محاسبات مجانبی*

عباس محرابیان

دانشگاه بریتیش کلمبیا

AbbasMehrabian@gmail.com

۷ تیر ۱۳۹۵

چکیده

نمادهای مجانبی مثل O, o, Ω در ریاضیات و علوم کامپیوتر بسیار استفاده می‌شوند و آشنایی با آن‌ها برای دانشجویان و پژوهش‌گران این رشته‌ها ضروری است. در این مقاله، که عمدتاً ترجمه‌ای است از [۵، ۴]، این نمادها را تعریف می‌کنیم و چندین مثال می‌زنیم. هدف اینست که خواننده پس از مطالعهٔ مثال‌ها و حل تمرین‌ها بتواند به راحتی با این نمادها کار کند.

۱ مقدمات.

همانند فیزیک، در نظریهٔ محاسبه[۱] هم فهمیدن رشد توابع خیلی مهم است، به خصوص این که می‌خواهیم بدانیم زمان لازم برای حل یک مسئله چگونه برحسب اندازهٔ مسئله رشد می‌کند. برای طبقه‌بندی توابع برحسب مرتبهٔ رشدشان از نمادهای مجانبی مثل O, Ω و o استفاده می‌کنیم.

مثلاً فرض کنید الگوریتمی را تحلیل کرده‌ایم و دیدیم زمان اجرایش برای ورودی به طول n برابر است با

$$T(n) = an^۲ + bn + cn\ ,$$

که در آن a, b, c اعداد ثابتی هستند (که به نحوهٔ پیاده‌سازی و زبان برنامه‌نویسی هم بستگی دارند) و $a > ۰$. مهم‌ترین نکته اینست که جملهٔ درجهٔ دوم است که رشد $T(n)$ را تعیین می‌کند. به عبارت دیگر، ثابت d وجود دارد که برای هر عدد n داریم

$$T(n) \leq dn^۲\ .$$

---

* asymptotic calculations

۱ complexity theory

ما این نامساوی را به صورت $T(n) = O\left(n^{۲}\right)$ نشان می‌دهیم. دقت کنید که وقتی از این نمادگذاری استفاده می‌کنیم معنی‌اش اینست که مقدار دقیق d برایمان مهم نیست.

تعریف. فرض کنید $f, g : \mathbb{N} \to \mathbb{R}$. می‌نویسیم $f(n) = O(g(n))$ هرگاه اعداد ثابت $C, n_{۰}$ وجود داشته باشند که

$$|f(n)| \leq C|g(n)| \qquad \forall n > n_{۰}$$

تعریف بالا معادلست با تعریف حدی زیر

$$f(n) = O(g(n)) \Leftrightarrow \exists C : \limsup_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| \leq C\,.$$

گاهی اوقات برای کوتاه کردن نمادگذاری‌ها n را حذف می‌کنیم و می‌نویسیم $f = O(g)$. یک حالت خاص بسیار مهم وقتی است که f تابعی کران‌دار باشد. در این صورت می‌نویسیم $f(n) = O\left(۱\right)$.

هم‌چنین می‌توانیم به کمک نماد O بیان کنیم که تابع f با سرعت مشخصی به صفر میل می‌کند. مثلاً $f(n) = O(۱/n)$ یعنی عدد ثابت C وجود دارد که

$$-C/n \leq f(n) \leq C/n\,.$$

تذکر. در این مقاله تأکید بر کاربرد محاسبات مجانبی در علوم کامپیوتر است، و بنابراین فقط با توابع از $\mathbb{N}$ به $\mathbb{R}$ کار می کنیم. در حالت کلی نماد O را می‌توان برای توابع دلخواه هم تعریف کرد، و هم‌چنین برای حالتی که پارامتر تابع به صفر (یا هر عدد ثابت دیگری) میل می کند. برای توضیح بیشتر [۳] را ببینید.

در این مقاله، n همواره عددی طبیعی را نشان می‌دهد که به بی‌نهایت میل می‌کند.

مثال ۱. ۱. برای هر $k \geq ۲$ داریم $an^{۲} + bn + c = O\left(n^{k}\right)$.

۲.

$$\sqrt{۳n + ۱۰} = O\left(\sqrt{n}\right)\,.$$

۳.

$$\ln(n^{۷}) = O\left(\ln n\right)\,.$$

۴. برای هر عدد ثابت k داریم $n^{k} = O\left(۲^{n}\right)$.

۵.

$$n! = O\left(n^{n}\right)\,.$$

۶.

$$e^{\sin n} = O(۱) .$$

۷. برای هر عدد ثابت k داریم $e^{-n} = O\left(n^{-k}\right)$.

وقتی عبارتی مثل $O(g(n))$ را داخل فرمولی دیگر به کار می‌بریم، مثلاً

$$f(n) = ۲^{O(g(n))} ,$$

منظورمان اینست که تابع $h(n)$ وجود دارد که $f(n) = ۲^{h(n)}$ و $h(n) = O(g(n))$، ولی مقدار دقیق $h(n)$ برایمان اهمیت ندارد. برای مثال، می‌توانیم بنویسیم

$$\sum_{i=۱}^{n} i = \frac{n^۲}{۲} + O(n) = \frac{n^۲}{۲}\left(۱ + O\left(\frac{۱}{n}\right)\right) .$$

تمرین ۱. ۱. اگر $f_۱ = O(g)$ و $f_۲ = O(g)$ ثابت کنید $f_۱ + f_۲ = O(g)$.

۲. اگر $f = O(g)$ و $g = O(h)$ ثابت کنید $f = O(h)$.

۳. اگر $f \leq g$ و $g = O(h)$ ثابت کنید $f = O(h)$.

۴. وقتی می‌نویسیم $f(n) = O(\log n)$، چرا لازم نیست مبنای لگاریتم را بیان کنیم؟

۵. اشکال موجود در برهان زیر را بیابید.

برای هر k داریم $kn = O(n)$. در نتیجه،

$$\sum_{k=۱}^{n} kn = \sum_{k=۱}^{n} O(n) = nO(n) = O\left(n^۲\right) .$$

۶. آیا $۲^{O(n)}$ همان $O(۲^n)$ است؟ اثبات کنید، یا مثال نقض بزنید.

در برخی از علوم، مثلاً فیزیک، منظور از $f = O(g)$ اینست که مرتبهٔ رشد f و g یکی است. ولی در علوم کامپیوتر، این فرمول یک کران بالا برای f ارائه می‌دهد: مثلاً $n^۲ = O\left(n^۳\right)$. به زبان فارسی، $f = O(g)$ بیان می‌کند که «f از g سریع‌تر رشد نمی‌کند» و یا «اگر ضرایب ثابت را نادیده بگیریم، $f \leq g$».

به همین ترتیب، می‌توان نمادهای مجانبی دیگری را هم تعریف کرد.

Ω: برعکس O است: $f = \Omega(g)$ یعنی f از g کندتر رشد نمی‌کند. به عبارت دقیق‌تر، اعداد ثابت $C, n_۰ > ۰$ هستند که

$$|f(n)| \geq C|g(n)| \qquad \forall n > n_۰ .$$

معادلاً،

$$f(n) = \Omega(g(n)) \Leftrightarrow \exists C > ۰ : \liminf_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| \geq C .$$

Θ: ترکیب O و Ω است: می‌نویسیم $f = \Theta(g)$ هرگاه اعداد ثابت $C_۱, C_۲ > ۰$ وجود داشته باشند که برای nهای بزرگ،

$$C_۱|g(n)| \leq |f(n)| \leq C_۲|g(n)| .$$

معادلاً،

$$f(n) = \Theta(g(n)) \Leftrightarrow \exists C_۱, C_۲ > ۰ : C_۱ \leq \liminf_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| \leq \limsup_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| \leq C_۲ .$$

معادلاً،

$$f(n) = \Theta(g(n)) \Leftrightarrow f(n) = O(g(n)) \text{ and } f(n) = \Omega(g(n)) .$$

در برخی کتاب‌های $f(n) = \Theta(g(n))$ را به صورت $f(n) \asymp g(n)$ هم نمایش می‌دهند.

o: قوی‌تر از O است: $f = o(g)$ یعنی «f اکیداً کندتر از g رشد می‌کند». به عبارت ریاضی،

$$f(n) = o(g(n)) \Leftrightarrow \lim_{n\to\infty} \frac{f(n)}{g(n)} = ۰ .$$

ω: برعکس o است: $f = \omega(g)$ یعنی f اکیداً سریع‌تر از g رشد می‌کند. به عبارت ریاضی،

$$f(n) = \omega(g(n)) \Leftrightarrow \forall C > ۰ : \liminf_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| \geq C \Leftrightarrow \lim_{n\to\infty} \frac{g(n)}{f(n)} = ۰ .$$

$\sim$: از Θ قوی‌تر است: $f(n) \sim g(n)$ یعنی

$$\lim_{n\to\infty} \left|\frac{f(n)}{g(n)}\right| = ۱ .$$

مثلاً داریم

$$\sum_{i=۱}^{n} i = \frac{n^۲ + n}{۲} \sim \frac{n^۲}{۲} .$$

poly : وقتی می‌نویسیم $f(n) = \mathrm{poly}(g(n))$ یعنی $f(n) = g(n)^{O(1)}$. به عبارت دیگر، ثابت C وجود دارد که برای nهای بزرگ داریم

$$|f(n)| \leq g(n)^C .$$

$\widetilde{O}$: وقتی می‌نویسیم $f = \widetilde{O}(g)$ یعنی

$$f(n) = O\left(g(n) \times \mathrm{poly}(\log g(n))\right) .$$

این نمادگذاری در مواقعی به کار می‌رود که نه تنها ضرایب ثابت برایمان مهم نیستند، بلکه عبارت‌هایی که برحسب $\log g(n)$ چندجمله‌ای هستند (و لذا بسیار از $g(n)$ کوچک‌ترند) هم برایمان مهم نیستند. برای مثال، داریم

$$n^۵ \times ۳^n = \widetilde{O}(۴^n), \qquad\qquad n^۵ \times ۳^n = \widetilde{O}(۳^n) ,$$

ولی

$$n^۵ \times ۳^n \neq \widetilde{O}(۲^n) ,$$

چرا که ۳^n به مراتب رشد سریع‌تری نسبت به ۲^n دارد. نمادهای $\widetilde{\Omega}$ و $\widetilde{\Theta}$ هم مشابهاً تعریف می‌شوند.

تمرین ۲. ۱. مثال ۱ را نگاه کنید. در کدام آن‌ها می‌توان O را با Θ جایگزین کرد؟ در کدام آن‌ها می‌توان O را با o جایگزین کرد؟

۲. توابع f و g را مثال بزنید که $f = \Theta(g)$ ولی $۲^f = o\left(۲^g\right)$.

۳. توابع f و g را مثال بزنید که $\log f = \Theta(\log g)$ ولی $f = o(g)$.

۴. برای هر زوج f, g از توابع زیر تعیین کنید که آیا $f = o(g)$ یا $f = \Theta(g)$ و یا $f = \omega(g)$؟

$$\begin{aligned} f &= \ln(n^۲), & g &= \ln\sqrt{n} \\ f &= ۲^n, & g &= ۳^{n/۲} \\ f &= n^{\ln n}, & g &= ۲^n \\ f &= ۲^n, & g &= ۲^{n+\ln n} . \end{aligned}$$

۵. چندین تابع مثل f مثال بزنید که برای هر عدد ثابت $c > ۰$، $f(n) = \omega(n^c)$ باشد و $f(n) = o\left(۲^{n^c}\right)$. (راهنمایی: ابتدا $\ln f$ را پیدا کنید!)

یک نکتهٔ بسیار مهم و کمی گمراه‌کننده دربارهٔ محاسبات مجانبی اینست که تساوی‌های مجانبی دوطرفه نیستند. مثلاً تساوی $\sum_{i=۱}^{n} i = \frac{n^۲+n}{۲}$ دوطرفه است یعنی می‌توان نوشت $\frac{n^۲+n}{۲} = \sum_{i=۱}^{n} i$. ولی تساوی

$$\frac{n^۲ + n}{۲} = O\left(n^۲\right)$$

دوطرفه نیست و عبارتی مثل

$$O\left(n^۲\right) = \frac{n^۲ + n}{۲}$$

اصولاً بی‌معناست! با این حال، می‌توانیم بنویسیم

$$\sum_{i=۱}^{n} i = \frac{n^۲ + n}{۲} = \frac{n^۲}{۲}\left(۱ + \frac{۱}{n}\right) = \frac{n^۲}{۲}\left(۱ + o(۱)\right) = O\left(n^۲\right)$$

و چنین عبارتی را باید «از چپ به راست» خواند و فهمید.

تمرین ۳. فرض کنید $x(n)$ و $y(n)$ توابعی دلخواه بر حسب n باشند و $x \sim y$. همچنین فرض کنید $a(n)$ و $b(n)$ توابعی دلخواه بر حسب n باشند و k یک عدد طبیعی ثابت باشد. هر یک از گزاره‌های زیر را اثبات کنید یا برایش مثال نقض بزنید.

۱. $y \sim x$

۲. $x - a \sim y - a$

۳. اگر $a \sim b$ آن‌گاه $xa \sim yb$.

۴. $x^k \sim y^k$

۵. $x^a \sim y^a$

۶. $k^x \sim k^y$

۲ چند تقریب مفید.

در این بخش از قضیهٔ تیلور[۲] استفاده می‌کنیم تا دو تقریب مهم را، که در بخش‌های بعدی بهشان نیاز داریم اثبات کنیم. شایان ذکر است که این قضیه گونه‌های مختلفی دارد و نسخه‌ای که ما در این‌جا بیان می‌کنیم، فرم لاگرانژی باقی‌مانده[۳] نام دارد [۸].

قضیهٔ ۲. فرض کنید تابع f دارای $k+۱$ مشتق پیوسته در بازهٔ $[a-r, a+r]$باشد، و فرض کنید عدد ثابت M وجود دارد که

$$\forall x \in [a-r, a+r] \qquad \left|f^{(k+۱)}(x)\right| \leq M\,.$$

[۲] Taylor's theorem
[۳] Lagrange form of the remainder

در این صورت، برای هر $x \in (a-r, a+r)$ داریم

$$\left| f(x) - f(a) - f'(a)(x-a) - \frac{f''(a)(x-a)^{۲}}{۲!} - \cdots - \frac{f^{(k)}(a)(x-a)^{k}}{k!} \right| \leq \frac{M|x-a|^{k+۱}}{(k+۱)!} .$$

با استفاده از این قضیه می‌توان سه نامساوی زیر را ثابت کرد.

$$۱ + x \leq \exp(x) \leq ۱ + x + x^{۲} \qquad \forall x \in (-۱, ۱) \qquad (۱)$$

$$\exp(x - x^{۲}) \leq ۱ + x \qquad \forall x \in (-\frac{۱}{۳}, \frac{۱}{۳}) \qquad (۲)$$

تذکر. شایان ذکر است که نامساوی‌های (۱) در حقیقت برای هر $x \in (-\infty, \frac{۷}{۴}]$ صادق هستند (نامساوی $۱ + x \leq e^{x}$ برای هر x حقیقی صادق است!) و نامساوی (۲) برای هر $x \in [-\frac{۲}{۳}, +\infty)$ صادق است. برای بررسی صحت این ادعاها کافیست نمودار این توابع را به کمک نرم‌افزار مورد علاقه‌تان رسم کنید.

قرار دهید $a = ۰, r = ۱, f(x) = e^{x}$. دقت کنید که برای هر $x \in [-۱, ۱]$ داریم

$$|f'''(x)| = |e^{x}| \leq e ,$$

در نتیجه طبق قضیهٔ ۲، برای هر $x \in [-۱, ۱]$ داریم

$$\left| e^{x} - ۱ - x - \frac{x^{۲}}{۲} \right| \leq \frac{ex^{۳}}{۶} \leq \frac{x^{۲}}{۲} ,$$

و (۱) ثابت می‌شود.

برای اثبات (۲)، قرار می‌دهیم $f(x) = \ln(x), a = ۱, r = ۱/۳$. برای هر $x \in [۲/۳, ۴/۳]$ داریم

$$|f'''(x)| = \left|x^{-۳}\right| \leq ۲۷/۸ .$$

پس طبق قضیهٔ ۲، برای هر $x \in (۲/۳, ۴/۳)$ داریم

$$\left| \ln(x) - (x-۱) + \frac{(x-۱)^{۲}}{۲} \right| \leq \frac{۲۷}{۸} \times \frac{(x-۱)^{۳}}{۶} \leq \frac{(x-۱)^{۲}}{۲} ,$$

و در نتیجه

$$\ln(x) \geq (x-۱) - (x-۱)^{۲} ,$$

پس برای هر $y = x - ۱ \in (-۱/۳, ۱/۳)$ داریم

$$\ln(۱ + y) \geq y - y^{۲} ,$$

و (۲) با رساندن e به توان دو طرف ثابت می‌شود.

تمرین ۴. به یاد بیاورید که برای هر عدد حقیقی x داریم

$$e^x = ۱ + x + \frac{x^۲}{۲!} + \frac{x^۳}{۳!} + \frac{x^۴}{۴!} + \cdots$$

با استفاده از این تساوی، ثابت کنید اولاً برای هر عدد حقیقی مثبت x داریم

$$e^{-x} \leq ۱ - x + x^۲/۲ \ ,$$

و ثانیاً برای هر عدد طبیعی k داریم

$$e^k > k^k/k! \ .$$

۳ اعداد همساز.

فرض کنید m سطل داریم و تعداد زیادی توپ را به صورت تصادفی بین آن‌ها پخش می‌کنیم. به عبارت دقیق‌تر، در هر گام یک توپ جدید برمی‌داریم، یک سطل را با توزیع یکنواخت انتخاب می‌کنیم و توپ را درون آن می‌اندازیم. به طور متوسط چند تا توپ باید بیندازیم تا هیچ سطلی خالی نماند؟ فرض کنیم X تعداد گام‌هایی را نشان بدهد که بعد از گام Xاُم، هیچ سطلی خالی نیست. توجه کنید که X متغیری تصادفی است و در این‌جا می‌خواهیم امیدریاضی‌اش را به دست بیاوریم.

در گام اول یکی از سطل‌ها صاحب توپ می‌شود. در گام دوم توپی که می‌اندازیم به احتمال $\frac{m-۱}{m}$ داخل یک سطل جدید می‌رود، و به احتمال $\frac{۱}{m}$ داخل همان سطل صاحب توپ می‌رود. پس تعداد گام‌هایی که طول می‌کشد تا سطل جدیدی صاحب توپ بشود، یک متغیر تصادفی هندسی با پارامتر $\frac{m-۱}{m}$ است. به طور کلی، اگر k تا سطل صاحب توپ داشته باشیم، آن‌گاه تعداد گام‌هایی که طول می‌کشد که سطل تازه‌ای صاحب توپ شود، متغیری هندسی با پارامتر $\frac{m-k}{m}$ است. به علاوه، این متغیرهای تصادفی از هم مستقلند. چون امید ریاضی چنین متغیری $\frac{m}{m-k}$ است، داریم

$$\mathbb{E}\,[X] = \sum_{k=۰}^{m-۱} \frac{m}{m-k} = m \times \sum_{k=۱}^{m} \frac{۱}{k} \ .$$

تعریف کنید

$$H_n = \frac{۱}{۱} + \frac{۱}{۲} + \frac{۱}{۳} + \cdots + \frac{۱}{n}.$$

عدد H_n را عدد همساز nاُم[۴] می‌نامیم. بنابراین پاسخ مسئلهٔ سطل‌ها و توپ‌ها دقیقاً mH_m است. در این بخش مقدار H_n را برای nهای بزرگ تقریب می‌زنیم.

[۴] n-th harmonic number

فرض بگیریم $k = \lceil \log_۲(n+۱) \rceil$. داریم $۲^k - ۱ \geq n$ ولذا

$$\begin{aligned}
H_n &= \frac{۱}{۱} + \frac{۱}{۲} + \frac{۱}{۳} + \cdots + \frac{۱}{n} \\
&\leq \frac{۱}{۱} + \frac{۱}{۲} + \frac{۱}{۳} + \cdots + \frac{۱}{۲^k - ۱} \\
&= \left(\frac{۱}{۱}\right) + \left(\frac{۱}{۲} + \frac{۱}{۳}\right) + \left(\frac{۱}{۴} + \frac{۱}{۵} + \frac{۱}{۶} + \frac{۱}{۷}\right) + \cdots + \left(\frac{۱}{۲^{k-۱}} + \cdots + \frac{۱}{۲^k - ۱}\right) \\
&\leq \left(\frac{۱}{۱}\right) + \left(\frac{۱}{۲} + \frac{۱}{۲}\right) + \left(\frac{۱}{۴} + \frac{۱}{۴} + \frac{۱}{۴} + \frac{۱}{۴}\right) + \cdots + \left(\frac{۱}{۲^{k-۱}} + \cdots + \frac{۱}{۲^{k-۱}}\right) \\
&= ۱ + ۱ + ۱ + \cdots + ۱ = k\,.
\end{aligned}$$

بنابراین، $H_n \leq k = \lceil \log_۲(n+۱) \rceil = O\left(\log_۲ n\right)$.

مشابهاً اگر $k = \lfloor \log_۲ n \rfloor$ آن‌گاه $۲^k \leq n$ ولذا

$$\begin{aligned}
H_n &= \frac{۱}{۱} + \frac{۱}{۲} + \frac{۱}{۳} + \cdots + \frac{۱}{n} \\
&\geq \frac{۱}{۱} + \frac{۱}{۲} + \frac{۱}{۳} + \cdots + \frac{۱}{۲^k} \\
&= \left(\frac{۱}{۱}\right) + \left(\frac{۱}{۲}\right) + \left(\frac{۱}{۳} + \frac{۱}{۴}\right) + \left(\frac{۱}{۵} + \frac{۱}{۶} + \frac{۱}{۷} + \frac{۱}{۸}\right) + \cdots + \left(\frac{۱}{۲^{k-۱}+۱} + \cdots + \frac{۱}{۲^k}\right) \\
&\geq \left(\frac{۱}{۱}\right) + \left(\frac{۱}{۲}\right) + \left(\frac{۱}{۴} + \frac{۱}{۴}\right) + \left(\frac{۱}{۸} + \frac{۱}{۸} + \frac{۱}{۸} + \frac{۱}{۸}\right) + \cdots + \left(\frac{۱}{۲^k} + \cdots + \frac{۱}{۲^k}\right) \\
&= ۱ + \frac{۱}{۲} + \frac{۱}{۲} + \cdots + \frac{۱}{۲} = ۱ + k/۲\,.
\end{aligned}$$

بنابراین، $H_n \geq ۱ + \lfloor \log_۲ n \rfloor / ۲ = \Omega(\log_۲ n)$. تا همین‌جا می‌توانیم نتیجه بگیریم که $H_n = \Theta(\log n)$.

چگونه می‌توانیم فرمول دقیق‌تری برای H_n بیابیم؟ برای این کار، سری $\sum_{i=۱}^{n} ۱/i$ را با انتگرال تقریب می‌زنیم. با توجه به شکل ۱ داریم

$$\frac{۱}{۲} + \frac{۱}{۳} + \cdots + \frac{۱}{n} \leq \int_{x=۱}^{n} \frac{۱}{x}\,\mathrm{d}x \leq ۱ + \frac{۱}{۲} + \frac{۱}{۳} + \cdots + \frac{۱}{n}\,. \tag{۳}$$

حال دقت کنید که

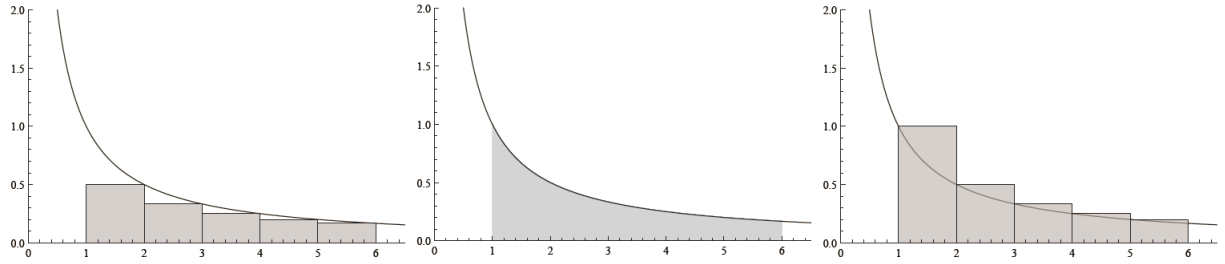
$$\int_{x=۱}^{n} \frac{۱}{x}\,\mathrm{d}x = \ln x|_{۱}^{n} = \ln n - \ln ۱ = \ln n\,,$$

پس

$$H_n - ۱ \leq \ln n \leq H_{n-۱}\,,$$

که نتیجه می‌دهد

$$\ln n < \ln(n+۱) \leq H_n \leq ۱ + \ln n\,,$$



(a) $\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}$ (b) $\int_1^n \frac{1}{x}\,dx$ (c) $1+\frac{1}{2}+\cdots+\frac{1}{n-1}$

شکل ۱: اثبات (۳)

پس $H_n \sim \ln n$. از حوصلهٔ این مقاله خارج است، ولی در مقالهٔ [۹] نشان داده شده است که در حقیقت

$$H_n = \ln n + \gamma + O(۱/n)\,,$$

که در آن $\gamma \approx ۰/۵۷۷$ به ثابت Euler-Mascheroni شهرت دارد [۱].

۴ پارادکس روز تولد.

همانند بخش قبل، فرض کنید m سطل داریم و n توپ را به صورت تصادفی بین آن‌ها پخش می‌کنیم. در این بخش به این پرسش می‌پردازیم: چقدر احتمال دارد که همهٔ توپ‌ها سطل‌های متفاوتی را انتخاب کنند؟ به عبارت دیگر، چقدر احتمال دارد که هیچ سطلی بیش از یک توپ دریافت نکند؟

اگر $m = ۳۶۶$، این مسئله به پارادکس روز تولد[۵] مشهور است: اگر n نفر آدم داشته باشیم، چقدر احتمال دارد که هیچ دو نفری روز تولدشان یکی نباشد؟ علت استفاده از کلمهٔ «پارادکس» اینست که این احتمال خیلی کم‌تر از آنست که در نگاه اول به نظر می‌آید: مثلاً اگر ۳۰ نفر داشته باشیم، این احتمال کم‌تر از ۳۰ درصد است!

حال این احتمال، که آن را $p_{n,m}$ می‌نامیم محاسبه می‌کنیم. برای این که همهٔ n توپ سطل‌های متفاوتی را انتخاب کنند، بدیهی است که باید $n \le m$ باشد. فرض کنید توپ‌ها یکی‌یکی سطل انتخاب می‌کنند. توپ اول سطل دلخواهی را انتخاب می‌کند. توپ دوم باید یکی از $m-۱$ سطل باقی‌مانده را انتخاب کند، که احتمالش $\frac{m-۱}{m}$ است. توپ سوم باید یکی از $m-۲$ سطل باقی‌مانده را انتخاب کند، که احتمالش $\frac{m-۲}{m}$ است. به همین ترتیب تا توپ nاُم، که باید یکی از $m-(n-۱)$ سطل باقی‌مانده را انتخاب کند، که احتمالش $\frac{m-(n-۱)}{m}$ است. در نتیجه، داریم

$$p_{n,m} = \left(\frac{m-۱}{m}\right)\left(\frac{m-۲}{m}\right)\cdots\left(\frac{m-(n-۱)}{m}\right)\,.$$

گرچه این عبارت فرمی بسته برای $p_{n,m}$ به ما می‌دهد، به خودی خود چندان خوش‌دست نیست. مثلاً فرض کنید سؤالمان این باشد که

[۵] the birthday paradox

«برای یک m داده شده، کوچک‌ترین مقدار n چیست به طوری که $p_{n,m} \leq ۱/۲$ باشد؟»

پاسخ این پرسش را با $f(m)$ نشان می‌دهیم. (دقت کنید که برای یک m داده شده، $p_{n,m}$ برحسب n نزولی است بنابراین $f(m)$ خوش‌تعریف است.) برای بسیاری از مسائل فقط مرتبهٔ $f(m)$ برایمان مهم است. در این بخش مرتبهٔ $f(m)$ را، برای mهای بزرگ، پیدا می‌کنیم.

قضیهٔ ۳. $f(m) = \Theta(\sqrt{m})$.

اثبات. ابتدا یک کران بالا برای $f(m)$ می‌یابیم. با استفاده از نامساوی $e^x \geq ۱ + x$ که در بخش ۲ دیدیم، داریم

$$p_{n,m} = \left(۱ - \frac{۱}{m}\right)\left(۱ - \frac{۲}{m}\right)\cdots\left(۱ - \frac{n-۱}{m}\right) \leq e^{-۱/m} \cdot e^{-۲/m} \cdots e^{-(n-۱)/m}$$
$$= \exp\left(-۱/m - ۲/m - \cdots - (n-۱)/m\right) = \exp\left(\frac{-n(n-۱)}{۲m}\right)$$

اگر $n \geq ۲\sqrt{m}$ آن‌گاه

$$\frac{n(n-۱)}{۲m} \geq \frac{n^۲}{۴m} \geq ۱\,,$$

ولذا

$$p_{n,m} \leq \exp\left(\frac{-n(n-۱)}{۲m}\right) \leq e^{-۱} < ۱/۲\,.$$

از آن‌جا که $p_{n,m}$ برحسب n نزولی است، نتیجه می‌گیریم

$$f(m) \leq ۲\sqrt{m} = O(\sqrt{m}). \qquad (۴)$$

حال کران پایینی برای $f(m)$ می‌یابیم. فرض می‌گیریم

$$n < \sqrt{۲\ln(۴/۳)m}$$

و ثابت می‌کنیم برای این n، $p_{n,m} > ۱/۲$ است. نتیجه می‌گیریم $f(m) = \Omega(\sqrt{m})$ و قضیه ثابت می‌شود!

توجه کنید که برای m های بزرگ داریم $n - ۱ \leq m/۳$ بنابراین از نامساوی (۲) داریم

$$۱ - \frac{n-۱}{m} \geq \exp\left(-\frac{n-۱}{m} - \left(-\frac{n-۱}{m}\right)^۲\right) = \exp\left(-\frac{n-۱}{m} - O\left(\frac{n-۱}{m}\right)^۲\right)\,,$$

بنابراین

$$\begin{aligned} p_{n,m} &= \left(۱ - \frac{۱}{m}\right)\left(۱ - \frac{۲}{m}\right)\cdots\left(۱ - \frac{n-۱}{m}\right) \\ &\geq \exp\left(-\frac{۱}{m} - O\left(\frac{۱}{m}\right)^{۲}\right)\cdots\exp\left(-\frac{n-۱}{m} - O\left(\frac{n-۱}{m}\right)^{۲}\right) \\ &= \exp\left(-\frac{۱}{m} - \frac{۲}{m} - \cdots - \frac{n-۱}{m}\right)\exp\left(-O\left(\frac{۱}{m^{۲}} + \cdots + \frac{(n-۱)^{۲}}{m^{۲}}\right)\right) \\ &= \exp\left(-\frac{n(n-۱)}{۲m}\right)\exp\left(-O\left(\frac{n^{۳}}{m^{۲}}\right)\right) \end{aligned} \tag{۵}$$

در تساوی آخر از رابطهٔ

$$۱^{۲} + ۲^{۲} + \cdots + (n-۱)^{۲} \leq (n-۱) \times (n-۱)^{۲} = O\left(n^{۳}\right)$$

استفاده کردیم. با توجه به فرض $n < \sqrt{۲\ln(۴/۳)m}$، داریم

$$\exp\left(-\frac{n(n-۱)}{۲m}\right) \geq \exp\left(-\frac{n^{۲}}{۲m}\right) > \exp\left(-\ln(۴/۳)\right) = ۳/۴$$

و هم‌چنین برای mهای بزرگ،

$$\exp\left(-O\left(\frac{n^{۳}}{m^{۲}}\right)\right) \geq ۱ - O\left(\frac{n^{۳}}{m^{۲}}\right) \geq ۱ - O\left(\frac{m^{۳/۲}}{m^{۲}}\right) = ۱ - O(۱/\sqrt{m}) > ۲/۳\,.$$

با کنار هم گذاشتن این نامساوی‌ها، می‌رسیم به

$$p_{n,m} \geq \exp\left(-\frac{n(n-۱)}{۲m}\right)\exp\left(-O\left(\frac{n^{۳}}{m^{۲}}\right)\right) > ۳/۴ \times ۲/۳ > ۱/۲\,,$$

و قضیه ثابت می‌شود.

شایان ذکر است که تمام این محاسبات را با اعداد «واقعی» به جای O هم می‌شد انجام داد. مزیت استفاده از این نماد برای نویسندهٔ اثبات اینست که لازم نیست درگیر محاسبهٔ مقادیر عددی بشود و می‌تواند وقتش را روی «هستهٔ اثبات» بگذارد، و در عین حال به نتیجهٔ لازم که مرتبهٔ $f(m)$ است دست بیابد. مزیت آن برای خوانندهٔ اثبات اینست که اثبات بدون مقادیر عددی تمیزتر است، و همچنین خواننده تمرکزش روی هستهٔ اثبات بیشتر می‌شود بدون این که درگیر شاخ‌وبرگ‌های اضافی بشود. □

تمرین ۵. اثبات کنید که برای هر عدد k ثابت، داریم

$$۱^{k} + ۲^{k} + \cdots + n^{k} \sim \frac{n^{k+۱}}{k+۱}\,.$$

۵ فرمول استرلینگ. می‌دانیم که

$$n! = n \times (n-۱) \times (n-۲) \times \cdots \times ۲ \times ۱ ,$$

ولی این تساوی دربارهٔ نحوهٔ رشد $n!$ برحسب n چیزی نمی‌گوید. در این بخش رشد $n!$ را بررسی می‌کنیم و به فرمول استرلینگ[۶] می‌رسیم.

یک کران بالای بدیهی برای $n!$ عبارتست از n^n. دو کران پایین ساده عبارتند از:

(۱) $n! \geq ۲^{n-۱}$ چون همهٔ ضرایب به غیر از ۱، از ۲ بزرگ‌ترمساوی‌اند.

(۲) برای nهای زوج، $\frac{n}{۲}+۱$ تا از ضرایب از $n/۲$ بزرگ‌ترمساوی‌اند، پس $n! \geq \left(\frac{n}{۲}\right)^{۱+n/۲}$.

کدام یک از $۲^{n-۱}$ و $\left(\frac{n}{۲}\right)^{۱+n/۲}$ کران پایین بهتری‌اند؟ سرراست‌ترین کار برای کار با چنین اعداد بزرگی کار با لگاریتم آن‌هاست. از آن‌جا که برای nهای بزرگ،

$$(n-۱)\ln ۲ \leq (۱+n/۲)\ln\left(\frac{n}{۲}\right) \leq n\ln n , \qquad (۶)$$

می‌رسیم به

$$۲^{n-۱} \leq \left(\frac{n}{۲}\right)^{۱+n/۲} \leq n! \leq n^n .$$

تا همین‌جا و با نگاه کردن به نامساوی‌های (۶) می‌توان دید که $\ln(n!) = \Theta(n\ln n)$ و در نتیجه $n! = ۲^{\Theta(n\ln n)}$ ولی ما کران دقیق‌تری می‌خواهیم.

تمرین ۶. تابع $f(n)$ را مثال بزنید که $f(n) = ۲^{\Theta(n\ln n)}$ ولی $f(n) \neq \Theta\left(۲^{n\ln n}\right)$. ثابت کنید که اگر $g(n) = \Theta\left(۲^{n\ln n}\right)$ آن‌گاه $g(n) = ۲^{\Theta(n\ln n)}$.

برای به دست آوردن کران دقیق‌تری برای $n!$، مجدداً از لگاریتم گرفتن کمک می‌گیریم:

$$\ln(n!) = \ln\left(n \times (n-۱) \times (n-۲)\ldots ۲ \times ۱\right) = \ln ۲ + \ln ۳ + \cdots + \ln n .$$

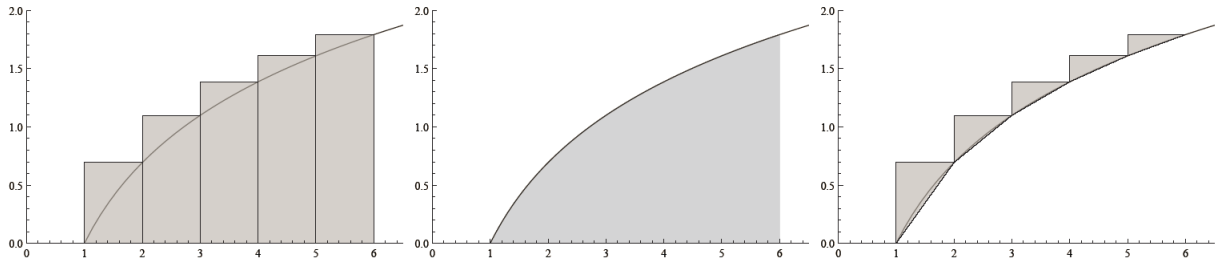
یک روش کارگشا در این مواقع (که در بخش ۳ هم دیدیم) تقریب زدن سری با انتگرال است. با توجه به شکل ۲(چپ) داریم

$$\begin{aligned}\ln(n!) &\geq \int_{۱}^{n} \ln x \mathrm{d}x = x\ln x - x\big|_{۱}^{n} \\ &= (n\ln n - n) - (۱\ln ۱ - ۱) = n\ln n - n + ۱ . \end{aligned} \qquad (۷)$$

برای دادن کران بالا برای $\ln(n!)$ به شکل ۲(راست) نگاه می‌کنیم. اگر مجموع مساحت مثلث‌ها را با S نشان دهیم، دقت کنید که

$$\ln(n!) \leq \int_{۱}^{n} \ln x \mathrm{d}x + S ,$$

[۶] Stirling's formula



(a) $\ln 2 + \cdots + \ln n$ (b) $\int_1^n \ln x \, dx$ (c) $\frac{1}{2} \ln n$

شکل ۲: اثبات (۹)

چرا که اگر مساحت زیر نمودار را با مساحت مثلث‌ها جمع بزنیم، قسمت‌های باریک بین نمودار و وترهای مثلث‌ها را دوبار حساب کرده‌ایم. حال دقت کنید که عرض هر یک از مثلث‌ها برابر واحد است، و مجموع ارتفاعات آن‌ها دقیقاً $\ln n$ است، در نتیجه $S = (\ln n)/۲$. نتیجه می‌گیریم

$$\ln(n!) \leq n \ln n - n + ۱ + (\ln n)/۲ \, . \tag{۸}$$

از ترکیب نامساوی‌های (۷) و (۸) و رساندن e به توان طرفین می‌رسیم به

$$e(n/e)^n \leq n! \leq e\sqrt{n}(n/e)^n \, . \tag{۹}$$

و در نتیجه

$$n! = \widetilde{\Theta}\left(\frac{n^n}{e^n}\right) \, .$$

فرمول استرلینگ (یا تقریب استرلینگ) نسخهٔ قوی‌تری از این رابطه است که بیان می‌کند

$$n! \sim \sqrt{۲\pi n}(n/e)^n \, . \tag{۱۰}$$

نسخهٔ قوی‌تری از این فرمول هم هست که بیان می‌کند

$$n! = \sqrt{۲\pi n}(n/e)^n \left(۱ + O(۱/n)\right) \, .$$

اگر تقریبی بخواهیم که برای هر عدد n درست باشد (نه فقط برای nهای بزرگ) می‌توانیم از تقریب بسیار دقیق زیر استفاده کنیم که برای هر عدد طبیعی $n > ۱$ صدق می‌کند:

$$\sqrt{۲\pi n}\left(\frac{n}{e}\right)^n < \sqrt{۲\pi n}\left(\frac{n}{e}\right)^n \exp\left(\frac{۱}{۱۲n+۱}\right) < n! < \sqrt{۲\pi n}\left(\frac{n}{e}\right)^n \exp\left(\frac{۱}{۱۲n}\right) < e\sqrt{n}\left(\frac{n}{e}\right)^n , \tag{۱۱}$$

برای اطلاعات بیشتر صفحهٔ ویکیپدیای مربوطه را ببینید [۷]. فرمول استرلینگ کاربردهای فراوانی در ترکیبیات دارد، که از جملهٔ آن‌ها می‌توان به تقریب زدن ضریب دوجمله‌ای اشاره کرد، که در بخش بعدی خواهیم دید.

۶ تقریب ضریب دوجمله‌ای. به یاد بیاورید که

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-۱)\cdots(n-k+۱)}{k!}$$

را ضریب دوجمله‌ای می‌گوییم. در این بخش از فرمول استرلینگ یعنی (۱۰) استفاده می‌کنیم تا مقدار این ضریب را برای nهای بزرگ تقریب بزنیم. دقت کنید که k می‌تواند ثابت باشد یا این که خودش تابعی از n باشد و به بی‌نهایت میل کند. برحسب سرعت رشد k برحسب n ما سه حالت را بررسی می‌کنیم.

قبل از آغاز، به عنوان دست‌گرمی تمرین زیر را حل کنید.

تمرین ۷. برای هر دو عدد طبیعی $k < n$ ثابت کنید

$$\left(\frac{n}{k}\right)^k \le \binom{n}{k} \le \frac{n^k}{k!} \le \left(\frac{en}{k}\right)^k .$$

در صورت لزوم از تقریب (۱۱) استفاده کنید.

۱.۶ $k = O(۱)$. اگر k عددی ثابت باشد (یا این که تابعی کران‌دار برحسب n باشد) آن‌گاه

$$\binom{n}{k} = \frac{n(n-۱)\cdots(n-k+۱)}{k!} \sim \frac{n^k}{k!}$$

۲.۶ $k = o(\sqrt{n})$. در این بخش فرض می‌کنیم $k = k(n)$ تابعی برحسب n است ولی $k = o(\sqrt{n})$ یعنی سرعت رشدش اکیداً کوچک‌تر از رادیکال n است. در این صورت،

$$\binom{n}{k} = \frac{n^k}{k!}\left(۱ - \frac{۱}{n}\right)\left(۱ - \frac{۲}{n}\right)\cdots\left(۱ - \frac{k-۱}{n}\right) = \frac{n^k}{k!}p_{k,n} ,$$

که $p_{k,n}$ را در بخش ۴ تعریف کرده بودیم. از نامساوی (۵) که در آن بخش ثابت کردیم، و با توجه به این که $k = o(\sqrt{n})$، داریم

$$p_{k,n} \ge \exp\left(-\frac{k(k-۱)}{۲n} - O\left(\frac{k^۳}{n^۲}\right)\right) = \exp(-o(۱)) \ge ۱ - o(۱) .$$

در نتیجه

$$\binom{n}{k} = \frac{n^k}{k!}p_{k,n} \ge (۱ - o(۱))\frac{n^k}{k!}$$

با ترکیب این نامساوی و $\binom{n}{k} \le \frac{n^k}{k!}$ که برای همهٔ k و nها درست است، می‌رسیم به

$$\binom{n}{k} \sim \frac{n^k}{k!} \sim \left(\frac{en}{k}\right)^k \Big/ \sqrt{۲\pi k} ,$$

که در رابطهٔ آخر از فرمول استرلینگ استفاده کردیم.

۳.۶ $k = \Theta(n)$. حال فرض کنید $k = k(n)$ به صورت خطی برحسب n رشد می‌کند، به عبارت دقیق‌تر $k = pn$ که $p \in (۰, ۱)$ عددی ثابت است. تعریف کنید $q = ۱ - p$. از فرمول استرلینگ داریم

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n!}{(pn)!(qn)!} \sim \frac{\sqrt{۲\pi n}(n/e)^n}{\sqrt{۲\pi pn}(pn/e)^{pn}\sqrt{۲\pi qn}(qn/e)^{qn}} = \frac{p^{-pn}q^{-qn}}{\sqrt{۲\pi pqn}} = \frac{۲^{H(p)n}}{\sqrt{۲\pi pqn}} ,$$

که در آن $H(p) = -p\log_۲ p - q\log_۲ q$ به آنتروپی دودویی p[۷] شهرت دارد.

به عنوان مثال، $H(۱/۲) = ۱$ و در نتیجه

$$\frac{\binom{n}{n/۲}}{۲^n} \sim \frac{۱}{\sqrt{۲\pi n/۴}} = \sqrt{\frac{۲}{\pi n}} = \Theta\left(\frac{۱}{\sqrt{n}}\right) .$$

پس اگر n عددی زوج باشد، وقتی n سکه را شیر یا خط کنیم، احتمال این که دقیقاً نصف آن‌ها شیر بیایند از مرتبهٔ $۱/\sqrt{n}$ است.

این مقاله را با تمرین زیر به پایان می‌رسانیم. خوانندهٔ علاقمند برای دیدن مثال‌ها و تمرین‌های بیشتر می‌تواند مقالهٔ [۳] و یا فصل نهم کتاب [۲] را ببیند.

تمرین ۸. فرض کنید $\pi(n)$ تعداد اعداد اول بین ۱ تا n را نشان بدهد. مثلاً داریم $\pi(۱) = ۰$، $\pi(۲) = ۱$، و $\pi(۳) = \pi(۴) = ۲$ و $\pi(۵) = ۳$. در این تمرین ثابت می کنیم $\pi(n) = \Omega(n/\ln n)$. قضیهٔ اعداد اول [۶] می گوید در حقیقت $\pi(n) \sim n/\ln n$. اثبات این که $\pi(n) = O(n/\ln n)$ هم خیلی سخت نیست. جالب است بدانید فرضیهٔ ریمان معادلست با

$$\pi(n) = \int_۲^n \frac{\mathrm{d}t}{\ln t} + O(\sqrt{n}\ln n) .$$

الف) تعریف کنید $C_n = \binom{۲n}{n}$. ثابت کنید $C_n \geq ۲^n$.

ب) فرض کنید p عددی اول باشد، و فرض کنید k بزرگ‌ترین توانی از p باشد که C_n را عاد می کند (یعنی C_n بر p^k بخش‌پذیر است). ثابت کنید

$$k = \left(\left\lfloor \frac{۲n}{p} \right\rfloor - ۲\left\lfloor \frac{n}{p} \right\rfloor\right) + \left(\left\lfloor \frac{۲n}{p^۲} \right\rfloor - ۲\left\lfloor \frac{n}{p^۲} \right\rfloor\right) + \left(\left\lfloor \frac{۲n}{p^۳} \right\rfloor - ۲\left\lfloor \frac{n}{p^۳} \right\rfloor\right) + \cdots$$

پ) ثابت کنید هر یک از پرانترهای بالا حاصلش یا ۰ است یا ۱. نتیجه بگیرید $p^k \leq ۲n$.

ت) نتیجه بگیرید $\pi(۲n) \geq n/\log_۲(۲n)$ و در نتیجه

$$\pi(n) \geq \frac{\ln ۲}{۲} \times \frac{n}{\ln n} \times \left(۱ - \frac{۱}{n}\right) = \Omega(n/\ln n) .$$

[۷] binary entropy of p

مراجع

[۱] Euler–Mascheroni constant. In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Euler-Mascheroni_constant

[۲] Ronald Graham, Donald Knuth, and Oren Patashnik. Concrete mathematics, Chapter 9: asymptotics. Addison-Wesley Publishing Company, Reading, 1994.

[۳] A.J. Hildebrand. Lecture 2: Asymptotic notations, In *Lecture notes for Asymptotic Methods in Analysis course*, University of Illinois at Urbana-Champaign, Fall 2009. http://www.math.illinois.edu/~ajh/595ama/ama-ch2.pdf

[۴] Cristopher Moore and Stephan Mertens. The Nature of Computation, Appendix A: mathematical tools. Oxford University Press, Oxford, 2011. pp. 911–914.

[۵] Ryan O'Donnell. Lecture 1: Asymptotics. In *Lecture notes for A Theorist's Toolkit course*, Carnegie Mellon University, September 2013. Scribe: Misha Lavrov. http://www.cs.cmu.edu/~odonnell/toolkit13/lecture01.pdf

[۶] Prime number theorem. In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Prime_number_theorem

[۷] Stirling's approximation. In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Stirling's_approximation

[۸] Taylor's theorem. In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Taylor's_theorem#Estimates_for_the_remainder

[۹] Robert M. Young. 75.9 euler's constant. *The Mathematical Gazette*, 75 (472): pp. 187–190, 1991.