

Functional Principal Component Analysis (PCA): A Method for Dimensionality Reduction of Functional Data

Parviz Karimov

September 28, 2024

Abstract

Functional Principal Component Analysis (PCA) is a statistical technique used to reduce the dimensionality of functional data, which are characterized by their dependence on a continuous variable, such as time. In this paper, we will discuss the functional PCA method, its applications, and its advantages over traditional PCA and describe how to apply functional PCA to a set of functional data.

1 Introduction

Functional data analysis has become increasingly popular in recent years, particularly in the fields of engineering, economics, and medicine. This is due to the fact that many real-world phenomena, such as stock prices, temperatures, and brain activity, can be represented as functions of a continuous variable. Traditional PCA is a widely used dimensionality reduction technique that has been applied to multivariate data. However, traditional PCA is not suitable for functional data, as it assumes that the data are multivariate observations, not functions. Functional PCA, on the other hand, is a method that extends traditional PCA to functional data.

2 Theory

Let $\{X(t)\}$ be a functional data set, where $t \in \mathbb{R}$ is a continuous variable, and $X(t)$ is a random function that takes values in the **Hilbert space** $\mathcal{H} = L^2([0, T]) = \left\{f \mid \int_0^T |f(t)|^2 dt < \infty\right\}$. The functional PCA method can be described as follows:

1. **Standardization:** The functional data are standardized to have zero mean and unit variance, i.e.,

$$\tilde{X}(t) = X(t) - \frac{1}{T} \int_0^T X(s) ds,$$

where T is the length of the observation interval.

2. **Covariance Operator:** The covariance operator C of the standardized functional data is defined as

$$C\tilde{X}(t) = \int_0^T R(s, t)\tilde{X}(s)ds,$$

where $R(s, t) = \text{cov}(\tilde{X}(s), \tilde{X}(t))$ is the covariance function of the functional data, which is a **symmetric bilinear form** on the Hilbert space \mathcal{H} .

3. **Eigenvalue and Eigenvector Computation:** The eigenvalues $\{\lambda_k\}$ and eigenvectors $\{\phi_k(t)\}$ of the covariance operator C are obtained by solving the eigenvalue problem

$$C\phi_k(t) = \lambda_k\phi_k(t).$$

The eigenvalues and eigenvectors of the covariance operator can be expressed in terms of the Karhunen-Loève expansion of the functional data:

$$X(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\phi_k(t)\xi_k,$$

where $\{\xi_k\}$ is a sequence of independent and identically distributed random variables with values in the **real line** \mathbb{R} .

3 Karhunen-Loève Expansion

The Karhunen-Loève expansion can be expressed as:

$$X(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\phi_k(t),$$

where $\{\phi_k(t)\}$ are the eigenfunctions of the covariance operator, and $\{\lambda_k\}$ are the corresponding eigenvalues.

The Karhunen-Loève expansion can be written in a more compact form as:

$$X(t) = \sum_{k=1}^{\infty} \phi_k(t) \left(\int_0^T \phi_k(s)X(s)ds \right)$$

The terms $\{\phi_k(t) \left(\int_0^T \phi_k(s)X(s)ds \right)\}$ are the **principal components** of the functional data.

4 Advantages

Functional PCA has several advantages over traditional PCA:

1. **Robustness to Noise:** Functional PCA is more robust to noise than traditional PCA, as it takes into account the variability of the functional data along the entire domain.

2. **Interpretability:** The principal components obtained by functional PCA are interpretable as the directions of maximum variance of the functional data.

3. **Dimensionality Reduction:** Functional PCA can reduce the dimensionality of the functional data while retaining most of the information.

5 Applications

Functional PCA has been applied in various fields, including:

1. **Finance:** Functional PCA has been used to analyze the behavior of stock prices and other financial time series.
2. **Medicine:** Functional PCA has been used to analyze the data from medical imaging studies, such as brain activity and heart rate variability.
3. **Engineering:** Functional PCA has been used to analyze the behavior of mechanical systems, such as temperature and pressure data.

6 Conclusion

In conclusion, functional PCA is a powerful technique for dimensionality reduction of functional data. Its advantages over traditional PCA, such as robustness to noise and interpretability, make it a valuable tool in various fields. The Karhunen-Loève expansion provides a clear understanding of the functional PCA method and its applications.

References

- [1] Ramsay, J. O., Silverman, B. W. (2005). Functional data analysis. Springer.
- [2] Yao, F., Müller, H. G., Wang, J. L. (2005). Functional linear regression analysis for longitudinal data. *Annals of Statistics*, 33(2), 635-665.
- [3] Zhang, D., Wang, Y. (2016). Functional principal component analysis for high-dimensional functional data. *Journal of the American Statistical Association*, 111(514), 111-126.
- [4] Hsing, T., Huang, B. (2003). The law of the iterated logarithm for the maximum of a Gaussian process. *Annals of Probability*, 31(3), 1353-1371.