دانشکده مهندسی برق

عنوان

زنجیره و فرآیند مارکوف

نام نگارنده: ساناز مطیع

استاد راهنما: دکتر سعید عبادالهی

شماره دانشجویی: ۹۹۴۱۳۰۸۲

زمستان ۱۴۰۱

بسم الله الرحمن الرحیم

چکیده

فرآیندهای تصادفی، یکی از تئوری‌های مدل‌سازی است که براساس آمار و احتمال شکل گرفته و برای تحلیل داده‌ها به کار می‌رود. در اکثر موارد فرآیندهای تصادفی برحسب زمان فهرست‌بندی شده‌اند. زنجیره مارکوف[1] یا فرآیند مارکوف[2]، مدلی برای نمایش دنباله‌ای از متغیرهای تصادفی است که در آن احتمال رویداد هر پیشامد فقط به پیشامد قبلی وابسته است. به این ترتیب احتمال رخداد پیشامدها در چنین مدلی فقط به زمان قبل وابسته بوده و بقیه پیشامدها در میزان احتمال دخالت نمی‌کنند. چنین وضعیتی را برای فرایند تصادفی گاهی خاصیت عدم حافظه نیز می‌نامند. این مدل به افتخار ریاضی‌دان روسی **آندری مارکوف** که در سال‌های اولیه قرن بیستم در این زمینه دست به نوآوری زده بود، مدل مارکوف یا زنجیره مارکوف نامیده می‌شود.

واژه‌های کلیدی: زنجیره مارکوف، فرآیند مارکوف، حالت مانا زنجیره مارکوف

---

[1] Markov Chain
[2] Markov Process

فهرست مطالب

فهرست تصاویر

فصل۱: مقدمه

۱.۱ مقدمه

در زندگی روزمره به مدل‌های زیادی برمی‌خوریم که احتمال شرایط اینده فقط به شرایط حال حاضر بستگی دارد و نه قبل‌تر از آن مانند هواشناسی، بازار بورس، میزان بهره‌برداری و آب موجود در سدها و از این دست در این شرایط برای محاسبه شانس وقوع پیشامدها لازم است مدل مناسبی انتخاب شود. در سال ۱۹۱۳، ریاضیدان روسی آندریویچ مارکوف این شاخه از احتمال را پایه گذاری کرد.

فرایندهای تصادفی که با اطلاع از وضعیت کنونی، وضعیت‌های گذشته آن‌ها اثری بر احتمال‌های شرطی پیشامدهایی که در آینده رخ می‌دهند، ندارد، فرایند مارکوف می‌گویند. فرایند مارکوف را که فضایی وضعیت آن متناهی یا نامتناهی شماراست، زنجیره مارکوف می‌گوییم. زنجیره مارکوف یک روش ریاضی برای مدل‌بندی فرایندهای تصادفی است. که توالی از مشاهدات را در طول زمان نشان می‌دهد. وابستگی این زنجیره به زمان یا از طریق ضرایب همبستگی و یا با استفاده از ماتریس‌های احتمال انتقال بیان می‌شود. تجزیه و تحلیل زنجیره مارکوف راه‌حلی است برای تحلیل کنونی متغیرهای معلوم تا پیش‌بینی حرکات آینده متغیرها امکان پذیر باشد و در صورتی مارکوف مرتبه اول نتواند پاسخگوی احتمالات دقیق آینده باشد باید از درجه بالاتر این تحلیل استفاده شود.

۱.۲ ساختار گزارش

در این مقاله ابتدا به تاریخچه پیدایش زنجیره مارکوف پرداخته می‌شود. سپس خواص آن گفته می‌شود و در آخر به اثبات یکی از مهم ترین ویژگی‌های آن یعنی یکتا بودن توزیع مانا پرداخته می‌شود.

فصل دوم: تاریخچه

۲.۱ مقدمه

در این بخش به‌طور خلاصه به بررسی تاریخچه پیدایش زنجیره مارکوف پرداخته می‌شود.

۲.۲ تاریخچه پیدایش زنجیره مارکوف

فرض کنید متنی به شما داده می‌شود و یک حرف به صورت تصادفی انتخاب می‌شود و از شما خواسته می‌شود حدس بزنید که آیا این حرف صامت است یا مصوت. بهترین حدس شما حرف صامت است زیرا تعداد حروف صامت بیشتر ازحروف مصوت است، در نتیجه احتمال اینکه درست گفته باشید بیشتر است. حال فرض کنید تصمیم داریم کمی اطلاعات بیشتری بدهیم و به شما بگوییم حرف قبل از حرفی که انتخاب کرده‌اید صامت است یا مصوت، آیا اکنون استراتژی بهتری وجود دارد که بتوانید دنبال کنید؟

در سال ۱۹۱۳، ریاضیدان روسی آندری آندریویچ مارکوف[۳] شاخه جدیدی از نظریه احتمال را با به کار بردن ریاضیات در شعر پایه‌گذاری کرد. او در حال تلاش برای پاسخ به پرسش فوق بود که به تجزیه و تحلیل در متن رمان الکساندر پوشکین در شعر یوجین اونگین پرداخت. او دریافت که ۴۳٪ درصد حروف مصوت و۵۷٪ صامت‌ها هستند پس در مسئله اول همیشه باید صامت را حدس زد و ۵۷٪ احتمال دارد که درست باشد.

با این حال در ۸۷٪ مواقع بعد از یک مصوت یک حرف صامت می‌آید و در ۶۶٪ پس از یک حرف صامت یک مصوت می‌آید. در نتیجه در پاسخ به پرسش دوم استراتژی بهتر این است که عکس حرف قبلی را حدس بزنیم در نتیجه دانستن حرف قبل احتمال درست حدس زدن ما را بیشتر می‌کند.

بینش واقعی زمانی حاصل شد که مارکوف تحلیل را یک قدم جلوتر برد. مارکوف تحقیق کرد که آیا دانش در مورد دو حرف قبل مزیت بیشتری به همراه دارد یا خیر. او دریافت که هیچ مزیت قابل توجهی برای دانستن دو حرف قبلی اضافی وجود ندارد. این منجر به ایده اصلی زنجیره مارکوف شد. در حالی که نتایج متوالی مستقل از هم نیستند، فقط جدیدترین نتیجه برای پیش بینی نتیجه بعدی استفاده می‌شود.

[3] Андре́й Андре́евич Ма́рков

تصویر ۲.۱ آندری آندریویچ مارکوف

۲.۳ جمع بندی و تحلیل این فصل

در این فصل به تاریخچه پیدایش زنجیره مارکوف اشاره کردیم. در فصل بعد به توضیح ریاضی زنجیره و فرآیند مارکوف پرداخته می‌شود.

فصل سوم: زنجیره مارکوف و فرآیند مارکوف زمان گسسته

۳.۱ مقدمه

در این فصل به تعریف ریاضی زنجیره و فرآیند مارکوف زمان گسسته پرداخته می‌شود. سپس برخی از مهم‌ترین خواص آن ذکرمی‌شود.

۳.۲ تعریف زنجیره مارکوف

دنباله‌ای از متغیرهای تصادفی ...X1,X2, را که احتمال تغییر وضعیت از زمان t به t+1 مستقل از وضعیت‌های قبلی باشد را یک زنجیره مارکوف می‌نامند. این گزاره را به بیان متغیرهای تصادفی و تابع احتمال به صورت زیر نشان می‌دهیم.

$$Pr(Xt+1=x|X1=x1,X2=x2,\ldots,Xn=xt)=Pr(Xt+1=x|Xt=xt) \qquad (۳.۱)$$

در محاسبه این احتمال شرطی باید Pr(X1=x1,…,Xn=xn)>0 باشد در غیر اینصورت امکان محاسبه احتمال شرطی وجود ندارد. این احتمال به معنی طی کردن مسیر (X1=x1,…,Xn=xn) با احتمال مثبت در فرآیند تصادفی است. بنابراین در فرآیند تصادفی مارکوف، امکان رسیدن به نقطه Xt=xt از مسیر یاد شده وجود دارد.

۳.۳ تعریف فرآیند مارکوف

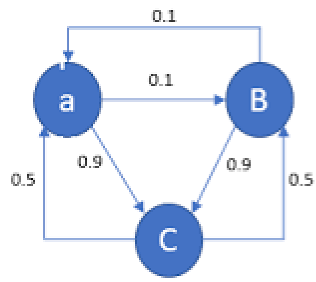
اگر یک فرآیند تصادفی به صورت دنباله‌ای نامتناهی از متغیرهای تصادفی معرفی شود به‌طوری که مقادیر قبلی بر مقادیر جدید تاثیر نگذارند گفته می‌شود که دارای خاصیت مارکوفی است و آن را فرآیند تصادفی مارکوف می‌نامند. اگر فضای حالت در این فرآیند متناهی بوده، فرآیند را متناهی و در غیر اینصورت نامتناهی می‌نامند.

۳.۴ ماتریس احتمال انتقال[۴]

پیاده‌روی تصادفی در زنجیره مارکوف از یک حالت شروع می‌شود. مثلا اگر در حالت x باشد، حالت بعدی y به طور تصادفی با احتمال p_{xy} انتخاب می‌شود. برای نمایش پیاده‌روی تصادفی در زنجیره مارکوفی معمولا از گراف جهت‌دار استفاده می‌شود. مقدار هر یک از یال‌های این گراف، احتمال انتقال از یک راس به راس دیگر را نشان می‌دهد.

از آنجایی این گراف را می‌توان به صورت یک ماتریس نمایش داد، ماتریس این زنجیره مارکوف نیز به صورت ماتریس احتمال انتقال قابل نمایش است. در این حالت عناصر ماتریس که به صورت p_{ij} نوشته می‌شوند، احتمال انتقال از نقطه i به j را نشان می‌دهد. واضح است که در اینجا فضای حالت دارای سه عنصر S={a,b,c} است.

[4] Transition Probability Matrix



$$P = \begin{bmatrix} 0 & 0.1 & 0.9 \\ 0.1 & 0.0 & 0.9 \\ 0.5 & 0.5 & 0.0 \end{bmatrix}$$

تصویر ۳.۱ زنجیره مارکوف و ماتریس انتقال با فضای حالت سه وضعیتی

در ماتریس انتقال با توجه به اصول احتمال باید $\sum_{i=0}^{k} pij = 1$ باشد. یعنی مجموع احتمالات هر سطر در ماتریس باید برابر با ۱ باشد.

بطور کلی در زنجیره مارکوف، احتمال اینکه از وضعیت i به j در گام n ام برسیم به صورت $p_{\mathrm{ij}}^{(\mathrm{n})}$ نشان داده می‌شود. به علت وجود خاصیت مارکوفی در زنجیره یا فرآیند مارکوفی داریم:

$$p_{ij}^{(n)} = \sum_{\mathbf{r}\in\mathbf{S}} p_{ir}^{(k)}\, p_{rj}^{(n-k)}, \forall \mathrm{k}, 0<\mathrm{k}<\mathrm{n} \qquad (۳.۲)$$

در حقیقت اگر بخواهیم احتمال گذر از یک نقطه و رسیدن به نقطه‌ای دیگر از فضای حالت را در n گام محاسبه کنیم باید ماتریس انتقال فرآیند را n بار در خودش ضرب کنیم.

۳.۵ خواص زنجیره مارکوف

۳.۵.۱ **ویژگی اساسی زنجیره مارکوف معادله زیر است.**

$$\vec{\mathrm{x}}\,(\mathrm{n}+1) = \mathrm{A}\vec{\mathrm{x}}(\mathrm{n}) \qquad (۳.۳)$$

یعنی تابع در هر مرحله فقط به مرحله قبلی خود وابسته است و نقطه شروع اهمیتی ندارد.

۳.۵.۲ **ماتریس انتقال یک ماتریس تصادفیست.**

ماتریس انتقال زنجیره مارکوف زیرمجموعه ماتریس‌هایی ست که به آن ماتریس تصادفی[5] می‌گویند. یک ماتریس تصادفی یک ماتریس $\mathrm{n} \times \mathrm{n}$ است که همه درایه‌ها آن غیر منفی است و مجموع مقادیر هر ردیف یک است. اگر

[5] stochastic matrix

P و Q هر دو ماتریس‌های تصادفی باشند، آنگاه P Q نیز یک ماتریس تصادفی‌ست، و اگر μ یک بردار احتمال باشد، μP نیز یک بردار احتمال است.

۳.۵.۳ زنجیره مارکوف مرتبه k- با حافظه

همانطور که اشاره شد، زنجیره مارکوف دارای خاصیت عدم حافظه است. اگر میزان حافظه زنجیره مارکوف به k مرحله یا زمان قبل محدود شود، به آن زنجیره مارکوف با حافظه یا مرتبه k گفته می‌شود. در این حالت رابطه زیر برقرار خواهد بود.

$$Pr(Xn=xn|Xn-1=xn-1,Xn-2=xn-2,\ldots,X1=x1)=Pr(Xn=xn|Xn-1=xn-1,Xn-2=xn-2,\ldots,Xn-k=xn-k), \text{ for } n>k \quad (۳.۴)$$

۳.۵.۴ زنجیره مارکوف تناوبی[۶]

اگر بتوان از وضعیت i با گام‌هایی از مضرب k مجدداً به نقطه i رسید، زنجیره مارکوف را تناوبی گویند. در چنین حالتی مقدار تناوب زنجیره به شکل زیر محاسبه می‌شود.

$$k=\gcd\{n>0: Pr(Xn=i| X_0=i) > 0\} \quad (۳.۵)$$

۳.۵.۵ زنجیره مارکوف تقلیل‌ناپذیر[۷]

یک زنجیره مارکوف تقلیل‌ناپذیر است اگر برای همه حالت‌ها $x, y \in \Omega$ یک $t \geq 0$ وجود داشته باشد به صورتی که $P^t(x, y) > 0$. به طور شهودی، این بدان معنی است که می‌توان از x به y برای هر $x , y \in \Omega$ در تعداد محدودی مرحله رسید. در واقع یعنی توالی ای از حالات وجود دارند که $x = x_0, x_1, \ldots , x_{t-1}, x_t = y$ (که به آن یک مسیرگفته می‌شود) و می‌تواند زنجیره را از x به y ببرد به صورتی که برای تمام $0 \leq i < t$ ، $P(xi , xi+1) > 0$

۳.۵.۶ توزیع مانا

توزیع π، توزیع مانا زنجیره مارکوفP نامیده می‌شود اگر $\pi P = \pi$. توزیع مانا زنجیره مارکوف برای تمام زنجیره مانا است یعنی اگر توزیع Xt توزیع مانا π باشد، آنگاه توزیع مانا X_{t+1} نیز π است.

---

[6] Periodic
[7] Irreducible

۳.۶ جمع بندی و تحلیل این فصل

در این فصل به توضیح دقیقی از زنجیره و فرآیند مارکوف پرداختیم و برخی ویژگی‌های مهم آن را بررسی کردیم در فصل بعد اثبات می‌شود که چرا توزیع مانا زنجیره مارکوف یکتاست.

فصل چهارم: یکتایی توزیع مانا زنجیره مارکوف

۴.۱ مقدمه

تمام زنجیره‌های مارکوف تقلیل‌ناپذیر داری توزیع مانا هستند در این فصل به اثبات این پرداخته می‌شود که در صورت وجود توزیع مانا یکتاست.

۴.۲ اثبات یکتایی توزیع مانا

تعریف: تابع $h : \Omega \to R$ در $x \in \Omega$ هارمونیک است اگر $h(x) = \sum_{y\in\Omega} P(x,y)h(y)$.

اگر h در تمام حالت‌های $\Omega = \{x1, x2, . . . , xn\}$ هارمونیک باشد گفته می‌شود h در Ω هارمونیک است. در نتیجه $Ph=h$ و $h=\begin{pmatrix} h(x1) \\ h(x2) \\ \vdots \\ h(xn) \end{pmatrix}$.

اگر P تقلیل‌ناپذیر باشد و h روی Ω هارمونیک باشد، h یک تابع مانا است.که به این صورت اثبات می‌شود.

از آنجایی که Ω متناهی است، h در $x_0 \in \Omega$ به حداکثر می‌رسد، به طوری که $h(x_0) \geq h(y)\ \forall y \in \Omega$. اگر $z\in\Omega$ و $P(x_0, z) > 0$ باشد و فرض کنید $h(z) < h(x_0)$. از آنجایی که h در x_0 هارمونیک است.

$$
\begin{aligned}
h(x_0) &= \sum_{y\in\Omega} P(x_0,y)h(y) \\
&= P(x_0,z)h(z) + \sum_{y\in\Omega, y\neq z} P(x_0,y)h(y) \\
&\leq P(x_0,z)h(z) + \sum_{y\in\Omega, y\neq z} P(x_0,y)h(x_0) \\
&< P(x_0,z)h(x_0) + \sum_{y\in\Omega, y\neq z} P(x_0,y)h(x_0) \\
&= \left(\sum_{y\in\Omega} P(x_0,y)\right) h(x_0) \\
&= h(x_0),
\end{aligned}
\qquad (۳.۵)
$$

آخرین نابرابری شروط $P(x_0, z) > 0$ و $h(z) < h(x_0)$ را دارد. و نتیجه می‌دهد $h(x_0) < h(x_0)$، که یک تناقض است در نتیجه $h(x_0) = h(z)$.

اکنون برای هر $y \in \Omega$، تقلیل‌ناپذیر بودن P به معنای وجود مسیری از x_0 به y است.

اگر این مسیر $x_0, x_1, \ldots, x_n = y$ باشد به صورتی که $P(x_i, x_{i+1}) > 0$ بنابراین، $h(x_0) = h(x_1)$، و h در x_1 نیز ماکزیمم است یعنی $h(x_1) = h(x_2)$. با ادامه این روند $h(x_0) = h(x_1) = \cdots = h(x_n) = h(y)$.

حال نشان داده می‌شود که اگر توزیع مانا وجود داشته باشد یکتاست.

۴.۲.۱ قضیه: اگر P تقلیل‌ناپذیر باشد و توزیع مانا π داشته باشد، آنگاه π تنها توزیع مانا آن است.

اثبات: تنها توابع h که هارمونیک هستند توابعی به فرم $h(x) = c\ \forall x \in \Omega$ برای هر مقدار مانا c هستند. اگر h را به صورت برداری بنویسیم، تنها راه حل معادله $Ph = h$ یا معادل آن $(P - I)h = 0$ به صورت زیر است.

$$h = c\begin{pmatrix} 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \qquad (۳.۶)$$

بنابراین، $\dim(\ker(P - I)) = 1$ ، رتبه $(P - I)$ برابر است با $|\Omega| - 1$ [8].

رتبه $(P - I)$ = رتبه $((P - I)^T)$ = رتبه $(P^T - I) = |\Omega| - 1$. در نتیجه $\dim(\ker(P^T - I)) = 1$. در نتیجه برای تمام بردارهای سطری $v \in R^{|\Omega|}$ معادله $(P^T - I)v^T = 0$ فقط جواب‌هایی در فضای یک‌بعدی دارد. این معادله با $vP = v$ برابر است بنابراین، با توجه به اینکه $\pi P = \pi$ یک جواب است، تمام جواب‌ها برای یک اسکالر λ باید به فرم $v = \lambda\pi$ باشند. اگر λ=1 را در نظر بگیریم از آنجایی که p یک ماتریس انتقال است مجموع درایه‌های هر سطر آن برابر با یک است در نتیجه سطرهای ماتریس $P - I$ صفر هستند. این بدان معناست که سطرهای آن بهم وابسته هستند، پس $P - I$ یک ماتریس منفرد است در نتیجه λ=1 مقدار ویژه ماتریس p است. پس تنها توزیع مانا موجود $v = \pi$ است.

۴.۳ جمع بندی و تحلیل این فصل

در این فصل به این موضوع پرداختیم که توزیع مانا زنجیره مارکوف یکتاست و در زمان‌های مختلف تغییری نمی‌کند. در فصل بعد به نتیجه گیری مقاله پرداخته می‌شود.

[8] طبق قضیه rank-nullity

فصل پنجم: نتیجه‌گیری

۵.۱ نتیجه‌گیری

روش زنجیره‌های مارکوف یکی از ابزارهای توانمند برای مدلسازی پدیده‌های احتمالی می‌باشد که تاکنون در زمینه‌های مختلف علمی از آن استفاده شده است. زنجیره مارکوف کاربردهای واضح و روشنی در زمینه‌های مختلف از جمله پژوهش کار، امور مالی، علوم سیاسی، مهندسی شیمی، جمعیت‌شناسی و غیره دارد. خاصیت بی‌حافظگی زنجیره مارکوف در مدل‌های مختلف مانند بازار بورس و پیش بینی قیمت سهام کاربرد دارد.

در این متن تلاش شد تا برخی ویژگی‌ها و خواص زنجیره مارکوف اثبات شود تا در درک بهتر زنجیره مارکوف کمک کند.

منابع و مراجع

[1] https://www.americanscientist.org/article/first-links-in-the-markov-chain
[2] David A. Levin, Yuval Peres and Elizabeth L. Wilmer. Markov Chains and Mixing Times.
[3] http://pages.uoregon.edu/dlevin/MARKOV/markovmixing.pdf