http://math-sci.iranjournals.ir

ریاضی و جامعه
شاپا (چاپی): ۶۴۹۳-۲۳۴۵، شاپا(الکترونیکی): ۶۵۰۷-۲۳۴۵
جلد ۳، شماره ۲ (۱۳۹۷)، صص. ۱۹-۱۱

http://www.ui.ac.ir

گوگل چگونه کار می‌کند؟

کریستین روسو
مترجم: مریم لطفی‌پور

چکیده. نوشته حاضر، ترجمه مقاله زیر است:

C. Roussau, How Google works? klein vignette (www.kleinproject.org)

در این نوشته، الگوریتم رتبه‌بندی گوگل مورد بررسی قرار می‌گیرد و روش‌های ریاضی به‌کار رفته در آن شرح داده می‌شود.
در اینجا کاربرد مهمی از قضیه نقطه ثابت باناخ بیان شده و با استفاده از تکنیك اثبات این قضیه، نقطه ثابت مورد نیاز در الگوریتم رتبه‌بندی گوگل به‌دست می‌آید.

۱. مقدمه مترجم

در جنبه‌های مختلف پژوهش و همچنین آموزش ریاضیات، آگاهی از کاربردهای این علم در زمینه‌های مختلف از اهمیت بسیاری برخوردار است. دانستن ارتباط و کاربرد ریاضیات در سایر علوم، با ایجاد انگیزه مضاعف در یادگیری آن می‌تواند به فهم بیشتر مسائل و ترویج این دانش بنیادی کمك شایانی نماید. در این نوشته تلاش شده است که از سویی نقش و کاربرد عمیق ریاضیات در پایه‌ریزی پروژه‌های بزرگ کاربردی مانند موتورهای جستجو نشان داده شده و از طرف دیگر چگونگی ادغام شاخه‌های گوناگون ریاضی مانند جبرخطی، نظریه احتمال و نقطه ثابت در راستای نیل به یك هدف مشترك به نمایش گذاشته شود.

به علاوه در این نوشته می‌توان دریافت که چگونه یك مفهوم ریاضی مانند نقطه ثابت می‌تواند فراتر از خاستگاه اصلی خود در سایر علوم کاربرد داشته و مورد استفاده قرار گیرد.

۲. متن اصلی

گوگل از همان ابتدای فعالیت خود به یك موتور جستجوی شناخته شده مبدل شد. این جایگاه به‌دلیل برتری الگوریتم رتبه‌بندی آن به‌دست آمده است: الگوریتم رتبه صفحه (الگوریتم پیج‌رنك). در حقیقت با توجه به حجم وسیع صفحات در شبکه گسترده جهانی، بسیاری از جستجوها به هزاران یا میلیون‌ها نتیجه می‌انجامد. حال از آنجا که هیچ کس قادر به بررسی میلیون‌ها ورودی نمی‌باشد، اگر ترتیب‌گذاری مناسب انجام نگردد، این جستجو نتیجه مفیدی در برنخواهد داشت.

عبارات و کلمات کلیدی. الگوریتم پیج‌رنك، نقطه ثابت باناخ، زنجیره مارکوف.
دبیرتخصصی رابط: مجید فخار
تاریخ دریافت: ۱۳۹۶/۱۲/۲۲ تاریخ پذیرش: ۱۳۹۷/۰۲/۲۲

http://dx.doi.org/10.22108/msci.2018.110135.1270

۱.۲. الگوریتم پیج رنك چگونه کار می‌کند؟ ما این موضوع را تشریح می‌کنیم. اما قبل از آن بیایید جستجویی در گوگل انجام دهیم. در تاریخ ۴ ژوئن ۲۰۱۰ با وجود اینکه پروژه کلاین به تازگی کار خود را آغاز نموده است، تعداد ۰۰۰، ۳۰۰، ۱۶ نتیجه برای آن به‌دست می‌آید. دقیقاً در آن تاریخ اولین نتیجه صفحه زیر است،

http://www.mathunion.org/icmi/other-activities/klein-project/introduction/

و نه این صفحه

http://www.kleinproject.org/

url اول مربوط به صفحه‌ای است که در وب سایت اتحادیه بین‌المللی ریاضیات قرار دارد،

http://www.mathunion.org

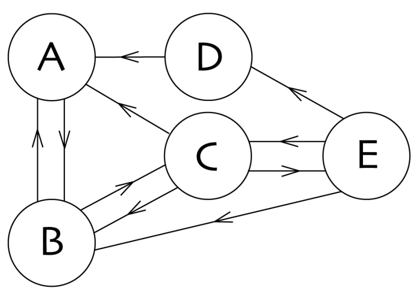
به‌دلیل اهمیت اتحادیه بین‌المللی ریاضیات، در زمان جستجوی «اتحادیه بین‌المللی ریاضیات» وب سایت رسمی آن در ابتدا قرار می‌گیرد. علاوه بر این، برخی موضوعات دارای اهمیت خود را به تمامی صفحاتش مرتبط می‌کند که یکی از آنها url زیر است،

http://www.mathunion.org/icmi/other-activities/klein-project/interoduction/

در چند ماه یا چند سال آینده می‌توان انتظار داشت که صفحه زیر به‌عنوان اولین نتیجه جستجوی پروژه کلاین ظاهر شود

http://www.kleinproject.org/

برای توصیف الگوریتم، شبکه را به‌صورت یك گراف جهت‌دار مدل‌سازی می‌نماییم. رئوس، صفحات و یال‌های جهت‌دار، لینك‌های بین صفحات هستند. همان‌طور که توضیح داده شد، هر صفحه به یك url متفاوت مربوط می‌شود. بنابراین یك وب‌سایت می‌تواند صفحات متعددی داشته باشد. این مدل تمایزی بین صفحات مجزای یك وب‌سایت و صفحه اول آن قائل نمی‌شود. اما این الگوریتم با احتمال زیاد به صفحه اول یك وب‌سایت مهم رتبه بهتری می‌دهد.



شکل ۱: یك شبکه ساده

۳. یك مثال ساده

شبکه ساده شکل (۱) با پنج صفحه به نام‌های A، B، C، D و E را در نظر بگیرید. این صفحه پنج لینك دارد. اگر ما در صفحه A باشیم آن‌گاه تنها یك لینك به صفحه B وجود دارد، در حالی که اگر در صفحه C باشیم، می‌توان سه لینك یافت و حرکت به یکی از صفحات A، B یا E را انتخاب کرد. توجه داشته باشید که از هر صفحه حداقل یك لینك وجود دارد.

ما یك بازی انجام می‌دهیم که به بیان ساده پیاده‌روی تصادفی بر یك گراف جهت‌دار است. با شروع از یك صفحه در هر گام به طور تصادفی یك لینك از صفحه‌ای که در آن قرار داریم انتخاب می‌کنیم و آن را دنبال می‌کنیم. به عنوان مثال

اگر از صفحه B شروع کنیم آنگاه می‌توان با احتمال $\frac{۱}{۲}$ به هر یك از صفحات A یا C رفت. در حالی‌که اگر از صفحه D حرکت کنیم، لزوماً با احتمال ۱ به صفحه A می‌رویم. بازی را تکرار می‌کنیم.

پس از n گام کجا خواهیم بود؟

برای انجام فرآیند به‌طور خودکار، شبکه را در ماتریس p به‌صورت زیر خلاصه می‌کنیم به‌طوری که هر ستون نشان‌دهنده صفحه خروج و هر سطر، صفحه‌ای باشد که به آن وارد می‌شویم،

$$P = \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} ۰ & \frac{۱}{۲} & \frac{۱}{۳} & ۱ & ۰ \\ ۱ & ۰ & \frac{۱}{۳} & ۰ & \frac{۱}{۳} \\ ۰ & \frac{۱}{۲} & ۰ & ۰ & \frac{۱}{۳} \\ ۰ & ۰ & ۰ & ۰ & \frac{۱}{۳} \\ ۰ & ۰ & \frac{۱}{۳} & ۰ & ۰ \end{bmatrix} \end{array} \begin{array}{c} A \\ B \\ C \\ D \\ E \end{array}$$

توجه داشته باشید که مجموع درایه‌های هر ستون P برابر با یك است و همه درایه‌ها بزرگتر یا مساوی صفر هستند. ماتریس‌هایی با این دو ویژگی اهمیت خاصی دارند. چنین ماتریسی مربوط به یك فرآیند زنجیره مارکوف است و ماتریس انتقال مارکوف نیز نامیده می‌شود. در این ماتریس همواره یکی از مقدارهای ویژه ۱ است و بردار ویژه‌ای متناظر با این مقدار ویژه ۱ وجود دارد که تمام درایه‌های آن کمتر یا مساوی ۱ و بزرگتر یا مساوی صفر با مجموع ۱ هستند. پیش از یادآوری تعاریف مقدار ویژه و بردار ویژه، مزیت نمایش ماتریسی گراف شبکه را بررسی می‌کنیم.

متغیر تصادفی X_n را با مقادیری از مجموعه $\{A, B, C, D, E\}$ که شامل N صفحه (در اینجا $N = ۵$) در نظر می‌گیریم. X_n صفحه‌ای را نشان می‌دهد که پس از n گام تصادفی در آن قرار داریم. بنابراین اگر p_{ij} درایه‌ای از ماتریس P باشد که در سطر iام و ستون jام قرار دارد، آنگاه p_{ij} احتمال شرطی این است که پس از $n+۱$ گام در صفحه iام قرار داشته باشیم با این فرض که در گام nام در صفحه jام بوده‌ایم،

$$p_{ij} = Prob(X_{n+۱} = i \mid X_n = j).$$

دقت کنید که این احتمال مستقل از n است! گوییم فرآیند زنجیره مارکوف فاقد حافظه‌ای از گذشته است. این مطلب که احتمالات بعد از دو گام را می‌توان در ماتریس $P^۲$ خلاصه کرد کار دشواری نیست.

این مطلب را اثبات می‌کنیم (می‌توانید از این اثبات بگذرید)

قضیه احتمال کل نتیجه می‌دهد

$$Prob(X_{n+۲} = i \mid X_n = j) = \sum_{k=۱}^{N} Prob(X_{n+۲} = i, X_{n+۱} = k \mid X_n = j).$$

با استفاده از تعریف احتمال شرطی داریم

$$Prob(X_{n+۲} = i \mid X_n = j) = \sum_{k=۱}^{N} \frac{Prob(X_{n+۲} = i, X_{n+۱} = k, X_n = j)}{Prob(X_n = j)}.$$

http://dx.doi.org/10.22108/msci.2018.110135.1270

یك ترفند آشنا را به کار می‌بریم: مقداری را ضرب و تقسیم می‌نماییم

$$\begin{aligned} &Prob(X_{n+۲}=i \mid X_n=j) \\ &\quad = \sum_{k=۱}^{N} \frac{Prob(X_{n+۲}=i, X_{n+۱}=k, X_n=j)}{Prob(X_{n+۱}=k,\ X_n=j)} \frac{Prob(X_{n+۱}=k,\ X_n=j)}{Prob(X_n=j)}. \end{aligned}$$

خارج‌قسمت کسر اول برابر است با

$$Prob(X_{n+۲}=i \mid X_{n+۱}=k,\ X_n=j) = Prob(X_{n+۲}=i \mid X_{n+۱}=k),$$

زیرا فرآیند زنجیره مارکوف فاقد حافظه‌ی گام گذشته است. بنابراین

$$\begin{aligned} Prob(X_{n+۲}=i \mid X_n=j) &= \sum_{k=۱}^{N} Prob(X_{n+۲}=i \mid X_{n+۱}=k) Prob(X_{n+۱}=k \mid X_n=j) \\ &= \sum_{k=۱}^{N} p_{ik}p_{kj} \\ &= (P^{۲})_{ij}. \end{aligned}$$

در این مثال

$$P^{۲} = \begin{array}{cc} \begin{array}{ccccc} A & B & C & D & E \end{array} & \\ \left[\begin{array}{ccccc} \frac{۱}{۲} & \frac{۱}{۶} & \frac{۱}{۶} & ۰ & \frac{۱۱}{۱۸} \\ ۰ & \frac{۲}{۳} & \frac{۴}{۹} & ۱ & \frac{۱}{۹} \\ \frac{۱}{۲} & ۰ & \frac{۵}{۱۸} & ۰ & \frac{۱}{۶} \\ ۰ & ۰ & \frac{۱}{۹} & ۰ & ۰ \\ ۰ & \frac{۱}{۶} & ۰ & ۰ & \frac{۱}{۹} \end{array}\right] & \begin{array}{c} A \\ B \\ C \\ D \\ E \end{array} \end{array}$$

با تکرار این ایده، واضح است که درایه $(P^m)_{ij}$ از ماتریس P^m، احتمال $Prob(X_{n+m}=i \mid X_n=j)$ را نشان می‌دهد. به عنوان مثال

$$P^{۳۲} = \begin{array}{cc} \begin{array}{ccccc} A & B & C & D & E \end{array} & \\ \left[\begin{array}{ccccc} ۰٫۲۹۳ & ۰٫۲۹۳ & ۰٫۲۹۳ & ۰٫۰۲۹۳ & ۰٫۲۹۳ \\ ۰٫۰۳۹۰ & ۰٫۳۹۰ & ۰٫۳۹۰ & ۰٫۳۹۰ & ۰٫۳۹۰ \\ ۰٫۲۲۰ & ۰٫۲۲۰ & ۰٫۲۲۰ & ۰٫۲۲۰ & ۰٫۲۲۰ \\ ۰٫۰۲۴ & ۰٫۰۲۴ & ۰٫۰۲۴ & ۰٫۰۲۴ & ۰٫۰۲۴ \\ ۰٫۰۷۳ & ۰٫۰۷۳ & ۰٫۰۷۳ & ۰٫۰۷۳ & ۰٫۰۷۳ \end{array}\right] & \begin{array}{c} A \\ B \\ C \\ D \\ E \end{array} \end{array}$$

اگر با دقت ۳ رقم اعشار کار کنیم تمام ستون‌های $P^{۳۲}$ یکسان می‌شوند و در مورد ستون‌های P^n که $n > ۳۲$ نیز اتفاق مشابهی رخ می‌دهد. اگر دقت بالاتری در نظر بگیریم برای nهای بزرگتر از ۳۲ به پایدارسازی نیز می‌رسیم. بنابراین پس از n گام، برای nهای به‌اندازه کافی بزرگ، احتمال بودن در یك صفحه، مستقل از نقطه شروع است.

علاوه بر این فرض می‌کنیم

$$\pi^t = (۰٫۲۹۳, ۰٫۳۹۰, ۰٫۲۲۰, ۰٫۰۲۴, ۰٫۰۷۳)$$

http://dx.doi.org/10.22108/msci.2018.110135.1270

(π یك بردار قائم و ترانهاده آن π^t یك بردار افقی است).

به‌آسانی می‌توان تحقیق کرد که $P\pi = \pi$. اگر مؤلفه iام بردار π^t را به‌عنوان احتمال بودن در صفحه i در زمان n و درنتیجه π^t را به‌عنوان توزیع احتمال صفحات در زمان n در نظر بگیریم، آنگاه توزیع احتمال در زمان $n+1$ نیز همین مقدار می‌باشد. به‌این دلیل، بردار π توزیع مانا نامیده می‌شود. این توزیع مانا رتبه‌بندی صفحات را میسر می‌سازد. در این مثال صفحات را به‌صورت B، A، C، E و D رتبه‌بندی و B را به‌عنوان مهم‌ترین صفحه اعلام می‌کنیم.

۴. حالت کلی

در حالت کلی دقیقاً مشابه مثال بالا عمل می‌کنیم. شبکه را به‌صورت یك گراف جهت‌دار نمایش می‌دهیم که N رأس آن معرف N صفحه شبکه و یال‌های جهت‌دار آن لینك‌های بین صفحات باشند. این گراف را در یك ماتریس $N \times N$ به نام P خلاصه می‌کنیم، به‌گونه‌ای که ستون jام آن صفحه خروجی jام و سطر iام آن iامین صفحه ورودی را نشان دهد. در مثال بالا برداری به نام π که $P\pi = \pi$ یافتیم. این بردار یك بردار ویژه برای مقدار ویژه ۱ است. در این قسمت تعریف مقدار ویژه و بردار ویژه را یادآوری می‌نماییم.

تعریف ۱.۴. *فرض کنید P یك ماتریس $N \times N$ باشد. عدد $\lambda \in C$ یك مقدار ویژه برای ماتریس P است هرگاه بردار غیرصفر $X \in \mathbb{C}^N$ وجود داشته باشد که $PX = \lambda X$. این بردار را یك بردار ویژه برای P می‌نامند.*

در اینجا روش یافتن مقدارهای ویژه و بردارهای ویژه را نیز یادآوری می‌کنیم.

فرض کنید P یك ماتریس $N \times N$ باشد. مقادیر ویژه ماتریس P ریشه‌های چندجمله‌ای مشخصه $det(\lambda I - P) = 0$ هستند که در آن I ماتریس همانی $N \times N$ است. بردارهای ویژه متناظر با یك مقدار ویژه λ جواب‌های غیرصفر دستگاه خطی همگن $(\lambda I - P)X = 0$ هستند.

در ادامه قضیه فروبنیوس تضمین می‌کند که همواره می‌توان برای ماتریس متناظر با گراف شبکه یك جواب مانا یافت.

قضیه ۲.۴. *(قضیه فروبنیوس) فرض کنید $P = (\rho_{ij})$ یك ماتریس انتقال مارکوف $N \times N$ باشد (برای هر i و j، $\rho_{ij} \in [0, 1]$ و مجموع درایه‌های هر ستون برابر با ۱ است یعنی $\sum_{i=1}^{N} \rho_{ij} = 1$) آنگاه*

- *$\lambda = 1$ یك مقدار ویژه برای P است.*
- *اگر λ مقدار ویژه برای P باشد آنگاه $|\lambda| \leq 1$.*
- *متناظر با مقدار ویژه ۱، بردار ویژه‌ای مانند π وجود دارد که تمام مؤلفه‌های آن بزرگ‌تر یا مساوی صفر هستند. بدون کاستن از کلیت می‌توان فرض کرد که مجموع مؤلفه‌های آن برابر با ۱ است.*

اکنون زمان آن رسیده است که اهمیت این قضیه را نشان دهیم. در این راستا برای ساده‌تر شدن مطلب، فرض می‌کنیم ماتریس P پایه‌ای از بردارهای ویژه مانند $B = \{v_1, \dots, v_n\}$ دارد و v_1 بردار π مربوط به قضیه فروبینوس است. برای هر v_i می‌توان λ_i را به‌گونه‌ای یافت که $Pv_i = \lambda_i v_i$. بردار غیرصفر X به‌صورت $X^t = (x_1, \dots, x_N)$ را که $x_i \in [0, 1]$ و $\sum_{i=1}^{N} x_i = 1$ در نظر می‌گیریم. X را به شکل ترکیبی از اعضای پایه B به‌صورت زیر نمایش می‌دهیم،

$$X = \sum_{i=1}^{N} a_i v_i.$$

http://dx.doi.org/10.22108/msci.2018.110135.1270

با اثباتی تکنیکی که در اینجا از آن صرف‌نظر می‌کنیم، می‌توان نشان داد که $a_۱ = ۱$. حال PX را محاسبه می‌نماییم،

$$PX = \sum_{i=۱}^{N} a_i P v_i = \sum_{i=۱}^{n} a_i \lambda_i v_i,$$

زیرا v_i یک بردار ویژه متناظر با مقدار ویژه λ_i است. با تکرار این عمل نتیجه می‌گیریم

$$P^n X = \sum_{i=۱}^{n} a_i \lambda_i^n v_i.$$

اگر برای هر λ_i که $i > ۱$ داشته باشیم $|\lambda| < ۱$، آن‌گاه

$$\lim_{n\longrightarrow\infty} P^n X = a_۱ v_۱ = \pi,$$

که این مورد دقیقاً در مثال ما دیده شد!

قابل ذکر است که این قضیه تضمین نمی‌کند که هر ماتریسی که در شرایط این قضیه صدق کند لزوماً این ویژگی را دارد. در این قسمت نقایص احتمالی و راه‌حل آن را بررسی می‌کنیم.

نقایص احتمالی

- مقدار ویژه ۱ برای چندجمله‌ای مشخصه $det(\lambda I - A) = ۰$ یک ریشه مکرر باشد.
- ماتریس P علاوه بر ۱ مقدار ویژه دیگری مانند λ با قدرمطلق ۱ داشته باشد.

در این شرایط چه می‌کنیم؟

ماتریس انتقال مارکوف بدون این نقایص را یک ماتریس انتقال مارکوف منظم می‌نامیم.

تعریف ۳.۴. *یک ماتریس انتقال مارکوف منظم است هرگاه*

- *مقدار ویژه‌ی ۱، ریشه ساده معادله مشخصه $det(\lambda I - A) = ۰$ باشد.*
- *سایر مقادیر ویژه P به جز ۱ دارای قدرمطلق کمتر از ۱ باشند.*

خاطرنشان می‌کنیم که اغلب ماتریس‌های P منظم هستند. بنابراین اگر ماتریس انتقال مارکوف متناظر یک شبکه، منظم نبود با تغییرات جزئی به یک ماتریس منظم تغییر شکل داده می‌شود.

راه‌حل. ماتریس $N \times N$ به‌صورت $Q = (q_{ij})$ که برای هر i و j داشته باشیم $q_{ij} = \frac{۱}{N}$ را در نظر می‌گیریم. ماتریس شبکه P را برای یک $\beta \in [۰, ۱]$ با ماتریس زیر جایگزین می‌کنیم،

$$P_\beta = (۱ - \beta)P + \beta Q, \tag{۱}$$

(مقدار β استفاده شده برای گوگل $\beta = ۰/۱۵$ است). توجه داشته باشید که باز هم درایه‌های ماتریس P_β نامنفی بوده و مجموع درایه‌های هر ستون برابر با ۱ است. بنابراین هنوز یک ماتریس انتقال مارکوف است. قضیه زیر وجود یک مقدار کوچک β جهت رفع نقایص موجود را تضمین می‌کند.

قضیه ۴.۴. *برای هر ماتریس انتقال P می‌توان عدد مثبت به اندازه دلخواه کوچک β یافت که P_β منظم باشد. هرگاه π بردار ویژه متناظر ۱ برای ماتریس P_β بوده که به‌گونه‌ای نرمال شده است که مجموع درایه‌های آن ۱ باشد، آن‌گاه برای این ماتریس P_β و برای هر بردار غیرصفر X که $X^t = (p_۱, \ldots, p_N)$، $p_i \in [۰, ۱]$ و $\sum_{i=۱}^{N} p_i = ۱$ داریم،*

$$\lim_{n\longrightarrow\infty} (P_\beta)^n X = \pi. \tag{۲}$$

http://dx.doi.org/10.22108/msci.2018.110135.1270

۵. ارتباط با قضیه نقطه ثابت باناخ

در این قسمت به قضیه نقطه ثابت باناخ می‌پردازیم. قضیه بالا را می‌توان به عنوان کاربردی از آن در نظر گرفت. اگر قسمت قبلی را نخوانده‌اید می‌توانید از این بخش صرف‌نظر کنید.

قضیه ۱.۵. *فرض کنید P یك ماتریس انتقال مارکوف منظم باشد. مجموعه*

$$S=\{X\mid X^t=(p_1,\dots,p_N), p_i\in[0,1], \sum_{i=1}^{N}p_i=1\},$$

به همراه یك فاصله مناسب $d(X,Y)$ را در نظر می‌گیریم (این فاصله به P بستگی دارد). عملگر خطی $L:S\longrightarrow S$ را به صورت $L(X)=PX$ روی S تعریف می‌کنیم. عملگر L یك انقباض روی S است یعنی عدد $c\in[0,1[$ وجود دارد که برای هر $X,Y\in S$

$$d(L(X),L(Y))\le cd(X,Y).$$

در این صورت بردار یکتای $\pi\in S$ موجود است که $L(\pi)=\pi$.

علاوه بر این هرگاه برای $X_0\in S$ با یك روند استقرایی دنباله $\{X_n\}$ را به صورت $X_{n+1}=L(X_n)$ تعریف کنیم آن‌گاه $\lim_{n\longrightarrow\infty}X_n=\pi$.

تعریف فاصله d. تعریف فاصله d کمی پیچیده است و می‌توان از آن عبور کرد اما در اینجا جهت تکمیل مطلب و اطلاع خواننده از جزئیات ذکر می‌شود. ما مساله را به حالتی که P قطری شدنی باشد محدود می‌کنیم. فرض کنیم $B=\{v_1=\pi,v_2,\dots,v_N\}$ پایه‌ای از بردارهای ویژه باشد. بردارهای $X,Y\in S$ را می‌توان با پایه β به‌صورت

$$X=\sum_{i=1}^{N}a_iv_i,\quad Y=\sum_{i=1}^{N}b_iv_i,$$

که $a_i,b_i\in\mathbb{C}$ نمایش داد. سپس تعریف می‌کنیم

$$d(X,Y)=\sum_{i=1}^{n}|a_i-b_i|.$$

با این فاصله، S یك فضای متریك کامل است، یعنی هر دنباله کوشی در آن همگرا است.

این قضیه نه تنها وجود π را تضمین می‌کند بلکه روشی برای ساخت آن به‌صورت حد دنباله $\{X_n\}$ نیز ارائه می‌کند. قبلاً در مثال ذکر شده نمونه‌ای از این همگرایی را ملاحظه کردیم. در حقیقت ستون jام P^n بردار P^ne_j است که $e_j^t=(0,0,\dots,1,0,\dots,0)$ بردار jام پایه‌ی متعارف است. البته در آن مثال توانستیم با حل سیستم $(I-P)X=0$ بردار π را با ماتریس زیر بیابیم،

$$I-P=\begin{pmatrix}1&-\frac{1}{2}&-\frac{1}{3}&-1&0\\-1&1&-\frac{1}{3}&0&-\frac{1}{3}\\0&-\frac{1}{2}&1&0&-\frac{1}{3}\\0&0&0&1&-\frac{1}{3}\\0&0&-\frac{1}{3}&0&1\end{pmatrix}$$

http://dx.doi.org/10.22108/msci.2018.110135.1270

به این نتیجه رسیدیم که تمام جواب‌ها برای $s \in \mathbb{R}$ به‌صورت $(۴s, \frac{۱۶}{۳}s, ۳s, \frac{۱}{۴}s, s)^t$ هستند. بنابراین جوابی که مجموع مؤلفه‌های آن ۱ باشد π است که

$$\pi^t = \left(\frac{۱۲}{۴۱}, \frac{۱۶}{۴۱}, \frac{۹}{۴۱}, \frac{۱}{۴۱}, \frac{۳}{۴۱}\right).$$

۶. محاسبات عملی توزیع مانا

ما ایده‌ی ساده‌ای که در الگوریتم نهفته است را دریافتیم. با این حال یافتن توزیع مانای π، یعنی بردار ویژه متناظر با مقدار ویژه ۱ برای ماتریس P_β در (۱) در حالتی‌که ماتریس، میلیاردها سطر و ستون دارد به هیچ وجه کار کوچك و پیش پا افتاده‌ای نیست. هم مدت زمان محاسبات و هم فضای حافظه مورد نیاز چالش‌های جدی پیش‌روی ما قرار می‌دهند. روش رایج حذفی گاوس در این حالت کاربردی ندارد. هم به دلیل میزان محاسبات و هم به این دلیل که نیاز به تقسیم ضرائب کوچك داریم. الگوریتم کارآمدتری که می‌توان به کار برد، استفاده از ویژگی (۲) را میسر می‌کند. (مرجع [۲] را ببینید)

در اینجا ارتباط مساله با قضیه نقطه ثابت باناخ مشخص می‌شود، به این ترتیب که اثبات قضیه نقطه ثابت باناخ الگوریتمی برای ساخت این نقطه ثابت ارائه می‌کند.

در حقیقت کافی است با یك $X_۰$ به‌صورت زیر

$$X_۰^t = \left(\frac{۱}{N}, \frac{۱}{N}, \dots, \frac{۱}{N}\right),$$

شروع کرده و $\lim_{n\longrightarrow\infty}(P_\beta)^n X_۰$ را محاسبه کنیم. معمولاً $P^n X_۰$ برای یك n بین ۵۰ تا ۱۰۰ تقریب خوبی از π به ما می‌دهد. با استفاده از استقراء $X_{n+۱} = P_\beta X_n$ محاسبه می‌شود. حتی این محاسبات نیز طولانی هستند. در واقع ماتریس P_β در (۱) با توجه به ساختارش هیچ درایه صفری ندارد. از طرف دیگر اغلب درایه‌های ماتریس P صفر هستند. بنابراین برای بهره بردن از این شرایط باید محاسبات را تفکیك کرد، به عبارت دیگر

$$X_{n+۱} = (۱-\beta)PX_n + \beta Q X_n.$$

به دلیل شکل خاص Q به راحتی می‌توان تحقیق کرد که اگر X برداری باشد که مجموع درایه‌های آن ۱ است آن‌گاه $QX = X_۰$. درنتیجه کافی است در محاسبات از دنباله زیر استفاده کنیم،

$$X_{n+۱} = (۱-\beta)PX_n + \beta X_۰.$$

نتیجه‌گیری

ما بخش عمومی الگوریتم پیج رنك گوگل را ارائه نمودیم. هم اکنون شما می‌توانید با شبکه‌های ساده آن را مورد آزمایش قرار داده و با افزودن لینك‌های داخلی و خارجی به صورت بهینه، ترفندهایی بیابید که رتبه‌بندی صفحه شخصی خود را بهبود بخشید. برخی قسمت‌های پیشرفته‌تر و خصوصی این الگوریتم همچنان در حال توسعه هستند. بخشی از آن شامل جایگزین نمودن ماتریس خنثی Q در (۱) با ماتریس‌هایی است که سلیقه کاربران شبکه را منعکس می‌کند. همچنین بخش‌های دیگر تضمین می‌کنند که دستکاری الگوریتم رتبه‌بندی توسط افرادی که قصد افزایش رتبه صفحه خود را دارند تأثیر چندانی در رتبه‌بندی ندارد.

به طور کلی به چه نتیجه‌ای رسیدیم؟ یك ایده‌ی ساده و زیرکانه منجر به یك جهش و پیشرفت فوق‌العاده در کارآیی موتورهای جستجو و موجب پیدایش امپراتوری تجارت گشت. هر چند پیاده‌سازی آن در حقیقت با استفاده از محاسبات انجام می‌گردد اما ایده‌ی اولیه آن نیازمند ریاضیات مقدماتی مانند جبر خطی و نظریه احتمال است. این ابزارهای نسبتاً

http://dx.doi.org/10.22108/msci.2018.110135.1270

استاندارد ریاضی مانند قطری‌سازی ماتریس‌ها زمانی که در جایی استفاده می‌شوند که فراتر از زمینه‌ی عادی به کار بردن آنهاست، قدرت و نفوذ خود را به طور کامل به نمایش می‌گذارند. علاوه بر این ما در اینجا با بیان اینکه قضیه نقطه ثابت باناخ می‌تواند کاربردهایی فراتر از منشاء اصلی خود داشته باشد، ایده‌هایی را مطرح کردیم که جنبه‌ی ادغام شدن بخش‌های مختلف علم را به نمایش می‌گذارد.

مراجع

[1] M. Eisermann, Comment Google classe les pages webb, http://images.math.cnrs.fr/comment-Google-classe-les-pages.html, 2009.

[2] A. N. Langville and C. D. Meyer, A Survey of Eigenvector Methods for Web Information Retrieval, *SIAM Review*, 47 (2005) 135–161.

[3] C. Rousseau and Y. Saint-Aubin, *Mathematics and technology, SUMAT Series, Sptinger-Verlag, (A French version of the book exists, published in the same series.) 2008.*

مریم لطفی‌پور
فارس، فسا، دانشگاه فسا، گروه ریاضیات و کاربردها
lotfipour@fasau.ac.ir

مریم لطفی‌پور کارشناسی و کارشناسی ارشد خود را در دانشگاه شیراز به پایان رساند. سپس دوره دکتری خود را در دانشگاه اصفهان تحت راهنمایی آقایان دکتر زعفرانی و دکتر فخار آغاز نموده و در سال ۹۲ موفق به اخذ مدرك دکتری خود در رشته ریاضی گردید. وی هم‌اکنون به عنوان عضو هیأت علمی گروه ریاضیات و کاربردها در دانشگاه فسا فعالیت می‌نماید.

http://dx.doi.org/10.22108/msci.2018.110135.1270