Chapter 4

Further Longitudinal Data Models

4.1 Growth Curve Model

4.1.1 Estimation and Prediction in One-Way Anova Model

– Consider the repeated measures type of random vector

$$\mathbf{y}_i = \begin{pmatrix} y_{it_1} \\ y_{it_2} \\ \vdots \\ y_{it_m} \end{pmatrix}$$

with assumption of the normally distributed

$$\mathbf{y}_i \sim N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}). \tag{4.1}$$

– Suppose that the expectation vector $\boldsymbol{\mu}_i$ depends on the category value j of the explanatory variable X through the growth curve model

$$\boldsymbol{\mu}_i = \begin{pmatrix} \mu_{i1} \\ \mu_{i2} \\ \mu_{i3} \\ \vdots \\ \mu_{im} \end{pmatrix} = \begin{pmatrix} 1 & t_1 & t_1^2 & \dots & t_1^r \\ 1 & t_2 & t_2^2 & \dots & t_2^r \\ 1 & t_3 & t_3^2 & \dots & t_3^r \\ & \vdots & & & \\ 1 & t_m & t_m^2 & \dots & t_m^r \end{pmatrix} \begin{pmatrix} \theta_{0_j} \\ \theta_{1_j} \\ \theta_{2_j} \\ \vdots \\ \theta_{r_j} \end{pmatrix} = (\mathbf{1} : \mathbf{t} : \mathbf{t}^2 : \dots : \mathbf{t}^r) \begin{pmatrix} \theta_{0_j} \\ \theta_{1_j} \\ \theta_{2_j} \\ \vdots \\ \theta_{r_j} \end{pmatrix} = \mathbf{T}\boldsymbol{\theta}_j = \mathbf{T}\boldsymbol{\Theta}'\mathbf{x}_i. \tag{4.2}$$

– In the growth curve model, a dummy vector $\mathbf{x}_i = (0, 0, \dots, 0, 1, 0, \dots, 0)'$ is denoting to which subpopulation j the sampling i is belonging. Furthermore,

$$\mathbf{\Theta} = \begin{pmatrix} \boldsymbol{\theta}_1' \\ \boldsymbol{\theta}_2' \\ \vdots \\ \boldsymbol{\theta}_k' \end{pmatrix}.$$

– The growth model can written as

$$\mathbf{y}_i = \mathbf{T}\mathbf{\Theta}'\mathbf{x}_i + \boldsymbol{\varepsilon}_i, \qquad \boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \mathbf{\Sigma}). \tag{4.3}$$

– Hence

$$\mathbf{Y}' = \mathbf{T}\mathbf{\Theta}'\mathbf{X}' + \mathbf{E}', \tag{4.4}$$

and

$$\mathbf{Y} = \mathbf{X}\mathbf{\Theta}\mathbf{T}' + \mathbf{E}, \tag{4.5}$$

where, in one-way anova situation,

$$\mathbf{X} = \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \dots & \mathbf{0} \\ & \vdots & & & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{1} \end{pmatrix}, \qquad \mathbf{E} = \begin{pmatrix} \boldsymbol{\varepsilon}_1' \\ \boldsymbol{\varepsilon}_2' \\ \vdots \\ \boldsymbol{\varepsilon}_n' \end{pmatrix}.$$

– Estimation of the parametric matrix $\boldsymbol{\Theta}$ can be based on multivariate linear model theory or on estimation theory of general linear models. We consider here multivariate linear model theory.

– Post multiplying the model equation (4.5) by the matrix $\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}$ gives

$$\begin{aligned}
\mathbf{Y}\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1} &= \mathbf{X}\boldsymbol{\Theta}\mathbf{T}'\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1} + \mathbf{E}\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1} \\
\mathbf{Y}\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1} &= \mathbf{X}\boldsymbol{\Theta} + \mathbf{E}\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1} \\
\mathbf{Y}_* &= \mathbf{X}\boldsymbol{\Theta} + \mathbf{E}_* \qquad (4.6)
\end{aligned}$$

– Hence

$$\widehat{\boldsymbol{\Theta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}, \qquad (4.7)$$

$$\hat{\boldsymbol{\mu}}_i = \mathbf{T}\widehat{\boldsymbol{\Theta}}'\mathbf{x}_i = \mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{Y}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i. \qquad (4.8)$$

– For the covariance matrix $\mathrm{Cov}(\hat{\boldsymbol{\mu}}_i)$ it holds

$$\begin{aligned}
\mathrm{Cov}(\hat{\boldsymbol{\mu}}_i) &= \mathrm{Cov}\left[(\mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \otimes \mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}')\,\mathrm{vec}(\mathbf{Y}')\right] \\
&= (\mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \otimes \mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}')(\mathbf{I} \otimes \boldsymbol{\Sigma})(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i \otimes \mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}') \\
&= (\mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i)\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\boldsymbol{\Sigma}\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'. \qquad (4.9)
\end{aligned}$$

– Note also that

$$\widehat{\mathbf{E}} = \mathbf{Y} - \mathbf{X}\widehat{\mathbf{\Theta}}\mathbf{T}', \tag{4.10}$$

$$\widehat{\mathbf{\Sigma}} = \frac{\widehat{\mathbf{E}}'\widehat{\mathbf{E}}}{n - \mathrm{rank}(\mathbf{X})}. \tag{4.11}$$

– The best empirical predictor for the new random vector $\mathbf{y}_f$ is

$$\widehat{\mathrm{BP}}(\mathbf{y}_f|\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n) = \widehat{\mathrm{BP}}(\mathbf{y}_f|\mathbf{Y}) = \hat{\boldsymbol{\mu}}_f = \mathbf{T}\widehat{\mathbf{\Theta}}'\mathbf{x}_f. \tag{4.12}$$

– The covariance matrix of the prediction error $\mathbf{y}_f - \widehat{\mathrm{BP}}(\mathbf{y}_f|\mathbf{Y})$ is

$$\begin{aligned}\mathrm{Cov}(\mathbf{y}_f - \widehat{\mathrm{BP}}(\mathbf{y}_f|\mathbf{Y})) &= \mathrm{Cov}(\mathbf{y}_f) + \mathrm{Cov}(\widehat{\mathrm{BP}}(\mathbf{y}_f|\mathbf{Y})) \\ &= \mathbf{\Sigma} + (\mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i)\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{\Sigma}\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'.\end{aligned} \tag{4.13}$$

– Simultaneous $100(1-\alpha)\%$ prediction intervals for each element z_{fk} of the random vector $\mathbf{y}_f$ are

$$\left[\widehat{\mathrm{BP}}(\mathbf{y}_f|\mathbf{Y}) \pm z_{\alpha/2}\,\mathbf{diag}(\widehat{\mathbf{\Sigma}} + (\mathbf{x}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i)\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\widehat{\mathbf{\Sigma}}\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}')^{1/2}\right]. \tag{4.14}$$

– Consider the prediction in a partitioned random vector $\mathbf{y}_f$ situation

$$\mathbf{y}_f = \begin{pmatrix} \mathbf{y}_{f_1} \\ \mathbf{y}_{f_2} \end{pmatrix} = \begin{pmatrix} \mathbf{T}_1 \\ \mathbf{T}_2 \end{pmatrix} \mathbf{\Theta}' \mathbf{x}_f + \begin{pmatrix} \boldsymbol{\varepsilon}_{f_1} \\ \boldsymbol{\varepsilon}_{f_2} \end{pmatrix}. \tag{4.15}$$

– Let the problem be the prediction of $\mathbf{y}_{f_2}$ given $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ and $\mathbf{y}_{f_1}$.

– Under normality

$$\begin{pmatrix} \mathbf{y}_{f_1} \\ \mathbf{y}_{f_2 2} \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{T}_1 \mathbf{\Theta}' \mathbf{x}_f \\ \mathbf{T}_2 \mathbf{\Theta}' \mathbf{x}_f \end{pmatrix}, \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix} \right), \tag{4.16}$$

the best empirical predictor is

$$\widehat{\mathrm{BP}}(\mathbf{y}_{f_2} | \mathbf{Y}, \mathbf{y}_{f_1}) = \mathbf{T}_2 \widehat{\mathbf{\Theta}}' \mathbf{x}_f + \widehat{\mathbf{\Sigma}}_{21} \widehat{\mathbf{\Sigma}}_{11}^{-1} (\mathbf{y}_{f_1} - \mathbf{T}_1 \widehat{\mathbf{\Theta}}' \mathbf{x}_f) \tag{4.17}$$

with the (asymptotic) estimated covariance matrix of the prediction error being

$$\widehat{\mathrm{Cov}}(\mathbf{y}_{f_2} - \widehat{\mathrm{BP}}(\mathbf{y}_{f_2} | \mathbf{Y}, \mathbf{y}_{f_1})) = \widehat{\mathbf{\Sigma}}_{22} - \widehat{\mathbf{\Sigma}}_{21} \widehat{\mathbf{\Sigma}}_{11}^{-1} \widehat{\mathbf{\Sigma}}_{12}. \tag{4.18}$$

– Simultaneous $100(1-\alpha)\%$ prediction intervals for each element of the random vector $\mathbf{y}_{f_2}$ are

$$\left[\widehat{\mathrm{BP}}(\mathbf{y}_{f_2} | \mathbf{Y}, \mathbf{y}_{f_1}) \pm z_{\alpha/2} \, \mathbf{diag}(\widehat{\mathbf{\Sigma}}_{22} - \widehat{\mathbf{\Sigma}}_{21} \widehat{\mathbf{\Sigma}}_{11}^{-1} \widehat{\mathbf{\Sigma}}_{12})^{1/2} \right]. \tag{4.19}$$

4.1.2 Testing in One-Way Anova Growth Model

– Let us consider testing the significance of the explanatory variable X. Under H_0 hypothesis, the growth model is

$$\mathbf{y}_i = \mathbf{T}\boldsymbol{\theta}_{H_0} + \boldsymbol{\varepsilon}_i, \qquad \boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \boldsymbol{\Sigma}). \tag{4.20}$$

– Testing the significance of the explanatory variable X is performed by comparing following two multivariate linear models

$$\begin{aligned} H_0 &: \mathbf{Y}\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1} = \mathbf{1}\boldsymbol{\theta}'_{H_0} + \mathbf{E}\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}, \\ H_1 &: \mathbf{Y}\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1} = \mathbf{X}\boldsymbol{\Theta} + \mathbf{E}\mathbf{T}(\mathbf{T}'\mathbf{T})^{-1}. \end{aligned}$$

– Let the matrix $\mathbf{T}$ be partitioned as $\mathbf{T} = (\mathbf{T}_1 : \mathbf{T}_2)$. Then testing the hypotheses

$$\begin{aligned} H_0 &: \boldsymbol{\mu}_i = \mathbf{T}_1\boldsymbol{\Theta}'_1\mathbf{x}_i \\ H_1 &: \boldsymbol{\mu}_i = \mathbf{T}_1\boldsymbol{\Theta}'_1\mathbf{x}_i + \mathbf{T}_2\boldsymbol{\Theta}'_2\mathbf{x}_i, \end{aligned}$$

can be performed by comparing following two growth curve models

$$\begin{aligned} \mathcal{M}_1 &: \mathbf{Y} = \mathbf{X}\boldsymbol{\Theta}_1\mathbf{T}'_1 + \mathbf{E}, \\ \mathcal{M}_2 &: \mathbf{Y} = \mathbf{X}\boldsymbol{\Theta}\mathbf{T}' + \mathbf{E}, \end{aligned}$$

then testing the hypotheses (with possible p-value obtained by a randomization test)

$$\begin{aligned} H_0 &: \operatorname{vec}(Y) = (\mathbf{T}_1 \otimes \mathbf{X})\operatorname{vec}(\boldsymbol{\Theta}_1) + \operatorname{vec}(\mathbf{E}), \\ H_1 &: \operatorname{vec}(Y) = (\mathbf{T} \otimes \mathbf{X})\operatorname{vec}(\boldsymbol{\Theta}) + \operatorname{vec}(\mathbf{E}). \end{aligned}$$

Example 4.1.

Consider the dataset knee2.txt:

	N	Th	Age	Sex	R1	R2	R3	R4
1	1	1	28	1	4	4	4	4
2	2	1	32	1	4	4	4	4
3	3	1	41	1	3	3	3	3
4	4	2	21	1	4	3	3	2
5	5	2	34	1	4	3	3	2
6	6	1	24	1	3	3	3	2
7	7	2	28	1	4	3	3	2
8	8	2	40	1	3	2	2	2

```
A data frame with 127 observations on the following 8 variables.
N- Patient's number
Th - Therapy ( placebo = 1, treatment = 2)
Age - Age in years
Sex -Gender (male = 0, female = 1)
R1 -Pain before treatment (no pain = 1, severe pain = 5)
R2 -Pain after three days of treatment
R3 -Pain after seven days of treatment
R4 -Pain after ten days of treatment
```

.

```
In a clinical study n=127 patients with sport related injuries have been treated with
two different therapies (chosen by random design).
After 3,7 and 10 days of treatment the pain occuring during knee movement was observed.
```

Model the data $\mathbf{y}_i \sim N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$ by the growth curve model

$$\mathcal{M}: \quad \begin{pmatrix} \mu_{it_1} \\ \mu_{it_2} \\ \mu_{it_3} \\ \mu_{it_4} \end{pmatrix} = \begin{pmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \\ 1 & t_4 & t_4^2 \end{pmatrix} \begin{pmatrix} \theta_{0_j} \\ \theta_{1_j} \\ \theta_{2_j} \end{pmatrix} = (\mathbf{1} : \mathbf{t} : \mathbf{t}^2) \begin{pmatrix} \theta_{0_j} \\ \theta_{1_j} \\ \theta_{2_j} \end{pmatrix} = \mathbf{T}\boldsymbol{\theta}_j,$$

where the time values are $\mathbf{t} = (t_1, t_2, t_3, t_4)' = (0, 3, 7, 10)'$ and where $X = \mathsf{Th}$.

(a) The model $\mathcal{M}$ can be written for the whole data as $\mathbf{Y} = \mathbf{X}\boldsymbol{\Theta}\mathbf{T}' + \mathbf{E}$, where, for each row of $\mathbf{E}$, it is assumed $\boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, and $\boldsymbol{\Theta} = \begin{pmatrix} \boldsymbol{\theta}_1' \\ \boldsymbol{\theta}_2' \end{pmatrix}$. Find an estimate $\widehat{\boldsymbol{\Theta}}$ for the parameters $\boldsymbol{\Theta}$.

(b) Predict the value of the new observation $\mathbf{y}_f$ in a situation of person getting real treatment. Create also 80 % simultaneous (asymptotic) prediction intervals for elements of the random vector $\mathbf{y}_f$.

(c) Test at 5% significance level, is the explanatory variable $X = \mathsf{Th}$ statistically significant variable in the model.

(d) Test the hypotheses

$$H_0 : \boldsymbol{\mu}_i = (\mathbf{1} : \mathbf{t}) \begin{pmatrix} \theta_{0_j} \\ \theta_{1_j} \end{pmatrix}, \qquad H_1 : \boldsymbol{\mu}_i = (\mathbf{1} : \mathbf{t} : \mathbf{t}^2) \begin{pmatrix} \theta_{0_j} \\ \theta_{1_j} \\ \theta_{2_j} \end{pmatrix}.$$

4.2 Multivariate Longitudinal Linear Mixed Model

4.2.1 Estimation in Mixed Model

– Consider the multivariate longitudinal linear mixed effects model

$$\mathbf{Y}_i = \mathbf{X}_i\mathbf{B} + \mathbf{Z}_i\mathbf{R}_i + \mathbf{E}_i, \qquad \mathrm{vec}(\mathbf{E}_i) \sim N(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I}), \tag{4.21}$$

where $\mathbf{R}_i$ is a matrix of random effects and $\mathbf{Z}_i$ is a design matrix associated to the random effects.

– For each sampling unit i, above model can be vectorized as

$$\begin{aligned} \mathrm{vec}(\mathbf{Y}_i) &= (\mathbf{I} \otimes \mathbf{X}_i)\,\mathrm{vec}(\mathbf{B}) + (\mathbf{I} \otimes \mathbf{Z}_i)\,\mathrm{vec}(\mathbf{R}_i) + \mathrm{vec}(\mathbf{E}_i), \\ \mathbf{y}_{i_\#} &= \mathbf{X}_{i_\#}\boldsymbol{\beta} + \mathbf{Z}_{i_\#}\mathbf{b}_{i_\#} + \boldsymbol{\varepsilon}_{i_\#} \qquad i = 1, 2, \ldots, N, \end{aligned} \tag{4.22}$$

with assumptions of

$$\mathrm{Cov}(\mathbf{b}_{i_\#}) = \mathbf{F}, \qquad \mathrm{Cov}(\boldsymbol{\varepsilon}_{i_\#}) = \boldsymbol{\Sigma} \otimes \mathbf{I}, \qquad \mathrm{Cov}(\mathbf{b}_{i_\#}, \boldsymbol{\varepsilon}_{i_\#}) = \mathbf{0}. \tag{4.23}$$

– Hence the marginal model is

$$\mathbf{y}_{i_\#} \sim N(\mathbf{X}_{i_\#}\boldsymbol{\beta}, \mathbf{V}_{i_\#}), \qquad \mathbf{V}_{i_\#} = \boldsymbol{\Sigma} \otimes \mathbf{I} + \mathbf{Z}_{i_\#}\mathbf{F}\mathbf{Z}'_{i_\#}. \tag{4.24}$$

– By stacking vectors $\mathbf{y}_{i_\#}$, the model for the whole data can be written as

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\varepsilon}, \qquad \mathrm{Cov}(\mathbf{b}) = \mathbf{G} = \mathbf{I} \otimes \mathbf{F}, \\ \mathbf{y} &\sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{V}), \qquad \mathbf{V} = \mathbf{I} \otimes \mathrm{Cov}(\boldsymbol{\varepsilon}_{i_\#}) + \mathbf{Z}\mathbf{G}\mathbf{Z}', \end{aligned} \tag{4.25}$$

where

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_{1\#} \\ \mathbf{y}_{2\#} \\ \vdots \\ \mathbf{y}_{N\#} \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} \mathbf{X}_{1\#} \\ \mathbf{X}_{2\#} \\ \vdots \\ \mathbf{X}_{N\#} \end{pmatrix}, \quad \mathbf{Z} = \begin{pmatrix} \mathbf{Z}_{1\#} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_{2\#} & \dots & \mathbf{0} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{Z}_{N\#} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} \mathbf{b}_{1\#} \\ \mathbf{b}_{2\#} \\ \vdots \\ \mathbf{b}_{N\#} \end{pmatrix}, \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \boldsymbol{\varepsilon}_{1\#} \\ \boldsymbol{\varepsilon}_{2\#} \\ \vdots \\ \boldsymbol{\varepsilon}_{N\#} \end{pmatrix}.$$

– Estimation and prediction of parameters $\boldsymbol{\beta}$ and b can be performed by the standard linear mixed effects methods. However, these methods do not take into account that $\mathrm{Cov}(\boldsymbol{\varepsilon}_{i_\#}) = \boldsymbol{\Sigma} \otimes \mathbf{I}$, which means that standard method values $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{b}}$ are not optimal estimates.

– The covariance matrix $\boldsymbol{\Sigma}$ can be estimated by

$$\widehat{\boldsymbol{\Sigma}} = \frac{\sum_{i=1}^{N} \widehat{\mathbf{E}}_i' \widehat{\mathbf{E}}_i}{n - \mathrm{rank}(\mathbf{X})},$$

where $\widehat{\mathbf{E}}_i = \mathbf{Y}_i - (\mathbf{X}_i \widehat{\mathbf{B}} + \mathbf{Z}_i \widehat{\mathbf{R}_i})$.

4.2.2 Prediction and Testing in Mixed Model

– Prediction of new observations (related to a new sampling unit f)

$$\mathbf{y}_f = \mathbf{B}'\mathbf{x}_f + \mathbf{R}'_f\mathbf{z}_f + \boldsymbol{\varepsilon}_f \tag{4.26}$$

can be done by the empirical BLUP

$$\hat{\mathbf{y}}_f = \widehat{\mathbf{B}}'\mathbf{x}_f + \widehat{\mathbf{R}}'_f\mathbf{z}_f. \tag{4.27}$$

– Simultaneous prediction intervals for elements of the random vector $\mathbf{y}_f$ can be calculated by the estimated quantiles method, i.e, by determining $\alpha/2$ and $1-\alpha/2$ the quantiles of the estimated distribution $\mathbf{y}_f \sim N(\widehat{\mathbf{B}}'\mathbf{x}_f + \widehat{\mathbf{R}}'_f\mathbf{z}_f, \widehat{\boldsymbol{\Sigma}})$.

– Consider the prediction in a partitioned random vector $\mathbf{y}_f$ situation

$$\mathbf{y}_f = \begin{pmatrix}\mathbf{y}_{f1}\\ \mathbf{y}_{f2}\end{pmatrix} = \begin{pmatrix}\mathbf{B}'_1\\ \mathbf{B}'_2\end{pmatrix}\mathbf{x}_f + \begin{pmatrix}\mathbf{R}'_{f1}\\ \mathbf{R}'_{f2}\end{pmatrix}\mathbf{z}_f + \begin{pmatrix}\boldsymbol{\varepsilon}_{f1}\\ \boldsymbol{\varepsilon}_{f2}\end{pmatrix}. \tag{4.28}$$

– The best empirical predictor is

$$\widehat{\mathrm{BP}}(\mathbf{y}_{f2}|\mathbf{Y}, \mathbf{y}_{f1}) = \widehat{\mathbf{B}}'_2\mathbf{x}_f + \widehat{\mathbf{R}}'_{f2}\mathbf{z}_f + \widehat{\boldsymbol{\Sigma}}_{21}\widehat{\boldsymbol{\Sigma}}_{11}^{-1}(\mathbf{y}_{f1} - \widehat{\mathbf{B}}'_1\mathbf{x}_f - \widehat{\mathbf{R}}'_{f1}\mathbf{z}_f) \tag{4.29}$$

with the (asymptotic) estimated covariance matrix of the prediction error being

$$\widehat{\mathrm{Cov}}(\mathbf{y}_{f2} - \widehat{\mathrm{BP}}(\mathbf{y}_{f2}|\mathbf{Y}, \mathbf{y}_{f1})) = \widehat{\boldsymbol{\Sigma}}_{22} - \widehat{\boldsymbol{\Sigma}}_{21}\widehat{\boldsymbol{\Sigma}}_{11}^{-1}\widehat{\boldsymbol{\Sigma}}_{12}. \tag{4.30}$$

– Simultaneous $100(1-\alpha)\%$ prediction intervals for each element of the random vector $\mathbf{y}_{f_2}$ are

$$\left[\widehat{\mathrm{BP}}(\mathbf{y}_{f2}|\mathbf{Y},\mathbf{y}_{f1}) \pm z_{\alpha/2}\,\mathbf{diag}(\widehat{\boldsymbol{\Sigma}}_{22} - \widehat{\boldsymbol{\Sigma}}_{21}\widehat{\boldsymbol{\Sigma}}_{11}^{-1}\widehat{\boldsymbol{\Sigma}}_{12})^{1/2}\right]. \tag{4.31}$$

– Linear hypotheses

$$\begin{aligned} &H_0 : \mathbf{K}'\boldsymbol{\beta} = \mathbf{0}, \\ &H_1 : \mathbf{K}'\boldsymbol{\beta} \neq \mathbf{0}, \qquad \mathbf{K}' \in \mathbb{R}^{q,(p+1)}, \end{aligned}$$

can be tested by the Wald statistic

$$W = (\mathbf{K}'\hat{\boldsymbol{\beta}})'(\mathbf{K}'\widehat{\mathrm{Cov}}(\hat{\boldsymbol{\beta}})\mathbf{K})^{-1}\mathbf{K}'\hat{\boldsymbol{\beta}}, \tag{4.32}$$

which follows asymptotically χ^2 distribution with degrees of freedom $\mathrm{rank}(\mathbf{K})$.

Example 4.2.

Consider the following dataset:

```
       CCNG1    TRAF5      CLU
0-1 16.24913 17.26195 16.18577
0-2 16.31433 17.31736 15.96045
0-3 16.24913 17.37948 15.96045
0-4 16.18577 17.22650 16.08763
0-5 16.31433 17.37948 16.08763
0-6 16.18577 17.09170 16.01778
.
.
72-33 15.90861 17.49438 15.75553
72-34 16.80062 17.59871 16.05501

[1] "cyclin G1"
[2] "TNF receptor-associated factor 5"
[3] "clusterin (complement lysis inhibitor, SP-40,40, sulfated glycoprotein 2,
testosterone-repressed prostate message 2, apolipoprotein J)"

The data result from two experiments investigating the expression response of human T cells to PMA and ionomicin treatment.

The data set (tcell.34) contains the temporal expression levels of 58 genes for 10 unequally spaced time points.
At each time point there are 34 separate measurements.

Rangel, C., Angus, J., Ghahramani, Z., Lioumi, M., Sotheran, E., Gaiba, A., Wild, D. L., and Falciani, F. (2004)
Modeling T-cell activation using gene expression profiling and state-space models. Bioinformatics, 20, 1361?1372.
```

Denote the variables as following

$$Y_1 = \mathsf{CCNG1}, \qquad Y_2 = \mathsf{TRAF5}, \quad Y_3 = \mathsf{CLU}.$$

For the experiment i, consider the multivariate linear mixed effects model

$$\begin{aligned}\mathcal{M}: \quad y_{i1t} &= \beta_{0_1} + \beta_{1_1} t_i + b_{i0_1} + \varepsilon_{i1t},\\ y_{i2t} &= \beta_{0_2} + \beta_{1_2} t_i + b_{i0_2} + \varepsilon_{i2t},\\ y_{i3t} &= \beta_{0_3} + \beta_{1_3} t_i + b_{i0_3} + \varepsilon_{i3t},\end{aligned}$$

which can be written also as

$$\mathcal{M}: \quad \mathbf{Y}_i = \mathbf{X}_i\mathbf{B} + \mathbf{Z}_i\mathbf{R}_i + \mathbf{E}_i, \qquad \operatorname{vec}(\mathbf{E}_i) \sim N(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I}),$$

where

$$\mathbf{Y}_i = (\mathbf{y}_{i1} : \mathbf{y}_{i2} : \mathbf{y}_{i3}), \quad \mathbf{X}_i = (\mathbf{1} : \mathbf{t}_i), \quad \mathbf{Z}_i = (\mathbf{1}), \quad \operatorname{vec}(\mathbf{R}_i) \sim N(\mathbf{0}, \mathbf{F}).$$

(a) Calculate the estimate for the fixed parameters $\mathbf{B}$.

(b) Calculate the estimate for the covariance matrix $\boldsymbol{\Sigma}$.

(c) Predict the value of the new observation $\mathbf{y}_f$, when $i = 34$ and $t = 80$. Create also 80 % simultaneous (quantile based) prediction intervals for elements of the random vector $\mathbf{y}_f$.

(d) Predict the value of the new observation y_{f3t}, i.e., predict the value of the variable $Y_3 = \texttt{CLU}$ when we know that $i = 34$, $t = 80$, and $y_{f1t} = 17$, $y_{f2t} = 18$.