Contents

List of Tables

List of Figures

Estimation and prediction in linear mixed-effects models with multivariate t errors

Student: Shu-Chen Chou   Advisor: Dr. Jack C. Lee
Dr. Tsung I. Lin

Institute of Statistics
National Chiao Tung University
Hsinchu, Taiwan

Abstract

In this thesis, we proposed a robust extension of linear mixed-effects models basing on the multivariate t distributions. In particular, the within-subject correlation is assumed to have a parsimonious AR(1) dependence structure which is very important in practice. Meanwhile, we offer a general approach of how to obtain the score test statistic for testing the existence of autocorrelation among the within-subject errors. We use the scoring method for estimation and present the approximation of t restricted maximum likelihood for estimating variance components. The technique of predicting the future response vector of a subject given past measurements is also investigated. Numerical results are illustrated with real data from a multiple sclerosis clinical trial.

1. INTRODUCTION

Multiple sclerosis (MS) is a chronic disease of the central nervous system, especially young adults easily fall victim. Genetic and environmental factors are known to contribute to MS, but its precise cause is still not identified. Abundant researches consider MS as an autoimmune disease in which the immune system attacks its own myelin, causing disruptions to the never transmissions. The presence of areas of demyelination and T-cell predominant perivascular inflammation in the brain white matter are the main pathology of MS and some axons may be spared. There are no drugs to cure MS, but some treatments are available to ease the symptoms. For example, Interferon beta-1b (IFNB) was approved by the US Food and Drug Administration in mid 1993 for use in early stage relapsing-remitting MS patients and can be used to reduce the frequency and severity of relapses. Cranial magnetic resonance imaging (MRI) is a technique to assess the response of MS patients and provides another descriptor of its natural history. The MRI data is used to be a quantitative outcome for the MS patients in clinical trials. We use the patient's burden of disease, the total area of MS lesions of the MRI scan, as an indicator.

In this thesis, we analyze data from a randomized study which included 372 patients with relapsing-remitting MS. It began in June 1988 and ended in May 1990. All of them are in ten different medical centers in the United States and one in Canada. It was a place-controlled trial of interferon beta-1b (IFNB) and randomized to either a placebo (PL), a low-dose (LD), or a high-dose (HD). Low-dose and high-dose mean a dose of 1.6 and 8 million international units (MIU) of interferon beta-1b(IFNB) every other day, respectively. We consider a sub-study clinical trial of 52 patients, who are administrated at the University of British Columbia (UBC).

In many longitudinal studies, the model must include between-subject variabilities. Random-effects models are the most widely used general approach (Laird and Ware, 1982). However, the distribution assumption of error terms are usually focused on independently multivariate normal. Since longitudinal data may have some variability within subjects, the within-subject errors should assume follow some specific time series models such as AR(1). Chi and Reinsel (1989) proposed a linear mixed model that contain both between-subject random effects and AR(1) dependence for the within-subject errors. Gill (2000) proposed a robust estimation procedure to bound the influence of outlying observations by using Huber ρ function. In this thesis, we shall employ an alternative robust approach to analyzing the MS data using the t linear mixed model, which was first considered by Welsh and Richardson (1997). Subsequently, Pinherio *et al.* (2001) provide some efficient EM-type algorithms for maximum likelihood (ML) estimation based on a hierarchical complete-data formulation.

In Section 2, we introduce the t linear mixed-effects models with AR(1) and employ the score test approach for testing whether autocorrelation exists. Section 3 describes the computational aspects for both maximum likelihood (ML) estimation and restricted maximum likelihood (REML) estimation in the t linear mixed model. Prediction of future observations are discussed in Section 4. In Section 5, we illustrate the proposed methodologies using the MS data. Finally, some discussions are given in Section 6.

2. The MODEL AND SCORE TEST FOR AUTOCORRELATION

2.1. t linear mixed model

Suppose that response measurements are collected on each of N subjects with the ith subject being observed on p_i time points. The model for subject i is

$$\begin{aligned} &\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{\varepsilon}_i, \quad \boldsymbol{b}_i \mid \tau_i \sim N_{m_2}\left(\boldsymbol{0}, \frac{\sigma^2}{\tau_i}\boldsymbol{\Gamma}\right) \\ &\boldsymbol{\varepsilon}_i \mid \tau_i \sim N_{p_i}\left(\boldsymbol{0}, \frac{\sigma^2}{\tau_i}\boldsymbol{C}_i\right), \quad \tau_i \sim \chi^2_\nu/\nu, \quad \boldsymbol{b}_i \mid \tau_i \perp \boldsymbol{\varepsilon}_i \mid \tau_i \end{aligned} \tag{1}$$

where $\boldsymbol{Y}_i = (Y_{i1}, Y_{i2}, \dots, Y_{ip_i})'$ is a vector of length p_i observable responses for subject i, $\boldsymbol{\beta}$ is an $m_1 \times 1$ vector of unknown but fixed parameters with full-rank design matrix $\boldsymbol{X}_i$ of dimension $p_i \times m_1$, $\boldsymbol{b}_i$ is an $m_2 \times 1$ vector of unobservable random effects with design matrix $\boldsymbol{Z}_i$ of dimension $p_i \times m_2$, $\boldsymbol{\varepsilon}_i$ is $p_i \times 1$ vector of residual errors and τ_i is an unknown scale assumed to be distributed as gamma with mean 1 and variance $2/\nu$. Furthermore, $\mathbf{\Gamma}$ and $\boldsymbol{C}_i$ are respectively $m_2 \times m_2$ and $p_i \times p_i$ scale matrices.

If $\boldsymbol{Y} \mid \tau \sim N_p(\boldsymbol{\mu}, \mathbf{\Sigma}/\tau)$ and $\tau \sim \chi^2_\nu/\nu$, we can conclude that $\boldsymbol{Y} \sim t_p(\boldsymbol{\mu}, \mathbf{\Sigma}, \nu)$. Here $t_p(\boldsymbol{\mu}, \mathbf{\Sigma}, \nu)$ denotes the p-variate t distribution with mean vector $\boldsymbol{\mu}$, scale matrix $\mathbf{\Sigma}$ and degree-of-freedom ν. A hierarchical version of the t linear mixed-effects model (1) can be expressed by:

$$\boldsymbol{Y}_i \mid \boldsymbol{b}_i, \tau_i \overset{ind}{\sim} N_{p_i}(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i, \frac{\sigma^2}{\tau_i}\boldsymbol{C}_i),$$
$$\boldsymbol{b}_i \mid \tau_i \overset{ind}{\sim} N_{m_2}(\mathbf{0}, \frac{\sigma^2}{\tau_i}\mathbf{\Gamma}), \qquad \tau_i \overset{iid}{\sim} \chi^2_\nu/\nu. \tag{2}$$

In longitudinal data, repeated measurements of some variable are made on a number of subjects over a period of time. Therefore, the model should incorporate autocorrelation for within-subject variabilities. Assume that for each subject p_i observations are taken at time $t_i = (t_{i1}, t_{i2}, \ldots, t_{ip_i})'$ and are equally spaced. The dimension of the scale correlation matrix for $\boldsymbol{\varepsilon}_i$ is $p_i \times p_i$. By parsimony, we concentrate on $\boldsymbol{C}_i$ having a simple AR(1) structure, that is,

$$\boldsymbol{C}_i(\rho) = [\rho^{|r-s|}], \qquad -1 < \rho < 1 \tag{3}$$

where $r, s = 1, \ldots, p_i$.

2.2. The Score test for the autocorrelation

For many longitudinal data, we don't know whether the autocorrelation in the within-individual errors exist or not. Therefore, a score test is used to check the presence of autocorrelation in the errors. The score vector is defined as

$$S(\boldsymbol{\theta}) = \frac{\partial}{\partial\boldsymbol{\theta}} \log f(\boldsymbol{Y} \mid \boldsymbol{\theta}) = \frac{\partial}{\partial\boldsymbol{\theta}} \log L(\boldsymbol{\theta} \mid \boldsymbol{Y}) \tag{4}$$

where $L(\boldsymbol{\theta}|\boldsymbol{Y})$ is likelihood function of $\boldsymbol{\theta}$, $\boldsymbol{\theta}$ is the vector of parameters in the model, and $\boldsymbol{Y}' = (\boldsymbol{Y}'_1, \boldsymbol{Y}'_2, \ldots, \boldsymbol{Y}'_N)'$. Under regularity conditions, we know that, for all $\boldsymbol{\theta}$, $E_{\boldsymbol{\theta}} S(\boldsymbol{\theta}) = \mathbf{0}$. In particular, if we are testing $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \in \mathbf{\Theta}_0$ against $H_1 : \boldsymbol{\theta} \in \mathbf{\Theta}_1$ and if H_0 is true, then $S(\boldsymbol{\theta}_0)$ has mean vector $\mathbf{0}$. Furthermore, the covariance of

$S(\boldsymbol{\theta})$ is

$$\mathrm{Cov}\left(S(\boldsymbol{\theta})\right) = -\mathrm{E}_{\boldsymbol{\theta}}\left(\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\log \mathrm{L}(\boldsymbol{\theta} \mid \boldsymbol{Y})\right) = \mathbf{J}(\boldsymbol{\theta}), \tag{5}$$

where $\mathbf{J}(\boldsymbol{\theta})$ is the Fisher information matrix. We denote the score test statistic

$$\lambda_s = S(\boldsymbol{\theta}_0)'\mathbf{J}(\boldsymbol{\theta}_0)^{-1}S(\boldsymbol{\theta}_0). \tag{6}$$

By asymptotic efficiency of MLEs, λ_s is asymptotically distributed as chi-squared with $\dim(\boldsymbol{\Theta}_1)$-$\dim(\boldsymbol{\Theta}_0)$ degrees-of-freedom under true H_0.

Our purpose is to test the presence of possible autocorrelation by using the score test. It means to test the null model $H_0 : \rho = 0$ in (3). If we reject the null model, the autocorrelation in the errors may exist.

The following is the marginal distribution of $\boldsymbol{Y}_i$ from model (2):

$$\begin{aligned} f(\boldsymbol{Y}_i) &= \int\int f(\boldsymbol{Y}_i \mid \boldsymbol{b}_i, \tau_i) f(\boldsymbol{Y}_i \mid \tau_i) f(\tau_i) d\boldsymbol{b}_i d\tau_i \\ &= \frac{\Gamma(\frac{\nu+p_i}{2})|\boldsymbol{\Lambda}_i|^{-1/2}}{\Gamma(\frac{\nu}{2})(\pi\nu\sigma^2)^{p_i/2}}\left(1 + \frac{(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta})'\boldsymbol{\Lambda}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta})}{\nu\sigma^2}\right)^{-(\nu+p_i)/2} \end{aligned} \tag{7}$$

where

$$\boldsymbol{\Lambda}_i = \boldsymbol{\Lambda}_i(\boldsymbol{\Gamma}, \rho) = \boldsymbol{Z}_i\boldsymbol{\Gamma}\boldsymbol{Z}_i' + \boldsymbol{C}_i(\rho).$$

Thus, $\boldsymbol{Y}_i$ is distributed as $t_{p_i}(\boldsymbol{X}_i\boldsymbol{\beta}, \sigma^2\boldsymbol{\Lambda}_i, \nu)$. We denote $\boldsymbol{e}_i = \boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}$ and $\boldsymbol{\Delta}_i(\boldsymbol{\beta}, \boldsymbol{\Gamma}, \rho) = \boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i$. We parameterize $\boldsymbol{\Gamma} = \mathbf{F}'\mathbf{F}$ by Cholesky decomposition, where $\mathbf{F}$ is an upper triangular matrix, for guaranteeing $\boldsymbol{\Gamma}$ to be positive definite. Let $\boldsymbol{\alpha} = (\boldsymbol{\beta}', \boldsymbol{\theta}', \rho)'$, with $\boldsymbol{\theta} = (\sigma^2, \boldsymbol{f}', \nu)'$, and $\boldsymbol{f} = (f_1, f_2, \ldots, f_k)'$, which is the vector of the distinct $k = m_2(m_2 + 1)/2$ components in $\mathbf{F}$. Let $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ be the ML estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ under the null model. Thus, we denote $\hat{\boldsymbol{\alpha}}_0 = (\hat{\boldsymbol{\beta}}', \hat{\boldsymbol{\theta}}', 0)'$. Let l denote the log-likelihood function of $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_N$. We have

$$\begin{aligned} l &= \text{const.} + \sum_{i=1}^{N}\left\{\log\left(\Gamma\Big(\frac{\nu+p_i}{2}\Big)\right) - \log\left(\Gamma\Big(\frac{\nu}{2}\Big)\right)\right\} - \frac{n}{2}\log(\nu\sigma^2) \\ &\quad -\frac{1}{2}\sum_{i=1}^{N}\log\mid\boldsymbol{\Lambda}_i\mid -\frac{1}{2}\sum_{i=1}^{N}(\nu+p_i)\log\left(1 + \frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta}, \boldsymbol{\Gamma}, \rho)}{\sigma^2\nu}\right), \end{aligned} \tag{8}$$

where $n = \sum_{i=1}^{N} p_i$ is the total number of observations from all the subjects.

The score vector $\partial l/\partial \boldsymbol{\alpha}$ has all components equal 0 except the derivative with respect to ρ, denoted by $(\partial l/\partial \rho)_0 = \partial l/\partial \rho|_{\hat{\boldsymbol{\alpha}}_0}$, when evaluated at $\hat{\boldsymbol{\alpha}}_0$. The information matrix is

$$\mathbf{J}_{\alpha\alpha} = \mathrm{E}(-\partial^2 l/\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}') = \begin{pmatrix} \mathbf{J}_{\beta\beta} & \mathbf{J}_{\beta\theta} & \mathbf{J}_{\beta\nu} \\ \mathbf{J}'_{\beta\theta} & \mathbf{J}_{\theta\theta} & \mathbf{J}_{\theta\rho} \\ \mathbf{J}'_{\beta\rho} & \mathbf{J}'_{\theta\rho} & \mathbf{J}_{\rho\rho} \end{pmatrix} = \begin{pmatrix} \mathbf{J}_{\beta\beta} & \mathbf{0} & \mathbf{0} \\ \mathbf{0}' & \mathbf{J}_{\theta\theta} & \mathbf{J}_{\theta\rho} \\ \mathbf{0}' & \mathbf{J}'_{\theta\rho} & \mathbf{J}_{\rho\rho} \end{pmatrix}, \quad (9)$$

where $\mathbf{J}_{\theta\theta} = \begin{pmatrix} \mathbf{J}_{\sigma^2\sigma^2} & \mathbf{J}_{\sigma^2 f} & \mathbf{J}_{\sigma^2\nu} \\ \mathbf{J}'_{\sigma^2 f} & \mathbf{J}_{ff} & \mathbf{J}_{f\nu} \\ \mathbf{J}'_{\sigma^2\nu} & \mathbf{J}'_{f\nu} & \mathbf{J}_{\nu\nu} \end{pmatrix}$ and $\mathbf{J}_{\theta\rho} = \begin{pmatrix} \mathbf{J}_{\sigma^2\rho} \\ \mathbf{J}_{f\rho} \\ \mathbf{J}_{\nu\rho} \end{pmatrix}$.

The score test statistic, denoted by λ_s, then

$$\lambda_s = \left(\frac{\partial l}{\partial \boldsymbol{\alpha}} \bigg|_{\hat{\boldsymbol{\alpha}}_0} \right)' (\mathbf{J}|_{\hat{\boldsymbol{\alpha}}_0})^{-1} \left(\frac{\partial l}{\partial \boldsymbol{\alpha}} \bigg|_{\hat{\boldsymbol{\alpha}}_0} \right) = \left(\frac{\partial l}{\partial \rho} \right)_0^2 / s, \quad (10)$$

where $s = \mathbf{J}^0_{\rho\rho} - \mathbf{J}^0_{\theta\rho}{}' \mathbf{J}^0_{\theta\theta}{}^{-1} \mathbf{J}^0_{\theta\rho}$ and the superscript 0 indicates the elements of $\mathbf{J}$ are evaluated at $\hat{\boldsymbol{\alpha}}_0$. The evaluation of s is detailed in Appendix **A**.

To calculate λ_s, we first have

$$\frac{\partial l}{\partial \rho} = -\frac{1}{2} \sum_{i=1}^{N} \mathrm{tr} \left(\boldsymbol{\Lambda}_i^{-1} \frac{\partial \boldsymbol{\Lambda}_i}{\partial \rho} \right) + \frac{1}{2} \sum_{i=1}^{N} (\nu + p_i) \frac{\boldsymbol{e}_i' \boldsymbol{\Lambda}_i^{-1} \frac{\partial \boldsymbol{\Lambda}_i}{\partial \rho} \boldsymbol{\Lambda}_i^{-1} \boldsymbol{e}_i}{\sigma^2 \nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta}, \boldsymbol{\Gamma}, \rho)}. \quad (11)$$

Note that by matrix inversion formula (See Rao (1973), page 33),

$$\boldsymbol{\Lambda}_i^{-1} (\boldsymbol{Y}_i - \boldsymbol{X}_i \boldsymbol{\beta}) = \boldsymbol{C}_i^{-1} (\boldsymbol{Y}_i - \boldsymbol{X}_i \boldsymbol{\beta} - \boldsymbol{Z}_i \tilde{\boldsymbol{b}}_i), \quad (12)$$

where $\tilde{\boldsymbol{b}}_i = (\boldsymbol{Z}_i{}' \boldsymbol{C}_i^{-1} \boldsymbol{Z}_i + \boldsymbol{\Gamma}^{-1})^{-1} \boldsymbol{Z}_i{}' \boldsymbol{C}_i^{-1} (\boldsymbol{Y}_i - \boldsymbol{X}_i \boldsymbol{\beta})$.

When evaluated at $\hat{\boldsymbol{\alpha}}_0$, we obtain $\boldsymbol{\Lambda}_i^{-1} (\boldsymbol{Y}_i - \boldsymbol{X}_i \boldsymbol{\beta})|_{\hat{\boldsymbol{\alpha}}_0} = \boldsymbol{Y}_i - \boldsymbol{X}_i \hat{\boldsymbol{\beta}} - \boldsymbol{Z}_i \hat{\boldsymbol{b}}_i = \hat{\boldsymbol{u}}_i$, where $\hat{\boldsymbol{b}}_i = (\boldsymbol{Z}_i{}' \boldsymbol{Z}_i + \hat{\boldsymbol{\Gamma}}^{-1})^{-1} \boldsymbol{Z}_i{}' (\boldsymbol{Y}_i - \boldsymbol{X}_i \hat{\boldsymbol{\beta}})$ and $\hat{\boldsymbol{u}}_i$ can be viewed as the vector of residuals from the model for individual i. In addition, $\boldsymbol{C}_i^{-1}|_{\hat{\boldsymbol{\alpha}}_0} = \boldsymbol{I}_{p_i \times p_i}$ and $\partial \boldsymbol{\Lambda}_i / \partial \rho|_{\hat{\boldsymbol{\alpha}}_0} = \partial \boldsymbol{C}_i / \partial \boldsymbol{\alpha}|_{\hat{\boldsymbol{\alpha}}_0} = \boldsymbol{L}_i' + \boldsymbol{L}_i$, where $\boldsymbol{L}_i'$ is a $p_i \times p_i$ matrix with the elements on the first super-diagonal being 1. By a matrix inversion formula, we can obtain $\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1} \frac{\partial \boldsymbol{\Lambda}_i}{\partial \rho})|_{\hat{\boldsymbol{\alpha}}_0} = -2\mathrm{tr}[(\boldsymbol{Z}_i{}' \boldsymbol{Z}_i + \hat{\boldsymbol{\Gamma}}^{-1})^{-1} \boldsymbol{Z}_i{}' \boldsymbol{L}_i \boldsymbol{Z}_i]$,

$$\left(\frac{\partial l}{\partial \rho} \right)_0 = \sum_{i=1}^{N} \mathrm{tr}[(\boldsymbol{Z}_i{}' \boldsymbol{Z}_i + \hat{\boldsymbol{\Gamma}}^{-1})^{-1} \boldsymbol{Z}_i{}' \boldsymbol{L}_i \boldsymbol{Z}_i] + \sum_{i=1}^{N} (\hat{\nu} + p_i) \frac{\hat{\boldsymbol{u}}_i' \boldsymbol{L}_i \hat{\boldsymbol{u}}_i}{\hat{\sigma^2} \hat{\nu} + \boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Gamma}}, 0)}, \quad (13)$$

where $\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Gamma}}, 0) = (\boldsymbol{Y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}})'\hat{\boldsymbol{u}}_i$. The score statistic λ_s can be calculated, from (10) and (13), as

$$\lambda_s = \left(\sum_{i=1}^{N} \mathrm{tr}[(\boldsymbol{Z}_i{}'\boldsymbol{Z}_i + \hat{\boldsymbol{\Gamma}}^{-1})^{-1}\boldsymbol{Z}_i{}'\boldsymbol{L}_i\boldsymbol{Z}_i] + \sum_{i=1}^{N}(\hat{\nu} + p_i)\frac{\hat{\boldsymbol{u}}_i'\boldsymbol{L}_i\hat{\boldsymbol{u}}_i}{\hat{\sigma}^2\hat{\nu} + \boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Gamma}}, 0)} \right)^2 / s. \quad (14)$$

3. PARAMETER ESTIMATION

3.1. Maximum likelihood estimation

In this subsection, we employ the scoring method for parameter estimation. Therefore, the score vector and the Fisher information matrix are needed. We consider the log-likelihood function of $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_N$ as given by (8). Let $\boldsymbol{\alpha} = (\boldsymbol{\beta}', \boldsymbol{\theta}', \rho)' = (\boldsymbol{\beta}', \boldsymbol{\theta}^{*\prime})'$ then the Fisher information matrix (9) can be partitioned as

$$\mathbf{J}_{\alpha\alpha} = \mathrm{E}(-\partial^2 l/\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha}') = \begin{pmatrix} \mathbf{J}_{\beta\beta} & \mathbf{0} & \mathbf{0} \\ \mathbf{0}' & \mathbf{J}_{\theta\theta} & \mathbf{J}_{\theta\rho} \\ \mathbf{0}' & \mathbf{J}_{\theta\rho}' & \mathbf{J}_{\rho\rho} \end{pmatrix} = \begin{pmatrix} \mathbf{J}_{\beta\beta} & \mathbf{0} \\ \mathbf{0}' & \mathbf{J}_{\theta^*\theta^*} \end{pmatrix}. \quad (15)$$

In order to estimate $\boldsymbol{\alpha}$, we start with any suitable initial value of $\boldsymbol{\theta}^*$ and $\boldsymbol{\beta}$ then we can conclude the new value of $\boldsymbol{\theta}^*$ and $\boldsymbol{\beta}$. The new ones are used to start the next iteration and continue the iteration process until the absolute difference of the new and previous estimators is less than some fixed tolerance. The score vector $\mathbf{s} = \partial l/\partial\boldsymbol{\alpha}$ and the expectation of the negative Hessian matrix are listed in the Appendix A. With the current estimates $\hat{\boldsymbol{\theta}}^{*(h)}$ and $\hat{\boldsymbol{\beta}}^{(h)}$ at the hth iteration step, the scoring procedure for estimation is obtained by the following recursive equation

$$\hat{\boldsymbol{\theta}}^{*(h+1)} = \hat{\boldsymbol{\theta}}^{*(h)} + \hat{\mathbf{J}}_{\theta^*\theta^*}^{(h)^{-1}} \hat{\mathbf{s}}_{\theta^*}^{(h)} \quad (16)$$

where $\hat{\mathbf{s}}_{\theta^*}^{(h)}$ and $\hat{\mathbf{J}}_{\theta^*\theta^*}^{(h)}$ denote $\mathbf{s}$ and $\mathbf{J}_{\theta^*\theta^*}$ evaluated at $\hat{\boldsymbol{\beta}}^{(h)}$ and $\hat{\boldsymbol{\theta}}^{*(h)}$ at the hth iteration step. Then the next estimate $\hat{\boldsymbol{\beta}}^{(h+1)}$ is obtained as

$$\hat{\boldsymbol{\beta}}^{(h+1)} = \hat{\boldsymbol{\beta}}^{(h)} + \hat{\mathbf{J}}_{\beta\beta}^{(h+1)^{-1}} \hat{\mathbf{s}}_{\beta}^{(h+0.5)} \quad (17)$$

where $\hat{\mathbf{J}}_{\beta\beta}^{(h+1)}$ denotes $\mathbf{J}_{\beta\beta}$ evaluated at $\hat{\boldsymbol{\theta}}^{*(h+1)}$ and $\hat{\mathbf{s}}_{\beta}^{(h+0.5)}$ denotes $\mathbf{s}_\beta$ evaluated at $\hat{\boldsymbol{\beta}}^{(h)}$ and evaluated at $\hat{\boldsymbol{\theta}}^{*(h+1)}$.

For the rate of convergence, a preferred method is given suitable initial values such as the ML or REML estimates of a normal linear mixed model with appropriate dependence structure. It is convenient for users to fit normal linear mixed models using commercial softwares such as SAS procedure PROC MIXED or S-PLUS function lme().

Under some regularity conditions, variance-covariance estimates can be computed by plugging the converged ML estimates $\hat{\boldsymbol{\alpha}} = (\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}}^*)$ into the inverse of the Fisher information matrix. Thus, we have

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}) = \hat{\sigma}^2 \left(\sum_{i=1}^{N} \frac{\hat{\nu} + p_i}{\hat{\nu} + p_i + 2} \boldsymbol{X}_i' \hat{\boldsymbol{\Lambda}}_i^{-1} \boldsymbol{X}_i \right)^{-1} \quad \text{and} \quad \mathrm{Cov}(\hat{\boldsymbol{\theta}}^*) = \hat{\mathbf{J}}_{\theta^*\theta^*}^{-1}. \tag{18}$$

The standard errors of the elements of $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ can be estimated by taking the square root of the corresponding diagonal elements in (18). Moreover, an approximate $100(1-\alpha)\%$ confidence ellipsoid for $\boldsymbol{\beta}$ is given by

$$(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})' \left(\frac{1}{\hat{\sigma}^2} \sum_{i=1}^{N} \frac{\hat{\nu} + p_i}{\hat{\nu} + p_i + 2} \boldsymbol{X}_i' \hat{\boldsymbol{\Lambda}}_i^{-1} \boldsymbol{X}_i \right) (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) \le \chi^2_{m_1}(\alpha),$$

where $\chi^2_{m_1}(\alpha)$ denotes the upper 100(1-α)% quantile of the chi-squared distribution with m_1 d.f..

3.2. Restricted maximum likelihood estimation

Restricted maximum likelihood estimation (REML) is a method of estimating variance components. The REML estimates are obtained by maximizing the likelihood of $\boldsymbol{\theta}$ based on some error contrasts. Harville (1974) showed that error contrasts to make Bayesian inferences for variance components were equivalent to ignoring the prior information on the fixed effects. The estimators of variance components obtained from maximum likelihood are usually biased downwards. Therefore, REML is often preferred to maximize likelihood estimation (ML) because it takes account of the loss of degrees of freedom in estimating the mean and produces unbiased estimating equations for the variance components. In this subsection, we discuss the Bayesian approach to integrate with respect to the fixed effects $\boldsymbol{\beta}$ for obtaining the marginal density function given the observation vector $\boldsymbol{Y}$. We have $\boldsymbol{Y}_i \sim t_{p_i}(\boldsymbol{X}_i\boldsymbol{\beta}, \sigma^2\boldsymbol{\Lambda}_i, \nu)$ and let $\boldsymbol{\theta}_R = (\sigma^2, \boldsymbol{f}', \nu, \rho)$ to obtain

$$\begin{aligned} L_R(\boldsymbol{\theta}_R|\boldsymbol{Y}) &= \int L(\boldsymbol{\beta}, \boldsymbol{\theta}_R|\boldsymbol{Y})d\boldsymbol{\beta} \\ &= (\sigma^2)^{-\frac{n}{2}} \prod_{i=1}^{N} \frac{\boldsymbol{\Gamma}((\nu+p_i)/2)}{\boldsymbol{\Gamma}(\nu/2)\sqrt{\pi\nu}} |\boldsymbol{\Lambda}_i|^{-\frac{1}{2}} \int \prod_{i}^{N} \left(1 + \frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta}, \boldsymbol{\Gamma}, \rho)}{\nu\sigma^2}\right)^{-(\nu+p_i)/2} d\boldsymbol{\beta}. \end{aligned}$$

Since it is not easy to evaluate, we use Laplace's method to get an approximation. Let $f(\boldsymbol{\beta}) = \prod_{i=1}^{N}\left(1 + \frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta}, \boldsymbol{\Gamma}, \rho)}{\nu\sigma^2}\right)^{-(\nu+p_i)/2}$ and $\log(\mathrm{f}(\boldsymbol{\beta})) = \mathrm{u}(\boldsymbol{\beta})$. The following is the approximation by Laplace's method,

$$\begin{aligned} \int f(\boldsymbol{\beta})d\boldsymbol{\beta} &= \int \exp\left(u(\boldsymbol{\beta})\right) d\boldsymbol{\beta} \\ &\approx \int \exp\left\{u(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})) + \frac{1}{2}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}(\boldsymbol{\theta}))' u''(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}))(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}(\boldsymbol{\theta}))\right\} d\boldsymbol{\beta} \\ &= (2\pi)^{-\frac{m_1}{2}} f(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}))|-u''(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}))|^{-\frac{1}{2}}, \end{aligned}$$

where

$$\begin{aligned} u''(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})) &= \mathbf{H}_{\beta\beta}|_{\beta=\hat{\beta}(\boldsymbol{\theta})} \\ &= \sum_{i=1}^{N}(\nu+p_i)\left\{\frac{-\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{X}_i}{\sigma^2\nu + \boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}), \boldsymbol{\Gamma}, \rho)} + \frac{2\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\hat{\boldsymbol{e}}_i(\boldsymbol{\theta})\hat{\boldsymbol{e}}_i{}'(\boldsymbol{\theta})\boldsymbol{\Lambda}_i^{-1}\boldsymbol{X}_i}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}), \boldsymbol{\Gamma}, \rho)]^2}\right\} \end{aligned}$$

and $\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})$ can be obtained by solving the following estimating equation

$$\sum_{i=1}^{N}(\nu+p_i)\frac{\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta})}{\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta}, \boldsymbol{\Gamma}, \rho)} = 0, \qquad (19)$$

with $\hat{\boldsymbol{e}}_i(\boldsymbol{\theta}) = \boldsymbol{Y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})$.

Hence, the approximation form of t restricted maximum likelihood is

$$\begin{aligned} &L_R(\boldsymbol{\theta}_R|\boldsymbol{Y}) \\ &\approx (\sigma^2)^{-\frac{n}{2}}|\sum_{i=1}^{N}\boldsymbol{X}_i'\mathbf{H}_i\boldsymbol{X}_i|^{-\frac{1}{2}} \prod_{i=1}^{N}\frac{\boldsymbol{\Gamma}((\nu+p_i)/2)}{\boldsymbol{\Gamma}(\nu/2)\sqrt{\pi\nu}}|\boldsymbol{\Lambda}_i|^{-\frac{1}{2}}\left(1 + \frac{\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}), \boldsymbol{\Gamma}, \rho)}{\nu\sigma^2}\right)^{-(\nu+p_i)/2}, \end{aligned}$$

where

$$\mathbf{H}_i = (\nu + p_i)\left\{\frac{\mathbf{\Lambda}_i^{-1}}{\sigma^2\nu + \mathbf{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}), \mathbf{\Gamma}, \rho)} - \frac{2\mathbf{\Lambda}_i^{-1}\hat{\boldsymbol{e}}_i(\boldsymbol{\theta})\hat{\boldsymbol{e}}_i{}'(\boldsymbol{\theta})\mathbf{\Lambda}_i^{-1}}{[\sigma^2\nu + \mathbf{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}), \mathbf{\Gamma}, \rho)]^2}\right\}, i = 1, \ldots, N$$

and $\hat{\boldsymbol{e}}_i(\boldsymbol{\theta}) = \boldsymbol{Y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})$.

Since the negative expectation of Hessian matrix of $L_R(\boldsymbol{\theta}_R|\boldsymbol{Y})$ is difficult to compute, we employ the Newton-Raphson (NR) algorithm for estimating. The REML estimate of $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}_R$, can be obtained by implementing the Newton-Raphson (NR) algorithm with ML estimates as the initial values. In the NR algorithm, the estimate of $\boldsymbol{\beta}$, $\hat{\boldsymbol{\beta}}_R(\hat{\boldsymbol{\theta}}_R)$, must be computed at each iteration by solving the estimating equation in (19).

The first derivatives of the restricted log-likelihood $L_R(\boldsymbol{\theta}_R|\boldsymbol{Y})$ are listed in Appendix **B**. The second derivatives of $L_R(\boldsymbol{\theta}_R|\boldsymbol{Y})$ are complicated, we can use the numerical approximation form for each component of the Hessian matrix by

$$\begin{aligned}\frac{\partial^2 \log L_R}{\partial\theta_i\partial\theta_j} \approx\ & \Big[\log L_R(\boldsymbol{\theta} + \delta_i\boldsymbol{e}_i + \delta_j\boldsymbol{e}_j) - \log L_R(\boldsymbol{\theta} + \delta_i\boldsymbol{e}_i - \delta_j\boldsymbol{e}_j) \\ & - \log L_R(\boldsymbol{\theta} - \delta_i\boldsymbol{e}_i + \delta_j\boldsymbol{e}_j) + \log L_R(\boldsymbol{\theta} - \delta_i\boldsymbol{e}_i - \delta_j\boldsymbol{e}_j)\Big]/(4\delta_i\delta_j),\end{aligned}$$

where δ_i and δ_j are small values that are chosen based on the scale of the problem, and $\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ are respectively the unit vector corresponding to the ith and jth components of $\boldsymbol{\theta}$.

4. EMPIRICAL BAYES ESTIMATION FOR RANDOM EFFECTS AND PREDICTION OF FUTURE VALUES

The estimation of the random effect $\boldsymbol{b}_i$ is also of interest, and we will use empirical Bayes to estimate it. Laird and Ware (1982) addressed the empirical Bayes of $\boldsymbol{b}_i$, which is the expectation of the posterior density of $\boldsymbol{b}_i$ given $\boldsymbol{Y}_i$ with the parameters replaced by their estimates and defined as $\hat{\boldsymbol{b}}_i = \mathrm{E}(\boldsymbol{b}_i \mid \boldsymbol{Y}_i)|_{\alpha=\hat{\alpha}}$. For evaluating the conditional expectation of $\boldsymbol{b}_i$ given $\boldsymbol{Y}_i$, we need the joint distribution of $(\boldsymbol{Y}_i, \boldsymbol{b}_i)$.

By integrating τ_i of equation (2), we get the joint distribution of $(\boldsymbol{Y}_i, \boldsymbol{b}_i)$, that is a (p_i+m_2)-variate t distribution with degree of freedom ν, which can be expressed as

$$\begin{pmatrix} \boldsymbol{Y}_i \\ \boldsymbol{b}_i \end{pmatrix} \sim t_{p_i+m_2}\left(\begin{pmatrix} \boldsymbol{X}_i\boldsymbol{\beta} \\ \boldsymbol{0} \end{pmatrix}, \sigma^2 \begin{pmatrix} \boldsymbol{\Lambda}_i & \boldsymbol{Z}_i\boldsymbol{\Gamma} \\ \boldsymbol{\Gamma}\boldsymbol{Z}_i' & \boldsymbol{\Gamma} \end{pmatrix}, \nu\right). \tag{20}$$

According to the posterior density of $\boldsymbol{b}_i$ given $\boldsymbol{Y}_i$, the conditional mean of $\boldsymbol{b}_i$ given $\boldsymbol{Y}_i$ is the minimum mean squared error (MSE) estimator of $\boldsymbol{b}_i$, which is given by

$$\hat{\boldsymbol{b}}_i(\boldsymbol{\theta}) = \boldsymbol{b}_i^*(\boldsymbol{\theta}) - \boldsymbol{W}_i(\boldsymbol{W}_i + \boldsymbol{\Gamma})^{-1}\boldsymbol{b}_i^*(\boldsymbol{\theta}) \tag{21}$$

where $\boldsymbol{b}_i^*(\boldsymbol{\theta}) = (\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)^{-1}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta})$ and $\boldsymbol{W}_i = (\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)^{-1}$. The error covariance matrix of $\tilde{\boldsymbol{b}}_i$ is

$$\mathrm{E}[(\hat{\boldsymbol{b}}_i(\boldsymbol{\theta}) - \boldsymbol{b}_i)(\hat{\boldsymbol{b}}_i(\boldsymbol{\theta}) - \boldsymbol{b}_i)'] = \frac{\nu}{\nu-2}\sigma^2[\boldsymbol{W}_i - \boldsymbol{W}_i(\boldsymbol{W}_i + \boldsymbol{\Gamma})^{-1}\boldsymbol{W}_i] \tag{22}$$

The calculation of equations (21) and (22) are shown in Appendix **C**. The empirical Bayes estimates of random effects $\hat{\boldsymbol{b}}_i = \hat{\boldsymbol{b}}_i(\hat{\boldsymbol{\theta}})$ is obtained by plugging $\hat{\boldsymbol{\theta}}$ in (21).

The following is the prediction of future values of a subject based on t linear mixed-effects model (2). Chi and Reinsel (1989) also have discussed the prediction of future observations for normal mixed-effects model with AR(1) errors. We consider the observations $\boldsymbol{Y}_i$ for a subject of length p_i which are available to predict the future values $\boldsymbol{y}_i$ for the same subject, that is a q ($q > 0$) dimension of values of measurements.

The future values $\boldsymbol{y}_i$ can be expressed as $\boldsymbol{y}_i = \boldsymbol{x}_i\boldsymbol{\beta} + \boldsymbol{z}_i\boldsymbol{b}_i + \boldsymbol{\varepsilon}_{iq}$, where $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$ denote $q \times m_1$ and $q \times m_2$ design matrices, respectively. Let the vector $\boldsymbol{Y}_i^* = (\boldsymbol{Y}_i, \boldsymbol{y}_i)'$ be represented as $\boldsymbol{Y}_i^* = \boldsymbol{X}_i^*\boldsymbol{\beta} + \boldsymbol{Z}_i^*\boldsymbol{b}_i + \boldsymbol{\varepsilon}_i^*$, where

$$\boldsymbol{X}_i^* = \begin{pmatrix} \boldsymbol{X}_i \\ \boldsymbol{x}_i \end{pmatrix}_{(p_i+q)\times m_1}, \quad \boldsymbol{Z}_i^* = \begin{pmatrix} \boldsymbol{Z}_i \\ \boldsymbol{z}_i \end{pmatrix}_{(p_i+q)\times m_2} \quad \boldsymbol{\varepsilon}_i^* = \begin{pmatrix} \boldsymbol{\varepsilon}_i \\ \boldsymbol{\varepsilon}_{iq} \end{pmatrix}_{(p_i+q)\times 1}.$$

Then we can get a (p_i+q)-dimensional random vector $\boldsymbol{Y}_i^*$ drawn from the multivariate t distribution with location vector $\boldsymbol{X}_i^*\boldsymbol{\beta}$, scale matrix $\boldsymbol{\Omega}_i^*$ and degree of freedom ν, which can be expressed as

$$\begin{pmatrix} \boldsymbol{Y}_i \\ \boldsymbol{y}_i \end{pmatrix} \sim t_{p_i+q}\left(\begin{pmatrix} \boldsymbol{X}_i\boldsymbol{\beta} \\ \boldsymbol{x}_i\boldsymbol{\beta} \end{pmatrix}, \sigma^2 \begin{pmatrix} \boldsymbol{\Omega}_{i11} & \boldsymbol{\Omega}_{i12} \\ \boldsymbol{\Omega}_{i21} & \boldsymbol{\Omega}_{i22} \end{pmatrix}, \nu\right). \tag{23}$$

The scale matrix of $\boldsymbol{Y}_i^*$ is $\boldsymbol{\Omega}_i^* = \boldsymbol{Z}_i^*\boldsymbol{\Gamma}\boldsymbol{Z}_i^{*\prime}+\boldsymbol{C}_i^*$ with $\boldsymbol{C}_i^* = [\rho^{|r-s|}]$ for $r, s = 1, \ldots, p_i+q$, $\boldsymbol{C}_i^*$ can be partitioned as

$$\boldsymbol{C}_i^* = \begin{bmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{bmatrix}.$$

It is noted that $\boldsymbol{C}_{11} = \boldsymbol{C}_i$ and $\boldsymbol{C}_{21} = \boldsymbol{C}_{12}'$. We will need the distribution of $\boldsymbol{y}_i$ given $\boldsymbol{Y}_i$. As the result of calculation by Bayesian approach

$$f(\boldsymbol{y}_i|\boldsymbol{Y}_i) \propto \int f(\boldsymbol{y}_i|\boldsymbol{Y}_i, \tau_i) f(\boldsymbol{Y}_i|\tau_i) g(\tau_i) d\tau_i,$$

we have

$$\boldsymbol{y}_i|\boldsymbol{Y}_i \sim t_q\left(\boldsymbol{\mu}_{i,2\cdot 1}, \omega_i \boldsymbol{\Sigma}_{i,22\cdot 1}, \nu + p_i\right),$$

where the location vector $\boldsymbol{\mu}_{i,2\cdot 1} = \boldsymbol{x}_i\boldsymbol{\beta} + \boldsymbol{\Omega}_{i21}\boldsymbol{\Omega}_{i11}^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta})$ and the scale matrix $\boldsymbol{\omega}_i\boldsymbol{\Sigma}_{i,22\cdot 1} = \sigma^2\boldsymbol{\omega}_i\left(\boldsymbol{\Omega}_{i22} - \boldsymbol{\Omega}_{i21}\boldsymbol{\Omega}_{i11}^{-1}\boldsymbol{\Omega}_{i12}\right)$ with $\omega_i = \frac{\nu+(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta})'(\sigma^2\boldsymbol{\Omega}_{i11})^{-1}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta})}{\nu+p_i}$. The detailed calculation are given in Appendix **D**. The minimized MSE predictor of future values $\boldsymbol{y}_i$ is $\hat{\boldsymbol{y}}_i$ which is the conditional expectation of $\boldsymbol{y}_i$ given $\boldsymbol{Y}_i$,

$$\hat{\boldsymbol{y}}_i = \hat{\boldsymbol{\mu}}_{i,2\cdot 1} = \boldsymbol{x}_i\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{\Omega}}_{i21}\hat{\boldsymbol{\Omega}}_{i11}^{-1}\left(\boldsymbol{Y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}}\right), \tag{24}$$

where $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\Omega}}_{i21}$, $\hat{\boldsymbol{\Omega}}_{i11}$ are replaced by their estimates, such as their ML or REML estimates, and the error covariance matrix of the predictor (26) is given by

$$\mathrm{E}[(\hat{\boldsymbol{y}}_i - \boldsymbol{y}_i)(\hat{\boldsymbol{y}}_i - \boldsymbol{y}_i)'] = \frac{(\nu + p_i)}{\nu + p_i - 2}\omega_i\boldsymbol{\Sigma}_{i,22\cdot 1}, \tag{25}$$

where

$$\begin{aligned} \boldsymbol{\Sigma}_{i,22\cdot 1} &= \sigma^2\Big(\boldsymbol{C}_{22\cdot 1} + (\boldsymbol{z}_i - \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i)\Big(\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\Big) \\ &\qquad (\boldsymbol{z}_i - \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i)'\Big), \quad \boldsymbol{W}_{11} = (\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i)^{-1}. \end{aligned}$$

The proof of (25) is shown in Appendix **E**.

5. EXAMPLE

We will next analyze the relapsing-remitting MS sub-study data which involved burden repeated measurements of 52 patients at the University of British Columbia

(UBC). The sample are taken approximately six weeks over two years at UBC. The patients are randomized to placebo (PL), low-dose (LD), and high-dose (HD) groups. Each group involved 17, 18, and 17 patients, respectively. Two patients, one of them is in the low-dose (LD) group and the other is in the high-dose (HD) group, are dropped out early in the study. Besides, one patient in the low-dose (LD) group had 3 measurements zero on MRI scans. Therefore, we only use the remaining data of 49 patients in this analysis.

For longitudinal study, we use the incomplete data. Most of the patients have 17 measurement points except 5 patients, one in placebo (PL) group dropped out after completing 14 visits, two in low-dose (LD) group dropped out after completing 13 visits, and two in high-dose (HD) group dropped out after completing 12 visits.

We use the patient's burden of disease, the total area of MS lesions of the MRI scan (mm^2), as an indicator to measure the disease of MS. The burden of disease at time point j is denoted as Area(j), and Area(0) means MRI area of the patient at the baseline time point. We use the log relative burden LRB(j)=log(Area(j)/Area(0)) to be the response variable ($Y_{ij},\ j = 1, \ldots, p_i$) as a measure of the severity with respect to the baseline. The clinical procedure and the usefulness of LRB are detailly described by Gill (2000) and D'yachkova *et al.* (1997).

Figure 1 depicts the LRB evolution over the 49 patients at each time point from various groups. Figure 2 shows the boxplots of average LRB for the three treatments, and indicates the average of the HD group is lower than the others. Broadly, the variability of LRB values for the PL is largest and is smallest for the HD group. Since IFNB can be used to reduce the frequency and severity of relapses, the patients in HD group with higher dose have remitted the frequency of relapses and have stable condition. Hence, the variance of LRB values for the HD group is smallest. The boxplots also clearly shows that there are three patients having larger LRB values than the other patients in the PL group, and there is one patient in the LD group having lower LRB values than the remaining patients.

Since the structure of the three treatments are different from each other, we analyze the MS data for each treatment group. The following is the form of the

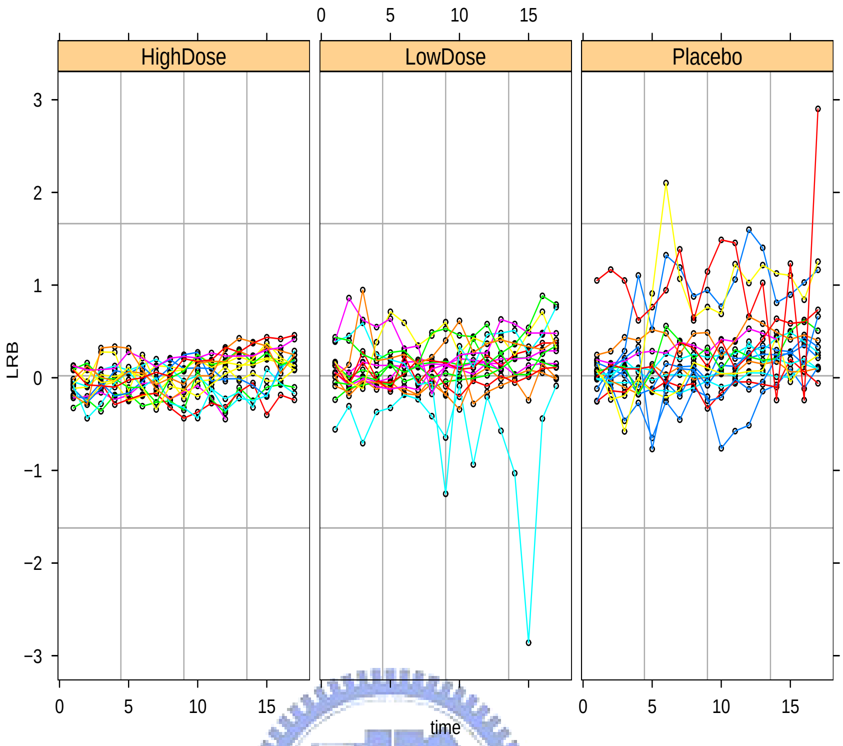

Figure 1: Trellis graphics of Multiple Sclerosis data

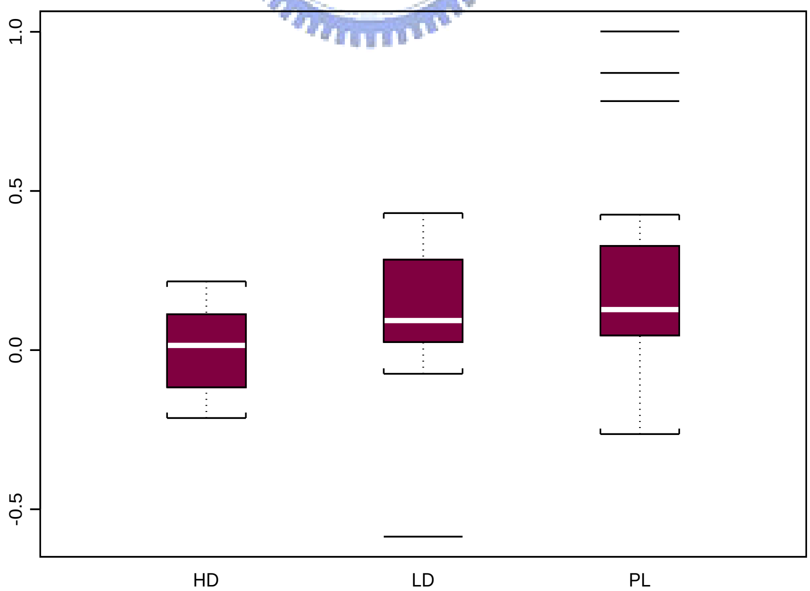

Figure 2: Box plot for average LRB

fixed-effects $\boldsymbol{\beta}$ and the design matrix of the fixed-effects $\boldsymbol{X}_i$.

$$\boldsymbol{X}_i = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & p_i \end{bmatrix}', \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix},$$

where β_0 and β_1 represent the fixed-effects intercept and slope, respectively. To explore the autocorrelation in the within-subject errors, we start by fitting a t linear mixed model with random intercept and white noise structure ($\rho = 0$). With these assumptions the model takes the following form:

$$\begin{aligned} \boldsymbol{Y}_i &= \boldsymbol{X}_i\boldsymbol{\beta} + \mathbf{1}_i b_i + \boldsymbol{\varepsilon}_i, \quad b_i \mid \tau_i \sim N(0, \sigma^2\gamma), \\ \boldsymbol{\varepsilon}_i \mid \tau_i &\sim N_{p_i}(\mathbf{0},\ \sigma^2\boldsymbol{I}_{p_i}), \quad \tau_i \sim \text{Gamma}(\nu/2, \nu/2), \end{aligned} \tag{26}$$

where $\mathbf{1}_i = (1, 1, \ldots, 1)'$, a $p_i \times 1$ vector with all elements 1, and the covariance matrix of $\boldsymbol{Y}_i$ is $\sigma^2(\mathbf{1}\boldsymbol{\Gamma}\mathbf{1}' + \boldsymbol{I}_{p_i})$. We fit (26) to the MS data for each group. We

Table 1: ML estimation results and score test statistics for model (26), where F is such that $\Gamma = F^2$

Group	β_0	β_1	σ^2	F	ν	AIC	λ_s
PL	-0.0076 (0.0460)	0.0185 (0.0015)	0.0143 (0.0042)	1.8290 (0.272)	1.82 (0.59)	-99.02	16.91
LD	-0.0048 (0.0325)	0.0122 (0.0015)	0.0132 (0.0039)	0.9462 (0.194)	2.08 (0.71)	-145.58	21.30
HD	-0.0846 (0.0356)	0.0143 (0.0015)	0.0128 (0.0027)	1.0856 (0.2185)	4.58 (1.87)	-251.44	49.58

have no evidence to prove the errors with an AR(1) process. Therefore, we consider model (26) to test whether the autocorrelation exits or not by score test. The score statistic under $H_0 : \rho = 0$ and $H_0 : \rho \neq 0$ is expressed as equation (14). Table 1 lists the ML estimates and Akaike's Information Criterions (AICs) along with the score test statistics λ_s. Here we define AIC to be -2(log-likelihood-the number of model parameters). As shown in the table, the score statistics are all highly significant compared with χ^2_1 distribution, indicating that there exists autocorrelation in the

Table 2: ML estimation results for model (27), where F is such that $\Gamma = F^2$

Group	β_0	β_1	σ^2	F	ρ	ν	AIC
PL	-0.0036	0.0176	0.0154	1.8290	0.3164	1.81	-116.26
	(0.0476)	(0.0020)	(0.0046)	(0.7322)	(0.0685)	(0.59)	
LD	-0.0040	0.0125	0.0137	0.8734	0.3584	1.97	-167.10
	(0.0343)	(0.0021)	(0.0043)	(0.1994)	(0.0712)	(0.67)	
HD	-0.0814	0.0137	0.0174	0.8348	0.5587	6.79	-310.38
	(0.0400)	(0.0027)	(0.0040)	(0.2192)	(0.0689)	(3.15)	

Table 3: ML estimation results for model (28), where $\mathbf{F}$ is the Cholesky decomposition of $\mathbf{\Gamma}$

Group	β_0	β_1	σ^2	F_{11},	F_{12},	F_{22}	ν	AIC
PL	-0.0139	0.0190	0.0129	1.1255	0.0352	0.0480	1.76	-108.20
	(0.0358)	(0.0023)	(0.0038)	(0.2629)	(0.1022)	(0.0323)	(0.57)	
LD	-0.0070	0.0125	0.0114	1.0115	-0.0264	0.0663	1.97	-153.68
	(0.0319)	(0.0025)	(0.0040)	(0.5310)	(0.2712)	(0.0344)	(0.67)	
HD	-0.0857	0.0140	0.0113	0.9412	-0.0136	0.0989	5.49	-273.26
	(0.0299)	(0.0032)	(0.0023)	(0.2665)	(0.1216)	(0.0253)	(2.39)	

within-subjects errors of model (26). It leads us to consider the alternative model,

$$\begin{aligned} \boldsymbol{Y}_i &= \boldsymbol{X}_i\boldsymbol{\beta} + \mathbf{1}_i b_i + \boldsymbol{\varepsilon}_i, \quad b_i \mid \tau_i \sim N(0, \sigma^2\gamma), \\ \boldsymbol{\varepsilon}_i \mid \tau_i &\sim N_{p_i}(\mathbf{0},\ \sigma^2\boldsymbol{C}_i(\rho)), \quad \tau_i \sim \text{Gamma}(\nu/2, \nu/2), \end{aligned} \tag{27}$$

with AR(1) structure for the within-subject errors. The covariance matrix of $\boldsymbol{Y}_i$ is $\sigma^2(\mathbf{1}\mathbf{\Gamma}\mathbf{1}' + \boldsymbol{C}_i(\rho))$. The resulting ML estimates with the associated standard errors in parentheses, along with AICs are shown in Table 2. As expected, this kind of model has lower AICs than model (26). In some cases, however, a more general random effects may be useful for the interpretation of autocorrelation in the within-subjects error. Therefore, we consider to fit the following model,

$$\begin{aligned} \boldsymbol{Y}_i &= \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{X}_i\boldsymbol{b}_i + \boldsymbol{\varepsilon}_i, \quad \boldsymbol{b}_i \mid \tau_i \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{\Gamma}), \\ \boldsymbol{\varepsilon}_i \mid \tau_i &\sim N_{p_i}(\boldsymbol{0},\ \sigma^2\boldsymbol{I}_{p_i}), \quad \tau_i \sim \text{Gamma}(\nu/2, \nu/2). \end{aligned} \tag{28}$$

with random intercept, random slope and white noise structure ($\rho = 0$). The covariance matrix of $\boldsymbol{Y}_i$ is $\sigma^2(\boldsymbol{X}_i\boldsymbol{\Gamma}\boldsymbol{X}_i' + \boldsymbol{C}_i(\rho))$. The ML estimation results are given in Table 3. The values of AICs were found to be -108.20 for PL, -153.68 for LD and -273,26 for HD, which did not improve model (27) with the corresponding AICs being -116.26, -167.10 and -310.38, respectively. In addition, the elements of F_{12} and F_{22} were relatively small compared with their respective standard errors, indicating the random slope effects should be extra. Overall, we conclude that model (27) is the most suitable fit for the MS data since it not only explains the autocorrelation but also with fewer parameters and lower AICs.

In order to see whether the model (27) is more robust or not, we compare the t and normal linear mixed model such as

$$\boldsymbol{Y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{1}_i b_i + \boldsymbol{\varepsilon}_i, \quad b_i \sim N(0, \sigma^2\gamma), \quad \boldsymbol{\varepsilon}_i \sim N_{p_i}(\boldsymbol{0},\ \sigma^2\boldsymbol{C}_i(\rho)). \tag{29}$$

Table 4 shows the log-likelihood and the values of AIC for model (27) and (29). The results reflect model (29) are not suitable for the PL and LD groups based on substantially larger AIC statistics, while it is comparable for the HD group. We can see clearly that no matter which treatment we take, the t linear mixed model has smaller values of AIC than the normal linear model. Therefore, the t linear mixed model with AR(1) owns the better explanation than the normal linear model for the MS data. Since the REML estimates are similar to the ML estimates, we just list the estimators by maximum likelihood method.

Figure 3 displays the corresponding normal quantile plots for the residuals from model (29). Obviously, the residuals of PL and LD seriously deviate from normality, confirming the presence of longer-than-normal tails. In contract, the departure of normality for HD is minor. Based on these findings, it appears that a normal model (29) might be adequate for HD.

We next compare the prediction accuracy between t model (27) and normal

Table 4: Estimated logarithm of maximum likelihood and AIC

Model	*number of parameters*	Placebo		Low dose		High dose	
		$\log_e L$	AIC	$\log_e L$	AIC	$\log_e L$	AIC
nlme	5	-65.67	141.34	-15.13	40.26	154.65	-299.30
tlme	6	64.13	**-116.26**	89.55	**-167.10**	161.19	**-310.38**

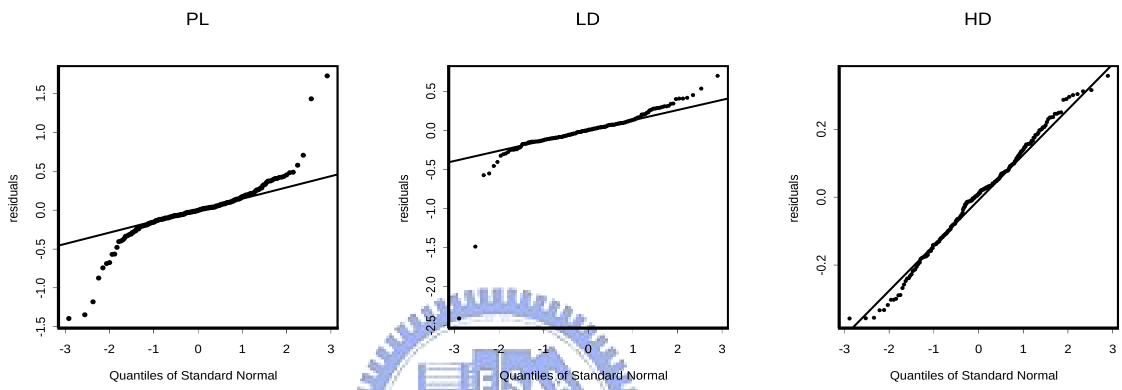

Figure 3: Normal quantile plot for residuals

model (29). We use the predictive sample reuse procedure described in Geisser (1975). The process is addressed that the last point of each vector measurement is taken out as the true value and use the rest to be predicted each time. As a measure of precision we use the mean absolute relative deviation (MARD), which is defined as

$$\text{MARD} = \frac{1}{N}\sum_{j=1}^{N}\left|\frac{y_{pj} - \hat{y}_{pj}}{y_{pj}}\right|$$

The MARD is a widely used criterion when applied to time series observations are monotonically increasing. We restrict our attention to the one-step-ahead forecasts by setting $p = 13, 14, 15, 16, 17$. We withheld one response vector $\boldsymbol{Y}_j$ and use the most recent $p-1$ measurements as the sample to predict the last component Y_{jp} using (24). The process is repeated for each of N subjects. Table 5 shows fixed effects estimates and prediction results from these two models. Comparatively, the

Table 5: Comparison fixed effects estimates and one-step-ahead forecast accuracy in terms of MARD

Group	Time of point being forecast	Normal linear mixed model			t linear mixed model			
		β_0	β_1	MARD_N	β_0	β_1	ν	MARD_t
PL	13	0.052	0.022	0.532	0.004	0.017	1.93	0.414
	14	0.047	0.023	1.186	0.004	0.018	1.89	1.020
	15	0.064	0.020	1.825	-0.003	0.019	1.86	1.610
	16	0.050	0.023	0.818	-0.003	0.019	1.93	0.702
	17	0.065	0.020	0.717	-0.006	0.019	1.88	0.675
	Average			1.016				0.884
LD	13	0.059	0.004	0.528	0.017	0.010	2.11	0.504
	14	0.053	0.006	1.044	0.016	0.009	2.25	1.010
	15	0.056	0.005	0.580	0.019	0.009	2.30	0.591
	16	0.079	0.001	0.716	0.016	0.010	1.49	0.484
	17	0.046	0.007	0.302	-0.002	0.012	1.50	0.287
	Average			0.634				0.572
HD	13	-0.067	0.006	0.803	-0.063	0.008	10.43	0.817
	14	-0.081	0.009	0.245	-0.071	0.010	8.94	0.220
	15	-0.078	0.008	0.414	-0.068	0.010	9.26	0.382
	16	-0.087	0.010	0.765	-0.080	0.013	8.10	0.741
	17	-0.089	0.011	0.412	-0.079	0.013	7.08	0.397
	Average			0.528				0.511
	Overall average			0.726				0.656

t-based model has much smaller MARD than those of normal for PL and LD. The percentages of improvement ($|\text{MARD}_N - \text{MARD}_t|/|\text{MARD}_N| \times 100\%$) using the t-based model are 13% and 9.3%, respectively. On the contrary, the prediction performances using the t-based model for HD is slightly better than the normal model with only 3.2% improvement. It is clear that the t-based model not only provides better model fittings, but also yields smaller forecast errors for the MS data. Regardless of the treatment groups, the overall average MARD for the t-based model is smaller than that of the normal model, with an improvement of about 9.6%.

6. DISCUSSION

We propose the t linear mixed model with AR(1) dependence for longitudinal data analysis. The heavy-tailed t linear mixed modelling provides an alternatively robust way of dealing with longitudinal data when some outlying observations are present. Besides, from our the illustrated MS data it is encouraging that the use of t linear mixed model coupling AR(1) structure offers better fitting as well as prediction performance than the use of normal counterpart. Inclusion of the simple AR(1) dependence could lead an appropriate representation of correlation structure. In the future tasks, one may consider higher order ARMA(p, q) dependence structures for the within-subject errors. A nature extension of the present research is to explore some potential advantages of the Markov Chain Monte Carlo in the Bayesian paradigm.

References

Chi, E.M. and Reinsel, G.C. (1989). Models for longitudinal data with random effects and AR(1) errors. *Journal of the American Statistical Association.* **84**, 452-459.

D'yachkova, Y.D., Petkau, J., and White, R. (1997). Longitudinal analysis for magnetic resonance imaging outcomes in multiple sclerosis clinical trials. *Journal of Biopharmaceutical Statistics* **7**, 501-531.

Geisser S. 1975. The predictive sample reuse method with applications. *Journal of the American Statistical Association.* **70**, 320-328.

Gill, P.S. (2000). A robust mixed linear model analysis for longitudinal data. *Statistics and Medicine,* **19**, 975-987.

Harville, D.A. (1974) Bayesian Inference for Variance Components Using Only Error Contrast. *Biometrika.* **61**, 383-385.

Laird, N.M. and Ware, J.H. (1982). Random effects models for longitudinal data. *Biometrics.* **38**, 963-974.

Pinheiro, J.C., Liu, C.H., & Wu, Y.N. (2001). Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate t distribution. *Journal of Computational and Graphical Statistics.* **10**, 249-276.

Rao C.R. (1973). Linear Statistical inference and Its Application. John Wiley: New York.

Welsh, A.H. and Richardson, A.M. (1997). Approaches to the robust estimation of mixed models. *Handbooks of Statistics.* **15**, 343-383.

APPENDIX **A**: The evaluation of $\mathbf{s}$

The following is the score vector $\mathbf{s}_\alpha$, which we define as the first derivatives of the log-likelihood function with respect to $\boldsymbol{\alpha}$:

$$\begin{aligned}
\mathbf{s}_\beta &= \sum_{i=1}^{N}(\nu+p_i)\frac{\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)},\\
\mathbf{s}_{\sigma^2} &= -\frac{n}{2\sigma^2}+\frac{1}{2\sigma^2}\sum_{i=1}^{N}(\nu+p_i)\left(\frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}{\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right),\\
\mathbf{s}_\nu &= \frac{1}{2}\sum_{i=1}^{N}\left\{\phi\left(\frac{\nu+p_i}{2}\right)-\phi\left(\frac{\nu}{2}\right)-\frac{p_i}{\nu}-\log\left(1+\frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}{\sigma^2\nu}\right)\right.\\
&\quad\left.+\frac{(\nu+p_i)}{\nu}\frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}{\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right\},\\
[\mathbf{s}_\omega]_r &= -\frac{1}{2}\sum_{i=1}^{N}\left\{\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}^{-1})-(\nu+p_i)\left(\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}^{-1}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right)\right\},
\end{aligned}$$

where $\boldsymbol{\omega}=(\mathrm{vech}(\mathbf{F}),\rho)$, $\dot{\boldsymbol{\Lambda}}_{ir}=\partial\boldsymbol{\Lambda}_i(\boldsymbol{\omega})/\partial\omega_r$, for $r=1,\ldots,k+1$ and $\phi(x)=d/dx\log(\Gamma(x))$.

To obtain the Fisher information matrix, we must get the Hessian matrix which is the second derivatives of the log-likelihood function.

We have

$$\begin{aligned}
\frac{d}{d\boldsymbol{\beta}}\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho) &= -\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i-\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{X}_i=-2\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{X}_i,\\
\frac{d}{ds}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}) &= -\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}+\boldsymbol{\Lambda}_i^{-1}[\ddot{\boldsymbol{\Lambda}}_{irs}\boldsymbol{\Lambda}_i^{-1}+\dot{\boldsymbol{\Lambda}}_{ir}(-\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1})]\\
&= -2\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}+\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs}\boldsymbol{\Lambda}_i^{-1},
\end{aligned}$$

then the elements of the Hessian matrix are given below:

$$\begin{aligned}
\mathbf{H}_{\beta\beta} &= \sum_{i=1}^{N}(\nu+p_i)\times\frac{\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{X}_i(\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho))-\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i(-2\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{X}_i)}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2}\\
&= \sum_{i=1}^{N}(\nu+p_i)\left\{\frac{-\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{X}_i}{\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}+\frac{2\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{X}_i}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2}\right\},\\
\mathbf{H}_{\beta\sigma^2} &= \sum_{i=1}^{N}(\nu+p_i)\times\frac{-\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i\nu}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2}=\sum_{i=1}^{N}-\nu(\nu+p_i)\frac{\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2},
\end{aligned}$$

$$
\begin{aligned}
\mathbf{H}_{\beta\nu} &= \sum_{i=1}^{N} \frac{(\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i)(\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)) - (\nu+p_i)(\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i)\sigma^2}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2} \\
&= \sum_{i=1}^{N} \left\{ \frac{\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)} - \sigma^2(\nu+p_i)\frac{\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2} \right\}, \\
[\mathbf{H}_{\beta\omega}]_r &= \sum_{i=1}^{N} \frac{(\nu+p_i)[-\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i](\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho))}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2} \\
&\quad - \frac{(\nu+p_i)(\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i)(-\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i)}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2} \\
&= -\sum_{i=1}^{N}(\nu+p_i) \left\{ \frac{\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)} - \frac{\boldsymbol{X}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2} \right\}, \\
\mathbf{H}_{\sigma^2\sigma^2} &= \frac{n}{2\sigma^4} - \frac{1}{2\sigma^4}\sum_{i=1}^{N}(\nu+p_i)\frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}{\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)} + \frac{1}{2\sigma^4}\sum_{i=1}^{N}(\nu+p_i)\frac{-\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)\nu}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2} \\
&= \frac{n}{2\sigma^4} - \frac{1}{2\sigma^4}\left\{ \sum_{i=1}^{N}(\nu+p_i)\left(\frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}{\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)} + \frac{\sigma^2\nu(\nu+p_i)\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2} \right) \right\}, \\
\mathbf{H}_{\sigma^2\nu} &= \frac{1}{2\sigma^2}\left\{ \sum_{i=1}^{N}\left(\frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}{\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)} + \frac{(\nu+p_i)(0 - \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)\sigma^2)}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2} \right) \right\} \\
&= \frac{1}{2\sigma^2}\sum_{i=1}^{N}\left\{ \frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}{\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)} - \frac{\sigma^2(\nu+p_i)\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2} \right\}, \\
[\mathbf{H}_{\sigma^2\omega}]_r &= \frac{1}{2\sigma^2}\sum_{i=1}^{N}\left\{ (\nu+p_i)\frac{-\sigma^2\nu(\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i)}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2} \right\} \\
&= -\frac{\nu}{2}\sum_{i=1}^{N}\left\{ (\nu+p_i)\frac{(\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i)}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2} \right\}, \\
\mathbf{H}_{\nu\nu} &= \frac{1}{2}\sum_{i=1}^{N}\left\{ \frac{1}{2}\psi\left(\frac{\nu+p_i}{2}\right) - \frac{1}{2}\psi\left(\frac{\nu}{2}\right) + \frac{p_i}{\nu^2} - \frac{\sigma^2\nu}{\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\left(-\frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}{\sigma^2\nu^2}\right) \right. \\
&\quad \left. + \frac{-\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)\sigma^2}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2} - \frac{p_i}{\nu^2}\frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}{\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)} + \frac{p_i}{\nu}\frac{-\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)\sigma^2}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2} \right\} \\
&= \frac{1}{2}\sum_{i=1}^{N}\left\{ \frac{1}{2}\psi\left(\frac{\nu+p_i}{2}\right) - \frac{1}{2}\psi\left(\frac{\nu}{2}\right) + \frac{p_i}{\nu^2} + \frac{1}{\nu}\frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}{\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)} \right. \\
&\quad \left. - \frac{p_i}{\nu^2}\frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}{\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)} - \frac{\nu+p_i}{\nu}\frac{\sigma^2\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)\sigma^2}{[\sigma^2\nu + \boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2} \right\},
\end{aligned}
$$

$$
\begin{aligned}
[\mathbf{H}_{\nu\omega}]_r &= \frac{1}{2}\sum_{i=1}^{N}\left\{\frac{-\sigma^2\nu(-\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i)}{\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\frac{1}{\sigma^2\nu}+\frac{\nu+p_i}{\nu}\frac{-\sigma^2\nu(-\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i)}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2}\right\}\\
&= \frac{1}{2}\sum_{i=1}^{N}\left\{\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}+\frac{\nu+p_i}{\nu}\frac{\sigma^2\nu(-\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i)}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2}\right\}\\
&= \frac{1}{2}\sum_{i=1}^{N}\left\{\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}-\frac{\sigma^2(\nu+p_i)(\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i)}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2}\right\},\\
[\mathbf{H}_{\omega\omega}]_{rs} &= -\frac{1}{2}\sum_{i=1}^{N}\Big\{\mathrm{tr}(-\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir})+\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs})-(\nu+p_i)\frac{1}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2}\\
&\quad\times\Big[\boldsymbol{e}_i'(-2\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}+\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs}\boldsymbol{\Lambda}_i^{-1})\boldsymbol{e}_i(\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho))\\
&\quad-\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i(-\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i)\Big]\Big\}\\
&= \frac{1}{2}\sum_{i=1}^{N}\Big\{\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir})-\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs})\\
&\quad-(\nu+p_i)\frac{\boldsymbol{e}_i'(2\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}-\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs}\boldsymbol{\Lambda}_i^{-1})\boldsymbol{e}_i}{\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\\
&\quad+(\nu+p_i)\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2}\Big\}.
\end{aligned}
$$

The Fisher information matrix is the negative expectation of the Hessian matrix. The following lists expectations which we need to obtain the Fisher information matrix.

The marginal distribution of $\boldsymbol{Y}_i$ is

$$
\begin{aligned}
f(\boldsymbol{Y}_i) &= \int\int f(\boldsymbol{Y}_i\mid \boldsymbol{b}_i,\tau_i)f(\boldsymbol{b}_i\mid\tau_i)f(\tau_i)d\boldsymbol{b}_i d\tau_i\\
&= \frac{\Gamma(\frac{\nu+p_i}{2})|\boldsymbol{\Lambda}_i|^{-1/2}}{\Gamma(\frac{\nu}{2})(\pi\nu\sigma^2)^{p_i/2}}\left(1+\frac{(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta})'\boldsymbol{\Lambda}_i^{-1}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta})}{\nu\sigma^2}\right)^{-(\nu+p_i)/2}\\
&= \frac{\Gamma(\frac{\nu+p_i}{2})|\boldsymbol{\Lambda}_i|^{-1/2}}{\Gamma(\frac{\nu}{2})(\pi\nu\sigma^2)^{p_i/2}}\left(1+\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{\nu\sigma^2}\right)^{-(\nu+p_i)/2},
\end{aligned}
$$

where $\boldsymbol{e}_i=\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta}$, and $\boldsymbol{\Lambda}_i=\boldsymbol{\Lambda}_i(\boldsymbol{\Gamma},\rho)=\boldsymbol{Z}_i\boldsymbol{\Gamma}\boldsymbol{Z}_i'+\boldsymbol{C}_i(\rho)$.

(a)

$$
\mathrm{E}\left(\frac{1}{\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right)
$$

$$
\begin{aligned}
&= \mathrm{E}\left(\frac{1}{\sigma^2\nu + (\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta})'\boldsymbol{\Lambda}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta})}\right) \\
&= \frac{1}{\sigma^2\nu}\mathrm{E}\left(\frac{1}{1+\frac{(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta})'\boldsymbol{\Lambda}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta})}{\sigma^2\nu}}\right) \\
&= \frac{1}{\sigma^2\nu}\int \frac{\Gamma(\frac{\nu+p_i}{2})|\boldsymbol{\Lambda}_i|^{-1/2}}{\Gamma(\frac{\nu}{2})(\pi\nu\sigma^2)^{p_i/2}}\left(1+\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{\nu\sigma^2}\right)^{-(\nu+2+p_i)/2} d\boldsymbol{Y}_i \\
&= \frac{1}{\sigma^2\nu}\frac{\boldsymbol{\Gamma}\left(\frac{\nu+p_i}{2}\right)\boldsymbol{\Gamma}\left(\frac{\nu+2}{2}\right)}{\boldsymbol{\Gamma}\left(\frac{\nu}{2}\right)\boldsymbol{\Gamma}\left(\frac{\nu+p_i+2}{2}\right)}\int \frac{\Gamma(\frac{\nu+p_i+2}{2})|\frac{\nu}{\nu+2}\boldsymbol{\Lambda}_i|^{-1/2}}{\Gamma(\frac{\nu+2}{2})[\pi(\nu+2)\sigma^2]^{p_i/2}}\left(1+\frac{\boldsymbol{e}_i'[\frac{\nu}{\nu+2}\boldsymbol{\Lambda}_i]^{-1}\boldsymbol{e}_i}{(\nu+2)\sigma^2}\right)^{-(\nu+2+p_i)/2} d\boldsymbol{Y}_i \\
&= \frac{1}{\sigma^2\nu}\frac{\frac{\nu}{2}}{\frac{\nu+p_i}{2}} = \frac{1}{\sigma^2(\nu+p_i)}.
\end{aligned}
$$

It is noted that

$$
\left|\frac{\nu}{\nu+2}\boldsymbol{\Lambda}_i\right|^{-1/2} = \frac{(\nu+2)^{p_i/2}}{\nu^{p_i/2}}|\boldsymbol{\Lambda}_i|^{-1/2}.
$$

(b)

$$
\begin{aligned}
&\mathrm{E}\left(\frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}{\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right) = \mathrm{E}\left(\frac{\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)+\sigma^2\nu-\sigma^2\nu}{\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right) = \mathrm{E}\left(1-\frac{\sigma^2\nu}{\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right) \\
&= 1-\sigma^2\nu\times\frac{1}{\sigma^2(\nu+p_i)} = 1-\frac{\nu}{\nu+p_i} = \frac{p_i}{\nu+p_i}.
\end{aligned}
$$

(c)

$$
\begin{aligned}
&\mathrm{E}\left(\frac{\boldsymbol{e}_i\boldsymbol{e}_i'}{\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right) \\
&= \frac{1}{\sigma^2\nu}\mathrm{E}\left(\frac{\boldsymbol{e}_i\boldsymbol{e}_i'}{1+\frac{(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta})'\boldsymbol{\Lambda}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta})}{\sigma^2\nu}}\right) \\
&= \frac{1}{\sigma^2\nu}\int \boldsymbol{e}_i\boldsymbol{e}_i'\frac{\Gamma(\frac{\nu+p_i}{2})|\boldsymbol{\Lambda}_i|^{-1/2}}{\Gamma(\frac{\nu}{2})[\pi\nu\sigma^2]^{p_i/2}}\left(1+\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{\nu\sigma^2}\right)^{-(\nu+2+p_i)/2} d\boldsymbol{Y}_i \\
&= \frac{1}{\sigma^2\nu}\frac{\boldsymbol{\Gamma}\left(\frac{\nu+p_i}{2}\right)\boldsymbol{\Gamma}\left(\frac{\nu+2}{2}\right)}{\boldsymbol{\Gamma}\left(\frac{\nu}{2}\right)\boldsymbol{\Gamma}\left(\frac{\nu+p_i+2}{2}\right)} \\
&\quad \int \boldsymbol{e}_i\boldsymbol{e}_i'\frac{\Gamma(\frac{\nu+p_i+2}{2})|\frac{\nu}{\nu+2}\boldsymbol{\Lambda}_i|^{-1/2}}{\Gamma(\frac{\nu+2}{2})[\pi(\nu+2)\sigma^2]^{p_i/2}}\left(1+\frac{\boldsymbol{e}_i'[\frac{\nu}{\nu+2}\boldsymbol{\Lambda}_i]^{-1}\boldsymbol{e}_i}{(\nu+2)\sigma^2}\right)^{-(\nu+2+p_i)/2} d\boldsymbol{Y}_i \\
&= \frac{1}{\sigma^2\nu}\frac{\frac{\nu}{2}}{\frac{\nu+p_i}{2}}\sigma^2\boldsymbol{\Lambda}_i = \frac{\boldsymbol{\Lambda}_i}{\nu+p_i}.
\end{aligned}
$$

It is noted that

$$\frac{\Gamma(\frac{\nu+p_i+2}{2})|\frac{\nu}{\nu+2}\mathbf{\Lambda}_i|^{-1/2}}{\Gamma(\frac{\nu+2}{2})[\pi(\nu+2)\sigma^2]^{p_i/2}}\left(1+\frac{\boldsymbol{e}_i'[\frac{\nu}{\nu+2}\mathbf{\Lambda}_i]^{-1}\boldsymbol{e}_i}{(\nu+2)\sigma^2}\right)^{-(\nu+2+p_i)/2}$$

$\boldsymbol{Y}_i$ is the distributed as $t_{p_i}(\boldsymbol{X}_i\boldsymbol{\beta}, \frac{\nu}{\nu+2}\mathbf{\Lambda}_i, \nu+2)$.

Then

$$\int \boldsymbol{e}_i\boldsymbol{e}_i' \frac{\Gamma(\frac{\nu+p_i+2}{2})|\frac{\nu}{\nu+2}\mathbf{\Lambda}_i|^{-1/2}}{\Gamma(\frac{\nu+2}{2})[\pi(\nu+2)\sigma^2]^{p_i/2}}\left(1+\frac{\boldsymbol{e}_i'[\frac{\nu}{\nu+2}\mathbf{\Lambda}_i]^{-1}\boldsymbol{e}_i}{(\nu+2)\sigma^2}\right)^{-(\nu+2+p_i)/2} d\boldsymbol{Y}_i$$

$$= \sigma^2 \times \frac{\nu+2}{(\nu+2)-2} \times \frac{\nu}{\nu+2}\mathbf{\Lambda}_i = \sigma^2\mathbf{\Lambda}_i.$$

(d)

To obtain $\mathrm{E}\left(\frac{\mathbf{\Delta}_i(\boldsymbol{\beta}, \mathbf{\Gamma}, \rho)}{[\sigma^2\nu+\mathbf{\Delta}_i(\boldsymbol{\beta}, \mathbf{\Gamma}, \rho)]^2}\right)$, we must compute the expectation of $\mathrm{E}\left(\frac{1}{[\sigma^2\nu+\mathbf{\Delta}_i(\boldsymbol{\beta}, \mathbf{\Gamma}, \rho)]^2}\right)$.

First

$$\begin{aligned}
&\mathrm{E}\left(\frac{1}{[\sigma^2\nu+\mathbf{\Delta}_i(\boldsymbol{\beta},\mathbf{\Gamma},\rho)]^2}\right)\\
=\ & \frac{1}{\sigma^4\nu^2}\mathrm{E}\left(\frac{1}{[1+\frac{(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta})'\mathbf{\Lambda}_i^{-1}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta})}{\sigma^2\nu}]^2}\right)\\
=\ & \frac{1}{\sigma^4\nu^2}\int \frac{\Gamma(\frac{\nu+p_i}{2})|\mathbf{\Lambda}_i|^{-1/2}}{\Gamma(\frac{\nu}{2})(\pi\nu\sigma^2)^{p_i/2}}\left(1+\frac{\boldsymbol{e}_i'\mathbf{\Lambda}_i^{-1}\boldsymbol{e}_i}{\nu\sigma^2}\right)^{-(\nu+4+p_i)/2} d\boldsymbol{Y}_i\\
=\ & \frac{1}{\sigma^4\nu^2}\frac{\mathbf{\Gamma}\left(\frac{\nu+p_i}{2}\right)\mathbf{\Gamma}\left(\frac{\nu+4}{2}\right)}{\mathbf{\Gamma}\left(\frac{\nu}{2}\right)\mathbf{\Gamma}\left(\frac{\nu+p_i+4}{2}\right)}\\
&\int \frac{\Gamma(\frac{\nu+p_i+4}{2})|\frac{\nu}{\nu+4}\mathbf{\Lambda}_i|^{-1/2}}{\Gamma(\frac{\nu+4}{2})[\pi(\nu+4)\sigma^2]^{p_i/2}}\left(1+\frac{\boldsymbol{e}_i'[\frac{\nu}{\nu+4}\mathbf{\Lambda}_i]^{-1}\boldsymbol{e}_i}{(\nu+4)\sigma^2}\right)^{-(\nu+4+p_i)/2} d\boldsymbol{Y}_i\\
=\ & \frac{1}{\sigma^4\nu^2}\frac{(\frac{\nu}{2})(\frac{\nu}{2}+1)}{(\frac{\nu+p_i}{2})(\frac{\nu+p_i}{2}+1)} = \frac{\nu+2}{\sigma^4\nu(\nu+p_i)(\nu+p_i+2)}.
\end{aligned}$$

Then

$$\begin{aligned}
&\mathrm{E}\left(\frac{\mathbf{\Delta}_i(\boldsymbol{\beta},\mathbf{\Gamma},\rho)}{[\sigma^2\nu+\mathbf{\Delta}_i(\boldsymbol{\beta},\mathbf{\Gamma},\rho)]^2}\right) = \mathrm{E}\left(\frac{\mathbf{\Delta}_i(\boldsymbol{\beta},\mathbf{\Gamma},\rho)+\sigma^2\nu-\sigma^2\nu}{[\sigma^2\nu+\mathbf{\Delta}_i(\boldsymbol{\beta},\mathbf{\Gamma},\rho)]^2}\right)\\
=\ & \mathrm{E}\left(\frac{1}{\sigma^2\nu+\mathbf{\Delta}_i(\boldsymbol{\beta},\mathbf{\Gamma},\rho)}\right) - \sigma^2\nu\mathrm{E}\left(\frac{1}{[\sigma^2\nu+\mathbf{\Delta}_i(\boldsymbol{\beta},\mathbf{\Gamma},\rho)]^2}\right)\\
=\ & \frac{1}{\sigma^2(\nu+p_i)} - \sigma^2\nu \times \frac{\nu+2}{\sigma^4\nu(\nu+p_i)(\nu+p_i+2)}\\
=\ & \frac{1}{\sigma^2(\nu+p_i)} - \frac{\nu+2}{\sigma^2(\nu+p_i)(\nu+p_i+2)} = \frac{p_i}{\sigma^2(\nu+p_i)(\nu+p_i+2)}.
\end{aligned}$$

(e)

$$
\begin{aligned}
& \mathrm{E}\left(\frac{\boldsymbol{e}_i\boldsymbol{e}_i'}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2}\right) \\
= & \frac{1}{\sigma^4\nu^2}\mathrm{E}\left(\frac{\boldsymbol{e}_i\boldsymbol{e}_i'}{[1+\frac{(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta})'\boldsymbol{\Lambda}_i^{-1}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta})}{\sigma^2\nu}]^2}\right) \\
= & \frac{1}{\sigma^4\nu^2}\int \boldsymbol{e}_i\boldsymbol{e}_i'\frac{\Gamma(\frac{\nu+p_i}{2})|\boldsymbol{\Lambda}_i|^{-1/2}}{\Gamma(\frac{\nu}{2})[\pi\nu\sigma^2]^{p_i/2}}\left(1+\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{\nu\sigma^2}\right)^{-(\nu+4+p_i)/2} d\boldsymbol{Y}_i \\
= & \frac{1}{\sigma^4\nu^2}\frac{\boldsymbol{\Gamma}\left(\frac{\nu+p_i}{2}\right)\boldsymbol{\Gamma}\left(\frac{\nu+4}{2}\right)}{\boldsymbol{\Gamma}\left(\frac{\nu}{2}\right)\boldsymbol{\Gamma}\left(\frac{\nu+p_i+4}{2}\right)} \\
& \int \boldsymbol{e}_i\boldsymbol{e}_i'\frac{\boldsymbol{\Gamma}(\frac{\nu+p_i+4}{2})|\frac{\nu}{\nu+4}\boldsymbol{\Lambda}_i|^{-1/2}}{\Gamma(\frac{\nu+4}{2})[\pi(\nu+4)\sigma^2]^{p_i/2}}\left(1+\frac{\boldsymbol{e}_i'[\frac{\nu}{\nu+4}\boldsymbol{\Lambda}_i]^{-1}\boldsymbol{e}_i}{(\nu+4)\sigma^2}\right)^{-(\nu+4+p_i)/2} d\boldsymbol{Y}_i \\
= & \frac{1}{\sigma^4\nu^2}\times\frac{(\frac{\nu}{2})(\frac{\nu}{2}+1)}{(\frac{\nu+p_i}{2})(\frac{\nu+p_i}{2}+1)}\times\sigma^2\times\frac{\nu+4}{(\nu+4)-2}\times\frac{\nu}{\nu+4}\boldsymbol{\Lambda}_i \\
= & \frac{\boldsymbol{\Lambda}_i}{\sigma^2(\nu+p_i)(\nu+p_i+2)}.
\end{aligned}
$$

(f)

We must use the fact that

$$
\mathrm{E}\left(\frac{\partial^2\log f(\boldsymbol{Y}_i)}{\partial\boldsymbol{\omega}_r\partial\boldsymbol{\omega}_s}\right)=-\mathrm{E}\left(\frac{\partial\log f(\boldsymbol{Y}_i)}{\partial\boldsymbol{\omega}_r}\frac{\partial\log f(\boldsymbol{Y}_i)}{\partial\boldsymbol{\omega}_s}\right).
$$

Hence, the following equations are needed to obtain $\mathrm{E}\left(\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{irs}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{[\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2}\right)$,

$$
\begin{aligned}
\frac{\partial^2\log f(\boldsymbol{Y}_i)}{\partial\boldsymbol{\omega}_r\partial\boldsymbol{\omega}_s} = & \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\right)-\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs}\right) \\
& -\frac{1}{2}(\nu+p_i)\left\{\frac{\boldsymbol{e}_i'(2\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}-\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs}\boldsymbol{\Lambda}_i^{-1})\boldsymbol{e}_i}{\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right\} \\
& +\frac{1}{2}(\nu+p_i)\left\{\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{[\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2}\right\}
\end{aligned}
$$

$$
\begin{aligned}
- & \left(\frac{\partial\log f(\boldsymbol{Y}_i)}{\partial\boldsymbol{\omega}_r}\right)\left(\frac{\partial\log f(\boldsymbol{Y}_i)}{\partial\boldsymbol{\omega}_s}\right) \\
= & -\frac{1}{4}\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir})\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is})+\frac{1}{4}(\nu+p_i)\left(\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right)\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is})
\end{aligned}
$$

$$
+\frac{1}{4}(\nu+p_i)\left(\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right)\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir})
$$

$$
-\frac{1}{4}(\nu+p_i)^2\left\{\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{[\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2}\right\}
$$

and

$$
\begin{aligned}
&-\frac{1}{2}(\nu+p_i)\mathrm{E}\left\{\frac{\boldsymbol{e}_i'(2\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}-\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs}\boldsymbol{\Lambda}_i^{-1})\boldsymbol{e}_i}{\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right\}\\
=\;&-\frac{1}{2}(\nu+p_i)\mathrm{Etr}\left\{\frac{\boldsymbol{e}_i'(2\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}-\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs}\boldsymbol{\Lambda}_i^{-1})\boldsymbol{e}_i}{\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right\}\\
=\;&-\frac{1}{2}(\nu+p_i)\mathrm{Etr}\left\{\frac{\boldsymbol{e}_i\boldsymbol{e}_i'(2\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}-\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs}\boldsymbol{\Lambda}_i^{-1})}{\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right\}\\
=\;&-\frac{1}{2}(\nu+p_i)\mathrm{trE}\left\{\frac{\boldsymbol{e}_i\boldsymbol{e}_i'(2\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}-\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs}\boldsymbol{\Lambda}_i^{-1})}{\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right\}\\
=\;&-\frac{1}{2}(\nu+p_i)\frac{1}{\nu+p_i}\mathrm{tr}(2\boldsymbol{\Lambda}_i\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}-\boldsymbol{\Lambda}_i\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs}\boldsymbol{\Lambda}_i^{-1})\\
=\;&-\mathrm{tr}(\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1})+\frac{1}{2}\mathrm{tr}(\ddot{\boldsymbol{\Lambda}}_{irs}\boldsymbol{\Lambda}_i^{-1})\\
=\;&-\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is})+\frac{1}{2}\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs})
\end{aligned}
$$

$$
\begin{aligned}
&\frac{1}{4}(\nu+p_i)\mathrm{E}\left(\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right)=\frac{1}{4}(\nu+p_i)\mathrm{Etr}\left(\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right)\\
=\;&\frac{1}{4}(\nu+p_i)\mathrm{Etr}\left(\frac{\boldsymbol{e}_i'\boldsymbol{e}_i\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}}{\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right)=\frac{1}{4}(\nu+p_i)\mathrm{trE}\left(\frac{\boldsymbol{e}_i'\boldsymbol{e}_i\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}}{\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right)\\
=\;&\frac{1}{4}(\nu+p_i)\frac{1}{\nu+p_i}\mathrm{tr}\left(\boldsymbol{\Lambda}_i\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}\right)=\frac{1}{4}\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}).
\end{aligned}
$$

Then

$$
\begin{aligned}
\mathrm{E}\left(\frac{\partial^2\log f(\boldsymbol{Y}_i)}{\partial\boldsymbol{\omega}_r\partial\boldsymbol{\omega}_s}\right)=\;&\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\right)-\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs}\right)\\
&-\frac{1}{2}(\nu+p_i)\mathrm{E}\left\{\frac{\boldsymbol{e}_i'(2\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}-\boldsymbol{\Lambda}_i^{-1}\ddot{\boldsymbol{\Lambda}}_{irs}\boldsymbol{\Lambda}_i^{-1})\boldsymbol{e}_i}{\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)}\right\}\\
&+\frac{1}{2}(\nu+p_i)\mathrm{E}\left\{\frac{\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i\boldsymbol{e}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{e}_i}{[\sigma^2\nu+\boldsymbol{\Delta}(\boldsymbol{\beta},\boldsymbol{\Gamma},\rho)]^2}\right\}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{2}\mathrm{tr}\left(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{ir}\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{is}\right) - \frac{1}{2}\mathrm{tr}\left(\mathbf{\Lambda}_i^{-1}\ddot{\mathbf{\Lambda}}_{irs}\right) \\
&\quad -\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{ir}\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{is}) + \frac{1}{2}\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\ddot{\mathbf{\Lambda}}_{irs}) \\
&\quad +\frac{1}{2}(\nu+p_i)\mathrm{E}\left\{\frac{\boldsymbol{e}_i'\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{ir}\mathbf{\Lambda}_i^{-1}\boldsymbol{e}_i\boldsymbol{e}_i'\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{is}\mathbf{\Lambda}_i^{-1}\boldsymbol{e}_i}{[\sigma^2\nu+\mathbf{\Delta}(\boldsymbol{\beta},\mathbf{\Gamma},\rho)]^2}\right\}
\end{aligned}
$$

$$
\begin{aligned}
- &\ \mathrm{E}\left(\frac{\partial \log f(\boldsymbol{Y}_i)}{\partial \boldsymbol{\omega}_r}\right)\left(\frac{\partial \log f(\boldsymbol{Y}_i)}{\partial \boldsymbol{\omega}_s}\right) \\
= &\ -\frac{1}{4}\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{ir})\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{is}) + \frac{1}{4}(\nu+p_i)\mathrm{E}\left(\frac{\boldsymbol{e}_i'\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{ir}\mathbf{\Lambda}_i^{-1}\boldsymbol{e}_i}{\sigma^2\nu+\mathbf{\Delta}(\boldsymbol{\beta},\mathbf{\Gamma},\rho)}\right)\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{is}) \\
&\ +\frac{1}{4}(\nu+p_i)\mathrm{E}\left(\frac{\boldsymbol{e}_i'\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{is}\mathbf{\Lambda}_i^{-1}\boldsymbol{e}_i}{\sigma^2\nu+\mathbf{\Delta}(\boldsymbol{\beta},\mathbf{\Gamma},\rho)}\right)\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{ir}) \\
&\ -\frac{1}{4}(\nu+p_i)^2\mathrm{E}\left\{\frac{\boldsymbol{e}_i'\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{ir}\mathbf{\Lambda}_i^{-1}\boldsymbol{e}_i\boldsymbol{e}_i'\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{is}\mathbf{\Lambda}_i^{-1}\boldsymbol{e}_i}{[\sigma^2\nu+\mathbf{\Delta}(\boldsymbol{\beta},\mathbf{\Gamma},\rho)]^2}\right\} \\
= &\ -\frac{1}{4}\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{ir})\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{is}) + \frac{1}{4}\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{ir})\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{is}) \\
&\ +\frac{1}{4}\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{is})\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{ir}) - \frac{1}{4}(\nu+p_i)^2\mathrm{E}\left\{\frac{\boldsymbol{e}_i'\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{ir}\mathbf{\Lambda}_i^{-1}\boldsymbol{e}_i\boldsymbol{e}_i'\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{is}\mathbf{\Lambda}_i^{-1}\boldsymbol{e}_i}{[\sigma^2\nu+\mathbf{\Delta}(\boldsymbol{\beta},\mathbf{\Gamma},\rho)]^2}\right\}.
\end{aligned}
$$

We conclude that

$$
\begin{aligned}
&\mathrm{E}\left(\frac{\boldsymbol{e}_i'\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{ir}\mathbf{\Lambda}_i^{-1}\boldsymbol{e}_i\boldsymbol{e}_i'\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{irs}\mathbf{\Lambda}_i^{-1}\boldsymbol{e}_i}{[\sigma^2\nu+\mathbf{\Delta}(\boldsymbol{\beta},\mathbf{\Gamma},\rho)]^2}\right) \\
&= \frac{1}{(\nu+p_i)(\nu+p_i+2)}\left(\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{ir})\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{is}) + 2\mathrm{tr}(\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{ir}\mathbf{\Lambda}_i^{-1}\dot{\mathbf{\Lambda}}_{is})\right).
\end{aligned}
$$

The following are the components of Fisher information matrix $\mathbf{J}$: Let $\boldsymbol{\omega} = (\boldsymbol{f}, \rho)'$, and $\boldsymbol{f} = (f_1, f_2, \ldots, f_k)'$

$$
\begin{aligned}
\mathbf{J}_{\beta\beta} &= \sum_{i=1}^{N} \frac{\nu+p_i}{\sigma^2(\nu+p_i+2)}\boldsymbol{X}_i{}'\mathbf{\Lambda}_i^{-1}\boldsymbol{X}_i, \\
\mathbf{J}_{\sigma^2\sigma^2} &= \frac{\nu}{2\sigma^4}\sum_{i=1}^{N}\frac{p_i}{(\nu+p_i+2)}, \\
\mathbf{J}_{\sigma^2\nu} &= \frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(\frac{p_i}{(\nu+p_i+2)} - \frac{p_i}{\nu+p_i}\right),
\end{aligned}
$$

$$
\begin{aligned}
[\mathbf{J}_{\sigma^2\boldsymbol{\omega}}]_r &= \frac{\nu}{2\sigma^2}\sum_{i=1}^{N}\frac{1}{(\nu+p_i+2)}\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}), \\
\mathbf{J}_{\nu\nu} &= \frac{1}{4}\sum_{i=1}^{N}\left\{\psi(\frac{\nu}{2})-\psi(\frac{\nu+p_i}{2})-\frac{2(\nu+p_i)}{\nu(\nu+p_i+2)}-\frac{2}{\nu}+\frac{4}{\nu+p_i}\right\}, \\
[\mathbf{J}_{\nu\boldsymbol{\omega}}]_r &= -\sum_{i=1}^{N}\frac{1}{(\nu+p_i)(\nu+p_i+2)}\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}), \\
[\mathbf{J}_{\boldsymbol{\omega}\boldsymbol{\omega}}]_{rs} &= \frac{1}{2}\sum_{i=1}^{N}\frac{1}{(\nu+p_i+2)}\left\{(\nu+p_i)\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is})-\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir})\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{is})\right\},
\end{aligned}
$$

where $\dot{\boldsymbol{\Lambda}}_{ir} = \partial\boldsymbol{\Lambda}_i(\boldsymbol{\omega})/\partial\boldsymbol{\omega}_r$,for $r = 1,\ldots,k+1$ and $\psi(x) = d^2/dx^2\log(\boldsymbol{\Gamma}(x))$ denotes the trigamma function. We note that

$$
\begin{aligned}
\left.\frac{\partial\boldsymbol{\Lambda}_i}{\partial f_j}\right|_{\hat{\boldsymbol{\alpha}}_0} &= \boldsymbol{Z}_i\frac{\partial\mathbf{F}}{\partial f_j}\hat{\mathbf{F}}\boldsymbol{Z}_i{}' + \boldsymbol{Z}_i\hat{\mathbf{F}}'\frac{\partial\mathbf{F}}{\partial f_j}\hat{\mathbf{F}}\boldsymbol{Z}_i{}' = \hat{\boldsymbol{A}}_{ij}, \\
\left.\frac{\partial\boldsymbol{\Lambda}_i}{\partial\rho}\right|_{\hat{\boldsymbol{\alpha}}_0} &= \boldsymbol{L}_i' + \boldsymbol{L}_i = \hat{\boldsymbol{B}}_i.
\end{aligned}
$$

¿From the above, as $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}$ are the estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ in the null model, we can conclude that

$$
\begin{aligned}
\mathbf{J}_{\rho\rho}|_{\hat{\boldsymbol{\alpha}}_0} &= \frac{1}{2}\sum_{i=1}^{N}\frac{1}{\hat{\nu}+p_i+2}\left\{(\hat{\nu}+p_i)\mathrm{tr}(\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{B}}_i\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{B}}_i)-2[\mathrm{tr}(\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{B}}_i)]^2\right\}, \\
\mathbf{J}_{\sigma^2\rho}|_{\hat{\boldsymbol{\alpha}}_0} &= \frac{\hat{\nu}}{2\hat{\sigma^2}}\sum_{i=1}^{N}\frac{1}{\hat{\nu}+p_i+2}\mathrm{tr}(\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{B}}_i), \\
\left.[\mathbf{J}_{f\rho}]_j\right|_{\hat{\boldsymbol{\alpha}}_0} &= \frac{1}{2}\sum_{i=1}^{N}\frac{1}{\hat{\nu}+p_i+2}\left\{(\hat{\nu}+p_i)\mathrm{tr}(\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{A}}_{ij}\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{B}}_i)-2tr(\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{A}}_{ij})\mathrm{tr}(\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{B}}_i)\right\}, \\
\mathbf{J}_{\nu\rho}|_{\hat{\boldsymbol{\alpha}}_0} &= -\sum_{i=1}^{N}\frac{1}{(\hat{\nu}+p_i)(\hat{\nu}+p_i+2)}\mathrm{tr}(\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{B}}_i), \\
\mathbf{J}_{\sigma^2\sigma^2}|_{\hat{\boldsymbol{\alpha}}_0} &= \frac{\hat{\nu}}{2\hat{\sigma^4}}\sum_{i=1}^{N}\frac{p_i}{\hat{\nu}+p_i+2}, \\
\left.[\mathbf{J}_{\sigma^2 f}]_j\right|_{\hat{\boldsymbol{\alpha}}_0} &= \frac{\hat{\nu}}{2\hat{\sigma^2}}\sum_{i=1}^{N}\frac{1}{\hat{\nu}+p_i+2}\mathrm{tr}(\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{A}}_{ij}),
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{J}_{\sigma^2\nu}\big|_{\hat{\boldsymbol{\alpha}}_0} &= \frac{1}{2\hat{\sigma^2}}\sum_{i=1}^{N}\left(\frac{p_i}{\hat{\nu}+p_i+2}-\frac{p_i}{\hat{\nu}+p_i}\right), \\
[\mathbf{J}_{ff}]_{jr}\Big|_{\hat{\boldsymbol{\alpha}}_0} &= \frac{1}{2}\sum_{i=1}^{N}\frac{1}{\hat{\nu}+p_i+2}\left\{(\hat{\nu}+p_i)\mathrm{tr}(\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{A}}_{ij}\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{A}}_{ir})-2tr(\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{A}}_{ij})\mathrm{tr}(\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{A}}_{ir})\right\}, \\
[\mathbf{J}_{\nu f}]_{j}\Big|_{\hat{\boldsymbol{\alpha}}_0} &= -\sum_{i=1}^{N}\frac{1}{(\hat{\nu}+p_i)(\hat{\nu}+p_i+2)}\mathrm{tr}(\hat{\boldsymbol{\Lambda}}_i^{-1}\hat{\boldsymbol{A}}_{ij}), \\
\mathbf{J}_{\nu\nu}\big|_{\hat{\boldsymbol{\alpha}}_0} &= \frac{1}{4}\sum_{i=1}^{N}\left\{\psi(\frac{\hat{\nu}}{2})-\psi(\frac{\hat{\nu}+p_i}{2})-\frac{2(\hat{\nu}+p_i)}{\hat{\nu}(\hat{\nu}+p_i+2)}-\frac{2}{\hat{\nu}}+\frac{4}{\hat{\nu}+p_i}\right\}.
\end{aligned}
$$

Therefore, the form of $\mathbf{s}$ is concluded.

APPENDIX B: The first derivatives of the restricted log-likelihood

The following is the log of restricted likelihood $L_R(\theta_R|\boldsymbol{Y})$:

$$
\begin{aligned}
l_R &= \log L_R(\boldsymbol{\theta}_R|\boldsymbol{Y}) \\
&= \text{constant}-\frac{n}{2}\log(\sigma^2)-\frac{1}{2}\log\left|\sum_{i=1}^{N}\boldsymbol{X}_i'\mathbf{H}_i\boldsymbol{X}_i\right|+\sum_{i=1}^{N}\left\{\log\left(\Gamma(\frac{\nu+p_i}{2})\right)-\log\left(\Gamma(\frac{\nu}{2})\right)\right\} \\
&\quad -\frac{N}{2}\log(\nu)-\frac{1}{2}\sum_{i=1}^{N}\log|\boldsymbol{\Lambda}_i|-\frac{1}{2}\sum_{i=1}^{N}(\nu+p_i)\log\left(1+\frac{\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)}{\sigma^2\nu}\right).
\end{aligned}
$$

We take the first derivatives of l_R with respect to σ^2, ν, and $\boldsymbol{\omega}$, which can be expressed as:

$$
\begin{aligned}
\frac{\partial l_R}{\partial\sigma^2} &= -\frac{n}{2\sigma^2}-\frac{1}{2}\mathrm{tr}\left\{\left(\sum_{i=1}^{N}\boldsymbol{X}_i'\mathbf{H}_i\boldsymbol{X}_i\right)^{-1}\left(\sum_{i=1}^{N}\boldsymbol{X}_i'\frac{\partial\mathbf{H}_i}{\partial\sigma^2}\boldsymbol{X}_i\right)\right\} \\
&\quad +\frac{1}{2\sigma^2}\sum_{i=1}^{N}(\nu+p_i)\left(\frac{\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)}{\sigma^2\nu+\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)}\right), \\
\frac{\partial l_R}{\partial\nu} &= -\frac{1}{2}\mathrm{tr}\left\{\left(\sum_{i=1}^{N}\boldsymbol{X}_i'\mathbf{H}_i\boldsymbol{X}_i\right)^{-1}\left(\sum_{i=1}^{N}\boldsymbol{X}_i'\frac{\partial\mathbf{H}_i}{\partial\nu}\boldsymbol{X}_i\right)\right\} \\
&\quad +\frac{1}{2}\sum_{i=1}^{N}\left\{\phi\left(\frac{\nu+p_i}{2}\right)-\phi\left(\frac{\nu}{2}\right)-\frac{1}{\nu}-\log\left(1+\frac{\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)}{\sigma^2\nu}\right)\right.
\end{aligned}
$$

$$
+\frac{(\nu+p_i)}{\nu}\frac{\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)}{\sigma^2\nu+\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)}\Bigg\},
$$

$$
\begin{aligned}
\left[\frac{\partial l_R}{\partial \omega}\right]_r &= -\frac{1}{2}\mathrm{tr}\left\{\left(\sum_{i=1}^{N}\boldsymbol{X}_i'\mathbf{H}_i\boldsymbol{X}_i\right)^{-1}\left(\sum_{i=1}^{N}\boldsymbol{X}_i'\frac{\partial \mathbf{H}_i}{\partial \omega_r}\boldsymbol{X}_i\right)\right\} \\
&\quad -\frac{1}{2}\sum_{i=1}^{N}\left\{\mathrm{tr}(\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}^{-1})-(\nu+p_i)\left(\frac{\hat{\boldsymbol{e}}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}^{-1}\boldsymbol{\Lambda}_i^{-1}\hat{\boldsymbol{e}}_i}{\sigma^2\nu+\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)}\right)\right\},
\end{aligned}
$$

where

$$
\begin{aligned}
\frac{\partial \mathbf{H}_i}{\partial \sigma^2} &= -\nu(\nu+p_i)\left\{\frac{\boldsymbol{\Lambda}_i^{-1}}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)]^2}-\frac{4\boldsymbol{\Lambda}_i^{-1}\hat{\boldsymbol{e}}_i\hat{\boldsymbol{e}}_i'\boldsymbol{\Lambda}_i^{-1}}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)]^3}\right\}, \\
\frac{\partial \mathbf{H}_i}{\partial \nu} &= \left\{\frac{\boldsymbol{\Lambda}_i^{-1}}{\sigma^2\nu+\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)}-\frac{2\boldsymbol{\Lambda}_i^{-1}\hat{\boldsymbol{e}}_i\hat{\boldsymbol{e}}_i'\boldsymbol{\Lambda}_i^{-1}}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)]^2}\right\} \\
&\quad -\sigma^2(\nu+p_i)\left\{\frac{\boldsymbol{\Lambda}_i^{-1}}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)]^2}-\frac{4\boldsymbol{\Lambda}_i^{-1}\hat{\boldsymbol{e}}_i\hat{\boldsymbol{e}}_i'\boldsymbol{\Lambda}_i^{-1}}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)]^3}\right\}, \\
\frac{\partial \mathbf{H}_i}{\partial \omega_r} &= (\nu+p_i)\left\{\frac{-\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}[\sigma^2\nu+\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)]+\boldsymbol{\Lambda}_i^{-1}\hat{\boldsymbol{e}}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\hat{\boldsymbol{e}}_i}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)]^2}\right. \\
&\quad \left.+\frac{2\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\hat{\boldsymbol{e}}_i\hat{\boldsymbol{e}}_i'\boldsymbol{\Lambda}_i^{-1}+2\boldsymbol{\Lambda}_i^{-1}\hat{\boldsymbol{e}}_i\hat{\boldsymbol{e}}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)]^2}-\frac{4\boldsymbol{\Lambda}_i^{-1}\hat{\boldsymbol{e}}_i\hat{\boldsymbol{e}}_i'\boldsymbol{\Lambda}_i^{-1}\hat{\boldsymbol{e}}_i'\boldsymbol{\Lambda}_i^{-1}\dot{\boldsymbol{\Lambda}}_{ir}\boldsymbol{\Lambda}_i^{-1}\hat{\boldsymbol{e}}_i}{[\sigma^2\nu+\boldsymbol{\Delta}_i(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}),\boldsymbol{\Gamma},\rho)]^3}\right\}.
\end{aligned}
$$

APPENDIX C: The Bayes estimate $\hat{\boldsymbol{b}}_i(\boldsymbol{\theta})$ and the MSE matrix of $\hat{\boldsymbol{b}}_i(\boldsymbol{\theta})$

The conditional mean of $\boldsymbol{b}_i$ given $\boldsymbol{Y}_i$ is the minimum squared error (MSE) estimator of $\boldsymbol{b}_i$ simultaneously, which is $\hat{\boldsymbol{b}}_i(\boldsymbol{\theta})=\boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{\Lambda}_i^{-1}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta})$. Since $\boldsymbol{\Lambda}_i=\boldsymbol{Z}_i\boldsymbol{\Gamma}\boldsymbol{Z}_i'+\boldsymbol{C}_i$, using the matrix inversion formula, we can get

$$
\begin{aligned}
\boldsymbol{\Lambda}_i^{-1} &= \boldsymbol{C}_i^{-1}-\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i\left[\boldsymbol{\Gamma}^{-1}+\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i\right]^{-1}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1} \\
&= \boldsymbol{C}_i^{-1}-\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i\left[\boldsymbol{\Gamma}^{-1}+\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i\right]^{-1}(\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)(\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)^{-1}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1} \\
&= \boldsymbol{C}_i^{-1}-\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i\left[(\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)^{-1}\boldsymbol{\Gamma}^{-1}+\boldsymbol{I}_{p_i}\right]^{-1}(\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)^{-1}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1} \\
&= \boldsymbol{C}_i^{-1}-\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_i\boldsymbol{\Gamma}^{-1}+\boldsymbol{I}_{p_i}\right]^{-1}(\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)^{-1}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1} \\
&= \boldsymbol{C}_i^{-1}-\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma}\left[\boldsymbol{W}_i+\boldsymbol{\Gamma}\right]^{-1}(\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)^{-1}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}
\end{aligned}
$$

where $\boldsymbol{W}_i = (\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)^{-1}$. Then

$$\begin{aligned}
\hat{\boldsymbol{b}}_i(\boldsymbol{\theta}) &= \boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{\Lambda}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}) \\
&= \boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}) - \boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma}\left[\boldsymbol{W}_i + \boldsymbol{\Gamma}\right]^{-1}(\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)^{-1}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}) \\
&= \boldsymbol{\Gamma}(\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)(\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)^{-1}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}) - \boldsymbol{\Gamma}\boldsymbol{W}_i^{-1}\boldsymbol{\Gamma}\left[\boldsymbol{W}_i + \boldsymbol{\Gamma}\right]^{-1}(\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)^{-1} \\
&\quad \times\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta}) \\
&= \boldsymbol{\Gamma}\boldsymbol{W}_i^{-1}\boldsymbol{b}_i^*(\boldsymbol{\theta}) - \boldsymbol{\Gamma}\boldsymbol{W}_i^{-1}\boldsymbol{\Gamma}(\boldsymbol{W}_i + \boldsymbol{\Gamma})^{-1}\boldsymbol{b}_i^*(\boldsymbol{\theta}) \\
&= \boldsymbol{\Gamma}\boldsymbol{W}_i^{-1}\boldsymbol{b}_i^*(\boldsymbol{\theta}) - \boldsymbol{\Gamma}\boldsymbol{W}_i^{-1}(\boldsymbol{W}_i + \boldsymbol{\Gamma} - \boldsymbol{W}_i)(\boldsymbol{W}_i + \boldsymbol{\Gamma})^{-1}\boldsymbol{b}_i^*(\boldsymbol{\theta}) \\
&= \boldsymbol{\Gamma}\boldsymbol{W}_i^{-1}\left[\boldsymbol{I}_{p_i} - \boldsymbol{I}_{p_i} + \boldsymbol{W}_i(\boldsymbol{W}_i + \boldsymbol{\Gamma})^{-1}\right]\boldsymbol{b}_i^*(\boldsymbol{\theta}) \\
&= \boldsymbol{\Gamma}(\boldsymbol{W}_i + \boldsymbol{\Gamma})^{-1}\boldsymbol{b}_i^*(\boldsymbol{\theta}) \\
&= \boldsymbol{b}_i^*(\boldsymbol{\theta}) - \boldsymbol{b}_i^*(\boldsymbol{\theta}) + \boldsymbol{\Gamma}(\boldsymbol{W}_i + \boldsymbol{\Gamma})^{-1}\boldsymbol{b}_i^*(\boldsymbol{\theta}) \\
&= \boldsymbol{b}_i^*(\boldsymbol{\theta}) - \boldsymbol{b}_i^*(\boldsymbol{\theta}) + (\boldsymbol{W}_i + \boldsymbol{\Gamma} - \boldsymbol{W}_i)(\boldsymbol{W}_i + \boldsymbol{\Gamma})^{-1}\boldsymbol{b}_i^*(\boldsymbol{\theta}) \\
&= \boldsymbol{b}_i^*(\boldsymbol{\theta}) - \left[\boldsymbol{I}_{p_i} - \boldsymbol{I}_{p_i} + \boldsymbol{W}_i(\boldsymbol{W}_i + \boldsymbol{\Gamma})^{-1}\right]\boldsymbol{b}_i^*(\boldsymbol{\theta}) \\
&= \boldsymbol{b}_i^*(\boldsymbol{\theta}) - \boldsymbol{W}_i(\boldsymbol{W}_i + \boldsymbol{\Gamma})^{-1}\boldsymbol{b}_i^*(\boldsymbol{\theta}),
\end{aligned}$$

where $\boldsymbol{b}_i^*(\boldsymbol{\theta}) = (\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)^{-1}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}(\boldsymbol{Y}_i - \boldsymbol{X}_i\boldsymbol{\beta})$, and the equation (22) holds.

The following is the calculation of the MSE matrix of $\hat{\boldsymbol{b}}_i(\boldsymbol{\theta})$. To conclude the MSE matrix of $\hat{\boldsymbol{b}}_i(\boldsymbol{\theta})$, we need the covariance matrices of $\boldsymbol{b}_i$ and $\hat{\boldsymbol{b}}_i(\boldsymbol{\theta})$, which are $\frac{\nu}{\nu-2}\sigma^2\boldsymbol{\Gamma}$ and $\boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{\Lambda}_i^{-1}(\frac{\nu}{\nu-2}\sigma^2\boldsymbol{\Lambda}_i)\boldsymbol{\Lambda}_i^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma}$, respectively. Hence,

$$\begin{aligned}
& \mathrm{E}\left[(\hat{\boldsymbol{b}}_i(\boldsymbol{\theta}) - \boldsymbol{b}_i)(\hat{\boldsymbol{b}}_i(\boldsymbol{\theta}) - \boldsymbol{b}_i)'\right] \\
=\; & \mathrm{Cov}\left(\hat{\boldsymbol{b}}_i(\boldsymbol{\theta}) - \boldsymbol{b}_i, \hat{\boldsymbol{b}}_i(\boldsymbol{\theta}) - \boldsymbol{b}_i\right) \qquad \text{(Since } \mathrm{E}(\hat{\boldsymbol{b}}_i(\boldsymbol{\theta}) - \boldsymbol{b}_i) = \boldsymbol{0}) \\
=\; & \mathrm{Cov}\left(\boldsymbol{b}_i, \boldsymbol{b}_i\right) + \mathrm{Cov}\left(\hat{\boldsymbol{b}}_i(\boldsymbol{\theta}), \hat{\boldsymbol{b}}_i(\boldsymbol{\theta})\right) - 2\,\mathrm{Cov}\left(\hat{\boldsymbol{b}}_i(\boldsymbol{\theta}), \boldsymbol{b}_i\right) \\
=\; & \frac{\nu}{\nu-2}\sigma^2\boldsymbol{\Gamma} + \boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{\Lambda}_i^{-1}\left(\frac{\nu}{\nu-2}\sigma^2\boldsymbol{\Lambda}_i\right)\boldsymbol{\Lambda}_i^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma} - 2\,\boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{Z}_i\left(\frac{\nu}{\nu-2}\sigma^2\right)\boldsymbol{\Gamma} \\
=\; & \frac{\nu}{\nu-2}\sigma^2\left[\boldsymbol{\Gamma} - \boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{\Lambda}_i^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma}\right] \\
=\; & \frac{\nu}{\nu-2}\sigma^2\left[\boldsymbol{\Gamma} - \boldsymbol{\Gamma}\boldsymbol{Z}_i'\left[\boldsymbol{C}_i^{-1} - \boldsymbol{C}_i^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma}\left[\boldsymbol{W}_i + \boldsymbol{\Gamma}\right]^{-1}(\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i)^{-1}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\right]\boldsymbol{Z}_i\boldsymbol{\Gamma}\right] \\
=\; & \frac{\nu}{\nu-2}\sigma^2\left[\boldsymbol{\Gamma} - \boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma} + \boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma}\left(\boldsymbol{W}_i + \boldsymbol{\Gamma}\right)^{-1}\boldsymbol{\Gamma}\right] \\
=\; & \frac{\nu}{\nu-2}\sigma^2\left[\boldsymbol{\Gamma} - \boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{C}_i^{-1}\boldsymbol{Z}_i\left(\boldsymbol{I}_{p_i} - \boldsymbol{\Gamma}\left(\boldsymbol{W}_i + \boldsymbol{\Gamma}\right)^{-1}\right)\boldsymbol{\Gamma}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\nu}{\nu-2}\sigma^2\left[\boldsymbol{\Gamma}-\boldsymbol{\Gamma}\boldsymbol{W}_i^{-1}\left(\boldsymbol{I}_{p_i}-\boldsymbol{\Gamma}\left(\boldsymbol{W}_i+\boldsymbol{\Gamma}\right)^{-1}\right)\boldsymbol{\Gamma}\right]\\
&= \frac{\nu}{\nu-2}\sigma^2\left[\boldsymbol{\Gamma}-\boldsymbol{\Gamma}\boldsymbol{W}_i^{-1}\left((\boldsymbol{W}_i+\boldsymbol{\Gamma})(\boldsymbol{W}_i+\boldsymbol{\Gamma})^{-1}-\boldsymbol{\Gamma}\left(\boldsymbol{W}_i+\boldsymbol{\Gamma}\right)^{-1}\right)\boldsymbol{\Gamma}\right]\\
&= \frac{\nu}{\nu-2}\sigma^2\left[\boldsymbol{\Gamma}-\boldsymbol{\Gamma}\boldsymbol{W}_i^{-1}\left(\boldsymbol{W}_i(\boldsymbol{W}_i+\boldsymbol{\Gamma})^{-1}\right)\boldsymbol{\Gamma}\right]\\
&= \frac{\nu}{\nu-2}\sigma^2\left[\boldsymbol{\Gamma}\left(\boldsymbol{I}_{p_i}-(\boldsymbol{W}_i+\boldsymbol{\Gamma})^{-1}\boldsymbol{\Gamma}\right)\right]\\
&= \frac{\nu}{\nu-2}\sigma^2\left[\boldsymbol{\Gamma}\left(\boldsymbol{I}_{p_i}-(\boldsymbol{W}_i+\boldsymbol{\Gamma})^{-1}(\boldsymbol{W}_i+\boldsymbol{\Gamma}-\boldsymbol{W}_i)\right)\right]\\
&= \frac{\nu}{\nu-2}\sigma^2\left[\boldsymbol{\Gamma}(\boldsymbol{W}_i+\boldsymbol{\Gamma})^{-1}\right]\boldsymbol{W}_i\\
&= \frac{\nu}{\nu-2}\sigma^2\left[(\boldsymbol{W}_i+\boldsymbol{\Gamma}-\boldsymbol{W}_i)(\boldsymbol{W}_i+\boldsymbol{\Gamma})^{-1}\right]\boldsymbol{W}_i\\
&= \frac{\nu}{\nu-2}\sigma^2\left[\boldsymbol{I}_{p_i}-\boldsymbol{W}_i(\boldsymbol{W}_i+\boldsymbol{\Gamma})^{-1}\right]\boldsymbol{W}_i\\
&= \frac{\nu}{\nu-2}\sigma^2\left[\boldsymbol{W}_i-\boldsymbol{W}_i(\boldsymbol{W}_i+\boldsymbol{\Gamma})^{-1}\boldsymbol{W}_i\right].
\end{aligned}
$$

APPENDIX D: The distribution of $\boldsymbol{y}_i$ given $\boldsymbol{Y}_i$

Since

$$
\begin{pmatrix}\boldsymbol{Y}_i\\ \boldsymbol{y}_i\end{pmatrix}\sim t_{p_i+q}\left(\boldsymbol{X}_i^*\boldsymbol{\beta},\sigma^2\boldsymbol{\Omega}_i,\nu\right),
$$

we can get

$$
\left.\begin{pmatrix}\boldsymbol{Y}_i\\ \boldsymbol{y}_i\end{pmatrix}\right|\tau_i\sim N_{p_i+q}\left(\boldsymbol{X}_i^*\boldsymbol{\beta},\frac{\sigma^2}{\tau_i}\boldsymbol{\Omega}_i\right).
$$

The conditional distribution of $\boldsymbol{y}_i$ given $\boldsymbol{Y}_i$, τ_i is normal with mean $\mu_{i,2\cdot 1}=\boldsymbol{x}_i\boldsymbol{\beta}+\boldsymbol{\Omega}_{i,22}\boldsymbol{\Omega}_{i,11}^{-1}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta})$ and covariance matrix $\boldsymbol{\Sigma}_{i,22.1}^*=\frac{\sigma^2}{\tau_i}(\boldsymbol{\Omega}_{i,22}-\boldsymbol{\Omega}_{i,21}\boldsymbol{\Omega}_{i,11}^{-1}\boldsymbol{\Omega}_{i,12})$, and note that $\boldsymbol{\Sigma}_{i,22\cdot 1}=\sigma^2(\boldsymbol{\Omega}_{i,22}-\boldsymbol{\Omega}_{i,21}\boldsymbol{\Omega}_{i,11}^{-1}\boldsymbol{\Omega}_{i,12})$. By Bayesian approach, we have

$$
\begin{aligned}
& f(\boldsymbol{y}_i|\boldsymbol{Y}_i)\\
\propto & \int f(\boldsymbol{y}_i|\boldsymbol{Y}_i,\tau_i)f(\boldsymbol{Y}_i|\tau_i)g(\tau_i)d\tau_i\\
= & (2\pi)^{-\frac{q}{2}}\left|\frac{1}{\tau_i}\boldsymbol{\Sigma}_{i,22\cdot 1}\right|^{-\frac{1}{2}}\exp\left\{-\frac{(\boldsymbol{y}_i-\mu_{i,2\cdot 1})'\left(\frac{1}{\tau_i}\boldsymbol{\Sigma}_{i,22\cdot 1}\right)^{-1}(\boldsymbol{y}_i-\mu_{i,2\cdot 1})}{2}\right\}
\end{aligned}
$$

$$
\begin{aligned}
& \times(2\pi)^{-\frac{p_i}{2}} \left|\frac{\sigma^2}{\tau_i}\boldsymbol{\Omega}_{i,11}\right|^{-\frac{1}{2}} \exp\left\{-\frac{(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta})'\left(\frac{\sigma^2}{\tau_i}\boldsymbol{\Omega}_{i,11}\right)^{-1}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta})}{2}\right\} \\
& \times\frac{\left(\frac{1}{2}\right)^{\frac{\nu}{2}}\nu^{\frac{\nu}{2}}\tau_i^{\frac{\nu}{2}-1}}{\boldsymbol{\Gamma}(\frac{\nu}{2})}\exp\left(-\frac{\nu}{2}\tau_i\right)d\tau_i \\
= & \ |\boldsymbol{\Sigma}_{i,22\cdot 1}|^{-\frac{1}{2}}\left|\sigma^2\boldsymbol{\Omega}_{i,11}\right|^{-\frac{1}{2}}\frac{(2\pi)^{-\frac{p_i+q}{2}}\left(\frac{1}{2}\right)^{\frac{\nu}{2}}\nu^{\frac{\nu}{2}}}{\boldsymbol{\Gamma}(\frac{\nu}{2})} \\
& \times\int\tau_i^{\frac{p_i+q+\nu}{2}-1}\exp\left\{-\frac{(\boldsymbol{y}_i-\mu_{i,2\cdot 1})'(\boldsymbol{\Sigma}_{i,22\cdot 1})^{-1}(\boldsymbol{y}_i-\mu_{i,2\cdot 1})+(\omega_i(\nu+p_i)-\nu)+\nu}{2}\tau_i\right\}d\tau_i \\
= & \ |\boldsymbol{\Sigma}_{i,22\cdot 1}|^{-\frac{1}{2}}\left|\sigma^2\boldsymbol{\Omega}_{i,11}\right|^{-\frac{1}{2}}\frac{(2\pi)^{-\frac{p_i+q}{2}}\left(\frac{1}{2}\right)^{\frac{\nu}{2}}\nu^{\frac{\nu}{2}}}{\boldsymbol{\Gamma}(\frac{\nu}{2})} \\
& \times\boldsymbol{\Gamma}(\frac{p_i+q+\nu}{2})\times\left(\frac{(\boldsymbol{y}_i-\mu_{i,2\cdot 1})'(\boldsymbol{\Sigma}_{i,22\cdot 1})^{-1}(\boldsymbol{y}_i-\mu_{i,2\cdot 1})+\omega_i(\nu+p_i)}{2}\right)^{-\frac{p_i+q+\nu}{2}} \\
\propto & \ \left(\frac{(\boldsymbol{y}_i-\mu_{i,2\cdot 1})'(\omega_i\boldsymbol{\Sigma}_{i,22\cdot 1})^{-1}(\boldsymbol{y}_i-\mu_{i,2\cdot 1})}{\nu+p_i}+1\right)^{-\frac{p_i+q+\nu}{2}}.
\end{aligned}
$$

Thus we have

$$
\boldsymbol{y}_i|\boldsymbol{Y}_i \sim t_q\left(\boldsymbol{\mu}_{i,2\cdot 1},\omega_i\boldsymbol{\Sigma}_{i,22\cdot 1},\nu+p_i\right),
$$

where

$$
\begin{aligned}
\boldsymbol{\mu}_{i,2\cdot 1} &= \boldsymbol{x}_i\boldsymbol{\beta}+\boldsymbol{\Omega}_{i21}\boldsymbol{\Omega}_{i11}^{-1}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta}), \\
\boldsymbol{\Sigma}_{i,22\cdot 1} &= \sigma^2\left(\boldsymbol{\Omega}_{i22}-\boldsymbol{\Omega}_{i21}\boldsymbol{\Omega}_{i11}^{-1}\boldsymbol{\Omega}_{i12}\right), \\
\omega_i &= \frac{\nu+(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta})'(\sigma^2\boldsymbol{\Omega}_{i11})^{-1}(\boldsymbol{Y}_i-\boldsymbol{X}_i\boldsymbol{\beta})}{\nu+p_i}.
\end{aligned}
$$

APPENDIX **E**: The equation of (25)

¿From Appendix **C**, we know that $\boldsymbol{\Sigma}_{i,22\cdot 1}=\sigma^2\left(\boldsymbol{\Omega}_{i22}-\boldsymbol{\Omega}_{i21}\boldsymbol{\Omega}_{i11}^{-1}\boldsymbol{\Omega}_{i12}\right)$, where

$$
\boldsymbol{\Omega}_i=\boldsymbol{Z}_i^*\boldsymbol{\Gamma}\boldsymbol{Z}_i^{*\prime}+\boldsymbol{C}_i^*=\begin{bmatrix}\boldsymbol{\Omega}_{i11} & \boldsymbol{\Omega}_{i12}\\ \boldsymbol{\Omega}_{i21} & \boldsymbol{\Omega}_{i22}\end{bmatrix}=\begin{bmatrix}\boldsymbol{Z}_i\boldsymbol{\Gamma}\boldsymbol{Z}_i'+\boldsymbol{C}_{11} & \boldsymbol{Z}_i\boldsymbol{\Gamma}\boldsymbol{z}_i'+\boldsymbol{C}_{12}\\ \boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{Z}_i'+\boldsymbol{C}_{21} & \boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{z}_i'+\boldsymbol{C}_{22}\end{bmatrix}.
$$

Using the matrix inversion formula, we get

$$
\boldsymbol{\Omega}_{i11}^{-1} = (\boldsymbol{Z}_i\boldsymbol{\Gamma}\boldsymbol{Z}_i'+\boldsymbol{C}_{11})^{-1}
$$

$$
\begin{aligned}
&= \boldsymbol{C}_{11}^{-1} - \boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{\Gamma}^{-1} + \boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\right]^{-1}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1} \\
&= \boldsymbol{C}_{11}^{-1} - \boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[(\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i)^{-1} - (\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i)^{-1}(\boldsymbol{\Gamma} + (\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i)^{-1})^{-1}(\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i)^{-1}\right] \\
&\qquad\qquad \times \boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1} \\
&= \boldsymbol{C}_{11}^{-1} - \boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}\left(\boldsymbol{\Gamma} + \boldsymbol{W}_{11}\right)^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1},
\end{aligned}
$$

where $\boldsymbol{W}_{11} = (\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i)^{-1}$. The following is the calculation of $\boldsymbol{\Omega}_{i21}\boldsymbol{\Omega}_{i11}^{-1}\boldsymbol{\Omega}_{i12}$:

$$
\begin{aligned}
& \boldsymbol{\Omega}_{i21}\boldsymbol{\Omega}_{i11}^{-1}\boldsymbol{\Omega}_{i12} \\
={}& \left[\boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{Z}_i' + \boldsymbol{C}_{21}\right]\left[\boldsymbol{C}_{11}^{-1} - \boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}\left(\boldsymbol{\Gamma} + \boldsymbol{W}_{11}\right)^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\right]\left[\boldsymbol{Z}_i\boldsymbol{\Gamma}\boldsymbol{z}_i' + \boldsymbol{C}_{12}\right] \\
={}& \boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma}\boldsymbol{z}_i' + \boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} \\
& -\boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma}\boldsymbol{z}_i' \\
& -\boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} \\
& +\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma}\boldsymbol{z}_i' + \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} \\
& -\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma}\boldsymbol{z}_i' \\
& -\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} \\
={}& \boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{W}_{11}^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' + \boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} \\
& -\boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{W}_{11}^{-1}\boldsymbol{W}_{11}\boldsymbol{W}_{11}^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' + \boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{W}_{11}^{-1}\boldsymbol{W}_{11}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\boldsymbol{W}_{11}^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' \\
& -\boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{W}_{11}^{-1}\boldsymbol{W}_{11}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} + \boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{W}_{11}^{-1}\boldsymbol{W}_{11}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} \\
& +\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma}\boldsymbol{z}_i' + \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} \\
& -\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\boldsymbol{W}_{11}\boldsymbol{W}_{11}^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' + \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\boldsymbol{W}_{11}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\boldsymbol{W}_{11}^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' \\
& -\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} \\
={}& \boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{W}_{11}^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' + \boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} - \boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{W}_{11}^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' + \boldsymbol{z}_i\boldsymbol{\Gamma}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' \\
& -\boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} + \boldsymbol{z}_i\boldsymbol{\Gamma}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} + \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma}\boldsymbol{z}_i' + \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} \\
& -\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\boldsymbol{\Gamma}\boldsymbol{z}_i' + \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\boldsymbol{W}_{11}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' \\
& -\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} \\
={}& \boldsymbol{z}_i\boldsymbol{\Gamma}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' + \boldsymbol{z}_i\boldsymbol{\Gamma}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} + \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} \\
& +\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\boldsymbol{W}_{11}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' \\
-{}& \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma} + \boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}.
\end{aligned}
$$

Then

$$
\begin{aligned}
& \boldsymbol{\Omega}_{i22} - \boldsymbol{\Omega}_{i21}\boldsymbol{\Omega}_{i11}^{-1}\boldsymbol{\Omega}_{i12} \\
= \; & \boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{z}_i' + \boldsymbol{C}_{22} - \boldsymbol{z}_i\boldsymbol{\Gamma}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' - \boldsymbol{z}_i\boldsymbol{\Gamma}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} - \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} \\
& -\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' \\
& +\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} \\
= \; & \boldsymbol{C}_{22} - \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} + \boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{z}_i' - \boldsymbol{z}_i\boldsymbol{\Gamma}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' \\
& -\boldsymbol{z}_i\boldsymbol{\Gamma}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} - \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' \\
& +\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}.
\end{aligned}
$$

We note that

$$
\begin{aligned}
& \boldsymbol{z}_i\boldsymbol{\Gamma}\boldsymbol{z}_i' - \boldsymbol{z}_i\boldsymbol{\Gamma}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' = \boldsymbol{z}_i\left[\boldsymbol{\Gamma} - \boldsymbol{\Gamma}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma}\right]\boldsymbol{z}_i' \\
= \; & \boldsymbol{z}_i\left[\boldsymbol{\Gamma} - \boldsymbol{\Gamma}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma} + \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma}\right]\boldsymbol{z}_i' \\
= \; & \boldsymbol{z}_i\left[\boldsymbol{\Gamma} - (\boldsymbol{\Gamma}+\boldsymbol{W}_{11})(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma} + \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma}\right]\boldsymbol{z}_i' \\
= \; & \boldsymbol{z}_i\left[\boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11} + \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma}\right]\boldsymbol{z}_i' \\
= \; & \boldsymbol{z}_i\left[\boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11}) - \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{z}_i' \\
= \; & \boldsymbol{z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{z}_i',
\end{aligned}
$$

and

$$
\begin{aligned}
\boldsymbol{z}_i\boldsymbol{\Gamma}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} &= \boldsymbol{z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right](\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i)' \\
\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{\Gamma}\boldsymbol{z}_i' &= \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{z}_i'.
\end{aligned}
$$

Therefore, we conclude that

$$
\begin{aligned}
& \boldsymbol{\Omega}_{i22} - \boldsymbol{\Omega}_{i21}\boldsymbol{\Omega}_{i11}^{-1}\boldsymbol{\Omega}_{i12} \\
= \; & \boldsymbol{C}_{22} - \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} + \boldsymbol{z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{z}_i' \\
& -\boldsymbol{z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right](\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i)' \\
& -\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{z}_i' \\
& +\boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right]\boldsymbol{Z}_i'\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12} \\
= \; & \boldsymbol{C}_{22\cdot 1} + (\boldsymbol{z}_i - \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i)\left[\boldsymbol{W}_{11} - \boldsymbol{W}_{11}(\boldsymbol{\Gamma}+\boldsymbol{W}_{11})^{-1}\boldsymbol{W}_{11}\right](\boldsymbol{z}_i - \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{Z}_i)'.
\end{aligned}
$$