یازدهمین سمینار

جبرخطی و کاربردهای آن

۱۴ و ۱۵ بهمن‌ماه ۱۴۰۰

دانشکده ریاضی و علوم کامپیوتر

دانشگاه حکیم سبزواری

چکیده مبسوط

فهرست مطالب

نمایه نویسندگان

سخنرانی‌ها

روش تکراری فوق‌تخفیف گاوس-سایدل بلوکی برای حل زوج معادلات شرودینگر غیرخطی مکان کسری

حامد اصلانی[۱] *، داود خجسته سالکویه[۱] و مهران تقی‌پور[۱]

[۱]دانشکده علوم ریاضی، دانشگاه گیلان، رشت، ایران

چکیده

در این مقاله، دستگاه حاصل از گسسته‌سازی زوج معادلات شرودینگر غیرخطی مکان کسری در نظر گرفته می‌شود. ماتریس ضرایب دستگاه حاصل از گسسته سازی، شامل دو بخش حقیقی و موهومی است. قسمت حقیقی از ماتریس توپلیتز و یک ماتریس قطری و قسمت موهومی از ماتریس همانی تشکیل می‌گردد. یک روش کارا، موسوم به روش فوق‌تخفیف گاوس-سایدل بلوکی برای حل دستگاه حاصل، بکار گرفته خواهد شد. نکته حائز اهمیت روش ارائه شده، این است که تنها شامل دو ضرب ماتریس-در-بردار می‌باشد. بنابراین، حجم محاسبات و میزان اشغال فضای حافظه بسیار اندک خواهد بود. در ادامه، همگرایی و محاسبه پارامتر بهینه فوق‌تخفیف ارائه می‌شود. در نهایت، یک مثال عددی برای کارایی روش مورد نظر عرضه خواهد شد.

واژه‌های کلیدی: معادلات شرودینگر کسری فضایی، ماتریس توپلیتز، روش فوق‌تخفیف گاوس-سایدل بلوکی، همگرایی.
کد موضوع‌بندی ریاضی [۲۰۱۰]: 65F10، 81Q05، 81V99.

۱ مقدمه

یکی از معادلات مهم در مکانیک کوانتوم معادله شرودینگر است. از معادله شرودینگر با نام هسته کوانتوم نیز یاد می‌شود. این معادله از انتگرال گیری سطح روی حرکات براونی[۱] حاصل می‌گردد.
زوج معادلات شرودینگر غیرخطی مکان کسری به صورت زیر تعریف می‌شوند:

$$\begin{cases} iu_t + \gamma(-\Delta)^{\frac{\alpha}{2}} u + \rho\left(|u|^2 + \beta|v|^2\right) u = 0, \\ iv_t + \gamma(-\Delta)^{\frac{\alpha}{2}} v + \rho\left(|v|^2 + \beta|u|^2\right) v = 0, \end{cases} \qquad a \le x \le b, 0 < t \le T, \tag{1}$$

به‌طوری‌که در شرایط اولیه

$$\begin{cases} u(x,0) = u_0(x), v(x,0) = v_0(x), & a \le x \le b, \\ u(a,t) = u(b,t) = 0, v(a,t) = v(b,t) = 0, & 0 \le t \le T, \end{cases}$$

صدق می‌کنند. در اینجا $1 < \alpha < 2$ و به‌علاوه $\beta \ge 0$، $\rho > 0$ و $\gamma > 0$. در [۴] ، لاپلاسین کسری[۲] به صورت

$$(-\Delta)^{\frac{\alpha}{2}} u(x,t) = \mathcal{F}^{-1}\left(|\xi|^{\alpha} \mathcal{F}(u(x,t))\right),$$

* حامد اصلانی. آدرس ایمیل: hamedaslani525@gmail.com

[۱] Brownian motion
[۲] Fractional Laplacian

تعریف شده است، که در آن $\mathcal{F}$ تبدیل فوریه[۱] اعمال شده بر متغیر مکانی x می‌باشد. فرض کنید

$$ {}_{-\infty}D_x^{\alpha}u(x,t) = \frac{1}{\Gamma(n-\alpha)}\frac{\partial^n}{\partial x^n}\int_{-\infty}^{x}(x-\tau)^{n-1-\alpha}u(\tau,t)d\tau, $$

$$ {}_{x}D_{+\infty}^{\alpha}u(x,t) = \frac{1}{\Gamma(n-\alpha)}\frac{\partial^n}{\partial x^n}\int_{x}^{+\infty}(\tau-x)^{n-1-\alpha}u(\tau,t)d\tau, $$

به ترتیب مشتقات کسری ریمان-لیوویل[۲] چپ و راست باشند. در این‌صورت مشتق کسری ریز[۳] به صورت

$$ \frac{\partial^{\alpha}}{\partial|x|^{\alpha}}u(x,t) = -(-\Delta)^{\frac{\alpha}{2}}u(x,t) = -\frac{1}{2\cos\frac{\pi\alpha}{2}}\left[{}_{-\infty}D_x^{\alpha}u(x,t) + {}_{x}D_{+\infty}^{\alpha}u(x,t)\right]. $$

خواهد بود.

به طور کلی بررسی جواب دقیق زوج معادلات شرودینگر غیرخطی مکان کسری، مسئله‌ای دشوار است. در سال‌های اخیر، چند روش عددی برای حل آن‌ها در نظر گرفته شده‌است. از جمله می توان به روش کرانک-نیکلسون[۴] [۱] ، روش کالوکیشن[۵] [۲] ، و روش تفاضلات متناهی[۶] [۵] اشاره کرد.

گسسته‌سازی زوج معادلات شرودینگر غیرخطی مکان کسری، منجر به حل دستگاه خطی متقارن و مختلط می‌شود. ماتریس ضرایب دستگاه بدست‌آمده از گسسته‌سازی، شامل دو بخش حقیقی و موهومی است. قسمت حقیقی از مجموع ماتریس‌های توپلیتز[۷] و یک ماتریس قطری با درایه‌های قطری مثبت تشکیل می‌شود. به‌علاوه، قسمت موهومی از ماتریس همانی تشکیل می‌شود. اخیراً، در [۳]، یک دستگاه دو-در-دو بلوکی متناظر با دستگاه اولیه ساخته و روش گاوس-سایدل برای حل آن بکار گرفته شده است. هم‌چنین همگرایی روش گاوس-سایدل برای حل دستگاه دو-در-دو بلوکی متناظر مورد بررسی قرار گرفته است. در این مقاله، یک روش کارا با نام روش تکراری فوق‌تخفیف گاوس-سایدل بلوکی برای حل دستگاه معادلات خطی دو-در-دو متناظر با دستگاه حاصل از گسسته سازی زوج معادلات شرودینگر غیرخطی مکان کسری ارائه خواهد گردید.

این مقاله شامل چهار بخش است. در بخش ۲ ، روش تکراری فوق‌تخفیف گاوس-سایدل بلوکی برای دستگاه حاصل از گسسته سازی زوج مسئله شرودینگر غیرخطی مکان کسری معرفی خواهد شد. در بخش ۳ ، با ارایه یک مثال به بررسی کارایی روش مودرنظر در مقایسه با سایر روش‌های موجود می‌پردازیم. در نهایت، نتیجه‌گیری در بخش ۴ عرضه خواهد شد.

۲ روش تکراری فوق‌تخفیف گاوس-سایدل بلوکی

فرض کنید بازه $[a,b]$ به $M+1$ قسمت مساوی و بازه زمانی $[0,T]$ به $4M$ قسمت مساوی تقسیم شود. دستگاه معادلات خطی حاصل از گسسته‌سازی زوج معادلات شرودینگر غیرخطی مکان کسری در هر گام زمانی به صورت

$$ AU = b, \quad A \in \mathbb{C}^{M\times M} \quad , U,b \in \mathbb{C}^{M}, \tag{۲} $$

است، به‌طوری‌که $A = iI + T + D$، یک ماتریس نامنفرد ، b بردار سمت راست و U بردار مجهولات است [۳] . به‌علاوه، $D = \mathrm{diag}(d_1, d_2, \ldots, d_M)$ که در آن $d_i \geq 0, \; i = 1,2,\ldots,M$. T ماتریس توپلیتز و معین مثبت متقارن (SPD) می‌باشد. قرار می‌دهیم $U = x + iy$ و $b = f + ig$ ، به‌طوری‌که $y, z, p, q \in \mathbb{R}^M$. بنابراین، دستگاه (۲)، به صورت دستگاه

$$ \mathcal{A}\mathbf{x} \equiv \begin{pmatrix} -I & W \\ W & I \end{pmatrix}\begin{pmatrix} y \\ x \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} \equiv \mathcal{P}, \tag{۳} $$

تبدیل می‌شود.

---

[۱] Fourier transform
[۲] Riemann-Liouville
[۳] Riesz
[۴] Crank-Nickelson
[۵] Collocation
[۶] Finite difference
[۷] Toeplitz

۱.۲ روش فوق تخفیف گاوس-سایدل بلوکی

ابتدا یک شکافت به صورت

$$\mathcal{A} = (\omega\mathcal{D} - \mathcal{E}) - (\mathcal{E}^T - (1-\omega)\mathcal{D}) =: \mathcal{M} - \mathcal{N}, \tag{۴}$$

را برای ماتریس ضرایب (۴) در نظر می‌گیریم، به‌طوری‌که

$$\mathcal{D} = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}, \qquad \mathcal{E} = \begin{pmatrix} 0 & 0 \\ -W & 0 \end{pmatrix}.$$

فرض کنید $\omega \neq 0$. پارامتر ω را پارامتر فوق‌تخفیف می‌نامیم. با استفاده از شکافت (۴) ، روش تکراری فوق‌تخفیف گاوس-سایدل بلوکی به صورت زیر خواهد بود:

$$\mathcal{M}z^{(k+1)} = \mathcal{N}z^{(k)} + \mathcal{P}, \quad k = 0, 1, 2, \cdots,$$

که $\mathcal{M}$ و $\mathcal{N}$ در (۴) تعریف شده‌اند. هم‌چنین $z^{(k)} = (y^{(k)}; x^{(k)})$، به‌طوری‌که $y^{(k)}$ و $x^{(k)}$ بردارهای M -تایی می‌باشند و $z^{(0)}$ یک حدس اولیه است. درنتیجه، شکل کلی روش مورد نظر به صورت زیر خواهد بود:

$$\begin{cases} y^{(k+1)} = \frac{1}{\omega}\left((\omega-1)y^{(k)} + Wx^{(k)} - f\right), \\ x^{(k+1)} = \frac{1}{\omega}\left((\omega-1)x^{(k)} + g - Wy^{(k+1)}\right). \end{cases} \tag{۵}$$

همانطور که مشاهده می‌شود، این روش شامل هیچ‌گونه حل دستگاهی نمی‌باشد. این موضوع بسیار حائز اهمیت است که حجم محاسبات تنها به دو ضرب ماتریس-در-بردار محدود می‌گردد. توجه می‌کنیم که به ازای $\omega = 1$، روش گاوس-سایدل بلوکی بدست می‌آید. بنابراین روش گاوس-سایدل حالت خاصی از روش ارائه‌شده می‌باشد. در ادامه به بررسی همگرایی روش و یافتن مقدار بهینه پارامتر فوق‌تخفیف می‌پردازیم. ابتدا یک لم مرور می‌گردد.

لم ۱.۲. *[۶] اندازه‌ی هر دو ریشه معادله درجه دوم $x^2 - px + q = 0$ کمتر از یک هستند، اگر و تنها اگر $|q| < 1$ و $|p| < 1 + q$.*

قضیه ۲.۲. *فرض کنید $A = iI + D + T \in \mathbb{R}^{M\times M}$، به‌طوری‌که D و T به ترتیب ماتریس‌های قطری و توپلیتز باشند. در این صورت، شرط لازم و کافی برای همگرایی روش فوق‌تخفیف گاوس-سایدل بلوکی (۵) برای حل دستگاه (۳)، عبارت است از $\omega < \frac{1+\mu_{\max}}{2}$، که در آن $\mu_{\max}$ بزرگترین مقدار ویژه W است.*

اثبات. فرض کنید $(\lambda, \mathbf{x} = (u; v))$ یک زوج ویژه ماتریس $\mathcal{G} = \mathcal{M}^{-1}\mathcal{N}$ باشد. بدون کاستن از کلیت مسأله فرض کنید $\lambda \neq 0$. بنابراین،

$$(\mathcal{D} - \omega\mathcal{E})^{-1}(\mathcal{E}^T - (1 - \alpha\mathcal{D}))\mathbf{x} = \lambda\mathbf{x},$$

به طور معادل،

$$\begin{cases} (1-\omega)u - Wv = -\lambda\omega u, & (۶) \\ (\omega-1)v = \lambda(Wu + \omega v). & (۷) \end{cases}$$

با محاسبه $v = ((\lambda-1)\omega + 1)W^{-1}u$ از تساوی (۶) و جایگذاری در تساوی (۷) خواهیم داشت:

$$-\lambda W^2 u = ((\lambda-1)\omega + 1)^2 u.$$

رابطه اخیر نشان می‌دهد که اگر μ یک مقدار ویژه W باشد، آنگاه

$$\lambda\mu^2 = -((\lambda-1)\omega + 1)^2 \tag{۸}$$
$$= -(\lambda^2\omega^2 + 2\omega(1-\omega)\lambda + (\omega-1)^2). \tag{۹}$$

با اندکی محاسبه می‌توان دریافت که تساوی (۹)، به صورت

$$\lambda^{۲} - \left(\frac{۲\omega^{۲} - ۲\omega - \mu^{۲}}{\omega^{۲}}\right)\lambda + (\frac{\omega - ۱}{\omega})^{۲} = ۰,$$

قابل بیان است. اکنون با توجه به لم ۱.۲، شرایط لازم و کافی برای همگرایی روش به شکل زیر خواهند بود:

$$\begin{cases} |\omega - ۱| < \omega, \\ |۲\omega^{۲} - ۲\omega - \mu^{۲}| < ۲\omega^{۲} - ۲\omega + ۱. \end{cases}$$

توجه می‌کنیم که شرط $|\omega - ۱| < \omega$، معادل با $\omega > \frac{۱}{۲}$، است. به‌علاوه، برقراری شرط دوم با برقراری رابطه

$$(۲\omega - ۱)^{۲} > \mu^{۲}, \tag{۱۰}$$

معادل است. برای برقراری رابطه (۱۰)، کافی است داشته باشیم $\omega > \frac{۱+\mu}{۲}$ و این اثبات را تکمیل می‌کند. □

در ادامه به یافتن پارامتر فوق‌تخفیف بهینه خواهیم پرداخت. برای این منظور بایستی ω به گونه‌ای محاسبه گردد که داشته باشیم:

$$\rho\left(\mathcal{G}_{\omega^{*}}\right) = \arg \min_{\omega > \frac{۱+\mu_{\max}}{۲}} \rho\left(\mathcal{G}_{\omega}\right).$$

حال قضیه بعدی را بیان می‌کنیم.

قضیه ۳.۲. *فرض کنید شرایط قضیه ۲.۲ برقرار باشند. مقدار بهینه پارامتر فوق‌تخفیف در روش فوق‌تخفیف گاوس-سایدل بلوکی به صورت*

$$\omega^{*} = \frac{۱}{۲}\left(۱ + \sqrt{۱ + \rho^{۲}(W)}\right), \tag{۱۱}$$

است. به‌علاوه

$$\rho\left(\mathcal{G}_{\omega^{*}}\right) = ۱ - \frac{۱}{\omega^{*}} = (\frac{\rho(W)}{۱ + \sqrt{۱ + \rho^{۲}(W)}})^{۲}.$$

اثبات. فرض کنید λ یک مقدار ویژه $\mathcal{G}_{\omega}$ باشد. در این صورت، با توجه به رابطه (۸)، خواهیم داشت: $\lambda < ۰$ یا $\lambda \in \mathbb{C}$. با محاسبه λ از رابطه (۹)، داریم

$$\lambda_{۱,۲}(\omega) = \frac{-۲\omega^{۲} + ۲\omega + \mu^{۲}}{۲\omega^{۲}} \pm \frac{\sqrt{\Delta}}{۲},$$

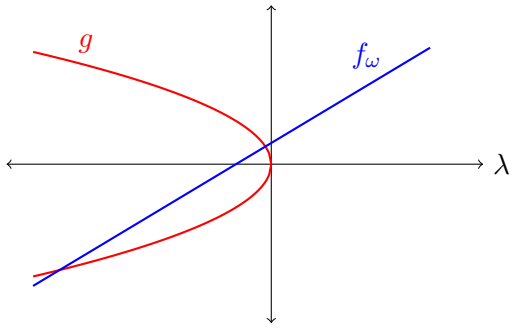
به‌طوری‌که

$$\Delta = \left(\frac{۲\omega^{۲} - ۲\omega - \mu^{۲}}{\omega^{۲}}\right)^{۲} - ۴\left(\frac{\omega - ۱}{\omega}\right)^{۲}.$$

همچنین با استفاده از رابطه (۹)، داریم

$$(\lambda - ۱)\omega + ۱ = \pm\mu\sqrt{-\lambda}. \tag{۱۲}$$

قرار می‌دهیم $f_{\omega}(\lambda) = (\lambda - ۱)\omega + ۱$ و $g(\lambda) = \pm\mu\sqrt{-\lambda}$ به ازای یک ω نمودار این دو تابع در شکل ۱ رسم شده‌اند. به وضوح $f(۱) = ۱$. توجه می‌کنیم که شیب تابع f با افزایش ω افزایش می‌یابد. با توجه به شکل ۱ ، با افزایش ω اندازه ریشه‌ها بزرگ می‌شوند، تاجایی که دو نمودار بر هم مماس گردند. در این صورت $\lambda_{۱} = \lambda_{۲}$ و بنابراین $\Delta = ۰$، که نتیجه می‌دهد $\mu = ۰$ یا $۴\omega^{۲} - ۴\omega - \mu^{۲} = ۰$. با توجه به اینکه W یک ماتریس SPD است، پس $\mu = ۰$ قابل قبول نمی‌باشد. در نتیجه $۴\omega^{۲} - ۴\omega - \mu^{۲} = ۰$. معادله درجه دوم اخیر شامل دو ریشه به صورت $\omega_{۱,۲} = \frac{۱}{۲}\left(۱ \pm \sqrt{۱ + \mu^{۲}}\right)$ است. با



شکل ۱: نمودار توابع $f_\omega(\lambda)$ و $g(\lambda)$

توجه به اینکه $w > \frac{۱+\mu}{۲}$، بنابراین ریشه قابل قبول به صورت $\tilde{\omega} = \frac{۱}{۲}\left(۱+\sqrt{۱+\mu^۲}\right)$، می‌باشد و به سادگی می‌توان دید $\lambda_{۱,۲}(\tilde{\omega}) = ۱ - \frac{۱}{\tilde{\omega}}$. در ادامه فرض می‌کنیم $\omega > \tilde{\omega}$. در این حالت، ریشه‌های معادله (۹)، مختلط و به صورت

$$\lambda_{۱,۲}(\omega) = \frac{(۲\omega^۲ - ۲\omega - \mu^۲)}{۲\omega^۲} \pm i\frac{\sqrt{\Delta'}}{۲},$$

خواهند بود، به طوری که

$$\Delta' = ۴\left(\frac{\omega - ۱}{\omega}\right)^۲ - \left(\frac{۲\omega^۲ - ۲\omega - \mu^۲}{\omega^۲}\right)^۲.$$

لذا $|\lambda_{۱,۲}| = ۱ - \frac{۱}{\omega}$. با توجه به اینکه $۱ - \frac{۱}{\tilde{\omega}} < ۱ - \frac{۱}{\omega}$، پس بهترین انتخاب برای ω مقدار $\tilde{\omega}$ می‌باشد. از سوی دیگر، منحنی g، یک کران بالا برای همه منحنی‌ها به شکل $\pm\mu\sqrt{-\lambda}, ۰ \leq \lambda \leq \rho(W)$ است. با جمع‌بندی مطالب فوق داریم:

$$\rho\left(\mathcal{G}_{\omega^*}\right) = \min_{\omega} \max_{\omega > \frac{۱+\mu_{\max}}{۲}} |۱ - \frac{۱}{\omega}| = ۱ - \frac{۱}{\omega^*} = \left(\frac{\rho(W)}{۱+\sqrt{۱+\rho^۲(W)}}\right)^۲,$$

که ω^* در رابطه (۱۱) تعریف شده است. □

۳ نتایج عددی

در این بخش، یک مثال عددی برای بررسی کارایی روش ارائه‌شده در نظر گرفته شده است. در این مثال، روش ارائه‌شده را با روش‌های گاوس-سایدل بلوکی و روش GMRES با شروع مجدد ۱۰، مقایسه می‌کنیم. بردار حدس اولیه را بردار تصادفی انتخاب می‌کنیم. قرار می‌دهیم $Res = \frac{\|r_k\|_۲}{\|r_۰\|_۲}$، به‌طوری‌که $r_k = \mathcal{P} - \mathcal{A}\mathbf{u}^{(k)}$ و $u^{(k)}$ جواب حاصل از تکرار $k-$ام می‌باشد. در صورتی‌که $Res < ۱۰^{-۹}$، یا تعداد تکرارها به ۱۰۰۰ برسد، تکرارها متوقف می‌شوند. نتایج عددی حاصل از اجرای برنامه‌ها در Matlab-R۲۰۱۷ بر لپ‌تاپ با مشخصات

intel (R), Core(TM) i5-8265U, CPU @ 1.60 GHz, 8 GB,

است.

مثال ۱.۳. زوج معادلات شرودینگر

$$\begin{cases} iu_t + (-\Delta)^{\frac{\alpha}{۲}}u + ۲\left(|u|^۲ + |v|^۲\right)u = ۰, \\ iv_t + (-\Delta)^{\frac{\alpha}{۲}}v + ۲\left(|v|^۲ + |u|^۲\right)v = ۰, \end{cases} \tag{۱۳}$$

را در نظر می‌گیریم، به‌طوری‌که

$$\begin{cases} u(x,۰) = \text{sech}(x+۱)\cdot\exp(۲ix), v(x,۰) = \text{sech}(x-۱)\cdot\exp(-۲ix), \\ u(-۲۰,۰) = u(۲۰,۰) = ۰, v(-۲۰,۰) = v(۲۰,۰) = ۰. \end{cases}$$

جدول ۱: مقادیر ω^* با فرض $\alpha = ۱/۳$ و $N = ۴M$، متناظر با ماتریس B.

M	۸۰۰	۱۶۰۰	۳۲۰۰	۶۴۰۰
ω^*	۱/۰۰۳	۱/۰۰۴	۱/۰۰۷	۱/۰۱۰

جدول ۲: نتایج عددی با فرض $\alpha = ۱/۳$ و $N = ۴M$.

Method	M	۸۰۰		۱۶۰۰		۳۲۰۰		۶۴۰۰	
		B	C	B	C	B	C	B	C
BGSOR	IT	۵	۵	۵	۵	۵	۵	۵	۵
	CPU	۰/۰۱۱	۰/۰۰۹	۰/۰۵۰	۰/۰۲۱	۰/۱۹۰	۰/۱۴۰	۰/۹۷۳	۰/۸۷۸
BGS	IT	۵	۵	۶	۶	۶	۶	۷	۷
	CPU	۰/۰۱۶	۰/۰۱۲	۰/۰۵۸	۰/۰۶۵	۰/۲۱۹	۰/۲۳۶	۱/۶۴۰	۱/۰۴۸
GMRES(۱۰)	IT	۶	۶	۷	۷	۷	۷	۸	۸
	CPU	۰/۰۵۷	۰/۰۱۴	۰/۰۷۲	۰/۰۳۹	۰/۱۱۵	۰/۱۰۹	۲/۳۳۵	۱/۹۱۴

فرض کنید $\gamma = ۱/۳$ و $\rho = ۱/۲$. زوج دستگاه (۱۳)، منجر به تولید دو دستگاه معادلات خطی می‌شود. در این مثال، دستگاه‌های حاصل از گسسته‌سازی زوج معادلات شرودینگر غیرخطی مکان کسری به ازای گام زمانی نهایی $t = ۲$، در نظر گرفته شده‌اند. ماتریس ضرایب این دستگاه‌ها را B و C می‌نامیم. این دو ماتریس ساختار مشابهی دارند. مقادیر بهینه پارامتر فوق‌تخفیف متناظر با ماتریس B، در جدول ۱ ارائه شده‌اند. با توجه به تشابه ساختاری دو ماتریس از ارائه ω^*، متناظر با ماتریس C خودداری می‌کنیم. نتایج عددی حاصل از سه روش GMRES(10)، گاوس-سایدل بلوکی (BGS) و فوق‌تخفیف گاوس-سایدل بلوکی (BGSOR) در جدول ۲، ارائه شده‌اند. در این جدول "IT" و "CPU" ، به ترتیب نشان‌دهنده زمان لازم برای همگرایی روش ها و تعداد تکرارها می‌باشند. با توجه به جدول ۲، می توان دریافت که روش ارائه شده بر روش‌های پیشین برتری مطلق دارد.

۴ نتیجه‌گیری

در این مقاله، روش تکراری فوق‌تخفیف گاوس-سایدل بلوکی برای حل زوج معادلات شرودینگر غیرخطی مکان کسری بکار گرفته شده است. همگرایی روش مذکور مورد بررسی قرار گرفت. همانطور که مشاهده شد، روش ارائه‌شده تحت شرایط مناسبی همگرا می‌شود. یکی از مزیت‌های این روش، حجم محاسبات اندک برای محاسبه جواب دستگاه مورد نظر می‌باشد. به‌علاوه، مقدار بهینه پارامتر فوق تخفیف محاسبه گردید. در نهایت، با یک مثال عددی برتری روش نسبت به سایر روش‌ها مورد بررسی قرار گرفت.

مراجع

[1] A. Atangana and A.H. Cloot, Stability and convergence of the space fractional variable-order Schrödinger equation, *Adv. Differential Equations,* 2013 (2013), 80.

[2] P. Amore, F.M. Fernández, C.P. Hofmann and R.A. Sáenz, Collocation method for fractional quantum mechanics, *J. Math. Phys.*, 51 (2010), 122101.

[3] P. Dai and Q. Wu, An efficient block Gauss–Seidel iteration method for the space fractional coupled nonlinear Schrödinger equations, *Appl. Math. Lett.*, 117 (2021), 107-116.

[4] F. Demengel and G. Demengel, Fractional sobolev spaces, in: *Functional Spaces for the Theory of Elliptic Partial Differential Equations,* Springer, London, (2012), 179–228.

[5] D. Wang, A. Xiao and W. Yang, A linearly implicit conservative difference scheme for the space fractional coupled nonlinear Schrödinger equations, *J. Comput. Phys.*, 272 (2014), 644-655.

[6] D. M. Young, *Iterative Solution or Large Linear Systems*, Academic Press, New York, 1971.

کاربرد تقریب کم-رتبه نیستروم در محاسبه هسته متعامد لژاندر جهت افزایش دقت تشخیص زودهنگام کووید-۱۹

علیرضا افضل آقائی نائینی [۱]، * و کورش پرند [۱]

[۱] گروه علوم داده‌ها و کامپیوتر، دانشکده ریاضی، دانشگاه شهید بهشتی، تهران، ایران

چکیده

همه‌گیری کووید-۱۹ به یکی از بزرگ‌ترین چالش‌های قرن بیست و یکم تبدیل شده است. علائم رایج این بیماری، شامل تب، سرفه و تنگی نفس می‌باشد. شباهت این علائم به آنفولانزای فصلی، امکان تشخیص سریع و کم‌هزینه کووید-۱۹ را از متخصصان سلب کرده است. در این پژوهش یک الگوریتم هوشمند جهت دسته‌بندی افراد مبتلا به کووید-۱۹ از میان افراد مشکوک ارائه می‌شود. برای این منظور، ابتدا کاربرد چندجمله‌ای‌های متعامد به عنوان توابع هسته بررسی شده و سپس جهت غلبه بر پیچیدگی‌های محاسباتی، از تقریب کم-رتبه نیستروم و تجزیه مقادیر منفرد استفاده می‌کنیم. در پایان به کمک دسته‌بند گرادیان کاهشی، یک مدل بلادرنگ جهت تشخیص بیماری ارائه خواهیم کرد. نتایج حاصله بیانگر دقت ۹۸ درصدی الگوریتم ارائه شده می‌باشد.

واژه‌های کلیدی: کووید-۱۹، تجزیه مقادیر منفرد، چندجمله‌ای‌های متعامد، یادگیری ماشین، هوش مصنوعی
کد موضوع‌بندی ریاضی [۲۰۱۰]: 15A23، 68T10

۱ مقدمه

کووید-۱۹ (COVID-19) یک بیماری عفونی ناشی از کروناویروس سندرم حاد تنفسی ۲ (SARS-CoV-2) است که در دسامبر سال ۲۰۱۹ در ووهان چین پدیدار شد. این بیماری واگیردار با تحت تاثیر قرار دادن سیستم تنفسی افراد، منجر به بروز علائمی همچون مشکلات تنفسی، تب و سرفه خشک می‌گردد. معمولا افرادی که سیستم ایمنی قوی دارند، پس از ابتلا به این بیماری، بدون انجام درمان خاصی در نهایت بهبود می‌یابند. با این حال، افراد مسن و کسانی که دارای بیماری‌های زمینه‌ای همچون بیماری‌های قلبی-عروقی، دیابت، بیماری‌های تنفسی و سرطان هستند، آسیب پذیری بیشتری دارند. بر اساس یافته‌های سازمان بهداشت جهانی علائم کمتر شایع این ویروس، شامل سردرد، گلودرد، اسهال، از دست دادن بویایی و از دست دادن گفتار می‌باشد.

علی‌رغم تلاش‌های جامعه برای مهار ویروس، کووید-۱۹ با جهش‌های متوالی همچنان در حال گسترش است، شیوع بیماری می‌تواند منجر به افزایش تقاضا در منابع بیمارستانی و کمبود تجهیزات پزشکی، کارکنان مراقبت‌های بهداشتی و البته کیت‌های تست کووید-۱۹ شود. دسترسی محدود به این کیت‌ها می‌تواند تشخیص زودهنگام بیماری را مختل کرده و ارائه بهترین مراقبت ممکن برای بیماران مشکوک را سخت کند. در نتیجه، وجود یک سیستم پیش‌بینی خودکار که هدف آن تعیین وجود ویروس کووید-۱۹ در یک فرد مشکوک است ضروریست.

الگوریتم‌های یادگیری ماشین، از ابزارهای مفید و کاربردی در تشخیص زودهنگام انواع بیماری‌های مختلف همچون دیابت، سرطان و... هستند. این الگوریتم‌ها با دریافت علائم بیماری، با دقت مناسب به پیش‌بینی وجود یا عدم وجود بیماری در بدن فرد می‌پردازند. دانشمندان این مدل‌ها را جهت تشخیص بیماری کووید-۱۹ نیز توسعه داده‌اند [۳، ۶]. در این رویکردها، الگوریتم به‌طور صریح تعریف نشده و به مدل اجازه داده می‌شود برخی از بخش‌های خود را با کسب تجربه به‌روز کرده و نتیجه نهایی را بهبود دهد. در اینجا منظور از تجربه یک مجموعه داده شامل چند نمونه مختلف بوده که هرکدام دارای چندین ویژگی مستقل و یک برچسب صحیح می‌باشند.

* سخنران. آدرس ایمیل: alirezaafzalaghaei@gmail.com

در میان رویکردهای حل مسائل یادگیری ماشین، روش‌های مبتنی بر هسته، دسته‌ای از الگوریتم‌ها هستند که مدل‌سازی خود را بر اساس شباهت میان داده‌ها انجام می‌دهند. در این روش‌ها، تابع هسته داده‌ها را به طور ضمنی به یک فضای با ابعاد بالا نگاشت کرده و عملیات یادگیری در آن فضا انجام می‌شود. از معروف ترین توابع هسته می‌توان به هسته خطی، چندجمله‌ای و تابع پایه شعاعی گوسی اشاره کرد. اخیرا نشان داده شده است که چندجمله‌ای‌های متعامد نیز می‌توانند به عنوان تابع هسته به کار گرفته شده و دقت مدل یادگیری ماشین را افزایش دهند [۵].

استفاده از توابع متعامد به عنوان هسته در مدل‌های یادگیری ماشین، منجر به افزایش پیچیدگی محاسباتی و زمان اجرا شده است. در این پژوهش به کمک روش‌های تقریب کم-رتبه، یک تقریب از ماتریس هسته را بدست آورده و به تعیین وجود یا عدم وجود ویروس کووید-۱۹ در افراد مشکوک می‌پردازیم. به این منظور در بخش بعد به معرفی پیش‌نیازهای لازم پرداخته و در بخش سوم مجموعه داده مورد استفاده تحقیق را شرح می‌دهیم. در بخش چهارم یک الگوریتم سریع برای تعیین وجود ویروس کووید-۱۹ در بدن افراد ارائه خواهیم کرد. در پایان نیز با ارائه نتایج و جدول یافته‌های تحقیق به نتیجه گیری می‌پردازیم.

۲ پیش نیازها

در این بخش به معرفی توابع متعامد و چندجمله‌ای‌های لژاندر، روش تقریب کم-رتبه نیستروم برای ماتریس‌های نیمه‌معین مثبت متقارن و همچنین دسته‌بند گرادیان کاهشی می‌پردازیم.

۱.۲ چندجمله‌ای‌های متعامد

چندجمله‌ای‌های متعامد دنباله‌ای نامتناهی از چندجمله‌ای‌هایی است که دو به دو بر یکدیگر عمودند. خاصیت عمود بودن توابع به کمک تعریف ضرب داخلی

$$\langle f, g \rangle_w = \int_a^b f(x)g(x)w(x)dx \tag{۱}$$

بیان می‌شود. چندجمله‌ای‌های لژاندر، چبیشف، ژاکوبی و گگنبایر از معروف ترین توابع متعامد بوده که در بازه‌ی $[-۱, ۱]$ براساس تابع وزن مربوطه خاصیت تعامد دارند. چندجمله‌ای‌های لژاندر به دلیل داشتن تابع وزن واحد $w(x) \equiv ۱$ در بین پژوهشگران از محبوبیت خاصی برخوردار است. این چندجمله‌ای‌ها با فرمول بازگشتی زیر تعریف می‌گردند.

$$\begin{aligned} &P_۰(x) = ۱, \quad P_۱(x) = x, \\ &(n+۱)P_{n+۱}(x) = (۲n+۱)xP_n(x) - nP_{n-۱}(x), \quad n \geq ۱. \end{aligned} \tag{۲}$$

جهت تعریف این رده از توابع در بازه‌ی متناهی دلخواه $[a, b]$ می‌توان از نگاشت $\varphi(x) = {}^{(۲x-a-b)}/_{(b-a)}$ استفاده کرد و چندجمله‌ای‌های لژاندر انتقال یافته را به صورت $\tilde{P}(x) = P(\varphi(x))$ محاسبه نمود. جهت مطالعه بیشتر در مورد این توابع خواننده را به مرجع [۱، ۵] ارجاع می‌دهیم. لازم به ذکر است این توابع در افزایش سرعت حل دستگاه‌های معادلات خطی نیز به کار گرفته شده‌اند [۲].

۲.۲ روش تقریب هسته نیستروم

در روش‌های مبتنی بر هسته برای یک مجموعه داده‌ی $D = \{(x_i, y_i)\}_{i=۱}^N$ ماتریس نیمه‌معین مثبت متقارن $\Omega_{N\times N}$ به صورت

$$\Omega_{i,j} = K(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle, \qquad i, j = ۱, ۲, \ldots, N. \tag{۳}$$

تعریف خواهد شد. در اینجا تابع $K(x, t)$ هسته‌ی مدل یادگیری ماشین است. این تابع، ضرب داخلی تابع گسترش ویژگی $\varphi : \mathbb{R}^d \to \mathbb{R}^k$، که در آن $k > d$ است را به صورت ضمنی محاسبه می‌کند. در مسائلی که تعداد داده‌های مسئله بزرگ باشد، بدست آوردن ماتریس Ω از نظر محاسباتی کارا نیست. از این رو دانشمندان رویکردهایی جهت تقریب این ماتریس، با تعداد محدودی از داده‌ها ارائه کرده‌اند. یکی از معروف ترین این روش‌ها، تقریب کم-رتبه نیستروم می‌باشد. در قضیه‌ی ۱.۲ به معرفی این روش می‌پردازیم.

قضیه ۱.۲. *ماتریس نیمه‌معین مثبت متقارن $A \in \mathbb{R}^{n\times n}$ را در نظر بگیرید. فرض کنید $C_{n\times m}$ یک ماتریس شامل $m \ll n$ ستون از ماتریس A باشد که به صورت تصادفی انتخاب شده است. در این صورت می‌توان ماتریس A و C را به صورت زیر بازنویسی کرد.*

$$C = \begin{bmatrix} W \\ S \end{bmatrix} \text{ و } A = \begin{bmatrix} W & S^T \\ S & B \end{bmatrix}. \tag{۴}$$

در اینجا $W \in \mathbb{R}^{m\times m}, S \in \mathbb{R}^{(n-m)\times m}$ و $B \in \mathbb{R}^{(n-m)\times(n-m)}$ می‌باشد. در این صورت ماتریس W نیز نیمه‌معین مثبت خواهد بود. حال برای هر $k \leq m$ تقریب نیستروم مرتبه k ماتریس A به صورت

$$\hat{A}_k = CW_k^{\dagger}C^T \tag{۵}$$

تعریف می‌شود. در اینجا W_k بهترین تقریب با رتبه k ماتریس W و $W_k^{\dagger}$ شبه‌معکوس مور-پنروز این ماتریس است.

اثبات. برای اثبات به مرجع [۴] مراجعه کنید. □

با استفاده از این قضیه می‌توان یک تقریب کم-رتبه از ماتریس هسته Ω ارائه کرد. برای این منظور ابتدا یک مجموعه اندیس $I \subseteq \{۱, ۲, \ldots, N\}$ انتخاب کرده و ماتریس

$$W_{i,j} = K(x_i, x_j), \quad i, j \in I \tag{۶}$$

را تشکیل می‌دهیم. جهت محاسبه بهترین تقریب مرتبه k ماتریس W از تجزیه مقادیر منفرد استفاده می‌شود.

$$W = U\Sigma U^T \tag{۷}$$

در اینجا $U, V \in \mathbb{R}^{m\times m}$ ماتریس‌های متعامد و Σ نیز یک ماتریس قطری است که عناصر آن شامل مقادیر منفرد ماتریس W بوده که به صورت نزولی مرتب شده‌اند. ماتریس کم-رتبه W_k با استفاده از این تجزیه تقریب زده می‌شود:

$$W_k = U_k\Sigma_k U_k^T. \tag{۸}$$

شبه‌معکوس ماتریس W نیز به کمک $r \leq k$ مقدار منفرد غیر صفر ماتریس W_k محاسبه می‌گردد:

$$W_k^{\dagger} = \sum_{i=۱}^{r} \sigma_i^{-۱} U_{:,i} U_{:,i}^T. \tag{۹}$$

در اینجا σ_i بیانگر $i-$امین بزرگترین مقدار منفرد ماتریس W و $U_{:,i}$ بیانگر $i-$امین ستون ماتریس متعامد U است. برای تقریب ماتریس هسته Ω لازم است ماتریس C نیز تعریف شود، برای این منظور

$$C_{i,j} = K(x_i, x_j), \quad i \in \{۱, ۲, \ldots, N, j \in I\} \tag{۱۰}$$

را توسط مجموعه اندیس‌های I به دست می‌آوریم. حال به کمک معادله (۵) می‌توان تقریب کم-رتبه ماتریس هسته را بدست آورد.

۳.۲ دسته‌بند خطی گرادیان کاهشی

دسته‌بند گرادیان کاهشی یک مدل خطی در یادگیری ماشین است که به جهت مقیاس پذیری و کارایی محاسباتی مورد توجه دانشمندان قرار گرفته است. در این مدل برای مجموعه داده‌ی $D = \{(x_i, y_i)\}_{i=۱}^{N}$ ، که در آن $x_i \in \mathbb{R}^m, y_i \in \{-۱, ۱\}$ است، مدل خطی

$$y(x) = sign(w^T x + b) \tag{۱۱}$$

با پارامترهای مجهول w, b در نظر گرفته می‌شود. با مفروض بودن تابع خطای محدب $L : \mathbb{R}^{۲} \to \mathbb{R}$ که با دریافت پاسخ صحیح و پیش‌بینی مدل، خطای پیش‌بینی را محاسبه می‌کند، الگوریتم گرادیان کاهشی به صورت

$$w_i^{\tau+۱} = w_i^{\tau} - \eta * \frac{\partial L}{\partial w_i} \tag{۱۲}$$

نوشته می‌شود. با اجرای این الگوریتم به تعداد تکرار مشخصی، پارامترهای بهینه برای پیش‌بینی دسته‌بند به دست می‌آید. انتخاب توابع خطای مختلف، مدل‌های خطی یادگیری ماشین همچون رگرسیون لاجستیک و یا ماشین بردار پشتیبان را شبیه سازی می‌کند.

۳ تشخیص کووید-۱۹

در این بخش پس از معرفی مجموعه داده‌ی مورد استفاده، به ارائه روش پیشنهادی بر اساس توابع متعامد لژاندر و الگوریتم دسته‌بند گرادیان کاهشی می‌پردازیم.

۱.۳ مجموعه داده

در این پژوهش از مجموعه داده متن باز Symptoms and COVID Presence [۱] استفاده می‌کنیم. این مجموعه شامل ۵۴۳۴ نمونه‌ی مختلف است که هرکدام شامل ۲۰ ویژگی مستقل و یک ویژگی وابسته است. از این تعداد ۴۳۸۳ نفر مبتلا به کووید-۱۹ بوده درحالی که ۱۰۵۱ نفر علائم این بیماری را داشته اما به کووید-۱۹ مبتلا نیستند. در جدول ۱ ویژگی‌های این مجموعه داده، شرح داده شده است.

نام مشخصه	درصد مبتلایان به کووید-۱۹	نام مشخصه	درصد مبتلایان به کووید-۱۹
مشکل تنفسی	۷۶٫۹٪	احساس خستگی	۵۰٫۸٪
تب	۸۵٫۷٪	مشکل گوارشی	۴۶٫۹٪
سرفه خشک	۸۸٫۵٪	سفر خارج از کشور	۵۵٫۹٪
گلو درد	۸۳٫۷٪	تماس با بیمار کووید-۱۹	۵۸٫۹٪
آبریزش بینی	۵۴٫۲٪	شرکت در گردهمایی بزرگ	۵۵٫۷٪
آسم	۴۸٫۵٪	بازدید از مکان‌های عمومی	۵۴٫۸٪
بیماری مزمن ریه	۴۵٫۸٪	اشتغال اعضای خانواده در مکان‌های عمومی	۴۵٫۵٪
سردرد	۴۹٫۷٪	استفاده از ماسک	–
بیماری قلبی	۴۷٫۱٪	استفاده از محصولات ضدعفونی	–
دیابت	۴۸٫۶٪	مبتلا به کووید-۱۹	۱۰۰٪
فشار خون بالا	۵۱٫۵٪		

جدول ۱: ویژگی‌های مجموعه داده و درصد افرادی که دارای هر یک از علائم فوق بوده و مبتلا به کووید-۱۹ می‌باشند.

پیش پردازش داده‌های ورودی جهت آموزش یک مدل یادگیری ماشین، تاثیر زیادی در بدست آوردن دقت مناسب مدل دارد. پس از انجام عملیات پیش پردازش این داده‌ها، دریافتیم که دو ویژگی استفاده از ماسک و استفاده از محصولات ضدعفونی، برای تمامی نمونه‌ها دارای مقدار خیر بوده است. از این رو این دو ویژگی از مجموعه داده حذف می‌گردد.

۲.۳ چندجمله‌ای‌های لژاندر به عنوان هسته

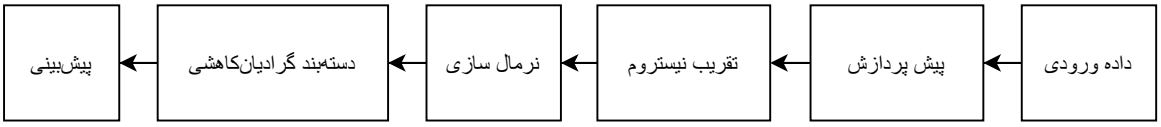
انتخاب درست توابع هسته در مدل‌های یادگیری ماشین می‌تواند تاثیر زیادی بر دقت مدل داشته باشد [۶]. یک شرط لازم و کافی برای قابل استفاده بودن تابع دلخواه $K(x,t)$ به عنوان تابع هسته مدل یادگیری ماشین، توسط مرسر [۲] ارائه شده است [۵]. توابع هسته رایج در زمینه یادگیری ماشین که در شرط مرسر نیز صدق می‌کنند عبارت‌اند از:

- هسته خطی: $K(x,t) = x^T t$
- هسته چندجمله‌ای: $K(x,t;d) = (۱ + x^T t)^d$
- هسته پایه شعاعی: $K(x,t;\gamma) = \exp(\gamma\|x-t\|)$

در این پژوهش چندجمله‌ای‌های متعامد لژاندر انتقال یافته به بازه $[۰, ۱]$ از درجه $q-۱$ به عنوان تابع هسته مدل‌های یادگیری ماشین به کار گرفته می‌شوند. این توابع طبق تعریف هسته، برای ورودی با تک ویژگی به‌صورت

$$K(x,t;q) = \langle \mathbf{P}_q(x), \mathbf{P}_q(t)\rangle = \sum_{i=۰}^{q} P_i(x)P_i(t), \qquad (۱۳)$$

[۱] https://www.kaggle.com/hemanthhari/symptoms-and-covid-presence
[۲] Mercer's theorem



شکل ۱: الگوریتم تشخیص ابتلا به کووید-۱۹

تعریف می‌شوند. همچنین برای مجموعه داده با چند ویژگی $x, t \in \mathbb{R}^d$ با استفاده از خواص تابع هسته، هسته‌ی جدید به صورت

$$K(x,t;q) = \prod_{i=1}^{d} K(x_i,t_i) = \prod_{i=1}^{d} \sum_{j=0}^{q} P_j(x_i)P_j(t_i),$$

فرموله بندی می‌شود. لازم به ذکر است ورودی‌های مسئله نیز باید در بازه $[0,1]$ مقیاس بندی شده باشد [۵].

۴ مدل پیشنهادی

در این پژوهش جهت تشخیص ابتلا به بیماری کووید-۱۹ الگوریتم ۱ را پیشنهاد می‌کنیم. این الگوریتم در فاز اولیه به کمک پارامترهای ذکر شده در جدول ۲ آموزش می‌بینند. برای این منظور داده‌های ورودی پس از انجام پیش‌پردازش و افزایش ویژگی نرمال شده و سپس به الگوریتم دسته‌بند گرادیان کاهشی داده می‌شوند. نرمال سازی ویژگی‌ها با حذف میانگین و تغییر مقیاس واریانس به یک انجام می‌شود. استفاده از تابع خطای hinge به همراه روش منظم سازی تیخونوف یا L2 در مدل دسته‌بند گرادیان کاهشی باعث می‌شود الگوریتم ارائه شده همانند فرم اولیه مسئله ماشین بردار پشتیبان عمل کند [۵]. تابع خطای hinge به صورت

$$L(y_i, f(x_i)) = \max(0, 1 - y_i f(x_i)), \tag{۱۴}$$

تعریف می‌گردد. در اینجا $y_i = \pm 1$ برچسب صحیح نمونه x_i و $f(x_i)$ پیش‌بینی دسته‌بند گرادیان کاهشی برای این نمونه می‌باشد.

در فاز آزمایش مدل، از پارامترهای بدست آمده در فاز آموزشی استفاده می‌کنیم. برای این منظور هر داده‌ی جدید پس از انجام پیش پردازش لازم و افزایش ویژگی به کمک تقریب نیستروم، به کمک پارامترهای میانگین و واریانس داده‌های آموزشی نرمال می‌شود. سپس با استفاده از وزن‌های بدست آمده توسط فاز اولیه الگوریتم گرادیان کاهشی، دسته‌بند پیش‌بینی خود را ارائه می‌کند.

در بخش بعد با ذکر روش تنظیم ابر-پارامترهای مسئله و معیارهای بررسی دقت مدل، به مقایسه روش ارائه شده با دیگر روش‌های موجود می‌پردازیم.

۱.۴ نتایج

یافتن ابر-پارامترهای بهینه برای مدل به‌وسیله‌ی اعتبارسنجی متقابل 5-Fold انجام شده است. پس از انجام آزمایشات مختلف، پارامترهای مدل پیشنهادی مطابق با جدول ۲ بدست آمد. جدول ۴ زمان لازم برای محاسبه ماتریس هسته به روش ساده و همچنین روش تقریب هسته نیستروم به ازای مقادیر متفاوتی از تعداد کامپوننت‌ها (نمونه‌های مورد استفاده جهت تقریب) آورده شده است. مشاهده می‌شود با انتخاب زیرمجموعه کوچکی از داده‌های آموزشی، می‌توان ماتریس هسته را با دقت مناسب در زمان کوتاه تر تقریب زد. در جدول ۵ با انتخاب ٪۲۰ از داده‌ها به عنوان مجموعه داده‌های آزمایشی، معیارهای دقت، صحت، بازخوانی و F1 گزارش شده است. با در نظر گرفتن مثبت صادق (TP) ، منفی صادق (TN) ،مثبت کاذب (FP) و منفی کاذب (FN) این معیارها به صورت زیر تعریف می‌شوند:

$$Precision = \frac{TP}{TP+FP}, \quad Recall = \frac{TP}{TP+FN}, \quad F1 = \frac{2*(Recall*Precision)}{Recall+Precision} \tag{۱۵}$$

در جدول ۳ دقت مدل ارائه شده با رویکردهای مختلف مقایسه شده است. مشاهده می‌شود مدل ارائه شده دقت بالایی در تشخیص افراد مبتلا به کووید-۱۹ در میان افراد مشکوک دارد. لازم به ذکر است تعیین وضعیت بیماری افراد در فاز تست بلادرنگ بوده و به کمتر از یک صدم ثانیه زمان نیاز دارد.

مدل	نام پارامتر	مقدار پارامتر
نیستروم	تعداد کامپونت‌ها	۵۰۰
دسته‌بند	منظم‌سازی	L2
دسته‌بند	ضریب منظم‌سازی	$۱۰^{-۶}$
دسته‌بند	تابع خطا	hinge
الگوریتم گرادیان کاهشی	طول گام	تطبیقی
الگوریتم گرادیان کاهشی	حداکثر تکرار	۱۰۰۰

جدول ۲: پارامترهای مورد استفاده برای آموزش مدل

هسته	پارامترها	دقت
خطی [۳]	-	۹۶٫۸۰٪
چندجمله‌ای [۶]	-	۹۵٫۴۸٪
چندجمله‌ای نرمال شده [۶]	-	۹۵٫۳۹٪
تابع پایه شعاعی	$\gamma = ۰٫۰۱$	۹۶٫۴۰٪
تابع پایه شعاعی	$\gamma = ۰٫۱$	۹۷٫۸۲٪
چندجمله‌ای لژاندر	$q = ۳$	۹۷٫۷۰٪

جدول ۳: تاثیر پارامتر هسته در دقت مدل ماشین بردار پشتیبان

روش	تعداد کامپونت‌ها	زمان(ثانیه)
محاسبه کل هسته	-	۲۱٫۱
نیستروم	۲۰۰	۱٫۶۵
نیستروم	۵۰۰	۴٫۳۸

جدول ۴: مقایسه سرعت اجرای محاسبه هسته و تقریب نیستروم

	صحت	بازخوانی	F1	تعداد داده‌ها
غیرکووید-۱۹	۰٫۹۳	۰٫۹۷	۰٫۹۵	۲۱۰
کووید-۱۹	۰٫۹۹	۰٫۹۸	۰٫۹۹	۸۷۷
دقت			۰٫۹۸	۱۰۸۷

جدول ۵: نتایج دسته‌بند پیشنهادی

۵ نتیجه‌گیری

در این تحقیق، با ترکیب روش‌های جبرخطی و توابع متعامد، یک روش کارآمد جهت حل مسائل یادگیری ماشین ارائه شد. در این رویکرد به کمک تجزیه مقادیر منفرد یک تقریب از ماتریس هسته بدست آمد. سپس به کمک روش دسته‌بند گرادیان کاهشی یک الگوریتم جهت تشخیص وجود یا عدم وجود ویروس کووید-۱۹ در افراد مشکوک ارائه شد. بررسی دقت مدل بر روی یک مجموعه داده شناخته شده انجام شد و نتایج بدست آمده بیانگر دقت بالای مدل در تشخیص این ویروس بود.

مراجع

[۱] ک. پرند، م. رزاقی، ج. امانی راد، م. دلخوش، م. معیری. رویکردهای نوین از روش‌های طیفی در محاسبات علمی: نظریه و کاربردها. دانشگاه شهید بهشتی؛ ۱۳۹۸.

[۲] م. حجاریان. ،نخستین درس در جبر خطی عددی، انتشارات شهید بهشتی، تهران، ۱۳۹۷.

[3] Aliyu S, Zakari AS, Adeyanju I, Ajoge NS. A Bayesian Network Model for the Prognosis of the Novel Coronavirus (COVID-19). In International Conference on Computational Science and Its Applications 2021 Sep 13; (pp. 127-140).

[4] Drineas P, Mahoney MW, Cristianini N. On the Nystrom Method for Approximating a Gram Matrix for Improved Kernel-Based Learning. In journal of machine learning research. 2005 Dec 1;6(12).

[5] Parand K, Aghaei AA, Jani M, Ghodsi A. A new approach to the numerical solution of Fredholm integral equations using least squares-support vector regression. In Mathematics and Computers in Simulation 2021 Feb 1;180(pp. 114-28).

[6] Villavicencio CN, Macrohon JJ, Inbaraj XA, Jeng JH, Hsieh JG. COVID-19 Prediction applying supervised machine learning algorithms with comparative analysis using WEKA. In Algorithms. 2021 Jul;14(7) No. 201.

استفاده از تجزیه مقادیر منفرد در شبکه عصبی پیچشی برای بهبود دقت در تشخیص تومور مغزی

مریم بابائی[۱*] ، زینب حاجی‌محمدی[۱] و کورش پرند[۱]

[۱]گروه علوم داده و کامپیوتر،دانشکده ریاضی، دانشگاه شهید بهشتی، تهران، ایران

چکیده

تومور مغزی تشکیل شده از سلول‌هایی است که نشان دهنده‌ی رشد بی‌رویه در مغز می‌باشند. آشکارسازي دقیق و به موقع ناحیه تومور مغزي در انتخاب نوع درمان و دنبال کردن روند بیماري در طول درمان تاثیر بسزایی دارد. در عین حال تصاویر ام‌آرآی مغزی با نویزهایی همراه هستند. از بین بردن نویزهای موجود می‌تواند در تشخیص بهتر تومورهای مغزی تاثیر بسزایی داشته باشد. در این پژوهش سعی شده‌است با استفاده از تجزیه مقادیر ویژه و انتخاب تعداد مشخصی از مقادیر منفرد، نویز های موجور در تصاویر را کاهش داده و سپس با استفاده از شبکه عصبی عمیق به تشخیص تومور در تصاویر بپردازیم. نتایج پیاده‌سازي ها نشان مي‌دهند که استفاده از تصاویر با مقادیر نویز کمتر دقت بدست آمده در شبکه های عصبی را افزایش می‌دهد.

واژه‌های کلیدی: تجزیه مقادیر منفرد، شبکه عصبی مصنوعی، تومور مغزی، ام‌آرآی مغز

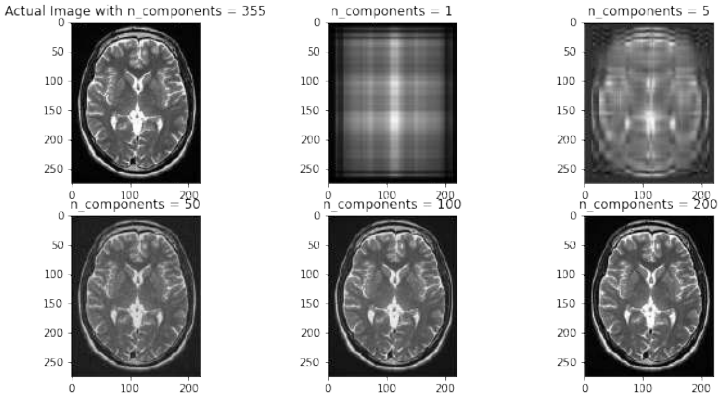
کد موضوع‌بندی ریاضی [۲۰۱۰]: 15A23

۱ مقدمه

تومور مغزی یک توده غیر طبیعی در مغز است که می‌تواند ماهیت سرطانی (بدخیم) یا غیرسرطانی (خوش‌خیم) داشته باشد. عواملی مانند نوع، محل، اندازه و نحوه‌ی گسترش تومور در میزان تهدیدکنندگی یک تومور تاثیر دارد. تشخیص تومور مغزی به دلیل اندازه، شکل و محل ظاهر شدن در مغز، کار پیچیده‌ای است. امروزه با پیشرفت الگوریتم‌های بخش‌بندی تصویر امکان تشخیص تومور مغزی به صورت خودکار نیز فراهم شده‌است. این روش ها امکان ذخیره سازی وضعیت رشد تومور در طول مدت درمان و کاهش خطای انسانی را فراهم می‌کنند. اما به دلیل گوناگی تومورها رسیدن به دقت بالا در تشخیص تومور بسیار دشوار است. یکی از روش‌های موجود برای تشخیص تومورهای مغزی در تصاویر ام‌آرآی (MRI) ،استفاده از شبکه عصبی پیچشی است. شبکه عصبی پیچشی (CNN) یکی از روش‌های عمیق یادگیری ماشین است که با توجه به کارایی بالای آن، استفاده از آن در استخراج ویژگی و دسته‌بندی تصاویر امر بسیار رایجی شده است. کیفیت تصاویر ورودی در دقت تشخیص تصاویر تاثیر بسزایی دارد. لذا حذف نویزهای موجود در تصاویر عملکرد الگوریتم‌های مورد استفاده را افزایش می‌دهد. یکی از روش‌های حذف نویز استفاده از تجزیه مقادیر منفرد است. فیلتر (k-SVD) در کنار حذف نویزهای تصویر، کیفیت تصاویر را نیز حفظ کند. از این رویکرد می‌توان در زمینه‌ای مختلف از جمله کاهش نویز تصاویر ام‌آرآی و سیتی اسکن استفاده کرد.

در این مقاله در ابتدا سعی شده است با انتخاب تعداد مناسب مقادیر منفرد در تجزیه مقادیر منفرد تصاویر ام‌آرآی مغزی، حجم تصاویر و همچنین نویزهای موجود را کاهش دهیم و سپس با استفاده از یک شبکه عصبی پیچشی به تشخیص تومور مغزی در تصاویر بپردازیم.

[*]سخنران. آدرس ایمیل: maryam.01997@gmail.com



شکل ۱: تأثیر استفاده از روش k-SVD بر روی تصاویر امآرآی با انتخاب مقادیر k متفاوت

۲ تجزیه مقادیر منفرد

تجزیه مقادیر منفرد [۱] نقش بسیار مهمی در زمینه‌های مختلف جبر خطی عددی ایفا می‌کند. به دلیل خواص و کارایی این تجزیه، بسیاری از الگوریتم‌ها و روش‌های مدرن مبتنی بر تجزیه مقادیر منفرد. تاکنون روش‌های مختلفی برای بیان تجزیه مقادیر منفرد ارائه شده‌است که در آن‌ها یک ماتریس را به سه ماتریس دیگر تجزیه می‌کند. برای مثال، ماتریس A را می‌توان به صورت زیر نوشت:

$$A = USV^T$$

که در آن $A \in \mathbb{R}^{m\times n}$؛ $V \in \mathbb{R}^{n\times n}$ و $U \in \mathbb{R}^{m\times m}$ ماتریس‌های متعامد و $S \in \mathbb{R}^{m\times m}$ یک ماتریس قطری به فرم زیر است:

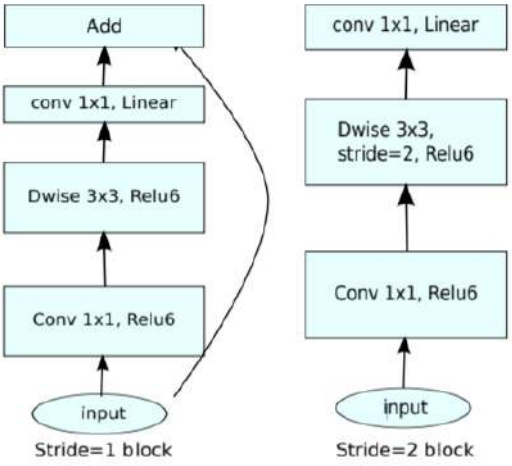
$$S = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & \sigma_n \end{bmatrix}$$

به طوریکه $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$. از دلایل کاربردهای متعدد SVD نیاز به شناخت رتبه و تقریب یک ماتریس با استفاده از ماتریس‌های با رتبه‌های کمتر و پایه‌های متعامد است. یکی از کاربردهای این تجزیه استفاده از این روش در پردازش تصویر برای فشرده‌سازی تصویر، حذف نویز و خش‌ها در یک عکس و ... است. معمولا ماتریس‌های متناظر با تصاویر دارای تعداد زیادی مقادیر تکین کوچک اند. فرض کنید $(n-k)$ مقدار تکین کوچک از ماتریس A وجود داشته باشد که قابل صرف نظر است. آنگاه ماتریس $\sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \ldots + \sigma_k u_k v_k^T$ یک تقریب خیلی خوب از A است، و یک چنین تقریبی می‌تواند در کاربردها کافی باشد. در شکل (۱) تاثیر اعمال این روش با انتخاب مقادیر k متفاوت بر روی تصاویر امآرآی مغزی نشان داده شده است.

۳ شبکه عصبی پیچشی

شبکه عصبی مصنوعی [۲] از بین روش‌های یادگیری ماشین روشی پرطرفدار است که مکانیسم یادگیری در موجودات بیولوژیکی را شبیه‌سازی می‌کند. در یک شبکه عصبی مصنوعی با ایجاد شبکه‌ای بین گره‌هایی به نام نورون و اعمال یک الگوریتم آموزشی به آن، شبکه را آموزش می‌دهند. در این سیستم، ورودی‌ها در شبکه به جریان افتاده و یک سری خروجی تولید می‌گردد. سپس خروجی‌ها با داده‌های معتبر مقایسه می‌گردند. سپس مسیرهایی که به تشخیص موارد درست منجر شده، از طرف شبکه تقویت خواهند شد و مسیرهایی با تشخیص اشتباه اصلاح می‌شوند. با تکرار این فرایند در دفعات زیاد، شبکه نهایتا قادر است به دقت بسیار خوبی در اجرای وظیفه موردنظر دست یابد. یکي از گام‌هاي مهم در توسعه شبکه‌های عصبی مصنوعی، طراحي معماری شبکه است که تاثیر

[۱] Singular value decomposition
[۲] Artificial neural network



شکل ۲: معماری کلی شبکه‌های عصبی موبایل نت نسخه دوم

زیادی بر عملکرد شبکه دارد. در طراحی معماری شبکه‌های عصبی مصنوعی، عواملي از قبیل تعداد لایه‌های پنهان، تعداد نورون‌ها در هر لایه، توابع تبدیل و الگوریتم آموزش باید تعیین شوند.

یکی از معماری‌های رایج شبکه عصبی پیچشی نام دارد. شبکه‌های عصبی پیچشی [۱] ، رده‌ای از شبکه‌های عصبی عمیق هستند که معمولاً برای انجام تحلیل‌های تصویری یا گفتاری در یادگیری ماشین استفاده می‌شوند. انگیزه بنیادی برای این شبکه‌ها از شناخت کارکرد بخش بینایی مغز گربه بدست آمد که در بخش‌های خاص میدان دید به نظر می‌رسید نورون‌های خاصی را تحریک می‌کنند. این قاعده به طور گسترده‌تر برای طراحی معماری این شبکه‌ها مورد استفاده قرار گرفت. در معماری شبکه عصبی پیچشی، هر لایه شبکه ۳-بعدی بوده، که محدوده فضایی و عمق متناظر به تعداد ویژگی دارد. از لحاظ تاریخی شبکه‌های عصبی پیچشی موفق‌ترین نوع شبکه‌های عصبی بوده‌اند. این شبکه‌ها به طور گسترده‌ای برای شناسایی تصویر، محلی‌سازی شیء و حتی پردازش متن مورد استفاده قرار گرفته‌اند. عملکرد این شبکه‌ها اخیرا از انسان‌ها در مسئله طبقه‌بندی تصویر فراتر رفته‌است.

۴ شبکه عصبی پیچشی در تشخیص تومور

در این پژوهش در ابتدا سعی شده است با استفاده از روش k-SVD نویزهای موجود در تصاویر را کاهش دهیم و تصاویری با اختلالات کمتر جهت پردازش و کلاس بندی در اختیار شبکه عصبی پیچشی طراحی شده قرار دهیم. سپس برای پردازش این تصاویر از شبکه عصبی موبایل نت نسخه دوم [۲] که در سال ۲۰۱۸ معرفی شده است، استفاده کرده‌ایم. این شبکه در کنار سایز کوچک سرعت بالایی در پردازش دارد و قابلیت استفاده در رباتیک، بردهای مینی کامپیوتری موبایل‌ها را فراهم می‌کند. شبکه عصبی موبایل نت بستر سخت افزاری سبک‌تر و پارامترهای کمتری دارد (در مقایسه با سایر شبکه های عمیق). این ویژگی ها در مصارف پزشکی که سخت افزار و داده‌های آموزشی محدود است؛ قابل توجه و مهم هستند، و همچنین از ویژگی بافت الگوی دودویی محلی که در دسته بندی تصاویر پزشکی و غیرپزشکی پرکاربرد می باشد، در این شبکه استفاده شده است. معماری شبکه عصبی موبایل نت در شکل (۲) آورده شده است.

به طور کلی ساختار پیشنهادی شامل یک بخش برای کاهش نویزهای تصویر و بخش دیگر یک شبکه عصبی موبایل نت برای دسته بندی تصاویر به کلاس‌های موارد دارای تومور مغزی و موارد فاقد تومور، می‌شود.

[۱] Convolutional neural network book
[۲] MobileNetV2

۵ نتایج عددی

برای بررسی کارایی روش پیشنهادی از تصاویر ام‌آرآی مغزی شامل ۳۰۰۰ تصویر با دسته‌بندی تصاویر همراه با تومور مغزی و فاقد تومور با اندازه‌ی ی (۲۵۸،۳۱۴) پیکسل، استفاده کرده‌ایم. پس از آن تصاویر بهبود یافته شده با استفاده از روش k-SVD با مقدار k= ۲۰۰ را با استفاده از شبکه عصبی موبایل نت با ۱۵ تکرار پردازش کرده‌ایم و دقت بدست آمده از آن را با دقت بدست آمده از آموزش شبکه با داده‌ای خام برای داده‌های آزمایشی و داده‌های آزمون در جدول ۱ مقایسه کرده‌ایم. معماری شبکه‌ی طراحی شده در شکل ۳ قابل مشاهده است.

Output size	Layer
224×224	Image
112×112	3×3 Conv, 32
56×56	3×3 DWConv, 32
	1×1 Conv, 128
28×28	3×3 DWConv, 64
	1×1 Conv, 128
	3×3 DWConv, 128
	1×1 Conv, 128
14×14	3×3 DWConv, 128
	1×1 Conv, 256
	3×3 DWConv, 256
	1×1 Conv, 256
7×7	3×3 DWConv, 256
	1×1 Conv, 512
	3×3 DWConv, 512 4× 1×1 Conv, 512
	3×3 DWConv, 512
	1×1 Conv, 1024
1×1	Global Maxpoling
	Softmax

شکل ۳: معماری شبکه موبایل نت طراحی شده برای تصاویر ام‌آرآی مغزی

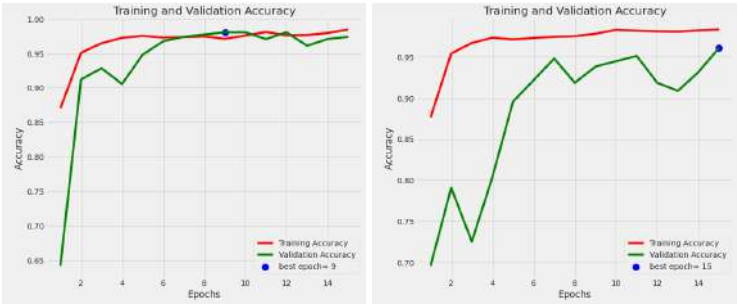
جدول ۱: مقایسه تابع هزینه و دقت بدست آمده از شبکه موبایل نت برای داده‌های خام و داده‌های پردازش شده توسط روش $k-SVD$

نوع داده	تابع هزینه	دقت داده‌های آزمون	دقت داده‌های آزمایشی
داده‌های خام	۰.۲۷۷	۰.۹۶۰۷	۰.۹۵۸۴
داده‌های پردازش شده توسط روش $k-SVD$	۰.۲۱۳	۰.۹۷۳۹	۰.۹۷۲۴

نمودارهای موجود در تصویر ۴ تاثیر پردازش تصاویر با میزان نویز کمتر بر همگرایی و دقت شبکه را نشان می‌دهد. مطابق تصویر زمانیکه از تصاویر بدست آمده توسط روش k-SVD استفاده می‌کنیم در تعداد دوره‌های کمتری به دقت بالا می‌رسیم. همچنین دقت شبکه برای داده‌های اعتبارسنجی نیز نسبت به مدل با ورودی‌های خام افزایش یافته‌است.

۶ نتیجه‌گیری

یکی از چالش‌های موجود در پردازش و تشخیص بیماری‌ها با توجه به تصاویر ام‌آرآی و سیتی اسکن، کاهش نویزهای موجود در تصاویر است که برای تحقق آن می‌توان از تجزیه مقادیر منفرد استفاده کرد. در این پژوهش تاثیر استفاده از روش k-SVD در سرعت و دقت تشخیص در یک شبکه عصبی پیچشی بررسی شده است. نتایج بدست‌آمده نشان می‌دهد، استفاده از تجزیه مقادیر منفرد برای از بین بردن نویز تصاویر زمان اجرای هر دوره از یادگیری شبکه عصبی کاهش داده است و در کنار آن سرعت همگرایی شبکه به میزان چشم گیری افزایش یافته است.



شکل ۴: مقایسه دقت و سرعت همگرایی در داده‌های آزمون و اعتبارسنجی برای شبکه با داده های پیش پردازش شده توسط روش k-SVD در سمت چپ و داده‌های بدون تغییرات در سمت راست

مراجع

[۱] م. حجاریان، نخستین درس در جبر خطی عددی، انتشارات شهید بهشتی، تهران؛ ۱۳۹۷.

[2] I. J. Goodfellow, Y. Bengio and A. Courville, Deep Learning ,MIT Press, 2016.

[3] T. Workalemahu, Singular Value Decomposition in Image Noise Filtering and Reconstruction, Thesis, Georgia State University, 2008.

[4] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, LC. Chen, MobileNetV2: Inverted Residuals and Linear Bottlenecks,Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 4510-4520.

کاربرد تجزیه مقادیر منفرد تصادفی در تحلیل مولفه‌های اصلی به منظور افزایش سرعت آموزش در روش‌های یادگیری ماشین

زهرا بهروزه، [۱] * مسعود حجاریان، [۲] علیرضا افضل آقائی نائینی [۱]
[۱] گروه علوم داده ها و کامپیوتر، دانشکده ریاضی، دانشگاه شهید بهشتی، تهران، ایران
[۲] گروه علوم ریاضی، دانشکده ریاضی، دانشگاه شهید بهشتی، تهران، ایران

چکیده

الگوریتم‌های یادگیری ماشین بر اساس مجموعه داده مسئله، یک مدل ریاضیاتی ایجاد کرده و پس از آموزش مدل، به پیش‌بینی داده‌های جدید می‌پردازند. افزایش ابعاد مجموعه داده، یادگیری این مدل‌ها را با مشکلاتی مواجه کرده است. تحلیل مولفه‌های اصلی یکی از راهکارهای رفع این مشکل می‌باشد. این الگوریتم بر اساس تجزیه مقادیر منفرد با ترکیب ویژگی‌های مجموعه داده یک تقریب از فضای اصلی مسئله ارائه می‌کند. محاسبه این تجزیه در مسائل با ابعاد بالا خود دارای مشکلاتی است. در این تحقیق پس از معرفی تجزیه مقادیر منفرد تصادفی به بررسی کاربردهای این الگوریتم در افزایش سرعت یادگیری مدل ماشین بردار پشتیبان در چند مجموعه داده شناخته شده می‌پردازیم.

واژه‌های کلیدی: یادگیری ماشین، تحلیل مولفه‌های اصلی، تجزیه مقادیر منفرد تصادفی، ماشین بردار پشتیبان
کد موضوع‌بندی ریاضی [۲۰۱۰]: 15A23، 68T10

۱ مقدمه

آرتور ساموئل در سال ۱۹۵۹ لفظ یادگیری ماشین را ابداع نمود. او معتقد بود یادگیری ماشین، حوزه‌ای از تحقیقات است که در آن کامپیوترها بدون برنامه‌نویسی توانایی یادگیری دارند. واقعیت امر به این صورت است که دانش یادگیری ماشین سعی بر این دارد تا با استفاده از الگوریتم‌ها، یک مدل را به گونه‌ای طراحی کند که این مدل بتواند آموزش ببیند و به درستی عمل کند. یادگیری ماشین به دسته‌های مختلفی همچون یادگیری نظارت شده، یادگیری بدون نظارت و یادگیری نیمه نظارتی تقسیم می‌شود. در این روش‌ها رایج است که یک مجموعه داده، مشخص شده و سپس الگوریتم‌هایی جهت یادگیری الگوهای پنهان در میان این داده‌ها استفاده می‌شود. این مجموعه شامل تعدادی نمونه بوده که هر کدام از آن‌ها دارای ویژگی‌هایی هستند. بدین ترتیب این مجموعه داده دارای دو شاخصه مهم، به نام تعداد نمونه‌ها و تعداد ویژگی‌ها می‌باشد.تعداد ویژگی در برخی مسائل بسیار بالا بوده و برای یافتن الگو در این داده‌ها نیازمند الگوریتم‌های خاصی هستیم تا بتوانیم مشکلاتی همچون کمبود حافظه و پیچیدگی محاسباتی را کاهش دهیم. علاوه بر مشکلات محاسباتی، در این مجموعه داده‌ها پدیده مشقت بُعدچندی یا نفرین ابعاد نیز ظاهر می‌گردد. این پدیده بیانگر آن است که با افزایش ابعاد، حجم فضا آنقدر سریع افزایش می‌یابد که داده‌های موجود پراکنده و تُنُک می‌شوند. همچنین با افزایش ابعاد لازم است داده‌های مورد نیاز به‌طور نمایی افزایش یابند تا نتیجه حاصله از نظر آماری معقول و معتبر باشد. برای رفع این مشکل دانشمندان بر اساس رویکردهای نظارت شده و بدون نظارت ایده‌های مختلفی جهت کاهش تعداد ویژگی‌ها ارائه کرده‌اند. برای حل این مشکل دانشمندان دو رویکرد استخراج و انتخاب ویژگی را ارائه دادند. در حالی که انتخاب ویژگی هدف انتخاب زیرمجموعه‌ای از ویژگی داده‌ها را دارد، استخراج ویژگی با ترکیب ویژگی‌های موجود، ویژگی‌های جدید را تولید می‌کند. در تمامی روش‌های انتخاب و استخراج ویژگی، رویکرد های جبرخطی به کاربرده میشود. روش تحلیل مولفه‌های اصلی یکی از پرکاربردترین روش‌های استخراج ویژگی می‌باشد. این رویکرد با یافتن یک پایه متعامد جدید برای فضای ویژگی‌ها، تعداد ویژگی‌های مدل را کاهش می‌دهد. یافتن این پایه متعامد به کمک روش تجزیه مقادیر منفرد انجام می‌گردد. تجزیه مقادیر منفرد یک رویکرد ماتریسی است که ماتریس را به سه ماتریس دیگر تجزیه

* سخنران. آدرس ایمیل: zahrabehrouzeh@gmail.com

می‌کند. محاسبه این تجزیه برای مسائل در مقیاس بزرگ از نظر زمانی بهینه نیست. الگوریتم تجزیه مقادیر منفرد تصادفی جهت رفع این مشکل ارائه شده است. در ادامه به بررسی روش‌های تحلیل مولفه‌های اصلی، تجزیه مقادیر منفرد و تجزیه مقادیر منفرد تصادفی می‌پردازیم و کاربرد تجزیه مقادیر منفرد تصادفی در تحلیل مولفه‌های اصلی برای چند مجموعه داده مختلف را به کمک ماشین بردار پشتیبان بررسی می‌نماییم.

۲ پیش نیازها

در این قسمت به بررسی مفاهیم به کاربرده شده در این پژوهش از جمله تحلیل مولفه‌های اصلی، تجزیه مقادیر منفرد، تجزیه مقادیر منفرد تصادفی و ماشین بردار پشتیبان می‌پردازیم.

۱.۲ تحلیل مولفه‌های اصلی (PCA)

تحلیل مولفه‌های اصلی یکی از روش‌های پرکاربرد استخراج ویژگی است. در این رویکرد تلاش می‌شود به کمک ترکیب خطی ویژگی‌های موجود، ویژگی‌های جدیدی تولید شود. این شیوه با یافتن جهت‌های بیشترین واریانس در داده‌ها، یک دستگاه مختصات جدید ارائه می‌کند. در این دستگاه جدید اولین بُعد متناظر با جهت بیشترین واریانس و اخرین بُعد بیانگر جهت کمترین واریانس در میان داده‌ها است. از این رو با انتخاب k بُعد ابتدای این دستگاه می‌توان ویژگی های استخراج شده و مفید کمتری بدست آورد. در ادامه به شرح این روش می‌پردازیم . [۴]

قضیه ۱.۲. *فرض کنید مجموعه داده X با ابعاد $n \times d$ با میانگین ستونی صفر داده شده است. در این صورت جهت‌های بیشترین واریانس بردارهای داده‌ها، توسط بردارهای ویژه ماتریس کواریانس ارائه می‌شود.*

اثبات. با فرض صفر بودن میانگین ویژگی‌های مجموعه داده، ماتریس کواریانس به صورت

$$C = \frac{1}{n}\sum_{i=0}^{n} x_i x_i^T$$

نمایش داده خواهد شد. فرض کنید بردار u_1 جهت بیشترین واریانس داده ها را نمایش دهد. برای یافتن این بردار مسئله بهینه‌سازی زیر را تشکیل می‌دهیم.

$$\begin{array}{ll} \max\limits_{u_1} & \frac{1}{n} u_1 C u_1 \\ \text{s.t.} & \|u_1\| = 1 \end{array} \qquad (1)$$

محاسبه فرم دوگان ۱ مسئله مقدار ویژه‌ی زیر را نتیجه می‌دهد.

$$C \cdot u_1 = \lambda \cdot u_1 \qquad (2)$$

بنابراین برای محاسبه جهت بیشترین واریانس، بزرگترین مقدار ویژه و بردارهای ویژه متناظر آن از ماتریس کوراپانس انتخاب می‌شوند. به منظور یافتن دومین جهت بیشترین واریانس، دومین بزرگترین مقدار ویژه محاسبه میگردد. لازم به ذکر است محاسبه مسئله مقدار ویژه ماتریس کواریانس برابر با تجزیه مقادیر منفرد ماتریس داده X می‌باشد. [۴] □

۲.۲ تجزیه مقادیر منفرد (SVD)

در بسیاری از مسائل یادگیری ماشین ما با ماتریس‌هایی سر و کار داریم که دارای سطر و ستون‌های بسیاری می‌باشند در واقع انجام عملیات برروی آن‌ها بسیار هزینه بالای دارد. روش تجزیه مقادیر منفر در جبر خطی پاسخی برای ساده‌سازی این مشکل می‌باشد. تجزیه مقادیر منفرد ماتریس A به روشی گفته می‌شود که ماتریس مذکور را به سه ماتریس بالا مثلثی، قطری و پایین مثلثی تجزیه می‌کند. [۱]حال به شرح مختصر این روش می‌پردازیم. فرض کنید ماتریس $A_{m\times n}$ را داریم و مقادیر m, n بسیار بزرگ می‌باشند. آنگاه تجزیه مقادیر منفرد ماتریس A به صورت زیر است:

$$A_{m\times n} = \mathrm{U}\Sigma\mathrm{V}^{\mathrm{T}}$$

U یک ماتریس $m \times m$ متعامد، Σ یک ماتریس قطری $m \times n$ و V^T یک ماتریس $n \times n$ می‌باشد. از آن‌جایی که ماتریس U و V متعامد می‌باشند داریم:

$$U^T U = V V^T = I$$

و هم‌چنین به سادگی ثابت می‌شود که مقادیر روی قطر اصلی Σبرابر است با مقادیر ویژه ماتریس $A^T A$ درنتیجه با معلوم بودن ماتریس Σو A می‌توان U،V را به سادگی با جای‌گذاری در فرمول‌های زیر محاسبه نمود :

$$A^T A V = V \Sigma^2 \qquad\qquad A A^T U = U \Sigma^2$$

۳.۲ تجزیه مقادیر منفرد تصادفی

رویکرد اصلی در روش تجزیه مقادیر منفرد تصادفی. [۳] به این صورت می‌باشد که ماتریس با ابعاد بزرگ‌تر را به ماتریس با ابعاد کوچک‌تر تقریب می‌زنیم و سپس این ماتریس تقریب را به کمک SVD تجزیه می‌کنیم و در نهایت مقدار SVD ماتریس بزرگتر را به کمک SVD ماتریس کوچک تقریب می‌زنیم. ماتریس $A_{m\times n}$ را درنظر بگیریم. می‌خواهیم ماتریس‌های $B_{m\times k}$ و$C_{k\times n}$ را به شرطی تعریف کنیم که $k \ll n$ و A≈BC باشند. برای تقریب این محاسبه با استفاده از الگوریتم های تصادفی، دانشمندان یک محاسبه دو مرحله ای را پیشنهاد می کنند:

قدم اول: یک تقریب پایه با برد ماتریس A محاسبه می‌کنیم. می‌خواهیم یک ماتریس Q متعامد و نرمال با l ستون تعریف کنیم به طوری که عملA را به تصویر می‌کشد. در در واقع $QQ^*A \simeq A$ و Q را به گونه‌ای تعریف می‌کنیم که نامساوی زیر به ازای ϵ مشخص برقرار باشد.

$$||I - QQ^*A|| \leq \epsilon$$

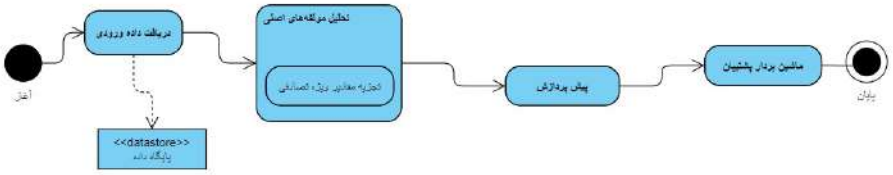
الگوریتم ۲.۲. الگوریتم قدم اول

- تعریف یک ابرپارامتر l برای نمونه‌برداری بیش از اندازه
- ساخت یک ماتریس Ω گاوسی تصادفی با ابعاد $n \times l$
- ایجاد ماتریس $Y = A\Omega$ با ابعاد $m \times l$
- ساخت یک ماتریس متعامد نرمال Q برای مثال با استفاده از روش فاکتورگیری QR ، $Y = QR$

قدم دوم: در این مرحله قصد داریم تجزیه مقادیر منفرد تصادفی ماتریس Q که از ماتریس A ابعاد کوچک‌تری دارد را محاسبه می‌کنیم. سپس به وسیله تجزیه مقادیر منفرد ماتریس Q به تجزیه مقادیر منفرد ماتریس A را به دست می‌آوریم.

الگوریتم ۳.۲. الگوریتم قدم دوم

- تعریف ماتریس B به صورت $B = Q^*A$
- محاسبه تجزیه مقادیر منفرد ماتریس $B = S\Sigma V^T$
- از آن‌جایی که A≈QQ^*A پس $A = QB = QS\Sigma V^T$
- تعریف U به طوری که: $U = QS$
- در نهایت U, Σ, V با تجزیه مقادیر منفرد ماتریس A .



شکل ۱: نمودار الگوریتم

۴.۲ ماشین بردار پشتیبان

ماشین بردار پشتیبان یکی از شناخته شده ترین الگوریتم های یادگیری ماشین در حوزه یادگیری نظارت شده و یادگیری بدون نظارت است. [۲]. این الگوریتم، رویکردی افتراقی دارد. مدل ماشین بردار پشتیبان که بر پایه ی تئوری یادگیری آماری و قضیه VC بنا شده است یکی از قدرتمند ترین مدل‌های یادگیری ماشین است.در این مدل فرض می شود هر نمونه داده ورودی به صورت $x_i \in \mathbb{R}^d$ و خروجی مطلوب آن برای مسائل رگرسیون $y \in \mathbb{R}$ و برای طبقه‌بندی $y \in \{-1, +1\}$ است. در مدل دسته بند، الگوریتم تلاش می کند معادله بهترین ابرصفحه جدا کننده را به گونه ای به دست آورد که جداکننده حاصل از داده های دو دسته بیشترین حاشیه را داشته باشد.. [۲] در ادامه ماشین بردار پشتیبان، با استفاده از ترفند هسته این ابرصفحه را به یک رویه تبدیل کرده و عملیات پیش بینی را انجام می دهد. به این منظور دسته بند به صورت:

$$y(x) = \text{sign}\left[w^T\varphi(x) + b\right]$$

در نظر گرفته شده و جهت به دست آوردن ابرصفحه بهینه، مسئله ی بهینه سازی محدب زیر تشکیل می‌شود:

$$\begin{aligned} \min w, b, \xi \quad & \frac{1}{2}w^T w + C\sum_{k=1}^{n}\xi_k \\ \text{s.t.} \quad & y_k\left[w^T\varphi(x_k) + b\right] \geq 1 - \xi_k, \qquad k = 1, \ldots, n. \\ & \xi_k \geq 0, \qquad k = 1, \ldots, n. \end{aligned}$$

فرم دوگان این مسئله بهینه سازی نیز به کمک ضرایب لاگرانژ محاسبه می‌شود که در آن Ω همان ماتریس هسته می‌باشد.

$$\begin{aligned} \max_{\alpha} \quad & \frac{1}{2}\alpha^T\Omega\alpha - 1^T\alpha \\ \text{s.t.} \quad & 0 \leq \alpha \leq C \\ & y^T\alpha = 0 \end{aligned}$$

$$\Omega_{i,j} = y_i y_j \varphi(x_i)^T\varphi(x_j), \quad i, j = 1, 2, \ldots, n.$$

در این صورت دسته بند مدل در فرم دوگان به صورت زیر خواهد بود:

$$y(x) = \text{sign}\left[\sum_{k=1}^{n}\alpha_k y_k \varphi(x_k)^T\varphi(x) + b\right]$$

۳ کاربرد تجزیه مقادیر منفرد تصادفی در الگوریتم تحلیل مولفه‌های اصلی

در این پژوهش برای محاسبه تجزیه مقادیر منفرد در الگوریتم تحلیل مولفه‌های اصلی از تجزیه مقادیر منفرد تصادفی استفاده می‌کنیم.در این بخش به معرفی این مدل و نتایج حاصل از به کارگیری این مدل بر روی مجموعه داده‌های انتخابی می‌پردازیم.

مجموعه داده	تعداد داده‌های آموزشی	تعداد داده‌های آزمایشی	تعداد ویژگی
Scania Trucks	۶۰۰۰۰	۱۶۰۰۰	۱۷۱
Fashion-MNIST	۶۰۰۰۰	۱۰۰۰۰	۷۸۴
MNIST	۶۰۰۰۰	۱۰۰۰۰	۷۸۴

جدول ۱: جدول داده‌ها

۱.۳ مدل پیشنهادی

حال می‌خواهیم به بررسی مدل به کار برده شده در این پژوهش بپردازیم. نمودار فعالیت الگوریتم این مدل در نمودار الگوریتم آمده است. ابتدا مجموعه داده ورودی از پایگاه داده خوانده می‌شود. سپس در مرحله بعدی الگوریتم تحلیل مولفه‌های اصلی که پیش از این بررسی نمودیم بر روی پایگاه داده اعمال می‌شود. الگوریتم تحلیل مولفه‌های اصلی که در این مدل به کار می‌بریم پایه متعامد را به کمک تجزیه مقادیر تصادفی محاسبه می‌نماید. پس از محاسبه فضای داده جدید توسط تحلیل مولفه‌های اصلی پیش پردازش‌های مورد نیاز برای عملکرد هرچه بهتر الگوریتم بر روی مجموعه داده فعلی اعمال می‌شود. حال پس از پیش‌پردازش الگوریتم SVM بر روی مجموعه داده اعمال می‌کنیم.

۲.۳ نتایج

به منظور سنجش عملکرد صحیح الگوریتم پیشنهادی در این تحقیق از سه مجموعه داده معتبر در سایت Kaggle استفاده نمودیم. ابتدا به شرح مختصر این مجموعه داده‌ها می پردازیم.

مجموعه داده Scania Trucks : این مجموعه داده شامل داده‌های جمع آوری شده مربوط به سیستم فشار هوای کامیون‌های اسکانیا می‌باشد. این سیستم در خودرو هوای مورد نیاز برای عملکردهای مختلف مانند ترمزگیری و تعویض دنده و غیره را تولید می‌کند. کلاس مثبت با ۱۰۰۰ رکورد شامل کامیون‌هایی با خرابی اجزا برای یک جزء خاص از سیستم فشار هوا است. کلاس منفی با ۵۹۰۰۰ رکورد شامل کامیون‌هایی با خرابی قطعات غیر مرتبط با سیستم فشار هوا است. هر رکورد نیز شامل ۱۷۱ ویژگی می‌باشد.

مجموعه داده Fashion-MNIST : این مجموعه داده عکس‌های مقاله زولاندو شامل ۶۰۰۰۰ داده آموزشی و ۱۰۰۰۰ داده آزمایشی می‌باشد. هر رکورد یک عکس با ابعاد ۲۸ × ۲۸ پیکسل است که هر کدام متعلق به یکی از ۱۰ کلاس برچسب‌ها هستند. این برچسب‌ها نام نوعی از لباس مانند : تیشرت، پلیور و ... هستند.

مجموعه داده MNIST : این مجموعه داده شامل داده‌های MNIST می‌باشد. شامل ۶۰۰۰۰ داده آموزشی و ۱۰۰۰۰ داده آزمایشی می‌باشند. هر رکورد تصویری ۲۸ × ۲۸ پیکسل می‌باشد که متعلق به کلاس یکی از اعداد ۰ الی ۹ می‌باشد.

نتایج حاصل از عملکرد مدل در سه حالت بدون استفاده از تحلیل مولفه‌های اصلی، با استفاده از تحلیل مولفه‌های اصلی بدون به کارگیری تجزیه مقادیر منفرد و با استفاده از تحلیل مولفه‌های اصلی با به کارگیری تجزیه مقادیر منفرد تصادفی را در جدول نتایج مشاهده می‌فرمایید. .همان‌طور که مشاهده می‌کنید در هر سه مجموعه داده بعد از استفاده از PCA با کاهش زیاد زمان رو به رو هستیم. هم‌چنین در زمان استفاده از تجزیه مقادیر منفرد تصادفی در PCA برای داده‌های با ابعاد بزرگتر این زمان باز هم کاهش می‌یابد و تغییر ناچیزی در دقت اتفاق می‌افتد که قابل چشم‌پوشی است.

۴ نتیجه‌گیری

هنگامی که تعداد ویژگی‌ها در برخی از مسائل بسیار بالا بوده دچار مشکلات زمان، حافظه و پدیده‌ای مانند مشقت چندبعدی یا نفرین ابعاد می‌شویم. پدیده مشقت چندبعدی نیز به همراه خود مشکلاتی همچون کاهش سرعت و دقت مدل‌های یادگیری ماشین را خواهد داشت. در این تحقیق تلاش نمودیم با استفاده از روش تجزیه مقادیر منفرد تصادفی در الگوریتم تحلیل مولفه‌های اصلی کاهش ابعاد انجام دهیم. نتایج بدست آمده حاکی از کارایی بالای این روش در سرعت و زمان با حفظ دقت می‌باشد..

مجموعه داده	بدون اعمال PCA		PCA بدون اعمال SVD تصادفی		PCA با اعمال SVD تصادفی	
	زمان	دقت	زمان	دقت	زمان	دقت
Trucks Scania	۳۳,۵۴	۹۱۸,۸,۹۸	۹۲,۰	۷۷۵۰,۹۸	۹۹,۰	۸۶۸۸,۹۸
MNIST	۵۹,۶۷۹	۱۹۰۰,۹۱	۰۱,۷	۳۲۰۰,۹۱	۷۴,۱	۶۲۰۰,۹۰
MNIST Fashion	۴۹,۷۳۸	۷۴۰۰,۸۳	۹۶,۷	۹۲۰۰,۸۴	۷۲,۱	۳۹۰۰,۸۳

جدول ۲: جدول نتایج

مراجع

[۱] م. حجاریان. ،نخستین درس در جبر خطی عددی، انتشارات شهید بهشتی، تهران، ۱۳۹۷.

[2] Cortes, Corinna, and Vladimir Vapnik. ”Support-vector networks.” Machine learning 20.3 (1995): 273-297.

[3] N. Halko, P. G. Martinsson, J. A. Tropp. Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions. SIAM Review.2011.

[4] S Wold, K Esbensen, P Geladi. Principal component analysis. Chemometrics and Intelligent Laboratory Systems.1987 August.

فشرده‌سازی تصویر با استفاده از تجزیه مقدار تکین، کدگذاری کوتاه کردن بلوک و فیلتر پایین‌گذر

ساجده خیراندیش[۱] *، سید ابوالفضل شاهزاده‌فاضلی[۱] و اعظم صادقیان[۱]

[۱]بخش علوم کامپیوتر، دانشگاه یزد، یزد، ایران

چکیده

در دنیای امروز به‌دلیل حجم عظیم اطلاعات، نیاز به فشرده‌سازی اطلاعات اهمیت بیشتری پیدا کرده است. در بحث فشرده‌سازی تصویر، فشرده‌سازی باید به‌گونه‌ای باشد که تصویر دچار افت کیفیت نشود. تجزیه مقدار تکین و کدگذاری کوتاه کردن بلوک، ازجمله روش‌های فشرده‌سازی تصویر به‌صورت بااتلاف هستند که در این مقاله توسط این دو روش و همچنین ترکیب این دو روش، تصویر فشرده می‌شود و در نهایت فیلتر پایین‌گذر به‌روی تصویر فشرده‌شده اعمال می‌شود. نتایج نشان می‌دهد که فیلتر پایین‌گذر در بهبود کیفیت تصاویر فشرده‌شده مؤثر است و می‌تواند میزان خطای تصویر فشرده‌شده را کاهش دهد.

واژه‌های کلیدی: تجزیه مقدار تکین، فشرده‌سازی تصویر بااتلاف، فیلتر پایین‌گذر، کدگذاری کوتاه‌کردن بلوک

کد موضوع‌بندی ریاضی [۲۰۱۰]: 68W99.

۱ مقدمه

فشرده‌سازی تصویر به‌عنوان یک نیاز اساسی در سیستم‌های بایگانی و ارسال تصویر مطرح می‌شود. روش‌های زیادی برای افزایش نرخ فشرده‌سازی تصویر وجود دارد اما افزایش نرخ فشرده‌سازی و حفظ کیفیت تصویر فشرده‌شده یک ضرورت محسوب می‌شود. هر یک از روش‌های فشرده‌سازی به‌ازای تصاویر مختلف، نرخ فشرده‌سازی یکسانی را ارائه نمی‌دهند. نرخ‌های فشرده‌سازی تصویر به عواملی چون تنوع سطوح خاکستری و یکسان بودن مقادیر پیکسل‌های مجاور در تصویر بستگی دارد. فشرده‌سازی تصویر روشی برای کاهش حجم داده‌های یک تصویر می‌باشد. اصل عملکرد فشرده‌سازی براساس حذف داده‌های اضافی و زائد صورت می‌گیرد، به‌طوریکه کیفیت تصویر نیز کاهش پیدا نکند.

El Asnaoui [۲] فشرده‌سازی تصویر با استفاده از روش توانی تجزیه مقدار تکین بلوکی را پیشنهاد داده است که با یافتن یک رتبه مناسب برای ماتریس، تصویرها را به‌صورتی فشرده می‌کند که علاوه‌بر کیفیت بصری، فضای کمتری را در حافظه اشغال کند.

تکنیک ارائه شده توسط El Asnaoui و همکاران [۳] دو روش جدید را معرفی می‌کند. روش اول بهبود روش کدگذاری کوتاه کردن بلوک است که بر معایب کدگذاری کوتاه کردن بلوک کلاسیک غلبه می‌کند، و روش دوم نحوه به‌دست آوردن رتبه جدید از روش تجزیه مقدار تکین[۱] را توصیف می‌کند؛ که فشرده‌سازی تصویر بهتری را نتیجه می‌دهد [۴].

Doaa و Chadi [۱] کدگذاری کوتاه کردن بلوک[۲] را پیشنهاد کردند، که یکی از ساده‌ترین و راحت‌ترین الگوریتم‌ها جهت فشرده‌سازی تصویر به صورت بااتلاف می‌باشد. همچنین یک روش کدگذاری تصویر است که برای به‌دست آوردن خصوصیات آماری یک بلوک در تصویر فشرده‌شده در نظر گرفته شده است.

Ruchika و Gaurav [۵] روش‌های مختلف فیلترکردن تصویر را پیشنهاد کردند که این روش‌ها، مهم‌ترین و پرکاربردترین عملیات در بحث پردازش تصویر هستند و می‌توانند در بهبود کیفیت تصویر مؤثر باشند.

* سخنران. آدرس ایمیل: sa.kherandish@gmail.com

[۱] Singular value decomposition (SVD)

[۲] Block truncation coding (BTC)

۲ ابزار و روش‌ها

در این مقاله از طریق کاهش رتبه در تجزیه مقدار تکین و همچنین ترکیب آن با کدگذاری کوتاه کردن بلوک، تصویر فشرده می‌شود. سپس با استفاده از فیلتر پایین‌گذر کیفیت تصویر فشرده‌شده افزایش می‌یابد. باتوجه به این‌که باید کیفیت تصویر فشرده‌شده مشخص شود، لازم است معیارهایی برای سنجش میزان کیفیت و خطای تصویر فشرده‌شده مشخص گردد.

۱.۲ تجزیه مقدار تکین

ماتریس A با m سطر و n ستون و رتبه r را در نظر بگیرید به‌طوری‌که $r \leq n \leq m$ باشد. در این‌صورت می‌توان A را به سه ماتریس V^T، $\sum$ و U تجزیه کرد به گونه‌ای که $A = U \sum V^T$ است و U و V ماتریس‌های متعامد هستند و $\sum$ نیز یک ماتریس قطری است که عناصر قطری آن همان مقادیر تکین ماتریس A هستند.

تصویر فشرده‌شده I_{comp} با استفاده از فرمول (۱) محاسبه می‌شود که K تعداد مقادیر تکین است که با افزایش تعداد مقدار تکین کیفیت تصویر نیز بیشتر می‌شود.

$$I_{comp} = U(:, ۱ : K) \times \sum(۱ : K, ۱ : K) \times \left(V(:, ۱ : K)^T\right) \tag{۱}$$

۲.۲ کدگذاری کوتاه کردن بلوک

کدگذاری کوتاه کردن بلوک نوعی روش فشرده‌سازی تصویر است که به‌صورت بااتلاف می‌باشد؛ به‌طوری‌که تصویر اصلی را به بلوک‌هایی تقسیم می‌کند و سپس تعداد سطوح خاکستری در هر بلوک را کاهش می‌دهد درحالی که میانگین و انحراف معیار را حفظ می‌کند.

کدگذاری کوتاه کردن بلوک به این صورت است که ابتدا تصویر را به بلوک‌های $B \times B$ تقسیم می‌کنیم سپس میانگین و انحراف معیار هر بلوک را محاسبه کرده و به ترتیب در متغیرهای $\bar{x}$ و σ قرار می‌دهیم. پیکسل‌های بزرگ‌تر از میانگین را یک و بقیه را صفر قرار می‌دهیم. تعداد کل پیکسل‌های یک بلوک را در متغیر m و تعداد یک‌های بلوک را در متغیر q قرار می‌دهیم. سپس مقادیر a و b را توسط فرمول‌های (۲) و (۳) به‌دست می‌آوریم و درنهایت اعداد یک بلوک مورد نظر را با مقدار a و اعداد صفر را با مقدار b جایگزین می‌کنیم.

$$a = \bar{x} - \sigma\sqrt{\frac{q}{m-q}} \tag{۲}$$

$$b = \bar{x} + \sigma\sqrt{\frac{m-q}{q}} \tag{۳}$$

۳.۲ فیلتر پایین‌گذر

هنگام استفاده از فیلتر پایین‌گذر به‌روی تصویر، پیکسل‌هایی که فرکانس آن‌ها کم است حفظ می‌شوند و بر روی پیکسل‌هایی با فرکانس بالا تغییراتی ایجاد می‌شود. منظور از پیکسلی که فرکانس آن کم باشد، پیکسلی است که اختلاف شدت روشنایی آن با پیکسل‌های همسایه‌اش کم باشد و در مقابل پیکسل با فرکانس بالا پیکسلی است که اختلاف شدت روشنایی آن با پیکسل‌های همسایه‌اش زیاد باشد. درنتیجه استفاده از فیلتر پایین‌گذر، باعث کاهش نویز و افزایش کیفیت تصویر می‌شود. در این مقاله از فیلتر پایین‌گذر گاوسی جهت افزایش کیفیت در تصویر فشرده‌شده استفاده شده است.

۴.۲ پارامترهای مقایسه‌ای

برای محاسبه کیفیت و ارزیابی عملکرد الگوریتم از معیار $PSNR$ که در فرمول (۴) آورده شده؛ استفاده می‌شود. $PSNR$ زیاد نشان می‌دهد که تخریب بصری در تصویر فشرده‌شده کمتر است. در این فرمول، max بیشترین مقدار پیکسل موجود در تصویر است.

$$PSNR = ۱۰ \log_{۱۰} \frac{max^۲}{e_{MSE}} \tag{۴}$$

e_{MSE} برای دو تصویر I و J که در آن I تصویر اصلی و J تصویر فشرده‌شده است و m و n نیز تعداد سطرها و ستون‌های ماتریس است، توسط فرمول (۵) تعریف می‌شود.

$$e_{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - J(i,j)]^2 \quad (۵)$$

استفاده از RMSE بسیار رایج است و برای پیش‌بینی‌های عددی، خطای کلی بسیار خوبی ایجاد می‌کند و به‌صورت فرمول (۶) تعریف می‌شود. RMSE کمتر نشان‌دهنده مدل مناسب‌تر است.

$$RMSE = \sqrt{e_{MSE}} \quad (۶)$$

۳ نتایج اصلی

در این قسمت دو روش جهت فشرده‌سازی تصویر ارائه می‌شود. روش اول از طریق کاهش رتبه در تجزیه مقدار تکین، تصویر را فشرده می‌کند و سپس با استفاده از فیلتر پایین‌گذر گاوسی، کیفیت تصویر را افزایش می‌دهد. در روش دوم از طریق کاهش رتبه در تجزیه مقدار تکین و کدگذاری کوتاه کردن بلوک، تصویر فشرده می‌شود و برای افزایش کیفیت تصویر از فیلتر پایین‌گذر گاوسی استفاده می‌شود.

ورودی‌های الگوریتم ۱.۳ که مربوط به فشرده‌سازی تصویر از طریق تجزیه مقدار تکین و فیلتر پایین‌گذر است، شامل ماتریس A مربوط به تصویر و تعداد مقادیر تکین K و خروجی الگوریتم نیز تصویر فشرده‌شده و کیفیت مربوط به آن می‌باشد.

الگوریتم ۱.۳. فشرده‌سازی تصویر با استفاده از تجزیه مقدار تکین و فیلتر پایین‌گذر.

۱. تجزیه مقدار تکین A را به‌دست آورید و در ماتریس‌های U و S و V قرار دهید.

۲. $U_k = U(:, 1:K)$ و $S_k = S(1:K, 1:K)$ و $V_k = V(:, 1:K)^T$

۳. $D_k = U_k \times S_k \times V_k$

۴. اعمال فیلتر پایین‌گذر گاوسی با کرنل‌های 3×3 به‌روی ماتریس D_k.

۵. تصویر فشرده‌شده را نمایش دهید و $PSNR$، $RMSE$ و مربوط به تصویر فشرده‌شده را محاسبه کنید.

ورودی الگوریتم ۲.۳ که مربوط به فشرده‌سازی تصویر از طریق ترکیب دو روش تجزیه مقدار تکین و کدگذاری کوتاه کردن بلوک و استفاده از فیلتر پایین گذر جهت افزایش کیفیت تصویر است، شامل ماتریس A مربوط به تصویر، تعداد مقادیر تکین K و ابعاد بلوک‌ها یعنی $B \times B$ می‌باشد. خروجی الگوریتم نیز تصویر فشرده‌شده و کیفیت مربوط به آن می‌باشد.

الگوریتم ۲.۳. فشرده‌سازی تصویر با استفاده از تجزیه مقدار تکین، کدگذاری کوتاه کردن بلوک و فیلتر پایین‌گذر.

۱. تجزیه مقدار تکین A را به‌دست آورید و در ماتریس‌های U و S و V قرار دهید.

۲. $U_k = U(:, 1:K)$ و $S_k = S(1:K, 1:K)$ و $V_k = V(:, 1:K)^T$

۳. $D_k = U_k \times S_k \times V_k$

۴. ماتریس D_k را به بلوک‌های $B \times B$ تقسیم کنید.

۵. میانگین هر بلوک را در متغیر $\bar{x}$ و انحراف معیار را در متغیر σ قرار دهید.

۶. پیکسل‌های بزرگ‌تر از میانگین را یک و بقیه را صفر قرار دهید.

۷. تعداد کل پیکسل‌های یک بلوک را در متغیر m و تعداد یک‌های بلوک را در متغیر q قرار دهید.

۸. $a = \bar{x} - \sigma\sqrt{\frac{q}{m-q}}$ و $b = \bar{x} + \sigma\sqrt{\frac{m-q}{q}}$

۹. اعداد یک را با مقدار a و اعداد صفر را با مقدار b جایگزین کنید.

۱۰. اعمال فیلتر پایین‌گذر گاوسی با کرنل‌های ۳ × ۳ به‌روی ماتریس مربوط به تصویر فشرده‌شده.

۱۱. تصویر فشرده‌شده را نمایش دهید و $PSNR$، $RMSE$ و مربوط به تصویر فشرده‌شده را محاسبه کنید.

۴ نتایج عددی

برای ارزیابی الگوریتم‌ها، از دو تصویر خاکستری ۱ (آ) و رنگی ۲ (آ) استفاده کرده‌ایم، سپس نتایج حاصل از اجرای الگوریتم‌ها به‌روی هر تصویر را باتوجه به کیفیت تصویر نشان می‌دهیم.

شکل ۱ نتایج حاصل از الگوریتم‌های ۱.۳ و ۲.۳ را به‌روی تصویر خاکستری نشان می‌دهد. در شکل ۱ (آ)، تصویر اصلی و در شکل ۱ (ب)، تصویر فشرده شده با استفاده از تجزیه مقدار تکین و فیلتر پایین‌گذر با ۴۰۰ مقدار تکین نشان داده شده و در شکل ۱ (ج)، تصویر فشرده شده با استفاده از تجزیه مقدار تکین، کدگذاری کوتاه کردن بلوک و فیلتر پایین‌گذر با ۴۰۰ مقدار تکین و بلوک‌هایی با ابعاد ۲ × ۲ نشان داده شده است.

(ج) (ب) (آ)

شکل ۱: تصاویر خاکستری فشرده‌شده توسط دو الگوریتم.

شکل ۲ نتایج حاصل از الگوریتم‌ها به‌روی تصویر رنگی را نشان می‌دهد. در شکل ۲ (آ)، تصویر اصلی و در شکل ۲ (ب)، تصویر فشرده شده با استفاده از تجزیه مقدار تکین و فیلتر پایین‌گذر با ۴۰۰ مقدار تکین نشان داده شده و در شکل ۲ (ج)، تصویر فشرده شده با استفاده از تجزیه مقدار تکین، کدگذاری کوتاه کردن بلوک و فیلتر پایین‌گذر با ۴۰۰ مقدار تکین و بلوک‌هایی با ابعاد ۲ × ۲ نشان داده شده است.

(ج) (ب) (آ)

شکل ۲: تصاویر رنگی فشرده‌شده توسط دو الگوریتم.

در جدول ۱ کیفیت و مقدار خطای مربوط به تصویرهای ۱ (آ) و ۲ (آ) که با استفاده از کاهش رتبه در تجزیه مقدار تکین و فیلتر پایین‌گذر و همچنین بدون فیلتر پایین‌گذر فشرده شده‌اند، نشان داده شده است. تعداد مقادیر تکین برای تصویرهای خاکستری و رنگی ۱۰۰، ۲۰۰ و ۴۰۰ در نظر گرفته شده است. با افزایش مقادیر تکین کیفیت تصویر افزایش و میزان خطا کاهش می‌یابد.

در جدول ۲ کیفیت و مقدار خطای مربوط به تصویرهای ۱ (آ) و ۲ (آ) که با استفاده از کاهش رتبه در تجزیه مقدار تکین، کدگذاری کوتاه کردن بلوک و فیلتر پایین‌گذر و همچنین بدون فیلتر پایین‌گذر فشرده شده‌اند، نشان داده شده است. تعداد مقادیر تکین ۵۰ و ۲۰۰ و ابعاد بلوک‌ها ۲ × ۲، ۴ × ۴ و ۸ × ۸ در نظر گرفته شده است. هرچه ابعاد بلوک کوچکتر باشد کیفیت تصویر افزایش می‌یابد.

در جدول ۳ تأثیر فیلتر پایین‌گذر به‌روی سه روش فشرده‌سازی تصویر که شامل تجزیه مقدار تکین، ترکیب تجزیه مقدار تکین و کدگذاری کوتاه کردن بلوک و کدگذاری کوتاه کردن بلوک هستند، با ۴۰۰ مقدار تکین و بلوک‌هایی با ابعاد ۲ × ۲ بررسی شده است.

جدول ۱: تأثیر فیلتر پایین‌گذر به‌روی کیفیت و خطای مربوط به تصاویر فشرده‌شده با روش تجزیه مقدار تکین.

نوع تصویر	تعداد مقادیر تکین	PSNR بدون فیلتر	RMSE بدون فیلتر	PSNR با فیلتر	RMSE با فیلتر
خاکستری	۱۰۰	۳۸.۸۵۱۶	۲.۹۱۰۴	۳۹.۰۲۰۴	۲.۸۵۴۴
خاکستری	۲۰۰	۴۹.۹۸۱۸	۰.۸۰۸۱	۵۰.۲۴۰۷	۰.۷۸۴۳
خاکستری	۴۰۰	۵۵.۸۸۵۱	۰.۴۰۹۵	۵۶.۰۸۳۹	۰.۴۰۰۳
رنگی	۱۰۰	۳۲.۶۴۲۰	۵.۹۴۸۹	۳۲.۶۴۲۳	۵.۹۴۸۷
رنگی	۲۰۰	۳۸.۱۶۱۶	۳.۱۵۱۱	۳۸.۱۶۱۷	۳.۱۵۱۰
رنگی	۴۰۰	۴۸.۲۸۰۵	۰.۹۸۲۹	۴۸.۳۳۴۵	۰.۹۷۶۷

جدول ۲: تأثیر فیلتر پایین‌گذر به‌روی کیفیت و خطای مربوط‌به تصاویر فشرده‌شده باروش تجزیه‌مقدارتکین و کدگذاری‌کوتاه‌کردن بلوک.

نوع تصویر	تعداد مقادیر تکین	ابعاد بلوک	PSNR بدون فیلتر	RMSE بدون فیلتر	PSNR با فیلتر	RMSE با فیلتر
خاکستری	۵۰	۲×۲	۱۲.۲۲۹۵	۶۲.۳۸۲۸	۱۲.۲۵۰۸	۶۲.۲۳۳۵
خاکستری	۵۰	۴×۴	۱۲.۰۹۳۲	۶۳.۳۶۹۲	۱۲.۱۴۳۵	۶۳.۰۰۳۳
خاکستری	۵۰	۸×۸	۱۱.۹۱۸۵	۶۴.۶۵۷۰	۱۲.۰۰۴۱	۶۴.۰۲۳۰
خاکستری	۲۰۰	۲×۲	۲۵.۸۷۳۵	۱۲.۹۹۶۸	۲۹.۲۲۵۵	۸.۸۱۵۹
خاکستری	۲۰۰	۴×۴	۱۹.۰۹۶۰	۲۸.۳۹۴۰	۲۱.۰۲۷۸	۲۲.۶۵۴۲
خاکستری	۲۰۰	۸×۸	۱۳.۳۹۷۲	۵۳.۲۵۶۳	۱۴.۶۵۴۹	۴۷.۱۸۳۹
رنگی	۵۰	۲×۲	۲۶.۶۳۷۵	۱۱.۸۷۵۸	۲۷.۶۰۱۸	۱۰.۶۲۷۹
رنگی	۵۰	۴×۴	۲۰.۵۳۶۲	۲۳.۹۷۳۵	۲۲.۲۷۷۶	۱۹.۶۱۸۲
رنگی	۵۰	۸×۸	۱۳.۲۲۰۰	۵۵.۶۵۹۴	۱۴.۱۸۳۳	۴۹.۸۱۶۵
رنگی	۲۰۰	۲×۲	۲۷.۰۶۳۳	۱۱.۳۰۷۷	۲۹.۹۷۰۰	۸.۰۹۱۷
رنگی	۲۰۰	۴×۴	۱۸.۷۹۵۶	۲۹.۲۹۲۸	۲۰.۸۴۳۳	۲۳.۱۴۰۷
رنگی	۲۰۰	۸×۸	۱۱.۹۱۸۸	۶۴.۶۵۴۷	۱۲.۹۷۸۸	۵۷.۲۲۷۱

جدول ۳: مقایسه سه روش جهت فشرده‌سازی تصویرهای خاکستری و رنگی.

نوع تصویر	روش	PSNR بدون فیلتر	RMSE بدون فیلتر	PSNR با فیلتر	RMSE با فیلتر
خاکستری	SVD	۵۵.۸۸۵۱	۰.۴۰۹۵	۵۶.۰۸۳۹	۰.۴۰۰۳
خاکستری	SVD & BTC	۲۵.۸۲۵۹	۱۳.۰۳۹۱	۲۹.۲۰۱۲	۸.۸۴۰۶
خاکستری	BTC	۲۵.۸۱۹۱	۱۳.۰۴۹۲	۲۹.۱۷۹۹	۸.۸۶۲۳
رنگی	SVD	۴۸.۲۸۰۵	۰.۹۸۲۹	۴۸.۳۳۴۵	۰.۹۷۶۷
رنگی	SVD & BTC	۲۶.۶۶۳۱	۱۱.۸۴۱۰	۲۹.۸۹۱۷	۸.۱۶۴۶
رنگی	BTC	۲۶.۶۰۹۷	۱۱.۹۱۳۹	۲۹.۸۶۷۱	۸.۱۸۸۱

۵ نتیجه‌گیری

در این مقاله دو روش جهت فشرده‌سازی تصویر ارائه شده است و سپس تأثیر فیلتر پایین‌گذر به‌روی تصویرهای فشرده‌شده توسط این دو روش مورد بررسی قرار گرفته است. باتوجه به نتایجی که به دست آمده می‌توان گفت که کیفیت حاصل از تصویر فشرده‌شده با استفاده از روش تجزیه مقدار تکین و فیلتر پایین‌گذر و همچنین روش ترکیبی تجزیه مقدار تکین و کدگذاری کوتاه کردن بلوک و استفاده از فیلتر پایین‌گذر، کیفیت مناسبی است و خطای کمتری حاصل می‌شود.

مراجع

[1] M. Doaa and A. F. Chadi, Image compression using block truncation coding, Cyber Journals: Multidisciplinary Journals in Science and Technology, *Journal of Selected Areas in Telecommunications (JSAT)*, February Edition, 2011.

[2] K. El Asnaoui, Image compression based on block svd power method, *Journal of Intelligent Systems*, 29(1):1345–1359, 2019.

[3] K. El Asnaoui and Y. Chaw ki. Youness, Two new methods for image compression, *International Journal of Imaging and Robotics*, 15(4):1–11, 2015.

[4] K. El Asnaoui, M. Ouhda, B. Aksasse and M. Ouanan, An application of linear algebra to image compression, *In Seminar on Algebra and its Applications*, pages 41–54, Springer, 2016.

[5] C. Ruchika and G. Gaurav, Image filtering algorithms and techniques: A review, *International Journal of Advanced Research in Computer Science and Software Engineering*, 2013.

معکوس ماتریس‌های ساختار یافته مختلط

فاطمه رضائیان *

گروه علوم ریاضی، دانشگاه آزاد اسلامی واحد اهواز، اهواز

چکیده

ماتریس‌هایی با ساختار خاص و کاربردهای فراوان آن در علوم مختلف، ریاضیدانان و صاحب نظران سایر علوم را مایل به ارائه الگوریتم‌هایی نموده تا معکوس آن‌ها را از روش‌های متعدد محاسبه نمایند. در این مقاله الگوریتمی را ارائه می‌دهیم تا معکوس بعضی از ماتریس‌های ساختار یافته‌ی نامتقارن را محاسبه نماییم.

واژه‌های کلیدی: ماتریس‌های ساختار یافته، معکوس، ماتریس‌های نامتقارن

کد موضوع‌بندی ریاضی [۲۰۱۰]: 15A03، 15A23، 15B36

۱ مقدمه

در جبر خطی ماتریس‌های توپلیتز و هنکل ماتریس‌های ساختار یافته‌ای هستند که دارای نظم خاصی در درایه‌های روی قطر اصلی، بالا و پایین قطر اصلی می‌باشند.
نویسندگان متفاوتی بر روی معکوس این ماتریس‌ها کار کرده‌اند و پایداری آن‌ها را بررسی کرده‌اند [۲–۸] به طوری که در [۱] معکوس ماتریس‌های توپلیتز متقارن بدست آمده است. ما این روش را برای ماتریس‌های نامتقارن تعمیم می‌دهیم و همچنین ثابت می‌کنیم می‌توانیم با این روش معکوس ماتریس‌های هنکل متقارن و نامتقارن را بدست بیاوریم و در ادامه خوش وضع بودن روش را بیان می‌کنیم.

تعریف ۱.۱. ماتریس T یک ماتریس توپلیتز $n \times n$ است.

$$T = \begin{bmatrix} a_0 & a_{-1} & a_{-2} & \cdots & a_{1-n} \\ a_1 & a_0 & a_{-1} & \cdots & a_{2-n} \\ a_2 & a_1 & a_0 & \cdots & a_{3-n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_0 \end{bmatrix}. \quad (1)$$

جایی که $a_{n-1}, \cdots, a_{-(n-1)}$ اعداد مختلط هستند ما از فرم کوتاه شده زیر برای نمایش T استفاده می‌کنیم.

$$T = \begin{pmatrix} a \\ p-q \end{pmatrix}_{p,q=1}^{n}. \quad (2)$$

در [۱] برای ماتریس‌های توپلیتز متقارن داریم.

$$KT - TK = fe_n^T - e_1 f^T J. \quad (3)$$

* سخنران. آدرس ایمیل: Rezaeianf@ymail.com

جایی که

$$K=\begin{bmatrix}0 & & & 1\\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & 0\end{bmatrix},\quad J=\begin{bmatrix} & & 1\\ & \iddots & \\ 1 & & \end{bmatrix},\quad e_1=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},\quad e_n=\begin{bmatrix}0\\\vdots\\0\\1\end{bmatrix},\quad f=\begin{bmatrix}0\\a_{n-1}-a_{-1}\\\vdots\\a_2-a_{-n+2}\\a_1-a_{-n+1}\end{bmatrix}.$$

بوضوح مشخص است که معادله‌ی (۱) برای ماتریس‌های توپلیتز نامتقارن صحیح نیست می‌توان با یک مثال نقض عدم صحت آن را برای ماتریس‌های اشاره شده را بررسی نمود.

۲ نتایج اصلی

در این بخش الگوریتم اصلاح شده‌ی مقاله‌ی [۱] را برای محاسبه‌ی T^{-1} در ماتریس‌های توپلیتز متقارن و نامتقارن را بیان می‌کنیم و سپس این روش را برای محاسبه‌ی معکوس ماتریس‌های هنکل متقارن و نامتقارن بسط می‌دهیم.

گزاره ۱.۲. *فرض کنید T و K و J مانند قبل باشند و f نیز به صورت زیر تعریف شود.*

$$f=\begin{bmatrix}f_1 & f_2 & \cdots f_n\end{bmatrix}^T.$$

بنابراین خواهیم داشت:

$$KT=\begin{bmatrix}a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_1 & a_0\\ a_0 & a_{-1} & a_{-2} & \cdots & a_{2-n} & a_{1-n}\\ a_1 & a_0 & a_{-1} & \cdots & a_{3-n} & a_{2-n}\\ a_2 & a_1 & a_0 & \cdots & a_{4-n} & a_{3-n}\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ a_{n-2} & a_{n-3} & a_{n-4} & \cdots & a_0 & a_{-1}\end{bmatrix},\quad TK=\begin{bmatrix}a_{-1} & a_{-2} & a_{-3} & \cdots & a_{1-n} & a_0\\ a_0 & a_{-1} & a_{-2} & \cdots & a_{2-n} & a_1\\ a_1 & a_0 & a_{-1} & \cdots & a_{3-n} & a_2\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ a_{n-2} & a_{n-3} & a_{n-4} & \cdots & a_0 & a_{n-1}\end{bmatrix}$$

$$KT-TK=\begin{bmatrix}a_{n-1}-a_{-1} & a_{n-2}-a_{-2} & a_{n-3}-a_{-3} & \cdots & a_1-a_{1-n} & 0\\ 0 & 0 & 0 & \cdots & 0 & a_{1-n}-a_1\\ 0 & 0 & 0 & \cdots & 0 & a_{2-n}-a_2\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 0 & a_{-1}-a_{n-1}\end{bmatrix}.$$

از طرفی:

$$fe_n^T-e_1f^TJ=\begin{bmatrix}-f_n & -f_{n-1} & -f_{n-2} & \cdots & -f_2 & 0\\ 0 & 0 & 0 & \cdots & 0 & f_2\\ 0 & 0 & 0 & \cdots & 0 & f_3\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 0 & f_n\end{bmatrix}.$$

بوضوح $f=\begin{bmatrix}f_1\\a_{1-n}-a_1\\\vdots\\a_{-2}-a_{n-2}\\a_{-1}-a_{n-1}\end{bmatrix}$.

هم‌اکنون می‌توانیم الگوریتم اصلاح شده را بیان کنیم.

لم ۲.۲. *فرض کنید T ماتریس توپلیتز $n \times n$ مانند (۱) باشد آنگاه معادله‌ی (۳) برای بدست آوردن معکوس ماتریس توپلیتز متقارن و نامتقارن استفاده می‌شود هرگاه:*

$$f = \begin{bmatrix} ۰ \\ a_{۱-n} - a_{۱} \\ \vdots \\ a_{-۲} - a_{n-۲} \\ a_{-۱} - a_{n-۱} \end{bmatrix}$$

قضیه ۳.۲. *فرض کنید T یک ماتریس توپلیتز است که مشابه (۲) تعریف می‌شود، اگر هر یک از معادلات $Ty = e_{۱}$ و $Tx = f$ معکوس پذیر باشند به طوری که $x = (x_{۱}, x_{۲}, \cdots, x_{n})^T$ و $y = (y_{۱}, y_{۲}, \cdots, y_{n})^T$ آنگاه:*

الف- T معکوس پذیر است.

ب- $T^{-۱} = T_{۱}U_{۱} + T_{۲}U_{۲}$ [۱].

جایی که

$$T_{۱} = \begin{bmatrix} y_{۱} & y_{n} & \cdots & y_{۲} \\ y_{۲} & y_{۱} & \ddots & \\ \vdots & \ddots & \ddots & y_{n} \\ y_{n} & \cdots & y_{۲} & y_{۱} \end{bmatrix}, \quad U_{۱} = \begin{bmatrix} ۱ & -x_{n} & \cdots & -x_{۲} \\ & ۱ & \ddots & \vdots \\ & & \ddots & -x_{n} \\ & & & ۱ \end{bmatrix} \quad T_{۲} = \begin{bmatrix} x_{۱} & x_{n} & \cdots & x_{۲} \\ x_{۲} & x_{۱} & \ddots & \\ \vdots & \ddots & \ddots & x_{n} \\ x_{n} & \cdots & x_{۲} & x_{۱} \end{bmatrix}, \quad U_{۲} = \begin{bmatrix} ۰ & y_{n} & \cdots & y_{۲} \\ & ۰ & \ddots & \vdots \\ & & \ddots & y_{n} \\ & & & ۰ \end{bmatrix}.$$

لم ۴.۲. *فرض کنید H یک ماتریس هنکل $n \times n$ باشد.*

$$H = \begin{bmatrix} a_{۱-n} & a_{۲-n} & \cdots & a_{-۱} & a_{۰} \\ a_{۲-n} & a_{۳-n} & \cdots & a_{۰} & a_{۱} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{-۱} & a_{۰} & \cdots & a_{n-۳} & a_{n-۲} \\ a_{۰} & a_{۱} & \cdots & a_{n-۲} & a_{n-۱} \end{bmatrix}.$$

بنابراین بوضوح مشابه لم ۲.۲ ثابت می‌شود:

$$KH - HK^T = fe_{n}^T - e_{۱}f^T. \qquad (۴)$$

اثبات.

$$KT - TK = fe_{n}^T - e_{۱}f^T J \Rightarrow KTJ - TKJ = fe_{n}^T J - e_{۱}f^T JJ$$
$$\Rightarrow KTJ - TJK^T = fe_{۱}^T - e_{۱}f^T \Rightarrow KT - HK^T = fe_{n}^T - e_{۱}f^T.$$

□

قضیه ۵.۲. *فرض کنید H یک ماتریس هنکل باشد $x, y \in C^n$ و $Hx = f$ و $Hy = e_{۱}$ به طوری که*

$$x = (x_{۱}, x_{۲}, \cdots, x_{n})^T, \quad y = (y_{۱}, y_{۲}, \cdots, y_{n})^T$$

آنگاه

الف- H معکوس پذیر است.

ب- $H^T = H_{۱}U_{۱} + H_{۲}U_{۲}$ جایی که:

$$H_{۱} = \begin{bmatrix} y_{۱} & & y_{n-۱} & y_{n} \\ & \iddots & \iddots & y_{۱} \\ y_{n-۱} & y_{n} & \iddots & \\ y_{n} & y_{۱} & & y_{n-۱} \end{bmatrix}, \quad U_{۱} = \begin{bmatrix} ۱ & -x_{۱} & \cdots & -x_{n-۱} \\ & ۱ & \ddots & \vdots \\ & & \ddots & -x_{۱} \\ & & & ۱ \end{bmatrix}, \quad H_{۲} = \begin{bmatrix} x_{۱} & & x_{n-۱} & x_{n} \\ & \iddots & \iddots & x_{۱} \\ x_{n-۱} & x_{n} & \iddots & \\ x_{n} & x_{۱} & & x_{n-۱} \end{bmatrix}, \quad U_{۲} = \begin{bmatrix} ۰ & y_{۱} & \cdots & y_{n-۱} \\ & ۰ & \ddots & \vdots \\ & & \ddots & y_{۱} \\ & & & ۰ \end{bmatrix}.$$

اثبات. الف-

$$Hx = f \Leftrightarrow HJJx = f \Leftrightarrow TJx = f.$$

ب- طبق قضیه‌ی ۳.۲ ما داریم:

$$T_1 = \begin{bmatrix} s_1 & s_n & \cdots & s_2 \\ s_2 & s_1 & \ddots & \\ \vdots & \ddots & \ddots & s_n \\ s_n & \cdots & s_2 & s_1 \end{bmatrix} = \begin{bmatrix} y_n & y_1 & \cdots & y_{n-1} \\ y_{n-1} & y_n & \ddots & \\ \vdots & \ddots & \ddots & y_1 \\ y_1 & \cdots & y_{n-1} & y_n \end{bmatrix}, \quad U_1 = \begin{bmatrix} 1 & -r_n & \cdots & -r_2 \\ & 1 & \ddots & \vdots \\ & & \ddots & -r_n \\ & & & 1 \end{bmatrix} = \begin{bmatrix} 1 & -x_1 & \cdots & -x_{n-1} \\ & 1 & \ddots & \vdots \\ & & \ddots & -x_1 \\ & & & 1 \end{bmatrix},$$

$$T_2 = \begin{bmatrix} r_1 & r_n & \cdots & r_2 \\ r_2 & r_1 & \ddots & \\ \vdots & \ddots & \ddots & r_n \\ r_n & \cdots & r_2 & r_1 \end{bmatrix} = \begin{bmatrix} x_n & x_1 & \cdots & x_{n-1} \\ x_{n-1} & x_n & \ddots & \\ \vdots & \ddots & \ddots & x_1 \\ x_1 & \cdots & x_{n-1} & x_n \end{bmatrix}, \quad U_2 = \begin{bmatrix} 0 & s_n & \cdots & s_2 \\ & 0 & \ddots & \vdots \\ & & \ddots & s_n \\ & & & 0 \end{bmatrix} = \begin{bmatrix} 0 & y_1 & \cdots & y_{n-1} \\ & 0 & \ddots & \vdots \\ & & \ddots & y_1 \\ & & & 0 \end{bmatrix}.$$

$$T^{-1} = T_1U_1 + T_2U_2 \overset{(TJ=H \Rightarrow H^{-1}=JT^{-1})}{\Longrightarrow} JT^{-1} = JT_1U_1 + JT_2U_2$$
$$\Longrightarrow H^{-1} = JT_1U_1 + JT_2U_2 = H_1U_1 + H_2U_2.$$

$$JT_1 = H_1 = \begin{bmatrix} y_1 & & y_{n-1} & y_n \\ & \iddots & \iddots & y_1 \\ y_{n-1} & y_n & \iddots & \\ y_n & y_1 & & y_{n-1} \end{bmatrix}.$$

□

۳ تحلیل پایداری ماتریس‌های هنکل

در این بخش ما پایداری معکوس ماتریس‌های هنکل نامتقارن در بخش ۲ را بررسی می‌کنیم. ما می‌دانیم الگوریتمی پایدار نامیده می‌شود که برای همه‌ی راه حل‌های x به جواب $\tilde{x}$ نزدیک باشد و رابطه‌ی $\frac{\|x-\tilde{x}\|_2}{\|x\|}$ کوچک است.

قضیه ۱.۳. *فرض کنید $\tilde{\varepsilon}$ و $\tilde{y}$، $\tilde{x}$ به ترتیب خطای ماشین و خطای محاسباتی قضیه ۵.۲ باشند آنگاه:*

$$\|\tilde{x}\|_2 \le \|x\|_2(1+\tilde{\varepsilon}), \quad \|\tilde{y}\|_2 \le \|y\|_2(1+\tilde{\varepsilon}), \quad \|H_1\|_F = \sqrt{n}\|y\|_2, \quad \|H_2\|_F = \sqrt{n}\|x\|_2$$
$$\|U_1\|_F \le \sqrt{n}\sqrt{1+\|x\|_2^2}, \qquad \|U_2\|_F \le \sqrt{n}\|y\|_2.$$

خواهیم داشت:

$$\begin{aligned} \tilde{H}^{-1} &= fl(\tilde{H}\tilde{U}_1 + \tilde{H}_2\tilde{U}_2) \\ &= fl\big((H_1+\Delta H_1)(U_1+\Delta U_1) + (H_2+\Delta H_2)(U_2+\Delta U_2)\big) \\ &= H^{-1} + \Delta H_1U_1 + H_1\Delta U_1 + \Delta H_2U_2 + H_2\Delta U_2 + E + F. \end{aligned}$$

از راه حل‌های $\tilde{x}$ و $\tilde{y}$ در فرمول‌های معکوس قضیه‌ی گفته شده استفاده می‌کنیم E خطای ماتریس و F خطای تفریق است. $\Delta H_1, \Delta U_1, \Delta H_2$ و ΔU_2 خطاهای ماتریس هستند.

$$\|\Delta H_1\|_F = \tilde{\varepsilon}\|H_1\|_F \le \tilde{\varepsilon}\|H_1\|_F = \tilde{\varepsilon}\sqrt{n}\|y\|_2, \quad \|\Delta H_2\|_F = \tilde{\varepsilon}\|H_2\|_F \le \tilde{\varepsilon}\|H_2\|_F = \tilde{\varepsilon}\sqrt{n}\|x\|_2,$$
$$\|\Delta U_1\|_F = \tilde{\varepsilon}\|U_1\|_F \le \tilde{\varepsilon}\sqrt{1+\|x\|_2^2}, \qquad \|\Delta U_2\|_F = \tilde{\varepsilon}\|U_2\|_F \le \tilde{\varepsilon}\sqrt{n}\|y\|_2.$$

به طوری که:

$$\begin{aligned}\|E\|_۲ &\le \|E\|_F \\ &\le \varepsilon n\Big(\|H_۱\|_F\|U_۱\|_F + \|H_۲\|_F\|U_۲\|_F\Big) \\ &\le \varepsilon n^۲\|y\|_۲\Big(\sqrt{۱+\|x\|_۲^۲}+\|x\|_۲\Big) \\ &\le \varepsilon n\|y\|_۲(۱+۲\|x\|_۲), \qquad \|F\|_۲ \le \varepsilon\sqrt{n}\|H^{-۱}\|_۲.\end{aligned}$$

آنگاه:

$$\|\tilde{H}^{-۱}-H^{-۱}\|_۲ \le n(۲\tilde{\varepsilon}+n\varepsilon)\|y\|_۲(۱+۲\|x\|_۲)+\varepsilon\sqrt{n}\|H^{-۱}\|_۲.$$

با در نظر گرفتن $Hx=f$ و $Hy=e_۱$ داریم:

$$\|y\|_۲ \le \|H^{-۱}\|_۲, \qquad \|x\|_۲ \le \|H^{-۱}\|_۲\|f\|_۲.$$

بنابراین:

$$\frac{\|\tilde{H}^{-۱}-H^{-۱}\|_۲}{\|H^{-۱}\|_۲} \le n(۲\tilde{\varepsilon}+n\varepsilon)(۱+۲\|H^{-۱}\|_۲)\|f\|_۲+\varepsilon\sqrt{n}.$$

ما نتیجه می‌گیریم که H خوش وضع است و جایی که $\|H^{-۱}\|_۲$ متناهی آنگاه $\|f\|_۲$ متناهی. بنابراین معکوس ماتریس‌های متقارن، نامتقارن توپلیتز و هنکل پایدار پیشرو است.

مراجع

[1] L.V. Xiao-Guang and T.Z. Huang, A note on inversion of Toeplitz matrices, *Applied Mathematics Letters*, 20 (2007), 1189–1193.

[2] K. Ye and L.H. Lim, Every Matrix is a Product of Toeplitz Matrices, *Found Comput Math.*, (2016), 16, 577–598.

[3] W.F. Trench, An algorithm for the inversion of finite Toeplitz matrix, *J. SIAM.*, 13 (3) (1964), 515–522.

[4] I. Gohberg and A. Semencul, On the inversion of finite Toeplitz matrices and their continuous analogues, *Mat. Issled.*, 7 (1972), 201–223 (in Russian).

[5] I. Gohberg and N. Krupnik, A formula for the inversion of finite Toeplitz matrices, *Mat. Issled.*, 7 (12) (1972), 272–283 (in Russian).

[6] G. Heinig and K. Rost, Algebraic methods for Toeplitz-like matrices and operators, in: Operator Theory: Advances and Applications, 13, *Birkhuser, Basel.*, (1984).

[7] M.K. Ng, K. Rost and Y.-W. Wen, On inversion of Toeplitz matrices, *Linear Algebra Appl.*, 348 (2002), 145-151.

[8] A. Ben-Artzi and T. Shalom, On inversion of Toeplitz and close to Toepaz matrices, *Linear Algebra Appl.*, 75 (1986), 173–192.

تحلیل سلسله مراتبی در انتخاب سرویس ابری

امیر حسین شاه‌بخش رضوی[۱]* و مرتضی جعفرزاده[۲]

[۱] دانشجوی کارشناسی ارشد علوم کامپیوتر، دانشگاه حکیم سبزواری، سبزوار، ایران
[۲] گروه ریاضی کاربردی، دانشکده‌ی ریاضی و علوم کامپیوتر، دانشگاه حکیم سبزواری، سبزوار، ایران

چکیده

تحلیل سلسله مراتبی یکی از روش‌های چند معیاره است که به منظور تصمیم‌گیری و انتخاب بین یک یا چند گزینه از میان گزینه‌های متعدد تصمیم با توجه به شاخص‌هایی که تصمیم‌گیرنده تعیین می‌کند در مسائل مختلف از جمله انتخاب سرویس‌های ابری مورد استفاده قرار می‌گیرد. در این مقاله بر اساس ویژگی‌های سرویس‌های ابری مقایسه‌هایی بین گزینه‌های مختلف براساس هر شاخص صورت می‌پذیرد و با استفاده از شاخص تصمیم و انجام مقایسه بین گزینه‌های تصمیم بهترین سرویس ابری انتخاب می‌گردد. در این پژوهش برای مقایسه سرویس‌های ابری مختلف از شاخص اندازه‌گیری خدمات [۱] که بر اساس ویژگی‌های مشترک سرویس‌های ابری تعیین می‌شود، استفاده شده است.

واژه‌های کلیدی: هوش مصنوعی، اینترنت اشیاء، مساله بهینه‌سازی چند معیاره
کد موضوع‌بندی ریاضی [۲۰۱۰]: 97Rxx، 90Bxx، 97R40

۱ مقدمه

مفهوم اینترنت اشیاء[۲] که در دنیای ارتباطات و فناوری اطلاعات مطرح است، مضامین گسترده‌ای را در بر می‌گیرد ولی با تعریف اتحادیه جهانی مخابرات[۳] اینترنت اشیاء، بخشی از اینترنت یکپارچه‌شده آینده به حساب می‌آید که ضمن برقراری ارتباط بین انسان‌ها و اشیاء فیزیکی و مجازی مختلف، ارتباط میان این اشیاء را از طریق شبکه‌های اینترنتی فراهم می‌آورد [۲]. شبکه ابری یک شبکه جهانی از مراکز داده‌هایی است که در سرتاسر کره زمین پخش شده‌اند. این شبکه‌ها سرویس‌هایی چون پردازش و ذخیره‌سازی را در اختیار کاربران قرار می‌دهند تا کاربران بتوانند از راه دور و از طریق اینترنت به آنها دسترسی داشته باشند. ابرها کاربرد و اهمیت خود را در حوزه‌های تجاری و علمی ثابت کرده‌اند؛ ارائه دهندگانی مانند (مایکروسافت، اورکل، Salesforce و...) ابعاد جدیدی از امکانات و فناوری را برای برنامه‌های علمی ایجاد کرده‌اند که پردازش مجموعه وسیعی از داده‌ها و محاسبات به صورت موازی و با کارایی بالا را در اختیار کاربر قرار می‌دهند. چند نمونه از شرکت‌های بزرگ ارائه دهنده سرویس‌های ابری به شرح ذیل هستند:

۱. گوگل: IoT Core یک سرویس کاملاً مدیریت شده است که این امکان را به کاربران می‌دهد تا به راحتی و در امنیت کامل، داده‌های میلیون‌ها دستگاه پراکنده در سراسر جهان را متصل، مدیریت و دریافت کنند.
۲. آمازون: خدمات ذخیره‌سازی خود را با نام خدمات ذخیره‌سازی ساده آمازون[۴] ارائه می‌کند.
۳. مایکروسافت یک پلتفرم ابری است که مجموعه‌ی گسترده‌ای از ابزارهای قدرتمند یکپارچه‌سازی خدمات و محصولات را در اختیار کسب و کارها و توسعه‌دهندگان قرار می‌دهد. این ابزارها به فضای ابری وابسته هستند و می‌توانند به بهبود امنیت و توان پردازشی کامپیوترها کمک کنند.

* سخنران. آدرس ایمیل: amirhosein.shahbakhsh@hsu.ac.ir

[۱] Service Measurement Index
[۲] Internet Of Things (IOT)
[۳] International Telecommunication Union (ITU)
[۴] Amazon Simple Storage Service

۲ تحلیل سلسله مراتبی[۱]

در بسیاری از موارد، نتیجه تصمیم‌گیری‌ها وقتی مطلوب و مورد رضایت تصمیم گیرنده است که تصمیم‌گیری بر اساس چندین معیار، بررسی و مورد تجزیه و تحلیل قرار گرفته شده باشد. دربعضی از مسائل، معیارها ممکن است با یکدیگر متضاد باشند، یعنی افزایش یک شاخص یا معیار موجب کاهش عمل و معیار دیگر می‌شود. مسائل تصمیم‌گیری با معیارهای چندگانه معمولا به دنبال گزینه‌ای هستند که بیشترین مزیت را برای تمامی معیارها داشته باشد. یک چارچوب کامل و منطقی برای ساختار یک مساله تصمیم‌گیری، نمایش و کمّی کردن عناصر آن، ارتباط آن عناصر با اهداف کلی و ارزیابی راه‌حل‌های جایگزین را فراهم می‌کند. روش استفاده از AHP را می‌توان به صورت زیر خلاصه کرد [۴، ۵]:

- مسئله به صورت سلسله مراتبی شامل هدف تصمیم گیری، گزینه‌های دستیابی به آن و شاخص‌هایی برای ارزیابی گزینه‌ها مدل می‌شود.
- با انجام یک سری مقایسه‌ها بر اساس مقایسه زوجی عناصر، اولویت‌ها در میان عناصر سلسله مراتبی تعیین می‌شود.
- مقایسه‌ها ترکیب می‌شوند تا مجموعه‌ای از اولویت‌های کلی برای سلسله مراتب به‌دست آورده شود.

هدف روش AHP تصمیم‌گیری و انتخاب یک گزینه از میان گزینه‌های متعدد تصمیم با توجه به شاخص‌هایی که تصمیم‌گیرنده تعیین می‌کند، است. به کارگیری این روش مستلزم پنج گام زیر است [۱].

۱. **مدل‌سازی:** در این قدم، مسأله و هدف تصمیم‌گیری به صورت سلسله مراتبی از عناصر تصمیم که با هم در ارتباط هستند، تعریف می‌شود. عناصر تصمیم شامل «شاخص‌های تصمیم‌گیری» و «گزینه‌های تصمیم» است. فرایند تحلیل سلسله مراتبی نیازمند شکستن یک مساله با چندین شاخص به سلسله مراتبی از سطوح است. سطح بالا بیانگر هدف اصلی فرایند تصمیم‌گیری است. سطح دوم، نشان‌دهنده شاخص‌های عمده و اساسی که ممکن است به شاخص‌های فرعی و جزئی‌تر در سطح بعدی شکسته شود، است. سطح آخر گزینه‌های تصمیم را ارائه می‌کند.

۲. **قضاوت ترجیحی:** قضاوت ترجیحی انجام مقایسه‌هایی بین گزینه‌های مختلف تصمیم بر اساس هر شاخص و قضاوت در مورد اهمیت شاخص تصمیم با انجام مقایسه‌های زوجی است. بعد از طراحی سلسله مراتبی مساله تصمیم، تصمیم‌گیرنده بایستی مجموعه ماتریس‌هایی که به طور عددی اهمیت یا ارجحیت نسبی شاخص‌ها را نسبت به یکدیگر و هر گزینه تصمیم را با توجه به شاخص‌ها نسبت به سایر گزینه‌ها اندازه‌گیری می‌نماید، ایجاد کند. این کار با انجام مقایسه‌های دو به دو بین عناصر تصمیم (مقایسه زوجی) و از طریق تخصیص امتیازات عددی که نشان‌دهنده ارجحیت یا اهمیت بین دو عنصر تصمیم است، صورت می‌گیرد. برای انجام این کار معمولا از مقایسه گزینه‌ها با شاخص‌های iام نسبت به گزینه‌ها یا شاخص‌های jام استفاده می‌شود.

۳. **محاسبه وزن‌های نسبی:** هدف تعیین وزن «عناصر تصمیم» نسبت به هم از طریق مجموعه‌ای از محاسبه‌های عددی است که اولویت و وزن هر یک از عناصر تصمیم با استفاده از اطلاعات ماتریس مقایسه‌های زوجی تعیین می‌گردد.

۴. **ادغام وزن‌های نسبی:** به منظور رتبه‌بندی گزینه‌های تصمیم، در این مرحله بایستی وزن نسبی هر عنصر را در وزن عناصر بالاتر ضرب کرد تا وزن نهایی آن به‌دست آید. با انجام این مرحله برای هر گزینه، مقدار وزن نهایی به‌دست می‌آید.

۵. **سازگاری در قضاوت‌ها:** به‌طور معمول تمامی محاسبه‌های مربوط به AHP بر اساس قضاوت اولیه تصمیم‌گیرنده که در قالب ماتریس مقایسه‌های زوجی ظاهر می‌شود، صورت می‌پذیرد و هر گونه خطا و ناسازگاری در مقایسه و تعیین اهمیت بین گزینه‌ها و شاخص‌ها، نتیجه نهایی به دست آمده از محاسبات را مخدوش می‌سازد. نرخ ناسازگاری[۲] ابزاری است که سازگاری را مشخص ساخته و نشان می‌دهد که تا چه حد می‌توان به اولویت‌های حاصل از مقایسه‌ها اعتماد کرد. شاید مقایسه دو گزینه امری ساده باشد، اما وقتی که تعداد مقایسه‌ها افزایش یابد اطمینان از سازگاری مقایسه‌ها به راحتی میسر نبوده و باید با به کارگیری نرخ سازگاری به این اعتماد دست یافت. اگر نرخ ناسازگاری کمتر از ۰,۱۰ باشد، سازگاری مقایسه‌ها قابل قبول بوده و در غیر این‌صورت مقایسه‌ها باید تجدید نظر شود. نرخ ناسازگاری بر اساس گام‌های زیر محاسبه می‌گردد که در آن ماتریس مقایسه‌های زوجی یک ماتریس $n \times n$، مثبت و معکوس[۳] است:

---

[۱] Analytical Hierarchy Process (AHP)

[۲] Inconsistency Ratio

[۳] درایه‌های آن نسبت به قطر اصلی معکوس هستند

- گام ۱ (محاسبه بردار مجموع وزنی[۱]): ماتریس مقایسه‌های زوجی را در بردار ستونی «وزن نسبی» ضرب می‌کنیم.
- گام ۲ (محاسبه بردار سازگاری[۲]): عناصر بردار مجموع وزنی را بر بردار وزن نسبی تقسیم می‌کنیم.
- گام ۳ (به‌دست آوردن λ_{max}[۳]): با محاسبه میانگین عناصر بردار سازگاری، λ_{max} به‌دست می‌آید.
- گام ۴ (محاسبه شاخص سازگاری[۴]): شاخص سازگاری بر اساس رابطه زیر محاسبه می‌گردد:

$$CI = \frac{\lambda_{max} - n}{n - ۱}$$

- گام ۵ (محاسبه نرخ سازگاری): نرخ سازگاری از تقسیم شاخص سازگاری بر شاخص تصادفی[۵] به‌دست می‌آید:

$$CR = \frac{CI}{RI}$$

۳ سیستم شاخص ارزیابی خدمات ابری

بر اساس ویژگی‌های سرویس‌های ابری، برخی از شاخص‌ها مانند امنیت، کیفیت خدمات، هزینه و اعتبار در انتخاب شرکت‌های ارائه دهنده سرویس‌های ابری مانند (مایکروسافت، اورکل، Salesforce و ...) مؤثر هستند [۴]. از این‌رو چهار شاخص زیر را در روش AHP در نظر می‌گیریم:

- **امنیت:** زمانی است که کاربران تصمیم می‌گیرند داده‌ها را در فضای ابری ذخیره کنند، به این معنی که کنترل امنیت داده‌ها را از دست داده‌اند و به سرویس امنیتی ارائه دهنده ابری، اطمینان می‌کنند. بنابراین، امنیت به عنوان یک نگرانی اصلی برای سرویس ابری در نظر گرفته می‌شود.
- **کیفیت خدمات:** به طور معمول برای توصیف سرویس‌های دریافت شده از اینترنت به‌کار می‌رود.
- **هزینه:** این شاخص، بیانگر هزینه‌هایی است که هنگام استفاده مشتری از خدمات ارائه شده توسط ارائه دهنده خدمات ابری پرداخت می‌شود.
- **اعتبار:** خدماتی است که سرویس‌های ابری در اختیار مشتریان خود قرار می‌دهند و باعث اعتبار بخشی به کار خود می‌شوند.

۴ مثال

یکی از کاربردهای سرویس ابری، مدیریت ارتباط با مشتری به صورت آنلاین[۶] است. دراین مثال هدف این است که از بین شرکت‌های ارائه دهنده خدمات ابری Oracle Salesforce، و Microsoft بهترین سرور ابری برای سیستم ارتباط با مشتری با روش سلسله مراتبی انتخاب شود.

۱.۴ ساختار سلسله مراتبی

با توجه به سیستم شاخص ارزیابی سرویس ابری در این مثال، می‌توان سلسله مراتب AHP را بر اساس شکل ۱ در نظر گرفت [۴].

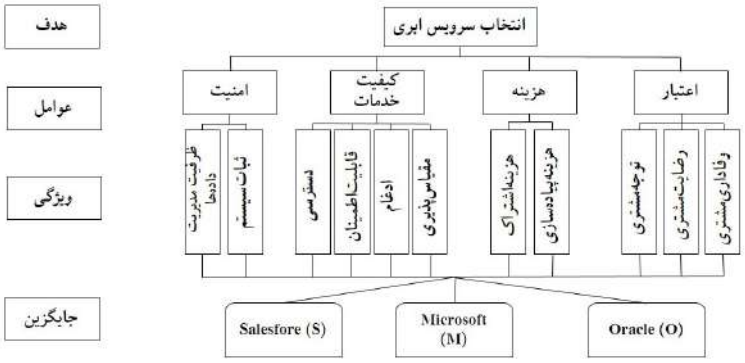
۲.۴ وزن‌دهی به ویژگی‌ها

در جدول ۱ ماتریس مقایسه زوجی با توجه به هدف مساله آمده است. درایه‌های قطر اصلی در ماتریس مقایسه زوجی همیشه یک است، یعنی هر گزینه نسبت به خودش برتری ندارد. تمامی اعداد ماتریس مقایسه زوجی به صورت متقارن و معکوس لحاظ می‌شوند. برای محاسبه وزن کافیست که مجموع هر ستون ماتریس را محاسبه کرده و سپس هر درایه را تقسیم بر مجموع ستونش کرد. در

[۱] Weighted sum vector
[۲] Consistency vector
[۳] بزرگترین مقدار ویژه ماتریس مقایسه‌های زوجی
[۴] Consistency Index (CI)
[۵] Random Index (RI)
[۶] Customer Relationship Management (CRM)

انتها از اعداد به‌دست آمده می‌توان میانگین سطری گرفت تا وزن‌های هر شاخص نسبت به هدف (W_i) به‌دست آید. به طور مشابه، می‌توان وزن هر ویژگی نسبت به هدف (W_j) را بر اساس رابطه (۱) به‌دست آورد. در جدول ۲ وزن‌های اصلی و فرعی آمده است.

$$W_j = W_i * W_{ij} \qquad (۱)$$



شکل ۱: سلسله مراتب AHP برای تصمیم خرید سرویس CRM آنلاین

جدول ۱: مقایسه زوجی شاخص‌ها با توجه به هدف

هدف	امنیت	کیفیت خدمات	هزینه	اعتبار	وزن
امنیت	۱	۲	۴	۳	۰/۴۷۲
کیفیت خدمات	$\frac{1}{2}$	۱	۲	۲	۰/۲۵۱
هزینه	$\frac{1}{4}$	$\frac{1}{2}$	۱	۲	۰/۱۵۷
اعتبار	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	۱	۰/۱۲۰

جدول ۲: وزن‌های اصلی و فرعی

شاخص‌ها	وزن اصلی (W_i)	ویژگی	وزن فرعی (W_{ij})	وزن W_j
امنیت	۰/۴۷۲	ظرفیت مدیریت داده	۰/۶۶۷	۰/۳۱۴۸
		پایداری سیستم	۰/۳۳۳	۰/۱۵۷۲
کیفیت خدمات	۰/۲۵۱	دسترسی	۰/۴۰۰	۰/۱۰۰۴
		قابلیت اطمینان	۰/۳۳۷	۰/۰۸۴۶
		ادغام	۰/۱۶۵	۰/۰۴۱۴
		مقیاس‌پذیری	۰/۰۹۸	۰/۰۲۴۶
هزینه	۰/۱۵۷	هزینه اشتراک	۰/۷۵	۰/۱۱۷
		هزینه پیاده‌سازی	۰/۲۵	۰/۰۳۹۲
اعتبار	۰/۱۲۰	قابلیت اطمینان	۰/۱۴	۰/۰۱۶۸
		ادغام	۰/۳۳	۰/۰۳۹۶
		مقیاس‌پذیری	۰/۵۳	۰/۰۶۳۶

۳.۴ مقیاس‌بندی داده‌ها

از آن‌جایی که برخی از ویژگی‌ها مثبت و برخی دیگر منفی هستند، بنابراین باید با اعمال دو مقیاس، همه داده‌ها را مقیاس‌بندی نماییم که مجموعه هدف Y به‌دست می‌آید. با نرمال کردن مجموعه هدف Y، امتیاز ویژگی هدف بدست می‌آید که آن را با Y_j نشان می‌دهیم. به دلیل عدم دسترسی به داده‌ها، امتیاز خام به صورت تصادفی تولید شده است [۴،۵]. در جدول ۳ امتیاز بندی ویژگی‌ها آمده است.

جدول ۳: امتیاز بندی ویژگی‌های سه شرکت Salesforce (S)، Oracle (O) و Microsoft (M)

	امتیاز خام شرکت‌ها			امتیاز هدف شرکت‌ها(Y_j)		
ویژگی	S	M	O	S	M	O
مدیریت ظرفیت داده‌ها	۲٫۵۱	۳٫۲۸	۲٫۷۵	۰	۱	۰٫۳۱۱۶
پایداری سیستم	۴٫۰۵	۴٫۱۸	۴٫۳۶	۰	۰٫۴۱۹۳	۱
دسترسی	۴٫۸۵	۴٫۷۳	۴٫۵۴	۱	۰٫۶۱۲	۰
قابلیت اطمینان	۳٫۰۷	۳٫۰۴	۲٫۹۵	۱	۰٫۷۵	۰
مقیاس‌پذیری	۴٫۵۴	۴٫۴۴	۴٫۴۹	۱	۰	۰٫۳۸۴۶
ادغام	۴٫۷۲	۴٫۷۷	۴٫۶۶	۰٫۵۴۵۴	۱	۰
هزینه اشتراک	۶۵	۴۹	۶۸	۰٫۱۵۷۸	۱	۰
هزینه مفهومی	۵۳	۴۷	۵۹	۰٫۵	۱	۰
توجه مشتری	۰٫۱۱	۰٫۲۲	۰٫۱۱	۰	۱	۰
رضایت مشتری	۰٫۲۵	۰٫۲۴	۰٫۲۱	۱	۰٫۷۵	۰
وفاداری مشتری	۰٫۸۴	۰٫۸۳	۰٫۸۰	۱	۰٫۷۵	۰

۴.۴ رتبه‌بندی

با استفاده از رابطه (۲) می‌توان امتیاز کلی هر شرکت ارائه دهنده خدمات ابری را محاسبه کرد که در آن n تعداد ویژگی‌ها و W_j وزن هر ویژگی با توجه به هدف است. رتبه‌بندی هر ویژگی برای سه شرکت ارائه دهنده خدمات ابری در جدول ۴ آمده است که ردیف آخر در آن امتیاز کلی ۳ شرکت را نشان می‌دهد. مشاهده می‌شود که Microsoft بالاترین امتیاز را نسب به دو شرکت دیگر دارد و به عنوان بهترین گزینه برای سیستم ارتباط با مشتری با استفاده از روش AHP پیشنهاد می‌شود.

$$Score_s = \sum_{j=1}^{n} (W_i * Y_j) \qquad (۲)$$

جدول ۴: رتبه‌بندی هر ویژگی برای سه شرکت Salesforce (S)، Oracle (O) و Microsoft (M)

ویژگی	S	M	O
مدیریت ظرفیت داده‌ها	۰	۰٫۳۱۴۸	۰٫۰۹۸۰
پایداری سیستم	۰	۰٫۰۶۵۹	۰٫۱۵۷۲
دسترسی	۰٫۱۰۰۴	۰٫۰۶۱۴۴	۰
قابلیت اطمینان	۰٫۰۸۴۶	۰٫۶۳۴۵	۰
مقیاس‌پذیری	۰٫۰۴۱۴	۰	۰٫۰۱۵۹
ادغام	۰٫۰۱۳۴	۰٫۰۲۴۶	۰
هزینه اشتراک	۰٫۰۱۹۵	۰٫۱۲۳۸	۰
هزینه پیامد	۰٫۰۲۰۶	۰٫۰۴۱۳	۰
توجه مشتری	۰	۰٫۰۱۳۷	۰
رضایت مشتری	۰٫۰۳۲۳	۰٫۲۴۲۲	۰
وفاداری مشتری	۰٫۰۵۱۹	۰٫۰۳۸۹	۰
امتیاز کلی	۰٫۳۶۴۱	۱٫۵	۰٫۲۷۱۱

۵ نتیجه‌گیری

با توجه به معیارهای مختلف سرویس‌های ارائه کننده ابری و استفاده روزافزون کاربران از این خدمات نیاز است، کاربران قبل از انتخاب سرویس ابری و با توجه به اهدافی که دارند چندین سرویس ابری را با روش تحلیل سلسله مراتبی مورد مقایسه قرار بدهند. با توجه به ویژگی‌ها و شاخص‌های مورد نظر کاربران، روش سلسله مراتبی بهترین گزینه انتخاب را در اختیار کاربران استفاده کننده از سرویس‌های ابری قرار می‌دهد.

مراجع

[۱] محمدرضا مهرگان، *پژوهش عملیاتی پیشرفته* ، چاپ اول، انتشارات کتاب دانشگاهی، ۱۳۸۳.

[2] Amir Masoud, Shabnam and Rahmani,Shadroo, *Systematic survey of big data and data mining in internet of things*, 139 (2018), 19–47.

[3] Ashkan Yousefpour and Caleb Fung and Tam Nguyen and Krishna Kadiyala and Fatemeh Jalali and Amirreza Niakanlahiji and Jian Kong and Jason P. Jue, *All one needs to know about fog computing and related edge computing paradigms: A complete survey*, 98 (2019), 289-330.

[4] Donglin, Fadika, Guihua and She, Qiping and Chen, *Evaluation index system of cloud service and the purchase decision-making process based on AHP* , (2011), 345–352.

[5] Donglin, Fadika, Guihua and She, Qiping and Chen, *A framework for ranking of cloud computing services* , 29 (2013), No.4, 1012–1023.

یازدهمین سمینار بین‌المللی
جبر خطی و کاربردهای آن
بهمن ۱۴۰۰

معکوس ماتریس‌های نواری

مریم شمس سولاری[۱]، * و مهران رسولی[۲]
دانشکده علوم، گروه ریاضی، دانشگاه پیام‌نور، تهران، ایران

چکیده

در این مقاله، الگوریتمی برای یافتن معکوس ماتریس‌های نواری ارائه می‌دهیم. ابتدا با تجزیه کردن ماتریس با روش دولیتل LU روابط و فرمول‌هایی را برای محاسبه معکوس ماتریس به‌دست آورده و در آخر الگوریتم را با استفاده از این روابط ارائه می‌کنیم. در پایان، هزینه محاسبات و مثال عددی برای اثبات کارایی الگوریتم در مقایسه با معکوس معمولی در نرم افزار متلب بیان می کنیم.

واژه‌های کلیدی: ماتریس معکوس، الگوریتم، تجزیهLU، ماتریس نواری
کد موضوع‌بندی ریاضی [۲۰۱۰]: 15A03، 15A23، 15B36

۱ مقدمه

شکل کلی یک ماتریس rـ قطری $n \times n$ که در آن $(r = ۲m + ۱ : m = ۱, ۲, \dots)$ و m تعداد قطرهای ناصفر بالا یا پایین قطر اصلی را نشان می‌دهد، به‌صورت زیر است:

$$G_{m,n} = \begin{bmatrix} d_۱ & a_{۱,۱} & a_{۲,۱} & \cdots & a_{m,۱} & ۰ & \cdots & ۰ \\ b_{۱,۱} & d_۲ & a_{۱,۲} & a_{۲,۲} & \cdots & a_{m,۲} & \ddots & \vdots \\ b_{۲,۱} & b_{۱,۲} & \ddots & \ddots & \ddots & \ddots & \ddots & ۰ \\ \vdots & b_{۲,۲} & \ddots & \ddots & \ddots & \ddots & \ddots & a_{m,n-m} \\ b_{m,۱} & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ ۰ & b_{m,۲} & \ddots & \ddots & \ddots & \ddots & \ddots & a_{۲,n-۲} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & a_{۱,n-۱} \\ ۰ & \cdots & ۰ & b_{m,n-m} & \cdots & b_{۲,n-۲} & b_{۱,n-۱} & d_n \end{bmatrix}. \quad (۱)$$

* سخنران. آدرس ایمیل: shamssolary@email.com

شکل کلی یک ماتریس (r,k)- قطری $n\times n$ که در آن $(r=۲m+۱ : m=۱,۲,\dots)$ که m تعداد قطرهای ناصفر بالا یا پایین قطر اصلی و k فاصله بین این قطرهای ناصفر را نشان می‌دهد، به‌صورت زیر است:

$$G^{(k)}_{m,n}=\begin{cases} G_{i,i}=d_i, & i=۱,۲,\dots,n,\\ G_{j,k+j}=a_{۱,j},\quad G_{k+j,j}=b_{۱,j}, & j=۱,\dots,n-k,\\ G_{j,۲k+j}=a_{۲,j},\quad G_{۲k+j,j}=b_{۲,j}, & j=۱,\dots,n-۲k,\\ \qquad\vdots & \\ G_{j,mk+j}=a_{m,j},\quad G_{mk+j,j}=b_{m,j}, & j=۱,\dots,n-mk, \end{cases} \tag{۲}$$

همچنین $a_{j,i}$ و $b_{j,i}$ دنباله‌های متناهی از اعداد ناصفر به‌ازای $(j=۱,۲,\dots,m)$ و $(i=۱,۲,\dots,n)$ هستند.

۲ نتایج اصلی

در این بخش الگوریتم اصلی را با استفاده از لم زیر و مراجع [۱، ۳، ۵] برای تعیین معکوس هر ماتریس (r,k)- قطری ارائه می‌کنیم.

لم ۱.۲. *([۴]) فرض کنید D_k، برای $k=۱,۲,\dots,n$ دترمینان زیرماتریس‌های حاصل از حذف $n-k$ سطر و ستون ماتریس $G^{(k)}_{m,n}$ باشند، اگر D_k-ها مخالف صفر باشند، آنگاه ماتریس $G^{(k)}_{m,n}$ تنها یک تجزیه دولیتل LU به‌صورت،*

$$G^{(k)}_{m,n}=L^{(k)}_{m,n}U^{(k)}_{m,n} \tag{۳}$$

دارد که در آن ماتریس‌های $L^{(k)}_{m,n}$ و $U^{(k)}_{m,n}$ به‌شکل،

$$L^{(k)}_{m,n}=\begin{cases} l_{i,i}=۱, & i=۱,۲,\dots,n,\\ l_{k+j,j}=l_{۱,j}, & j=۱,\dots,n-k,\\ l_{۲k+j,j}=l_{۲,j}, & j=۱,\dots,n-۲k,\\ \qquad\vdots & \\ l_{mk+j,j}=l_{m,j}, & j=۱,\dots,n-mk, \end{cases} \tag{۴}$$

و

$$U^{(k)}_{m,n}=\begin{cases} u_{i,i}=t_i, & i=۱,۲,\dots,n,\\ u_{j,k+j}=u_{۱,j}, & j=۱,\dots,n-k,\\ u_{j,۲k+j}=u_{۲,j}, & j=۱,\dots,n-۲k,\\ \qquad\vdots & \\ u_{j,mk+j}=u_{m,j} & j=۱,\dots,n-mk, \end{cases} \tag{۵}$$

هستند و $det(G^{(k)}_{m,n})=\prod_{i=۱}^{n}t_i$. حال با جایگذاری (۲)، (۴) و (۵) در (۳) روابط زیر را به‌دست می‌آوریم،

$$t_i=\begin{cases} d_i, & i=۱,۲,\dots,k,\\ d_i-\sum_{q=۱}^{p}l_{q,i-qk}u_{q,i-qk}, & p=۱,\dots,m-۱, i=pk+۱,\dots,(p+۱)k,\\ d_i-\sum_{q=۱}^{m}l_{q,i-qk}u_{q,i-qk}, & i=mk+۱,\dots,n. \end{cases} \tag{۶}$$

برای $j = ۱, \ldots, m-۱$ داریم،

$$u_{j,i} = \begin{cases} a_{j,i}, & i = ۱, \ldots, k, \\ a_{j,i} - \sum_{q=۱}^{p} l_{q,i-qk} u_{j+q,i-qk}, & p = ۱, \ldots, m-j-۱, i = pk+۱, \ldots, (p+۱)k, \\ a_{j,i} - \sum_{q=۱}^{p} l_{q,i-qk} u_{j+q,i-qk}, & p = m-j, i = (m-j)k+۱, \ldots, n-jk, \end{cases} \quad (۷)$$

$$u_{m,i} = a_{m,i} \qquad i = ۱, \ldots, n-mk.$$

$$l_{j,i} = \begin{cases} \frac{b_{j,i}}{t_i}, & i = ۱, \ldots, k, \\ \frac{b_{j,i} - \sum_{q=۱}^{p} l_{j+q,i-qk} u_{q,i-qk}}{t_i}, & p = ۱, \ldots, m-j-۱, i = pk+۱, \ldots, (p+۱)k, \\ \frac{b_{j,i} - \sum_{q=۱}^{p} l_{j+q,i-qk} u_{q,i-qk}}{t_i}, & p = m-j, i = (m-j)k+۱, \ldots, n-jk, \end{cases} \quad (۸)$$

$$l_{m,i} = \frac{b_{m,i}}{t_i} \qquad i = ۱, \ldots, n-mk.$$

با استفاده از روابط (۶)، (۷) و (۸) می‌توانیم روابط مربوط به تجزیه دولیتل LU ماتریس $G_{m,n}^{(k)}$ را به‌ازای هر m دلخواه تعیین کنیم. حال اگر ماتریس (۲) ناتکین باشد آنگاه $S = (S_۱, S_۲, \ldots, S_n)^t = (G_{m,n}^{(k)})^{-۱}$ که S_j، j-امین سطر ماتریس $(G_{m,n}^{(k)})^{-۱}$ به‌ازای $j = ۱, ۲, \ldots, n$ است. با استفاده از $G_{m,n}^{(k)}(G_{m,n}^{(k)})^{-۱} = E$ (یک ماتریس همانی $n \times n$ است) داریم، $L_{m,n}^{(k)}(U_{m,n}^{(k)} S) = E$. قرار دهید،

$$L_{m,n}^{(k)} X = E. \quad (۹)$$

در این صورت

$$U_{m,n}^{(k)} S = X, \quad (۱۰)$$

که $X = (X_۱, \ldots, X_n)^t$، $E = (E_۱, \ldots, E_n)^t$ و $S = (S_۱, \ldots, S_n)^t$.
به ترتیب با استفاده از (۹) و (۱۰) داریم،

$$X_i = \begin{cases} E_i, & i = ۱, ۲, \ldots, k, \\ E_i - \sum_{q=۱}^{p} l_{q,i-qk} X_{i-qk}, & p = ۱, \ldots, m-۱, i = pk+۱, \ldots, (p+۱)k, \\ E_i - \sum_{q=۱}^{m} l_{q,i-qk} X_{i-qk}, & i = mk+۱, \ldots, n. \end{cases} \quad (۱۱)$$

و

$$S_i = \begin{cases} \frac{X_i}{t_i}, & i = n, n-۱, \ldots, n-k+۱, \\ \frac{X_i - \sum_{q=۱}^{p} u_{q,i} S_{i+qk}}{t_i}, & p = ۱, \ldots, m-۱, i = n-pk, \ldots, n-(p+۱)k+۱, \\ \frac{X_i - \sum_{q=۱}^{m} u_{q,i} S_{i+qk}}{t_i}, & i = n-mk, \ldots, ۱. \end{cases} \quad (۱۲)$$

اکنون با استفاده از (۶)، (۷)، (۸)، (۱۱) و (۱۲) به ساخت الگوریتم می‌پردازیم.

۱.۲ الگوریتمی برای تعیین معکوس یک ماتریس (r,k)-قطری:

الگوریتم ۲.۲. ورودی: اعداد mو k و ماتریس $G_{m,n}^{(k)}$.
خروجی: ماتریس معکوس، W.
گام اول:

For $p=0,1,\ldots,m$ **do**
For $i=pk+1,\ldots,(p+1)k$ **do**

$$t_i = G_{i,i} - \sum_{q=1}^{p} l_{q,i-qk}u_{q,i-qk}$$

If $t_i=0$ **then** $t_i=z$ **end if.**
For $j=1,2,\ldots,m-p$ **do**

$$u_{j,i} = G_{i,jk+i} - \sum_{q=1}^{p} l_{q,i-qk}u_{j+q,i-qk}$$

$$l_{j,i} = \frac{G_{jk+i,i} - \sum_{q=1}^{p} l_{j+q,i-qk}u_{q,i-qk}}{t_i}$$

end do
if $p\geq 1$**then**
For $j=m,m-1,\ldots,m-p+1$ **do**

$$u_{j,i} = G_{i,jk+i} - \sum_{q=1}^{m-j} l_{q,i-qk}u_{j+q,i-qk}$$

$$l_{j,i} = \frac{G_{jk+i,i} - \sum_{q=1}^{m-j} l_{j+q,i-qk}u_{q,i-qk}}{t_i}$$

end do
end if
end do
end do.

گام دوم:
$det(G_{m,n}^{(k)}) = \prod_{i=1}^{n} t_i$را محاسبه کن، سپس $z=0$را جایگذاری کن. اگر $det(G_{m,n}^{(k)})=0$ ماتریس $G_{m,n}^{(k)}$ تکین است. در اینصورت عبارت (ماتریس تکین است) را در خروجی چاپ کن و توقف الگوریتم.
گام سوم:

For $p=0,1,\ldots,m$ **do**
For $i=pk+1,pk+2,\ldots,(p+1)k$ **do**

$$X_i = E_i - \sum_{q=1}^{p} l_{q,i-qk}X_{i-qk}$$

end do
end do.

گام چهارم:

For $p=0,1,\ldots,m$ **do**
For $i=n-pk,n-pk-1,\ldots,n-(p+1)k+1$ **do**

$$S_i = \frac{X_i - \sum_{q=1}^{p} u_{q,i}S_{i+qk}}{t_i}$$

end do
end do.

گام پنجم:

ماتریس معکوس، $W|_{z=0}$ است.

۳ نتایج عددی

در این بخش با مثال عددی میزان خطای نسبی الگوریتم را نشان می‌دهیم:

مثال ۱.۳. معکوس ماتریس یازده‌قطری ۱۱×۱۱ زیر را بیابید.

$$G^{(۲)}_{۵,۱۱} = \begin{bmatrix} ۲ & ۰ & ۱ & ۰ & -۲ & ۰ & -۱ & ۰ & ۳ & ۰ & ۱ \\ ۰ & -۱ & ۰ & -۱ & ۰ & ۳ & ۰ & -۱ & ۰ & -۱ & ۰ \\ -۱ & ۰ & -۱ & ۰ & -۱ & ۰ & ۲ & ۰ & ۲ & ۰ & ۲ \\ ۰ & ۲ & ۰ & ۳ & ۰ & ۲ & ۰ & ۱ & ۰ & ۱ & ۰ \\ ۱ & ۰ & ۳ & ۰ & ۲ & ۰ & ۱ & ۰ & -۱ & ۰ & -۱ \\ ۰ & ۱ & ۰ & ۴ & ۰ & -۲ & ۰ & ۲ & ۰ & ۱ & ۰ \\ ۳ & ۰ & ۱ & ۰ & ۱ & ۰ & ۱ & ۰ & ۱ & ۰ & ۲ \\ ۰ & ۳ & ۰ & ۱ & ۰ & -۱ & ۰ & ۳ & ۰ & ۱ & ۰ \\ ۴ & ۰ & ۲ & ۰ & -۱ & ۰ & ۳ & ۰ & -۱ & ۰ & -۲ \\ ۰ & ۱ & ۰ & ۲ & ۰ & ۱ & ۰ & ۲ & ۰ & ۳ & ۰ \\ -۱ & ۰ & ۱ & ۰ & ۲ & ۰ & ۱ & ۰ & ۲ & ۰ & ۳ \end{bmatrix}$$

حل: با استفاده از الگوریتم ۲.۲ داریم:

- $det(G^{(۲)}_{۵,۱۱}) = \prod_{i=۱}^{۱۱} t_i = ۵۲۵۰,$

$$W = \begin{bmatrix} -\frac{۳}{۳۵} & ۰ & -\frac{۷۲}{۳۵} & ۰ & -\frac{۱۱}{۵} & ۰ & -\frac{۳۲}{۳۵} & ۰ & \frac{۸}{۵} & ۰ & \frac{۸۲}{۳۵} \\ ۰ & -\frac{۱۱}{۲۵} & ۰ & \frac{۱۱}{۲۵} & ۰ & -\frac{۹}{۲۵} & ۰ & \frac{۲}{۲۵} & ۰ & -\frac{۱}{۵} & ۰ \\ \frac{۳}{۷} & ۰ & \frac{۴۴}{۷} & ۰ & ۷ & ۰ & \frac{۲۵}{۷} & ۰ & -۵ & ۰ & -\frac{۵۴}{۷} \\ ۰ & \frac{۱}{۵۰} & ۰ & \frac{۱۱}{۷۵} & ۰ & \frac{۱۶}{۷۵} & ۰ & -\frac{۷}{۵۰} & ۰ & -\frac{۱}{۱۵} & ۰ \\ -\frac{۱۶}{۳۵} & ۰ & -\frac{۲۷۹}{۳۵} & ۰ & -\frac{۴۲}{۵} & ۰ & -\frac{۱۵۹}{۳۵} & ۰ & \frac{۳۱}{۵} & ۰ & \frac{۳۴۴}{۳۵} \\ ۰ & \frac{۷}{۲۵} & ۰ & \frac{۴}{۷۵} & ۰ & -\frac{۱}{۷۵} & ۰ & \frac{۱}{۲۵} & ۰ & \frac{۱}{۱۵} & ۰ \\ -\frac{۶}{۳۵} & ۰ & -\frac{۳۹}{۳۵} & ۰ & -\frac{۷}{۵} & ۰ & -\frac{۲۹}{۳۵} & ۰ & \frac{۶}{۵} & ۰ & \frac{۵۹}{۳۵} \\ ۰ & \frac{۳۳}{۵۰} & ۰ & -\frac{۳۷}{۷۵} & ۰ & \frac{۲۸}{۷۵} & ۰ & \frac{۱۹}{۵۰} & ۰ & \frac{۲}{۱۵} & ۰ \\ -\frac{۸}{۳۵} & ۰ & -\frac{۳۳۲}{۳۵} & ۰ & -\frac{۵۱}{۵} & ۰ & -\frac{۲۰۲}{۳۵} & ۰ & \frac{۳۸}{۵} & ۰ & \frac{۴۱۷}{۳۵} \\ ۰ & -\frac{۲}{۵} & ۰ & \frac{۱}{۱۵} & ۰ & -\frac{۴}{۱۵} & ۰ & -\frac{۱}{۵} & ۰ & \frac{۱}{۳} & ۰ \\ \frac{۱۲}{۳۵} & ۰ & \frac{۳۲۳}{۳۵} & ۰ & \frac{۴۹}{۵} & ۰ & \frac{۱۹۸}{۳۵} & ۰ & -\frac{۳۷}{۵} & ۰ & -\frac{۳۹۸}{۳۵} \end{bmatrix},$$

و میزان خطای نسبی برابر است با،

$$\frac{\|GW - I\|_F}{\|I\|_F} = ۲٫۹۲۴۶ \times ۱۰^{-۱۵}.$$

الگوریتم ۲.۲ را برای n، m و kهای متفاوت اجرا و نتایج زیر را بدست آوردیم. در جدول ۱ زمان برحسب ثانیه و الگوریتم با نرم‌افزار $Matlab$(۲۰۱۹) انجام شده است و نتایج عددی بدست آمده با نتایج عددی دستور inv در نرم‌افزار $Matlab$ از لحاظ زمان اجرا مقایسه شده‌اند. همانطورکه نتایج این بخش نشان می‌دهند الگوریتم ارائه شده در این مقاله برای ماتریس‌هایی با بعد بالا خصوصا زمانی که $m \ll n$ باشد، نتایج قابل قبولی را نشان می‌دهد.

جدول ۱: مقایسه زمان الگوریتم۲.۲ با دستور invدر $Matlab$ و محاسبه خطا

n	m	k	زمان الگوریتم جدید	زمان دستور inv	$\frac{\|GW-I\|_F}{\|I\|_F}$
۳۰۰۰	۹	۶	۰٫۷۰۳۷	۰٫۷۸۴۲	۳٫۳۶۸۳ × ۱۰$^{-۱۲}$
۴۰۰۰	۱۰	۷	۱٫۱۶۹۴	۱٫۷۸۲۱	۵٫۶۸۳۸ × ۱۰$^{-۱۱}$
۵۰۰۰	۲۰	۱۰	۳٫۱۰۱۸	۳٫۳۷۴۶	۳٫۹۰۵۶ × ۱۰$^{-۱۱}$
۶۰۰۰	۲۰	۸	۴٫۲۰۵۶	۵٫۶۶۰۹	۳٫۱۳۹۶ × ۱۰$^{-۱۱}$
۱۰۰۰۰	۳۰	۱۵	۱۷٫۰۹۶۷	۳۱٫۶۴۴۳	۲٫۷۳۱۳ × ۱۰$^{-۱۱}$
۱۲۰۰۰	۵۰	۲۰	۳۴٫۸۳۴۳	۵۹٫۴۰۵۴	۱٫۱۹۹۱ × ۱۰$^{-۱۰}$

مراجع

[۱] م. شمس سولاری، م. رسولی، *الگوریتمی برای محاسبه معکوس هر ماتریس r-قطری*، موجک ها و جبرخطی (۷)۳ ۷۹-۷۴، ۱۴۰۰.

[2] J. Jia, S. Li, Symbolic algorithms for the inverses of general k-tridiagonal matrices, *Computers and Mathematics with Applications*, 70, 3032–3042, 2015.

[3] M. El-Mikkawy, A generalized symbolic Thomas algorithm, *Applied Mathematics*, 3, 342-345, 2012.

[4] M. El-Mikkawy, F. Atlan, A novel algorithm for inverting a general k-tridiagonal matrix, *Applied Mathematics Letters*, 32, 41-47, 2014.

[5] F. Diele, L. Lopez, The use of the factorization of five-diagonal matrices by tridiagonal Toeplitz matrices, *Applied Mathematics Letters*, 11, 61-69, 1998.

[6] Y. Lin, X. Lin, A novel algorithm for inverting a k-pentadiagonal matrix, *Journal of Shaanxi University of Science and Technology*, 1, 578-582, 2016.

یازدهمین سمینار بین‌المللی
جبر خطی و کاربردهای آن
بهمن ۱۴۰۰

استفاده از رگرسیون خطی در حل مدل طراحی شبکه دوسطحی

محمود صدرا[۱] و مهدی زعفرانیه[۲,*]

[۱] دانشجوی ارشد علوم کامپیوتر، دانشگاه حکیم سبزواری، سبزوار، ایران

[۲] گروه ریاضی کاربردی، دانشگاه حکیم سبزواری، سبزوار، ایران

چکیده

از جمله راه های کنترل ترافیک شهری، احداث خیابانهای جدید می باشد. اما با توجه به محدودیت همیشگی بودجه، همواره پاسخ این سوال که ازمیان چندین خیابان کاندیدا برای احداث، کدام موارد انتخاب شود تا بیشترین اثر را در کاهش بار ترافیکی معابر داشته باشد، بسیار حیاتی است. این مساله، مساله "طراحی شبکه حمل و نقل گسسته "[۱] DNDP نامیده می شود. در این مقاله یک مساله دو سطحی معرفی می شود که هدف آن انتخاب بهینه یالهای جدید برای احداث در یک شبکه حمل و نقل است. از یک روش ترکیبی یادگیری ماشین نظارت شده برای حل این مساله دوسطحی استفاده شده است. مثالهای عددی در محیط پایتون پیاده سازی شده است. برای آماده سازی شبکه شهری از نرم افزار QGIS و پلاگین AequilibraE آن بهره برده ایم.

واژه‌های کلیدی: مساله طراحی شبکه حمل و نقل گسسته(DNDP)،برنامه ریزی دوسطحی،مساله تخصیص ترافیک(TAP)، ،رگرسیون خطی،یادگیری ماشین نظارت شده

کد موضوع‌بندی ریاضی [۲۰۱۰]: 68T05، 91A80، 90C35

۱ مقدمه

مسائل طراحی شبکه حمل و نقل (TNDP) به سه گروه عمده تقسیم می شوند که عبارتند از مسائل طراحی شبکه حمل و نقل گسسته(DNDP) که با پارامترهای گسسته شبکه همچون اضافه کردن یا کم کردن یک یال مرتبط هستند، مسائل طراحی شبکه حمل و نقل پیوسته(CNDP) که با پارامترهای پیوسته شبکه همچون افزایش یا کاهش ظرفیت یک یال مرتبط هستند، مسائل طراحی شبکه حمل و نقل مخلوط(MNDP) که بطور همزمان به هر دو نوع ویژگی می پردازند. یکی از مسائل معروف DNDP مساله انتخاب یک سبد بهینه از میان چندین خیابان کاندیدا برای احداث است تا بیشترین اثر را در کاهش بار ترافیکی معابر داشته باشد. مساله DNDP یک مساله دو سطحی است که در آن ابتدا مدیران شهری اقدام به انتخاب خیابانهای جدید برای احداث می کنند و سپس مسافران با توجه به شرایط جدید و براساس اصول دوگانه وردراپ و مساله تخصیص ترافیک (TAP) اقدام به انتخاب مسیرها نموده و جریان ترافیک معابر را شکل می دهند. در مساله سطح بالا هدف بیشترین کاهش زمان کل سفر شبکه و هزینه لازم برای احداث خیابانهای جدید است.

۲ مساله تخصیص ترافیک

طراحی یک سیستم حمل و نقل شهری در چهار مرحله مطالعه و انجام می شود که شامل این مراحل است [۴]: تولید سفر، توزیع سفر، تنوع سفر، تخصیص سفر. با توجه به تقاضاهای سفری که در مرحله تولید سفر بدست آورده ایم، میزان سفر بین تمام نقاط را محاسبه می کنیم. به اینصورت که تمام نقاط را بصورت زوجهای مبدا-مقصد (OD) در نظر می گیریم و حجم سفر مورد تقاضا میان

* سخنران. آدرس ایمیل: sadra@hsu.ac.ir

[۱] discrete network design problem

آنها را مشخص می کنیم. منظور از نوع سفر، وسیله نقلیه ای است که مسافر از آن استفاده می کند. با استفاده از اطلاعات سیستم حمل و نقل شهری، میزان ترافیکی که از هر خیابان و جاده عبور می کند را مشخص می کنیم و آن را "مساله تخصیص ترافیک" می نامیم. این کار بر اساس اصول دوگانه وردراپ انجام می شود. این دو اصل را وردراپ در سال ۱۹۵۲ معرفی کرد که بر اساس آنها سیستم در حالت پایدار هم از دید مسافران و هم از دید مدیر سیستم بهینه عمل می کند [۴].

اصل اول: در حالتی که سیستم پایدار شده است همواره مسافران از مسیر که کمترین هزینه را دارد سفر خود را انجام می دهند و هیچ مسافری نمی تواند با انتخاب مسیر دیگری هزینه سفرش را کمتر کند.

اصل دوم: در حالتی که سیستم پایدار شده است متوسط هزینه سفر در کل سیستم کمترین مقدار می باشد

۱.۲ تعریف مساله تخصیص ترافیک:

مساله تخصیص ترافیک بصورت مساله بهینه سازی محدب زیر فرمول نویسی می شود(نمادهای مورد استفاده در ضمیمه مقاله آمده است):

$$min \sum_{a\in A} \int_{\circ}^{x_a} s_a(t)dt,$$

subject to

$$\sum_{p\in P_k} h_{pk} = r_k, \forall k$$

$$x_a = \sum_k \sum_{p\in P_k} \delta_{kap} h_{pk}, \forall a \in A$$

$$h_{pk} \geq \circ, \forall a \in A, \forall k$$

ما در این مقاله برای حل مدل ریاضی مساله از بسته pyomo و حل کننده های ipopt و glpk در پایتون استفاده و به کمک آن مساله تخصیص ترافیک را بعنوان یک مساله بهینه سازی ریاضی حل کردیم.برای آشنایی با روشهای حل Tap می توانید به [۴] نگاه کنید

۳ مسائل برنامه ریزی دو سطحی

مساله بهینه سازی دو سطحی در واقع مربوط به برنامه اقتصادی استکلبرگ در نظریه بازیهاست. بازیکن اول که تابع هدف سطح اول(پیشرو) را دارد، سعی می کند با انتخاب مقدار برای متغیر تصمیم ،y مقدار تابع F(x،y) را کمینه کند، بازیکن دوم که تابع هدف سطح دوم (پیرو) را دارد با اطلاع کامل از تصمیم بازیکن اول سعی می کند با انتخاب مقدار برای متغیر تصمیم ،x تابع هدف f(x،y) را کمینه کند. به این ترتیب تصمیم بازیکن دوم وابسته است به تصمیم بازیکن اول [۱]. ساختار کلی برنامه ریزی دوسطحی بصورت زیر می باشد:

$$min \quad F(x,y) \quad x \in X, y \in Y$$

$$s.t \quad G(x,y) \leq \circ$$

$$min \quad f(x,y) \quad y \in Y$$

$$s.t \quad g(x,y) \leq \circ$$

برخی از روشهای حل مسائل دو سطحی عبارتند از: روش نقطه راسی، روش شاخه و کران، روش محور گیری مکمل، روش تابع جریمه، روش ناحیه ایمن، روش رویکرد کاهشی، روش KKT ، یادگیری ماشین [۳].

۴ مساله طراحی شبکه حمل و نقل گسسته DNDP

مساله NDP به روش مشخص کردن تعدادی از یالهای بهینه بر روی یک شبکه اطلاق می شود. در این مقاله یک مساله DNDP دو سطحی بررسی می شود. در سطح اول دیدگاه مدیران شهری جهت انتخاب یک سبد بهینه از خیابانهای(یالها) جدید مورد بررسی قرار می گیرد. در سطح دوم دیدگاه کاربران شبکه جهت انتخاب بهترین مسیرهای ممکن بین مبدا و مقصدهای متفاوت در قالب مساله TAP بررسی می شود. در مدل کلاسیک DNDP هدف انتخاب یک سبد بهینه از خیابانهای جدید با توجه به یک بودجه محدود است. در این مقاله محدودیت بودجه بعنوان قسمتی از تابع هدف سطح اول درنظر گرفته شده است.

$$min_{x_a,y_a} Z(x_a,y_a) = \sum_{a\in A\cup A'} x_a.t_a(x_a) + \sum_{a\in A'} c_a.y_a$$

subject to

$$y_a = \{۰,۱\}, \qquad a \in A'$$

$$min_x \sum_{a\in A\cup A'} \int_{۰}^{x_a} t_a(x_a)dx,$$

subject to

$$\sum_{k\in P_i} h_k = q_i, \qquad i \in I$$

$$x_a = \sum_{i\ inI} \sum_{k\in P_i} h_k \delta_{a,p}, \qquad a \in A\cup A'$$

$$x_a \geq ۰, \qquad a \in A\cup A'$$

$$x_a \leq M.y_a, \qquad a \in A'$$

تابع هدف عبارت از مجموع زمان سفر کل شبکه (T۱) و هزینه کل احداث یالهای انتخابی (T۲) است. برای متناسب سازی تاثیر دو جمله تابع هدف، از شکل نرمال شده آنها استفاده کرده ایم.

$$T۱ = \sum_{a\in A\cup A'} x_a.t_a(x_a)$$

$$T۲ = \sum_{a\in A'} c_a.y_a$$

$$min_{x_a,y_a} Z(x_a,y_a) = \frac{\sum_{a\in A\cup A'} x_a.t_a(x_a) - Min(T۱)}{Max(T۱) - Min(T۱)} + \frac{\sum_{a\in A'} c_a.y_a - Min(T۲)}{Max(T۲) - Min(T۲)}$$

۵ روش حل

ما در حل مساله از رگرسیون خطی که یکی از کاربردهای مهم جبر خطی در هوش مصنوعی و یادگیری ماشین است استفاده کرده ایم. رگرسیون خطی نوعی تابع پیش‌بینی‌کننده خطی است که در آن متغیر وابسته (متغیری که قرار است پیش‌بینی شود) به صورت ترکیبی خطی از متغیرهای مستقل پیش‌بینی می‌شود. بدین معنی که هر کدام از متغیرهای مستقل در ضریبی که در فرایند تخمین برای آن متغیر به‌دست آمده ضرب می‌شودو جواب نهائی مجموع حاصل‌ضرب‌ها به علاوه یک مقدار ثابت خواهد بود که آن هم در فرایند تخمین به‌دست آمده‌است.

ساده‌ترین نوع رگرسیون خطی، رگرسیون خطی ساده است که بر خلاف رگرسیون خطی چندگانه، تنها یک متغیر مستقل دارد. نوع دیگر رگرسیون خطی، رگرسیون خطی چندمتغیره است که در آن به جای پیش‌بینی یک متغیر وابسته چندین متغیر وابسته پیش‌بینی می‌شود. ما از رگرسیون خطی چندگانه استفاده کرده ایم وبرای این منظور تمام داده ها را بصورت بردار و ماتریس در آورده ایم تا بتوانیم

از الگوریتمهای جبر خطی برای پیش بینی مقادیر متغیرهای وابسته استفاده کنیم.متغیر وابسته مقدار تابع هدف است و متغیرهای مستقل همان یالهای کاندیدا می باشند.

ابتدا مقادیر کمینه و بیشنه هر کدام از جملات تابع هدف به تنهایی را محاسبه می کنیم که مربوط می شود به حالاتی که تمام یالها انتخاب شوند یا هیچکدام از یالها انتخاب نشوند. سپس به سراغ یک فرآیند تکراری محاسبه مقدار تابع هدف با توجه به یالها انتخاب شده می رویم به اینصورت که ابتدا یک بردار شدنی از یالهای کاندیدا (Y) در نظر می گیریم که ما بردار Y تمام صفر را در نظر گرفته ایم و با استفاده از آن تابع سطح پایین را که TAP است حل می کنیم.سپس جریانهای بدست آمده را به تابع سطح بالا می فرستیم تا مقدار تابع هدف (z) را بدست آید. حالا یک مجموعه از یالها داریم وبه ازای آنها یک مقدار برای تابع هدف.حال به کمک این دو مقدار یک تابع رگرسیون می سازیم. و از روی آن، مقدار جدیدی برای Y حدس می زنیم و آنرا به تابع سطح پایین TAP می فرستیم تا جریانهای جدید را محاسبه کند و به تابع بالا بفرستد و مقدار جدیدی برای تابع هدف بدست آورد حالا به کمک این دو مقدار و مقادیر قبلی Y و z یک تابع رگرسیون جدید می سازیم و دوباره مقدار جدیدی برای Y حدس می زنیم و کار در قالب یک حلقه ادامه پیدا می دهیم.

شرط خروج از حلقه را هم می توان کم شدن اختلاف میان دو z متوالی از یک حد خاص در نظر گرفت و هم تعداد تکرار. در مثالهای عملی شرط تعداد تکرار بهتر وموثرتر عمل کرده است.نمودار روش در ضمیمه مقاله آمده است.

۶ نتایج اصلی

در این مقاله یک مدل دو سطحی طراحی شبکه مورد بررسی قرار گرفته است که یک مدل np-hard است و حتی بدست آوردن یک جواب بهینه محلی برای آن نیز np-hard است. ما به کمک یک روش یادگیری ماشین نظارت شده، این مساله را حل کردیم و بر روی یک شبکه نسبتا بزرگ جوابهای مناسب در زمانها بسیار کوتاه بدست آوردیم.

۷ نتایج عددی

برای مطالعه موردی شبکه شهر SiouxFalls که یک الگوی استاندارد و شناخته شده در حوزه حمل و نقل است، استفاده شده است که در این آدرس (https://github.com/bstabler/TransportationNetworks/) در دسترس است. نقشه شهر SiouxFalls که با نرم افزار QGIS تهیه شده است و در آن تمام یالها یک طرفه هستند در ضمیه مقاله امده است.

مساله این است که ۱۰ یال با هزینه های مشخص، کاندیدای احداث هستند. هدف یافتن سبدی از یالها است که با کمترین هزینه، بیشترین تاثیر را در کاهش ترافیک داشته باشند. جدول یالهای کاندیدا و هزینه آنها در ضمیمه آمده است.

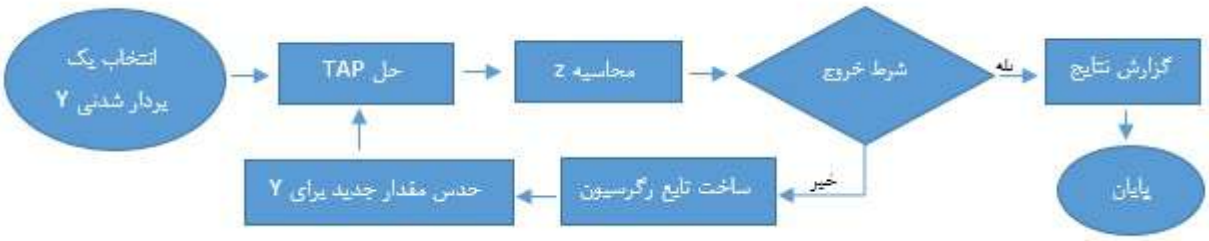
ما از یک سیستم با ۶ گیگا بایت حافظه و پردازنده Core i5 نسل ۵ برای اجرای برنامه استفاده کرده ایم و پس از ورود اطلاعات ، برنامه را اجرا کردیم که در یک حلقه با ۲۰ تکرار مساله را حل کرد. مدت زمان اجرا ۲۳ دقیقه و مقدار بیشینه تابع هدف ۱,۲۰۰۵۹۷۳ و مقدار کمینه تابع هدف ۰,۴۳۷۷۴۸۷۳ گزارش شد که در تکرار ۱۷ام اتفاق افتاد.

جدول ۱: هزینه احداث یالها.

شماره یال	هزینه احداث(میلیون دلار)
۱	۴۴
۳	۵۹
۱۲	۷۹
۱۵	۸۶
۳۳	۵۸
۳۶	۴۴
۴۱	۵۹
۴۶	۷۹
۵۰	۸۶
۵۵	۵۸

یک شبکه با مجموعه N از گره ها و مجموعه A از یالها	$G = (A,N)$
مجموعه یالهای کاندیدا برای ایجاد	A'
مجموعه زوجهای مبدا-مقصد شبکه G	$K \subseteq N \times N$
مجموعه مسیرهای ساده بین نود ابتدایی و انتهایی kامین زوج	P_k
جریان موجود در مسیر pام در مجموعه P_k	h_{pk}
جریان یال a	x_a
زمان سفری که کاربر در زمان سفر در یال a با آن روبرو می شود	$s_a(x_a)$
تقاضای سفر بین Kامین زوج	r_k
ماتریس واقع شدن یک یال در یک مسیر که در آن δ_{kap}، یک است اگر مسیر pام از kامین زوج شامل یال a باشد و در غیر اینصورت صفر	$\Delta = [\delta_{kap}]$
بردار یالهای کاندیدا برای اضافه شدن به شبکه (1: یال اضافه شود. 0: یال اضافه نشود)	Y
هزینه احداث یال a	C_a
کل بودجه	B
تقاضای سفر زوج iام	q_i
تابع تقاضای سفر روی زوج مبدا-مقصد kام	$g_k(\pi)$
عدد بسیار بزرگ	M

شکل ۱: جدول نمادهای بکار رفته در مقاله .



شکل ۲: نمودار روش حل .

شماره تکرار	مقدار تابع هدف	یالهای انتخاب شده برای احداث									
		55	50	46	41	36	33	15	12	3	1
1	1.1031999										*
2	1.0168463					*					
3	1.10262456	*									
4	1.02006345						*				
5	0.93859915					*	*				
6	1.2005973					*	*			*	
7	1.07822271	*				*	*				
8	0.90122969				*	*	*				
9	1.01080675				*	*					
10	0.93173014				*		*				
11	0.93362567				*						
12	0.49705059			*	*	*	*				
13	0.48706806		*	*	*	*	*	*	*		
14	0.46927982		*	*	*		*	*	*		
15	0.49038387		*	*	*		*	*			
16	0.49293671		*	*	*		*		*		
17	0.43774873			*	*		*	*	*		
18	0.44185862			*	*	*	*	*	*		
19	0.47827769			*	*		*	*			
20	0.49926913			*	*		*		*		

شکل ۳: نتایج ۲۰ تکرار.

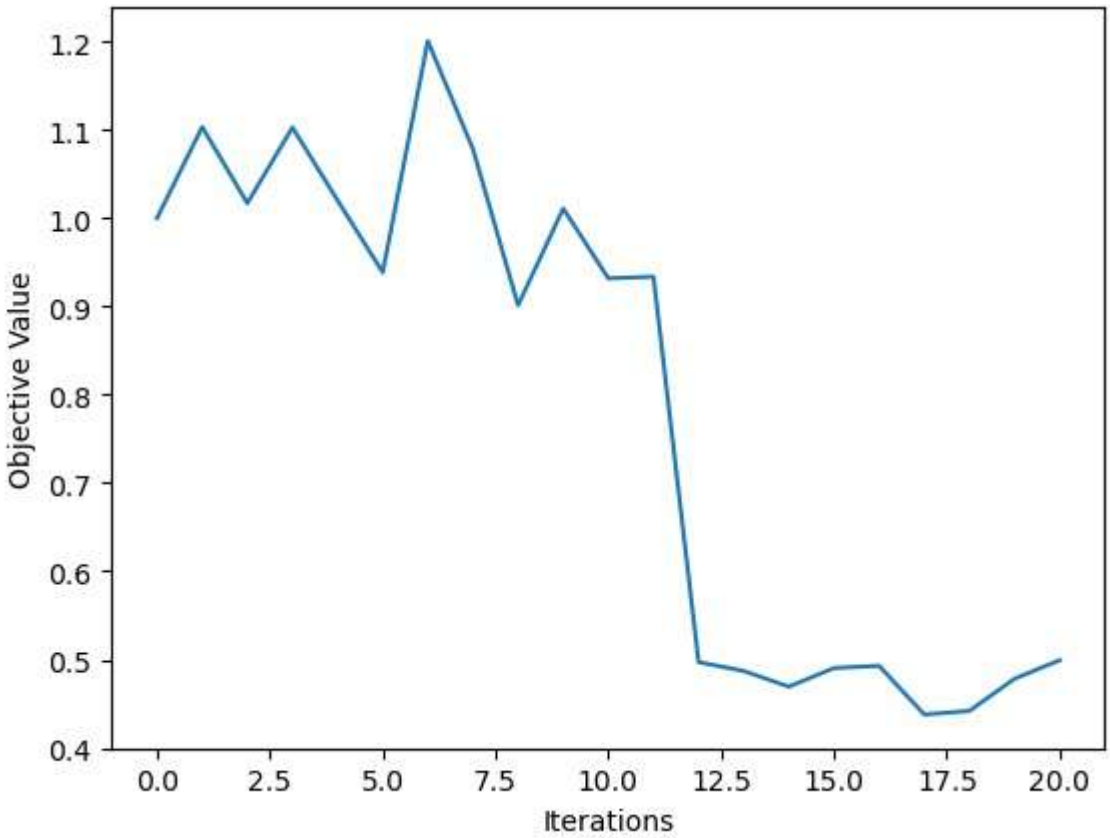
مراجع

[1] Abareshi, Maryam, and Mehdi Zaferanieh. ”A bi-level capacitated P-median facility location problem with the most likely allocation solution.” Transportation Research Part B: Methodological 123 (2019): .1-20

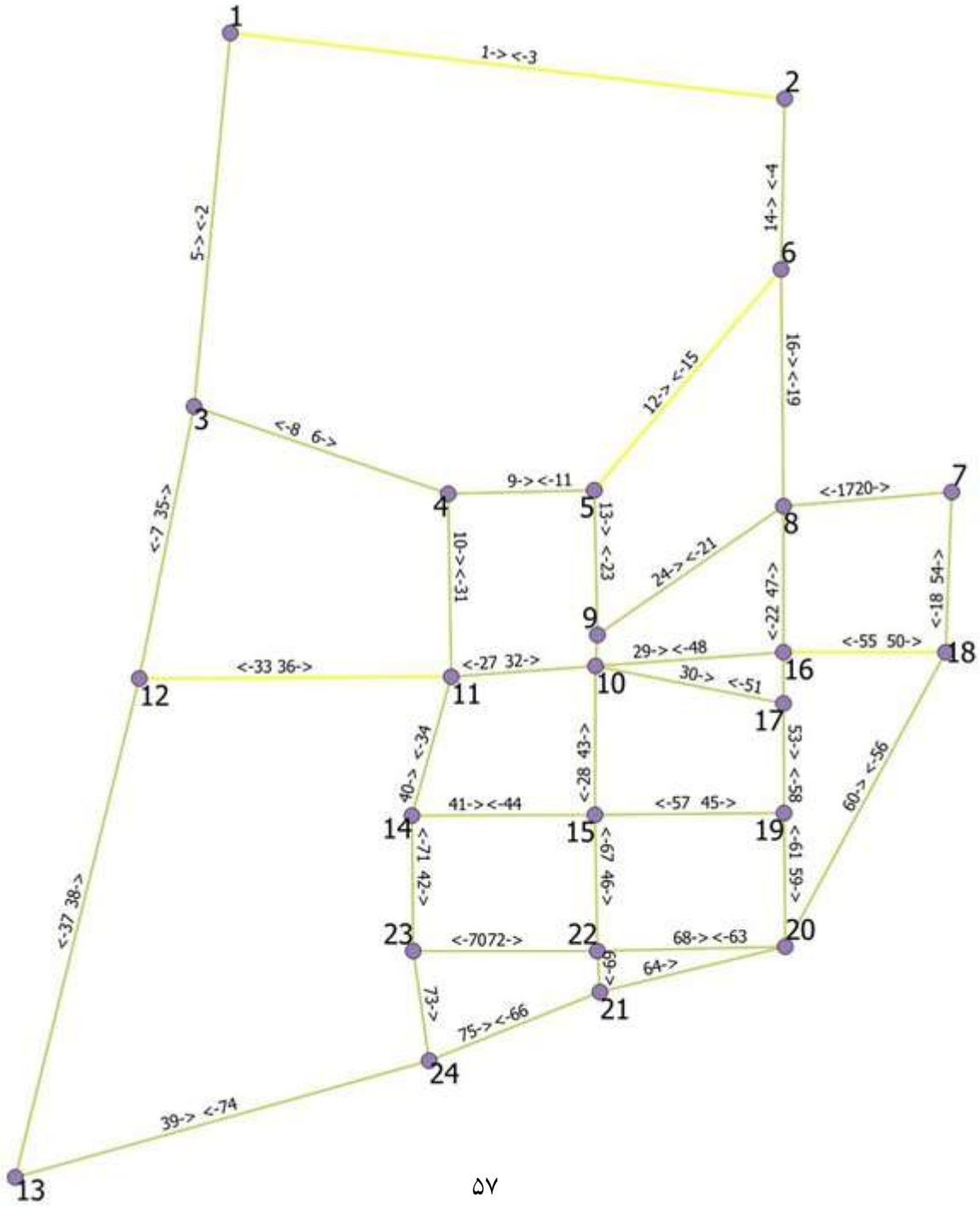
[2] Bagloee, Saeed Asadi, et al. ”A hybrid machine-learning and optimization method to solve bi-level problems.” Expert Systems with Applications 95 (2018): .142-152

[3] Dempe, Stephan, and Alain Zemkoho. Bilevel optimization. Springer, .2020

[4] Patriksson, Michael. The traffic assignment problem: models and methods. Courier Dover Publications, 2015



شکل ۴: نمودار نتایج ۲۰ تکرار .



شکل ۵: شبکه شهر SiouxFalls که با نرم افزار QGIS تهیه شده است .

پنهان‌نگاری تصویر با استفاده از تبدیل قابک و تجزیه مقادیر تکین دوقطری

عفت گلپر رابوکی[۱] و معصومه صفری سیاهگورابی[۲,*]

[۱]بخش ریاضی، دانشگاه قم، قم، ایران

[۲]بخش ریاضی، دانشگاه قم، قم، ایران

چکیده

پنهان‌نگاری و آنالیز آن از زمینه‌های مهم تحقیقاتی برای مخفی کردن اطلاعات است، که در آن تصویر رمز را در یک تصویر میزبان پنهان می‌کنیم. در این مقاله طرح پنهان‌نگاری تصویر با استفاده از تبدیل قابک و مقادیر تکین دوقطری را بیان می کنیم. بازسازی کامل، پایداری و ثبات از جمله مزایای تبدیل قابک می‌باشند. درنتیجه در این مقاله از تبدیل قابک به‌عنوان یک تبدیل مناسب برای تجزیه تصویر میزبان و بدست آوردن ضرایب استفاده شده است. اطلاعات محرمانه یا تصویر رمز در مقادیر تکین ضرایب قابک پنهان می‌شوند و تصویر پوششی بدست می‌آید. ارزیابی‌های گوناگونی برای اثربخش بودن این طرح صورت گرفته‌است که نشان می‌دهد تصاویر رمز کیفیت مناسبی دارند و غیرقابل تشخیص هستند.

واژه‌های کلیدی: تصویر پنهان‌نگاری شده، مقادیر تکین دوقطری، مقادیر تکین قطری، تبدیل قابک، تبدیل موجک.

کد موضوع‌بندی ریاضی [۲۰۱۰]: 94A12، 65T60، 15A15

۱ مقدمه

امروزه با پیشرفت سریع فناوری اطلاعات و شبکه‌های اجتماعی دیگر مفهوم ((آنچه می‌بینید همان چیزی است که دریافت می‌کنید)) صادق نیست. زیرا افراد می‌توانند به‌راحتی اطلاعات مخفی را در شکل‌های مختلف با استفاده از اینترنت به هر نقطه از جهان ارسال یا دریافت کنند. همچنین با رشد سریع اطلاعات چندرسانه‌ای، شبکه‌های کامپیوتری تحت حملات مختلفی قرار می‌گیرند. پس حفاظت و امنیت اطلاعات دیجیتال یکی از مسائل مهم می‌باشد. پنهان‌نگاری[۱] روشی است که می‌توان داده‌های محرمانه را با استفاده از محصولات چند رسانه‌ای صوتی، تصویری و متنی انتقال داد به‌طوریکه قابل تشخیص نباشد. پنهان‌کاوی هنر شکستن پنهان‌نگاری است یعنی سعی دارد وجود اطلاعات پنهان شده را تشخیص دهد. هدف پنهان‌نگاری این است که تصویر میزبان را به‌گونه‌ای تغییر دهد که اطلاعات پنهان شده در آن قابل تشخیص نباشد. به دلیل افزایش استفاده از اینترنت، تصاویر به‌طور گسترده‌ای برای پنهان‌نگاری مورد استفاده قرار می‌گیرند. هدف از پنهان‌نگاری تصویر به حداکثر رساندن ظرفیت جاسازی و در عین حال به حداقل رساندن قابلیت تشخیص تصویر رمز است. در پنهان‌نگاری تصویر ابتدا تصویر میزبان را با استفاده از یک تبدیل مناسب تجزیه کرده و سپس پیام مخفی را در ضرایب تبدیل پنهان می‌کنیم. از ویژگی‌های مهم طراحی الگوریتم پنهان‌نگاری تصویر، نامحسوس بودن، ظرفیت جاسازی، استحکام در برابر حملات و مقاومت در برابر پنهان‌کاوی است. الگوریتم پنهان‌کاوی خاصیت‌های تصویر میزبان و رمز را مورد مطالعه قرار می‌دهد و سعی در تشخیص اطلاعات پنهان شده دارد. در این‌صورت باید بین ظرفیت جاسازی و غیرقابل تشخیص بودن یک تعادل برقرار باشد به‌گونه‌ای که اطلاعات بیشتری پنهان شود ولی قابل تشخیص نباشد. مراحل پنهان‌نگاری تصویر به‌طور کلی برای جاسازی و استخراج به‌صورت زیر است.

۱. یک تصویر میزبان(اصلی) و یک تصویر رمز را انتخاب نمایید.

۲. تصویر میزبان را با استفاده از یک تبدیل مناسب تجزیه کنید و ضرایب تبدیل را انتخاب نمایید.

[*]سخنران. آدرس ایمیل: safary.masomeh@yahoo.com

[۱]Steganograpgy

۳. ضرایب تبدیل که قرار است در آن جاسازی انجام شود را انتخاب کنید و تصویر رمز را در آن به‌گونه‌ای مخفی می‌کنید که در تصویر میزبان تغییری محسوس حاصل نشود.

۴. ضرایب تبدیل را با استفاده از تعبیه مناسب تغییر دهید تا داده مخفی پنهان شوند.

۵. با استفاده از عکس تبدیل مناسب استفاده شده، تصویر پوششی را بدست آورید.

۶. فردی که تصویر پوششی را می‌گیرد آن را با تبدیل مناسب تجزیه می‌نماید و ضرایب اصلاح شده را شناسایی می‌کند.

۷. با استفاده از ضرایب تبدیل، اطلاعات پنهان شده را بدست می‌آورد.

هنگامی که تصویر پوششی ارسال می شود، مزاحم سعی می‌کند اطلاعات پنهان شده را شناسایی کند یا از بین ببرد. در این‌صورت یک طرح پنهان‌نگاری با وجود حملات باید غیرقابل تشخیص باشد. مطالعات مختلفی برای ارائه الگوریتم پنهان‌نگاری انجام شده‌است که هر کدام سعی در بهبود شفافیت و غیرقابل تشخیص بودن دارد. پنهان‌نگاری در دو حوزه زمان و فرکانس انجام می‌شود. در مقاله‌ای از الگوریتم پنهان‌نگاری بر اساس روش کم اهمیت‌ترین بیت LBS و الگوریتم ژنتیک استفاده شده‌است که در حوزه مکان صورت گرفته‌است. پنهان‌نگاری تصویر در حوزه فرکانس با استفاده از تبدیل کسینوسی (DCT) شروع شد. ربیع و همکاران پنهان‌نگاری با استفاده از DCT و فرمت $JPEG$ را مطرح کردند. سعیدی و همکاران الگوریتم دیگری برای پنهان‌نگاری با استفاده از DCT بیان کرده‌اند. با توسعه فناوری اطلاعات، استفاده از تبدیل موجک گسسته (DWT) در پردازش تصویر مورد بررسی قرار گرفت. الگوریتم‌های پنهان‌نگاری با استفاده از DWT به‌گونه‌ای است که اطلاعات محرمانه در هر یک از زیرباندها با روش‌های مختلفی جاسازی می‌شوند. یک الگوریتم پنهان‌نگاری تصویر براساس DWT توسط آتابی و همکاران پیشنهاد شد. مشابه طرح‌های پنهان‌نگاری براساس DCT، طرح‌های پنهان‌نگاری مختلفی براساس DWT وجود دارد. طرح‌های پنهان‌نگاری با استفاده از تبدیل موجک گسسته DWT در مقالات مختلف مورد بررسی قرار گرفته‌است. فخردانش و همکاران الگوریتم پنهان‌نگاری براساس DWT که سازگاری بیشتری با بینایی انسان HVS دارد را بیان کرده‌اند. سابهدار و مانکار در مقاله از تجزیه مقادیر تکین SVD و تبدیل موجک گسسته DWT استفاده کرده‌اند [۴]. **تصویر میزبان:** در پنهان‌نگاری تصویر، تصویر میزبان همان تصویر اصلی است که قرار است در آن تصویر دیگری پنهان شود. **تصویر رمز:** در پنهان‌نگاری تصویر، تصویر رمز تصویری است که جزء اطلاعات محرمانه می‌باشد و هدف پنهان‌نگاری تصویر، مخفی کردن این تصویر است. **تصویر پوششی:** تصویر پوششی، تصویری است که تصویر رمز در آن مخفی شده‌است برای مخفی کردن تصویر رمز در تصویر میزبان، باید قسمتی از تصویر رمز به‌گونه‌ای در تصویر میزبان قرار گیرد که با مشاهده تصویر پوششی حاوی تصویر رمز، تغییرات قابل تشخیص نباشد.

۲ تبدیل قابک

قابک[۱] بسیار شبیه موجک است اما تفاوت مهمی دارد. موجک دارای یک عملگر مقیاس‌بندی $\Phi(t)$ و عملگر موجک $\Psi(t)$ است ولی قابک دارای یک عملگر مقیاس‌بندی $\Phi(t)$ و دو عملگر قابک $\Psi_1(t)$ و $\Psi_2(t)$ است [۲]. تابع مقیاس‌گذاری و موجک‌ها در قابک به‌صورت زیر تعریف می‌شوند

$$\Phi(t) = \sqrt{2}\sum_n h_0(n)\Phi(2t-n), \quad \Psi_i(t) = \sqrt{2}\sum_n h_i(n)\Phi(2t-n), \quad i = 1,2$$

که در آن $h_0(n)$ فیلتر پایین‌گذر(مقیاس‌گذاری) و $h_i(n)$ فیلترهای بالاگذر هستند. تابع $f(t)$ را می‌توان به‌صورت زیر تقریب زد

$$f(t) \simeq \sum_{k=-\infty}^{\infty} c(k)\Phi_k(t) + \sum_{j=0}^{+\infty}\sum_{k=-\infty}^{k=+\infty} d_1(j,k)\Psi_{1,j,k}(t) + d_2(j,k)\Psi_{2,j,k}(t) \qquad (1)$$

به‌طوری‌که

$$c(k) = \int f(t)\Phi_k(t)\mathrm{d}t, \quad d_i(j,k) = \int f(t)\Psi_{i,j,k}(t)\mathrm{d}t, \quad i = 1,2.$$

مجموعه توابع $\{\Phi_k(t), \Psi_{i,j,k}(t) \mid j,k \in \mathbb{Z}, \quad j \geq 0 \quad i \in \{1,2\}\}$ تشکیل قابک برای $L^2(\mathbb{R})$ می‌دهد [۱].

[۱] Framelet

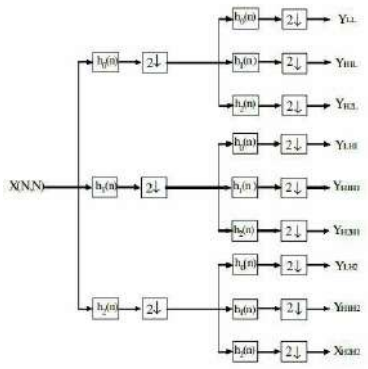
فیلترهای متعامد $h_0(n)$، $h_1(n)$ و $h_2(n)$ باید شرایط بازسازی را به‌طور کامل برآورده کنند. شرایط بازسازی کامل در حوزه فرکانس روابط (۲) و (۳) می‌باشند.

$$H_0(Z)H_0(\frac{1}{Z}) + H_1(Z)H_1(\frac{1}{Z}) + H_2(Z)H_2(\frac{1}{Z}) = 2 \quad (2)$$

$$H_0(-Z)H_0(\frac{1}{Z}) + H_1(-Z)H_1(\frac{1}{Z}) + H_2(-Z)H_2(\frac{1}{Z}) = 0 \quad (3)$$

۱.۲ تبدیل قابک دوبعدی

یک تبدیل قابک ۲-بعدی معادل دو تبدیل قابک یک‌بعدی می‌باشد. برای این کار ابتدا قابک یک‌بعدی را روی سطرها اعمال کرده و سپس تبدیل قابک یک‌بعدی را روی ستون داده‌های حاصل از تبدیل سطری اعمال می‌کنیم. شکل ۱ نشان‌دهنده تبدیل قابک ۲-بعدی یک سیگنال ورودی $X(N,N)$ می‌باشد [۱].



شکل ۱: تبدیل قابک ۲-بعدی سیگنال ورودی $X(N,N)$ [۱]

۲.۲ مراحل تبدیل یک سطح از قابک برای سیگنال ۲-بعدی X

۱- سیگنال ورودی باید به طول $N \times N$ باشد که N زوج است و طول فیلترهای آنالیز $N >=$ نیز می‌باشد.

۲- ماتریس عملیاتی W را با توجه به N بسازید [۱].

$$W=\begin{bmatrix}
h_{\circ}(\circ) & h_{\circ}(۱) & h_{\circ}(۲) & h_{\circ}(۳) & h_{\circ}(۴) & h_{\circ}(۵) & h_{\circ}(۶) & h_{\circ}(۷) & h_{\circ}(۸) & h_{\circ}(۹)\\
\circ & \circ & \dots & \circ & \circ & & & & & \\
\circ & \circ & h_{\circ}(\circ) & h_{\circ}(۱) & h_{\circ}(۲) & h_{\circ}(۳) & h_{\circ}(۴) & h_{\circ}(۵) & h_{\circ}(۶) & h_{\circ}(۷)\\
h_{\circ}(۸) & h_{\circ}(۹) & \dots & \circ & \circ & & & & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
\vdots & \vdots & \dots & \vdots & \vdots & & & & & \\
h_{\circ}(۲) & h_{\circ}(۳) & h_{\circ}(۴) & h_{\circ}(۵) & h_{\circ}(۶) & h_{\circ}(۷) & h_{\circ}(۸) & h_{\circ}(۹) & \circ & \circ\\
\circ & \circ & \dots & h_{\circ}(\circ) & h_{\circ}(۱) & & & & & \\
h_{۱}(\circ) & h_{۱}(۱) & h_{۱}(۲) & h_{۱}(۳) & h_{۱}(۴) & h_{۱}(۵) & h_{۱}(۶) & h_{۱}(۷) & h_{۱}(۸) & h_{۱}(۹)\\
\circ & \circ & \dots & \circ & \circ & & & & & \\
\circ & \circ & h_{۱}(\circ) & h_{۱}(۱) & h_{۱}(۲) & h_{۱}(۳) & h_{۱}(۴) & h_{۱}(۵) & h_{۱}(۶) & h_{۱}(۷)\\
h_{۱}(۸) & h_{۱}(۹) & \dots & \circ & \circ & & & & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
\vdots & \vdots & \dots & \vdots & \vdots & & & & & \\
h_{۱}(۲) & h_{۱}(۳) & h_{۱}(۴) & h_{۱}(۵) & h_{۱}(۶) & h_{۱}(۷) & h_{۱}(۸) & h_{۱}(۹) & \circ & \circ\\
\circ & \circ & \dots & h_{۱}(\circ) & h_{۱}(۱) & & & & & \\
h_{۲}(\circ) & h_{۲}(۱) & h_{۲}(۲) & h_{۲}(۳) & h_{۲}(۴) & h_{۲}(۵) & h_{۲}(۶) & h_{۲}(۷) & h_{۲}(۸) & h_{۲}(۹)\\
\circ & \circ & \dots & \circ & \circ & & & & & \\
\circ & \circ & h_{۲}(\circ) & h_{۲}(۱) & h_{۲}(۲) & h_{۲}(۳) & h_{۲}(۴) & h_{۲}(۵) & h_{۲}(۶) & h_{۲}(۷)\\
h_{۲}(۸) & h_{۲}(۹) & \dots & \circ & \circ & & & & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
\vdots & \vdots & \dots & \vdots & \vdots & & & & & \\
h_{۲}(۲) & h_{۲}(۳) & h_{۲}(۴) & h_{۲}(۵) & h_{۲}(۶) & h_{۲}(۷) & h_{۲}(۸) & h_{۲}(۹) & \circ & \circ\\
\circ & \circ & \dots & h_{۲}(\circ) & h_{۲}(۱) & & & & &
\end{bmatrix}_{\frac{۳N}{۲}\times N}$$

۳- تبدیل سطرهای سیگنال ورودی را با استفاده از ضرب ماتریس عملیاتی در سیگنال ورودی به‌صورت $Y=[W]_{\frac{۳N}{۲}\times N}[X]_{N\times N}$ است که Y در آن یک ماتریس انتقال با ابعاد $\frac{۳N}{۲}\times N$ است.

۴- تبدیل ستون‌های سیگنال ورودی که به‌صورت زیر می‌باشد.

(الف) با استفاده از ماتریس عملیاتی که در گام ۳ به‌دست آمده است سطرها را به‌صورت زیر می‌توان جابه‌جا کرد، یعنی ترانهاده ماتریس انتقال Y را به‌دست آورده $[Y']_{N\times\frac{۳N}{۲}}=[Y]^T_{N\times\frac{۳N}{۲}}$.

(ب) ماتریس عملیاتی را در ماتریس فوق ضرب کرده و ستون‌های ماتریس اصلی(سیگنال ورودی) را به‌دست آورده $YY=[W]_{\frac{۳N}{۲}\times N}[Y']_{N\times\frac{۳N}{۲}}$ در انتها ماتریس تبدیل شده نهایی به‌صورت $Y_{\circ}=[YY]'_{\frac{۳N}{۲}\times\frac{۳N}{۲}}$ است که به‌طور خلاصه می‌توان نوشت $Y_{\circ}=W.X.W^T$ [۳].

۳.۲ معکوس تبدیل قابک۲-بعدی

برای بازسازی سیگنال اصلی از سیگنال حاصل از تأثیر قابک مراحل زیر را طی کرد [۳].

۱- فرض کنید $Y_{\circ}$ با ابعاد $\frac{۳N}{۲}\times\frac{۳N}{۲}$ ماتریس تبدیل شده قابک باشد.

۲- ماتریس بازسازی را با استفاده از ماتریس‌های عملیاتی به‌صورت $T=[W]^T_{N\times\frac{۳N}{۲}}$ بدست آورید.

۳- ستون‌ها را با استفاده از ضرب ماتریس بازسازی در ماتریس تبدیل شده قابک به‌صورت $YY = [T]_{N\times\frac{۳N}{۲}}[Y_۰]_{\frac{۳N}{۲}\times\frac{۳N}{۲}}$ بدست آورید.

۴- بازسازی سطرها به‌صورت زیر انجام دهید:

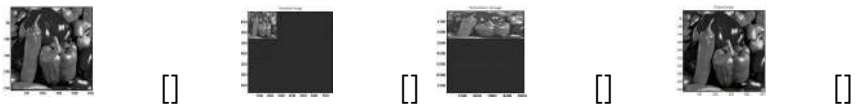
(آ) با استفاده از بازسازی ستون‌ها در مرحله سوم خواهیم داشت $Y = [YY]'_{\frac{۳N}{۲}\times N}$.

(ب) با استفاده از ماتریس بازسازی و ماتریس Y، سیگنال ورودی را بازسازی می‌کنیم $X = [W^T]_{N\times\frac{۳N}{۲}}[Y']_{\frac{۳N}{۲}\times N}$.

که به‌طور خلاصه برابر است با $X = TY_۰T^T$ یا $X = W^TY_۰W$. حال با توجه به مطالب بیان شده در این بخش، هر تصویر که دارای ۲ بعد است ابتدا تبدیل قابک یک‌بعدی را روی سطرهای آن انجام داده و سپس روی ستون‌های آن اعمال می‌کنیم که در این‌صورت به ۲ بعد افزایش می‌یابد. در نتیجه با اعمال این فیلترها و تکرار آن‌ها، تصویر به ۹ زیر باند تجزیه می‌شود، که در تصویر ۳ نشان داده شده است.

LL	$LH_۱$	$LH_۲$
$H_۱L$	$H_۱H_۱$	$H_۱H_۲$
$H_۲L$	$H_۲H_۱$	$H_۲H_۲$

شکل ۲: تبدیل قابک بر روی تصویر



شکل ۳: تصویر زیر باندهای قابک تصویرمیزبان (آ) [۱]

۳ طرح پنهان‌نگاری تصویر با استفاده از تبدیل قابک و $BSVD$

در این بخش طرح پیشنهادی براساس تبدیل قابک و مقادیر ویژه دو قطری که در مقاله [۴] بیان شده‌است ارائه می‌شود. مراحل مربوط به پنهان‌نگاری و استخراج تصویر رمز به‌صورت زیر می‌باشد. فرض کنید $C = f(x,y),\ ۰ \le x \le M, ۰ \le y \le N$ تصویر میزبان اصلی است که در آن $f(x,y)$ مقادیر تصویر در نقطه (x,y) است. به همین ترتیب فرض کنید $H = f(z,w)$، $۰ \le z \le M$، $۰ \le w \le N$ تصویر رمز باشد.

۱.۳ مراحل پنهان کردن تصویر رمز

۱- از یک تصویر خاکستری به‌عنوان تصویر رمز استفاده شده‌است.

۲- به‌عنوان اولین لایه امنیتی از تبدیل آرنولد استفاده می‌شود.

۳- اکنون تبدیل قابک را بر تصویر میزبان اعمال کنید. ضرایب فرکانس بالا انتخاب نمایید. $\Im(C) \Longleftarrow C^\theta$ به‌طوریکه $\Im(.)$ تجزیه قابک FT را نشان می‌دهد.

۴- تجزیه $BSVD$ روی باندهای فرکانسی موردنظر برای مثال فرکانس بالا حاصل از تبدیل قابک اعمال می‌کنیم و مقادیر تکین را به‌دست آورید. سپس روی زیرباندهای موردنظر تصویر میزبان ابتدا یک تجزیه SVD دوقطری و سپس بر روی ماتریس دوقطری حاصل یک تجزیه SVD می‌زنیم. در واقع به‌صورت زیر عمل می کنیم. فرض کنید A زیرباند موردنظر باشد. با استفاده از الگوریتم کلوب-کاهان تجزیه $BSVD$ زیرباند A را به‌دست می‌آوریم.

$$A = U_A \times B \times V_A^T \qquad (۴)$$

که در آن U_A و V_A ماتریس‌های متعامد هستند و B یک ماتریس دوقطری است. حال تجزیه SVD را بر ماتریس B اعمال می کنیم.

$$B = U_B \times S \times V_B^T \qquad (۵)$$

باجایگزین کردن (۵) در (۴) خواهیم داشت $A = U_A \times U_B \times S \times V_B^T \times V_A^T$. ماتریس S را برای مخفی کردن تصویر رمز انتخاب می‌کنیم. زیرا باعث می شود تا شانس بازیابی اطلاعات تصویر رمز کاهش یابد.

$$SV = BSVD[C^\theta]$$

۵- با استفاده از فاکتور تعبیه مناسب α، تصویر رمز را در مقادیر تکین زیرباند تصویر میزبان به‌صورت $SV_{new} = SV + \alpha * H$ قرار دهید.

۶- معکوس SVD را انجام دهید و ضرایب اصلاح شده را به‌صورت $SV' = SVD[SV_{new}]$ به‌دست آورید.

۷- برای به‌دست آوردن تصویر رمز از معکوس تبدیل قابک (IFT) $C' \Longleftarrow \Im^{-1}(C^\theta)$ استفاده کنید.

۲.۳ مراحل استخراج تصویر پنهان‌شده

روش استخراج شامل مراحل مشابهی است که باید با تصویر رمز انجام دهید [۴].

۱- با استفاده از تبدیل قابک، تصویر رمز را تجزیه کنید. ضرایب قابک زیرباند اصلاح‌شده را استخراج کنید $C'^\theta \Longleftarrow \Im(C')$.

۲- $BSVD$ را بر روی ضرایب قابک اعمال کنید $SV' = BSVD(C'^\theta)$.

۳- اطلاعات تعبیه شده را با استفاده از همان فاکتور مناسب تعبیه α، بازیابی کنید $H' = \frac{1}{\alpha}(SV_{new} - SV')$.

۴- با استفاده از کلید مخفی به‌عنوان دوره تحول آرنولد، تصویر رمز H را بازیابی کرده و با تصویر رمز اصلی H مقایسه کنید.

شکل ۴: تصویر پنهان‌نگاری شده [۴]

شکل ۴ نتیجه پنهان‌نگاری حاصل از بکارگیری الگوریتم تبدیل قابک با استفاده از $BSVD$ را نشان می دهد.

۳.۳ معیارهای ارزیابی

برای سنجش شفافیت تصاویر حاصل از الگوریتم پیشنهادی پنهان‌نگاری از پایگاه داده تصاویر $USC - SIPI$ استفاده شده‌است. از شکل ۴ مشخص شد تصاویر پنهان‌نگاری شده بدست آمده از کیفیت تصویری بالاتری برخوردار است. جدول ۱ ارزیابی عینی با استفاده از معیاراهای مختلف کیفیت تصویر را نشان می‌دهد. تمام معیارهای کیفیت دارای مقادیر قابل قبولی هستند و عدم محسوس بودن را تأیید می‌کنند.

۴ نتیجه‌گیری

این مقاله یک طرح پنهان‌نگاری تصویری مبتنی بر $BSVD$ و تبدیل قابک را ارائه می‌دهد. این طرح از مخفی‌سازی فرکانس زمان بهتر و تغییر شکل قالب استفاده می‌کند. استفاده از مقادیر تکین دوقطری باعث افزایش ظرفیت و امنیت می‌شود. روش پیشنهادی با انجام آزمایشات مختلف برای بررسی عدم محسوس بودن استحکام در برابر انواع حملات تصویر رمز که شامل عملیات پردازش تصویر است، اعتبارسنجی می‌شود. در نتیجه کیفیت بصری تصاویر رمز به‌دست آمده با طرح پیشنهادی در مقایسه با آثار موجود برتر است.

جدول ۱: معیارهای کیفیت تصویر برای تصاویر رمز

Stegeo image	PSNR	SSIM	IF	UQI	FSIM
Splash	۶۹,۵۹۲۴	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۷	۰,۹۹۹۹
Tiffany	۶۹,۴۷۹۶	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۶	۰,۹۹۹۹
Baboon	۶۹,۵۶۸۶	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۹
Lena	۶۹,۴۴۱۳	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۹
Airplane	۶۹,۶۱۰۲	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۴	۰,۹۹۹۹
Lake	۶۹,۵۱۸۶	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۹
Peppers	۶۹,۵۳۰۱	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۹
Pirate	۶۹,۵۲۴۵	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۷	۰,۹۹۹۹
Living Room	۶۹,۵۶۷۳	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۹
Boat	۶۹,۵۲۹۶	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۹	۰,۹۹۹۹

مراجع

[1] Al-Taai, H. N. (2008). A novel fast computing method for framelet coefficients. American Journal of Applied Sciences, 5(11), 1522-1527.

[2] Selesnick, I. W. (2001). Smooth wavelet tight frames with zero moments.

[3] Sultan, N. H., Khammas, N. H. A., & Najm, Z. H. (2021). Image watermarking based on framelet transform. Periodicals of Engineering and Natural Sciences (PEN), 9(1), 37-47.

[4] Subhedar, M. S., & Mankar, V. H. (2020). Secure image steganography using framelet transform and bidiagonal SVD. Multimedia Tools and Applications, 79(3), 1865-1886

ایده سازی و ارائه یک روش تکراری برای محاسبه جواب یک دستگاه خطی مقید مبتنی بر روش تکرار نقطه ثابت

مجید ادیب[۱] * و زهرا عبادی[۱]†

[۱] دانشکده علوم، گروه ریاضی، دانشگاه زنجان، زنجان، ایران

چکیده

یکی از مهم‌ترین و اساسی‌ترین مباحث ریاضیات، به دلیل کاربرد وسیع آن در اکثر شاخه‌های ریاضی و حتی علوم دیگر، دستگاه معادلات خطی هستند. به همین دلیل از دیرباز تاکنون محققان و پژوهشگران مطالعات خود را نه تنها در این زمینه متوقف نکرده‌اند بلکه همچنان این‌گونه پژوهش‌ها با هدف ارائه الگوریتم‌های کاراتر و کاربردی‌تر ادامه دارد. در این مقاله بر مبنای روش تکرار نقطه ثابت یک روش تکراری برای حل دستگاه معادلات خطی معرفی و آن را به دستگاه معادلات خطی مقید نیز تعمیم می‌دهیم و شرایط لازم و کافی همگرایی آن را نیز بیان می‌کنیم.

واژه‌های کلیدی: دستگاه معادلات خطی، روش تکرار نقطه ثابت

کد موضوع‌بندی ریاضی [۲۰۱۰]: 15A09، 65F05، 65F015

۱ مقدمه

همان‌طور که میدانیم در روش تکرار نقطه ثابت برای حل یک معادله غیرخطی $f(x) = 0$ ابتدا معادله با انتخاب یک تابع $g(x)$ مناسب، به فرم $x = g(x)$ تبدیل می‌شود و سپس در صورت مناسب بودن g دنباله $x_{n+1} = g(x_n)$ به نقطه ثابت g و یا به‌طور معادل به ریشه f همگرا خواهد شد [۳].

در این مقاله بر مبنای روش تکرار نقطه ثابت و با انتخاب $f(x) = Ax - b$ روشی تکراری و کارآمد که به جواب دستگاه مقید $Ax = b$، $x \in S$ که در آن S زیر فضایی از $\mathbb{C}^n$ همگرا گردد، معرفی می‌کنیم. ابتدا توجه کنید که اگر A وارون‌پذیر باشد و $B = A^{-1}$، آنگاه $BA = I$ و یا به طور معادل $I = 2I - BA$ و در نتیجه

$$B = B(2I - AB). \tag{۱}$$

روش تکرار نقطه ثابت و رابطه (۱) می‌تواند ایده‌ای برای تعریف یک دنباله تکراری با هدف محاسبه تقریبی از وارون ماتریس A به صورت زیر در اختیار ما قرار می‌دهد:

$$B_{k+1} = B_k(2I - AB_k), \tag{۲}$$

که در صورت یک انتخاب مناسب برای تقریب آغازین B_0 می‌تواند به دنباله‌ای همگرا تبدیل شود. بعلاوه می‌توان از این ایده برای محاسبه تقریب‌هایی از وارون‌های خاص ماتریس‌هایی که لزوما وارون طبیعی آن‌ها وجود ندارد نیز استفاده کرد.

همچنین توجه کنید که اگر A وارون پذیر باشد جواب دستگاه $Ax = b$ عبارت است از $x = Bb$ و لذا

$$x = x + B(b - Ax). \tag{۳}$$

با به‌کارگیری روش تکرار نقطه ثابت روی رابطه (۳) می‌توان دنباله تکراری زیر را برای محاسبه جواب دستگاه معادلات خطی $Ax = b$ و یا تقریبی از آن، حتی در حالتی‌که A وارون پذیر نباشد، تعریف کرد:

$$x_{k+1} = x_k + B_{k+1}(b - Ax_k). \tag{۴}$$

* آدرس ایمیل: madib@znu.ac.ir

† سخنران. آدرس ایمیل: ebadizahra1112@gmail.com

۲ یک روش تکراری برای حل دستگاه معادلات خطی مقید

دستگاه معادلات مقید زیر را در نظر بگیرید:

$$Ax = b, \quad A \in \mathbb{C}^{m\times n}, \quad b \in \mathbb{C}^m, \quad x \in S \tag{۵}$$

با توجه به ایده و مقدمات مذکور در بخش قبل یک روش تکراری برای حل دستگاه معادلات مقید (۵) به صورت زیر تعریف می‌کنیم:

$$B_{k+1} = B_k(2I - AB_k), \tag{۶}$$

$$x_{k+1} = x_k + B_{k+1}(b - Ax_k). \tag{۷}$$

توجه داشته باشید که اگر A، $m \times n$ باشد، A دارای شبه وارون‌های مختلفی از جمله وارون مور پنروز و یا وارون درازین می‌باشد. همچنین توجه دارید که شبه وارون $A \in \mathbb{C}^{m\times n}$ ماتریسی است مانند $B \in \mathbb{C}^{n\times m}$ که دارای خواص خاصی است و ما در این مقاله قصد ورود به آن‌ها را نداریم، در این خصوص خواننده علاقه‌مند می‌تواند به مرجع [۲] رجوع کند. تنها این نکته را متذکر می‌شویم که در استفاده از رابطه (۶) در الگوریتم (۷)-(۶)، تقریب اولیه B_0 می‌بایست تقریبی از شبه وارون مورد نظر کاربر در مسئله استفاده گردد. قضیه زیر شرایط لازم و کافی برای همگرایی الگوریتم مذکور را فراهم می‌کند:

قضیه ۱.۲. *فرض کنید $B_0 = \beta Y$ یک تقریب اولیه برای وارون ماتریس A باشد که در آن β یک عدد حقیقی غیر صفر و $Y \in \mathbb{C}^{n\times m}$ در رابطه‌ی $R(Y) \subseteq S$ صدق کنند(فضای برد $R(Y)=Y$).*
الف)- اگر $b \in AS$ آنگاه دنباله تکراری حاصل از رابطه (۷) با تقریب اولیه $x_0 \in S$، به یک جواب دستگاه معادلات خطی (۵) همگراست اگر و تنها اگر $\rho(P_{AT} - \beta AY) < 1$ و بعلاوه داریم

$$\lim_{k\to\infty} x_k = x_0 + \beta Y(P_{(AS)^\perp} + \beta AY)^{-1}(b - Ax_0) \tag{۸}$$

که در آن P_{AS}، نشان دهنده ماتریس تصویر متعامد روی AS و $\rho(.)$ نمایانگر شعاع طیفی است

ب) اگر دستگاه دارای جواب منحصر بفرد باشد در این صورت الگوریتم فوق به جواب زیر همگراست:

$$x = \beta Y(P_{N(Y)} + \beta AY)^{-1}b,$$

که در آن $N(Y)$ نمایانگر فضای پوچ Y است. □

۳ کاربرد روش ارائه شده در خصوص حل مسئله کمترین مربعات

فرض کنید $A \in \mathbb{C}^{m\times n}$. توجه کنید که اگر دستگاه معادلات $Ax = b$ ناسازگار باشد، در این صورت مسئله یافتن جواب کمترین مربعات دستگاه $Ax = b$ مطرح می‌شود. مسئله کمترین مربعات روشی است که بردار x، بردار باقیمانده را به حداقل می‌رساند. به عبارت دیگر مسئله کمترین مربعات یافتن جواب مسئله زیر است:

$$min\|Ax - b\|_2 \tag{۹}$$

که در آن $\|.\|$ نشان‌دهنده نرم اقلیدسی یک بردار است. همان‌طور که می‌دانیم حل مسئله (۹) معادل است با حل دستگاه خطی و سازگار:

$$A^*Ax = A^*b. \tag{۱۰}$$

بنابراین برای محاسبه جواب کمترین مربعات دستگاه

$$Ax = b, \quad A \in \mathbb{C}^{m\times n}, \quad m \geq n, \quad b \in \mathbb{C}^n$$

می‌توان روش ارائه شده توسط روابط (۷)-(۶) را روی دستگاه سازگار و نامقید $A^*Ax = A^*b$ اعمال کرد.

۴ کاربرد روش ارئه شده در خصوص یافتن جواب با کمترین طول اقلیدسی یک دستگاه معادلات خطی

فرض کنید $A \in \mathbb{C}^{m\times n}$، $m \leq n$ و دستگاه $Ax = b$ سازگار باشد. در این صورت مسئله یافتن جواب با کمترین طول اقلیدسی دستگاه با حل مسئله زیر معادل است:

$$\begin{cases} min & \|x\|_۲ \\ s.t & Ax = b \end{cases}$$

همان‌طور که می‌دانیم اگر جواب این مسئله را x_+ بنامیم این جواب در $R(A^*)$ قرار دارد [۱]. لذا برای محاسبه جواب با کمترین طول اقلیدسی یک دستگاه خطی با تعریف $S = R(A^*)$ می‌توان روش (۷)-(۶) را روی مسئله مقید زیر اعمال نمود.

$$Ax = b, \quad x \in S$$

۵ نتیجه‌گیری

در این مقاله ابتدا یک روش تکراری برای حل دستگاه معادلات خطی ارائه شد و سپس به محاسبه جواب یک دستگاه معادله خطی روی یک زیر فضای دلخواه S تعمیم داده شد. بعلاوه در حالت خاص وقتی‌که $S = R(A^*)$ این الگوریتم در عمل مسئله جواب با کمترین طول اقلیدسی را حل می‌کند. لازم به ذکر است که می‌توان پارامتر β مذکور در قضیه ۲,۱ را چنان تعیین کرد که الگوریتم دارای سرعت همگرایی مطلوب باشد.

مراجع

[1] P.E. Gill, W. Murray, *Numerical linear algebra and optimization*, Avalon Publishing, 1991.

[2] K. Manjunatha Prasad, R. B. Bapat, The generalized moore-penrose inverse, *Elsevier Science Publishing Co.*, 655 (1992), 59-69.

[3] E. Popova, Generalization of a parametric fixed-point iteration , *Apple Math Mech.*, 4 (2004), 680-681.

برخی از ویژگی‌های مجموعه‌ی $M_n(B)$

لیلا فضل‌پر[۱] * و علی آرمندنژاد[۱]

[۱]گروه ریاضی، دانشگاه ولی‌عصر (عج) رفسنجان، رفسنجان، ایران

چکیده

فرض کنید B نیم‌حلقه‌ی بولی دوتایی باشد. به عبارت دیگر $B=\{۰,۱\}$ به‌طوری که $۱+۱=۱$ و سایر عملیات جمع و ضرب در آن مشابه اعداد حقیقی است. فرض کنیم $M_n(B)$ مجموعه‌ی همه‌ی ماتریس‌های $n\times n$ را نشان دهد به‌طوری که درایه‌های آن متعلق به B باشد. در این مقاله، ماتریس‌های معکوس‌پذیر متعلق به $M_n(B)$ را شناسایی و معرفی می‌کنیم.

واژه‌های کلیدی: نیم‌حلقه‌ی بولی، ماتریس‌های معکوس‌پذیر.

کد موضوع‌بندی ریاضی [۲۰۱۰]: 16Y60، 15A09.

۱ مقدمه

این مقاله به ماتریس‌های متعلق به مجموعه‌ی $M_n(B)$ می‌پردازد. این قبیل ماتریس‌ها دارای کاربرد مهمی در نظریه‌ی مدار الکتریکی هستند. اکنون تعدادی از تعاریف اولیه و مورد نیاز در مقاله را معرفی می‌کنیم.

یک نیم‌حلقه‌ی سه‌تایی $(S,+,\cdot)$ است به‌طوری که $(S,+)$ و $(S,\cdot)$ نیم‌گروه باشند و برای هر $x,y,z\in S$

۱) $x\cdot(y+z)=x\cdot y+x\cdot z$

۲) $(y+z)\cdot x=y\cdot x+z\cdot x$.

یک نیم‌حلقه‌ی $(S,+,\cdot)$ به ترتیب تعویض پذیر جمعی و یا ضربی نامیده می‌شود، اگر

$$x+y=y+x \qquad \text{or}$$
$$x\cdot y=y\cdot x$$

$۱\in S$ همانی سه‌تایی $(S,+,\cdot)$ نامیده می‌شود اگر

$$x\cdot ۱=۱\cdot x=x, \qquad \forall x\in S$$

عضو صفر و عضو یک نیم‌حلقه منحصر به فرد هستند.

عنصر $x\in S$ به ترتیب وارون‌پذیر جمعی و یا ضربی نامیده می‌شود اگر $y\in S$ وجود داشته باشد که

$$x+y=y+x=۰ \qquad \text{or}$$
$$xy=yx=۱.$$

در این مقاله نیم‌حلقه‌ی بولی $B=\{۰,۱\}$ را تعریف کرده و ماتریس‌های معکوس‌پذیر $M_n(B)$ را معرفی خواهیم کرد.

* سخنران. آدرس ایمیل: fazlparleila65@gmail.com

۲ قضایا و مثال ها

تعریف ۱.۲. نیم‌حلقه‌ی بولی یک نیم‌حلقه‌ی جابه‌جایی دارای صفر است به‌طوری که

$$x^۲ = x, \qquad \forall x \in S.$$

مثال ۲.۲. فرض کنیم X یک مجموعه‌ی ناتهی باشد. تعریف می‌کنیم

$$A \oplus B = A \cup B, \quad A \cdot B = A \cap B, \qquad \forall A, B \in P(X)$$

پس $(P(X), \oplus, \cdot)$ یک نیم‌حلقه‌ی بولی است و $\emptyset$ به عنوان صفر و X به عنوان عضو همانی آن می‌باشد و $\emptyset$ تنها عضو وارون‌پذیر جمعی $(P(X), \oplus, \cdot)$ و $A \oplus A = A$.

مثال ۳.۲. فرض کنید

$$S = \{۰\} \cup \left[\frac{۱}{۲}, ۱\right],$$

و تعریف می‌کنیم

$$\begin{aligned} x \oplus ۰ = ۰ \oplus x = x, &\qquad \forall x \in S \\ x \oplus y = \frac{۱}{۲}, &\qquad \forall x, y \in \left[\frac{۱}{۲}, ۱\right] \\ x \circ y = \min\{x, y\}, &\qquad \forall x, y \in S \end{aligned}$$

$(S, \oplus, \cdot)$ یک نیم‌حلقه‌ی بولی با عضو خنثی ۰ و عضو همانی ۱ می‌باشد و ۰ تنها عضو وارون‌پذیری جمعی برای $(S, \oplus, \cdot)$ می‌باشد.

یکی از ساختارهای جبری روی $\{۰, ۱\}$ به صورت زیر است:

$+$	۰	۱
۰	۰	۱
۱	۱	۱

$\cdot$	۰	۱
۰	۰	۰
۱	۰	۱

که در مرجع [۶]، نیم‌حلقه‌ی بولی نامیده می‌شود و با نماد B نمایش داده می‌شود. نیم‌حلقه‌ی B معادل با جبر بولی زیرمجموعه‌های مجموعه‌ی تک عضوی $\{a\}$ است و $۰ = \emptyset$ و $۱ = \{a\}$ و $+$ اجتماع و $\times$ اشتراک است.

توجه: فرض کنیم $M_n(B)$ مجموعه‌ی ماتریس‌های $n \times n$ را نشان می‌دهد به‌طوری که درایه‌های آن‌ها از مجموعه‌ی B باشد.

قضیه ۴.۲ (لوس). *یک ماتریس بولی وارون‌پذیر است اگر و تنها اگر متعامد باشد.*

قضیه ۵.۲ (رادرفورد). *اگر یک ماتریس بولی وارون یک طرفه داشته باشد در این صورت وارون دو طرفه خواهد داشت. به علاوه این وارون در صورت وجود منحصر به فرد است و برابر B^T می‌باشد.*

قضیه ۶.۲. *$B \in M_n(B)$ یک ماتریس معکوس‌پذیر است اگر و تنها اگر B یک ماتریس جایگشت باشد.*

اثبات. فرض کنیم B یک ماتریس جایگشت باشد. چون هر ماتریس جایگشت یک ماتریس متعامد است پس وارون‌پذیر است و $B^T = B^{-۱}$.
اکنون فرض کنیم که B یک ماتریس معکوس‌پذیر باشد. نشان می‌دهیم B یک ماتریس جایگشت است. چون B معکوس‌پذیر است پس در هر سطر یک درایه‌ی غیر صفر دارد. فرض کنیم که درایه‌ی $(۱, ۱)$ ماتریس B ناصفر باشد، این فرض با استفاده از ماتریس جایگشت P، قابل توجیه است. فرض کنیم

$$A = [a_{ij}] = PB,$$

و نشان می‌دهیم که A یک ماتریس جایگشت است. در این صورت داریم

$$A = PB \Longrightarrow B = P^{-۱}A,$$

پس B یک ماتریس جایگشت است.
پس بدون از دست رفتن کلیت فرض می‌کنیم $A = [a_{ij}]$ ماتریس معکوس‌پذیر است و $a_{11} \neq 0$. نشان می‌دهیم که

$$\begin{cases} a_{12} = \cdots = a_{1n} = 0 \\ a_{21} = \cdots = a_{n1} = 0 \end{cases}$$

چون طبق قضیه ۲، $A^{-1} = A^T$ پس داریم $AA^T = I$.

$$\begin{bmatrix} 1 & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} 1 & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \dots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix},$$

$$\begin{aligned} (AA^T)_{12} = 0 &\Longrightarrow a_{21} + \cdots = 0 \Longrightarrow a_{21} = 0, \\ (AA^T)_{13} = 0 &\Longrightarrow a_{31} + \cdots = 0 \Longrightarrow a_{31} = 0, \\ &\vdots \\ (AA^T)_{1n} = 0 &\Longrightarrow a_{n1} + \cdots = 0 \Longrightarrow a_{n1} = 0. \end{aligned}$$

به‌طور مشابه چون $A^TA = I$ داریم

$$\begin{bmatrix} 1 & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \dots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} 1 & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix},$$

$$\begin{aligned} (A^TA)_{12} = 0 &\Longrightarrow a_{12} + \cdots = 0 \Longrightarrow a_{12} = 0, \\ &\vdots \\ (A^TA)_{1n} = 0 &\Longrightarrow a_{1n} + \cdots = 0 \Longrightarrow a_{1n} = 0. \end{aligned}$$

اکنون با استفاده از استقرا حکم را اثبات می‌کنیم.
فرض قضیه برای $M_2(B)$ برقرار است. فرض کنیم قضیه برای ماتریس‌های $(n-1) \times (n-1)$ برقرار باشد. اکنون حکم را برای ماتریس‌های $n \times n$ ثابت می‌کنیم.

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{bmatrix}_{n \times n}, \qquad A = 1 \oplus C,$$

□

ماتریس C یک ماتریس $(n-1) \times (n-1)$ است و درایه های آن متعلق به B است پس طبق فرض استقرا ماتریس جایگشت است و $A = 1 \oplus C$ و چون جمع مستقیم ماتریس های جایگشت ماتریس جایگشت است پس A یک ماتریس جایگشت است.

۳ نتیجه گیری

ماتریس های بولی کاربرد مهمی در نظریه ی مدارهای الکتریکی دارند یک زیر مجوعه خاص از این ماتریس ها مجموعه ی $M_2(B)$ است و ماتریس های متعلق به این مجموعه در صورت وارون پذیری ماتریس های جایگشت هستند.

مراجع

[1] F. Fazlpar and A. Armandnejad, *Some notes on* $M_n(B)$, (preprint).

[2] F. Guzman, The variety of Boolean semirings, *Journal of Pure and Applied Algebra*, 78 (1992), 253–270.

[3] R. D. Luce, A note on Boolean matrix theory, *Proc. Amer. Math. Soc.*, 6 (1952), 382–388.

[4] W. Kuich and A. Salomaa, *Semirings, Automata, languages*, Springer, Berlin, 1986.

[5] D. E. Rutherfod, Inverses of Boolean matrices, *Proc. Glasgow Math. Assoc.*, 6 (1963), 49–53.

[6] N.Sirasuntorn,S.Sombatboriboon,N.udomsub,Inversion of Matrices over booleansemirings, Thai Journal of Mathematics 7(2009),105-113.

بررسی چند جمله ای های روش $DGMRES$

فرنگیس کیانفر *

بخش ریاضی کاربردی، دانشگاه شهید باهنر کرمان، کرمان، ایران

چکیده

یکی ازروشهای رایج برای پیداکردن جواب دستگاه خطی منفرد $Ax = b$ روش تکراری درازین مانده می نیمال $DGMRES$ است . این روش براساس فرایندمتعامدسازی آرنولدی برای بدست آوردن جواب معکوس درازین یک سیستم خطی منفرد سازگاریا ناسازگاراست. چند جمله ایهای مانده وچند جمله ایهای ایده ال این روش در این مقاله مورد بررسی قرار می گیرد. با استفاده از نرم این چند جمله ایها می توان کرانهایی برای نرم بردار باقی مانده بدست آورد.

واژه‌های کلیدی: کلید واژه : مقادیرریتز، مقادیرهارمونیک ریتز، معکوس درازین ، روش مانده می نیمال درازین

کد موضوع‌بندی ریاضی [۲۰۱۰]: 15A06، 15A09، 65F10

۱ مقدمه

روش تکراری مانده مینیمال تعمیم یافته $(GMRES)$ یک روش زیرفضای کریلف است که برمبنای متعامدسازی آرنولدی ، تقریبی ازجواب دستگاه خطی $Ax = b$ رابدست می آورد که درآن ماتریس A یک ماتریس نامنفرداست [۵] . دراین روش بردار $x_\circ$ برداراولیه و $r_\circ = b - Ax_\circ$ بردارمانده اولیه می باشد. فضای کریلف مرتبه k نسبت به ماتریس A وبردار $r_\circ$ بصورت زیرتعریف می شود.

$$\mathbf{K}_k(A, r_\circ) = span\{r_\circ, Ar_\circ, A^{2}r_\circ, \cdots, A^{k-1}r_\circ\}.$$

جواب تقریبی x_k^G درمرحله k به صورت $x_k^G \in x_\circ + \mathbf{K}_k(A, r_\circ)$ بیان می شود، بطوریکه k -امین مانده ، r_k^G باشرط تعامد $r_k^G = b - Ax_k^G \perp A\mathcal{K}_k(A, r_\circ)$ مینییم سازی شود. براساس روابط فوق می توان نرم بردارمانده مرحله $k-$ ام رابصورت معادله زیر نشان داد.

$$\|r_k^G\| = \min_{p\in\pi_k(\circ)}\|p(A)r_\circ\| = \min_{q\in\pi_{k-1}}\|(I - Aq(A))r_\circ\|. \qquad (۱)$$

که درآن $\pi_k(\circ) = \{p(x) \in \pi_k : p(\circ) = 1\}$ می باشد. بسهولت می توان نشا ن دادکه نرم بردارهای مانده درروش مانده مینیمال تعمیم یافته درشرط $r_\circ \geq r_1 \geq r_2 \geq \cdots r_{n-1} \geq r_n = \circ$ صدق می کنند وروش فوق باحداکثر n تکراردرمرحله n به جواب دقیق دستگاه خطی دسترسی پیدا می کند. درواقع هدف تقریب درروش مانده مینیمال تعمیم یافته مشخص کردن چندجمله ای $p^G(z)$ که درآن سمت راست معادله فوق به مقدارمی نیمم می رسد. این چندجمله ای ، چندجمله ای مانده مینیمال تعمیم یافته نامیده می شود. بعلاوه اگر

$$\frac{\|r_k^G\|}{\|r_\circ\|} \leq \min_{p\in\pi_k(\circ)}\|p(A)\| = \min_{q\in\pi_{k-1}}\|(I - Aq(A))r_\circ\|. \qquad (۲)$$

مسئله ایده ال روش مانده مینیمال تعمیم یافته ($GMRES$) حاصل می شود. به چندجمله ای که سمت راست معادله فوق را می نیمم می کند، چندجمله ای ایده ال مانده مینیمال تعمیم یافته نامیده می شود.

*آدرس ایمیل سخنران : kyanfar@uk.ac.ir

فرض کنید ماتریس A نامنفرد باشد، گرین باوم وترفتن درمقاله [۳] نشان دادند، اگر $\min_{p\in\pi_k(0)}\|p(A)\|>0$ انگاه یک چندجمله ای $p^{IG}(x)\in\pi_k(0)$ ایده ال روش مانده مینیمال تعمیم یافته به صورت منحصربفرد وجوددارد‌بطوریکه

$$\min_{p\in\pi_{k-1}}\|I-Ap(A)\|=\min_{p\in\pi_k(0)}\|p(A)\|=\|p^{IG}(A)\| \qquad (۳)$$

برقرااست. بعلاوه اگر $\pi_k(1)=\{p(x)\in\pi_k:\ p \text{ is monic }\}$ باشد و $\min_{p\in\pi_k(1)}\|p(A)\|>0$ آنگاه یک چندجمله ای منحصربفرد ایده ال آرنولدی $p^{IA}(x)\in\pi_k(1)$ وجوددارد بطوریکه دررابطه زیرصدق می کند.

$$\min_{p\in\pi_{k-1}}\|A^k-p(A)\|=\min_{p\in\pi_k(1)}\|p(A)\|=\|p^{IA}(A)\|. \qquad (۴)$$

ریشه های این چندجمله ایها مقادیر ریتز و هارمونیک ریتز می باشند که در زیر تعاریف مربوطه ارایه می شوند.

تعریف ۱.۱. اگر $\mathcal{V}_k$ یک زیرفضایی ازفضای برداری $\mathbb{C}^n$ باشد، انگاه θ_k یک مقدار ریتز ماتریس A نسبت به $\mathcal{V}_k$ نامیده می شود اگر

$$Av_k-\theta_kv_k\perp\mathcal{V}_k,\quad for\ some\ v_k\in\mathcal{V}_k,\ \ v_k\neq 0.$$

تعریف ۲.۱. اگر $\mathcal{W}_k$ یک زیرفضایی ازفضای برداری $\mathbb{C}^n$ باشد، انگاه μ_k یک مقدار هارمونیک ریتز ماتریس A نسبت به $\mathcal{W}_k$ نامیده می شود اگر

$$A^{-1}w_k-\mu_kw_k\perp\mathcal{W}_k,\quad for\ some\ w_k\in\mathcal{W}_k,\ \ w_k\neq 0.$$

کیم-چن توی درپایان نامه دکتری خودریشه های چندجمله ایهای ایده ال ارنولدی و ایده ال مانده مینیمال راموردبررسی قراردادند.

قضیه ۳.۱. *[۷، قضیه۵,۱۰] فرض کنید A یک ماتریس نامنفردباشد و چندجمله ای ایده ال آرنولدی ماتریس A بصورت $p^{IA}(x)\in\pi_k(1)$ بیان شده باشد. آنگاه اگر $\|p^{IA}(A)\|>0$ آنگاه ریشه های چندجمله ای ایده ال آرنولدی $\|p^{IA}(x)\|$ مقادیرریتزماتریس A هستندکه درمحدوده بردعددی ماتریس A قراردارند*

$$\{x\ :\ p^{IA}(x)=0\}\subset\mathcal{F}(A).$$

قضیه ۴.۱. *[۷، قضیه۵,۱۱] فرض کنید چندجمله ای $k-$ ام ایده ال $GMRES$ ماتریس A باشدو $p^{IG}(x)\in\pi_k(0)$ یک ماتریس نامنفردباشد. آنگاه اگر $\|p^{IG}(A)\|>0$ آنگاه ریشه های چندجمله ای ایده ال آرنولدی $\|p^{IG}(x)\|$ مقادیرریتزماتریس A هستندکه درمحدوده بردعددی ماتریس A^{-1} قراردارند*

$$\{1/x\ :\ p^{IG}(x)=0\}\subset\mathcal{F}(A^{-1}).$$

۲ نتایج اصلی

در این بخش در صدد هستیم چندجمله ایهای $DGMRES$ و $DArnoldi$ را موردمطالعه قراردهیم. فرض کنید $A\in\mathbb{C}^{n\times n}$ یک ماتریس منفردبااندیس α باشد. ماتریس یکانی Q وجوددارد بطوریکه تجزیه ماتریس A براساس فضای پوچ $\mathcal{N}(A^\alpha)$ وبرد $\mathcal{R}(A^\alpha)$ بصورت زیر می باشد [۲].

$$A=Q\begin{bmatrix}B & *\\ 0 & N\end{bmatrix}Q^*, \qquad (۵)$$

که درآن ماتریس $B\in\mathbb{C}^{m\times m}$ یک ماتریس معکوس پذیراست و ماتریس $N\in\mathbb{C}^{n-m\times n-m}$ یک ماتریس پوچ توان بااندیس α می باشد. فرض کنید بردار x_0 یک حدس اولیه ازجواب دستگاه خطی $Ax=b$ داده شده باشد وبردار $r_0=b-Ax_0$ بردارمانده اولیه باشد. درروش $DGMRES$ بردارجواب تقریبی x_k^D در مرحله $k-$ ام بصورت زیر بیان می شود [۶].

$$x_k^D\in\ x_0+span\{A^\alpha r_0,\dots,A^{k-1}r_0\},\quad k=\alpha+1,\dots,n$$

بطوریکه نرم بردارمانده درمرحله k -ام ، $r_k^D=b-Ax_k^D$ درروش فوق بایددرمعادله زیر صدق کند.

$$\|A^\alpha r_k\|=\min_{p\in\pi_k^\alpha(0)}\left\|Q\begin{bmatrix}p(B) & *\\ 0 & I_{n-m}\end{bmatrix}Q^*A^\alpha r_0\right\|, \qquad (۶)$$

نتیجه می گیریم که k-امین مانده روش مینمال تعمیم یافته درازین بصورت

$$\varphi_k^D = \|A^\alpha r_k\| = \min_{p\in\pi_k^\alpha(\circ)}\|p(A)A^\alpha r_\circ\| \qquad (۷)$$

که درآن $\pi_k^\alpha(\circ) = \{p \in \pi_k(\circ) \ : \ p^{(i)}(\circ) = \circ, \ \ 1 \le i \le \alpha\}$، $p^D(z)$، k -امین چندجمله ای مینمال تعمیم یافته درازین نامیده می شود که سمت راست معادله فوق رامی نیم می کند. وجود و یکتایی این چند جمله ای در قضایای زیر بیان می شود.

با استفاده از قضیه زیر می توان نشان داد که k -امین چندجمله ای مینمال تعمیم یافته درازین وجود دارد.

قضیه ۱.۲. *[۴، قضیه ۲,۳] فرض کنید $B \in \mathbb{C}^{J\times J}$ ماتریس نامنفرد داده شده باشد، $m \ge \circ$ و $l \ge \circ$ اعدادصحیح باشند. اگرعبارت $\min_{p\in\Omega_{l,m}}\|p(B)\| > \circ$، انگاه مسئله دارای جواب یکتاکمینه است. بطوریکه*

$$\Omega_{l,m} := \{z^{m+1}q + 1 \ : \ q \in \pi_l\}.$$

فرض کنید $l = k - \alpha - 1 \ge \circ$ و $m = \alpha$ ، دراین صورت $\Omega_{l,m} = \pi_k^\alpha(\circ)$

چون $Q^*A^\alpha r_\circ = (\tilde{r}_\circ, \circ)^t$ لذا با استفاده از (۶) و (۷)

$$\|A^\alpha r_k\| = \min_{p\in\pi_k^\alpha(\circ)}\|p(B)\tilde{r}_\circ\| \qquad (۸)$$

قضیه ۲.۲. *فرض کنید $A \in \mathbb{C}^{n\times n}$ با اندیس $\alpha = ind(A) > \circ$ همانند رابطه (۵) و (۸) بیان شده باشد. اگرمعادله $\min_{p\in\pi_k^\alpha(\circ)}\|p(B)\| > \circ$ انگاه یک چندجمله ای یکتا $p^*(z) \in \pi_k^\alpha(\circ)$ وجوددارد بطوریکه*

$$\min_{p\in\pi_k^\alpha(\circ)}\|p(B)\| = \|p^*(B)\| \qquad (۹)$$

تعریف ۳.۲. بااستفاده ازقضیه ۲ .۲ ، اگر $\min_{p\in\pi_k^\alpha(\circ)}\|p(B)\| > \circ$ k -امین ایده ال مینمال تعمیم یافته درازین برای دستگاه منفرد $Ax = b$ به صورت زیرتعریف می شود.

$$\varphi_k^{ID} = \min_{p\in\pi_k^\alpha(\circ)}\max_{\|r\|=1, r\in\mathcal{R}(A^\alpha)}\|p(A)r\|. \qquad (۱۰)$$

تعریف ۴.۲. $-k$ امین مینمال تعمیم یافته درازین دربدترین حالت برای دستگاه منفرد $Ax = b$ به صورت زیرتعریف می شود.

$$\varphi_k^{WD} = \max_{\|r\|=1, r\in\mathcal{R}(A^\alpha)}\min_{p\in\pi_k^\alpha(\circ)}\|p(A)r\|. \qquad (۱۱)$$

گزاره ۵.۲. *فرض کنید φ_k^{WD}, ϕ_k^D و ϕ_k^{ID} ، به ترتیب مقادیرتعریف شده دررواط (۷) ، (۱۰) و (۱۱) باشند ، آنگاه*

$$\varphi_k^D \le \varphi_k^{WD} \le \varphi_k^{ID}.$$

بعبارت دیگرایده ال روش مانده می نیمال درازین کرانی برای بدترین حالت مانده می نیمال درازین می باشد.

درروش درازین آرنولدی $(DArnoldi)$ هدف پیداکردن تقریبی است که بتواند عبارت زیررابه می نیم برساند

$$\min_{p\in\pi_{k-1}}\|(A^k - p(A))A^\alpha r_\circ\| = \min_{p\in\pi_k(1)}\|p(A)A^\alpha r_\circ\|.$$

بااستفاده ازرابطه (۶)

$$\min_{p\in\pi_k(1)}\|p(A)A^\alpha r_\circ\| = \min_{p\in\pi_k(1)}\left\|Q\begin{bmatrix} p(B) & * \\ \circ & p(N)\end{bmatrix}Q^*A^\alpha r_\circ\right\|,$$

ازانجاکه $Q^*A^\alpha r_\circ = (\tilde{r}_\circ, \circ)^t$، رابطه زیربدست می آید

$$\begin{aligned}\|p^{DArnoldi}(A)A^\alpha r_\circ\| &= \min_{p\in\pi_k(1)}\|p(A)A^\alpha r_\circ\| \\ &= \min_{p\in\pi_k(1)}\|p(B)\tilde{r}_\circ\| \le \min_{p\in\pi_k(1)}\|p(B)\|\|\tilde{r}_\circ\|\end{aligned}$$

چون $\|A^\alpha r_\circ\| = \|\tilde{r}_\circ\|$ بنابراین

$$\min_{p\in\pi_k(1)}\frac{\|p(A)A^\alpha r_\circ\|}{\|A^\alpha r_\circ\|} \le \min_{p\in\pi_k(1)}\|p(B)\| = \|p^{IDA}(B)\|$$

چندجمله ای $p^{IDA}(z)$ چندجمله ای ایده ال درازین آرنولدی نامیده می شود. بنابراین چندجمله ای ایده ال درازین ماتریس A با چندجمله ای ایده ال آرنولدی برای بخش نامنفرد ماتریس A باهم برابرهستند.

مراجع

[1] R. Carden and M. Embree, *Ritz value localization for non-Hermitian matrices*, SIAM J. Matrix Anal. Appl. **33** (2012), pp. 1320-1338.

[2] A. Greenbaum, F. Kyanfar and A. Salemi *On the Convergence Rate of DGMRES*, Lin. Alg. Appl. **552** (2018), pp. 219-238.

[3] A. Greenbaum and L. N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems*, SIAM J. Sci. Comput. **15** (1994), pp. 359-368.

[4] J. Liesen and P. Tichý, *On Best approximation of polynomials in matrices in the matrix 2-norm*, SIAM J. Matrix Anal. Appl., **31-2** (2009) 853-863.

[5] Y. Saad and M.H. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. Statist. Comput., 1986; **7**: 856-869.

[6] A. Sidi, *DGMRES: A GMRES-type algorithm for Drazin-inverse solution of singular nonsymmetric linear systems*, Lin. Alg. Appl. **335** (2001), pp. 189-204.

[7] K. C. Toh, *Matrix approximation problems and nonsymmetric iterative methods*, PhD dissertation, Cornell University, Ithaca, NY, 1996.

پرتفوی مماس و بهینه درمدل مارکویتز

سید علی محمد محسنی الحسینی *

دانشگاه ولی عصر(عج) رفسنجان، دانشکده علوم ریاضی و کامپیوتر

چکیده

مدل مارکوویتز [۱] از جمله مدلهای ریاضی است که توانسته بهینه سازی پرتفوی را تحت تاثیر دهد. در اینجا نقش این مدل را درحالات حداقل واریانس پرتفوی، پرتفوی مماس، پرتفوی بهینه مورد بررسی قرار خواهیم داد. ضمنا با بکارگیری مدل مارکوویتز فرمول میانگین برای مرزهای کارآمد و نرخ بازدهی پرتفوی مورد انتظاررا بدست می آوریم.

واژه‌های کلیدی: کلید واژه : ضرائب لاگرانژ، مدل مارکوویتز، پرتفوی مماس، پرتفوی بهینه

کد موضوع‌بندی ریاضی [۲۰۱۰]: G11 ، C61

۱ مقدمه

در تئوری بهینه سازی سبد سهام، اندازه گیری ریسک انحراف استاندارد بسیار مشهور است. هری مارکوویتز در سال ۱۹۵۲مدل اساسی پرتفوی که مشهور به تئوری مدرن پرتفوی است را ارائه کردو [۵]. جایی که او یک روش بهینه سازی برای سرمایه گذاران ریسک گریز را توضیح داد. وی به دلیل اهمیت این موضوع در سال ۱۹۹۰ جایزه نوبل را از آن خود کرد. لذا از آن زمان به بعد تجزیه و تحلیل واریانس میانگین وی در بسیاری از مقالات استفاده می شود. ایده اصلی مارکوویتز یافتن بهترین بازده مورد انتظار و ریسک برای هر سرمایه گذار است. درسال ۱۹۵۹ شارپ و لیتنر [۲] زمینه را براي مدل قيمت گذاري داراييهاي سرمايهاي (CAPM)[۳] با استفاده از مدل هری مارکوویتز فراهم نمودند. التون، گروبر[۴] در سال ۱۹۸۱ نظریه مارکوویتز را مورد بررسی قرار داد [۴]. در سال ۲۰۰۱ براندت[۵] متغیر های انتخابی در پرتفویی را مورد بررسی قرار داد [۳]. رضا راعی و سعید فلاح پوردر سال ۱۳۹۰مدلی را برای مدیریت فعال پرتفوی با استفاده ازارزش در معرض ریسک والگوریتم ژنتیک طراحی نمودند [۱]. مریم کشتکاری در سال ۱۳۸۷ روش های انتخاب اوراق بهادر وتشکیل پر تفوی بهینه را مورد بررسی قرارداد [۲].

۲ مدل مارکوویتز

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{pmatrix} \; ; \; \bar{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \; ; \; \bar{\mathbf{r}} = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} \; ; \; \bar{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} \; ; \; \sum = \begin{pmatrix} \sigma_{11} & \dots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \dots & \sigma_{nn} \end{pmatrix}$$

*آدرس ایمیل سخنران : amah@vru.ac.ir;mohsenialhosseini@gmail.com

[۱] Harry Markowitz
[۲] Sharpe Linter
[۳] Capital Asset Pricing Model
[۴] Elton, Gruber
[۵] Brandt

T :ترانهاده، θ_i :مبلغ سرمایه گذاری شده در دارایی i ، $C_0 = \bar{1}^T\theta$:سرمایه اولیه ، C_{end} :،سرمایه نهایی r :نرخ بهره بانکی، μ : میانگین دارایی‌ها، $\mu_p = \mu^T\theta$:نرخ بازدهی پرتفوی مورد انتظار، $\sum$:ماتریس کواریانس از rاست، σ_{ij} :کواریانس بازدهی پرتفوی دارایی بین iو j.
در این مدل برای محاسبه مرز کارآمد، باید خطر (انحراف استاندارد) را با توجه به بازده مورد انتظار به حداقل برسانیم. در این مدل تابع هدف زیر را داریم:

$$min \qquad \theta^T \sum \theta$$

$$s.t. \qquad A^T\theta = B$$

که در آن $A = (\mu, \bar{1}), \mathbf{B} = \begin{pmatrix} \mu_p \\ C_0 \end{pmatrix}$.

$$V(R_p) = \theta^T \sum \theta = \theta^T \sum \sum^{-1} A\lambda = (A^T\theta)^T H^{-1} B = B^T H^{-1} B$$

با توجه به اینکه H یک ماتریس متقارن است، فرض می کنیم به صورت زیر باشد:

$$H = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow H^{-1} = \frac{1}{ac - b^2} \begin{pmatrix} c & -b \\ -b & a \end{pmatrix} \qquad (1)$$

چون $H = A^T \sum^{-1} A$ واضح است که:

$a = \mu^T \sum \mu$
$b = \mu^T \sum^{-1} \bar{1} = \bar{1}^T \sum^{-1} \mu$
$c = \bar{1}^{-1} \sum^{-1} \bar{1}$
$d = ac - b^2$

با توجه به رابطه (۱) فرمول واریانس به صورت زیر در می آید:

$$V(R_p) = \frac{1}{d}(c\mu_p^2 - 2b\mu_p C_0 + ac_0^2$$

این فرمول به ما میانگینی برای مرزهای کارآمد در یک چارچوب ریسک خنثی را می‌دهد. به عبارتی فرمول مرزکارآمد به صورت زیر است:

$$\sigma_p^2 = \frac{1}{d}(c\mu_p^2 - 2b\mu_p C_0 + ac_0^2 \qquad (2)$$

با جذر گرفتن از رابطه (۲) انحراف استاندارد بدست میآید. بعد از انجام محاسبات لازم بر روی رابطه (۲) می توان به فرمول نرخ بازدهی پرتفوی مورد انتظار دست یافت.

$$\mu_p = \frac{b}{c} C_0 \pm \sqrt{\frac{d}{c}} \sigma_p$$

تخصیص پرتفویی θ_{EF}که از رابطه زیر بدست می آید متعلق به مرزهای کارآمدی است که سرمایه گذار باید در دارایی های منفرد سرمایه گذاری کند تا بازده و ریسک در مرز کارآمدرا بدست آورد.

$$\theta_{EF} = \sum^{-1} A\lambda = \frac{1}{d}((a\bar{1} - b\mu)C_0 - (c\mu - b\bar{1})\mu_p) \qquad (3)$$

بنابراین برای هر مقدار دلخواه بازده پرتفوی می توان حداقل انحراف استاندارد و تخصیص مربوطه را به ترتیب با استفاده از (۲) و (۳) محاسبه کرد.

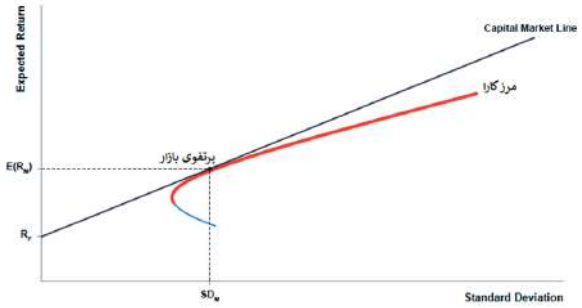
۳ پرتفوی مماس

فرض کنید یک سرمایه گذار ترجیحات دیگری غیر از کمترین میزان خطر پذیری و در نتیجه سرمایه گذاری در پرتفوی با حداقل واریانس داشته باشد.

نمونه ای از اولویت های دیگر سرمایه گذاری در پرتفوی با حداکثر نسبت شارپ (نسبت میانگین به انحراف استاندارد) است. نسبت شارپ به عنوان نسبت ریسک بازگشت تعریف می شود. این بازده مورد انتظار برای هر واحد ریسک را نشان می دهد، بنابراین پرتفوی با حداکثر نسبت شارپ بالاترین بازده مورد انتظار برای هر واحد ریسک را دارند و بنابراین کارآمدترین پرتفوی هستند.

از لحاظ گرافیکی، پرتفوی مماس نسبت به فضای حاصل از میانگین و انحراف استاندارد ، نقطه تقاطع خطی است که از مبدا شروع و با منحنی مرز کارآ مماس می شود، زیرا این نقطه از ویژگی مهمی برخوردار است که دارای بالاترین میانگین ممکن نسبت به انحراف استاندارد می باشد.

اما از نظر بازار سرمایه نسبت شارپ درنمودار (۱)نشان داده شده است. خط بازار سرمایه، تمام ترکیب های محتمل پرتفوی که شامل نسبتهای متفاوت بین دارایی بدون ریسک و پرتفوی بازار است را نشان می دهد. از سوی دیگر، مرز کارا نشاندهنده تمام ترکیب های محتمل پرتفوی های کارا است. اما این پرتفوی ها فقط شامل نسبتهای مختلف از دارایی های ریسکی می باشند و نه دارایی بدون ریسک.



۱. نمودار پرتفوی مماس(بازار)

نقطه تقاطع خط بازار سرمایه با منحنی مرز کارا، پرتفوی بازار یا پرتفوی مماس نامیده می شود. اگر یک سرمایه گذار، منطقی و ریسک گریز باشد تنها زمانی ریسک بالاتر را می پذیرد که بازدهی متناسب با آن افزایش یابد. از این نقطه نظر، پرتفوی مماس بعنوان کاراترین پرتفوی می باشد.

برای محاسبه پرتفوی مماس ما به فرمولی برای مرز کارآمد نیاز داریم. با جذر گرفتن انحراف استاندارد رابطه (۲) داریم:

$$\sigma_p = \sqrt{\frac{1}{d}(c\mu_p^2 - 2b\mu_p C_0 + aC_0^2} \qquad (۴)$$

$$\frac{\Delta\sigma_p}{\Delta\mu_p} = \frac{\sqrt{\frac{1}{d}(c\mu_p^2 - 2b\mu_p C_0 + aC_0^2} - 0}{\mu_{tg} - 0}$$

شیب مرز کارآ در نقطه مماس از مشتق مرز کارآ در آن نقطه بدست میآید:

$$\frac{\partial\sigma_p}{\partial\mu_p} = \frac{1}{2}(\frac{1}{d}(c\mu_p^2 - 2b\mu_p C_0 + aC_0^2))^{-\frac{1}{2}}\frac{1}{d}(2c\mu_p - 2bC_0)|_{\mu_p=\mu_{tg}} = \frac{c\mu_p - bC_0}{d\sqrt{\frac{1}{d}(c\mu_p^2 - 2b\mu_p C_0 + aC_0^2)}}$$

با توجه به اینکه در نقطه مماس دو شیب برابراند، پس داریم:

$$\frac{\sqrt{\frac{1}{d}(c\mu_p^2 - 2b\mu_p C_0 + aC_0^2)}}{\mu_{tg}} = \frac{c\mu_p - bC_0}{d\sqrt{\frac{1}{d}(c\mu_p^2 - 2b\mu_p C_0 + aC_0^2)}} \Rightarrow \mu_{tg} = \frac{a}{b}C_0$$

با جانشین کردن μ_{tg}در فرمول مرز کارا فرمول انحراف استاندارد σ_{tg} متناظر با آن به صورت زیر بدست می آید:

$$\sigma_{tg} = \sqrt{\frac{1}{d}(c\frac{a^2}{c^2C_\circ^2} - \frac{2ab}{b}C_\circ^2 + aC_\circ^2)} = \sqrt{\frac{a}{b}}C_\circ$$

بطوریکه $d = ac - b^2$می باشد. تخصیص دارایی ها θ در مرحله مماس، با استفاده از فرمول (۳) به صورت زیر بدست می آید.

$$\theta_{tg} = \frac{c\frac{a}{b}C_\circ - bC_\circ}{d}\sum{}^{-1}\mu + \frac{aC_\circ - b\frac{a}{b}C_\circ}{d}\sum{}^{-1}\bar{1} = \sum{}^{-1}\mu\frac{C_\circ}{b}$$

بنابراین

$$\theta_{tg} = \sum{}^{-1}\mu\frac{C_\circ}{b} \qquad (۵)$$

۴ پرتفوی بهینه

تاکنون، شاهد دو اوراق بهادار هستیم که یک سرمایه گذار می تواند ترجیح دهد. اگر او خواهان حداقل میزان خطر باشد، حداقل پرتفوی واریانس را به عهده می گیرد. اگر هدف به حداکثر رساندن نسبت شارپ پرتفوی باشد، پرتفوی مماس در نظر گرفته میشود. با این حال، نظریه مارکوویتز نوع دیگری از اولویت را برای سرمایه گذار فرض میکند. در این مدل هدف سرمایه گذاران به حداکثر رساندن سود است، بطوریکه از رابطه زیر بدست میآید:

$$u = E(C_{end}) - \frac{1}{2}\gamma var(C_{end}) \qquad (۶)$$

$$u = E(C_{end}) - \frac{1}{2}(C_{end}) = E(E_\circ + R_p) - \frac{1}{2}\gamma var(E_\circ + R_p) = C_\circ + \mu_p - \frac{1}{2}\gamma var(R_p) \qquad (۷)$$

$$= C_\circ + \mu^T\theta - \frac{1}{2}\gamma\sigma_p^2 = C_\circ + \mu^T\theta - \frac{1}{2}\gamma\theta^T\sum\theta$$

از لحاظ گرافیکی، پرتفوی با حداکثر سودآوری با حرکت دادن منحنی سودآوری در بالاترین حد ممکن به دست می آید. منحنی سودآوری منحنی است که ترکیبات احتمالی میانگین و انحراف معیار را نشان میدهد که نتیجه آن سودمندی یکسان است. بنا به رابطه (۷)، $\mu_p = u - C_\circ + \frac{1}{2}\gamma\sigma_p^2$ در فضای حاصل از میانگین و انحراف استاندارد یک سهمی است.
برای محاسبه پرتفوی بهینه، ما باید حداکثر سودمندی را با محدودیت بودجه به حداکثر برسانیم. در این مدل هدف:

$$max \qquad C_\circ + \mu^T\theta + \frac{1}{2}\gamma\sigma_p^2$$

$$s.t \qquad \bar{1}^T\theta = C_\circ$$

حال با استفاده از روش ضرایب لاگرانژ، مقدار بهینه تخصیص دارایی ها به صورت زیر بدست می آید:

$$\theta_{opt} = \frac{\sum^{-1}\mu}{\gamma} + \frac{\sum^{-1}\bar{1}}{\gamma}(\frac{\gamma C_\circ - b}{c}) = \frac{1}{\gamma}\sum^{-1}+(\mu + \bar{1}(\frac{\circ - b}{c}))$$

اگر سرمایه گذار بخواهد سودمندی خود را به حداکثر برساند، وی بایست مبالغی را در هر دارایی سرمایه گذاری کند. ما می توانیم این عبارت را با استفاده از حداقل واریانس پرتفوی و پرتفوی مماس ساده کنیم. بابکارگیری آنها داریم:

$$\theta_{mv} = \sum{}^{-1}\bar{1}\frac{C_\circ}{b}, \qquad \theta_{tg} = \sum{}^{-1}\mu\frac{C_\circ}{b}$$

$$\theta_{opt} = \frac{b}{C_\circ\gamma}\theta_{tg} + \frac{c}{C_\circ}(\frac{C_\circ - \frac{b}{\gamma}}{c})\theta_{mv} = \frac{b}{C_\circ\gamma}\theta_{tg} + (1 + \frac{b}{C_\circ\gamma})\theta_{mv} \qquad (8)$$

ما می بینیم که پرتفوی بهینه ترکیبی از پرتفوی حداقل واریانس و پرتفوی مماس است. مقادیر مربوط بهμ_p و σ_p^2برابر است با:

$$\mu_{opt} = \frac{d}{c\gamma} + \mu_{mv} \qquad , \qquad \sigma_{opt}^2 = \frac{d}{c\gamma^2} + \sigma_{mv}^2$$

بنابراین میانگین و واریانس پرتفوی بهینه با مقادیر حداقل واریانس پرتفوی به علاوه مبلغی که به ضریب ریسک پذیری (γ)بستگی دارد، تعیین می شود.
هنگامی که یک سرمایه گذار از ریسک مطلق متنفر است، بنابراین نمی خواهد ریسکی را بپذیرد، (γ)به بی نهایت خواهد رفت و پرتفوی بهینه حداقل واریانس پرتفوی خواهد بود. بنابراین یک سرمایه گذار با پارامتر نامحدود ریسک گریزی با حداقل واریانس پرتفوی سرمایه گذاری خواهد کرد.

مراجع

[۱] ر. راعی، و س. فلاح پور، ،طراحی مدلی برای مدیریت فعال پرتفوی با استفاده ازارزش در معرض ریسک، ۱۳۹۰.

[۲] م. کشتکاری، راهبرهای مدیریت پرتفوی وروش های انتخاب اوراق بهادار و تشکیل پرتفوی بهینه، ۱۳۸۷.

[3] Y. Ait-Sahali, M. W. Brandt, *Variable Selection for Portfolio Choice*, Journal of Finance, **56** (2001) 1297–1351.

[4] Elton. Gruber, *Modern Portfolio Theory and Investment Analysis*, John Wiley Sons Inc, 1981.

[5] H. Markowitz, *Portfolio Selection*, Journal of Finance,**7** (1952) 77–91.

[6] H. Pham, *Continuous-time stochastic control and optimization with financial applications*, springer, 2009.

پوسترها

تحلیل پوششی داده ها DEA [۱] و کاربرد آن در عملکردموسسات بانکی با رویکرد تحلیل مولفه اصلی PCA [۲]

محمدرضا پارساصفت [۱] * و مهدی زعفرانیه[۲،۱]†

[۱]دانشکده ریاضی وعلوم کامپیوتر، دانشگاه حکیم سبزواری،سبزوار، ایران

چکیده

بانکها ازجمله مهمترین شرکتهای مالی در هر کشوری هستند که به دلیل بین المللی شدن و آزادسازی بانکداری جهانی، به سرعت طی دو دهه گذشته گسترش یافته اند. برای مقابله با این محیط رقابتی، بسیاری از مسئولان بانکها و پژوهشگران دانشگاهی برای یافتن راههایی جهت بهبود عملکرد بانکها تلاش کرده اند .از جمله روش های مناسب جهت ارزیابی عملکرد بانک ها استفاده از روش تحلیل پوششی داده ها می باشد.در این پژوهش ۱۸ شعبه بانک شهر در منطقه شرق کشور مورد بررسی قرار گرفته است .به علت آنکه افزایش تعداد متغییر های ورودی و خروجی باعث کاهش کیفیت ارزیابی عملکرد شعب می گردد، از روش PCA برای کاهش تعداد متغییر های ورودی و خروجی استفاده گردیده است.سپس داده ها توسط مدل های تحلیل پوششی داده (AP)،(BCC)،(CCR) و دو مرحله ایی مورد ارزیابی قرار گرفته است .

واژه‌های کلیدی: تحلیل پوششی داده ها، تحلیل مولفه اصلی،کارآمدی شعب بانک ها

کد موضوع‌بندی ریاضی [۲۰۱۰]: 90Bxx، 90C08، 62H25

۱ مقدمه وتاریخچه

در اقتصادهای بانک محور، نظام بانکی مسئولیت بسیار سنگینی بر عهده دارد و در واقع یکی ازمهمترین اجزای اقتصادی کشور، نظام بانکی است. بخش بانکی از طریق تبدیل سپرده به سرمایه گذاری مولد نقش مهمی را به عنوان واسطه مالی انجام می دهد. برخلاف کشورهای توسعه یافته که بازارهای مالی و بخش بانکی به صورت هماهنگ در انتقال منابع مالی همکاری می کنند، در کشورهای درحال توسعه، بازارهای مالی به صورت ضعیف عمل می کنند و اغلب به صورت کامل و کارا عمل نمی کنند . تاریخچه بانکداری به قرنها قبل از میلاد مسیح برمیگردد. آغازگر حرفه بانکداری در جهان صرافانی بودند که با تعیین عیار فلزات قیمتی موجب سهولت مبادله آنها با کالاها شده و با جلب اعتماد مردم و صدور اسناد تعهد توانستند امانتدار اموال تجاری شوند که تداوم تجارت مردم به یاری و حمایت همان صرافان امکانپذیر شد.

سالها است که محققان از روش تحلیل مولفه هاي اساسي در علوم مختلف استفاده نموده اند. اولين كاربردهاي روش مذكور در علوم روانشناسي بود كه به تدريج به ساير حوزهها از علوم طبيعي و پزشكي تا علوم اقتصادي و اجتماعي راه يافت. از اين رهيافت در حوزههاي اقتصادي استفادههاي فراواني شده است.در سال ۱۹۹۸ كينگ بانايان [۳] و همكاران مولفه هاي اساسي استقلال بانک مركزي را از اين روش استخراج نمودند.جيمز [۴] و وارويک [۵] از دانشگاه ملي استراليا در سال ۲۰۰۵ تاثير آزادسازي مالي را بر رشد اقتصادي در كشور مالزي بررسي كردند. ايشان براي بررسي توسعه مالي شاخص عمق مالي را با استفاده از روش PCA استخراج

[۱]Data Envelopment Analysis
[۲]Principal component analysis

*سخنران. آدرس ایمیل: mparsasefat3@gmail.com
†سخنران. آدرس ایمیل: m.zaferanieh@hsu.ac.ir

[۳]King Banaian
[۴]James B. Ang
[۵]Warwick J.Mckibbin

نمودند. يكي از انتشارات دانشگاه آكسفورد در سال ۲۰۰۶ مطالعه اي بود كه توسط يک تيم پژوهشي براي ساختن شاخص وضعيت اقتصادي و اجتماعي انجام شده بود . آنها براي ساختن اين شاخص از اطلاعات و روش PCA استفاده نمودند. در سال ۲۰۰۶ در دانشكده مهندسي رياضي و دانشكده فني هلسينكي DHS [۱] طرحي مطالعاتي براي ساختار زماني نرخ بهره با استفاده از روش PCA انجام شد.

۲ پارامترها و شاخص‌های مدل

داده‌های این پژوهش از طریق منابع اطلاعاتی بانک شهر بدست آمده است و محقق طی بررسی مقالات ومراجع مرتبط به موضوع ، سه ورودی و چهارخروجی را انتخاب نمود. سپس برای هر شاخص(ورودی و خروجی) چند زیر شاخص انتخاب گردید وبا استفاده از روش PCA و کاهش ابعاد مقداری مناسب برای هر شاخص انتخاب گردید.

۱.۲ روش(PCA)برای انتخاب ورودی و خروجی مطلوب

روش PCA ،تکنیکی برای کاهش تعداد زیاد متغییر های وابسته می باشد. در این روش یک مجموعه بزرگ از متغییر ها تبدیل به مجموعه ایی کوچکتر می شود ، به طوری که شامل اطلاعاتی مفید از مجموعه اولیه نیز باشد. هدف کم کردن تعداد متغییر ها به m مولفه مهمتر می باشد.هر مولفه از ترکیبی خطی k متغییر تخمین می خورد. پس داریم:

$$F_i = W_{i,1}X_1 + W_{i,2}X_2 + ... + W_{i,k}X_k$$

روش کار به این صورت می باشد که :

۱. ابتدا در روش PCA دو زیر شاخص را به عنوان ورودی انتخاب می کنیم.مثلا زیر شاخص X وY برای متغییر ورودی .

۲. مقدار میانگین هر زیر شاخص X وY را محاسبه می کنیم.

۳. ماتریس W=[X،Y] متشکل از دو زیر شاخص XوY می باشد . سپس کواریانس ماتریس راW حساب می کنیم.

$$C = COV(W) = \begin{bmatrix} COV(X,X) & COV(X,Y) \\ COV(Y,X) & COV(Y,Y) \end{bmatrix} \quad (۱)$$

۴. مقادیر ویژه وبردار های ویژه ماتریس W را حساب می کنیم:

$$|S - \lambda(I)| = EigenValue[\lambda_1, \lambda_2] \quad (۲)$$

حال می دانیم کهλ_1 و λ_2 بزرگترین مقادیر ویژه ماتریس W می باشند.سپس سعی می کنیم دو بعد ماتریس w را به تنها یک بعد کاهش دهیم . می توان این کاهش بعد را برای چندین زیر شاخص از شاخص اصلی نیز بسط داد.در این جا به علت انجام محاسبات زیاد روش PCA برای شاخص های دو بعدی بیان شده است

۵. بردار های ویژه متناظر ماتریس W را حساب می کنیم:

$$(S - \lambda(I))U = (S - \lambda(I)) \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \quad (۳)$$

۶. حال به محاسبه مقادیر جدید که از روش (PCA) بدست می آید می پردازیم:

$$P(ij) = U_1^T \begin{bmatrix} X_i - \bar{X} \\ Y_j - \bar{Y} \end{bmatrix} \quad (۴)$$

با روش PCA برای همه متغییر های ورودی و خروجی مقادیر جدید را با کاهش ابعاد بدست می آوریم . سپس با توجه به مقادیر بدست آمده برای شاخص های اصلی به بررسی مدل های تحلیل پوششی داده می پردازیم.

[۱] Demographic Health Survey

متغیر های خروجی		متغیر های ورودی	
شاخص اصلی	زیر شاخص	شاخص اصلی	زیر شاخص
منابع	۱-منابع ارزان قیمت	تعداد کارکنان	۱-تعداد کارکنان کاردانی و پایین تر
			۲-تعداد کارکنان کارشناسی و بالاتر
	۲-منابع گران قیمت	درآمد تسهیلات	۱- درآمد حاصل از عقود غیرمبادله ای
			۲-درآمد حاصل از عقود مبادله ای
			۳-درآمد حاصل مشارکت حقوقی
			۴-درآمد حاصل سرمایه گذاری مستقیم
			۵- درآمد حاصل عقود قرض الحسنه
		سود پرداختی	۱-سود سپرده کوتاه مدت
			۲-سود سپرده بلند مدت
تسهیلات	۱-عقود مبادله ای		
	۲-عقود مشارکتی		
	۳-عقود قرض الحسنه		
ضمانت نامه	۱-ضمانت نامه شرکت در مناقصه یا مزایده		
	۲-ضمانت نامه حسن انجام کار وتعهدات و پیش پرداخت وکسور وجه الضمان		
	۳-ضمانت نامه گمرکی و پیمان		
مطالبات	۱-مطالبات سررسید گذشته		
	۲-مطالبات معوق		
	۳-مطالبات مشکوک الوصول		

شکل ۱: جدول زیر شاخص های استفاده شده در روش PCA

۲.۲ متغییر های ورودی ومتغییر های خروجی

به این نکته اساسی توجه کرد که هدف اصلی بانک را می توان در سه حوزه تجهیز منابع، تخصیص منابع و خدمات خلاصه کرد. متغییر های ورودی برای مدل های مورد استفاده ،تعداد کارکنان،درآمد تسهیلات و سود پرداختی می باشند. متغییر های خروجی برای مدل های مورد استفاده ،منابع بانک ،تسهیلات ،مطالبات وضمانت نامه می باشند.

۳ تحلیل پوششی داده ها

فارل در سال ۱۹۵۷ اولین بار با استفاده از روش غیرپارامتری اقدام به اندازه‌گیری کارایی یک واحد تولیدی نمود. در ادامه محققانی نظیر چارلز و همکارانش دیدگاه فارل را توسعه دادند را ارائه نمودند که تحت عنوان تحلیل پوششی داده ها نام گرفت. در ادامه به انواع مختلف مدل های خروجی محور تحلیل پوششی داده می پردازیم.

۱.۳ مدل خروجی محور CCR:

$$\max \sum_{r=۱}^{s} u_r y_{ro} \qquad (۵)$$

s.t.

$$\sum_{i=1}^{m} v_i x_{io} = 1 \tag{۶}$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0 \qquad j = 1, ..., n \tag{۷}$$

$$u_r, v_i \geq 0 \tag{۸}$$

۲.۳ مدل خروجی محور BCC:

$$\max \sum_{r=1}^{s} u_r y_{ro} - u_0 \tag{۹}$$

s.t.

$$\sum_{i=1}^{m} v_i x_{io} = 1 \tag{۱۰}$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} - u_0 \leq 0 \qquad j = 1, ..., n \tag{۱۱}$$

$$u_r, v_i \geq 0 \tag{۱۲}$$

۳.۳ مدل خروجی محوراندرسون- پیترسون

$$h_k = \max \sum_{r=1}^{s} u_r y_{rk} \tag{۱۳}$$

s.t.

$$\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \geq 0 \qquad for \qquad j = 1, ..., n \qquad j \neq k \tag{۱۴}$$

$$\sum_{i=1}^{m} v_i x_{ik} = 1 \tag{۱۵}$$

$$u_r \geq 0 \qquad for \qquad r = 1, ..., s \tag{۱۶}$$

$$v_i \geq 0 \qquad for \qquad i = 1, ..., m \tag{۱۷}$$

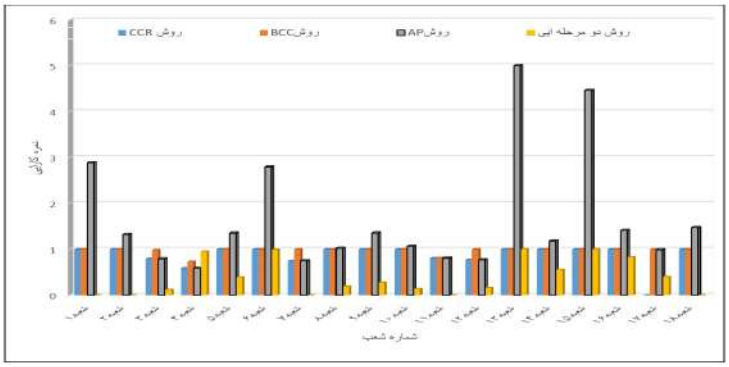
۴.۳ مدل خروجی محور دو مرحله ایی

$$\max \sum_{r=1}^{s} u_r y_{rp} \tag{۱۸}$$

s.t.

$$\sum_{i=1}^{m} v_i x_{ip} = 1 \tag{۱۹}$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0 \tag{۲۰}$$



شکل ۲: نمودار مقایسه ایی مدل های چهارگانه تحلیل چوششی داده ها

$$\sum_{b=1}^{o} w_b z_{bj} - \sum_{i=1}^{m} v_i x_{ij} \leq 0 \tag{۲۱}$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{b=1}^{o} w_b z_{bj} \leq 0 \tag{۲۲}$$

$$u_r, v_i, w_b \geq 0 \qquad j = 1, ..., n \tag{۲۳}$$

۴ جداول و نمودار

در نمودار بالا برای آنکه مقایسه میزان کارایی شعب ملموس تر باشد ،مقدار کارایی شعبه شماره ۱۳ به جای عدد ۶۳۲۴عدد ۵ در نظر گرفته شده است .

۵ نتیجه‌گیری

با بررسی نمودار مقایسه ایی بالا نتایج زیر حاصل می گردد:

۱. استفاده از روشPCAدر راستی آزمایی کارایی شعب بسیار موثر می باشد.بدین صورت که استفاده از این روش باعث می شود که شعبی که کارایی حداکثری را دارا نمی باشند ولی جزء شعب کارا محسوب می شوند، از شعب کارا تفکیک گردند.

۲. شعب شماره(۱-۲-۵-۶-۸-۹-۱۰-۱۳-۱۴-۱۵-۱۶-۱۸) در دو روش CCR و BCC دارای نمره کارایی حداکثر برابر یک را دارا می باشند.لذا این ۱۲شعبه به عنوان شعب ورودی ، کاندیدای حضور در روش اندرسون- پترسون معرفی گردیدند.

۳. با بررسی ۱۲ شعبه بررسی شده با مدل اندرسون -پترسون شعبه شماره ۱۳ از میان سایر شعب بیشترین نمره کارایی را اخذ نمود .

۴. به منظور اطمینان از نتایج روش های BCC و CCR و AP مدل دو مرحله ایی نیز بررسی گردید.که در این مرحله شعبه شماره ۱۳ و۱۵ حداکثر نمره کارایی را اخذ نمودند.

۵. با جمع بندی نتایج روش های چهارگانه شعبه شماره ۱۳ بهترین عملکرد شعب را در بین ۱۸ شعبه شرق کشور اخذ نموده است . با بررسی عینی شعبه شماره ۱۳ به این مهم دست یافتیم که شعبه مذکور بهترین عملکرد را در منطقه اخذ نموده .ارتقا شعبه در سال گذشته در منطقه نیز به وضوح به درستی بررسی نتایج ما اذعان دارد.

جدول ۱: در جدول زیر نمره کارایی شعب در چهار روش بررسی شده نشان داده شده است.

شعبه	BCC	CCR	AP	دو مرحله ایی
۱	۱	۱	۲/۸۸۷۵۵۵	۰/۴۵۹۰۸۱۹
۲	۱	۱	۱/۳۲۷۸۷۵	۰/۶۹۵۸۴۹۱
۳	۰/۹۸۵۷۹۱۲	۰/۷۹۲۹۹۹۳	۰/۷۹۲۹۹۹۳	۰/۱۱۶۰۹۶۵
۴	۰/۷۳۳۵۹۳۳۸	۰/۶۱۰۴۱۱۷	۰/۵۶۵۷۲۷۸	۰/۶۳۷۵۸
۵	۱	۱	۱/۳۵۴۵۰۶	۰/۳۹۰۲۷۶۱
۶	۱	۱	۲/۷۹۵۵۲۹	۰/۹۹۱۴۴۱۱
۷	۱	۰/۷۵۰۷۰۹۳	۰/۷۵۰۷۰۹۳	۰/۲۶۰۶۱۱
۸	۱	۱	۱,۰۲۲۶۴۰	۰/۱۹۱۷۱۹۸
۹	۱	۱	۱/۳۵۶۷۶	۰/۲۷۰۶۹۸۸
۱۰	۱	۱	۱/۰۶۸۵۵۸	۰/۱۳۴۱۰۳۳
۱۱	۰/۸۱۰۳۴۴	۰/۸۰۷۵۷۵۹	۰/۸۰۷۵۷۵۹	۰,۴۸۹۰۵۹۷
۱۲	۱	۰/۷۷۳۰۲۸۳	۰/۷۷۳۰۲۸۳	۰/۱۵۷۱۰۶۹
۱۳	۱	۱	۶۳۲۴	۱
۱۴	۱	۱	۱/۱۸۱۰۳۸	۰/۵۵۵۶۰۶۵
۱۵	۱	۱	۴/۴۶۱۶۳۹	۱
۱۶	۱	۱	۱,۴۱۳۰۱۲	۰/۸۲۵۴۶۷۸
۱۷	۱	۰/۹۹۰۴۵۶۷	۰/۸۹۱۷۹۰۲	۰/۴۰۳۰۸۰۸
۱۸	۱	۱	۱/۴۷۶۶۷۹	۰/۴۴۶۹۱۲۷

مراجع

[۱] مهرگان,محمدرضا, *تحلیل پوششی داده ها مدل های کمی در ارزیابی عملکرد سازمان ها*, نشر کتاب دانشگاهی،تهران، ۱۳۹۱.

[۲] آذر,عادل, نوبهار , عماد, *ارزیابی عملکرد شعب بانک با استفاده از رویکرد ترکیبی تحلیل مولفه اصلی و تحلیل پوششی داده ها PCA-DEA*, پژوهش های مدیریت منابع انسانی, دوره ۵، شماره ۳، پاییز ۹۴.

[۳] ویلیام کوپر، لورنس سیفورد.کوراتن. *تحلیل پوششی داده ها مدل ها وکاربرد ها*. ترجمه‌ی دکتر سید علی میر حسنی، نشر دانشگاه صنعتی امیر کبیر،تهران، تابستان ۹۶.

[4] L.Seiford, M.ZhuJ , Profitability and marketability of the top 55 US commercial banks. Management science. 2015;45(9):1270-1288 .

[5] Abdi H, Williams LJ.، Principal component analysis. Wiley interdisciplinary reviews: computational statistics.، . 2010 Jul. ;2(4):433-59. ,

[6] Mehrizi, Seyyed Mohammad Tabatabaei, and Mohsen Shafiei , Ranking of Faculties by Double Frontiers DEA and Traditional DEA (AP) Methods. Management science. 2015 .

یادگیری عمیق: چالش ها، کاربردها و فرصت ها

شکوفه خوش‌نظر *

بخش علوم کامپیوتر، دانشگاه ولایت، ایرانشهر، ایران

چکیده

یادگیری عمیق، شاخه‌ای از یادگیری ماشین است که هدف آن نزدیک‌تر شدن به هوش مصنوعی است. این مقاله عمدتاً روش‌های یادگیری عمیق را شرح می‌دهد. اولاً، توسعه جهانی و وضعیت فعلی یادگیری عمیق را معرفی می‌کند. ثانیاً، اصل ساختاری، ویژگی‌ها و انواع مدل‌های کلاسیک یادگیری عمیق، مانند رمزگذار خودکار پشته‌ای، شبکه باور عمیق، ماشین بولتزمن عمیق را توصیف می‌کند. ثالثاً، آخرین پیشرفت‌ها و کاربردهای یادگیری عمیق را در بسیاری از زمینه‌ها مانند پردازش گفتار، بینایی ماشین، پردازش زبان طبیعی و کاربردهای پزشکی ارائه می‌دهد. در نهایت، چالش‌ها، مسائل و جهت‌گیری‌های تحقیقاتی آینده یادگیری عمیق را مطرح می‌کند.

واژه‌های کلیدی: یادگیری عمیق، رمزگذار خودکار پشته‌ای، شبکه‌های باور عمیق، ماشین بولتزمن عمیق، کاربردها و فرصت‌ها.
کد موضوع‌بندی ریاضی [۲۰۱۰]: 68T01، 68T30، 68T45

۱ مقدمه

یادگیری عمیق چیزی نیست جز طبقه‌بندی‌کننده‌های زیادی که با هم کار می‌کنند و براساس رگرسیون خطی و به دنبال آن برخی از توابع فعال‌سازی هستند. اساس آن همان روش رگرسیون خطی آماری سنتی $W^T X + b$ است. تنها تفاوت این است که گره‌های عصبی زیادی در یادگیری عمیق به جای تنها یک گره وجود دارد که در یادگیری آماری سنتی، رگرسیون خطی نامیده می‌شود. این گره‌های عصبی به عنوان یک شبکه عصبی نیز شناخته می‌شوند و یک گره طبقه‌بندی کننده به عنوان واحد عصبی یا ادراک شناخته می‌شود. نکته متضاد دیگری که باید به آن توجه کرد این است که در یادگیری عمیق لایه‌های زیادی بین ورودی و خروجی وجود دارد. یک لایه می‌تواند صدها یا حتی هزاران واحد عصبی داشته باشد. لایه‌هایی که بین ورودی و خروجی قرار دارند به عنوان لایه‌های پنهان و گره‌ها به عنوان گره‌های پنهان شناخته می‌شوند. اشکال طبقه‌بندی‌کننده‌های یادگیری ماشین سنتی این است که خودمان باید یک فرضیه پیچیده بنویسیم، در حالی که در شبکه عصبی عمیق توسط خود شبکه تولید می‌شود که آن را به ابزاری قدرتمند برای یادگیری موثر روابط غیرخطی تبدیل می‌کند.

یادگیری ماشینی را می توان به دو فرآیند توسعه تقسیم کرد: یادگیری سطحی و یادگیری عمیق. در سال ۲۰۰۶، قبل از اینکه یادگیری عمیق در پژوهش‌ها پرطرفدار شود، عمدتاً بر ساختار یادگیری کم عمق برای پردازش داده‌ها متمرکز بودند. رایج ترین ساختارهای کم عمق عبارتند از رگرسیون لجستیک، ماشین های بردار پشتیبان، مدل های مخلوط گاوسی و غیره. تا کنون، یادگیری سطحی تنها می‌تواند به سرعت و کارآمد مسائل را با محدودیت‌های متعدد حل کند، اما نمی تواند مسائل پیچیده در دنیای واقعی، مانند صدای انسان، تصاویر طبیعی، صحنه های بصری و غیره را حل کند. یادگیری سطحی دارای محدودیت‌هایی است به طوری‌که هرگز نمی‌توان مانند مغز انسان برای کسب اطلاعات با آن رفتار کرد.

مطالعه یادگیری عمیق عمدتاً در برگزاری کنفرانس‌های مختلف هوش مصنوعی در سطح جهانی، تأسیس گروه تحقیقاتی نخبگان جهانی، تأسیس تیم تحقیقاتی سازمانی و کاربردهای مستمر یادگیری عمیق در هوش مصنوعی تجسم می‌یابد. الگوریتم‌های یادگیری عمیق به طور مداوم پیشنهاد می‌شوند و رکوردهای جدید به طور مداوم در بسیاری از مجموعه های داده ایجاد می‌شوند. به عنوان مثال،

* سخنران. آدرس ایمیل: Sh.khoshnazar@velayat.ac.ir

در فرآیند آزمایش طبقه‌بندی تصویر برای ۱۰۰۰ نوع تصویر، در مدت پنج سال، از طریق بهبود مستمر مدل یادگیری عمیق، میزان خطای طبقه بندی تصویر به ۳/۵% کاهش یافت که بالاتر از دقت افراد عادی است. در واقع، این موفقیت در استفاده از یادگیری عمیق برای فعال کردن ماشین‌ها برای یادگیری نحوه شناسایی و دسته‌بندی موفق تصاویر بود. توسعه علم و فناوری به طور مداوم شناخت انسان را تازه می‌کند و مدل یادگیری عمیق به عنوان مدل فناوری هسته‌ای هوش مصنوعی در محیط کلان داده به طور مداوم به روز می‌شود و منعکس کننده آخرین پیشرفت‌های تحقیقاتی علم و فناوری فعلی است.

۲ مدل‌های یادگیری عمیق

در این بخش، ما اصول مدل‌های مبتنی بر یادگیری عمیق را بررسی می‌کنیم و معماری و ویژگی‌های آن‌ها را مورد بحث قرار می‌دهیم. در حال حاضر، یادگیری عمیق عمدتا شامل رمزگذار خودکار پشته‌ای، شبکه باور عمیق، ماشین بولتزمن عمیق، شبکه عصبی پیچشی و غیره است. در ادامه به معرفی مختصری از مدل‌های پایه می‌پردازیم.

۱.۲ رمزگذار خودکار

رمزگذار خودکار [۲] عمدتاً از رمزگذار، رمزگشا و لایه پنهان تشکیل شده است. فرآیند کار در شکل ۱ نشان داده شده‌است. یک رمزگذار خودکار ابتدا سیگنال ورودی را رمزگذاری می‌کند و سپس از سیگنال کدگذاری شده برای بازسازی سیگنال اولیه استفاده می‌کند. این سیگنال کدگذاری شده می‌تواند خطای بین سیگنال اولیه و سیگنال بازسازی شده را به حداقل برساند. در فرآیند رمزگذاری و بازسازی، رمزگذار دادههای ورودی را به یک فضای ویژگی خاص نگاشت می‌کند. ویژگی‌های کدگذاری شده می‌تواند سیگنال اولیه را به شکلی متفاوت بازیابی کند. بنابراین، ویژگی‌ها استخراج می‌شوند تا بتوانند به یادگیری خودکار دست یابند.

۲.۲ رمزگذار خودکار پشته‌ای

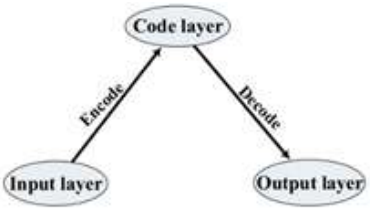
در سال ۲۰۰۶، پس از اینکه محققان به تدریج رمزگذار اتوماتیک بهینه‌شده، انکودر خودکار انقباضی، رمزگذار خودکار یدکی، رمزگذار خودکار کانولوشن و غیره را ارائه کردند، ساختار رمزگذار را برای بهبود رمزگذار خودکار پیشنهادی بدون نویز ارتقا داده شد. رمزگذارهای خودکار فوق، رمزگذارهای خودکار پشته‌ای هستند. رمزگذارهای خودکار پشته‌ای ساختارهای شبکه عمیقی هستند که توسط برهم نهی n بار ساختارهای کدگذاری خودکار ساده تشکیل شده‌اند. همانطور که در شکل ۲ نشان داده شده است، n رمزگذار خودکار که از پایین به بالا آموزش داده شده‌اند. ابتدا، اولین رمزگذار خودکار آموزش داده می‌شود و خطای بازسازی اولیه به حداقل می‌رسد. سپس خروجی رمزگذار خودکار اول به عنوان ورودی رمزگذار دوم آموزش داده می‌شود و تا آخرین لایه کار می‌کند. سپس از خروجی آخرین لایه به عنوان داده ورودی استفاده می‌شود. خروجی آخرین لایه به عنوان داده ورودی طبقه‌بندی کننده استفاده می‌شود و پارامترهای آن مجدداً مقداردهی اولیه می‌شوند. در نهایت بر اساس استاندارد نظارت، فقط قسمت بالایی تنظیم می‌شود و یا تمام لایه‌ها به درستی تنظیم می‌شوند.

۳.۲ ماشین بولتزمن محدود

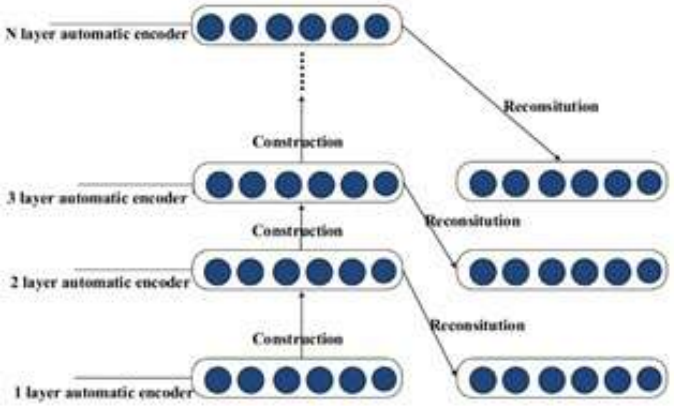
برای یک گراف دوبخشی، اگر بین لایه اول و لایه دوم پیوندی وجود نداشته باشد، لایه اول به عنوان لایه ورودی در نظر گرفته می شود (یعنی لایه بصری) و لایه دوم به عنوان لایه پنهان در نظر گرفته می‌شود. فرض کنید همه گره‌ها متغیرهای باینری تصادفی هستند. علاوه بر این، توزیع احتمال کامل $p(v, h)$ در معرض توزیع بولتزمن است که ماشین بولتزمن نامیده می‌شود. با توجه به مشخصه ماشین بولتزمن محدود، شرایط فعال سازی لایه‌های پنهان برای یک حالت معین از لایه قابل مشاهده (یعنی داده‌های ورودی) مستقل هستند. بنابراین، برای وضعیت یک لایه مخفی داده شده، شرایط فعالسازی لایه‌های قابل مشاهده مستقل است. اگرچه توزیع ماشین محدود شده بولتزمن را نمی‌توان به طور موثر محاسبه کرد، اما یک نمونه تصادفی را می‌توان با نمونه گیری گیبس به دست آورد. این نمونه تصادفی در معرض ماشین بولتزمن محدود قرار می‌گیرد. تا زمانی که تعداد لایه‌های پنهان کافی باشد، ماشین بولتزمن محدود می‌تواند هر توزیع گسسته را متناسب کند. از نظر کاربرد، مدل محدود شده بولتزمن با موفقیت برای حل مسائل مختلف یادگیری ماشین مانند رگرسیون، طبقه‌بندی، کاهش ابعاد، مدل‌سازی سری زمانی، فیلتر مشترک تصویر و استخراج ویژگی استفاده شده است [۱].

۴.۲ شبکه باور عمیق

شبکه باور عمیق از برهم نهی مدلهای محدود بولتزمن با شبکه‌های عصبی عامل توضیحی پنهان چند لایه تشکیل شده است. در فرآیند آموزش شبکه باور عمیق، پیش آموزش با یک روش حریصانه بدون نظارت انجام می‌شود تا مقادیر ویژه مدل لایه به لایه به‌دست آید.



شکل ۱: نمودار شماتیک رمزگذار خودکار



شکل ۲: رمزگذار خودکار پشته‌ای

روش لایه حریص بدون نظارت، واگرایی متضاد نامیده می‌شود و به عنوان یک روش معتبر ثابت شده است. در طول فرآیند آموزش، لایه بصری یک بردار v تولید می‌کند. داده‌ها را از طریق بردار v به لایه پنهان ارسال می‌کند. برعکس، ورودی لایه بصری به‌طور تصادفی برای تلاش برای بازسازی سیگنال ورودی اصلی انتخاب می‌شود. پس از آن، واحد فعال‌سازی نورون واحد بصری جدید می‌تواند به ارسال ورودی انتقال به منظور بازسازی واحد فعال‌سازی لایه پنهان ادامه دهد و سپس بردار h را بدست آورد. این فرآیندهای تکراری با نام نمونه‌گیری گیبس نیز شناخته می‌شوند. همبستگی بین ورودی لایه بصری و واحد فعال‌سازی لایه پنهان مهمترین مبنای اندازه‌گیری و به روز رسانی وزن است. ماشین‌های بولتزمن محدود برای هر لایه از پایین به بالا آموزش داده می‌شوند [۳].

۵.۲ ماشین بولتزمن عمیق

ماشین بولتزمن عمیق نیز توسط پشته ماشین بولتزمن محدود تشکیل شده است که شبیه به شبکه باور عمیق است. تفاوت بین ماشین بولتزمن عمیق و شبکه باور عمیق در این است که لایه قبلی و لایه فعلی بین اتصالات غیر جهتی قرار دارند و هیچ پارامتر بازخوردی از بالا به پایین وجود ندارد. روش آموزش عمیق ماشین بولتزمن ابتدا از پیش آموزش بدون نظارت برای بدست آوردن اقتدار اولیه مورد نظر استفاده می‌کند و سپس از الگوریتم میانگین میدانی استفاده می‌کند. در نهایت، تنظیم دقیق نظارت شده انجام می‌شود. دستگاه بولتزمن عمیق با مدل‌های دیگر متفاوت است. اولاً، ماشین بولتزمن عمیق توانایی یادگیری بازنمایی‌های درونی پیچیده‌تری را دارد که روش جدیدی برای تشخیص گفتار و تشخیص اشیا است. یادگیری عمیق می‌تواند به طور قابل توجهی عملکرد را در زمینه تشخیص صدا بهبود بخشد. ثانیا، ماشین بولتزمن عمیق می‌تواند نمایش بالاتری در تعداد زیادی از داده‌های بدون برچسب ایجاد کند. برای دستیابی به مقدار مورد نظر، بولتزمن عمیق از داده‌های برچسب مصنوعی شناخته شده برای تنظیم دقیق مدل استفاده می‌کند. همچنین، دستگاه بولتزمن عمیق می‌تواند برای مقابله با اطلاعات مبهم ورودی داده قوی‌تر باشد و میتواند بهتر پخش شود که باعث کاهش خطا در فرآیند انتشار می‌شود [۴].

۳ کاربردهای یادگیری عمیق

در ادامه کاربردهای مختلف یادگیری عمیق آورده شده است.

۱.۳ پردازش زبان طبیعی

در زبان طبیعی، یادگیری عمیق در بسیاری از زمینه‌ها مانند ترجمه صوتی، ترجمه ماشینی، درک معنایی کامپیوتری و غیره کاربرد دارد. در سال ۲۰۱۵، گوگل فناوری Word Lens را بر اساس یادگیری عمیق معرفی کرد که در ترجمه تماس بلادرنگ و ترجمه ویدیویی استفاده می‌شد. این فناوری نه تنها می‌تواند کلمات را در زمان واقعی بخواند، بلکه می‌تواند آن کلمات را به زبان مقصد مورد نظر ترجمه کند. فناوری فعلی را می‌توان در بیش از یک ترجمه تصویری از ۲۰ زبان به کار برد. علاوه بر این، گوگل یک تابع پاسخ خودکار به ایمیل در جیمیل پیشنهاد کرد که از یک مدل یادگیری عمیق برای استخراج محتوای ایمیل و تجزیه و تحلیل معنایی آن استفاده می‌کرد. در نهایت، یک پاسخ بر اساس تجزیه و تحلیل معنایی تولید می‌شود. این تکنیک اساساً با عملکرد سنتی پاسخ خودکار ایمیل متفاوت است.

۲.۳ تشخیص گفتار

به منظور تحقق تعامل انسان و رایانه، محققان تلاش زیادی کردند. تحقیق در مورد فناوری تشخیص گفتار چند دهه سابقه دارد و پیشرفت چشمگیری داشته است. با بهبود مستمر مدل یادگیری عمیق، تعداد زیادی از دستگاه‌ها یا برنامه‌های تشخیص گفتار از آزمایشگاه به سمت تجاری شدن، حرکت کرده‌اند.

۳.۳ کاربردهای پزشکی

عملکرد پیش‌بینی یادگیری عمیق و شناسایی خودکار ویژگی‌ها، آن را به روش محبوب در تشخیص بیماری نیز تبدیل می‌کند. در سال ۲۰۱۶، گوگل یک سیستم بینایی برای شناسایی بیماری‌های چشمی در مراحل اولیه ایجاد کرد. یک ماه بعد، گوگل از تکنیک‌های یادگیری عمیق برای طراحی روش پرتودرمانی سرطان سر و گردن استفاده کرد که کنترل موثری بر زمان رادیوتراپی بیمار داشت و توانست آسیب رادیوتراپی بیمار را به حداقل برساند.

۴.۳ بینایی ماشین

یکی از کاربردهای اساسی هوش مصنوعی است که می‌تواند از رایانه‌ها و دوربین‌ها برای جایگزینی چشم انسان برای تشخیص هدف، ردیابی، اندازه‌گیری و سایر مسائل بصری استفاده کند و حتی با قابلیت‌های پردازش تصویر به توانایی فراتر از چشم دست یابد. بینایی کامپیوتر یک اصطلاح گسترده است که جهات تحقیقاتی زیادی را به وجود می‌آورد. برخی از جوانب شناخته شده هستند که زیر چتر بینایی رایانه قرار می‌گیرند، عبارتند از: تقسیم بندی تصویر، تشخیص چهره، تشخیص اشیا، تقسیم بندی معنایی تصویر، تقسیم بندی اشیاء ویدیویی، جداسازی پس زمینه/پیش زمینه.

۵.۳ یادگیری عمیق روی گرافها

در سال‌های اخیر، محققان در تلاش برای توسعه تکنیک‌های جدیدی هستند که می‌توانند به طور موثر الگوها را از داده‌های ساخت -یافته گراف یاد بگیرند. بسیاری از مسائل وجود دارد که با استفاده از یادگیری عمیق بر روی گراف‌ها حل شده‌اند.

۶.۳ سیستم حمل و نقل هوشمند

سیستم های حمل و نقل هوشمند در قلب شهرهای هوشمند قرار دارند که تمرکز تحقیقاتی قرن ۲۱ است. سیستم‌های حمل و نقل، ستون فقرات هر ملتی در طول تاریخ تمدن بوده‌است. حمل و نقل هوشمند یکی از خواسته‌های تمام شهرهای بزرگ در سراسر جهان است. داده‌های حمل و نقل می‌تواند از حروف و اعداد گرفته تا تصاویر صوتی و فیلم‌ها متفاوت باشد. برای مثال یک شمارنده خودکار مسافر که منجر به پیش‌بینی درآمدزایی می‌شود، نیاز به تشخیص تصویر و نظارت تصویری دارد. همراه با شمارنده خودکار مسافر، همچنین باید در تحلیل مسیرهایی که مردم بیشتر و در چه ساعتی پیمودند، بپردازیم. به اطلاعات GPS و نقشه راه نیاز دارد. همچنین گاهی اوقات به داده‌های تولید شده غیر انسانی مانند آب و هوا نیاز دارد. این داده‌های ناهمگون از حسگرهای مختلفی به دست می‌آیند که در مکان‌های مختلف نصب می‌شوند، مثلاً در علائم راهنمایی و رانندگی، در اتومبیل‌ها و غیره. مشکلات اصلی که سیستم

های حمل و نقل هوشمند بر آن تمرکز دارد عبارتند از: پیش‌بینی مقصد، کنترل سیگنال ترافیک، پیش‌بینی تقاضا، پیش‌بینی جریان ترافیک، حالت حمل‌ونقل و بهینه‌سازی ترکیبی [۵].

۴ چالش‌های یادگیری عمیق

۱.۴ عدم وجود نوآوری در ساختار مدل

از سال ۲۰۰۶، مدل یادگیری عمیق عمدتاً به عنوان چندین مدل کلاسیک معرفی شد. آخرین معرفی از مدل‌های یادگیری عمیق در این مدل‌های سنتی بر اساس تکامل بیش از یک دهه بود. در گذشته اکثر مدل‌ها روی مدل‌های ساده چیده می‌شدند و به دلیل این انباشته شدن، افزایش کارایی پردازش داده‌ها دشوارتر می‌شد. با این حال، عمق مزایای فناوری یادگیری هنوز به طور کامل اجرا نشده است.

۲.۴ به روز رسانی روشهای آموزشی

یادگیری تحت نظارت و بدون نظارت دو روش آموزشی برای مدل‌های یادگیری عمیق فعلی هستند. روش‌های آموزشی تحت نظارت، ماشین بولتزمن محدود و رمزگذار خودکار به عنوان مدل اصلی و برای پیش‌آموزش اصلی، مانند تعداد زیادی روش آموزشی راه یادگیری بدون نظارت است. هیچ مفهوم واقعی برای انجام آموزش کامل بدون نظارت وجود ندارد. بنابراین، چگونگی دستیابی به آموزش کامل بدون نظارت، مسیر مطالعه آینده مدل یادگیری عمیق است.

۳.۴ چالش‌های یادگیری پارامتر

چالش‌های زیادی برای یادگیری پارامتر در شبکه‌های عصبی عمیق وجود دارد که به طور خلاصه عبارتند از: میزان یادگیری، بهینه محلی، نقاط زینی، ناپدید شدن و انفجار شیب.

۴.۴ کاهش زمان تمرین

در حال حاضر، تشخیص انواع مدل‌های یادگیری عمیق بیشتر در محیط ایده‌آل انجام می‌شود. در محیط پیچیده واقعی، فناوری فعلی هنوز قادر به دستیابی به نتایج مطلوب نیست. همچنین مدل یادگیری عمیق یا از مدل ساده یا چند مدل تشکیل شده است. از آنجایی که پیچیدگی مسائل بیشتر است، میزان اطلاعات پردازش شده قابل توجه است، به این معنی که نیاز به زمان آموزش بسیار زیاد مدل یادگیری عمیق وجود دارد. چگونگی تغییر مدل یادگیری عمیق بدون هیچ گونه انعطاف در سخت‌افزار برای بهبود دقت و سرعت پردازش داده‌ها، تحقیقات آینده فناوری یادگیری عمیق است.

۵.۴ یادگیری آنلاین

پیش‌آموزش بدون نظارت و تنظیم دقیق نظارت شده، روش‌های اصلی آموزشی برای تکنیک‌های یادگیری عمیق امروزی هستند. با این حال، آموزش یادگیری آنلاین نیاز به تنظیم دقیق سراسری دارد که باعث می‌شود خروجی به بهینه محلی برسد. بنابراین، آموزش فعلی برای تحقق یادگیری آنلاین مساعد نیست. باید با پیشرفت‌های یادگیری آنلاین مبتنی بر یک مدل یادگیری عمیق خلاقانه رو به رو شد.

۶.۴ غلبه بر نمونه متخاصم

اگر نمونه ورودی تعمداً اضافه شود، تداخل ظریف در مجموعه داده‌ها می‌تواند باعث شود که مدل خروجی اشتباه را با اطمینان بالا ارائه دهد. با این حال، نمونه متخاصم یک مشکل بزرگ در یادگیری عمیق فعلی است. این افزودن نمونه ورودی نه تنها نمی‌تواند به طور موثر از مشکلات امنیتی احتمالی در هنگام بررسی چگونگی غلبه بر مساله نمونه متخاصم جلوگیری کند، بلکه می‌تواند به بهبود مدل یادگیری عمیق برای حل مشکل دقت کمک کند. به یک معنا، یک تضاد اساسی بین ایجاد یک مدل خطی برای آموزش آسان و ایجاد یک مدل غیر خطی که بتواند در برابر نمونه مقاومت کند، وجود دارد. با این حال، از توسعه بلندمدت یادگیری عمیق، ایجاد روش‌های بهینه‌سازی قوی‌تر و مدل‌های غیرخطی بیشتر آموزش، مسیر آینده در این زمینه تحقیقاتی است.

۵ نتیجه‌گیری

در این مقاله برخی از پیشرفت‌های یادگیری عمیق و کاربردهای آن در زمینه‌های مختلف مورد بحث قرار داده شد. مسائل علمی بسیاری وجود دارد که روز به روز در حال حل شدن است، بنابراین گاهی اوقات می‌توان با یادگیری عمیق در بسیاری از زمینه ها به عملکرد بهتری دست یافت. در آینده نزدیک، یادگیری عمیق برای داده‌های غیر اقلیدسی مانند یادگیری عمیق برای گراف‌ها، یادگیری عمیق هندسی، شبکه‌های عصبی هایپربولیک، و چگونگی بهبود ساختارها و الگوریتم‌های یک مدل شبکه عصبی عمیق [۶] و غیره مورد توجه قرار خواهند گرفت.

مراجع

[1] D.H. Ackley, G.E. Hinton, T.J. Sejnowski, *A learning algorithm for boltzmann machines*, Cognitive Science 9 (1985), No. 1, 147–169.

[2] Y. Bengio, *Learning Deep Architectures for AI*, Now Publishers Inc, 2009.

[3] S. Rifai, Y. Bengio, A. Courville, P. Vincent, M. Mirza, Disentangling factors of variation for facial expression recognition, in: European Conference on Computer Vision, Springer, (2012), 808–822.

[4] R. Salakhutdinov, G. Hinton, Deep boltzmann machines, in: Artificial Intelligence and Statistics, (2009), 448–455.

[5] M. Veres, M. Moussa, *Deep learning for intelligent transportation systems: a survey of emerging trends*, IEEE Transactions on Intelligent Transportation Systems 21 (2019), No. 8, 3152–3168.

[6] S. Wang, J. Cao, P. Yu, *Deep learning for spatio-temporal data mining: a survey, IEEE Transactions on Knowledge and Data Engineering*, (2020) 1-1 http://dx.doi.org/10.1109/TKDE.2020.3025580.

موازی سازی الگوریتم انتخاب موارد آزمون رگرسیون سیستم های نرم افزاری با استفاده از تجزیه مقادیر منفرد

مهدی موحدیان مقدم[۱] * و دکتر محمود فضلعلی[۱] †

[۱]گروه علوم کامپیوتر، دانشکده ریاضی، دانشگاه شهید بهشتی، تهران، ایران

چکیده

هنگام توسعه یک سیستم نرم افزاری تغییر در یک قسمت سیستم ممکن است منجر به ایجاد تغییرات ناخواسته در بخش های دیگر سیستم شود. این بخش های متاثر ممکن است عملکرد سیستم را دچار اختلال کند، بنابراین از آزمون رگرسیون جهت مقابله با این اختلالات استفاده می شود. این آزمون درصدد سنجش مجدد این بخش ها برآمده تا از این اختلالات جلوگیری کند، اما تشخیص این بخش ها جهت آزمون مجدد کار مشکلی است. تلاش ما بر این است که به کمک خوشه بندی تغییرات سیستم نرم افزاری خود براساس توابع سازنده سیستم به کمک تجزیه مقادیر منفرد ، بتوان برای تشخیص این بخش ها در طی یک تغییر جدید، جهت انجام مجدد آزمون، استفاده نمود. به جهت افزایش کارایی، محاسبات ما به صورت موازی روی سیستم های حافظه مشترک صورت گرفت تا با افزایش مقیاس سیستم های نرم افزاری ، بتوان پاسخی بهینه دریافت نمود.

واژه‌های کلیدی: تجزیه مقادیر منفرد ، آزمون نرم افزاری رگرسیون ، کاهش موارد آزمون موازی ، خوشه بندی موازی موارد آزمون

کد موضوع‌بندی ریاضی [۲۰۱۰]: 62P30

۱ مقدمه

امروزه در توسعه سیستم های نرم افزاری پس از یک دوره توسعه، ممکن است نیاز به ایجاد تغییرات در بخش های توسعه داده شده قبلی باشد. این تغییرات ممکن است موجب تاثیر بر قسمت های دیگر نرم افزار شود. این تاثیرات ممکن است موجب ایجاد اختلال جدیدی در عملکرد سیستم شود. جهت جلوگیری از این اختلالات ناخواسته نوعی از انواع آزمون های نرم افزاری به نام آزمون رگرسیون [۱] به کار گرفته می شود. تست رگرسیون از موارد آزمون [۲] موجود برای بررسی اینکه آیا تغییرات ایجاد شده خطای جدیدی ایجاد نکرده و خطاهای جانبی ناخواسته ای نداشته است، استفاده می کند. به عبارت دیگر، هدف این است که اطمینان حاصل شود که بخش‌هایی از یک سیستم که تغییر نکرده‌اند، همچنان مانند قبل از انجام تغییرات کار می‌کنند [۵]. آزمون رگرسیون درصدد اجرای مجدد مجموعه موارد آزمون برخواهد آمد که بتوان پس از اجرای موفقیت آمیز همه آن ها به طور قطعی نسبت به صحت عملکرد سیستم نظر داد. به دست آوردن این مجموعه موارد آزمون کار چندان آسانی نخواهد بود طوری که از روش های برنامه ریزی خطی صحیح نیز امروزه در جهت یافتن این مجموعه موارد آزمون بهره گرفته شده است [۲]. ما بر این تلاش هستیم تا این مجموعه موارد آزمون را به کمک خوشه بندی توسط تجزیه مقادیر منفرد [۳] بر روی ماتریسی از توابع [۴] به کار رفته در سیستم که از نظر تغییرات زمانی به یکدیگر نزدیک هستند

* سخنران. آدرس ایمیل: m.movahedian@mail.sbu.ac.ir
† سخنران. آدرس ایمیل: fazlali@sbu.ac.ir

[۱] Regression Test
[۲] Test cases
[۳] Singular value decomposition
[۴] Functions

به دست آوریم. این اطلاعات به کمک ابزارهای کنترل منبع [۱] که امروزه در اکثر شرکت های تولید کننده سیستم های نرم افزاری به کار گرفته می شود، به سادگی به دست خواهد آمد. به دلیل وجود حجم بالای توابع در سیستم های نرم افزاری تجاری ما در طی این خوشه بندی از یک رویکرد موازی [۲] استفاده نمودیم تا به این طریق امکان استفاده از این رویکرد در سیستم های تجاری بزرگ نیز در کوتاه ترین زمان ممکن، فراهم شود.
این پژوهش در بخش دوم به معرفی روش های مورد استفاده می پردازد و در بخش سوم از این روش ها جهت انجام خوشه بندی اطلاعات و در پی آن استخراج این داده ها از یک سیستم نرم افزاری به صورت موازی و یافتن مجموعه مواردی که باید به ازای یک بردار تغییر خاص مورد آزمون مجدد قرار بگیرند استفاده شده است، در بخش چهارم نتایج تجربی این پژوهش و بخش پنجم جهت نتیجه گیری نهایی تدوین شده است.

۲ کارهای مرتبط

در این بخش به کارهای مرتبط انجام شده در حوزه انتخاب موارد آزمون رگرسیون و همچنین خوشه بندی کد منبع [۳] و تجزیه مقادیر منفرد پرداخته می شود.

۱.۲ انتخاب موارد آزمون رگرسیون

تست رگرسیون فرآیند تست مجدد نرم افزار است ، که اصلاح شده می باشد. سیستم‌های بزرگ معمولاً مجموعه‌های تست رگرسیون بزرگی دارند. در یک سیستم نرم‌افزاری، تغییرات کوچک در یک بخش از یک سیستم اغلب باعث ایجاد مشکلاتی در بخش‌های دوردست سیستم می‌شود. برای یافتن این نوع مشکلات از تست رگرسیون استفاده می شود [۱]. هدف از تکنیک‌های انتخاب آزمون رگرسیون ، جداسازی تست‌هایی است که به صورت احتمالی پس از اصلاح سیستم، خطاهای ایجاد شده را کشف می‌کنند. ساده‌ترین تکنیک انتخاب آزمون رگرسیون، استراتژی همه‌آزمایی مجدد است که در آن کل مجموعه موارد آزمایشی مورد بررسی قرار می گیرند [۴].

۲.۲ خوشه بندی کد منبع

امروزه خوشه بندی کد منبع به جهت کاربرد آن در سیستم های تشخیص سرقت ادبی اهمیت بخصوصی پیدا کرده است،از جمله این موارد می توان تلاش (Duracik و همکاران) را بدین صورت ذکر کرد که سعی بر خوشه بندی خط به خط کد منبع یک سیستم نرم افزاری، براساس الگوریتم خوشه بندی k-means داشته اند، که سپس با در پیش گرفتن یک رویکرد افزایشی جهت این خوشه بندی، عمل جستجو را به صورت برداری بر روی کد منبع به سرانجام رسانده اند [۳]. در این پژوهش، ما خوشه بندی تغییرات کد منبع را براساس تغییرات توابع توسعه داده ایم.

۳.۲ تجزیه مقادیر منفرد موازی

در سال های اخیر موازی سازی های متعددی بر روی الگوریتم تجزیه مقادیر منفرد صورت گرفته است، در تلاش (Xiao و همکاران) شاهد موازی سازی الگوریتم مذکور براساس موازی سازی بلوکی ژاکوبی بوده ایم [۶]. ما در این پژوهش علاوه بر موازی سازی الگوریتم تجزیه مقادیر منفرد خویش، استخراج اطلاعات خوشه های خود را به صورت موازی توسعه دادیم. این توسعه موازی برای استفاده توسط زیرساخت های شرکت های نرم افزاری تجاری بسیار مناسب تر می باشد، زیرا شرکت های تجاری بزرگ که دارای زیرساخت های سخت افزاری قوی تری می باشند عموما در حال توسعه سیستم های نرم افزاری هستند که به این نوع آزمون بسیار نیازمندند.

۳ انتخاب موازی موارد آزمون رگرسیون براساس تجزیه مقادیر منفرد

این پژوهش الگوریتمی را ارائه می کند که براساس تغییرات توابع در یک کد منبع سیستم نرم افزاری، بتوان مجموعه توابعی که روی یکدیگر تاثیر بیشتری دارند را شناسایی کرد و به کمک تجزیه مقادیر منفرد، به نوعی خوشه بندی دست یافت که از طریق این خوشه

[۱] Source control
[۲] Parallel
[۳] Source code

بندی بتوان مجموعه خلاصه تری از مجموعه موارد آزمون رگرسیون را تولید نمود. تمامی الگوریتم عنوان شده به صورت موازی روی سیستم های حافظه مشترک جهت افزایش کارایی توسعه داده شده است.الگوریتم فوق ۱.۳ به این شرح می باشد.

الگوریتم ۱.۳. انتخاب موازی موارد آزمون رگرسیون براساس تجزیه مقادیر منفرد

۱. ایجاد ماتریس F براساس تغییرات توابع نسبت به یکدیگر

۲. ایجاد ماتریس T براساس هر مورد آزمون نسبت به توابع مربوطه به خود

۳. ایجاد ماتریس R جهت نگهداری خوشه ها

۴. ایجاد بردار X جهت نگهداری توابع تغییر یافته

۵. $[U, S, V] = PSVD(F)$.

۶. برای $i = ۱, \ldots, size(U)$ به صورت موازی انجام دهید:

۷. برای $j = ۱, \ldots, size(U)$ به صورت موازی انجام دهید:

۸. اگر $|U[j, i]| > threshold$ انجام دهید:

۹. شروع ناحیه بحرانی

۱۰. افزودن داده های مثبت خوشه $U[i]$ به ماتریس R

۱۱. افزودن داده های منفی خوشه $U[i]$ به ماتریس R

۱۲. پایان ناحیه بحرانی

۱۳. پایان حلقه

۱۴. پایان حلقه

۱۵. $P = T * R$.

۱۶. $Y = R * P^T$.

۱۷. $S = Y * X$.

۱.۳ داده های اولیه

در سیستم های کنترل منبع، تمامی تغییرات کد منبع یک سیستم نرم افزاری در حین توسعه توسط سیستم نرم افزاری، ثبت می شود. این تغییرات در قالب یک تاریخچه در سطوح متفاوتی از قبیل فایل ها و توابع و یا حتی خطوط تغییر یافته، قابل رویت می باشد. ما این تغییرات را در سطح توابع رصد نموده و اجرای موارد آزمون سیستم را به توابع محدود نمودیم.این کار به دو دلیل صورت گرفت که مورد اول را می توان اینگونه بیان کرد که اگر در سطح فایل تغییرات را بررسی کنیم ممکن است به دلیل وجود خطوط زیادی از کدهای منبع موجود در هر فایل، دقت لازم حاصل نشود و دلیل دوم اینکه تغییرات کد منبع را در سطح خطوط تغییر یافته محدود نساختیم زیرا ابعاد ماتریس های تغییرات ما به صورت سرسام آوری افزایش می یافت و باعث تنگی زیاد آن ها در مسئله می شد. به دلایل ذکر شده تصمیم بر این شد که تغییرات کد منبع در سطح تغییرات توابع مورد بررسی قرار گیرد.

جهت آماده سازی داده های مورد نیاز، یک ماتریس مربعی از مرتبه تعداد توابع مورد بررسی که هر درایه آن نشان دهنده وجود تغییرات دو تابع با یکدیگر می باشد را از طریق این سیستم کنترل منبع می سازیم. این ماتریس، F نامیده می شود. این ارتباط به صورت دوطرفه می باشد یعنی تغییرات تابع i و j ام اگر صورت گرفته باشد هر دو درایه مرتبط با F(i,j) و F(j,i) باید مقداردهی شوند، پس ماتریس F حالت متقارن دارد. این ماتریس دارای ابعاد M * M خواهد بود که M به عنوان تعداد توابع مورد بررسی می باشد.

ماتریس دیگری که مورد نیاز است ماتریس T بوده که هر درایه آن نشان گر ارتباط بین هر مورد آزمون با توابع آن سیستم نرم افزاری می باشد. این ماتریس لزوما متقارن نیست چون لزوما تعداد موارد آزمون با تعداد توابع یکسان نخواهد بود و ممکن است یک مورد آزمون با تعدادی تابع ارتباط داشته باشد یا بالعکس.این ماتریس نیز دارای ابعاد TC * M خواهد بود که TC بیان گر تعداد موارد آزمون می باشد.

۲.۳ خوشه بندی موازی توسط تجزیه مقادیر منفرد

همانگونه که ذکر شد، ماتریس F به دلیل متقارن و مثبت معین بودن دارای تجزیه مقادیر منفردی با مقادیر U و V یکسان می باشد. پس از اجرای تجزیه مقادیر منفرد بر روی ماتریس F خروجی مطلوب سه ماتریس V،U،Σ است.

$$F = \begin{pmatrix} ۳ & ۱ & ۳ & ۴ & ۲ \\ ۱ & ۰ & ۰ & ۴ & ۰ \\ ۳ & ۰ & ۱ & ۱ & ۰ \\ ۴ & ۴ & ۱ & ۲ & ۰ \\ ۲ & ۰ & ۰ & ۰ & ۱ \end{pmatrix}, T = \begin{pmatrix} ۱ & ۰ & ۰ & ۰ & ۰ \\ ۰ & ۱ & ۰ & ۰ & ۰ \\ ۰ & ۰ & ۱ & ۰ & ۰ \\ ۰ & ۰ & ۱ & ۱ & ۰ \\ ۰ & ۰ & ۱ & ۱ & ۱ \\ ۰ & ۰ & ۰ & ۱ & ۱ \end{pmatrix} \quad (۱)$$

پس از تجزیه مقادیر منفرد و اجرای موازی دو حلقه موجود در الگوریتم جهت تولید خوشه های مد نظر به ماتریسی دست خواهیم یافت که در ابعاد M * C و بدین صورت خواهد بود. جهت جلوگیری از نوشتن همزمان مقادیر توسط دو واحد اجرایی موازی(مانند نخ یا فرآیند) از نواحی بحرانی استفاده نمودیم تا از نوشته شدن مقادیر نامعتبر در یک درایه از ماتریس R جلوگیری شود.

$$R = \begin{pmatrix} -۰/۶۴ & ۰/۴ & ۰/۰۰ & ۰/۰۰ & -۰/۳۶ & ۰/۰۰ & -۰/۵۳ & ۰/۰۰ & ۰/۰۰ \\ -۰/۳۳ & ۰/۵۹ & ۰/۰۰ & ۰/۵۳ & ۰/۰۰ & ۰/۴۷ & ۰/۰۰ & ۰/۱۵ & ۰/۰۰ \\ -۰/۳۱ & ۰/۰۰ & -۰/۱۱ & ۰/۰۰ & -۰/۴۱ & ۰/۵۸ & ۰/۰۰ & ۰/۰۰ & ۰/۶ \\ -۰/۵۹ & ۰/۰۰ & -۰/۶۶ & ۰/۴۴ & ۰/۰۰ & ۰/۰۰ & -۰/۰۹ & ۰/۰۳ & ۰/۰۰ \\ -۰/۱۵ & ۰/۰۰ & -۰/۱۶ & ۰/۰۰ & -۰/۴۵ & ۰/۳۶ & ۰/۰۰ & ۰/۷۷ & ۰/۰۰ \end{pmatrix} \quad (۲)$$

هر ستون این ماتریس یا به عبارت دیگر هر خوشه از این مجموعه بیان گر توابعی هستند که نسبت به هم وابستگی تغییرات بیشتری داشته اند. تنظیم این حد از وابستگی بین توابع به تنظیم مقدار آستانه در الگوریتم وابسته است و به نوعی با آن رابطه مستقیم دارد، طوری که هر چقدر آستانه را افزایش دهیم خوشه هایی خواهیم داشت که دارای عناصری با وابستگی های بیشتر هستند.

۳.۳ تولید مجموعه کاهش یافته موارد آزمون

پس از به دست آمدن ماتریس R به کمک ماتریس T باید نوعی ارتباط میان موارد آزمون رگرسیون و این خوشه های حاصل، ایجاد نمود. پس به کمک ترانهاده کردن ضرب آن دو ماتریس به ماتریسی دست می یابیم که به صورت خوشه بندی شده موارد آزمونی که به ازای هر تغییر تابع باید مورد بررسی قرار گیرند، را نشان خواهد داد. این ماتریس حاصل را Y می نامیم.

$$Y = R * (T * R)^T = \begin{pmatrix} -۰/۶۴ & ۰/۴ & ۰/۰۰ & ۰/۰۰ & -۰/۳۶ & ۰/۰۰ & -۰/۵۳ & ۰/۰۰ & ۰/۰۰ \\ -۰/۳۳ & ۰/۵۹ & ۰/۰۰ & ۰/۵۳ & ۰/۰۰ & ۰/۴۷ & ۰/۰۰ & ۰/۱۵ & ۰/۰۰ \\ -۰/۳۱ & ۰/۰۰ & -۰/۱۱ & ۰/۰۰ & -۰/۴۱ & ۰/۵۸ & ۰/۰۰ & ۰/۰۰ & ۰/۶ \\ -۰/۵۹ & ۰/۰۰ & -۰/۶۶ & ۰/۴۴ & ۰/۰۰ & ۰/۰۰ & -۰/۰۹ & ۰/۰۳ & ۰/۰۰ \\ -۰/۱۵ & ۰/۰۰ & -۰/۱۶ & ۰/۰۰ & -۰/۴۵ & ۰/۳۶ & ۰/۰۰ & ۰/۷۷ & ۰/۰۰ \end{pmatrix} *$$

$$\begin{pmatrix} -۰/۶۴ & -۰/۳۳ & -۰/۳۱ & -۰/۹۰ & -۱/۰۵ & -۱/۰۵ \\ ۰/۴۰ & ۰/۵۹ & ۰/۰۰ & ۰/۰۰ & ۰/۰۰ & ۰/۰۰ \\ ۰/۰۰ & ۰/۰۰ & -۰/۱۱ & -۰/۷۷ & -۰/۹۳ & -۰/۹۳ \\ ۰/۰۰ & ۰/۵۳ & ۰/۰۰ & ۰/۴۴ & ۰/۴۴ & ۰/۴۴ \\ -۰/۳۶ & ۰/۰۰ & -۰/۴۱ & -۰/۴۱ & -۰/۸۶ & -۰/۸۶ \\ ۰/۰۰ & ۰/۴۷ & ۰/۵۸ & ۰/۵۸ & ۰/۹۴ & ۰/۹۴ \\ -۰/۵۳ & ۰/۰۰ & ۰/۰۰ & -۰/۰۹ & -۰/۰۹ & -۰/۰۹ \\ ۰/۰۰ & ۰/۱۵ & ۰/۰۰ & ۰/۰۳ & ۰/۰۸ & ۰/۰۸ \\ ۰/۰۰ & ۰/۰۰ & ۰/۶۰ & ۰/۶۰ & ۰/۶۰ & ۰/۶۰ \end{pmatrix} = \begin{pmatrix} ۰/۹۸ & ۰/۴۴ & ۰/۳۴ & ۰/۷۷ & ۱/۰۲ & ۱/۰۲ \\ ۰/۴۴ & ۰/۹۸ & ۰/۳۷ & ۰/۸۰ & ۱/۱۴ & ۱/۱۴ \\ ۰/۳۴ & ۰/۳۷ & ۰/۹۷ & ۱/۲۲ & ۱/۶۸ & ۱/۶۸ \\ ۰/۴۲ & ۰/۴۳ & ۰/۲۵ & ۱/۲۴ & ۱/۴۵ & ۱/۴۵ \\ ۰/۲۵ & ۰/۳۳ & ۰/۴۵ & ۰/۶۷ & ۱/۶۴ & ۱/۶۴ \end{pmatrix} \quad (۳)$$

بنابراین به ازای هر بردار تغییر X دارای ابعادی به صورت M * 1 که فقط نشان دهنده توابع تغییر یافته می باشد. با اعمال عمل ضرب در ترانهاده ماتریس Y به دست آمده، می توان مجموعه موارد آزمونی که باید مورد بررسی مجدد در این نسخه از نرم افزار قرار

گیرند مشخص شود.

$$X = \begin{pmatrix} ۰ \\ ۱ \\ ۱ \\ ۰ \\ ۱ \end{pmatrix}, Y^T * X = \begin{pmatrix} ۱/۰۵ \\ ۱/۶۹ \\ ۱/۸۰ \\ ۲/۷۱ \\ ۴/۴۷ \\ ۴/۴۷ \end{pmatrix} \quad (۴)$$

بردار حاصل نشان دهنده مجموعه موارد آزمونی است که باید مجدد در ازای این تغییرات مورد بررسی قرار گیرند. این نتیجه نشان دهنده آن است که موارد آزمون ۵ و ۶ دارای اولویت بالاتری جهت اجرای مجدد نسبت به اجرای موارد آزمون دیگر می باشد.

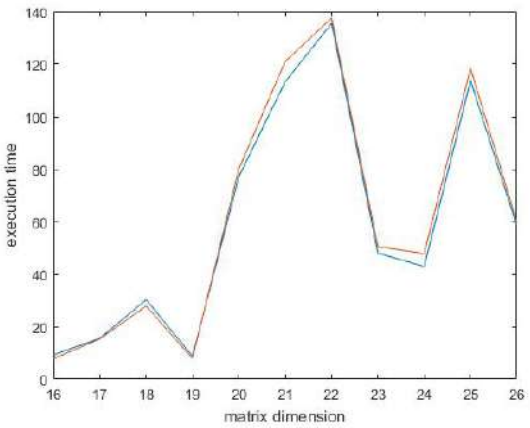
۴ نتایج تجربی

در این بخش بررسی نمودیم چنانچه به جای تغییرات فایل های یک سیستم نرم افزاری، جهت افزایش دقت، تغییرات توابع سیستم را بررسی نماییم با چه چالش هایی رو به رو خواهیم بود.
در بررسی تغییرات توابع با یکدیگر با ابعاد ماتریس های ورودی خیلی بزرگتری رو به رو خواهیم بود ، زیرا تعداد توابع یک سیستم بسیار زیادتر از تعداد فایل های آن سیستم است،این رویکرد برای توسعه دهندگان یک سیستم نرم افزاری چالش محاسباتی ایجاد می کند. بدین جهت رویکردی موازی برای الگوریتم فوق به جای رویکردی ترتیبی ارائه دادیم که در این بخش سعی بر مقایسه آن دو و تحلیل نقاطی داریم که برای اجرای موازی الگوریتم مناسب تر می باشد.
جهت انجام آزمایش خود، از یک سیستم چند پردازنده حافظه مشترک با ۴ هسته پردازشی نسل سوم با سرعت ساعت 2.9 گیگاهرتز

جدول ۱: مقایسه زمان اجرای الگوریتم به صورت موازی و ترتیبی بر حسب ثانیه.

بعد	۱۶	۱۷	۱۸	۱۹	۲۰	۲۱	۲۲	۲۳	۲۴	۲۵	۲۶
موازی	9.27	15.55	30.25	8.71	77.08	113.32	135.53	48.00	42.91	113.80	58.32
ترتیبی	7.81	15.35	27.84	7.88	79.98	121.077	137.53	50.45	47.84	118.35	59.77

ساخته شرکت اینتل، به همراه پیاده سازی دو نسخه موازی و غیرموازی از الگوریتم فوق، استفاده نمودیم. جهت جمع آوری داده های مربوط به تغییرات سیستم نرم افزاری از سیستم های کنترل منبع استفاده می نماییم. این سیستم ها در هر بازه زمانی قادر به ثبت تغییرات محصول نرم افزاری به ازای هر تغییر ایجاد شده توسط هر یک از اعضای تیم توسعه نرم افزار می باشند.
همان طور که در جدول ۱ مشاهده می کنید برای ابعاد ماتریس کمتر از ۲۰ روش موازی شده دارای زمان اجرای بیشتری نسبت به



شکل ۱: مقایسه زمان اجرای الگوریتم به صورت موازی و ترتیبی بر حسب ثانیه.

روش ترتیبی می باشد، اما برای ابعاد بزرگتر از این مقدار روش موازی شده دارای سرعت اجرای کمتری می باشد. بدین جهت که

پژوهش ما به دلیل بررسی تغییرات توابع با یکدیگر دارای ماتریس هایی با ابعاد ورودی بسیار بالا می باشد، ما استفاده از رویکرد موازی را پیشنهاد می نماییم.
با توجه به شکل ۱ این موازی سازی انجام شده بر روی ۱۱ مورد از ابعاد ماتریس های متفاوت حاصل از تغییرات یک سیستم نرم افزاری مورد بررسی قرار گرفته است که اختلالات ناشی از آن به دلیل مقدار تنک بودن هر ماتریس می باشد که این اختلالات، تاثیری روی مقایسه دو رویکرد موازی و ترتیبی و نتیجه حاصل از این دو رویکرد نداشته است. این تغییرات ناشی از این می باشد که در یک نسخه از نرم افزار تغییرات کمتری نسبت به نسخه های دیگر وجود داشته است.

۵ نتیجه‌گیری

در رویکرد معرفی شده توسط (Sherriff و همکاران) به دلیل بررسی تغییرات فایل ها و در پی آن پایین بودن اندازه مسئله، نیازی به موازی سازی نبود [۴]. اما در رویکرد معرفی شده توسط ما به دلیل افزایش دقت و بررسی توابع تغییر یافته در فایل ها، با افزایش اندازه مسئله رو به رو شدیم که با رویکرد موازی ارائه شده توانستیم به این چالش محاسباتی غلبه نماییم. در نتیجه؛ این پژوهش که به کمک رویکرد موازی خوشه بندی کد منبع براساس تغییرات توابع توسعه یافته است، نسبت به پژوهش هایی که براساس تغییرات فایل های کد منبع انجام شده بود، دقت عمل بسیار بالاتری به دست آمده است.

تشکر و قدردانی

در آخر از دکتر کوروش پرند، هیئت علمی دانشکده ریاضی دانشگاه شهید بهشتی، که اینجانب را در این مسیر بسیار یاری نموده اند، کمال تشکر را دارم.

مراجع

[1] P. Ammann, J. Offutt, *Introduction to Software Testing*, Cambridge University Press, Cambridge, 2017.

[2] C. Chi-Lun ,H. Chin-Yu ,C. Chang-Yu ,C. Kai-Wen and L. Chen-Hua, Analysis and assessment of weighted combinatorial criterion for test suite reduction, *Quality and Reliability Engineering International*, (2021), 1–31.

[3] M. Duracik,E. Krsak and P. Hrkut, Searching source code fragments using incremental clustering, *Concurrency and Computation Practice and Experience*, 32 (2020), No. 13

[4] M. Sherriff, M. Lake and L. Williams, Prioritization of Regression Tests using Singular Value Decomposition with Empirical Change Records, The 18th IEEE International Symposium on Software Reliability (ISSRE '07), (2007).

[5] A. Spillner, T. Linz,*Software Testing Foundations A Study Guide for the Certified Tester Exam*, Rocky Nook, 2021.

[6] P. Xiao, Z. Wang and S. Rajasekaran, Novel Speedup Techniques for Parallel Singular Value Decomposition, 20th International Conference on High Performance Computing and Communications, 2018.