Task 2, List 4

Task

Consider any dataset $X \in R_n^d$, where d is the number of columns, and n is the number of samples. Let $\sum_{i=1}^{d} \mathcal{D}^2 X^i$ be the sum of variances of all the columns. Let $X_{\text{pca}} \in R_n^d$ be the PCA transformed X dataset. Show that

$$\sum_{i=1}^{d} \mathcal{D}^2 X^i = \sum_{i=1}^{d} \mathcal{D}^2 X_{\text{pca}}^i.$$

0. Introduction

We prove that Principal Component Analysis (PCA) preserves the total variance of a dataset. We intuitively know that to be true - we can imagine PCA as looking at the parts of data from different side - it does not change the distance between those datapoints. We will start by explaining a few concepts (and introducing variables), and then prove the equation.

0.1. PCA Transformation

PCA is an orthogonal transformation. Let P be the transformation matrix (it's orthogonal, so $PP^T = I$, which will be useful later). Then we can denote the transformed data with:

$$X_{\text{pca}} = XP.$$

0.2. Trace of a Matrix

The trace of a square matrix $A \in R^{d \times d}$, denoted as $\text{tr}(A)$, is the sum of its diagonal elements:

$$\text{tr}(A) = \sum_{i=1}^{d} A_{ii}.$$

where

$$\text{tr}(AB) = \text{tr}(BA).$$

This property will be useful in proving the variance preservation under PCA.

1. Proving the equation

1.1. Total Variance in the Original Dataset

The variance of each column i in the original dataset is given by:

$$\mathcal{D}^2 X^i = \frac{1}{n} \sum_{j=1}^{n} \left(x_j^i - \bar{x}^i \right)^2,$$

where $\bar{x}^i$ is the mean of column i.

The total variance of the dataset is the sum of variances over all d features, which can be rewritten in terms of the trace of the covariance matrix:

$$\sum_{i=1}^{d} \mathcal{D}^2 X^i = \operatorname{tr}(\operatorname{Cov}(X)),$$

where the covariance matrix of X is:

$$\operatorname{Cov}(X) = \frac{1}{n} X^T X.$$

1.2. Total Variance in the PCA-Transformed Dataset

Since $X_{\text{pca}} = XP$, its covariance matrix is:

$$\operatorname{Cov}(X_{\text{pca}}) = \frac{1}{n} (XP)^T (XP).$$

Using properties of transpose and last equation of 1.1., we can write this as:

$$\operatorname{Cov}(X_{\text{pca}}) = P^T \operatorname{Cov}(X) P.$$

Leaving us only Covariance term to deal with (using last equation of 0.2.):

$$\operatorname{tr}(\operatorname{Cov}(X_{\text{pca}})) = \operatorname{tr}(P^T \operatorname{Cov}(X) P) = \operatorname{tr}(\operatorname{Cov}(X)).$$

Which proves that the total variance is unchanged (and original equation is true)

$$\sum_{i=1}^{d} \mathcal{D}^2 X_{\text{pca}}^i = \sum_{i=1}^{d} \mathcal{D}^2 X^i.$$