

ESTIMATION DANS UN MODÈLE DÉFINI PAR DES ÉQUATIONS ESTIMANTES CONDITIONNELLES POUR DES DONNÉES FONCTIONNELLES

Matthieu Saumard & Valentin Patilea

*INSA DE RENNES
20 AVENUE DES BUTTES DE COESMES
CS 70 839
35 708 RENNES CEDEX 7*

Résumé : Nous considérons ici un modèle défini par des restrictions de moments conditionnels dans lequel les valeurs de la variable dans le conditionnement ainsi que le paramètre d'intérêt sont des éléments d'un espace fonctionnel. Pour estimer le paramètre d'intérêt, nous utiliserons une technique de troncature afin d'appliquer et généraliser la théorie existante dans les modèles finies-dimensionnelles.

mots clefs : données fonctionnelles, équations estimantes conditionnelles, méthode de moments généralisés

Abstract : We consider a conditional moment restrictions model for functional data. The response variable is a vector in a finite dimensional space while the conditioning variable as well as the parameter to be estimated are elements of a functional space. To estimate the parameter we first decompose the conditioning variable and the parameter in a basis of the functional space. For a given sample size we truncate their decompositions to a finite number of coefficients and we apply the classical approach for estimating equations. For the asymptotic results we let this finite number of coefficients to grow to infinity at a suitable rate.

keywords : functional data, conditional estimating equations, generalized method of moments

1 Introduction

Plusieurs études impliquent des mesures étroitement répétés dans le temps sur le même individu ou un direct enregistrement d'une courbe. Si les mesures longitudinales sont faites sur une grille suffisamment dense, ces données sont souvent regardées comme une courbe ou comme des données fonctionnelles. Il est clairement d'intérêt d'étendre la notion de modèles définis par des équations estimantes conditionnelles au cas de données fonctionnelles. L'extension des modèles classiques de régression au cas des données fonctionnelles et des paramètres fonctionnels implique des techniques d'estimation spécifiques. Voir par exemple Cardot et Sarda (2009). Voir aussi le modèle linéaire généralisé fonctionnel étudié par Müller et Stadtmüller (2005).

Nous proposons ici un cadre plus général d'un modèle défini par des équations estimantes conditionnelles dans lequel le vecteur X intervenant dans le conditionnement ainsi que le paramètre θ identifié par les équations sont de dimension infinie. Pour l'instant, nous ne considérons que le cas où les équations estimantes font apparaître une seule variable 'explicative' Y et elle dépendent de X et θ seulement à travers le produit scalaire entre X et θ . L'idée principale pour construire un estimateur du paramètre fonctionnel θ suit l'approche proposée par Müller et Stadtmüller qui consiste à utiliser un développement de Karhunen et Loève (d'autres type de decompositions dans une base orthonormale peuvent être utilisés) de la fonction aléatoire X et du paramètre θ avec le but de réduire la dimension aux premières composantes de ces decompositions. Les decompositions sont ainsi tronquées à un nombre fini de termes mais ce nombre croît avec la taille de l'échantillon.

Une fois que pour une taille d'échantillon n la dimension de X et de θ est réduite à un nombre fini de composantes p_n , nous nous retrouvons dans le cas classique des équations estimantes conditionnelles avec des vecteurs aléatoires et de paramètres de taille finie. On peut alors appliquer l'approche habituelle par la méthode des moments généralisée (GMM en anglais), voir par exemple Newey (1993). Ensuite il reste à étudier le comportement de l'estimateur ainsi obtenu lorsque la dimension p_n augmente avec la taille de l'échantillon n .

2 Le modèle

On dispose de n observations i.i.d. $z_i = (x_i, y_i)$ $i = 1, \dots, n$ où $x_i \in L^2([0, 1])$ sont des courbes aléatoires et $y_i \in \mathbb{R}^d$ sont des vecteurs aléatoires de taille fixe. Soit g une fonction connue à valeurs en \mathbb{R}^s avec $s \geq 1$. Notre objectif est d'estimer la vraie valeur θ_0 d'un paramètre d'intérêt $\theta \in L^2([0, 1])$ en supposant que

$$\mathbb{E}[g(z, \theta)|x] = 0 \text{ p.s.} \iff \theta = \theta_0.$$

Pour l'instant, nous allons nous limiter au cas où la dépendance par rapport à θ est linéaire, plus précisément nous avons

$$\mathbb{E}[g(y, \int_{[0,1]} x(t)\theta(t)dt)|x] = 0 \text{ p.s.} \iff \theta = \theta_0$$

Soit ρ_j , $j = 1, 2, \dots$, une base orthonormale de l'espace $L^2([0, 1])$, on a donc

$$\int_{[0,1]} \rho_j(t)\rho_k(t)dt = \delta_{jk}, \quad j, k = 1, 2, \dots$$

Ainsi $x(t)$ et $\theta(t)$, $t \in [0, 1]$ peuvent se décomposer en :

$$x(t) = \sum_{j=1}^{\infty} \epsilon_j \rho_j(t), \quad \theta(t) = \sum_{j=1}^{\infty} \theta_j \rho_j(t)$$

Les coefficients ϵ_j et θ_j sont données par $\epsilon_j = \int x(t)\rho_j(t)dt$ et par $\theta_j = \int \theta(t)\rho_j(t)dt$ respectivement. L'orthonormalité de la base ρ_j à pour conséquence que

$$\int x(t)\theta(t)dt = \sum_{j=1}^{\infty} \epsilon_j\theta_j.$$

Pour contrer la difficulté causé par la dimension infinie, on approche le modèle défini plus haut par un modèle avec tronqué à $p = p_n$ variables et paramètres avec p_n qui croit quand n tend vers l'infini. On peut alors utiliser l'approche type GMM pour estimer les vraies valeurs $\theta_{0,1}\dots\theta_{0,p_n}$. Ayant un estimateur $(\hat{\theta}_1, \dots, \hat{\theta}_{p_n})$ de $(\theta_{0,1}\dots\theta_{0,p_n})$, on estime la fonction $\theta_0(t)$ par

$$\hat{\theta}(t) = \sum_{j=1}^{p_n} \hat{\theta}_j \rho_j(t), \quad t \in [0, 1].$$

3 Estimation

Pour décrire l'estimateur, introduisons une matrice de 'variables instrumentales' $A(x)$ de dimension $p_n \times s$. Alors $\mathbb{E}[A(x)g(z, \theta_0)] = 0$. En fait, on utilise la matrice 'optimale' $A(x) = D(x)'\Omega(x)^{-1}$ qui est cependant inconnue, où

$$D(x) = \mathbb{E}\left[\frac{\partial g}{\partial \theta}(z, \theta_0)|x\right]$$

et $\Omega(x)$ est une matrice $s \times s$ donnée par

$$\Omega(x) = \mathbb{E}[g(z, \theta_0)g(z, \theta_0)'|x].$$

Ici l'écriture $\partial g/\partial \theta$ doit être comprise comme la matrice $s \times p_n$ de dérivées partielles de la fonction g par rapport aux p_n premières composantes de θ . De même θ_0 n'est que θ_0 arrêté aux p_n premières composantes. Soit \hat{P} une matrice $p_n \times p_n$ semi-définie positive qui peut-être aléatoire ou éventuellement égale à l'identité, et considérons l'estimateur

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \hat{g}_n(\theta)' \hat{P} \hat{g}_n(\theta)$$

où

$$\hat{g}_n(\theta) = \frac{1}{n} \sum_{i=1}^n A(x_i)g(y_i, \sum_{j=1}^{p_n} \theta_j \epsilon_j^i)$$

Par abus de notation on a écrit $A(x_i)$ au lieu de $A(\epsilon_1^i, \dots, \epsilon_{p_n}^i)$ où ϵ_j^i représente le $j^{\text{ème}}$ élément dans la décomposition de x_i suivant la base $\{\rho_k, k = 1, 2, \dots\}$ et $\sum_{j=1}^{p_n} \theta_j \epsilon_j^i$ n'est autre que le produit scalaire entre θ et x_i tronqué à p_n .

Sous certaines hypothèses dont entre autre l'idenfiabilité du paramètre θ_0 , la régularité de la fonction connue g , ainsi que $p_n \rightarrow \infty$ et que $p_n = o(n^{1/4})$, nous obtenons la convergence en loi

$$\frac{n(\hat{\theta} - \theta_0)' \Lambda^{-1} (\hat{\theta} - \theta_0) - p_n}{\sqrt{2p_n}} \Longrightarrow N(0, 1),$$

quand n tend vers l'infini, avec $\Lambda = \Lambda_{p_n} = (\mathbb{E}[D(x)' \Omega(x)^{-1} D(x)])^{-1}$, ici Λ étant une matrice de taille $p_n \times p_n$.

Bibliographie

- [1] Cardot, H. et Sarda, P. (2009), Functional Linear Regression, dans *Handbook of Functional Data Analysis*, Ferraty, F. and Romain, Y. (Eds.), Oxford University Press.
- [2] Müller, H.G. et Stadtmüller, U. (2005), Generalized functional linear models, *the Annals of Statistics*, 33, 774–805.
- [3] Newey, W.K. (1993) Efficient estimation of models with conditional moment restrictions, *Handbook of statistics*, 11, 419–454.
- [4] Ramsay, J.O. et Silverman, B.W. (1997) Functional data analysis, *Springer*, New York.