Preliminary – Linear Algebra & Calculus

Vectors

Inner product

$\langle u, v \rangle = u^T v = \sum_{i=1}^{n} u_i v_i, \quad \forall u, v \in \mathbb{R}^n$

* Two vectors are called orthogonal if their inner product is 0.

Norms

Euclidian norm : $\|v\| = \langle v, v \rangle^{1/2} = \left(\sum_{i=1}^{n} v_i^2 \right)^{1/2}$

l_p-norm : $\|v\|_p = \left(\sum_{i=1}^{n} v_i^p \right)^{1/p}$

* When $p = \infty$, $\|v\|_\infty = \max_{i \in [n]} |v_i|$
* When $p = 1$, $\|v\|_1 = \sum_{i=1}^{n} |v_i|$

Inequalities

Cauchy-Schwarz inequality : $\langle u, v \rangle \le \|u\| \cdot \|v\|$

Hölder inequality : $\langle u, v \rangle \le \|u\|_p \cdot \|v\|_q$, $\forall p, q \in [1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$

Matrices

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix}_{m \times n}, \quad A \in \mathbb{R}^{m \times n} : m \text{ rows}, n \text{ columns}$$

Trace

$Tr(A) = \sum_{i=1}^{n} A_{ii}$, $A = (A_{ii}) \in \mathbb{R}^{n \times n}$, which is the sum of its diagonal entries.

Cyclicity

For every pair of matrices $A, B \in \mathbb{R}^{m \times n}$, $Tr(AB^T) = Tr(B^T A)$.

* Connection between traces and Euclidian norms : for $v \in \mathbb{R}^n$, we have
  $\|v\|^2 = Tr(v^T v) = Tr(v v^T)$

Trace & inner product

For $A, B \in \mathbb{R}^{m \times n}$, we have

$\langle A, B \rangle = Tr(AB^T) = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} B_{ij}$

Trace & expectation

vector $z \in \mathbb{R}^n$, symmetric matrix $A \in \mathbb{R}^{n \times n}$

$\mathbb{E}[z^T A z] = Tr(A \cdot \mathbb{E}[z z^T])$

Orthogonal matrices

$U \in \mathbb{R}^{n \times n}$ is orthogonal if its columns and rows are orthogonal unit vectors – also called orthonomal.

* $U^T U = U U^T = I_n \implies U^{-1} = U^T$
* For $U \in \mathbb{R}^{n \times n}$ and $v \in \mathbb{R}^n$, we have $\|Uv\| = \|v\|$.

Positive semidefinite matrices

A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is called positive semidefinite (PSD), denoted as $A \succeq 0$, iff $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$. A satisfies

- All eigenvalues of A are non-negative.
- $\langle A, M \rangle \geq 0$ for all PSD matrices M
- There is a unique PSD matrix $C \in \mathbb{R}^{n \times n}$ such that $A = CC$ (non-negative square root)

<u>Positive definite matrices</u>

A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is called positive definite, denoted as $A \succ 0$, iff $x^T A x > 0$ for all non-zero $x \in \mathbb{R}^n$. A satisfies

- All eigenvalues of A are positive
- $\langle A, M \rangle > 0$ for all positive definite matrices M
- There is a unique positive def. matrix $C \in \mathbb{R}^{n \times n}$ such that $A = CC$ (positive square root).

<u>Singular-value decomposition</u>

For a matrix $A \in \mathbb{R}^{m \times n}$, a singular-value decomposition (SVD) is a factorization as

$$A = \sum_{i=1}^{\min\{m,n\}} \sigma_i \cdot u_i \cdot v_i^T$$

- $\sigma_1 \geq \cdots \sigma_n \geq 0$ are scalar (singular values)
- $u_1, \cdots, u_n \in \mathbb{R}^m$ are orthonormal vectors (left-singular vectors)
- $v_1, \cdots, v_n \in \mathbb{R}^n$ are orthonormal vectors (right-singular vectors)

* Such a decomposition always exists but not necessarily unique.

* Rank of a matrix is the number of non-zero, i.e., strictly positive, singular values.

* SVD of A gives rise to best low-rank approximations for A. For $r \leq \min\{m,n\}$,
$A' = \sum_{i=1}^{r} \sigma_i \cdot u_i v_i^T$, satisfies $A' = \operatorname{argmin} \{ \|M - A\|_F^2 \mid \operatorname{rank}(M) < r \}$

<u>example</u> $A = U \Sigma V^T$

$$A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \quad A^T = \begin{bmatrix} 2 & -1 \\ 2 & 1 \end{bmatrix}, \text{ then } A^T A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & -1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} (4+1) & (4-1) \\ (4-1) & (4+1) \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$$

$$\det(A^T A - \lambda I) = (5-\lambda)(5-\lambda) - 3 \cdot 3 = 0$$
$$\Rightarrow 25 - 10\lambda + \lambda^2 - 9 = 0$$
$$\Rightarrow \lambda^2 - 10\lambda + 16 = 0$$
$$\Rightarrow \lambda_1 = 8, \lambda_2 = 2$$

$$\sigma_1 = \sqrt{8} = 2\sqrt{2}, \quad \sigma_2 = \sqrt{2}$$
$$\Rightarrow \Sigma = \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}$$

With $\lambda_1 = 8$, we have

$$\begin{bmatrix} 5-8 & 3 \\ 3 & 5-8 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} -3 & 3 \\ 3 & -3 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\Rightarrow x = y, \; v_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

With $\lambda_2 = 2$, we have

$$\begin{bmatrix} 5-2 & 3 \\ 3 & 5-2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\Rightarrow x = -y, \; v_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$u_1 = \frac{1}{2\sqrt{2}} \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{2\sqrt{2}} \begin{bmatrix} 4/\sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$u_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 2/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$V = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$
$$\Downarrow$$
$$V^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

<u>Eigen decomposition</u>

Suppose A is a real symmetric matrix, a vector-scalar pair $(w, \lambda) \in \mathbb{R}^n \times \mathbb{R}$ is called an eigenvalue-eigenvector-pair of the matrix A if $Aw = \lambda w$.

- λ is eigenvalue
- w is eigenvector

* For a real symmetric matrix, there always exists n orthogonal eigenvectors $w_1, \cdots, w_n$ and $\lambda_1, \cdots, \lambda_n$ eigenvalues.

* Eigenvalue decomposition is a factorization of $A = \sum_{i=1}^{n} \lambda_i w_i w_i^T$.

example: $A = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}$, $A - \lambda I = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{bmatrix}$

$$\det(A - \lambda I) = (4-\lambda)(3-\lambda) - 1 \cdot 2 = 0$$
$$\Rightarrow 12 - 4\lambda - 3\lambda + \lambda^2 - 2 = 0$$
$$\Rightarrow \lambda^2 - 7\lambda + 10 = 0 \Rightarrow \lambda = 5 \text{ or } 2$$

$\Lambda = \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}$ eigenvalue found

$$AQ = \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} \overset{5}{\cancel{4+1}} & \overset{2}{\cancel{4-2}} \\ \underset{5}{\cancel{2+3}} & \underset{-4}{\cancel{2-6}} \end{bmatrix}$$

$$Q\Lambda = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ 5 & -4 \end{bmatrix}$$

With $\lambda_1 = 5$, we have

$$\begin{bmatrix} 4-5 & 1 \\ 2 & 3-5 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\Rightarrow x = y \Rightarrow V_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

With $\lambda_2 = 2$, we have

$$\begin{bmatrix} 4-2 & 1 \\ 2 & 3-2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\Rightarrow y = -2x \Rightarrow V_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$A = Q \Lambda Q^{-1}$$

$$Q = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \Rightarrow Q^{-1} = \frac{1}{\det(Q)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

eigen vector

$$= \frac{1}{1 \cdot (-2) - 1 \cdot 1} \begin{bmatrix} -2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{bmatrix}$$

<u>Matrix norms and inequalities</u>

Matrix norm	Symbol	Func. of singular value	Vector equivalence
Spectral norm	$\|X\|$	σ_1	L_∞ - norm
Frobenius norm	$\|X\|_F$	$\sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$	L_2 - norm
Nuclear norm	$\|X\|_*$	$\sigma_1 + \cdots + \sigma_r$	L_1 - norm

Suppose $A \in \mathbb{R}^{m \times n}$, then

- The spectral norm of A is $\|A\| = \sigma_1$
- The nuclear norm of A is $\|A\|_* = \sum_{i=1}^{r} \sigma_i$
- The Frobenius norm of A is $\|A\|_F = \left(\sum_{i=1}^{n} \sigma_i^2 \right)^{1/2}$

Some properties of these norms:

- $\|A\| \leq \|A\|_F \leq \|A\|_*$
- $\|A\|_* \leq \sqrt{r} \cdot \|A\|_F$
- $\langle A, B \rangle \leq \|A\| \cdot \|B\|_*$ (Hölder's inequality)

<u>Series</u>

<u>Taylor series</u>

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots$$

<u>Exponential series</u>

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

Logarithmic expansion

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \quad , \quad |x| < 1$$

Geometric series

$$\sum_{k=0}^{\infty} ar^k = a + ar + ar^2 + ar^3 + \cdots = \frac{a}{1-r} \quad , \quad |r| < 1$$

Bound for binomial term

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \leq \frac{n^k}{k!}$$ * upper bound

$$e^k = \sum_{i=0}^{\infty} \frac{k^i}{i!} = 1 + k + \frac{k^2}{2!} + \cdots + \frac{k^k}{k!} + \cdots > \frac{k^k}{k!} \longrightarrow e^k > \frac{k^k}{k!}$$ * lower bound for $k!$

Together: $\binom{n}{k} \leq \frac{n^k}{k!} < \frac{n^k}{\left(\frac{k}{e}\right)^k} = \left(\frac{en}{k}\right)^k$ $\qquad \Longrightarrow k! > \left(\frac{k}{e}\right)^k$