

# A Note on Hilbertian Elliptically Contoured Distributions

Yehua Li

Department of Statistics, University of Georgia, Athens, GA 30602, USA

**Abstract.** In this paper, we discuss elliptically contoured distribution for random variables defined on a separable Hilbert space. It is a generalization of the multivariate elliptically contoured distribution to distributions on infinite dimensional spaces. Some theoretical properties of the Hilbertian elliptically contoured distribution are discussed, examples on functional data are investigated to illustrate the applications of such distributions.

*Keywords:* Elliptically contoured distribution, functional data, Hilbertian random variable.

## 1 Introduction

Elliptically contoured distribution is an important class of distribution in Multivariate Analysis, with some very nice symmetry properties. It is widely used in statistical practices, for example in dimension reduction (Li, 1991, Cook and Weisberg 1991, Schott, 1994) and regression graphs (Cook, 1998). The most important example of this class of distribution is of course the multivariate Gaussian distribution. Properties of multivariate elliptically contoured distribution are well studied, see for example Cambanis, Huang, and Simons (1981), Eaton (1986).

Recent developments in statistics has lead us to look beyond random vectors on the Euclidian space. Statistical models for random vectors defined on infinite dimensional Hilbert spaces are in demand. One important example is the functional data analysis (Ramsay and Silverman, 2005), where the data are vectors in a functional space, e.g. the  $L^2$  space.

Among Hilbertian distributions, the Gaussian distribution is still the most well-understood. For example, in functional data analysis, the random functions are

---

<sup>1</sup>E-mail addresses: yehuali@uga.edu

usually modeled as a Gaussian processes. The class of elliptically contoured distributions is an important generalization from Gaussian. It has important applications in dimension reduction, see recent literature on functional sliced inverse regression (Ferré and Yao, 2002, Li and Hsing, 2007), yet its theoretical properties are not well studied. The goal of this paper is to fill in this gap.

The rest of the paper is organized as the following. We introduce some backgrounds and definitions regarding linear operators and Hilbertian random variables in Section 2. A random representation for Hilbertian elliptically contoured random variables is introduced in Section 3. Some theoretical properties of Hilbertian elliptically contoured distribution are discussed in Section 4, including marginal and conditional distributions of a random variable  $X$  when it is mapped into different Hilbert spaces. Finally, we give some examples in Section 5 to illustrate the applications of the theories derived in the previous sections, especially their application in functional data analysis.

## 2 Definitions and Backgrounds

### 2.1 Linear operators

We first introduce some notation and backgrounds for linear operators on Hilbert spaces. More theories on linear operators can be found in Dunford and Schwartz (1988).

We will restrict our discussion to separable Hilbert spaces. A separable Hilbert space  $\mathcal{H}$  is a Hilbert space with a countable basis,  $\{e_1, e_2, \dots\}$ . For two Hilbert spaces  $\mathcal{H}$  and  $\mathcal{H}'$ , a linear operator  $T : \mathcal{H} \rightarrow \mathcal{H}'$  is a linear map from  $\mathcal{H}$  to  $\mathcal{H}'$ , i.e.

$$T(ax) = a \times (Tx), \quad T(x + y) = Tx + Ty,$$

for any  $x, y \in \mathcal{H}$  and any scalar value  $a$ .  $T$  is *bounded* if

$$\|Tx\|_{\mathcal{H}'} \leq M\|x\|_{\mathcal{H}}, \quad \forall x \in \mathcal{H},$$

for some non-negative real number  $M$ . Denote the class of bounded linear operator from  $\mathcal{H}$  to  $\mathcal{H}'$  as  $\mathcal{B}(\mathcal{H}, \mathcal{H}')$ , when  $\mathcal{H}' = \mathcal{H}$  it is simplified as  $\mathcal{B}(\mathcal{H})$ .

The *adjoint* of an operator  $T \in \mathcal{B}(\mathcal{H}, \mathcal{H}')$ , denoted as  $T^*$ , is an operator mapping from  $\mathcal{H}'$  to  $\mathcal{H}$ , with

$$\langle y, Tx \rangle_{\mathcal{H}'} = \langle T^*y, x \rangle_{\mathcal{H}}, \quad \forall x \in \mathcal{H}, \forall y \in \mathcal{H}'.$$

When  $\mathcal{H} = \mathcal{H}'$ ,  $T$  is called *self-adjoint* if  $T^* = T$ .

## 2.2 Hilbertian random variables

Let  $\mathcal{H}$  be a separable Hilbert space with inner product  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ ,  $(\Omega, \mathcal{F}, P)$  be a probability space, then a Hilbertian valued random variable is a map  $X : (\Omega, \mathcal{F}, P) \rightarrow \mathcal{H}$ . Since finite dimensional Hilbert space are isomorphic to the Euclidian space, and theories in Multivariate Analysis apply, we are generically interested in random variables on an infinite dimensional Hilbert space. In functional data analysis,  $\mathcal{H}$  could be  $L^2$  functional space, or Sobolev space, etc.

The mean of  $X$ , if exists, is defined as  $\mu_X = EX = \int X(\omega)dP(\omega)$ , which is an element in  $\mathcal{H}$  satisfying  $\langle b, EX \rangle = E\langle b, X \rangle$  for all  $b \in \mathcal{H}$ . The variance of  $X$  is an operator on  $\mathcal{H}$ , defined as

$$\begin{aligned} V_X(g) &= E\{(X - EX) \otimes (X - EX)\}(g) \\ &= E\{\langle X - EX, g \rangle (X - EX)\}, \quad \text{for } \forall g \in \mathcal{H}. \end{aligned}$$

It is easy to show that  $V_X$  is a self-adjoint, nonnegative definite operator.

The characteristic function of a Hilbertian random variable is

$$\phi_X(f) = E\{\exp(i\langle f, X \rangle)\}, \quad (1)$$

for all  $f \in \mathcal{H}$ .

For a separable Hilbert space, there is a countable basis,  $\{e_1, e_2, \dots\}$ . Define  $x_j = \langle e_j, X \rangle$ , they are univariate random variables. Then  $X$  has the coordinate  $(x_1, x_2, \dots)$ , which is a  $\ell^2$  random variable. Denote the space of  $\mathcal{H}$ -valued random variables as  $\mathcal{X}$ , then  $\mathcal{X}$  is isomorphic to the space of  $\ell^2$  random variables.

An operator  $T$  is *nuclear* if the trace of it is finite and independent of the choice of the basis. The trace of an operator is defined as

$$\text{tr}(T) = \sum_{i=1}^{\infty} \langle e_i, Te_i \rangle,$$

for any complete orthonormal basis of  $\mathcal{H}$ .

The covariance operator of  $X$  considered in this paper is a self-adjoint, non-negative definite, nuclear operator, denoted as  $\Gamma$ .  $\Gamma$  has a spectrum decomposition

$$\Gamma = \sum_j \lambda_j \psi_j \otimes \psi_j, \quad (2)$$

where  $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$  are eigenvalues of  $\Gamma$ ,  $\psi_j$  are corresponding eigenvectors. If  $\{\psi_j\}$  is incomplete, we can always make them into a complete basis for  $\mathcal{H}$  by including basis of the null space of  $\Gamma$ .  $\psi_j$ 's are the principal components of the random vector  $X$ .

**DEFINITION 1** *A Hilbertian random variable  $X$  has an elliptically contoured distribution if the characteristic function of  $X - \mu$  has the form  $\phi_{X-\mu}(f) = \phi_0(\langle f, \Gamma f \rangle)$  for a univariate function  $\phi_0$ , where  $\Gamma$  is a self-adjoint, non-negative definite, nuclear operator on  $\mathcal{H}$ . Denote the distribution of  $X$  as  $\text{HEC}_{\mathcal{H}}(\mu, \Gamma, \phi)$ .*

One important example of elliptically contoured distribution is the Gaussian distribution, whose characteristic function has the form  $\phi_{X-\mu}(f) = \exp(-\langle f, \Gamma f \rangle/2)$ .

### 3 Random representation for Hilbertian valued elliptically contoured random variables

For a fixed self-adjoint, non-negative definite, nuclear operator  $\Gamma$ , we can define a metric in  $\mathcal{H}$ ,  $d_{\Gamma}(x, y) = \|x - y\|_{\Gamma} = \langle x - y, \Gamma(x - y) \rangle_{\mathcal{H}}^{1/2}$ .

**LEMMA 2** *Suppose  $\phi_{X-\mu}(f) = \phi_0(\langle f, \Gamma f \rangle)$  is the characteristic function of an elliptically contoured distribution, then  $\phi_0(\cdot)$  is a non-negative definite function on  $\mathcal{H}$  with respect to the metric  $d_{\Gamma}(\cdot, \cdot)$ , i.e.*

$$\sum_{j=1}^n a_i a_j \phi_0\{d_{\Gamma}^2(f_i, f_j)\} \geq 0,$$

for any finite collections of  $\{f_i; i = 1, \dots, n\} \subset \mathcal{H}$  and for any real values  $\{a_i; i = 1, \dots, n\}$ .

Lemma 2 is a straight forward application of the Sazanov theorem (Vakhania, Tarieladze and Chobemyan, 1987), which is a generalization of the Bochner's Theorem to the infinity dimensional Hilbert space.

By the definition of the characteristic function, one can easily see that  $\phi_{X-\mu}(f-g) = \overline{\phi_{X-\mu}(g-f)}$ , which leads to  $\phi_0(\|f-g\|_\Gamma^2) = \overline{\phi_0(\|f-g\|_\Gamma^2)}$ . This means  $\phi_0(\cdot)$  must be real valued.

**THEOREM 3**  $X \sim \text{HEC}_{\mathcal{H}}(\mu, \Gamma, \phi)$ , if and only if

$$X \stackrel{d}{=} \mu + RU, \quad (3)$$

where  $U \sim \text{Gaussian}(0, \Gamma)$ ,  $R$  is a nonnegative univariate random variable independent of  $U$ .

**Proof:** We first prove the "if" part. Suppose (3) is true, and let  $F$  be the distribution function of  $R$ , then

$$\phi_{X-\mu}(f) = E \exp(iR\langle f, U \rangle) = \int_{[0, \infty)} \exp(-r^2\langle f, \Gamma f \rangle/2) dF(r). \quad (4)$$

By Definition 1,  $X$  is a Hilbertian-valued elliptically contoured random variable.

Conversely, suppose  $X$  is an elliptically contoured Hilbertian random variable with the characteristic function  $\phi_{X-\mu}(f) = \phi_0(\|f\|_\Gamma)$ , by Lemma 2  $g(t) = \phi_0(t^2)$  is a positive definite function. By Theorem 2 in Schoenberg (1938),

$$g(t) = \int_0^\infty \exp(-t^2 u^2) d\alpha(u),$$

for some bounded non-decreasing function  $\alpha(u)$ . By the definition of the characteristic function, we have  $1 = \phi_0(0) = \int_0^\infty d\alpha(u)$ . Therefore,  $\alpha(\cdot)$  is the cumulative distribution function of a non-negative random variable. We now change variable, let  $t = \|f\|_\Gamma$  and define a random variable  $R$ , such that  $2^{-1/2}R$  has distribution function  $\alpha(\cdot)$ . Let  $F$  be the distribution function of  $R$ , then  $F(r) = \alpha(2^{-1/2}r)$ . We have

$$\phi_{X-\mu}(f) = \phi_0(\langle f, \Gamma f \rangle) = \int_0^\infty \exp(-r^2\langle f, \Gamma f \rangle/2) dF(r),$$

Therefore,  $X$  has the stochastic representation (3).

## 4 Properties of elliptically contoured distribution

We first discuss moment properties of elliptically contoured distribution. Suppose the first two moments of  $X$  exist. By (3),

$$EX = \mu, \quad V(X) = E(R^2)\Gamma.$$

On the other hand, if we start from the characteristic function, assuming that  $\phi$  is twice differentiable,

$$V(X) = \phi_{X-\mu}^{(2)}(0) = 2\phi_0'(\langle f, \Gamma f \rangle)\Gamma + 4\phi_0^{(2)}(\langle f, \Gamma f \rangle)(\Gamma f) \otimes (\Gamma f)|_{f=0} = 2\phi_0'(0)\Gamma.$$

To make  $\Gamma$  identifiable, we can let  $E(R^2) = 2\phi_0'(0) = 1$ , then  $\Gamma = V(X)$  is the covariance operator of  $X$ .

**THEOREM 4** *Let  $\mathcal{H}$ ,  $\mathcal{H}_1$  and  $\mathcal{H}_2$  be separable Hilbert spaces, suppose  $X \sim \text{HEC}_{\mathcal{H}}(\mu, \Gamma, \phi_0)$ ,  $P_1 \in \mathcal{B}(\mathcal{H}, \mathcal{H}_1)$ ,  $P_2 \in \mathcal{B}(\mathcal{H}, \mathcal{H}_2)$  are two bounded operators. Define  $X_i = P_i X$ ,  $\mu_i = P_i \mu$ ,  $\Gamma_{ij} = P_i \Gamma P_j^*$ , for  $i, j = 1, 2$ . Suppose  $\Gamma_{12} = 0$ . Then  $X_1 \sim \text{HEC}_{\mathcal{H}_1}(\mu_1, \Gamma_{11}, \phi_0)$ .*

*If  $\Gamma_{22}$  is a finite dimensional operator, then*

$$X_1|X_2 \sim \text{HEC}_{\mathcal{H}_1}(\mu_1, \Gamma_{11}, \phi_{T(X_2)}), \quad (5)$$

*where  $\phi_{T(X_2)}(t^2)$  is a non-negative definite function depends on  $T(X_2) = \langle X_2 - \mu_2, \Gamma_{22}^-(X_2 - \mu_2) \rangle_{\mathcal{H}_2}^{1/2}$ ,  $\Gamma_{22}^-$  is a generalized inverse of  $\Gamma_{22}$ .*

*If  $\Gamma_{22}$  is an infinite dimensional operator, then*

$$X_1|X_2 \sim \text{Gaussian}_{\mathcal{H}_1}(\mu_1, r^2(X_2)\Gamma_{11}), \quad (6)$$

*where  $r(\cdot)$  is a deterministic function given in (7).*

**Proof:** By (3),  $P_i X = P_i \mu + R U_i$ , where  $U_i = P_i U \sim \text{Gaussian}(0, \Gamma_{ii})$ , for  $i = 1, 2$ . Therefore,  $X_1 \sim \text{HEC}_{\mathcal{H}_1}(\mu_1, \Gamma_{11}, \phi_0)$  by Theorem 3.

Since  $\text{Cov}(U_1, U_2) = P_1 \Gamma P_2^* = 0$ , by the property of Gaussian variables,  $U_1$  is independent of  $U_2$ .  $X_1$  depends on  $X_2$  only though the information on  $R$  provided by  $X_2$ .

Suppose  $\Gamma_{22}$  is finite dimensional, then

$$X_2|R \sim \text{Gaussian}_{\mathcal{H}_2}(\mu_2, R^2\Gamma_{22}).$$

Notice that  $\Gamma_{22}$  is a finite dimensional operator (matrix) with a generalized inverse  $\Gamma_{22}^-$ . From the theory of finite dimensional Gaussian,  $T(X_2)$  is a sufficient statistic for  $R$ , i.e.  $X_2|T(X_2)$  is independent of  $R$ . Therefore,

$$X_1|X_2 \stackrel{d}{=} P_1\mu + U_1 \times \{R|T(X_2)\}.$$

(5) is obtained by Theorem 3.

On the other hand, if  $\Gamma_{22}$  is infinite dimensional, we claim that  $X_2$  provide all information about  $R$ . It is easy to see that  $\Gamma_{22}$  is self-adjoint, non-negative definite and nuclear, therefore it has a spectrum decomposition,

$$\Gamma_{22} = \sum_{j=1}^{\infty} \lambda_j \psi_j \otimes \psi_j,$$

where  $\lambda_j$ 's are the positive eigenvalues of  $\Gamma_{22}$ . Define  $r_n(X_2) = \frac{1}{n} \sum_{j=1}^n \lambda_j^{-1/2} \langle X_2 - \mu_2, \psi_j \rangle_{\mathcal{H}_2}$ . Notice that  $X_2 - \mu_2 = RU_2$ , and therefore  $\lambda_j^{-1/2} \langle X_2 - \mu_2, \psi_j \rangle_{\mathcal{H}_2} \stackrel{d}{=} RU_{2j}$ , where  $U_{2j}$  are i.i.d Normal(0, 1) independent of  $R$ . By Law of Large Numbers,

$$r(X_2) = \lim_{n \rightarrow \infty} r_n(X_2) = R \quad (7)$$

with probability 1. Therefore,  $X_1|X_2 \stackrel{d}{=} \mu_1 + r(X_2)U_1$  which is the Gaussian distribution in (6).

Theorem 4 gives the conditional distribution of  $X_1$  given  $X_2$  when they are uncorrelated, i.e.  $\Gamma_{12} = \text{Cov}(P_1X, P_2X) = 0$ . The following corollary gives the condition distribution for the more general cases.

**COROLLARY 5** *Let  $\mathcal{H}$ ,  $\mathcal{H}_1$  and  $\mathcal{H}_2$  be separable Hilbert spaces, suppose  $X \sim \text{HEC}_{\mathcal{H}}(\mu, \Gamma, \phi_0)$ ,  $P_1 \in \mathcal{B}(\mathcal{H}, \mathcal{H}_1)$ ,  $P_2 \in \mathcal{B}(\mathcal{H}, \mathcal{H}_2)$ , define  $X_i = P_iX$ ,  $\mu_i = P_i\mu$ ,  $\Gamma_{ij} = P_i\Gamma P_j^*$ , for  $i, j = 1, 2$ . Define  $\mu_1^* = \mu_1 + \Gamma_{12}\Gamma_{22}^-(X_2 - \mu_2)$  and  $\Gamma_{11}^* = \Gamma_{11} - \Gamma_{12}\Gamma_{22}^-\Gamma_{21}$ .*

*If  $\Gamma_{22}$  is a finite dimensional operator,*

$$X_1|X_2 \sim \text{HEC}_{\mathcal{H}_1}(\mu_1^*, \Gamma_{11}^*, \phi_{T(X_2)}), \quad (8)$$

*where  $\phi_{T(X_2)}(t^2)$  is a non-negative definite function depends on  $T(X_2) = \langle X_2 - \mu_2, \Gamma_{22}^-(X_2 - \mu_2) \rangle_{\mathcal{H}_2}^{1/2}$ .*

If  $\Gamma_{22}$  is an infinite dimensional operator, then

$$X_1|X_2 \sim \text{Gaussian}_{\mathcal{H}_1}(\mu_1^*, r^2(X_2)\Gamma_{11}^*), \quad (9)$$

where  $r(\cdot)$  is a deterministic function given in (7).

**Proof:** First of all,  $\Gamma_{ij} = \text{Cov}(X_i, X_j)$  are bounded operators, like in multivariate analysis, by Cauchy's inequality, we have  $\Gamma_{11}^*$  bounded and positive semidefinite. Also,  $\mu_1^*$  is well defined, since  $\text{Cov}(\mu_1^*) = \Gamma_{12}\Gamma_{22}^-\Gamma_{21}$  is bounded by  $\Gamma_{11}$  and therefore  $P(\mu_1^* \in \mathcal{H}_1) = 1$ .

Let  $U_i = P_i U$ ,  $i = 1, 2$ . Since  $U_i$ 's are Gaussian, it is easy to check through moment calculations that

$$(U_1, U_2) \stackrel{d}{=} (Z_1 + \Gamma_{12}\Gamma_{22}^-Z_2, Z_2),$$

where  $Z_1 \sim \text{Gaussian}_{\mathcal{H}_1}(0, \Gamma_{11}^*)$ ,  $Z_2 \sim \text{Gaussian}_{\mathcal{H}_2}(0, \Gamma_{22})$ , and they are independent. Therefore

$$\begin{aligned} X_1|X_2 &\stackrel{d}{=} \{\mu_1 + RZ_1 + \Gamma_{12}\Gamma_{22}^-(RZ_2)\}|X_2 \\ &\stackrel{d}{=} \mu_1 + \Gamma_{12}\Gamma_{22}^-(X_2 - \mu_2) + RZ_1|X_2 \\ &\stackrel{d}{=} \mu_1^* + Z_1 \times (R|X_2). \end{aligned} \quad (10)$$

Here we used the fact that  $Z_1$  is independent of  $X_2$ .

When the range of the operator  $\Gamma_{22}$  is finite dimensional,  $X_2$  is also finite dimensional. Use the arguments like in the proof of Theorem 4, we can show that  $(R|X_2)$  is a nonnegative random variable, which depends on value of  $X_2$  only through the statistic  $T(X_2)$ . Then (8) follows from a direct application of Theorem 3.

When  $\Gamma_{22}$  is of infinite dimension, like in the proof of Theorem 4,  $(R|X_2)$  is a deterministic function  $r(X_2)$  given by (7). (9) is proved.

**Remark:** Although, in this paper, we are only interested in the case that  $X$  is defined on an infinite dimensional Hilbert space  $\mathcal{H}$ , we do allow the Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  that  $P_1$  and  $P_2$  mapped into to be finite dimensional. For example, we allow  $\mathcal{H}_1$  and  $\mathcal{H}_2$  to be the Euclidian space  $\mathcal{R}^m$ . See our examples in Section 5.



## 5 Applications

To show the usefulness of our theory in statistical practices, we will provide a few examples which are direct results of the theorems in Section 4.

**Example 1: (Principal Component Analysis)** Suppose  $X \sim \text{HEC}_{\mathcal{H}}(\mu, \Gamma, \phi_0)$ , and the covariance operator  $\Gamma$  has the spectrum decomposition as in (2).  $\psi_j$ 's are the principal components of  $X$ . Define the principal component score  $\xi_j = \langle \psi_j, X \rangle_{\mathcal{H}}$  for  $j = 1, 2, \dots$ , then  $X$  has the following decomposition

$$X = \mu + \sum_{j=1}^{\infty} \xi_j \psi_j. \quad (11)$$

Such decomposition in functional analysis is also called the Karhunen-Loéve decomposition (Ash and Gardner, 1975).

For any finite collection,  $\{\psi_{j_1}, \dots, \psi_{j_m}\}$ , define an operator  $P : \mathcal{H} \rightarrow \mathcal{R}^m$  by  $P = \psi_{j_1} \otimes e_1 + \dots + \psi_{j_m} \otimes e_m$ , where  $e_j$  are the  $j^{\text{th}}$  column vector of the identity matrix. Then by Theorem 4,

$$(\xi_{j_1}, \dots, \xi_{j_m})^{\text{T}} = P(X - \mu) \sim \text{HEC}_{\mathcal{R}^m}\{0, \text{diag}(\lambda_{j_1}, \dots, \lambda_{j_m}), \phi_0\}.$$

In other words, any finite collection of principal component scores of  $X$  follows a multivariate elliptically contoured distribution.

This example also suggest a way to simulate a  $\text{HEC}_{\mathcal{H}}(\mu, \Gamma, \phi)$  random variable. Since  $\lambda_j \rightarrow 0$  as  $j \rightarrow \infty$ , we can truncate the series in (11) at a large number  $m$  and simulate the first  $m$  principal component scores from a multivariate elliptical contoured distribution. This is very useful in simulating functional data.

**Example 2: (Conditional moments)** Let  $X \sim \text{HEC}_{\mathcal{H}}(0, \Gamma, \phi)$ ,  $P_1 \in \mathcal{B}(\mathcal{H}, \mathcal{H}_1)$ ,  $P_2 \in \mathcal{B}(\mathcal{H}, \mathcal{H}_2)$ . By Corollary 5,

$$E(P_1 X | P_2 X) = \mu_1 + \Gamma_{12} \Gamma_{22}^{-1} (X_2 - \mu_2).$$

Suppose  $\mu = 0$ ,  $P_1 \Gamma P_2^* = 0$ , then  $\Gamma_{12} = 0$  and we have

$$E(P_1 X | P_2 X) = 0.$$

On the other hand, by the random representation (10),

$$\text{Var}(P_1 X | P_2 X) = \Gamma_{11}^* E(R^2 | X_2).$$

When  $\Gamma_{22}$  is a finite dimensional operator,  $E(R^2 | X_2) = g\{T(X_2)\}$  for some univariate function  $g$  depends on the elliptically contoured distribution,  $T(X_2)$  is defined in Theorem 4. Therefore

$$\text{Var}(P_1 X | P_2 X) = g\{T(X_2)\} \Gamma_{11}^*.$$

When  $\Gamma_{22}$  is an infinite dimensional operator, by (9),

$$\text{Var}(P_1 X | P_2 X) = r^2(X_2) \Gamma_{11}^*.$$

If  $P_1 \Gamma P_2^* = 0$ ,  $\Gamma_{11}^* = \Gamma_{11}$ .

**Example 3: (Functional sliced inverse regression)** Suppose the Hilbert space  $\mathcal{H}$  is the  $L^2[0, 1]$  functional space,  $X \in \mathcal{H}$  are random functions defined on the  $[0, 1]$  interval. A general functional regression model is given by

$$Y = f(\langle \beta_1, X \rangle, \langle \beta_2, X \rangle, \dots, \langle \beta_K, X \rangle, \epsilon), \quad (12)$$

where  $Y$  is a scalar response variable,  $\epsilon$  is the error term independent with  $X$ ,  $\beta_1, \dots, \beta_K$  are linearly independent coefficient function,  $f$  is a nonparametric link function. Model (12) is very general, it can be very useful in many applications, see the discussion in Ferré and Yao (2003) and Li and Hsing (2007). Since we do not impose any structure on the link function  $f$ , the coefficient functions  $\beta_k$ 's are usually unidentifiable, but the subspace spanned by these function is. This subspace is called the Effective Dimension Reduction space, or the EDR space. We can chose any  $K$  orthonormal basis functions in the EDR space as  $\beta_k$ 's, these functions are also called the EDR directions.

The functional sliced inverse regression (FSIR) approach can be used to estimate the EDR directions and to decide the dimension of EDR space. We will show that the class of process  $X$  with a Hilbertian Elliptically Contoured distribution satisfies a key assumption for (FSIR), and we will discuss an important result for elliptically contoured functional predictor, which is useful for FSIR. For a more comprehensive

account for the method and theory of FSIR, we refer to Ferré and Yao (2003), Li and Hsing (2007).

One key assumption for FSIR is that, for any  $\beta_0 \in \mathcal{H}$ , then

$$E(\langle \beta_0, X \rangle | \langle \beta_1, X \rangle, \langle \beta_2, X \rangle, \dots, \langle \beta_K, X \rangle) = c_0 + \sum_{k=1}^K c_k \langle \beta_k, X \rangle \quad (13)$$

for some constants  $c_0, \dots, c_K$ , see Ferré and Yao (2003). We will show this assumption is satisfied if  $X$  is elliptically contoured.

Define operators  $P_1x = \langle \beta_0, x \rangle$ ,  $P_2x = (\langle \beta_1, x \rangle, \langle \beta_2, x \rangle, \dots, \langle \beta_K, x \rangle)^T$  for  $\forall x \in \mathcal{H}$ . Notice that  $\mathcal{H}_2 = \mathcal{R}^K$ , and  $P_2$  is clearly a finite dimensional operator. For any vector  $v = (v_1, \dots, v_K)^T$ ,

$$\langle P_2^*v, x \rangle_{\mathcal{H}} = \langle v, P_2x \rangle_{\mathcal{H}_2} = \left\langle \sum_{k=1}^K v_k \beta_k, x \right\rangle, \quad \forall x \in \mathcal{H}.$$

Therefore,  $P_2^*v = (\beta_1, \dots, \beta_K)v$ . One can also show that  $\Gamma_{22}$  is a  $K \times K$  matrix with the  $(j, k)^{\text{th}}$  entry equal to  $\langle \beta_j, \Gamma \beta_k \rangle$ . Similarly,  $\Gamma_{12} = (\langle \beta_0, \Gamma \beta_1 \rangle, \dots, \langle \beta_0, \Gamma \beta_K \rangle)$  is a  $1 \times K$  matrix. By Corollary 5,

$$E(P_1X | P_2X) = P_1\mu + \Gamma_{12}\Gamma_{22}^-(P_2X - P_2\mu).$$

Therefore, assumption (13) is satisfied.

**Example 4: (Functional sliced inverse regression, continued)**

Suppose  $X$  is elliptically contoured, with mean  $\mu = 0$ . Let  $P_2$  be the operator defined as in the previous example, define the operator  $P_1x = (\langle \gamma_1, x \rangle, \dots, \langle \gamma_m, x \rangle)^T$  for a set of orthonormal vectors in  $\mathcal{H}$ ,  $\{\gamma_1, \dots, \gamma_m\}$ . Suppose  $P_1\Gamma P_2^* = 0$ , i.e.  $\langle \gamma_j, \Gamma \beta_k \rangle = 0$  for  $j = 1, \dots, m$ ,  $k = 1, \dots, K$ .

By model (12), all the information in  $Y$  about  $X$  are contained in  $P_2X$ , we have

$$E(P_1X | Y) = E\{E(P_1X | P_2X) | Y\}.$$

By Example 2,  $E(P_1X | P_2X) = 0$ , therefore

$$P_1E(X | Y) = 0. \quad (14)$$

Equation (14) provides information about the shape of the inverse regression curve  $E(X|Y)$ .

On the other hand,

$$\begin{aligned}\text{Var}(P_1X|Y) &= E\{\text{Var}(P_1X|P_2X)|Y\} + \text{Var}\{E(P_1X|P_2X)|Y\} \\ &= E\{\text{Var}(P_1X|P_2X)|Y\}. \quad (\text{by } E(P_1X|P_2X) = 0)\end{aligned}$$

Again, by Example 2,  $\text{Var}(P_1X|P_2X) = g\{T(P_2X)\}\Gamma_{11}$ , therefore

$$\text{Var}(P_1X|Y) = E[g\{T(P_2X)\}|Y]\Gamma_{11}.$$

Since  $\Gamma_{11}$  is the marginal covariance for  $P_1X$ , this result shows that the conditional covariance of  $P_1X$  given  $Y$  is proportional to the the marginal covariance. This result is important to constructing tests for FSIR.

#### REFERENCES

- [1] Ash, R. B. and Gardner, M. F. (1975). *Topics in stochastic processes*, Academic press.
- [2] Cambanis, S., Huang, S. and Simons, G. (1981). On the theory of elliptically contoured distributions, *Journal of Multivariate Analysis*, 11, 368-385.
- [3] Cook, D. R. and Weisberg, S. (1991). Comments on "Sliced Inverse Regression for Dimension Reduction," by K. C. Li, *Journal of the American Statistical Association*, 86, 328-332.
- [4] Eaton, M. L. (1986). A characterization of spherical distributions', *Journal of Multivariate Analysis*, 20, 272-276.
- [5] Ferré, L. and Yao, A. (2003). Functional sliced inverse regression analysis. *Statistics*, 37, 475-488.
- [6] Li, K. (1991). Sliced inverse regression for dimension reduction. *Journal of the American Statistical Association*, 86, 414, 316-327.
- [7] Li, Y. and Hsing, T. (2007). Determination of the Dimensionality in Functional Sliced Inverse Regression, *manuscript*.

- [8] Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis, 2nd Edition*. Springer-Verlag, New York.
- [9] Schoenberg, I. J. (1938). Metric spaces and completely monotone functions, *Annals of Mathematics*, 39, No.4, 811-841.
- [10] Vakhania, N. N., Tarieladze, V. I. and Chobemyan, S.A. (1987). *Probability Distributions on Banach Spaces*. D. Reidel, Dordrecht.