Advanced Studies In Mathematics Exercise

Hwijae Son

December 3, 2024

1. A kernel is a non-negative integrable function K such that $K(x) = K(-x)$, and

$$\int K(x)dx = 1.$$

   Show that

$$\hat{f}(x) = \frac{1}{Nh}\sum_{i=1}^{N} K(\frac{x - x_i}{h})$$

   is a probability density function for $h > 0$.

2. Prove that the Kullback-Leibler divergence of P from Q is not symmetric, nonnegative, zero only when $P = Q$, and convex in both P and Q.

3. Compute the gradient of the cost function

$$C = \sum_i \sum_j P_{ij} \log \frac{P_{ij}}{Q_{ij}},$$

   of the t-SNE algorithm.

4. Let $\{x_i \in \mathbb{R}^d\}_{i=1}^N$ be given. Prove that the sample covariance matrix

$$C = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(x_i - \bar{x})^T,$$

   is symmetric, and positive semidefinite.

5. Let $A \in \mathbb{R}^{m\times n}$ be a rank r matrix, and $A = U\Sigma V^T = \sum_{i=1}^r \sigma_i u_i v_i^T$ be its singular value decomposition. Prove that for each i, $u_i v_i^T$ is a rank 1 matrix.

6. Derive the gradient descent updates for the nonnegative matrix factorization algorithm.

7. (Python) Implement your own t-SNE algorithm to perform dimensionality reduction to MNIST dataset. Visualize the result in 2-dimensional space.

8. (Python) Implement your own PCA and Autoencoder algorithms to perform dimensionality reduction to MNIST dataset. Compare the results with reduced dimension $k = 30$.