Statistica Sinica **17**(2007), 1683-1687

REPLY TO THE PAPER BY LILIANA FORZANI AND R. DENNIS COOK : "A NOTE ON SMOOTHED FUNCTIONAL INVERSE REGRESSION"

L. Ferré and A. F. Yao

Université Toulouse Le Mirail and Université de la Méditerranée

Abstract: Ferré and Yao (2005) proposed an approach to the estimation of the Effective Dimension Reduction space in functional sliced inverse regression. This did not require the inversion of the variance-covariance operator of the explanatory variable, and thus circumvent the main drawback of SIR in the functional case. Forzani and Cook (2007) argued that our result is possible only under additional conditions. Although our discussion may have created ambiguity, the result is correct and does not require further assumptions. In this note, we give the additional details needed to make our purpose clear.

Key words and phrases: Dimension reduction, functional data analysis, inverse regression.

1. Introduction

In Ferré and Yao (2005), we addressed the issue of the estimation of the Effective Direction Reduction space in functional Sliced Inverse Regression. As in many functional statistical problems, one of main drawback is that, while the covariance operator of the explanatory functional variable X is invertible, it has an unbounded inverse so that its estimator is ill-conditioned. We assume that X is square integrable so that, for the particular problem, the covariance operators mentioned are well-defined. Then, we proposed a solution where the covariance operator need not be inverted. Indeed, if Γ (respectively Γ_e) denotes the covariance operator of X (resp. of $E(X|Y)$, where Y is the real response variable), while the EDR space is usually derived from the eigen decomposition of $\Gamma^{-1}\Gamma_e$, we use $\Gamma_e^+\Gamma$, where Γ_e^+ is the Moore-Penrose generalized inverse of Γ_e.

Forzani and Cook (2007) claim that our result does not hold without additional assumptions. For illustration purpose, they give two examples to suggest that, in general, "the sentence the eigenvectors of $\Gamma^{-1}\Gamma_e$ are the ones of $\Gamma_e^+\Gamma$ as stated in Ferré and Yao (2005) is not correct". We agree with them concerning the general case but, in the particular context of Sliced Inverse Regression, this assertion is true as it will be proved in the next section.

Let us first focus on Forzani and Cook's finite dimensional example.

Recall that we are dealing with the (multivariate or functional) SIR model

$$Y = g(\langle b_1,\, X\rangle, \ldots, \langle b_D,\, X\rangle, \varepsilon), \tag{1}$$

and assume the linearity condition of Ferré and Yao (2005).

For any real response Y, and particularly under the model above, one has $\Gamma = \Gamma_e + \Lambda$, where $\Lambda = E(\mathrm{Var}\,(X|Y))$ is a positive definite operator (in either the finite or infinite dimensional case).

The finite-dimensional example of Forzani and Cook is not compatible with the SIR model, since we have : $\Gamma = \begin{pmatrix} 2\,1 \\ 1\,4 \end{pmatrix}$, $\Gamma_e = \begin{pmatrix} 2\,0 \\ 0\,0 \end{pmatrix}$, and then $\Lambda = \Gamma - \Gamma_e = \begin{pmatrix} 0\,1 \\ 1\,4 \end{pmatrix}$ is not positive definite.

This illustrates only the known fact that the proposal is uncorrect in general, but it does not prove that it is uncorrect under (1).

On the other hand, the examples treated (simulated or real) in Ferré and Yao (2005), Ferré and Villa (2005) and in Amato, Antoniadis and De Feis (2006) suggest that our approach leads to convenient solutions. In Amato et al. (2006) a similar solution is proposed except that the conditional covariance operator is estimated from wavelets instead of kernels. Concerning their simulated example based on Brownian motion, the results are not surprising since their EDR space is generated by some eigenfunctions of the covariance operator of the Brownian motion and, thus, the assumption added by Forzani and Cook is straightforwardly satisfied. This is no longer true for our simulations and we have a priori no idea for real data, but it works. Thus, the underlying properties of the model are responsible for the possibility of estimating the EDR space from $\Gamma_e^{+}\Gamma$.

In the following section, we present lemmas and a theorem to this effect in the context of multivariate and functional SIR, and without any additional assumptions. Moreover, we prove that Condition A of Forzani and Cook is satisfied under Model (1). The results of Ferré and Yao (2005) are thus clarified.

2. Additional Results

We first introduce some notations. Let

- $\beta_1, \ldots, \beta_D$ be the Γ-orthonormed eigenvectors of $\Gamma^{-1}\Gamma_e$ associated with strictly positive eigenvalues $(\lambda_i)_{i=1,\ldots,D}$;
- $\Pi^{\Gamma}_{S_D} = \sum_{i=1}^{D} \Gamma\beta_i \otimes \beta_i$ be the $\Gamma-$orthogonal projector onto the space $S_D = R(\Gamma^{-1}\Gamma_e)$;
- $\Pi^{\Gamma *}_{S_D}$ be the adjoint of $\Pi^{\Gamma}_{S_D}$, *i.e.*, $\Pi^{\Gamma *}_{S_D} = \sum_{i=1}^{D} \langle \beta_i, .\rangle \Gamma\beta_i$;
- $\Pi_{R(\Gamma_e)}$ be the orthogonal projector onto $R(\Gamma_e)$.

If we complete the basis $(\beta_i)_{i=1,\ldots,D}$ to obtain a basis of $\Gamma-$orthonormed vectors of H, we have $\Pi^{\Gamma}_{S_D^{\perp_\Gamma}} = \sum_{j>D} \Gamma\beta_j \otimes \beta_j$ (where $S_D^{\perp_\Gamma}$ is the space $\Gamma-$orthogonal to S_D) . Let $\Pi^{\Gamma *}_{S_D^{\perp_\Gamma}}$ be its adjoint.

Lemma 2.1. *For all* $b \in H$*, we have* $\Pi_{R(\Gamma_e)} b = \Pi_{R(\Gamma_e)}.\Pi^{\Gamma *}_{S_D} b$.

Proof. Let $b \in H$, so $b = \Pi^{\Gamma *}_{S_D} b + \Pi^{\Gamma *}_{S_D^{\perp_\Gamma}} b$ (by identifying H and its dual). Thus, $\Pi_{R(\Gamma_e)} b = \Pi_{R(\Gamma_e)}.\Pi^{\Gamma *}_{S_D} b + \Pi_{R(\Gamma_e)}.\Pi^{\Gamma *}_{S_D^{\perp_\Gamma}} b$.

Since for a fixed, $j > D$, $\langle \beta_i, \Gamma\beta_j\rangle = 0$ for all $i \in \{1,\ldots,D\}$, we have $\mathbf{E}(\langle \beta_j, X\rangle|Y) = 0$ a.s. (see Dauxois, Ferré and Yao (2001)). So, $\mathbf{E}(\Pi^{\Gamma *}_{S_D^{\perp_\Gamma}} X\,|Y) = 0$ a.s. and $\mathbf{E}(X|Y) = \mathbf{E}(\Pi^{\Gamma *}_{S_D} X|Y)$ a.s.

Then,

$$\Gamma_e = \mathbf{var}(\mathbf{E}(X|Y)) = \mathbf{var}(\mathbf{E}(\Pi^{\Gamma *}_{S_D} X|Y)) = \Pi^{\Gamma}_{S_D} \Gamma_e \Pi^{\Gamma *}_{S_D}, \tag{2}$$

which implies that $R(\Gamma_e) \subset S_D$ and, since $\Pi^{\Gamma *}_{S_D^{\perp_\Gamma}} b = \sum_{i>D} \langle \beta_j, b\rangle \Gamma\beta_j$, we have $\Pi_{R(\Gamma_e)}.\Pi^{\Gamma *}_{S_D^{\perp_\Gamma}} b = 0$ so $\Pi_{R(\Gamma_e)} b = \Pi_{R(\Gamma_e)}.\Pi^{\Gamma *}_{S_D} b$.

Remark. This result is interesting in that the second example (in infinite dimensions) of Forzani and Cook cannot hold under the SIR model.

Indeed, suppose that Model (1) holds in the context of their example. Then we are dealing with a single index model (*i.e.*, $i = 1$). Now, according to their assumptions, we get $\beta_1 = f/\sqrt{\langle \Gamma f, f\rangle}$ and, as they suggest, $\Gamma_e = \Pi_{R(\Gamma_e)} = h \otimes h/\|h\|^2$.

But, by using the previous Lemma (more precisely equality (2) in the proof), we have $\Gamma_e = \Pi_{R(\Gamma_e)} = (\|h\|^2/\langle \Gamma f, f\rangle)\langle f, .\rangle f$, which implies that there exists a constant c such that $h = c\,f$ and then, necessarily as they show, if $f = \sum_{i=1}^{\infty} a_i\phi_i$ with $\sum_{i=1}^{\infty} a_i^2 < \infty$, there exists an $a_i = 0$, which is a condition required for Model (1) to hold, in contradiction to the Forzani and Cook statement.

Lemma 2.2. *For* $i = 1,\ldots,D$, $\langle \beta_i, X\rangle = \langle \Pi_{R(\Gamma_e)}\beta_i, X\rangle$.

Proof. Since for all $i = 1,\ldots,D$, $\Gamma^{-1}\Gamma_e\beta_i = \lambda_i\beta_i$, $\Gamma_e\beta_i = \lambda_i\Gamma\beta_i$, and then $\Gamma\beta_i \in R(\Gamma_e)$ (and $\Gamma\beta_i = \Pi_{R(\Gamma_e)}\Gamma\beta_i$). So we have $\beta_i = \Pi^{\Gamma}_{S_D}\beta_i = \sum_i^D \langle \Gamma\beta_i, \beta_i\rangle\beta_i = \sum_i^D \langle \Gamma\beta_i, \Pi_{R(\Gamma_e)}\beta_i\rangle\beta_i = \Pi^{\Gamma}_{S_D}\Pi_{R(\Gamma_e)}\beta_i$ and

$$\langle \beta_i, X\rangle = \langle \Pi^{\Gamma}_{S_D}\Pi_{R(\Gamma_e)}\beta_i, X\rangle = \langle \beta_i, \Pi_{R(\Gamma_e)}.\Pi^{\Gamma *}_{S_D} X\rangle.$$

Now, applying the previous lemma to X, for all $i \in \{1,\ldots,D\}$, $\langle \beta_i, X\rangle = \langle \beta_i, \Pi_{R(\Gamma_e)}\Pi^{\Gamma *}_{S_D^{\Gamma}} X\rangle = \langle \Pi_{R(\Gamma_e)}\beta_i, X\rangle$, and the proof is complete.

Theorem 2.3. *For* $i = 1,\ldots,D$, β_i *is an eigenvector of* $\Gamma_e^{+}\Gamma\beta_i$.

Proof. From $\langle \beta_i, X\rangle = \langle \Pi_{R(\Gamma_e)}\beta_i, X\rangle$, we get $\Gamma\beta_i = E(\langle \Pi_{R(\Gamma_e)}\beta_i, X\rangle X) = \Gamma\,\Pi_{R(\Gamma_e)}\beta_i$, and then $\beta_i = \Pi_{R(\Gamma_e)}\beta_i$ which implies that $\beta_i \in R(\Gamma_e)$. Now, we have $\Gamma^{-1}\Gamma_e\beta_i = \lambda_i\beta_i$, and then $\Pi_{R(\Gamma_e)}\beta_i = \lambda_i\Gamma_e^+\Gamma\beta_i$.

This achieves the proof since $\Pi_{R(\Gamma_e)}\beta_i = \beta_i$.

The above theorem shows that, for $i = 1,\ldots,D$, the vector β_i is the eigenvector of $\Gamma_e^+\Gamma$ associated with eigenvalue $1/\lambda_i$. Since $R(\Gamma^{-1}\Gamma_e)$ and $R(\Gamma_e^+\Gamma)$ are, under Model (1), both $D-$dimensional subspaces that contain the D independent vectors $\beta_1,\ldots,\beta_D$, we have proved that in the framework of (multivariate or functional) SIR,

$$R(\Gamma^{-1}\Gamma_e) = R(\Gamma_e^+\Gamma).$$

Then, unlike the Forzani and Cook claim, this property holds without additional assumptions under Model (1). Moreover, $R(\Gamma_e)$ is also a $D-$dimensional subspace that contains the β_i's so that

$$R(\Gamma^{-1}\Gamma_e) = R(\Gamma_e^+\Gamma) = R(\Gamma_e),$$

and we find again the result obtained by Forzani and Cook.

In practice, since β_i belongs to $R(\Gamma_e)$ and $\Gamma_e\beta_i = \lambda_i\Gamma\beta_i$, we have $\Gamma_e^+\Gamma\beta_i = 1/\lambda_i\beta_i$. This leads to $\Gamma^{1/2}\Gamma_e^+\Gamma^{1/2}\nu_i = 1/\lambda_i\nu_i$ by letting $\nu_i = \Gamma^{1/2}\beta_i$. Then, the vectors ν_i are derived from the eigenvalue decomposition of $\Gamma^{1/2}\Gamma_e^+\Gamma^{1/2}$, and we have $\lambda_i\Gamma_e^+\Gamma^{1/2}\nu_i = \lambda_i\Gamma_e^+\Gamma\beta_i = \beta_i$. Now let Γ_n be the usual (empirical) estimate of Γ, and $\hat{\Gamma}_e$ a root-n consistent estimates of Γ_e. Then, for $i = 1,\ldots,n$, the estimation of β_i is achieved by $\hat{\beta}_i = \hat{\lambda}_i\hat{\Gamma}_e^+\Gamma_n^{1/2}\hat{\nu}_i$, where $\hat{\nu}_i$ is the eigenvector of $\Gamma_n^{1/2}\hat{\Gamma}_e^+\Gamma_n^{1/2}$ associated with eigenvalue $1/\hat{\lambda}_i$. This is the estimator proposed in Ferré and Yao (2005). Finally, we have $\Gamma_n^{1/2}\hat{\Gamma}_e^+\Gamma_n^{1/2} - \Gamma^{1/2}\Gamma_e^+\Gamma^{1/2} = O_p(1/\sqrt{n})$ and, by applying the results of the perturbation theory of linear operators, we achieve the root-n consistency of the estimates of the β's as indicated in Ferré and Yao (2005).

Acknowledgement

The authors are grateful to the Co-Editors and an associated editor for their valuable comments and suggestions.

References

Amato, U., Antoniadis, A. and De Feis, I. (2006). Dimension Reduction in Functional Regression with Applications. *CSDA* **50**, 2422-2446.

Dauxois, J., Ferré, L. and Yao, A. F. (2001) Un modèle semi-paramétrique pour variable aléatoire hilbertienne. *C. R. Acad. Sci. Paris* **t.327**, série I, 947-952.

Ferré L. and Villa N. (2005) Discrimination de Courbes par Regression Inverse Fonctionnelle. *Rev. Statistique Appliqués* **LIII**, 39-57.

Ferré, L. and Yao, A. F. (2005). Smoothed Functional Sliced Inverse Regression. *Statist. Sinica* **15**, 665-683.

Forzani, L. and Cook, R. D. (2007) A Note on Smoothed Functional Inverse Regression. *Statist. Sinica* to appear.

Université Toulouse Le Mirail, Institut de Mathématiques de Toulouse, UMR CNRS 5219, 5, allées Antonio Machado, 31058 Toulouse, Cedex 09, France.

E-mail: loferre@univ-tlse2.fr

Université de la Méditerranée, Centre d'Océanologie de Marseille, Campus de Luminy, Case 901, 13288 MARSEILLE, Cedex 09, France.

E-mail: anne-francoise.yao@univmed.fr

(Received June 2007; accepted June 2007)