Research on dimensionality reduction method with PCA and some applications

Vu Thi Nguyet

Abstract: This study investigates the application of Principal Component Analysis (PCA) as a dimensionality reduction technique and explores its practical implementations across various domains. Dimensionality reduction is a critical preprocessing step in data analysis, addressing challenges posed by high-dimensional datasets, including computational inefficiency, overfitting, and the curse of dimensionality. The research systematically examines the mathematical foundations of PCA, focusing on eigenvalue decomposition and covariance matrix analysis to identify principal components that retain maximum variance. Experimental evaluations are conducted on benchmark datasets to assess PCA's effectiveness in reducing feature space while preserving essential data structure. The study further demonstrates PCA's utility in three key applications: image compression, where significant storage reduction is achieved with minimal perceptual loss; noise filtering, which enhances signal-to-noise ratios in sensor data; and feature extraction for classification tasks, improving model accuracy and training speed. Comparative analyses with alternative methods, such as t-SNE and autoencoders, highlight PCA's computational efficiency and interpretability. The findings confirm that PCA remains a robust, versatile tool for dimensionality reduction, offering a balance between simplicity and performance, particularly in scenarios requiring linear feature extraction and rapid processing.

Keywords: PCA, dimensionality reduction, application, principal component analysis