In this paper we develop a method for estimating varying coefficients on effects of covariates without modeling the shape of the spatial-temporal baseline trend. We consider the situation where primary interest is in the effects of covariates and the spatial-temporal baseline trend, though non-negligible, is of secondary interest. This is similar to the situation with the Cox proportional hazards model in survival analysis. Basis functions are used to model the shapes of the varying coefficients, but no particular shape is assumed for the spatial-temporal baseline trend. After the effects of covariates are evaluated, estimates of the spatial-temporal baseline trend can be obtained nonparametrically.

Key words and phrases: Extended Growth curve model, mixture of GMANOVA and MANOVA model, nuisance parameter, simultaneous confidence interval, spatial-temporal baseline trend, varying coefficient.

1. Introduction

Let \( y(t \mid u, v) \) \((t = t_1, \ldots, t_p)\) be longitudinal measurements at the location coordinate \((u, v)\) and \( \mathbf{\beta}(t) = (\beta_1(t), \ldots, \beta_k(t))' \) be a vector of regression coefficients for covariates \( \mathbf{a} = (a_1, \ldots, a_k)' \). The \( \mathbf{\beta}(t) \), generally called varying coefficients, expresses time-varying effects of the covariates (see Hastie and Tibshirani (1993)). For describing a relationship between longitudinal measurements and covariates, we consider the varying coefficient model:

\[
(1.1) \quad y(t \mid u, v) = g(u, v) + f_j(t) + a' \mathbf{\beta}(t) + \varepsilon(t),
\]

where \( g(u, v) \) and \( f_j(t) \) \((j = 1, \ldots, r)\) describe spatial and temporal baseline trends, respectively, and \( \varepsilon(t) \) is an error term. We assume that there are \( r \) groups or clusters of temporal baseline trends and \( \varepsilon = (\varepsilon(t_1), \ldots, \varepsilon(t_p))' \sim N_p(\mathbf{0}, \Sigma) \), where \( \Sigma \) is a \( p \times p \) positive definite matrix.

If there is no baseline spatial variation, i.e., \( g(u, v) \equiv 0 \), model (1.1) is equivalent to the growth curve model proposed by Potthoff and Roy (1964). In that case, Satoh and Yanagihara (2010) proposed a method for estimating the varying coefficients \( \mathbf{\beta}(t) \) while specifying the shape of the temporal baseline trend \( f_j(t) \) by a particular parametric function of \( t \). Regarding \( f_j(t) \) as a nuisance...
parameter, Satoh and Tonda (2016) estimated $\beta(t)$ without specifying the shape of $f_j(t)$.

In this paper, we consider the situation where primary interest is in the effects of the covariates, whereas the spatial-temporal baseline trend, though perhaps non-negligible, is only of secondary interest. Under this situation, we develop an approach to statistical inference on the varying coefficients $\beta(t)$ without specifying the spatial-temporal baseline trends $g(u,v)$ and $f_j(t)$. This is analogous to the proportional hazards model of Cox (1972) in survival analysis, which estimates regression coefficients without specifying the shape of the baseline hazard function.

This paper is organized as follows. In Section 2, we derive an estimator of the varying coefficients $\beta(t)$ representing the varying coefficient model (1.1) with the extended growth curve model. Simultaneous confidence intervals and tests of the varying coefficients are constructed in Section 3. We apply the proposed method to spatial-temporal data on the number of burglaries in 368 residential areas in Philadelphia, Pennsylvania from 1998 to 2009 in Section 4, and a simulation study is presented in Section 5. We conclude with a discussion in Section 6.

2. Estimation of the varying coefficients

We assume that the varying coefficients have the following linear structure:

\begin{align}
\beta(t) = \Theta x(t),
\end{align}

where $\Theta = (\theta_1, \ldots, \theta_k)'$ is a $k \times q$ matrix of unknown parameters and $x(t) = (x_1(t), \ldots, x_q(t))'$ is a vector of $q$ basis functions to describe the shape of $\beta(t)$. If we fit a polynomial curve of degree $q - 1$ in terms of time $t$, we have $x(t) = (1, t, \ldots, t^{q-1})'$.

We partition the measurement locations into $s + 1$ distinct regions, such as a mesh grid, and regard one of them as the reference region. Let $c = (c_1, \ldots, c_s)$ denote a vector of class variables that indicate the region to which the location belongs, i.e., $c_\ell$ takes value 1 if location $(u, v)$ is in the $\ell$th region and 0 if not. Note that $c = (0, \ldots, 0)'$ if the location $(u, v)$ is in the reference region. Letting $g_\ell (\ell = 1, \ldots, s)$ be an unknown parameter for the difference between the $j$th region and the reference region, the spatial baseline trend can be expressed by $g(u,v) = c'g$. $g = (g_1, \ldots, g_s)'$.

The $i$th subject ($i = 1, \ldots, n$) has longitudinal measurements $y_i(t \mid u_i, v_i)$ ($t = t_1, \ldots, t_p$) at the location coordinate $(u_i, v_i)$, a $k$-dimensional vector of covariates $a_i = (a_{i1}, \ldots, a_{ik})'$, and an $s$-dimensional vector of class variables on measurement location $c_i = (c_{i1}, \ldots, c_{is})$. Note that all measurement time points are common among subjects, i.e., this is a balanced design. Letting $y_i = (y_i(t_1 \mid u_i, v_i), \ldots, y_i(t_p \mid u_i, v_i))'$, $f_j = (f_j(t_1), \ldots, f_j(t_p))'$, and $\varepsilon_i = (\varepsilon_1(t_1), \ldots, \varepsilon_1(t_p))'$, the model (1.1) can be expressed as

\begin{align}
y_i' = u_i'g 1_p' + f_j' + a_i'\Theta X + \varepsilon_i',
\end{align}

where $1_p$ is a $p$-dimensional vector in which each element is one, $X = (x(t_1), \ldots, x(t_p))$ is a $q \times p$ design matrix, and the $\varepsilon_i'$s are independently and identically
VARYING COEFFICIENT MODEL WITH SPATIAL-TEMPORAL BASELINE

distributed as a $p$-dimensional normal random vector with mean zero and covariance matrix $\Sigma$. Using an $n \times r$ class variable matrix on the measurement time point defined by

$$
D = 
\begin{pmatrix}
1_{n_1} & \cdots & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & 1_{n_j} & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \cdots & 1_{n_r}
\end{pmatrix}, \quad \sum_{j=1}^{r} n_j = n,
$$

(2.3)

an $n \times p$ observation matrix: $Y = (y_1, \ldots, y_n)'$, an $n \times k$ covariate matrix: $A = (a_1, \ldots, a_n)'$, an $n \times s$ class variable matrix on the measurement location: $C = (c_1, \ldots, c_n)'$, an $r \times p$ nuisance parameter matrix on baseline time trends: $F = (f_1, \ldots, f_r)'$, and an $n \times p$ error matrix: $\mathcal{E} = (\varepsilon_1, \ldots, \varepsilon_n)'$, we have the matrix expression for the model (1.1):

$$
Y = W\Lambda + A\Theta X + \mathcal{E}, \quad W = (C \; D), \quad \Lambda = \begin{pmatrix} g_{1:p} \end{pmatrix},
$$

(2.4)

where $\mathcal{E} \sim N_{n \times p}(0_{n \times p}, \Sigma \otimes I_n)$. Note that $C$ and $D$ are of full rank, i.e., rank($C$) = $s$ and rank($D$) = $r$. The equation (2.4) is a special case of the extended Growth curve model (e.g. see, Kollo and von Rosen (2005), Chapter 4), which is also called a mixture of the GMANOVA and MANOVA models (Chinchilli and Elswick (1985)). In the equation (2.4), the part of GMANOVA corresponds to varying coefficients of covariate effect and the part of MANOVA captures the spatial-temporal baseline trend.

Regarding $\Lambda$ as a nuisance parameter, we first construct the estimator of $\Theta$. Using the result on the maximum likelihood estimator of $\Theta$ in the extended growth curve model (e.g., see Kollo and von Rosen (2005), Chapter 4), if $W$, $A$, $X$, $A'Q_WA$ are of full rank, i.e., rank($W$) = $s + r$, rank($A$) = $k$, rank($X$) = $q$, rank($A'Q_WA$) = $k$, and $n - s - r$ ≥ $p$, we can obtain the estimator of $\Theta$ as follows:

$$
\hat{\Theta} = (A'Q_WA)^{-1}A'Q_WYS^{-1}X'(XS^{-1}X')^{-1},
$$

(2.5)

where $S = Y'Q_WY$ and $Q_W = I_n - W(W'W)^{-1}W'$. Thus we can estimate the time varying coefficients by $\hat{\beta}(t) = \hat{\Theta}x(t)$.

Because $E(Y) = W\Lambda + A\Theta X$, and $Q_WW = O_{n \times (s+r)}$, we have an unbiased estimator:

$$
E(\hat{\Theta}) = \Theta.
$$

If $n - s - r - p + k - 1 > 0$, then the covariance matrix of $\hat{\Theta}$ can be given by

$$
\Omega = \text{cov}(\text{vec}(\hat{\Theta})) = c_1(X\Sigma^{-1}X')^{-1} \otimes (A'Q_WA)^{-1},
$$

(2.6)
where $c_1 = (n - s - r - 1)/(n - s - r - p + k - 1)$. Note that $E((X'S^{-1}X)^{-1}) = (n - s - r - p + k)(X\Sigma^{-1}X')^{-1}$. An unbiased estimator of $\Omega$ can be obtained by

$$
\hat{\Omega} = c_2(XS^{-1}X')^{-1} \otimes (A'W A)^{-1},
$$

(2.7) where $c_2 = c_1/(n - s - r - p + k)$.

Once the estimate $\hat{\Theta}$ is obtained, we can estimate the matrix of nuisance parameters $g$ and $F$. From equation (2.4), we obtain

$$
\hat{\Lambda} = \begin{pmatrix} \hat{g}^1_p \\ \hat{F} \end{pmatrix} = (W'W)^{-1}W'(Y - A\hat{\Theta}X).
$$

(2.8) Note that $W = (C' D)$. We can obtain the estimators of $g$ and $F$,

$$
\hat{F} = (D'Q_C D)^{-1}D'Q_C(Y - A\hat{\Theta}X),
$$

$$
\hat{g} = (C'C)^{-1}C(Y - D\hat{F} - A\hat{\Theta}X)1_p(1'_p 1_p)^{-1},
$$

(2.9) which are unbiased.

3. Simultaneous confidence intervals and tests of the varying coefficients

The estimator $\hat{\beta}(t) = \hat{\Theta}x(t)$ is asymptotic normal because of the asymptotic properties of estimator $\hat{\Theta}$ (e.g., see Kollo and von Rosen (2005), Chapter 4). Using this asymptotic property, we can construct simultaneous confidence intervals and perform tests of the varying coefficients $\beta(t) = (\beta_1(t), \ldots, \beta_k(t))'$ by applying the method in Satoh and Yanagihara (2010).

Let $\hat{\Omega}_j$ be the submatrix of $\hat{\Omega}$ corresponding to the estimates of $\Omega_j = \text{cov}(\theta_j)$. Using the property of matrix algebra (e.g., see Schott (2005)), $\hat{\Omega}_j$ can be expressed by

$$
\hat{\Omega}_j = c_2[(A'Q_W A)^{-1}]_{jj}(XS^{-1}X')^{-1},
$$

(3.1) where $[(A'Q_W A)^{-1}]_{jj}$ denotes the $j$th diagonal element of matrix $(A'Q_W A)^{-1}$. Then, if $\hat{\theta}_j \sim N_q(\theta_j, \Omega_j)$, $\hat{\beta}_j(t) = \hat{\theta}_j'x(t) \sim N(\beta_j(t), \sigma^2_j(t))$, where $\sigma^2_j(t) = x(t)'\Omega_j x(t)$. Letting $\hat{\sigma}^2_j = x(t)'\hat{\Omega}_j x(t)$, we construct a confidence interval for varying coefficient $\beta_j(t)$ given by

$$
\mathcal{I}_{j,\alpha}(t \mid u_\alpha) = [\hat{\beta}_j(t) - u_\alpha \hat{\sigma}_j(t), \hat{\beta}_j(t) + u_\alpha \hat{\sigma}_j(t)].
$$

(3.2) The coverage probability of $\mathcal{I}_{j,\alpha}(t \mid u_\alpha)$ depends on $u_\alpha$. For example, the pointwise confidence interval at a fixed time point $t$ can be constructed by letting $u_\alpha = z_\alpha/2$, where $z_\alpha$ denotes the upper 100$\alpha$ percentage point of $N(0, 1)$. Note that the confidence interval $\mathcal{I}_{j,\alpha}(t \mid z_\alpha/2)$ satisfies $\Pr(\beta_j(t) \in \mathcal{I}_{j,\alpha}(t \mid z_\alpha/2)) \approx 1 - \alpha$ for a fixed time point $t$.

To construct a simultaneous confidence interval as a function of $t$, it is necessary to evaluate the distribution of the supremum of the Wald type statistic
VARYING COEFFICIENT MODEL WITH SPATIAL-TEMPORAL BASELINE

$T_j(t) = \{ \hat{\beta}_j(t) - \beta_j(t) \}/\hat{\sigma}_j(t)$, but it is difficult to derive the explicit distribution of the supremum statistic in general. We here evaluate the upper bound statistic of the supremum of $T_j(t)$ in the same manner as Satoh and Yanagihara (2010). From the inequality in Rao (1973, p. 60), $\hat{\beta}_j(t)$ asymptotically satisfies the following equation:

$$\sup_{t \in \mathbb{R}} T_j(t)^2 = \sup_{x \in \mathbb{R}^p} \left\{ x'(\hat{\theta}_j - \theta_j) \right\}^2 x'\hat{\Omega}_j x \leq \sup_{x \in \mathbb{R}^q} \left\{ x'(\hat{\theta}_j - \theta_j) \right\}^2 x'\hat{\Omega}_j x = (\hat{\theta}_j - \theta_j)'\hat{\Omega}_j^{-1}(\hat{\theta}_j - \theta_j) \sim \chi^2_q, \quad n \to \infty.$$  

(3.3)

Note that the asymptotic distribution of the upper bound statistic is $\chi^2_q$ for any time point $t$. Let $c_{q,\alpha}$ be the upper $100\alpha$ percentage point of $\chi^2_q$. Using $u_\alpha = \sqrt{c_{q,\alpha}}$, the coverage probability of the confidence interval $I_{j,\alpha}(t | \sqrt{c_{q,\alpha}})$ satisfies

$$\Pr(\beta_j(t) \in I_{j,\alpha}(t | \sqrt{c_{q,\alpha}}) | \forall t \in \mathbb{R}) \geq 1 - \alpha.$$  

(3.4)

Based on equation (3.3), we construct test statistics for some null hypotheses on the varying coefficient $\beta_j(t)$. We first consider the hypothesis that the varying coefficient is uniformly zero over time, i.e.,

Uniformly zero: $H_0: \beta_j(t) = 0 \quad \forall t \in \mathbb{R}.$  

(3.5)

The uniformly zero hypothesis in (3.5) is equivalent to $\theta_j = 0$. Using the equation (3.3) under the uniformly zero hypothesis, the upper bound of the supremum of $T_j(t)^2$ is $W_j = (\hat{\theta}_j)'\hat{\Omega}_j^{-1}\hat{\theta}_j$, whose asymptotic distribution is $\chi^2_q$. Note that $W_j$ can be expressed by

$$W_j = \frac{((A'QWA)^{-1}A'QWYPX_{t}'S^{-1}Y'QWA(A'QWA)^{-1})_{jj}}{c_2((A'QWA)^{-1})_{jj}},$$  

(3.6)

where $P_{X_{t}'S^{-1}} = S^{-1}X'(XS^{-1}X')^{-1}XS^{-1}$. Hence $W_j$ can be utilized as a test statistic for the null hypothesis $H_0$. The uniformly zero hypothesis is rejected when $W_j > c_{q,\alpha}$, and the $p$-value can be calculated by $\Pr(\chi^2_q > W_j)$.

Next, we consider the hypothesis that the varying coefficient is unchanged over the time, i.e.,

$$\text{Uniformly constant: } H_0: \beta_j(t) = \text{const.} \quad \forall t \in \mathbb{R}.  

(3.7)$$

Note that the uniformly constant hypothesis in (3.7) is equivalent to $\theta_j^{(-1)} = 0$, where $\theta_j^{(-1)}$ is a $(q - 1)$-dimensional vector excluding the first elements of $\theta_j$, because the first element of $x(t)$ is 1. Analogous to the test for uniformly zero hypothesis, we can construct a test statistic and derive an asymptotic null distribution for the uniformly constant hypothesis.
4. Application to real data

In this section, we illustrate the method with an application to data on the number of burglaries in 368 residential areas in Philadelphia, Pennsylvania from 1998 to 2009. The dataset is open to the public and is located at the Philadelphia Neighborhood Information System web site: www.cml.upenn.edu/nis/. In order to explain the number of burglaries, we selected median household income and population size in each area, both of which were recorded in 2000.

Let $y_i(t \mid u_i, v_i)$ be the number of burglaries at time $t$ ($t = 1998, \ldots, 2009$) in the $i$th residential area ($i = 1, \ldots, 368$), whose location coordinate is $(u_i, v_i)$, and $\mathbf{a}_i = (a_{i1}, a_{i2})'$ the vector of covariates with $a_{i1}$ the median household income ($\times 10^3$) and $a_{i2}$ the population size ($\times 10^3$). Thus, this dataset has $n = 368$ samples measured at $p = 22$ time points with $k = 2$ covariates. Figure 1 shows the map of Philadelphia, Pennsylvania with the 368 residential areas. The center of gravity in each area, marked by gray point in Figure 1, is used as the location coordinate $(u_i, v_i)$. We partitioned the residential areas into 57 mesh grids based on 11 vertical and horizontal grid lines. In Figure 1, the region enclosed by the thick line is used as the reference region for describing the spatial baseline trend.

Our primary interest is in the time varying effect of household income and population size. We assume that the varying coefficients describing the effects of covariates are expressed as quadratic functions of time, and the dimension of the temporal baseline is $r = 1$. Then, the design vector $\mathbf{x}(t)$ is given by $\mathbf{x}(t) = (1, t, t^2)'$ with length $q = 3$ and $D = 1_n$. Using equation (2.5), the
(a) (b)

Figure 2. Estimated varying coefficients and 95% simultaneous confidence intervals: (a) effect of household income, (b) effect of population size.

| Covariates       | Uniformly zero | Uniformly constant |
|------------------|----------------|-------------------|
|                  | test stat.     | p-value           | test stat. | p-value |
| household income | 13.56          | 0.004             | 0.12       | 0.939   |
| population size  | 404.04         | <0.001            | 37.90      | <0.001  |

estimate of $\Theta$ is

\[
\hat{\Theta} = \begin{pmatrix} -0.1401 & -0.0035 & 0.0001 \\ 6.8630 & -0.5586 & 0.0386 \end{pmatrix}.
\]

Using $\hat{\Theta}$ in the equation (4.1), the time varying effect of household income and population size can be described by $\hat{\beta}_1(t) = -0.1401 - 0.0035t + 0.0001t^2$ and $\hat{\beta}_2(t) = 6.8630 - 0.5586t + 0.0386t^2$, respectively. The estimated curves of the varying coefficients and their 95% simultaneous confidence intervals are shown in Figure 2. Household income has a negative effect which seems slightly decreasing with time, and the effect of population size is positive and decreasing with time. From the results of tests of hypotheses of the varying coefficients in Table 1, the effect of household income is significant but not significantly varying over time.

After the varying coefficients were estimated, we estimated the spatial-temporal baseline trend using equation (2.9), with results shown in Figure 3. From Figure 3, we can see that there is a high number of burglaries in the center areas and low number in the north east areas, and the number of burglaries tends to decrease at first but remains roughly flat after that.
Figure 3. Estimated spatial-temporal baseline trends: (a) estimates of spatial baseline trend which denote the regional difference compared with the reference region, (b) estimates of temporal baseline trend.

Figure 4. The histogram of $\sup_{t \in \mathbb{R}} T_2(t)^2$ by Monte Carlo method and approximate distribution of upper bound statistic.

5. Simulation study

Using the upper bound statistic instead of the supremum statistic in the inequality (3.3), we constructed the statistical inferences on varying coefficient. The gap between the supremum and upper bound statistics might affect accuracy of statistical inference.

We here investigate actual covering probability of the proposed simultaneous
confidence intervals in Section 3 for finite sample sizes by simulation studies. 100,000 simulated data are generated by the varying coefficients model (1.1) or its matrix expression (2.4), and we set the estimates for Philadelphia burglaries data to the true values of unknown parameters. We here deal with the accuracy on the simultaneous confidence interval of $\beta_2(t)$, effect of population size. The distribution of supremum statistic can be obtained by Monte Carlo method.

Figure 4 shows the distribution of supremum statistic (white histogram) and approximate distribution of upper bound statistic (black solid curve). The short line on the horizontal axes is its upper 95% percentage point of $\chi^2_3$, i.e., $c_{3,0.05} = 7.81$. Then, the actual covering probability of the 95% simultaneous confidence interval was 98.43%. From this simulation study, the simultaneous confidence interval would be conservative even for large sample size.

6. Discussion

In this paper, we dealt with a varying coefficient model that has spatial-temporal baseline trends, and estimated the varying coefficients on the effects of covariates without specifying the shapes of the spatial-temporal baseline trends using the extended growth curve model. The proposed method is particularly useful when primary interest is in the effects of covariates. In general, it is not easy to appropriately model the spatial-temporal baseline patterns if the number of measurement locations or time points is large. However, the proposed method can produce estimates of the varying coefficients without the need to model the spatial-temporal baseline trends. This is often an advantage in practical situations.

Figure 5. Map of Philadelphia, Pennsylvania and partitioned regions. The 368 residential areas are partitioned into 17 mesh grids. The region enclosed by the thick line is used as the reference region.
If the spatial-temporal baseline trend is also of interest, we can estimate it after the effects of covariates have been evaluated. As mentioned above, it is sometimes difficult to determine an appropriate function to model the spatial-temporal baseline trend. Therefore, we consider that it is convenient to obtain estimates of the spatial-temporal baseline trends without making assumptions about their shapes, i.e., to obtain them nonparametrically.

We partitioned the measurement locations into $s + 1$ distinct regions. An important question is whether the number of partitions affects the results. Therefore, we applied the proposed method using alternative partitions. Figure 5 shows the case where the 368 residential areas were partitioned into 17 mesh grids based on 6 vertical and horizontal grid lines. Figure 6 shows the estimated varying coefficients and their 95% simultaneous confidence intervals based on the partitions in Figure 5. Comparing Figure 2 with Figure 6, there is little change in the estimates and the interpretation remains unchanged. Figure 7 shows the estimated spatial-temporal baseline trend. The rough tendency of Figure 7(a) is similar to that of Figure 3(a) and the differences between Figures 3(a) and 7(a) are only in terms of resolution. Comparing Figure 7(b) with Figure 3(b), the resolution of estimates of the spatial baseline has an effect on estimates of the temporal baseline, but the overall tendency is similar. From these results we consider that there is little impact of the number of partitions.

In summary, we conclude that the proposed method is particularly useful in practice, when the main interest is in the effects of covariates, because modeling of the baseline is not required and the number of partitions does not affect the results. Furthermore, the spatial-temporal baseline trend can be evaluated
Figure 7. Estimated spatial-temporal baseline trends using the partitions in Figure 5: (a) estimates of spatial baseline trend which denote the regional difference compared with the reference region, (b) estimates of temporal baseline trend.

nonparametrically, if needed.

Acknowledgements

The authors are grateful to the referee for his/her valuable comments and suggestions which improved an earlier version of the manuscript. Part of this research was supported by the Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Scientific Research (C) # 26330043 (2014–2017), # 26460754 (2014–2017), and the Program of the network-type joint Usage/Research Center for Radiation Disaster Medical Science of Hiroshima University, Nagasaki University and Fukushima Medical University.

References

Chinchilli, V. M. and Elswick, R. K. (1985). A mixture of the MANOVA and GMANOVA models, *Comm. Statist. Theory Methods*, **14**, 3075–3089.

Cox, D. R. (1972). Regression models and life-tables, *J. Roy. Statist. Soc. Ser. B*, **34**, 187–220.

Hastie, T. and Tibshirani, R. (1993). Varying–coefficient models, *J. Roy. Statist. Soc. Ser. B*, **55**, 757–796.

Kollo, T. and von Rosen, D. (2005). *Advanced Multivariate Statistics with Matrices*, Springer, Netherlands.

Potthoff, R. F. and Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems, *Biometrika*, **51**, 313–326.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, John Wiley, Inc., New York.

Satoh, K. and Tonda, T. (2016). Estimating regression coefficients for balanced growth curve model when time trend of baseline is not specified, *Amer. J. Math. Management Sci.*, **35**, 183–193.
Satoh, K. and Yanagihara, H. (2010). Estimation of varying coefficients for a growth curve model, *Amer. J. Math. Management Sci.*, **30**, 243–256.

Schott, J. R. (2005). *Matrix Analysis for Statistics, 2nd Edition*, John Wiley & Sons, Inc., New York.