال فعلی روی حالت‌ها $\vec{\pi}$ است. بنابراین، معادله می‌گوید که اگر توزیع را بر اساس $\vec{\pi}$ شروع کنیم، یک مرحله بعد احتمال قرار گرفتن ما در هر حالت همچنان از توزیع $\vec{\pi}$ پیروی می‌کند. بنابراین از آن به بعد ما همیشه توزیع احتمال یکسانی را روی حالت‌ها خواهیم داشت. از این رو توزیع را «ایستا» می‌نامیم.

۳.۱.۵ حالت حدی و پایدار

قضیه ۱. توزیع احتمال حدی برابر است با توزیع احتمال ایستا

با توجه به یک $DTMC$ حالت محدود با تعداد حالات M :

$$\pi_j = \lim_{n\to\infty} P_{ij}^n > ۰$$

احتمال حدی در حالت j باشد و :

$$\vec{\pi} = (\pi_۰, \pi_۱, ..., \pi_{M-۱}), \quad \sum_{i=۰}^{M-۱} \pi_i = ۱$$

توزیع حدی باشد. با فرض وجود حدی، $\vec{\pi}$ نیز یک توزیع ایستا است و هیچ توزیع ایستای دیگری وجود ندارد.

اثبات. بخش اول: ابتدا اثبات می‌کنیم که توزیع حدی یک توزیع ایستا است:

$$\pi_j = \lim_{n\to\infty} P_{ij}^{n+۱} = \lim_{n\to\infty} \sum_{k=۰}^{M-۱} P_{ik}^n P_{kj} = \sum_{k=۰}^{M-۱} \lim_{n\to\infty} P_{ik}^n P_{kj} = \sum_{k=۰}^{M-۱} \pi_k P_{kj}$$

بخش دوم: اثبات می‌کنیم هر توزیع ایستا باید برابر با توزیع حدی باشد .

فرض کنید $\vec{\pi}'$ یک توزیع ایستا باشد. $\vec{\pi}$ نمایش دهنده توزیع حدی باشد، داریم:

$$\pi_j' = \mathbb{P}\{X_۰ = j\} = \mathbb{P}\{X_n = j\}$$

چرا که $\vec{\pi}'$ توزیع ایستاست.

$$\pi_j' = \mathbb{P}\{X_n = j\}, \quad \forall n$$

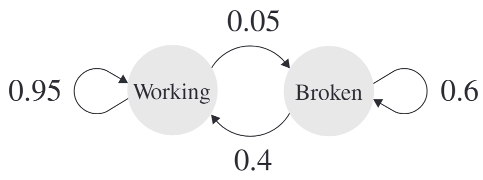
$$= \sum_{i=۰}^{M-۱} \mathbb{P}\{X_n = j | X_۰ = i\} \cdot \mathbb{P}\{X_۰ = i\}, \quad \forall n$$

$$= \sum_{i=۰}^{M-۱} P_{ij}^n \pi', \quad \forall n$$

بنابراین:

$$\pi_j' = \lim_{n\to\infty} \pi_j' = \lim_{n\to\infty} \sum_{i=۰}^{M-۱} P_{ij}^n \pi_i' = \sum_{i=۰}^{M-۱} \lim_{n\to\infty} P_{ij}^n \pi_i' = \sum_{i=۰}^{M-۱} \pi_j \pi_i' = \pi_j \sum_{i=۰}^{M-۱} \pi_i' = \pi_j$$

□



شکل ۲: دیاگرام زنجیره مارکف مسئله خرابی دستگاه

۴.۱.۵ حل معادلات حالت پایدار

دوباره مسئله خرابی دستگاه را در نظر بگیرید که توسط DTMC حالت محدود نشان داده شده است:
سوال: مرکز تعمیر درتلاش است تا بفهمد برای نگهداری از دستگاه چقدر از ما هزینه بگیرد. آنها متوجه می‌شوند که هر روز که دستگاه ما در حال تعمیر است، ۳۰۰ دلار هزینه دارد. صورتحساب تعمیر سالیانه ما چقدر خواهد بود؟
برای پاسخ به این سوال، ابتدا توزیع محدود $\vec{\pi} = (\pi_W, \pi_B)$ را برای این زنجیره استخراج می‌کنیم. معادلات ثابت را برای بدست آوردن $\vec{\pi}$ به صورت زیر حل می‌کنیم:

$$\vec{\pi} = \vec{\pi} \times P, \quad P = \begin{bmatrix} ۰/۹۵ & ۰/۰۵ \\ ۰/۴ & ۰/۶ \end{bmatrix}$$

$$\pi_W + \pi_B = ۱$$

سوال: در مورد دو معادله اول بالا چه چیزی متوجه می‌شوید؟
پاسخ : آنها یکسان هستند! به طور کلی، اگر $\vec{\pi} = \vec{\pi} \cdot P$ به معادلات M منجر شود، تنها $M - ۱$ از این معادلات خطی مستقل خواهند بود. با حل معادلات، $\pi_W = \frac{۸}{۹}, \pi_B = \frac{۱}{۹}$ بدست می‌آوریم. با قضیه ۱ ، توزیع ایستا نیز توزیع احتمال حدی را نشان می‌دهد. بنابراین دستگاه ما به طور متوسط ۱ روز از هر ۹ روز خراب می‌شود. هزینه روزانه مورد انتظار $۳۳/۳۳ = ۳۰۰ \cdot \frac{۱}{۹}$ دلار است (با هزینه سالانه بیش از ۱۲۰۰۰ دلار).

۵.۱.۵ وجود حالت مجانبی

تحت چه شرایطی این توزیع حدی وجود دارد؟
تحت دو شرایط تحلیل‌ها ما بسیار راحت خواهد بود و این توزیع حدی به نقطه شروع ما بستگی نخواهد داشت.
تعریف :
تناوب حالت j، بزرگترین مقسوم علیه مشترک مجموعه اعداد صحیح n است، به طوری که $P^n_{jj} > ۰$. حالتی غیرتناوبی است اگر دوره تناوب ۱ داشته باشد. به زنجیره‌ای غیرمتناوب گفته می‌شود که تمام حالت‌های آن غیرتناوبی است.
تعریف :
حالت j از حالت i قابل دسترسی است اگر $P^n_{ij} > ۰$ برای برخی $n > ۰$ باشد. حالت‌های i و j ارتباط برقرار می‌کنند اگر i از j قابل دسترسی باشد و بالعکس.
تعریف :
زنجیره مارکوف تقلیل‌ناپذیر[۲۰] است، اگر تمام حالات آن با یکدیگر ارتباط برقرار کنند.

پرسش تئوری ۳۳. چرا تقلیل ناپذیری برای مستقل بودن احتمالات محدود کننده از حالت شروع لازم است؟

قضیه ۲. یک DTMC با تعداد حالات محدود ، غیرمتناوب و غیرقابل تقلیل، با ماتریس انتقالی P، $\lim_{n\to\infty} P^n = L$ ، که در آن L یک ماتریس حدی است که همه ردیف های آن بردار یکسانی هستند $(\vec{\pi})$. بردار $\vec{\pi}$ دارای درایه‌ها مثبت است و مجموع آن‌ها برابر ۱ است.

---

[۲۰] irreducible

اثبات. فرض کنید $\vec{e_j}$ بردار ستونی با همان ابعاد P را نشان دهد که j امین درایه آن ۱ است و اجزای باقی‌مانده آن همه ۰ هستند.

$$\vec{e_j} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

ما سعی می‌کنیم این را نشان دهیم:

$$P^n \vec{e_j}$$

به یک بردار همگرا می‌شود که همه اجزای آن یکسان هستند.
اگر M_n بزرگترین درایه $P^n \vec{e_j}$ و m_n کوچکترین درایه $P^n \vec{e_j}$ باشد، نشان خواهیم داد که:

$$M_n - m_n \leq (1 - 2s)(M_{n-1} - m_{n-1})$$

که s کوچکترین درایه در P است.
برای اثبات این موضوع یک کران بالا برای M_n و یک کران پایین برای m_n می‌یابیم و با استفاده از این دو، یک کران پایین برای تغییرات مقدار اختلاف بیشنه و کمینه در هر مرحله می‌یابیم.
فرض کنید $(y = P^{n-1}\vec{e_j})$ ، یک کران بالا برای بزرگترین درایه $(Py = P^n \vec{e_j})$ برابر است با :

$$M_n \leq s \cdot m_{n-1} + (1 - s) \cdot M_{n-1}$$

پرسش تئوری ۳۴. چرا به این کران می‌رسیم؟

همچنین با استدلال مشابه به کران پایین برای کوچک ترین درایه $(P^n \vec{e_j})$ می‌رسیم:

$$m_n \geq s \cdot M_{n-1} + (1 - s) \cdot m_{n-1}$$

با کم کردن این دو از همدیگر، داریم:

$$M_n - m_n \leq (1 - 2s)(M_{n-1} - m_{n-1})$$

پرسش تئوری ۳۵. بزرگترین مقدار برای s برابر ۰.۵ است. چرا؟

همانطور که مشاهده می‌کنید، در صورتیکه s برابر با ۰ باشد، همگرایی تضمین نمی‌شود. همچنین از فرض‌ها تقلیل‌ناپذیری و متناوب بودن استفاده نکردیم.
چون P نامتناوب و تقلیل ناپذیر است، n_0 وجود دارد که تمام درایه‌ها P^{n_0} بزرگتر از ۰ است. چون زنجیره مارکف تقلیل ناپذیر است، n_1 و n_2 وجود دارد که $P_{ij}^{n_1}$ و $P_{ij}^{n_2}$ بزرگتر از ۰ است و چون زنجیره مارکف غیرمتناوب است، ک.م.م این دو عدد برابر ۱ است، n_0' وجود دارد که تمام اعداد بزرگتر مساوی n_0'، با ترکیب خطی n_1 و n_2 با ضرایب طبیعی ساخته می‌شوند.

پرسش تئوری ۳۶. ادعا بالا را ثابت کنید.

$$gcd(n_1, n_2) = 1 \Rightarrow \forall n \geq n_0' \quad n = \alpha \cdot n_1 + \beta \cdot n_2 \quad \alpha, \beta \in \mathbb{N}$$

همین استدلال را برای P_{ii}^{α} برقرار است. با انتخاب $n_0 = \max\{n_{00}', ..., n_{mm}'\}$ تمام درایه‌ها P^{n_0} بزرگتر از ۰ خواهد بود. حال با استفاده از بخش قبل و انتخاب گام n_0 ، می‌توان همگرایی را تضمین کرد.

$$P' = P^{n_0} \Rightarrow P^n \vec{e_j} = (P^{n_0})^{n/n_0} \vec{e_j} = (P')^{n/n_0} \vec{e_j}$$

باتوجه به اینکه P یک ماتریس احتمال است و P^n نیز یک ماتریس احتمال است، مجموع درایه‌ها موجود در یک سطر برابر ۱ خواهد بود. $\square$

خلاصه : ما ثابت کردیم که برای هر زنجیره مارکف نامتناوب، تقلیل‌ناپذیر و حالت محدود، توزیع حدی وجود دارد.

۶.۱.۵ مقدار زمان متوسط بین بازدید یک حالت

فرض کنید m_{ij} تعداد گام های زمانی مورد نیاز برای اولین بار رسیدن به حالت j را نشان دهد، با توجه به اینکه ما در حال حاضر در حالت i هستیم. به همین ترتیب، m_{jj} تعداد مراحل مورد انتظار بین بازدیدها از حالت j را نشان دهد.

قضیه ۳. برای یک زنجیره مارکف تقلیل‌ناپذیر و غیرمتناوب با حالت محدود با ماتریس انتقال P:

$$m_{ij} = \frac{1}{\pi_j}$$

اثبات. با شرطی کردن اولین قدم، داریم:

$$\begin{aligned} m_{ij} &= P_{ij} \cdot 1 + \sum_{k \neq j} P_{ik}(1 + m_{kj}) \\ &= 1 + \sum_{k \neq j} P_{ik} m_{kj} \end{aligned}$$

همچنین:

$$\begin{aligned} m_{jj} &= P_{jj} \cdot 1 + \sum_{k \neq j} P_{jk}(1 + m_{kj}) \\ &= 1 + \sum_{k \neq j} P_{jk} m_{kj} \end{aligned}$$

اگر ضرایب m_{ij}ها را در یک ماتریس M قرار دهیم، ماتریس M را می‌توان به صورت $M = E + PN$ تجزیه کرد که D ماتریس قطری است که عناصر آن برابر است با ماتریس M $(d_{jj} = m_{jj})$. ماتریس N نیز برابر $N = M - D$، تمام عناصر قطری آن صفر است و عناصر غیر قطری آن برابر با ماتریس M است. بنابرین داریم:

$$\begin{aligned} M &= E + PN \\ N + D &= E + PN \\ (I - P) \cdot N &= E - D \end{aligned}$$

حال طرفین را در $\vec{\pi}$ ضرب می‌کنیم:

$$\begin{aligned} \vec{\pi} \cdot (I - P) \cdot n &= \vec{\pi}(E - D) \\ \vec{\pi}P = \vec{\pi} : \quad (\vec{\pi} - \vec{\pi})N &= 0 = \vec{\pi}(E - D) \\ &\Rightarrow \vec{\pi}E = \vec{\pi}D \\ \Rightarrow (1, 1, \dots, 1) &= (\pi_0 m_{00}, \pi_1 m_{11}, \dots, \pi_{M-1} m_{M-1}) \\ &\Rightarrow \pi_i m_{ii} = 1, \quad \forall i. \end{aligned}$$

$\square$

۷.۱.۵ زنجیره مارکف با تعداد حالت نامحدود

تا کنون ما فقط در مورد DTMC های حالت محدود با حالت های M صحبت کرده‌ایم. اکنون به سراغ DTMC ها با تعداد حالت بی نهایت می‌رویم. برای یک زنجیره مارکوف با تعداد بی نهایت حالت، می توان یک ماتریس احتمال انتقال، P را تصور کرد، اگرچه ماتریس دارای بیشمار درایه است.

ما توزیع احتمال حدی در حالت ها را اینگونه نشان می‌دهیم:

$$\vec{\pi} = (\pi_۰, \pi_۱, \pi_۲, ...) : \quad \pi_j = \lim_{n\to\infty} P_{ij}^n \quad and \quad \sum_{j=۰}^{\infty} \pi_j = ۱$$

ما دیدیم که برای یک DTMC حالت محدود، اگر توزیع حدی وجود داشته باشد، توزیع حدی و توزیع ایستا معادل هستند (قضیه ۱). همین نتیجه برای DTMC های حالت نامحدود صدق می‌کند.

اثبات. ما دو چیز را در مورد این توزیع حدی با فرض وجود آن اثبات خواهیم کرد:

۱. ثابت خواهیم کرد که $\{\pi_j, j = ۰, ۱, ۲, ...\}$ یک توزیع ایستا است. از این رو حداقل یک توزیع ایستا وجود دارد.

۲. ما ثابت خواهیم کرد که هر توزیع ایستا باید برابر با توزیع حدی باشد.

برای اثبات بخش اول باید نشان دهیم توزیع حدی، یک توزیع ایستا است. متأسفانه، به طور کلی نمی‌توانیم حد و مجموع را وقتی که مجموع نامحدود است، با هم عوض کنیم. بنابراین کاری که ما در این نوع اثبات‌ها باید انجام دهیم این است که مجموع نامحدود را به مجموع نامتناهی تبدیل کنیم، جابه‌جایی را انجام دهیم و سپس به مجموع نامحدود تبدیل کنیم:

$$\begin{aligned} \pi_j &= \lim_{n\to\infty} P_{ij}^{n+۱} \\ &= \lim_{n\to\infty} P_{ik}^n P_{kj} \\ &= \lim_{n\to\infty} \lim_{M\to\infty} \sum_{k=۰}^{M} P_{ik}^n P_{kj}, \quad \forall M \\ &\geq \lim_{n\to\infty} \sum_{k=۰}^{M} P_{ik}^n P_{kj}, \quad \forall M \\ &= \sum_{k=۰}^{M} \lim_{n\to\infty} P_{ik}^n P_{kj}, \quad \forall M \\ &= \sum_{k=۰}^{M} \pi_k P_{kj}, \quad \forall M \end{aligned}$$

بنابرین:

$$\pi_j \geq \sum_{k=۰}^{M} \pi_k P_{kj} \quad \forall M$$

$$\Rightarrow \pi_j \geq \lim_{M\to\infty} \sum_{k=۰}^{M} \pi_k P_{kj}$$

$$\Rightarrow \pi_j \geq \sum_{k=۰}^{\infty} \pi_k P_{kj}$$

ما می‌خواهیم ثابت کنیم که نابرابری فوق در واقع یک برابری است. فرض کنید، با تناقض، l طوری که:

$$\pi_j > \sum_{k=0}^{\infty} \pi_k P_{kj}$$

$$\sum_{j=0}^{\infty} \pi_j > \sum_{j=0}^{\infty} (\sum_{k=0}^{\infty} \pi_k P_{kj}) = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \pi_k P_{kj} = \sum_{k=0}^{\infty} \pi_k \cdot 1 = 1$$

بنابرین، مجموع احتمالات حدی بزرگتر از ۱ است، که غیرممکن است و در نتیجه یک تناقض است.

$$\pi_j = \sum_{k=0}^{\infty} \pi_k P_{kj}$$

برای اثبات اینکه هر توزیع ایستا باید برابر با توزیع حدی باشد، فرض کنید $\vec{\pi}'$ یک توزیع احتمالی ایستا باشد. طبق معمول، $\vec{\pi}$ نمایانگر توزیع احتمالی حدی است. هدف ما اثبات این است که برای هر j داشته باشیم: $\pi'_j = \pi_j$.

فرض کنیم از همان ابتدا توزیع $\vec{\pi}'$ را داریم. بنابراین:

$$\pi'_j = P\{X_0 = j\} = P\{X_n = j\}$$

زیرا π'_j ایستا است. پس داریم:

$$\pi'_j = P\{X_n = j\} = \sum_{i=0}^{\infty} P\{X_n = j | X_0 = i\} \cdot P\{X_0 = i\} = \sum_{i=0}^{\infty} P_{ij}^n \pi'_i$$

حال می‌توان نوشت:

$$\pi'_j = \sum_{i=0}^{M} P_{ij}^n \pi'_i + \sum_{i=M+1}^{\infty} P_{ij}^n \pi'_i \quad \forall M)$$

باتوجه به اینکه $0 \leq P_{ij}^n \leq 1$ ، اکنون از **قضیه ساندویچ** برای این معادله استفاده می‌کنیم تا نشان دهیم π'_j بین کران‌های π_j محدود است.

$$\sum_{i=0}^{M} P_{ij}^n \pi'_i \leq \pi'_j \leq \sum_{i=0}^{M} P_{ij}^n \pi'_i + \sum_{i=M+1}^{\infty} \pi'_i$$

حال با گرفتن حد در $n \to \infty$:

$$\lim_{n\to\infty} \sum_{i=0}^{M} P_{ij}^n \pi'_i \leq \lim_{n\to\infty} \pi'_j \leq \lim_{n\to\infty} \sum_{i=0}^{M} P_{ij}^n \pi'_i + \lim_{n\to\infty} \sum_{i=M+1}^{\infty} \pi'_i$$

به دلیل محدود بودن Σ امکان جابه‌جایی $\lim$ وجود دارد و سپس با استفاده از همگرایی حد داریم:

$$\sum_{i=0}^{M} \pi_j \pi'_i \leq \pi'_j \leq \sum_{i=0}^{M} \pi_j \pi'_i + \sum_{i=M+1}^{\infty} \pi'_i$$

$$\pi_j \sum_{i=0}^{M} \pi'_i \leq \pi'_j \leq \pi_j \sum_{i=0}^{M} \pi'_i + \sum_{i=M+1}^{\infty} \pi'_i$$

با گرفتن حد در $M \to \infty$:

$$\lim_{M\to\infty} \pi_j \sum_{i=0}^{M} \pi'_i \leq \lim_{M\to\infty} \pi'_j \leq \lim_{M\to\infty} \pi_j \sum_{i=0}^{M} \pi'_i + \lim_{M\to\infty} \sum_{i=M+1}^{\infty} \pi'_i$$

از آنجا که مقادیر انتهایی به صفر می‌رسند:

$$\pi_j \leq \pi'_j \leq \pi_j$$

بنابراین اثبات شد که $\pi'_j = \pi_j$ و توزیع ایستا با توزیع حدی برابر است. □

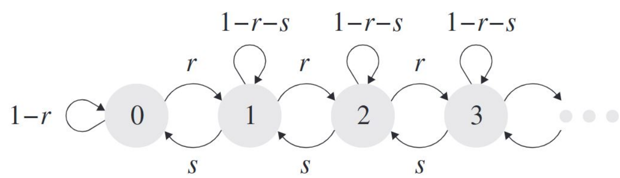
۸.۱.۵ حل زنجیره مارکف با تعداد حالت نامحدود

برای به دست آوردن توزیع احتمال حدی $\vec{\pi}$ ، با فرض اینکه وجود داشته باشد، فقط نیاز است که معادلات ایستا را حل کنیم. اما بی‌نهایت معادله ایستا وجود دارد! بیایید ببینیم چگونه آن‌ها را حل می‌کنیم.

یک صف بدون محدودیت را در نظر بگیرید (تعداد کارها در صف محدودیتی ندارند). فرض کنید کارها به یک سرور می‌رسند. این کارها در صف سرور قرار می‌گیرند. سرور بر روی کاری که در سر صف قرار دارد کار می‌کند و وقتی آن کار تمام شد، به کار بعدی می‌پردازد. کارها تا زمانی که نوبت آن‌ها نشده در صف باقی می‌مانند و سرور آن‌ها را حذف نمی‌کند. فرض کنید در هر گام زمانی، با احتمال $p = \frac{۱}{۴۰}$ یک کار وارد شود و به‌طور مستقل، با احتمال $q = \frac{۱}{۳۲}$ یک کار انجام شود. توجه کنید که در طول یک گام زمانی، ممکن است هم ورود و هم خروج داشته باشیم، یا هیچ‌کدام رخ ندهد.

تعداد متوسط کارها در سیستم چقدر است؟

ابتدا نگاهی به ماتریس انتقال و نمودار حالت کنیم:



شکل ۳: دیاگرام زنجیره مارکف صف بی‌نهایت

$$\boldsymbol{P} = \begin{bmatrix} ۱-r & r & ۰ & ۰ & \dots \\ s & ۱-r-s & r & ۰ & \dots \\ ۰ & s & ۱-r-s & r & \dots \\ ۰ & ۰ & s & ۱-r-s & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

$$r = p(۱-q), \quad s = q(۱-p)$$

اگر معادلات حالت ایستا را بنویسیم:

$$\pi_۰ = \pi_۰(۱-r) + \pi_۱ s$$
$$\pi_۱ = \pi_۰ r + \pi_۱(۱-r-s) + \pi_۲ s$$
$$\pi_۲ = \pi_۱ r + \pi_۲(۱-r-s) + \pi_۳ s$$
$$\pi_۳ = \pi_۲ r + \pi_۳(۱-r-s) + \pi_۴ s$$
$$\dots$$
$$\pi_۰ + \pi_۱ + \pi_۲ + \dots = ۱$$

پرسش تئوری ۳۷. ثابت کنید:

$$\pi_i = (\frac{r}{s})^i \cdot (۱ - \frac{r}{s})$$

پرسش تئوری ۳۸. با استفاده از رابطه‌ای که در سوال قبل یافتید، متوسط تعداد کارها در صف را بیابید.

تعریف حالت‌های بازگشتی و گذرا

اگر f_j برابر احتمال اینکه زنجیره‌ای که از حالت j شروع می‌شود، دوباره به همان حالت برگردد باشد، یک حالت j یا بازگشتی [۲۱] یا گذرا [۲۲] است:

۱. اگر $f_j = 1$، حالت j بازگشتی است.
۲. اگر $f_j < 1$، حالت j گذرا است.

پرسش تئوری ۳۹. میانگین تعداد بازدید از یک حالت گذرا چگونه است؟

پرسش تئوری ۴۰. میانگین تعداد بازدید از یک حالت بازگشتی چگونه است؟

پرسش تئوری ۴۱. روابط زیر را اثبات کنید.

$$\mathbb{E}[\text{تعداد بازدید از حالت } i \text{ در } s \text{ مرحله زمانی} \mid \text{شروع از } i] = \sum_{n=0}^{s} P_{ii}^n$$

$$\mathbb{E}[\text{تعداد کل بازدید از حالت } i \mid \text{شروع از } i] = \sum_{n=0}^{\infty} P_{ii}^n$$

ارتباط بین حالات بازگشتی

۱. اگر حالت i بازگشتی باشد و با حالت j ارتباط داشته باشد (یعنی $i \leftrightarrow j$) ، آنگاه j نیز بازگشتی است.

۲. اگر حالت i گذرا باشد و با حالت j ارتباط داشته باشد، آنگاه j نیز گذرا است.

با شهود برای اثبات شروع می‌کنیم. فرض کنید که ما بی‌نهایت بار به حالت i بازمی‌گردیم. با توجه به تعریف **ارتباط داشتن**، هر بار که در حالت i هستیم، با احتمال مشخصی ممکن است راهی به سوی حالت j را انتخاب کنیم، و هنگامی که در حالت j هستیم، با احتمال مشخصی ممکن است راهی به سوی حالت i انتخاب شود. بنابراین، برای هر بازدید از حالت i، یک احتمال غیرصفر وجود دارد که از حالت j هم بازدید کنیم. از این رو، تعداد بازدیدها از حالت j متناسب با تعداد بازدیدها از حالت i است. چون تعداد بازدیدها از حالت i بی‌نهایت است، پس تعداد بازدیدها از حالت j نیز بی‌نهایت خواهد بود.

پرسش تئوری ۴۲. اثبات شهودی بالا را بصورت ریاضی بیان کنید.

عدم وجود توزیع ایستا برای زنجیره‌های گذرا

باتوجه به بخش قبل، در یک زنجیره مارکف تقلیل‌ناپذیر، یا همه حالات گذرا هستند یا همه بازگشتی هستند. برای یک زنجیره ماركوف گذرا، احتمال قرار گرفتن در حالت j پس از n گام، با افزایش n به سمت صفر میل می‌کند:

$$\lim_{n\to\infty} P_{ij}^n = 0, \forall j$$

$$\sum_{j=0}^{\infty} \pi_j = 0$$

بنابراین توزیع حدی وجود ندارد.

پرسش تئوری ۴۳. با توجه به اینکه توزیع حدی وجود ندارد، برای یک زنجیره مارکف غیرمتناوب تقلیل‌ناپذیر، نشان دهید که توزیع ایستا وجود ندارد.

---

[۲۱] recurrent
[۲۲] transient

۱۰.۱.۵ بازگشتی مثبت در مقابل بازگشتی تهی

زنجیره‌های مارکف بازگشتی به دو دسته تقسیم می‌شوند: بازگشتی مثبت[۲۳] و بازگشتی تهی[۲۴]. در یک زنجیره بازگشتی مثبت، میانگین زمان بین بازگشت‌ها محدود است، اما در یک زنجیره بازگشتی تهی، میانگین زمان بین بازگشت‌ها بی‌نهایت است.

۱. اگر حالت i بازگشتی مثبت باشد و با حالت j ارتباط داشته باشد (یعنی $i \leftrightarrow j$)، آنگاه j نیز بازگشتی مثبت است.

۲. اگر حالت i بازگشتی تهی باشد و با حالت j ارتباط داشته باشد، آنگاه j نیز بازگشتی تهی است.

زنجیره‌های بازگشتی تهی شبیه به یک تناقض است: برای یک حالت بازگشتی تهی j، میانگین زمان بین بازدیدها از حالت j برابر ∞ است، و با این حال، حالت j بی‌نهایت بار بازدید می‌شود.
برای درک بهتر زنجیره‌ها بازگشتی و گذرا، به مسئله زیر توجه کنید:

قدم زدن تصادفی بی‌نهایت
یک محور مختصات یک بعدی را در نظر بگیرید، یک مورچه در مبدا مختصات قرار دارد و در هر گام زمانی، با احتمال p به سمت راست و با احتمال q به سمت چپ حرکت می‌کند.
از آنجا که تمامی حالات با یکدیگر در ارتباط هستند، همه حالات گذرا هستند یا همه بازگشتی. بنابراین، برای تعیین اینکه زنجیره بازگشتی است یا گذرا، کافیست حالت مبدا مختصات را بررسی کنیم.
برای بررسی اینکه آیا حالت مبدا مختصات گذراست یا بازگشتی، فرض می‌کنیم:

$$V = \sum_{n=۱}^{\infty} P_{\circ\circ}^{n}$$

نمایانگر تعداد مورد انتظار بازدیدها از حالت مبدا مختصات باشد. اگر V محدود باشد، آنگاه حالت صفر گذرا است و در غیر این صورت بازگشتی خواهد بود.
از آنجا که نمی‌توان در تعداد فردی از مراحل از مبدا به مبدا بازگشت، داریم:

$$V = \sum_{n=۱}^{\infty} P_{\circ\circ}^{۲n} = \sum_{n=۱}^{\infty} \binom{۲n}{n} p^n q^n$$

اکنون این معادله را با استفاده از لم لاوروف [۲۵] ساده می‌کنیم.

لم ۴. لم لاوروف برای $n \geq ۱$

$$\frac{۴^n}{۲n+۱} < \binom{۲n}{n} < ۴^n$$

پرسش تئوری ۴۴. لم لاوروف را اثبات کنید.

با استفاده از لم لاوروف داریم:

$$\sum_{n=۱}^{\infty} \frac{۴^n}{۲n+۱} p^n q^n < V < \sum_{n=۱}^{\infty} ۴^n p^n q^n$$

[۲۳] positive recurrent
[۲۴] null recurrent
[۲۵] Lavrov

اگر $p = q = \frac{1}{2}$ را در سمت چپ جایگذاری کنیم، به این نتیجه می‌رسیم که:

$$V > \sum_{n=1}^{\infty} \frac{4^n}{2n+1} \cdot \frac{1}{4^n} = \sum_{n=1}^{\infty} \frac{1}{2n+1} = \infty$$

در مقابل، اگر فرض کنیم $p \neq q$ و به سمت راست نگاه کنیم، خواهیم داشت:

$$V < \sum_{n=1}^{\infty} (4pq)^n < \infty \quad (4pq < 1)$$

بنابراین مشاهده می‌کنیم که $V = \sum_{n=1}^{\infty} P_{00}^n$ تنها در صورتی بی‌نهایت است که $p = \frac{1}{2}$ باشد. در نتیجه، زنجیره بازگشتی است اگر و تنها اگر $p = \frac{1}{2}$.

پرسش تئوری ۴۵. نشان دهید احتمال بازگشت به حالت مبدا، برابر $2\min(p, q)$ است.

برای حرکت تصادفی متقارن روی محور مختصات با فرض $p = \frac{1}{2}$ ، میانگین تعداد مراحل بین بازدیدها از حالت مبدا بی‌نهایت است.

اثبات. با روش برهان خلف اثبات می‌کنیم. فرض می‌کنیم $m_{0,0}$ محدود است. در این صورت، $m_{1,0}$ نیز باید محدود باشد. این نتیجه از رابطه زیر حاصل می‌شود:

$$m_{0,0} = \frac{1}{2}(m_{1,0} + 1) + \frac{1}{2}(m_{-1,0} + 1)$$

بنابراین، با توجه به تقارن مسئله نمی‌توان $m_{1,0}$ را بی‌نهایت در نظر گرفت.

حال که $m_{1,0}$ محدود است، آن را با شرطی سازی روی حالت بعدی محاسبه می‌کنیم:

$$m_{1,0} = \frac{1}{2} \cdot 1 + \frac{1}{2}(m_{2,0} + 1)$$

اما می‌دانیم که:

$$m_{2,0} = 2m_{1,0}$$

زیرا زمان متوسط برای رفتن از ۲ به ۰، با زمان متوسط برای رفتن از ۲ به ۱ به‌علاوه زمان متوسط برای رفتن از ۱ به ۰ برابر است. و از آنجا که زنجیره از نظر مکان مستقل است، داریم $m_{2,1} = m_{1,0}$. بنابراین داریم:

$$m_{1,0} = \frac{1}{2} \cdot 1 + \frac{1}{2}(m_{2,0} + 1) = 1 + \frac{1}{2} \cdot 2m_{1,0}$$

$$m_{1,0} = 1 + m_{1,0}$$

اما این نتیجه با فرض محدود بودن $m_{1,0}$ تناقض دارد و اثبات ما کامل می‌شود. □

۲.۵ ارگادیک بودن زنجیره‌های مارکف

یک زنجیره مارکف گسسته‌زمان ارگادیک[۲۶] زنجیره‌ای است که دارای سه ویژگی مهم است:

- غیرتناوبی بودن
- تجزیه‌ناپذیر بودن
- بازگشتی مثبت بودن

[۲۶] Ergodic

پرسش تئوری ۴۶. نشان دهید برای زنجیره‌های مارکوف با تعداد حالات محدود، بازگشتی مثبت نتیجه تجزیه‌ناپذیر بودن است. بنابراین، برای زنجیره‌های با تعداد حالات محدود، دو ویژگی غیرتناوبی و تجزیه‌ناپذیر بودن برای ارگادیک بودن کافی است.

قضیه ۴. یک زنجیره مارکوف گسسته‌زمان غیرتناوبی و تجزیه‌ناپذیر به یکی از دو دسته زیر تعلق دارد:

۱. تمام حالت‌ها، گذرا یا بازگشتی تهی هستند. در این صورت $\pi_j = \lim_{n\to\infty} P_{ij}^n = 0, \forall j$ و هیچ توزیع ایستایی وجود ندارد.

۲. تمام حالت‌ها بازگشتی مثبت هستند. در این صورت توزیع حدی $(\pi_0, \pi_1, \pi_2, \dots)$ $= \vec{\pi}$ وجود دارد و احتمال بودن در هر حالت وجود دارد.

$$\pi_j = \lim_{n\to\infty} P_{ij}^n > 0, \forall i$$

π_j احتمال حدی حضور در حالت j است. در این صورت، $\vec{\pi}$ یک توزیع ایستا است و توزیع ایستای دیگری وجود ندارد. همچنین π_j برابر با $\frac{1}{m_{jj}}$ است، که m_{jj} میانگین تعداد مراحل بین بازدیدها از حالت j است.

پرسش تئوری ۴۷. قضیه بالا را ثابت کنید.
(راهنمایی: از تمام قضایایی که یاد گرفتیم استفاده کنید.)

۳.۵ میانگین زمانی

فرض کنید $N_j(t)$ تعداد دفعاتی باشد که در t–امین مرحله زمانی زنجیره مارکف وارد حالت j می‌شود. نسبت میانگین زمانی که زنجیره مارکف در حالت j قرار دارد برابر است با:

$$p_j = \lim_{t\to\infty} \frac{N_j(t)}{t}$$

توجه کنید که p_j در واقع به عنوان میانگین بر روی یک مسیر نمونه خاص ω، تعریف شده است. می‌خواهیم بگوییم که اولا این میانگین با احتمال ۱ همگرا می‌شود (یعنی تقریباً برای همه مسیرهای نمونه همگرا است)، دوما همیشه به یک مقدار واحد همگرا می‌شود و سوما این مقدار مثبت است. قضیه زیر بیان می‌کند که اگر زنجیره مثبت بازگشت و تجزیه‌ناپذیر باشد، همه این ویژگی‌ها تحقق می‌یابند.

قضیه ۵. برای یک زنجیره مارکوف بازگشتی مثبت، تجزیه‌ناپذیر، با احتمال ۱ داریم:

$$p_j = \lim_{t\to\infty} \frac{N_j(t)}{t} = \frac{1}{m_{jj}} > 0$$

که در آن m_{jj} میانگین تعداد مراحل بین بازدیدها از حالت j است.

برای اثبات این موضوع، به قضایای زیر نیاز داریم:

قضیه ۶. SLLN[27] : فرض کنید $X_1, X_2, \dots$ یک دنباله از متغیرهای تصادفی مستقل و هم‌توزیع با میانگین $\mathbb{E}[X]$ باشد. بگذارید $S_n = \sum_{i=1}^n X_i$ باشد. در این صورت، با احتمال ۱ داریم:

$$\lim_{n\to\infty} \frac{S_n}{n} = \mathbb{E}[X]$$

[27] Strong Law of Large Numbers

با داشتن SLLN ، آماده هستیم که یک فرایند تجدید را تعریف کنیم.
فرایند تجدید[۲۸] به فرایندی اطلاق می‌شود که زمان بین رخدادها متغیرهای تصادفی مستقل و هم‌توزیع با توزیع F باشد.

قضیه ۷. قضیه تجدید : برای یک فرایند تجدید، اگر $\mathbb{E}[X]$ میانگین زمان بین تجدیدها باشد، آنگاه با احتمال ۱ داریم:

$$\lim_{t\to\infty}\frac{N(t)}{t}=\frac{1}{\mathbb{E}[X]}$$

اثبات :
برای اثبات قضیه، صرفاً از قضیه SLLN استفاده می‌کنیم، فرض کنید S_n زمان n امین بازگشت به حالت مفروض باشد، داریم:

$$\forall t: \quad S_{N(t)} \le t < S_{N(t)+1}$$

$$\frac{S_{N(t)}}{N(t)} \le \frac{t}{N(t)} < \frac{S_{N(t)+1}}{N(t)}$$

اما داریم:

$$\frac{S_{N(t)}}{N(t)}=\frac{\sum_{i=1}^{N(t)}X_i}{N(t)}\to\mathbb{E}[X] \quad \text{as} \quad t\to\infty \quad \text{w.p.1}$$

همچنین:

$$\frac{S_{N(t)+1}}{N(t)}=\frac{S_{N(t)+1}}{N(t)+1}\cdot\frac{N(t)+1}{N(t)}\to\mathbb{E}[X] \quad \text{as} \quad t\to\infty \quad \text{w.p.1}$$

با استفاده از قضیه ساندویچ داریم:

$$\frac{t}{N(t)}\to\mathbb{E}[X] \quad \text{w.p.1}$$

$$\Rightarrow \frac{N(t)}{t}\to\frac{1}{\mathbb{E}[X]} \quad \text{as} \quad t\to\infty \quad \text{w.p.1}$$

پرسش تئوری ۴۸. با استفاده از قضیه تجدید، قضیه میانگین زمانی را اثبات کنید.

۴.۵ توزیع حدی بعنوان نرخ

تا اینجا دیده‌ایم که برای یک زنجیره مارکوف ارگادیک، داریم:

نسبت بلندمدت زمانی که فرآیند در حالت i است = احتمال حدی اینکه زنجیره مارکوف در حالت i باشد $=\pi_i$

اکنون مشاهده می‌کنیم که:

نرخ انتقال‌ها از حالت i به حالت j $=\pi_i P_{ij}$

برای درک این موضوع، توجه کنید که زنجیره مارکوف در π_i کسر از کل مراحل زمانی در حالت i قرار دارد. علاوه بر این، در کسر P_{ij} از این مراحل، زنجیره به حالت j حرکت می‌کند. بنابراین، در کسر $\pi_i P_{ij}$ از کل مراحل زمانی، زنجیره مارکف در حالت i است و مرحله بعدی آن حرکت به j است. پس اگر به t مرحله زمانی (با t بزرگ) نگاه کنیم، $\pi_i P_{ij} t$ تعداد انتقال‌ها از i به j خواهد بود. با تقسیم بر t، می‌بینیم که نرخ انتقال‌ها (تعداد انتقال‌ها در هر مرحله زمانی) که از i به j است، برابر با $\pi_i P_{ij}$ است.

[۲۸] renewal process

پرسش تئوری ۴۹. $\sum_j \pi_i P_{ij}$ چه چیزی را نشان می‌دهد؟
پرسش تئوری ۵۰. $\sum_j \pi_j P_{ji}$ چه چیزی را نشان می‌دهد؟

یادآوری می‌کنیم که معادله حالت ایستا برای حالت i به صورت زیر است:

$$\pi_i = \sum_j \pi_j P_{ji}$$

همچنین می‌دانیم که $\pi_i = \pi_i \sum_j P_{ij} = \sum_j \pi_i P_{ij}$. بنابراین داریم:

$$\pi_i = \sum_j \pi_i P_{ij} = \sum_j \pi_j P_{ji}$$

معادلات ایستا نرخ کل انتقال خارج از حالت i را با نرخ کل انتقال به حالت i مرتبط می‌کنند:

نرخ کل ورود به حالت i $=$ نرخ کل خروج از حالت i

چرا منطقی است که نرخ کل انتقال‌های خروجی از حالت i باید برابر با نرخ کل انتقال‌های ورودی به حالت i باشد؟
هر بار که یک انتقال از حالت i خارج می‌شود، نمی‌توان انتقال دیگری را از حالت i داشت تا زمانی که انتقالی به حالت i وارد شود. بنابراین تعداد انتقال‌های خروجی از حالت i حداکثر یک واحد با تعداد انتقال‌های ورودی به i تفاوت دارد. اکنون نرخ انتقال‌های خروجی از حالت i، با تعداد کل انتقال‌ها در یک دوره طولانی t ، تقسیم بر t برابر است. چون t بزرگ است، اختلاف ۱ در تعداد انتقال‌های خروجی و ورودی به i در هنگام تقسیم بر t از بین می‌رود و بنابراین نرخ‌ها برابر می‌شوند.
می‌توانیم معادلات ایستا را به‌صورت زیر و بدون در نظر گرفتن خودحلقه‌ها نیز بنویسیم:

$$\sum_{j \neq i} \pi_i P_{ij} = \sum_{j \neq i} \pi_j P_{ji}$$

که از کم کردن $\pi_i P_{ii}$ از هر دو طرف معادله ایستا به دست می‌آید. مجموعه معادلات برای تمام حالات i اغلب به‌عنوان معادلات ایستا شناخته می‌شوند، زیرا نرخ خروج از حالت i برای رفتن به حالتی غیر از i را با نرخ ورود به حالت i از حالتی غیر از i برابر می‌کنند. بنابراین می‌توانیم همیشه خودحلقه‌ها را نادیده بگیریم و معادلات ایستا را بنویسیم. معادلات ایستا می‌توانند هم برای یک مجموعه از حالات و هم برای یک حالت منفرد اعمال شوند. برای مثال، اگر زنجیره مارکف به دو مجموعه از حالات تقسیم شود که آن‌ها را S و S^c می‌نامیم (اینجا S^c مکمل S را نشان می‌دهد). آنگاه می‌توانیم نرخ انتقال‌ها از S به S^c را با نرخ انتقال‌ها از S^c به S برابر کنیم.
چرا منطقی است که کل جریان از S به S^c برابر با جریان از S^c به S باشد؟
این استدلال مشابه چیزی است که برای یک حالت منفرد مشاهده کردیم. هر بار که انتقالی از S به S^c صورت می‌گیرد، ما از حالات S خارج می‌شویم. بنابراین نمی‌توان انتقال دیگری از S به S^c داشت تا زمانی که مجدداً به حالات S بازگردیم، و این نیازمند یک انتقال از S^c به S است.

۵.۵ زنجیره مارکف پیوسته در زمان

تعریف زنجیره مارکوف گسسته DTMC را به یاد آورید :

زنجیره مارکوف گسسته DTMC یک فرآیند تصادفی است $\{X_n, n = ۰, ۱, ۲, \ldots\}$، که در آن X_n نشان‌دهنده حالت در زمان گسسته n است به‌گونه‌ای که، $\forall n \geq ۰$، $\forall i, j$ و $\forall i_۰, \ldots, i_{n-۱}$،

$$\mathbb{P}\{X_{n+۱} = j | X_n = i, X_{n-۱} = i_{n-۱}, \ldots, X_۰ = i_۰\} = \mathbb{P}\{X_{n+۱} = j | X_n = i\} = P_{ij}$$

که در آن P_{ij} مستقل از گام زمانی و تاریخچه گذشته است.

سه ویژگی زنجیره مارکوف گسسته را در نظر بگیرید:

۱. انتقال‌ها همیشه در گام‌های زمانی گسسته انجام می‌شوند، $n = 0, 1, 2, \dots$

۲. گذشته اهمیت ندارد. فقط حالت فعلی مهم است. به‌ویژه مهم نیست که زنجیره مارکوف چه مدت در حالت i بوده است. این ویژگی مارکوفی است.

۳. احتمال‌های انتقال «ایستا» هستند، یعنی مستقل از گام زمانی n می‌باشند.

زنجیره‌های مارکوف پیوسته CTMC معادل پیوسته‌ی زنجیره‌های مارکوف گسسته هستند. بنابراین ما ویژگی‌های ۲ و ۳ زنجیره مارکوف گسسته را حفظ می‌کنیم. اما ویژگی ۱ را به این شکل تغییر می‌دهیم: «انتقال‌ها بین حالت‌ها می‌توانند در هر لحظه اتفاق بیفتند.»

زنجیره مارکوف پیوسته CTMC یک فرآیند تصادفی پیوسته است $\{X(t), t \geq 0\}$ به‌گونه‌ای که $\forall s, t \geq 0$ و $\forall i, j, x(u)$:

$$\mathbb{P}\{X(t+s) = j | X(s) = i, X(u) = x(u), 0 \leq u \leq s\} = \mathbb{P}\{X(t+s) = j | X(s) = i\}$$
$$= \mathbb{P}\{X(t) = j | X(0) = i\} = P_{ij}(t)$$

ما فرض می‌کنیم که فضای حالت قابل شمارش است.

تعریف :

τ_i را به عنوان زمانی که زنجیره مارکوف پیوسته حالت i را ترک می‌کند، تعریف کنید، به شرط آنکه زنجیره مارکوف پیوسته در حال حاضر در حالت i باشد.

با توجه به ویژگی‌های مارکوفی و ایستایی زنجیره مارکوف پیوسته، احتمال اینکه زنجیره مارکوف پیوسته حالت i را در t ثانیه بعدی ترک کند، مستقل از مدت زمان حضور زنجیره در حالت i است. به عبارت دیگر،

$$\mathbb{P}\{\tau_i > t + s | \tau_i > s\} = \mathbb{P}\{\tau_i > t\}.$$

پرسش تئوری ۵۱. سوال: این چه چیزی را درباره توزیع احتمالاتی τ_i بیان می‌کند؟

دیدگاه ۱

یک زنجیره مارکوف پیوسته CTMC یک فرآیند تصادفی است که با این ویژگی تعریف می‌شود که هر بار وارد حالت i می‌شود، موارد زیر برقرار هستند:

۱. مدت زمانی که فرآیند در حالت i پیش از انجام یک انتقال می‌ماند، به صورت نمایی با نرخ ν_i توزیع شده است.

۲. هنگامی که فرآیند حالت i را ترک می‌کند، به احتمال p_{ij} (مستقل از زمان سپری شده در حالت i است)، وارد حالت j می‌شود.

در این دیدگاه، ما به مدت زمان $\tau_i \sim \exp(\nu_i)$ در حالت i می‌مانیم. وقتی i را ترک می‌کنیم، با احتمال p_{ij} به j منتقل می‌شویم، که در آن $\sum_j p_{ij} = 1$.

توجه کنید که p_{ij}، احتمال اینکه وقتی i را ترک کردیم به حالت j برویم، یک ثابت است. این احتمال مستقل از زمان t است (به دلیل ایستایی) و همچنین مستقل از مدت زمان سپری شده در حالت i، یعنی τ_i است (به دلیل ویژگی مارکوفی).

لحظه‌ای را درست پیش از ترک حالت i در نظر بگیرید، در این لحظه، مدت زمانی که در حالت i گذرانده‌ایم، بی‌اهمیت نسبت به گذشته است (فرضیه مارکوفی). تنها نکته مهم این است که در لحظه s در حالت i قرار داریم. زمان خاص s نیز بی‌اهمیت است (فرضیه ایستایی).

دیدگاه ۲

بیایید $X_j \sim \exp(\nu_i p_{ij})$ را به عنوان زمان انتقال از i به j تعریف کنیم، برای $\forall j \neq i$. همچنین $\tau_i = \min_j\{X_j\}$ را به عنوان زمان تا ترک حالت i در نظر بگیرید. حالت بعدی را m بگذارید که در آن $m = \arg\min_j\{X_j\}$، یعنی X_m کمترین مقدار را در میان $\{X_j\}$ دارد.

این تعریف ممکن است گیج‌کننده به نظر برسد. مثالی می‌زنیم تا بفهمیم که دیدگاه ۲ چه چیزی را بیان می‌کند: تصور کنید که شما در حالت i در شب جمعه نشسته‌اید و منتظر تماس تلفنی هستید که شما را دعوت کنند. فرض کنید سه دوست دارید، پس سه امکان برای دعوت وجود دارد: j، k، یا l (این‌ها نام سه نفری هستند که ممکن است با شما تماس بگیرند). شما سه خط تلفن دارید. خط اول مستقیماً به حالت j متصل است تا j بتواند با شما تماس بگیرد و شما را دعوت کند. خط دوم مستقیماً به حالت k متصل است. خط سوم مستقیماً به حالت l متصل است. شما دعوت اولین کسی که تماس بگیرد را می‌پذیرید. زمان تا دریافت تماس از حالت j یک متغیر تصادفی $X_j \sim \exp(\nu_i p_{ij})$ است. به طور مستقل، زمان تا دریافت تماس از حالت k یک متغیر تصادفی $X_k \sim \exp(\nu_i p_{ik})$ است. زمان تا دریافت تماس از حالت l یک متغیر تصادفی $X_l \sim \exp(\nu_i p_{il})$ است. تمام X_j ، X_k ، و X_l به طور همزمان رخ می‌دهند. به محض اینکه تماسی دریافت کنید، حالت i را ترک می‌کنید و به مکانی که شما را دعوت کرده می‌روید، و فرآیند در آن حالت مجدداً آغاز می‌شود.

اثبات اینکه دیدگاه ۲ $\rightarrow$ دیدگاه ۱

با استفاده از دیدگاه ۲، زمان صرف شده در حالت i، یعنی τ_i را در نظر بگیرید.

$$\tau_i = \min_j\{X_j\} \sim \exp\left(\sum_j \nu_i p_{ij}\right) = \exp(\nu_i).$$

علاوه بر این، احتمال اینکه از حالت i به حالت m منتقل شویم برابر است با:

$$\mathbb{P}\{X_m = \min_j\{X_j\}\} = \frac{\nu_i p_{im}}{\sum_j \nu_i p_{ij}} = p_{im}.$$

اثبات شهودی اینکه دیدگاه ۱ $\rightarrow$ دیدگاه ۲

با استفاده از دیدگاه ۱، زنجیره مارکوف را به گونه‌ای توصیف می‌کنیم که در حالت i به مدت زمان $\exp(\nu_i)$ بماند. پس از آن، یک «سکه جهت‌دار» پرتاب می‌شود، با احتمال p_{ij}، سکه جهت j، با احتمال p_{ik}، سکه جهت k و با احتمال p_{il}، سکه جهت l را نشان می‌دهد. برای مشخص‌تر کردن این موضوع، فرض کنید که $p_{ij} = \frac{۱}{۲}$، $p_{ik} = \frac{۱}{۳}$، و $p_{il} = \frac{۱}{۶}$ باشد.

حال بیایید این موضوع را به معادل هندسی آن بیان کنیم: زنجیره مارکوف در حالت i قرار دارد و هر δ گام یک سکه را پرتاب می‌کند. سکه با احتمال $\nu_i\delta$ موفقیت دارد. زنجیره مارکوف منتظر یک موفقیت است. به محض اینکه موفقیتی به دست آید، سکه جهت‌دار پرتاب می‌شود تا نوع موفقیت را تعیین کند.

این توصیف معادل توصیف زیر است: زنجیره مارکوف در حالت i قرار دارد و هر δ گام یک سکه را پرتاب می‌کند. با احتمال $\nu_i \cdot \frac{۱}{۲}\delta$، سکه موفقیت از نوع j است. با احتمال $\nu_i \cdot \frac{۱}{۳}\delta$، سکه موفقیت از نوع k است. با احتمال $\nu_i \cdot \frac{۱}{۶}\delta$، سکه موفقیت از نوع l است. با احتمال $۱ - \nu_i\delta$، سکه موفقیت ندارد. زنجیره مارکوف منتظر موفقیت است (از هر نوع) و به مقصد بعدی مناسب (بر اساس نوع موفقیت) حرکت می‌کند.

با در نظر گرفتن این پرتاب‌های سکه تنها از دیدگاه اینکه آیا منجر به موفقیت از نوع j می‌شود یا خیر، می‌بینیم که زمان تا موفقیت از نوع j برابر $\exp(\nu_i p_{ij})$ است. به همین ترتیب، به طور مستقل و همزمان، می‌بینیم که زمان تا موفقیت از نوع k برابر $\exp(\nu_i p_{ik})$ است. و باز هم به طور مستقل و همزمان، می‌بینیم که زمان تا موفقیت از نوع l برابر $\exp(\nu_i p_{il})$ است.

۱.۵.۵ حل کردن زنجیره مارکوف پیوسته CTMC

فرض کنید $\pi_j = \lim_{t\to\infty} P_{ij}(t)$ احتمال حدی برای بودن در حالت j است. چگونه می‌توانیم π_jها را تعیین کنیم؟

۱. فرض کنید یک زنجیره مارکوف گسسته DTMC داشتید. در این صورت، می‌دانستید که چگونه π_i ها را بدست آورید: کافی است که تناوب‌ناپذیری و تجزیه‌ناپذیری را بررسی کنید و سپس سعی کنید معادلات ایستا را حل کنید. اگر یک راه‌حل برای معادلات ایستا وجود داشته باشد، آن راه‌حل احتمال حدی خواهد بود. اگر راه‌حلی برای معادلات ایستا وجود نداشته باشد، تمامی احتمالات حدی صفر خواهند بود.

۲. اما ما یک DTMC نداریم؛ ما یک CTMC داریم.

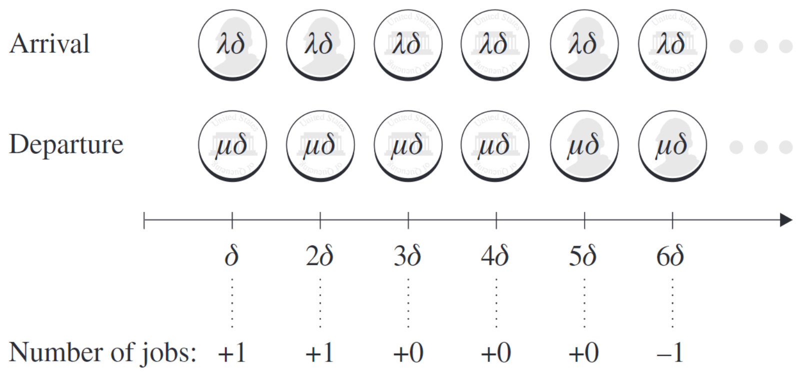
۳. با این حال، می‌توانیم CTMC را با یک DTMC مدل‌سازی کنیم. در CTMC، اگر در حالت i باشیم،

$$\text{زمان تا رسیدن به ورودی بعدی} \sim \exp(\lambda)$$

$$\text{زمان تا خروج بعدی} \sim \exp(\mu)$$

$$\text{زمان تا اولین وقوع این موارد (یعنی خروج از حالت } i) \sim \exp(\lambda+\mu).$$

می‌توانیم این حالت را با پرتاب دو سکه به صورت همزمان در هر گام δ مدل‌سازی کنیم، به طوری که $\delta \to ۰$، همانطور که در شکل ۴ نشان داده شده است.



شکل ۴: پرتاب دو سکه به صورت همزمان در هر گام δ ، یکی با احتمال $\lambda\delta$ شیر و دیگری با احتمال $\mu\delta$ شیر.

سکه اول نشان‌دهنده‌ی ورودها است. وقتی سکه‌ی اول در هر گام δ انداخته می‌شود، با احتمال $\lambda\delta$ ورود و با احتمال $۱-\lambda\delta$ هیچ اتفاقی را نشان می‌دهد.

سکه دوم نشان‌دهنده‌ی خروج‌ها است. وقتی سکه‌ی دوم در هر گام δ انداخته می‌شود، با احتمال $\mu\delta$ خروج و با احتمال $۱-\mu\delta$ هیچ اتفاقی را نشان می‌دهد.

وضعیت گام بعدی به نتیجه سکه‌ها وابسته است. با گسترش چهار حالت ممکن، داریم:

- با احتمال $\lambda\delta(۱-\mu\delta)$، ورود و بدون خروج رخ می‌دهد .
- با احتمال $(۱-\lambda\delta)\mu\delta$، خروج و بدون ورود رخ می‌دهد.
- با احتمال $\lambda\delta\mu\delta$، ورود و خروج رخ می‌دهد.
- با احتمال (همه موارد بالا -۱)، نه ورود و نه خروج رخ می‌دهد.

اما این دقیقاً شبیه یک زنجیره مارکوف گسسته DTMC به نظر می‌رسد که در هر گام δ با احتمالات نشان داده شده در شکل ۵ یک انتقال انجام می‌دهد.

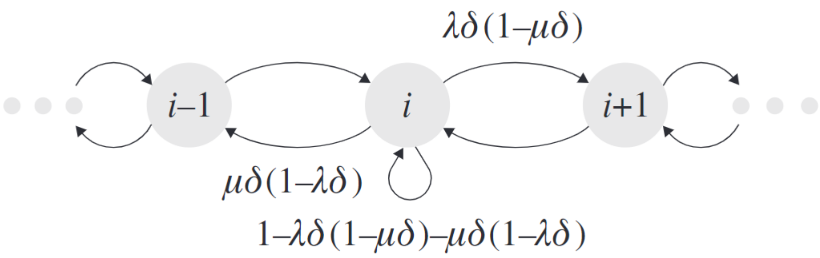
۴. بنابراین، راه‌حل CTMC، برابر با راه‌حل DTMC در شکل ۵ است که در آن انتقال‌ها در هر گام δ اتفاق می‌افتند. بیایید ببینیم آیا می‌توانیم این DTMC را ساده‌تر کنیم.

$$\lambda\delta(۱-\mu\delta) \to \lambda\delta + o(\delta)$$

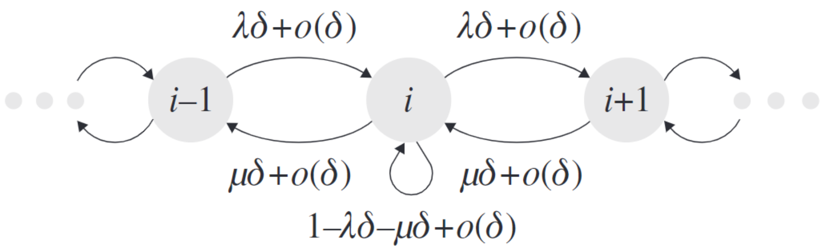
$$\mu\delta(۱-\lambda\delta) \to \mu\delta + o(\delta)$$

$$\lambda\delta\mu\delta \to o(\delta)$$

$$(\text{همه موارد بالا} - ۱) \to ۱ - \lambda\delta - \mu\delta + o(\delta)$$



شکل ۵: DTCM معادل



شکل ۶: DTMC معادل با صرف‌نظر از δ های مراتب بالاتر.

که به ما امکان می‌دهد DTMC را به شکل ۶ نمایش دهیم. توجه کنید انتقال‌ها در هر گام δ رخ می‌دهند. همچنین توجه کنید که اگر DTMC در حالت i باشد، با احتمال بالا در گام δ بعدی همچنان در حالت i خواهد بود.

۵. در اینجا به حل یک DTMC با حد گرفتن از $\delta \to 0$ می‌پردازیم و معادلات تعادل را با فرض ساده‌سازی خود-حلقه‌ها به دست می‌آوریم.

- **معادله تعادل برای حالت ۰:** نرخ خروج از حالت ۰ برابر است با نرخ ورود به حالت ۰:

$$\pi_0(\lambda\delta + o(\delta)) = \pi_1(\mu\delta + o(\delta))$$

با تقسیم بر δ و حد گرفتن از $\delta \to 0$:

$$\lim_{\delta\to 0} \pi_0 \left(\frac{\lambda\delta + o(\delta)}{\delta}\right) = \lim_{\delta\to 0} \pi_1 \left(\frac{\mu\delta + o(\delta)}{\delta}\right)$$

ساده‌سازی می‌شود به:

$$\pi_0 \cdot \lambda = \pi_1 \cdot \mu \Rightarrow \pi_1 = \frac{\lambda}{\mu}\pi_0$$

- **معادله تعادل برای حالت ۱:** نرخ خروج از حالت ۱ برابر است با نرخ ورود به حالت ۱:

$$\pi_1(\lambda\delta + o(\delta) + \mu\delta + o(\delta)) = \pi_0(\lambda\delta + o(\delta)) + \pi_2(\mu\delta + o(\delta))$$

با تقسیم بر δ و حد گرفتن از $\delta \to 0$، به دست می‌آوریم:

$$\pi_1(\lambda + \mu) = \pi_0(\lambda) + \pi_2(\mu)$$

با جایگذاری $\pi_1 = \frac{\lambda}{\mu}\pi_0$:

$$\frac{\lambda}{\mu}\pi_0(\lambda + \mu) = \lambda\pi_0 + \mu\pi_2 \Rightarrow \pi_2 = \left(\frac{\lambda}{\mu}\right)^2 \pi_0$$

• **معادله تعادل برای حالت ۲:** نرخ خروج از حالت ۲ برابر است با نرخ ورود به حالت ۲:

$$\pi_2(\lambda\delta + o(\delta) + \mu\delta + o(\delta)) = \pi_1(\lambda\delta + o(\delta)) + \pi_3(\mu\delta + o(\delta))$$

با حد گرفتن از $\delta \to 0$ و ساده‌سازی، به دست می‌آوریم:

$$\pi_2(\lambda + \mu) = \pi_1(\lambda) + \pi_3(\mu) \Rightarrow \pi_3 = \left(\frac{\lambda}{\mu}\right)^3 \pi_0$$

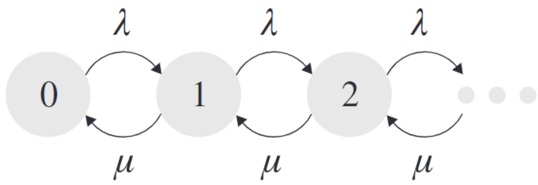
بنابراین، حل عمومی به شکل زیر خواهد بود:

$$\pi_k = \left(\frac{\lambda}{\mu}\right)^k \pi_0$$

۶. به فرم معادلاتی که برای DTMC باقی مانده است نگاه کنید. این معادلات به نظر شبیه **معادلات تعادل** برای CTMC اصلی می‌آیند:

$$\pi_0\lambda = \pi_1\mu$$
$$\pi_1(\lambda + \mu) = \pi_0\lambda + \pi_2\mu$$
$$\pi_2(\lambda + \mu) = \pi_1\lambda + \pi_3\mu$$

این معادلات را در چارچوب CTMC اصلی در نظر بگیرید. در این معادلات، λ و μ احتمالات نیستند، اما در ساختن معادلات به شکلی عمل می‌کنند که گویی هستند.



شکل ۷: CTMC اصلی ما.

۷. بنابراین، به نظر می‌رسد که برای حل یک CTMC ، کافی است معادلات تعادل را برای آن CTMC بنویسیم و آن‌ها را حل کنیم. این راه‌حل با حل DTMC گام-دلتا معادل است، که به خودی خود معادل CTMC است. توجه داشته باشید که ما فقط نشان داده‌ایم این گزاره برای این مثال خاص صادق است. در ادامه زنجیره‌های عمومی‌تری را بررسی می‌کنیم.

۸. برای یافتن احتمالات حدی از طریق معادلات تعادل (یا ایستا)، همچنان باید مطمئن شویم که DTMC تجزیه‌ناپذیر و تناوب‌ناپذیر است تا بتوانیم نظریه DTMC ها را اعمال کنیم. برای این منظور، کافی است بررسی کنیم که زنجیره مارکوف تجزیه‌ناپذیر است (نامتناوبی در حالت پیوسته مشکلی ایجاد نمی‌کند).

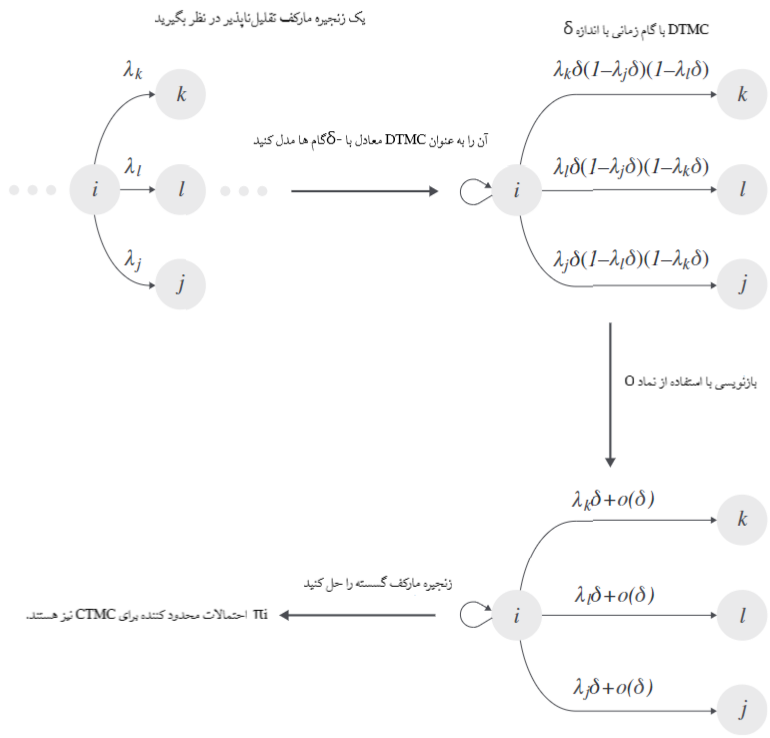
۲.۵.۵ تعمیم

دلیلی که برای تبدیل زنجیره مارکوف پیوسته CTMC به زنجیره مارکوف گسسته DTMC ارائه کردیم، برای یک مثال خاص اعمال شد. شکل ۸ برخی از شهودهای مربوط به دلیل کارکرد این روش برای یک CTMC کلی را نشان می‌دهد.

فرض کنید که با یک CTMC عمومی شروع می‌کنیم. یک حالت منفرد i را در نظر بگیرید، همان‌طور که در شکل ۸ نشان داده شده است. کمان‌های خروجی از حالت i نرخ‌های نمایی را نشان می‌دهند. می‌توانیم CTMC را به عنوان یک DTMC مدل‌سازی کنیم که در آن انتقالات در هر گام δ اتفاق می‌افتد. احتمال‌های انتقال در شکل ۸ نشان داده شده‌اند. در اینجا یک حلقه بازگشتی در حالت i وجود دارد که حاوی احتمال باقیمانده برای باقی ماندن در حالت i است.

توجه کنید که در تفسیر DTMC ، احتمال کمی وجود دارد که در حالتی که در حالت i هستیم، بیش از یک مقصد در یک گام مشخص δ انتخاب شوند (همزمان بیش از دو برنولی فعال شوند). با این حال، احتمال وقوع چنین رویدادی تنها در $o(\delta)$ است. مهم نیست که چگونه این رویداد را تفسیر کنیم، زیرا $o(\delta)$ ناپدید می‌شود. به این ترتیب در صورت انتخاب همزمان بیش از یک مقصد در یک گام δ ، با ماندن در حالت i به این رویداد پاسخ می‌دهیم.

اکنون با نوشتن معادلات موازنه برای DTMC ، آن را حل می‌کنیم. احتمالات حدی که برای DTMC به‌دست می‌آوریم، همان احتمالات حدی CTMC اصلی هستند. توجه کنید که در DTMC نشان داده شده در شکل ۸ ، اگر در حالت i باشیم، در بیشتر انتقالات گام‌های δ به سادگی به حالت i بازمی‌گردیم. این دقیقاً همان مدل‌سازی CTMC است برای مدتی که در حالت i می‌ماند.



شکل ۸: حل یک CTMC دلخواه با استفاده از روش ما.

نکته‌ای که باید از این بحث برداشت شود این است که در عمل نیازی نیست به یک DTMC با گام‌های δ فکر کنیم. فقط باید آن بخش را در نظر بگیریم و معادلات موازنه را برای CTMC مستقیماً بنویسیم و از آن‌ها برای حل π_i ها استفاده کنیم.

تفسیر معادلات تعادل برای CTMC

به معادلاتی که در انتهای تبدیل به آن‌ها می‌رسیم، معادلات تعادل می‌گوییم. این معادلات برای هر حالت j نرخ خروج کارها از حالت j در CTMC را با نرخ ورود کارها به حالت j در CTMC برابر می‌کنند. اینجا نمادگذاری استاندارد برای CTMC ها آورده شده است:

$$\pi_j \nu_j = \sum_i \pi_i q_{ij}$$

سمت چپ معادله حاصلضرب احتمال حدی حضور در حالت j و نرخ خروج زنجیره مارکوف از حالت j به شرط حضور در حالت j (ν_j) را نشان می‌دهد. بنابراین سمت چپ نمایانگر نرخ کل انتقالات خروجی از حالت j است.

ترم i در جمع سمت راست معادله نمایانگر حاصلضرب احتمال حدی حضور در حالت i ، و نرخ خروج زنجیره مارکوف از حالت i به حالت j به شرط حضور در حالت i (q_{ij}) است. بنابراین ترم i در جمع سمت راست نمایانگر

نرخ انتقالات خروجی از حالت i به حالت j است. بنابراین سمت راست معادله نشان‌دهنده نرخ کل انتقالات ورودی به حالت j از هر حالت دیگر است.

توجه کنید که $q_{ij} = \nu_i \cdot p_{ij}$ است. از این رو، می‌توانیم معادله را به‌صورت معادل زیر بنویسیم:

$$\pi_j \sum_i q_{ji} = \sum_i \pi_i q_{ij}$$

خلاصه برای CTMC ها

به دلیل این‌که CTMC ها به طور کلی به عنوان DTMC هایی در نظر گرفته می‌شوند که گام زمانی به صفر میل می‌کند، تمام نظریه ارگادیک که برای DTMC ها توسعه داده شده بود، به CTMC ها نیز اعمال می‌شود. به‌ویژه، قضیه خلاصه زیر را داریم که مشابه قضیه است.

قضیه ۸. قضیه (قضیه خلاصه برای CTMCها) فرض کنید یک CTMC تجزیه‌ناپذیر داده شده است. فرض کنید مقادیری π_i وجود داشته باشند به طوری که برای همه jها:

$$\pi_j \nu_j = \sum_i \pi_i q_{ij} \quad \text{و} \quad \sum_i \pi_i = ۱.$$

در این صورت، π_iها احتمالات حدی برای CTMC هستند و CTMC ارگادیک است.

۶ صف های M/M/1

ساده‌ترین سیستم صف، صفی است که شامل یک سرور است. زمان‌های سرویس‌دهی در این سیستم، متغیرهای تصادفی نمایی مستقل با میانگین $\frac{۱}{\mu}$ هستند و کارها مطابق با یک فرایند پواسون با نرخ λ وارد سیستم می‌شوند. چنین سیستمی به عنوان سیستم صف M/M/1 شناخته می‌شود.

نماد M/M/1، که به آن نماد Kendall نیز می‌گویند، ساختار سیستم صف را توصیف می‌کند. اولین بخش از این نماد نشان‌دهنده توزیع زمان‌های بین‌ورود در فرایند ورود است. حرف «M» در این بخش نشان‌دهنده «بی‌حافظه» بودن و توزیع نمایی زمان‌های بین‌ورود است. بخش دوم، توزیع زمان‌های سرویس‌دهی را مشخص می‌کند و حرف «M» در این بخش بیانگر آن است که زمان‌های سرویس‌دهی نیز بی‌حافظه و دارای توزیع نمایی هستند. بخش سوم نشان‌دهنده تعداد سرورها در سیستم است که در اینجا برابر با ۱ است. در سیستم‌های پیچیده‌تر، این بخش می‌تواند شامل مقادیر دیگری باشد. در برخی مواقع، از بخش چهارم برای تعیین حداکثر ظرفیت سیستم یا سیاست صف‌بندی استفاده می‌شود. عدم وجود بخش چهارم نشان‌دهنده آن است که صف بدون محدودیت است. سیاست صف‌بندی به صورت FCFS است.

تعداد کارها در یک سیستم M/M/1 یک زنجیره مارکوف پیوسته‌زمان (CTMC) را تشکیل می‌دهد، به‌طوری که حالت سیستم نشان‌دهنده تعداد کارها در سیستم است. ساختار این زنجیره به عنوان یک فرایند تولد_مرگ شناخته می‌شود، به طوری که نرخ λ نشان‌دهنده «تولدها» و نرخ μ نشان‌دهنده «مرگ‌ها» هستند. به طور کلی، در فرایند تولد_مرگ، انتقالات تنها بین حالت‌ها متوالی تعریف شده‌اند، اما نرخ‌های تولد و مرگ لزوماً یکسان نیستند.

احتمالات حدی برای حالت‌های مختلف سیستم M/M/1 را می‌توان با حل معادلات موازنه به دست آورد. معادلات موازنه، نرخ خروج از حالت j را با نرخ ورود به حالت j برابر قرار می‌دهند.

سوال: نرخ انتقالات از حالت ۱ به حالت ۲ چقدر است؟

پاسخ: $\lambda \pi_۱$. دقت کنید که λ نشان‌دهنده نرخ انتقالات (تعداد انتقالات در واحد زمانی) از هر حالت i به حالت بعدی، یعنی $i + ۱$ است. با ضرب در $\pi_۱$، تنها انتقالات به سمت بالا که در حالت ۱ رخ می‌دهند را محدود می‌کنیم.

معادلات موازنه به صورت زیر هستند:

$$\text{حالت } ۰ : \quad \pi_۰ \lambda = \pi_۱ \mu \Rightarrow \pi_۱ = \frac{\lambda}{\mu} \pi_۰$$

$$\text{حالت ۱}: \quad \pi_1(\lambda+\mu) = \pi_0\lambda + \pi_2\mu \Rightarrow \pi_2 = \left(\frac{\lambda}{\mu}\right)^2 \pi_0$$

$$\text{حالت ۲}: \quad \pi_2(\lambda+\mu) = \pi_1\lambda + \pi_3\mu \Rightarrow \pi_3 = \left(\frac{\lambda}{\mu}\right)^3 \pi_0$$

$$i \text{ حالت}: \quad \pi_i(\lambda+\mu) = \pi_{i-1}\lambda + \pi_{i+1}\mu$$

در فصل قبل دیدیم که رابطه زیر بین احتمال حدی حالت‌ها با حالت ۰ برقرار است:

$$\pi_i = \left(\frac{\lambda}{\mu}\right)^i \pi_0$$

باتوجه به اینکه $\vec{\pi}$ یک بردار احتمال است و باید رابطه $\sum_{i=0}^{\infty} \pi_i = 1$ را برقرار کند، می توان مقدار π_0 را پیدا کرد:

$$\sum_{i=0}^{\infty} \pi_i = 1 \Rightarrow \sum_{i=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^i \pi_0 = 1$$

$$\Rightarrow \pi_0 \left(\frac{1}{1-\frac{\lambda}{\mu}}\right) = 1 \Rightarrow \pi_0 = 1 - \frac{\lambda}{\mu}$$

بنابراین، با جایگذاری مقدار π_0 در π_i، نتیجه می‌گیریم که:

$$\pi_i = \left(\frac{\lambda}{\mu}\right)^i (1 - \frac{\lambda}{\mu}) = \rho^i(1-\rho)$$

که در آن $\rho = \frac{\lambda}{\mu}$ نشان‌دهنده بهره‌برداری از سرور است. بدیهی است که π_0، احتمال بیکاری سیستم، برابر با $1-\rho$ خواهد بود.

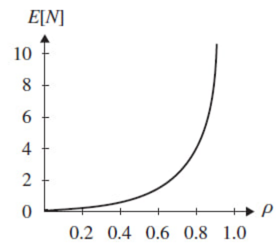
توجه کنید که شرط $\rho < 1$ باید برقرار باشد تا سیستم پایدار بماند و تعداد کارها در سیستم به بی‌نهایت نرود. برای این شرط، باید $\lambda < \mu$ باشد.

میانگین تعداد مشتریان در سیستم را می‌توان با شرط‌بندی روی حالت‌ها به دست آورد:

$$\mathbb{E}[N] = \sum_{i=0}^{\infty} i\pi_i = \sum_{i=1}^{\infty} i\pi_i = \sum_{i=1}^{\infty} i\rho^i(1-\rho) = \frac{\rho}{1-\rho}\sum_{i=1}^{\infty} i\rho^{i-1}$$

$$= \frac{\rho}{1-\rho}\frac{d}{d\rho}\left(\sum_{i=0}^{\infty} \rho^i\right) = \frac{\rho}{1-\rho}\frac{d}{d\rho}\left(\frac{1}{1-\rho}\right) = \frac{\rho}{(1-\rho)^2}.$$

شکل ۹ نمودار $\mathbb{E}[N] = \frac{\rho}{1-\rho}$ را نشان می‌دهد. مشاهده می‌شود که برای $\rho < 0.5$ یا حتی $\rho < 0.6$، میانگین تعداد مشتریان در سیستم به‌آهستگی افزایش می‌یابد. اما پس از این نقطه، به‌شدت افزایش می‌یابد. همچنین، افزایش ρ از ۰٫۸ به ۰٫۹ تأثیر بیشتری نسبت به افزایش آن از ۰٫۷ به ۰٫۸ دارد. می‌توان نشان داد که واریانس



شکل ۹: نمودار میانگین تعداد مشتری‌ها بر حسب ρ در صف M

تعداد مشتریان در سیستم برابر است با:

$$Var(N) = \frac{\rho}{(۱-\rho)^۲}.$$

این مقدار سریع‌تر از میانگین تعداد کارها رشد می‌کند.

میانگین زمان حضور در سیستم و میانگین زمان در صف را می‌توان با استفاده از Little's Law به‌دست آورد:

$$\mathbb{E}[T] = \frac{\mathbb{E}[N]}{\lambda} = \frac{۱}{\mu-\lambda},$$

$$\mathbb{E}[T_Q] = \mathbb{E}[T] - \frac{۱}{\mu} = \frac{\rho}{\mu-\lambda}.$$

سوال : نرخ انتقالات از حالت ۱ به حالت ۲ چقدر است؟

پاسخ: $\pi_۱\lambda$. دقت کنید که λ نشان‌دهنده نرخ انتقالات (تعداد انتقالات در ثانیه) از هر حالت i به حالت بعدی، یعنی $i+۱$ است. با ضرب در $\pi_۱$، تنها انتقالات به سمت بالا که در حالت ۱ رخ می‌دهند را محدود می‌کنیم.

سوال : در یک سرور M/M/1، بیشترین نرخ ورود مجاز برای کارها چقدر است اگر اندازه میانگین کار (تقاضای سرویس) ۳ دقیقه باشد و زمان انتظار میانگین ($\mathbb{E}[T_Q]$) باید کمتر از ۶ دقیقه باقی بماند؟

پاسخ: داده شده $\mu = \frac{۱}{۳}$ کار در دقیقه. همچنین زمان مورد انتظار در سیستم باید کمتر از ۹ دقیقه باشد ($\mathbb{E}[T] = \mathbb{E}[T_Q] + \mathbb{E}[S]$). بنابراین:

$$\frac{۱}{\mu-\lambda} \leq ۹ \Rightarrow \lambda \leq \mu - \frac{۱}{۹} = \frac{۲}{۹}\text{ کار در دقیقه }.$$

سوال : فرض کنید در یک سیستم M/M/1 (با $\lambda < \mu$)، نرخ ورود λ و نرخ سرویس μ را با ضریب k افزایش می‌دهیم. چگونه موارد زیر تحت تأثیر قرار می‌گیرند؟

- بهره‌برداری، ρ؟
- توان عملیاتی، X؟
- تعداد میانگین در سیستم، $\mathbb{E}[N]$؟
- زمان میانگین در سیستم، $\mathbb{E}[T]$؟

پاسخ :

$$\lambda_{new} = k\lambda, \quad \mu_{new} = k\mu.$$

این نتایج را خواهیم داشت:

$$\rho_{\text{new}} = \frac{k\lambda}{k\mu} = \frac{\lambda}{\mu} = \rho_{old},$$

$$X_{new} = \lambda_{new} = k\lambda_{old},$$

$$\mathbb{E}[N_{new}] = \frac{\rho_{new}}{۱-\rho_{new}} = \frac{\rho_{old}}{۱-\rho_{old}} = \mathbb{E}[N_{old}],$$

$$\mathbb{E}[T_{new}] = \frac{۱}{\mu_{new}-\lambda_{new}} = \frac{۱}{k(\mu-\lambda)} = \frac{۱}{k}\mathbb{E}[T_{old}].$$

بنابراین، بهره‌برداری سیستم بدون تغییر باقی می‌ماند و توان عملیاتی به ضریب k افزایش می‌یابد، اما زمان پاسخ میانگین به ضریب k کاهش می‌یابد.

۱.۶ قانون PASTA

در مثال‌های بعدی، اغلب به بررسی وضعیت سیستم از دیدگاه یک ورود علاقه‌مند خواهیم بود. برای روشن‌تر شدن این بحث، فرض کنید شما یک شبیه‌سازی از سیستمی را اجرا می‌کنید. هدف شما این است که سهم زمانی که سیستم دارای n کار است را تعیین کنید. برای این کار، هر ورود به سیستم را بررسی کرده و می‌پرسید آیا تعداد n کار را می‌بیند یا خیر. سپس سهم ورودهایی که شاهد n کار در زمان ورودشان به سیستم بوده‌اند را دنبال می‌کنید. سوال این است که آیا این روش جواب درست به شما می‌دهد یا خیر.

فرض می‌کنیم سیستم پیوسته و ارگادیک است. بگذارید $\pi_n = p_n$ احتمال حدی وجود n کار در سیستم (یا معادل آن سهم زمانی طولانی‌مدت که سیستم دارای n کار است) باشد. همچنین a_n را به عنوان احتمال حدی که یک ورود، n کار را در سیستم می‌بیند (یا معادل آن، سهم ورودهایی که n کار را در سیستم مشاهده می‌کنند) تعریف می‌کنیم. همچنین d_n را به عنوان احتمال حدی که یک خروج، n کار را پس از خود در سیستم باقی می‌گذارد (یا معادل آن سهم خروج‌هایی که n کار را در سیستم باقی می‌گذارند) تعریف می‌کنیم.

سوال: آیا $a_n = p_n$ است؟

پاسخ: خیر

اثبات از طریق مثال: یک صف را در نظر بگیرید که در آن کارها زمان بین‌ورودی یکنواختی بین ۱ و ۲ دارند (یعنی $\sim U(۱,۲)$). فرض کنید همه زمان‌های خدمت ثابت و برابر ۱ هستند. در این حالت $a_۰ = ۱$ و $d_۰ = ۱$ زیرا مشتری قبل از رسیدن مشتری بعدی سرویس خود را تمام می‌کند. اما $p_۰ \neq ۱$ زیرا سیستم همیشه خالی نیست.

سوال: آیا $a_n = d_n$ است؟

پاسخ: بله

اثبات. زمانی که تعداد کارهای داخل سیستم از n به $n+۱$ تغییر کند، ورودی متناظر تعداد n کار را در صف می‌بیند. حال درصورتی که تعداد کارهای داخل سیستم از $n+۱$ به n کاهش یابد، کار خروجی متناظر تعداد n کار را در سیستم خواهد دید.

از آنجا که کارها یکی یکی وارد و خارج می‌شوند، این رویدادها به تعداد مساوی اتفاق می‌افتند (با حداکثر یک اختلاف). به دلیل بازگشتی بودن تمامی حالات، هر یک از این رویدادها به طور نامحدود تکرار می‌شوند. بنابراین تعداد ورودهایی که n کار در سیستم می‌بینند برابر با تعداد خروجی‌هایی است که n کار باقی مانده در سیستم می‌بینند. زیرا نرخ کلی ورودها برابر با نرخ کلی خروج‌ها است. □

اکنون برای معرفی قانون PASTA آماده هستیم «ورودهای پواسون زمان متوسط را مشاهده می‌کنند» . این قانون بیان می‌کند که $a_n = p_n$ و بر اساس ، $d_n = p_n$ است، زمانی که ورودها از فرآیند پواسون پیروی کنند:

اگر فرآیند ورود به سیستم پواسون باشد، آنگاه $a_n = p_n$.

اثبات. با در نظرگرفتن p_n و a_n به عنوان احتمال‌های حدی، داریم:

$$p_n = \lim_{t\to\infty} \mathbb{P}\{N(t) = n\}.$$

$$a_n = \lim_{t\to\infty} \mathbb{P}\{N(t) = n \mid \text{یک ورود در زمان t رخ دهد}\}.$$

نشان می‌دهیم که $a_n = p_n$. بگذارید $A(t, t+\delta)$ رویدادی باشد که یک ورود بلافاصله پس از زمان t رخ می‌دهد. پس داریم:

$$\begin{aligned}
a_n &= \lim_{t\to\infty}\lim_{\delta\to ۰} P\{N(t) = n \mid A(t, t+\delta)\} \\
&= \lim_{t\to\infty}\lim_{\delta\to ۰} \frac{P\{N(t) = n, A(t,t+\delta)\}}{P\{A(t,t+\delta)\}} \\
&= \lim_{t\to\infty}\lim_{\delta\to ۰} \frac{P\{A(t,t+\delta) \mid N(t) = n\}P\{N(t)=n\}}{P\{A(t,t+\delta)\}} \quad (*) \\
&= \lim_{t\to\infty}\lim_{\delta\to ۰} \frac{P\{A(t,t+\delta)\}P\{N(t)=n\}}{P\{A(t,t+\delta)\}} \quad (**)
\end{aligned}$$

$$= \lim_{t\to\infty} P\{N(t) = n\} = p_n.$$

سوال: چرا این اثبات برای مثال فرآیند ورود یکنواخت با زمان‌های خدمت قطعی برقرار نیست؟
پاسخ: در این حالت، نمی‌توانیم از $(*)$ به $(**)$ منتقل شویم زیرا اطلاع از $N(t)$ بر وقوع یک ورود در چند ثانیه آینده تأثیر دارد. در حالیکه در فرآیند ورود پواسون، ورودها به صورت مستقل از $N(t)$ رخ می‌دهند.
$\square$

۷ نصیحتی کُنَمَت بشنو و بهانه مَگیر![۲۹]

لطفاً به نکات زیر دقّت کنید:

۱. این پروژه بخشی از نمره‌ی شما در این درس را تشکیل خواهد داد.

۲. به زودی کانالی در تلگرام برای پاسخگویی به سوالات احتمالی شما ایجاد خواهد شد.

۳. می‌توانید پروژه را فردی یا در قالب گروه‌های ۲ نفره انجام دهید. گروه‌هایی با تعداد اعضای بیشتر از ۲ مجاز نیست. فرمی برای ثبت گروه‌ها در اختیار شما قرار خواهد گرفت. دقّت داشته باشید که در هنگام تحویل پروژه باید تمامی اعضای گروه به تمامی بخش‌ها مسلّط باشند و در نهایت همه‌ی اعضای یک گروه نمره‌ی واحدی را دریافت خواهند کرد.

۴. تحویل پروژه به صورت گزارش و پاسخ‌های نوشته‌شده است. گزارش باید شامل پاسخ پرسش‌ها، تصاویر و نمودارها و نتیجه‌گیری‌های لازم باشد. همچنین تمیزی گزارش بسیار مهم است. گزارش را در یک فایل فشرده‌شده در سامانه‌ی درس‌افزار آپلود کنید.

۵. اگر برای پاسخ به پرسش‌ها، از منبعی (کتاب، مقاله، سایت و...) کمک گرفته‌اید، حتماً به آن ارجاع دهید.

۶. نوشتن گزارش کار با LaTeX نمره‌ی امتیازی دارد. علّت این نکته تشویق شما به استفاده از سیستم حروف‌چینی LaTeX است و استفاده از LaTeX موضوعیت دارد و نه طریقیّت! برای گرفتن نمره‌ی امتیازی لازم است که فایل `tex.` را نیز در کنار فایل اصلی گزارش ارسال کنید.

۷. و پرسش‌های تئوری با رنگ آبی مشخص شده‌اند همچنین قضایا با رنگ مشخص شده‌اند توجه کنید اثبات قضایا در صورتی که خواسته نشده اجباری نیست.

۸. بخش‌های تئوری گزارش که در قالب پرسش‌ها طرح شده‌اند را می‌توانید روی کاغذ بنویسید و تصویر آن‌ها را در گزارش خود بیاورید، ولی توصیه‌ی برادرانه می‌کنم که این کار را نکنید! اوّلین علّت این توصیه این است که اگر چنین کاری کنید، تحت هیچ شرایطی نمره‌ی امتیازی LaTeX را نخواهید گرفت!

۹. درصورت مشاهده‌ی تقلّب، نمره‌ی هردو فرد صفر منظور خواهد شد.

موفّق باشید!

[۲۹] نصیحتی کُنَمَت بشنو و بهانه مَگیر/ هر آنچه ناصحِ مُشفِق بگویَدَت بپذیر [حافظ]