نابرابریهای دُمی از طریق مارتینگیل

امید اعتصامی
پژوهشگاه دانشهای بنیادی، پژوهشکده ریاضیات

۱ انگیزه

با وجود این که تعریف مارتینگیل بر اساس برابری امید ریاضی‌هاست، می‌توان از طریق مارتینگیل‌ها نابرابریهای دُمی [1] اثبات کرد، یعنی می‌توان برای برخی متغیرهای تصادفی نشان داد که احتمال اینکه خیلی بزرگ یا خیلی کوچک باشند کم است (یا به عبارتی دُم توزیع احتمالاتی این متغیرها کوچک است). این نابرابری‌ها برای تحلیل الگوریتم‌ها مفیدند.

۲ تعریف زیرمارتینگیل و فوق‌مارتینگیل

دنباله‌ی

$$\langle Z_n \rangle = Z_0, Z_1, Z_2, \ldots$$

از متغیرهای تصادفی را زیرمارتینگیل می‌نامیم اگر برای هر $n \geq 0$

$$\mathbb{E}(Z_{n+1}|Z_0, \ldots, Z_n) \geq Z_n.$$

آن را فوق‌مارتینگیل می‌نامیم اگر برای هر $n \geq 0$

$$\mathbb{E}(Z_{n+1}|Z_0, \ldots, Z_n) \geq Z_n.$$

یک مارتینگیل همزمان زیرمارتینگیل و فوق‌مارتینگیل است. اگر $\langle Z_n \rangle$ زیرمارتینگیل باشد داریم

$$\mathbb{E}Z_0 \leq \mathbb{E}Z_1 \leq \mathbb{E}Z_2 \leq \ldots$$

و اگر فوق‌مارتینگیل باشد داریم

$$\mathbb{E}Z_0 \geq \mathbb{E}Z_1 \geq \mathbb{E}Z_2 \geq \ldots.$$

۳ روش ساخت زیرمارتینگیل

قضیه ۱ اگر $\langle Z_n \rangle$ مارتینگیل و f تابعی محدب باشد، آنگاه $f(Z_n)$ زیرمارتینگیل است.

اثبات: چون $\mathbb{E}(Z_{n+1}|Z_0, \ldots, Z_n) = Z_n$، طبق نابرابری ینسن داریم

$$\mathbb{E}(f(Z_{n+1})|Z_0, \ldots, Z_n) \geq f(Z_n).$$

بنابراین

$$\mathbb{E}(f(Z_{n+1})|f(Z_0), \ldots, f(Z_n)) \geq f(Z_n).$$

□

پس اگر $\langle Z_n \rangle$ مارتینگیل باشد، آنگاه $\langle |Z_n| \rangle$ و $\langle \max(Z_n, c) \rangle$ و $\langle Z_n^2 \rangle$ و $\langle e^{Z_n} \rangle$ همگی زیرمارتینگیل هستند. اگر همچنین $Z_n \geq 0$ آنگاه $\langle 1/Z_n \rangle$ و $\langle \ln(1/Z_n) \rangle$ و $\langle Z_n \ln Z_n \rangle$ هم زیرمارتینگیل هستند.

[1] tail inequalities

۴ توقف زیرمارتینگیل

اگر زیرمارتینگیل $\langle Z_n \rangle$ را در زمان N متوقف می‌کنیم، دنباله‌ی متغیرهای تصادفی به دست آمده $\langle Z'_n \rangle$ هم زیرمارتینگیل است. اگر N دارای کران بالای m باشد، آنگاه

$$\mathbb{E}Z'_m = \mathbb{E}Z_N \leq \mathbb{E}Z_m.$$

۵ نابرابری ماکزیمال

قضیه ۲ اگر $\langle Z_n \rangle$ زیرمارتینگیل نامنفی باشد و $x > 0$، آنگاه

$$\Pr(\max(Z_0, \ldots, Z_n) \geq x) \leq \mathbb{E}Z_n / x.$$

اثبات: اگر $\max(Z_0, \ldots, Z_n) \geq x$ آنگاه N را برابر کوچکترین i قراردهید به‌طوری که $Z_i \geq x$، و در غیر این صورت N را برابر n قرار دهید. به وضوح N یک قاعده‌ی توقف برای دنباله‌ی $\langle Z_n \rangle$ است. بنابر نابرابری مارکوف داریم

$$\mathbb{E}Z_n \geq \mathbb{E}Z_N \geq x \Pr(\max(Z_0, \ldots, Z_n) \geq x).$$

$\square$

به عنوان مثال اگر $\langle Z_n \rangle$ مارتینگیل باشد و $x > 0$، آنگاه

$$\Pr(\max(|Z_0|, \ldots, |Z_n|) \geq x) \leq \mathbb{E}|Z_n| / x.$$

۶ نابرابری کولموگوروف

به عنوان نتیجه‌ی نابرابری ماکزیمال، اگر $\langle Z_n \rangle$ مارتینگیل باشد و $x > 0$، آنگاه

$$\Pr(\max(Z_0^2, \ldots, Z_n^2) \geq x) \leq \mathbb{E}Z_n^2 / x.$$

کولموگورف [۵] حالت خاص این نابرابری را وقتی $Z_n = Y_1 + \ldots + Y_n$ که در آن Y_iها متغیرهای تصادفی مستقل با متوسط ۰ هستند، اثبات کرد: اگر $\mathrm{var} Z_n = \sigma^2$، آنگاه

$$\Pr(|Y_1| < t\sigma, |Y_1 + Y_2| < t\sigma, \ldots, |Y_1 + \ldots + Y_n| < t\sigma) \geq 1 - 1/t^2.$$

این در حالیست که نابرابری چبیشف تنها

$$\Pr(|Y_1 + \ldots + Y_n| < t\sigma) \geq 1 - 1/t^2$$

را می‌دهد.

۷ یک لم تکنیکی

برای اثبات نابرابری هفدینگ-آزوما به لم زیر نیاز داریم.

لم ۳ اگر $t \geq 0$ و $-p \leq y \leq q$ که در آن $0 \leq p \leq 1$ و $q = 1 - p$، آنگاه تابع h وجود دارد به طوری که

$$e^{yt} \leq e^{t^2/8} + yh(t).$$

اثبات: به ازای $x \in [0,1]$ داریم

$$e^{xt} - xe^t - 1 + x \geq 0$$

چون این نابرابری به ازای $t = 0$ برقرار است و

$$\frac{d}{dt}(e^{xt} - xe^t - 1 + x) = xe^{xt} - xe^t \geq 0.$$

بنابراین

$$e^{(y+p)t} - (y+p)e^t - 1 + (y+p) \geq 0.$$

پس

$$e^{yt} \leq e^{-pt}(q + pe^t) + ye^{-pt}(e^t - 1).$$

قرار دهید $h(t) = e^{-pt}(e^t - 1)$ و $f(t) = -pt + \ln(q + pe^t)$. فرمول تیلور (با فرم انتگرالی برای باقیمانده) برای f حول $t = 0$، به ما می‌گوید که

$$f(t) = f(0) + f'(0)t + \int_0^t \frac{x^1}{1!} f''(t-x)dx$$

که در آن

$$f(0) = 0, \quad f'(0) = 0$$

و

$$f''(t) = \frac{pqe^t}{(q + pe^t)^2} \leq 1/4$$

چون میانگین حسابی q و pe^t از میانگین هندسی‌شان کمتر نیست. بنابراین

$$f(t) \leq 0 + 0 + \int_0^t (x/4)dx \leq t^2/8.$$

□

۸ نابرابری هفدینگ-آزوما

برای درک بهتر صورت قضیه‌ی هفدینگ-آزوما [۱،۴]، به یاد آورید که هر مارتینگیل $\langle Z_n \rangle$ را می‌توان به صورت $Z_n = Y_0 + \ldots + Y_n$ نوشت که در آن $\langle Y_n \rangle$ یک دنباله‌ی منصفانه است.

قضیه ۴ اگر $\langle Y_n \rangle$ یک دنباله‌ی منصفانه باشد به طوری که $a_n \leq Y_n \leq b_n$، آنگاه به ازای $x \geq 0$، داریم

$$\Pr(Y_1 + \ldots + Y_n \geq x) \leq e^{-2x^2/((b_1-a_1)^2 + \ldots + (b_n-a_n)^2)}.$$

اثبات: با توجه به لم ۳ به ازای $t \geq 0$ داریم

$$e^{tY_i} = e^{(b_i-a_i)t(Y_i/(b_i-a_i))} \leq e^{(b_i-a_i)^2t^2/8} + \frac{Y_i}{b_i - a_i}h(t(b_i - a_i)).$$

طبق خاصیت منصفانه بودن

$$\mathbb{E}(e^{tY_i}|Y_1, \ldots, Y_{i-1}) \leq e^{(b_i-a_i)^2t^2/8}.$$

بنابراین

$$\mathbb{E}e^{t(Y_1+\ldots+Y_n)} \leq e^{t^2\sum_{i=1}^n (b_i-a_i)^2}.$$

طبق قضیه‌ی مارکوف داریم

$$\Pr(Y_1 + \ldots + Y_n \geq x) \leq \mathbb{E}e^{t(Y_1+\ldots+Y_n)}/e^{xt}.$$

اگر بهترین t انتخاب کنیم، آنگاه حکم قضیه اثبات می‌شود. □

در قضیه‌ی فوق چون $-b_n \leq -Y_n \leq -a_n$، بنابراین

$$\Pr(Y_1 + \ldots + Y_n \leq -x) \leq e^{-2x^2/((b_1-a_1)^2+\ldots+(b_n-a_n)^2)}.$$

پس

$$\Pr(|Y_1 + \ldots + Y_n| \geq x) \leq 2e^{-2x^2/((b_1-a_1)^2+\ldots+(b_n-a_n)^2)}.$$

همچنین در اثبات نابرابری هفدینگ-آزوما می‌توان دید که $b_n - a_n$ باید تابعی از n باشد، ولی a_n و b_n می‌توانند به Y_1،...،Y_{n-1} بستگی داشته باشند.

۹ مارتینگیل دوب

فرض کنید $\langle X_n \rangle$ دنباله‌ای از متغیرهای تصادفی و Q متغیر تصادفی دیگری باشد. آنگاه

$$Z_n = \mathbb{E}(Q|X_0, \ldots, X_n)$$

مارتینگیل است (و به آن مارتینگیل دوب [۲، ۳] می‌گوییم.)

۱۰ الگوریتم درهم‌سازی

یک الگوریتم درهم‌سازی[۲] در نظر بگیرید که t عنصر را در m لیست می‌گذارد به طوری که عنصر nام در لیست X_nام قرار می‌گیرد. فرض کنید X_n هر کدام مستقلا یکی از مقادیر 1، ...، m را به طور یکنواخت می‌پذیرد. متغیر تصادفی Q را تعداد لیست‌های خالی پس از گذاردن تمام t عنصر تعریف کنید. می‌دانیم که

$$\mathbb{E}Q = m(\frac{m-1}{m})^t.$$

Q چه مقدار از مقدار متوسطش فاصله دارد؟ اگر Z_0،...،Z_t را مارتینگیل دوب متناظر با متغیر تصادفی Q و دنباله‌ی متغیرهای X_1،...،X_t در نظر بگیرید، آنگاه $Z_0 = \mathbb{E}Q$ و $Z_n = Q$.

اگر تنها یک متغیر از متغیرهای X_1،...،X_t را عوض کنیم آنگاه Q حداکثر به اندازه‌ی ۱ تغییر می‌کند، بنابراین به ازای هر q_1، ...، q_{n-1} وجود دارد a_n و b_n با $b_n - a_n \leq 1$ به طوری که وقتی به ازای $1 \leq i \leq n-1$ داشته باشیم $Q_i = q_i$ آنگاه

$$a_i \leq Q_i - Q_{i-1} \leq b_i.$$

پس اگر نابرابری هفدینگ-آزوما را روی این مارتینگیل دوب اعمال کنیم، داریم

$$\Pr(|Q - m(\frac{m-1}{m})^t| \geq \sqrt{t\ln(f(t))} \leq 2/f(t)^2$$

هرگاه $f(t) > 0$.

۱۱ بسته بندی سطلی

برای تعریف مساله‌ی بسته بندی سطلی [۳]، همانند مساله‌ی درهم‌سازی فرض کنید X_1،...،X_t متغیرهای تصادفی صحیح مستقل بین ۱ و m هستند، ولی این بار اجازه می‌دهیم هر یک از این t متغیر توزیع خود را داشته باشد. تصور کنید که t شیء داریم به طوری که حجم شیء nام برابر X_n است.

Q را تعریف کنید کمترین تعداد سطل هر کدام به حجم m به طوری که بتوان t شیء را در این Q سطل جاسازی داد به طوری که مجموع حجم اشیا درون هر سطل حداکثر m باشد (و هیچ شیء را به چند قسمت تقسیم نمی‌کنیم.) تحلیل Q برای مساله‌ی بسته‌بندی سطلی سختتر از

[۲] hashing algorithm
[۳] bin packing

مساله‌ی قبلی درهم‌سازی است: مثلا محاسبه‌ی Q بر حسب X_nها NP-کامل [۴] است. اما هم‌چنان اگر تنها یک متغیر از متغیرهای X_1،...،X_t را عوض کنیم آنگاه Q حداکثر به اندازه‌ی ۱ تغییر می‌کند، بنابراین طبق استدلال قبل

$$\Pr(|Q - \mathbb{E}Q| \geq \sqrt{t \ln(f(t))}) \leq 2/f(t)^2.$$

این در حالیست که ما مقدار $\mathbb{E}Q$ را نمی‌دانیم.

وقتی X_1، ...، X_t همگی توزیع یکسانی دارند و ما تعریف کنیم

$$\beta_t = Q(X_1, \ldots, X_t)$$

آنگاه

$$\beta_{t+t'} \leq \beta_t + \beta_{t'}$$

چون می‌توان t و t' شیء را به صورت مجزا بسته بندی کرد. بنابراین طبق لم فکته [۵] β_t/t دارای حد است. اما چون محاسبه Q از روی X_1،...،X_t سخت است، نمی‌توان مقدار این حد را با تکرار جاگذاری مقادیر تصادفی برای X_1،...،X_t و میانگین گرفتن از Q متناظر راحت بدست آورد.

مراجع

[1] Kazuoki Azuma (1967). Weighted Sums of Certain Dependent Random Variables. Tôhoku Mathematical Journal 19 (3): 357–367.

[2] Joseph L. Doob (1953). Stochastic Processes, New York: Wiley.

[3] Joseph L. Doob (1940). Regularity properties of certain families of chance variables. Transactions of the American Mathematical Society 47, 486.

[4] Wassily Hoeffding (1963). Probability inequalities for sums of bounded random variables. Journal of the American Statistical Association 58 (301): 13–30.

[5] Andrei Nikolaevich Kolmogorov (1928). Über die Summen durch den Zufall bestimmter unabhängiger Größen. Mathematische Annalen 99, 309–319.

[۴] NP-complete

[۵] Fekete's lemma