تحلیل الگوریتم پیدا کردن ماکزیمم

امید اعتصامی
پژوهشگاه دانشهای بنیادی، پژوهشکده ریاضیات

۱ انگیزه

در تحلیل الگوریتم‌ها، تحلیل احتمالاتی الگوریتم‌ها روشی برای تخمین پیچیدگی محاسباتی یک الگوریتم یا مساله‌ی محاسباتی در حالت متوسط [۱] می‌باشد. در این نوع تحلیل ما درباره‌ی توزیع احتمالاتی ورودی الگوریتم فرض‌هایی می‌کنیم و سپس تحلیل را بر پایه‌ی این فرض انجام می‌دهیم. این با تحلیل بدترین حالت [۲] به وضوح متفاوت است.

در تحلیل احتمالاتی خود الگوریتم غیرتصادفی است و تنها ورودی تصادفی است. این تحلیل کمی با تحلیل الگوریتم‌های تصادفی فرق می‌کند: در آنجا ما اجرای تصادفی الگوریتم را روی یک ورودی غیرتصادفی تحلیل می‌کنیم.

۲ الگوریتم پیدا کردن ماکزیمم

در الگوریتم ساده‌ی زیر، ماکزیمم مقادیر $A[1]$، ...، $A[n]$ محاسبه شده و مقدار آن $A[j]$ در متغیر m ذخیره می‌گردد.

$m \leftarrow -\infty$
for $i \leftarrow 1$ to n **do**
  if $A[i] > m$ **then**
    $j \leftarrow i$
    $m \leftarrow A[i]$
  end if
end for

۳ مساله‌ی استخدام

زمان اجرای الگوریتم پیدا کردن ماکزیمم در بهترین و بدترین حالت $\Theta(n)$ می‌باشد، پس به نظر می‌رسد که چیز خاصی برای تحلیل وجود ندارد. اما X را برابر تعداد بار تغییر دادن مقدار m تعریف کنید. مقدار X حداقل ۱ و حداکثر n می‌باشد. می‌خواهیم استدلال کنیم که برای حداقل یک مساله، تحلیل مقدار X مفید است.

مساله‌ی مورد نظر مساله‌ی استخدام منشی[۳] است: فرض کنید می‌خواهیم از بین n متقاضی برای شغل منشی‌گری یکی را انتخاب کنیم. ما متقاضیان را یکی یکی مصاحبه می‌کنیم. فرض کنید در انتهای مصاحبه هر متقاضی، اگر ما خواهان آن متقاضی هستیم باید دقیقا پس از مصاحبه این را به متقاضی اعلام کنیم. در انتهای مصاحبه‌ی هر متقاضی، اگر آن متقاضی بهتر از تمام متقاضیان مصاحبه شده‌ی قبلی بود ما خواهان آن متقاضی هستیم. در این روش استخدام، بهترین متقاضی به عنوان منشی انتخاب می‌شود، اما ما مجبور هستیم که X بار پیشنهاد کار دهیم و $X-1$ بار نظرمان را عوض کنیم. چون پیشنهاد و عوض کردن نظر هزینه‌ی زیادی دارد، دانستن مقدار X مهم است.

۴ توزیع ورودی

همانطور که گفتیم مقدار X در بهترین حالت (وقتی $A[1]$ ماکزیمم باشد) برابر ۱، و در بدترین حالت (وقتی آرایه‌ی A صعودی باشد) برابر n می‌باشد. برای ساده‌سازی فرض می‌کنیم مقادیر $A[1]$، ...، $A[n]$ همگی متمایز باشند. برای به دست آوردن مقدار متوسط X، باید روی توزیع

[۱] average-case
[۲] worst-case analysis
[۳] hiring problem,secretary problem

ورودی فرض‌هایی داشته باشیم. فرضی که ما می‌کنیم این است که تمام $n!$ جایگشت از مجموعه‌ی $\{A[1], \ldots, A[n]\}$ دارای احتمال برابر برای دنباله‌ی $A[1]$، ...، $A[n]$ هستند.

۵ الگوریتم تصادفی

فرض کنید $A[1]$، ...، $A[n]$ همگی متمایز باشند، اما بعضی جایگشت‌ها از $\{A[1], \ldots, A[n]\}$ دارای احتمال بیشتری باشند. ما می‌توانیم با اجرای الگوریتم تصادفی زیر که یک جایگشت کاملا تصادفی تولید می‌کند، مطمئن باشیم تمام جایگشت‌ها به احتمال مساوی ظاهر می‌شوند:

for $i \leftarrow 1$ to n **do**
  set k = 1, 2, …, or i uniformly at random
  swap $A[k]$ with $A[i]$
end for

۶ متغیر تصادفی نشانگر

گزاره‌ی S_i را تعریف کنید «$A[i]$ ماکزیمم $A[1]$، ...، $A[i]$ می‌باشد.» متغیر تصادفی X_i را برابر $[S_i]$ تعریف کنید که در این‌جا ما از نمادگذاری ایورسون [۴] استفاده می‌کنیم، یعنی $[S_i] = 1$ هر گاه گزاره‌ی S_i صحیح باشد و $[S_i] = 0$ هرگاه گزاره‌ی S_i غلط باشد. به متغیر X_i که تنها مقادیر صفر و یک می‌پذیرد و این چنین تعریف می‌شود، یک متغیر تصادفی نشانگر [۵] می‌گوییم.

داریم

$$X = X_1 + \ldots + X_n$$

و هم‌چنین

$$\Pr(X_i = 1) = 1/i.$$

۷ خطی بودن امید ریاضی و گاهی واریانس

چون امید ریاضی خطی است داریم

$$\mathbb{E}X = \sum_{i=1}^{n} \mathbb{E}X_i = \sum_{i=1}^{n} \Pr(X_i = 1) = \sum_{i=1}^{n} 1/i = H_n.$$

عدد هارمونیک nام H_n دارای تقریب

$$\ln(n) + \gamma + O(1/n)$$

می‌باشد که در آن γ ثابت اویلر است.

همچنین وقتی تعدادی متغیر تصادفی از هم مستقل باشند، واریانس مجموع این متغیرهای تصادفی برابر مجموع واریانس آنهاست. توجه کنید که متغیرهای X_1، ...، X_n از هم مستقلند، بنابراین

$$\mathrm{var}X = \sum_{i=1}^{n} \mathrm{var}X_i = \sum_{i=1}^{n} \Pr(X_i = 1)(1 - \Pr(X_i = 1)) = \sum_{i=1}^{n} (1/i - 1/i^2) = \ln(n) + O(1).$$

۸ تعداد دورهای یک جایگشت تصادفی

هر جایگشت را می‌توان به طور یکتا به صورت اجتماع چند دور نمایش داد. مثلا

$$(2, 5, 4)(3, 1)$$

---

[۴] Iverson’s notation
[۵] indicator random variable

جایگشتی است که ۲ را به ۵، ۵ را به ۴، ۴ را به ۲، ۳ را به ۱، و ۱ را به ۳ می‌برد. هر دور را با بزرگترین عضوش آغاز کنید، مثلا جایگشت قبلی تبدیل به

$$(5,4,2)(3,1)$$

می‌شود. سپس دورها را بر حسب عضو اولشان به صورت صعودی مرتب کنید، مثلا

$$(3,1)(5,4,2)$$

و سپس پرانتزها را بردارید، مثلا

$$31542$$

با کمی فکر متوجه می‌شوید که در این حالت می‌توان بدون پرانتز هم دورها را تشخیص داد. هم چنین تعداد دورها برابر با مقدار X است وقتی

$$(A[1],\ldots,A[5])=(3,1,5,4,2).$$

بنابراین متوسط و واریانس تعداد دورهای یک جایگشت تصادفی همان متوسط و واریانس محاسبه شده برای X در بخش قبلی است.

۹ تابع مولد احتمال

اگر p_{nk} تعداد جایگشتهای n عضوی باشد به طوری که $X=k$، آنگاه

$$p_{nk}=p_{(n-1)(k-1)}+(n-1)p_{(n-1)k}$$

که در آن جمله‌ی $p_{(n-1)(k-1)}$ مربوط به حالتیست که $A[n]$ عنصر ماکزیمم باشد و $(n-1)p_{(n-1)k}$ مربوط به حالتیست که $A[n]$ عنصر ماکزیمم نباشد. بنابراین

$$\Pr(X=k)=p_{nk}/n!=\frac{1}{n}p_{(n-1)(k-1)}/(n-1)!+\frac{n-1}{n}p_{(n-1)k}/(n-1)!$$

حال می‌خواهیم همان متغیر تصادفی X را با روش تابع مولد احتمال تحلیل کنیم. به طور کلی اگر X متغیر تصادفی با مقادیر صحیح نامنفی باشد، تابع مولد احتمال آن برابر است با

$$G(z)=\sum_{k=0}^{\infty}\Pr(X=k)z^k=\mathbb{E}\,z^X.$$

در مساله‌ی ما که X به n ربط دارد، $G_n(z)$ را تابع مولد برای X تعریف کنید وقتی جایگشت n عضوی باشد. آنگاه بنابر رابطه‌ی بازگشتی بالا، داریم

$$G_n(z)=\frac{z+n-1}{n}G_{n-1}(z),\qquad G_0(z)=1.$$

بنابراین

$$G_n(z)=\prod_{k=1}^{n}(z+k-1)/n!.$$

۱۰ محاسبه‌ی امید ریاضی و واریانس از روی تابع مولد

آنها که با توابع مولد مختلف آشنا هستند شاید بدانند که

$$\prod_{k=1}^{n}(z+k-1)=\sum_{k=1}^{n}\begin{bmatrix}n\\k\end{bmatrix}z^k,$$

که در آن $\left[{n \atop k}\right]$ عدد استرلینگ نوع اول است. اما ما می‌خواهیم امید ریاضی و واریانس را مستقیما از خود تابع مولد بدست بیاوریم. اگر $G(z)$ تابع مولد برای متغیر X باشد، آنگاه

$$\mathbb{E}X = G'(1), \qquad \operatorname{var}(X) = G''(1) + G'(1) - G'(1)^2.$$

می‌توان به سادگی دید که اگر تابع مولد متغیرهای X، Y، و Z به ترتیب $G(z)$، $H(z)$، و $G(z)H(z)$ باشد آنگاه

$$\mathbb{E}Z = (\mathbb{E}X) + (\mathbb{E}Y), \quad \operatorname{var}(Z) = \operatorname{var}(X) + \operatorname{var}(Y).$$

چون متوسط و واریانس متغیر با تابع مولد $(z+k-1)/k$ برابر $1/k$ و $1/k - 1/k^2$ می‌باشد، داریم

$$\mathbb{E}X = \sum_{k=1}^{n} 1/k, \qquad \operatorname{var}(X) = \sum_{k=1}^{n} (1/k - 1/k^2).$$

۱۱ گشتاورهای بالاتر X

در این قسمت ما گشتاور rام متغیر تصادفی X را به ازای $r \geq 1$ تخمین می‌زنیم.

قضیه ۱ به ازای $r \geq 1$ صحیح و ثابت داریم

$$\mathbb{E}X^r = (\ln n)^r (1 + o_n(1)).$$

اثبات: با استقرا روی r اثبات می‌کنیم. داریم

$$X = X_1 + \ldots + X_n.$$

بنابراین

$$\mathbb{E}X^r = \sum_{1 \leq i_1, \ldots, i_r \leq n} \mathbb{E}(X_{i_1} \ldots X_{i_s}).$$

وقتی i_1، ...، i_r از هم متمایز هستند داریم

$$\mathbb{E}(X_{i_1} \ldots X_{i_r}) = \mathbb{E}(X_{i_1}) \ldots \mathbb{E}(X_{i_r})$$

چون X_1، ...، X_n از هم مستقلند. در حالت کلی

$$\mathbb{E}(X_{i_1} \ldots X_{i_r}) \geq \mathbb{E}(X_{i_1}) \ldots \mathbb{E}(X_{i_r}).$$

بنابراین یک کران پایین برای $\mathbb{E}X^r$ عبارتست از

$$\sum_{1 \leq i_1, \ldots, i_r \leq n} \mathbb{E}(X_{i_1}) \ldots \mathbb{E}(X_{i_r}) = \left(\sum_{i=1}^{n} \mathbb{E}X_i\right)^r = H_n^r.$$

میزان اختلاف این کران با مقدار واقعی گشتاور حداکثر برابر است با

$$\sum_{\substack{1 \leq i_1, \ldots, i_r \leq n \\ i_1, \ldots, i_r \text{not distinct}}} \mathbb{E}(X_{i_1} \ldots X_{i_r}) \leq \binom{r}{2} \mathbb{E}X^{r-1}.$$

چون اگر مثلا $i_1 = i_2$ آنگاه

$$X_{i_1} X_{i_2} \ldots X_{i_s} = X_{i_2} \ldots X_{i_s}.$$

طبق فرض استقرا

$$\mathbb{E}X^{r-1} = O(\ln^{r-1} n).$$

بنابراین

$$\mathbb{E}X^r = H_n^r + \binom{r}{2} O(\ln^{r-1} n) = (1+o(1)) \ln^r(n).$$

$\square$

توزیع متغیر تصادفی X در [۱، ۲] مطالعه شده است.

مراجع

[1] F. G. Foster and A. Stuart (1954). Distribution-Free Tests in Time-Series Based on the Breaking of Records. Journal of the Royal Statistical Society. Series B (Methodological) 1-22.

[2] P. B. M. Roes (1966). A note on linear programming algorithm design: a combinatorial problem. Communications of the ACM, 9, page 342.