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
- xA
- $x^T y$
- 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 $A \to Tr(A)$ 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:

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

Positive semi-definite matrices

1. 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.

2. For any general $X \in \mathbb{R}^{m\times n}$, show that
   a. $X^T X$ and XX^T are symmetric, p.s.d.
   b. the non-zero eigenvalues of $X^T X$ and XX^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