

یادآوری جبر خطی و حسابان

نمادها

تعاریف

بردار - $x \in \mathbb{R}^n$ یک بردار با n درایه است، که $x_i \in \mathbb{R}$ درایه i ام می باشد:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

ماتریس - $A \in \mathbb{R}^{m \times n}$ یک بردار با m سطر و n ستون است، که در آن $A_{i,j} \in \mathbb{R}$ درایه ای است که در سطر i ام و ستون j ام قرار دارد:

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} \in \mathbb{R}^{m \times n}$$

نکته: بردار x که در بالا تعریف شد را می‌توان به صورت یک ماتریس $1 \times n$ در نظر گرفت که به طور خاص به آن بردار ستونی گویند.

ماتریس‌های اصلی:

ماتریس همانی - ماتریس همانی $I \in \mathbb{R}^{n \times n}$ یک ماتریس مربعی است که درایه‌های قطری آن همه مقدار ۱ و بقیه‌ی درایه‌ها مقدار ۰ دارند:

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

نکته: برای همه‌ی ماتریس‌های $A \in \mathbb{R}^{n \times n}$ داریم $A \times I = I \times A = A$.

ماتریس قطری - ماتریس $D \in \mathbb{R}^{n \times n}$ یک ماتریس مربعی است که درایه‌های قطری آن مقادیر غیرصفر دارند و بقیه‌ی درایه‌ها صفر هستند:

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{pmatrix}$$

نکته: D همچنین به صورت $D = \text{diag}(d_1, \dots, d_n)$ هم نمایش داده می‌شود.

عملیات ماتریسی

ضرب

بردار با بردار - دو نوع عملیات ضرب بردار با بردار وجود دارد:

ضرب داخلی: برای هر $x, y \in \mathbb{R}^n$ داریم:

$$x^T y = \sum_{i=1}^n x_i y_i \in \mathbb{R}$$

ضرب خارجی: برای هر $x \in \mathbb{R}^m$ و $y \in \mathbb{R}^n$ داریم:

$$xy^T = \begin{pmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & & \vdots \\ x_m y_1 & \cdots & x_m y_n \end{pmatrix} \in \mathbb{R}^{m \times n}$$

ماتریس با بردار - ضرب ماتریس $A \in \mathbb{R}^{m \times n}$ و بردار $x \in \mathbb{R}^n$ برداری با اندازه‌ی m است به طوری که:

$$Ax = \begin{pmatrix} a_{r,1}^T x \\ \vdots \\ a_{r,m}^T x \end{pmatrix} = \sum_{i=1}^n a_{c,i} x_i \in \mathbb{R}^m$$

که $a_{r,i}^T$ بردارهای سطری و $a_{c,j}$ بردارهای ستونی A ، و x_i درایه‌های x هستند.

ماتریس با ماتریس - ضرب ماتریس‌های $A \in \mathbb{R}^{m \times n}$ و $B \in \mathbb{R}^{n \times p}$ ماتریسی با اندازه‌ی $n \times p$ است که:

$$AB = \begin{pmatrix} a_{r,1}^T b_{c,1} & \cdots & a_{r,1}^T b_{c,p} \\ \vdots & & \vdots \\ a_{r,m}^T b_{c,1} & \cdots & a_{r,m}^T b_{c,p} \end{pmatrix} = \sum_{i=1}^n a_{c,i} b_{r,i}^T \in \mathbb{R}^{n \times p}$$

که $a_{r,i}^T, b_{r,i}^T$ بردارهای سطری و $a_{c,j}$ و $b_{c,j}$ بردارهای ستونی A و B هستند.

دیگر عملیات

ترانهاده - ترانهادی ماتریس $A \in \mathbb{R}^{m \times n}$ که با A^T نمایش داده می‌شود، ماتریسی است که مکان درایه‌های آن نسبت به قطر ماتریس برعکس شده‌اند:

$$\forall i, j, \quad A_{i,j}^T = A_{j,i}$$

نکته: برای ماتریس‌های A و B ، داریم $(AB)^T = B^T A^T$.

معکوس - معکوس یک ماتریس مربعی معکوس‌پذیر A که با A^{-1} نمایش داده می‌شود، تنها ماتریسی است که:

$$AA^{-1} = A^{-1}A = I$$

نکته: همه‌ی ماتریس‌های مربعی معکوس‌پذیر نیستند. همچنین، برای ماتریس‌های مربعی معکوس‌پذیر A و B داریم $(AB)^{-1} = B^{-1}A^{-1}$.

اثر - اثر ماتریس مربعی A که با $\text{tr}(A)$ نمایش داده می‌شود، مجموع همه‌ی درایه‌های قطری ماتریس است.

$$\text{tr}(A) = \sum_{i=1}^n A_{i,i}$$

نکته: برای ماتریس‌های A و B داریم $\text{tr}(A^T) = \text{tr}(A)$ و $\text{tr}(AB) = \text{tr}(BA)$.

دترمینان - دترمینان یک ماتریس مربعی $A \in \mathbb{R}^{n \times n}$ که با $|A|$ یا $\det(A)$ نمایش داده می‌شود، به صورت یک عبارت بازگشتی بر روی $A_{\setminus i, \setminus j}$ ، که ماتریس A بدون سطر i -ام و ستون j -ام است، به صورت زیر تعریف می‌شود:

$$\det(A) = |A| = \sum_{j=1}^n (-1)^{i+j} A_{i,j} |A_{\setminus i, \setminus j}|$$

نکته: A معکوس‌پذیر است اگر و فقط اگر $|A| \neq 0$. همچنین $|AB| = |A||B|$ و $|A^T| = |A|$.

ویژگی‌های ماتریس‌ها

تعاریف

تجزیه‌ی متقارن - یک ماتریس دلخواه A را می‌توان با استفاده از اجزای متقارن و غیرمتقارن آن به صورت زیر نشان داد:

$$A = \underbrace{\frac{A + A^T}{2}}_{\text{Symmetric}} + \underbrace{\frac{A - A^T}{2}}_{\text{Antisymmetric}}$$

نرم - نرم تابع $N : V \rightarrow [0, +\infty[$ است که V یک فضای برداری است، و به گونه‌ای است که برای هر $x, y \in V$ داریم:

- $N(x + y) \leq N(x) + N(y)$ •
- $N(ax) = |a|N(x)$ برای عدد اسکالر •
- اگر $N(x) = 0$ باشد در این صورت $x = 0$ •

برای $x \in V$ ، نرم‌هایی که بیشتر استفاده می‌شوند در جدول زیر آمده‌اند:

کاربرد	تعریف	نماد	نرم
LASSO	$\sum_{i=1}^n x_i $	$\ x\ _1$	Manhattan, L^1
Ridge	$\sqrt{\sum_{i=1}^n x_i^2}$	$\ x\ _2$	Euclidean, L^2
Inégalité de Hölder	$\left(\sum_{i=1}^n x_i^p \right)^{\frac{1}{p}}$	$\ x\ _p$	p -norme, L^p

Infini, L^∞	$\ x\ _\infty$	$\max_i x_i $	Convergence uniforme
--------------------	----------------	----------------	----------------------

وابستگی خطی - مجموعه‌ای از بردارها وابستگی خطی دارند اگر یکی از بردارهای مجموعه را بتوان به صورت ترکیب خطی دیگر بردارها تعریف کرد.
 نکته: اگر نتوان هیچ برداری را به این شکل تعریف کرد، در این صورت بردارها استقلال خطی دارند.

رتبه ماتریس - رتبه‌ی یک ماتریس A که با $\text{rank}(A)$ نمایش داده می‌شود، تعداد ابعاد فضایی است که توسط ستون‌های آن ایجاد می‌شود. این مقدار برابر است با حداکثر تعداد ستون‌های A که استقلال خطی داشته باشند.

ماتریس مثبت نیمه‌معین - ماتریس $A \in \mathbb{R}^{n \times n}$ یک ماتریس مثبت نیمه‌معین است که با $A \succeq 0$ نمایش داده می‌شود اگر داشته باشیم:

$$\boxed{A = A^T} \quad \text{و} \quad \boxed{\forall x \in \mathbb{R}^n, \quad x^T A x \geq 0}$$

نکته: به طور مشابه، یک ماتریس A مثبت معین است ($A \succ 0$)، اگر یک ماتریس مثبت نیمه‌معین باشد که برای هر بردار غیر صفر x داشته باشیم $x^T A x > 0$.

مقدار ویژه، بردار ویژه - برای یک ماتریس $A \in \mathbb{R}^{n \times n}$ ، گوئیم λ یک مقدار ویژه ماتریس A است اگر وجود داشته باشد بردار $z \in \mathbb{R}^n \setminus \{0\}$ که یک بردار ویژه نام دارد، به طوری که:

$$Az = \lambda z$$

قضیه‌ی طیفی - فرض کنید $A \in \mathbb{R}^{n \times n}$ باشد. اگر A متقارن باشد، در این صورت A توسط یک ماتریس حقیقی متعامد $U \in \mathbb{R}^{n \times n}$ قطری‌پذیر است. با نمایش $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ داریم:

$$\exists \Lambda \text{ diagonal, } A = U\Lambda U^T$$

تجزیه‌ی مقدار منفرد - برای یک ماتریس A با ابعاد $m \times n$ ، تجزیه‌ی مقدار منفرد یک تکنیک تقسیم‌بندی است که تضمین می‌کند یک ماتریس یکانی $U \in \mathbb{R}^{n \times n}$ ، یک ماتریس قطری $\Sigma \in \mathbb{R}^{m \times n}$ ، و یک ماتریس یکانی $V \in \mathbb{R}^{m \times m}$ وجود دارند، به طوری که:

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

حسابان ماتریسی

گرادیان - فرض کنید $f : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ یک تابع و $A \in \mathbb{R}^{m \times n}$ یک ماتریس باشد. گرادیان f نسبت به A یک ماتریس با ابعاد $m \times n$ است و با $\nabla_A f(A)$ نمایش داده می‌شود، به طوری که:

$$\left(\nabla_A f(A) \right)_{i,j} = \frac{\partial f(A)}{\partial A_{i,j}}$$

نکته: گرادینان f تنها زمانی تعریف شده است که f تابعی باشد که یک عدد/اسکالر خروجی دهد.

هسیان - فرض کنید $f: \mathbb{R}^n \rightarrow \mathbb{R}$ یک تابع و $x \in \mathbb{R}^n$ یک بردار باشد. هسیان f نسبت به x یک ماتریس متقارن با ابعاد $n \times n$ است و با $\nabla_x^2 f(x)$ نمایش داده می‌شود، به طوری که:

$$\left(\nabla_x^2 f(x) \right)_{i,j} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$$

نکته: هسیان تابع f تنها زمانی تعریف شده است که f تابعی با خروجی اسکالر باشد.

عملیات گرادینانی - برای ماتریس‌های A ، B ، و C ، ویژگی‌های زیر را به خاطر داشته باشید:

$$\nabla_A \text{tr}(AB) = B^T$$

$$\nabla_{A^T} f(A) = (\nabla_A f(A))^T$$

$$\nabla_A \text{tr}(ABA^T C) = CAB + C^T AB^T$$

$$\nabla_A |A| = |A| (A^{-1})^T$$