دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

دانشکده ریاضی و علوم کامپیوتر
استاد درس: دکتر مهدی دهقان

پاییز ۱۴۰۳

مقدمه‌ای بر SVD و کاربرد آن در دسته‌بندی داده‌ها

هلیا صاحب‌قدم | ۹۹۲۵۰۲۳ | جبر خطی عددی

فهرست مطالب

۱ چکیده **۳**

۲ مقدمه **۴**
۱.۲ تجزیه مقادیر منفرد ۴
۲.۲ محاسبه ی تجزیه مقادیر منفرد ۵
۳.۲ کاربرد های تجزیه مقادیر منفرد ۹

۳ دسته بندی ارقام دست نویس **۱۱**
۱.۳ شرح مساله ۱۱
۲.۳ تنسور ۱۱
۳.۳ دیتاست ۱۲
۴.۳ یک الگوریتم ساده برای طبقه‌بندی ۱۳
۵.۳ دسته بندی داده ها با پایه های SVD ۱۳

۴ فاصله تانژانت **۱۹**
۱.۴ تبدیلات ۲۱

۵ نتیجه‌گیری **۲۴**

۱ چکیده

در دنیای یادگیری ماشین و علوم داده، تجزیه مقدار منفرد به عنوان یکی از تکنیک‌های کلیدی برای کاهش ابعاد داده و استخراج ویژگی‌های اساسی استفاده می‌شود. این روش با کاهش نویز و حفظ اطلاعات مهم، ابزاری قدرتمند برای بهبود عملکرد مدل‌های طبقه‌بندی است. در این پژوهش، ابتدا اصول و مبانی تجزیه مقدار منفرد معرفی شده و سپس نحوه محاسبه و کاربردهای آن بررسی می‌شود.

یکی از کاربردهای مهم آن در دسته‌بندی ارقام دست‌نویس است. در این مقاله، یک الگوریتم ساده برای طبقه‌بندی این داده‌ها با استفاده از پایه‌های تجزیه مقادیر منفرد ارائه شده است. نتایج آزمایشات نشان می‌دهد که این تکنیک نه تنها پیچیدگی محاسباتی را کاهش می‌دهد، بلکه عملکرد مدل‌های طبقه‌بندی را نیز بهبود می‌بخشد. علاوه بر این، تاثیر تغییرات و تداخلات بر فاصله تانژانت بررسی شده است.

واژه‌های کلیدی: تجزیه مقدار منفرد، کاهش ابعاد، استخراج ویژگی، طبقه‌بندی ارقام دست‌نویس، فاصله تانژانت، تنسور

۲ مقدمه

تجزیه مقادیر منفرد[1] یک تکنیک پایه‌ای در فاکتورگیری ماتریس است که ساختارهای پنهان را استخراج، ابعاد داده‌ها را کاهش و تفسیرپذیری داده‌ها را افزایش می‌دهد. این روش ماتریس A را به صورت $U\Sigma V^T$ تجزیه می‌کند، که در آن U و V ماتریس‌های متعامد هستند و Σ شامل مقادیر منفرد[2] مرتب شده است که ویژگی‌های غالب داده را نشان می‌دهند. برخلاف تجزیه QR، SVD ردیف‌ها و ستون‌ها را به طور متقارن مدیریت کرده و بینش عمیق‌تری ارائه می‌دهد و تجزیه خود را به مسائل خاص تطبیق می‌دهد. SVD در روش‌هایی مانند تحلیل مؤلفه‌های اصلی[3] و حل مسائل کمترین مربعات[4] برای ماتریس‌های غیرمربعی نقش کلیدی دارد. تطبیق‌پذیری و توانایی آن در اولویت‌بندی مؤلفه‌های تأثیرگذار، آن را به ابزاری بی‌بدیل در داده‌کاوی، پردازش سیگنال و علوم محاسباتی تبدیل کرده است.

۱.۲ تجزیه مقادیر منفرد

هر ماتریس A با ابعاد $m \times n$ که $m \geq n$، می‌تواند به صورت زیر تجزیه شود:

$$A = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^T,$$

که در آن $U \in \mathbb{R}^{m\times m}$ و $V \in \mathbb{R}^{n\times n}$ ماتریس‌های متعامد هستند و $\Sigma \in \mathbb{R}^{n\times n}$ ماتریسی قطری است به صورت زیر:

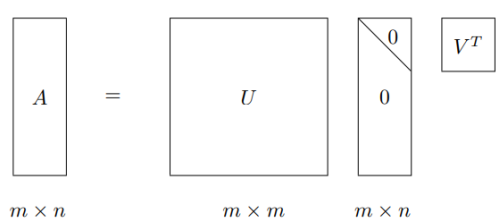
$$\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n),$$

که در آن:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0.$$

ستون‌های U و V را به ترتیب بردارهای تکین چپ و راست می‌نامند و عناصر قطری σ_i را مقادیر تکین مثبت می‌نامند.

تجزیه مقادیر منفرد را بصورت نمادین در شکل زیر نمایش می دهیم.



شکل ۱: نمایش SVD

[1] Singular Value Decomposition (SVD)
[2] Singular Values
[3] Principal Component Analysis (PCA)
[4] Least-Squares Problems

۲.۲ محاسبه ی تجزیه مقادیر منفرد

در اینجا نحوه محاسبه SVD را بیان می کنیم. (برابر با تعداد مقادیر تکین مثبت است (r

گام ۱: ماتریس A^TA را تشکیل دهید و جذر مقادیر ویژه‌ی آن را با $\sigma_1 = \sqrt{\lambda_1}, \ldots, \sigma_r = \sqrt{\lambda_r}, \sigma_{r+1} = \cdots = \sigma_n = 0$ قرار می‌دهیم (مقادیر تکین به صورت نزولی در ماتریس Σ قرار گرفته‌اند):

$$\Sigma_1 = \begin{bmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_r \end{bmatrix}_{r\times r} \quad , \quad \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}_{m\times n}$$

گام ۲: بردارهای ویژه‌ی یکه شده A^TA را با $v_1, v_2, \ldots, v_n$ نمایش داده و ماتریس V را به صورت زیر تشکیل دهید:

$$V = [v_1, v_2, \ldots, v_n]$$

گام ۳: برای $i = 1, 2, \ldots, r$، بردارهای u_i را به صورت زیر می‌سازیم:

$$u_i = \frac{1}{\sigma_i} A v_i$$

گام ۴: بردارهای $u_{r+1}, u_{r+2}, \ldots, u_m$ را طوری بدست می آوریم که ماتریس U متعامد باشد.

$$U = [u_1, u_2, \ldots, u_r, u_{r+1}, u_{r+2}, \ldots, u_m]$$

گام ۵: ماتریس‌های V, Σ, U تشکیل شده‌اند، لذا تجزیه‌ی SVD ماتریس A به صورت زیر خواهد بود:

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

برای اشاره به مفید بودن تجزیه، یک مسئله مینیمم سازی را در نظر بگیرید:

$$\min_X \|F - X\|_2$$

که در آن $\text{rank}(X) = r$ و $r < n = \text{rank}(F)$. حل این مسئله بهترین ماتریس رتبه r را که تقریباً ماتریس F را برآورد می‌کند، به دست می‌دهد. تجزیه مقادیر منفرد (SVD) راه‌حل را به صورت زیر ارائه می‌دهد:

$$X = U \cdot \Sigma_r \cdot V^H$$

که در آن Σ_r فقط شامل r مقدار منفرد اول از Σ اصلی است.

مثال ۱.۲: تجزیه مقادیر منفرد (SVD)

برای درک بهتر الگوریتم SVD به مثال زیر توجه کنید. فرض کنید ماتریس زیر داده شده است:

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ 0 & 3 \end{bmatrix}$$

گام ۱: تشکیل A^TA و AA^T

ابتدا ماتریس‌های A^TA و AA^T را محاسبه می‌کنیم:

$$A^TA = \begin{bmatrix} 5 & 1 \\ 1 & 11 \end{bmatrix}, \quad AA^T = \begin{bmatrix} 2 & 1 & -3 \\ 1 & 5 & 3 \\ -3 & 3 & 9 \end{bmatrix}$$

گام ۲: یافتن مقادیر و بردارهای ویژه AA^T

برای AA^T مقادیر ویژه (λ) و بردارهای ویژه (u) به صورت زیر محاسبه می‌شوند:

$$\lambda_1 \approx 11.1623, \quad u_1 = \begin{bmatrix} -0.2475 \\ 0.3913 \\ 0.8863 \end{bmatrix}$$

$$\lambda_2 \approx 4.8377, \quad u_2 = \begin{bmatrix} 0.5216 \\ 0.8247 \\ -0.2185 \end{bmatrix}$$

$$\lambda_3 \approx 0, \quad u_3 = \begin{bmatrix} -0.8165 \\ 0 \\ -0.4082 \end{bmatrix}$$

پس ماتریس U به صورت زیر خواهد بود:

$$U = \begin{bmatrix} -0.2475 & 0.5216 & -0.8165 \\ 0.3913 & 0.8247 & 0 \\ 0.8863 & -0.2185 & -0.4082 \end{bmatrix}$$

گام ۳: یافتن مقادیر و بردارهای ویژه A^TA

برای A^TA مقادیر و بردارهای ویژه به صورت زیر هستند:

$$\lambda_1 \approx 11.1623, \quad v_1 = \begin{bmatrix} 0.1602 \\ 0.9871 \end{bmatrix}$$

$$\lambda_2 \approx 4.8377, \quad v_2 = \begin{bmatrix} -0.9871 \\ 0.1602 \end{bmatrix}$$

پس ماتریس V به صورت زیر است:

$$V = \begin{bmatrix} 0.1602 & -0.9871 \\ 0.9871 & 0.1602 \end{bmatrix}$$

گام ۴: محاسبه مقادیر منفرد و تشکیل Σ

مقادیر منفرد ماتریس A با جذر مقادیر ویژه A^TA به دست می‌آیند:

$$\sigma_1 = \sqrt{\lambda_1} \approx 3.3410, \quad \sigma_2 = \sqrt{\lambda_2} \approx 2.1995$$

ماتریس قطری Σ به صورت زیر است:

$$\Sigma = \begin{bmatrix} 3.3410 & 0 \\ 0 & 2.1995 \\ 0 & 0 \end{bmatrix}$$

گام ۵: بازسازی A با $U\Sigma V^T$

در نهایت، ماتریس A با ضرب U، Σ و V^T به صورت زیر بازسازی می‌شود:

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

$$\begin{bmatrix} 1 & -1 \\ 2 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} -0.2475 & 0.5216 & -0.8165 \\ 0.3913 & 0.8247 & 0 \\ 0.8863 & -0.2185 & -0.4082 \end{bmatrix} \begin{bmatrix} 3.3410 & 0 & 0 \\ 0 & 2.1995 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.1602 & 0.9871 \\ -0.9871 & 0.1602 \end{bmatrix}^T$$

$$A = U \cdot \Sigma \cdot V^T$$

ماتریس های پایه ی عمود

ماتریس F را می‌توان به صورت مجموع ماتریس‌های رتبه یک نوشت:

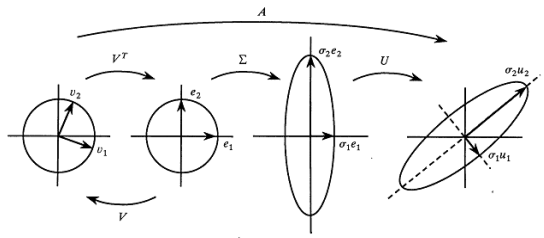
$$F = U \cdot \Sigma \cdot V^T = \sum_{i=1}^{n} \sigma_i u_i v_i^T$$

که در آن σ_i مقادیر منفرد هستند و u_i و v_i به ترتیب ستون‌های ماتریس‌های U و V هستند.

شهود هندسی

در تصویر زیر می توان شهود هندسی تجزیه مقادیر منفرد را مشاهده نمود.

چرخش یا بازتاب اولیه (V^T)
این ماتریس، فضای ورودی را به‌گونه‌ای تغییر می‌دهد که محورهای مختصات جدید با جهت‌های خاص ماتریس A هماهنگ شوند.
مقیاس‌گذاری (Σ)
در این مرحله، هر محور جدید با یک مقدار مشخص (مقدار منفرد) مقیاس می‌شود. این مقادیر نشان‌دهنده میزان کشیدگی یا فشردگی در آن جهت هستند.



شکل ۲: شهود هندسی SVD

چرخش یا بازتاب نهایی (U)
این ماتریس، فضای تغییر یافته را دوباره چرخانده یا بازتاب می‌دهد تا به فضای خروجی منتقل شود.

مثال فرض کنید A یک ماتریس 2×2 باشد. این ماتریس می‌تواند یک بیضی را از یک دایره واحد ایجاد کند:

- V^T دایره را در فضای ورودی به‌گونه‌ای می‌چرخاند که محورهای بیضی با محورهای مختصات هماهنگ شوند.
- Σ شعاع‌های دایره را در جهت‌های مختلف کشیده یا فشرده می‌کند، به طوری که یک بیضی حاصل می‌شود.
- U این بیضی را به فضای خروجی می‌چرخاند.

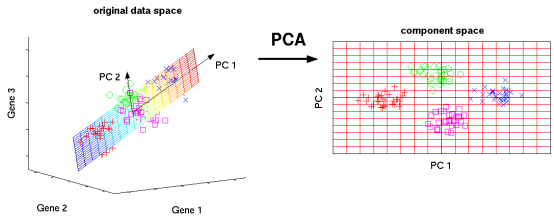
به طور خلاصه، SVD نشان می‌دهد که هر تبدیل خطی می‌تواند به سه مرحله هندسی قابل فهم تقسیم شود: چرخش، مقیاس‌گذاری، و چرخش مجدد.

۳.۲ کاربرد های تجزیه مقادیر منفرد

تجزیه مقادیر منفرد یکی از ابزارهای قدرتمند در جبر خطی است که کاربردهای گسترده‌ای در حوزه‌های مختلف دارد. در اینجا پنج کاربرد متداول آن را نام می‌بریم:

۱. کاهش ابعاد داده

در مجموعه‌های داده با ابعاد بالا، تحلیل و پردازش داده‌ها می‌تواند پیچیده و زمان‌بر باشد. SVD با کاهش ابعاد داده‌ها به فضای با بعد کمتر کمک می‌کند، به طوری که اطلاعات اصلی حفظ می‌شوند. این روش به ویژه در تحلیل مولفه‌های اصلی (PCA) به کار می‌رود. کاهش ابعاد به بهبود عملکرد الگوریتم‌های یادگیری ماشین و کاهش هزینه‌های محاسباتی منجر می‌شود. به عنوان مثال، در داده‌های تصویری، می‌توان تعداد پیکسل‌ها را بدون از دست دادن اطلاعات مهم کاهش داد.



شکل ۳: تحلیل مولفه‌های اصلی

۲. فشرده‌سازی تصویر

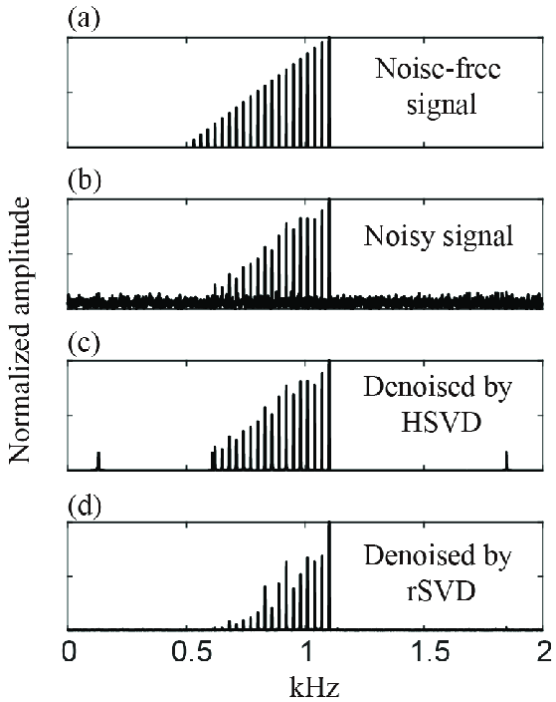
SVD برای فشرده‌سازی تصاویر به کار می‌رود. در این روش، ماتریس تصویر با استفاده از مقادیر منفرد بزرگ‌تر بازسازی می‌شود. با انتخاب تعداد محدودی از این مقادیر، می‌توان حجم داده‌ها را کاهش داد و تصویر را فشرده کرد. این فرآیند باعث صرفه‌جویی در فضای ذخیره‌سازی و پهنای باند انتقال می‌شود. اگرچه برخی از جزئیات تصویر از دست می‌روند، اما برای بسیاری از کاربردها، این کاهش کیفیت قابل قبول است.

۳. حل مسائل معکوس و سیستم‌های خطی

در بسیاری از مسائل علمی و مهندسی، باید معادلات خطی را حل کرد. اما هنگامی که این معادلات به صورت بدوضع (ill-posed) باشند، روش‌های معمول ممکن است دقت کافی نداشته باشند. SVD یک راه حل پایدار برای چنین مسائل فراهم می‌کند. با استفاده از SVD، می‌توان به طور موثری سیستم‌های معادلات خطی با ماتریس‌های نامتعادل یا بزرگ را حل کرد. همچنین، این روش برای یافتن تقریب بهترین جواب در مسائل معکوس، جایی که راه‌حل دقیقی وجود ندارد، مفید است.

۴. کاهش نویز

در پردازش سیگنال و تصویر، کاهش نویز یکی از چالش‌های مهم است. SVD می‌تواند با حذف مقادیر منفرد کوچک‌تر که معمولاً نشان‌دهنده نویز هستند، به کاهش نویز کمک کند. این روش به طور موثر ساختارهای اصلی داده را حفظ کرده و نویز را حذف می‌کند. در پردازش تصاویر، این فرآیند می‌تواند به تصاویر واضح‌تر و سیگنال‌های با کیفیت‌تر منجر شود.

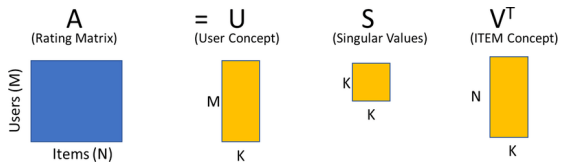


شکل ۴: کاهش نویز

۵. توصیه‌گرها

سیستم‌های توصیه‌گر برای پیشنهاد آیتم‌هایی مانند فیلم، کتاب، یا محصولات به کاربران استفاده می‌شوند. SVD به عنوان یکی از تکنیک‌های اصلی در این سیستم‌ها، با تجزیه ماتریس تعاملات کاربر–آیتم، الگوهای پنهان را استخراج می‌کند. با کاهش رتبه ماتریس و فاکتورگیری، سیستم می‌تواند پیش‌بینی‌های دقیقی برای علایق کاربران ارائه دهد. این روش به ویژه در سیستم‌های بزرگی که تعداد کاربران و آیتم‌ها زیاد است، به کار می‌رود و به بهبود دقت و کارایی سیستم‌های توصیه‌گر کمک می‌کند.

به طور خلاصه، SVD به دلیل توانایی خود در تحلیل و پردازش داده‌ها به روشی بهینه و موثر، در بسیاری از کاربردهای علمی و صنعتی به کار گرفته می‌شود.



شکل ۵: تحلیل مولفه‌های اصلی

۳ دسته بندی ارقام دست نویس

۱.۳ شرح مساله

طبقه‌بندی فرآیندی است که طی آن داده‌ها یا اشیاء بر اساس ویژگی‌ها یا خصوصیاتشان به کلاس‌ها یا دسته‌های از پیش تعریف‌شده دسته‌بندی می‌شوند. مسئله‌ی طبقه‌بندی کاراکترها در این زمینه به عنوان یکی از مسائل بنیادی در نظر گرفته شده و الگوریتم‌های طبقه‌بندی متعددی وجود دارند که در حوزه‌های خاص به صورت کم و بیش عملکرد مناسبی دارند.

راه‌های مختلفی برای تقسیم‌بندی این مسئله به دسته‌بندی‌های اصولاً متفاوت وجود دارد. از این رو، طبیعی است که مناسب‌ترین روش‌ها برای حل مسئله در شرایط مختلف ممکن است متفاوت باشند. بسیاری از الگوریتم‌ها بر اساس نظریه‌های ریاضی مختلف ابداع شده‌اند. جبر خطی، تحلیل تابعی و آمار تنها چند مورد از این نظریه‌ها هستند.

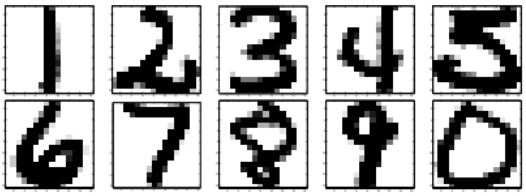
یکی از راه‌های تمایز بین الگوریتم‌ها، بررسی نحوه‌ی نمایش کاراکترها است. به عنوان مثال، وقتی با قلم می‌نویسیم، یک نماد شکلی بر روی کاغذ ثبت می‌شود. این شکل را می‌توان حداقل به دو روش نمایش داد - به صورت یک تصویر پیکسلی یا به صورت منحنی‌هایی در یک صفحه. بنابراین، ساختار داده‌های ورودی تأثیر عمده‌ای بر نحوه‌ی ساخت الگوریتم دارد. ممکن است در این مورد نمایش منحنی را ترجیح دهیم زیرا شکل‌های اصلی که قصد طبقه‌بندی آن‌ها را داریم، منحنی هستند و با حرکت منحنی‌وار دست نوشته می‌شوند و نه به صورت نقاط یا پیکسل‌هایی با شدت‌های مختلف. با داشتن مجموعه‌ای از اعداد دست‌نویس که به‌صورت دستی طبقه‌بندی شده‌اند (مجموعه آموزش)، هدف طبقه‌بندی مجموعه‌ای از اعداد ناشناخته (مجموعه تست) است.

۲.۳ تنسور

ساختارها یا مقادیر ریاضی مانند بردارها و ماتریس‌ها گاهی اوقات برای توصیف ماهیت واقعی داده‌ها در زمینه‌های مختلف پردازش سیگنال مناسب نیستند. این ساختارها به‌راحتی می‌توانند به ساختارهای عمومی‌تر و با مرتبه بالاتر گسترش یابند که به آن‌ها آرایه‌های چندبعدی، ماتریس‌های چندبعدی یا، همان‌طور که در این گزارش به آن‌ها اشاره خواهد شد، به سادگی تنسورها گفته می‌شود. بردارها و ماتریس‌ها به ترتیب معادل تنسورهای مرتبه اول و دوم هستند. الگوریتم ساخته‌شده عمدتاً بر اساس حالت خاصی از این نظریه برای تنسورهای مرتبه سوم است. بنابراین، بیشتر نظریه‌های ارائه‌شده در این بخش به تنسورهای مرتبه سوم مربوط می‌شود.

۳.۳ دیتاست

دیتاستی که در این مقاله مورد بحث می باشد [۳]ارقام دست نویس از ۰ تا ۹ است که شامل ۷۲۹۱ رقم است که هرکدام یک عکس خاکستری ۱۶ در ۱۶ پیکسلی است. مقادیر پیکسل ها نرمال سازی شده اند تا در بازه ی [۱، ۱-] باشند.



شکل ۶: اعداد دست‌نویس از پایگاه داده خدمات پستی ایالات متحده

ما ارقام را به سه فرمت مختلف اما معادل بررسی خواهیم کرد:

۱. به عنوان تصاویر خاکستری مقیاس 16×16،

۲. به عنوان توابعی از دو متغیر، $s = s(x, y)$ ؛

۳. به عنوان بردارهایی در R^{256}.

برای طبقه‌بندی یک عدد ناشناخته، نیاز داریم فاصله آن را با اعداد شناخته‌شده محاسبه کنیم. انواع مختلفی از معیارهای فاصله را می‌توان استفاده کرد، اما شاید طبیعی‌ترین آن‌ها فاصله اقلیدسی [۵] .باشد ستون‌های تصویر را در یک بردار قرار دهید و هر رقم را به عنوان یک بردار در R^{256} شناسایی کنید. سپس تابع فاصله را به صورت زیر تعریف کنید:

$$(x, y) = \|x - y\|_2.$$

ساختن یک پایه [۶]

برای ساختن یک پایه و الگوریتمی که هدف آن تشخیص یک رقم ناشناخته به یکی از کلاس‌های از پیش تعریف شده است، باید اطلاعاتی در مورد هر کلاس در الگوریتم گنجانده شود. این اطلاعات سپس در فرآیند طبقه‌بندی مورد استفاده قرار می‌گیرند. در مورد خاص ما، الگوریتم باید شامل اطلاعاتی درباره هر نوع رقم باشد و این اطلاعات از مجموعه آموزش استخراج می‌شود.

بردارسازی تصویر:

ستون‌های تصویر 16×16 را روی هم قرار دهید تا هر تصویر به یک بردار در فضای $\mathbb{R}^{256}$ تبدیل شود. هر عدد به شکل برداری $\mathbf{x} \in \mathbb{R}^{256}$ نمایش داده می‌شود.

[5] Euclidean Distance
[6] Basis

تابع فاصله:

برای استفاده از فاصله اقلیدسی از رابطه زیر استفاده می کنیم.

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2 = \sqrt{\sum_{i=1}^{256} (x_i - y_i)^2}$$

تابع فاصله جایگزین:

به‌جای فاصله اقلیدسی، می‌توان از **فاصله کسینوسی** استفاده کرد. این معیار زاویه بین دو بردار را می‌سنجد و برای داده‌های نرمال‌شده مناسب است:

$$d_{\text{cosine}}(\mathbf{x}, \mathbf{y}) = 1 - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}$$

این تابع مقدار ۰ را نشان می‌دهد اگر بردارها کاملاً هم‌جهت باشند، و مقدار ۱ اگر کاملاً نامرتبط باشند.

اگر مجموعه آموزش اعداد را به‌صورت بردارها یا نقاط در نظر بگیریم، منطقی است که فرض کنیم تمام اعداد از یک نوع خاص، یک **خوشه** از نقاط را در فضای برداری اقلیدسی ۲۵۶ ـ بعدی تشکیل می‌دهند.

در حالت ایده‌آل، این خوشه‌ها باید به‌خوبی از یکدیگر جدا باشند (در غیر این صورت، کار طبقه‌بندی اعداد ناشناخته بسیار دشوار خواهد بود).

۴.۳ یک الگوریتم ساده برای طبقه‌بندی

آموزش:

با داشتن مجموعه آموزشی که به‌صورت دستی طبقه‌بندی شده است، میانگین‌ها (یا مراکز خوشه‌ها) $\mathbf{m}_i$، برای $i = 0, \ldots, 9$ را برای تمام ۱۰ کلاس محاسبه کنید.

طبقه‌بندی:

برای هر عدد در مجموعه تست، آن را به کلاس k نسبت دهید اگر $\mathbf{m}_k$ نزدیک‌ترین میانگین باشد.

دلیل عملکرد نسبتاً ضعیف این الگوریتم این است که هیچ اطلاعاتی درباره تغییرات درون هر کلاس از اعداد را در نظر نمی‌گیرد.

۵.۳ دسته بندی داده ها با پایه های SVD

در این بخش، یک الگوریتم طبقه‌بندی را توضیح خواهیم داد که بر اساس مدل‌سازی تغییرات درون هر کلاس عددی با استفاده از بردارهای پایه عمود بر هم که با استفاده از تجزیه مقادیر منفرد محاسبه می‌شوند، عمل می‌کند.

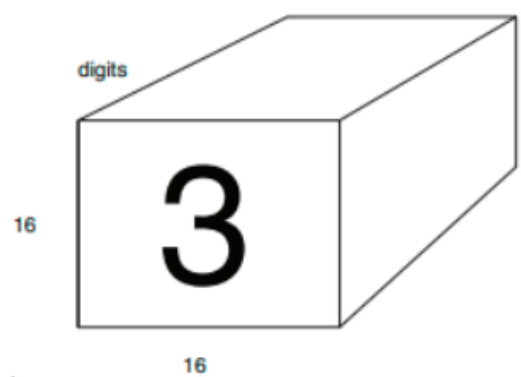
فرض کنید قصد داریم تغییرات درون یک کلاس خاص از تصاویر دست‌نویس (مانند عدد ۳) را مدل‌سازی کنیم. هر تصویر یک ماتریس 16×16 است که نمایانگر پیکسل‌های تصویر است. هدف ما این است که تفاوت‌های موجود در دست‌نویس‌های مختلف عدد ۳ را شناسایی کرده و مدل کنیم.

ساخت ماتریس A

ابتدا تمام تصاویر عدد ۳ را که در اختیار داریم، به صورت ستون به ستون در یک ماتریس بزرگ A قرار می‌دهیم. هر ستون از این ماتریس نمایانگر یک تصویر از عدد ۳ است. فرض کنید تعداد کل تصاویر N عدد ۳ در اختیار داریم. از آنجا که هر تصویر یک ماتریس 16×16 است، هر تصویر به یک بردار ستونی 256×1 تبدیل می‌شود. بنابراین، ماتریس A که شامل تمام این تصاویر است، به ابعاد $256 \times N$ خواهد بود، به این صورت:

$$A = \begin{bmatrix} | & | & | & \dots & | \\ a_1 & a_2 & a_3 & \dots & a_N \\ | & | & | & \dots & | \end{bmatrix}$$

در اینجا، a_i نمایانگر یک تصویر از عدد ۳ است که به یک بردار ستونی تبدیل شده.



شکل ۷: مجموعه ای از ارقام همنوع تشکیل یک تنسور را می دهند.

برای درک این فرایند به عکس بالا دقت کنید.

فضای زیرفضای خطی

ستون‌های ماتریس A یک زیرفضای خطی در فضای $\mathbb{R}^{256}$ را می‌سازند. به عبارت دیگر، این ستون‌ها نشان‌دهنده‌ی مجموعه‌ای از الگوهای اصلی و ساختارهای موجود در دست‌نویس‌های مختلف عدد ۳ هستند. این زیرفضای خطی معمولاً کوچکتر از فضای کل خواهد بود، چرا که داده‌ها (تصاویر) ساختارمند و وابسته به یکدیگر هستند.

استفاده از تجزیه مقادیر منفرد (SVD)

با استفاده از تجزیه مقادیر منفرد ،(SVD) می‌توان ماتریس A را به صورت جمعی از ماتریس‌های رتبه‌ـ ۱ نوشت:

$$A = \sum_{i=1}^{m} \sigma_i u_i v_i^T$$

در اینجا:

- u_i: بردارهای منفرد چپ که پایه‌های زیرفضای خطی هستند. این بردارها الگوهای اصلی تغییرات رقم را مدل می کند و به آن تصاویر منفرد گفته می شود.
- σ_i: مقادیر منفرد که میزان اهمیت هر بردار منفرد را نشان می‌دهند. درواقع نشان دهنده میزان اهمیت هر الگو می باشد.
- v_i^T: بردارهای منفرد راست که نشان‌دهنده‌ی این هستند که هر تصویر چگونه با این الگوها ترکیب می شود.

مختصات تصویر j در ماتریس A بر اساس این پایه برابر با $\sigma_i v_{ij}$ است. با توجه به خواص تقریب ماتریس در تجزیه SVD ، می‌دانیم که اولین بردار منفرد u_1 نمایانگر «جهت غالب» در ماتریس داده است. بنابراین، اگر بردارهای u_i را به تصاویر بازگردانیم، انتظار داریم که اولین بردار منفرد شبیه به یک رقم ۳ باشد، و تصاویر منفرد بعدی تغییرات غالب مجموعه آموزشی را حول اولین تصویر منفرد نشان دهند.

مدل‌سازی تغییرات درون کلاس

بردارهای u_i به عنوان پایه‌های زیرفضای مربوط به کلاس خاصی از اعداد (مثلاً عدد ۳) در نظر گرفته می‌شوند. این بردارها به ما کمک می‌کنند تا تغییرات درون یک کلاس خاص (عدد ۳) را مدل کنیم. به عبارت دیگر، این روش تفاوت‌ها و تنوع‌های موجود در دست‌نویس‌های مختلف عدد ۳ را به صورت مؤثری نشان می‌دهد.

کاربرد در تشخیص

فرض کنید که ماتریس‌های پایه برای تمام کلاس‌ها ساخته شده و در دسترس هستند. در این صورت، کدام مجموعه از پایه‌ها، یک رقم ناشناخته را توصیف می‌کند؟

برای طبقه‌بندی یک عدد ناشناخته، باید بررسی کنیم که چگونه این عدد می‌تواند در ۱۰ پایه مختلف (مربوط به ارقام مختلف) نمایش داده شود. این کار از طریق محاسبه بردار باقی‌مانده در مسئله کمترین مربعات به صورت زیر انجام می‌شود:

$$\min_{\alpha_i} \|z - \sum_{i=1}^{k} \alpha_i u_i\|_2,$$

که در آن:

- z: عدد ناشناخته است.
- u_i: تصاویر منفرد مربوط به هر کلاس هستند.
- α_i: ضرایب ترکیب خطی تصاویر منفرد.

می‌توان این مسئله را به صورت ماتریسی بازنویسی کرد:

$$\min_{\alpha} \|z - U_k\alpha\|_2,$$

که در آن:

$$U_k = [u_1\ u_2\ \cdots\ u_k].$$

از آنجا که ستون‌های ماتریس U_k متعامد هستند، راه‌حل این مسئله به صورت زیر داده می‌شود:

$$\alpha = U_k^T z.$$

نرم بردار باقی‌مانده مسئله کمترین مربعات برابر است با:

$$\|(I - U_kU_k^T)z\|_2,$$

که این مقدار بیانگر نرم تصویر عدد ناشناخته z بر زیرفضای متعامد با $\text{span}(U_k)$ است.

بردار باقی‌مانده نشان‌دهنده مقدار خطایی است که هنگام نمایش عدد ناشناخته در زیرفضای $\text{span}(U_k)$ ایجاد می‌شود. برای هر کدام از ۱۰ کلاس (ارقام ۰ تا ۹)، این مقدار محاسبه شده و کلاس با کمترین مقدار بردار باقی‌مانده به عنوان پیش‌بینی طبقه‌بندی عدد ناشناخته انتخاب می‌شود.

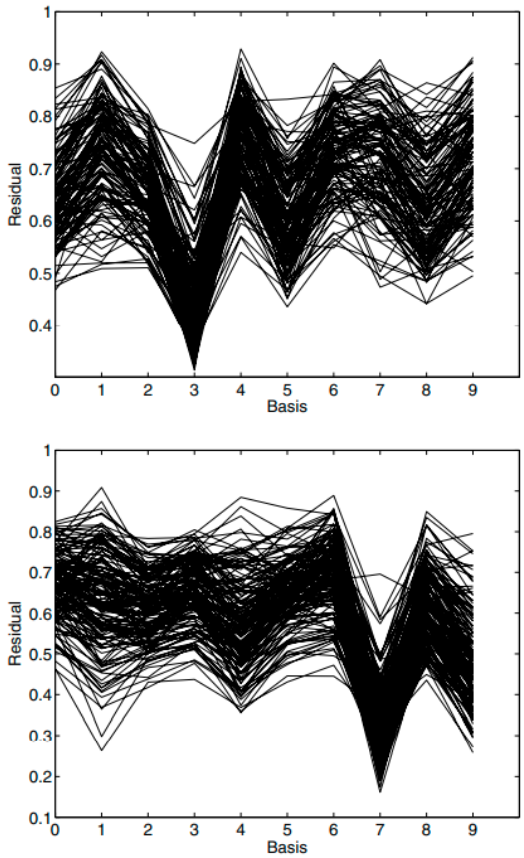
در واقع، این روش کمک می‌کند تا تشخیص دقیق‌تری از اعداد مختلف داشته باشیم، چرا که نه تنها شکل کلی عدد، بلکه تفاوت‌های جزئی در دست‌نویس‌های مختلف را نیز در نظر می‌گیرد.

برای نشان دادن اینکه فرضیات قبلی منطقی هستند، در شکل ۴ نُرم باقی‌مانده نسبی برای تمامی اعداد ناشناخته ۳ و ۷ در ارتباط با ۱۰ پایه مختلف نمایش داده شده است. در این نمودارها، برای هر عدد ناشناخته یک منحنی رسم شده است. به طور طبیعی، مشاهده جزئیات تک‌تک منحنی‌ها ممکن نیست، اما می‌توان دید که بیشتر اعداد ناشناخته ۳ و ۷ به بهترین شکل ممکن در پایه مربوط به خود نمایش داده شده‌اند.

این نمودارها همچنین اطلاعاتی درباره خطاهای احتمالی طبقه‌بندی ارائه می‌دهند. به عنوان مثال، اعداد ۳ و ۵ به یکدیگر شباهت بیشتری دارند، در حالی که اعداد ۳ و ۴ تفاوت‌های قابل توجهی دارند.

الگوریتم طبقه‌بندی مبتنی بر پایه‌های SVD

با الگوریتم های متفاوتی می توان این دسته بندی را پیاده سازی نمود. در زیر یک مدل ساده از آن را آوردیم.



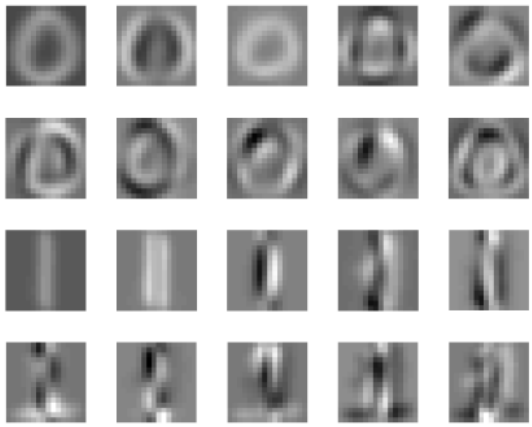
شکل ۸: نمودار نُرم باقی‌مانده نسبی برای اعداد ناشناخته ۳ و ۷ در ارتباط با ۱۰ پایه مختلف.

Algorithm ۱ الگوریتم طبقه‌بندی مبتنی بر پایه‌های SVD

آموزش:

for هر کلاس از ارقام شناخته‌شده **do**
محاسبه تجزیه مقادیر منفرد (SVD) برای ماتریس تصاویر ارقام آن کلاس
طبقه‌بندی:

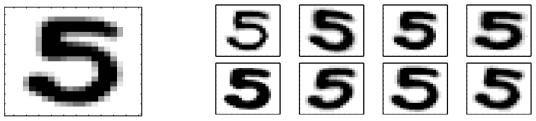
for عدد آزمون ناشناخته **do**
محاسبه باقی‌مانده نسبی در هر یک از ۱۰ پایه
باقی‌مانده یکی از پایه‌ها به طور معناداری کوچک‌تر از سایر پایه‌ها باشد آنگاه عدد آزمون به کلاس متناظر تخصیص یابد.



شکل ۹: U۱۰ برای ارقام صفر و یک

۴ فاصله تانژانت

یک الگوریتم طبقه‌بندی خوب باید قادر باشد ارقام ناشناخته‌ای را که به‌طور نسبتاً خوبی نوشته شده‌اند اما هنوز هم انحرافات قابل توجهی از رقم ایده‌آل در فاصله اقلیدسی دارند، طبقه‌بندی کند. برخی از این انحرافات وجود دارند که انسان‌ها می‌توانند به راحتی از پس آن‌ها برآیند و در واقع این نوع انحرافات کاملاً رایج و قابل قبول هستند. در شکل، چند نمونه از این تغییرات را نشان داده‌ایم.



شکل ۱۰: نمونه های دست نوشته

این نوع تغییرات برای یک خواننده انسانی مشکلی ایجاد نمی‌کند و در حالت ایده‌آل باید به راحتی در شناسایی خودکار ارقام نیز قابل مدیریت باشند. به عبارت دیگر، الگوریتم‌های شناسایی ارقام باید قادر باشند تغییرات کوچک و رایج در نحوه نوشتن ارقام را درک کرده و به درستی آن‌ها را شبیه‌سازی کنند، مشابه نحوه برخورد انسان‌ها با این تغییرات.

تصاویر 16×16 را می‌توان به عنوان نقاطی در $\mathbb{R}^{256}$ تفسیر کرد. فرض کنید p یک الگوی ثابت در یک تصویر باشد. ابتدا حالت فقط یک تغییر مجاز، یعنی جابجایی الگو (رقم) در جهت محور x، را در نظر می‌گیریم. این جابجایی را می‌توان به عنوان حرکت الگو در امتداد یک منحنی در $\mathbb{R}^{256}$ در نظر گرفت. فرض کنید منحنی با یک پارامتر حقیقی α پارامترسازی شود، به طوری که منحنی به شکل $s(p,\alpha)$ داده شده و به گونه‌ای باشد که $s(p,0) = p$. به طور کلی، منحنی غیرخطی است و می‌توان آن را با دو جمله اول در توسعه تیلور تقریب زد:

$$s(p,\alpha) = s(p,0) + \frac{ds}{d\alpha}(p,0)\alpha + O(\alpha^2) \approx p + t_p\alpha,$$

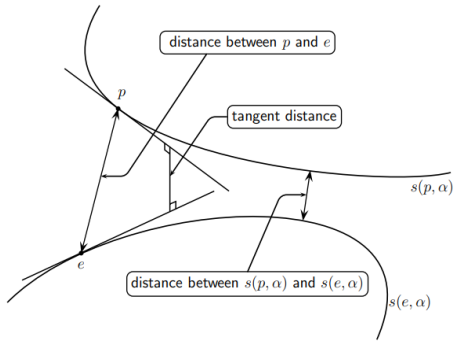
که در آن $t_p = \frac{ds}{d\alpha}(p,0)$ یک بردار در $\mathbb{R}^{256}$ است. با تغییر اندک α در اطراف ۰، یک حرکت کوچک از الگو در امتداد مماس در نقطه p روی منحنی انجام می‌دهیم.

حال فرض کنید که یک الگوی دیگر به نام e به طور مشابه تقریب زده شده باشد:

$$s(e,\alpha) \approx e + t_e\alpha.$$

از آنجا که ما جابجایی‌های کوچک در امتداد منحنی‌ها را مجاز می‌دانیم، این جابجایی‌های کوچک نباید بر روی تابع فاصله تأثیر بگذارند. بنابراین، به طور ایده‌آل، ما می‌خواهیم معیار نزدیکی بین p و e را به عنوان نزدیک‌ترین فاصله بین دو منحنی تعریف کنیم.

اما از آنجا که به طور کلی نمی‌توانیم فاصله بین منحنی‌ها را محاسبه کنیم، می‌توانیم از تقریب‌های مرتبه اول استفاده کرده و نزدیک‌ترین فاصله بین دو مماس را در نقاط p و e محاسبه کنیم. بنابراین، الگوها را به طور مستقل در امتداد مماس‌های خود حرکت می‌دهیم تا کوچک‌ترین



شکل ۱۱: فاصله بین نقاط p و e و فاصله تانژانت

فاصله را پیدا کنیم. اگر این فاصله را با استفاده از معیاری اقلیدسی معمولی اندازه‌گیری کنیم، مسأله کمینه مربعات به شکل زیر حل می‌شود:

$$\min_{\alpha_p,\alpha_e} \|p + t_p\alpha_p - e - t_e\alpha_e\|^2 = \min_{\alpha_p,\alpha_e} \left\|(p-e) - \begin{pmatrix} -t_p & t_e \end{pmatrix} \begin{pmatrix} \alpha_p \\ \alpha_e \end{pmatrix}\right\|^2 .$$

حال، فرض کنید که مجاز هستیم الگو p را در l منحنی مختلف در $\mathbb{R}^{256}$ حرکت دهیم که با پارامتر $\alpha = (\alpha_1, \ldots, \alpha_l)^T$ پارامترسازی شده‌اند. این معادل با حرکت الگو در یک سطح lـ بعدی در $\mathbb{R}^{256}$ است. فرض کنید دو الگوی p و e داریم که هر یک مجاز به حرکت در سطح خود از تغییرات هستند. ایده‌آل این است که ما نزدیک‌ترین فاصله بین سطوح را پیدا کنیم، اما از آنجا که این محاسبه غیرممکن است، اکنون یک معیار فاصله تعریف می‌کنیم، جایی که فاصله بین دو صفحه مماس در نقاط p و e را محاسبه می‌کنیم.

همانطور که قبل‌تر اشاره شد، صفحه مماس با دو جمله اول در توسعه تیلور تابع $s(p,\alpha)$ داده می‌شود:

$$s(p,\alpha) = s(p,0) + \sum_{i=1}^{l} \frac{ds}{d\alpha_i}(p,0)\alpha_i + O(\alpha^2) \approx p + T_p\alpha,$$

که در آن T_p ماتریس است:

$$T_p = \begin{pmatrix} \frac{ds}{d\alpha_1} & \frac{ds}{d\alpha_2} & \cdots & \frac{ds}{d\alpha_l} \end{pmatrix},$$

و مشتقات همگی در نقطه $(p,0)$ محاسبه می‌شوند.

بنابراین، فاصله مماس بین نقاط p و e به صورت کمترین باقی‌مانده در مسأله کمینه مربعات تعریف می‌شود:

$$\min_{\alpha_p,\alpha_e} \|p + T_p\alpha_p - e - T_e\alpha_e\|^2 = \min_{\alpha_p,\alpha_e} \left\|(p-e) - \begin{pmatrix} -T_p & T_e \end{pmatrix} \begin{pmatrix} \alpha_p \\ \alpha_e \end{pmatrix}\right\|^2 .$$

مسأله کمینه مربعات را می‌توان با استفاده از تجزیه مقدار منفرد (SVD) ماتریس $A = \begin{pmatrix} -T_p & T_e \end{pmatrix}$ حل کرد.

توجه داشته باشید که ما به دنبال حل خود مسأله نیستیم بلکه تنها به معیار باقی‌مانده علاقه داریم. مسأله کمینه مربعات را می‌توان به صورت زیر نوشت:

$$\min_{\alpha} \|b - A\alpha\|^2, \quad b = p - e, \quad \alpha = \begin{pmatrix} \alpha_p \\ \alpha_e \end{pmatrix}.$$

اگر از تجزیه QR استفاده کنیم:

$$A = Q \begin{pmatrix} R \\ 0 \end{pmatrix} = (Q_1\, Q_2) \begin{pmatrix} R \\ 0 \end{pmatrix} = Q_1 R,$$

معیار باقی‌مانده به شکل زیر است:

$$\min_{\alpha} \|b - A\alpha\|^2 = \min_{\alpha} \left\|Q_1^T b - R\alpha\right\|^2 + \|Q_2^T b\|^2 = \|Q_2^T b\|^2.$$

موردی که در آن ماتریس A ممکن است رتبه کامل نداشته باشد به راحتی با استفاده از SVD قابل حل است؛ برای جزئیات بیشتر به بخش ۷.۶ مراجعه کنید. احتمالاً وقتی دو الگو به یکدیگر نزدیک هستند، ستون‌های ماتریس مماس تقریباً وابسته خطی خواهند بود.

مهم‌ترین ویژگی این تابع فاصله این است که تحت جابجایی الگوها در صفحات مماس تغییر نخواهد کرد. به عنوان مثال، اگر یک جابجایی کوچک در جهت محور x از یک الگو انجام دهیم، با این معیار، فاصله‌ای که الگو جابجا شده برابر با صفر خواهد بود.

۱.۴ تبدیلات

در این بخش، الگوی تصویر را به‌عنوان یک تابع از دو متغیر $p = p(x, y)$ در نظر می‌گیریم و نشان می‌دهیم که مشتق هر تبدیل را می‌توان به‌صورت یک عملگر مشتق‌گیری که ترکیب خطی از مشتقات $p_x = \frac{\partial p}{\partial x}$ و $p_y = \frac{\partial p}{\partial y}$ است، بیان کرد.

انتقال

ساده‌ترین تبدیل، تبدیلی است که در آن الگو به میزان α_x در جهت x جابجا می‌شود، یعنی

$$s(p, \alpha_x)(x, y) = p(x + \alpha_x, y).$$

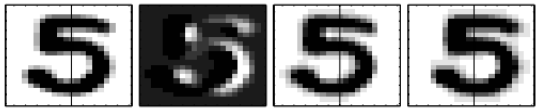
با استفاده از قانون زنجیره، داریم

$$\frac{d}{d\alpha_x} \left(s(p, \alpha_x)(x, y)\right)\Big|_{\alpha_x=0} = \frac{d}{d\alpha_x} p(x + \alpha_x, y)\Big|_{\alpha_x=0} = p_x(x, y).$$

در شکل، یک الگو و مشتق آن در جهت x را نشان می‌دهیم. سپس نشان می‌دهیم که با افزودن یک ضرب کوچک از مشتق، الگو می‌تواند به چپ یا راست منتقل شود.

به‌طور مشابه، برای ترجمه در جهت y، داریم

$$\frac{d}{d\alpha_y} \left(s(p, \alpha_y)(x, y)\right)\Big|_{\alpha_y=0} = p_y(x, y).$$



شکل ۱۲: الگو، مشتق آن در جهت x و ترجمه‌های x الگو

چرخش

چرخش. چرخش یک الگو به زاویه α_r با جایگزینی مقدار p در نقطه (x, y) با مقدار در نقطه

$$\begin{pmatrix} \cos\alpha_r & \sin\alpha_r \\ -\sin\alpha_r & \cos\alpha_r \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

انجام می‌شود. بنابراین ما تابع

$$s(p, \alpha_r)(x, y) = p(x\cos\alpha_r + y\sin\alpha_r, -x\sin\alpha_r + y\cos\alpha_r)$$

را تعریف می‌کنیم و مشتق آن را به صورت

$$\frac{d}{d\alpha_r}\left(s(p, \alpha_r)(x, y)\right) = (-x\sin\alpha_r + y\cos\alpha_r)p_x + (-x\cos\alpha_r - y\sin\alpha_r)p_y$$

به دست می‌آوریم. با قرار دادن $\alpha_r = 0$، داریم

$$\frac{d}{d\alpha_r}\left(s(p, \alpha_r)(x, y)\right)\Big|_{\alpha_r=0} = yp_x - xp_y$$

که در آن مشتقات در نقطه (x, y) ارزیابی می‌شوند.

شکل ۱۳: الگو، مشتق دورانی آن، و چرخش الگو

مقیاس بندی

یک مقیاس بندی از الگو با تعریف تابع زیر به دست می‌آید:

$$s(p, \alpha_s)(x, y) = p\left((1 + \alpha_s)x, (1 + \alpha_s)y\right),$$

و مشتق آن به صورت زیر است:

$$\frac{d}{d\alpha_s}\left(s(p, \alpha_s)(x, y)\right)\Big|_{\alpha_s=0} = x\frac{\partial p}{\partial x} + y\frac{\partial p}{\partial y}.$$

شکل ۱۴: الگو، مشتق اسکیلینگ آن و "اسکیلینگ به بالا" الگو

ضخیم‌سازی

الگو می‌تواند با استفاده از تکنیک‌های مشابه نازک‌تر یا ضخیم‌تر شود؛ برای جزئیات بیشتر، به [۸۷] مراجعه کنید. مشتق "ضخیم‌سازی" به صورت زیر است:

$$\left(\frac{\partial p}{\partial x}\right)^2 + \left(\frac{\partial p}{\partial y}\right)^2 .$$

Algorithm ۲ الگوریتم طبقه‌بندی مبتنی بر فاصله تانژانت

آموزش:

for برای هر رقم در مجموعه آموزشی **do**
ماتریس تانژانت Tp را محاسبه کنید
طبقه‌بندی:

for برای هر رقم آزمایشی **do**
ماتریس مماس آن را محاسبه کنید؛ فاصله مماس را با تمام ارقام آموزشی محاسبه کرده و رقم آزمایشی را به عنوان نزدیک‌ترین رقم آموزشی دسته‌بندی کنید.

۵ نتیجه‌گیری

در این پژوهش، تجزیه مقادیر منفرد (SVD) به‌عنوان یکی از روش‌های کلیدی برای بهبود پردازش داده‌ها مورد بررسی قرار گرفت. این روش با تجزیه یک ماتریس به سه مولفه‌ی اصلی U، Σ و V^T، امکان کاهش پیچیدگی داده‌ها و تمرکز بر اطلاعات مهم و ضروری را فراهم می‌کند. از جمله مزایای این روش می‌توان به حفظ ویژگی‌های اصلی داده، کاهش نویز و حذف اطلاعات کم‌اهمیت اشاره کرد.

در روند این تحقیق، کاربردهای SVD در تحلیل داده‌های پیچیده و بزرگ، خصوصاً در زمینه‌های یادگیری ماشین و هوش مصنوعی، برجسته شد. این تکنیک علاوه بر کاهش ابعاد، نقش کلیدی در افزایش دقت مدل‌های یادگیری، کاهش سربار محاسباتی و بهبود عملکرد کلی سیستم دارد. همچنین، استفاده از فاصله تانژانت به‌عنوان یک معیار در ارزیابی شباهت و تفکیک‌پذیری داده‌ها نشان داد که این متریک می‌تواند در بسیاری از سناریوها منجر به ارتقاء دقت طبقه‌بندی و تحلیل داده شود.

از دیگر دستاوردهای این پژوهش، امکان بازسازی دقیق ماتریس اصلی از مولفه‌های U، Σ و V^T بود که نشان‌دهنده قدرت بالای SVD در حفظ ساختار و اطلاعات داده است. علاوه بر این، تحلیل رفتار داده‌ها پس از کاهش ابعاد و ارزیابی اثرات نویز در داده‌های اصلی و کاهش‌یافته، اهمیت این روش را در کاربردهای متنوع از جمله بینایی ماشین، تحلیل متون، و پردازش سیگنال بیش از پیش نشان داد.

به طور کلی، نتایج این پژوهش بر اهمیت این روش و مزایای آن در پردازش داده‌ها تأکید دارد. این روش نه‌تنها موجب کاهش پیچیدگی داده‌ها می‌شود، بلکه با حفظ اطلاعات کلیدی، در بهبود عملکرد مدل‌های یادگیری و تحلیل داده‌ها نقش بسزایی ایفا می‌کند. آینده این روش در تلفیق با دیگر تکنیک‌های پیشرفته، مانند شبکه‌های عصبی و یادگیری عمیق، نویدبخش پیشرفت‌های چشمگیری در حوزه تحلیل داده و هوش مصنوعی است.

مراجع

[1] Eldén, L., Matrix Methods in Data Mining and Pattern Recognition

[2] Berkant Savas, Analyses and Tests of Handwritten Digit Recognition Algorithms

[3] http://www.gaussianprocess.org/gpml/data/