Theoretical Considerations for Multivariate Functional Data Analysis

Yoshiharu Sato
Hokkaido University, Sapporo, Japan yosiharu@ss.iij4u.or.jp

Multivariate functional data is defined as an element of a direct sum of Hilbert spaces, $\boldsymbol{\mathcal{H}}^{(p)} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \cdots \oplus \mathcal{H}_p$, where each $\mathcal{H}_k$ $(k = 1, 2, \ldots, p)$ is a real separable Hilbert space. In this paper, we consider a Gaussian measures on $\boldsymbol{\mathcal{H}}^{(p)}$ as its probability structure. That is, $\boldsymbol{\mathcal{H}}^{(p)}$-valued Gaussian random variable is defined for a measureable space.

Main objective of multivariate analysis for the finite dimensional real random variables is an interpretation of the covariance structure through the several models. In traditional functional data analysis, the covariance between the pair of random variables is the same as the finite dimensional case. Then there are no essential differences in the covariance structures between the finite dimensional case and infinite case for finite samples.

We will discuss the multivariate analysis just like the classical multivariate analysis under the joint Gaussian probability measure on $\boldsymbol{\mathcal{H}}^{(p)}$. For this purpose, we shall extend the theory of finite dimensional multivariate normal distribution to $\boldsymbol{\mathcal{H}}^{(p)}$-valued random variables. Using the properties of $\boldsymbol{\mathcal{H}}^{(p)}$-valued Gaussian measure, it is shown that a concept of regression can be given by the conditional expectation and the concept of principal components is given by the use of eigen structure of covariance operator.

Keywords: conditional expectation, direct sum of Hilbert spaces, joint Gaussian measure, principal components.