

معرفی روش انقباض

رامین کاظمی^۱

تاریخ دریافت: ۱۳۹۸/۱/۲۸

تاریخ پذیرش: ۱۳۹۸/۷/۲۳

چکیده:

هدف این مقاله، معرفی روش انقباض برای تحلیل الگوریتم‌ها است. بر اساس این روش، چندین رده از روابط بازگشتی می‌توانند به‌عنوان حالت‌های خاص چارچوب کلی بیان شده تحلیل شوند. گام‌های اصلی این فن بر اساس ویژگی‌های انقباض الگوریتم نسبت به مترهای احتمالی مناسب پایه‌ریزی می‌شوند. نوعاً توزیع حدی به‌عنوان نقطه ثابت یک عملگر حدی روی رده توزیع‌های احتمال مشخص‌سازی می‌شود.

واژه‌های کلیدی: الگوریتم‌های بازگشتی، متر احتمال، عملگر حدی، نقطه ثابت، توزیع حدی.

۱ مقدمه

L_p برای هر $p > 0$ نتیجه می‌شود.

روش دیگری موسوم به روش انقباض وجود دارد. اما روش انقباض چیست؟ روش انقباض یکی از مفیدترین روش‌ها برای تحلیل مجانبی الگوریتم‌ها است. این روش تا حدودی شبیه به روش گشتاوری است. روش انقباض برای تحلیل مجانبی یک الگوریتم بازگشتی شامل مراحل زیر است: ابتدا با مطالعه اولین گشتاورها مجبور به نرمال‌سازی صحیح الگوریتم هستیم. نسخه نرمال شده مجدداً در یک شکل بازگشتی ارائه می‌شود. سپس اغلب یک معادله نقطه ثابت برای معادله حدی الگوریتم حدس زده می‌شود. این معادله نقطه ثابت حدی توسط یک معادله به شکل $K\mu = \mu$ برای یک عملگر حدی K روی مجموعه همه توزیع‌های احتمال بیان می‌شود. در گام بعدی، مجبور به اثبات وجود یک حل از این معادله نقطه ثابت هستیم. این کار توسط انتخاب یک متر احتمال مناسب روی فضای اندازه‌های احتمال انجام می‌شود به‌طوری‌که K یک نگاشت انقباضی نسبت به این متر است و قضیه نقطه ثابت باناخ به‌کار گرفته می‌شود. در گام آخر، مجبور به اثبات همگرایی الگوریتم (در توزیع) به این نقطه ثابت هستیم. این کار نیز با معرفی دنباله‌های همراه توزیع‌ها که الگوریتم را همانندسازی و همچنین ساختار معادله حدی را

روش‌های اصلی برای تحلیل مجانبی الگوریتم‌ها روش‌های تبدیل (به‌ویژه، توابع مولد گشتاور) [۳]، روش مارتینگلی [۴] و روش فرایندهای شاخه‌ای [۲] از طریق تقریب به‌وسیله برون‌گشت‌های براونی هستند. برای یک رده محدودتر از الگوریتم‌های تصادفی روش پایه‌ریزی شده روی تقریب‌های تصادفی نیز در این میان قرار می‌گیرند. روش گشتاورها، یکی از کلاسیک‌ترین روش‌های دستیابی توزیع حدی، در مسائل گوناگون در حوزه‌های مختلف است [۱۰]. این روش شامل گام‌های زیر است:

- ۱- میانگین و واریانس محاسبه می‌شوند.
- ۲- متغیر تصادفی به‌طور سَره‌ای مقیاس‌ده می‌شود.
- ۳- با استفاده از استقراء گشتاورهای بالاتر متغیر تصادفی مقیاس‌ده محاسبه می‌شوند.
- ۴- معیار کارلمان برای تصدیق یکتایی قاعده حدی به‌کار گرفته می‌شود.
- ۵- بنا بر قضیه همگرایی گشتاوری فرشه-شوهات، همگرایی در توزیع و همگرایی در همه گشتاورها (یا همگرایی در هیأت علمی گروه آمار، دانشگاه بین‌المللی امام خمینی، قزوین، ایران

منعکس می‌کنند، انجام می‌شود.

می‌شود. امید ریاضی در معادله بازگشتی

$$a_n = \sum_{i \in \mathbb{N}} \mathbb{E}(T_{n,i}) a_{Z_{n,i}} + c_n \quad (2)$$

صدق می‌کند به‌طوری‌که $n_* \leq n$ و $c_n = \mathbb{E}(C_n)$. شکل استاندارد معادله به‌صورت

$$a_n = \sum_{j=0}^{n-1} \nu_n(j) a_j + c_n \quad (3)$$

است. در این جا ν_n یک اندازه علامت‌دار روی مجموعه $\{0, \dots, n-1\}$ و نه لزوماً یک اندازه احتمال است. در بیشتر مواقع، به رفتار مجانبی a_n در n علاقه‌مند هستیم. در مثال‌های استاندارد، نه جایی که $a_n \rightarrow 0$ ، a_n از مرتبه

$$n \ln n + (\text{ثابت}) n + o(n) \quad (\text{ثابت})$$

یا مرتبه

$$n + o(n) \quad (\text{ثابت})$$

و یا از مرتبه $\ln n$ است.

روش‌های متعدد و فن‌های گوناگونی از جمله یک برهان با یک حدس، توابع مولد و چند معادله جبری، معادلات دیفرانسیلی و معادله تفاضلی در حد، تفسیر احتمالی از طریق فرایندهای تجدید یا مارکف با استفاده از مسئله دریکله و کوشی وجود دارند. یک روش کلی از نظریه تجدید برای این معادلات برخاسته از تحلیل الگوریتم‌ها در روسلا [۱۱] وجود دارد.

اگر ν_n تنها جرمی در نقطه $n-1$ داشته باشد، آنگاه معادله

$$a_n = \nu_n(n-1) a_{n-1} + c_n \quad (4)$$

ساده‌تر (با تکرار یا از طریق معادلات تفاضلی یا معادلات دیفرانسیل حدی) حل خواهد شد.

گام دوم (نرمال کردن): نرمال کردن درست حدس زده می‌شود به‌طوری‌که نسخه نرمال شده Y_n از X_n به یک حد W همگرا باشد. به نسخه

$$Y_n := \frac{X_n - a_n}{b_n}$$

برای $b_n > 0$ توجه می‌کنیم. قاعده کلی و دقیقی برای یک انتخاب مناسب b_n نمی‌توان ارائه کرد. یک روش، انتخاب b_n به اندازه ممکن کوچک است به‌طوری‌که نسخه نرمال شده C_n^* از

اولین مثال که در آن از روش انقباض استفاده شد در مقاله نوشته شده توسط روسلا و در مطالعه الگوریتم مرتب‌سازی سریع بود [۱۱]. در شکل کلی، این روش به‌طور مستقل توسط روسلا [۱۲] و راشف و روشندورف [۱۰] بهبود داده شده است. بعد از این، در بسیاری از مطالعه‌های این روش بر روش‌های احتمالی و روش‌های تحلیلی و ترکیبیاتی ترجیح داده شده است. یک موضوع کلاسیک نظریه احتمال، قضیه حدی مرکزی برای دنباله‌هایی از مجموع‌هایی از متغیرهای تصادفی وابسته است. این دنباله‌های تجمعی یک ساختار بازگشتی ساده دارند. مترهای احتمال و فن‌های انقباض در این حالت، به‌ویژه در مدرسه روسی احتمال برای اثبات قضیه حدی مرکزی معرفی شده‌اند. در این مدرسه، مترهای احتمال مفید، ویژگی‌های نظم آنها، و ارتباط بین آنها به‌طور جدی مطالعه شده است (برای اطلاعات بیشتر به [۹، ۱۰، ۱۵] مراجعه کنید). این ویژگی‌های نظم در تحلیل الگوریتم‌های بازگشتی کلی مفید هستند.

۲ گام‌های ترتیبی

فرض کنید $(X_n)_{n \in \mathbb{N}}$ دنباله‌ای از متغیرهای تصادفی باشد که برای هر n ، $n \geq n_*$ ، در برابری توزیعی

$$X_n \stackrel{d}{=} \sum_{i \in \mathbb{N}} T_{n,i} X_{Z_{n,i}}^{(i)} + C_n \quad (1)$$

صدق می‌کند. عدد n_* یک عدد طبیعی و متغیرهای تصادفی (C_n, Z_n, T_n) و $X_j^{(i)} \in \mathbb{N}$ برای $0 \leq j \leq n-1$ مستقل هستند.

از نمادگذاری $Z_n = (Z_{n,1}, Z_{n,2}, \dots)$ و $T_n = (T_{n,1}, T_{n,2}, \dots)$ استفاده می‌کنیم. متغیرهای تصادفی $Z_{n,i}$ مقادیر خود را در مجموعه $\{0, 1, \dots, n-1\}$ اتخاذ می‌کنند. توزیع X_j برای $0 \leq j \leq n_* - 1$ داده می‌شود و توزیع $X_j^{(i)}$ همان توزیع X_j است. همچنین، قرار می‌دهیم $X_* \equiv 0$. تحلیل X_n در گام‌های ترتیبی زیر صورت می‌گیرد.

گام اول (امید ریاضی): امید ریاضی $a_n := \mathbb{E}(X_n)$ محاسبه

داریم. این جا (C, T) و W_i برای $i \in \mathbb{N}$ مستقل هستند و همه W_i ها همان توزیع W را دارند. توزیع W باید یک حل معادله نقطه ثابت بالا باشد. برخی اوقات، این معادله نقطه ثابت، توزیع حدی را مشخص می کند. حال، مایل به فرمول بندی این معادله بر حسب توزیع ها و یک عملگر هستیم. عملگر K را روی یک زیرمجموعه M از اندازه های احتمال به خودش از طریق

$$K(\mu) := \mathcal{L}\left(\sum_{i \in \mathbb{N}} T_i Y_i + C\right) \quad (6)$$

تعریف می کنیم. این جا (C, T) و Y_i برای هر $i \in \mathbb{N}$ مستقل هستند و Y_i ها دارای توزیع μ هستند، و توزیع (C, T) از معادله حدی می آید. عملگر K را عملگر حدی می نامیم. یک نقطه ثابت K یک اندازه احتمال ν است که در شرط $K(\nu) = \nu$ صدق می کند. این شرط معادل با

$$W \stackrel{d}{=} \sum_{i \in \mathbb{N}} T_i W_i + C \quad (7)$$

است. توزیع W یک نقطه ثابت است. معادله (۷) را معادله نقطه ثابت و حل آن را یک نقطه ثابت گوئیم. گام چهارم (روش انقباض): این روش ابزاری برای اثبات همگرایی دنباله نرمال شده Y_n به W است؛ ایده پایه یک بحث انقباض روی فضای توزیع ها است. نگاشت های K_n را برای هر $n \geq 0$ روی M^n به M از طریق

$$K_n(\mu_0, \dots, \mu_{n-1}) := \mathcal{L}\left(\sum_{i \in \mathbb{N}} T_{n,i}^* Y_{Z_{n,i}}^{(i)} + C_n^*\right) \quad (8)$$

تعریف کنید. در این جا (C_n^*, Z_n, T_n^*) و متغیرهای $Y_j^{(i)}$ برای $i \in \mathbb{N}$ و $0 \leq j \leq n$ مستقل هستند، $Y_j^{(i)}$ همان توزیع μ_j را دارد و توزیع (C_n^*, Z_n, T_n^*) به صورت بالا داده می شود. با شروع از اندازه های احتمال $\mu_0, \dots, \mu_{n-1}$ ، یک دنباله به صورت

$$\begin{aligned} \mu_{n_0} &= K_{n_0}(\mu_0, \dots, \mu_{n_0-1}), \\ \mu_{n_0+1} &= K_{n_0+1}(\mu_0, \dots, \mu_{n_0}), \\ &\vdots \\ \mu_n &= K_n(\mu_0, \dots, \mu_{n-1}) \end{aligned}$$

به دست می آید. توزیع μ_n همان توزیع Y_n است هرگاه $\mu_0, \dots, \mu_{n-1}, X_{n-1}, \dots, X_0$ باشند.

C_n تعریف شده در زیر به یک حد و در صورت امکان به یک حد ناتهیده همگرا باشد. اگر واریانس متناهی باشد و قابل محاسبه، در این صورت $b_n = \sqrt{\text{var}(X_n)}$ گزینه مناسبی است. مثال های نوعی b_n از مرتبه n یا $\sqrt{n}$ یا $\ln n$ است. معادله پایه (۱) به صورت

$$\begin{aligned} Y_n &\stackrel{d}{=} \sum_{i \in \mathbb{N}} T_{n,i} \frac{b_{Z_{n,i}}}{b_n} Y_{Z_{n,i}}^{(i)} \\ &+ \sum_{i \in \mathbb{N}} \frac{T_{n,i} a_{Z_{n,i}} - \mathbb{E}(T_{n,i}) a_{Z_{n,i}}}{b_n} + C_n - \mathbb{E}(C_n) \end{aligned} \quad (5)$$

قابل بازنویسی است. این معادله را معادله پایه ای نرمال شده گوئید. این معادله ساختاری مشابه با معادله (۱) با $T_{n,i}$ و C_n جایگزین شده توسط

$$\begin{aligned} T_{n,i}^* &:= T_{n,i} \frac{b_{Z_{n,i}}}{b_n}, \\ C_n^* &:= \sum_{i \in \mathbb{N}} \frac{T_{n,i} a_{Z_{n,i}} - \mathbb{E}(T_{n,i}) a_{Z_{n,i}}}{b_n} + C_n - \mathbb{E}(C_n) \end{aligned}$$

دارد. اختلاف اصلی با معادله (۱) این است که همه Y_i ها دارای میانگین صفر هستند.

گام سوم (توزیع حدی). یک توزیع حدی W از Y_n حدس زده می شود. در بسیاری از وضعیت ها، توزیع حدی می تواند به عنوان نقطه ثابت یک عملگر K حدس زده شود. قضیه نقطه ثابت نقش مهمی در کاربردهای شاخه های مختلف ریاضی بازی می کند. اصطلاح قضیه نقطه ثابت متریک به آن دسته از نتایج نظری نقطه ثابت اشاره دارد که در شرایط هندسی روی فضا های اساسی و نگاشت ها نقش حیاتی دارند. باناخ ریاضی دان لهستانی، نتیجه ای بسیار مهم در مورد نگاشت های انقباضی ثابت کرد که به عنوان اصل انقباض باناخ شناخته می شود. این قضیه فنی برای حل مسائل کاربردی گوناگون در علوم ریاضی و مهندسی است.

در معادله نرمال شده (۵)، اگر Y_n به W ، $Z_{n,i}$ به ∞ و همچنین (C_n^*, T_n^*) به (C, T) همگرا باشد، آنگاه یک معادله حدی به صورت

$$\begin{aligned} Y_n &\stackrel{d}{=} \sum_{i \in \mathbb{N}} T_{n,i}^* Y_{Z_{n,i}}^{(i)} + C_n^* \\ &\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\ W &\stackrel{d}{=} \sum_{i \in \mathbb{N}} T_i W_i + C, \end{aligned}$$

است. این جا U به طور یکنواخت روی بازه $[0, 1]$ توزیع می شود و F_μ تابع توزیع μ است. متر ℓ_p دارای ویژگی های

$$\begin{aligned}\ell_p(aX, aY) &\leq |a| \ell_p(X, Y), \\ \ell_p\left(\sum_i X_i, \sum_i Y_i\right) &\leq \sum_i \ell_p(X_i, Y_i), \\ \ell_p(AX, AY) &\leq \|A\|_p \ell_p(X, Y)\end{aligned}$$

است که در آن $a \in \mathbb{R}$ و برای هر $i \in \mathbb{N}$ ، (X_i, Y_i) مستقل هستند و A مستقل از (X, Y) است. فرض کنید $\mathcal{M}_p$ فضای همه توزیع های μ با گشتاورهای قدر مطلق متناهی باشد، یعنی

$$\int |x|^p \mu(dx) < \infty.$$

فضای $(\mathcal{M}_p, \ell_p)$ یک فضای متریک تفکیک پذیر کامل است [۱]. در ادامه، به حالت $p = 2$ توجه می کنیم.

قضیه ۱.۳. [۱۰] فرض کنید $\mu, v \in \mathcal{M}_2$ دو اندازه احتمال با

$$\int x \mu(dx) = 0 = \int x v(dx)$$

باشند. در این صورت،

$$\ell_2(K(\mu), K(v)) \leq \sqrt{\mathbb{E}\left(\sum_j T_j^2\right) \ell_2(\mu, v)}. \quad (9)$$

نابرابری (۹) که به لم انقباض مشهور است، نقش اساسی در روش انقباض بازی می کند. قضیه زیر که به قضیه نقطه ثابت باناخ مشهور است، وجود و یکتایی حل معادله نقطه ثابت را تضمین می کند.

قضیه ۲.۳. (قضیه نقطه ثابت باناخ). فرض کنید (E, d) یک فضای متریک کامل باشد. فرض کنید $K : E \rightarrow E$ یک انقباض اکید با ثابت انقباض $k < 1$ باشد. در این صورت، دنباله $K^n(e)$ برای $n \in \mathbb{N}$ به طور نمایی با مرتبه k^n به نقطه ثابت یکتای K برای هر $e \in E$ همگرا است.

قضیه نقطه ثابت باناخ به قضیه بنیادی زیر در فضای متر ℓ_2 منجر می شود.

قضیه ۳.۳. فرض کنید $\mathbb{E}(C^2) < \infty$ ، $\mathbb{E}(C) = 0$ و $\sum_i \mathbb{E}(T_i^2) < 1$ برای هر $n \in \mathbb{N}$

روی یک زیرفضای M از اندازه های احتمال به متر d توجه کنید (برای سادگی از نماد $d(X, Y) := d(\mathcal{L}(X), \mathcal{L}(Y))$ استفاده می شود). فرض کنید v_n دنباله ای از توزیع ها و همگرا به توزیع معلوم v باشد. معمولاً v_n در یک معادله بازگشتی مشابه با μ_n صدق می کند که بر حسب یک عملگر L_n فرمول بندی می شود. دنباله v_n را دنباله همراه می نامند. روش انقباض روشی برای نشان دادن این است که اختلاف بین μ_n و v_n به صفر همگرا است که البته بعد از حدس زدن یک L_n و دنباله v_n اتفاق می افتاد. این ثابت می کند که توزیع Y_n همگرا به توزیع v است. اغلب دنباله v_n یکسان با نقطه ثابت v برای عملگر حدی است. چون با عملگرهای خطی K سروکار داریم مترهای با ویژگی های خاص می توانند بهتر از دیگر مترها برای نشان دادن اجرای روش انقباض مورد استفاده قرار گیرند. به طور مشخص متر واسراشتاین و متر زلتارف از این دسته هستند.

۳ متر واسراشتاین

فرض کنید $f : [a, b] \rightarrow \mathbb{R}$ تابعی تعریف شده روی بازه $[a, b]$ باشد به طوری که انتگرال $\int_a^b |f(x)|^p dx$ متناهی است. برای هر $p \geq 1$ ، p -نرم به صورت

$$\|f\|_p = \left(\int_a^b |f(x)|^p dx \right)^{\frac{1}{p}}$$

تعریف می شود. متر ℓ_p با $1 \leq p < \infty$ که متر مالوس [۱] نیز نامیده می شود به صورت

$$\ell_p(\mu, v) = \inf \{ \|X - Y\|_p \}$$

تعریف می شود. اینفیم روی همه متغیرهای تصادفی X با توزیع μ و متغیرهای تصادفی Y با توزیع v روی فضای احتمال غیراتی گرفته می شود. علامت $\|\cdot\|_p$ بیان کننده نرم L_p معمول است. نمادگذاری مرسوم برای این متر، ℓ_p است. اینفیم به صورت

$$\ell_p(\mu, v) = \|F_\mu^{-1}(U) - F_v^{-1}(U)\|_p$$

و $\mu \in \mathcal{M}^\circ$ به طور نمایی در متر ℓ_2 به نقطه ثابت یکتای K در به وضوح، $\mathcal{M}^\circ$ همگرا است.

$$\mathbb{E}(Y_n) = \mathbb{E}(Y_{I_n}) + \mathbb{E}(Y_{n-1-I_n}^*) + n - 1$$

$$= \frac{1}{n} \sum_{j=1}^n \mathbb{E}(Y_{j+1}) + \frac{1}{n} \sum_{j=1}^n \mathbb{E}(Y_{n-j}) + n - 1 \quad \square$$

اثبات. به [۱۴] مراجعه کنید.

$$= 2(n+1) \sum_{h=1}^{n+1} \frac{1}{h} - 2(n+1) + 2$$

$$= 2n \log n + n(2\gamma - 4) + 2 \log n + 2\gamma + 1$$

$$+ O\left(\frac{\log n}{n}\right),$$

زیرا

$$H_n = \sum_{h=1}^n \frac{1}{h} = \log n + \gamma - \frac{1}{2n} + O(n^{-2}),$$

که در آن γ ثابت اویلر است. حال، دنباله متغیرهای تصادفی $\frac{Y_n - \mathbb{E}(Y_n)}{n}$ را در نظر بگیرید. ادعا می‌کنیم که

$$\frac{Y_n - \mathbb{E}(Y_n)}{n} \xrightarrow{d} X,$$

که در آن متغیر تصادفی X در معادله نقطه ثابت

$$X \stackrel{d}{=} UX + (1-U)X^* + b(U) \quad (10)$$

صدق می‌کند به طوری که $U \sim U(0, 1)$ ، $X \stackrel{d}{=} X^*$ و X و X^* و U مستقل هستند و

$$b(x) = 2x \log x + 2(1-x) \log(1-x) + 1.$$

قبل از اثبات این ادعا توجه کنید که

$$\mathbb{E}(X) = \mathbb{E}(UX + (1-U)X^* + b(U))$$

$$= 0.$$

و

$$\text{var}(X) = \int_0^1 b^2(x) dx = \gamma - \frac{2}{3}\pi^2$$

بنا بر این،

$$\text{var}(Y_n) \sim n^2 \text{var}(X) = \left(\gamma - \frac{2}{3}\pi^2\right) n^2.$$

برای اثبات ادعا، فرض کنید $\mathcal{M}_2$ معرف فضای اندازه‌ها روی $\mathbb{R}$ با گشتاور دوم متناهی و میانگین صفر باشد. در این صورت متر واسراشتاین یا (متر L_2) که با ℓ_2 نشان داده می‌شود به صورت

$$\ell_2(\mu, \nu) = \inf \|X - Y\|_2$$

مثال ساده زیر را در نظر بگیرید. فرض کنید $(X_n)_{n \geq 1}$ دنباله‌ای از متغیرهای تصادفی و $\frac{X_n - \mathbb{E}(X_n)}{s_n}$ نسخه نرمال شده آن باشد. یادآور می‌شویم که s_n^2 دقیقاً واریانس نیست. به طور مشخص فرض کنید به هر روشی توانسته‌ایم $\mathbb{E}(X_n)$ را به دست آوریم ولی واریانس X_n تعیین نشده است. آیا می‌توان تقریب مناسبی برای واریانس ارائه کرد؟ عموماً s_n^2 به عنوان مقیاسی مشخص پیشنهاد می‌شود که مرتبه آن توان دوم مرتبه میانگین است. برای مثال اگر $\mathbb{E}(X_n) \sim \log n$ ، آنگاه می‌توان $s_n = \sqrt{\log n}$ اختیار کرد. دلیل این انتخاب واضح است [۱۴]. اما پرسش اصلی این است که آیا این امکان وجود دارد که متغیری تصادفی چون X پیدا کرد که در یک معادله نقطه ثابت صدق کند که جوابی یکتا دارد، امید آن صفر است و واریانس آن قابل محاسبه است به طوری که $\frac{X_n - \mathbb{E}(X_n)}{s_n} \xrightarrow{d} X$. برای یک لحظه فرض کنید واریانس X تعیین شود. در این صورت،

$$\text{var}(X_n) \sim s_n^2 \text{var}(X).$$

روسلا [۱۱] اولین بار روش انقباض را در مطالعه تعداد مقایسه‌های مورد نیاز به وسیله الگوریتم مرتب‌سازی سریع برای مرتب کردن یک فهرست از n عدد درون ترتیب‌های طبیعی آنها به کار گرفت. در ادامه، به این مسئله و چگونگی حل آن اشاره می‌شود. فرض کنید توانسته‌ایم نشان دهیم متغیر تعریف شده در بالا در رابطه هم‌توزیعی زیر صدق می‌کند (صرف نظر از چگونگی آن): برای هر $n \geq 2$ ،

$$Y_n \stackrel{d}{=} Y_{I_n} + Y_{n-1-I_n}^* + n - 1,$$

که در آن $Y_0 = Y_1 = 0$ ، $Y_2 = 1$ و I_n دارای توزیع یکنواخت گسسته روی مجموعه $\{0, 1, 2, \dots, n-1\}$ است. همچنین، Y_j و I_n و Y_j^* برای $1 \leq j \leq n$ مستقل هستند.

در نتیجه،

$$\begin{aligned} & \ell_2^*(T(\mu), T(\nu)) \\ & \leq \|UX + (1-U)X^* - UY - (1-U)Y^*\|_2^2 \\ & = (\mathbb{E}(U^2) + \mathbb{E}(1-U)^2)\mathbb{E}(X-Y)^2 \\ & = \frac{2}{3}\mathbb{E}(X-Y)^2, \end{aligned}$$

زیرا $\mathbb{E}(U) = 1/2$ و $\mathbb{E}(U^2) = 1/3$. با گرفتن اینفیم روی همه X و Y های ممکن،

$$\ell_2(T(\mu), T(\nu)) \leq \sqrt{\frac{2}{3}}\ell_2(\mu, \nu),$$

که ادعا را ثابت می‌کند. گام بعدی نشان دادن این مطلب است که X_n واقعاً به X همگرا است. یادآوری می‌کنیم که $X_n \xrightarrow{d} X$ بر این دلالت دارد که $\ell_2(\mathcal{L}(X_n), \mathcal{L}(X)) \rightarrow 0$. فرض کنید $\nu_n = \mathcal{L}(X_n)$ و X^* مستقل باشند و $\mathcal{L}(X) = \mu$. $\mathcal{L}(X^*) = \mu$ که در آن μ نقطه ثابت $T(\mu) = \mu$ است. نسخه‌های X_j, X_j^* که برای $1 \leq j \leq n-1$ مستقل هستند را با

$$\text{var}(X_j - X) = \ell_2^*(\nu_j - \mu),$$

$$\text{var}(X_j^* - X^*) = \ell_2^*(\nu_j - \mu)$$

انتخاب کنید. برای $x \in \left(\frac{j}{n}, \frac{j+1}{n}\right]$ قرار دهید $V_x = X_j$ و $V_x^* = X_j^*$. در این صورت، برای U مستقل از X_j و X_j^*

$$\nu_n = \mathcal{L}\left(\frac{[nU] - 1}{n}V_U + \frac{n - [nU]}{n}V_{1-U}^* + b_n([nU])\right).$$

همچنین،

$$\sup_{0 \leq x \leq 1} |b_n([nx]) - b(x)| = \mathcal{O}\left(\frac{\log n}{n}\right). \quad (12)$$

تعریف می‌شود که در آن $\|\cdot\|_2$ معرف L_2 -نرم است و اینفیم روی همه متغیرهای تصادفی X با قانون $\mu = \mathcal{L}(X)$ و همه Y ها با قانون $\nu = \mathcal{L}(Y)$ گرفته می‌شود. روی فضای $(\mathcal{M}_2, \ell_2)$ ، دنباله μ_n به μ همگراست در $\mathcal{M}_2$ اگر و تنها اگر μ_n در توزیع به μ همگرا باشد و اگر گشتاورهای دوم μ_n به گشتاور دوم μ همگرا باشد [۱]. اگر قرار دهیم $X_n = \frac{Y_n - \mathbb{E}(Y_n)}{n}$ ، آن‌گاه

(۱۱)

$$X_n \stackrel{d}{=} X_{I_n} \frac{I_n}{n} + X_{n-1-I_n}^* \frac{n-1-I_n}{n} + b_n(I_n) \quad n \geq 2$$

که در آن $X_0 = X_1 = 0$ و $X_j \stackrel{d}{=} X_j^*$ و I_n و X_j مستقل هستند. دقت می‌کنیم که $U = \frac{I_n}{n}$ دارای توزیع یکنواخت روی بازه $(0, 1)$ است. به علاوه،

$$\begin{aligned} b_n(j) &= \frac{n-1}{n} + \frac{1}{n}(\mathbb{E}(Y_j) + \mathbb{E}(Y_{n-1-j}) - \mathbb{E}(Y_n)) \\ &= 1 + \frac{2}{n} \log \frac{j}{n} + 2\left(1 - \frac{j}{n}\right) \log \left(1 - \frac{j}{n}\right) \\ &\quad + \mathcal{O}\left(\frac{\log n}{n}\right). \end{aligned}$$

بنا بر این، اگر X_n دارای یک توزیع حدی به X باشد، آن‌گاه باید در رابطه (10) صدق کند. اولین گام نشان دادن این است که رابطه (10) دارای یک حل یکتا با $\mathbb{E}(X) = 0$ است. فرض کنید $T: \mathcal{M}_2 \rightarrow \mathcal{M}_2$ ، نگاشتی تعریف شده به صورت

$$T(\mu) = \mathcal{L}(UX + (1-U)X^* + b(U))$$

باشد که در آن X, X^*, U مستقل هستند $\mathcal{L}(X) = \mathcal{L}(X^*) = \mu$ و $0 < u < 1$ و $U \sim f(u) = 1$. در این صورت، T یک انقباض نسبت به متر واسراشتاین ℓ_2 است و بنا بر این یک نقطه ثابت $T(\mu) = \mu$ با $\mu \in \mathcal{M}_2$ است. برای این فرض کنید $\mu, \nu \in \mathcal{M}_2$ و این‌که $\mathcal{L}(X) = \mathcal{L}(X^*) = \mu$ ، $\mathcal{L}(Y) = \mathcal{L}(Y^*) = \nu$ به طوری‌که U, X, X^* و U, Y, Y^* مستقل هستند. در این صورت،

$$T(\mu) = \mathcal{L}(UX + (1-U)X^* + b(U)),$$

$$T(\nu) = \mathcal{L}(UY + (1-U)Y^* + b(U)).$$

بنا بر تعریف متر واسراشتاین، در نهایت برای عبارت پنجم (و به طور مشابه برای عبارت ششم)،

$$\begin{aligned} & \mathbb{E}\left(\left(\frac{\lceil nU \rceil - 1}{n} V_U - UX\right)\right. \\ & \times \left.(b_n(\lceil nU \rceil) - b(U))\right) \\ & = O\left(\frac{\log n}{n^{\frac{1}{2}}} \sum_{j=1}^n \frac{j-1}{n} \ell_{\frac{1}{2}}(v_{j-1}, \mu)\right) + O\left(\frac{\log n}{n^{\frac{1}{2}}}\right). \end{aligned}$$

بنا بر این، با انتخاب $a_j = \ell_{\frac{1}{2}}(v_j, \mu)$

$$\begin{aligned} a_n &= \frac{1}{n} \sum_{j=1}^n \left(\frac{j-1}{n}\right)^{\frac{1}{2}} a_{j-1} \\ &+ O\left(\frac{\log n}{n^{\frac{1}{2}}}\right) \sum_{j=1}^n \frac{j-1}{n} \sqrt{a_{j-1}} + O\left(\frac{(\log n)^{\frac{1}{2}}}{n^{\frac{1}{2}}}\right), \end{aligned}$$

و لذا برای مقادیر مثبت C_1 و C_2

$$a_n \leq \frac{1}{2} \max_{0 \leq j \leq n-1} a_j + C_1 \frac{\log n}{n} \max_{0 \leq j \leq n-1} \sqrt{a_j} + C_2 \frac{\log^{\frac{1}{2}} n}{n^{\frac{1}{2}}}.$$

به آسانی می توان نشان داد که $a_n \rightarrow 0$. بنا بر این،

$$\lim_{n \rightarrow \infty} \ell_{\frac{1}{2}}(\mathcal{L}(X_n), \mathcal{L}(X)) = 0.$$

برای جزئیات بیشتر به [۱۱] مراجعه کنید.

۱.۳ روش انقباض روی L_2

روشی که در بخش ۳ معرفی شد را می توان به چندین روش تعمیم داد. برای رسیدن به این هدف، یک رده از دنباله های بازگشتی از توزیع ها که می توانند توسط ویژگی های فضای L_2 تحلیل شوند، مشخص سازی می شوند. به منظور ارائه یک حالت کلی تر به بردارهای تصادفی در $\mathbb{R}^d$ توجه می کنیم.

فرض کنید یک دنباله از بردارهای d -بُعدی $(Y_n)_{n \geq 1}$ در

رابطه بازگشتی

$$Y_n \stackrel{d}{=} \sum_{r=1}^K A_r(n) Y_{I_r(n)}^{(r)} + b_n, \quad n \geq n_0. \quad (13)$$

صدق می کنند که در آن

$$(Y_n^{(1)}), \dots, (Y_n^{(k)})$$

و

$$(A_1(n), \dots, A_K(n), b_n, I^{(n)})$$

$$\begin{aligned} \ell_{\frac{1}{2}}(v_n, \mu) &\leq \mathbb{E}\left(\frac{\lceil nU \rceil}{n} V_U + \frac{n - \lceil nU \rceil}{n} V_{1-U}^* + b_n(\lceil nU \rceil)\right. \\ &- \left.UX - (1 - U)X^* - b(U)\right) \\ &= \mathbb{E}\left(\frac{\lceil nU \rceil - 1}{n} V_U - UX\right)^{\frac{1}{2}} \\ &+ \mathbb{E}\left(\frac{n - \lceil nU \rceil}{n} V_{1-U}^* - (1 - U)X^*\right)^{\frac{1}{2}} \\ &+ \mathbb{E}(b_n(\lceil nU \rceil) - b(U))^{\frac{1}{2}} \\ &+ \frac{1}{2} \mathbb{E}\left(\left(\frac{\lceil nU \rceil - 1}{n} V_U - UX\right)\right. \\ &\times \left.(b_n(\lceil nU \rceil) - b(U))\right) \\ &+ \frac{1}{2} \mathbb{E}\left(\left(\frac{n - \lceil nU \rceil}{n} V_{1-U}^* - (1 - U)X^*\right)\right. \\ &\times \left.(b_n(\lceil nU \rceil) - b(U))\right) \\ &+ \frac{1}{2} \mathbb{E}\left(\left(\frac{\lceil nU \rceil - 1}{n} V_U - UX\right)\right. \\ &\times \left.\left(\frac{n - \lceil nU \rceil}{n} V_{1-U}^* - (1 - U)X^*\right)\right). \end{aligned}$$

از طرفی،

$$\begin{aligned} & \mathbb{E}\left(\frac{\lceil nU \rceil - 1}{n} V_U - UX\right)^{\frac{1}{2}} \\ &= \mathbb{E}\left(\sum_{j=1}^n I\left(\frac{j-1}{n}, \frac{j}{n}\right)(U) \left(\frac{j-1}{n} X_{j-1} - UX\right)\right)^{\frac{1}{2}} \\ &= \frac{1}{n} \sum_{j=1}^n \frac{(j-1)^{\frac{1}{2}}}{n^{\frac{1}{2}}} \ell_{\frac{1}{2}}(v_{j-1}, \mu) \\ &+ O\left(\frac{1}{n^{\frac{1}{2}}} \sum_{j=1}^n \frac{j-1}{n} \ell_{\frac{1}{2}}(v_{j-1}, \mu)\right) + O\left(\frac{1}{n^{\frac{1}{2}}}\right). \end{aligned}$$

برای عبارت دوم نیز به طریق مشابه خواهد بود. عبارت سوم با استفاده از رابطه (۱۲) به صورت

$$\mathbb{E}(b_n(\lceil nU \rceil) - b(U))^{\frac{1}{2}} = O\left(\frac{(\log n)^{\frac{1}{2}}}{n^{\frac{1}{2}}}\right),$$

برآورد می شود. بنا بر فرض استقلال و این که $\mathbb{E}(X_j) = 0$ داریم:

$$\begin{aligned} & \mathbb{E}\left(\left(\frac{\lceil nU \rceil - 1}{n} V_U - UX\right)\right. \\ & \times \left.\left(\frac{n - \lceil nU \rceil}{n} V_{1-U}^* - (1 - U)X^*\right)\right) \\ &= O\left(\frac{1}{n^{\frac{1}{2}}} \sum_{j=1}^n \frac{j-1}{n} \ell_{\frac{1}{2}}(v_{j-1}, \mu)\right) + O\left(\frac{1}{n^{\frac{1}{2}}}\right). \end{aligned}$$

برای بازفرمول‌بندی ویژگی‌های نقطه ثابت، فضای همه اندازه‌های احتمال روی $\mathbb{R}^d$ را با $\mathcal{M}^d$ نشان می‌دهیم و قرار می‌دهیم:

$$T: \mathcal{M}^d \rightarrow \mathcal{M}^d, \mu \mapsto \mathcal{L} \left(\sum_{r=1}^K A_r^* Z^{(r)} + b^* \right), \quad (20)$$

که در آن $(A_1^*, \dots, A_K^*, b^*)$ و $Z^{(1)}, \dots, Z^{(K)}$ مستقل هستند و برای $r = 1, \dots, K$ ، $\mathcal{L}(Z^{(r)}) = \mu$ ، در این صورت، X یک حل نقطه ثابت (۱۹) است اگر و تنها اگر توزیع $\mathcal{L}(X)$ یک نقطه ثابت نگاشت T باشد.

نگاشتی از نوع (۲۰) اغلب نقاط ثابت یکتا در فضای همه توزیع‌های احتمال ندارد و مشخص‌سازی مجموعه همه نقاط ثابت یک مسئله مهم و باز است. به‌منظور یک مطالعه دقیق‌تر زیرمجموعه‌هایی از $\mathcal{M}^d$ به‌صورت زیر تعریف می‌شوند:

$$\mathcal{M}_s^d = \{\mu \in \mathcal{M}^d : \|\mu\|_s < \infty\}, \quad s > 0 \quad (21)$$

$$(22)$$

$$\mathcal{M}_s^d(M) = \{\mu \in \mathcal{M}_s^d : \mathbb{E}(\mu) = M\}, \quad s \geq 1$$

$$(23)$$

$$\mathcal{M}_s^d(M, C) = \{\mu \in \mathcal{M}_s^d(M) : \text{cov}(\mu) = C\}, \quad s \geq 2$$

که در آن $M \in \mathbb{R}^d$ و C یک ماتریس معین مثبت و متقارن $d \times d$ است و $\|\mu\|_s$ ، $\mathbb{E}(\mu)$ و $\text{cov}(\mu)$ به‌ترتیب بیان‌کننده s -امین گشتاور قدر مطلق، امید ریاضی و ماتریس کوواریانس یک بردار تصادفی با توزیع μ هستند.

هسته این روش اعطای یک زیرفضای مناسب $M^* \in \mathcal{M}^d$ روی یکی از مجموعه‌های ارائه شده در (۲۱) تا (۲۳) با یک متر کامل δ است به‌طوری‌که قید T به M^* یک انقباض روی فضای متریک (M^*, δ) به معنای اشاره شده در قضیه نقطه ثابت باناخ است. این نشان می‌دهد که یک نقطه ثابت موجود $\mathcal{L}(X)$ از T در M^* یکتا می‌شود.

در گام دوم هدف نشان دادن همگرایی کمیت‌های بازمقیاسیده $\mathcal{L}(X_n)$ به $\mathcal{L}(X)$ در متر δ است. یعنی $\delta(\mathcal{L}(X_n), \mathcal{L}(X)) \rightarrow 0$ وقتی $n \rightarrow \infty$ که روی یک همگرایی مناسب ضرایب بیان شده در (۱۸) پایه‌ریزی می‌شود. اگر δ

مستقل هستند، $A_1(n), \dots, A_K(n)$ ماتریس‌های تصادفی $d \times d$ ، b_n یک بردار d -بعدی تصادفی، $I^{(n)}$ برداری از صحیح‌های تصادفی $I_r^{(n)}$ متعلق به $\{0, 1, 2, \dots, n\}$ و $(Y_n^{(1)}), \dots, (Y_n^{(K)})$ دارای توزیع (Y_n) هستند. برای $n_0 \geq 1$ ، بردارهای $Y_{n_0}, \dots, Y_{n-1}$ بردارهای تصادفی آغازین هستند. تعداد $K \geq 1$ تعیینی است. بردار تصادفی (Y_n) در رابطه (۱۳) توسط

$$X_n = (Y_n - M_n)C_n^{-1/2}, \quad (14)$$

که در آن $M_n \in \mathbb{R}^d$ و C_n یک ماتریس $d \times d$ معین-مثبت متقارن است، بازمقیاسیده می‌شود. اگر دو گشتاور اول Y_n متناهی باشند، آنگاه M_n و C_n نوعاً به‌ترتیب از مرتبه امید ریاضی و ماتریس کوواریانس Y_n هستند. از رابطه بازگشتی (۱۳) برای X_n رابطه بازگشتی بهبودیافته

$$X_n \stackrel{d}{=} \sum_{r=1}^K A_r^{(n)} X_{I_r^{(n)}}^{(n)} + b^{(n)}, \quad n \geq n_0. \quad (15)$$

با

$$A_r^{(n)} = C_n^{-1/2} A_r(n) C_{I_r^{(n)}}^{1/2}, \quad (16)$$

$$b^n = C_n^{-1/2} \left(b_n - M_n + \sum_{r=1}^K A_r(n) M_{I_r^{(n)}} \right) \quad (17)$$

به دست می‌آیند که استقلال‌ها مانند رابطه (۱۳) است. روش انقباض به‌صورت زیر دنبال می‌شود.

فرض می‌کنیم که X_n توسط (۱۵) مشخص‌سازی می‌شود. در این صورت، همگرایی ضرایب

$$A_r^{(n)} \xrightarrow{d} A_r^*, \quad b^{(n)} \xrightarrow{d} b^*, \quad (18)$$

برای همگرایی توزیع کمیت‌های (X_n) به یک حد X دلالت می‌کنند. توزیع حدی $\mathcal{L}(X)$ ، توسط یک معادله نقطه ثابت، مشخص‌سازی می‌شود که از رابطه بازگشتی بهبودیافته با قرار دادن $n \rightarrow \infty$ به دست می‌آید:

$$X \stackrel{d}{=} \sum_{r=1}^K A_r^* X^{(r)} + b^*. \quad (19)$$

در این جا $(A_1^*, \dots, A_K^*, b^*)$ و $X^{(1)}, \dots, X^{(K)}$ مستقل هستند و برای هر $r = 1, \dots, K$

$$X^{(r)} \stackrel{d}{=} X.$$

T همان باشد که در رابطه (۲۰) معرفی شد. در این صورت، قید T به $\mathcal{M}_p^d(\circ)$ لپشیتس پیوسته در ℓ_p است و برای ثابت لپشیتس $lip(T)$ داریم:

$$lip(T) \leq \left\| \mathbb{E} \left(\sum_{r=1}^K (A_r^*)^T A_r^* \right) \right\|_{op}^{\frac{1}{p}}.$$

گام بعدی روش انقباض نشان دادن همگرایی در ℓ_p برای دنباله $(\mathcal{L}(X_n))$ ارائه شده در (۱۵) است. قضیه زیر شرطهای کافی برای این موضوع را نشان می‌دهد. برای برهان و جزئیات آن به قضیه ۱۰۳ در [۱۳] و قضیه ۱۰۴ در [۷] مراجعه کنید.

قضیه ۵.۳. فرض کنید (X_n) دنباله‌ای از بردارهای تصادفی d -بعدی انتگرال‌پذیر در L^2 باشد که در شرط (۲۰) صدق می‌کند و در آن ماتریس‌های $d \times d$ تصادفی و L^2 -انتگرال‌پذیر $A_r^{(n)}$ برای $1 \leq r \leq K$ و بردار تصادفی L^2 -انتگرال‌پذیر $b^{(n)}$ و بردار مقادیر صحیح تصادفی $I^{(n)} = (I_1^{(n)}, \dots, I_K^{(n)})$ با $I_r^{(n)} \in \{0, 1, \dots, n\}$ معرفی شدند. فرض کنید که برای $\ell \geq 0$ و $r = 1, \dots, K$

$$(A_1^{(n)}, \dots, A_K^{(n)}, b^{(n)}) \xrightarrow{L^2} (A_1^*, \dots, A_K^*, b^*), \quad (24)$$

$$\mathbb{E} \left(\sum_{r=1}^K \| (A_r^*)^T A_r^* \|_{op} \right) < 1, \quad (25)$$

$$\mathbb{E} \left(I_{\{I_r^{(n)} \leq \ell\}} \| (A_r^{(n)})^T A_r^{(n)} \|_{op} \right) \rightarrow 0, \quad n \rightarrow \infty. \quad (26)$$

در این صورت، وقتی $n \rightarrow \infty$

$$\ell_p(\mathcal{L}(X_n), \mathcal{L}(X)) \rightarrow 0,$$

که در آن $\mathcal{L}(X)$ نقطه ثابت نگاشت T تعریف شده در (۲۰) است که در $\mathcal{M}_p^d(\circ)$ یکتا است.

شرط (۲۴) بدین معنا است که همگرایی ضرایب معرفی شده در (۱۸) باید در L^2 برقرار باشد. برای این، مجاز به قید $(A_1^*, \dots, A_K^*, b^*)$ مطابق با $(A_1^{(n)}, \dots, A_K^{(n)}, b^{(n)})$ روی فضای احتمال توأم هستیم. یعنی (۲۴) بدین معنا است که $\ell_p(\mathcal{L}(A_1^{(n)}, \dots, A_K^{(n)}, b^{(n)}), \mathcal{L}(A_1^*, \dots, A_K^*, b^*)) \rightarrow 0$.

شرط (۲۵) بنا بر نابرابری پِرسون قوی‌تر از شرط انقباض

$$\left\| \mathbb{E} \left(\sum_{r=1}^K (A_r^*)^T A_r^* \right) \right\|_{op} < 1 \quad (27)$$

به‌گونه‌ای انتخاب شود که همگرایی در δ بر همگرایی ضعیف دلالت کند، آنگاه همگرایی در توزیع نتیجه می‌شود.

نکته حیاتی انتخاب یک متر سَره برای نشان دادن این مطلب است که نگاشت T یک انقباض است. یک انتخاب طبیعی، متر ℓ_p برای $p > 0$ است که به‌صورت

$$\ell_p(\mu, v) := \inf \{ \|X - Y\|_p; \mathcal{L}(X) = \mu, \mathcal{L}(Y) = v \}$$

برای $\mu, v \in \mathcal{M}_p^d$ تعریف می‌شود که در آن

$$\|X\|_p = (\mathbb{E}(\|X\|^p))^{\min\{\frac{1}{p}, 1\}},$$

بیان‌کننده نرم L_p برای یک بردار تصادفی X و $\|X\|$ بیان‌کننده نرم اقلیدسی آن است. متر واستراشتاین ℓ_2 یک حالت خاص و البته مهم‌ترین حالت است. به‌علاوه فضای $(\mathcal{M}_p^d, \ell_p)$ برای $p > 0$ به‌خوبی $(\mathcal{M}_p^d(M), \ell_p)$ برای $M \in \mathbb{R}^d$ و $p \geq 1$ فضای متریک کامل است و همگرایی در ℓ_p معادل همگرایی ضعیف و همچنین همگرایی گشتاور قدر مطلق p ام است.

برای $\mu, v \in \mathcal{M}_p^d$ همیشه بردارهای X و Y روی یک فضای احتمال توأم با $\mathcal{L}(X) = \mu$ و $\mathcal{L}(Y) = v$ و $\ell(\mu, v) = \|X - Y\|_p$ وجود دارند به‌طوری‌که این بردارها مفصل‌های بهینه μ و v نامیده می‌شوند. برای این مطلب و ویژگی‌های متر ℓ_p به [۶] مراجعه کنید.

به‌منظور دستیابی به ویژگی‌های انقباض نگاشت (۲۰) نرم عملگر یک ماتریس مربع A را با

$$\|A\|_{op} = \sup_{\|x\|=1} \|Ax\|$$

و ترانواده آن را با A^T و ثابت لپشیتس T را در صورت متناهی بودن با

$$lip(T) = \inf_{\substack{\mu, v \in \mathcal{M}^* \\ \mu \neq v}} \frac{\ell_p(T(\mu), T(v))}{\ell_p(\mu, v)}$$

نشان می‌دهیم. در حالت $p = 2$ ثابت لپشیتس به آسانی برآورد می‌شود.

لم ۴.۳. فرض کنید $(A_1^*, \dots, A_K^*, b^*)$ یک بردار L^2 -انتگرال‌پذیر از ماتریس‌های $d \times d$ تصادفی $A_1^*, \dots, A_K^*$ و یک بردار d -بعدی تصادفی b^* با $\mathbb{E}(b^*) = 0$ باشد و فرض کنید

در لم ۴.۳ است. شرط (۲۶) یک شرط فنی است که معمولاً برای دستیابی در کاربرد آسان است. در برخی مواقع بررسی همگرایی $X_n \xrightarrow{d} X$ به طور مستقیم صورت می گیرد.

برای کاربرد قضیه ۵.۳ برای دنباله های بازگشتی (Y_n) در (۱۳) توجه به نکات زیر ضروری است. برای مقیاس بندی در (۱۴) مجبور به انتخاب $M_n = \mathbb{E}(Y_n)$ به منظور تعیین شرط های $\mathbb{E}(X_n) = 0$ و $\mathbb{E}(b^{(n)}) = 0$ هستیم. چون $b^{(n)}$ شامل کمیت های M_n است نیازمند به دست آوردن یک حد برای $b^{(n)}$ هستیم. این دلالت بر این دارد که برای کاربرد قضیه ۵.۳ یک بسط مجانبی میانگین $\mathbb{E}(Y_n)$ باید مشخص شود. در مقابل ماتریس کوواریانس $cov(Y_n)$ باید در بسط مجانبی مرتبه اول آن حدس زده شود. چون همگرایی در ℓ_2 بر همگرایی گشتاور دوم دلالت دارد، قضیه ۵.۳ به طور خودکار بر یک بسط مجانبی ماتریس کوواریانس $cov(Y_n)$ دلالت می کند.

مجدداً مثال ارائه شده توسط روسلا در بخش ۳ را در نظر بگیرید. قرار دهید:

$$d = 1,$$

$$K = 2,$$

$$A_1^{(n)} = A_2^{(n)} = 1,$$

$$I_1^{(n)} \sim DU\{0, \dots, n-1\},$$

$$I_2^{(n)} = n-1 - I_1^{(n)},$$

$$b_n = n-1.$$

یادآوری می کنیم که مجبور به کار در فضای $\mathcal{M}_2(0)$ هستیم. یعنی مجبور به داشتن اطلاع روی اولین گشتاور به منظور تقلیل (۱۳) به (۱۵) هستیم. با نرمال کردن توسط $1/n$ ، داریم:

$$X_n = \frac{Y_n - \mathbb{E}(Y_n)}{n}.$$

با بحث مشابه در قسمت ۳، به رابطه بازگشتی (۱۱) می رسم که به شکل (۱۵) است. معادله حدی برای X توسط (۱۰) داده

می شود. بنا بر این،

$$A_1^* = U,$$

$$A_2^* = 1 - U,$$

$$b^* = 1 + 2U \log U + 2(1-U) \log(1-U).$$

شرایط قضیه ۵.۳ می توانند به طور مستقیم بررسی شوند. برای مثال، ثابت لیپشیتس می تواند توسط

$$\begin{aligned} \left(\mathbb{E} \| (A_1^*)^\top \|_{op} + \mathbb{E} \| (A_2^*)^\top \|_{op} \right)^{1/2} &= \left(2 \int_0^1 u^\top du \right)^{1/2} \\ &= \sqrt{\frac{2}{3}} \end{aligned}$$

برآورد شود.

حال، فرض کنید دو دنباله (W_n) و (Y_n) در برابری های توزیعی زیر صدق کنند:

$$\begin{aligned} W_n &\stackrel{d}{=} W_{I_n}^{(1)} + (n + I_n) Y_{I_n} + W_{n-1-I_n}^{(2)} + (I_n + 1) Y_{n-I_n-1} \\ &\quad + 2I^{(n)}(n-1-I^{(n)}) + n-1, \end{aligned}$$

$$Y_n \stackrel{d}{=} Y_{I_n}^{(1)} + Y_{n-1-I_n}^{(2)} + n-1.$$

این بدان معنا است که

$$d = 2,$$

$$K = 2,$$

$$I_1^{(n)} \sim DU\{0, \dots, n-1\},$$

$$I_2^{(n)} = n-1 - I_1^{(n)},$$

$$A_1^{(n)} = \begin{pmatrix} 1 & n - I_1^{(n)} \\ 0 & 1 \end{pmatrix},$$

$$A_2^{(n)} = \begin{pmatrix} 1 & n - I_2^{(n)} \\ 0 & 1 \end{pmatrix},$$

$$b_n = \begin{pmatrix} 2I_1^{(n)} I_2^{(n)} + n-1 \\ n-1 \end{pmatrix}.$$

بنا بر این،

$$\begin{aligned} \mathbb{E}(W_n) &= 2n^\top H_n - 2n^\top + 2nH_n - 1 \circ n + 2H_n \\ &= 2n^\top \log n + (2\gamma - 2)n^\top + o(n^\top). \end{aligned}$$

اگر به بردار نرمال شده

تباهیده است. برای مثال، اگر رابطه بازگشتی

$$Y_n \stackrel{d}{=} Y_{I_n} + b_n, \quad n \geq n_0.$$

$$X_n = \left(\frac{W_n - \mathbb{E}(W_n)}{n^2}, \frac{Y_n - \mathbb{E}(Y_n)}{n} \right)$$

توجه کنیم، آن‌گاه

را بازمقیاسیده کنیم ممکن است به معادله حدی تباهیده‌ای به صورت $X \stackrel{d}{=} X$ برسیم. این معادله حدی هیچ اطلاعی در خصوص توزیع حدی ندارد و مفهوم روش انقباض نیازمند توسیعی برای چنین مواردی است.

$$A_1^* = \begin{pmatrix} (1-U)^2 & U(1-U) \\ 0 & 1-U \end{pmatrix},$$

$$A_r^* = \begin{pmatrix} (1-U)^2 & U(1-U) \\ 0 & 1-U \end{pmatrix},$$

$$b^* = \begin{pmatrix} 6U(1-U) + 2U \log U + 2(1-U) \log(1-U) \\ 1 + 2U \log U + 2(1-U) \log(1-U) \end{pmatrix}$$

۴ متر زلثارف

به منظور بی پاسخ نگذاشتن مسائلی از معادلات حدی که حل یکتایی ندارند مجبور به انتخاب یک زیرفضای سره هستیم. برای مثال، اگر با $\mathcal{M}_s^d(M, C)$ به جای $\mathcal{M}_s^d(M)$ کار کنیم، آن‌گاه کوواریانس ثابت است و همچنین معادله $X \stackrel{d}{=} \sum_{r=1}^K A_r^* X^{(r)}$ تحت قید $\sum_{r=1}^K (A_r^*)^2 = 1$ یک حد یکتا خواهد داشت. به علاوه، مجبور به استفاده از متر دیگری که باعث می شود که نگاشت T یک انقباض روی زیرفضا داشته باشد، هستیم. این موضوع توسط نینینگر و روشندورف [۸] بررسی شد. آنها از متر زلثارف ξ_s روی بردارهای تصادفی X و Y در $\mathbb{R}^d$ که به صورت

که در آن U به طور یکنواخت روی بازه $[0, 1]$ توزیع می شود. متغیر W_n شاخص وینر و Y_n طول مسیر داخلی درخت های جستجوی دودویی هستند (برای جزئیات [۷] را ببینید).

استفاده از L_2 نسبتاً ساده است اما چندین محدودیت در استفاده از آن وجود دارد. برای مثال، اگر

$$\mathbb{E} \left(\sum_{r=1}^K (A_r^*)^2 \right) = 1, \quad (28)$$

آن‌گاه L_2 را نمی توان به کار گرفت. یعنی، شرط (۲۷) و به ترتیب (۲۵) برقرار نیستند. به ویژه، کاربردهای فراوانی وجود دارد که به معادله حدی

$$X \stackrel{d}{=} \sum_{r=1}^K A_r^* X^{(r)} \quad (29)$$

با

$$\sum_{r=1}^K (A_r^*)^2 = 1 \quad (30)$$

منجر می شود. از ویژگی پیش توزیع نرمال نتیجه می شود که متغیرهای تصادفی نرمال با قاعده $N(0, \sigma^2)$ حل های معادله حدی (۲۹) تحت قید (۳۰) هستند. بنا بر این، به نظر می رسد که نمی توان یک قضیه حدی مرکزی را حتی اگر وجود داشته باشد ثابت کرد [۱۰]. چون معادله حدی، نقطه ثابت یکتا ندارد بدون شک یک متر برای این که T یک انقباض باشد نیز وجود ندارد. مشکل اصلی دیگر روش انقباض، ظهور معادله های حدی

$$\xi_s(X, Y) = \sup_{f \in \mathcal{F}_s} |\mathbb{E}(f(X) - f(Y))|$$

تعریف می شود، استفاده کردند که در آن $s = m + \alpha$ با $0 < \alpha \leq 1$ است، $m \in \mathbb{N}$ و

$$\mathcal{F}_s = \{f \in C^m(\mathbb{R}^d, \mathbb{R}) : \|f^{(m)}(x) - f^{(m)}(y)\| \leq \|x - y\|^\alpha\},$$

بیانگر فضای توابع m بار مشتق پذیر پیوسته از $\mathbb{R}^d$ به $\mathbb{R}$ است به طوری که m -امین مشتق پیوسته هولدر از مرتبه α است. از نماد کوتاه

$$\xi_s(\mathcal{L}(X), \mathcal{L}(Y)) = \xi_s(X, Y)$$

استفاده می کنیم. داریم که $\xi_s(X, Y) < \infty$ هرگاه همه گشتاورهای آمیخته مرتبه های $1, \dots, m$ از X و Y برابر باشند و s -امین گشتاورهای قدر مطلق X و Y متناهی باشند. به علاوه $(\mathcal{M}_s^d, \xi_s)$ برای $0 < s \leq 1$ ، $(\mathcal{M}_s^d(M), \xi_s)$ برای $1 < s \leq 2$ و $(\mathcal{M}_s^d(M, C), \xi_s)$ برای $2 < s \leq 3$ فضاهای متریک کامل با

$M \in \mathbb{R}^d$ و ماتریس $d \times d$ معین-مثبت و متقارن C هستند. ثابت لیپشیتس نگاشت T می‌تواند توسط

$$\text{lip}(T) \leq \mathbb{E} \left(\sum_{r=1}^k \|A_r^*\|_{op}^s \right)$$

برآورد شود. برای مثال، اگر $\sum_{r=1}^K (A_r^*)^2 = 1$ ، آن‌گاه

$$T : \mathcal{M}_s^1(\circ, 1) \rightarrow \mathcal{M}_s^1(\circ, 1)$$

یک انقباض برای هر $2 < s \leq 3$ است. یک توسیع مستقیم از قضیه ۵.۳ به متر زلتارف با کاربردهای بسیاری در آمار و الگوریتم‌های بازگشتی و موضوعات مرتبط وجود دارد که منتهی به قضایای حدی مرکزی تحت شرایط کلی‌تر می‌شود. حتی کار کردن در یک فضای باناخ مناسب نیز امکان‌پذیر است. اولین مسئله‌ای که به‌کمک متر زلتارف مورد بررسی قرار گرفت به شرح زیر است: فرض کنید $L_0 = L_1 = 1$ و

$$L_n \stackrel{d}{=} 1 + L_{I_n}^{(1)} + L_{n-I_n}^{(2)},$$

که در آن $I_n \sim \text{Bin}(n, p)$ و $L_k^{(1)} \stackrel{d}{=} L_k^{(2)}$. همچون گذشته شرط استقلال برقرار است. ثابت شد که تحت شرایط خاصی،

$$Y_n = \frac{L_n - \mathbb{E}(L_n)}{\sqrt{n}} \rightarrow N(\circ, \sigma^2),$$

۵ نتیجه‌گیری

در این مقاله، یکی از مهم‌ترین روش‌های تحلیل روابط بازگشتی تصادفی که اغلب در تحلیل الگوریتم‌ها و ساختارهای بازگشتی رخ می‌دهد، معرفی شد. معادله بازگشتی زیر را با شرایط استقلال بیان‌شده در طول مقاله و شرایط آغازین دلخواه در نظر بگیرید:

$$X_n \stackrel{d}{=} X_{I_n}^{(1)} + X_{n-I_n}^{(2)} + T_n.$$

بخش غیروابسته به متغیر اصلی تحت بررسی که در معادله بازگشتی به‌صورت T_n ظاهر می‌شود تابع عوارض نامیده می‌شود. تابع عوارض یا به I_n وابسته است یا وابسته نیست. در مثال حل‌شده توسط روسلا این تابع برابر با $n-1$ و در مثال بخش ۴ برابر با ۱ است. اگر تابع عوارض به I_n وابسته باشد تحلیل پیچیده‌تر خواهد شد. هوانگ و نینینگر [۵] این موضوع را مورد بررسی دقیق قرار داده‌اند. برای مثال‌های گوناگون در حوزه‌های مختلف [۵، ۸، ۱۰] را ببینید.

مراجع

- [1] Bickel, P., and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap. *Annals of Statistics*, **9**, 1196-1217.
- [2] Gutjahr, W. and Pflug, G. Ch. (1992). The asymptotic contour process of a binary tree is a Brownian excursion, *Stochastic Processes and their Applications*, **41(1)**, 69-89.
- [3] Hofri, M. (1987). *Probabilistic Analysis of Algorithms*, Springer Verlag, New York.
- [4] Maejima, M., and Rachev, S. (1987). An ideal metric and the rate of convergence to a self-similar process, *Annals of Probability*, **14**, 708-727.
- [5] Hwang, H-K., and Neininger, R. (2002). Phase change of limit laws in the quicksort recurrence under varying toll functions, *SIAM Journal on Computing (SICOMP)*, **31(6)**, 1687-1722.

- [6] Major, P. (1978). On the invariance principle for sums of independent identically distributed random variables. *Journal of Multivariate Analysis*, **8(4)**, 487-517.
- [7] Neininger, R. (2001). On a multivariate contraction method for random recursive structures with applications to Quicksort. *Random Structures Algorithms*, **19(3-4)**, 498-524.
- [8] Neininger, R., and Ruschendorf, L. (2004). A general limit theorem for recursive algorithms and combinatorial structures. *Annals of Applied Probability*, **14(1)**, 378-418.
- [9] Rachev, S. (1991). *Probability Metrics and the Stability of Stochastic Models*, Wiley, New York.
- [10] Rachev, S. and Rschendorf, L. (1995). Probability metrics and recursive algorithms, *Advances in Applied Probability*, **27**, 770-799
- [11] Roesler, U. (1991). A limit theorem for "Quicksort", *RAIRO-Theoretical Informatics and Applications*, **25(1)**, 85-100.
- [12] Roesler, U. (1992). A fixed point theorem for distributions, *Stochastic Processes and their Applications*, **37**, 95-214.
- [13] Roesler, U. (2001). The analysis of stochastic divide and conquer algorithms, *Algorithmica*, **29(1-2)**, 238-261
- [14] Roesler, U., and Rüschemdorf, L. (2001). The contraction method for recursive algorithms, *Algorithmica*, **29**, 3-33.
- [15] Zolotarev, V. M. (1997). *Modern Theory of Summation of Random Variables*, VSP, Utrecht.