پردیس علوم
دانشکده ریاضی، آمار و علوم کامپیوتر

پروژه دوره کارشناسی

گراف‌های خیلی بزرگ

ریحانه درفشی

استاد راهنما : دکتر مرتضی محمد نوری

بهمن ۹۹

چکیده

بسیاری از پدیده‌ها و ساختارها در دنیای واقعی توسط گراف‌ها و شبکه‌ها قابل شبیه‌سازی هستند. بعضی از این گراف‌ها آنقدر بزرگ اند که نگه‌داری و استخراج اطلاعات از ساختار کلی آن‌ها دشوار است، در بعضی موارد حتی گراف مورد نظر دائم در حال تغییر است یا به طور کامل تعریف نشده است یا در دسترس نیست. در این بحث ابتدا کلیتی درمورد این گراف‌ها ارائه داده سپس چند مدل حدی برای مدل کردن این گراف‌ها و پیشبینی رفتارهایشان معرفی شده است.

فهرست مطالب

فصل ۱

مقدمه‌ای بر گراف‌ها و مفاهیم اولیه

در این بخش به ارائه تعریف‌ها و مفاهیمی که در ادامه بحث به کار برده شده اند میپردازیم.

۱.۱ گراف‌ها

تعریف ۱.۱ (گراف). [۱] یک زوج مرتب از دو مجموعه V راس ها و G یال ها است مانند $G = (V, E)$ که در آن E زیرمجموعه ای از دوتایی های (غیرمرتب) راس‌ها است.

تعریف ۲.۱ (گراف ساده). [۲] گرافی است که یال چندگانه و طوقه ندارد.

تعریف ۳.۱ (گراف جهت‌دار). [۳] گرافی است که یال‌ها در آن دارای جهت هستند.

تعریف ۴.۱ (گراف وزن‌دار). [۴] گرافی است که به هر یک از یال‌ها یا به هریک از راس‌های آن عددی نسبت داده شده‌است که وزن آن یال یا راس می‌باشد.

[۱] graph
[۲] simple graph
[۳] directed graph
[۴] weighted graph

تعریف ۵.۱ (گراف همبند). [۵] گراف G را همبند می نامیم هرگاه بین هر دو رأسِ آن حداقل یک مسیر وجود داشته باشد

تعریف ۶.۱ (گراف کامل K_n). [۶] گرافی‌است که بین هر دو راس آن دقیقاً یک یال وجود داشته باشد. یک گراف کامل از مرتبه n را با K_n نشان میدهند .

تعریف ۷.۱ (زیرگراف القایی). [۷] زیرگرافی مانند H که اگر $v_1, v_2 \in V(H)$ و داشته باشیم $v_1v_2 \in E(G)$ انگاه $v_1v_2 \in E(H)$

تعریف ۸.۱ (همسایگی یک راس). [۸] همسایگی یک راس v در یک گراف G زیرگراف G است که توسط تمام رئوس مجاور V القا می شود .

تعریف ۹.۱ (درجه رأس). [۹] درجه یک راس به تعداد یال‌های متصل به آن راس گفته می‌شود.

۱.۱.۱ یک‌ریختی [۱۰] و هم‌ریختی [۱۱]

تابع همریختی f از گراف $G = (V(G), E(G))$ به گراف $H = (V(H), E(H))$ که به صورت زیر نوشته میشود

$$f : G \to H$$

[۵] connected graph
[۶] Complete graph
[۷] induced subgraph
[۸] neighbourhood of a vertex v
[۹] degree of a vertex
[۱۰] isomorphism
[۱۱] homomorphism

یک تابع از $V(G)$ به $V(H)$ است که دو سر هر یال از G را به دو سر یک یال از V میبرد. به عبارتی برای هر جفت $u,\ v \in V(G)$:

$$\{u, v\} \in E(G) \to \{f(u), f(v)\} \in E(H)$$

یک هم‌ریختی یک‌به‌یک از G به H وجود دارد اگر و تنها اگر G زیرگرافی از H باشد.

اگر هم‌ریختی $f : G \to H$ یک تابع یک‌به‌یک و پوشا باشد که معکوس آن نیز یک هم‌ریختی است، آنگاه f یک یک‌ریختی است.

۲.۱ مفاهیمی در احتمال

تعریف ۱۰.۱ (قانون اعداد بزرگ). [۱۲] احتمالاً معروفترین نتیجه در نظریهٔ احتمالات است که برای توصیف نتیجهٔ تکرار یک آزمایش به دفعات زیاد به کار می‌رود. بر طبق این قانون هر قدر تعداد دفعات تکرار آزمایش بیشتر شود، میانگین نتایج به امید ریاضی آن نزدیک‌تر می‌شود.

تعریف ۱۱.۱ (پیوستار). [۱۳] توالی پیوسته که در آن عناصر مجاور به طور محسوسی با یکدیگر متفاوت نیستند ، گرچه در فواصل زیاد کاملاً قابل تمایز هستند.

تعریف ۱۲.۱ (توزیع احتمال). [۱۴] تابع توزیع احتمال بیانگر احتمال هر یک از مقادیر متغیر تصادفی (در مورد متغیر گسسته) یا احتمال قرار گرفتن متغیر در یک بازه مشخص (در مورد متغیر تصادفی پیوسته) میباشد. توزیع تجمعی احتمال یک متغیر تصادفی تابعی است از دامنهٔ آن متغیر بر بازهٔ $[0, 1]$ به طوری که احتمال رخدادن پیشامدهای با مقدار عددی کمتر از آن را نمایش می‌دهد. و به صورت

[۱۲] Law of large numbers

[۱۳] Continuum

[۱۴] Probability distribution

دقیق به شکل زیر تعریف می‌شود:

$$F_X(x) = Pr[X \leq x]$$

تعریف ۱.۱۳ (توزیع یکنواخت گسسته). [۱۵] توزیع یکنواخت گسسته یک توزیع احتمال گسسته است که احتمال مشاهده تعداد محدودی پیشامد را یکسان گزارش می دهد. احتمال هر کدام از n شامد قابل مشاهده ای که از این توزیع پیروی می کنند، برابر با $\frac{1}{n}$ است.

تعریف ۱.۱۴ (توزیع پوآسون). [۱۶] یک توزیع احتمالی گسسته است که احتمال اینکه یک حادثه به تعداد مشخصی در فاصلهٔ زمانی یا مکانی ثابتی رخ دهد را شرح می‌دهد گفته می شود که یک متغیر تصادفی گسسته X دارای توزیع پواسون با پارامتر $\lambda \geq 0$ است اگر برای $k = 0, 1, 2, \ldots$ تابع توزیع احتمال X با استفاده ازرابطه زیر به دست آید:

$$f(k; \lambda) = Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

۱.۳ مفاهیمی در جبر

تعریف ۱.۱۵ (جبر سیگما). [۱۷] جبر سیگما بر روی مجموعه X ، به مجموعه‌ای از زیرمجموعه‌های X گفته می‌شود که تحت انجام تعداد شمارایی از جبر مجموعه‌ای (مانند اجتماع، اشتراک یا متمم)، بسته بماند. یعنی تعداد شمارایی از انجام این گونه جبرها بر روی اعضای جبر سیگما، باز همواره عضوی از آن خواهد بود.

تعریف ۱.۱۶ (اندازه). [۱۸] فرض کنید X یک مجموعه و σ یک جبر سیگما روی X باشد. تابعی چون μ از Σ به خط اعداد حقیقی توسعه یافته را اندازه گویند اگر در خواص زیر صدق کند:

[۱۵] Discrete uniform distribution

[۱۶] Poisson distribution

[۱۷] $\sigma - algebra$

[۱۸] measure

۱. نامنفی بودن: برای تمام E ها در Σ ، داریم $0 \leq \mu(E)$.

۲. مجموعه تهی: $\mu(\emptyset) = 0$.

۳. جمع‌پذیر شمارا: برای تمام گردایه های شمارایی چون $\{E_i\}_{i=1}^{\infty}$ از مجموعه های مجزای درون Σ داریم:

$$\mu(\bigcup_{k=1}^{\infty} E_k) = \sum_{k=1}^{\infty} \mu(E_k)$$

تعریف ۱.۱۷ (نگاشت حافظ اندازه). [۱۹] اگر $(X, \mathcal{A}, \mu)$ یک فضای اندازه باشد میگوییم نگاشت قابل اندازه‌گیری $T : X \to X$ یک نگاشت حافظ اندازه است و μ زیر T ثابت است اگر برای هر $A \in \mathcal{A}$ داشته باشیم:

$$\mu(A) = \mu(T^{-1}(A))$$

رفتار پویای نگاشت‌های حافظ اندازه موضوع نظریه ارگودیک است.

[۱۹] Measure-Preserving Map

فصل ۲

شبکه‌های عظیم و نحوه مطالعه آن‌ها

۱.۲ شبکه‌های عظیم [۱]

در دهه گذشته مشخص شد که تعداد زیادی از جالب‌ترین ساختارها و پدیده‌های جهان را می‌توان توسط شبکه‌ها توصیف کرد: عناصر قابل جدا شدن ، با اتصالات (یا فعل و انفعالات) بین جفتهای خاصی از آنها.

- در میان اینگونه از شبکه‌ها، شناخته‌شده‌ترین آن‌ها اینترنت است.همچنین اینترنت باعث به جود آمدن بسیاری از شبکه‌های دیگر میشود، مانند شبکه های اجتماعی مبتنی بر اینترنت، پایگاه داده های توزیع شده و غیره.
- شبکه های اجتماعی ابزاری اساسی در بسیاری از مطالعات در زمینه جامعه شناسی ، تاریخ ، اپیدمیولوژی و اقتصاد است.بزرگ ترین شبکه اجتماعی گراف آشنایی همه افراد کره زمین با حدود هفت میلیارد گره است.

[۱] Huge networks

- میتوان گفت کل جهان یک شبکه بسیار عظیم و احتمالا بینهایت است که گره های آن رویدادها (فعل و انفعالات بین ذرات بنیادی) و یال‌های آن خود گره‌ها هستند. این شبکه احتمالا حدود 10^{80} گره دارد.

در مسائل نظریه گراف سنتی ، کل گراف دقیقاً داده شده است ، و ما به دنبال روابط بین پارامترهای آن یا الگوریتم های کارآمد برای محاسبه پارامترهای آن هستیم. از طرف دیگر ، شبکه های بسیار بزرگ (مانند اینترنت) هرگز کاملاً شناخته شده نیستند، در بیشتر موارد حتی به خوبی تعریف نشده اند. اطلاعات مربوط به آنها فقط با استفاده از روشهای غیر مستقیم مانند نمونه برداری محلی تصادفی یا با نظارت بر رفتارهای مختلف جهانی فرایندها قابل جمع آوری است. شبکه های متراکم [۲] (که در آنها یک گره با درصد مثبتی از گره های دیگر مجاور است) و شبکه های پراکنده [۳] (که در آن یک گره تعداد محدودی همسایه دارد) رفتار بسیار متنوعی از خود نشان می دهند. از نظر عملی ، شبکه های پراکنده از اهمیت بیشتری برخوردار هستند ، اما در حال حاضر از نتایج نظری کامل تری برای شبکه های متراکم برخورداریم.

۲.۲ نمونه‌برداری محلی [۴]

در مورد گراف های متراکم G نمونه برداری بسیار ساده است : به تعداد k تا گره تصادفی انتخاب میکنیم، و یال های بین آن‌ها را مشخص میکنیم تا یک زیرگراف القایی تصادفی به دست آوریم. این زیرگراف در صورتی که k به اندازه کافی بزرگ باشد اطلاعات زیادی در مورد گراف موردنظر به ما میدهد. در مورد گرافهای پراکنده با درجه رئوس محدود اما این روش یک نتیجه بدیهی دارد : زیرگراف نمونه‌برداری‌شده به احتمال بالایی فاقد یال است.احتمالا بهنرین روش برای حل کردن این

[۲] dense networks
[۳] sparse networks
[۴] local sampling

موضوع استفاده از نمونه‌برداری از همسایگی است .

به این صورت که اگر G یک گراف با درجه راس محدود d باشد، یک راس تصادفی از آن انتخاب کرده و تا عمق m از در گراف پیش میرویم و نمونه برداری میکنیم. البته این روش نسبت به روش قبل کمتر مطلوب است و ویژگی هایی مانند پراکنش گراف را نشان نمیدهد. روش دیگری که وجود دارد انتخاب دو یا چند گره تصادفی و تعیین مقادیر ساده مرتبط با آنها ، مانند فاصله های دوتایی ، حداکثر جریان ، مقاومت الکتریکی ، تعداد دیده شدن در گشت های تصادفی است.

به جای این ، ما می توانیم تعداد همریختی ها (یا همریختی های یک به یک) از گراف‌های کوچک به گراف را حساب کنیم. اغلب تعداد همریختی ها از نظر جبری رفتار بهتری دارند و همچنین این مزیت را دارند که رویکردهای "دوگانه" و متفاوت با معکوس کردن پیکان ها در همریختی های گراف را پیشنهاد می کنند.

۳.۲ گراف‌های تصادفی [۵]

ساده‌ترین مدل گراف‌های تصادفی توسط Erdos و Renyi و Gilbert در سال ۱۹۵۹ معرفی شد. عدد صحیح مثبت n و عدد حقیقی $0 \leq p \leq 1$ داده شده است، گراف تصادفی $G(n,p)$ را با گرفتن n گره، به عنوان مثال $[n] = \{1,\ldots,n\}$ و وصل کردن هر دوتای آن ها با احتمال p و گرفتن این تصمیم به طور مستقل برای هر جفت میسازیم.

مدل‌های جایگزین معادل نیز وجود دارند: میتوانیم تعداد یال‌ها را عدد ثابت m فرض کنیم، و یک زیرمجموعه m تایی از جفت های روی $[n]$ به صورت تصادفی انتخاب کنیم. گراف تصادفی $G(n,m)$ بسیار شبیه به $G(n,p)$ وقتی $m = p\binom{n}{2}$ است، میشود.

مدل دیگری که به برخی از تحولات اخیر نزدیکتر است ، گراف های تصادفی در حال تکامل

[۵] random graphs

[۶] است ، به طوری که یال ها یکی یکی اضافه می شوند و همیشه از مجموعه جفت های غیرمتصل به طور یکنواخت انتخاب می شوند و پس از m مرحله متوقف میشود و گراف تصادفی $G(n,m)$ به دست می‌آید.

گراف‌های تصادفی Erdos − Renyi دارای بسیاری از ویژگی های جالب ، اغلب غافلگیر کننده و ادبیات عظیم هستند. یک نظر مرسوم در مورد گراف های تصادفی با کران درجه معین این است که همه آنها یکسان هستند. پارامترهای اساسی آنها ، مانند عدد رنگی ، بزرگترین خوشه، تراکم مثلث و غیره بسیار نزدیک هم هستند. این واقعیت انگیزه مهمی در تعریف معیار درست تشابه جهانی گراف ها خواهد بود.

۴.۲ گراف‌هایی که به صورت تصادفی در حال رشد هستند [۷]

مدل های گراف‌های تصادفی در یک مجموعه ثابت از گره ها ، که در بالا بحث شد ، قادر به تولید خصوصیات مهم شبکه های واقعی نیستند. به عنوان مثال ، درجه گراف‌های تصادفی Erdos − Renyi از توزیع دو جمله ای پیروی می کنند ، بنابراین در صورت ثابت بودن احتمال وجود یال p به صورت مجانبی نرمال هستند و اگر درجه راس ها ثابت باشد به صورت مجانبی پوآسون هستند (به طور مثال $p = p(n) \sim c/n$) .

در هر صورت ، درجه ها تقریباً در حدود میانگین متمرکز هستند ، در حالی که درجه های شبکه های زندگی واقعی تمایل به پیروی از قانون زیپف [۸] دارند ، این بدان معنی است که توزیع به صورت نمایی کاهش می یابد.

در سال ۲۰۰۲ Albert و Barabasi یک مدل شبکه تصادفی ایجاد کردند که طبق قوانین

---

[۶] evolving random graphs

[۷] randomly growing graphs

[۸] zipf law

طبیعی رشد می کند و می تواند رفتار آن را تولید کند. از آن زمان تغییرات زیادی در شبکه های در حال رشد ایجاد شد. روند تولید گراف معمولاً شامل مراحل تصادفی و پیروی از برخی قوانین محلی است. این شاید اولین نکته ای باشد که یکی از ابزارهای اصلی ما را تعیین می کند ، یعنی تعیین حد برای توالی گراف‌ها. دقیقاً همانطور که قانون اعداد بزرگ [۹] به ما می گوید با اضافه کردن متغیرهای تصادفی بیشتر و مستقل تری ، یک نتیجه رفتاری قطعی‌تری به دست می آوریم ، این توالی های گرافی در حال رشد ، مستقل از انتخاب های تصادفی که در این راه انجام می شود ، دارای ساختاری کاملاً مشخص هستند. در حد ، تصادف از بین می رود ، و رفتار مجانبی دنباله را می توان با یک شئ حدی کاملاً مشخص توصیف کرد.

[۹] law of large numbers

فصل ۳

پارامترهای گراف

یک پارامتر گراف یک تابع با مقادیر حقیقی است که روی یک‌ریختی‌های گراف تعریف میشود. یعنی برای گراف‌های یک‌ریخت مقدار یکسانی دارد. (که شامل گراف k_0 که فاقد یال و راس است نیزمیشود). پارامتر گراف f ضرب‌پذیر [۱] است اگر $f(G) = f(G_1)f(G_2)$ به طوری که G اجتماع جدا از هم [۲] G_1 و G_2 باشد. میگوییم یک پارامتر گراف نرمال‌سازی شده است اگر مقدار آن روی K_1 ، 1 باشد.

۱.۳ هم‌ریختی‌های از سمت چپ [۳]

به جای آزمایش ، معمولاً راحت تر است که در مورد هم‌ریختی های بین گراف‌ها (نگاشت با حفظ مجاورت) [۴] صحبت کنیم. اگر یک گراف (بزرگ) G به ما داده شود، ما میتوانیم ساختارهای محلی

[۱] multiplicative
[۲] disjoint union
[۳] Homomorphisms from the left
[۴] adjacency-preserving maps

آن را با شمردن هم‌ریختی‌ها از گراف‌های مختلف (کوچک) F به G مطالعه کنیم. همچنین میتوانیم ساختار کلی آن را با استفاده از شمردن هم‌ریختی‌های آن به گراف‌های مختلف کوچک مطالعه کنیم. نوع اول اطلاعات به دست آمده بسیار نزدیک(در بعضی مواقع برابر با) اطلاعات به دست آمده از نمونه برداری است. در حالی که نوع دوم اطلاعات به فیزیک آماری مرتبط است. (که در این بحث من وارد آن نشده ام)

۱.۱.۳ تعداد هم‌ریختی‌ها

برای دو گراف متناهی F و G ، $hom(F,G)$ تعداد هم‌ریختی‌های موجود از F به G ، $inj(F,G)$ تعداد هم‌ریختی‌های یک‌به‌یک، و $ind(F,G)$ تعداد دفعاتی که F به عنوان زیرگراف در G ظاهر شده است. [۵] را نشان میدهد.

این مقادیر به هم مرتبط هستند:

$$inj(F,G) = \sum_{F' \supseteq F} ind(F',G)$$

به طوری که F' گراف‌هایی است که از اضافه کردن راس به F تولید میشود .

$$hom(F,G) = \sum_{F''} inj(F'',G)$$

به طوری که F'' شامل تمام گراف هایی است که با نامگذاری رئوس F به دست می‌آیند.

[۵] the number of embedding of F into G

۳.۱.۲ چگالی هم‌ریختی[۶] در گراف‌های متراکم

این مقادیر را نرمال‌سازی کرده و چگالی‌های هم‌ریختی را به دست می‌آوریم:

$$t(F,G) = \frac{hom(F,G)}{\mid V(G) \mid^{|V(G)|}}$$

که احتمال هم‌ریختی بودن یک نگاشت تصادفی از F به G را نشان میدهد.

همچنین:

$$t_{inj}(F,G) = \frac{inj(F,G)}{n(n-1)\ldots(n-k+1)}$$

و

$$t_{ind}(F,G) = \frac{ind(F,G)}{n(n-1)\ldots(n-k+1)}$$

۳.۱.۳ چگالی هم‌ریختی در گراف های پراکنده

$$s(F,G) = \frac{hom(F,G)}{\mid V(G) \mid}$$

که برای گراف‌های هم‌بند درنظر میگیریم. این فرمول را به صورت مقابل میتوانیم توجیه کنیم. برای هر $u \in V(F)$ و $v \in V(G)$ ، فرض کنید $hom_{v \to u}$ تعداد هم‌ریختی‌های φ از F به G را نشان دهد، به‌طوری که $\varphi(u) = v$. حال یک راس u از F را ثابت در نظر گرفته و راس تصادفی v از G را انتخاب میکنیم. در این صورت مقدار $s(F,G)$ برابر با $hom_{v \to u}(F,G)$ می‌شود. به طور مشابه :

$$s_{inj}(F,G) = \frac{inj(F,G)}{\mid V(G) \mid}$$

[۶] homomorphism density

و

$$s_{ind}(F,G) = \frac{ind(F,G)}{\mid V(G) \mid}$$

فصل ۴

ساختارهای گراف‌مانند روی فضای احتمال

هدف در ادامه این بحث معرفی چند مدل تحلیلی است که به عنوان شئ حدی در گراف‌ها و به طور جدا برای موارد پراکنده و متراکم نیز قابل استفاده هستند. در مورد گراف‌های متراکم چندین تا از این مدل ها به صورت قابل اثبات معادل هستند. درمورد گراف‌های پراکنده چند نمونه مرتبط اما غیر معادل نیز معرفی شده است . در این فصل مقدمات و توضیحاتی درمورد این مدل ها و در فصل‌های بعد توضیحات مفصل‌تری آورده شده است.

۱.۴ تقریب با بی نهایت: همگرایی و حدود

این ایده می تواند ناشی از نگاه ما به یک قطعه بزرگ فلز باشد. این یک کریستال است ، یعنی گراف واقعاً بزرگی متشکل از اتم ها و پیوندهای بین آنها. اما از خیلی نظرها (به عنوان مثال ، استفاده از فلز در ساختن پل) ، مفیدتر است که آن را به عنوان یک پیوستار [۱] با چند پارامتر مهم (چگالی ، کشش و غیره) در نظر بگیریم که رفتار آن توسط معادلات دیفرانسیل اداره می شود. آیا می توان یک گراف

[۱] Continuum

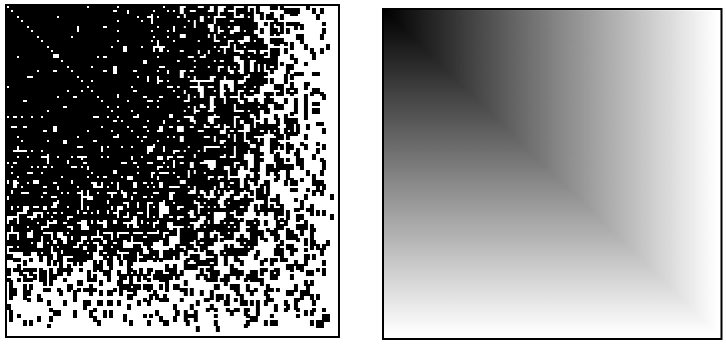
بسیار بزرگ و کلی را به عنوان نوعی پیوستار در نظر گرفت؟

یکی از راه های دقیق سازی این شهود ، در نظر گرفتن یک توالی در حال رشد G_n از گراف‌ها است که تعداد گره های آنها به بی نهایت متمایل است ، و تعریف اینکه چه زمانی چنین توالی ای همگرا است. بحث ما در مورد نمونه برداری یک اصل کلی را ارائه می دهد که منجر به این تعریف می شود: ما نمونه هایی از اندازه ثابت k را از G_n و توزیع آنها در نظر می گیریم .

می گوییم در صورتی که $n \to \infty$ و توزیع برای هر k ثابتی همگرا شود ، توالی به صورت محلی همگراست (با توجه به روش نمونه گیری داده شده). خانواده توزیع های حد دار (برای هر k) می تواند به عنوان یک شئ حدی از دنباله در نظر گرفته شود. تعریف فوق نشان دهنده حد یک توالی گراف به عنوان مجموعه ای از توزیع های احتمال روی گراف‌ها ، یک عدد برای هر اندازه نمونه است. این همیشه نمایش مفیدی از شی حدی نیست و توصیف صریح تر آن مطلوب است. گام بعدی نمایش خانواده توزیع ها روی گراف‌های متناهی (نمونه ها) توسط یک توزیع احتمال واحد بر روی گراف‌های قابل شمارش است. همچنین می توان توصیف صریح تری از این اشیا حدی را بیان کرد.

با مورد متراکم شروع میکنیم. در اینجا شی حدی را می توان به عنوان یک تابع قابل اندازه گیری دو متغیره $W : [0,1]^2 \to [0,1]$ که گرافون نامیده میشود، توصیف کرد . این اشیا حدی را می توان به عنوان گراف‌های وزن‌دار با یک مجموعه زیرین پیوستار در نظر گرفت ، یا (در صورت تمایل) به عنوان گراف با یک مدل نمایش غیر استاندارد از فاصله واحد.

اجازه دهید مثالی را در اینجا شرح دهیم. تصویر سمت چپ شکل ۱.۴ ماتریس مجاورت گراف G با 100 گره است ، 1 ها با مربع های سیاه و 0 ها با مربع های سفید نشان داده شده اند. . گراف خود توسط یک قاعده رشد تصادفی ساده ساخته شده است: با شروع یک گره ، ما یک گره جدید یا یک یال جدید اضافه می کنیم. اگر n تعداد فعلی گره ها باشد ،گره جدیدی با احتمال $\frac{1}{n}$ متولد می شود .



شکل ۱.۴: یک گراف به طور یکنواخت و تصادفی رشد یافته با ۱۰۰ گره

تصویر سمت راست تصویری از عملکرد تابع $U(x,y) = 1 - max(x,y)$ است.

شباهت با تصویر سمت چپ مشهود است. و نشان می دهد که حد توالی گراف در سمت چپ برابر با این تابع است. به نظر می رسد این به معنای کاملاً مشخصی باشد. از این رو می توان برای محاسبه تقریبی پارامترهای مختلف گراف در سمت چپ ، پارامترهای مربوط به تابع نشان داده شده در سمت راست را محاسبه کرد.

به عنوان مثال ، چگالی تعدا مثلث ها درگراف سمت چپ ($n \to \infty$) به انتگرال زیر میل میکند:

$$\int_{[0,1]^3} U(x,y)U(y,z)U(z,x)\ dx\ dy\ dz.$$

دو نکته دیگر در مورد پرونده متراکم وجود دارد. البته ، یک نمودار می تواند بی نهایت پیچیده باشد، اما در بسیاری از موارد حد توالی های گراف‌های در حال رشد معادل یک گرافون حدی است که برابر با یک تابع پیوسته که با یک فرمول ساده نمایش داده میشود است.

به جای فاصله $[0,1]$ هر فضای احتمالی $(\Omega, \pi, \mathcal{A})$ به همراه یک تابع متقارن قابل اندازه‌گیری $W : \omega \times \omega \to [0,1]$ استفاده کرد. این امر کلیت بیشتری به ما نمی دهد ، اما گاهی اوقات مفید است که شئ حدی را با سایر فضاهای احتمال ارائه دهیم.

در حالت پراکنده شئ حدی را میتوان به عنوان گرافینگ (در نظریه گروه ها معرفی شده است) یا به عنوان یک گراف حافظ اندازه، یا به عنوان یک توزیع روی گراف‌های شمارای ریشه‌دار [۲] با ویژگی‌های خاص توصیف کرد.

به جای نمونه گیری ، می توان از اندازه گیری های دوگانه (جهانی) ، به طور دقیق تر ، هم‌ریختی به گراف‌های کوچک ثابت ، برای تعریف همگرایی استفاده کرد. واقعیت قابل توجه این است که در شرایط مناسب ، این منجر به یک مفهوم معادل می شود.

---

[۲] countable rooted graphs

فصل ۵

گرافون‌ها [1]

فرض کنید $\mathcal{W}$ فضای تمام توابع قابل اندازه‌گیری متقارن کران‌دار $W : [0,1]^2 \to \Re$ باشد . فرض کنید $\mathcal{W}_0$ مجموعه تمام توابع $W \in \mathcal{W}$ باشد به طوری که $0 \leq W \leq 1$ باشد. تابع $W \in \mathcal{W}$ یک تابع پله ای است اگر یک افراز به شکل $S_1 \cup S_2 \cup \cdots \cup S_k$ از $[0,1]$ به شکل مجموعه‌های قابل اندازه‌گیری وجود داشته باشد به طوری که مقدار W روی هر $S_i \times S_j$ ثابت باشد .

برای هر گراف وزن‌دار G تابع پله‌ای $W_G \in \mathcal{W}_0$ را به اینگونه تعریف میکنیم :

فرض کنید $V(G) = [n]$.

بازه $[0,1]$ را به n بازه $J_1, \ldots, J_2$ با طول $\lambda(J_i) = \alpha_i \backslash \alpha_G$ تقسیم کنید. .

برای هر $x \in J_i$ و $y \in J_j$ فرض کنید :

$$W_G(x, y) = \beta_{i,j}(G)$$

فرض کنید $W \in \mathcal{W}$ و فرض کنید $\varphi : [0,1] \to [0,1]$ یک نگاشت حافظ اندازه باشد. تابع

---

[1] Graphons

W^φ را میتوانیم به صورت زیر معرفی کنیم:

$$W^\varphi(x,y) = W(\varphi(x), \varphi(y))$$

اگر بخواهیم صرفا به چشم یک آنالوگ پیوسته از گراف‌ها به این توابع نگاه کنیم، توابع W و W^φ فرق اساسی ای ندارند. (تفاوت آن‌ها همانند دو گراف یک‌ریخت است که راس‌های آن‌ها به طور متفاوتی برچسب گذاری شده اند.) البته باید دقت داشت که نگاشت‌های حافظ اندازه لزوما برگشت‌پذیر نیستند.بنابراین رابطه بین این دو متقارن نیست. دو گرافون W و W' را یک‌ریخت ضعیف مینامیم اگر گرافون سومی مانند U و نگاشت های حافظ اندازه ای مانند $\varphi, \varphi' : [0,1] \to [0,1]$ وجود داشته باشند به طوری که در همه جا $W = U^\varphi$ و $W' = U^{\varphi'}$. اثبات اینکه رابطه یک‌ریختی ضعیف یک رابطه برابری است سخت نیست.

به کلاس‌های هم‌ارزی توابع عضو $\mathcal{W}_0$ زیر رابطه هم‌ارزی ضعیف گرافون میگوییم. (البته گاهی به تابع $W \in \mathcal{W}_0$ گرافون میگوییم، این توابع با توجه به شباهتشان به گراف‌ها، گرافون برچسبدار نیز صدا زده میشوند.)

۱.۵ تعداد هم‌ریختی‌ها به گرافون‌ها و از گرافون‌ها

شمردن تعداد یک‌ریختی‌ها به گراف‌ها به شمردن تعداد یک‌ریختی‌ها به گرافون‌ها با استدلال زیر تعمیم پیدا میکند:

برای هر $W \in \mathcal{W}$ و گراف ساده $F = (V,E)$ تعریف میکنیم :

$$t(F,W) = \int_{[0,1]^V} \prod_{ij \in E} W(x_i, x_j) \prod_{i \in V} dx_i$$

حال برای هر گراف ساده G ساده است که نشان دهیم:

$$t(F,G) = t(F, W_G)$$

قضیه ۱.۵. دو گرافون هم‌ریخت ضعیف هستند اگر و تنها اگر برای هر گراف ساده F :

$$t(F,W) = t(F,W')$$

۲.۵ گراف‌های تصادفی W[۲]

گرافون W باعث ایجاد روشی برای ایجاد گراف‌های تصادفی می شود که عمومیت بیشتری نسبت به گراف‌های Erdos – Renyi دارند. این روش توسط Lovasz و Szegedy و Bollobas ,Janson ,Riordan معرفی شده است.

گرافون W و عدد طبیعی $n \leq 0$ داده شده است: گراف تصادفی $G(n,W)$ را میتوانیم روی مجموعه رئوس $[n]$ به صورت زیر تولید کنیم: n تا عدد تصادفی $X_1,\ldots,X_n$ با توزیع یکنواخت روی بازه ی $[0,1]$ تولید میکنیم. سپس راس‌های i و j را با احتمال $W(X_i,X_j)$ به هم وصل میکنیم و برای هر زوج (i,j) این تصمیم را مستقلا میگیریم.

در موارد خاص اگر تابع W عینا یک تابع p باشد، ما یک گراف رندم معمولی $G(n,p)$ به دست می‌آوریم.

این روش را میتوانیم برای ساختن گراف‌های تصادفی شمارای $G(W)$ روی $\mathbb{N}$ تعمیم دهیم:

دنباله بینهایت تصادفی را $X_1,X_2,\ldots$ از یک توزیع یکنواخت تصادفی روی بازه $[0,1]$ تولید میکنیم و هر دو راس i و j را با احتمال $W[X_i,X_j]$ به هم وصل میکنیم. سپس یک یال تصادفی از ریشه آن انتخاب میکنیم .

[۲] W-random graphs

فصل ۶

گرافینگ‌ها [۱]

۱.۶ گراف‌های حافظ اندازه [۲]

فرض کنید G یک گراف با مجموعه رئوس $[0,1]$ با کران درجه d باشد. G قابل اندازه گیری [۳] است اگر برای هر مجموعه قابل اندازه‌گیری B همسایگی $N(B)$ در G نیز قابل اندازه‌گیری باشد.

برای هر مجموعه $A \subseteq [0,1]$ و هر $x \in [0,1]$ ، فرض کنید $d_A(x)$ تعداد همسایه های x در A را نشان دهد.با استفاده از قابل اندازه گیری بودن G میتواند نشان داد $d_A(x)$ یک تابع قابل اندازه‌گیری از x است.

میگوییم G حافظ اندازه است اگر و فقط اگر قابل اندازه‌گیری باشد و برای هر دو مجموعه قابل اندازه‌گیری A و B :

$$\int_A d_B(x)d_x = \int_B d_A(x)d_x \tag{۱.۶}$$

[۱] Graphings
[۲] measure-preserving graphs
[۳] measurable

۲.۶ گرافینگ‌ها

فرض کنید $A_1, \ldots, A_d, B_1, \ldots, B_d$ زیرمجموعه‌های قابل اندازه‌گیری از بازه $[0,1]$ باشند، و فرض کنید $\varphi_i : A_i \to B_i$ یک نگاشت دوطرفه حافظ اندازه باشد. $H = ([0,1], \varphi_1, \ldots, \varphi_d)$ یک گرافینگ نامیده میشود.

از هر گرافینگ H یک گراف جهت‌دار $\overrightarrow{G}$ روی بازه $[0,1]$ با وصل کردن x و y در بازه $[0,1]$ اگر i ای وجود داشته باشد به طوری که $y = \varphi_i(x)$ به دست می‌آید.

رئوس این گراف جهت‌دار با d رنگ رنگ‌آمیزی میشود و هر کلاس رنگی نشان‌دهنده یک تابع یک‌به‌یک و پوشا و حافظ اندازه روی دو زیرمجموعه از $[0,1]$ است.

اگر رنگ‌ها و جهت‌ها در این گراف را کنار بگذاریم یک گراف حافظ اندازه با کران درجه رئوس $2d$ به دست می‌آوریم.

یک گراف حافظ اندازه که رئوسش رنگ آمیزی شده و جهت دار است و هر رنگ یک نشان‌دهنده یک نگاشت یک‌به‌یک و پوشا و حافظ اندازه است، معادل با یک گرافینگ است.

احتمالا طبیعی تر است که فرض کنیم نگاشت های $\varphi_1, \ldots, \varphi_n$ برگشت [۴] هستند. که برای هر کدام میتوانیم یک گراف بدون جهت به دست آوریم، و میتوانیم آن هارا به برگشت‌های حافظ اندازه $[0,1] \to [0,1]$ گسترش دهیم.

این یک گزاره صحیح است که برای هر گرافینگ یک گرافینگ پیچشی [۵] که گراف حافظ اندازه یکسانی را توصیف میکند وجود دارد، ولی تعداد نگاشت ها ممکن است بسیار بیشتر باشد .

در آخر،هر گراف حافظ اندازه از یک گرافینگ به دست می آید .

قضیه ۱.۶. فرض کنید G یک گراف حافظ اندازه با کران درجه d باشد. در این صورت گرافینگ $H = ([0,1], \varphi_1, \ldots, \varphi_r)$ وجود دارد به طوری که $r \leq d^2$ که گراف معادل آن G است.

[۴] involutions
[۵] involutive graphing

یک راه دیگر نگاه کردن به نمایش گراف های حافظ اندازه با عنوان گرافینگ این است که یک مدرک برای حافظ اندازه بودن گراف فراهم میشود. گرافینگ معادل با یک گراف حافظ اندازه ممکن است یکتا نباشد.

قابل ذکر است که بازه $[0,1]$ با هر بازه دیگری قابل جایگزینی است، که البته به کلیت موضوع چیزی اضافه نمیکند.

۳.۶ گراف‌های ریشه‌دار شمارای تصادفی [۶]

فرض کنید G یک گراف حافظ اندازه باشد. نقطه تصادفی و یکنواخت $x \in [0,1]$ را انتخاب کنید. مولفه همبندی G_x از G که شامل x است یک گراف شمارا با کران درجه d و ریشه x است.

فرض کنید G_d مجموعه تمام گراف های همبند شمارا با کران درجه d و یک راس به عنوان ریشه است. فرض کنید A_d یک جبر σ [۷] روی G_d باشد. که توسط زیرمجموعه هایی که با ثابت کردن همسایگی متناهی ریشه به دست می آیند، به دست آمده باشد. نگاشت $x \to G_x$ به اندازه نگاشت $(G_d, A_d) \to [0,1]$ قابل اندازه‌گیری است، بنابراین هر گراف حافظ اندازه G یک توزیع احتمال π روی (G_d, A_d) را توصیف میکند.

وضعیت ۱.۶ ویژگی مقابل را برای اندازه π به دست میدهد. یک گراف ریشه دار از π انتخاب میکنیم سپس یک یال تصادفی از ریشه آن انتخاب میکنیم. در این صورت یک توزیع احتمال π^* روی مجموعه G'_d از گراف های ریشه دار در G_d با یک یال از ریشه به دست می آوریم.

میگوییم π تک‌ماژوله است اگر نگاشت $G'_d \to G'_d$ که با استفاده از جابه‌جا کردن ریشه به راس مجاور ریشه در یال مورد نظر ایجاد میشود، با توجه به π حافظ اندازه باشد.

اندازه ای که روی G_d با استفاده از گراف حافظ اندازه به دست آمد تک ماژوله است. به طور

[۶] Random countable rooted graphs

[۷] σ-algebra

برعکس هر اندازه ای به این شکل از یک گرافیگ به دست می آید.

کتاب‌نامه

[۱] Laszlo Lovasz.، "Large networks and graph limits". ۲۰۱۰

[۲] Billingsley، Patrick .(۲۰۱۲) Probability and Measure (Anniversary ed.). Wiley. ISBN ۲–۱۲۲۳۷–۱۱۸–۱–۹۷۸

[۳] Bender، Edward A.؛ Williamson، S. Gill .(۲۰۱۰) Lists، Decisions and Graphs. With an Introduction to Probability

Abstract

In the last decade it became apparent that a large number of the most interesting structures and phenomena of the world can be described by networks: separable elements, with connections (or interactions) between certain pairs of them. These huge networks pose exciting challenges for the mathematician. Graph Theory (the mathematical theory of networks) has been one of the fastest developing areas of mathematics in the last decades; with the appearance of the Internet, however, it faces fairly novel, unconventional problems. In traditional graph theoretical problems the whole graph is exactly given, and we are looking for relationships between its parameters or efficient algorithms for computing its parameters. On the other hand, very large networks (like the Internet) are never completely known, in most cases they are not even well defined. Data about them can be collected only by indirect means like random local sampling or by monitoring the behavior of various global processes.

Very Large Graphs

Reyhaneh Derafshi

Under supervision of :

prof. Morteza Mohammad Nouri

February 2021