machine learning prerequisites workshop

Linear Algebra

Hamid R Rabiee – Zahra Dehghanian

Spring 2025

Sharif University of Technology

فضاهای برداری (Vector Spaces)

فضاهای برداری مجموعه‌ای هستند که عناصر آن **بردار** نامیده می‌شوند. در این مجموعه، دو عمل تعریف شده است:

- جمع بردارها
- ضرب بردارها در اعداد حقیقی که به آنها **اسکالر** گفته می‌شود.

یک فضای برداری باید شرایط زیر را داشته باشد:

۱. وجود یک عضو همانی جمع (عدد 0) به طوری که برای هر بردار x داشته باشیم: $x + 0 = x$
۲. برای هر بردار x وجود یک عکس جمعی $-x$ به طوری که $x + (-x) = 0$
۳. وجود عضو همانی ضرب (عدد 1) در مجموعه اعداد حقیقی که برای هر بردار x داشته باشیم: $1x = x$
۴. جابه‌جایی: $x + y = y + x$ برای همه‌ی x, y
۵. شرکت‌پذیری: $(x + y) + z = x + (y + z)$ و $\alpha(\beta x) = (\alpha\beta)x$ برای همه‌ی x, y, z و $\alpha, \beta \in \mathbb{R}$
۶. توزیع‌پذیری: $\alpha(x + y) = \alpha x + \alpha y$ و $(\alpha + \beta)x = \alpha x + \beta x$

استقلال خطی (Linear Independence)

مجموعه‌ای از بردارها $v_1, v_2, \ldots, v_n$ در فضای برداری V به صورت **خطی مستقل** تعریف می‌شوند اگر:

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0 \quad \Rightarrow \quad \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$$

یعنی ترکیب خطی آنها فقط زمانی صفر شود که همه ضرایب صفر باشند.

گستره (Span)

گستره‌ی یک مجموعه از بردارها $v_1, v_2, \ldots, v_n$ به صورت زیر تعریف می‌شود:

$$\text{span}\{v_1, \ldots, v_n\} = \{v \in V : \exists \alpha_1, \ldots, \alpha_n \text{ که } v = \alpha_1 v_1 + \cdots + \alpha_n v_n\}$$

مجموعه‌ای از بردارها که مستقل خطی هستند و گستره‌ی آنها تمام V را پوشش می‌دهد، **پایه** (Basis) نامیده می‌شود.

بعد (Dimension)

اگر یک فضای برداری توسط مجموعه‌ای متناهی از بردارها پوشش داده شود، **بعد متناهی** (Finite-Dimensional) نام دارد. در غیر این صورت، **بعد نامتناهی** (Infinite-Dimensional) دارد.

تعداد بردارهای یک پایه برای فضای برداری V، **بعد فضای برداری** نامیده می‌شود و به صورت $\dim V$ نمایش داده می‌شود.

فضای اقلیدسی یک فضای برداری است که با $\mathbb{R}^n$ نمایش داده می‌شود. بردارهای این فضا شامل nــتایی از اعداد حقیقی هستند:

$$\mathbf{x} = (x_1, x_2, \dots, x_n)$$

در بسیاری از موارد، می‌توان این بردارها را به عنوان ماتریس‌های $n \times 1$ یا **بردارهای ستونی** (Column Vectors) در نظر گرفت:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

جمع و ضرب اسکالر در $\mathbb{R}^n$

جمع دو بردار و ضرب اسکالر در بردارها به صورت مؤلفه‌ای تعریف می‌شود:

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}, \quad \alpha \mathbf{x} = \begin{bmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{bmatrix}$$

زیر فضاها (Subspaces)

فضاهای برداری می‌توانند شامل فضاهای برداری دیگر باشند. اگر V یک فضای برداری باشد، آنگاه $S \subseteq V$ یک **زیر فضا** از V است اگر:

۱. $0 \in S$
۲. S تحت جمع بسته باشد: اگر $\mathbf{x}, \mathbf{y} \in S$ باشد، آنگاه $\mathbf{x} + \mathbf{y} \in S$
۳. S تحت ضرب اسکالر بسته باشد: اگر $\mathbf{x} \in S$ و $\alpha \in \mathbb{R}$ باشد، آنگاه $\alpha\mathbf{x} \in S$

نکته: فضای V همواره یک زیر فضا از خودش است. همچنین، فضای برداری بدیهی که فقط شامل $\mathbf{0}$ باشد، یک زیر فضا محسوب می‌شود.

جمع مستقیم (Direct Sum)

اگر U و W دو زیر فضای V باشند، آنگاه **مجموع آنها** به صورت زیر تعریف می‌شود:

$$U + W = \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \in U, \mathbf{w} \in W\}$$

مجموعه‌ی بالا نیز یک زیر فضا از V است. اگر $U \cap W = \{0\}$، مجموع، **جمع مستقیم** نامیده شده و به صورت $U \oplus W$ نوشته می‌شود.

هر بردار در $U \oplus W$ را می‌توان به طور یکتا به صورت $\mathbf{u} + \mathbf{w}$ نوشت، که در آن:

$$\mathbf{u} \in U, \quad \mathbf{w} \in W$$

این یک شرط لازم و کافی برای جمع مستقیم است.

رابطه‌ی ابعاد زیر فضاها

ابعاد مجموع زیر فضاها از رابطه‌ی زیر پیروی می‌کند:

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$$

اگر جمع، مستقیم باشد، آنگاه:

$$\dim(U \oplus W) = \dim U + \dim W$$

زیرا:

$$\dim(U \cap W) = \dim(\{0\}) = 0$$

نگاشت‌های خطی (Linear Maps)

یک **نگاشت خطی** تابعی از V به W است که در آن V و W فضاهای برداری هستند و دو شرط زیر را برآورده می‌کند:

۱. $T(\mathbf{x}+\mathbf{y}) = T\mathbf{x} + T\mathbf{y} \quad \forall \mathbf{x}, \mathbf{y} \in V$
۲. $T(\alpha \mathbf{x}) = \alpha T\mathbf{x} \quad \forall \mathbf{x} \in V, \alpha \in \mathbb{R}$

در نگارش رایج، پرانتزهای غیرضروری حذف می‌شوند و به جای $T(\mathbf{x})$ از $T\mathbf{x}$ استفاده می‌شود، مشروط بر اینکه ابهامی ایجاد نشود.

عملگر خطی (Linear Operator)

یک نگاشت خطی که از V به خودش باشد، یک **عملگر خطی** نامیده می‌شود.

این تعریف ساختار فضاهای برداری را حفظ می‌کند، زیرا دو عمل اصلی این فضاها، یعنی جمع و ضرب اسکالر، را حفظ می‌کند.
در اصطلاحات جبری، یک نگاشت خطی را **همریختی فضاهای برداری (Homomorphism)** می‌نامند.

اگر یک همریختی خطی معکوس‌پذیر باشد، آن را **هم‌ریختی (Isomorphism)** می‌نامند.

همریختی (Isomorphism)

اگر همریختی‌ای از V به W وجود داشته باشد، آنگاه V و W را **همریخت** (ایزومورفیک) می‌نامیم و می‌نویسیم:

$$V \cong W$$

فضاهای برداری همریخت در واقع از نظر ساختار جبری "یکسان" هستند.
یک نتیجه جالب این است که فضاهای برداری **با بعد یکسان** همیشه همریخت هستند. اگر V و W فضاهای برداری حقیقی با ابعاد یکسان باشند ($\dim V = \dim W = n$)، آنگاه همریختی طبیعی زیر را داریم:

$$\varphi : V \to W$$

$$\alpha_1 v_1 + \cdots + \alpha_n v_n \mapsto \alpha_1 w_1 + \cdots + \alpha_n w_n$$

که در آن $v_1, \ldots, v_n$ و $w_1, \ldots, w_n$ پایه‌هایی برای V و W هستند.

نتیجه‌گیری

این هم‌ریختی **یک‌به‌یک** و **پوشا** است، زیرا هر بردار در V را می‌توان به طور یکتا به عنوان ترکیبی خطی از $v_1, \ldots, v_n$ نوشت.
بنابراین، به راحتی می‌توان تأیید کرد که φ یک هم‌ریختی است، و در نتیجه داریم:

$$V \cong W$$

به طور خاص، هر فضای برداری حقیقی nـبعدی با $\mathbb{R}^n$ هم‌ریخت است.

ماتریس یک نگاشت خطی

فضاهای برداری به صورت انتزاعی تعریف می‌شوند. برای نمایش و انجام عملیات بر روی بردارها و نگاشت‌های خطی در رایانه، معمولاً از آرایه‌های مستطیلی اعداد که به نام **ماتریس (Matrix)** شناخته می‌شوند، استفاده می‌کنیم.
فرض کنید که V و W فضاهای برداری متناهی‌بعدی با پایه‌های

$$v_1, \ldots, v_n \quad \text{و} \quad w_1, \ldots, w_m$$

باشند و $T : V \to W$ یک نگاشت خطی باشد. آنگاه ماتریس T که درایه‌های آن A_{ij} هستند و برای $i = 1, \ldots, m$ و $j = 1, \ldots, n$ تعریف می‌شود، به صورت زیر خواهد بود:

$$Tv_j = A_{1j}w_1 + \cdots + A_{mj}w_m$$

این بدان معناست که ستون jـام ماتریس A شامل مختصات Tv_j در پایه انتخاب شده برای W است.

القای نگاشت خطی توسط ماتریس

به طور عکس، هر ماتریس $A \in \mathbb{R}^{m\times n}$ یک نگاشت خطی را القا می‌کند:

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

که به صورت

$$T\mathbf{x} = A\mathbf{x}$$

تعریف می‌شود. ماتریس این نگاشت نسبت به پایه‌های استاندارد $\mathbb{R}^n$ و $\mathbb{R}^m$، همان ماتریس A است.

ترانهاده‌ی ماتریس (Transpose of a Matrix)

اگر $A \in \mathbb{R}^{m\times n}$ باشد، **ترانهاده‌ی (Transpose)** آن، که با $A^T \in \mathbb{R}^{n\times m}$ نشان داده می‌شود، به صورت زیر تعریف می‌شود:

$$(A^T)_{ij} = A_{ji} \quad \text{برای هر}\,(i,j)$$

به عبارت دیگر، ستون‌های A تبدیل به سطرهای A^T و سطرهای A تبدیل به ستون‌های A^T می‌شوند.

ویژگی‌های ترانهاده‌ی ماتریس

ترانهاده‌ی یک ماتریس دارای چندین خاصیت مهم و کاربردی است که به راحتی از تعریف آن نتیجه می‌شوند:

۱. $(A^T)^T = A$
۲. $(A+B)^T = A^T + B^T$
۳. $(\alpha A)^T = \alpha A^T$
۴. $(AB)^T = B^T A^T$

فضاهای نرم‌دار (Normed Spaces)

نرم (Norm)، تعمیم مفهوم طول از فضای اقلیدسی است.
یک **نرم** روی فضای برداری حقیقی V تابعی است به صورت:

$$\|\cdot\| : V \to \mathbb{R}$$

که ویژگی‌های زیر را دارد:

۱. $\|x\| \geq 0$، و برابری تنها زمانی برقرار است که $x = 0$
۲. $\|\alpha x\| = |\alpha|\|x\|$
۳. $\|x + y\| \leq \|x\| + \|y\|$ این خاصیت با نام **نامساوی مثلثی (Triangle Inequality)** شناخته می‌شود

برای همه‌ی $x, y \in V$ و همه‌ی $\alpha \in \mathbb{R}$.
یک فضای برداری که همراه با یک نرم مجهز شده باشد، به عنوان **فضای برداری نرم‌دار (Normed Vector Space)** شناخته می‌شود.

نرم اقلیدسی

$$\|x\|_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}$$

ضرب داخلی (Inner Product) روی یک فضای برداری حقیقی V، تابعی است:

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$$

که خواص زیر را داراست:

۱. $\langle x, x \rangle \geq 0$، و برابری تنها زمانی برقرار است که $x = 0$.
۲. خاصیت خطی در متغیر اول:

$$\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle, \quad \langle \alpha x, y \rangle = \alpha \langle x, y \rangle$$

۳. خاصیت متقارن بودن:

$$\langle x, y \rangle = \langle y, x \rangle$$

برای همه‌ی $x, y, z \in V$ و $\alpha \in \mathbb{R}$، یک فضای برداری که همراه با یک ضرب داخلی مجهز شده باشد، به عنوان **فضای ضرب داخلی (Inner Product Space)** شناخته می‌شود.
هر ضرب داخلی روی V، یک **نرم (Norm)** روی V القا می‌کند:

$$\|x\| = \sqrt{\langle x, x \rangle}$$

می‌توان نشان داد که اصول مربوط به نرم‌ها تحت این تعریف برقرار هستند. بنابراین، هر فضای ضرب داخلی، یک فضای نرم‌دار است (و در نتیجه یک فضای متریک نیز محسوب می‌شود).

دو بردار x و y در فضای V، **متعامد (Orthogonal)** نامیده می‌شوند، اگر:

$$\langle x, y \rangle = 0$$

که معمولاً به صورت $y \perp x$ نمایش داده می‌شود.
بردارهای متعامد یکه (Orthonormal)، بردارهایی هستند که علاوه بر متعامد بودن، طول واحد نیز دارند:

$$\|x\| = \|y\| = 1$$

ضرب داخلی استاندارد در $\mathbb{R}^n$

ضرب داخلی استاندارد در $\mathbb{R}^n$ به صورت زیر تعریف می‌شود:

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i = x^T y$$

در واقع، این ضرب داخلی یک حالت خاص از ضرب ماتریسی است که نتیجه‌ی آن یک ماتریس 1×1 در نظر گرفته می‌شود که مقدار آن به عنوان یک عدد اسکالر تفسیر می‌شود.

قضیه فیثاغورس (Pythagorean Theorem)

قضیه معروف فیثاغورس به صورت طبیعی به تمام فضاهای ضرب داخلی تعمیم می‌یابد.

قضیه ۱. اگر $y \perp x$، آنگاه:

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2$$

اثبات: فرض کنید $y \perp x$، یعنی $\langle x, y\rangle = 0$. بنابراین:

$$\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x, x\rangle + \langle y, x\rangle + \langle x, y\rangle + \langle y, y\rangle = \|x\|^2 + \|y\|^2$$

که اثبات را کامل می‌کند.

این نابرابری گاهی برای اثبات کران‌های مختلف مفید است:

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

برای تمام $x, y \in V$. برابری زمانی برقرار است که x و y مضارب عددی از یکدیگر باشند (یا معادل آن، وقتی که آن‌ها **وابسته خطی** (**linearly dependent**) باشند).

متمم متعامد و تصویرها (Orthogonal Complements and Projections)

اگر $S \subseteq V$ که در آن V یک فضای ضرب داخلی است، آنگاه **متمم متعامد (Orthogonal Complement)** مجموعه‌ای از تمام بردارهای V است که بر هر عضو از S متعامد هستند:

$$S^{\perp} = \{v \in V \mid v \perp s \text{ all for } s \in S\}$$

بردارهای ویژه و مقادیر ویژه

برای یک **ماتریس مربعی (square matrix)** $A \in \mathbb{R}^{n \times n}$، ممکن است بردارهایی وجود داشته باشند که هنگام اعمال A به آنها، فقط در یک مقدار ثابت ضرب شوند.

تعریف بردار ویژه (Eigenvector) و مقدار ویژه (Eigenvalue)

یک بردار غیر صفر $x \in \mathbb{R}^n$ **بردار ویژه (eigenvector)** از A با مقدار ویژه λ است اگر:

$$A\mathbf{x} = \lambda\mathbf{x}$$

بردار صفر در این تعریف لحاظ نمی‌شود، زیرا برای هر λ، داریم:

$$A\mathbf{0} = 0 = \lambda\mathbf{0}$$

کنید x یک **بردار ویژه (eigenvector)** از A با مقدار ویژه λ باشد. در این صورت:

1. برای هر $\gamma \in \mathbb{R}$، x یک بردار ویژه از $A + \gamma I$ با مقدار ویژه $\gamma + \lambda$ است.
2. اگر A **وارون‌پذیر (invertible)** باشد، آنگاه x یک بردار ویژه از A^{-1} با مقدار ویژه λ^{-1} است.
3. برای هر $k \in \mathbb{Z}$، داریم:
$$A^k \mathbf{x} = \lambda^k \mathbf{x}$$
(که در آن $A^0 = I$ بر اساس تعریف است).

ماتریس (Trace)

Trace یک ماتریس مربعی برابر با مجموع درایه‌های قطری آن است:

$$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} A_{ii}$$

Trace دارای خواص جبری زیر است:

۱. $\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$
۲. $\mathrm{tr}(\alpha \mathbf{A}) = \alpha \mathrm{tr}(\mathbf{A})$
۳. $\mathrm{tr}(\mathbf{A}^T) = \mathrm{tr}(\mathbf{A})$

ویژگی‌های بیشتر Trace

$$\operatorname{tr}(\mathbf{ABCD}) = \operatorname{tr}(\mathbf{BCDA}) = \operatorname{tr}(\mathbf{CDAB}) = \operatorname{tr}(\mathbf{DABC})$$

این خاصیت به عنوان **ناوردایی تحت جایگشت‌های چرخه‌ای (invariance under cyclic permutations)** شناخته می‌شود.

نکته: ماتریس‌ها را نمی‌توان به‌طور دلخواه جایگشت داد. به عنوان مثال، در حالت کلی داریم:

$$\operatorname{tr}(\mathbf{ABCD}) \neq \operatorname{tr}(\mathbf{BACD})$$

Trace و مقادیر ویژه

ویژگی مهم: Trace یک ماتریس برابر با مجموع مقادیر ویژه آن است:

$$\mathrm{tr}(\mathbf{A}) = \sum_i \lambda_i(\mathbf{A})$$

(مقادیر ویژه بر اساس تکرارهایشان در نظر گرفته می‌شوند.)

دترمینان (Determinant)

دترمینان یک ماتریس مربعی را می‌توان به روش‌های مختلفی تعریف کرد، اما مهم‌ترین ویژگی‌های آن به صورت زیر هستند:

۱. $\det(\mathbf{I}) = 1$
۲. $\det(\mathbf{A}^T) = \det(\mathbf{A})$
۳. $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$
۴. $\det(\mathbf{A}^{-1}) = \det(\mathbf{A})^{-1}$
۵. $\det(\alpha\mathbf{A}) = \alpha^n \det(\mathbf{A})$

ویژگی مهم: دترمینان یک ماتریس برابر با حاصل‌ضرب مقادیر ویژه آن است:

$$\det(\mathbf{A}) = \prod_{i} \lambda_i(\mathbf{A})$$

(مقادیر ویژه بر اساس تکرارهایشان در نظر گرفته می‌شوند.)

یک ماتریس $\mathbf{Q} \in \mathbb{R}^{n\times n}$ **متعامد** نامیده می‌شود اگر ستون‌های آن دو به دو متعامد یکه باشند. این تعریف به این معنی است که:

$$\mathbf{Q}^\top\mathbf{Q} = \mathbf{Q}\mathbf{Q}^\top = \mathbf{I}$$

یا به عبارتی:

$$\mathbf{Q}^\top = \mathbf{Q}^{-1}$$

ویژگی مهم ماتریس‌های متعامد این است که حاصل‌ضرب داخلی را حفظ می‌کنند:

$$(\mathbf{Qx})^\top(\mathbf{Qy}) = \mathbf{x}^\top\mathbf{Q}^\top\mathbf{Qy} = \mathbf{x}^\top\mathbf{Iy} = \mathbf{x}^\top\mathbf{y}$$

حفظ نرم دو (۲-Norm)

نتیجه مستقیم خاصیت قبل این است که این ماتریس‌ها نرم دو را نیز حفظ می‌کنند:

$$\|\mathbf{Qx}\|_2 = \sqrt{(\mathbf{Qx})^\top(\mathbf{Qx})} = \sqrt{\mathbf{x}^\top\mathbf{x}} = \|\mathbf{x}\|_2$$

تفسیر هندسی

بنابراین، ضرب یک ماتریس متعامد در یک بردار را می‌توان به عنوان یک تبدیل در نظر گرفت که طول بردار را حفظ می‌کند، اما ممکن است آن را دوران (Rotation) داده یا نسبت به مبدأ بازتاب (Reflection) دهد.

تعریف ماتریس‌های متقارن

یک ماتریس $\mathbf{A} \in \mathbb{R}^{n\times n}$ **متقارن (Symmetric)** نامیده می‌شود، اگر با ترانهاده‌ی خود برابر باشد:

$$\mathbf{A} = \mathbf{A}^\top$$

این یعنی که:

$$A_{ij} = A_{ji} \quad \forall(i,j)$$

این تعریف دارای پیامدهای قوی است.

قضیه : (Spectral Theorem) اگر $\mathbf{A} \in \mathbb{R}^{n\times n}$ **متقارن** باشد، آنگاه یک **پایه‌ی متعامد (Orthonormal Basis)** از **بردارهای ویژه (Eigenvectors)** برای $\mathbb{R}^n$ وجود دارد که شامل بردارهای ویژه‌ی $\mathbf{A}$ است.

این قضیه منجر به نوع خاصی از تجزیه‌ی ماتریس‌های متقارن می‌شود که به آن:

- **تجزیه‌ی مقدار ویژه (Eigen Decomposition)**
- **تجزیه‌ی طیفی (Spectral Decomposition)**

گفته می‌شود.

اگر $\mathbf{q}_1, \ldots, \mathbf{q}_n$ یک **پایه‌ی متعامد** از بردارهای ویژه باشند و $\lambda_1, \ldots, \lambda_n$ مقادیر ویژه‌ی متناظر، آنگاه $\mathbf{Q}$ یک ماتریس متعامد شامل بردارهای $\mathbf{q}_i$ به عنوان ستون‌های آن است و:

$$= \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$$

رابطه‌ی اساسی تجزیه طیفی

با توجه به تعریف:

$$\mathbf{A}\mathbf{q}_i = \lambda_i \mathbf{q}_i$$

برای هر i، داریم:

$$\mathbf{AQ} = \mathbf{Q}$$

که با ضرب طرف راست در $\mathbf{Q}^\top$، به فرم تجزیه مقدار ویژه می‌رسیم:

$$\mathbf{A} = \mathbf{Q}\;\mathbf{Q}^\top$$

SVD ابزاری مهم در جبر خطی است که قدرت آن به این واقعیت بازمی‌گردد که **هر ماتریس** $\mathbf{A} \in \mathbb{R}^{m \times n}$ دارای SVD است. SVD به صورت زیر تعریف می‌شود:

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^\top$$

که در آن:

- $\mathbf{U} \in \mathbb{R}^{m \times m}$ و $\mathbf{V} \in \mathbb{R}^{n \times n}$ **ماتریس‌های متعامد (Orthogonal Matrices)** هستند.
- $\Sigma \in \mathbb{R}^{m \times n}$ **ماتریس قطری (Diagonal Matrix)** است که **مقادیر منفرد (Singular Values)** $\mathbf{A}$ روی قطر آن قرار دارند.

ترتیب مقادیر منفرد

فقط اولین $r = \text{rank}(\mathbf{A})$ مقدار منفرد مقدار غیرصفر دارند، و بر اساس قرارداد، این مقادیر به صورت نزولی مرتب می‌شوند:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_{\min(m,n)} = 0$$

نمایش دیگر SVD

راه دیگری برای نوشتن SVD (با استفاده از **مجموع حاصل‌ضرب خارجی (sum-of-outer-products)**) به صورت زیر است:

$$\mathbf{A} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$$

که در آن $\mathbf{u}_i$ و $\mathbf{v}_i$ به ترتیب ستون‌های iـام از $\mathbf{U}$ و $\mathbf{V}$ هستند.

عوامل تجزیه SVD **تجزیه مقدار ویژه** (Eigen Decomposition) برای $\mathbf{A}^\top\mathbf{A}$ و $\mathbf{A}\mathbf{A}^\top$ ارائه می‌کنند:

$$\mathbf{A}^\top\mathbf{A} = (\mathbf{U}\Sigma\mathbf{V}^\top)^\top(\mathbf{U}\Sigma\mathbf{V}^\top)$$

$$= \mathbf{V}\Sigma^\top\mathbf{U}^\top\mathbf{U}\Sigma\mathbf{V}^\top = \mathbf{V}\Sigma^\top\Sigma\mathbf{V}^\top$$

$$\mathbf{A}\mathbf{A}^\top = \mathbf{U}\Sigma\mathbf{V}^\top(\mathbf{U}\Sigma\mathbf{V}^\top)^\top$$

$$= \mathbf{U}\Sigma\mathbf{V}^\top\mathbf{V}\Sigma^\top\mathbf{U}^\top = \mathbf{U}\Sigma\Sigma^\top\mathbf{U}^\top$$

بردارهای ویژه و single values

- ستون‌های $\mathbf{V}$ (بردارهای ویژه راست - **Right Singular Vectors**) مقادیر ویژه‌ی $\mathbf{A}^\top\mathbf{A}$ هستند.
- ستون‌های $\mathbf{U}$ (بردارهای ویژه چپ - **Left Singular Vectors**) مقادیر ویژه‌ی $\mathbf{A}\mathbf{A}^\top$ هستند.
- ماتریس‌های $\Sigma^\top\Sigma$ و $\Sigma\Sigma^\top$ هر دو قطری هستند و مقدارهای روی قطر آن‌ها برابر با مربع مقادیر منفرد σ_i^2 است.
- بنابراین، مقادیر منفرد $\mathbf{A}$ ریشه‌های دوم مقادیر ویژه‌ی $\mathbf{A}^\top\mathbf{A}$ (یا معادل آن $\mathbf{A}\mathbf{A}^\top$) هستند.

قضیه

اگر $\mathbf{A} \in \mathbb{R}^{m\times n}$ باشد، آنگاه:

۱. فضای $(\textbf{Null space})$ از $\mathbf{A}$ برابر است با متمم متعامد تصویر $\mathbf{A}^\top$:

$$\mathrm{null}(\mathbf{A}) = \mathrm{range}(\mathbf{A}^\top)^\perp$$

۲. جمع مستقیم (Direct Sum) نول و تصویر ماتریس برابر با $\mathbb{R}^n$ است:

$$\mathrm{null}(\mathbf{A}) \oplus \mathrm{range}(\mathbf{A}^\top) = \mathbb{R}^n$$

۳. بعد فضای تصویر و بعد فضای نول با هم جمع می‌شوند تا مقدار n را بدهند:

$$\dim \mathrm{range}(\mathbf{A}) + \dim \mathrm{null}(\mathbf{A}) = n$$

SVD و زیرفضاهای اساسی

اگر تجزیه مقدار منفرد $\mathbf{A}$ به صورت زیر باشد:

$$\mathbf{A} = \mathbf{U}\ \mathbf{V}^\top$$

آنگاه ستون‌های $\mathbf{U}$ و $\mathbf{V}$ به ترتیب پایه‌های متعامد چهار **زیرفضای اساسی (Fundamental Subspaces)** از $\mathbf{A}$ را تشکیل می‌دهند.

زیرفضا (Subspace)	**ستون‌ها (Columns)**
$\text{range}(\mathbf{A})$	اولین r ستون‌های $\mathbf{U}$
$\text{range}(\mathbf{A}^\top)$	اولین r ستون‌های $\mathbf{V}$
$\text{null}(\mathbf{A}^\top)$	آخرین $m - r$ ستون‌های $\mathbf{U}$
$\text{null}(\mathbf{A})$	آخرین $n - r$ ستون‌های $\mathbf{V}$

که در آن: $r = \text{rank}(\mathbf{A})$

فرم‌های درجه دوم

فرض کنید $\mathbf{A} \in \mathbb{R}^{n\times n}$ یک ماتریس متقارن باشد. عبارت $\mathbf{x}^\top \mathbf{A}\mathbf{x}$ را یک **فرم درجه دوم** (Quadratic Form) از $\mathbf{A}$ می‌نامند.

گاهی بازنویسی فرم درجه دوم برحسب مؤلفه‌های منفرد ماتریس $\mathbf{A}$ و بردار $\mathbf{x}$ سودمند است:

$$\mathbf{x}^\top \mathbf{A}\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}x_i x_j$$

ویژگی‌های فرم‌های درجه دوم

- این رابطه برای هر ماتریس مربعی معتبر است، حتی اگر متقارن نباشد.
- با این حال، فرم‌های درجه دوم معمولاً در زمینه‌ی **ماتریس‌های متقارن** مورد بحث قرار می‌گیرند.

منابع و مراجع

- Linear Algebra - Hamid R. Rabiee , Maryam Ramezani
- Mathematics for Machine Learning - Garrett Thomas