
بسم الله الرحمن الرحیم
الحمد لله رب العالمین
الرحمن الرحیم
مالک یوم الدین
ایاک نعبد و ایاک نستعین
اهدنا الصراط المستقیم
صراط الذین انعمت علیهم
غیر المغضوب علیهم
و لا الضالین

دانشگاه پیام نور مرکز تهران

گروه آمار

پایان نامه برای دریافت درجه کارشناسی ارشد در رشته آمار ریاضی

آنالیز بیزی از مدل های مخفی مارکوف چند متغیره گوسی

نگارش:

محمد دهستانی اردکانی

استاد راهنما:

دکتر پرویز نصیری

استاد مشاور:

دکتر علی شادرخ

پاییز ۱۳۹۰

تقدیم به:

پدر و مادر عزیز،

و

همسر صبور و فداکارم

تقدیر و تشکر

اینک که این مرحله از زندگی تحصیل خود را به پایان بردم، بر خود لازم می دانم شکرگذار نعمات خداوند بی همتا باشم. چرا که می دانم "لا مانع لما اعطیت و لا معطی لما منعت" و نیز از منظر "لم یشکر المخلوق، لم یشکر الخالق". برخود لازم می دانم از زحمات عزیزانی که در این مسیر یاری گر من بوده اند تقدیر و تشکر نمایم.

از آقایان دکتر پرویز نصیری که راهنمایی این پایان نامه را تقبل نمودند و در این مسیر همواره یاری گر من بوده و اساسی ترین نقش را در انجام این تحقیق داشته اند.

از آقای دکتر علی شادرخ که علاوه بر تقبل زحمت مشاوره این پایان نامه، در طول تحصیل از وجودشان بهره فراوان بردم.

از تمامی اساتیدم، مخصوصاً دکتر بهنام زرپاک که زحمت داوری را عهده دار بودند

از برادر و دوست گرامی،آقای امیرانیشکانی به خاطر همکاری و راهنمایی‌های ارزشمندشان سپاسگزارم

از تمامی دوستان و عزیزانی که در طی این مسیر یاری گر و مشوق من بوده اند، نهایت تقدیر و تشکر را داشته و در پایان از پدر، مادر، همسر و همه فرشتگانی که بالهای محبت خود را گسترانیدند و با تحمل دشواری‌ها، سبب شدند تا در کمال آسودگی خیال و فراغت بال، شوق آموختن در من زنده بماند صمیمانه سپاسگزارم و این نیست جز جلوه‌ای از لطف و رحمت پرودگاری که از ادای شکر حتی یک نعمت او ناتوانم.

چکیده:

یکی از مسائلی که در پردازش سیگنال توجهات را به خود معطوف نموده است، مدل سازی سیگنال است. انتخاب های مختلفی برای مدل کردن سیگنال و خصوصیات آن وجود دارد. از یک دیدگاه می توان مدل های سیگنال را به دو دسته مدل های معین و مدل های آماری تقسیم بندی نمود.

مدل های معین عمدتا برخی خواص شناخته شده سیگنال را مورد استفاده قرار می دهند. در این حالت تشکیل مدل سیگنال سر راست است و تنها کافی است مقادیر پارامترهای مدل تخمین زده شود. در مدل های آماری سعی در ایجاد مدل با استفاده از خواص آماری سیگنال است. مدل های گوسی، زنجیره مارکوف و مدل مخفی مارکوف از جمله این روش ها هستند.

مدل های مخفی مارکوف چند متغیره گوسی با تعدادی مجهول از وضعیت ها مرتبط است در این جا آنالیز و الگوریتم های مونت کارلو زنجیره مارکوف بازگشت پذیر موثر جدیدی را برای برآورد پارامتر های مجهول از مدل ارائه می شود مدل های مخفی مارکوف یک بسط از مدل های ترکیب که در سری های زمانی می توانند بکار روند هستند. بنابراین مشاهدات در یک تعداد از گروه ها طبقه بندی می شوند.

در مدل سازی ما توجه خود را معطوف به دنباله ی مشاهدات کرده و به دنبال بهترین ابزار برای تولید مجدد داده های اصلی می گردیم تا دنباله مناسب با مقدار واقعی بدست آید. و در پیش بینی ما به دنبال این هستیم که چه اتفاقی می افتد اگر ما وضعیت فعلی و حال را برای آینده در نظر بگیریم تا احتمال توزیع برای رسیدن به مقادیر آتی را بیابیم.

واژگان کلیدی: مدل مخفی مارکوف، پیش بینی، توزیع احتمال، مدل سازی، الگوریتم مونت کارلو

فهرست مطالب

عنوان صفحه

فصل اول

زنجیر مارکوف

۱–۱–مقدمه

آندری مارکوف در ژوئن ۱۸۵۶ در روسیه دیده به جهان گشود او فارغ التحصیل دانشگاه سنت پترزبورگ در سال ۱۸۷۸ بود وی در سال ۱۸۸۶ مدرک پروفسوری خود را دریافت کرد. کارهای زود هنگام مارکوف در تئوری اعداد ، آنالیز، حدود انتگرال ها، همگرایی سری ها، دنباله کسرها و بسیار اساسی بود و بعد از سال ۱۹۰۰، مارکوف تحت تأثیر استاد خود چبیشف، از روش دنباله های کسرها در تئوری احتمالات استفاده کرد. وی هم چنین در مورد رشته های متغیرهای وابسته متقابل، مطالعاتی انجام داد. با امید ثابت کردن قوانین حدی در احتمالات در حالات کلی آنها، او قضیه حد مرکزی را با در نظر گرفتن فرض های کامل آن، اثبات کرد.مارکوف به دلیل مطالعاتش پیرامون زنجیرهای مارکوف که رشته هایی از متغیرهای تصادفی هستند، معروف است. در سال ۱۹۲۳ «نوربرت واینر» اولین کسی بود که پیرامون یک سلسله از این مراحل مارکوف شروع به بحثی جدی کرد.اساس یک تئوری اصلی در سال ۱۹۳۰ توسط کولموگروف فراهم شد وتا کنون زنجیره مارکوف در حال پیشرفت است. در ایران هم رضوان صحرائی و عباس عباس پور در سال ۲۰۰۹ میلادی به منظور بررسی چگونگی پیش بینی عرضه نیروی انسانی در سازمان ها ازمدل مارکوف استفاده شده است.

۱–۲– تعریف زنجیره مارکوف

یک فرآیند تصادفی در بازه‌ی زمانی گسسته $\{X_n,\ n\geq 0\}$ که تعدادی حالت شما را به خود می گیرد (تعداد حالات یک عدد صحیح نا منفی می باشد) را در صورتی که مشخصات زیر داشته باشد، زنجیره ی مارکوف مرتبه اول می نامند: برای همه ی مقادیر $i_0, i_1, \ldots, i_{n-1}, i, j$ و به ازای همه ی مقادیر نا منفی n

$$P\{X_{n+1}=j|X_n=i, X_{n-1}=i_{n-1},\ldots,X_1=i_1, X_0=i_0\}$$
$$=P\{X_{n+1}=j|X_n=i\}$$

با فرض اینکه احتمال شرطی بالا از مقدار n مستقل است:

$$P_{ij} = p\{X_{n+1} = j|X_n = i\}, i,\ j \geq 0$$

را ، احتمال انتقال (یک مرحله ای) گویند که دارای مشخصه های زیر می باشد:

$$P_{ij} \geq ۰, \qquad i, j \geq ۰$$

$$\sum_{j=۰}^{\infty} P_{ij} = ۱, \qquad i = ۰,۱,...$$

به طور مشابه، این مطلب نیز می تواند به شکل ماتریس بیان شود که به آن ماتریس احتمال انتقال گویند.

$$P = \left(P_{ij}\right) = \begin{bmatrix} P_{۰۰} & P_{۰۱} & P_{۰۲} & \cdots \\ P_{۱۰} & P_{۱۱} & P_{۱۲} & \cdots \\ \vdots & \vdots & & \\ P_{i۰} & P_{i۱} & \cdots & P_{ij} \end{bmatrix}$$

به طوری که مجموع سطر ها از ماتریس برابر یک می باشد که در چنین حالتی می‌گوییم P تصادفی است.

چنانچه مجموع ستون ها نیز برابر با یک باشد و گوییم که P تصادفی مضاعف گویند.

مثال ۱–۱–(قدم زدن تصادفی عمومی) فرض کنید فرآیند $\{z_i\}_۱^{\infty}$ را یک iid گسسته تصادفی با pmf زیر می‌باشد

$$P\{Z_i = k\} = p_k, k = ۰,\pm۱,...$$

همچنین در نظر می گیریم $X_۰ = ۰$ و $X_n = \sum_{i=۱}^{n} Z_i$.

آنگاه $\{X_n, n \geq ۰\}$ یک سری مارکوف می **باشد** که آن را قدم زدن تصادفی عمومی با احتمال انتقال

$$P_{ij} = P\{X_{n+۱} = j | X_n = i\} = P\{Z_{n+۱} = j - i\} = p_{j-۱} \quad i, j = ۰,\pm۱,...$$

نکته : اگر $P\{Z_i = ۱\} = p = ۱ - P\{Z_i = -۱\}$ برای بعضی مقادیر $۰ < p < ۱$ برقرار باشد آنگاه $\{X_n, n \geq ۰\}$، یک قدم زدن تصادفی ساده خوانده می شود. به طور اختصاصی زمانی که $p=۱/۲$، آن را قدم زدن تصادفی متناسب گوییم.

مثال ۱–۲–هوای یک روز می تواند ابری و یا آفتابی باشد که این امر تنها به هوای دو روز متوالی گذشته بستگی دارد. احتمال اینکه روز n(ام) آفتابی باشد از این قرار است

احتمال اینکه روز n(ام) آفتابی باشد	روز ۱-n(ام)	روز ۲-n(ام)
۰.۹	آفتابی	آفتابی
۰.۷	آفتابی	ابری
۰.۵	ابری	آفتابی
۰.۲	ابری	ابری

سری مارکوف مربوطه را تشکیل دهید.

حل) $X_n(n \geq ۰)$ را برای روز آفتابی صفر و برای روز ابری ۱ در نظر می گیریم. آنگاه مسلماً $(X_n, n \geq ۰)$ یک سری مارکوف نمی باشد.

تعریف می کنیم:

$$y_n = \begin{cases} ۰ & \text{اگر روز ۱-}n\text{ آفتابی و روز }n\text{ آفتابی} \\ ۱ & \text{اگر روز ۱-}n\text{ ابری و روز }n\text{ آفتابی} \\ ۲ & \text{اگر روز ۱-}n\text{ آفتابی و روز }n\text{ ابری} \\ ۳ & \text{اگر روز ۱-}n\text{ ابری و روز }n\text{ ابری} \end{cases}$$

آنگاه $\{Y_n, n \geq ۱\}$ یک زنجیره مارکوف با ماتریس تغییروضعیت زیر است که :

$$P = \begin{bmatrix} ۰.۹ & ۰ & ۰.۱ & ۰ \\ ۰.۷ & ۰ & ۰.۳ & ۰ \\ ۰ & ۰.۵ & ۰ & ۰.۵ \\ ۰ & ۰.۲ & ۰ & ۰.۸ \end{bmatrix}$$

که بعنوان مثال در بالا $P_{۱۲}$ از روش زیر محاسبه شده است:

$$P_{۱۲} = P(Y_{n+۱} = ۲ | Y_n = ۱) =$$

P(روز ۱-n(ام) ابری و روز n(ام) آفتابی | روز n(ام) آفتابی و روز ۱+n(ام) ابری)

۰.۳=۰.۷-۱ = (روز n-۱(ام) ابری و روز n(ام) آفتابی باشد | روز n+۱(ام) ابری باشد) P =

توجه کنید که زمانی که ما یک فرآیند تصادفی دو بعدی $\{(X_n, X_{n+1}), n = 0,1,...\}$ تعریف می کنیم، این می تواند معادل یک سری مارکوف مانند بالا در نظر گرفته شود.

مثال ۱–۳–: مشتریان بر اساس یک فرایند تکراری با توزیع زمان بین دو ورود از نوع G به سرویس دهنده ای که زمان سرویس آن دارای توزیع نمایی با میانگین $\frac{1}{\mu}$ است رجوع می کنند.

$\{X_n, n \geq 1\}$ را برابر تعداد مشتریانی که تا قبل از n امین ورود در سیستم مشاهده شده‌اند در نظر بگیرید همچنین $\{Z_n, n \geq 0\}$ را برابر تعداد مشتریانی که بین n امین و n+۱ امین ورود سرویس گرفته‌اند در نظر بگیرید، آنگاه :

$$X_{n+1} = \max\{0, X_n + 1 - Z_n\}$$

توجه داشته باشید که Z_n با هر طولی از بازه ی زمانی t دارای توزیع پواسون با میانگین μt می‌باشد.

مشاهده می شود که X_{n+1} تنها به X_n وابسته است. بنابراین $\{X_n, n = 1,2,...\}$ یک سری مارکوف خواهد بود که به آن سری مارکوف ادغام شده‌ی صف $G/M/1$ گویند.

احتمالات انتقال این گونه بیان می‌شوند :

$$\begin{aligned} P_{i,i+1-j} &= P\{X_{n+1} = i+1-j \mid X_n = i\} = P\{Z_n = j\} \\ &= \int_0^\infty P\{Z_n = j \mid \text{interarrival time} = t\} dG(t) \\ &= \int_0^\infty \frac{e^{-\mu t}(\mu t)^j}{j!} dG(t),\ j = 0,1,...,i \end{aligned}$$

$$\begin{aligned} P_{i,0} &= P\{X_{n+1} = 0 \mid X_n = i\} = P\{Z_n \geq i+1\} \\ &= \int_0^\infty \sum_{k=i+1}^{\infty} \frac{e^{-\mu t}(\mu t)^k}{k!} dG(t) \end{aligned}$$

مثال ۱–۴– ۲۱مشتریان با یک فرآیند پوا سن و با سرعت λ به سرویس دهنده رجوع می کنند و زمان سرویس از توزیع G تبعیت می کند. یک سری مارکوف ادغام شده ارائه دهید.

حل) : همانگونه که برای صف $M/G/1$ داشتیم، ترتیب تعداد مشتریان در سیستم که تا n امین ورود مشاهده شده‌اند می تواند یک سری مارکوف برای این مثال باشد. ولی اگر چنین کاری را انجام دهیم، محاسبه ی احتمال گذر بسیار مشکل خواهد شد

در عوض، X_n را برابر تعداد مشتریانی که تا قبل از n امین خروج، از سیستم خارج شده اند و Y_n را برابر تعداد مشتریانی که در طول زمان سرویس n امین مشتری وارد می شوند در نظر می گیریم. داریم:

$$X_{n+1} = \begin{cases} X_n - 1 + Y_{n+1} & if \ X_n > 0 \\ Y_{n+1} & if \ X_n = 0 \end{cases}$$

بنابراین X_{n+1} تنها وابسته به X_n می باشد و لذا $\{X_n, n = 1,2,...\}$ یک سری مارکوف ادغام شده خواهد بود. توزیع Y_n عبارت خواهد بود از

$$P\{Y_n = k\} = \int_0^\infty \frac{e^{-\lambda t}(\lambda x)^k}{k!} dG(x)$$

احتمال گذر ها از این قرار خواهند بود :

$$P_{0j} = P\{X_{n+1} = j \mid X_n = 0\} = P\{Y_{n+1} = j\} = \int_0^\infty \frac{e^{-\lambda t}(\lambda x)^j}{j!} dG(x), j \geq 0$$

$$P_{ij} = P\{X_{n+1} = j \mid X_n = i\} = P\{Y_{n+1} = j - i + 1\} = \int_0^\infty \frac{e^{-\lambda t}(\lambda x)^{j-i+1}}{(j-i+1)!} dG(x),$$

$$j \geq i - 1 \quad , \quad i \geq 1.$$

۱–۳– معادلات چپمن – کولموگورف[1]

احتمال انتقال n مرحله ای را به صورت زیر تعریف می کنیم.

$$P_{ij}^{(n)} = P\{X_{m+n} = j \mid X_m = i\}, \quad i \geq 0, j \geq 0$$

معادله‌ی چپمن -کولموگورف به ما اجازه می‌دهد تا احتمال انتقال n مرحله‌ای را تنها از احتمال انتقال یک مرحله محاسبه کنیم.

۱_ chapman-rovkolmogo

قضیه ۱–۱– (معادله‌ی معادلات چپمن-کولموگورف)

$$P_{ij}^{(m+n)} = \sum_{k=0}^{\infty} P_{ik}^{(m)} P_{kj}^{(n)}$$

اثبات : اصلاح X_m نتیجه می دهد:

$$\begin{aligned} P_{ij}^{(m+n)} &= P\{X_{m+n} = j \mid X_0 = i\} \\ &= \sum_{0}^{\infty} P\{X_{m+n} = j \mid X_m = k, X_0 = i\} P\{X_m = k \mid X_0 = i\} \\ &= \sum_{0}^{\infty} P\{X_{m+n} = j \mid X_m = k\} P\{X_m = k \mid X_0 = i\} \\ &= \sum_{0}^{\infty} P_{kj}^{(n)} P_{ik}^{(m)} \end{aligned}$$

که ماتریس احتمال انتقال n مرحله‌ای را با $P^{(n)}$ نشان دهیم، آنگاه معادله ی کولموگورف–چپمن بصورت زیر بازنویسی می شود.

$$P^{(m+n)} = P^{(m)} P^{(n)} \qquad m, n \geq 0$$

بنابراین ماتریس انتقال n مرحله ای به شکل زیر خلاصه می شود :

$$P^{(n)} = P^{(1)} P^{(n-1)} = P P^{(n-1)} = P P P^{(n-2)} = \dots = P^{n}$$

یعنی ماتریس احتمال انتقال n مرحله ای را می توان با n بار ضرب کردن ماتریس احتمال انتقال یک مرحله ای در خودش به دست آورد.

مثال ۱–۵– : (قدم زدن تصادفی) : فرض کنید $\{X_n, n = 1,2,\dots\}$ یک قدم زدن تصادفی باشد که در آن :

$$P_{i,i+1} = p = 1 - P_{i,i-1}, \qquad i = 0, \pm 1, \dots$$

با دانستن $X_n = 0$ احتمال $X_{n+4} = 0$ را حساب کنید.

حل : اولاً

$$P_{0,-2}^{(2)} = (1-p)^2 = P_{0,2}^{(2)}, P_{0,0}^{(2)}, = 2p(1-p), P_{0,2}^{(2)} = p^2 = P_{-2,0}^{(2)}$$

بنابراین :

$$P_{0,0}^{(4)}=\sum_{k=-\infty}^{\infty}P_{0,k}^{(2)}P_{k,0}^{(2)}=P_{0,-2}^{(2)}P_{-2,0}^{(2)}+P_{0,0}^{(2)}P_{0,0}^{(2)}+P_{0,2}^{(2)}P_{2,0}^{(2)}$$
$$=(1-p)^2p^2+4p^2(1-p)^2+p^2(1-p)^2=6p^2(1-p)^2$$

توجه کنید که زمانی از $X_n=0$ ، $X_{n+4}=0$ نتیجه می شود که دو حرکت به راست و دو حرکت به چپ داشته باشیم پس :

$$P_{0,0}^{(4)}=\binom{4}{2}p^2(1-p)^2=6p^2(1-p)^2.$$

۱–۴– طبقه‌بندی وضعیت ها

هدف اصلی در زنجیره ی مارکوف عبارتست از یافتن احتمال اینکه وضعیت نهایی j باشد زمانی که وضعیت اولیه ی ما i بوده است درواقع هدف محاسبه $\lim_{n\to\infty} P\{X_n=j \mid X_0=i\}$ می باشد.

سری ها ی مارکوف غیر قابل تقلیل

تعریف ۱–۲–: حالت j را قابل دسترسی توسط حالت i گویند اگر به ازای بعضی از مقادیر $n>0$ داشته باشیم $P_{ij}^{(n)}>0$ و می‌نویسیم $(i->j)$

تعریف ۱–۳– دو حالت i,j را مرتبط گوییم هر گاه از یکی بتوانیم به دیگری برسیم و می‌نویسیم $(i<->j)$.

ارتباط اولین خصوصیتی است که در دسته بندی حالت های مورد بررسی قرار می گیرد.

قضیه ۱–۲– (خواص ارتباط):

۱) $i<->j$

2) $if \ \ i<->j \ , \ then \ \ j<->i.$

۳) $if \ \ i<->j \ \ and \ \ j<->k \ , \ then \ \ i<->k.$

اثبات: خواص ۱ و۲ واضح هستند

۳) فرض کنید $i->j$ و $j->k$. از آنجا که $i->j$ و $j->k$ مقداری مانند $m \geq ۰$ و $n \geq ۰$ وجود دارد که به ازای آن ها $P_{ij}^{(m)} > ۰, P_{jk}^{(n)} > ۰$ پس :

$$P_{ik}^{(m+n)} = \sum P_{ir}^{(m)} P_{rk}^{(n)} \geq P_{ij}^{(m)} P_{kk}^{(n)} > ۰$$

و این یعنی $S \geq ۰$ وجود دارد که $P_{ik}^{(s)} > ۰$ بنابراین $i->k$ به طور مشابه می توانیم نشان دهیم که $t \geq ۰$ وجود دارد که $P_{ik}^{(t)} > ۰$ که این به معنی $K->i$ می باشد پس نتیجه می گیریم که $i <-> k$

تعریف ۱–۴– دو حالت (وضعیت) که با یکدیگر ارتباط دارند را در یک کلاس در نظر می گیریم.

تعریف ۱–۵– وضعیتی مثل i که برای آن داشته باشیم $P_{ii} = ۱$ را یک وضعیت جذب کننده گوییم. یک وضعیت جذب کننده به تنهایی بعنوان یک کلاس در نظر گرفته می شود.

تعریف ۱–۶– اگر تنها یک کلاس داشته باشیم، سری مارکوف را غیر قابل تقلیل گوییم.

توجه داشته باشید که کلیه ی وضعیت های موجود در یک سری مارکوف غیر قابل تقلیل با یکدیگر ارتباط دارند. دیاگرام انتقال (وضعیت) برای مشاهده ی نظری غیر قابل تقلیل بودن یک سری مارکوف و نحوه ی تقسیم بندی وضعیت ها به کلاس های مختلف مفید خواهد بود. در دیاگرام انتقال وضعیت ها لیست می شوند و در صورتی که بین دو وضعیت احتمال انتقال مقدار مثبتی داشته باشد یک کمان جهت دار بین آن ها رسم می کنیم.

مثال ۱–۶– یک سری مارکوف با ماتریس احتمال انتقال زیر در نظر بگیرید:

$$P = \begin{bmatrix} ۰ & ۱ & ۰ \\ \frac{۱}{۲} & \frac{۱}{۴} & \frac{۱}{۴} \\ \frac{۱}{۳} & ۰ & \frac{۲}{۳} \end{bmatrix}$$

مشاهده می کنیم که تمام وضعیت ها با یکدیگر ارتباط دارند و سری مارکوف غیر قابل تقلیل است.

مثال ۱–۷– سری مارکوف زیر را کلاس‌بندی نمائید:

$$P=\begin{bmatrix} \frac{۱}{۲} & \frac{۱}{۲} & ۰ & ۰ \\ \frac{۱}{۳} & \frac{۲}{۳} & ۰ & ۰ \\ \frac{۱}{۴} & \frac{۱}{۴} & \frac{۱}{۴} & \frac{۱}{۴} \\ ۰ & ۰ & ۰ & ۱ \end{bmatrix}$$

حل) سه کلاس وجود دارد و عبارتند از : $\{۰,۱\},\{۲\},\{۳\}$

تعریف ۱–۷–گوییم وضعیت i دارای دوره‌ی تناوب d است چنانچه $P_{ii}^{(n)}=۰$ و n بر d بخش‌پذیر نبوده و d بزرگترین عدد صحیحی باشد که چنین خاصیتی دارد. (دوره ی i را با $d(i)$ نشان می‌دهیم.)

تعریف ۱–۸– وضعیتی که دارای دوره‌ی تناوب ۱ باشد را غیر متناوب گویند.

قضیه ۱–۳– اگر $i <-> j$ آنگاه $d(i)=d(j)$

مثال ۱–۸– یک سری مارکوف با ماتریس احتمال انتقال زیر در نظر بگیرید:

در این حالت گوییم سری مارکوف غیر قابل تقلیل

$$P=\begin{pmatrix} ۰ & ۱ & ۰ \\ ۰ & ۰ & ۱ \\ ۱ & ۰ & ۰ \end{pmatrix}$$

است و دوره‌ی تناوب هر یک از وضعیت ها برابر ۳ می‌باشد. توجه داشته باشید که:

$$P^{(۳n)}=\begin{pmatrix} ۱ & ۰ & ۰ \\ ۰ & ۱ & ۰ \\ ۰ & ۰ & ۱ \end{pmatrix} \qquad n=۱,۲...$$

وضعیت های بازگشت پذیر و گذرا

T_{ij} را بعنوان تعداد انتقال ها یی که فرآیند قبل از اولین ورود به وضعیت j و با شروع از وضعیت i $(X_0 = i)$ پشت سر نهاده است در نظر بگیرید که این مقدار بعنوان زمان اولین گذر از i به j شناخته می‌شود. در یک حالت خاص، T_{jj} را زمان تکرار وضعیت j گویند. همچنین $f_{ij}^{(n)}$ را احتمال اینکه فرآیند پس از شروع از وضعیت i در n امین انتقال به وضعیت j برسد در نظر بگیرید.

$$f_{ij}^{(n)} = P\{T_{ij} = n\} = P\{X_n = j, X_k \neq j, k = 1,\ldots,n-1 \mid X_0 = i\}, n \geq 1$$

همچنان f_{ij} را بعنوان احتمال هر بار رسیدن از i به j در نظر بگیرید. داریم :

$$f_{ij} = P\{T_{ij} < \infty\} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$$

تعریف ۱–۹:

۱) وضعیت j بازگشت پذیر گوییم اگر $f_{jj} = 1$

۲) وضعیت j را گذرا گوییم چنانچه $f_{jj} < 1$

توجه داشته باشید که چنانچه وضعیت j، بازگشت پذیر باشد آنگاه $f_{jj}^{(n)}$ ، تابع چگالی احتمال T_{jj} خواهد بود. برای وضعیت بازگشت پذیر j، میانگین کمترین زمان تکرار با μ_{jj} نشان داده می شود و برابر است با

$$\mu_{jj} = E[T_{jj}] = \sum_{n=1}^{\infty} n f_{jj}^{(n)}$$

حتی اگر وضعیت j بازگشت پذیر باشد، هیچ الزامی وجود ندارد که μ_{jj} محدود باشد، آن هم زمانی که فضای یک وضعیت نا متناهی است.

وضعیت‌ها ی بازگشت پذیر نیز به دو دسته تقسیم می شوند :

تعریف ۱–۱۰–

۱) وضعیت j را بازگشت پذیر مثبت گوییم چنانچه $\mu_{jj} < \infty$

۲) وضعیت j را بازگشت پذیر خنثی گوئیم چنانچه $\mu_{jj} = \infty$

مثال ۱–۹– سری مارکوف مثال ۱-۷را در نظر بگیرید

$$P\begin{bmatrix} ۱/۲ & ۱/۲ & ۰ & ۰ \\ ۱/۳ & ۲/۳ & ۰ & ۰ \\ ۱/۴ & ۱/۴ & ۱/۴ & ۱/۴ \\ ۰ & ۰ & ۰ & ۱ \end{bmatrix}$$

آیا وضعیت صفر بازگشت پذیر است یا گذرا؟ اگر بازگشت پذیر است میانگین زمان تکرار را محاسبه کنید. در مورد وضعیت ۲ چه می‌توان گفت؟

حل) $f_{۰۰}^{(n)} =, n \geq ۱$ از روش زیر بدست می‌آیند :

$$f_{۰۰}^{(۱)} = P\{X_۱ = ۰ \mid X_۰ = ۰\} = P_{۰۰} = ۱/۲$$
$$f_{۰۰}^{(۲)} = \{X_۲ = ۰, X_۱ \neq ۰ \mid X_۰ = ۰\} = P_{۰۱}P_{۱۰} = (۱/۲)(۱/۳)$$
$$f_{۰۰}^{(n)} = \frac{۱}{۲}\left(\frac{۲}{۳}\right)^{n-۲}\left(\frac{۱}{۳}\right), n \geq ۲$$

بنابراین :

$$f_{۰۰} = \sum_{n=۱}^{\infty} f_{۰۰}^{(n)} = \frac{۱}{۲} + \frac{۱}{۶}\sum_{n=۲}^{\infty}\left(\frac{۲}{۳}\right)^{n-۲} = ۱.$$

و لذا وضعیت صفر بازگشت پذیر است.

میانگین زمان تکرار وضعیت صفر با فرمول زیر محاسبه می‌شود:

$$\mu_{۰۰} = \sum_{n=۱}^{\infty} n f_{۰۰}^{(n)} = \frac{۱}{۲} + \frac{۱}{۶}\sum_{n=۲}^{\infty} n\left(\frac{۲}{۳}\right)^{n-۲} = \frac{۱}{۲} + \frac{۱}{۴}\sum_{n=۲}^{\infty} n\left(\frac{۲}{۳}\right)^{n-۱}$$
$$= \frac{۱}{۲} + \frac{۱}{۴}\frac{۲(۲/۳) - (۲-۳)^۲}{(۱-۲/۳)^۲} = \frac{۱}{۲} + \frac{۱}{۴}(۸) = ۲.۵$$

بنابراین وضعیت صفر بازگشت پذیر مثبت است. میانگین زمان تکرار وضعیت صفر همچنین می تواند با اعمال اصلاحات روی اولین وضعیت مشاهده شده به دست آید.

$$\mu_{۰۰} = (۱/۲)E[T_{۰۰} \mid X_۱ = ۰] + (۱/۲)E[T_{۰۰} \mid X_۱ = ۱]$$
$$= ۱/۲ + (۱/۲)(۱ + E[T_{۱۰}])$$

از آنجا که $T_{۱۰}$ توزیع هندسی با میانگین ۳ دارد، نتیجه ی یکسانی به دست می آید. به طور مشابه برای وضعیت ۲، $f_{۲۲}$ می تواند این گونه برآورد شود:

$f_{22} = f_{22}^{(1)} = P_{22} = 1/4$

بنابراین وضعیت ۲ گذرا است. اگر وضعیت j بازگشت پذیر باشد، تعداد مراجعات به این وضعیت باید نامحدود باشد و اگر وضعیت j گذرا باشد، تعداد مراجعان به آن نیز محدود خواهد بود. برای یک وضعیت گذرای j داریم:

$$P(\text{تعداد مراجعات به } j = n \mid X_0 = j) = (f_{jj})^{n-1}(1 - f_{jj})$$

و به این بدان معنی است که تعداد مراجعات به وضعیت j دارای توزیع هندسی با میانگین $\frac{1}{(1-f_{jj})}$ می باشد.

قضیه ۱–۴– وضعیت j بازگشت پذیر است اگر و تنها اگر $\sum_{n=1}^{\infty} P_{jj}^{(n)} = \infty$.

اثبات: برای اثبات کافی بودن فرض کنید وضعیت j بازگشت پذیر است. آنگاه تعداد مراجعات به وضعیت j نامحدود خواهد بود و بنابراین مقدار مورد انتظار برای تعداد مراجعات به وضعیت j نیز نامحدود خواهد بود. یعنی :

$$E(\text{تعداد مراجعات به وضعیت } j \mid X_0 = j) = \infty$$

اگر یک متغیر نشانه‌گر تعریف کنیم:

$$I_n = \begin{cases} 1 \longleftrightarrow if \ X_n = j \\ 0 \longleftrightarrow otherwise \end{cases}$$

بنابراین $\sum_{n=1}^{\infty} I_n$ تعداد مراجعات به وضعیت j را کاهش خواهد داد و از اینجا داریم:

$$E[\text{تعداد مراجعات به } j \mid X_0 = j] = E\left[\sum_{n=1}^{\infty} I_n \mid X_0 = j\right]$$

$$= \sum_{n=1}^{\infty} E[I_n \mid X_0 = j]$$

$$= \sum_{n=1}^{\infty} P\{X_n = j \mid X_0 = j\}$$

$$= \sum_{n=1}^{\infty} P_{ij}^{(n)}$$

شرط التزام نیز بلافاصله بعد از آن می آید.

قضیه ۱–۵– : در یک سری مارکوف با تعداد متناهی از وضعیت ها، همه ی وضعیت ها گذرا نیستند.

اثبات) از تناقض استفاده کنید.بنابراین در یک سری مارکوف با تعداد وضعیت های متناهی، حداقل یک وضعیت تکرار شونده وجود دارد.

قضیه ۱–۶–: اگر وضعیت i بازگشت پذیر باشد و $i \leftrightarrow j$ آنگاه j نیز بازگشت پذیر است.

اثبات: فرض کنید i بازگشت پذیر است و $i \leftrightarrow j$. بنابراین $m>۰$ و $n>۰$ وجود دارند به گونه‌ای که $P_{ij}^{(m)} > ۰, P_{ji}^{(n)} > ۰$.

همچنین از رابطه ی کولموگورف–چپمن و به ازای $S>۰$ نتیجه‌ی زیر حاصل می شود :

$$P_{jj}^{(m+n+s)} \geq P_{ji}^{(m)} P_{ii}^{(n)} P_{ij}^{(s)}.$$

بنابراین :

$$\sum_{s=۱}^{\infty} P_{jj}^{(m+n+s)} \geq P_{ji}^{(m)} P_{ij}^{(n)} \sum_{s=۱}^{\infty} P_{ii}^{(s)} = \infty$$

و لذا نتیجه حاصل می شود.

بنابر این وضعیت های موجود در یک کلاس همگی یا گذراهستند و یا بازگشت پذیر. همچنین مشاهده می شود که همه‌ی وضعیت های یک سری مارکوف غیر قابل تقلیل با تعداد محدود از وضعیت ها، بازگشت پذیر هستند.

مثال ۱–۱۰–: یک قدم زدن تصادفی ساده دارای وضعیت های گذرا یا بازگشت پذیر می باشد

وضعیت صفر را در نظر بگیرید مقدار $\sum_{n=۱}^{\infty} P_{..}^{(n)}$ برابر است با.

$$P_{..}^{(۲m+۱)} = ۰, \quad m = ۰,۱,...$$

با استفاده از تخمین استرلینگ[۲] برای $n!$ به شکل

۲-*stiling*