Kernel Methods AMMI 2024

Matrix Operations

In what follows, $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$ and $x, y \in \mathbb{R}^n$.

1. Compute the following, giving conditions on m, n, p for the formulas to be well-defined. When the result is a matrix, give its value at position (i, j). When it is a vector, give its value at position i.

- AB
- $x^T A$
- $x^T y$
- xy^T
- $y^T Ax$
- $Tr(A^T A)$

2. Show that

- $(AB)^T = B^T A^T$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^T)^{-1} = (A^{-1})^T$
- The *trace* operator $A \rightarrow Tr(A)$ is a *linear operator* on $\mathbb{R}^{m\times n}$.
- $Tr(AB) = Tr(BA)$.

3. Given a matrix $A \in \mathbb{R}^{n\times n}$ and a vector $r \in \mathbb{R}^n$, how to multiply each row of A by the elements of r?

4. Given $x_i \in \mathbb{R}^d,\ i = 1, ..., n$, how to compute the mean of $\{||x_i - x_j||_2;\ i, j = 1, ..., n\}$ in Python, without a loop?

Gradients

1. Let $A \in \mathbb{R}^{m\times n}$. In the following, give the function's output dimension and compute the gradient:

- $\nabla_x\ (xA), \quad x \in \mathbb{R}$
- $\nabla_{\mathrm{x}}\ (A\mathrm{x}), \quad \mathrm{x} \in \mathbb{R}^n$

- $\nabla_{\mathrm{x}}\ (\mathrm{y}^T A\mathrm{x}), \nabla_{\mathrm{y}}\ (\mathrm{y}^T A\mathrm{x}), \quad \mathrm{x}, \mathrm{y} \in \mathbb{R}^n$
- $\nabla_{\mathrm{x}}\ (\mathrm{x}^T A\mathrm{x}), \quad \mathrm{x} \in \mathbb{R}^n$
- $\nabla_{\mathrm{x}}\ ||\mathrm{y} - A\mathrm{x}||_2^2, \quad \mathrm{x} \in \mathbb{R}^n$
- $\nabla_X\ Tr(A^T X), \quad X \in \mathbb{R}^{m\times n}$

2. Compute the following gradients using the chain rule:

- $\nabla_z\ Ax(z)$
- $\nabla_z\ y(z)^T Ax(z)$

Symmetric matrices

Reminders

Def (Eigenvector Equation). Let $A \in \mathbb{R}^n$. If $Au = \lambda u$ with $\lambda \in \mathbb{R}$ and $u \in \mathbb{R}^n$, then λ is called an *eigenvalue* of A and u an *eigenvector*.

Def (Orthogonal matrix). $U \in \mathbb{R}^n$ is orthogonal $\iff$ U is inversible and $U^{-1} = U^T$.

Spectral theorem. Any (real) symmetric matrix $A \in \mathbb{R}^{n\times n}$ can be decomposed as

$$A = U\Lambda U^T$$

where U is an orthogonal matrix and $\lambda = \mathrm{diag}(\lambda_1, ..., \lambda_n)$.

Exercises

1. Show that for an orthogonal matrix U,

$$u_i^T u_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

2. Let $A \in \mathbb{R}^{n\times n}$ a (real) symmetric matrix and $A := U\Lambda U^T$ with U orthogonal and $\Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_n)$. Show that the columns of U are eigenvectors of A and $\lambda_i, i = 1, ..., n$ are eigenvalues.

3. Let $A = U\Lambda U^T$ a symmetric matrix. Show that:
   a. A is inversible $\iff$ $\lambda_i \neq 0, \forall i$
   b. $A^{-1} = U\Lambda^{-1}U^T$ with $\Lambda^{-1} = \mathrm{diag}(\lambda_1^{-1}, ..., \lambda_n^{-1})$
   c. $\mathrm{Tr}(A) = \sum_i \lambda_i,\ \mathrm{Tr}(A^{-1}) = \sum_i \lambda_i^{-1}$

Positive semi-definte matrices

Reminders

Def (Positive (semi-)definite matrix). Let $A \in \mathbb{R}^{n\times n}$.

$$\begin{array}{lcl} A \;\; \text{p.s.d.} & \iff & x^T A x \geq 0, \forall x \in \mathbb{R}^n \\ A \;\; \text{p.d.} & \iff & x^T A x > 0, \forall x \in \mathbb{R}^n \end{array}$$

Exercies

1. Let $A \in \mathbb{R}^{n\times n}$. Show that
   a. A p.s.d. and symmetric $\iff$ A has n eigenvalues $\lambda_i \geq 0$
   b. A p.d. and symmetric $\iff$ A has n eigenvalues $\lambda_i > 0$

2. Let $X \in \mathbb{R}^{n\times p},\ \lambda > 0$. Prove that $M = X^T X + \lambda I_p$ is inversible.
   Hint: Prove that the eigenvalues of M are larger than λ.

3. For any general $X \in \mathbb{R}^{m\times n}$, show that
   a. $X^T X$ and $X X^T$ are symmetric, p.s.d.
   b. the non-zero eigenvalues of $X^T X$ and $X X^T$ are the same, that are $\{\sigma_i^2, \sigma_i \neq 0, i = 1, ..., \min(m, n)\}$ where σ_i's are singular values of X