몬테카를로법을 이용한 비선형 확률계수모형의 추정*

Estimation Using Monte Carlo Methods in Nonlinear Random Coefficient Models

김성연**
Seongyeon Kim

Abstract

Repeated measurements on units under different conditions are common in biological and biomedical studies. In a number of growth and pharmacokinetic studies, the relationship between the response and the covariates is assumed to be nonlinear in some unknown parameters and the form remains the same for all units. Nonlinear random coefficient models are used to analyze such repeated measurement data. Extended least squares methods are proposed in the literature for estimating the parameters of the model. However, neither objective function has closed form expression in practice. This paper proposes Monte Carlo methods to estimate the objective functions and the corresponding estimators. A simulation study that compare various methods is included.

* 이 논문은 1998학년도 동아대학교 학술연구조성비(공모과제)에 의하여 연구되었음.
** 동아대학교 경영정보과학부

1. Introduction

Repeated measurements on units under different conditions or at different time points are common in biological and biomedical studies. In a plasma concentration study, a dose of size D of a drug is injected periodically into several patients and the drug concentration in the blood is measured over time. In most such applications, it is assumed that the form of the relationship between a response and covariates is nonlinear in unknown parameters and remains the same for all units. However, the parameters in the relationship may vary from individual to individual. A model that takes into account the variability among measurements within a given experimental unit and random variation among units is the nonlinear random coefficient model and is given

$$y_{ij}=f(\mathbf{x}_{ij},\ \boldsymbol{\beta}_i)+\varepsilon_{ij}\ , \qquad (1.1)$$
$$i=1,2,\cdots,n \text{ units}$$
$$j=1,2,\cdots,r_i \text{ measurements}$$

where $\boldsymbol{\beta}_i \sim NID(\boldsymbol{\beta},\Psi)$, $\varepsilon_{ij} \sim NID(0,\sigma^2)$ and { $\boldsymbol{\beta}_i$ } and { ε_{ij} } are independent. Here, $\boldsymbol{\beta}_i$ is p×1 vector of random coefficients that vary with the unit.

For example, to study the functional relationship between the ultrafiltration rate (UFR) and the transmembrane pressure (TMP) among a class of high flux dializers, Ramos(1993) considered the following model, which is a reparameterization of the model used by Vonesh and Cater(1992),

$$y_{ij}=\beta_{1i}+\beta_{3i}\ e^{-\beta_{2i}t_{ij}}+\varepsilon_{ij}\ , \qquad (1.2)$$

where y_{ij}'s are measurements of UFR on each dialyzer, t_{ij}'s are the measured TMP's, $\boldsymbol{\beta}_i=(\beta_{1i},\beta_{2i},\beta_{3i})' \sim NID(\boldsymbol{\beta},\Psi)$ and $\varepsilon_{ij} \sim NID(0,\sigma^2)$.

In linear random coefficient models and nonlinear mixed effects models where the random coefficients enter only as linear functions, we have $E(y_{ij})=f(\mathbf{x}_{ij},\boldsymbol{\beta})$. That is, $f(\cdot,\cdot)$ represents the population average response function. However, for model (1.2),

$$E(y_{ij})=\beta_1+[\beta_3-\psi_{23}t_{ij}]\ e^{-\beta_2 t_{ij}+0.5\psi_{22}t_{ij}^2} \neq \beta_1+\beta_3\ e^{-\beta_2 t_{ij}} \qquad (1.3)$$

unless $\psi_{22}=\psi_{23}=0$. That is, $E(y_{ij})\neq f(\mathbf{x}_{ij},\boldsymbol{\beta})$ unless β_{2i} is fixed, in which case (1.2) reduces to a nonlinear mixed effects model. In general, for model (1.1), $E(y_{ij})\neq f(\mathbf{x}_{ij},\boldsymbol{\beta})$. That is, the nonlinear function $f(\cdot,\cdot)$ represents a subject specific relationship rather than a population average. Since $E(y_{ij})\neq f(\mathbf{x}_{ij},\boldsymbol{\beta})$, methods that provide consistent estimators in linear random coefficient models and nonlinear mixed effects models, do not extend to nonlinear random coefficient models.

In Section 2, we focus the extended least squares estimator using Monte Carlo method (MCELS) that provides consistent estimators for nonlinear random coefficient models. In Section 3, using a simulation study, we compare our method with commonly used methods. We analyze the dializer data of Vonesh and Carter(1992) in Section 4. We conclude the paper with a brief discussion.

2. Estimation methods

In this section, we consider the extended least squares method proposed by Beal and Sheiner (1980) and the extended least squares method using Monte Carlo method (MCELS) for estimating $\boldsymbol{\theta} = (\boldsymbol{\beta}', \mathrm{vech}(\Psi)', \sigma^2)'$ of model(1.1), where vech(·) of a symmetric matrix is a vector obtained by stacking the elements of the columns on or below the diagonal.

2.1 Extended Least Squares Estimation

Sheiner and Beal(1982) consider the extended least squares (ELS) estimator which minimizes

$$Q_{\mathrm{ELS}}(\boldsymbol{\theta}) = \sum_{i=1}^{n} \log|H_i(\boldsymbol{\theta})| + \sum_{i=1}^{n} [\boldsymbol{y}_i - \boldsymbol{g}_i(\boldsymbol{\theta})]' H_i^{-1}(\boldsymbol{\theta}) [\boldsymbol{y}_i - \boldsymbol{g}_i(\boldsymbol{\theta})] \quad (2.1)$$

where

$$\boldsymbol{g}_i(\boldsymbol{\theta}) = E[\boldsymbol{y}_i] = E[\boldsymbol{f}_i(X_i, \boldsymbol{\beta}_i)]$$

$$H_i(\boldsymbol{\theta}) = \mathrm{Var}(\boldsymbol{y}_i) = \mathrm{Var}[\boldsymbol{f}_i(X_i, \boldsymbol{\beta}_i) + \sigma^2 I_{r_i}]$$

$$\boldsymbol{y}_i = (y_{i1}, \ldots, y_{ir_i})'$$

and

$$\boldsymbol{f}_i(X_i, \boldsymbol{\beta}_i) = (f(\boldsymbol{x}_{i1}, \boldsymbol{\beta}_i), \ldots, f(\boldsymbol{x}_{ir_i}, \boldsymbol{\beta}_i))'.$$

Note that (2.1), except for a constant, is -2 times the log likelihood function of a normal random vector with mean $\boldsymbol{g}_i(\boldsymbol{\theta})$ and variance $H_i(\boldsymbol{\theta})$. Even when $\boldsymbol{y}_i$'s are not normal, like in our model (1.1), the estimators minimizing the objective function (2.1) have certain desirable properties. Beal (1984) gives some regularity conditions under which the ELS estimators are consistent and asymptotically normal.

For model (1.2), Ramos(1993) show that the j^{th}element of $\boldsymbol{g}_i(\boldsymbol{\theta})$ is given by (1.3) and the $(j, l)^{th}$element of $H_i(\boldsymbol{\theta})$ is given by

$$\begin{aligned} H_{ijl}(\boldsymbol{\theta}) &= \psi_{11} + \beta_1^2 \\ &+ [\psi_{13} + \beta_1\beta_3 - (\beta_1\psi_{32} + \beta_3\psi_{12})t_{ij} + \psi_{12}\psi_{23}t_{ij}^2]a_{ij}^* \\ &+ [\psi_{13} + \beta_1\beta_3 - (\beta_1\psi_{32} + \beta_3\psi_{12})t_{ij} + \psi_{12}\psi_{23}t_{ij}^2]a_{il}^* \\ &+ [\psi_{33} + \beta_3 - \psi_{32}(t_{ij} + t_{il})^2]a_{ij}^* a_{il}^* e^{\phi_{22}t_{ij}t_{il}} \\ &+ \sigma^2 I_{jl} \end{aligned} \quad (2.2)$$

where

$$a_{ij}^* = e^{-\beta_2 t_{ij} + 0.5\phi_{22}t_{ij}^2}.$$

Even though for model (1.2), fortunately, a closed form expression is derivable, in general it is not possible to give expressions for $\boldsymbol{g}_i(\boldsymbol{\theta})$ and $H_i(\boldsymbol{\theta})$. Because of this difficulty, Sheiner and Beal(1980) and Beal and Sheiner(1985) propose approximate ELS estimators (APELS) based on first- or second-order Taylor series expansion of $\boldsymbol{f}_i(X_i, \boldsymbol{\beta}_i)$ around $\boldsymbol{\beta}$. Ramos and Pantula (1995) give an example to show that such approximate ELS estimators are inconsistent.

2.2 Extended Least Squares Estimation Using Monte Carlo Method

We propose approximating $\boldsymbol{g}_i(\boldsymbol{\theta})$ and

$H_i(\theta)$ based on a Monte Carlo estimation method that leads to consistent estimators. The algorithm for finding the Monte Carlo Extended Least Square (MCELS) estimator, $\hat{\theta}^{(k)}_{MCELS_n}$, based on k Monte Carlo replications in model (1.1) is given as follows:

(a) For each i, generate k random vectors, $\{Z_i^{(1)}, \ldots, Z_i^{(k)}\}$, distributed as $NID(0, I_p)$ and calculate

$$\beta_i^{(l)} = \beta + \Psi^{\frac{1}{2}} Z_i^{(l)}$$

and

$$f_i(\beta_i^{(l)}) = f_i(X_i, \beta_i^{(l)}).$$

(b) For large k, $g_i(\theta)$ and $H_i(\theta)$ can be approximated by

$$\overline{g}_i^{(k)}(\theta) = \frac{1}{k} \sum_{l=1}^{k} f_i(X_i, \beta_i^{(l)})$$

and

$$\overline{H}_i^{(k)}(\theta) = \frac{1}{k} \sum_{l=1}^{k} f_i(X_i, \beta_i^{(l)}) f_i(X_i, \beta_i^{(l)})' - \overline{g}_i^{(k)}(\theta) \overline{g}_i^{(k)}(\theta)' + \sigma^2 I_{r_i}.$$

We replace $g_i(\theta)$ and $H_i(\theta)$ with $\overline{g}_i^{(k)}(\theta)$ and $\overline{H}_i^{(k)}(\theta)$ in the ELS objective function in (2.1)

(c) Find the MCELS estimate, $\hat{\theta}^{(k)}_{MCELS_n}$, which minimizes the approximate MCELS objective function

$$\hat{\theta}^{(k)}_{MCELS_n}(\theta) = \sum_{i=1}^{n} \overline{u}_i^{(k)}(\theta)$$

where $\overline{u}_i^{(k)}(\theta) = \log|\overline{H}_i^{(k)}(\theta)| + [y_i - \overline{g}_i^{(k)}(\theta)]^{-1}[y_i - \overline{g}_i^{(k)}(\theta)].$

(d) Approximate variance covariance matrix for $\hat{\theta}^{(k)}_{MCELS_n}$ can be obtained using the empirical information sandwich.

$$\frac{1}{n} \hat{A}_n^{-1} \hat{B}_n \hat{A}_n^{-1}$$

where

$$\hat{A}_n = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{\partial^2 \overline{u}_i^{(k)}(\theta)}{\partial \theta \partial \theta'} \right]_{\theta = \hat{\theta}^{(k)}_{MCELS_n}}$$

and

$$\hat{B}_n = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{\partial \overline{u}_i^{(k)}(\theta)}{\partial \theta} \frac{\partial \overline{u}_i^{(k)}(\theta)'}{\partial \theta} \right]_{\theta = \hat{\theta}^{(k)}_{MCELS_n}}$$

Under some regularity conditions, Kim (1997) shows that $\hat{\theta}^{(k)}_{MCELS_n}$ converges to the ELS estimator $\hat{\theta}_{ELS_n}$ of Beal and Sheiner(1985) as k tends to infinity and is consistent and asymptotically normal as n and k tend to infinity. That is, we can expect MCELS estimates to approximate ELS estimates and to keep the desirable properties of ELS by choosing a sufficiently large k while other estimation procedures do not have such desirable properties (see Ramos and Pantula(1995)).

3. Simulation Study

To study the small sample performance of our procedures, we consider model (1.2) that was considered by Vonesh and Carter (1992) and Ramos (1993). Consider the model

$$y_{ij}=\beta_{1i}+\beta_{3i}\,e^{-\beta_{2i}t_{ij}}+\varepsilon_{ij}\,, \qquad (3.1)$$
$$i=1,2,\cdots,n$$
$$j=1,2,\cdots,r$$
$$\boldsymbol{\beta}_i=(\beta_{1i},\beta_{2i},\beta_{3i})'\sim NID(\boldsymbol{\beta},\Psi)$$
$$\varepsilon_{ij}\sim NID(0,\sigma^2)$$

and { $\boldsymbol{\beta}_i$ } and { ε_{ij} } are independent. We use the values used by Ramos (1993) and these are obtained from an analysis of the data presented in Vonesh and Carter (1992). We used

$$\boldsymbol{\beta}=(125,\ 0.00875,\ -131.163)'$$

$$\Psi=\begin{bmatrix} 100 & -0.00375\lambda & -102.225 \\ -0.00375\lambda & (800)^{-2}\lambda^2 & 0.0028077\lambda \\ -102.225 & 0.0028077\lambda & 106.556 \end{bmatrix}$$

$\sigma^2=1.96$, $n=50$, $r=7$ and $\lambda=1$ for Design 1 and $\lambda=\sqrt{3}$ for Design 2. Design 2 is considered to study the effect of increased variability of the nonlinear random coefficient β_{2i}, which appears in the exponent. The values of t_{ij} correspond to the transmembrane pressure values of the first 10 units, replicated 5 times, given in the example used by Vonesh and Carter (1992).

For each case, 100 data sets are generated. For model (3.1), it is possible to obtain closed form expressions for $g_i(\boldsymbol{\theta})$ (see (1.3)) and $H_i(\boldsymbol{\theta})$ (see (2.2)) and hence can compute the exact extended least squares estimator (EXELS), $\widehat{\boldsymbol{\theta}}_{ELS_n}$. Also, using the first order Taylor series approximation

$$f(t_{ij},\boldsymbol{\beta}_i)\fallingdotseq[\beta_1+\beta_3 e^{-\beta_2 t_{ij}}]+(\beta_{1i}-\beta_1)$$
$$+e^{-\beta_2 t_{ij}}(\beta_{3i}-\beta_3)$$
$$+\beta_3 e^{-\beta_2 t_{ij}}(-t_{ij})(\beta_{2i}-\beta_2)$$

one can obtain approximate expressions for $g_i(\boldsymbol{\theta})$ and $H_i(\boldsymbol{\theta})$ given by

$$g_{ij}(\boldsymbol{\theta})\fallingdotseq\beta_1+\beta_3 e^{-\beta_2 t_{ij}}$$

and

$$H_{ijl}(\boldsymbol{\theta})\fallingdotseq\psi_{11}-\beta_3(t_{il}e^{-\beta_2 t_{il}}+t_{ij}e^{-\beta_2 t_{ij}})\psi_{12}$$
$$+(e^{-\beta_2 t_{il}}+e^{-\beta_2 t_{ij}})\psi_{13}$$
$$+\beta_3^2 t_{ij}t_{il}e^{-\beta_2(t_{ij}+t_{il})}\psi_{22}-\beta_3 e^{-\beta_2(t_{ij}+t_{il})}\psi_{23}$$
$$+e^{-\beta_2(t_{ij}+t_{il})}\psi_{33}+\sigma^2 I_{jl}.$$

This one of the approximations considered by Sheiner and Beal (1980) and is commonly used in some software packages. We denote the approximate ELS(APELS) estimator obtained by using the first order approximation to $g_i(\boldsymbol{\theta})$ and $H_i(\boldsymbol{\theta})$, as $\widehat{\boldsymbol{\theta}}_{APELS_n}$. For each data set, we compute $\widehat{\boldsymbol{\theta}}_{ELS_n}$, $\widehat{\boldsymbol{\theta}}_{APELS_n}$, and $\widehat{\boldsymbol{\theta}}^{(k)}_{MCELS_n}$ with $k=1000$.

All simulation programs are written in FORTRAN 90 using IMSL subroutines. Powell's method was used for minimizing the objective functions. We used 10^{-10} as the tolerance level for the Powell's method. We have used the parameterization $\Psi=\Gamma'\Gamma$ where Γ is a lower triangular matrix with nonnegative elements on the diagonal and obtained an estimate of Γ first. This imposes the condition that Ψ is a nonnegative definite matrix.

We summarize the empirical mean, bias, estimated standard error, empirical standard error and the root mean square error in Table 3.1 and 3.2. We also report the coverage rate, the percentage of times that an asymptotic 95% confidence interval includes the true value of the mean parameters. For Design 1, where the variance of β_{2i} is not large, the three ELS estimates of the mean parameters are similar, though the APELS estimates of β_1 and β_3 tend to be somewhat biased. For Design 2, where the variance of β_2 is large, APELS performs poorly. APELS estimates of β_1 and β_3 are seriously biased and the coverage rate is quite below 95%. We also notice that the MCELS estimator behaves as good as the EXELS estimator in both designs.

To see the effect of a reparameterization and misspecified distributional assumption, we generate the data from a reparameterization of the model (3.1) given by

$$y_{ij} = 10\alpha_{1i}[1 - e^{-0.0125\alpha_{2i}(t_{ij}-\alpha_{3i})}] + \varepsilon_{ij}, \quad (3.2)$$

$$i = 1, 2, \cdots, n$$

$$j = 1, 2, \cdots, r$$

$$\boldsymbol{\alpha}_i = (\alpha_{1i}, \alpha_{2i}, \alpha_{3i})' \sim NID(\boldsymbol{\alpha}, \Sigma)$$

$$\varepsilon_{ij} \sim NID(0, \sigma^2)$$

and { $\boldsymbol{\beta}_i$ } and { ε_{ij} } are independent. Note that, comparing (3.1) with (3.2), $\beta_{1i} = 10\alpha_{1i}$, $\beta_{2i} = 0.0125\,\alpha_{2i}$ and $\beta_{3i} = -10\,\alpha_{1i}e^{0.00125\alpha_{2i}\alpha_{3i}}$. So, clearly if $\boldsymbol{\alpha}_i$ is normal then $\boldsymbol{\beta}_i$ is not normal. Though we generate the data from (3.2), we fit model (3.1) with the incorrect assumption that $\boldsymbol{\beta}_i$ is normally distributed. We took

$$\boldsymbol{\alpha} = (12.5,\ 7.0,\ 5.5)'$$

$$\Sigma = \begin{bmatrix} 1.0 & -0.3 & 0.0 \\ -0.3 & 1.0 & 0.0 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

and $\sigma^2 = 1.96$.

The choice of $\boldsymbol{\beta}$ and Ψ in Design 1 is obtained by calculating the approximate expectations and variances using the first order linear approximation of $\boldsymbol{\beta}_i$ in terms of $\boldsymbol{\alpha}_i$. Results for this design are summarized in Table 3.3. We notice that, even though the normality assumption is misspecified, both EXELS and MCELS provide estimates similar to those in Design 1. APELS still gives biased estimates of β_1 and β_3 .

To summarize, we notice that the APELS estimation based on linearization may have serious bias in mean parameter estimates, especially when the variability in random coefficients is large. Also, APELS method may lead to inconsistent estimates. On the other hand, MCELS estimation produces estimates close to that of EXELS and both procedures yield good estimates and confidence intervals for the mean parameters.

Table 3.1 Estimates, error measures and coverage rates for Design 1

Parameter True value	Method	Mean	Bias	Est'ed $\sqrt{Var}$	Std Error	$\sqrt{MSE}$	Coverage Rate
β_1	EXELS	124.84	-.16	1.48	1.55	1.55	.93
125.0	APELS	124.12	-.88	1.45	1.52	1.75	.91
	MCELS	124.85	-.15	1.48	1.54	1.54	.98
β_2	EXELS	87.73	.23	1.91	2.11	2.11	.90
87.50	APELS	87.27	-.23	1.88	2.10	2.10	.91
	MCELS	87.71	.21	1.91	2.11	2.11	1.00
β_3	EXELS	-131.04	.13	1.50	1.57	1.57	.98
-131.163	APELS	-130.08	1.08	1.49	1.55	1.88	.88
	MCELS	-131.05	.11	1.50	1.56	1.56	.98
σ^2	EXELS	1.99	.03	.20	.18	.18	
1.96	APELS	2.00	.04	.20	.18	.18	
	MCELS	1.97	.01	.19	.18	.17	
ψ_{11}	EXELS	101.79	1.79	21.51	19.56	19.55	
100.00	APELS	99.89	-.11	20.13	18.93	18.83	
	MCELS	102.18	2.18	21.22	19.62	19.65	
ψ_{12}	EXELS	-37.78	-.28	23.51	21.33	21.22	
-37.50	APELS	-31.75	5.75	19.13	19.33	20.07	
	MCELS	-38.34	-.84	22.38	21.51	21.42	
ψ_{13}	EXELS	-103.85	-1.63	21.24	19.59	19.56	
-102.225	APELS	-102.45	-.22	20.26	19.13	19.04	
	MCELS	-104.05	-1.82	21.18	19.51	19.50	
ψ_{22}	EXELS	151.30	-4.95	36.54	30.46	30.71	
156.25	APELS	146.44	-9.81	33.55	28.32	29.83	
	MCELS	152.53	-3.72	36.81	30.38	30.46	
ψ_{23}	EXELS	28.91	.83	23.24	20.25	20.16	
28.077	APELS	23.11	-4.97	19.06	18.56	19.12	
	MCELS	28.95	.87	21.93	20.48	20.40	
ψ_{33}	EXELS	107.89	1.33	21.71	20.69	20.63	
106.556	APELS	107.02	.47	21.08	20.35	20.25	
	MCELS	108.17	1.62	21.87	20.46	20.42	

β_2, ψ_{12}, and ψ_{23} should be rescaled by 10^{-4} and ψ_{22}, by 10^{-8}.

Table 3.2 Estimates, error measures and coverage rates for Design 2

Parameter True value	Method	Mean	Bias	Est'ed $\sqrt{Var}$	Std Error	$\sqrt{MSE}$	Coverage Rate
β_1	EXELS	124.93	-.07	1.54	1.52	1.52	.98
125.0	APELS	122.45	-2.55	1.48	1.51	2.96	.61
	MCELS	125.18	.18	1.51	1.72	1.72	.97
β_2	EXELS	87.31	-.19	3.23	3.78	3.77	.91
87.50	APELS	86.55	-.95	3.01	3.50	3.61	.90
	MCELS	86.96	-.54	3.18	4.17	4.18	.92
β_3	EXELS	-131.08	.09	1.58	1.58	1.57	.96
-131.163	APELS	-127.94	3.23	1.57	1.63	3.61	.43
	MCELS	-131.33	-.17	1.56	1.77	1.77	.96
σ^2	EXELS	1.97	.01	.20	.18	.18	
1.96	APELS	2.00	.04	.20	.19	.19	
	MCELS	1.95	-.01	.21	.18	.18	
ψ_{11}	EXELS	89.95	-10.05	29.05	31.16	32.59	
100.00	APELS	104.06	4.06	22.61	23.21	23.45	
	MCELS	81.93	-18.07	27.34	34.41	38.71	
ψ_{12}	EXELS	-66.22	-1.27	41.29	42.84	42.64	
-64.952	APELS	-23.22	41.73	31.94	35.27	54.53	
	MCELS	-74.63	-9.68	47.91	42.49	43.37	
ψ_{13}	EXELS	-92.05	10.17	30.54	33.50	34.85	
-102.225	APELS	-110.60	-8.38	23.81	25.27	26.50	
	MCELS	-82.91	19.32	27.55	37.02	41.59	
ψ_{22}	EXELS	456.15	-12.60	103.15	109.73	109.90	
468.750	APELS	408.50	-60.25	79.54	79.53	99.46	
	MCELS	484.20	15.45	97.96	118.85	119.26	
ψ_{23}	EXELS	49.44	.81	41.82	44.39	44.17	
48.631	APELS	8.83	-39.80	34.32	37.95	54.86	
	MCELS	55.97	7.34	49.83	44.15	44.54	
ψ_{33}	EXELS	96.28	-10.27	32.83	36.55	37.79	
106.556	APELS	119.69	13.14	25.87	28.21	30.99	
	MCELS	86.39	-20.17	29.33	39.73	44.38	

β_2, ψ_{12}, and ψ_{23} should be rescaled by 10^{-4} and ψ_{22}, by 10^{-8}.

Table 3.3 Estimates, error measures and coverage rates for Design 3

Parameter True value	Method	Mean	Bias	Est'ed $\sqrt{Var}$	Std Error	$\sqrt{MSE}$	Coverage Rate
β_1 125.0	EXELS	124.78	-.22	1.44	1.24	1.25	.98
	APELS	124.02	-.98	1.42	1.23	1.57	.95
	MCELS	124.78	-.22	1.45	1.24	1.25	.98
β_2 87.50	EXELS	87.51	.01	1.93	2.00	1.99	.97
	APELS	87.06	-.44	1.91	2.01	2.05	.93
	MCELS	87.50	.00	1.94	2.02	2.01	.96
β_3 -131.163	EXELS	-130.91	.25	1.49	1.21	1.23	.99
	APELS	-129.91	1.26	1.47	1.21	1.74	.93
	MCELS	-130.92	.25	1.49	1.20	1.22	.99
σ^2 1.96	EXELS	1.93	-.03	.19	.19	.19	
	APELS	1.94	-.02	.19	.20	.20	
	MCELS	1.92	-.04	.19	.18	.19	
ψ_{11} 100.00	EXELS	97.67	-2.33	20.22	25.57	25.55	
	APELS	95.83	-4.17	19.24	24.40	24.64	
	MCELS	98.32	-1.68	20.37	25.79	25.72	
ψ_{12} -37.50	EXELS	-35.32	2.18	22.81	23.73	23.71	
	APELS	-28.65	8.85	19.34	21.02	22.71	
	MCELS	-36.16	1.34	21.96	23.58	23.50	
ψ_{13} -102.225	EXELS	-100.74	1.49	20.53	25.43	25.35	
	APELS	-99.46	2.77	19.73	24.37	24.41	
	MCELS	-101.20	1.03	20.75	25.69	25.58	
ψ_{22} 156.25	EXELS	157.91	1.66	38.50	35.19	35.05	
	APELS	151.96	-4.29	35.05	31.95	32.08	
	MCELS	159.49	3.24	38.82	35.13	35.11	
ψ_{23} 28.077	EXELS	25.99	-2.09	22.40	24.61	24.58	
	APELS	19.61	-8.47	19.32	22.19	23.64	
	MCELS	26.29	-1.78	21.33	24.52	24.46	
ψ_{33} 106.556	EXELS	105.87	-.69	21.75	26.13	26.01	
	APELS	105.18	-1.37	21.04	25.19	25.10	
	MCELS	106.42	-.14	21.93	26.39	26.26	

β_2, ψ_{12}, and ψ_{23} should be rescaled by 10^{-4} and ψ_{22}, by 10^{-8}.

4. An Example

In this section, we analyze the dataset considered by Vonesh and Carter(1992). Several nonlinear models are fit to identify the underlying variance-covariance structure. We obtain MCELS estimates for each model. Estimates, plots, and the Akaike Information Criterion (AIC) are used to choose a final model.

4.1 Models

Standard low flux membrane dialyzers are used in hemodialysis to treat patients with ending stage renal disease. Their water transport kinetics are analyzed by linear relationship between the ultrafiltration rate (UFR in ml/hr) at which water is removed and the transmembrane pressure (TMP in mm Hg) which is exerted on the dialyzer membrane (Vonesh and Carter (1987)). After that, high flux membrane dialyzers have been introduced to reduce the time spent by patients on hemodialysis. Unlike their low flux dialyzers the water transport kinetics of high flux dialyzers are characterized by a nonlinear relationship between UFR and TMP.

Vonesh and Carter (1992) describe the relationship between UFR and TMP by the nonlinear function:

$$UFR = \alpha_1\{1 - \exp[-\alpha_2(TMP - \alpha_3)]\} \qquad (4.1)$$

where α_1 is the maximum UFR one can attain due to protein polarization, α_2 is a hydraulic permeability transport rate, and α_3 is the transmembrane pressure required to offset patient oncotic pressure.

We also consider a simple nonlinear model given by

$$UFR = \beta_1 + \beta_3\exp(-\beta_2 TMP). \qquad (4.2)$$

Here, we know that $\beta_1 = \alpha_1$, $\beta_2 = \alpha_2$ and $\beta_3 = \alpha_1 e^{\alpha_2\alpha_3}$.

These two response functions are used to characterize the water transport characteristics of 20 high flux membrane dialyzers. The dialyzers are evaluated *in vitro* using bovine blood at blood flow rates (Qb) of either 200 ml/min or 300 ml/min. UFR was measured at seven different TMPs per dialyzer.

To analyze the data, we use the nonlinear random coefficient model

$$UFR_{ij} = \delta_i(\alpha_{11i}\{1 - \exp[-\alpha_{21i}(TMP_{ij} - \alpha_{31i})]\} + \varepsilon_{1ij}) + (1-\delta_i)(\alpha_{12i}\{1 - \exp[-\alpha_{22i}(TMP_{ij} - \alpha_{32i})]\} + \varepsilon_{2ij}) \qquad (4.3)$$

where $\delta_i = 1$ if Qb = 200 ml/min and 0 if Qb = 300 ml/min,

$$\boldsymbol{\alpha}_{1i} = (\alpha_{11i}, \alpha_{21i}, \alpha_{31i})' \sim NID(\boldsymbol{\alpha}_1, \Sigma_1),$$
$$\boldsymbol{\alpha}_{2i} = (\alpha_{12i}, \alpha_{22i}, \alpha_{32i})' \sim NID(\boldsymbol{\alpha}_2, \Sigma_2),$$
$$\varepsilon_{1ij} \sim NID(0, \sigma_1^2), \quad \varepsilon_{2ij} \sim NID(0, \sigma_2^2)$$

and $\{\boldsymbol{\alpha}_{1i}\}, \{\boldsymbol{\alpha}_{2i}\}, \{\varepsilon_{1ij}\}$ and $\{\varepsilon_{2ij}\}$ are independent.

We also consider the model derived from model (4.2)

$$UFR_{ij} = \delta_i[\beta_{11i} + \beta_{31i}\exp(-\beta_{21i}TMP_{ij}) + \varepsilon_{1ij}] + (1-\delta_i)[\beta_{12i} + \beta_{32i}\exp(-\beta_{22i}TMP_{ij}) + \varepsilon_{2ij}] \qquad (4.4)$$

where $\delta_i = 1$ if Qb = 200 ml/min and 0 if Qb = 300 ml/min,

$$\boldsymbol{\beta}_{1i} = (\beta_{11i}, \beta_{21i}, \beta_{31i})' \sim NID(\boldsymbol{\beta}_1, \Gamma_1),$$
$$\boldsymbol{\beta}_{2i} = (\beta_{12i}, \beta_{22i}, \beta_{32i})' \sim NID(\boldsymbol{\beta}_2, \Gamma_2),$$
$$\varepsilon_{1ij} \sim NID(0, \sigma_1^2), \quad \varepsilon_{2ij} \sim NID(0, \sigma_2^2)$$

and $\{\boldsymbol{\beta}_{1i}\}, \{\boldsymbol{\beta}_{2i}\}, \{\varepsilon_{1ij}\}$ and $\{\varepsilon_{2ij}\}$ are independent.

By imposing assumptions, the following 6 models are considered in the data analysis.

(a) Model Ⅰ: model (4.3)
(b) Model Ⅱ: model (4.4)
(c) Model Ⅲ: model (4.3) with the assumption that $\sigma_1^2 = \sigma_2^2$
(d) Model Ⅳ: model (4.4) with the assumption that $\sigma_1^2 = \sigma_2^2$
(e) Model Ⅴ: model (4.3) with the assumption that $\Sigma_1 = \Sigma_2$ and $\sigma_1^2 = \sigma_2^2$
(f) Model Ⅵ: model (4.4) with the assumption that $\Gamma_1 = \Gamma_2$ and $\sigma_1^2 = \sigma_2^2$.

Each coefficient can be either random or fixed. There are $2^3 = 8$ cases for each model in classification of 3 coefficients as random or fixed. With these models, we will try to find the best model in the next section.

4.2 Model Selection

In this section, we use plots and the Akaike Information Criterion (AIC) to choose the best model.

First of all, we obtain estimates for Model Ⅰ-Ⅵ assuming that all coefficients are random, and plot the predicted values based on the estimates. We obtain the predicted values and prediction limits using the Monte Carlo method. For example, the procedure for Model Ⅰ with Qb=200 is :

(a) Compute $f(t_i, \boldsymbol{\alpha}_1^{(l)})$ for t_i =20, 21,⋯, 310 since t_{ij}'s of real data are in (20, 310) and l =1, 2, ⋯, 5000 as follows.

(a.1) Generate the random coefficients from a normal distribution with mean and variance-covariance parameters given by corresponding estimates.

$$\boldsymbol{\alpha}_1^{(l)} = \hat{\boldsymbol{\alpha}}_1 + \hat{\Sigma}_1^{1/2} \boldsymbol{z}^{(l)}$$

where $\boldsymbol{z}^{(l)} \sim NID(0, I)$.

(a.2) With generated random coefficients, $\boldsymbol{\alpha}_1^{(l)}$, compute the function values at each t_i as

$$f(t_i, \boldsymbol{\alpha}_1^{(l)}) = \alpha_{11}^{(l)}[1 - \exp\{-\alpha_{21}^{(l)}(t_i - \alpha_{31}^{(l)})\}].$$

(b) At each t_i, compute the predicted value $\bar{f}(t_i)$, as the mean of these values and prediction variance Var_i, using the sample variance and $\hat{\sigma}_1^2$ as

$$\bar{f}(t_i) = (1/5000) \sum_{l=0}^{5000} f(t_i, \boldsymbol{\alpha}_1^{(l)})$$

and

$$Var_i = (1/5000) \sum_{l=0}^{5000} f(t_i, \boldsymbol{\alpha}_1^{(l)})^2 - \bar{f}(t_i)^2 + \hat{\sigma}_1^2.$$

Prediction limits at each t_i are obtained as

$(\bar{f}(t_i) \pm t(7, 0.025)\sqrt{Var_i})$.

From Figures 4.1 - 4.3, we see unacceptable fitted lines and prediction limits for Model Ⅲ-Ⅳ. Their deviation from observations is due to the fact that the same variance assumption between two groups may not be correct. Table 4.1 shows significant differences between variances, σ_1^2 and σ_2^2 in both Model Ⅰ and Ⅱ.

Table 4.2 Variance Estimates in Model Ⅰ and Ⅱ

	Model Ⅰ		Model Ⅱ	
	estimate	std.error	estimate	std.error
σ_1^2	79975.0	20840.0	79998.4	23105.8
σ_2^2	39257.5	12290.8	39263.4	11981.8

From Figures 4.1 - 4.3, we find that reparameterized models have milder prediction limits, especially for smaller TMP values. It shows that the original model specification gives a better description of data variability which increases with level of TMP.

We also observed that some estimates of the parameters in variance covariance matrices for the two groups can not be regarded as the same for both groups. Thus, we conclude that neither the same variance nor the same variance-covariance matrix assumption is appropriate for this dataset.

It is interesting to observe that Figure 4.3 shows very good fitted prediction lines for Model Ⅴ and Ⅵ. Even though the assumption of equal variance-covariance parameters may not be valid, prediction based on these models seem to do well in our examples. Note that Vonesh and Carter (1992) assume the variance- covariance matrices are the same for both groups.

To help in choosing the best model, we compute the AIC of each possible combination of random coefficients for Model Ⅰ and Ⅱ. Here, we define the AIC function based on the MCELS objective function instead of based on the actual log likelihood function :

$$AIC = Q_{MCELS}(\hat{\theta}_{MCELS}) + 2m$$

where m is the number of parameters in the model.

Table 4.3 AIC for different models

random coefficient	Model Ⅰ AIC	random coefficient	Model Ⅱ AIC
$\alpha_1, \alpha_2, \alpha_3$	1772.63	$\beta_1, \beta_2, \beta_3$	1772.27
α_1, α_2	1761.59	β_1, β_2	1771.21
α_1, α_3	1766.29	β_1, β_3	1766.29
α_2, α_3	1778.89	β_2, β_3	1779.79
α_1	1759.62	β_1	1778.34
α_2	1758.98	β_2	1786.24
α_3	1817.14	β_3	1817.12
none	1813.74	none	1813.74

From Table 4.3, we see that the AIC attains its smallest value for Model Ⅰ when α_1 or α_2 is selected as a random coefficient. So, both coefficients could be regarded as random. We also find that the AIC has almost the same value when both α_1 and α_2 are selected as random coefficients. Thus, based on the AIC, Model Ⅰ with random coefficients α_1 and α_2 and remaining coefficient α_3 fixed is the best model. The MCELS estimates and the corresponding standard errors for this model are given by

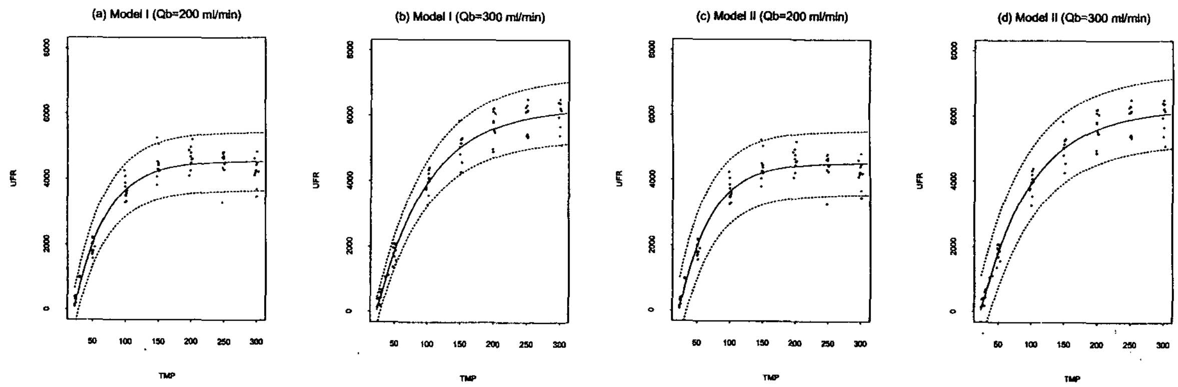

Figure 4.1 Mean response and prediction limits for Model Ⅰ and Ⅱ

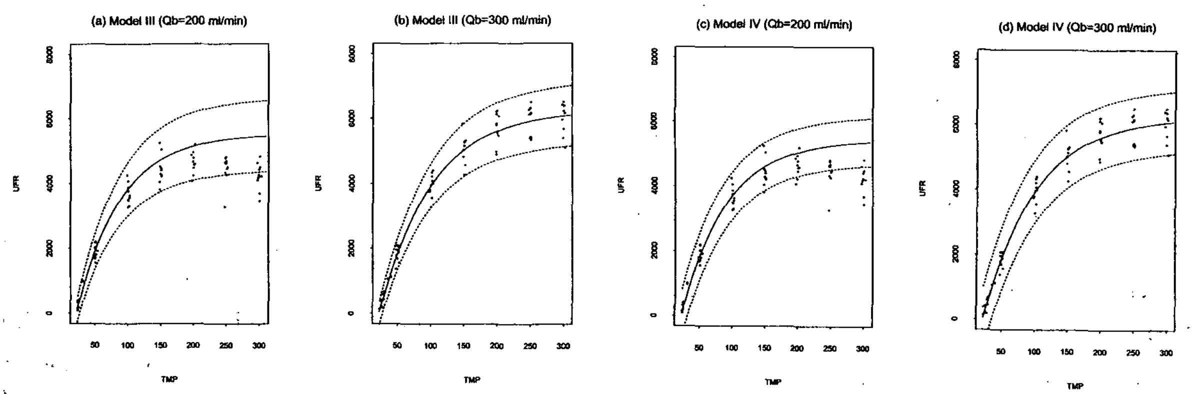

Figure 4.2 Mean response and prediction limits for Model Ⅲ and Ⅳ

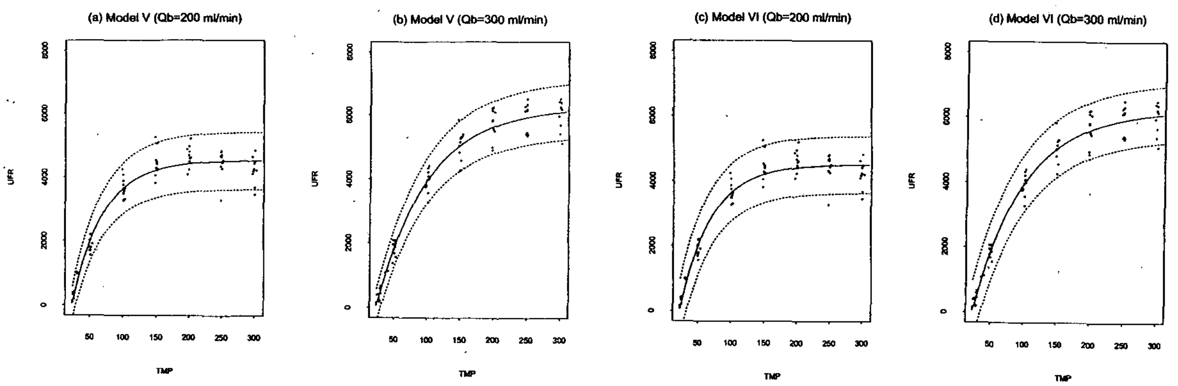

Figure 4.3 Mean response and prediction limits for Model Ⅴ and Ⅵ

$$\hat{\alpha}_1=\begin{bmatrix} 4534 \\ (118) \\ .0213 \\ (.0020) \\ 22.55 \\ (.383) \end{bmatrix},\quad \hat{\Sigma}_1=\begin{bmatrix} 84450 & -.4531 \\ (63960) & (.689) \\ -.4331 & .0000105 \\ (.689) & (.0000118) \end{bmatrix},$$

$$\hat{\sigma}_1^2=\begin{matrix} 81120 \\ (23200) \end{matrix}$$

$$\hat{\alpha}_2=\begin{bmatrix} 6230 \\ (137) \\ .0129 \\ (.00054) \\ 22.41 \\ (.698) \end{bmatrix},\quad \hat{\Sigma}_2=\begin{bmatrix} 170200 & -.1747 \\ (55710) & (.230) \\ -.1747 & .0000021 \\ (.230) & (.0000017) \end{bmatrix},$$

$$\hat{\sigma}_2^2=\begin{matrix} 40320 \\ (11940) \end{matrix}.$$

5. Discussion

Nonlinear random coefficient models are being used in various fields. Several different estimation methods are implemented by different software packages and being used by practitioners. Some of the estimation procedures, like estimates based on pooling individual estimates of $\boldsymbol{\beta}_i$ or those based on first or second order Taylor expansion of $f(\boldsymbol{a}_{ij}, \boldsymbol{\beta}_i)$, lead to inconsistent estimators. The extended least squares estimation method gives consistent estimates of parameters. However, in most applications, the objective function for ELS method does not have closed form expressions and hence obtain approximation to ELS estimates. Based on our simulations and the theoretical results discussed in Kim (1997), we recommend the MCELS estimators.

The MCELS methods depends on the normality assumption of the random coefficients $\boldsymbol{\beta}_i$. For the model considered in our simulation study, the MCELS estimation is robust to misspecification of the distribution. This is because the ELS method is one of the general class of generalized estimating equations (GEE) and has the properties of GEE estimators. Hartford and Davidian (1999) study the consequences of nonnormality of the random coefficients on approximate maximum likelihood estimates based on first order and Laplace approximations of the likelihood functions.

Model selection in nonlinear random coefficients is still in fledgeling stage. Davidian and Giltinan (1995) and Pinheiro and Bates (1995) discuss some general guidelines for model selection. We have used plots and AIC type criterion to select a model. In addition, it would be helpful to develop test criteria for model selection and study their performance in finite samples.

References

[1] Beal, S.L., Asymptotic properties of optimization estimators for the independent not identically distributed case with application to extended least squares estimators, Technical Report. Division of Clinical Pharmacology. University of California, San Francisco. 1984.

[2] Davidian, M and D.M. Giltinan, "Some general estimation methods for nonlinear mixed effects models", *Journal of Biophar -maceutical Statistics*, Vol3(1993), pp23-55.

[3] Davidian,M and D.M.Giltinan, Nonlinear Models for Repeated Measurement Data, Chapman & Hall., New York, 1995.

[4] Hartford, A. and M. Davidian, "Consequences of misspecifying assumptions in nonlinear mixed effects", Technical Report, Department of Statistics, North Carolina State University, Raleigh, N.C.

1999.
[5] Kim, S. Extended least squares estimation using Monte Carlo method in nonlinear random coefficient models, Ph.D. Dissertation, Department of Statistics, North Carolina State University, Raleigh, N.C. 1997
[6] Pinheiro, J.C. and D.M. Bates, "Approximations to the log-likelihood function in nonlinear mixed effects model". *Journal of Computational Graphical Statistics*, Vol.4(1995), pp12-35.
[7] Ramos, R.Q., Estimation of nonlinear mixed effects and random coefficient models. Ph.D. Dissertation, North Carolina State University, Raleigh, N.C. 1993.
[8] Ramos, R.Q. and S.G. Pantula, "Estimation of nonlinear random coefficient models", *Statistics & Probability Letters*, Vol. 24(1995), pp49-56.
[9] Sheiner, L.B. and S.L. Beal, "Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis -Menten model: Routine clinical pharmacokinetic data", *Journal of Pharmacokinetics and Biopharmaceutics*, Vol. 8(1980), pp553-571.
[10] Sheiner, L.B. and S.L. Beal, "Phamacokinetic parameter estimates from several least squares procedures: Superiority of extended least squares". *Jouranl of Pharmacokinetics and Biophamaceutics*, Vol.13, NO2(1985), pp185-201.
[11] Vonesh, E.F. and R.L. Carter, "Mixed effects nonlinear regression for unbalanced repeated measures", *Biometrics*, Vol. 48(1992), pp1-17.

● 저자소개 ●

김성연

1981 서울대학교 계산통계학과 (이학사)
1983 서울대학교 대학원 통계학전공 (이학석사)
1997 North Carolina State University (통계학박사)
현재 동아대학교 경영정보과학부 교수
관심분야 선형모형, 비선형모형, 혼합모형