Statistica Sinica **20** (2010), 235-238

NECESSARY AND SUFFICIENT CONDITIONS FOR CONSISTENCY OF A METHOD FOR SMOOTHED FUNCTIONAL INVERSE REGRESSION

R. D. Cook, L. Forzani and A. F. Yao

University of Minnesota, Instituto de Matemática Aplicada del Litoral and Université de la Méditerranée

Abstract: Ferré and Yao (2005, 2007) proposed a method to estimate the Effective Dimension Reduction space in functional sliced inverse regression. Their approach did not require the inversion of the variance-covariance operator of the explanatory variables, and it allowed them to get $\sqrt{n}$ consistent estimators in the functional case. In those papers there is a mistake. In this note we show that, in general, the approach does not give an estimator of the SIR subspace. We also give necessary and sufficient conditions for this to be true.

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

1. Introduction

For notation and background we refer to Ferré and Yao (2005, 2007) and Forzani and Cook (2007). Recall that we are dealing with the (multivariate or functional) SIR model:

$$Y = g(\boldsymbol{\beta}_1^T \mathbf{X}, \ldots, \boldsymbol{\beta}_d^T \mathbf{X}, \boldsymbol{\varepsilon}), \tag{1.1}$$

and that the linearity condition is assumed. It is also known that under the linearity condition the SIR subspace, the span of the $\boldsymbol{\Gamma}$-orthonormed non-zero eigenvectors of $\boldsymbol{\Gamma}^{-1}\boldsymbol{\Gamma}_e$, where $\boldsymbol{\Gamma}_e = \mathrm{Cov}\,(\mathrm{E}\,(\mathbf{X}|Y))$ and $\boldsymbol{\Gamma} = \mathrm{Cov}\,(\mathbf{X})$, form a subspace of the SIR subspace (Li (1991)) for the finite-dimensional case and Dauxois, Ferré and Yao (2001) for the functional case). Here $\boldsymbol{\Gamma}$ is a Hilbert-Schmidt operator and the inverse can be defined in its range. From now on $R(B)$ denotes the range of an operator B, which is the set of functions $B(f)$ with f belonging to the domain $T(B)$ of the operator B. Ferré and Yao (2007) claimed that under (1.1),

$$R(\boldsymbol{\Gamma}^{-1}\boldsymbol{\Gamma}_e) = R(\boldsymbol{\Gamma}_e^{-}\boldsymbol{\Gamma}). \tag{1.2}$$

We will prove in Theorem 1 that (1.2) is equivalent to the requirement that

$$\boldsymbol{\Gamma} = P_{\boldsymbol{\Gamma}_e}\boldsymbol{\Gamma}P_{\boldsymbol{\Gamma}_e} + Q_{\boldsymbol{\Gamma}_e}\boldsymbol{\Gamma}Q_{\boldsymbol{\Gamma}_e}, \tag{1.3}$$

and we will show with an example in Section 3 that (1.3) is not in general true under the model (1.1).

We prove later that (1.2) implies $R(\mathbf{\Gamma}^{-1}\mathbf{\Gamma}_e) = R(\mathbf{\Gamma}_e)$ and therefore if (1.2) were true, it would be not necessary to invert any matrix to get a consistent estimator of the SIR subspace and the problem of consistency would be, using the result by Dauxois et al. (2001), almost trivial even for the infinite dimensional case. The authors argued that the examples they treated (simulated or real) suggest that their approach leads to convenient solutions. But it should be noted that even if the central subspace is far away from the approximated one, the angle between the linear combination of the predictors and the estimate one is not necessarily large, and that may be the case for their examples.

Following Shao, Cook and Weisberg (2007) the SIR subspace is equal to the central subspace $\mathcal{S}_{Y|\mathbf{X}}$ (the central subspace is the smallest subspace S such that Y and X are independent conditionally to the projection of X on S), if the linearity of the predictors and an associated coverage condition both hold. Under these two conditions and as a consequence of the results of this paper, for model (1.1), $\mathcal{S}_{Y|\mathbf{X}} = \mathrm{span}(\mathrm{Cov}\,(\mathrm{E}\,(\mathbf{X}|Y)))$ if and only if (1.3) is true.

2. Equivalence

The following lemma gives a necessary and sufficient condition for (1.2).

Theorem 1. *Given a Hilbert-Schmidt operator $\mathbf{\Gamma}$ and a finite rank operator $\mathbf{\Gamma}_e$, (1.2) is equivalent to (1.3).*

That (1.3) implies (1.2) is trivial. The other implication follows from Lemmas 1, 2 and 3.

Lemma 1. *If (1.2) holds then $R(\mathbf{\Gamma}_e) = R(\mathbf{\Gamma}^{-1}\mathbf{\Gamma}_e)$.*

Proof. Given a set $B \subset L^2[0,1]$, denote by $B^{\perp}$ its orthogonal complement using the usual interior product in $L^2[a,b]$. The closure of the set B, denoted by $\bar{B}$, will be the smallest closed set (using the topology defined through the usual interior product) containing B. For an operator B from $L^2[a,b]$ into itself, let B^* denote its adjoint operator, again using the usual interior product.

Let $\{\beta_1,\ldots,\beta_D\}$ denote the D eigenfunctions, with eigenvalues nonzero, of $\mathbf{\Gamma}_e^{-}\mathbf{\Gamma}$. If (1.2) is true then $\mathrm{span}(\beta_1,\ldots,\beta_D) = R(\mathbf{\Gamma}^{-1}\mathbf{\Gamma}_e) = R(\mathbf{\Gamma}_e^{-}\mathbf{\Gamma}) \subset R(\mathbf{\Gamma}_e^{-})$. By definition of the generalized inverse (Groetsch (1977)) we have $R(\mathbf{\Gamma}_e^{-}) = N(\mathbf{\Gamma}_e)^{\perp} = \overline{R(\mathbf{\Gamma}_e^{*})} = \overline{R(\mathbf{\Gamma}_e)} = R(\mathbf{\Gamma}_e)$ where A^* is the adjoint of A and we use the facts that $\mathbf{\Gamma}_e$ is self-adjoint and that $R(\mathbf{\Gamma}_e)$ has dimension D and therefore is closed. Since $R(\mathbf{\Gamma}_e)$ has dimension D, the result follows.

Lemma 2. *Under (1.2) we have $R(\mathbf{\Gamma}\mathbf{\Gamma}_e) \subset R(\mathbf{\Gamma}_e)$.*

Proof. Since $\mathbf{\Gamma}$ is one to one, $\overline{R(\mathbf{\Gamma})} = L^2[a,b]$. On the other hand, by hypothesis, $R(\mathbf{\Gamma}_e) \subset T(\mathbf{\Gamma}^{-1})$. From the definition of the inverse of an operator (Groetsch (1977)) we have that $\mathbf{\Gamma}\mathbf{\Gamma}^{-1} = \mathcal{I}_d$ in $T(\mathbf{\Gamma}^{-1})$, where $\mathcal{I}_d$ indicates the identity operator. Now, take $v \in R(\mathbf{\Gamma}\mathbf{\Gamma}_e)$. Then $v = \mathbf{\Gamma}\mathbf{\Gamma}_e w$ for some $w \in L^2[a,b]$ and therefore $\mathbf{\Gamma}^{-1}v = \mathbf{\Gamma}_e w = \mathbf{\Gamma}^{-1}\mathbf{\Gamma}_e h$ for some $h \in L^2[a,b]$ (this last follows from Lemma 1). Since $\mathbf{\Gamma}^{-1}$ is one to one (in its domain) we get $v = \mathbf{\Gamma}_e h \in R(\mathbf{\Gamma}_e)$.

In mathematical terms, $R(\mathbf{\Gamma}\mathbf{\Gamma}_e) \subset R(\mathbf{\Gamma}_e)$ implies that $R(\mathbf{\Gamma}_e)$ is an invariant subspace of the operator $\mathbf{\Gamma}$ (see Conway (1990, p.39)). This implies that $\mathbf{\Gamma}$ has a spectral decomposition with eigenfunctions that live in $R(\mathbf{\Gamma}_e)$ or its orthogonal complement, as indicated formally in the following lemma. The finite-dimensional form of the lemma was stated by Cook, Li and Chiaromonte (2007).

Lemma 3. *Suppose* (1.2) *is true. Then* $\mathbf{\Gamma}$ *has a spectral decomposition with eigenfunctions on* $R(\mathbf{\Gamma}_e)$ *or* $R(\mathbf{\Gamma}_e)^{\perp}$.

Proof. Let v be an eigenvector of $\mathbf{\Gamma}$ associated to the eigenvalue $\lambda > 0$. Since $R(\mathbf{\Gamma}_e)$ is closed (being finite-dimensional), $v = u + w$ with $u \in R(\mathbf{\Gamma}_e)$ and $w \in R(\mathbf{\Gamma}_e)^{\perp}$. Since from Lemma 2, $\mathbf{\Gamma}u \in R(\mathbf{\Gamma}_e)$ and $\mathbf{\Gamma}w \in R(\mathbf{\Gamma}_e)^{\perp}$, we have that u and w are also eigenvectors of $\mathbf{\Gamma}$ if both u and w are different from zero. Otherwise v belongs to $R(\mathbf{\Gamma}_e)$ or $R(\mathbf{\Gamma}_e)^{\perp}$.

Now, let $\{v_i\}_{i=1}^{\infty}$ be a spectral decomposition of $\mathbf{\Gamma}$, countable since $\mathbf{\Gamma}$ is compact in $L^2[0,1]$. From what was said above, $v_i = u_i + w_i$ with u_i and w_i eigenvectors in $R(\mathbf{\Gamma}_e)$ and $R(\mathbf{\Gamma}_e)^{\perp}$, respectively. Now consider $\{u_i : u_i \neq 0\}$ and $\{w_i : w_i \neq 0\}$. Clearly they form a spectral decomposition of $\mathbf{\Gamma}$ with eigenfunctions on $R(\mathbf{\Gamma}_e)$ or $R(\mathbf{\Gamma}_e)^{\perp}$.

3. Example

Let us take the model considered by Cook (2007):

$$\mathbf{X}|(Y = y) = \mu + \boldsymbol{\alpha}\nu_y + \mathbf{\Delta}^{1/2}\boldsymbol{\varepsilon}, \tag{3.1}$$

where $\mu = \mathrm{E}(\mathbf{X})$, $\boldsymbol{\alpha} \in \mathbb{R}^{p\times d}$ and $\mathbf{\Delta}$ is any positive definite matrix $p \times p$. This is a particular case of (1.1). Using their notation we have $\mathbf{\Gamma}_e = \boldsymbol{\alpha}\,\mathrm{Cov}\,(\nu_Y)\boldsymbol{\alpha}^T$ and $\mathbf{\Gamma} = \mathrm{Var}\,(\mathbf{X}) = \mathbf{\Delta} + \boldsymbol{\alpha}\,\mathrm{Cov}\,(\nu_Y)\boldsymbol{\alpha}^T$. In this case the SIR subspace is $R(\mathbf{\Gamma}^{-1}\mathbf{\Gamma}_e) = R(\mathbf{\Delta}^{-1}\boldsymbol{\alpha})$, and it was proven by Cook (2007) that for (3.1) the minimal sufficient reduction is given by $R(\mathbf{\Delta}^{-1}\boldsymbol{\alpha}) = R(\mathbf{\Gamma}^{-1}\boldsymbol{\alpha})$. Therefore the central subspace and the SIR subspace coincide giving that $R(\boldsymbol{\beta}) = R(\mathbf{\Gamma}^{-1}\mathbf{\Gamma}_e) = R(\mathbf{\Delta}^{-1}\boldsymbol{\alpha})$. If Ferré and Yao's conclusion were true we get for model (3.1) that $R(\mathbf{\Gamma}^{-1}\boldsymbol{\alpha}) = R(\mathbf{\Gamma}^{-1}\mathbf{\Gamma}_e) = R(\mathbf{\Gamma}_e) = R(\boldsymbol{\alpha})$, but if we take $p = 2$, $\boldsymbol{\alpha} = (1,0)^T$, $\mathrm{Cov}\,(\nu_y) = 1$ and $\mathbf{\Delta} = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}$ in (3.1), we have $R(\mathbf{\Gamma}^{-1}\mathbf{\Gamma}_e) = \mathrm{span}((1,-0.25)^T)$ and $R(\mathbf{\Gamma}_e) =$

span($(1,0)^T$), clearly two different subspaces. This example shows that (1.2) it is not necessarily true for (1.1). Moreover, (1.2) is true for this model if and only if $R(\mathbf{\Gamma}^{-1}\boldsymbol{\alpha}) = R(\boldsymbol{\alpha})$ and, using (1.3), that implies $\mathbf{\Gamma} = \boldsymbol{\alpha}\boldsymbol{\alpha}^T\mathbf{\Sigma}\boldsymbol{\alpha}\boldsymbol{\alpha}^T + \boldsymbol{\alpha}_0\boldsymbol{\alpha}_0^T\mathbf{\Sigma}\boldsymbol{\alpha}_0\boldsymbol{\alpha}_0^T$, where $\boldsymbol{\alpha}_0$ is an orthogonal completion of $\boldsymbol{\alpha}$. Since $\mathbf{\Delta} = \mathbf{\Gamma} - \boldsymbol{\alpha}\,\mathrm{Cov}\,(\boldsymbol{\nu}_y)\boldsymbol{\alpha}^T$, this is equivalent to saying that (1.2) is true if and only if $\mathbf{\Delta}$ reduces $\boldsymbol{\alpha}$, and this implies a very specific structure for $\mathbf{\Delta}$, and not any positive definite matrix.

Acknowledgements

Part of this work was done while the first two authors were visiting the Isaac Newton Institute for Mathematical Sciences, Cambridge, UK.

References

Conway, J. B. (1990). *A Course in Functional Analysis.* 2nd ed. Springer, New York.

Cook, R. D. (2007). Fisher lecture: Dimension reduction in regression. *Statist. Sinica* **22**, 1-26.

Cook, R. D., Li, B. and Chiaromonte, F. (2007). Dimension reduction in regression without matrix inversion. *Biometrika* **94**, 569-584.

Dauxois, J., Ferré, L. and Yao, A. (2001). Un modèle semi-paramétrique pour variables aléatoires. *C. R. Acad Paris* **333**, 947-952.

Ferre, L. and Yao, A. F. (2005). Smoothed functional inverse regression. *Statist. Sinica* **15**, 665-683.

Ferre, L. and Yao, A. F. (2007). Reply to the paper by Liliana Forzani and R. Dennis Cook: "A note on smoothed functional inverse regression". *Statist. Sinica* **17**, 1683-1687.

Forzani, L. and Cook, R. D. (2007). A note on smoothed functional inverse regression. *Statist. Sinica* **17**, 1677-1681.

Groetsch, C. W. (1977). *Generalized Inverses of Linear Operators.* Marcel Dekker Inc., New York.

Li, K. C. (1991). Sliced inverse regression for dimension reduction (with discussion). *J. Amer. Statist. Assoc.* **86**, 316-342.

Shao, Y., Cook, R. D. and Weisberg, S. (2007). Marginal tests with sliced average variance estimation. *Biometrika* **94**, 285-296.

School of Statistics, University of Minnesota, 313 Ford Hall, 224 Church Street S.E., Minneapolis, Minnesota 55455, United State.

E-mail: dennis@stat.umn.edu

Facultad de Ingeniería Química, Universidad Nacional del Litoral and Instituto Matemática Aplicada del Litoral, CONICET, Gëmes 3450, (3000) Santa Fe, Argentina.

E-mail: liliana.forzani@gmail.com

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

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

(Received February 2008; accepted November 2008)