应用概率统计　第 34 卷
第 2 期　2018 年 4 月

Chinese Journal of Applied Probability and Statistics
Apr., 2018, Vol. 34, No. 2, pp. 156-168
doi: 10.3969/j.issn.1001-4268.2018.02.004

Testing for Homogeneity of Exponential Correlation Nonlinear Mixed Models Based on M-estimation *

SUN HuiHui

(*School of Mathematics and Statistics, Yancheng Teachers University, Yancheng, 224002, China*)

LIN JinGuan⋆

(*School of Science, Nanjing Audit University, Nanjing, 211815, China*)

Abstract: Homogeneity of variance and correlation coefficients is one of assumptions in the analysis of longitudinal data. However, the assumption can be challenged. In this paper, we mainly propose and analyze nonlinear mixed effects models for longitudinal data with exponential correlation covariance structure, intend to introduce Huber's function in the log likelihood function and get robust estimation (M-estimation) by Fisher scoring method. Score test statistics for homogeneity of variance and correlation coefficient based on M-estimation are then studied. A simulation study is carried to assess the performance of test statistics and the method we proposed in the paper is illustrated by an actual data example.

Keywords: nonlinear mixed models; exponential correlation; M-estimation; hypothesis testing

2010 Mathematics Subject Classification: 62J20

Citation: SUN H H, LIN J G. Testing for homogeneity of exponential correlation nonlinear mixed models based on M-estimation [J]. Chinese J Appl Probab Statist, 2018, 34(2): 156–168.

§1. Introduction

Mixed effects models offer a flexible framework by which to model the sources of variation and correlation that arise from grouped data. It can be used to model both linear and nonlinear relationships between dependent and independent variables. Linear mixed effects models can be used to express linear relationships between sets of variables, nonlinear models can model mechanistic relationships between independent and dependent variables and can estimate more physically interpretable parameters [1]. Nonlinear mixed effects models are important to the analysis of longitudinal data, multi-level data and repeated survey data, which widely exist in the field of economics, bio-pharmaceuticals,

*The project was supported by the National Statistical Science Research Project of China Foundation (Grant Nos. 2014LZ14; 2015LZ27) and the National Social Science Foundation of China (Grant No. 17BTJ026).

⋆Corresponding author, E-mail: jglin@nau.edu.cn.

Received March 7, 2016. Revised June 12, 2016.

agriculture and so on. Lin and Wei [2] and Lin [3] respectively studied testing for heteroscedasticity and autocorrelation in mixed effects nonlinear models with AR(1) errors and varying dispersion in generalized nonlinear models for longitudinal data. Pinheiro and Bates [1] considered linear models with random effects for longitudinal data and made some statistical diagnosis under the assumption of homogeneity of variance. Sun [4] made score test of correlation coefficients in uniform correlation mixed effects linear models based on robust estimation. Diggle et al. [5] discussed longitudinal data models with exponential correlation structure in detail, which is another correlation structure.

As is known to all, statistical inferences based on mixed models with normal random effects and residual errors are vulnerable to outliers. While suppose that random effects and errors follow distributions with heavy tails will enable the model to produce more robust estimates against outliers [6,7]. Pinheiro et al. [8], Lin and Lee [9] and Staudenmayer et al. [10] researched robust estimation of models which both random effects and errors have multivariate Student-t distributions. In this work, normality is still assumed for random effects and residual errors and we will improve the robustness of maximum likelihood estimation (MLE) through robust maximum likelihood estimation (RMLE). M-estimation is the most widely used robust estimation method, which was firstly introduced by Huber [11] on regression. Some recent works can be found in the literature involving robust estimation. Huggins [12] introduced a robust approach to the analysis of repeated measures. Sinha [13] developed a robust quasi-likelihood method, which appears to be useful for down-weighting any influential data points when estimating parameters in generalized linear mixed models. Yeap and Davidian [14] introduced a robust two-stage procedure for robust estimation in nonlinear mixed effects models. Gill [15] applied scoring method to obtain robust estimations of linear mixed model for longitudinal data. Most of the above researches focused on algorithms and robustness of the robust estimation while statistical diagnosis have received limited attention in the robust estimation context.

The purpose of this paper is to propose score-type test statistics in order to assess the homogeneity of variance and correlation coefficients in nonlinear mixed effects models with exponential correlation covariance structure. The structure of this paper is as follows. Section 2 introduces the exponential correlation nonlinear mixed models and uses Fisher scoring method to get M-estimation of parameters, which is called robust maximum likelihood estimation (RMLE). In Section 3, we derive score test statistics based on M-estimation, the properties of test statistics are investigated in Section 4 through Monte Carlo simulations. An example is analyzed in Section 5 to illustrate the proposed methodology.

§2. Model and Estimation

Suppose that response measurements are collected on m subjects and the ith subject being observed on n_i time points, thus $N = \sum\limits_{i=1}^{m} n_i$ is the total number of measurements. In the matrix notation, the model for measurements from subject i is

$$\boldsymbol{y}_i = f(\boldsymbol{X}_i, \boldsymbol{\beta}) + \boldsymbol{C}_i\boldsymbol{\tau}_i + \boldsymbol{e}_i, \qquad i = 1, 2, \ldots, m, \tag{1}$$

where $\boldsymbol{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in_i})^{\mathsf{T}}$ is a vector of length n_i containing observable response variable from subject i and $t_{i1}, t_{i2}, \ldots, t_{in_i}$ is the observation time. $f(\cdot,\cdot)$ is a known twice differentiable nonlinear function of the regression vector $\boldsymbol{\beta}$, which is a vector of p unknown but fixed parameters with known design matrix $\boldsymbol{X}_k$, and $\boldsymbol{X}_i = (\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \ldots, \boldsymbol{x}_{in_i})^{\mathsf{T}}$; $\boldsymbol{C}_i$ is the $n_i \times r$ design matrix for the random effects of subject i, $\boldsymbol{\tau}_i$ is an $r \times 1$ vector of unobservable random efforts assumed to be sampled from $\mathrm{N}(0, \sigma^2\Gamma_i)$; the random error vector $\boldsymbol{e}_i \sim \mathrm{N}(0, \sigma^2 V_i)$, where $V_i = (v_{jk})_{m\times n_i}$ and $v_{jk} = \mathsf{Cov}\,(y_{ij}, y_{ik}) = \exp\{-\phi_i|t_j - t_k|\}$ characterizes the exponential correlation structure. And $\boldsymbol{\tau}_i$ and $\boldsymbol{e}_i$ are independent from each other. Then

$$\mathsf{Cov}\,(\boldsymbol{y}_i) = \sigma^2\Sigma_i = \sigma^2\boldsymbol{C}_i\Gamma_i\boldsymbol{C}_i^{\mathsf{T}} + \sigma^2 V_i.$$

Let $\boldsymbol{\alpha}$ denote the vector of unknown parameters in Σ_i, and the log-likelihood for the nonlinear mixed model is

$$l(\boldsymbol{\beta}, \boldsymbol{\alpha}, \sigma^2 \,|\, \boldsymbol{y}) = \text{constant} - \frac{1}{2}N\ln\sigma^2 - \frac{1}{2}\sum_{i=1}^{m}\ln|\Sigma_i| - \sum_{i=1}^{m}\frac{1}{2}\boldsymbol{\varepsilon}_i^{\mathsf{T}}\boldsymbol{\varepsilon}_i, \tag{2}$$

where $\boldsymbol{\varepsilon}_i = \sigma^{-1}\Sigma_i^{-1/2}(\boldsymbol{y}_i - f(\boldsymbol{X}_i, \boldsymbol{\beta}))$. Note that the last term of (2) is a half sum of squares and grows quickly. If it is replaced by a function that grows more slowly, the robustness can be improved and then the influence of outliers on the estimate can be limited. The algorithm for robust estimation based on the above idea is mostly M-estimation, it is the most widely used robust estimation method.

In this paper, Huber's ρ function is chosen to bound the influence of outlying observations on the estimation.

$$\rho(\varepsilon) = \begin{cases} \varepsilon^2/2 & \text{if } |\varepsilon| \leqslant c; \\ c|\varepsilon| - c^2/2 & \text{if } |\varepsilon| > c, \end{cases}$$

where c is some fixed constant and usually $c \in [0.7, 2]$, here $c = 1.345$ [16]. For this function

$$\psi(\varepsilon) = \partial\rho(\varepsilon)/\partial\varepsilon = \begin{cases} \varepsilon & \text{if } |\varepsilon| \leqslant c; \\ c\,\mathrm{sign}(\varepsilon) & \text{if } |\varepsilon| > c. \end{cases}$$

Therefore, robustified version of (2) is given by

$$\eta(\boldsymbol{\beta},\boldsymbol{\alpha},\sigma^2)=\text{constant}-\frac{1}{2}\kappa_1 N\ln\sigma^2-\frac{1}{2}\kappa_1\sum_{i=1}^{m}\ln|\Sigma_i|-\sum_{i=1}^{m}\sum_{j=1}^{n_i}\rho(\varepsilon_{ji}),\tag{3}$$

where $\kappa_1=\mathsf{E}(\varepsilon\psi(\varepsilon))=\mathsf{P}(|\varepsilon|\leqslant c)$ is the consistency correction factor, the expectation being taken over the distribution of ε [12].

We will obtain robust estimation of regression parameters and covariance parameters through Fisher scoring method [15,17] based on (3).

First, score equation and the expected Hessian matrix for estimating $\boldsymbol{\beta}$ are given by

$$\frac{\partial\eta}{\partial\boldsymbol{\beta}}=\sigma^{-1}\sum_{i=1}^{m}\dot{f}_{ik}^{\mathsf{T}}\Sigma_i^{-1/2}\psi[\sigma^{-1}\Sigma_i^{-1/2}(\boldsymbol{y}_i-f(\boldsymbol{X}_i,\boldsymbol{\beta}))]$$

and

$$\frac{\partial^2\eta}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^{\mathsf{T}}}=-\sigma^{-2}\sum_{i=1}^{m}\dot{f}_{ik}^{\mathsf{T}}\Sigma_i^{-1/2}\Lambda\Sigma_i^{-1/2}\dot{f}_{il},$$

$$H_{\boldsymbol{\beta}\boldsymbol{\beta}^{\mathsf{T}}}=\mathsf{E}\Big(-\frac{\partial^2\eta}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^{\mathsf{T}}}\Big)=\nu\sigma^{-2}\sum_{i=1}^{m}\dot{f}_{ik}^{\mathsf{T}}\Sigma_i^{-1}\dot{f}_{il},$$

where $\dot{f}_{ik}=\partial f_i/\partial\beta_k$, $\dot{f}_{il}=\partial f_i/\partial\beta_l$. Λ is a diagonal matrix and if $|\varepsilon|\leqslant c$, $\Lambda_{jj}=\partial\psi(\varepsilon)/\partial\varepsilon=1$ otherwise, $\Lambda_{jj}=0$. $\nu=\mathsf{E}(\Lambda_{jj})=\mathsf{P}(|\varepsilon|\leqslant c)=\int_{-c}^{c}(2\pi)^{-1/2}\mathrm{e}^{-1/2\varepsilon^2}\mathrm{d}\varepsilon=\kappa_1$.

With the current estimates $\widehat{\boldsymbol{\beta}}^{(h)}$ at the hth iteration step, the next iteration of the scoring procedure is

$$\widehat{\boldsymbol{\beta}}^{(h+1)}=\widehat{\boldsymbol{\beta}}^{(h)}+\big(\widehat{H}_{\boldsymbol{\beta}\boldsymbol{\beta}^{\mathsf{T}}}^{(h)}\big)^{-1}\frac{\partial\eta^{(h)}}{\partial\boldsymbol{\beta}}.$$

Suppose the sequence converges to $\widehat{\boldsymbol{\beta}}$ which is the robust maximum likelihood estimation (RMLE) of $\boldsymbol{\beta}$. Note that the robust estimator $\widehat{\boldsymbol{\beta}}$ reduces to the classical ML form when $c=\infty$. With the same procedure we can get the robust estimation of covariance parameters. Domowitz and White [18] and Sinha [19] proved that the RMLE for parameters shares the same property as MLE, such as consistency and asymptotic normality for large sample sizes.

§3. Test of Homogeneity of Exponential Correlation Coefficients and Variance

3.1 Test of Homogeneity of Exponential Correlation Coefficients

In this subsection, we discuss test of homogeneity of exponential correlation coefficients ϕ_i under the assumption of homogeneity of variance, i.e. $\Gamma_i=\Gamma$ for all i. Parameterize ϕ_i as the form according to [20],

$$\phi_i = \phi\omega_i = \phi\omega(\upsilon_i, \boldsymbol{\gamma}),$$

where υ_i is a covariate and the range is $[-1, 1]$ for ϕ_i, $i = 1, 2, \ldots, m$. Furthermore, we assume that there exists a $\boldsymbol{\gamma}_0$ such that $\omega(\upsilon_i, \boldsymbol{\gamma}_0) = c \neq 0$ for all i, where c is a constant that is independent of i. Test of homogeneity of exponential correlation coefficients reduce to the test of hypothesis

$$H_0 : \boldsymbol{\gamma} = \boldsymbol{\gamma}_0 \qquad H_1 : \boldsymbol{\gamma} \neq \boldsymbol{\gamma}_0. \tag{4}$$

Denote $\boldsymbol{\theta}_1 = (\boldsymbol{\gamma}^\mathsf{T}, \boldsymbol{\beta}^\mathsf{T}, \phi, \sigma^2, \boldsymbol{\delta}^\mathsf{T})^\mathsf{T}$, where $\boldsymbol{\gamma}^\mathsf{T}$ is the parameter of interest and $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_{r'})^\mathsf{T} = (d_{11}, d_{12}, \ldots, d_{1r}, d_{22}, \ldots, d_{rr})^\mathsf{T}$ is an $r' = r(r+1)/2$ vector, d_{ij} is the (i, j)th element of Γ. The robustified version of the log-likelihood function of $\boldsymbol{\theta}_1$ can be expressed in the form of (3), next we will study the score test statistic of hypothesis problem (4) based on (3). By some tedious calculations, the Fisher information matrix of the parameters in $\boldsymbol{\theta}_1$ under the null hypothesis in block partitioned form is found to be

$$I(\boldsymbol{\theta}_1) = \begin{bmatrix} I_{\gamma\gamma} & 0 & I_{\gamma\phi} & I_{\gamma\sigma^2} & I_{\gamma\boldsymbol{\delta}} \\ 0 & I_{\boldsymbol{\beta\beta}} & 0 & 0 & 0 \\ I_{\phi\gamma} & 0 & I_{\phi\phi} & I_{\phi\sigma^2} & I_{\phi\boldsymbol{\delta}} \\ I_{\sigma^2\gamma} & 0 & I_{\sigma^2\phi} & I_{\sigma^2\sigma^2} & I_{\sigma^2\boldsymbol{\delta}} \\ I_{\boldsymbol{\delta}\gamma} & 0 & I_{\boldsymbol{\delta}\phi} & I_{\boldsymbol{\delta}\sigma^2} & I_{\boldsymbol{\delta\delta}} \end{bmatrix},$$

nonzero subblocks in Fisher information matrix $I(\boldsymbol{\theta}_1)$ are described in the Appendix.

Cox and Hinkley [21] considered the score test in the likelihood setting and Rotnitzky and Jewell [22] introduced a score-type statistic to testing hypothesis in the context of inference functions. Under some regular conditions, the score test statistic for the test of hypothesis $H_0 : \boldsymbol{\gamma} = \boldsymbol{\gamma}_0$ is as follows

$$\mathrm{SC}_1 = \Big\{ \Big(\frac{\partial \eta}{\partial \boldsymbol{\gamma}} \Big)^\mathsf{T} (m I^{\gamma\gamma} (J_m^{\gamma\gamma})^{-1} I^{\gamma\gamma}) \Big(\frac{\partial \eta}{\partial \boldsymbol{\gamma}} \Big) \Big\}_{\widehat{\boldsymbol{\theta}}_{10}},$$

where $\partial\eta/\partial\boldsymbol{\gamma}$ is the score function for null hypothesis and

$$\frac{\partial \eta}{\partial \boldsymbol{\gamma}} = \frac{1}{2} \Big\{ \sum_{i=1}^m \psi(\widehat{\boldsymbol{\varepsilon}}_i)^\mathsf{T} \widehat{\Sigma}_i^{-1} \frac{\partial V_i}{\partial \gamma_k} \widehat{\boldsymbol{\varepsilon}}_i - \kappa_1 \sum_{i=1}^m \mathrm{tr} \Big(\widehat{\Sigma}_i^{-1} \frac{\partial V_i}{\partial \gamma_k} \Big) \Big\}_{q \times 1}.$$

And $I^{\gamma\gamma}$ denotes the subblock in the inverse of $I(\boldsymbol{\theta}_1)$ corresponding to $I_{\gamma\gamma}$. By the asymptotic normal properties of the RMLE of parameters under some regularity conditions [23], we can calculate that

$$J_m(\boldsymbol{\theta}_1) = m^{-1} I(\boldsymbol{\theta}_1).$$

Then $J_m^{\gamma\gamma}$ can be calculated by the same procedure as $I^{\gamma\gamma}$. And $\widehat{\boldsymbol{\theta}}_{10}$ denotes the RMLE of $\boldsymbol{\theta}_1$ under H_0.

3.2 Test of Homogeneity of Variance

Longitudinal data are usually complexity and the assumption that homogeneity of variance is not necessarily appropriate. Therefore, testing of heteroscedasticity in mixed model with exponential correlation is very necessary. We will consider test of heteroscedasticity under the assumption of homogeneity of exponential correlation coefficients, i.e. $\phi_i = \phi$ for all i and V_i is denoted by V_{i0} here. And

$$\Sigma_i = V_{i0} + C_i \Gamma_i C_i^{\mathsf{T}}, \qquad \Gamma_i = \begin{pmatrix} \Gamma_{11i} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix},$$

where $\Gamma_{11i} = h_i \Gamma_{11}$, $h_i = h(\varsigma_i, \boldsymbol{\gamma})$, and there is a unique value $\boldsymbol{\gamma}_0$ of $\boldsymbol{\gamma}$ such that $h(\varsigma_i, \boldsymbol{\gamma}_0) = 1$, for all $i = 1, 2, \ldots, m$. Then the test of homogeneity of variance also reduce to the test of hypothesis $H_0 : \boldsymbol{\gamma} = \boldsymbol{\gamma}_0$. Correspondingly, we partition $C_i = \begin{pmatrix} C_i^{(1)} & C_i^{(2)} \end{pmatrix}$. Denote $\boldsymbol{\theta}_2 = (\boldsymbol{\gamma}^{\mathsf{T}}, \boldsymbol{\beta}^{\mathsf{T}}, \phi, \sigma^2, \boldsymbol{\delta}^{\mathsf{T}})^{\mathsf{T}}$. Fisher information matrix of the parameters in $\boldsymbol{\theta}_2$ under the null hypothesis in block partitioned form is found to be

$$I(\boldsymbol{\theta}_2) = \begin{bmatrix} I_{\boldsymbol{\gamma}\boldsymbol{\gamma}} & 0 & I_{\boldsymbol{\gamma}\phi} & I_{\boldsymbol{\gamma}\sigma^2} & I_{\boldsymbol{\gamma}\boldsymbol{\delta}} \\ 0 & I_{\boldsymbol{\beta}\boldsymbol{\beta}} & 0 & 0 & 0 \\ I_{\phi\boldsymbol{\gamma}} & 0 & I_{\phi\phi} & I_{\phi\sigma^2} & I_{\phi\boldsymbol{\delta}} \\ I_{\sigma^2\boldsymbol{\gamma}} & 0 & I_{\sigma^2\phi} & I_{\sigma^2\sigma^2} & I_{\sigma^2\boldsymbol{\delta}} \\ I_{\boldsymbol{\delta}\boldsymbol{\gamma}} & 0 & I_{\boldsymbol{\delta}\phi} & I_{\boldsymbol{\delta}\sigma^2} & I_{\boldsymbol{\delta}\boldsymbol{\delta}} \end{bmatrix}.$$

Nonzero subblocks in Fisher information matrix $I(\boldsymbol{\theta}_2)$ are given in the Appendix. Similar to the above subsection,

$$J_m(\boldsymbol{\theta}_2) = m^{-1} I(\boldsymbol{\theta}_2).$$

Then we obtain $I^{\boldsymbol{\gamma}\boldsymbol{\gamma}}$ and $J_m^{\boldsymbol{\gamma}\boldsymbol{\gamma}}$ from $I(\boldsymbol{\theta}_2)$ and $J_m(\boldsymbol{\theta}_2)$. Under some regular conditions, the score test statistic of homogeneity of variance, denotes by SC_2 is as follows

$$\mathrm{SC}_2 = \Big\{ \Big(\frac{\partial \eta}{\partial \boldsymbol{\gamma}}\Big)^{\mathsf{T}} (m I^{\boldsymbol{\gamma}\boldsymbol{\gamma}} (J_m^{\boldsymbol{\gamma}\boldsymbol{\gamma}})^{-1} I^{\boldsymbol{\gamma}\boldsymbol{\gamma}}) \Big(\frac{\partial \eta}{\partial \boldsymbol{\gamma}}\Big) \Big\}_{\widehat{\boldsymbol{\theta}}_{20}}$$

with score function for the null hypothesis

$$\frac{\partial \eta}{\partial \boldsymbol{\gamma}} = \frac{1}{2} \Big(\sum_{i=1}^m \psi(\widehat{\boldsymbol{\varepsilon}}_i)^{\mathsf{T}} \widehat{\Sigma}_i^{-1} \dot{h}_{ik} C_i^{(1)} \widehat{\Gamma}_{11} C_i^{(1)\mathsf{T}} \widehat{\boldsymbol{\varepsilon}}_i - \kappa_1 \sum_{i=1}^m \mathrm{tr}\big(\widehat{\Sigma}_i^{-1} \dot{h}_{ik} C_i^{(1)} \widehat{\Gamma}_{11} C_i^{(1)\mathsf{T}} \big) \Big)_{q \times 1}$$

and $\widehat{\Sigma}_i = C_i \widehat{\Gamma} C_i^{\mathsf{T}} + V_{i0}$, $\dot{h}_{ik} = \partial h_i / \partial \gamma_k$.

By the above results, we can calculate the score test statistic for homogeneity of variance, denotes by SC_2.

3.3 Joint Test of Homogeneity of Correlation Coefficient and Variance

In this case, we assume that both the exponential correlation coefficient and variance may be variable, and $\boldsymbol{\gamma}$ is the common parameter of heteroscedasticity and non-homogeneity correlation coefficients for uniformity. And $\Sigma_i = V_i + C_i\Gamma_i C_i^{\top}$. The test for homogeneity of correlation coefficient and variance reduce to the test of hypothesis (4).

Let $\boldsymbol{\theta}_3 = (\boldsymbol{\gamma}^{\top}, \boldsymbol{\beta}^{\top}, \phi, \sigma^2, \boldsymbol{\delta}^{\top})^{\top}$. The score function under the null hypothesis is given by

$$\begin{aligned}\frac{\partial\eta}{\partial\boldsymbol{\gamma}} = \frac{1}{2}\Big(&\sum_{i=1}^{m}\psi(\widehat{\boldsymbol{\varepsilon}}_i)^{\top}\widehat{\Sigma}_i^{-1}\big(\dot{h}_{ik}C_i^{(1)}\widehat{\Gamma}_{11}C_i^{(1)\top} + \dot{V}_{ik}\big)\widehat{\boldsymbol{\varepsilon}}_i \\ &- \kappa_1\sum_{i=1}^{m}\operatorname{tr}\big(\widehat{\Sigma}_i^{-1}\big(\dot{h}_{ik}C_i^{(1)}\widehat{\Gamma}_{11}C_i^{(1)\top} + \dot{V}_{ik}\big)\big)\Big)_{q\times 1},\end{aligned}$$

where $\dot{V}_{ik} = \partial V_i/\partial\gamma_k$.

Fisher information matrix of parameters in $\boldsymbol{\theta}_3$ under H_0 has the same form as $I(\boldsymbol{\theta}_2)$ and different subblocks are as follows

$$\begin{aligned}I_{\gamma\gamma} &= \Big(\frac{1}{2}\kappa_1\sum_{i=1}^{m}\operatorname{tr}\big(\widehat{\Sigma}_i^{-1}\big(\dot{h}_{ik}C_i^{(1)}\widehat{\Gamma}_{11}C_i^{(1)\top} + \dot{V}_{ik}\big)\big)^2 \\ &\qquad + \widehat{\Sigma}_i^{-1}\big(\dot{h}_{ik}\dot{h}_{il}C_i^{(1)}\widehat{\Gamma}_{11}C_i^{(1)\top} + \dot{V}_{ik}\dot{V}_{il}\big)\Big)_{q\times q}, \\ I_{\gamma\phi} &= \Big(\frac{1}{2}\kappa_1\sum_{i=1}^{m}\operatorname{tr}\big(\widehat{\Sigma}_i^{-1}\dot{V}_{i0}\widehat{\Sigma}_k^{-1}\big(\dot{h}_{ik}C_i^{(1)}\widehat{\Gamma}_{11}C_i^{(1)\top} + \dot{V}_{ik}\big) + \widehat{\Sigma}_k^{-1}\dot{V}_{i0}^{\top}\big)\Big)_{q\times 1}, \\ I_{\gamma\sigma^2} &= \Big(\frac{1}{4}\kappa_1(1+\kappa_2)\widehat{\sigma}^{-2}\sum_{i=1}^{m}\operatorname{tr}\big(\widehat{\Sigma}_i^{-1}\big(\dot{h}_{ik}C_i^{(1)}\widehat{\Gamma}_{11}C_i^{(1)\top} + \dot{V}_{ik}\big)\big)\Big)_{q\times 1}, \\ I_{\gamma\boldsymbol{\delta}} &= \Big(\frac{1}{2}\kappa_1\sum_{i=1}^{m}\operatorname{tr}\big(\widehat{\Sigma}_i^{-1}C_iE_{ab}^{i}C_i^{\top}\Sigma_i^{-1}\big(\dot{h}_{ik}C_i^{(1)}\widehat{\Gamma}_{11}C_i^{(1)\top} + \dot{V}_{ik}\big)\big)\Big)_{q\times r'}.\end{aligned}$$

The score test statistic for homogeneity of correlation coefficient and variance has the same form as SC_2 and can be obtained by the same procedure as SC_2, written as SC_3.

Now consider the asymptotic properties of the proposed score test statistics. Here "asymptotic" refers to the number of clusters $m \to \infty$ with cluster sizes n_i bounded. Under some regularity conditions, the asymptotic distribution of the above score test statistic is $\chi^2(q)$ when H_0 is true [23].

§4. Simulation Studies

In this section, the performance of the proposed test statistics is examined by Monte Carlo simulations. The model considered here is

$$y_{ij} = \beta_1\exp(\beta_2 x_{ij}) + \tau_i + e_{ij}, \qquad i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n,$$

where $e_i \sim \mathrm{N}(0, \sigma^2 V_i)$, $\tau_i \sim \mathrm{N}(0, \sigma^2 \Gamma_i)$ and $V_i = (v_{jk})_{m\times n}$ and $v_{jk} = \exp\{-\phi_i |t_j - t_k|\}$. Suppose $\Gamma_i = \Gamma \exp(z_i\gamma)$ and $\phi_i = \phi \exp(z_i\gamma)/[1+\exp(z_i\gamma)]$. Assume that the observation time intervals are $1, 2, 3, \ldots$

The covariates x_{ij}, z_i are, respectively, generated from discrete uniform distribution $[0, 20]$ and $[1, 15]$. The true parameter values are $\beta_1 = 1$, $\beta_2 = 1.5$, $\sigma^2 = 0.05$, $\Gamma = 0.1$, $\phi = 0.05$. For the given γ, together with x_{ij}, z_i, we get e_{ij} and τ_i. Furthermore, y_{ij} is obtained. The simulations are performed for different sample sizes (m, n) and different values of γ, each simulated case is replicated 2 000 times. Then the proportion of times which rejected the null hypothesis is just the simulated value of power. The nominal sizes of the tests are set to be 0.05.

(1) Testing for homogeneity of exponential correlation coefficients. Assume that the variances are homogeneous, the power for score test statistic is shown in Table 1.

Table 1 Simulation results of the test of homogeneity of exponential correlation coefficients

(m,n)	$\gamma=-0.3$	$\gamma=-0.2$	$\gamma=-0.1$	$\gamma=0$	$\gamma=0.1$	$\gamma=0.2$	$\gamma=0.3$
(10,20)	0.841	0.692	0.415	0.041	0.382	0.583	0.837
	(0.892)	(0.710)	(0.509)	(0.043)	(0.440)	(0.602)	(0.912)
	(0.964)	(0.781)	(0.567)	(0.044)	(0.512)	(0.699)	(0.947)
(20,20)	0.925	0.765	0.573	0.040	0.508	0.692	0.914
	(0.947)	(0.801)	(0.654)	(0.048)	(0.631)	(0.834)	(0.965)
	(0.987)	(0.841)	(0.694)	(0.045)	(0.701)	(0.897)	(0.979)
(30,25)	1.000	0.901	0.712	0.044	0.695	0.799	0.987
	(1.000)	(0.973)	(0.806)	(0.047)	(0.726)	(0.984)	(1.000)
	(1.000)	(0.990)	(0.878)	(0.046)	(0.843)	(1.000)	(1.000)

(2) Testing for homogeneity of variance. The powers for score test statistic are shown in Table 2.

Table 2 Simulation results of the test of homogeneity of variance

(m,n)	$\gamma=0$	$\gamma=0.01$	$\gamma=0.02$	$\gamma=0.03$	$\gamma=0.04$	$\gamma=0.05$	$\gamma=0.06$
(20,15)	0.042	0.274	0.426	0.597	0.789	0.891	0.930
	(0.041)	(0.305)	(0.597)	(0.735)	(0.826)	(0.963)	(1.000)
	(0.043)	(0.482)	(0.704)	(0.901)	(0.972)	(0.990)	(1.000)
(30,20)	0.041	0.281	0.512	0.683	0.833	0.972	1.000
	(0.041)	(0.326)	(0.643)	(0.798)	(0.915)	(1.000)	(1.000)
	(0.042)	(0.479)	(0.818)	(0.885)	(0.998)	(1.000)	(1.000)
(40,30)	0.043	0.311	0.609	0.857	0.964	1.000	1.000
	(0.045)	(0.446)	(0.842)	(0.947)	(0.991)	(1.000)	(1.000)
	(0.045)	(0.624)	(0.899)	(0.997)	(1.000)	(1.000)	(1.000)

(3) Joint testing for homogeneity of correlation coefficient and variance. The powers for score test statistic are shown in Table 3.

Table 3 Simulation results of the test for homogeneity of correlation coefficient and variance

(m,n)	$\gamma=0$	$\gamma=0.05$	$\gamma=0.1$	$\gamma=0.15$	$\gamma=0.2$	$\gamma=0.25$	$\gamma=0.3$
(20,20)	0.040	0.214	0.407	0.634	0.797	0.895	0.969
	(0.043)	(0.269)	(0.484)	(0.759)	(0.884)	(0.937)	(0.989)
	(0.044)	(0.417)	(0.698)	(0.907)	(0.972)	(0.991)	(1.000)
(20,30)	0.043	0.265	0.479	0.630	0.841	0.911	0.989
	(0.043)	(0.349)	(0.662)	(0.848)	(0.921)	(0.989)	(1.000)
	(0.045)	(0.438)	(0.827)	(0.908)	(0.992)	(1.000)	(1.000)
(30,30)	0.042	0.312	0.716	0.882	0.994	1.000	1.000
	(0.044)	(0.495)	(0.826)	(0.972)	(0.998)	(1.000)	(1.000)
	(0.053)	(0.634)	(0.902)	(0.997)	(1.000)	(1.000)	(1.000)

The results in Tables 1 − 3 show that the empirical sizes of the tests are very close to 0.05 under $\gamma = 0$. The power of the test increases and approaches 1 as γ or sample size increases. In the above three tables, figures in brackets represent the simulated values when $c = 2$ and $c = \infty$ (i.e. nonrobust situation) in Huber's function respectively, which are bigger than that discussed in this paper. These findings are consistent with theoretical results.

§5. An Illustrative Example

In this section, plasma drug penetration data is used to illustrate the application of the method. The data of the the example are taken form the following experiment: six volunteers were injected the same dose of a drug through veins. Then the drug concentration in their plasma was measured at eleven times within eight hours. The following double exponential model [24] was applied to describe the data:

$$y_{ij} = \mathrm{e}^{\beta_1}\exp(-\mathrm{e}^{\beta_2}x_{ij}) + \mathrm{e}^{\beta_3}\exp(-\mathrm{e}^{\beta_4}x_{ij}) + \tau_i + e_{ij}, \qquad i=1,2,\ldots,6;\ j=1,2,\ldots,11,$$

where y_{ij} denotes the drug concentration of the i-th volunteer at the j-th time; x_{ij} is the time interval of the i-th volunteer measure the j-th time drug concentration; τ_i is the drug effect of the i-th volunteer and $\tau_i \sim \mathrm{N}(0,\sigma^2\delta_i)$; the random error e_{ij} is the element of the vector e_i and $e_i \sim \mathrm{N}(0,\sigma^2 V_i)$, where $V_i = (v_{jk})_{6\times 11}$ and $v_{jk} = \mathsf{Cov}\,(y_{ij}, y_{ik}) = \exp\{-\phi_i|t_j - t_k|\}$ characterizes the exponential correlation structure. Design matrix $X_i =$

(0.25, 0.50, 0.75, 1.00, 1.25, 2.00, 3.00, 4.00, 5.00, 6.00, 8.00)$^\top$. Let $h_i = \exp(\upsilon_i\gamma)$ and $\omega_i = \exp(\upsilon_i\gamma)/[1+\exp(\upsilon_i\gamma)]$, where υ_i is the attribute variable for the i-th volunteer. Here $h_i = 1$ and $\omega_i = 1/2$ when $\gamma = 0$, both have no correlation with i.

First, we get M-estimation according to the algorithm proposed in Section 2.

$$\widehat{\boldsymbol{\beta}} = (1.6450,\ 0.8384,\ -0.4697,\ -1.0392)^\top,$$
$$\widehat{\sigma}^2 = 0.07831, \qquad \widehat{\delta} = 0.4608, \qquad \widehat{\phi} = 0.638.$$

Now, we study the following cases. (a) Test for homogeneity of exponential correlation coefficients. (b) Test for homogeneity of variance. (c) Joint test for homogeneity of correlation coefficient and variance. All the tests reduce to test $H_0 : \gamma = 0$.

Compute out the score test statistic and the results are listed in Table 4.

Table 4 Score test for cases (a) – (c)

c	(a)	(b)	(c)
1.345	22.3646	27.3448	19.2701
	(0.00)	(0.00)	(0.00)
0.7	18.6270	25.1392	15.1810
	(0.00)	(0.00)	(0.00)
∞	25.5291	30.8604	21.7691
	(0.00)	(0.00)	(0.00)

The table lists score test statistics calculated when $c = 1.345$ (the situation considered in this work), $c = 0.7$ and $c = \infty$ (i.e. nonrobust situation) of Huber's function respectively. And the corresponding p-values are listing in the parenthesis. The small p-values of the test statistics suggest strongly rejecting the null hypothesis of $\gamma = 0$. The results indicates that there may exits nonhomogeneity of variance and exponential correlation coefficients at the same time for plasma drug penetration data. Moreover, we find from the results that because of the restriction of Huber's function, score test statistics is smaller than that in nonrobust situation. And the smaller c is, the smaller of the statistics.

§6. Conclusion

In this article, Fisher scoring method is applied to get M-estimation of parameters in exponential correlation nonlinear mixed models for longitudinal data. Then we propose several score test statistics for heterogeneity of exponential correlation coefficient and variance. We have used Huber's ρ function to bound the influence of outlying observations.

In fact, there are alternative choices such as Hampel function, Tukey's bisquare function and so on. Making some kind of comparison of theoretical or empirical efficiency of various choices would be worthwhile. In addition, when the measured time interval is equal, the exponential correlation model discussed in the paper is equivalent to AR(1) model according to [5], which is another correlation structure.

Appendix

Nonzero subblocks of Fisher information matrix $I(\boldsymbol{\theta}_1)$ are as follows

$$
\begin{aligned}
I_{\gamma\gamma} &= \Big(\frac{1}{2}\kappa_1 \sum_{i=1}^{m} \mathrm{tr}\Big(\widehat{\Sigma}_i^{-1}\frac{\partial V_i}{\partial \gamma_k}\widehat{\Sigma}_i^{-1}\frac{\partial V_i}{\partial \gamma_l}\Big)\Big)_{q\times q},\\
I_{\gamma\sigma^2} &= \Big(\frac{1}{4}\kappa_1(1+\kappa_2)\widehat{\sigma}^{-2} \sum_{i=1}^{m} \mathrm{tr}\Big(\widehat{\Sigma}_i^{-1}\frac{\partial V_i}{\partial \gamma_k}\Big)\Big)_{q\times 1},\\
I_{\gamma\phi} &= \Big(\frac{1}{2}\kappa_1 \sum_{i=1}^{m} \mathrm{tr}\Big(\widehat{\Sigma}_i^{-1}\frac{\partial V_i}{\partial \phi}\widehat{\Sigma}_k^{-1}\frac{\partial V_i}{\partial \gamma_k}\Big)\Big)_{q\times 1},\\
I_{\gamma\boldsymbol{\delta}} &= \Big(\frac{\kappa_1}{2} \sum_{i=1}^{m} \mathrm{tr}\Big(\widehat{\Sigma}_i^{-1}C_iE_{ab}^iC_i^{\mathsf{T}}\widehat{\Sigma}_i^{-1}\frac{\partial V_i}{\partial \gamma_k}\Big)\Big)_{q\times r'},\\
I_{\boldsymbol{\beta}\boldsymbol{\beta}} &= \nu\widehat{\sigma}^{-2}\sum_{i=1}^{m}\dot{f}_{ik}^{\mathsf{T}}\widehat{\Sigma}_i^{-1}\dot{f}_{il}, \qquad I_{\sigma^2\sigma^2} = \frac{1}{2}N\kappa_1\widehat{\sigma}^{-4},\\
I_{\phi\phi} &= \frac{1}{2}\kappa_1 \sum_{i=1}^{m} \mathrm{tr}\Big(\widehat{\Sigma}_i^{-1}\frac{\partial V_i}{\partial \phi}\widehat{\Sigma}_i^{-1}\frac{\partial V_i}{\partial \phi}\Big),\\
I_{\phi\sigma^2} &= \frac{1}{4}\kappa_1(1+\kappa_2)\widehat{\sigma}^{-2} \sum_{i=1}^{m} \mathrm{tr}\Big(\widehat{\Sigma}_i^{-1}\frac{\partial V_i}{\partial \phi}\Big),\\
I_{\phi\boldsymbol{\delta}} &= \Big(\frac{\kappa_1}{2} \sum_{i=1}^{m} \mathrm{tr}\Big(\widehat{\Sigma}_i^{-1}\frac{\partial V_i}{\partial \phi}\widehat{\Sigma}_i^{-1}C_iE_{ab}^iC_i^{\mathsf{T}}\Big)\Big)_{1\times r'},\\
I_{\sigma^2\boldsymbol{\delta}} &= \Big(\frac{1}{4}\kappa_1(1+\kappa_2)\widehat{\sigma}^{-2} \sum_{i=1}^{m} \mathrm{tr}\big(\widehat{\Sigma}_i^{-1}C_iE_{ab}^iC_i^{\mathsf{T}}\big)\Big)_{1\times r'},\\
I_{\boldsymbol{\delta}\boldsymbol{\delta}} &= \Big(\frac{1}{2} \sum_{i=1}^{m} \mathrm{tr}\big(\big(\widehat{\Sigma}_i^{-1}C_iE_{ab}^iC_i^{\mathsf{T}}\big)^2\big)\Big)_{r'\times r'},
\end{aligned}
$$

where E_{ab}^i is a $n_i\times n_i$ matrix with the elements 1 at (a,b) and (b,a) and 0 otherwise, $\kappa_2 = \int_{-c}^{c}(2\pi)^{-1/2}\varepsilon^2\mathrm{e}^{-1/2\varepsilon^2}\mathrm{d}\varepsilon$.

The nonzero subblocks in Fisher information matrix $I(\boldsymbol{\theta}_2)$ are given

$$
\begin{aligned}
I_{\gamma\gamma} &= \Big(\frac{1}{2}\kappa_1 \sum_{i=1}^{m} \dot{h}_{ik}\dot{h}_{il}\mathrm{tr}\big(\widehat{\Sigma}_i^{-1}C_i^{(1)}\widehat{\Gamma}_{11}C_i^{(1)\mathsf{T}}\big)^2\Big)_{q\times q},\\
I_{\gamma\phi} &= \Big(\frac{1}{2}\kappa_1 \sum_{i=1}^{m} \dot{h}_{ik}\mathrm{tr}\big(\widehat{\Sigma}_i^{-1}\dot{V}_{i0}\widehat{\Sigma}_k^{-1}C_i^{(1)}\widehat{\Gamma}_{11}C_i^{(1)\mathsf{T}}\big)\Big)_{q\times 1},\\
I_{\gamma\sigma^2} &= \Big(\frac{1}{4}\kappa_1(1+\kappa_2)\widehat{\sigma}^{-2} \sum_{i=1}^{m} \dot{h}_{ik}\mathrm{tr}\big(\widehat{\Sigma}_i^{-1}C_i^{(1)}\widehat{\Gamma}_{11}C_i^{(1)\mathsf{T}}\big)\Big)_{q\times 1},
\end{aligned}
$$

$$I_{\gamma\delta} = \Big(\frac{1}{2}\kappa_1 \sum_{i=1}^m \dot{h}_{ik}\mathrm{tr}\big(\widehat{\Sigma}_i^{-1}C_iE_{ab}^iC_i^{\mathsf{T}}\widehat{\Sigma}_i^{-1}C_i^{(1)}\widehat{\Gamma}_{11}C_i^{(1)\mathsf{T}}\big)\Big)_{q\times r'},$$

$$I_{\beta\beta} = \nu\widehat{\sigma}^{-2}\sum_{i=1}^m \dot{f}_{ik}^{\mathsf{T}}\widehat{\Sigma}_i^{-1}\dot{f}_{il}, \qquad I_{\sigma^2\sigma^2} = \frac{1}{2}N\kappa_1\widehat{\sigma}^{-4},$$

$$I_{\phi\phi} = \frac{1}{2}\kappa_1\sum_{i=1}^m \mathrm{tr}\big(\widehat{\Sigma}_i^{-1}\dot{V}_{i0}\widehat{\Sigma}_k^{-1}\dot{V}_{i0}\big),$$

$$I_{\phi\sigma^2} = \frac{1}{4}\kappa_1(1+\kappa_2)\widehat{\sigma}^{-2}\sum_{i=1}^m \mathrm{tr}\big(\widehat{\Sigma}_i^{-1}\dot{V}_{i0}\big),$$

$$I_{\phi\delta} = \Big(\frac{\kappa_1}{2}\sum_{i=1}^m \mathrm{tr}\big(\widehat{\Sigma}_k^{-1}\dot{V}_{i0}\widehat{\Sigma}_i^{-1}C_iE_{ab}^iC_i^{\mathsf{T}}\big)\Big)_{1\times r'},$$

$$I_{\sigma^2\delta} = \Big(\frac{1}{4}\kappa_1(1+\kappa_2)\widehat{\sigma}^{-2}\sum_{i=1}^m \mathrm{tr}\big(\widehat{\Sigma}_i^{-1}C_iE_{ab}^iC_i^{\mathsf{T}}\big)\Big)_{1\times r'},$$

$$I_{\delta\delta} = \Big(\frac{1}{2}\sum_{i=1}^m \mathrm{tr}\big(\big(\widehat{\Sigma}_i^{-1}C_iE_{ab}^iC_i^{\mathsf{T}}\big)^2\big)\Big)_{r'\times r'},$$

where $\dot{h}_{ik} = \partial h_i/\partial\gamma_k$ and $\dot{h}_{il}$ has the same meaning.

References

[1] PINHEIRO J C, BATES D M. *Mixed-Effects Models in S and S-PLUS* [M]. New York: Springer, 2000.

[2] LIN J G, WEI B C. Testing for homogeneity of variance and autocorrelation coefficients in nonlinear models of longitudinal data with AR(1) errors [J]. *Acta Math Appl Sin*, 2004, **27(3)**: 466–480. (in Chinese)

[3] LIN J G. Asymptotic properties of the score test for varying dispersion in exponential family nonlinear models [J]. *Statistics*, 2009, **43(2)**: 105–119.

[4] SUN H H. Testing for correlation coefficients in uniform correlation longitudinal mixed effect linear models based on M-estimation [J]. *Chinese J Appl Probab Statist*, 2013, **29(5)**: 469–479. (in Chinese)

[5] DIGGLE P J, HEAGERTY P, LIANG K Y, et al. *Analysis of Longitudinal Data* [M]. Oxford, UK: Oxford University Press, 2002.

[6] MEZA C, OSORIO F, DE LA CRUZ R. Estimation in nonlinear mixed-effects models using heavy-tailed distributions [J]. *Stat Comput*, 2012, **22(1)**: 121–139.

[7] RUSSO C M, PAULA G A, AOKI R. Influence diagnostics in nonlinear mixed-effects elliptical models [J]. *Comput Statist Data Anal*, 2009, **53(12)**: 4143–4156.

[8] PINHEIRO J C, LIU C H, WU Y N. Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate t distribution [J]. *J Comput Graph Statist*, 2001, **10(2)**: 249–276.

[9] LIN T I, LEE J C. A robust approach to t linear mixed models applied to multiple sclerosis data [J]. *Stat Med*, 2006, **25(8)**: 1397–1412.

[10] STAUDENMAYER J, LAKE E E, WAND M P. Robustness for general design mixed models using the t-distribution [J]. *Stat Model*, 2009, **9(3)**: 235–255.

[11] HUBER P J. *Robust Statistics* [M]. New York: Wiley, 1981.

[12] HUGGINS R M. A robust approach to the analysis of repeated measures [J]. *Biometrics*, 1993, **49(3)**: 715–720.

[13] SINHA S K. Robust inference in generalized linear models for longitudinal data [J]. *Canad J Statist*, 2006, **34(2)**: 261–278.

[14] YEAP B Y, DAVIDIAN M. Robust two-stage estimation in hierarchical nonlinear models [J]. *Biometrics*, 2001, **57(1)**: 266–272.

[15] GILL P S. A robust mixed linear model analysis for longitudinal data [J]. *Stat Med*, 2000, **19(7)**: 975–987.

[16] VENABLES W N, RIPLEY B D. Robust Statistics [J]. *Stat Comput*, 1997, **52(279)**: 77–92.

[17] WILLIAMS J D, BIRCH J B, ABDEL-SALAM A G. Outlier robust nonlinear mixed model estimation [J]. *Stat Med*, 2015, **34(8)**: 1304–1316.

[18] DOMOWITZ I, WHITE H. Misspecified models with dependent observations [J]. *J Econometrics*, 1982, **20(1)**: 35–58.

[19] SINHA S K. Robust analysis of generalized linear mixed models [J]. *J Amer Statist Assoc*, 2004, **99(466)**: 451–460.

[20] NÚÑEZ-ANTÓN V, ZIMMERMAN D L. Modeling nonstationary longitudinal data [J]. *Biometrics*, 2000, **56(3)**: 699–705.

[21] COX D R, HINKLEY D V. *Theoretical Statistics* [M]. London: Chapman and Hall, 1974.

[22] ROTNITZKY A, JEWELL N P. Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster correlated data [J]. *Biometrika*, 1990, **77(3)**: 485–497.

[23] SUN H H, LIN J G. Testing for heteroscedasticity in mixed effect linear models based on M-estimation [J]. *Math Appl*, 2011, **24(4)**: 798–805.

[24] DAVIDIAN M, GILTINAN D M. *Nonlinear Models for Repeated Measurement Data* [M]. London: Chapman and Hall, 1995.

基于 M 估计的指数相关非线性混合效应模型的齐性检验

孙慧慧 林金官

(盐城师范学院数学与统计学院, 盐城, 224002) (南京审计大学理学院, 南京, 211815)

摘　要: 方差和相关系数的齐性是纵向数据分析中常用假设之一, 然而, 这些假设未必合适. 本文主要研究的是具有指数相关结构的纵向数据非线性混合效应模型, 首先将 Huber 函数引入模型的对数似然函数中, 利用 Fisher 得分迭代法得到模型参数的稳健估计 (M 估计), 然后基于 M 估计对模型的方差和相关系数的齐性进行了 Score 检验, 并给出了检验统计量的 Monte-Carlo 模拟结果. 最后用一个实例说明了本文的方法.

关键词: 非线性混合模型; 指数相关; M 估计; 假设检验

中图分类号: O212.2