International Journal of Pure and Applied Mathematics
Volume 98 No. 1 2015, 1-10
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
doi: http://dx.doi.org/10.12732/ijpam.v98i1.1

ijpam.eu

ON PARAMETER ESTIMATION OF SEMI-VARYING COEFFICIENT MODELS WITH CORRELATED RANDOM ERRORS

Yafeng Xia[1 §], Ling Zhang[2]
[1,2]School of Sciences
Lanzhou University of Technology
Lanzhou, 730050, P.R. CHINA

Abstract: This paper deals with estimation of semi-varying coefficient models with correlated random errors. The estimations of the function coefficients are are given by the use of generalized weighted least squares. The theorem of Gauss-Markov was very well known, but moreover we present an estimation of parameter σ^2.

AMS Subject Classification: 62G05, 62G08
Key Words: semi-varying coefficient models, the generalized weighted least squares

1. Introduction

Fan et al (see [1]) put forward the the semi-varying coefficient models when he studied whether the function coefficient of the varying coefficient really change,

$$Y = X\alpha(U) + Z\beta + \varepsilon. \tag{1.1}$$

Received: September 11, 2013

§Correspondence author

Suppose that we have a response variable Y and regressors $X = (X_1, X_2, \ldots X_p)$, $Z = (Z_1, Z_2, \ldots, Z_q)$ as well as another variable U. ε is a random error with $E(\varepsilon) = 0$ and it is independent of (U, Z, X); $\beta = (\beta_1, \beta_2, \ldots, \beta_q)^T$ is a q-dimensional vector of an unknown parameters and $\alpha(\cdot) = (\alpha_1(\cdot), \alpha_2(\cdot), \ldots, \alpha_p(\cdot))^T$ is a p-dimensional vector of an unknown coefficient function.

Most researches have focused on the case in which random error are independently and identically distributed in a model.

In practice, sometimes the random errors are correlated, so it is been necessary to discuss the question.

In this paper, we discuss the semi-varying coefficient models with correlated random errors, which can be defined as follows:

$$\begin{cases} Y = X\alpha(U) + Z\beta + e, \\ e \sim N(0, \sigma^2\Sigma), \quad \Sigma = diag(\omega_1, \omega_2, \ldots, \omega_n). \end{cases} \tag{1.2}$$

2. Estimation

Suppose that we have a random sample of size n. $\{Y_r, X_r, Z_r, U_r\}(r = 1, 2, \ldots, n)$ is a sample from the model (1.2), $e \sim N(0, \sigma^2\Sigma)$, $e = (e_1, e_2, \ldots, e_n)^T$, Σ is a given positive matrices, (1.2) can be written as

$$\begin{cases} Y^* = X\alpha(U) + e, \\ e \sim N(0, \sigma^2\Sigma). \end{cases} \tag{2.1}$$

where $Y^* = Y - Z\beta$,and β known,

$$X = (X_1^T, X_2^T, \ldots, X_n^T)^T, \quad X_r^T = (X_{r1}, X_{r2}, \ldots, X_{rp}),$$

$$Z = (Z_1^T, Z_2^T, \ldots, Z_n^T)^T, \quad Z_r^T = (Z_{r1}, Z_{r2}, \ldots, Z_{rq}).$$

As Σ is a given positive definite matrix,there must exist a orthogonal matrix. Hence, $\Sigma = P^T\Lambda P$, $\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$, $\lambda_r > 0$ $(r = 1, 2, \ldots, n)$ are characteristic roots of Σ.Use P^T left multiplied by (2.1).

So

$$\begin{cases} \widetilde{Y} = \widetilde{X}\alpha(U) + \varepsilon, \\ e \sim N(0, \sigma^2 I_n). \end{cases} \tag{2.2}$$

So (2.2) is the common varying coefficient model.

Let

$$\widetilde{Y} = (\widetilde{Y}_1, \widetilde{Y}_2, \dots, \widetilde{Y}_n)^T = \Sigma^{\frac{-1}{2}} Y^*, \quad \overline{Y} = \Sigma^{\frac{-1}{2}} Y,$$

$$\overline{Z} = \Sigma^{\frac{-1}{2}} Z, \quad \widetilde{X} = (\widetilde{X}_1^T, \widetilde{X}_2^T, \dots, \widetilde{X}_n^T)^T = \Sigma^{\frac{-1}{2}} X, \quad \widetilde{X}_r^T = (\widetilde{X}_{r1}, \widetilde{X}_{r2}, \dots, \widetilde{X}_{rp}),$$

$$\overline{Z} = (\overline{Z}_1^T, \overline{Z}_2^T, \dots, \overline{Z}_n^T)^T, \quad \overline{Z}_r^T = (\overline{Z}_{r1}, \overline{Z}_{r2}, \dots, \overline{Z}_{rq}),$$

$$\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T = \Sigma^{\frac{-1}{2}} e, \quad W = diag(W_{11}, W_{22}, \dots, W_{nn})$$

$(r = 1, 2, \dots, n)$. The coefficient function $\alpha(\cdot) = (\alpha_1(\cdot), \alpha_2(\cdot), \dots, \alpha_p(\cdot))^T$ is estimated by minimizing

$$Q(\alpha(u)) = (\widetilde{Y} - \widetilde{X}\alpha(U))^T W (\widetilde{Y} - \widetilde{X}\alpha(U)). \tag{2.3}$$

According to the principle of weighted least squares, if we assume that the inverse of the matrix $\widetilde{X}^T W \widetilde{X}$ exists for any u, then the estimated coefficient function at u can be expressed as

$$\begin{aligned}\widehat{\alpha}(u) &= (\widehat{\alpha}_1(u), \widehat{\alpha}_2(u), \dots, \widehat{\alpha}_p(u))^T \\ &= [X^T \Sigma^{\frac{-1}{2}} W \Sigma^{\frac{-1}{2}} X]^{-1} X^T \Sigma^{\frac{-1}{2}} W \Sigma^{\frac{-1}{2}} Y^*. \end{aligned} \tag{2.4}$$

Then (2.2) can be written as

$$\overline{Y} - \overline{Z}\beta = M + \varepsilon. \tag{2.5}$$

The estimator for M is then

$$\widehat{M} = \widetilde{X}[\widetilde{X}^T W \widetilde{X}]^{-1} \widetilde{X}^T W (\overline{Y} - \overline{Z}\beta) = S(\overline{Y} - \overline{Z}\beta).$$

The matrix S is a smoothing matrix and dependent only on the observations (U_r, X_r^T), $r = 1, 2, \dots, n$.Substituting $\widehat{M}$ into (2.5),we obtain

$$(I - S)\bar{Y} = (I - S)\bar{Z}\beta + \varepsilon. \tag{2.6}$$

Applying least squares to the linear model (2.6), we obtain

$$\widehat{\beta} = \{\overline{Z}^T (I - S)^T (I - S)\overline{Z}\}^{-1} \overline{Z}^T (I - S)^T (I - S)\overline{Y}.$$

Moreover

$$\widehat{M} = S(\overline{Y} - \overline{Z}\widehat{\beta}). \tag{2.7}$$

On the basis of $\widehat{\alpha}(u)$, we inferred from $\widehat{\beta}$, hence $C^T\beta$ is two-step estimation method of $C^T\alpha(U)$, due to the complexity of local fitting, so it's difficult to analyze the statistical inference properties.thus, we mainly focus on some properties of $\widehat{\alpha}(u)$.

Suppose that $EY^* = X\alpha(u)$, then:

$$E[\widehat{\alpha}(u)] = [X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}}X]^{-1}X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}}X\alpha(u),$$
$$Var[\widehat{\alpha}(u)] = \sigma^2[X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}}X]^{-1}[X^T\Sigma^{\frac{-1}{2}}W^2\Sigma^{\frac{-1}{2}}X][X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}}X].$$

Theorem 1. *Consider model (2.2), $\widetilde{\alpha}(u)$ is the generalized weighted least squares estimation*

$$\widetilde{\alpha}(u) = [X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}}X]^{-1}X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}}Y^*.$$

Suppose that $EY^* = X\alpha(u)$. Then:

(1) When X is a matrix with full column rank and the inverse of the matrix $X^T\Sigma^{\frac{-1}{2}}\ W\Sigma^{\frac{-1}{2}}X$ exist, then $\widehat{\alpha}(u)$ must be unique.

$$E[\widehat{\alpha}(u)] = \alpha(u).$$

Hence $\widehat{\alpha}(u)$ is an estimation of $\alpha(u)$, and we called $\widehat{\alpha}(u)$ the weighted linear unbiased estimate of $\alpha(u)$.

(2) When X is a matrix with full column rank and W is idempotent matrix,then

$$E[\widehat{\alpha}(u)] = \alpha(u).$$
$$Var[\widehat{\alpha}(u)] = \sigma^2[X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}}X]^{-1}.$$

(3) When X is a matrix with full column rank, W is idempotent matrix and $\lim\limits_{n\to\infty}(X^T\Sigma^{\frac{-1}{2}}\ W\Sigma^{\frac{-1}{2}}X)^{-1} = 0$. hence $\widehat{\alpha}(u)$ is the consistent estimate of $\alpha(u)$.

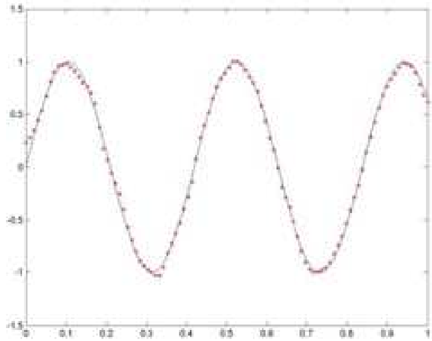
3. Simulation

In this section, we use Matlab conduct some research to estimator.

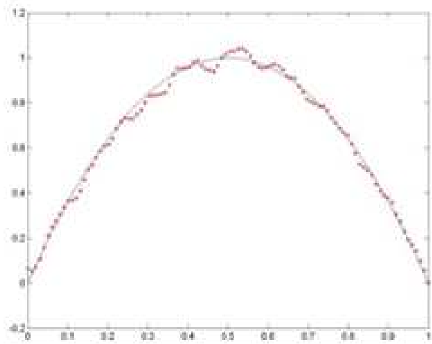
Case 1. $Y = sin(15U)X_1 + 4U(1-U)X_2 + 8Z_1 + \varepsilon$.

Case 2. $Y = cos(5piU)X_1 + exp(4U)X_2 + 8Z_1 + \varepsilon$.

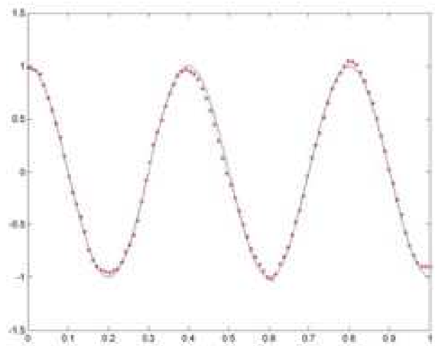
In these cases,where $U \sim N[0.1]$, $X_1 \sim N[1,3]$, $X_2 \sim N[5,10^{-5}]$,$Z_1 \sim N[0,10^{-5}]$, $\varepsilon \sim N(0,A)$, A is a random number in $[0,1.1]$. Using *Epanechnikov kernels* window width $h_n = \frac{1}{40}$, sample data randomly generated by $Matlab$. The sample sizes is 1000. In Case 1 and Case 2 we use $*$ indicate estimated value, and use solid-line curve indicate actual value. Then, we found estimated results is fine.

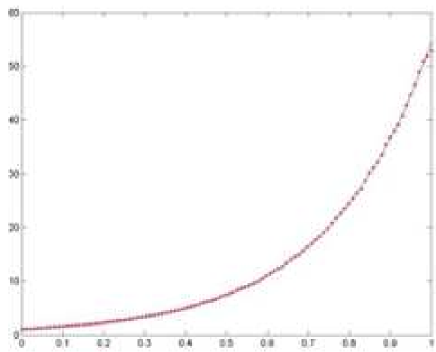


[t!]



The estimate of $Sin(15U)$ The estimate of $4U(1-U)$.





The estimate of $Cos(5piU)$ The estimate of $exp(4U)$.

4. The Estimation of the Linear Estimable Function and Optimality

Theorem 2. (The Generalized Weighted Gauss-Markov Theorem) *In model* (2.2), $\widehat{\alpha}(u)$ *is a generalized weighted least squares estimate , Suppose that* $EY^* = X\alpha(u)$. *Then:*

(1) $\widehat{\alpha}(u)$ *is the unique weighted linear unbiased estimation of* $\alpha(u)$.

(2) when W is an idempotent matrix, then $\widehat{\alpha}(u)$ is the unique best weighted linear unbiased estimation of $\alpha(u)$.

(3) when X is a matrix with full row rank,W is an idempotent matrix and $\lim\limits_{n\to\infty} c^T(X^T\Sigma^{\frac{-1}{2}} W\Sigma^{\frac{-1}{2}}X)^{-1}c = 0$. Then $c^T\widehat{\alpha}(u)$ is the consistent estimate of $c^T\alpha(u)$.

In order to prove Theorem 2, we give some definitions and lemmas are defined in this paper.

Definition 1. (see [2]) In model (2.2) , $c \in R^p$, linear function $\alpha(u)$ is weighting estimation. If we assume that $b^TW\widetilde{Y}$ is the weighted linear unbiased estimation of $c^T\alpha(u)$ and $b^TW\widetilde{Y}$ exists , $b \in R^n$,then

$$E(b^TW\widetilde{Y}) = c^T\alpha(u),$$

for any $\alpha(u)$.

Lemma 1. (see [2]) *$c^T\alpha(u)$ is the weighting estimation $\Leftrightarrow c^T \in R(\widetilde{X}^TW\widetilde{X}) = R(\widetilde{X}^TW) = R(\widetilde{X}^T)$.*

Lemma 2. (see [2]) *If $c^T\alpha(u)$ is the weighting estimation ,then $c^T\widehat{\alpha}(u)$ must be unique. since $b \in R_n$, we have*

$$c^T\widehat{\alpha}(u) = b^T\widetilde{X}^TW\widetilde{Y}.$$

Definition 2. (see [2]) Considering model(2.2) and $c \in R^p$,$b \in R^n$,so $b^TW\widetilde{Y}$ is the best weighted linear unbiased estimate.

If:

(1) $b^TW\widetilde{Y}$ is the weighted linear unbiased estimate of $c^T\alpha(u)$.

(2) Any $g^TW\widetilde{Y}$ is the weighted linear unbiased estimate of $c^T\alpha(u)$, we have

$$Var(b^TW\widetilde{Y}) \le Var(g^TW\widetilde{Y}).$$

Proof of Theorem 2. (1) $c^T\alpha(u)$ is the weighting estimation.By Lemma 1, $c \in R(\widetilde{X}^T)$ and $b \in R^n$.

We have

$$c = \widetilde{X}^Tb = X^T\Sigma^{\frac{-1}{2}}b.$$

then $c^T\widehat{\alpha}(u) = b^T\widetilde{X}(\widetilde{X}^TW\widetilde{X})^{-1}\widetilde{X}^T\widetilde{Y}$. so

$$E(c^T\widehat{\alpha}(u)) = b^T\Sigma^{\frac{-1}{2}}X\alpha(u).$$

By Lemma 2, we know $c^T\alpha(u)$ must be unique. So $\widehat{\alpha}(u)$ is the unique weighted linear unbiased estimate of $\alpha(u)$.

(2) W is an idempotent matrix, any $g^T W\widetilde{Y}$ is the weighted linear unbiased estimate of $c^T\alpha(u)$. So

$$c^T\alpha(u) = E(g^T W\widetilde{Y}) = g^T W\widetilde{X}\alpha(u),$$

for any $\alpha(u)$.

We have $c^T = g^T W\widetilde{X}$, then $c = \widetilde{X}^T W g$,

$$\begin{aligned} &Var(g^T W\widetilde{Y}) - Var(c^T\alpha(u)) \\ &\qquad = \sigma^2 g^T W^{\frac{1}{2}}[I_n - W^{\frac{1}{2}}\Sigma^{\frac{-1}{2}}X(X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}}X)^{-1}X^T\Sigma^{\frac{-1}{2}}W^{\frac{1}{2}}]W^{\frac{1}{2}}g, \end{aligned}$$

where $P_{W_{\frac{1}{2}}} = W^{\frac{1}{2}}\Sigma^{\frac{-1}{2}}X(X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}}X)^{-1}X^T\Sigma^{\frac{-1}{2}}W^{\frac{1}{2}}$ is an Projection matrix.

Therefore

$$Var(g^T W\widetilde{Y}) - Var(c^T\widehat{\alpha}(u)) = \sigma^2 g^T W^{\frac{1}{2}}[I_n - P_{W_{\frac{1}{2}}}]W^{\frac{1}{2}}g \geq 0. \tag{3.1}$$

From Definition 2, we have that $\widehat{\alpha}(u)$ is the weighted linear unbiased estimate of $\alpha(u)$. The conditions of the holding water of sign of equality in (3.1) is $(I_n - P_{W_{\frac{1}{2}}})g = 0$, then

$$g^T W\widetilde{Y} = g^T W\widetilde{X}(\widetilde{X}^T W\widetilde{X})^{-1}\widetilde{X}^T W\widetilde{Y} = c^T\widehat{\alpha}(u).$$

Hence $\widehat{\alpha}(u)$ is the unique best weighted linear unbiased estimate of $\alpha(u)$.

(3) We have $E(c^T\widehat{\alpha}(u)) = c^T\alpha(u)$, X is a matrix with full column rank, W is idempotent matrix and $\lim_{n\to\infty} c^T(X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}}X)^{-1}c = 0$.

By Theorem 1, we have

$$Var(c^T\widehat{\alpha}(u)) = c^T Var(\widehat{\alpha}(u))c = \sigma^2(X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}}X)^{-1}c \to 0 \quad (n\to\infty).$$

Therefore, $c^T\widehat{\alpha}(u)$ is the consistent estimate of $c^T\alpha(u)$. □

5. The Estimation of σ^2 and Optimality

Let us set:

$$Rss_e^2 = (Y^* - X\alpha(u))^T(Y^* - X\alpha(u)) = \|Y^* - X\alpha(u)\|^2.$$

$$e = Y^* - X\alpha(u) = (I_n - X(X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}}X)^{-1}X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}})Y^*.$$

If

$$P = X(X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}}X)^{-1}X^T\Sigma^{\frac{-1}{2}}W\Sigma^{\frac{-1}{2}}$$

and

$$P^2 = P,$$

then

$$e = (I_n - P)Y^*.$$

Setting $M = I_n - P$, $MX = X - PX = 0$, we obtain $Rss_e^2 = (Y^*)^T M^T MY^*$.

As before, suppose that $E(Y^*) = X\alpha(u)$, then

$$E(Rss_e^2) = E((Y^*)^T M^T MY^*) = tr(M^T Mcov(Y^*)) = \sigma^2 tr(\Sigma^{\frac{-1}{2}}M^T M\Sigma^{\frac{-1}{2}}).$$

Theorem 3. *Considering model (2.2) $c \in R^p$,suppose that$EY^* = X\alpha(u)$, then*

$$\widehat{\sigma}^2 = tr[(\Sigma^{\frac{-1}{2}})MM^T\Sigma^{\frac{-1}{2}}]^{-1}(Y^*)^T M^T MY^*.$$

is the unbiased estimation of σ^2.

In order to determine the probability distribution of σ^2, we have to modify the construction of σ^2.

As before, suppose that $E(Y^*) = X\alpha(u)$ $|M| \neq 0$ and $e \sim N(0, \Sigma)$, then $\Sigma^{\frac{-1}{2}}e \sim N(0, I_n)$;

$$\begin{aligned}
Rss_{e^*}^2 =& [(M^T\Sigma M)^{\frac{-1}{2}}(Y^* - X\alpha(u))]^T[(M^T\Sigma M)^{\frac{-1}{2}}(Y^* - X\alpha(u))] \\
=& (Y^*)^T M^T(M\Sigma M^T)^{-1}MY^*, \\
E[Rss_{e^*}^2] =& E[Y^*)^T M^T(M\Sigma M^T)^{-1}MY^*] \\
=& E[e^T M^T(M\Sigma M^T)^{-1}Me] \\
=& E[(\Sigma^{\frac{-1}{2}}e)^T\Sigma^{\frac{1}{2}}M^T(M\Sigma M^T)^{-1}M\Sigma^{\frac{1}{2}}(\Sigma^{\frac{-1}{2}}e)] \\
=& \sigma^2 tr[(M\Sigma M^T)^{-1}(M\Sigma M^T)] \\
=& \sigma^2 tr(I_n) \\
=& n\sigma^2.
\end{aligned}$$

Then $\widehat{\sigma}^2 = \frac{1}{n}(Y^*)^T M^T(M\Sigma M^T)^{-1}MY^*$ is the unbiased estimation of σ^2.

But $\Sigma^{\frac{-1}{2}}e \sim N(0, I_n)$ and $\Sigma^{\frac{1}{2}}M^T(M\Sigma M^T)^{-1}M\Sigma^{\frac{1}{2}}$ is an idempotent and symmetric matrix, so

$$tr[\Sigma^{\frac{1}{2}}M^T(M\Sigma M^T)^{-1}M\Sigma^{\frac{1}{2}}] = rk[\Sigma^{\frac{1}{2}}M^T(M\Sigma M^T)^{-1}M\Sigma^{\frac{1}{2}}] = n.$$

Hence

$$\frac{\Sigma^{\frac{-1}{2}}}{\sigma}e \sim N(0, I_n).$$

Therefore

$$\begin{aligned}\frac{Rss^2_{e^*}}{\sigma^2} =& \frac{1}{\sigma^2}(Y^*)^T M^T (M\Sigma M^T)^{-1} MY^* \\ =& \frac{1}{\sigma^2}(e^T M^T (M\Sigma M^T)^{-1} Me \\ =& \frac{1}{\sigma^2}(\Sigma^{\frac{-1}{2}}e)^T \Sigma^{\frac{1}{2}} M^T (M\Sigma M^T)^{-1} M\Sigma^{\frac{1}{2}}(\Sigma^{\frac{-1}{2}}e).\end{aligned}$$

$\frac{Rss^2_{e^*}}{\sigma^2}$ follows the χ^2 distribution with n degree of freedom and $Var(Rss^2_{e^*}) = 2n\sigma^2$.

Theorem 4. *Consider model (2.1). Suppose that* $EY^* = X\alpha(u)$,

$$\widehat{\sigma}^2 = \frac{1}{n}(Y^*)^T M^T (M\Sigma M^T)^{-1} MY^*$$

is the unbiased estimation of σ^2. *If* $\frac{\Sigma^{\frac{-1}{2}}}{\sigma}e \sim N(0, I_n)$, *then we have*

$$\frac{Rss^2_{e^*}}{\sigma^2} = \frac{1}{\sigma^2}(Y^*)^T M^T (M\Sigma M^T)^{-1} MY^*$$

which follows the χ^2 *distribution with* n *degree of freedom and* $Var(\widehat{\sigma}^2) = \frac{2\sigma^4}{n}$.

Acknowledgments

This paper is supported by Science and Technology Program for Guangdong Province (Grant No. 2012B010100044) and Science and Technological Program for Dongguans Higher Education, and Science and Research Institutions (Grant No. 2012108102031).

References

[1] J. Fan, W. Zhang, Semultaneous Cconfidence bands are and hypothsis testing in varying-coefficient models, *S.J. Statistics*, **27** (2000), 715-731.

[2] G. Ouyang, On perameter estimation for varying-coefficient regression models, *Journal of Xiangnan University*, **26** (2005), 15-22.

[3] C. Fatherinfham, C. Brunsdon, A. Fatheringham, Geographically Weighted Regressin: A method for exploring spatial nonstationarity, *Geographial*, **28** (1996), 281-298.

[4] C. Mei, W. Zhang, Statistical Inferences for Varying-coefficient models based on locally weighted regression technique, *Acta Mathematicae Applicatae sinica Engling Series*, **17** (2001), 407-417.

[5] Q. Tang, Componentwise estimation for varying coefficient models, *China ACademic Journal Electronic Publishing House*, **24**, No. 1 (2009), 32-38.

[6] J. Meng, R. Zhang, Estimation of semivarying-coefficient models with the correlated random errors, *Chinese Journal of Applied Probability and Statistics*, **26**, No. 5 (2010), 477-484.

[7] Y. Wu, J. Fan, Varying-coefficient functional linear regression, *Bernoulli*, **16**, No. 3 (2010), 730-758.

[8] C. Wei, X. Wu, Error variance estimation in partially linear varying, *Mathematica Applicata*, **22**, No. 2 (2008), 378-383.