A C T A U N I V E R S I T A T I S L O D Z I E N S I S
FOLIA OECONOMICA 132, 1993

*Dietrich von Rosen**

COMBINING RANDOM COEFFICIENT REGRESSION WITH THE POTTHOFF AND ROY MODEL IN GROWTH CURVE ANALYSIS

Abstract. In this paper ideas when modelling growth curve data with a random coefficient regression model are put together with ideas from the application of the Growth Curve model. In general, results from [L u n d b y e-C h r i s t e n s e n (1988)] are extended. It is shown that when there exist certain connections between the mean structure and the covariance structure it is possible to obtain straightforward algorithms, leading to estimators without any heavy calculations.

Key words: Growth curve analysis, Potthoff Roy model, random coefficient regression.

1. INTRODUCTION

This paper will treat some extensions of the ordinary Growth Curve model, (GC), which was introduced by [P o t t h o f f and R o y (1964)], and extensions of a random coefficient regression model, (RCR), both applicable to the analysis of growth curves. Data which usually are to be analysed by the models consists of several short time series, each containing repeated measurements on some "individual". For example, when studying the growth rate of a cohort of children during puberty, body height may have been

* Professor at the Department of Mathematical Statistics of the University of Stockholm.

measured every half year from an age of 11 up to an age of 18. The problem with this type of data is that the repeated measurements for various reasons, often are correlated.

In order to apply the GC model the repeated measurements have to be sampled at the same "time" points. In the above example this condition would have been satisfied if each child would have been sampled at an age of 11, 11.5, ..., 18. Note that no missing values are allowed in the ordinary GC model.

Over the years several extensions of the model have been put forward which allow for missing values as well as data which are sampled at different time points. For a review of the model see von [R o s e n (1989a)]. Data with some missing values and irregularly dispersed sample points is also a type of data which is much more frequent than data suitable for the Potthoff and Roy model. When studying growth curves, in order to overcome these difficulties with data, one is often thinking in terms of individual growth curves. The methodology is then to estimate the whole bunch of individual growth curves, one by one, and thereafter pool together the information from each single curve. In this spirit of estimating individual growth curves, a common approach to model data is by assuming that there exists a separate parameter for each individual but where the parameter is sampled from a population of parameters. This immediately leads us to a RCR model. Moreover, note that in comparison with the GC model the RCR model is more flexible when modelling the mean structure. On the other, one usually has to suppose that the measurement errors within each individual are independently distributed or have a known covariance matrix, which in many applications is not realistic. For the ordinary GC model we do not have to suppose independence within individuals since an arbitrary covariance matrix is assumed for the repeated measurements. To clearly see the difference between the RCR and GC models we define these models in the next. In the definitions as well as in the rest of this paper p.d. stands for positive definite (p.s.d. positive semi definite), $\rho(\cdot)$ for the rank and $|$ for a conditional distribution.

Definition 1.1. Growth Curve model: Let $Y: p \times n$, $A: p \times q$ $q \leq p$, $B: q \times k$, $C: k \times n$, $\Sigma: p \times p$ p.d., where $\rho(C) + p \leq n$ holds. The columns of Y are independently p-variate normally

distributed with an unknown dispersion matrix Σ and $E[Y] = ABC$ where A and C are known design matrices and B is an unknown parameter matrix.

Definition 1.2. RCR model: Let $Y_i: 1 \times n_i$, $B_i: 1 \times k$, $C_i: k \times n_i$, σ and ψ are scalars, the C_i are known, σ and ψ are unknown parameters and B_i stochastic variables $i = 1, 2, \ldots, m$. The elements in $Y_i|B_i$ are normally distributed, $E[Y_i|B_i] = B_iC_i$, $D[Y_i|B_i] = \psi I$, and B_i is normally distributed with an unknown mean B and covariance matrix $\sigma^2 I$.

Observe that elements in Y in both models are correlated. Furthermore, in general one may assume that each individual follows a growth process, especially some stationary or non-stationary Markov process or a Wiener process with drift, which in many realistic situations is independent of the measurement errors. In these cases we have a model with the same mean as in definition 1.2 but with a covariance structure

$$D[Y_i] = \sigma^2 H_i + \psi I \qquad (1.1)$$

where H_i is a prespecified p.s.d. matrix whose structure depends on the assumptions on the growth process. However, when $H_i = C_iC_i$ we have the RCR model. In the sequel we do not make any destinction between a model with covariance structure given by (1.1) and a model with the same covariance structure as in definition 1.2 and we will call both RCR models. In fact, any model of the form $D[Y_i] = \sigma^2 H_i + \psi I$; $E[Y_i] = BC_i$ can be obtained by aid of a mixed linear model.

Hitherto we have supposed that correlated observations appear in time but we can also measure not one, but several characteristics at each time point. For example in the cohort of children, previously mentioned, body height, body weight and some hormones may have been sampled. Then, presumably we also have a correlation between body height, body weight and hormone levels. Thus, we have two different correlation structures which must be taken into consideration; (i) the repeated measurements over "time" and (ii) the repeated measurements within each time point.

These two correlation structures have simultaneously been mo-

delled successfully by [L u n d b y e-C h r i s t e n s e n (1988) and the main purpose with this note is to show some immediate extensions which all have natural applications. As an example suppose that for the above mentioned children every half year a circadian rhythm is measured, for instance, ten times during a 24 hours period. Then we have to combine information from a long--range time serie with a short-range time serie. This can be done with the help of the extensions in section 3.

2. A MODEL WITH A UNIVARIAT GROWTH FACTOR

In this section we will see how Lundbye-Christensen modelled the correlation structures (i) and (ii) in the previous section and we restate some of his results. The idea put forward by Lundbye-Christensen was to use a hybrid between the GC and RCR models. It will be supposed that the growth process can be described with the help of a linear model. Let 1 stand for the column vector of ones and let $\mathcal{C}(\bullet)$ signify a column vector space. Furthermore, as a convention in this paper, $Y: p \times n$ is said to be matrix normally distributed if $vec(Y) \sim N_{pn}(E[vec(Y)], I \otimes \Sigma)$ and denoted $Y \sim N_{p,n}(E[Y], I \otimes \Sigma)$ where $vec(\bullet)$ is the vec operator and $\otimes$ the right Kronecker product.

Definition 2.1. A univariate growth factor: Lundbye-Christensen.

Let $Y_i: p \times n_i$, $B: 1 \times k$, $C_i: k \times n_i$, $H_i: n_i \times n_i$ p.s.d., $\Sigma: p \times p$ p.d. and σ a scalar, $i = 1, 2, \ldots, m$. C_i and Hi are known and B, σ and Σ are unknown parameters. The elements in Y_i are matrix normally distributed with $E[Y_i] = 1BC_i$, $D[Y_i] = \sigma^2 H_i \otimes 11' + I \otimes \Sigma$ and Y_i is independent of Y_j $i \neq j$.

From a view of applications this formulation with a separate model for each individual is the natural one. However, we will estimate the parameters simultaneously and as a basis for our discussion we will in the subsequent utilize an equivalent formulation of the model. Define partitioned matrices $Y = Y_1: \ldots: Y_m$, $C = C_1: \ldots: C_m$ and $H = diag(H_1, \ldots, H_m)$, $\Psi = \Sigma/\sigma^2$ and then we get ($n = \Sigma_i n_i$)

$$Y \sim N_{p,n}(1BC, \sigma^2(H \otimes 11' + I \otimes \Psi)).$$

From now on we are going to derive a canonical version of this model. Without loss of generality we assume that

$$Y \sim N_{p,n}(1BC, \sigma^2(I \otimes 11' + D \otimes \Psi)) \tag{2.1}$$

where D is a diagonal matrix. Let $\Gamma = (\Gamma_1 : \Gamma_2)$ be an orthogonal matrix where $\mathcal{C}(\Gamma_1) = \mathcal{C}(1)$ and $\mathcal{C}(\Gamma_2) = \mathcal{C}(1)^{\perp}$, and set $Z = \Gamma' Y$. The equation (2.1) is equivalent to (set $k = \Gamma_1 11' \Gamma_1$, $\Omega = \Gamma' \Psi \Gamma$)

$$Z = (Z_1' : Z_2')' \sim N_{p,n}\left(\begin{pmatrix} \eta C \\ 0 \end{pmatrix}, \sigma^2\left(I \otimes \begin{pmatrix} k & 0 \\ 0 & 0 \end{pmatrix} + D \otimes \Omega\right)\right)$$

where η stands for the new parameters after reparametrization. Since $E[Z_2] = 0$ we condition Z_1 with respect to Z_2. Hence, as an inference basis we take

$$Z_1 | Z_2 \sim N_{1,n}(\eta C + \delta' Z_2, \sigma^2(kI + D\tau)) \tag{2.2}$$

$$Z_2 \sim N_{p-1,n}(0, \sigma^2 D \otimes \Delta) \tag{2.3}$$

where the parameters τ, Δ and δ are defined through

$$\Omega = \begin{pmatrix} \tau + \delta' \Delta \delta & \delta' \Delta \\ \Delta \delta & \Delta \end{pmatrix}.$$

Note that the reparametrization is one to one. Set $\beta = (\eta : \delta')$ and $T = (C' : Z_2')'$. Now $\sigma^2 \Delta$ is estimated marginally from (2.3) and from (2.2) follows that (suppose that T is of full rank)

$$\hat{\beta}(\tau) = Z_1 (kI + \tau D)^{-1} T (T(kI + \tau D)^{-1} T')^{-1} \tag{2.4}$$

$$n\hat{\sigma}^2(\tau) = (Z_1 - \hat{\beta}(\tau)T)(kI + \tau D)^{-1}(Z_1 - \hat{\beta}(\tau)T)' \tag{2.5}$$

and $\hat{\tau}(\beta, \sigma)$ is obtained by maximizing

$$\hat{\sigma}^2(\tau)^{-n/2} |kI + \tau D|^{-1/2} \exp(-1/2\, 1/\hat{\sigma}^2(\tau)(Z_1 - \hat{\beta}(\tau)T)(kI + \tau D)^{-1}(Z_1 - \hat{\beta}(\tau)T)') \tag{2.6}$$

where $\hat{\sigma}^2(\tau)$ and $\hat{\beta}(\tau)$ are regarded as fixed. Hence, by aid of (2.4) - (2.6) we have constructed an iteration scheme including fairly simple calculations. In (2.6) we are searching for an one-dimensional parameter estimator and since $kI + \tau D$ is a diagonal matrix there will be no numeric problems. At least as far as we

have relevant data, which do not imply that estimators are found on the boundary of the parameter space, e.g. $\sigma^2 = 0$, or that data indicates that their is no clear maximum.

3. USEFUL EXTENSIONS

In this section we will consider some very natural extensions of the model given in section 2. The first which we are thinking of, is that in definition 2.1 there is one growth factor which influences, in the same manner, each at the different time points observed variables (remember the vector 1). Before presenting various extensions we will fix a terminology which we hope will illuminate the extensions. In principle we will discuss three different mean structures:

$E[X] = 1BC$ called a univariate model and presented in section 2 (3.1)

$E[X] = ABC$ called a multivariate model (3.2)

$E[X] = A_1B_1C_1 + A_2B_2C_2$, $\mathcal{C}(A_2) \subseteq \mathcal{C}(A_1)$ or $\mathcal{C}(C_2') \subseteq \mathcal{C}(C_1')$ (3.3)

called a multivariate model with different responses.

The mean structure in (3.2) is identical to the mean structure in the GC model given by definition 1.1. The extension in (3.3) is inspired by a straightforward extension of the GC model presented by [von Rosen (1989b)]. One drawback with the ordinary GC model is that each individual is supposed to have a growth response of the same type, e.g. the means follow a polynomial of the same order. The extension in (3.3) is designed to handle situations where individuals, to some extent, have different growth responses, e.g. follow polynomials of different orders. Observe that in this paper it means that the repeated measurements within each time point follow, for example, different polynomials. Especially the model may be useful when considering the combination of short-range time series with long-range time series, mentioned in the introduction.

Extension 1. A multivariate model.

Let $Y: p \times n$, $A: p \times q$, $B: q \times k$, $C: k \times n$, $E: p \times m$, $H: n \times n$ p.s.d., $\Psi: p \times p$ p.d. and σ a scalar where A, C, E and H are known and B, σ and Ψ are unknown parameters.

$$Y \sim N_{p,n}(ABC, \sigma^2(H \otimes EE' + I \otimes \Psi))$$

where $\mathcal{C}(E) \subseteq \mathcal{C}(A)$.

In comparison with definition 2.1 we have a more general mean structure, as well as a more general covariance structure. As observed above the mean is identical to the mean in the GC model given by definition 1. The extended covariance structure is motivated since the effects on the covariance structure from the growth process may differ between those variables observed at each time point.

By aid of some matrix manipulations we describe now how to reduce extension 1 to a similar model as the model in (2.2) and (2.3). First note that the extension is equivalent to

$$Y \sim N_{p,n}(ABC, \sigma^2(I \otimes EE' + D \otimes \Psi))$$

where D as before is diagonal and then we set $Z = (Z_1' : Z_2')' = \Gamma'Y$ where $\Gamma = (\Gamma_1 : \Gamma_2)$ is orthogonal and $\mathcal{C}(\Gamma_1) = \mathcal{C}(A)$, $\mathcal{C}(\Gamma_2) = \mathcal{C}(A)^{\perp}$. Hence we get

$$Z = (Z_1' : Z_2')' \sim N_{p,n}\left(\begin{pmatrix}\eta C \\ 0\end{pmatrix}, \sigma^2\left(I \otimes \begin{pmatrix}K & 0 \\ 0 & 0\end{pmatrix} + D \otimes \Omega\right)\right)$$

where $K = \Gamma_1 EE' \Gamma_1$ and new parameters η and $\Omega = \Gamma'\Psi\Gamma$.

Now we have reduced the model to be of the same form as in section 2 and obviously we obtain, similarly to (2.4) - (2.5), equations which give estimators. However, instead of an one-dimensional maximization procedure, as for (2.6), we have to work with a $\rho(A)$-dimensional. In practice $\rho(A)$ is fairly small and from a view of computations the model is still not too difficult. The equations which correspond to (2.4) - (2.6) are obtained by conditioning Z_1 with respect to Z_2, i.e.

$$Z_1 | Z_2 \sim N_{\rho(A_1),n}(BT, \sigma^2(I \otimes K + D \otimes \tau))$$

where $B = (\eta : \delta')$, $T = (C' : Z_2')'$ and

$$\Omega = \begin{pmatrix}\tau + \delta'\Delta\delta & \delta'\Delta \\ \Delta\delta & \Delta\end{pmatrix}.$$

Thus,

$$\mathrm{vec}(\hat{B}(\tau)) = ((T \otimes I)(I \otimes K + D \otimes \tau)^{-1}(T \otimes I))^{-1}(T' \otimes I)(I \otimes K + D \otimes \tau)^{-1}\mathrm{vec}(Z_1),$$

$$n\hat{\sigma}^2(\tau) = (\mathrm{vec}(Z_1 - \hat{B}(\tau)T))'(I \otimes K + D \otimes \tau)^{-1}\mathrm{vec}(Z_1 - \hat{B}(\tau)T)$$

and $\hat{\tau}(B, \sigma)$ is obtained by maximizing

$$\hat{\sigma}^2(\tau)^{-np/2}|I \otimes K + D \otimes \tau|^{-1/2}\exp(-1/2\ 1/\hat{\sigma}^2(\tau)(\mathrm{vec}(Z_1 - \hat{B}(\tau)T))'(I \otimes K + D \otimes \tau)^{-1}\mathrm{vec}(Z_1 - \hat{B}(\tau)T)).$$

Extension 2. A multivariate model with different responses

Let Y: $p \times n$, A_i: $p \times q_i$, B_i: $q_i \times k_i$, C_i: $k_i \times n$, $i = 1, 2$, E: $p \times m$, H: $p \times p$ p.s.d., Ψ: $p \times p$ p.d. and σ a scalar where A_i, C_i, E and H are known and B_i, σ and Ψ are unknown parameters.

$$Y \sim N_{p,n}(A_1B_1C_1 + A_2B_2C_2,\ \sigma^2(H \otimes EE' + I \otimes \Psi)),\quad \mathcal{C}(A_2) \subseteq \mathcal{C}(A_1)$$

Case (i): $\mathcal{C}(E) \subseteq \mathcal{C}(A_2)$

Case (ii): $\mathcal{C}(E) \subseteq \mathcal{C}(A_1)$ but $\mathcal{C}(A_2) \not\supseteq \mathcal{C}(E)$.

The derivation of the estimators is similar to the one for extension 1. We are going to utilize the decomposition

$$\mathcal{C}(A_1) = \mathcal{C}(A_2) \oplus \mathcal{C}(A_1) \cap \mathcal{C}(A_2)^{\perp},$$

which is valid since $\mathcal{C}(A_2) \subseteq \mathcal{C}(A_1)$, and construct an orthogonal matrix $\Gamma = (\Gamma_1 : \Gamma_2 : \Gamma_3)$ with the following property: $\mathcal{C}(\Gamma_1) = \mathcal{C}(A_2)$, $\mathcal{C}(\Gamma_2) = \mathcal{C}(A_1) \cap \mathcal{C}(A_2)^{\perp}$ and $\mathcal{C}(\Gamma_3) = \mathcal{C}(A_1)^{\perp}$. Now, as previously we reformulate the model. Let $Z = (Z_1' : Z_2' : Z_3')' = \Gamma'Y$ implying that

$$E[Z] = \begin{pmatrix} \eta_1 & \eta_2 \\ \eta_3 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$$

and in case (i)

$$D[Z] = \sigma^2(I \otimes \begin{pmatrix} K & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + D \otimes \Omega)$$

whereas in case (ii)

$$D[Z] = \sigma^2(I \otimes \begin{pmatrix} K & 0 \\ 0 & 0 \end{pmatrix} + D \otimes \Omega),$$

where D is diagonal the η's and Ω are new parameters and $K = \Gamma_1' EE' \Gamma_1$ in case (i) and in case (ii) $K = (\Gamma_1' : \Gamma_2')EE'(\Gamma_1' : \Gamma_2')'$. Now, for case (i)

$$E[Z_1 : Z_2 | Z_3] = \begin{pmatrix} \eta_1 & \eta_2 \\ \eta_3 & 0 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} + \delta Z_3 = \begin{pmatrix} \eta_1 & \delta \\ \eta_3 & \end{pmatrix} \begin{pmatrix} C_1 \\ Z_3 \end{pmatrix} + \begin{pmatrix} \eta_2 \\ 0 \end{pmatrix} C_2 \tag{3.4}$$

$$D[Z_1 : Z_2 | Z_3] = \sigma^2 (I \otimes \begin{pmatrix} K & 0 \\ 0 & 0 \end{pmatrix} + D \otimes \tau) \tag{3.5}$$

where as before

$$\Omega = \begin{pmatrix} \tau + \delta' \Delta \delta & \delta' \Delta \\ \Delta \delta & \Delta \end{pmatrix}$$

and the conditional distribution is of course belonging to the class of normal distributions. In order to obtain a computational feasible estimating procedure we will present estimators which are based on a marginal likelihood. The approach is identical to an approach Lundby-Christensen put forward when discussing a univariate model with an intercept. Let Q be aby matrix spanning $\mathcal{C}(C_1' : Z_3')^{\perp}$ such that $Q'Q = I$ and $Q'DQ = D_1$ (D_1 is another diagonal matrix). Then set $R = (Z_1' : Z_2')'Q$ and

$$R | Z_3 \sim N_{p(A_1),n}(\begin{pmatrix} \eta_2 C_2 Q \\ 0 \end{pmatrix}, \ \sigma^2 (I \otimes \begin{pmatrix} K & 0 \\ 0 & 0 \end{pmatrix} + D_1 \otimes \tau)). \tag{3.6}$$

The distribution in (3.6) is of the same form as the one in extension 1 and thus we can obtain estimators in the same manner as for extension 1. In order to obtain estimators for η_1, η_3 and δ we maximize the likelihood after having inserted the estimators for η_2, σ^2 and τ. Hence, since σ^2 and τ are estimated it follows from (3.5) that the variance now is completely specified (denote the estimated variance $\hat{V}$) and we can immediately write down the estimators;

$$\operatorname{vec}(\begin{pmatrix} \hat{\eta}_1 & \hat{\delta} \\ \hat{\eta}_3 & \end{pmatrix}) = ((T' \otimes I)\hat{V}^{-1}(T \otimes I))^{-1}(T' \otimes I)\hat{V}^{-1} \operatorname{vec}((Z_1' : Z_2')' - \begin{pmatrix} \hat{\eta}_2 C_2 \\ 0 \end{pmatrix}).$$

For case (ii) we obtain instead of (3.6)

$$R | Z_3 \sim N_{p(A_1),n}(\begin{pmatrix} \eta_2 C_2 Q \\ 0 \end{pmatrix}, \ \sigma^2 (I \otimes K + D_1 \otimes \tau))$$

but unfortunately we can then not rely on extension 1 since there is no nice structure in the covariance matrix (the K matrix).

4. DISCUSSION

In the previous sections we have indicated how to extend a model for handling a univariate growth process. The emphasizes have been to simplify the models so that algorithms can be constructed which do not include too heavy calculations. However, before applying the results we need to derive some statistical properties for the estimators. This task is difficult if we just assume that the number of individuals is increasing. Already in the RCR model there exist great difficulties [see J o h a n s e n 1982]. A very crude approach is to use the inverse information matrix but to show some correct asymptotic or/and approximative results is not easy. If we have many observations within each individual the problem is not too difficult, but unfortunately this is a rare event.

Many of the results in this paper are fairly straightforward to extend. For example, instead of dealing with the mean structure in (3.3) we could have obtained results for

$$E[Y] = \sum_{i=1}^{r} A_i B_i C_i, \qquad \mathcal{C}(A_1) \subseteq \mathcal{C}(A_2) \subseteq , \ldots , \subseteq \mathcal{C}(A_r)$$

and then the estimating procedure in principle consists of maximizing likelihoods and marginal likelihoods. The difficulties with the model depend on how $\mathcal{C}(A_i)$ is related to the covariance structure.

Hitherto we have only modelled situations where we have one growth process. If there exist several growth processes it is possible to discuss

$$Y \sim N_{p,n}(ABC, \sigma^2(H \otimes E\Omega E' + I \otimes \Psi))$$

where B, Ω and Ψ are unknown parameters. It is, however, difficult to derive any nice algorithms for this general situation. On the other hand, if the growth processes are independent Ω is diagonal and then it is possible to copy the arguments presented in sections 2 and 3. A still more general covariance structure is the multivariate variance components structure $\Sigma_{i=1}^{r} H_i \otimes \Psi_i$ where Ψ_i is unknown and H_i is known.

ACKNOWLEDGEMENTS

I wish to thank the University of Łódź for inviting me and the Swedish Natural Science Research Council for supporting the stay at the University of Łódź.

REFERENCES

J o h a n s e n S. (1982): *Asymptotic inference in random coefficient regression models*. Scand. J. Statist. 9, p. 201-207.

L u n d b y e-C h r i s t e n s e n S. (1988): *Modelling and Monitoring Pregnancy*. Thesis, Institute of Electronic Systems, Åalborg University Centre, Åalborg, Denmark.

P o t t h o f f R. F., R o y S. N. (1964): *A generalized multivariate analysis of variance model useful especially for growth curve problems*. "Biometrika", 51, p. 313-326.

V o n R o s e n D. (1989a): *The growth curve model*. A review. Submitted for publication.

V o n R o s e n D. (1989b): *Maximum likelihood estimators in multivariate linear normal models*. J. Multivar. Analysis, 31, p. 187-200.

Dietrich von Rosen

ŁĄCZENIE MODELU REGRESJI Z LOSOWYMI WSPÓŁCZYNNIKAMI Z MODELEM POTTHOFFA I ROYA W ANALIZIE KRZYWEJ WZROSTU

W pracy tej połączono koncepcję modelowania danych generowanych wg krzywej wzrostu z modelami regresji o losowych współczynnikach. Zasadniczo, zostają rozszerzone wyniki pracy dyplomowej [L u n d b y e-C h r i s t i a n s e n a (1988)]. Pokazano, że kiedy występują pewne powiązania pomiędzy strukturą średniej i strukturą kowariancji, możliwe jest uzyskanie prostych algorytmów, prowadzących do estymatorów bez żadnych uciążliwych obliczeń.