Original paper

Two-phase Multiple Regression Model with Some Order Autoregressive Processes

Motoi TSUCHIYA*, Yataro HORIKAWA** and Tairai MINO***

Abstract

The purpose of this paper is to present a two-phase multiple regression model containing some order auto-correlated error terms under a continuity constraint and to estimate the parameters and the change line of the model by a least squares method. The asymptotic distribution theory of the least squares estimators under the continuity constraint is obtained from the asymptotic property of data, making the residual sum of squares locally smooth enough to apply the classical techniques. Further, the model is applied to analyzing a set of long-term annual average temperature data in Tokyo.

Keywords. two-phase multiple regression, autoregressive process, change line, continuity constraint

1. Introduction

The aim of this paper is to consider a two-phase multiple regression model under a continuity constraint having error terms with some order autoregressive processes. The model is an extension of a two-phase simple linear regression model of which regression line bends at a change point[3)6)7)]. For simplicity, we treat the case where the model has two explanatory variables. In this model, therefore, a regression plane is exactly separated into two-phase ones by a change line. We here note that the continuity constraint imposing a change line plays an essential role in practical applications. Without this constraint, an estimated change line often appears in the interior of a phase and not on the boundary of the phases. In such cases this estimated change line does not show the threshold of the expected phases at all.

In Section 2, we describe the two-phase multiple regression model composed of two functions under a continuity constraint and a sequence of error term expressed as an sth order autoregressive process with some regularity conditions. In Section 3, applying the transformation method suggested by Cochrane and Orcutt[2)] and Kadiyala[5)] to the model having independently and identically distributed (i.i.d.) error terms, we derive least squares estimators (LSE) of the parameters and the change line for a fixed order in an autoregressive process together with two information criteria. In Section 4, we give main theorems, showing the consistencies and the asymptotic distribution theory of LSEs of the parameters and the change line. Section 5 is devoted to applying the model to an analysis of a set of annual average temperatures in Tokyo. The two-phase bent plane is estimated by a newly developed computational procedure.

2. Setting a two-phase multiple regression model

We set a two-phase multiple regression model $\boldsymbol{Y}=\tau(\boldsymbol{x}, \boldsymbol{\beta})+\boldsymbol{u}$ with stationary correlated error terms as an sth order autoregressive process under a continuity constraint in the following way. Let $\boldsymbol{Y}=(Y_1, Y_2, \cdots, Y_n)^t$ denote a vector of observations expressed as $Y_i=\tau(\boldsymbol{x}_i, \boldsymbol{\beta})+u_i$, $i=1, 2, \cdots, n$. The superfix t stands for a transposition of a vector or a matrix. Here τ is a two-phase multiple

* 健康管理学研究室 Seminar of Health Care and Administration
** 物理学研究室 Seminar of Physics
*** 数学研究室 Seminar of Mathematics

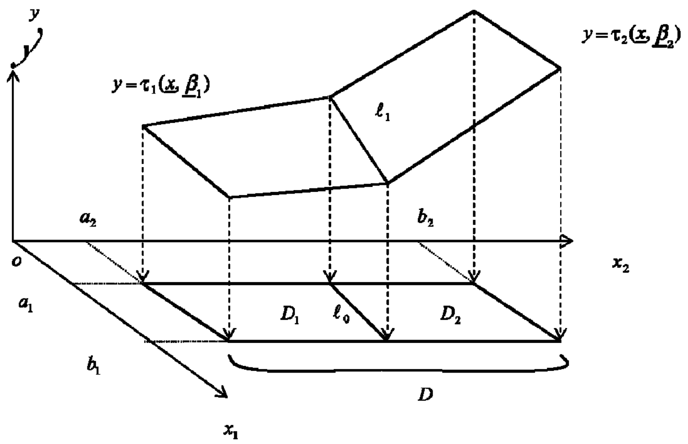

Fig. 1 The planes τ_1, τ_2, the change line l_1 and the domain D

regression function of the form

$$\tau \equiv \begin{cases} \tau_1(\boldsymbol{x}, \boldsymbol{\beta}_1) = \beta_0 + \beta_1 x_1 + \beta_2 x_2, \ \boldsymbol{x} \in D_1 \\ \tau_2(\boldsymbol{x}, \boldsymbol{\beta}_2) = \beta_0^+ + \beta_1^+ x_1 + \beta_2^+ x_2, \ \boldsymbol{x} \in D_2 \end{cases}$$

where $\boldsymbol{x} = (x_1, x_2)^t$ denotes an explanatory nonrandom variable taking values $\boldsymbol{x}_i = (x_{1i}, x_{2i})^t$, as observation points on a known direct product of two intervals $D = [a_1, b_1] \times [a_2, b_2]$. The direct product D is separated as $D = D_1 \oplus D_2$ with an unknown linear boundary l_0: $c_0^{(0)} + c_1^{(0)} x_1 + c_2^{(0)} x_2 = 0$ which passes through the points $(x_{10}, 0, 0)$, $(0, x_{20}, 0)$. That is, $D_1 = \{\boldsymbol{x} \mid \boldsymbol{x} \in D_1, c_0^{(0)} + c_1^{(0)} x_1 + c_2^{(0)} x_2 \leq 0\}$, $D_2 = \{\boldsymbol{x} \mid \boldsymbol{x} \in D_2, c_0^{(0)} + c_1^{(0)} x_1 + c_2^{(0)} x_2 > 0\}$. The vector $\boldsymbol{\beta} = (\boldsymbol{\beta}_1{}^t, \boldsymbol{\beta}_2{}^t)^t = (\beta_0, \beta_1, \beta_2; \beta_0^+, \beta_1^+, \beta_2^+)^t$ is that of unknown regression parameters.

As shown in Figure 1, the planes $y = \tau_1(\boldsymbol{x}, \boldsymbol{\beta}_1)$ and $y = \tau_2(\boldsymbol{x}, \boldsymbol{\beta}_2)$ are intersected with an unknown change line

$$l_1 : \frac{y}{c_0^{(1)}} = \frac{x_1 - d_1}{c_1^{(1)}} = \frac{x_2 - d_2}{c_2^{(1)}}$$

where

$$c_0^{(1)} = \beta_1 \beta_2^+ - \beta_1^+ \beta_2, \quad c_1^{(1)} = \beta_2^+ - \beta_2, \quad c_2^{(1)} = \beta_1 - \beta_1^+,$$

$$d_1 = \frac{\beta_2 \beta_0^+ - \beta_2^+ \beta_0}{\beta_1 \beta_2^+ - \beta_1^+ \beta_2}, \quad d_2 = \frac{\beta_1^+ \beta_0 - \beta_1 \beta_0^+}{\beta_1 \beta_2^+ - \beta_1^+ \beta_2}.$$

We here note that whole parameters of the change line are described in terms of the regression parameters; the discussion of the regression parameters such as estimation or asymptotic property results in that of the change line.

The function $\tau(\boldsymbol{x}, \boldsymbol{\beta})$ is continuously constrained by the line of intersection of l_1. We assume that the interior of l_0 is located in that of D. The line l_0 is the projection of l_1 to the x_1 x_2 plane.

In the model, $\boldsymbol{u} = (u_1, u_2, \cdots, u_n)^t$, $\boldsymbol{u}^- = (u_i, u_{i+1}, \cdots, u_{i-s+1})^t$ denotes the vector of error terms u_i which are assumed to form the sth order autoregressive process:

$$u_i = \rho_1 u_{i-1} + \rho_2 u_{i-2} + \cdots + \rho_s u_{i-s} + \varepsilon_i, \ i = s+1, s+2, \cdots, n,$$

where $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)^t$ stands for a vector of random variables ε_j independently and identically distributed (i.i.d.) with the conditions, $E(\varepsilon_j) = 0$, $E(\varepsilon_j^2) = \sigma^2 < \infty$, and $E(\varepsilon_j^{2(1+\delta)}) < \infty$, $j = 1, 2, \cdots, n$, for some $\delta > 0$. Here σ^2 is an unknown parameter and $\boldsymbol{\rho} = (\rho_1, \rho_2, \cdots, \rho_s)^t$ is an unknown vector of ρ_i satisfying the stationarity condition that the solutions of a polynomial equation $\lambda^s - \rho_1 \lambda^{s-1} - \rho_2 \lambda^{s-2} - \cdots - \rho_s = 0$ are all less than one in absolute value.

3. Estimation of the parameter true values

We consider least squares estimators of the parameters of the model. For every fixed $m \mid s+1 \leq m < n$, we define a matrix X with submatrices X_1, X_2:

$$X=\begin{pmatrix} X_1 & O \\ O & X_2 \end{pmatrix},\quad X_1=\begin{pmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{1m} & x_{2m} \end{pmatrix},\quad X_2=\begin{pmatrix} 1 & x_{1,m+1} & x_{2,m+1} \\ 1 & x_{1,m+2} & x_{2,m+2} \\ \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} \end{pmatrix}.$$

Using the matrix X, the model is rewritten as $Y = X\boldsymbol{\beta} + \boldsymbol{u}$. Further we define a transformation matrix[4)]

$$T=\begin{pmatrix} T_1 & O \\ O & T_2 \end{pmatrix}$$

with submatrices

$$T_1=(\underset{(m-s)\times s}{U} \;\vdots\; \underset{(m-s)\times(m-s)}{V}),\quad T_2=(\underset{(n-m-s)\times s}{U} \;\vdots\; \underset{(n-m-s)\times(n-m-s)}{V}),$$

where

$$U=\begin{pmatrix} -\rho_s & -\rho_{s-1} & \cdots & \cdots & -\rho_1 \\ 0 & -\rho_s & -\rho_{s-1} & \cdots & -\rho_2 \\ \vdots & 0 & -\rho_s & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\rho_s \end{pmatrix},$$

$$V=\begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ -\rho_1 & 1 & 0 & \cdots & \cdots & 0 \\ -\rho_2 & -\rho_1 & 1 & \cdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -\rho_{s-1} & \cdots & \cdots & \cdots & -\rho_1 & 1 \end{pmatrix}.$$

Premultiply the matrix T by the rewritten model, we have $TY = T(X\boldsymbol{\beta}+\boldsymbol{u}) = TX\boldsymbol{\beta}+\boldsymbol{\varepsilon}^-$, where $\boldsymbol{\varepsilon}^- = (\varepsilon_{s+1}, \varepsilon_{s+2}, \cdots, \varepsilon_m, \varepsilon_{m+s+1}, \varepsilon_{m+s+2}, \cdots, \varepsilon_n)^t$ is a random vector of $n-2s$ variables i.i.d. Now, fixing every $m \mid s+1 \leq m < n$, we first consider estimates $\hat{\boldsymbol{\beta}}_{m,n}$, $\hat{\boldsymbol{\rho}}_{m,n}$ of true parameter values $\boldsymbol{\beta}^*$, ρ^* as values minimizing the transformed residual sum of squares $S_{br}{}^2 \equiv S^2(\boldsymbol{\beta}, \boldsymbol{r}, m, n) = (T(Y-X\boldsymbol{\beta}))^t(T(Y-X\boldsymbol{\beta})) = (Y-X\boldsymbol{\beta})^t T^t T(Y-X\boldsymbol{\beta})$. We next obtain LSEs, $\hat{\boldsymbol{\beta}}_n$, $\hat{\boldsymbol{\rho}}_n$ of $\boldsymbol{\beta}^*$, $\boldsymbol{\rho}^*$ as those satisfying $S_e{}^2 \equiv S^2(\hat{\boldsymbol{\beta}}_n, \hat{\boldsymbol{\rho}}_n, \hat{m}_n, n) = \min\{S^2(\hat{\boldsymbol{\beta}}_{m,n}, \hat{\boldsymbol{\rho}}_{m,n}, m, n) \mid s+1 \leq m < n\}$ for an estimate $\hat{m}_n \mid s+1 \leq \hat{m}_n < n$ of the true value $m_n{}^*$.

Further, we obtain an LSE $\hat{\sigma}^2 = S_e{}^2/(\mathrm{n}-2s-6)$ together with a maximum likelihood estimator (MLE) $\tilde{\sigma}^2 = (n-2s-6)/(n-2s)\hat{\sigma}^2 = S_e{}^2/(n-2s)$.

For the selection of optimum order s among those of autoregressive processes, we adopt AIC and AICC, a corrected AIC[1)]:

$$\mathrm{AIC} = (n-2s)\log\left(\frac{2\pi S_e{}^2}{n-2s}\right) + n + 14,$$

$$\mathrm{AICC} = (n-2s)\log\left(\frac{2\pi S_e{}^2}{n-2s}\right) + (n-2s) + \frac{2(n-2s)(s+7)}{n-3s-8}.$$

4. Asymptotic properties of the estimators of the parameters

We briefly mention main theorems regarding the consistencies and the asymptotic distributions of the estimators of the parameters gained in the preceding section. The assumptions under which the theorems are formulated, and their proofs, are omitted.

Theorem 1 Under some assumptions, $\hat{\boldsymbol{\beta}}_n$, $\hat{\sigma}_n{}^2$ and $\hat{\boldsymbol{\rho}}_n$ converge in probability to $\boldsymbol{\beta}^*$, σ^{2*} and $\boldsymbol{\rho}^*$

respectively as $n \to \infty$.

Theorem 2 Assume that $\boldsymbol{\beta}^*$ is an interior point in the parameter space of admissible vector $\boldsymbol{\beta}$ satisfying the continuity constraint. Then, under several assumptions,

$$(n-2s)^{1/2}\begin{pmatrix}\hat{\boldsymbol{\beta}}_n-\boldsymbol{\beta}^*\\ \hat{\boldsymbol{\rho}}_n-\boldsymbol{\rho}^*\end{pmatrix}\sim N_{s+6}(\mathbf{0},\ \sigma^{*2}[I(\boldsymbol{\beta}^*,\ \boldsymbol{\rho}^*]^{-1}),$$

where

$$I(\boldsymbol{\beta}^*,\ \boldsymbol{\rho}^*)=\begin{pmatrix} I_1(\boldsymbol{\beta}_1) & O & O\\ O & I_2(\boldsymbol{\beta}_2) & O\\ O & O & E(\boldsymbol{u}^-\boldsymbol{u}^{-t})\end{pmatrix}$$

as $n \to \infty$.

Here $I_1(\boldsymbol{\beta}_1)$ and $I_2(\boldsymbol{\beta}_2)$ are third order square matrices whose elements are composed of definite integral values.

The latter theorem suggests that the estimators $\hat{\boldsymbol{\beta}}_1$, $\hat{\boldsymbol{\beta}}_2$ and $\hat{\boldsymbol{\rho}}_n$ are asymptotically independent.

5. Applied example

We apply the model to a set of data composed of the annual average temperatures y in degrees Centigrade, the cumulative number of cars $x_1{}'$ registered at the Ministry of Land, Infrastructure and Transport and the annual amount of water supply $x_2{}'$ in cubic meters for 1907–1999 in Tokyo, as shown in Table 1.

For the sake of variance stabilizing, we first carry out a common logarithmic transformation of $x_1{}'$ and $x_2{}'$ into x_1 and x_2, respectively. These x_1 and x_2 are regarded hereafter as the explanatory values for the data analysis.

We search for coordinates of two points $P_1=\{l_1\cap x_1\ y \text{ plane}\}$ and $P_2=\{l_1\cap x_2\ y \text{ plane}\}$: (1) fit a fairly high order polynomial regression curve to y against x_1 and y against x_2, respectively. (2) compute the respective radius of curvature r_c along each curve to find $P_1(x^{(1)}, y^{(1)})$ and $P_2(x^{(2)}, y^{(2)})$ at which r_c attains the minimum value.

To obtain the approximate values of the parameter, we repeat the following steps:

1° Fix a number $m(=2, 3, \cdots, n-1)$ corresponding to (x_{1m}, x_{2m}).

2° By the method of least squares, find initial values of $\hat{\boldsymbol{\beta}}$ minimizing

$$S_0{}^2=\sum_{i=1}^{m}\left[y_i-\left\{\beta_0+\beta_1\left(x_{1i}+\frac{x^{(1)}}{x^{(2)}}x_{2i}\right)+\left(\frac{y^{(2)}-y^{(1)}}{x^{(2)}}x_{2i}\right)\right\}\right]^2$$

$$+\sum_{i=m+1}^{n}\left[y_i-\left\{\beta_0{}^++\beta_1{}^+\left(x_{1i}+\frac{x^{(1)}}{x^{(2)}}x_{2i}\right)+\left(\frac{y^{(2)}-y^{(1)}}{x^{(2)}}x_{2i}\right)\right\}\right]^2.$$

Table 1　Annual average temperature, number of cars and total amount of water supply in Tokyo

t year	y average temperature (℃)	$x_1{}'$ cars registered	$x_2{}'$ water supply (m³)
1907	13.5	16	48,214,603
1908	13.2	44	49,869,666
⋮	⋮	⋮	⋮
1947	14.1	28,893	415,233,349
⋮	⋮	⋮	⋮
1998	16.7	4,654,196	1,672,471,000
1999	17.0	4,644,544	1,670,562,000

(omitted: 1909–1946, 1948–1997)

Incidentally,

$$\beta_2=\frac{y^{(2)}-y^{(1)}+\beta_1 x^{(1)}}{x^{(2)}}, \quad \beta_2{}^+=\frac{y^{(2)}-y^{(1)}+\beta_1{}^+ x^{(1)}}{x^{(2)}}.$$

3° Designate the order s concerning autoregressive correlated error terms.

4° Obtain $r=\hat{\rho}$, which minimizes

$$S_{br}{}^2=\sum_{i=s+1}^{m}\left(\tilde{u}_i-\sum_{t=1}^{s} r_t\tilde{u}_{i-t}\right)^2+\sum_{i=m+s+1}^{n}\left(\tilde{u}_i{}^+-\sum_{t=1}^{s} r_t\tilde{u}^*_{i-t}\right)^2$$

where

$$\tilde{u}_i=Y_i-(\hat{\beta}_0+\hat{\beta}_1 x_{1i}+\hat{\beta}_2 x_{2i}), \quad \tilde{u}_i{}^+=Y_i-(\hat{\beta}_0{}^+ +\hat{\beta}_1{}^+ x_{1i}+\hat{\beta}_2{}^+ x_{2i}).$$

5° Compute $\hat{\boldsymbol{\beta}}$ which minimizes

$$S_e^2=\sum_{i=s+1}^{m}\left[\left\{Y_i-\sum_{t=1}^{s} r_t Y_{i-t}\right\}-\left\{(\hat{\beta}_0+\hat{\beta}_1 x_{1i}+\hat{\beta}_2 x_{2i})-\sum_{t=1}^{s} r_t(\hat{\beta}_0+\hat{\beta}_1 x_{1,\,i-t}+\hat{\beta}_2 x_{2,\,i-t})\right\}\right]^2$$

$$+\sum_{i=m+s+1}^{n}\left[\left\{Y_i-\sum_{t=1}^{s} r_t Y_{i-t}\right\}-\left\{(\hat{\beta}_0{}^+ +\hat{\beta}_1{}^+ x_{1i}+\hat{\beta}_2{}^+ x_{2i})\right.\right.$$

$$\left.\left.-\sum_{t=1}^{s} r_t(\hat{\beta}_0{}^+ +\hat{\beta}_1{}^+ x_{1,\,i-t}+\hat{\beta}_2{}^+ x_{2,\,i-t})\right\}\right]^2.$$

6° Calculate AIC and AICC.

7° Return to step 2° until having the least AIC(AICC) for the fixed m.

8° Go back to step 1° until AIC(AICC) attains throughout the whole m.

Table 2 lists AIC(AICC), residual sums of squares $S_e{}^2$ and residual sums of squares and coefficients for the correlated error terms for each order s of the autoregressive process in the two-phase multiple regression model ('2' in column p, P2–model for short) together with those in the one-phase multiple regression model ('1' in column p, P1–model) for reference. From this analysis, the P2–model, with correlated error terms of the first order autoregressive process under the continuity constraint, is found to be the best one, since AIC and AICC take minimum values. Further, Table 3 displays the estimates of parameters of the best P2–model, in addition to those of the P1–

Table 2　The value of AIC (AICC) for model selection and the related

p	s	AIC	(AICC)	S_e^2	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\rho}_3$
2	0	122.248	(123.565)	17.438	—	—	—
	1	111.486	(113.243)	15.215	0.01164	—	—
	2	111.964	(114.242)	14.977	−0.01004	−0.00360	—
	3	113.421	(116.316)	14.907	−0.00936	0.01833	−0.00641
1	0	142.286	(142.741)	23.073	—	—	—

Table 3　Resultant estimates of parameters of the best model and estimates of P1–model

p	s	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_0{}^+$	$\hat{\beta}_1{}^+$	$\hat{\beta}_2{}^+$
2	1	6.091	0.088	0.955	48.483	7.018	−8.538
1	0	6.085	0.207	0.901	6.085	0.207	0.901

model.

Acknowledgments

This research was supported in part by Juntendo University, under joint research grant number 1–2003. Thanks are due to anonymous referees for their valuable comments. We are grateful to Professor Bruce Allen who made suggestions to aid the clarity of this paper.

References

1) Brockwell, P. J. and Davis, R. A. (1991) *Time Series: Theory and Methods*. New York: Springer-Verlag.
2) Cochrane, P. and Orcutt, G. H. (1949) Application of least squares regression to relationships containig auto-correlated error terms. *J. Amer. Statist. Assoc*. 44, 32–61.
3) Csörgo, M. and Hörvath, L. (1997) *Limit Theorems in Change-Point Analysis*. New York: John Wiley and Sons.
4) Johnston, J. (1984) *Econometric Methods. 3rd ed*. Singapore: McGraw-Hill, 321–325.
5) Kadiyala, K. R. (1968) A transformation used to circumvent the problem of autocorrelation. *Econometrika* 36, 93–96.
6) Krishnaiah, P. R. and Miao, B. Q. (1988) Review about estimation of change points. In *Handbook of Statistics* (P. R. Krishnaiah and C. R. Rao, eds). 7, 375–402.
7) Shaban, S. A. (1980) Change point problem and two-phase regression. *Inter. Statist. Review* 48, 83–93.

（平成15年11月10日　受付
　平成16年１月13日　受理）