UNIVERSITY OF CALIFORNIA
RIVERSIDE

Bayesian Analysis of Errors-in-Variables Growth Curve Models

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Applied Statistics

by

Hung-Jen Huang

August 2010

Dissertation Committee

Dr. Ashis SenGupta, Co-Chairperson

Dr. Keh-Shin Lii, Co-Chairperson

Dr. Laosheng Wu

The Dissertation of Hung-Jen Huang is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgments

I am very grateful to my dissertation committee, especially Dr. Lii and Dr. SenGupta, for their valuable advice and support.

Dedication

To my dearest mother Mei-Chiou Huang and to my wife.

ABSTRACT OF THE DISSERTATION

Bayesian Analysis of Errors-in-Variables Growth Curve Models

by

Hung-Jen Huang

Doctor of Philosophy, Graduate Program in Applied Statistics
University of California, Riverside, August 2010
Dr. Ashis SenGupta, Dr. Keh-Shin Lii, Co-Chairpersons

We propose analyzing our data with a model that exhibits errors-in-variables (EIV) in auxiliary information and which has an autoregressive covariance structure using a Bayesian methodology. The incorporation of these components into a model is often necessary and realistic in the study of many statistical problems. However, such an analysis usually mandates many simplifying and restrictive assumptions, especially when using a traditional probabilistic approach. Much research has been accumulated in this area. We will show that by taking a Bayesian approach, we can effectively deal with the complexity of these types of models. Such an approach cannot be found in the literature.

In fitting statistical models for the analysis of growth data, many curves and/or models have been proposed. We have collected an exhaustive list of the most important and frequently used growth curves, some of which are used in our model analysis. A motivating example is presented to show the applicability of our general approach. In

addition, auxiliary covariates, both qualitative and quantitative, can be added into our model as an extension. These EIV growth curves with auxiliary covariates provide a very general framework for practical application.

We give several illustrative examples demonstrating how a Bayesian approach using MCMC (Metropolis Hastings in Gibbs sampler) techniques and goodness of fit statistics for model selections can be utilized in our analysis. Highest Density Regions (HDR's) are also used to facilitate Bayesian estimation and inference in these examples. Multivariate growth curve models are presented and detailed to study the relationship of the variables in models.

In the final chapter, we present growth curve models with auxiliary variables containing uncertain data distributions (i.e. mixtures of standard components, such as normal distributions) using Dirichlet Process Priors (DPP, which are composed of discrete and continuous components). We show that DPPs are appropriate in determining the number of components and in estimating the parameters simultaneously, and are especially useful and advantageous in the aforementioned multimodal scenario with respect to the goodness of fit of the model.

Table of Contents

Table of Contents

Tables

Figures

Chapter 1 Introduction

1.1 Growth Curves

Growth curves have been extensively studied and used in a wide range of disciplines such as biology, engineering, economics, crop science, fishery research, etc. For example, within food microbiology, scientists use growth curves to describe the behavior of microorganisms under different physical and chemical conditions. Using growth curves, they can predict microbial safety levels or the shelf life of food products, detect the critical parts or weak links of the production and distribution process, as well as optimize food distribution chains (Zwietering et al, 1990).

Growth curves arise from repeated observations on a number of individuals in an orderly fashion, usually over time. Growth curve models are more general than repeated measurements (see section 1.3) in that the former models the functional relationship between the responses (or dependent variables) and the regressors (or explanatory variables, such as time, capital, income, etc.) These functional relationships can be approximated and represented either by linear or nonlinear curves, and the coefficients therein are of interest in growth curve models.

Growth curve models were first introduced by Potthoff and Roy (1964), and later analyzed by Khatri (1966), and Grizzle and Allen (1969) and many others. The principal objectives of a statistical analysis of such a growth curve model is to estimate the growth curve coefficients and to obtain the variance and covariance matrix of the estimated

growth curve coefficients. A helpful illustration of the analysis of a growth curve was made available by C. R. Rao (1959), who examined the growth of one group of rats under a control treatment.

1.2 Bayesian Growth Curves

Not much literature can be found discussing growth curves under Bayesian formulation. C. R. Rao (1987) discussed a Bayesian and empirical Bayesian methods in the prediction of future observations in growth curve models. He described how to derive a Bayesian prediction probability. He also developed some growth models and studied their usefulness in prediction. Some of the growth curves he mentioned include the Jenss, Winsor, Wright and Bock, and Thissen growth curves (see C. R. Rao, 1987). These curves are used to provide an individuals growth trend over a long time period. Kshirsagar and Smith (1995) gave some details about the pioneering work by Geisser (1970, 1980) and Lee and Geisser (1972) discussed details regarding Bayesian growth curves.

Barry (1995) used a Gaussian process growth function for Bayesian model analysis. Robinson and Crowder (2000) presented Bayesian methods for a growth-curve degradation model with repeated measures. The application of this kind of analysis is mainly for the reliability of manufactured products. Zhang et al. (2007) used growth curve models to analyze longitudinal data under Bayesian methodology. Their statistical inference on posteriors is based on point estimation and credible intervals. They also discussed the application of latent (unobserved) variables in growth curves, using

WinBugs for their simulation. Our models are different from these previous works in that we use nonlinear growth curves, along with errors-in-variables, auxiliary variables and autocorrelations, in the models.

In Table 1.1, we list more than a dozen frequently used growth curves with different parameter characteristics, taken from the literature (combined mainly from Berkey, 1987, Rao, 1987 and Zwietering et al, 1990). We also chose some growth curves (e.g., Gompertz curve) that better fit organism growth patterns than logistic, linear, quadratic and other exponential models (see Zwietering et al, 1990) for our models and illustrative examples.

1.3 Growth Curves, Repeated Measurements and Longitudinal Studies

The following key points about repeated measures (Kshirsagar and Smith, 1995) will be useful in defining and differentiating them from growth curves:

1. Repeated measures are statistical methods used for the analysis of a designed experiment where responses are observed at each repetition.
2. The fundamental goal of a repeated measures design is to study the effects of treatments or to detect differential treatment effects (or factors, or a combination of factor levels) at different times rather than estimating and predicting the effects of treatments at any future time. That is, whether the treatments differ as a whole over the entire time period is of interest to the investigators, and attention is

focused on tests of significance between treatments rather than the relationship between the effects at the times when the treatments are applied.

3. The basic premise in growth curve models is that there is a functional relationship (like a polynomial) between a treatment effect and its time of application and that this relationship may be modeled. There are usually physical interpretations of growth curves model coefficients and their estimates. Tests of significance, associated hypothesis testing, as well as prediction are all of interest to the investigator.
4. The repeated measures model is a particular case of a growth curve model. An analogy is the analysis of variance model (stated in forms of class effects and interactions) to the regression model (stated in linear or polynomial functions with unknown coefficients but with known values of the regressor variables).

In addition to the above, repeated measures can model both increase and decrease in the whole period of experiment, while growth curve models mainly deal with increases during the process of growth and occasionally for situations where growth may decrease, such as body weight loss or the shrinkage of tumors.

Kshirsagar and Smith (1995) provide a good example in the study of an experiment on pupil sizes in the eyes of 30 dogs using five treatments (different drugs, including one control group and four other groups). Each treatment group has 6 dogs and their pupil sizes are observed over a period of 60 minutes (6 time points). In a repeated measurements model, the emphasis is to test whether the mean pupil sizes for each drug

are equal in contrast to five different growth curves in functional forms, where the results of the estimates are coefficients.

Finally, regarding longitudinal research studies, according to Lavrakas (2008), their significance generally stems from the fact that the knowledge, skills, attitudes, perceptions, and behaviors of individual subjects usually develop, grow, and change in essential ways over a period of time. Longitudinal studies require formulating longitudinal research questions and hypotheses, using longitudinal data collection methods (e.g., panel surveys) and using longitudinal data analysis methods.

Researchers across different disciplines have used different terms to describe longitudinal studies that involve repeated observations and measurements of the same individual subjects (respondents) over time; some of the terms used are included in repeated-measures designs, growth modeling, etc., so it's not easy to draw a clear distinction between growth curve models and the longitudinal studies as previously defined. However it would be appropriate to say that the growth curve models place a special emphasis on the functional relationships between time and the dependent variable, as well as "growth" (or increase), whereas longitudinal studies place more emphasis on providing information about change.

1.4 Errors-in-variables in Model

In studying experimental problems in biology (Dellaportas, 1995), in engineering (Jitjareonchai, 2006) and in economics (Edgerton and Jochumzen), etc., the use of errors-in-variables (EIV) models has been shown to be necessary and realistic.

When data, which is contaminated with some measurement error, is used to estimate the parameters of any mathematical model by methods such as least squares, it is more the rule than the exception that at least some of the contaminated data (quantities) can be treated as independent variables, contrary to the basic requirement that independent variables be perfectly known. While useful results can often be obtained by neglecting some errors in some of the variables, it is true that to some degree, often unknown, errors of approximation are still present. The use of EIV eliminates this approximation.

A large amount of statistical literature has been devoted to the classical analysis of EIV models. Maximum likelihood methods and least-squares solutions are widely used and discussed in solving different parameter estimation problems involving EIV (see for example, Madansky, 1959, Solari, 1969, Britt and Leucke, 1973, Fuller, 1987, Schafer, 1987, Whittemore and Keler, 1988, Caroll, 1989, Carroll, Gail and Lubin, 1993). In our paper, EIV will become an integral part and a significant element of our growth curve models. EIV is important especially when we deal with measurement-error-contaminated data and want to reduce or eliminate the bias in growth curve model parameter estimations.

1.5 Bayesian Approach for Complex Models

We quote statements in Press (2003) regarding the importance of applying Bayesian approach to scientific inference as follows:

Today, scientists schooled in the Bayesian approach to scientific inference have been challenging the way statistical methodology itself has been developing. Many believe that a paradigm shift has been taking place in the way scientific inference is carried out ... The so-called objective information is summarized in the likelihood function. But the likelihood function is of course almost invariably based upon data that has been influenced by the subjectivity of the observer. Moreover, in small or often in even moderate size samples its structural form is not very well determined. So the likelihood function will almost invariably contain substantial subjective influences and uncertainty.

He uses a convincing example to illustrate his statement beginning with the question:

If 100 observations are collected from an experiment replicated 100 times; there is one observation from each replication. These data are sent to five scientists located in different parts of the world. All five scientists examine the same data set, that is the same 100 data points (all subjectivity involved in deciding what data to collect and in making the observations are eliminated). Should we expect all five of the scientists to draw the same conclusions from these data? And the answer to this very question is a definite "No".

Wolpert (1992) also wrote:

The idea of scientific objectivity has only limited value, for the way in which scientific ideas are generated can be highly subjective, and scientists will defend their views vigorously...It is, however, an illusion to think that scientists are unemotional in their attachment to their scientific views-- they may fail to give them up even in the face of evidence against them... scientific theories involve a continual interplay with other scientists and previously acquired knowledge... and an explanation which the other scientists would accept.

As researchers make every effort to make their models more realistic, more elements have to be incorporated such that the models become increasingly complicated and the exact solutions often become difficult to come by. However, under a Bayesian paradigm, we have the computational advantages in handling complicated models

without having to make many simplifying assumptions (unlike in the classical approach), to obtain reasonable estimates for the parameters of interest.

A Bayesian approach was proposed by Lindley and El-Sayyad (1968) and was shown to yield more sensible results than the maximum likelihood method. They point out that parameter estimation for the usual models, even with normal errors, presents difficulties and the results of the maximum likelihood method in EIV problems may be misleading. (Only in exceptional cases where the ratio of the errors of observations in both dependent and independent variables are known are the parameters estimates reliable).

The Bayesian method for EIV problems was further developed by Reilly and Patino-Leal (1981) and later extended to cases where the error covariance matrix is unknown (Keeler and Reilly, 1991). They use conventional optimization techniques to obtain point estimates of the parameters and to calculate approximate confidence limits. Dellaportas and Stephens (1995) use a Bayesian approach and Gibbs sampler technique to solve examples of nonlinear curve EIV problems. Nummi (2000) proposed a Bayesian growth curve model in which the observed time intervals (the independent variable) are subject to measurement errors. Jitjareonchai et al (2006) further pointed out that traditional techniques consist of complicated matrix manipulation and sometimes lead to convergence problems. To counter this, they propose and implement Gibbs Sampler techniques, managing to solve EIV problems with greater ease and obtaining results with an ideal degree of accuracy.

1.6 Multivariate Analysis

Multivariate analysis is included in our research (Chapter 7) because it enables us to study the effect of several variables acting simultaneously. It provides a closer resemblance to our own intuition, as well as a better understanding about the relationship among the variables. When more variables are analyzed simultaneously, we obtain greater statistical power as well as easier visualization and interpretation of the data through graphical measures (such as scatter plot or higher dimensional plots) to facilitate our study of the relationship among the variables.

1.7 Dirichlet Process Prior

In our study of growth curve models with auxiliary variables, we observe that there are situations when auxiliary variables could come from multimodal distributions. If we could somehow estimate the number of components, as well as the parameters of the mixtures of normals, then our model could be a better fit to that specific type of data. Through a literature survey, we find that the model by Bhattacharya and SenGupta (2009) for circular data is a good reference. We use this as our base framework, extending and applying it to growth curve models on the basis of a similar model to get better results.

1.8 Structure of the Dissertation

Our dissertation is structured as follows: in chapter 2 we will give the motivation for using errors-in-variables in our model and the disadvantages of getting biased estimates if not incorporating them in model when data are subject to this type of errors;

chapter 3-5 are stepping stones to bridge our readers to the main thrust of our research in chapters 6-8, so in chapter 3 we will look at the Bayesian analysis of some commonly used growth curves, the reason for adopting them and the method to use them in our models; in chapter 4 we will address Bayesian analysis of growth curves with AR(1) covariance structure and show that it will produce better estimates for parameters of interest; in chapter 5 we will discuss growth curves under Bayesian formulation with auxiliary information and AR(1) covariance structure because generally speaking, auxiliary variables in model will provide additional information and the estimates will spontaneously be less biased; chapter 6-8 are the major contributions of our work, in chapter 6 we will address Bayesian analysis of growth curve models with errors-in-variables in auxiliary information and AR(1) covariance structure, a complex scenario (a combination of all the previously mentioned situations), which has not been well addressed in literature; in chapter 7 we will discuss Bayesian analysis of multivariate growth curve models to study the relationships among the variables; and finally in chapter 8 we will look at the application of Dirichlet mixture of normals in growth curve models in which their auxiliary variables contain data from mixtures of standard normal distributions. We will show that by implementing our work in this chapter (along with the algorithm), all the parameters of interest can be adequately estimated and a best fit model can also be determined. We have also made a list (not exhaustive) of some of the most important growth curves in Table 1.1, from which our four commonly used curves are chosen for our subsequent analysis.

Table 1.1: Growth Models

Model		Number of parameters	Linear in parameters
Polynomial	*Quadratic* $Y = a + bX + cX^2$	**3**	**Yes**
	Cubic $Y = a + bX + cX^2 + dX^3$	**4**	**Yes**
Exponential	*Jenss* $Y = a + bX - \exp(c + dX)$	**4**	**No**
	Logistics $Y = a + \frac{b}{1 + \exp(c - dX)}$	**4**	**No**
	Gompertz $Y = a \cdot \exp[-\exp(b - cx)]$	**3**	**No**
	Winsor $Y = kexp(-\exp(c - dX))$	**3**	**No**
	Wright $Y = k[1 + \exp(a + bX)]^{-1}$	**3**	**No**
	Richards $y = a\{1 + \nu \cdot \exp[k(\tau - X)]\}^{(-1/\nu)}$	**4**	**No**
	Bock , Thissen $Y = \sum_{i=1}^{3} k_i[1 + \exp(a_i + b_i X)]^{-1}$	**9**	**No**
	Stannard $y = a\left\{1 + \exp[\frac{-(1 + kX)}{p}]\right\}^{(-p)}$	**3**	**No**
Reciprocal	*Count* $Y = a + bX + c(lnX)$	**3**	**Yes**
	Reed first order $Y = a + bX + c(lnX) + d\frac{1}{X}$	**4**	**Yes**
	Reed second order $Y = a + bX + c(lnX) + d_1\frac{1}{X} + d_2(\frac{1}{X})^2$	**5**	**Yes**

Take four commonly used growth curves ($j=1,\ 2,\ 3,\ 4$) for example:

1. Jenss growth curve $f(w_j,\theta)=a+b\,w_j-\exp(c+d\,w_j)$, a vector of parameters $\theta'=(a\ \ b\ \ c\ \ d)$, where $a,b>0$, c and $d\in\mathbb{R}$.
2. Gompertz growth curve $f(w_j,\theta)=a\cdot\exp[-\exp(b-c\,w_j)]$, a vector of parameters be $\theta'=(a\ \ b\ \ c)$, where $a,\ c>0$, $b\in\mathbb{R}$.
3. Richards growth curve $f(w_j,\theta)=a\cdot\left\{1+b\exp[c(d-w_j)]\right\}^{-1/b}$, a vector of parameters be $\theta'=(a\ \ b\ \ c\ \ d)$, where $a,b>0$, $c,d\in\mathbb{R}$ (strictly speaking, b can also be negative as long as $b>-\left\{\exp\left[c(d-w_j)\right]\right\}^{-1}$).
4. Polynomial growth curve $f(w_j,\theta)=a+b\,w_j+c\,w_j^2+d\,w_j^3$, a vector of parameters be $\theta'=(a\ \ b\ \ c\ \ d)$, where where a , b , c and $d\in\mathbb{R}$.

The various forms of growth curves that we have listed here are treated as expectation of the random response variable y_j (i.e., $E(y_j)=f(w_j,\theta)$), for which we have used an underlying normal distribution assumption throughout in our dissertation.

To give our readers an idea (through visualization) about how these four curves would look and behave when their parameters change, we pick the parameters c , b and adjust them independently (they carry different significances in different curves, so they are not used for comparisons according to scale) in the following Figures 1.1 and 1.2.

Figure 1.1 Growth Curves (1)

If c is increased from 0.1 to 0.7 with increment 0.2, the y values for Jenss and Gompertz curve will decrease but those values in Richards and Cubic curve will increase.

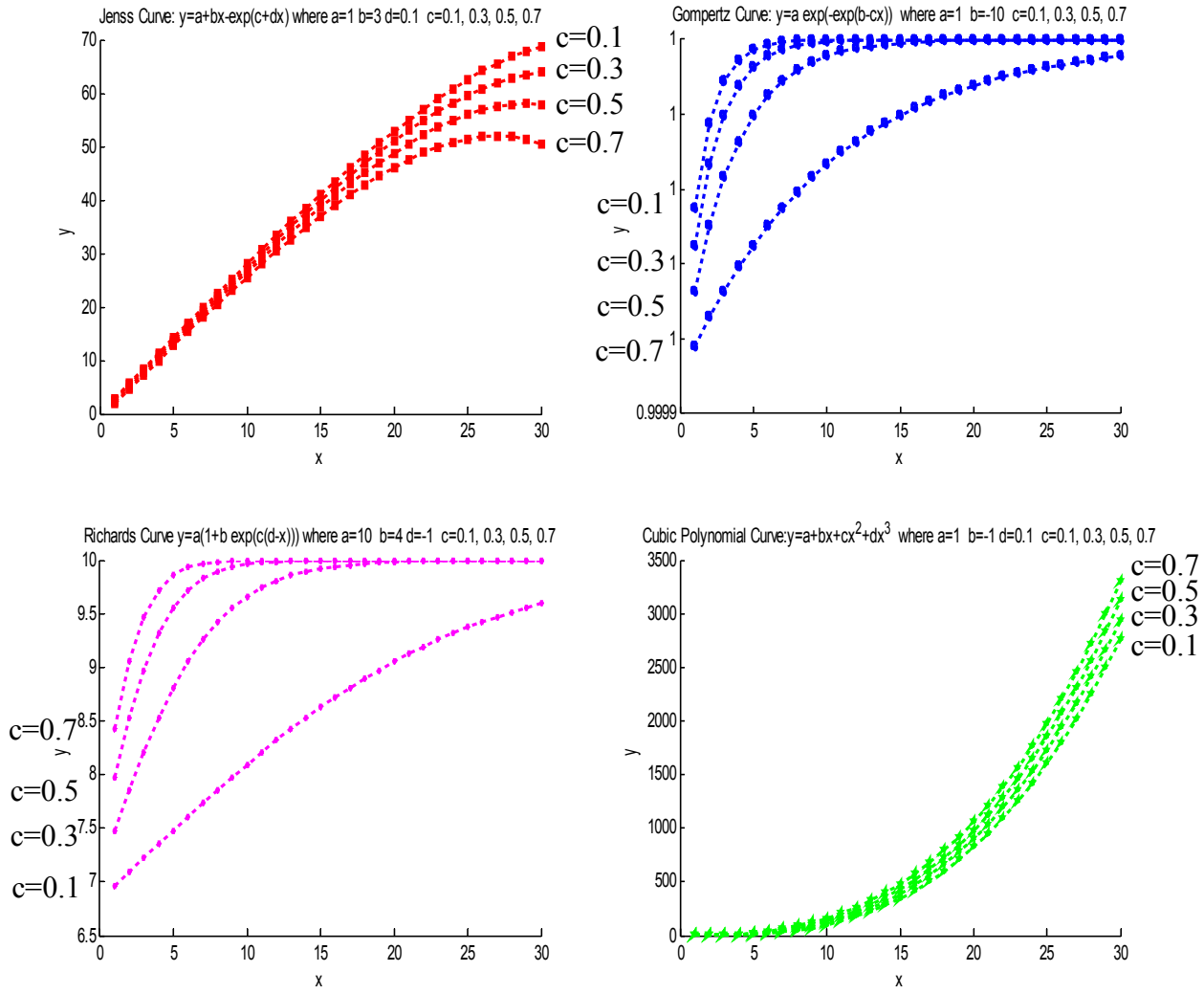
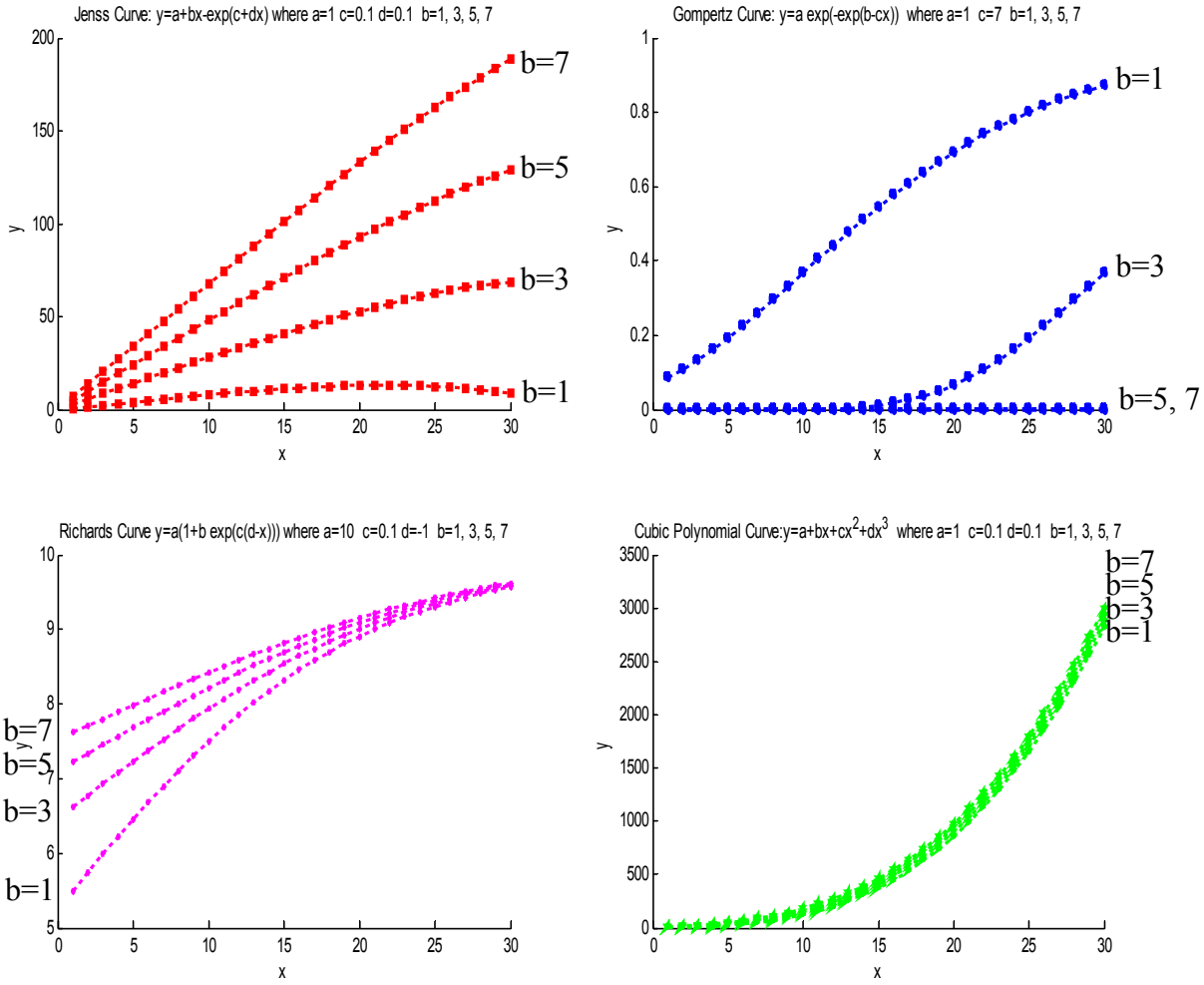


Figure 1.2 Growth Curves (2)

If b increased from 1 to 7 with increment 2, the y values for Gompertz curve will decrease but those values in Gompertz, Richards and Cubic curve will increase.



References

Agresti, Alan (2002). Categorical Data Analysis, 2nd edition. *Wiley*, New York.

Barry, Daniel (1995). A Bayesian Model for Growth Curve Analysis. *Biometrics*, 51, pp. 639-655.

Berkey, C. S. (1982). Bayesian approach for a nonlinear growth model. *Biometrics*, **38**, (1982: Dec.), pp.953-961.

Berkey, C. S. (1982). Comparison of two longitudinal growth models for preschool children. *Biometrics*, **38**, pp. 221-234.

Berkey, C. S. and R. B. Reed (1987). A Model for describing normal and abnormal growth in early childhood. *Human Biology*, **59:**6, p. 973.

Bhattacharya, S. and SenGupta, A. (2009). Bayesian Analysis of Semiparametric Linear-Circular Models. *Journal of Agricultural Biological Environmental Statistics*, **14**, pp. 33-65.

Britt, H. I. and R. H. Leucke (1973). The estimation of parameters in nonlinear, implicit models. *Technometrics* 15, pp. 233-247.

Cameron, A.C. and P.K. Trivedi (1998). Regression Analysis of Count Data. *Cambridge University Press, Cambridge*.

Carroll, R. J., Gail, M. H. and Lubin, J. H. (1993). Case-control studies with errors in covariates. *Journal of the American Statistical Association,* **88**, pp. 421, 185-199.

Carroll, R.J., D. Ruppert and L.A. Stefanski (1995). Measurement Error in Nonlinear Models. *Chapman & Hall*, London.

Diggle, P. J. (1988). An Approach to the Analysis of Repeated Measurements. *Biometrics*, **44**, pp. 959-971.

Dellaportas, P. and D. A. Stephens (1995). Bayesian analysis of errors-in-variables regression models. *Biometrics* **51**, pp. 1085-1095.

Edgerton, D. and P. Jochumzen (2003). Estimation in models with measurement errors and limited dependent variables. *Department of Economics, Lund University.*

Fuller, W. A. (1987). Measurement error Models. *Wiley*, New York.

Geisser, S. (1970). Bayesian analysis of growth curves. *Sankhya A*, **32**, pp. 53-64.

Geisser, S. (1980). Growth curve analysis. *Handbook of Statistics*, **1**, Krishnaiah, P. R. (E.d.), pp. 89-115. North Holland, Amsterdam.

Grizzle, J. E. and Allen, D. M. (1969). Analysis of growth and dose response curves. *Biometrics*, **25**, pp. 357-381.

Hyndman, R.J. (1996). Computing and graphing highest density regions. *American Statistician*, **50**, pp. 120-126.

Jenss, R.M. and N. Barley (1937). A mathematical method for studying the growth of a child. *Human Biology,* **9**, pp. 556-563.

Jitjareonchai, J. J., P. M. Reilly, T. A. Duever and D. B. Chambers (2006). Parameter estimation in the error-in-variables models using the Gibbs Sampler. *The Canadian Journal of Chemical Engineering*, **84**, pp.125-138.

Keeler, S. E. and P. M. Reilly (1991). The error-invariables model applied to parameter estimation when the error covariance matrix is unknown. *The Canadian Journal of Chemical Engineering*, **69**, pp. 27-34.

Khatri, C. A. (1966). A note on a MANOVA model applied to problems in growth curves. *Annals of the Institute of Statistical Mathematics*, **18**, pp. 75-86.

Koziol, J. A., D. A. Maxwell, M. Fukushima, M.E.M. Colmerauer and Y. H. Pilch (1981). A distribution-free test for tumor-growth curve analyses with application to an animal tumor immunotherapy experiment. *Biometrics*, **37**, pp. 383-390.

Kshirsagar, A. M., W. B. Smith (1995). Growth Curves. *Marcel Dekker.*

Kukush, A., H. Schneeweiss and R. Wolf (2004). Three estimators for the Poisson regression model with measurement errors. *Statistical Papers,* 45351–368.

Lee, J. C. and Geisser S. (1972). Growth curve prediction. *Sankhya A*, **34**, pp. 393-412.

Lindley, D. V. and G. M. El-Sayyad (1968). The Bayesian estimation of a linear functional relationship. *Journal of Royal Statistical Society*, B **30**, pp. 190-202.

Ma, Y. and M. G. Genton (2004). Flexible class of skew-symmetric distributions. *Scandinavian Journal of Statistics*, **31**, pp. 459-468.

Madansky, A. (1959). The fitting of straight lines when both variables are subject to error. *Journal of the American Statistical Association,* **54**, pp. 173-205.

Manwani, A. H. and K. N. Agarwal (1973). The growth pattern of Indian infants during the first year of life. *Human Biology*, **45**, p.341.

Nummi, T. (2000). Analysis of growth curves under measurement errors. *Journal of Applied Statistics*, **27**, pp. 235-243.

Oyhenart, E. E. et al. (2003). Sexual dimorphism and postnatal growth of intrauterine growth retarded rats. *Growth, Development and Aging*, **67**, pp. 73-83.

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

Racine-Poon, A. , C. Withhs and A. F. M. Smith (1991). Estimation of relative Potency with sequential dilution errors in radioimmunoassay. *Biometrics*, **47**, pp. 1235-1246.

Rao, C. R. (1959). Some problems involving linear hypothesis in multivariate analysis. *Biometrika*, **46**, pp. 49-58.

Rao, C. R. (1987). Prediction of future observations in growth curve models. *Statistical Science*, **2**, pp. 434-447.

Reilly, P. M. and H. Patino-Leal (1981). A Bayesian study of the error-in variables model. *Technometrics,* **23**, pp. 221-232.

Robinson, M. E. and M. J. Crowder (2000). Bayesian methods for a growth-curve degradation model with repeated measures. *Lifetime Data Analysis*, **6**, pp. 357-374.

Rubin, D. B. (1980). Using empirical Bayes techniques in the law school validity studies. *Journal of the American Statistical Association,* **75**, pp. 801-816.

Schafer, D. W. (1987). Covariate measurement error in generalized linear models. *Biometrica,* **74**, pp. 385-391.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, pp. 461-464.

Solari, M. E. (1969). The maximum likelihood solution of the problem of estimating a linear functional relationship. *Journal of Royal Statistical Society*, B **31**, pp. 372-375.

Spiegelhalter, D. J., N. G. Best, B.P. Carlin, and A. van der Linde (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of Royal Statistical Society B*, **64**, pp. 583-639.

Stuart, H. C. and R. B. Reed (1959). Longitudinal studies of child health and development. Harvard School of Public Health, Series II, No. 1. *Description of project, Pediatrics,* **23**, pp. 875-885.

Tiwari, C.T., K.A. Cronin, W. Davis, E. J. Feuer, B. Yu, and S. Chib (2005). Bayesian model selection for join regression with application to age-adjusted cancer rates. *Applied Statistics*, **54**, part 5, pp. 919-939.

Vonesh, E. F. and V. M. Chinchilli (1996). Linear and Nonlinear Models for the Analysis of Repeated Measurements. *Marcel Dekker.*

Whittemore, A. S. and J. B. Keller (1988). Approximation for regression with covariate measurement error. *Journal of the American Statistical Association,* **83**, pp. 404, 1057-1066.

Zhang, Z. et al. (2007). Bayesian analysis of longitudinal data using growth curve models. *International Journal of Behavioral Development*, **31** (4), pp. 374-383.

Zweitering, M. H., I. Jongenburger, F. M. Rombouts and K. Van't Riet (1990). Modeling of bacterial growth curve. *Applied and Environmental Microbiology*, pp. 1875-1881.

Chapter 2 Motivating Examples

2.0 Abstract

We use two illustrative examples to show the importance of incorporating errors-in-variables in a model when variables are contaminated with errors. The first example we use is a linear model (with fixed and random effects) in which we can see that the coefficient for the random effect is actually accounting for the errors in variable. In the second cubic polynomial growth curve example, we see an obvious difference in terms of bias in the parameter estimations when variables have measurement errors but the model does not consider them.

2.1 The Importance of EIV in A Model

A simple linear model example can show the importance of incorporating errors-in-variables (EIV) into a model. Let y_i, be the degree of sickness $y_i = a + b w_i + c z_i + \varepsilon_i$ where $i = 1, \ldots, T$, w_i is the amount of tablets taken, z_i is the body temperature, b is the coefficient of the fixed effect and c is the coefficient of the random effect.

The second model is one with EIV, it has $y_i = a + b w_i + c z_i + \varepsilon_i$ where z_i is the observed body temperature (taken with errors), $x_i = z_i + e_i$ is the true (unobserved) body temperature with error terms e_i for $i = 1, \ldots, T$, w_i is the number of tablets taken, so we see that w_i is without error ("fixed effect") and z_i is with error ("random effect"). An alternate way to denote a model with EIV is $z_i = x_i + e_i$ (we will use the former notation to be consistent with that in the second example. There's no difference between these two notations under Bayesian formulation).

By comparing these two models, we see that both have fixed and random parts, the former uses coefficients whereas the latter uses variables to account for the two "effects". In other words, people who use linear models deal with the problem of errors and through the coefficient c, but EIV in model approaches the same problem using the variable z, which, in a real sense, is a more explicit, direct and logical way of addressing the EIV problem.

2.2 Polynomial Growth Curve Simulation Example

While the first example is presented with appeal to intuition for understanding the importance of EIV in a model, this second example demonstrates convincing numerical evidence for the same purpose. We assume a cubic polynomial growth curve to display how EIV in model works:

$y_t = f(z_t;\theta) + \varepsilon_t,\ t = 1, \ldots, T$, $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$,

where $f(z_t;\beta) = \beta_0 + \beta_1 z_t + \beta_2 z_t^2 + \beta_3 z_t^3$ and $x_t = z_t + e_t$, $e_t \sim N(0, \sigma_e^2)$

assume that e_i are uncorrelated, and let $\beta = (\beta_0 \ \beta_1 \ \beta_2 \ \beta_3)$, $X = [x_1, x_2, \ldots, x_T]'$,

$Y = [y_1, y_2, \ldots, y_T]'$, $Z = [z_1, z_2, \ldots, z_T]'$.

We quote Moon and Gunst's (1995) example and their simulation results to show the importance of incorporating EIV into the model. They assumed $\beta = (0 \ \ 1 \ \ 0 \ \ -1)$, i.e., $f(z_t;\beta) = 0 + 1 z_t + 0 z_t^2 - 1 z_t^3$ and $\sigma_\varepsilon^2 = \sigma_e^2 = 0.1$, $T = 100$, $Z = (z_1, z_{2,} .., z_{100})$ chosen from among $(-5, -4, \ldots, 5)$ and errors are added to X and Y. Then they ran a simulation of 100 iterations and obtained the results in the following Table 2.1.

From these results we can see that model considering EIV gives closer estimates to true parameters as well as smaller TSE values, .i.e., not including EIV in models and simply use least squares will generate estimates that are more biased due to the introduced EIV in the regressor, and this is especially noticeable in the estimates of the intercept and the linear coefficient.

Table 2.1 Comparison of Parameter Estimations from Two Models

Parameters	Estimates	Standard error	TSE
No EIV (Least Squares)			
β_0	0.04	0.84	0.72
β_1	-0.95	0.85	4.70
β_2	-0.02	0.19	0.04
β_3	-0.85	0.08	0.03
EIV			
β_0	-0.01	0.10	0.01
β_1	1.16	0.19	0.06
β_2	-0.01	0.12	0.01
β_3	-1.02	0.04	0.00

$TSE=(\bar{\beta}_i-\beta_i)^2+\left[\sum(\hat{\beta}_i^{(j)}-\bar{\beta}_i)^2/99\right]$, $\bar{\beta}_i=\sum\hat{\beta}_i^{(j)}/100$, $bias=\bar{\beta}_i-\beta_i$.

2.3 Conclusion

From the previous two sections we can conclude that it is important to include EIV in our growth curve models.

References

Moon, M.S. & R. F. Gunst (1995). Polynomial measurement error modeling. *Computational Statistics and Data Analysis*, **19**, pp. 1-21.

Chapter 3 Bayesian Analysis of Commonly Used Growth Curves

3.0 Abstract

In classical inference problems, the emphasis is often placed on getting optimal or analytical solutions. However, it is usually difficult to obtain such solutions, especially when distributions are complicated. Bayesian methodology provides a simpler way to obtain estimates through computation, especially when these solutions are not attainable.

In our research, four commonly used growth curves under Bayesian formulation are proposed. These models will serve as a foundation for further development of our growth curves with EIV in models under an autocorrelated covariance structure in the following chapters.

The purpose of this chapter (along with chapters four and five) is to show the readers the progressive development of our research which leads to the models and results in chapter six through eight (in which our major contributions and the novelty of our research are detailed).

3.1 Introduction

Bayesian methodology has received extensive attention and has been a very useful tool for estimating models in many different disciplines and applications, especially for complex models. These include growth curves models, which are difficult or almost impossible to estimate in the current MLE and MLE-based software (Lee & Chang, 2000). Since Bayesian inference is not based on the asymptotic nature of the estimators as MLE is (Casella & Berger, 2001), Rindskopf (2006) argued that it would be more plausible to analyze small sample data sets. Zhang et al. (2007) used Bayesian method in analyzing longitudinal data with linear growth curve models. They use both the MLE method and the Bayesian approach to analyze their data set and concluding that there are the following advantages of using the latter:

1. The capacity for implementing Bayesian estimation procedures is acceptable in terms of the ease, flexibility and computation time. (Which generally can not be done in MLE).

2. Progress in scientific research rests on accumulated knowledge. Bayesian methodology provides a natural “chain” for incorporating previous findings with current findings through the use of informative priors and thus informative priors should be used if reliable and available.

3. The results show that Bayesian methods can obtain similar (or exactly the same if no data is missing) parameter estimates to those from MLE when using noninformative priors.

4. Bayesian methods provide a more appealing interpretation. Bayesian methods have unique strengths, such as the intuitive interpretation of the results (e.g., the credible Bayesian interval: "the probability that a parameter lies in the credible interval given the data is at least $100(1-\alpha)\%$ " vs. the confidence interval in frequentist statistics that "If the experiment is repeated many times and the confidence interval is calculated each time, then overall $100(1-\alpha)\%$ of them contain the true parameter") .

5. The complexities of models make the programming and computational demands for Bayesian methods simpler than those of the traditional MLE methods (McArdle and Wang, 2008).

We compare Bayesian and MLE estimators using the following linear growth model as an example,

$y_t = \beta_0 + \beta_1 z_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \quad , \quad t = 1, \ldots, T \quad ,$

where y_i is the dependent variable and z_i is the regressor, $Y = \{y_1, y_2, \ldots, y_T\}$ and $Z = \{z_1, z_2, \ldots, z_T\}$. We assume priors for $\beta_0 \sim N(b_1, \sigma_0^2) \quad , \quad \beta_1 \sim N(b_0, \sigma_1^2)$, then there are three possible cases:

Case 1: Noninformative priors, $\sigma_1^2 \to \infty$ and $\sigma_0^2 \to \infty$,

In this case, Bayesian estimates are equivalent to those of MLE as both methods obtain their estimates by optimizing the same likelihood (posterior) function.

Case 2: Full informative priors, $\sigma_1^2 = 0$ and $\sigma_0^2 = 0$,

If the priors are perfectly correct, then Bayesian estimates are better but if the priors are wrong, then Bayesian estimates may not be better.

Case 3: Partial informative priors, $0 < \sigma_1^2 < \infty \quad , \quad 0 < \sigma_0^2 < \infty$,

Bayesian estimates are better if we are judicious in choosing priors, i.e. we have better reliable prior information than the results of MLE estimators. Otherwise, Bayesian estimates are not better if the prior information is unreliable or misleading.

3.2 Model Formulation

Assume a growth model for a single experimental unit of the form $y_j = f(w_j, \theta) + \varepsilon_j$, $j = 1, \ldots, n$, n is the total number of observations; θ is a vector of parameters; w_j is the independent variable for the j^{th} observation and $\varepsilon_j \sim N(0,\ \sigma^2)$ is independent random error of y_j ; here $f(w_j, \theta)$ is one of the different types of growth curves. From this chapter and on (through chapter 7), we use these four most commonly used growth curves in our models as illustrative examples:

1. Jenss growth curve $f_1(w_j, \theta_1) = a_1 + b_1 w_j - \exp(c_1 + d_1 w_j)$, and a vector of parameters $\theta_1' = (a_1 \ \ b_1 \ \ c_1 \ \ d_1)$, where $a_1, b_1 > 0$, c_1 and $d_1 \in \mathbb{R}$.
2. Gompertz growth curve $f_2(w_j, \theta_2) = a_2 \cdot \exp[-\exp(b_2 - c_2 w_j)]$ and a vector of parameters $\theta_2' = (a_1 \ \ b_2 \ \ c_2)$, where $a_2,\ c_2 > 0$, $b_2 \in \mathbb{R}$.
3. Richards growth curve $f_3(w_j, \theta_3) = a_3 \cdot \left\{1 + b_3 \exp[c_3(d_3 - w_j)]\right\}^{-1/b_3}$, and a vector of parameters $\theta_3' = (a_3 \ \ b_3 \ \ c_3 \ \ d_3)$, where $a_3, b_3 > 0$, $c_3, d_3 \in \mathbb{R}$.
4. Polynomial growth curve $f_4(w_j, \theta_4) = a_4 + b_4 w_j + c_4 w_j^2 + d_4 w_j^3$, and a vector of parameters $\theta_4' = (a_4 \ \ b_4 \ \ c_4 \ \ d_4)$, where a_4 , b_4 , c_4 and $d_4 \in \mathbb{R}$.

Now let these $f_l(w_j, \theta_l)$, $l = 1, 2, 3, 4$ be substituted into the growth model and denote $\boldsymbol{y} = \{y_j,\ j = 1, \ldots, n\}$ and $\boldsymbol{w} = \{w_j,\ j = 1, \ldots, n\}$ then we get their likelihood functions, respectively, as follows:

$$g(\boldsymbol{y}|\theta_l,\sigma^{2,}\boldsymbol{w})=\frac{1}{(2\pi)^{n/2}\sigma^n}\exp\left[-\frac{1}{2\sigma^2}\sum_{j=1}^{n}\left\{y_j-f_l(w_j,\theta_l)\right\}^2\right], \tag{3.1}$$

where $f_l(w_j,\theta_l)=E_{\theta_l}(y_j)$, $l=1,2,3,4$.

Now we propose the following distributions for the four growth curve priors (where $Expon(\cdot)$ denotes an exponential distribution):

1. Jenss curve:

   $a_1\sim Expon(\frac{1}{\tilde{a}_1})$, $b_1\sim Expon(\frac{1}{\tilde{b}_1})$ and $\begin{pmatrix}c_1\\d_1\end{pmatrix}\sim N_2\left(\begin{pmatrix}\tilde{c}_1\\\tilde{d}_1\end{pmatrix},\begin{pmatrix}\tilde{\sigma}_c^2 & \tilde{\rho}\,\tilde{\sigma}_d\tilde{\sigma}_c\\\tilde{\rho}\,\tilde{\sigma}_c\tilde{\sigma}_d & \tilde{\sigma}_d^2\end{pmatrix}\right)$,

   then the empirical Bayes estimate (denoted by $\tilde{}$) vector $\tilde{\theta}_1=(\tilde{a}_1,\tilde{b}_1,\tilde{c}_1,\tilde{d}_1,\tilde{\sigma}_c^2,\tilde{\sigma}_d^2,\tilde{\rho})$ (note that the empirical estimates for the means in $\tilde{\theta}_1$ for Jenss and $\tilde{\theta}_i, i=2,3,4$ for other three curves alike can be obtained from nonlinear least square regression results by using MATLAB *nlinfit* function to fit nonlinear Jenss, Gompertz and Richards curves and *polyfit* function to fit polynomial curves. For details, please refer to MATLAB Help and Seber and Wild, 1989. The empirical estimates for the variances and correlation coefficient is assumed to be some reasonably small values). So the prior distribution is:

$$h(\theta_1|\tilde{\theta}_1)\propto\frac{1}{\tilde{a}_1\tilde{b}_1}\frac{1}{\tilde{\sigma}_c\tilde{\sigma}_d\sqrt{(1-\tilde{\rho}^2)}}\exp\left\{-\frac{1}{2(1-\tilde{\rho}^2)}\cdot\left[\frac{(c_1-\tilde{c}_1)^2}{\tilde{\sigma}_c^2}-\frac{2\tilde{\rho}(c_1-\tilde{c}_1)(d_1-\tilde{d}_1)}{\tilde{\sigma}_c\tilde{\sigma}_d}+\frac{(d_1-\tilde{d}_1)^2}{\tilde{\sigma}_d^2}\right]-\left[\frac{a_1}{\tilde{a}_1}+\frac{b_1}{\tilde{b}_1}\right]\right\}. \tag{3.2}$$

2. Gompertz growth curve:

$a_2 \sim Expon(\frac{1}{\tilde{a}_2})$, $c_2 \sim Expon(\frac{1}{\tilde{c}_2})$ and $b_2 \sim N(\tilde{b}_2, \tilde{\sigma}_b^2)$ then

$\tilde{\theta}_2 = (\tilde{a}_2, \tilde{b}_2, \tilde{c}_2, \tilde{\sigma}_b^2)$. So the prior distribution is:

$$h(\theta_2|\tilde{\theta}_2) \propto \frac{1}{\tilde{a}_2 \tilde{c}_2} \frac{1}{\tilde{\sigma}_b} \exp\left\{ -\frac{1}{2} \frac{(b_2 - \tilde{b}_2)^2}{\tilde{\sigma}_b^2} - \left[\frac{a_2}{\tilde{a}_2} + \frac{c_2}{\tilde{c}_2} \right] \right\} . \quad (3.3)$$

3. Richards curve:

$a_3 \sim Expon(\frac{1}{\tilde{a}_3})$, $b_3 \sim Expon(\frac{1}{\tilde{b}_3})$, $c_3 \sim N(\tilde{c}_3, \tilde{\sigma}_c^2)$ and $d_3 \sim N(\tilde{d}_3, \tilde{\sigma}_d^2)$,

where a_3 , b_3 , c_3 and d_3 are independent, the vector

$\tilde{\theta}_3 = (\tilde{a}_3, \tilde{b}_3, \tilde{c}_3, \tilde{d}_3, \tilde{\sigma}_c^2, \tilde{\sigma}_d^2)$. So the prior distribution is:

$$h(\theta_3|\tilde{\theta}_3) \propto \frac{1}{\tilde{a}_3 \tilde{b}_3} \frac{1}{\tilde{\sigma}_c \tilde{\sigma}_d} \exp\left\{ -\frac{1}{2} \left[\frac{(c_3 - \tilde{c}_3)^2}{\tilde{\sigma}_c^2} + \frac{(d_3 - \tilde{d}_3)^2}{\tilde{\sigma}_d^2} \right] - \left[\frac{a_3}{\tilde{a}_3} + \frac{b_3}{\tilde{b}_3} \right] \right\} . \quad (3.4)$$

4. Polynomial curve:

$a_4 \sim N(\tilde{a}_4, \tilde{\sigma}_a^2)$, $b_4 \sim N(\tilde{b}_4, \tilde{\sigma}_b^2)$, $c_4 \sim N(\tilde{c}_4, \tilde{\sigma}_c^2)$ and $d_4 \sim N(\tilde{d}_4, \tilde{\sigma}_d^2)$,

where a_4 , b_4 , c_4 and d_4 are independent, the vector

$\tilde{\theta}_4 = (\tilde{a}_4, \tilde{b}_4, \tilde{c}_4, \tilde{d}_4, \tilde{\sigma}_a^2, \tilde{\sigma}_b^2, \tilde{\sigma}_c^2, \tilde{\sigma}_d^2)$. So the prior distribution is:

$$h(\theta_4|\tilde{\theta}_4) \propto \frac{1}{\tilde{\sigma}_a \tilde{\sigma}_b} \frac{1}{\tilde{\sigma}_c \tilde{\sigma}_d} \exp\left\{ -\frac{1}{2} \left[\frac{(a_3 - \tilde{a}_3)^2}{\tilde{\sigma}_a^2} + \frac{(b_3 - \tilde{b}_3)^2}{\tilde{\sigma}_b^2} + \frac{(c_3 - \tilde{c}_3)^2}{\tilde{\sigma}_c^2} + \frac{(d_3 - \tilde{d}_3)^2}{\tilde{\sigma}_d^2} \right] \right\} . \quad (3.5)$$

For σ^2 we assume Jeffrey's prior and thus have density $p(\sigma^2) \propto 1/\sigma^2$. So the posterior distributions for a single experimental unit i are as follows:

1. Jenss growth curve:

$$\pi_1(\theta_1, \sigma^2 \mid \boldsymbol{y}, \boldsymbol{w}) \propto \frac{1}{\sigma^{n+2}} \exp\left\{ -\frac{1}{2}\left(\frac{\sum_{j=1}^{n} \left\{ y_j - f_1(w_j, \theta_1) \right\}^2}{\sigma^2} \right) - \frac{1}{2(1-\tilde{\rho}^2)} \right.$$
$$\left. \left[\frac{(c_1 - \tilde{c}_1)^2}{\tilde{\sigma}_c^2} - \frac{2\tilde{\rho}(c_1 - \tilde{c}_1)(d_1 - \tilde{d}_1)}{\tilde{\sigma}_c \tilde{\sigma}_d} + \frac{(d_1 - \tilde{d}_1)^2}{\tilde{\sigma}_d^2} \right] - \left[\frac{a_1}{\tilde{a}_1} + \frac{b_1}{\tilde{b}_1} \right] \right\}.$$

Integrating out σ^2 we have

$$\pi_1(\theta_1 \mid \boldsymbol{y}, \boldsymbol{w}) \propto \exp\left\{ -\left(\frac{a_1}{\tilde{a}_1} + \frac{b_1}{\tilde{b}_1} + \frac{1}{2(1-\tilde{\rho}^2)} \left[\frac{(c_1 - \tilde{c}_1)^2}{\tilde{\sigma}_c^2} - \right. \right. \right.$$
$$\left. \left. \left. \frac{2\tilde{\rho}(c_1 - \tilde{c}_1)(d_1 - \tilde{d}_1)}{\tilde{\sigma}_c \tilde{\sigma}_d} + \frac{(d_1 - \tilde{d}_1)^2}{\tilde{\sigma}_d^2} \right] \right) \right\} \left[\sum_{j=1}^{n} \left\{ y_j - f_1(w_j, \theta_1) \right\}^2 \right]^{-\frac{n}{2}}. \quad (3.6)$$

2. Gompertz growth curve:

$$\pi_2(\theta_2, \sigma^2 \mid \boldsymbol{y}, \boldsymbol{w}) \propto \frac{1}{\sigma^{n+2}} \exp\left\{ -\frac{\sum_{j=1}^{n} \left\{ y_j - f_2(w_j, \theta_2) \right\}^2}{2\sigma^2} - \left[\frac{1}{2} \frac{(b_2 - \tilde{b}_2)^2}{\tilde{\sigma}_b^2} + \frac{a_2}{\tilde{a}_2} + \frac{c_2}{\tilde{c}_2} \right] \right\}.$$

Integrating out σ^2 we have

$$\pi_2(\theta_2 \mid \boldsymbol{y}, \boldsymbol{w}) \propto \exp\left\{ -\left[\frac{1}{2} \frac{(b_2 - \tilde{b}_2)^2}{\tilde{\sigma}_b^2} + \frac{a_2}{\tilde{a}_2} + \frac{c_2}{\tilde{c}_2} \right] \right\} \left[\sum_{j=1}^{n} \left\{ y_j - f_2(w_j, \theta_2) \right\}^2 \right]^{-\frac{n}{2}}. \quad (3.7)$$

3. Richards growth curve:

$$\pi_3(\theta_3, \sigma^2 \mid \boldsymbol{y}, \boldsymbol{w}) \propto \frac{1}{\sigma^{n+2}} \exp\left\{ -\frac{\sum_{j=1}^{n} \left\{ y_j - f_3(w_j, \theta_3) \right\}^2}{2\sigma^2} - \left[\frac{(c_3 - \tilde{c}_3)^2}{2\tilde{\sigma}_c^2} + \frac{(d_3 - \tilde{d}_3)^2}{2\tilde{\sigma}_d^2} + \frac{a_3}{\tilde{a}_3} + \frac{b_3}{\tilde{b}_3} \right] \right\}.$$

Integrating out σ^2 we have

$$\pi_3(\theta_3 \mid \boldsymbol{y}, \boldsymbol{w}) \propto \exp\left\{ -\left[\frac{(c_3 - \tilde{c}_3)^2}{2\tilde{\sigma}_c^2} + \frac{(d_3 - \tilde{d}_3)^2}{2\tilde{\sigma}_d^2} + \frac{a_3}{\tilde{a}_3} + \frac{b_3}{\tilde{b}_3} \right] \right\} \left[\sum_{j=1}^{n} \left\{ y_j - f_3(w_j, \theta_3) \right\}^2 \right]^{-\frac{n}{2}}. \tag{3.8}$$

4. Polynomial growth curve:

$$\pi_4(\theta_4, \sigma^2 \mid \boldsymbol{y}, \boldsymbol{w}) \propto \frac{1}{\sigma^{n+2}} \exp\left\{ -\frac{\sum_{j=1}^{n} \left\{ y_j - f_4(w_j, \theta_4) \right\}^2}{2\sigma^2} - \left[\frac{(a_4 - \tilde{a}_4)^2}{2\tilde{\sigma}_a^2} + \frac{(b_4 - \tilde{b}_4)^2}{2\tilde{\sigma}_b^2} + \frac{(c_4 - \tilde{c}_4)^2}{2\tilde{\sigma}_c^2} + \frac{(d_4 - \tilde{d}_4)^2}{2\tilde{\sigma}_d^2} \right] \right\}.$$

Integrating out σ^2 we have

$$\pi_4(\theta_4 \mid \boldsymbol{y}, \boldsymbol{w}) \propto \exp\left\{ -\left[\frac{(a_4 - \tilde{a}_4)^2}{2\tilde{\sigma}_a^2} + \frac{(b_4 - \tilde{b}_4)^2}{2\tilde{\sigma}_b^2} + \frac{(c_4 - \tilde{c}_4)^2}{2\tilde{\sigma}_c^2} + \frac{(d_4 - \tilde{d}_4)^2}{2\tilde{\sigma}_d^2} \right] \right\} \cdot \left[\sum_{j=1}^{n} \left\{ y_j - f_4(w_j, \theta_4) \right\}^2 \right]^{-\frac{n}{2}}. \tag{3.9}$$

Now we continue to describe the sampling procedure for the Metropolis Hastings Algorithm within the Gibbs sampler and in our growth model. The basic idea behind

these algorithms is that the full conditional distributions for each parameter are obtained by substituting the current values of the other parameters. I.e., all the full conditional distributions can be obtained from algebraic and matrix manipulations of the posterior distribution in equation (3.6)-(3.9).

Suppose we take the Gompertz growth curve as an example. First, we need to get the full conditionals for the parameters from (3.7) as follows:

$$\pi_2(a_2 \mid \cdot) \propto \exp\left\{-\frac{a_2}{\tilde{a}_2}\right\}\left[\sum_{j=1}^{n}\left\{y_j - f_2(w_j,\theta_2)\right\}^2\right]^{-\frac{n}{2}}, \tag{3.10}$$

$$\pi_2(b_2 \mid \cdot) \propto \exp\left\{-\frac{1}{2}\frac{(b_2-\tilde{b}_2)^2}{\tilde{\sigma}_b^2}\right\}\left[\sum_{j=1}^{n}\left\{y_j - f_2(w_j,\theta_2)\right\}^2\right]^{-\frac{n}{2}}, \tag{3.11}$$

$$\pi_2(c_2 \mid \cdot) \propto \exp\left\{-\frac{c_2}{\tilde{c}_2}\right\}\left[\sum_{j=1}^{n}\left\{y_j - f_2(w_j,\theta_2)\right\}^2\right]^{-\frac{n}{2}}. \tag{3.12}$$

Now to define the MH algorithm, let $\varphi(a_2^{(old)}, a_2^{(new)})$ denote a source density for a candidate and draw $a_2^{(new)}$ given the current value $a_2^{(old)}$ in the sampled sequence. The density $\varphi(a_2^{(old)}, a_2^{(new)})$ is referred to as the proposal or candidate generating density. Then, the MH algorithm is defined by two steps: a first step in which a proposal value is drawn from the candidate generating density and a second step in which the proposal value is accepted as the next iterate in the Markov chain according to the probability:

$$\alpha\left(a_2^{(old)}, a_2^{(new)}\right)=min\left\{\frac{\pi_2(a_2^{(new)})\varphi\left(a_2^{(new)}, a_2^{(old)}\right)}{\pi_2(a_2^{(old)})\varphi\left(a_2^{(old)}, a_2^{(new)}\right)}, 1\right\} \quad , \quad \text{if} \quad \pi_2(a_2^{(old)})\varphi\left(a_2^{(old)}, a_2^{(new)}\right)>0$$

(otherwise $\alpha\left(a_2^{(old)}, a_2^{(new)}\right)=1$).

If the proposal value is rejected, then the next sampled value is taken to be the current value. Let's follow this recursive procedure for the MH algorithm:

1. Specify an initial value $a_2^{(0)}$.
2. Repeat for $j=1,2,\ldots,M$:

   a) Propose $a_2^{(new)}\sim\varphi(a_2^{(j)},\cdot)$.

   b) Let $a_2^{(j+1)}=a_2^{(new)}$ if $U(0,1)\leqslant\alpha(a_2^{(j)}, a_2^{(new)})$ otherwise $a_2^{(j+1)}=a_2^{(j)}$.
3. Return the values $a_2^{(1)}, a_2^{(2)}, \ldots, a_2^{(M)}$.

MH in Gibbs Sampling Algorithm:

We can take samples of θ_2 by using (3.10)-(3.12) through the following steps:

1. Set $j=0$ and select a set of starting parameter values $\theta_2^{(0)}$ and let $\theta_2^{(0)}=(a_2^{(0)}, b_2^{(0)}, c_2^{(0)})$.
2. Sample $a_2^{(j+1)}$ from $\pi_2(a_2^{(j+1)}|\cdot)$ (using (3.10) and MH algorithm).
3. Sample $b_2^{(j+1)}$ from $\pi_2(b_2^{(j+1)}|\cdot)$ (using (3.11) and MH algorithm).
4. Sample $c_2^{(j+1)}$ from $\pi_2(c_2^{(j+1)}|\cdot)$ (using (3.12) and MH algorithm).
5. Replace $\theta_2^{(j)}$ with $\theta_2^{(j+1)}$.
6. Set $j=j+1$ and repeat steps 2 through 5.

7. Stop when $j = N + B$, (B is burn-in samples to be dropped and N is sample size).

In the long run, the N samples taken from these full conditionals will be a sample drawn approximately from the above posterior distribution. In addition to the above analysis and estimation of parameters, we also generate highest density regions for the estimated parameters. Then, 90% confidence limits for the best-fit growth curve can be fitted under this Bayesian formulation by using the 5% and 95% percentiles of $\boldsymbol{y}$ calculated by substituting the N samples of θ_2 at a given w_j .

3.3 Summary

We have shown the reason and the way in using growth curves under Bayesian analysis in this chapter. These models serve as our foundations and their further developments will be detailed in the following chapters.

References

Casella, G. Berger and R. L. Berger (2001). *Statistical Inference (2nd ed.)* Pacific Grove, CA: Duxbury Press.

Lee, J. C. and C.H. Chang (2000). Bayesian analysis of a growth curve model with a general autoregressive covariance structure. *Scandinavian Journal of Statistics*, **27**, pp. 703-713.

McArdle, J.J. and L. Wang (2008). Modeling age-based turning points in longitudinal life-span growth curves of cognition. In P. Cohen (Ed.), *Turning Points Research*. Mahwah, NJ: Erlbaum.

Ratkowsky, D.A. (1983). *Nonlinear regression modeling: a unified practical approach,* Marcel Dekker Inc., NY, USA.

Rindskopf, D. (2006). Some neglected relationships between cognitive psychology and statistics: How people try to think using advanced statistical methods without realizing it. Paper presented at the 71st *Annual Meeting of the Psychometric Society,* Montreal, Canada.

Seber, G. A. F. and C. J. Wild (1989). *Nonlinear Regression*. John Wiley & Sons.

Chapter 4 Bayesian Analysis of Commonly Used Growth Curves with AR(1) Covariance Structure

4.0 Abstract

We demonstrate the need of including an autocorrelative relationship into our growth curve models. A simulation study shows that models lacking this kind of relationship will be subject to a larger bias than those with autocorrelation in covariance structure. We apply our models to a mice cancer growth data set, comparing goodness of fit statistics for the four commonly used growth curves.

The purpose of this chapter is to show our readers how Bayesian analysis works in scenarios of some commonly used growth curves with AR(1) covariance structure so they may be bridged to chapter six through eight, in which our major contributions are detailed.

4.1 Introduction

The existence of an autocorrelation relationship and the need for incorporating an autocorrelation structure, such as AR(1), into growth models can become self-evident by highlighting the word "growth". Take the growth curve relationship between weight and height as an example, and we assume $y_t = f(t) + c\,x_t + \varepsilon_t$, $t = 1, \ldots, T$, where t is time, $f(t) = a + bt$, y_t is weight, x_t is height. Then, although $f(t)$ may draw out most of the relationship between the weight y_t and time t , the relationship between height x_t and weight y_t is still autocorrelated.

To demonstrate the need of considering autocorrelation, we assume the following model:

$y_i = f(w_i; \theta) + \varepsilon_i$, $i = 1, \ldots, m$, $\varepsilon_i = \rho\,\varepsilon_{i-1} + \eta_i$, $\eta_i \sim N(0, \sigma_\eta^2)$, where

$f(z_i;\theta)=\theta_0+\theta_1 z_i+\theta_2 z_i^2+\theta_3 z_i^3$, ε_i is random error with autocorrelation coefficient ρ and m is the number of observations; also let $\theta=(\theta_0\ \ \theta_1\ \ \theta_2\ \ \theta_3)$, $Y=[y_1,\ \ y_2,\ldots,\ \ y_m]'$, $W=[w_1,w_2,\ldots,w_m]'$.

We run a simulation of 1000 iterations using $\theta=(1\ \ 1\ \ 1\ \ 1)$, $\sigma_\eta^2=1$ and $W=[0.1,\ \ 0.2,\ldots,1]$ to compare whether there are significant differences between the least square and Bayesian estimates of the coefficients.

From the results shown in Table 4.1, we see that for a smaller variance in the random error (σ_η^2 = 1, signal to noise ratio is between (2, 5)), the errors in estimates of the coefficients can be as high as almost 40% for $\hat{\theta}_2$ when the autocorrelation coefficient is increased to $\rho=0.7$. If variance is increased to σ_η^2 = 3, then the bias in estimates could go as high as 100% for both $\hat{\theta}_1$ and $\hat{\theta}_2$. Therefore this shows us that, it's necessary and important to take the effect of autocorrelation into consideration for our models.

Table 4.2 shows the results of the estimates using the least square method (without considering ρ) and Bayesian estimates (ρ included in model) in different scenarios (σ_η^2 and ρ combinations). We see that in both scenarios (including both $\rho>0$ and $\rho=0$ cases), Bayesian estimates appear to be more consistent in terms of 1) smaller standard deviations, 2) less biased estimates for the coefficients and 3) reasonable estimates for autocorrelation coefficients embedded in simulated data.

Table 4.1 Least Square Estimates in Different Autocorrelation Scenarios

σ_η^2	ρ	$\hat{\theta}_0$	std	$\hat{\theta}_1$	std	$\hat{\theta}_2$	std
1.0	0.0	1.022	1.210	0.926	4.970	1.054	4.342
	0.4	1.044	1.449	0.836	6.150	1.131	5.421
	0.7	0.991	1.547	1.368	7.006	0.606	6.257
3.0	0.7	1.166	3.563	-0.079	16.367	2.115	14.590

*Signal to noise ratio for $\sigma_\eta^2 = 1.0$ is within (min=1 , max=5), and is (min=0.7, max=2) for $\sigma_\eta^2 = 3.0$.

Table 4.2 Least Squares (LS) vs. Bayesian (BAY) Estimates in Different Autocorrelation Scenarios

Method	σ_η^2	ρ	$\hat{\theta}_0$	std	$\hat{\theta}_1$	std	$\hat{\theta}_2$	std	$\hat{\rho}$	std
BAY	100.0	0.0	0.987	1.070	0.976	0.154	0.960	0.090	0.001	0.005
LS	100.0	0.0	0.961	1.172	1.262	4.908	0.768	1.172	-	-
BAY	10.0	0.4	1.009	0.031	0.994	0.031	0.994	0.052	0.335	0.005
LS	10.0	0.4	1.042	1.449	0.672	6.114	1.343	1.449	-	-

*Signal to noise ratio for $\sigma_\eta^2 = 10.0$ is within (min=1 , max=3), and is (min=0.1, max=0.3) for $\sigma_\eta^2 = 100.0$. Using least square estimates as priors in Bayesian analysis.

4.2 Growth Curve Models with AR(1) Covariance Structure

We adopt the same notations as in previous chapters and assume a growth model for a single experimental unit of the form $y_j = f(w_j, \theta) + \varepsilon_j$, $j = 1, \ldots, n$, n is the total number of observations; θ is a vector of parameters; w_j is the independent variable for the j^{th} observation and $\varepsilon_j \sim N(0,\ \sigma^2)$ is independent random error of

y_j . Here $f(w_j,\theta)$ can be one of the different types of growth curves. We assume that the covariance for the j^{th} and k^{th} observation is $\Sigma=\sigma^2\Omega$, where

$\Omega = \begin{bmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ \rho & 1 & \cdots & \rho^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & 1 \end{bmatrix}_{n\times n}$ and ρ is the autocorrelation coefficient.

Let $f_l(w_j,\ \theta_l), l=1,2,3,4$ be the four different types of growth curves. In this chapter, we use four most commonly used growth curves in our model:

1. Jenss growth curve $f_1(w_j,\theta_1)=a_1+b_1w_j-\exp(c_1+d_1w_j)$, a vector of parameters $\theta_1'=(a_1\ \ b_1\ \ c_1\ \ d_1)$, where $a_1,b_1>0$, c_1 and $d_1\in\mathbb{R}$.
2. Gompertz growth curve $f_2(w_j,\theta_2)=a_2\cdot\exp[-\exp(b_2-c_2w_j)]$ and a vector of parameters $\theta_2'=(a_1\ \ b_2\ \ c_2)$, where $a_2,\ c_2>0$, $b_2\in\mathbb{R}$.
3. Richards growth curve $f_3(w_j,\theta_3)=a_3\cdot\{1+b_3\exp[c_3(d_3-w_j)]\}^{-1/b_3}$, and a vector of parameters $\theta_3'=(a_3\ \ b_3\ \ c_3\ \ d_3)$, where $a_3,b_3>0$, $c_3,d_3\in\mathbb{R}$.
4. Polynomial growth curve $f_4(w_j,\theta_4)=a_4+b_4w_j+c_4w_j^2+d_4w_j^3$, and a vector of parameters $\theta_4'=(a_4\ \ b_4\ \ c_4\ \ d_4)$, where a_4 , b_4 , c_4 and $d_4\in\mathbb{R}$.

Similar to previous chapter, we will let these $f_l(w_j,\theta_l)$, $l=1,2,3,4$ be substituted into the growth model and denote $\boldsymbol{y}=\{y_j,\ j=1,\ldots,n\}$, $\boldsymbol{w}=\{w_j,\ j=1,\ldots,n\}$ and $A_l=[\boldsymbol{y}-f_l(\boldsymbol{w},\theta_l)]'_{n\times1}\Omega^{-1}_{n\times n}[\boldsymbol{y}-f_l(\boldsymbol{w},\theta_l)]_{n\times1}$. Then we get their likelihood functions, respectively, as follows:

$$g(\boldsymbol{y}|\theta_l,\sigma^2,\rho,\boldsymbol{w})=\frac{1}{(2\pi)^{n/2}\sigma^n|\Omega|}\exp\left\{-\frac{1}{2\sigma^2}A_l\right\}, \tag{4.1}$$

where $f_l(w_j,\theta_l)=E_{\theta_i}(y_j)$ for $l=1,2,3,4$.

Now we propose the following distributions for the four growth curve priors ($Expon(\cdot)$ denotes an exponential distribution):

1. Jenss curve:

   $a_1\sim Expon(\frac{1}{\tilde{a}_1})$, $b_1\sim Expon(\frac{1}{\tilde{b}_1})$ and $\begin{pmatrix}c_1\\d_1\end{pmatrix}\sim N_2\left(\begin{pmatrix}\tilde{c}_1\\\tilde{d}_1\end{pmatrix},\tilde{\Sigma}_1\right)$, where c_1 and d_1 follow a bivariate normal distribution with Bayes estimate (denoted by ~), mean vector $\begin{pmatrix}\tilde{c}_1\\\tilde{d}_1\end{pmatrix}$, and variance covariance matrix is $\tilde{\Sigma}_1$. Then the empirical vector $\tilde{\theta}_1=(\tilde{a}_1,\tilde{b}_1,\tilde{c}_1,\tilde{d}_1,\tilde{\Sigma}_1)$ (note that the empirical estimates for the means in $\tilde{\theta}_1$ for Jenss and $\tilde{\theta}_i, i=2,3,4$ for other three curves alike can be obtained from nonlinear least square regression results by using MATLAB *nlinfit* function to fit nonlinear Jenss, Gompertz and Richards curves and *polyfit* function to fit polynomial curves. For details, please refer to MATLAB Help and Seber and Wild, 1989. The estimates for variances and covariances in $\tilde{\Sigma}_i$ are assumed to be some reasonably small values). So the prior distribution is:

$$h(\theta_1|\tilde{\theta}_1)\propto\frac{1}{\tilde{b}_1}\frac{1}{|\tilde{\Sigma}_1|^{1/2}}\exp\left\{\frac{-1}{2}\begin{pmatrix}a_1-\tilde{a}_1\\c_1-\tilde{c}_1\\d_1-\tilde{d}_1\end{pmatrix}'\tilde{\Sigma}_1^{-1}\begin{pmatrix}a_1-\tilde{a}_1\\c_1-\tilde{c}_1\\d_1-\tilde{d}_1\end{pmatrix}-\frac{b_1}{\tilde{b}_1}\right\}. \tag{4.2}$$

2. Gompertz growth curve:

$a_2 \sim Expon(\frac{1}{\tilde{a}_2})$, $c_2 \sim Expon(\frac{1}{\tilde{c}_2})$, $b_2 \sim N(\tilde{b}_2, \tilde{\sigma}_b^2)$ then the vector

$\tilde{\theta}_2 = (\tilde{a}_2, \tilde{b}_2, \tilde{c}_2, \tilde{\sigma}_b^2)$. So the prior distribution is:

$$h(\theta_2|\tilde{\theta}_2) \propto \frac{1}{\tilde{a}_2 \tilde{c}_2} \frac{1}{\tilde{\sigma}_b} \exp\left\{ -\frac{1}{2} \frac{(b_2 - \tilde{b}_2)^2}{\tilde{\sigma}_b^2} - \left[\frac{a_2}{\tilde{a}_2} + \frac{c_2}{\tilde{c}_2} \right] \right\} . \quad (4.3)$$

3. Richards curve:

$a_3 \sim Expon(\frac{1}{\tilde{a}_3})$, $b_3 \sim Expon(\frac{1}{\tilde{b}_3})$, $\begin{pmatrix} c_3 \\ d_3 \end{pmatrix} \sim N_2\left(\begin{pmatrix} \tilde{c}_3 \\ \tilde{d}_3 \end{pmatrix}, \tilde{\Sigma}_3 \right)$, where c_3 and d_3 are bivariate normally distributed with mean vector $\begin{pmatrix} \tilde{c}_3 \\ \tilde{d}_3 \end{pmatrix}$, and variance covariance matrix $\tilde{\Sigma}_3$. The vector $\tilde{\theta}_3 = (\tilde{a}_3, \tilde{b}_3, \tilde{c}_3, \tilde{d}_3, \tilde{\Sigma}_3)$. So the prior distribution is:

$$h(\theta_3|\tilde{\theta}_3) \propto \frac{1}{\tilde{a}_3 \tilde{b}_3} \frac{1}{|\tilde{\Sigma}_3|^{1/2}} \exp\left\{ -\frac{1}{2} \begin{pmatrix} c_3 - \tilde{c}_3 \\ d_3 - \tilde{d}_3 \end{pmatrix}' \tilde{\Sigma}_3^{-1} \begin{pmatrix} c_3 - \tilde{c}_3 \\ d_3 - \tilde{d}_3 \end{pmatrix} - \left[\frac{a_3}{\tilde{a}_3} + \frac{b_3}{\tilde{b}_3} \right] \right\} . \quad (4.4)$$

4. Polynomial curve:

$\begin{pmatrix} a_4 \\ b_4 \\ c_4 \\ d_4 \end{pmatrix} \sim N_4\left(\begin{pmatrix} \tilde{a}_4 \\ \tilde{b}_4 \\ \tilde{c}_4 \\ \tilde{d}_4 \end{pmatrix}, \tilde{\Sigma}_4 \right)$, where $(a_4, b_4, c_4, d_4)'$ are multivariate normally distributed with mean vector $(\tilde{a}_4, \tilde{b}_4, \tilde{c}_4, \tilde{d}_4)'$, variance and covariance matrix is $\tilde{\Sigma}_4$. The vector $\tilde{\theta}_4 = (\tilde{a}_4, \tilde{b}_4, \tilde{c}_4, \tilde{d}_4, \tilde{\Sigma}_4)$. So the prior distribution is:

$$h(\theta_4|\tilde{\theta}_4)\propto\frac{1}{|\tilde{\Sigma}_4|^{1/2}}\exp\left\{-\frac{1}{2}\begin{pmatrix}a_4-\tilde{a}_4\\ b_4-\tilde{b}_4\\ c_4-\tilde{c}_4\\ d_4-\tilde{d}_4\end{pmatrix}'\tilde{\Sigma}_4^{-1}\begin{pmatrix}a_4-\tilde{a}_4\\ b_4-\tilde{b}_4\\ c_4-\tilde{c}_4\\ d_4-\tilde{d}_4\end{pmatrix}\right\}. \qquad (4.5)$$

We assume Jeffrey's prior for σ^2 and thus have density $p(\sigma^2)\propto\frac{1}{\sigma^2}$. But for the prior for ρ, we assume that $\rho^*=\frac{\rho+1}{2}\sim Beta(\tilde{\alpha},\tilde{\beta})$, where $\tilde{\alpha}$ and $\tilde{\beta}$ are some known constant for the beta prior distribution. After transformation we get the prior distribution for ρ, $p(\rho)\propto(\rho+1)^{\tilde{\alpha}-1}(1-\rho)^{\tilde{\beta}-1}$.

So the posterior distributions for a single experimental unit i are as follows:

1. Jenss growth curve:

$$\pi_1(\theta_1,\sigma^2,\rho_1\ |\boldsymbol{y},\boldsymbol{w})\propto\frac{(\rho_1+1)^{\tilde{\alpha}-1}(1-\rho_1)^{\tilde{\beta}-1}}{\sigma^{n+2}|\Omega|}$$
$$\exp\left\{-\frac{1}{2\sigma^2}A_1-\frac{1}{2}\begin{pmatrix}a_1-\tilde{a}_1\\ c_1-\tilde{c}_1\\ d_1-\tilde{d}_1\end{pmatrix}'\tilde{\Sigma}_1^{-1}\begin{pmatrix}a_1-\tilde{a}_1\\ c_1-\tilde{c}_1\\ d_1-\tilde{d}_1\end{pmatrix}-\frac{b_1}{\tilde{b}_1}\right\}.$$

Integrating out σ^2 we have

$$\pi_1(\theta_1,\rho_1|\boldsymbol{y},\boldsymbol{w})\propto\frac{(\rho_1+1)^{\tilde{\alpha}-1}(1-\rho_1)^{\tilde{\beta}-1}}{|\Omega|}$$
$$\exp\left\{-\left[\frac{1}{2}\begin{pmatrix}a_1-\tilde{a}_1\\ c_1-\tilde{c}_1\\ d_1-\tilde{d}_1\end{pmatrix}'\tilde{\Sigma}_1^{-1}\begin{pmatrix}a_1-\tilde{a}_1\\ c_1-\tilde{c}_1\\ d_1-\tilde{d}_1\end{pmatrix}+\frac{b_1}{\tilde{b}_1}\right]\right\}\cdot A_1^{-\frac{n}{2}}. \qquad (4.6)$$

2. Gompertz growth curve:

$$\pi_2(\theta_2, \sigma^2, \rho_2 \mid \boldsymbol{y}, \boldsymbol{w}) \propto \frac{(\rho_2+1)^{\tilde{\alpha}-1}(1-\rho_2)^{\tilde{\beta}-1}}{\sigma^{n+2}|\Omega|} \cdot \exp\left\{-\frac{1}{2\sigma^2}A_2 - \left[\frac{1}{2}\frac{(b_2-\tilde{b}_2)^2}{\tilde{\sigma}_b^2} + \frac{a_2}{\tilde{a}_2} + \frac{c_2}{\tilde{c}_2}\right]\right\} .$$

Integrating out σ^2 we have

$$\pi_2(\theta_2, \sigma^2, \rho_2 \mid \boldsymbol{y}, \boldsymbol{w}) \propto \frac{(\rho_2+1)^{\tilde{\alpha}-1}(1-\rho_2)^{\tilde{\beta}-1}}{|\Omega|} \cdot \exp\left\{-\left[\frac{1}{2}\frac{(b_2-\tilde{b}_2)^2}{\tilde{\sigma}_b^2} + \frac{a_2}{\tilde{a}_2} + \frac{c_2}{\tilde{c}_2}\right]\right\} A_2^{-\frac{n}{2}} \cdot \tag{4.7}$$

3. Richards growth curve:

$$\pi_3(\theta_3, \sigma^2, \rho_3 | \boldsymbol{y}, \boldsymbol{w}) \propto \frac{(\rho_3+1)^{\tilde{\alpha}-1}(1-\rho_3)^{\tilde{\beta}-1}}{\sigma^{n+2}|\Omega|} \exp\left\{-\frac{1}{2\sigma^2}A_3 - \frac{1}{2}\begin{pmatrix} c_3-\tilde{c}_3 \\ d_3-\tilde{d}_3 \end{pmatrix}' \tilde{\Sigma}_3^{-1} \begin{pmatrix} c_3-\tilde{c}_3 \\ d_3-\tilde{d}_3 \end{pmatrix} - \left[\frac{a_3}{\tilde{a}_3} + \frac{b_3}{\tilde{b}_3}\right]\right\} .$$

Integrating out σ^2 we have

$$\pi_3(\theta_3, \rho_3 \mid \boldsymbol{y}, \boldsymbol{w}) \propto \frac{(\rho_3+1)^{\tilde{\alpha}-1}(1-\rho_3)^{\tilde{\beta}-1}}{|\Omega|} \exp\left\{-\frac{1}{2}\begin{pmatrix} c_3-\tilde{c}_3 \\ d_3-\tilde{d}_3 \end{pmatrix}' \tilde{\Sigma}_3^{-1} \begin{pmatrix} c_3-\tilde{c}_3 \\ d_3-\tilde{d}_3 \end{pmatrix} - \left[\frac{a_3}{\tilde{a}_3} + \frac{b_3}{\tilde{b}_3}\right]\right\} A_3^{-\frac{n}{2}} \cdot \tag{4.8}$$

4. Polynomial growth curve:

$$\pi_4(\theta_4, \sigma^2, \rho_4 \mid \boldsymbol{y}, \boldsymbol{w}) \propto \frac{(\rho_4+1)^{\tilde{\alpha}-1}(1-\rho_4)^{\tilde{\beta}-1}}{\sigma^{n+2}|\Omega|} \exp\left\{-\frac{1}{2\sigma^2}A_4 - \frac{1}{2}\begin{pmatrix} a_4-\tilde{a}_4 \\ b_4-\tilde{b}_4 \\ c_4-\tilde{c}_4 \\ d_4-\tilde{d}_4 \end{pmatrix}' \tilde{\Sigma}_4^{-1} \begin{pmatrix} a_4-\tilde{a}_4 \\ b_4-\tilde{b}_4 \\ c_4-\tilde{c}_4 \\ d_4-\tilde{d}_4 \end{pmatrix}\right\} .$$

Integrating out σ^2 we have

$$\pi_4(\theta_4, \rho_4 \mid \boldsymbol{y}, \boldsymbol{w}) \propto \frac{(\rho_4+1)^{\tilde{\alpha}-1}(1-\rho_4)^{\tilde{\beta}-1}}{|\Omega|} \exp\left\{-\frac{1}{2}\begin{pmatrix} a_4-\tilde{a}_4 \\ b_4-\tilde{b}_4 \\ c_4-\tilde{c}_4 \\ d_4-\tilde{d}_4 \end{pmatrix}' \tilde{\Sigma}_4^{-1} \begin{pmatrix} a_4-\tilde{a}_4 \\ b_4-\tilde{b}_4 \\ c_4-\tilde{c}_4 \\ d_4-\tilde{d}_4 \end{pmatrix}\right\} A_4^{-\frac{n}{2}}. \quad (4.9)$$

Using the MH algorithm and taking the Gompertz growth curve for an example, we first need to get the full conditionals for the parameters from (4.7) as follows:

$$\pi_2(a_2 \mid \cdot) \propto \exp\left\{-\frac{a_2}{\tilde{a}_2}\right\} A_2^{-\frac{n}{2}}, \quad (4.10)$$

$$\pi_2(b_2 \mid \cdot) \propto \exp\left\{-\frac{1}{2}\frac{(b_2-\tilde{b}_2)^2}{\tilde{\sigma}_b^2}\right\} A_2^{-\frac{n}{2}}, \quad (4.11)$$

$$\pi_2(c_2 \mid \cdot) \propto \exp\left\{-\frac{c_2}{\tilde{c}_2}\right\} A_2^{-\frac{n}{2}}, \quad (4.12)$$

$$\pi_2(\rho_2 \mid \cdot) \propto \frac{(\rho_2+1)^{\tilde{\alpha}-1}(1-\rho_2)^{\tilde{\beta}-1}}{|\Omega|} A_2^{-\frac{n}{2}}. \quad (4.13)$$

We follow the same Metropolis-Hastings in Gibbs Sampling Algorithm as mentioned in the previous chapter, taking samples of θ_2 and ρ_2 by using (4.10)-(4.13) through the following steps:

1. Set $j=0$ and select a set of starting parameter values for $\rho_2^{(0)}$ and $\theta_2^{(0)}=(a_2^{(0)}, b_2^{(0)}, c_2^{(0)})$.
2. Sample $a_2^{(j+1)}$ from $\pi_2(a_2^{(j+1)}|\cdot)$ (using (4.10) and MH algorithm).
3. Sample $b_2^{(j+1)}$ from $\pi_2(b_2^{(j+1)}|\cdot)$ (using (4.11) and MH algorithm).
4. Sample $c_2^{(j+1)}$ from $\pi_2(c_2^{(j+1)}|\cdot)$ (using (4.12) and MH algorithm).

5. Sample $\rho_2^{(j+1)}$ from $\pi_2(\rho_2^{(j+1)}|\cdot)$ (using (4.13) and MH algorithm).
6. Replace $\theta_2^{(j)}$ with $\theta_2^{(j+1)}$ and $\rho_2^{(j)}$ with $\rho_2^{(j+1)}$.
7. Set $j = j+1$ and repeat steps 2 through 6.
8. Stop when $j = N + B$, (B is the number of burn-in samples to be dropped and N is sample size).

In the long run, the N samples taken from these full conditionals will be a sample drawn approximately from the above posterior distribution. In addition to this analysis and estimation of parameters, we can also generate highest density regions for the estimated parameters. The 90% confidence limits for the best-fit growth curve can be obtained by using the 5% and 95% percentiles of $\boldsymbol{y}$ (calculated by substituting the N samples of θ_2 at a given w_j).

4.3 Example: Cancer Growth in Mice

We use data (Koziol et al, 1981) collected from an experiment conducted at the Cancer Center of the University of California at San Diego to study the growth curve for colon cancer in a group of five mice. In this experiment, each of a homogeneous population of 30 mice was injected with 1000 CT26 (mouse colon carcinoma) tumor cells; five days later the population was randomly divided into three groups. We selected data from five mice (to simplify the analysis, the other mice that had at least one missing data were not included) from the group A (of 10 mice) that received only the injections of

tissue culture medium around the growing tumor to illustrate the application of our growth curve models.

The data set is listed in Table 4.3. We take the log of tumor size for the analysis, using the least square estimates as the empirical Bayes estimates for the priors of the four different models, and then conducted Bayesian analysis using the above mentioned MH algorithm within Gibbs sampling. The results of parameter estimates are given in Tables 4.4 to 4.5.

Based on the results shown in Table 4.5 and Figure 4.1, we could determine which curve to choose by looking at goodness of fit statistics (such as adjusted R^2) and credible intervals (the narrowest among all four models). Note that all the parameter estimates include estimates for autocorrelation coefficients which calibrate the other parameters estimates (leading to smaller biases) in the Bayesian sampling process as they adjust for the autocorrelating effects. Although the degree of the effects of autocorrelation cannot be actually or correctly measured (as illustrated in the simulation example in the beginning of this chapter), their influence on our parameter estimations actually follows the same principle.

4.4 Summary

In summary, we provide the Bayesian analysis of four commonly used growth curves with autocorrelation in these models. The simulation and real data example display the importance as well as the practical application for this kind of model.

Table 4.3 Colon Cancer Tumor Size (Growth) in A Group of Mice

Days	Tumor Size
7	151.9
11	618.9
12	762.3
13	1107.3
14	1418.1
15	1627.1
17	1965.8
18	2999.1
19	3306.6
20	3146.9
21	3501.5

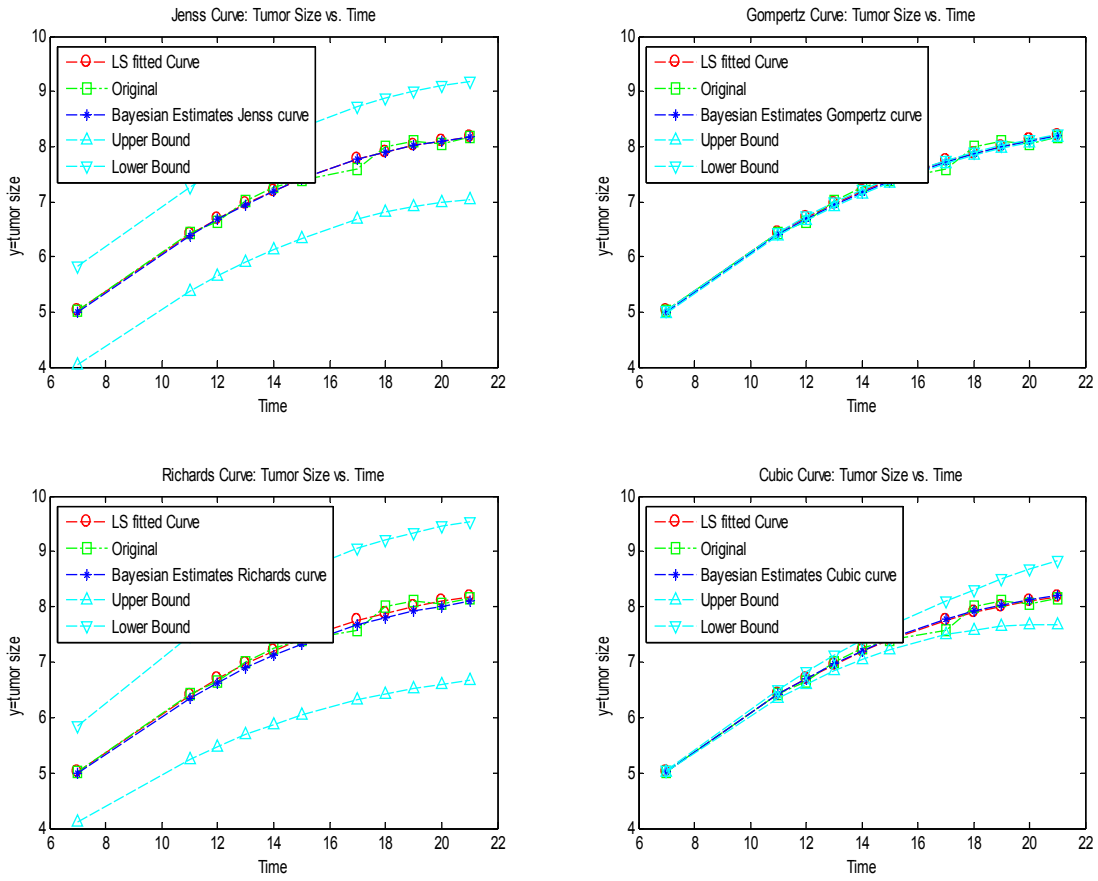
Table 4.4 Bayesian Parameter Estimates from Different Models

Parameters	Jenss Curve	Gompertz Curve	Richards Curve	Cubic Curve
a	115.705	8.847	8.471	1.659
b	2.067	0.448	1.605	0.565
c	4.736	0.143	0.205	-0.012
d	0.013	-	6.134	0.000
ρ	0.116	0.123	0.082	0.116

Table 4.5 BIC Values for Different Growth Curves

Growth Curves	BIC Values	Adjusted R^2
Jenss Growth Curve	0.4282	0.990
Gompertz Growth Curve	0.2693	0.989
Richards Growth Curve	0.4839	0.979
Cubic Polynomal Growth Curve	0.4277	0.989

Figure 4.1 90% Credible Intervals for the Four Different Curves



References

Koziol, J. A., D. A. Maxwell, M. Fukushima, M.E.M. Colmerauer and Y. H. Pilch (1981). A distribution-free test for tumor-growth curve analyses with application to an animal tumor immunotherapy experiment. *Biometrics*, **37**, pp. 383-390.

Seber, G. A. F. and C. J. Wild (1989). *Nonlinear Regression*. John Wiley & Sons.

Chapter 5 Bayesian Analysis of Commonly Used Growth Curves with Auxiliary Information and AR(1) Covariance Structure

5.0 Abstract

Generally, auxiliary variables may provide additional information for the dependent variables and one can fit a better model by incorporating auxiliary variables. We exhibit the need and importance of considering auxiliary variables in growth curve models with simulation examples.

The purpose of this chapter is to show our readers how Bayesian analysis can be applied to situations where growth curves have both auxiliary information and AR(1) covariance structure with a view to lead them to chapter six through eight, where the novelty of our research is detailed.

5.1 Introduction and Motivation

In addition to AR(1) correlation structure in models as shown in previous chapter, we will demonstrate the utility of including auxiliary variables if they are available and into our growth curve model in this chapter by assuming the following,

$y_i = f(w_i;\theta) + q(z_i,\phi) + \varepsilon_i,\ i = 1,\ldots,m$, $\varepsilon_i = \rho\,\varepsilon_{i-1} + \eta_i,\ \ \eta_i \sim N(0,\sigma_\eta^2)$,

where $f(w_i;\beta) = \theta_0 + \theta_1 w_i + \theta_2 w_i^2 + \theta_3 w_i^3$ in which $\theta = (\theta_0\ \ \theta_1\ \ \theta_2\ \ \theta_3)$, ε_i is the error term for the i^{th} observation and ρ is the correlation coefficient. Let the m points of observation for the variable time be $W = [w_1, w_2, \ldots, w_m]'$ and the auxiliary variable vector be $Z = [z_1, z_2, \ldots, z_m]'$. We assume a linear function for the auxiliary

variable, i.e., $q(z_i, \phi) = \phi z_i$, where ϕ is the coefficient, to study the effect of the auxiliary variable z_i on the model.

We ran a simulation of 3000 iterations according to different scenarios, comparing whether there are significant differences in the estimates (using least square estimates) of the coefficients or not. Please refer to Table 5.1 and 5.2 for details of the simulation.

From the results shown in Table 5.1 and 5.2, we observed the influences the auxiliary variables (assuming linear functions) exerted on the models. Basically, this simulation shows the need of including auxiliary variables in the model (in addition to the adverse effect of not considering autocorrelation in the model, which was already detailed in previous chapter).

When auxiliary variable(s) are present, then ignoring them will decrease the goodness of fit of the model with respect to the simulated (or real) data (and this is dependent on the magnitude of the auxiliary variables). We also noticed that the biasedness of the estimates of the parameters from a model without considering auxiliary variable(s) is not insignificant and it will be even more when the variances of the error terms become larger. Therefore we see the importance of including auxiliary variable(s) in models from these simulation results.

Table 5.1 Auxiliary Variable vs. No Auxiliary Variable (1)

Aux.	ρ	$Adj.\ R^2$	SSE	θ_0	std	θ_1	std	θ_2	std	θ_3	std	ϕ	std
No	0.0	0.9990	20.94	0.510	0.267	-2.337	0.300	2.652	0.100	0.830	0.010		
Yes	0.0	0.9997	2.84	1.034	0.268	1.077	0.323	1.006	0.103	1.002	0.010	1.002	0.020
No	0.4	0.9986	26.94	0.299	0.246	-2.249	0.283	2.873	0.091	0.788	0.009		
Yes	0.4	0.9997	2.61	1.006	0.246	1.092	0.275	1.002	0.092	1.003	0.009	1.004	0.012

Note:

1. σ_η^2 =1, $Z=[z_1, z_2, ..., z_6]'$ are random samples from $N(0,1)$ with 3000 iterations using data simulated from a given parameter vector $\theta=(1\ \ 1\ \ 1\ \ 1)$, $\phi=1$ and $W=[w_1, w_2, ..., w_6]'=[1,\ 1.1,\ ...,\ 1.5]'$.
2. The column label "Aux." denotes the results without (No) or with (Yes) auxiliary variable(s) in a model. The results show that the parameter estimates for models without auxiliary variables are biased and those for models with auxiliary variable(s) are better (unbiased).

Table 5.2 Auxiliary Variable vs. No Auxiliary Variable (2)

Aux.	ρ	$Adj.\ R^2$	SSE	θ_0	std	θ_1	std	θ_2	std	θ_3	std	ϕ	std
No	0.0	0.9958	74.09	-1.763	0.827	7.973	0.899	-1.634	0.289	1.232	0.027		
Yes	0.0	0.9966	29.93	1.152	0.830	1.194	0.923	1.042	0.305	1.006	0.029	1.004	0.033
No	0.4	0.9960	70.88	-7.985	0.712	13.570	0.889	-1.838	0.277	1.143	0.026		
Yes	0.4	0.9994	24.76	1.080	0.788	1.184	0.908	1.045	0.286	1.004	0.026	1.002	0.025

Note:

This scenario is similar to that in Table 5.1 only with σ_η^2 =9 and that $Z=[z_1, z_2, ..., z_6]'$ are random samples from $N(0,9)$. So we observed similar results except that the standard deviations and SSE are a little bit larger than those in previous scenario because their variances in error terms are larger.

5.2 Model Formulation

We adopt the same notations as in previous chapter and assume a growth model for a single experimental unit of the form $y_j=f(w_j,\theta)+q(z_j,\phi)+\varepsilon_j$, $j=1,...,n$,

n is the total number of observations; θ is a vector of parameters; w_j is the independent variable for the j^{th} observation; q , with a vector of coefficients ϕ , denotes the function of a vector of auxiliary variable $z=\{z_j,\ j=1,\ldots,n\}$; $\varepsilon_j \sim N(0,\ \sigma^2)$ is independent random error of y_j . Here $f(w_j,\theta)$ can be one of the different types of growth curves. We assume that the covariance for the j^{th} and k^{th} observation is $\Sigma=\sigma^2\Omega$, where

$$\Omega = \begin{bmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ \rho & 1 & \cdots & \rho^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & 1 \end{bmatrix}_{n\times n}$$, ρ is the autocorrelation coefficient.

Let $f_l(w_j,\ \theta_l),\ l=1,2,3,4$ be four different types of growth curves. Recall four commonly used growth curves from previous chapters:

1. Jenss growth curve $f_1(w_j,\theta_1)=a_1+b_1w_j-\exp(c_1+d_1w_j)$, a vector of parameters $\theta_1'=(a_1\ \ b_1\ \ c_1\ \ d_1)$, where $a_1,b_1>0$, c_1 and $d_1\in\mathbb{R}$.
2. Gompertz growth curve $f_2(w_j,\theta_2)=a_2\cdot\exp[-\exp(b_2-c_2w_j)]$ and a vector of parameters $\theta_2'=(a_1\ \ b_2\ \ c_2)$, where $a_2,\ c_2>0$, $b_2\in\mathbb{R}$.
3. Richards growth curve $f_3(w_j,\theta_3)=a_3\cdot\{1+b_3\exp[c_3(d_3-w_j)]\}^{-1/b_3}$, and a vector of parameters $\theta_3'=(a_3\ \ b_3\ \ c_3\ \ d_3)$, where $a_3,b_3>0$, $c_3,d_3\in\mathbb{R}$.
4. Polynomial growth curve $f_4(w_j,\theta_4)=a_4+b_4w_j+c_4w_j^2+d_4w_j^3$, and a vector of parameters $\theta_4'=(a_4\ \ b_4\ \ c_4\ \ d_4)$, where a_4 , b_4 , c_4 and $d_4\in\mathbb{R}$.

Similar to previous chapter, we will let this $f_l(w_j, \theta_l)$, $l=1,2,3,4$ be substituted into the growth model and denote $\boldsymbol{y}=\{y_j,\ j=1,\ldots,n\}$, $\boldsymbol{w}=\{w_j,\ j=1,\ldots,n\}$, $\boldsymbol{z}=\{z_j, j=1,\ldots,n\}$ and

$A_l=[\boldsymbol{y}-f_l(\boldsymbol{w},\theta_l)-q(\boldsymbol{z},\phi_l)]'_{1\times n}\Omega^{-1}_{n\times n}[\boldsymbol{y}-f_l(\boldsymbol{w},\theta_l)-q(\boldsymbol{z},\phi_l)]_{n\times 1}$ for $l=1,2,3,4$.

Then we get their likelihood functions, respectively, as follows:

$$g(\boldsymbol{y}|\theta_l,\phi_l,\sigma^2,\rho,\boldsymbol{w},\boldsymbol{z})=\frac{1}{(2\pi)^{n/2}\sigma^n|\Omega|}\exp\left\{-\frac{1}{2\sigma^2}A_l\right\}. \qquad (5.1)$$

Now we propose the following distributions for the four growth curve priors (note that empirical Bayes estimate vectors $\tilde{\theta}_i, i=1,\ldots,4$ in the following four curves can be obtained from nonlinear least square regression results by using MATLAB *nlinfit* function to fit nonlinear Jenss, Gompertz and Richards curves and *polyfit* function to fit polynomial curves. For details, please refer to MATLAB Help and Seber and Wild, 1989):

1. Jenss curve with auxiliary variable:

   $a_1\sim Expon(\frac{1}{\tilde{a}_1})$, $b_1\sim Expon(\frac{1}{\tilde{b}_1})$ $\begin{pmatrix}c_1\\ d_1\\ \phi_1\end{pmatrix}\sim N_3\left(\begin{pmatrix}\tilde{c}_1\\ \tilde{d}_1\\ \tilde{\phi}_1\end{pmatrix},\tilde{\Sigma}_1\right)$ where

   $(c_1,d_1,\phi_1)'$ are multivariate normally distributed with mean vector

   $(\tilde{c}_1,\tilde{d}_1,\tilde{\phi}_1)'$, and variance covariance matrix is $\tilde{\Sigma}_1$. The vector

   $\tilde{\theta}_1=(\tilde{a}_1,\tilde{b}_1,\tilde{c}_1,\tilde{d}_1,\tilde{\phi}_1,\tilde{\Sigma}_1)$. So the prior distribution can be written as follows:

$$h(\theta_1|\tilde{\theta}_1)\propto\frac{1}{\tilde{a}_1}\frac{1}{\tilde{b}_1}\frac{1}{|\tilde{\Sigma}_1|^{1/2}}\exp\left\{\frac{-1}{2}\begin{pmatrix}c_1-\tilde{c}_1\\ d_1-\tilde{d}_1\\ \phi_1-\tilde{\phi}_1\end{pmatrix}'\tilde{\Sigma}_1^{-1}\begin{pmatrix}c_1-\tilde{c}_1\\ d_1-\tilde{d}_1\\ \phi_1-\tilde{\phi}_1\end{pmatrix}-\frac{a_1}{\tilde{a}_1}-\frac{b_1}{\tilde{b}_1}\right\} \quad . \qquad (5.2)$$

2. Gompertz growth curve with auxiliary variable:

$a_2\sim Expon(\frac{1}{\tilde{a}_2})$, $c_2\sim Expon(\frac{1}{\tilde{c}_2})$, $b_2\sim N(\tilde{b}_2,\tilde{\sigma}_b^2)$, $\phi_2\sim N(\tilde{\phi}_2,\tilde{\sigma}_\phi)$ then

$\tilde{\theta}_2=(\tilde{a}_2,\tilde{b}_2,\tilde{c}_2,\tilde{\phi}_2,\tilde{\sigma}_b^2,\tilde{\sigma}_\phi^2)$. So the prior distribution is:

$$h(\theta_2|\tilde{\theta}_2)\propto\frac{1}{\tilde{a}_2\tilde{c}_2}\frac{1}{\tilde{\sigma}_b\tilde{\sigma}_\phi}\exp\left\{-\frac{1}{2}\left[\frac{(b_2-\tilde{b}_2)^2}{\tilde{\sigma}_b^2}+\frac{(\phi_2-\tilde{\phi}_2)^2}{\tilde{\sigma}_\phi^2}\right]-\left[\frac{a_2}{\tilde{a}_2}+\frac{c_2}{\tilde{c}_2}\right]\right\} . \qquad (5.3)$$

3. Richards curve with auxiliary variable:

$a_3\sim Expon(\frac{1}{\tilde{a}_3})$, $b_3\sim Expon(\frac{1}{\tilde{b}_3})$, $\begin{pmatrix}c_3\\ d_3\\ \tilde{\phi}_3\end{pmatrix}\sim N_3\left(\begin{pmatrix}\tilde{c}_3\\ \tilde{d}_3\\ \tilde{\phi}_3\end{pmatrix},\tilde{\Sigma}_3\right)$, where

$(c_3,d_3,\phi_3)'$ are multivariate normally distributed with mean vector

$(\tilde{c}_3,\tilde{d}_3,\tilde{\phi}_3)'$, and variance covariance matrix $\tilde{\Sigma}_3$. The vector

$\tilde{\theta}_3=(\tilde{a}_3,\tilde{b}_3,\tilde{c}_3,\tilde{d}_3,\tilde{\phi}_3,\tilde{\Sigma}_3)$. So the prior distribution is:

$$h(\theta_3|\tilde{\theta}_3)\propto\frac{1}{\tilde{a}_3\tilde{b}_3}\frac{1}{|\tilde{\Sigma}_3|^{1/2}}\exp\left\{-\frac{1}{2}\begin{pmatrix}c_3-\tilde{c}_3\\ d_3-\tilde{d}_3\\ \phi_3-\tilde{\phi}_3\end{pmatrix}'\tilde{\Sigma}_3^{-1}\begin{pmatrix}c_3-\tilde{c}_3\\ d_3-\tilde{d}_3\\ \phi_3-\tilde{\phi}_3\end{pmatrix}-\left[\frac{a_3}{\tilde{a}_3}+\frac{b_3}{\tilde{b}_3}\right]\right\} . \qquad (5.4)$$

4. Polynomial curve with auxiliary variable:

$$\begin{pmatrix} a_4 \\ b_4 \\ c_4 \\ d_4 \\ \phi_4 \end{pmatrix} \sim N_5 \left(\begin{pmatrix} \tilde{a}_4 \\ \tilde{b}_4 \\ \tilde{c}_4 \\ \tilde{d}_4 \\ \tilde{\phi}_4 \end{pmatrix}, \tilde{\Sigma}_4 \right)$$, where $(a_4, b_4, c_4, d_4, \phi_4)'$ are multivariate normally

distributed with mean vector $(\tilde{a}_4, \tilde{b}_4, \tilde{c}_4, \tilde{d}_4, \tilde{\phi}_4)'$, variance and covariance matrix is $\tilde{\Sigma}_4$. The vector $\tilde{\theta}_4 = (\tilde{a}_4, \tilde{b}_4, \tilde{c}_4, \tilde{d}_4, \tilde{\phi}_4, \tilde{\Sigma}_4)$. So the prior distribution is:

$$h(\theta_4 | \tilde{\theta}_4) \propto \frac{1}{|\tilde{\Sigma}_4|^{1/2}} \exp\left\{ -\frac{1}{2} \begin{pmatrix} a_4 - \tilde{a}_4 \\ b_4 - \tilde{b}_4 \\ c_4 - \tilde{c}_4 \\ d_4 - \tilde{d}_4 \\ \phi_4 - \tilde{\phi}_4 \end{pmatrix}' \tilde{\Sigma}_4^{-1} \begin{pmatrix} a_4 - \tilde{a}_4 \\ b_4 - \tilde{b}_4 \\ c_4 - \tilde{c}_4 \\ d_4 - \tilde{d}_4 \\ \phi_4 - \tilde{\phi}_4 \end{pmatrix} \right\} . \qquad (5.5)$$

We assume Jeffrey's prior for σ^2 and thus have density $p(\sigma^2) \propto \frac{1}{\sigma^2}$. But for the prior for ρ , we assume that $\rho^* = \frac{\rho + 1}{2} \sim Beta(\tilde{\alpha}, \tilde{\beta})$, where $\tilde{\alpha}$ and $\tilde{\beta}$ are some known constant for the beta prior distribution. After transformation we get the prior distribution for ρ , $p(\rho) \propto (\rho + 1)^{\tilde{\alpha}-1} (1 - \rho)^{\tilde{\beta}-1}$. We further assume that $q = \phi z_j, j = 1, \ldots, n$ and prior for $\phi \sim N(0, \tilde{\sigma}_\phi^2)$ (where $\tilde{\sigma}_\phi^2$ is a known constant). So the posterior distributions are as follows:

1. Jenss growth curve (with auxiliary variable):

$$\pi_1(\theta_1,\sigma^2,\rho_1,\phi_1\ |\boldsymbol{y},\boldsymbol{w},\boldsymbol{z})\propto\frac{(\rho_1+1)^{\tilde{\alpha}-1}(1-\rho_1)^{\tilde{\beta}-1}}{\sigma^{n+2}|\Omega|}$$
$$\exp\left\{-\frac{1}{2\sigma^2}A_1-\frac{1}{2}\begin{pmatrix}c_1-\tilde{c}_1\\ d_1-\tilde{d}_1\\ \phi_1-\tilde{\phi}_1\end{pmatrix}'\tilde{\Sigma}_1^{-1}\begin{pmatrix}c_1-\tilde{c}_1\\ d_1-\tilde{d}_1\\ \phi_1-\tilde{\phi}_1\end{pmatrix}-\frac{a_1}{\tilde{a}_1}-\frac{b_1}{\tilde{b}_1}\right\}\ .$$

Integrating out σ^2 we have

$$\pi_1(\theta_1,\rho_1,\phi_1\ |\boldsymbol{y},\boldsymbol{w},\boldsymbol{z})\propto\frac{(\rho_1+1)^{\tilde{\alpha}-1}(1-\rho_1)^{\tilde{\beta}-1}}{|\Omega|}$$
$$\exp\left\{-\left(\frac{a_1}{\tilde{a}_1}+\frac{b_1}{\tilde{b}_1}+\frac{1}{2}\begin{pmatrix}c_1-\tilde{c}_1\\ d_1-\tilde{d}_1\\ \phi_1-\tilde{\phi}_1\end{pmatrix}'\tilde{\Sigma}_1^{-1}\begin{pmatrix}c_1-\tilde{c}_1\\ d_1-\tilde{d}_1\\ \phi_1-\tilde{\phi}_1\end{pmatrix}\right)\right\}\cdot A_1^{-\frac{n}{2}}\ . \quad (5.6)$$

2. Gompertz growth curve (with auxiliary variable):

$$\pi_2(\theta_2,\sigma^2,\rho_2,\phi_2\ |\boldsymbol{y},\boldsymbol{w},\boldsymbol{z})\propto\frac{(\rho_2+1)^{\tilde{\alpha}-1}(1-\rho_2)^{\tilde{\beta}-1}}{\sigma^{n+2}|\Omega|}$$
$$\exp\left\{-\frac{A_2}{2\sigma^2}-\frac{1}{2}\left[\frac{(\phi_2-\tilde{\phi}_2)^2}{\tilde{\sigma}_\phi^2}+\frac{(b_2-\tilde{b}_2)^2}{\tilde{\sigma}_b^2}\right]-\left[\frac{a_2}{\tilde{a}_2}+\frac{c_2}{\tilde{c}_2}\right]\right\}\ .$$

Integrating out σ^2 we have

$$\pi_2(\theta_2,\rho_2,\phi_2\ |\boldsymbol{y},\boldsymbol{w},\boldsymbol{z})\propto\frac{(\rho_2+1)^{\tilde{\alpha}-1}(1-\rho_2)^{\tilde{\beta}-1}}{|\Omega|}$$
$$\exp\left\{-\frac{1}{2}\left[\frac{(\phi_2-\tilde{\phi}_2)^2}{\tilde{\sigma}_\phi^2}+\frac{(b_2-\tilde{b}_2)^2}{\tilde{\sigma}_b^2}\right]-\left[\frac{a_2}{\tilde{a}_2}+\frac{c_2}{\tilde{c}_2}\right]\right\}\cdot A_2^{-\frac{n}{2}}\ . \quad (5.7)$$

3. Richards growth curve (with auxiliary variable):

$$\pi_3(\theta_3, \sigma^2, \rho_3, \phi_3 \mid \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{z}) \propto \frac{(\rho_3+1)^{\tilde{\alpha}-1}(1-\rho_3)^{\tilde{\beta}-1}}{\sigma^{n+2}|\Omega|} \exp\left\{-\frac{1}{2\sigma^2}A_3 - \frac{1}{2}\begin{pmatrix} c_3-\tilde{c}_3 \\ d_3-\tilde{d}_3 \\ \phi_3-\tilde{\phi}_3 \end{pmatrix}' \tilde{\Sigma}_3^{-1} \begin{pmatrix} c_3-\tilde{c}_3 \\ d_3-\tilde{d}_3 \\ \phi_3-\tilde{\phi}_3 \end{pmatrix} - \left[\frac{a_3}{\tilde{a}_3}+\frac{b_3}{\tilde{b}_3}\right]\right\} \cdot$$

Integrating out σ^2 we have

$$\pi_3(\theta_3, \rho_3, \phi_3 \mid \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{z}) \propto \frac{(\rho_3+1)^{\tilde{\alpha}-1}(1-\rho_3)^{\tilde{\beta}-1}}{|\Omega|} \exp\left\{-\frac{1}{2}\begin{pmatrix} c_3-\tilde{c}_3 \\ d_3-\tilde{d}_3 \\ \phi_3-\tilde{\phi}_3 \end{pmatrix}' \tilde{\Sigma}_3^{-1} \begin{pmatrix} c_3-\tilde{c}_3 \\ d_3-\tilde{d}_3 \\ \phi_3-\tilde{\phi}_3 \end{pmatrix} - \left[\frac{a_3}{\tilde{a}_3}+\frac{b_3}{\tilde{b}_3}\right]\right\} \cdot A_3^{-\frac{n}{2}} \cdot$$

(5.8)

4. Polynomial growth curve (with auxiliary variable):

$$\pi_4(\theta_4, \sigma^2, \rho_4, \phi_4 \mid \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{z}) \propto \frac{(\rho_4+1)^{\tilde{\alpha}-1}(1-\rho_4)^{\tilde{\beta}-1}}{\sigma^{n+2}|\Omega|} \exp\left\{-\frac{A_4}{2\sigma^2} - \frac{1}{2}\begin{pmatrix} a_4-\tilde{a}_4 \\ b_4-\tilde{b}_4 \\ c_4-\tilde{c}_4 \\ d_4-\tilde{d}_4 \\ \phi_4-\tilde{\phi}_4 \end{pmatrix}' \tilde{\Sigma}_4^{-1} \begin{pmatrix} a_4-\tilde{a}_4 \\ b_4-\tilde{b}_4 \\ c_4-\tilde{c}_4 \\ d_4-\tilde{d}_4 \\ \phi_4-\tilde{\phi}_4 \end{pmatrix}\right\} \cdot$$

Integrating out σ^2 we have

$$\pi_4(\theta_4,\rho_4,\phi_4|\boldsymbol{y},\boldsymbol{w},\boldsymbol{z})\propto\frac{(\rho_4+1)^{\tilde{\alpha}-1}(1-\rho_4)^{\tilde{\beta}-1}}{|\Omega|}$$
$$\exp\left\{-\frac{1}{2}\begin{pmatrix}a_4-\tilde{a}_4\\ b_4-\tilde{b}_4\\ c_4-\tilde{c}_4\\ d_4-\tilde{d}_4\\ \phi_4-\tilde{\phi}_4\end{pmatrix}'\tilde{\Sigma}_4^{-1}\begin{pmatrix}a_4-\tilde{a}_4\\ b_4-\tilde{b}_4\\ c_4-\tilde{c}_4\\ d_4-\tilde{d}_4\\ \phi_4-\tilde{\phi}_4\end{pmatrix}\right\}A_4^{-\frac{n}{2}}. \tag{5.9}$$

Using a similar MH algorithm and taking the Gompertz growth curve for example, we first need to get the full conditionals for the parameters from (5.7) as follows:

$$\pi_2(a_2\ |\cdot)\propto\exp\left\{-\frac{a_2}{\tilde{a}_2}\right\}A_2^{-\frac{n}{2}}, \tag{5.10}$$

$$\pi_2(b_2\ |\cdot)\propto\exp\left\{-\frac{1}{2}\frac{(b_2-\tilde{b}_2)^2}{\tilde{\sigma}_b^2}\right\}A_2^{-\frac{n}{2}}, \tag{5.11}$$

$$\pi_2(c_2\ |\cdot)\propto\exp\left\{-\frac{c_2}{\tilde{c}_2}\right\}A_2^{-\frac{n}{2}}, \tag{5.12}$$

$$\pi_2(\rho_2\ |\cdot)\propto\frac{(\rho_2+1)^{\tilde{\alpha}-1}(1-\rho_2)^{\tilde{\beta}-1}}{|\Omega|}A_2^{-\frac{n}{2}}, \tag{5.13}$$

$$\pi_2(\phi_2\ |\cdot)\propto\exp\left\{-\frac{1}{2}\frac{(\phi_2-\tilde{\phi}_2)^2}{\tilde{\sigma}_\phi^2}\right\}A_2^{-\frac{n}{2}}. \tag{5.14}$$

We follow the same Metropolis-Hastings Algorithm as mentioned in previous chapter in taking samples of θ_2 , ρ_2 and ϕ_2 by using (5.10)-(5.14) through the following steps:

1. Set $j=0$ and select a set of starting parameter values for $\rho_2^{(0)}$ and $\theta_2^{(0)}=(a_2^{(0)}, b_2^{(0)}, c_2^{(0)})$.
2. Sample $a_2^{(j+1)}$ from $\pi_2(a_2^{(j+1)}|\cdot)$ (using (5.10) and MH algorithm).
3. Sample $b_2^{(j+1)}$ from $\pi_2(b_2^{(j+1)}|\cdot)$ (using (5.11) and MH algorithm).
4. Sample $c_2^{(j+1)}$ from $\pi_2(c_2^{(j+1)}|\cdot)$ (using (5.12) and MH algorithm).
5. Sample $\rho_2^{(j+1)}$ from $\pi_2(\rho_2^{(j+1)}|\cdot)$ (using (5.13) and MH algorithm).
6. Sample $\phi_2^{(j+1)}$ from $\pi_2(\phi_2^{(j+1)}|\cdot)$ (using (5.14) and MH algorithm).
7. Replace $\theta_2^{(j)}$ with $\theta_2^{(j+1)}$, $\rho_2^{(j)}$ with $\rho_2^{(j+1)}$ and $\phi_2^{(j)}$ with $\phi_2^{(j+1)}$.
8. Set $j=j+1$ and repeat steps 2 through 7.
9. Stop when $j=N+B$, (B is burn-in samples to be dropped and N is sample size).

In the long run, the N samples taken from these full conditionals will be a sample drawn approximately from the above posterior distribution. In addition to the analysis and estimation of parameters, we can also generate highest density regions for the estimated parameters. The 90% confidence limits for the best-fit growth curve can be obtained by using the 5% and 95% percentiles of $\boldsymbol{y}$ (calculated by substituting the N samples of θ_2 at given $\boldsymbol{w}$).

5.3 An Illustrative Example

We use a data set with $m=8$ (time points) and the same assumptions as in Table 5.1 ($\rho=0.4$) to illustrate how our model can be used to find a best fit model. They are plotted as the dashed line (with small circles) at each time point as in Figure 5.1.

Table 5.3 displays the results of Bayesian estimates of the parameters according to the model presented in this chapter. Firstly, the comparison of the adjusted R^2 values (and the corresponding sum of squares of errors SSE as well) between "with" and "without" auxiliary variables shows that the inclusion of an auxiliary variable into the model is reasonable, necessary and justifiable. Thus, ignoring or excluding auxiliary variables may result in models that are not reliable.

Secondly, the three major criteria we look at in determining which curve (model) is the best fit are 1) goodness of fit statistics 2) BIC and 3) credible intervals (90%). It is obvious that the Cubic curve model is better than the other models in terms of having the largest adjusted R^2(=1), the smallest SSE (=40.190) and the narrowest confidence intervals (almost overlapping and hard to identify from Figure 5.1). Gompertz curve model also fit the data quite well (R^2 is almost equal 1) because it uses fewer parameters (and that's also why its BIC values is the smallest among the four models) but it's credible interval is wider as time progresses.

Table 5.3

Different Growth Curves with Auxiliary Variable AR(1) in Model

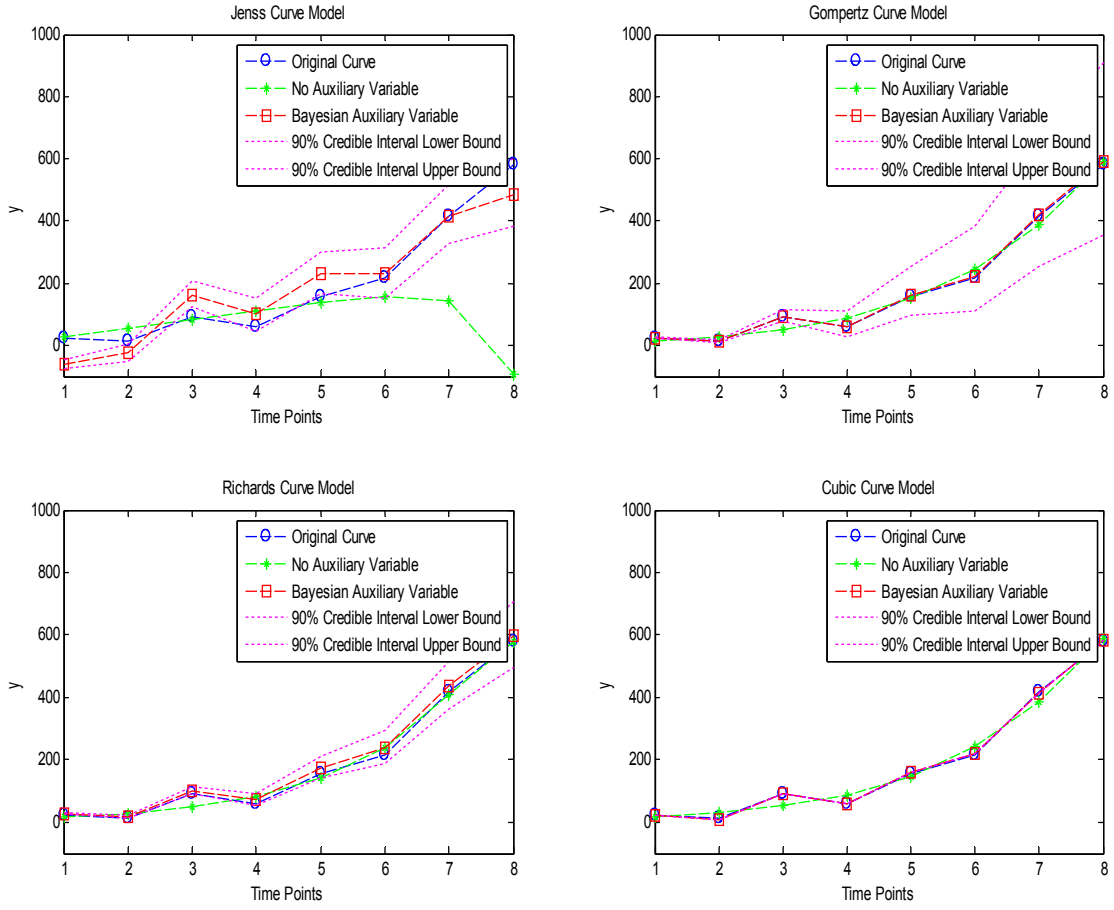
Parameters	Jenss	Gompertz	Richards	Cubic
a	174.912	3367.326	3527.398	7.804
b	79.445	2.050	0.008	7.475
c	*	0.185	0.177	3.499
d	*	-	11.358	0.775
ϕ	1.901	1.035	1.040	1.055
ρ	0.379	0.376	0.380	0.377
Goodness of Fit Statistics				
BIC	13.247	4.592	4.801	6.150
Adj. R^2 (Auxiliary)	0.764	0.999	0.987	1.000
Adj. R^2 (No Auxiliary)	0.714	0.984	0.978	0.973
SSE(Auxiliary)	*	154.780	1590.650	40.190
SSE(No Auxiliary)	*	4770.860	3608.580	4557.560

* Note: because the two parameters (c and d , marked by *) of the Jenss curve can not be very well estimated in this specific data set, its goodness of fit statistics is poorer than those of the other three curves.

5.4 Summary

Auxiliary variables have been shown to be a crucial component in a growth curve model as ignoring its effects will results in a model with lower goodness of fit scores and biased estimates for the parameters. Also, based on some criteria, such as goodness of fit statistics (including credible intervals) we can fit a commonly used growth curve for our model to better describe the data with different combinations of parameters.

Figure 5.1 Credible Intervals for the Four Different Curves



References

Seber, G. A. F. and C. J. Wild (1989). *Nonlinear Regression*. John Wiley & Sons.

Chapter 6 Bayesian Analysis of Growth Curve Models with Errors-in-variables in Auxiliary Information and AR(1) Covariance Structure

6.0 Abstract

We propose to analyze our data using models incorporating errors-in-variables (EIV) in auxiliary information and with autoregressive covariance structure through Bayesian methodology. The incorporation of these components in our growth model is necessary and realistic in the study of many statistical problems. The approach of classical frequentist analysis usually mandates many simplifying assumptions to reduce the complexity of the problems, or else analytic solutions will be impossible. In contrast, a Bayesian approach, with its computational advantages, can be effective in dealing with the complexity of these types of models. Though much research (especially using traditional approach) has been clustered in this area, models similar to the ones proposed here have been non-existent in literature.

6.1 Motivation

In this section, we demonstrate the need for combining and including EIV in auxiliary variables with AR(1) autocorrelation structure in our growth curve model.

Let the simulation model which includes EIV, auxiliary variables and AR(1) be as follows:

$$y_i = f(w_i;\theta) + q(z_i,\phi) + \varepsilon_i,\ i = 1,\dots,m \quad , \quad \varepsilon_i = \rho\,\varepsilon_{i-1} + \eta_i, \quad \eta_i \sim N(0,\sigma_\eta^2) \quad ,$$

$$x_i = z_i + e_i, \quad e_i \sim iid\ N(0,\sigma_e^2) \quad ,$$

where $f(w_i;\beta) = \theta_0 + \theta_1 w_i + \theta_2 w_i^2 + \theta_3 w_i^3$, $\theta = (\theta_0 \ \ \theta_1 \ \ \theta_2 \ \ \theta_3)$, ε_i is the error term for the i^{th} observation and ρ is the correlation coefficient, $Y = \left[y_1, y_2, \dots, y_m\right]'$. Let

the m points of observation for the variable time be $W=[w_1, w_2, \ldots, w_m]'$ and $Z=[z_1, z_2, \ldots, z_m]'$. We assume the linear function for q to be $q(z_i, \phi)=\phi z_i$. We then run simulations of 3000 iterations according to different scenarios, and compare whether or not there are significant differences in the estimates (using least square estimation) of the coefficients.

From the results shown in Table 6.1-6.2, we observed the influences EIV in auxiliary variable exerted on the models. Basically, these simulation results show that including EIV in auxiliary variables with AR(1) in the model will have advantages (mainly in terms of goodness of fit statistics) over models without them.

In Table 6.2, a scenario in which the simulation data are contaminated by EIV in its auxiliary variable, we see the advantages of incorporating EIV and auxiliary variable in model. We could say that in terms of goodness of fit statistics, a model with EIV in auxiliary variables is better than one with auxiliary variables only (no EIV), and the latter is better than a model without EIV and auxiliary variables. In addition, if ρ increases, then most of the standard deviations increase accordingly.

In summary, our simulation results support the incorporation of EIV and auxiliary variables in models when data are contaminated by EIV in auxiliary variables.

Table 6.1 Auxiliary Variable vs. No Auxiliary Variable

Type	*Adj. R^2*	SSE	θ_0	std	θ_1	std	θ_2	std	θ_3	std	ϕ	std
$\rho=0$												
No Aux.	0.9966	6.61	0.098	0.305	3.423	0.803	0.404	0.571	1.085	0.118		
Aux. Only	0.9989	1.89	0.998	0.319	1.089	0.833	1.027	0.576	1.007	0.118	1.008	0.098
EIV in Aux.	0.9989	1.89	0.998	0.319	1.089	0.833	1.027	0.576	1.007	0.118	1.008	0.098
$\rho=0.6$												
No Aux.	0.9935	12.06	-0.885	0.397	7.795	1.132	-3.339	0.853	1.884	0.174		
Aux. Only	0.9983	2.78	1.001	0.400	1.024	1.193	1.097	0.880	0.998	0.181	1.003	0.056
EIV in Aux.	0.9983	2.78	1.001	0.400	1.024	1.193	1.097	0.880	0.998	0.181	1.003	0.056

Note:

1. σ_η^2 =1, σ_e^2 =0, $Z=\left[z_1, z_2, \ldots, z_{15}\right]'$ are random samples from $N(1,1)$ with 3000 iterations using data simulated from a given parameter vector $\theta=(1\ \ 1\ \ 1\ \ 1)$, $\phi=1$ and $W=\left[w_1, w_2, \ldots, w_{15}\right]'=\left[0.2,\ 0.4,\ \ldots,\ 3.0\right]'$.
2. Since σ_e^2 =0 (which means no EIV in auxiliary variables) so the estimates for the rows in aux. only and EIV in aux. are identical.

Table 6.2 EIV in Auxiliary Variable vs. Auxiliary Variable and No Auxiliary Variable

Type	*Adj. R^2*	SSE	θ_0	std	θ_1	std	θ_2	std	θ_3	std	ϕ	std
$\rho=0.1$												
No Aux.	0.9998	18.10	0.743	0.030	3.515	0.356	0.121	0.130	1.088	0.014		
Aux. Only	1.0000	1.84	1.012	0.262	1.025	0.376	1.009	0.135	1.000	0.014	1.000	0.052
EIV in Aux.	1.0000	1.81	1.015	0.262	1.003	0.377	1.017	0.135	1.000	0.014	1.009	0.052
$\rho=0.6$												
No Aux.	0.9999	13.56	2.017	0.361	3.119	0.525	0.113	0.196	1.089	0.021		
Aux. Only	1.0000	3.08	1.035	0.363	1.033	0.578	1.014	0.222	1.001	0.023	1.000	0.077
EIV in Aux.	1.0000	3.05	1.023	0.363	1.008	0.578	1.025	0.222	1.000	0.023	1.011	0.078

Note:

1. σ_η^2 =1, σ_e^2 =0.01, $Z=\left[z_1, z_2, \ldots, z_{20}\right]'$ are random samples from $N(1,1)$ with 3000 iterations using data simulated from a given parameter vector $\theta=(1\ \ 1\ \ 1\ \ 1)$, $\phi=1$ and $W=\left[w_1, w_2, \ldots, w_{20}\right]'=\left[0.3,\ 0.6,\ \ldots,\ 6\right]'$.
2. We observed that as ρ increases, most of the standard deviations also increase. In addition, in terms of goodness of fit statistics (a*dj. R^2* and SSE), EIV in aux. outperforms the other two options (smallest SSE and almost unbiased estimates as well). These results are supportive of incorporating both EIV and auxiliary information in models.

6.2 Introduction

In fitting statistical models for the analysis of growth data, there are many different types of curves proposed in the literature. In our research, we select some important and frequently used growth curves and use them in our model analysis. In addition, we consider auxiliary covariates with AR(1) structure in our model. Furthermore, the EIV generalization for auxiliary variables of growth curves with AR(1) structural models renders a general framework for the practical application of growth. A real-life illustrative example is presented to demonstrate how a Bayesian approach through MCMC and a Bayesian Information Criterion (BIC) for model selections can be utilized in the analysis of our models.

Growth curves have been extensively studied and used in a wide range of disciplines such as biology, engineering, economics, crop science, fishery research, etc. For example, within food microbiology, scientists use growth curves to describe the behavior of microorganisms under different physical and chemical conditions. Using growth curves, they can predict microbial safety levels or shelf life of food products, detect the critical parts or weak links of the production and distribution process, as well as optimize food distribution chains (Zwietering et al, 1990).

Growth curves arise from repeated observations on a number of individuals in an orderly fashion, usually over time. Growth curve models are more general than repeated measurement alone in that the former models the functional relationship between the responses (or dependent variables) and the regressors (or explanatory variables, such as

time, capital, income, etc.) These functional relationships can be approximated and represented either by linear or nonlinear curves, and the coefficients therein are of interest in growth curve models.

Growth curve models were first introduced by Potthoff and Roy (1964), and later analyzed by Khatri (1966), Grizzle and Allen (1969) and many others. However, Not much literature can be found discussing growth curves under Bayesian formulation. C. R. Rao (1987) discussed a Bayesian and empirical Bayesian methods in the prediction of future observations in growth curve models. He described how to derive a Bayesian prediction probability. He also developed some growth models and studied their usefulness in prediction. Some of the growth curves he mentioned include the Jenss, Winsor, Wright and Bock, and Thissen growth curves. These curves are used to provide an individuals growth trend over a long time period. Kshirsagar and Smith (1995) gave some details about the pioneering work by Geisser (1970, 1980) and Lee and Geisser (1972) discussed details regarding Bayesian growth curves.

Barry (1995) used a Gaussian process growth function for Bayesian model analysis. Robinson and Crowder (2000) presented Bayesian methods for a growth-curve degradation model with repeated measures. The application of this kind of analysis is mainly for the reliability of manufactured products. Zhang et al. (2007) used growth curve models to analyze longitudinal data under Bayesian methodology. Their statistical inference on posteriors is based on point estimation and credible intervals. They also

discussed the application of latent (unobserved) variables in growth curves, using WinBugs for their simulation.

Regarding errors-in-variables (EIV), these arise in many experimental problems, such as in biology (Dellaportas, 1995), in engineering (Jitjareonchai, 2006) and in economics (Edgerton and Jochumzen). The use of errors-in-variables (EIV) models has been shown to be indispensable and requisite. If we let

$\boldsymbol{y}=\{y_j, j=1,\ldots,n\}$ a set of n measured (observed) values in the y-coordinate,

$\boldsymbol{x}=\{x_j, j=1,\ldots n\}$ a set of n measured (observed) values in the x-coordinate,

$\boldsymbol{w}=\{w_j, j=1,\ldots,n\}$ a set of n true (actual or unobserved) x – coordinate.

Then a simple definition of EIV is as follows:

Classical Model:

$y_j=f(w_j)+\varepsilon_j, j=1,\ldots,n$, $x_j=w_j+e_j, j=1,\ldots,n$,

Berkson-type Model:

$y_j=f(w_j)+\varepsilon_j, j=1,\ldots,n$, $w_j=x_j+e_j, j=1,\ldots,n$,

where ε_j and e_j are random errors, $f(w_j)$ is the growth curve. $\theta=(\theta_1,\ldots,\theta_p)$ is the parameters in $f(w_j)$, where p is the number of parameters in θ. The distinction between the two types of EIV models is ignorable under a Bayesian formulation except purely for reasons of model specification (Dellaportas and Stephens, 1995).

When data, which are contaminated with some measurement error, are used to estimate the parameters of any mathematical model by methods such as least squares, it is more the rule than the exception that at least some of the contaminated data (quantities)

can be treated as independent variables, contrary to the basic requirement that independent variables be perfectly known. While useful results can often be obtained by neglecting errors in some of the variables, it is true that some degree of approximation is still present, usually unknown. EIV is used in eliminating this approximation.

Much statistical literature has been devoted to the classical analysis of EIV models. Maximum likelihood methods and least-squares solutions are widely used and discussed in solving the parameter estimation problems involving EIV (see for example, Madansky, 1959, Solari, 1969, Britt and Leucke, 1973, Fuller, 1987, Schafer ,1987, Whittemore and Keler, 1988, Caroll, 1989, Carroll, Gail and Lubin, 1993). Solari (1969) pointed out that the maximum likelihood solution for linear EIV problems is a saddle point (a point of a function or surface which is a stationary point but not an extremum) instead of a simple maximum in the case of unknown error variances. Since a saddle point is not a true maximum, the maximum likelihood solutions in this case may be misleading. Schafer (1987) further indicated that maximum likelihood or least-squares estimation is far from straightforward, especially in the non-linear case, where numerical stability problems might arise.

A Bayesian approach was first proposed by Lindley and El-Sayyad (1968) and was shown to yield more sensible results than maximum likelihood method. They pointed out that the parameter estimation for the usual models, even with normal errors, present difficulties and the results of the maximum likelihood method in EIV problems may be misleading except for exceptional cases where at least the ratio of the errors of

observations in Y and X is known and only then are the parameters estimates reliable.

The Bayesian method for EIV problems was further developed by Reilly and Patino-Leal (1981) and later extended to case where error covariance matrix is unknown (Keeler and Reilly, 1991). They use conventional optimization techniques to obtain point estimates of the parameters and to calculate approximate confidence limits. Dellaportas and Stephens (1995) use a Bayesian approach and Gibbs sampler technique to solve examples of nonlinear curve EIV problems. Nummi (2000) proposed a growth curve model in which the observed time intervals (the regressor variable) are subject to measurement errors. Jitjareonchai, Reilly, Duever and Chambers (2006) further pointed out that these traditional techniques consist of complicated matrix manipulation and sometimes lead to convergence problems. So, they also proposed and implemented Gibbs Sampler techniques and succeeded to solve EIV problems with much ease and obtain results with satisfactory degree of accuracy.

In our study, we incorporate auxiliary information with errors-in-variables (EIV) and an autoregressive covariance structure, AR(1), into our models and use a Bayesian approach in our analysis, as this methodology is more effective in dealing with models of larger complexity. In spite of a large number of studies in this area, models similar to the ones in our study were not found in the existing literature.

6.3 Model Formulation

We adopt the same notations as in previous chapter and assume a growth model for a single experimental unit of the form, $y_j = f(w_j, \theta) + q(z_j, \phi) + \varepsilon_j$, $j = 1, \ldots, n$, n is the total number of observations, $x_j = z_j + e_j$, θ is a vector of parameters; w_j is the independent variable for the j^{th} observation; q , with a vector of coefficients ϕ , denotes the function of a vector of auxiliary variable $\boldsymbol{z} = \{z_j, j = 1, \ldots, n\}$; $x = \{x_j, j = 1, \ldots, n\}$ is a vector of "observed" (not true) auxiliary variables and therefore there exists an assumed vector of random errors $\boldsymbol{e} = \left\{e_j \overset{iid}{\sim} N(0, \sigma_e^2), j = 1, \ldots, n\right\}$; $\varepsilon_j \sim N(0, \sigma^2)$ is independent random error of y_j , here $f(w_j, \theta)$ is one of the different types of growth curves. We assume that the covariance for the j^{th} and k^{th} observation is $\Sigma = \sigma^2 \Omega$, where

$$\Omega = \begin{bmatrix} 1 & \rho & \cdots & \rho^{n-1} \\ \rho & 1 & \cdots & \rho^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \cdots & 1 \end{bmatrix}_{n \times n} , \quad \rho \text{ is the autocorrelation coefficient.}$$

To simplify the model, we further assume, as is often done in existing literature on EIV, that $\boldsymbol{y}$ given $\boldsymbol{w}$ and $\boldsymbol{z}$ is independent of $\boldsymbol{x}$. Let $f_l(w_j, \theta_l), l = 1, 2, 3, 4$ be four different types of growth curves. We use the same four commonly used growth curves for our model:

1. Jenss growth curve $f_1(w_j,\theta_1)=a_1+b_1 w_j-\exp(c_1+d_1 w_j)$, a vector of parameters $\theta_1'=(a_1\ \ b_1\ \ c_1\ \ d_1)$, where $a_1,b_1>0$, c_1 and $d_1\in\mathbb{R}$.
2. Gompertz growth curve $f_2(w_j,\theta_2)=a_2\cdot\exp[-\exp(b_2-c_2 w_j)]$ and a vector of parameters $\theta_2'=(a_1\ \ b_2\ \ c_2)$, where a_2, $c_2>0$, $b_2\in\mathbb{R}$.
3. Richards growth curve $f_3(w_j,\theta_3)=a_3\cdot\left\{1+b_3\exp[c_3(d_3-w_j)]\right\}^{-1/b_3}$, and a vector of parameters $\theta_3'=(a_3\ \ b_3\ \ c_3\ \ d_3)$, where $a_3,b_3>0$, $c_3,d_3\in\mathbb{R}$.
4. Polynomial growth curve $f_4(w_j,\theta_4)=a_4+b_4 w_j+c_4 w_j^2+d_4 w_j^3$, and a vector of parameters $\theta_4'=(a_4\ \ b_4\ \ c_4\ \ d_4)$, where where a_4 , b_4 , c_4 and $d_4\in\mathbb{R}$.

Now let this $f_l(w_j,\theta_l)$, $l=1,2,3,4$ be substituted into the growth model. We denote $\boldsymbol{y}=\{y_j,\ \ j=1,\ldots,n\}$, $\boldsymbol{w}=\{w_j,\ \ j=1,\ldots,n\}$, $\boldsymbol{z}=\{z_j,j=1,\ldots,n\}$, $B_l=(\boldsymbol{x}-\boldsymbol{z})_{1\times n}(\boldsymbol{x}-\boldsymbol{z})'_{n\times 1}$ and

$A_l=\left[\boldsymbol{y}-f_l(\boldsymbol{w},\theta_l)-q(\boldsymbol{z},\phi_l)\right]'_{1\times n}\Omega^{-1}_{n\times n}\left[\boldsymbol{y}-f_l(\boldsymbol{w},\theta_l)-q(\boldsymbol{z},\phi_l)\right]_{n\times 1}$ for $l=1,2,3,4$.

Then we get their likelihood functions, respectively, as follows:

$$g(\boldsymbol{y}|\theta_l,\phi_l,\sigma^2,\sigma_e^2,\rho,\boldsymbol{w},\boldsymbol{x},\boldsymbol{z})=\frac{1}{(2\pi)^n(\sigma_e^2)^{n/2}(\sigma^2)^{n/2}|\Omega|}\exp\left\{-\frac{A_l}{2\sigma^2}-\frac{B_l}{2\sigma_e^2}\right\}. \qquad (6.1)$$

Now we propose the following distributions for the four growth curve priors (note that empirical Bayes estimate vectors $\tilde{\theta}_i, i=1,\ldots,4$ in the following four curves can be obtained from nonlinear least square regression results by using MATLAB *nlinfit* function to fit nonlinear Jenss, Gompertz and Richards curves and *polyfit* function to fit

polynomial curves. For details, please refer to MATLAB Help and Seber and Wild, 1989):

1. Jenss curve with auxiliary variable:

$$a_1 \sim Expon(\frac{1}{\tilde{a}_1}) \quad , \quad b_1 \sim Expon(\frac{1}{\tilde{b}_1}) \quad , \quad \begin{pmatrix} c_1 \\ d_1 \\ \phi_1 \end{pmatrix} \sim N_4 \left(\begin{pmatrix} \tilde{c}_1 \\ \tilde{d}_1 \\ \tilde{\phi}_1 \end{pmatrix}, \tilde{\Sigma}_1 \right) \quad \text{where}$$

$(c_1, d_1, \phi_1)'$ are multivariate normal distributed with mean vector $(\tilde{c}_1, \tilde{d}_1, \tilde{\phi}_1)'$, and variance covariance matrix is $\tilde{\Sigma}_1$. The vector $\tilde{\theta}_1 = (\tilde{a}_1, \tilde{b}_1, \tilde{c}_1, \tilde{d}_1, \tilde{\phi}_1, \tilde{\Sigma}_1)$. So the prior distribution can be written as follows:

$$h(\theta_1|\tilde{\theta}_1) \propto \frac{1}{\tilde{a}_1}\frac{1}{\tilde{b}_1}\frac{1}{|\tilde{\Sigma}_1|^{1/2}} \exp\left\{ \frac{-1}{2} \begin{pmatrix} c_1-\tilde{c}_1 \\ d_1-\tilde{d}_1 \\ \phi_1-\tilde{\phi}_1 \end{pmatrix}' \tilde{\Sigma}_1^{-1} \begin{pmatrix} c_1-\tilde{c}_1 \\ d_1-\tilde{d}_1 \\ \phi_1-\tilde{\phi}_1 \end{pmatrix} - \frac{a_1}{\tilde{a}_1} - \frac{b_1}{\tilde{b}_1} \right\} \quad . \qquad (6.2)$$

2. Gompertz growth curve with auxiliary variable :

$$a_2 \sim Expon(\frac{1}{\tilde{a}_2}) \quad , \quad c_2 \sim Expon(\frac{1}{\tilde{c}_2}) \quad , \quad b_2 \sim N(\tilde{b}_2, \tilde{\sigma}_b^2) \quad , \quad \phi_2 \sim N(\tilde{\phi}_2, \tilde{\sigma}_\phi) \quad \text{then}$$

$\tilde{\theta}_2 = (\tilde{a}_2, \tilde{b}_2, \tilde{c}_2, \tilde{\phi}_2, \tilde{\sigma}_b^2, \tilde{\sigma}_\phi^2)$. So the prior distribution is:

$$h(\theta_2|\tilde{\theta}_2) \propto \frac{1}{\tilde{a}_2\tilde{c}_2}\frac{1}{\tilde{\sigma}_b\tilde{\sigma}_\phi} \exp\left\{ -\frac{1}{2}\left[\frac{(b_2-\tilde{b}_2)^2}{\tilde{\sigma}_b^2} + \frac{(\phi_2-\tilde{\phi}_2)^2}{\tilde{\sigma}_\phi^2} \right] - \left[\frac{a_2}{\tilde{a}_2} + \frac{c_2}{\tilde{c}_2} \right] \right\} . \qquad (6.3)$$

3. Richards curve with auxiliary variable:

$$a_3 \sim Expon(\frac{1}{\tilde{a}_3}) \quad , \quad b_3 \sim Expon(\frac{1}{\tilde{b}_3}) \quad , \quad \begin{pmatrix} c_3 \\ d_3 \\ \tilde{\phi}_3 \end{pmatrix} \sim N_3 \left(\begin{pmatrix} \tilde{c}_3 \\ \tilde{d}_3 \\ \tilde{\phi}_3 \end{pmatrix}, \tilde{\Sigma}_3 \right) \quad , \text{where}$$

$(c_3, d_3, \phi_3)'$ are multivariate normally distributed with mean vector $(\tilde{c}_3, \tilde{d}_3, \tilde{\phi}_3)'$, and variance covariance matrix $\tilde{\Sigma}_3$. The vector $\tilde{\theta}_3=(\tilde{a}_3, \tilde{b}_3, \tilde{c}_3, \tilde{d}_3, \tilde{\phi}_3, \tilde{\Sigma}_3)$. So the prior distribution is:

$$h(\theta_3|\tilde{\theta}_3) \propto \frac{1}{\tilde{a}_3 \tilde{b}_3} \frac{1}{|\tilde{\Sigma}_3|^{1/2}} \exp\left\{ -\frac{1}{2} \begin{pmatrix} c_3-\tilde{c}_3 \\ d_3-\tilde{d}_3 \\ \phi_3-\tilde{\phi}_3 \end{pmatrix}' \tilde{\Sigma}_3^{-1} \begin{pmatrix} c_3-\tilde{c}_3 \\ d_3-\tilde{d}_3 \\ \phi_3-\tilde{\phi}_3 \end{pmatrix} - \left[\frac{a_3}{\tilde{a}_3} + \frac{b_3}{\tilde{b}_3} \right] \right\}. \qquad (6.4)$$

4. Polynomial curve with auxiliary variable:

$$\begin{pmatrix} a_4 \\ b_4 \\ c_4 \\ d_4 \\ \phi_4 \end{pmatrix} \sim N_5 \left(\begin{pmatrix} \tilde{a}_4 \\ \tilde{b}_4 \\ \tilde{c}_4 \\ \tilde{d}_4 \\ \tilde{\phi}_4 \end{pmatrix}, \tilde{\Sigma}_4 \right)$$, where $(a_4, b_4, c_4, c_4, \phi_4)'$ are multivariate normally distributed with mean vector $(\tilde{a}_4, \tilde{b}_4, \tilde{c}_4, \tilde{d}_4, \tilde{\phi}_4)'$, variance and covariance matrix is $\tilde{\Sigma}_4$. The vector $\tilde{\theta}_4=(\tilde{a}_4, \tilde{b}_4, \tilde{c}_4, \tilde{d}_4, \tilde{\phi}_4, \tilde{\Sigma}_4)$. So the prior distribution is:

$$h(\theta_4|\tilde{\theta}_4) \propto \frac{1}{|\tilde{\Sigma}_4|^{1/2}} \exp\left\{ -\frac{1}{2} \begin{pmatrix} a_4-\tilde{a}_4 \\ b_4-\tilde{b}_4 \\ c_4-\tilde{c}_4 \\ d_4-\tilde{d}_4 \\ \phi_4-\tilde{\phi}_4 \end{pmatrix}' \tilde{\Sigma}_4^{-1} \begin{pmatrix} a_4-\tilde{a}_4 \\ b_4-\tilde{b}_4 \\ c_4-\tilde{c}_4 \\ d_4-\tilde{d}_4 \\ \phi_4-\tilde{\phi}_4 \end{pmatrix} \right\}. \qquad (6.5)$$

We assume Jeffrey's prior for σ^2 and σ_e^2 and thus have density $p(\sigma^2) \propto \frac{1}{\sigma^2}$,

$p(\sigma_e^2) \propto \frac{1}{\sigma_e^2}$. But for the prior for ρ , we assume that $\rho^* = \frac{\rho+1}{2} \sim Beta(\tilde{\alpha}, \tilde{\beta})$, $0 \leqslant \rho^* \leqslant 1$, where $\tilde{\alpha}$ and $\tilde{\beta}$ are some known constants for the beta prior distribution. After transformation we get the prior distribution for ρ , $p(\rho) \propto (\rho+1)^{\tilde{\alpha}-1}(1-\rho)^{\tilde{\beta}-1}$. We further assume that $z_j \overset{iid}{\sim} N(\bar{x}, \tilde{\sigma}_e^2)$, $j = 1, \ldots, n$, where $\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$ and $\tilde{\sigma}_z^2 = var(\boldsymbol{x})$. One more assumption is to let function $q(\boldsymbol{z}, \phi) = \phi z_j$, $j = 1, \ldots, n$ be linear and the prior for the coefficient ϕ be $\phi \sim N(0, \tilde{\sigma}_\phi^2)$ (we can estimate $\tilde{\sigma}_\phi^2$ using the results of the least square method for ϕ). So the posterior distributions are as follows:

1. Jenss growth curve:

$$\pi_1(\theta_1, \sigma^2, \sigma_e^2, \rho_1, \phi_1, \boldsymbol{z} \mid \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{x}) \propto \frac{(\rho_1+1)^{\tilde{\alpha}-1}(1-\rho_1)^{\tilde{\beta}-1}}{\sigma^{n+2}\sigma_e^{n+2}|\Omega|} \exp\left\{ -\frac{A_1}{2\sigma^2} - \frac{B_1}{2\sigma_e^2} - \frac{1}{2}\begin{pmatrix} c_1-\tilde{c}_1 \\ d_1-\tilde{d}_1 \\ \phi_1-\tilde{\phi}_1 \end{pmatrix}' \tilde{\Sigma}_1^{-1} \begin{pmatrix} c_1-\tilde{c}_1 \\ d_1-\tilde{d}_1 \\ \phi_1-\tilde{\phi}_1 \end{pmatrix} - \left[\frac{a_1}{\tilde{a}_1} + \frac{b_1}{\tilde{b}_1} + \frac{1}{2\tilde{\sigma}_e^2}\sum_{j=1}^{n}\left(z_j - \bar{x}\right)^2 \right] \right\} .$$

Integrating out σ^2 and σ_e^2 , respectively, we have

$$\pi_1(\theta_1, \rho_1, \phi_1, \boldsymbol{z} \mid \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{x}) \propto \frac{(\rho_1+1)^{\tilde{\alpha}-1}(1-\rho_1)^{\tilde{\beta}-1}}{|\Omega|} \exp\left\{ -\left[\frac{a_1}{\tilde{a}_1} + \frac{b_1}{\tilde{b}_1} + \frac{1}{2\tilde{\sigma}_e^2}\sum_{j=1}^{n}\left(z_j - \bar{x}\right)^2 \right] - \frac{1}{2}\begin{pmatrix} c_1-\tilde{c}_1 \\ d_1-\tilde{d}_1 \\ \phi_1-\tilde{\phi}_1 \end{pmatrix}' \tilde{\Sigma}_1^{-1} \begin{pmatrix} c_1-\tilde{c}_1 \\ d_1-\tilde{d}_1 \\ \phi_1-\tilde{\phi}_1 \end{pmatrix} \right\} \cdot A_1^{-\frac{n}{2}} \cdot B_1^{-\frac{n}{2}} . \quad (6.6)$$

2. Gompertz growth curve:

$$\pi_2(\theta_2,\sigma^2,\sigma_e^2,\rho_2,\phi_2,\boldsymbol{z}\ |\boldsymbol{y},\boldsymbol{w},\boldsymbol{x})\propto\frac{(\rho_2+1)^{\tilde{\alpha}-1}(1-\rho_2)^{\tilde{\beta}-1}}{\sigma^{n+2}\sigma_e^{n+2}|\Omega|}\exp\left\{-\frac{A_2}{2\sigma^2}-\frac{B_2}{2\sigma_e^2}\right.$$
$$\left.-\frac{1}{2}\left[\frac{(\phi_2-\tilde{\phi}_2)^2}{\tilde{\sigma}_\phi^2}+\frac{(b_2-\tilde{b}_2)^2}{\tilde{\sigma}_b^2}+\frac{\sum_{j=1}^{n}\left(z_j-\bar{x}\right)^2}{\tilde{\sigma}_e^2}\right]-\left[\frac{a_2}{\tilde{a}_2}+\frac{c_2}{\tilde{c}_2}\right]\right\}.$$

Integrating out σ^2 and σ_e^2, respectively, we have

$$\pi_2(\theta_2,\rho_2,\phi_2,\boldsymbol{z}\ |\boldsymbol{y},\boldsymbol{w},\boldsymbol{x})\propto\frac{(\rho_2+1)^{\tilde{\alpha}-1}(1-\rho_2)^{\tilde{\beta}-1}}{|\Omega|}$$
$$\exp\left\{-\frac{1}{2}\left[\frac{(\phi_2-\tilde{\phi}_2)^2}{\tilde{\sigma}_\phi^2}+\frac{(b_2-\tilde{b}_2)^2}{\tilde{\sigma}_b^2}+\frac{\sum_{j=1}^{n}\left(z_j-\bar{x}\right)^2}{\tilde{\sigma}_e^2}\right]-\left[\frac{a_2}{\tilde{a}_2}+\frac{c_2}{\tilde{c}_2}\right]\right\}\cdot A_2^{-\frac{n}{2}}\cdot B_2^{-\frac{n}{2}}. \quad (6.7)$$

3. Richards growth curve:

$$\pi_3(\theta_3,\sigma^2,\sigma_e^2,\rho_3,\phi_3,\boldsymbol{z}\ |\boldsymbol{y},\boldsymbol{w},\boldsymbol{x})\propto\frac{(\rho_3+1)^{\tilde{\alpha}-1}(1-\rho_3)^{\tilde{\beta}-1}}{\sigma^{n+2}\sigma_e^{n+2}|\Omega|}$$
$$\exp\left\{-\frac{A_3}{2\sigma^2}-\frac{B_3}{2\sigma_e^2}-\frac{\sum_{j=1}^{n}\left(z_j-\bar{x}\right)^2}{2\tilde{\sigma}_e^2}-\frac{1}{2}\begin{pmatrix}c_3-\tilde{c}_3\\ d_3-\tilde{d}_3\\ \phi_3-\tilde{\phi}_3\end{pmatrix}'\tilde{\Sigma}_3^{-1}\begin{pmatrix}c_3-\tilde{c}_3\\ d_3-\tilde{d}_3\\ \phi_3-\tilde{\phi}_3\end{pmatrix}-\left[\frac{a_3}{\tilde{a}_3}+\frac{b_3}{\tilde{b}_3}\right]\right\}.$$

Integrating out σ^2 and σ_e^2, respectively, we have

$$\pi_3(\theta_3,\rho_3,\phi_3\ |\boldsymbol{y},\boldsymbol{w},\boldsymbol{z})\propto\frac{(\rho_3+1)^{\tilde{\alpha}-1}(1-\rho_3)^{\tilde{\beta}-1}}{|\Omega|}$$
$$\exp\left\{\frac{-\sum_{j=1}^{n}\left(z_j-\bar{x}\right)^2}{2\tilde{\sigma}_e^2}-\frac{1}{2}\begin{pmatrix}c_3-\tilde{c}_3\\ d_3-\tilde{d}_3\\ \phi_3-\tilde{\phi}_3\end{pmatrix}'\tilde{\Sigma}_3^{-1}\begin{pmatrix}c_3-\tilde{c}_3\\ d_3-\tilde{d}_3\\ \phi_3-\tilde{\phi}_3\end{pmatrix}-\left[\frac{a_3}{\tilde{a}_3}+\frac{b_3}{\tilde{b}_3}\right]\right\}\cdot A_3^{-\frac{n}{2}}\cdot B_3^{-\frac{n}{2}}. \quad (6.8)$$

4. Polynomial growth curve:

$$\pi_4(\theta_4, \sigma^2, \sigma_e^2, \rho_4, \phi_4, \boldsymbol{z} \mid \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{x}) \propto \frac{(\rho_4+1)^{\tilde{\alpha}-1}(1-\rho_4)^{\tilde{\beta}-1}}{\sigma^{n+2}\sigma_e^{n+2}|\Omega|}$$
$$\exp\left\{-\frac{A_4}{2\sigma^2}-\frac{B_4}{2\sigma_e^2}-\frac{\sum_{j=1}^{n}(z_j-\bar{x})^2}{2\tilde{\sigma}_e^2}-\frac{1}{2}\begin{pmatrix} a_4-\tilde{a}_4 \\ b_4-\tilde{b}_4 \\ c_4-\tilde{c}_4 \\ d_4-\tilde{d}_4 \\ \phi_4-\tilde{\phi}_4 \end{pmatrix}' \tilde{\Sigma}_4^{-1} \begin{pmatrix} a_4-\tilde{a}_4 \\ b_4-\tilde{b}_4 \\ c_4-\tilde{c}_4 \\ d_4-\tilde{d}_4 \\ \phi_4-\tilde{\phi}_4 \end{pmatrix}\right\}.$$

Integrating out σ^2 and σ_e^2, respectively, we have

$$\pi_4(\theta_4, \rho_4, \phi_4, \boldsymbol{z} \mid \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{x}) \propto \frac{(\rho_4+1)^{\tilde{\alpha}-1}(1-\rho_4)^{\tilde{\beta}-1}}{|\Omega|} \cdot$$
$$\exp\left\{-\frac{\sum_{j=1}^{n}(z_j-\bar{x})^2}{2\tilde{\sigma}_e^2}-\frac{1}{2}\begin{pmatrix} a_4-\tilde{a}_4 \\ b_4-\tilde{b}_4 \\ c_4-\tilde{c}_4 \\ d_4-\tilde{d}_4 \\ \phi_4-\tilde{\phi}_4 \end{pmatrix}' \tilde{\Sigma}_4^{-1} \begin{pmatrix} a_4-\tilde{a}_4 \\ b_4-\tilde{b}_4 \\ c_4-\tilde{c}_4 \\ d_4-\tilde{d}_4 \\ \phi_4-\tilde{\phi}_4 \end{pmatrix}\right\} \cdot A_4^{-\frac{n}{2}} \cdot B_4^{-\frac{n}{2}}. \quad (6.9)$$

We now take Gompertz growth curve to illustrate how we can apply this model. In order to use MH algorithm, we first need to get the full conditionals for the parameters from (6.7). These are presented in the following equations (6.10)-(6.15):

$$\pi_2(a_2 \mid \cdot) \propto \exp\left\{-\frac{a_2}{\tilde{a}_2}\right\} A_2^{-\frac{n}{2}}, \quad (6.10)$$

$$\pi_2(b_2 \mid \cdot) \propto \exp\left\{-\frac{1}{2}\frac{(b_2-\tilde{b}_2)^2}{\tilde{\sigma}_b^2}\right\} A_2^{-\frac{n}{2}}, \quad (6.11)$$

$$\pi_2(c_2 \mid \cdot) \propto \exp\left\{-\frac{c_2}{\tilde{c}_2}\right\} A_2^{-\frac{n}{2}}, \quad (6.12)$$

$$\pi_2(\rho_2 \mid \cdot) \propto \frac{(\rho_4+1)^{\tilde{\alpha}-1}(1-\rho_4)^{\tilde{\beta}-1}}{|\Omega|} A_2^{-\frac{n}{2}} \quad , \tag{6.13}$$

$$\pi_2(\phi_2 \mid \cdot) \propto \exp\left\{-\frac{1}{2}\frac{(\phi_2-\tilde{\phi}_2)^2}{\tilde{\sigma}_\phi^2}\right\} A_2^{-\frac{n}{2}} \quad , \tag{6.14}$$

$$\pi_2(z \mid \cdot) \propto \exp\left\{\frac{-1}{2\tilde{\sigma}_z^2}\sum_{j=1}^{n}(z_j-\bar{x})^2\right\} A_2^{-\frac{n}{2}} B_2^{-\frac{n}{2}} \quad . \tag{6.15}$$

MH Algorithm:

Let's take the sampling of a_2 in Gompertz curve as an example. To define the algorithm, let $\varphi(a_2^{(old)}, a_2^{(new)})$ denote a source density for a candidate draw $a_2^{(new)}$ given the current value $a_2^{(old)}$ in the sampled sequence. The density $\varphi(a_2^{(old)}, a_2^{(new)})$ is referred to as the proposal or candidate generating density. Then, the MH algorithm is defined by two steps: a first step in which a proposal value is drawn from the candidate generating density and a second step in which the proposal value is accepted as the next iterate in the Markov chain according to the probability:

$$\alpha\left(a_2^{(old)}, a_2^{(new)}\right) = min\left\{\frac{\pi_2(a_2^{(new)})\varphi\left(a_2^{(new)}, a_2^{(old)}\right)}{\pi_2(a_2^{(old)})\varphi\left(a_2^{(old)}, a_2^{(new)}\right)}, 1\right\} \quad ,$$

if $\pi_2(a_2^{(old)})\varphi\left(a_2^{(old)}, a_2^{(new)}\right) > 0$ (otherwise $\alpha\left(a_2^{(old)}, a_2^{(new)}\right) = 1$).

If the proposal value is rejected, then the next sampled value is taken to be the current value. Let's follow this recursive procedure:

1. Specify an initial value $a_2^{(0)}$.
2. Repeat for $j=1,2,\ldots,M$:

   a) Propose $a_2^{(new)} \sim \varphi(a_2^{(j)},\cdot)$, and

   b) Let $a_2^{(j+1)}=a_2^{(new)}$ if $U(0,1)\leqslant\alpha(a_2^{(j)},a_2^{(new)})$ otherwise $a_2^{(j+1)}=a_2^{(j)}$.
3. Return the values $a_2^{(1)},a_2^{(2)},\ldots,a_2^{(M)}$.

We take samples of θ_2 , ρ_2 , ϕ_2 and z by using the full conditionals from (6.10) to (6.15) and MH algorithm within Gibbs Samplers through the following steps:

1. Set $j=0$ and select a set of starting parameter values for $\rho_2^{(0)}$, $\phi_2^{(0)}$, $\theta_2^{(0)}=(a_2^{(0)},b_2^{(0)},c_2^{(0)})$ and $z^{(0)}$.
2. Sample $a_2^{(j+1)}$ from $\pi(a^{(j+1)}|\cdot)$.
3. Sample $b_2^{(j+1)}$ from $\pi_2(b_2^{(j+1)}|\cdot)$.
4. Sample $c_2^{(j+1)}$ from $\pi_2(c_2^{(j+1)}|\cdot)$.
5. Sample $\phi_2^{(j+1)}$ from $\pi_2(\phi_2^{(j+1)}|\cdot)$.
6. Sample $z^{(j+1)}$ from $\pi_2(z^{(j+1)}|\cdot)$.
7. Replace $\theta_2^{(j)}\Leftarrow\theta_2^{(j+1)}$, $\rho_2^{(j)}\Leftarrow\rho_2^{(j+1)}$, $\phi_2^{(j)}\Leftarrow\phi_2^{(j+1)}$ and $z^{(j)}\Leftarrow z^{(j+1)}$.
8. Set $j=j+1$ and repeat steps 2 through 8.
9. Stop when $j=N+B$, (B is burn-in samples to be dropped and N is sample size).

In the long run, the N samples taken from these full conditionals will form a sample drawn approximately from the above posterior distribution. In addition to the

above estimation of parameters, we also generate highest density regions for the estimated parameters. Further, 90% confidence limits for the best-fit growth curve can be obtained by using the 5% and 95% percentiles of $\boldsymbol{y}$ (calculated by substituting the N samples of θ_2, given $\boldsymbol{w}$ and $\boldsymbol{z}$).

6.4 An Illustrative Example

We use similar simulation data (see Table 6.3) to demonstrate our Bayesian analysis EIV in auxiliary variable with AR(1) in covariance structure model. The advantages of incorporating EIV and auxiliary variable into models have already been demonstrated. So in this section, we further explore the goodness of fit statistics as criteria in determining a proper growth curve model.

We can use Bayesian analysis, in conjunction with least squares method, to obtain estimates of the parameters of interest and then use goodness of fit statistics to select the best model. The results of the Bayesian estimates for the parameters are shown in Table 6.3. From the goodness of fit statistics therein we see that although Cubic growth curve appears to have the best fit model (SSE almost equals to zero and R^2 almost one), we also have Gompertz curve that yields a very high R^2 (adjusted R^2 is almost the same) value but employs a smaller number of parameters. This may explain why (in Table 6.4) Gompertz has higher BIC score than the other three curves (even though the simulated data is actually based on an original cubic polynomial growth curve).

In summary, we have demonstrated that EIV in auxiliary variables is important

when data is subject to autocorrelation. We proposed a Bayesian growth curve model through goodness of fit statistics. We determined that we would choose a Cubic model according to the goodness of fit statistics. This conclusion is consistent with our expectation, as the data was simulated using a cubic polynomial curve. Note however that the Gompertz curve is also competitive as it has the advantage of getting almost the same goodness of fit statistics using one less parameter in the curve.

Table 6.3 Bayesian Analysis of EIV in Auxiliary Variable with Autocorrelation

Type	R^2	*Adj.* R^2	SSE	θ_0	std	θ_1	std	θ_2	std	θ_3	std	ϕ	std	ρ	std
$\rho=0.4$															
Jenss	0.92	0.90	6.10	2.86	0.013	4.520	1.881	0.737	0.018	-0.002	0.014	0.909	0.016	0.372	0.064
Gompertz	0.99	0.99	0.93	5409.06	0.009	2.092	0.012	0.151	0.012	NA	NA	1.004	0.016	0.374	0.062
Richards	0.68	0.60	24.93	64.48	0.292	5.595	0.648	7.111	0.043	2.986	0.014	0.983	0.011	0.364	0.065
Cubic	1.00	1.00	0.00	2.06	0.044	1.007	0.053	1.002	0.039	1.007	0.040	1.004	0.039	0.379	0.072

σ_η^2 =10, σ_e^2 =1, Z are $m=15$ random samples from $N(1,1)$, 1000 iterations using data simulated from a given parameter vector $\theta=(1\ \ 1\ \ 1\ \ 1)$, $\phi=1$ and $W=[0.1,\ 0.2,\ ...,\ 1.5]$. Note that the large estimate value for θ_0 of Gompertz curve should not be a surprise because there are two exponential functions in the curve.

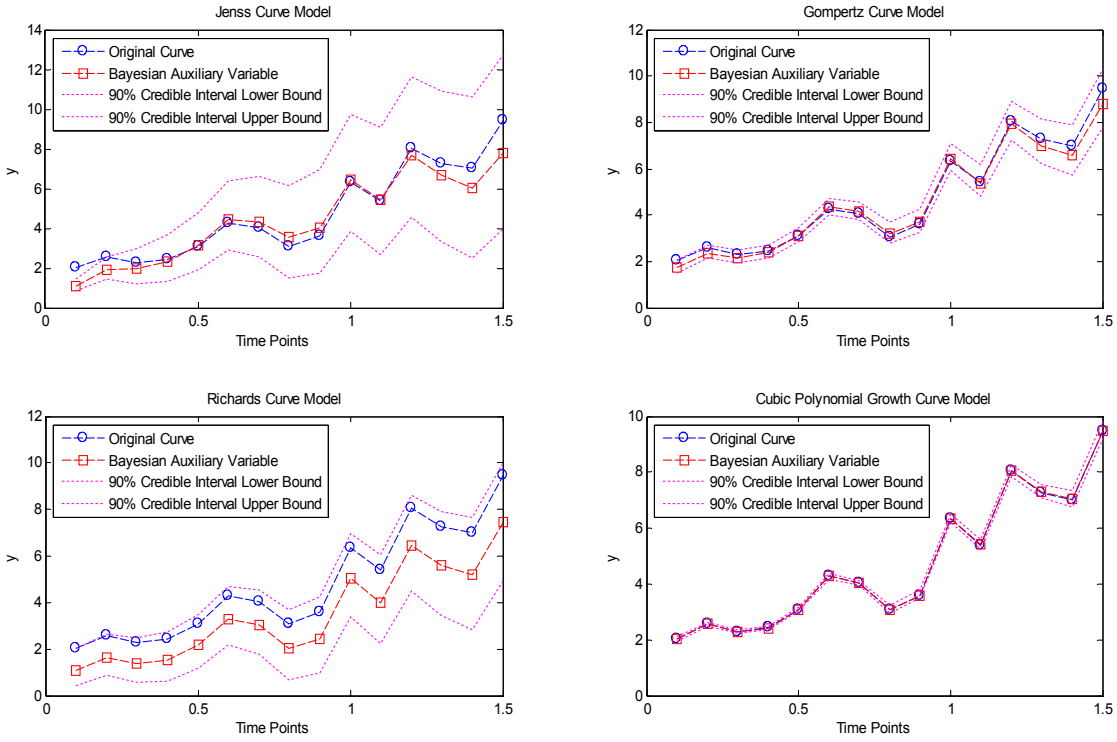
Table 6.4 BIC for Different Growth Curves

Jenss Growth Curve	6.73
Gompertz Growth Curve	6.53*
Richards Growth Curve	6.74
Cubic Polynomial Growth Curve	6.81

* Gompertz has the smallest BIC value

In the following Figure 6.1 we see the comparison of the four growth curve models and their 90% credible intervals. We can see that the figure confirms our conclusion in determining a model with the best fit (the best one is Cubic and then Gompertz).

Figure 6.1 Credible Intervals for the Four Different Curves



6.5 Conclusion

We have presented some very practical but complex growth curve models under Bayesian methodology and have shown that they are implementable and useful in terms of obtaining better goodness of fit statistics in different scenarios (such scenarios include various combinations of correlation coefficients, with or without auxiliary variables and using different auxiliary variable functions contaminated by EIV) versus models ignoring some or all of those components. This is the main thrust of our contribution, as such models are currently non-existent in the literature.

References

Agresti, Alan (2002). *Categorical Data Analysis*, 2nd edition. Wiley, New York.

Barry, Daniel (1995). A Bayesian Model for Growth Curve Analysis. *Biometrics*, **51**, pp. 639-655.

Berkey, C. S. (1982). Bayesian approach for a nonlinear growth model. *Biometrics*, **38**, pp.953-961.

Berkey, C. S. (1982). Comparison of two longitudinal growth models for preschool children. *Biometrics*, **38** (Mar., 1982), pp. 221-234.

Berkey, C. S. and R. B. Reed (1987). A Model for describing normal and abnormal growth in early childhood. *Human Biology*, **59**, p. 973.

Britt, H. I. and R. H. Leucke (1973). The estimation of parameters in nonlinear, implicit models. *Technometrics,* **15**, pp. 233-247.

Cameron, A.C. and P.K. Trivedi (1998). *Regression Analysis of Count Data*. Cambridge University Press, Cambridge.

Carroll, R. J., M. H. Gail and J. H. Lubin (1993). Case-control studies with errors in covariates. *Journal of the American Statistical Association,* **88**, pp. 421, 185-199.

Carroll, R.J., D. Ruppert and L.A. Stefanski (1995). *Measurement Error in Nonlinear Models*. Chapman & Hall, London.

Diggle, P. J. (1988). An Approach to the Analysis of Repeated Measurements. *Biometrics,* **44**, pp. 959-971.

Dellaportas, P. and D. A. Stephens (1995). Bayesian analysis of errors-in-variables regression models. *Biometrics,* **51**, pp. 1085-1095.

Edgerton, D. and P. Jochumzen (2003). *Estimation in models with measurement errors and limited dependent variables*. Department of Economics, Lund University.

Fuller, W. A. (1987). *Measurement error models*. Wiley, New York.

Geisser, S. (1970). Bayesian analysis of growth curves. *Sankhya A*, **32**, pp. 53-64.

Geisser, S. (1980). Growth curve analysis. In *Handbook of Statistics*, **1** ,Krishnaiah, P. R. (E.d.), pp. 89-115. North Holland, Amsterdam.

Grizzle, J. E. and Allen, D. M. (1969). Analysis of growth and dose response curves. *Biometrics*, **25**, pp. 357-381.

Hyndman, R.J. (1996). Computing and graphing highest density regions. *American Statistician*, **50**, pp. 120-126.

Jenss, R.M. and N. Barley (1937). A mathematical method for studying the growth of a child. *Human Biology,* **9**, pp. 556-563.

Jitjareonchai, J. J., P. M. Reilly, T. A. Duever and D. B. Chambers (2006). Parameter estimation in the error-in-variables models using the Gibbs Sampler. *The Canadian Journal of Chemical Engineering*, **84**, pp.125-138.

Keeler, S. E. and P. M. Reilly (1991). The error-invariables model applied to parameter estimation when the error covariance matrix is unknown. *The Canadian Journal of Chemical Engineering*, **69**, pp. 27-34.

Khatri, C. A. (1966). A note on a MANOVA model applied to problems in growth curves. *Annals of the Institute of Statistical Mathematics*, **18**, pp. 75-86.

Kshirsagar, A. M., W. B. Smith (1995). *Growth Curves*. Marcel Dekker.

Kukush, A., H. Schneeweiss and R. Wolf (2004). Three estimators for the Poisson regression model with measurement errors. *Statistical Papers,* 45351–368.

Lee, J. C. and Geisser S. (1972). Growth curve prediction. *Sankhya A*, **34**, pp. 393-412.

Lindley, D. V. and G. M. El-Sayyad (1968). The Bayesian estimation of a linear functional relationship. *Journal of Royal Statistical Society B*, **30**, pp. 190-202.

Ma, Y. and M. G. Genton (2004). Flexible class of skew-symmetric distributions. *Scandinavian Journal of Statistics*, **31**, pp. 459-468.

Madansky, A. (1959). The fitting of straight lines when both variables are subject to error. *Journal of the American Statistical Association,* **54**, pp. 173-205.

Manwani, A. H. and K. N. Agarwal (1973). The growth pattern of Indian infants during the first year of life. *Human Biology*, **45** , p.341.

Nummi, T. (2000). Analysis of growth curves under measurement errors. *Journal of Applied Statistics*, **27**, pp. 235-243.

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

Racine-Poon, A. , C. Withhs and A. F. M. Smith (1991). Estimation of relative Potency with sequential dilution errors in radioimmunoassay. *Biometrics*, **47**, pp. 1235-1246.

Rao, C. R. (1959). Some problems involving linear hypothesis in multivariate analysis. *Biometrika*, **46**, pp. 49-58.

Rao, C. R. (1987). Prediction of future observations in growth curve models. *Statistical Science*, **2**, pp. 434-447.

Reilly, Pun. M. and H. Patino-Leal (1981). A Bayesian study of the error-in variables model. *Technometrics,* **23**, pp. 221-232.

Robinson, M. E. and M. J. Crowder (2000). Bayesian methods for a growth-curve degradation model with repeated measures. *Lifetime Data Analysis*, **6**, pp. 357-374.

Rubin, D. B. (1980). Using empirical Bayes techniques in the law school validity studies. *Journal of the American Statistical Association,* **75**, pp. 801-816.

Schafer, D. W. (1987). Covariate measurement error in generalized linear models. *Biometrica,* **74**, pp. 385-391.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, pp. 461-464.

Seber, G. A. F. and C. J. Wild (1989). *Nonlinear Regression*. John Wiley & Sons.

Solari, M. E. (1969). The maximum likelihood solution of the problem of estimating a linear functional relationship. *Journal of Royal Statistical Society B*, **31**, pp. 372-375.

Spiegelhalter, D. J., N. G. Best, B.P. Carlin, and A. van der Linde (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of Royal Statistical Society B*, **64**, pp. 583-639.

Stuart, H. C. and R. B. Reed (1959). Longitudinal studies of child health and development. Harvard School of Public Health, Series II, No. 1. *Description of Project, Pediatrics,* **23**, pp. 875-885.

Tiwari, C.T., K.A. Cronin, W. Davis, E. J. Feuer, B. Yu, and S. Chib (2005). Bayesian model selection for join regression with application to age-adjusted cancer rates. *Applied Statistics*, **54**, pp. 919-939.

Vonesh, E. F. and V. M. Chinchilli (1996). *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. Marcel Dekker.

Whittemore, A. S. and J. B. Keller (1988). Approximation for regression with covariate measurement error. *Journal of the American Statistical Association,* **83**, pp. 404, 1057-1066.

Zhang, Z. et al. (2007). Bayesian analysis of longitudinal data using growth curve models. *International Journal of Behavioral Development*, **31**, pp. 374-383.

Zweitering, M. H., I. Jongenburger, F. M. Rombouts and K. Van't Riet (1990). Modeling of bacterial growth curve. *Applied and Environmental Microbiology*, pp. 1875-1881.

Chapter 7 Bayesian Analysis of Multivariate Growth Curve Models

7.0 Abstract

Growth curve models have been widely studied and extensively applied in many areas because they are useful in situations when time (an important factor) is involved. Researchers have considered growth curves (mainly linear) in conjunction with different covariance structures for numerous applications. In this paper, our interest is to use some commonly used nonlinear growth curves to describe (in terms of time) each variable in a multivariate dataset in the presence of random error covariance structures with autocorrelation dependence. No similar attempts have been found in the literature because under this complex scenario, the models become too complicated for classical analysis without making a lot of compromising assumptions.

In our paper, we show that under a Bayesian formulation, by judicious choice of priors, one can obtain the full conditionals and this allows one to conveniently implement the Metropolis Hastings algorithm to sample/generate observations from the conditional (posterior) observations. This makes Bayesian approach a simpler but useful alternative to classical analysis. We use intrauterine growth retardation in rats data as an illustrative example for our model.

7.1 Introduction

The motivation for conducting multivariate analysis in our research is that we are able to study the effect of several variables acting simultaneously. This gives a closer resemblance to our intuition as well as better understanding about the relationship between the variables. When more variables are analyzed simultaneously, greater statistical power will be obtained and we gain easier visualization and interpretation of the data through graphical measures, such as scatter plots or higher dimensional plots (e.g. 3D plots). So our focus is also spontaneously shifted from individual or isolated factors to the relationships among several variables of interest in a data set.

Growth curve models, which are useful especially for studying growth behavior of short time series in economics, biology, medical research and epidemiological problems (Grizzle and Allen, 1969, Lee and Geisser 1972), have a long history. Their initiation may be attributed to Potthoff and Roy (1964), who introduced their formulation and then studied the growth curve problems. Then subsequently, Rao (1965), Khatri (1966), Geisser (1970) and von Rosen (1989, 1990, 1991) became the primary researchers in analyzing the growth curve models. However, it took nearly a decade before the Bayesian approach (including predictive problem from a Bayesian perspective) was applied to the analysis of growth curve models and different assumptions about covariance matrices were also made accordingly. Lindley and Smith (1972) and Geisser (1980) assumed that covariance matrices were known, Fearn (1975) assumed that they were identity matrices with unknown variances. Barry (1995) gave a different treatment of the problem under Bayesian approach but also assumed identity matrix for covariances.

In our research, we approach similar general multivariate growth problems by assuming that the multivariate dependent variables (such as weight, height, etc.) can be described by some commonly used nonlinear growth curves in terms of the independent variable (time) with a certain correlation (dependence) relationship in the covariance matrix. So the multivariate growth curve models we propose will include nonlinear growth curves with autocorrelated errors in their covariance structures. The classical analysis for these types of realistic models becomes either too complicated to obtain

analytical solutions or may require a lot of simplifying assumptions, thus becoming unrealistic. Bayesian analysis, including experts' opinions, can help us computationally to get to the estimates of the parameters for growth curve models and thus become more appealing, as well as important to researchers. To the best of our knowledge, we have found no similar models which consider such complex scenarios (multivariate nonlinear growth curves with autocorrelation in covariance matrix) available in the literature.

We will detail the model formulation and then use the Gompertz growth curve as an example in first two sections. Then in the last two sections, we will demonstrate the applications of the models using a bivariate growth data set and discuss the simulation results as our conclusion.

7.2 Model Formulation

We consider a single subject of n observations. $\underset{(p\times 1)}{Y_j}$, for $j=1,\ldots,n$, is a vector of p multivariate correlated dependent variables. We define our model as,

$\boldsymbol{Y}=M+E$, where $E\sim N_p(0,\Omega)$, Ω is a $p\times p$ variance covariance matrix,

$$\underset{(p\times n)}{\boldsymbol{Y}}=\begin{pmatrix} y_1' \\ \vdots \\ y_p' \end{pmatrix}_{(p\times n)}=\left(\underset{(p\times 1)}{Y_1},\ldots,\underset{(p\times 1)}{Y_n}\right), \text{ where } y_k=\begin{pmatrix} y_{1k} \\ \vdots \\ y_{nk} \end{pmatrix} \text{ for } k=1,\ldots,p,$$

and

$$\underset{(p\times n)}{M}=\begin{pmatrix} \mu_1' \\ \vdots \\ \mu_p' \end{pmatrix}_{(p\times n)}=\left(\underset{(p\times 1)}{M_1},\ldots,\underset{(p\times 1)}{M_n}\right), \text{ where } \mu_k=\begin{pmatrix} \mu_{1k} \\ \vdots \\ \mu_{nk} \end{pmatrix} \text{ for } k=1,\ldots,p.$$

$\boldsymbol{W}=(w_1,\ldots,w_n)$ is a vector of independent variables (i.e., time) and $\underset{q\times p}{\Theta}=(\underset{q\times 1}{\theta_1},\ldots,\underset{q\times 1}{\theta_p})$, where θ_k, $k=1,..,p$ is a vector of coefficients (parameters) for growth curves and q is the number of coefficients for the specific growth curve in that model (e.g., $q=3$ in a Gompertz curve). Let $f(W|\theta_k)$, $k=1,\ldots,p$ be the growth curve. So for $k=1,\ldots,p$, $\underset{n\times 1}{\mu_k}=\underset{n\times 1}{f(W|\theta_k)}$, and $\underset{p\times n}{M}=\left[f(W|\Theta)_{n\times p}\right]'$.

Our model considers a covariance structure between weight and length in that, under normal conditions, the lengthier the subject grows, the weightier it becomes and vice versa. We assume that $\boldsymbol{Y}$ follows a $p\times n$ matrix normal distribution, which is actually a special case of the pn-variate multivariate normal distribution when the covariate matrix is separable. If we denote a pn-variate normal distribution with pn-dimensional mean μ and $pn\times pn$ covariance matrix Ω, then the p.d.f. function is as follows:

$$g(y|\mu,\Omega)=(2\pi)^{-\frac{np}{2}}|\Omega|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(y-\mu)'\Omega^{-1}(y-\mu)\right\}, \qquad (7.1)$$

where $\underset{(pn\times 1)}{y}=vect(\boldsymbol{Y}')=(y_1',\ldots,y_p')'$, $\mu=vect(M')=(\mu_1',\ldots,\mu_p')'$, in which the operator $vect(\cdot)$ stacks the columns of its matrix argument from left to right in a single vector. The separable matrix $\Omega=\Sigma\otimes\Phi$, where $\otimes$ is the Kronecker product which multiplies every entry of its first matrix argument by its entire second matrix argument, can be written as:

$$\Sigma \otimes \Phi = \begin{pmatrix} \sigma_{11}\Phi \cdots \sigma_{1p}\Phi \\ \vdots \quad\quad \vdots \\ \sigma_{p1}\Phi \cdots \sigma_{pp}\Phi \end{pmatrix} \quad .$$

Also we know that $\Omega^{-1} = (\Sigma \otimes \Phi)^{-1} = \Sigma^{-1} \otimes \Phi^{-1}$ and $|\Sigma \otimes \Phi|^{-\frac{1}{2}} = |\Sigma|^{-\frac{n}{2}} |\Phi|^{-\frac{p}{2}}$. Then we have

$$g(y|\mu, \Sigma, \Phi) = (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{n}{2}} |\Phi|^{-\frac{p}{2}} \exp\left\{ -\frac{1}{2}(y-\mu)'(\Sigma \otimes \Phi)^{-1}(y-\mu) \right\} \quad .$$

Note that also with the matrix identity, we have

$$(y-\mu)'_{1\times np} (\Sigma \otimes \Phi)^{-1}_{np\times np} (y-\mu)_{np\times 1} \quad = \quad tr\, \Sigma^{-1}_{p\times p} (\boldsymbol{Y}-M)_{p\times n} \Phi^{-1}_{n\times n} (\boldsymbol{Y}-M)'_{n\times p} \quad ,$$

$$g(\boldsymbol{Y}|M, \Sigma, \Phi) = (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{n}{2}} |\Phi|^{-\frac{p}{2}} \exp\left\{ -\frac{1}{2} tr\, \Sigma^{-1}_{p\times p} (\boldsymbol{Y}-M)_{p\times n} \Phi^{-1}_{n\times n} (\boldsymbol{Y}-M)'_{n\times p} \right\} \quad . \tag{7.2}$$

So $\boldsymbol{Y}$ is a random variable that follows a $p \times n$ matrix normal distribution and can be denoted as:

$$\boldsymbol{Y}|(M, \Sigma, \Phi) \sim N_{p\times n}(M, \Sigma \otimes \Phi) \quad ,$$

where (M, Σ, Φ) parameterize the above distribution with $\boldsymbol{Y} \in \mathbb{R}^{p\times n}$, $M \in \mathbb{R}^{p\times n}$ and $\Sigma, \Phi > 0$ (Σ and Φ are commonly referred to as the within and between covariance matrices). Recall that M is a function of Θ and assume that Φ is a function of correlation coefficient ρ and that, for simplicity, Θ , Σ and Φ are independent and adopt vague prior distributions for (Θ, Φ, Σ) . Then we have $h(\Theta, \Phi, \Sigma) = h(\Theta)h(\Phi)h(\Sigma)$ and because Φ is a function of ρ , their prior distribution assumptions are as follows: $h(\Theta) \propto constant$, $\rho \propto (1+\rho)^{\tilde{\alpha}-1}(1-\rho)^{\tilde{\beta}-1}$ for $-1 < \rho < 1$ (i.e., $(1+\rho)/2 \sim Beta(\tilde{\alpha}, \tilde{\beta})$, where $\tilde{\alpha}$ and $\tilde{\beta}$ can be chosen such that the mean $\tilde{\alpha}/(\tilde{\alpha}+\tilde{\beta})$ is

consistent with the empirical value for ρ) and $h(\Sigma)\propto\frac{1}{|\Sigma|^{(p+1)/2}}$.

So the prior distribution is $h(\Theta,\rho,\Sigma)\propto\frac{(1+\rho)^{\tilde{\alpha}-1}(1-\rho)^{\tilde{\beta}-1}}{|\Sigma|^{(p+1)/2}}$, and the joint posterior distribution of the parameters follows:

$$\pi(\Sigma,\Phi(\rho),\Theta|\boldsymbol{W},\boldsymbol{Y})=(2\pi)^{-\frac{np}{2}}|\Sigma|^{-\frac{n+p+1}{2}}|\Phi|^{-\frac{p}{2}}(1+\rho)^{\tilde{\alpha}-1}(1-\rho)^{\tilde{\beta}-1}\cdot \exp\left\{-\frac{1}{2}tr\,\Sigma^{-1}{}_{p\times p}(\boldsymbol{Y}-M)_{p\times n}\Phi^{-1}_{n\times n}(\boldsymbol{Y}-M)'_{n\times p}\right\} \quad . \tag{7.3}$$

Let $G=(\boldsymbol{Y}-M)\Phi^{-1}(\boldsymbol{Y}-M)'$ then (7.3) becomes

$$\pi(\Sigma,\Phi(\rho),\Theta|\boldsymbol{W},\boldsymbol{Y})\propto\left[|\Sigma|^{-\frac{n+p+1}{2}}\exp\left\{-\frac{1}{2}tr\,\Sigma^{-1}G\right\}\right]\cdot|\Phi|^{-\frac{p}{2}}(1+\rho)^{\tilde{\alpha}-1}(1-\rho)^{\tilde{\beta}-1} \quad .$$

This can be reduced to the joint distribution of Φ and Θ if we integrate out Σ ,

$$\pi(\Phi(\rho),\Theta|\boldsymbol{W},\boldsymbol{Y})\propto\left[\int_{\Sigma>0}|\Sigma|^{-\frac{n+p+1}{2}}\exp\left\{-\frac{1}{2}tr\,\Sigma^{-1}G\right\}d\,\Sigma\right]\cdot|\Phi|^{-\frac{p}{2}}(1+\rho)^{\tilde{\alpha}-1}(1-\rho)^{\tilde{\beta}-1} \quad ,$$

The integration can be easily worked out by recognizing that if Σ^{-1} follows a Wishart distribution (we follow J. Press, 2002, see sections 12.3.5-6 therein for the justification of making the assumptions of vague prior for Σ and Wishart distribution for Σ^{-1} in posterior distribution to simplify the integration) then it can be written as:

$$\pi(\Theta,\Phi(\rho)|W,\boldsymbol{Y})\propto\frac{|\Phi|^{-\frac{p}{2}}(1+\rho)^{\tilde{\alpha}-1}(1-\rho)^{\tilde{\beta}-1}}{|G|^{n/2}} \quad . \tag{7.4}$$

Assume an autocorrelation matrix for Φ with correlation coefficient ρ as follows:

$$\Phi=\begin{pmatrix} 1 & \rho & \rho^2 & \cdots \rho^{n-1} \\ \rho & 1 & \rho & \cdots \rho^{n-2} \\ \vdots & \vdots & \vdots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots \ 1 \end{pmatrix} \quad , \tag{7.5}$$

where ρ is the correlation coefficient. Then we can substitute the results that $|\Phi|=(1-\rho^2)^{n-1}$, into (7.4) and get the posterior function

$$\pi(\Theta,\Phi(\rho)|W,\boldsymbol{Y})\propto \frac{(1+\rho)^{(\tilde{\alpha}-1)-p(n-1)/2}(1-\rho)^{(\tilde{\beta}-1)-p(n-1)/2}}{|G|^{n/2}} \quad . \tag{7.6}$$

7.3 Gompertz Curve as An Illustrative Example

If we take the Gompertz curve as an illustrative example in fitting a bivariate data set which has weight and length as the variables and the following priors for

$\Theta=\begin{pmatrix}\theta_1 & \theta_2\end{pmatrix}=\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{pmatrix}$,where $a_1\sim Expon(\frac{1}{\tilde{a}_1})$, $a_2\sim Expon(\frac{1}{\tilde{a}_2})$,

$\begin{Bmatrix} b_1 \\ b_2 \end{Bmatrix}\sim N_2\left(\begin{Bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{Bmatrix},\tilde{\Sigma}_b\right)$, $c_1\sim Expon(\frac{1}{\tilde{c}_1})$ and $c_2\sim Expon(\frac{1}{\tilde{c}_2})$.

Let $\tilde{\Theta}=(\tilde{a}_1,\tilde{a}_2,\tilde{b}_1,\tilde{b}_2,\tilde{c}_1,\tilde{c}_2,\tilde{\Sigma}_b)$ (these can be estimated through empirical Bayes estimates from nonlinear least square regression method and results can be easily obtained using MATLAB *nlinfit* function to fit nonlinear Jenss, Gompertz and Richards curves and *polyfit* function to fit polynomial curves. For details, please refer to MATLAB Help and Seber and Wild, 1989). Depending on the data, although sometimes we could

get estimates of $\tilde{\Sigma}_b$, most of the time we have to assume them to be equal to some proper value for our Bayesian analysis. So the prior distributions are:

$$h(\Theta|\tilde{\Theta}) \propto \frac{1}{\tilde{a}_1\tilde{a}_2\tilde{c}_1\tilde{c}_2\left|\tilde{\Sigma}_b\right|^{1/2}} \exp\left\{-\frac{1}{2}\begin{pmatrix} b_1-\tilde{b}_1 \\ b_2-\tilde{b}_2 \end{pmatrix}' \tilde{\Sigma}_b^{-1} \begin{pmatrix} b_1-\tilde{b}_1 \\ b_2-\tilde{b}_2 \end{pmatrix} - \left(\frac{a_1}{\tilde{a}_1}+\frac{a_2}{\tilde{a}_2}+\frac{c_1}{\tilde{c}_1}+\frac{c_2}{\tilde{c}_2}\right)\right\} . \quad (7.7)$$

Let (7.7) be substituted into (7.6), then it becomes

$$\pi(\Theta,\Phi(\rho)|W,\boldsymbol{Y}) \propto \frac{(1+\rho)^{(\tilde{\alpha}-1)-p(n-1)/2}(1-\rho)^{(\tilde{\beta}-1)-p(n-1)/2}}{\left|(\boldsymbol{Y}-M)\Phi^{-1}(\boldsymbol{Y}-M)'\right|^{n/2}} \cdot \exp\left\{-\left(\frac{a_1}{\tilde{a}_1}+\frac{c_1}{\tilde{c}_1}+\frac{a_2}{\tilde{a}_2}+\frac{c_2}{\tilde{c}_2}\right)-\frac{1}{2}\begin{pmatrix} b_1-\tilde{b}_1 \\ b_2-\tilde{b}_2 \end{pmatrix}' \tilde{\Sigma}_b^{-1} \begin{pmatrix} b_1-\tilde{b}_1 \\ b_2-\tilde{b}_2 \end{pmatrix}\right\} . \quad (7.8)$$

Then we get the full conditionals of the parameters as follows:

$$\pi(\rho|\cdot) \propto \frac{(1+\rho)^{(\tilde{\alpha}-1)-p(n-1)/2}(1-\rho)^{(\tilde{\beta}-1)-p(n-1)/2}}{\left|(\boldsymbol{Y}-M)\Phi^{-1}(\boldsymbol{Y}-M)'\right|^{n/2}} , \quad (7.9)$$

$$\pi(a_1|\cdot) \propto \frac{\exp\left\{-a_1/\tilde{a}_1\right\}}{\left|(\boldsymbol{Y}-M)\Phi^{-1}(\boldsymbol{Y}-M)'\right|^{n/2}} , \quad (7.10)$$

$$\pi(a_2|\cdot) \propto \frac{\exp\left\{-a_2/\tilde{a}_2\right\}}{\left|(\boldsymbol{Y}-M)\Phi^{-1}(\boldsymbol{Y}-M)'\right|^{n/2}} , \quad (7.11)$$

$$\pi(b_1|\cdot) \propto \frac{1}{\left|(\boldsymbol{Y}-M)\Phi^{-1}(\boldsymbol{Y}-M)'\right|^{n/2}} \exp\left\{-\frac{1}{2}\begin{pmatrix} b_1-\tilde{b}_1 \\ b_2-\tilde{b}_2 \end{pmatrix}' \tilde{\Sigma}_b^{-1} \begin{pmatrix} b_1-\tilde{b}_1 \\ b_2-\tilde{b}_2 \end{pmatrix}\right\} , \quad (7.12)$$

$$\pi(b_2|\cdot) \propto \frac{1}{\left|(\boldsymbol{Y}-M)\Phi^{-1}(\boldsymbol{Y}-M)'\right|^{n/2}} \exp\left\{-\frac{1}{2}\begin{pmatrix} b_1-\tilde{b}_1 \\ b_2-\tilde{b}_2 \end{pmatrix}' \tilde{\Sigma}_b^{-1} \begin{pmatrix} b_1-\tilde{b}_1 \\ b_2-\tilde{b}_2 \end{pmatrix}\right\} , \quad (7.13)$$

$$\pi(c_1|\cdot) \propto \frac{\exp\left\{-c_1/\tilde{c}_1\right\}}{\left|(\boldsymbol{Y}-M)\Phi^{-1}(\boldsymbol{Y}-M)'\right|^{n/2}} , \quad (7.14)$$

$$\pi(c_2|\cdot) \propto \frac{\exp\left\{-c_2/\tilde{c}_2\right\}}{\left|(\boldsymbol{Y}-M)\boldsymbol{\Phi}^{-1}(\boldsymbol{Y}-M)'\right|^{n/2}} \quad . \tag{7.15}$$

Regarding the MH Algorithm:

Let's take the sampling of a_2 in Gompertz curve as an example. To define the algorithm, let $\varphi(a_2^{(old)}, a_2^{(new)})$ denote a source density for a candidate draw $a_2^{(new)}$ given the current value $a_2^{(old)}$ in the sampled sequence. The density $\varphi(a_2^{(old)}, a_2^{(new)})$ is referred to as the proposal or candidate generating density. Then, the MH algorithm is defined by two steps: a first step in which a proposal value is drawn from the candidate generating density and a second step in which the proposal value is accepted as the next iterate in the Markov chain according to the probability:

$$\alpha\left(a_2^{(old)}, a_2^{(new)}\right) = min\left\{\frac{\pi_2(a_2^{(new)})\varphi\left(a_2^{(new)}, a_2^{(old)}\right)}{\pi_2(a_2^{(old)})\varphi\left(a_2^{(old)}, a_2^{(new)}\right)}, 1\right\} \quad ,$$

if $\pi_2(a_2^{(old)})\varphi\left(a_2^{(old)}, a_2^{(new)}\right) > 0$ (otherwise $\alpha\left(a_2^{(old)}, a_2^{(new)}\right) = 1$).

If the proposal value is rejected, then the next sampled value is taken to be the current value. Let's follow this recursive procedure:

Metropolis-Hastings Algorithm:

1. Specify an initial value $a_2^{(0)}$.
2. Repeat for $j = 1,2,\ldots,M$:

   a) Propose $a_2^{(new)} \sim \varphi(a_2^{(j)}, \cdot)$, and

b) Let $a_2^{(j+1)} = a_2^{(new)}$ if $U(0,1) \leqslant \alpha(a_2^{(j)}, a_2^{(new)})$ otherwise $a_2^{(j+1)} = a_2^{(j)}$.

3. Return the values $a_2^{(1)}, a_2^{(2)}, \ldots, a_2^{(M)}$.

We follow the above Metropolis-Hastings Algorithm in taking samples of Θ and ρ by using (7.9)-(7.15) through the following steps:

1. Set $j = 0$ and select a set of initial parameter values for $\Theta^{(0)}$, $B^{(0)}$ and $\rho^{(0)}$.
2. Sample $\rho^{(j+1)}$ from (7.9) (using MH algorithm).
3. Sample $\Theta^{(j+1)}$ from (7.10)-(7.15) (using MH algorithm).
4. Replace $\rho^{(j)}$ by $\rho^{(j+1)}$, $\Theta^{(j)}$ by $\Theta^{(j+1)}$ and $B^{(j)}$ by $B^{(j+1)}$.
5. Set $j = j+1$ and repeat steps 2 through 4.

Drop the initial burn-in sets and retain the rest of the data for marginal distribution analysis. This analysis includes highest density regions for the estimated parameters.

In addition to this analysis of parameters, we can also generate 90% Credible Intervals (CIs or HDR's, Highest Density Regions) for the best-fit growth curve under this Bayesian formulation by using the 5% and 95% percentiles of $\boldsymbol{y}$ calculated by substituting the M samples of Θ (a vector of growth curve parameters, see section 7.2 for definition) at a given w_j .

7.4 Example: Using Intrauterine Growth Retardation in Rats Data

We use the intrauterine growth retarded rats data (Oyhenart et al, 2003) in this section as an example to demonstrate how to apply our approach to Bayesian analysis of multivariate growth curve model in a bivariate data setting (weight and length). In their

experiment, Oyhenart et al. chose fifty female rats that were mated overnight with ten adult males and then divided the pregnant female rats into three groups (control group, intrauterine growth control group and sham-operated group). They then measured body weight, body length, and other facial characteristics of the rats that were in those three groups, respectively, every four days for twenty days. For illustrative purposes and for simplifying our analysis, we chose the control group and only use the bivariates body weight and body length in our growth curve model. The dataset is as follows (units for ages: days, weight: gram and length: mm):

Table 7.1 Rats Growth Data

Age (Days)	Weight (g)	Length (mm)
1	6.6	54.5
5	10.4	65.6
9	16.3	77.2
13	23.2	87.5
17	28.6	94.6
21	38.4	110.4

We present below the growth curve models explicitly for this specific example:

1. Jenss growth curve: $f(w)=a+bw-\exp(c+dw)$.
2. Gompertz growth curve: $f(w)=a\exp\left[-\exp(b+c\,w)\right]$.
3. Richards growth curve: $f(w)=a\left\{1+b\exp[c(d-w)]\right\}^{-1/b}$.
4. Polynomial growth curve: $f(w)=a+bw+cw^2+dw^3$.

In this data set, we assume that $\tilde{\Sigma}_b$ is equal to $s^2\begin{bmatrix} 1 & 0.1 \\ 0.1 & 1 \end{bmatrix}$ for Gompertz curve (the same or similar assumption is made for the other growth curves for comparison. Here the constant s^2 could be quite small if prior knowledge is reliable, and the number we assume would allow some moderate correlation relationship between the covariates length and weight).

Bayesian estimates are displayed in Table 7.2. Using BIC, in conjunction with the graphs and CIs, it seems natural to say that the Cubic growth curve is the model of selection (better than other curves) for this specific bivariate intrauterine growth retarded rats data.

Table 7.2 Bayesian Estimates of Parameters and BIC

	Jenss		Gompertz		Richards		Cubic Polynomial	
(they are sample means of posterior distribution through MH within Gibbs Sampler)								
Estimates	length vs. time	weight vs. time	length vs. time	weight vs. time	length vs. time	weight vs. time	length vs. time	weight vs. time
a	52.299	21.536	275.231	146.321	966.688	747.626	50.226	5.325
b	2.674	1.575	0.501	1.167	-0.744	-0.393	3.998	1.051
c	0.115	2.906	0.027	0.041	0.005	0.010	-0.162	0.012
d	-0.861	0.000	-	-	37.458	76.668	0.005	0.001
ρ	-0.336		-0.361		-0.360		-0.369	
BIC	7.73		5.74		5.59		5.57	

Note: Take the numbers in the two columns under Gompertz as examples. They are the estimates of the parameters (coefficients) of the bivariate growth curves (for length and weight, respectively), where

$$length = f(w) = 275.231\exp(-\exp(0.501 + 0.027w))$$
$$weight = f(w) = 146.321\exp(-\exp(1.167 + 0.041w)) \quad .$$

Figure 7.1 Weight vs. Time Credible Intervals for the Four Different Growth Curves

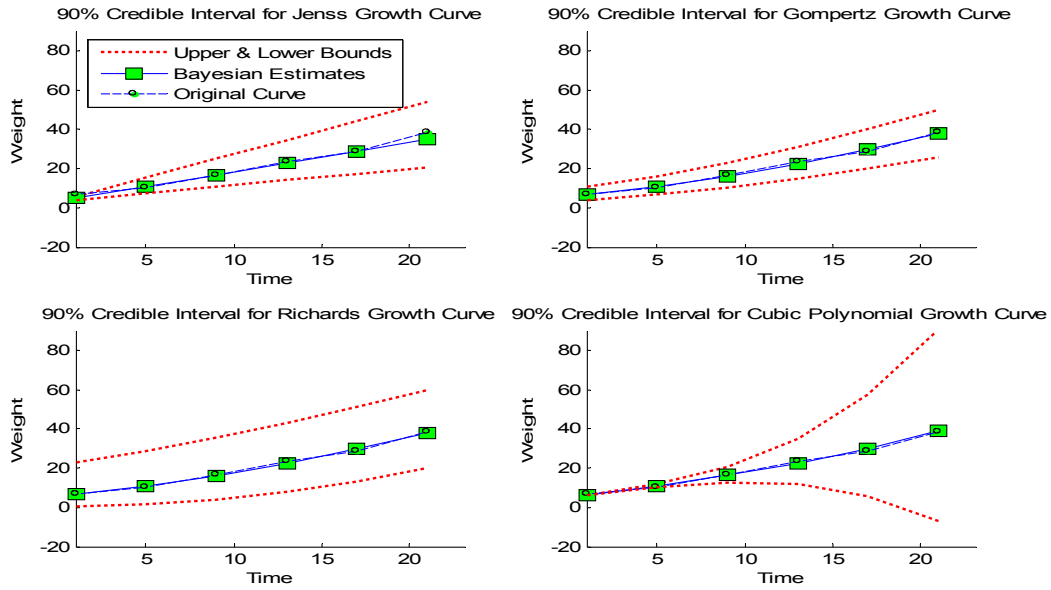


Figure 7.2 Length vs. Time Credible Intervals for the Four Different Growth Curves

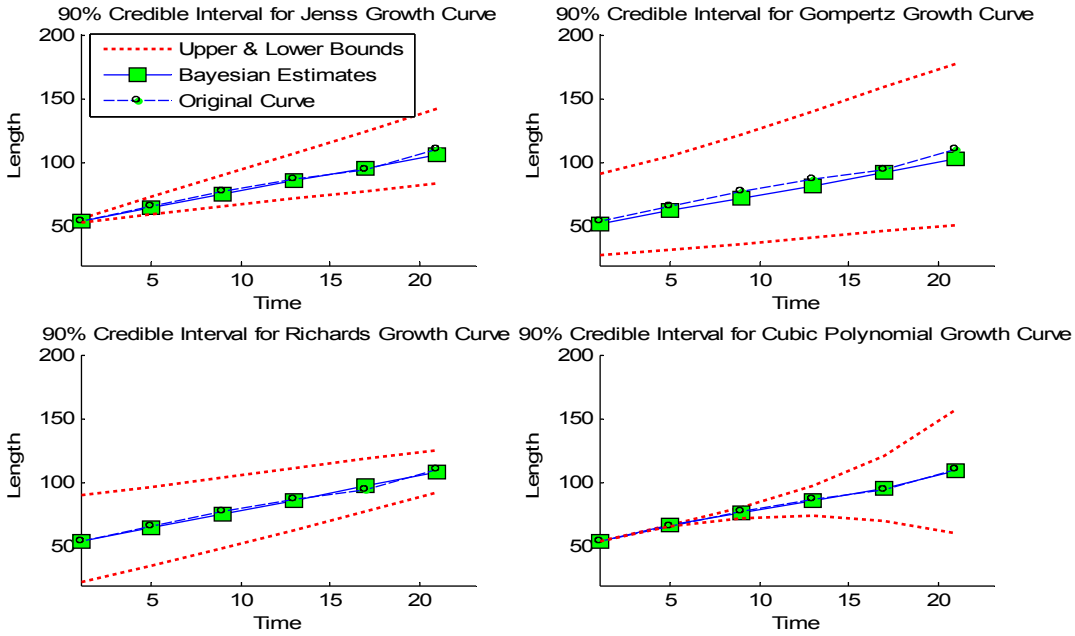


Figure 7.3 Weight vs. Length for the Four Different Growth Curves

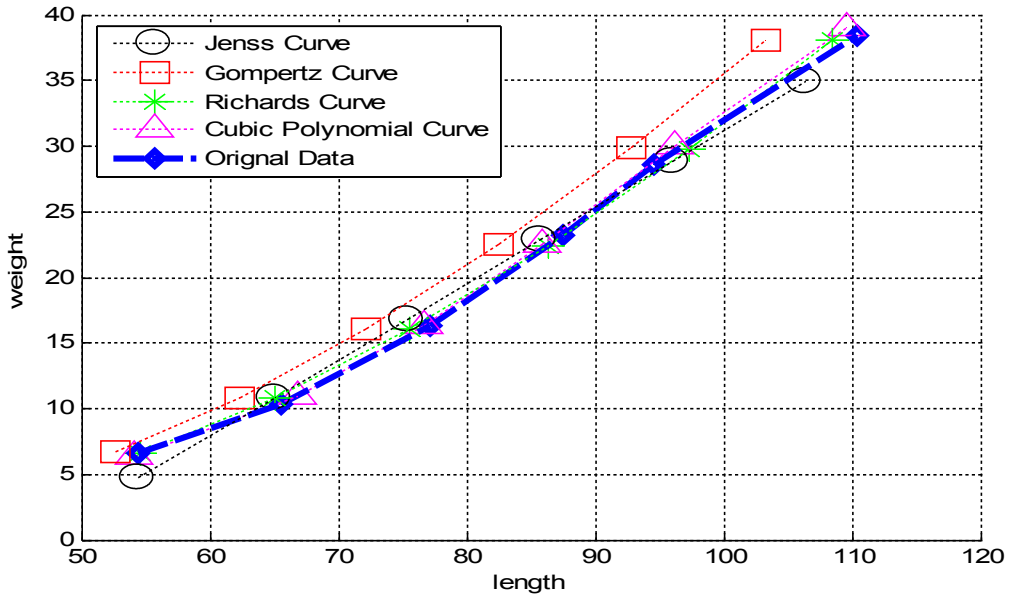
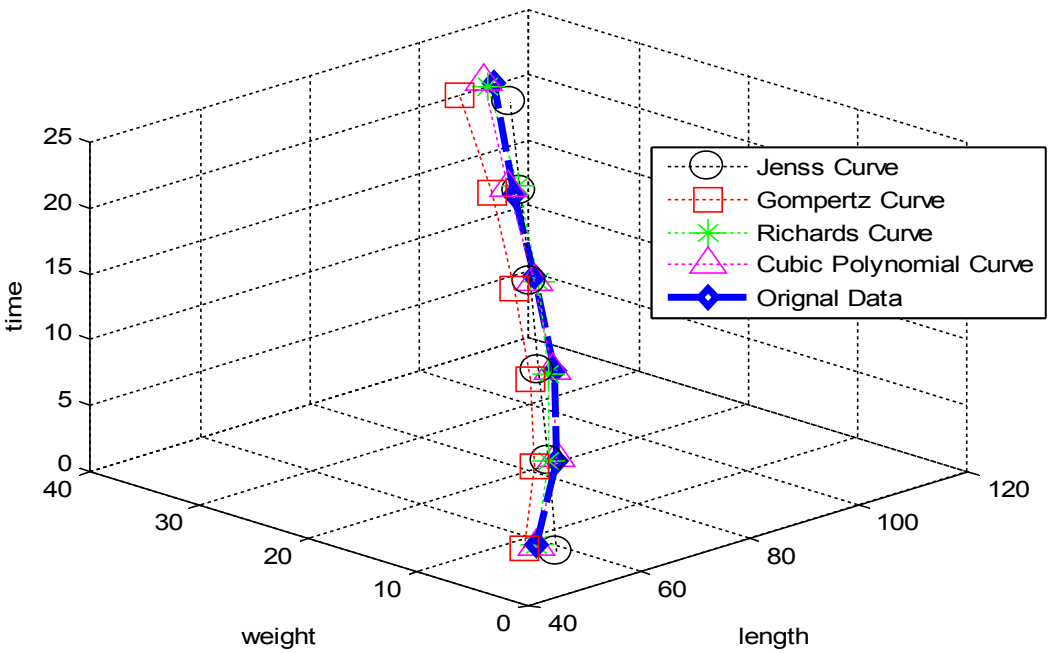


Figure 7.4 Three Dimensional Plot (Time, Weight and Length)



7.5 Discussion

We have exhibited the 90% Credible Intervals, the fitted curves, and the estimates of the model parameters for the four growth curves in Figures 7.1-4 and Table 7.1-2. There, we see that for weight versus time (Figure 7.1) and length vs. time (Figure 7.2), the 90% CI of Cubic curve is the narrowest among all four curves when time is small (but diverges like a funnel shape as time increases to approximately more than 15 days); for weight vs. time in Figure 7.1, Jenss and Gompertz curves both have relatively narrow 90% CIs, whereas for length vs. time in Figure 7.2, Jenss curve has smaller 90% CI. In Figures 7.3-7.4, we observed that on the one hand, the data display a positive trend that as length increases, the rate of change in weight also increases; on the other hand, when weight increases, the rate of change in length decreases.

Although all four curves fit the data reasonably well, the Cubic curve is apparently the best fit curve among them. In addition, we see that, as time increases (approximately before the fifteenth day), the rate of change in length and in weight both increase as weight and length increase. Here we observe again that the Cubic curve appears to be the best fit curve for the given data (BIC can be one of the useful criteria in model selection because the smaller value it is, the better curve fitting it will be, and this is consistent with our observations in those Figures).

In summary, our Bayesian analysis of multivariate growth curve models provides a formulation for generating Bayesian estimates as well as describing the dependence

relationship between variables with a certain autocorrelation relationship under consideration.

References

Barry, D. (1995). A Bayesian model for growth curves analysis. *Biometrics*, **51**, pp. 639-655.

Fearn, T. (1975). A Bayesian approach to growth curves. *Biometrika*, **62**, pp. 89-100.

Geisser, S. (1970). Bayesian analysis of growth curves. *Sankhya A*, **32,** pp. 53-64.

Geisser, S. (1980). Growth curve analysis. In *Handbook of Statistics* (P.R. Krishnaiah ed.) North-Holland, Amsterdam, pp. 89-115.

Grizzle, J. E. and D. M. Allen (1969). Analysis of growth curves and responses. *Biometrics*, **25**, pp. 357-381.

Khatri, C. G. (1966). A note on a MANOVA model applied to problems in growth curves. A*nnals of Institute of Statistical Mathematics*, **18**, pp. 75-86.

Lee, J. C. and S. Geisser (1975). Growth curve prediction. *Sankhya A*, **34**, pp. 393-412.

Lindley, D. V. and A.F.M. Smith (1972). Bayes estimates for the linear model. *Journal of the Royal Statistical Society B*, **34**, pp. 1-41.

Press, S. J. (2002). *The Subjective and Objective Bayesian Statistics Principles, Models and Applications.* 2nd edition, John Wiley & Sons.

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

Rao, C. R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika*, **52**, pp. 447-458.

Seber, G. A. F. and C. J. Wild (1989). *Nonlinear Regression*. John Wiley & Sons.

von Rosen, D. (1989). Maximum likelihood estimates in multivariate linear normal model. *Journal of Multivariate Analysis*, **31**, pp. 187-200.

von Rosen, D. (1990). Moments for a multivariate linear normal models with application to growth curve model. *Journal of Multivariate Analysis*, **35**, pp. 243-259.

von Rosen, D. (1991). The growth curve model: a review. *Communication in Statistics: Theory and Methods*, **20**, pp. 2791-2882.

Chapter 8 Application of Dirichlet Mixture of Normals in Growth Curve Models

8.0 Abstract

In this chapter, we present growth curve models with an auxiliary variable which contains an uncertain data distribution based on mixtures of standard components, such as normal distributions. The multimodality of the auxiliary random variable motivates and necessitates the use of mixtures of normal distributions in our model. We have observed that Dirichlet process priors, composed of discrete and continuous components, are appropriate in addressing the two problems of determining the number of components and estimating the parameters simultaneously and are especially useful and advantageous in the aforementioned multimodal scenario.

The application of Dirichlet mixture of normals (DMN) in growth curve models under Bayesian formulation is presented and algorithms for computing the number of components, as well as estimating the parameters are also rendered. The simulation results show that our model gives improved goodness of fit statistics over models without DMN and the estimates for the number of components and for parameters are reasonably accurate.

8.1 Introduction

In our study of growth curve models with auxiliary variables, we observed that there are situations when auxiliary variables could come from multimodal distributions. For example, the distribution of the weekly average oxygen as an auxiliary variable on the fish weight growth (Dellaportas and Stephens, 1995) exhibits some multimodality, i.e., the auxiliary variable may come from uncertain data distributions or mixtures of some standard components, such as normal distributions. So, if we could somehow estimate the number of components as well as the parameters of the mixtures of normals, then our model would be a better fit to that specific type of data. In our literature survey, we find that the framework of the Dirichlet process by Ferguson (1983) for deriving

mixture normals is useful in addressing our problem. There are, of course, subsequent developments along this line by West (1990) concerned with the algorithm for predictive distributions; Escobar and West (1995) developed a method to compute and evaluate the parameters of posterior and predictive distributions. MacEachern and Muller (2000) used Dirichlet process priors to extend the parametric models into the nonparametric models. Muller and Quintana (2004) reviewed some Bayesian inference problems, including Dirichlet process models. Bhattacharya and SenGupta (2009), using Dirichlet process and simple Gibbs sampling algorithm for simulating the mixture components involved in their multivariate linear-circular model, were able to handle a varieties of uncertainties arising in a linear-circular data set. We refer mainly to Bhattacharya and SenGupta (2009) for our base framework and extend the application to growth curve models. Thus the novelty of this paper lies in applying the Dirichlet process to growth curve models with auxiliary variables, handling uncertainties from mixtures of normals, and deriving algorithms to generate estimates for the number of components as well as for the parameters of the model.

What motivates us in using Dirichlet Process Priors (DPP) for mixture normals is that, given the distribution of the error terms of the data, the model is appropriate and useful for determining the number of components and in estimating the parameters. If there are p components in the mixture normals from which the data was obtained, we begin by assuming the maximum number of distinct components p_0 ($p_0 \geqslant p$). Then through DPP and the use of appropriate algorithms, we can simultaneously get an

estimate $p*$ for p (the discrete part of DPP) and obtain reasonable estimates for the parameters (the continuous part of DPP).

This chapter is organized in the following order: the first section is the introduction and literature review. In the second section, we will introduce the model under two scenarios. The third section will describe the algorithm for sampling from the model for inferences and the simulation results generated. Lastly, we will conclude with some discussion.

8.2 Model Formulation

The growth curve with auxiliary variable model is assumed to be

$y_t = f(t,\theta) + q(z_t,\phi) + \varepsilon_t$, $t = 1,\ldots,T$ where $\varepsilon_t \sim N(0,\sigma_t^2)$ so

$y_t \sim N(f(t,\theta) + q(z_t,\phi), \sigma_t^2)$ (σ_t^2 will be assumed to be equal to σ^2 and that σ^2 is either a constant or follows a certain prior distribution later) and $y_t's$ are independent. In addition, we assume a linear relationship for the function of auxiliary variable, which is $q(z_t,\phi) = \phi z_t$, and the mixture normals from which the auxiliary variable is sampled is given by:

$$z_t \sim \sum_{i=1}^{p} \pi_i N(\mu_i, \sigma_z^2) \quad , \tag{8.1}$$

where $\sum_{i=1}^{p} \pi_i = 1$ and p is the unknown number of normal components and must be inferred from the prior and the data set. This can also be expressed as an average of a fixed number of components, as shown in this alternate form:

$$z_t \sim \frac{1}{p_0}\sum_{i=1}^{p_0} N(\mu_i, \sigma_z^2) \quad , \tag{8.2}$$

where $p_0 (\geqslant p)$ can be defined as the maximum number of distinct mixture components of z_t, with $\pi_i = \frac{1}{p_0}$ for each i. Based on the implicit assumption that π_i can be reasonably approximated by the ratio of two integers, we now show in the following that under a nonparametric prior assumption for $\{\mu_i : 1 \leqslant i \leqslant p_0\}$, (8.2) boils down to the form of (8.1).

We assume that the parameters (means of the mixture normals) $\{\mu_i : 1 \leqslant i \leqslant p_0\}$ are samples drawn from a Dirichlet Process (see Escobar and West, 1995). In other words, we assume that $\mu_1, \ldots, \mu_{p_0}$ are samples from some unknown prior distribution $G(\cdot)$ on the real line. We further assume $G \sim D(\alpha G_0)$, a Dirichlet process defined by α, a positive scalar, and $G_0(\cdot)$, a specified univariate distribution function over the real line. Put more simply, we assume that

$\mu_1, \ldots, \mu_{p_0} | G_0 \overset{iid}{\sim} G$ and $G \sim D(\alpha G_0)$.

A larger α value here indicates a higher probability of sampling μ_i from the base distribution G_0 than from $\boldsymbol{\mu}^{(\neg i)}$ (define $\boldsymbol{\mu} = \{\mu_1, \ldots, \mu_{p_0}\}$ and $\boldsymbol{\mu}^{(\neg i)} = \{\mu_1, \ldots, \mu_{i-1}, \mu_{i+1}, \ldots, \mu_{p_0}\}$) because the full conditional priors for μ_i is given by

$$(\mu_i | \boldsymbol{\mu}^{(\neg i)}) \sim (\alpha + p_0 - 1)^{-1} \alpha G_0(\mu_i) + (\alpha + p_0 - 1)^{-1} \sum_{j \neq i,\, j=1}^{p_0} \delta_{\mu_j}(\mu_i) \quad , \tag{8.3}$$

(8.3) in fact shows that the joint distribution of $(\mu_1, \ldots, \mu_{p_0})$ is given by the following:

$\mu_1 \sim G_0$ and for $i = 2, \ldots, p_0$,

$\mu_i | \mu_1, \ldots, \mu_{i-1} \sim (\alpha + i - 1)^{-1} \alpha G_0(\mu_i) + (\alpha + i - 1)^{-1} \sum_{j=1}^{i-1} \delta_{\mu_j}(\mu_i)$. Thus, given a sample $\mu_1, \ldots, \mu_{i-1}$, μ_i is drawn from G_0 with probability $\alpha(\alpha + i - 1)^{-1}$ and is otherwise drawn uniformly with probability $(\alpha + i + 1)^{-1}$ from among the sample $\{\mu_1, \ldots, \mu_{i-1}\}$.

Now, suppose that a sample from the joint distribution of $\boldsymbol{\mu} = \{\mu_1, \ldots, \mu_{p_0}\}$ yields p^* distinct realizations given by $\boldsymbol{\mu}^*$ and let p_i denote the number of times the i^{th} member of $\boldsymbol{\mu}^*$ appears in $\boldsymbol{\mu}$. Then $\pi_i = p_i / p_0$, so

$\sum_{i=1}^{p} \pi_i = \sum_{i=1}^{p_0} \frac{1}{p_0} = \sum_{i=1}^{p^*} \frac{p_i}{p_0} = 1$, where $\sum_{i=1}^{p^*} p_i = p_0$.

In other words, we get an estimate p^* for p by assuming a maximum number of distinct realizations of samples p_0 and then collapse p_0 to only p^* . This explains how (8.2) reduces to the form of (8.1).

In our simulation example, we assume p_0 =5 to be the maximum number of components, $\alpha \sim Gamma(1, 1)$ and the base prior for Dirichlet process is $G_0(\mu_i) \equiv N(\tilde{\mu}_i, {\sigma_0}^2)$, where $\tilde{\mu}_i$ and σ_0^2 are some known constants. We can get the distribution for (8.2) as follows:

For $t = 1, \ldots, T$ and $i = 1, \ldots, p_0$,

$$p(Z|\mu,\sigma_z^2)=\frac{1}{(2\pi\sigma_z^2)^{-T/2}p_0^T}\sum_{i=1}^{p_0}\exp\left\{-\frac{\sum_{t=1}^{T}(z_t-\mu_i)^2}{2\sigma_z^2}\right\} \quad , \tag{8.4}$$

we can rearrange the terms and rewrite it as

$$(\mu_i|Z,\sigma_z^2)\sim N\left(\bar{Z},\frac{\sigma_z^2}{T}\right) \text{ , where } \bar{Z}=\frac{1}{T}\sum_{t=1}^{T}z_t \quad . \tag{8.5}$$

To simplify our analysis, we let σ_z^2 be a known constant, then we get the full conditional of μ_i given the rest in conjunction with (8.3) as follows

$$\begin{aligned}(\mu_i|Z,\sigma_z^2,\mu^{(\neg i)})\sim&(\alpha+p_0-1)^{-1}\alpha G_0(\mu_i)\,p(\mu_i|Z,\sigma_z^2)\\&+(\alpha+p_0-1)^{-1}\sum_{l\neq i,l=1}^{p_0}p(\mu_l|Z,\sigma_z^2)\delta_{\mu_i}(\mu_l) \quad .\end{aligned} \tag{8.6}$$

Thus by substituting $G_0(\mu_i)\equiv N(\tilde{\mu}_0,{\sigma_0}^2)$ into (8.6) and reorganizing the equation, we obtain the full conditional distribution for μ_i as follows (we use $\cdot$ to denote the given conditions and let $\lambda=\frac{\sigma_z^2}{\sigma_z^2+T\sigma_0^2}$ to simplify the notations):

$$(\mu_i|\cdot)\begin{cases}\sim\tilde{G}(\mu_i)=N\left(\lambda\tilde{\mu}_i+(1-\lambda)\bar{Z},\lambda\sigma_0^2\right) \text{ with probability } c\,q_{0i}\\=\mu_l,\ l=1,\ldots,p_0,\ \ l\neq i \quad \text{with probability } c\,q_{li}\end{cases} \quad , \tag{8.7}$$

where $q_{oi}=\frac{\alpha}{\sigma_0\sigma_z}\varphi\left(\frac{\bar{Z}-\tilde{\mu}_i}{\sqrt{(\sigma_z^2/T+\sigma_0^2)}}\right)$, $q_{li}=\varphi\left(\frac{\mu_l-\bar{Z}}{\sigma_z/\sqrt{T}}\right)$, $\varphi(\cdot)$ is the standard normal probability density function and subject to the constraint

$c\,q_{0i}+\sum_{l\neq i,l=1}^{p_0}c\,q_{li}=1$, i.e.

$$c=\left[\alpha q_{0i}+\sum_{l\neq i,\, l=1}^{p_0} q_{li}\right]^{-1} \quad .$$

For simplicity and illustrative purposes, we assume $f(t,\theta)=\theta_1+\theta_2 t$, $q(z_t,\phi)=\phi z_t$ and the priors for $\theta_i \sim N(\tilde{\theta}_i, s^2)$, $i=1,2$ and $\phi \sim N(\tilde{\phi}, s^2)$ where $\tilde{\theta}_i, i=1,2$ and $\tilde{\phi}$ are given constants. Now we will see the full conditional distribution for θ and ϕ by considering two different cases for the prior variances s^2 :

Case A: $s^2=\sigma^2$

For θ_1 we have $p(\theta_1|\cdot)\propto p(\theta_1)\left[p(Y|f(\theta,\boldsymbol{t}),q(\phi,Z),\sigma^2)\,p(\sigma^2)\right]$, where $\boldsymbol{t}=(1,2,...,T)$. We substitute in the Jeffrey prior for σ^2 , integrate out σ^2 , reorganize the terms in the brackets (the likelihood function), and let $Q_t=y_t-\theta_2 t-\phi z_t$ we get $(\theta_1|\cdot)\sim t_T(\mu_\theta)$, a $T-1$ degree of freedom t distribution, where the location parameter and the scale parameter are

$$\mu_\theta \;=\; \frac{1}{(T+1)}\left(\sum_{t=1}^{T} Q_t+\tilde{\theta}_1\right) \quad \text{, and}$$

$$\sigma_\theta^2 \;=\; \frac{1}{(T+1)(T-1)}\left[\left(\sum_{t=1}^{T} Q_t^2+\tilde{\theta}_1^2\right)-\frac{1}{(T+1)}\left(\sum_{t=1}^{T} Q_t+\tilde{\theta}_1\right)^2\right] \quad \text{, respectively.}$$

Like θ_1 , the full conditional distribution for θ_2 can be obtained in a similar way and is $(\theta_2|\cdot)\sim t_T(\mu_\theta)$, a *t*-distribution with $T-1$ degree of freedom, where the location parameter and the scale parameter are

$\mu_\theta = \frac{1}{\sum t^2+1}\left(\sum_{t=1}^{T} R_t+\tilde{\theta}_2\right)$ and

$\sigma_\theta^2 = \frac{1}{\left(\sum t^2+1\right)(T-1)}\left[\left(\sum_{t=1}^{T} R_t^2+\tilde{\theta}_2^2\right)-\frac{1}{\left(\sum t^2+1\right)}\left(\sum_{t=1}^{T} R_t+\tilde{\theta}_2\right)^2\right]$, respectively, where

$R_t=\left(y_t-\theta_1-\phi z_t\right)t$.

Like θ_1 and θ_2 , the full conditional distribution for ϕ can be obtained in a similar way and is $(\phi|\cdot)\sim t_T(\mu_\phi)$, a t-distribution with $T-1$ degree of freedom, where the location parameter and the scale parameter are

$\mu_\phi = \frac{1}{\sum z_t^2+1}\left(\sum_{t=1}^{T} O_t+\tilde{\phi}\right)$ and

$\sigma_\phi^2 = \frac{1}{(T-1)\left(\sum z_t^2+1\right)}\left[\left(\sum_{t=1}^{T} O_t^2+\tilde{\phi}^2\right)-\frac{1}{\sum z_t^2+1}\left(\sum_{t=1}^{T} O_t+\tilde{\phi}\right)^2\right]$, respectively and

$O_t=\left(y_t-\theta_1-\theta_2 t\right)z_t$.

Case B: $s^2\neq\sigma^2$

For θ_1 we have $p(\theta_1|\cdot)\propto p(\theta_1)\left[p(Y|f(\theta,W),q(\phi,Z),\sigma^2)\,p(\sigma^2)\right]$. Substituting in the Jeffrey prior for σ^2 , integrating out σ^2 , reorganizing the terms in the brackets (actually the likelihood function), and letting $Q_t=y_t-\theta_2 t-\phi z_t$ we get a t distribution with $T-1$ degree of freedom, where the location parameter and the scale parameter are $\frac{1}{T}\sum_{t=1}^{T} Q_t$ and $\frac{1}{T}\sum_{t=1}^{T} Q_t^2-\frac{1}{T^2}(\sum_{t=1}^{T} Q_t)^2$,

respectively. So the full conditional distribution for θ_1 is proportional to the product of the density function of the prior of $N(\tilde{\theta}_1, s^2)$ and the density function of this t distribution.

Like θ_1, the full conditional distribution for θ_2 can be obtained in a similar way and is the product of a density function of the prior of $N(\tilde{\theta}_2, s^2)$ and the density function of a t distribution with $T-1$ degree of freedom, location parameter $\frac{1}{\sum t^2}\sum_{t=1}^{T} R_t$ and scale parameter $\frac{1}{\sum t^2}\sum_{t=1}^{T} R_t^2 - \frac{1}{\left(\sum t^2\right)^2}\left(\sum_{t=1}^{T} R_t\right)^2$, where

$R_t = \left(y_t - \theta_1 - \phi z_t\right)t$.

Further, the full conditional distribution for ϕ can be obtained in a similar way and is the product of a density function of the prior of $N(\tilde{\phi}, s^2)$ and the density function of a t distribution with $T-1$ degree of freedom, location parameter is $\frac{1}{\sum z_t^2}\sum_{t=1}^{T} O_t$ and scale parameter is $\frac{1}{\sum z_t^2}\sum_{t=1}^{T} O_t^2 - \frac{1}{\left(\sum z_t^2\right)^2}\left(\sum_{t=1}^{T} O_t\right)^2$ where

$O_t = \left(y_t - \theta_1 - \theta_2 t\right)z_t$.

8.3 Algorithm for Implementing the Model

Escobar and West (1995) elucidated that estimating the number of components in their Dirichlet mixtures of normals (which is equivalent to the $p*$, an estimate of p in our model) could be as straightforward as using a histogram approximation. In our

research, we present a way to estimate $\boldsymbol{\mu}$ using an algorithm which combines the calculation and estimation of the "discrete random variable"(p) and the "continuous random variable" ($\boldsymbol{\mu}$) in one setting. In our algorithm, we use a similar but yet more complicated approach to obtain the estimate of p (i.e., p^*) by collapsing the same $\mu_i's$ in $\boldsymbol{\mu}$ in the iteration (composed of modes) from iterations which have both the same number of modes (from the histogram) and the identical combination (whose frequency is higher than those of all other combinations). After collapsing, the distinct $\boldsymbol{\mu}^*$ spontaneously become the estimates of the means of the mixture normals and simultaneously producing estimates for $\boldsymbol{\pi}$ as well.

We use the following Metropolis Hastings within Gibbs sampling algorithm to sample the parameters from the previous full conditional distributions:

- Step 1: Obtain initial values for $\mu^{(0)}$, $\theta^{(0)}$, $\phi^{(0)}$ (e.g., from the least square estimates) and $\alpha^{(0)}$.
- Step 2: Let j be the j^{th} step in the total N iterations, and denote $p^{(j)}$ to be the number of distinct $\boldsymbol{\mu}$ in the j^{th} step, then do

For j in $1:N$

sample $\mu_1^{(j)}$ from $\mu_1 | \mu_2 = \mu_2^{(j-1)}, \mu_3 = \mu_3^{(j-1)}, \ldots, \mu_{p_0} = \mu_{p_0}^{(j-1)}, \theta_1 = \theta_1^{(j-1)}, \theta_2 = \theta_2^{(j-1)}, \phi = \phi^{(j-1)}, \alpha = \alpha^{(j-1)}$,

sample $\mu_2^{(j)}$ from $\mu_2 | \mu_1 = \mu_1^{(j)}, \mu_3 = \mu_3^{(j-1)}, \ldots, \mu_{p_0} = \mu_{p_0}^{(j-1)}, \theta_1 = \theta_1^{(j-1)}, \theta_2 = \theta_2^{(j-1)}, \phi = \phi^{(j-1)}, \phi = \phi^{(j-1)}, \alpha = \alpha^{(j-1)}$,

$\vdots$

$$\text{sample } \mu_{p_0}^{(j)} \text{ from } \mu_{p_0}|\mu_1=\mu_1^{(j)},\mu_2=\mu_2^{(j)},\ldots,\mu_{p_0-1}=\mu_{p_0-1}^{(j)},\ \theta_1=\theta_1^{(j-1)},\theta_2=\theta_2^{(j-1)},\phi=\phi^{(j-1)},\alpha=\alpha^{(j-1)},$$

$$\text{sample } \theta_1^{(j)} \text{ from } \theta_1|\mu_1=\mu_1^{(j)},\mu_2=\mu_2^{(j)},\ldots,\mu_{p_0}=\mu_{p_0}^{(j)},\ \theta_2=\theta_2^{(j-1)},\phi=\phi^{(j-1)},\alpha=\alpha^{(j-1)},$$

$$\text{sample } \theta_2^{(j)} \text{ from } \theta_2|\mu_1=\mu_1^{(j)},\mu_2=\mu_2^{(j)},\ldots,\mu_{p_0}=\mu_{p_0}^{(j)},\ \theta_1=\theta_1^{(j)},\phi=\phi^{(j-1)},\alpha=\alpha^{(j-1)},$$

$$\text{sample } \phi^{(j)} \text{ from } \phi|\mu_1=\mu_1^{(j)},\mu_2=\mu_2^{(j)},\ldots,\mu_{p_0}=\mu_{p_0}^{(j)},\ \theta_1=\theta_1^{(j)},\theta_2=\theta_2^{(j)},\alpha=\alpha^{(j-1)},$$

$$\text{sample } \alpha^{(j)} \text{ from } \alpha|\mu_1=\mu_1^{(j)},\mu_2=\mu_2^{(j)},\ldots,\mu_{p_0}=\mu_{p_0}^{(j)},\ \theta_1=\theta_1^{(j)},\theta_2=\theta_2^{(j)},\phi=\phi^{(j)},$$

$$p^{(j)} = \#\text{ of distinct } \mu_i^{(j)} \text{ for } i=1,\ldots,p_0 \text{ in this } j^{th} \text{ step.}$$

End

- Step 3: Construct a histogram for $p^{(1)},\ldots,p^{(N)}$ and its mode p^* is the estimate of p, the number of distinct realizations of $\boldsymbol{\mu}$.
- Step 4: Choose only the iterations that have exactly p^* distinct realizations and count the frequencies of the (at most) $C_{p^*}^{p_0}$ number of different combinations.
- Step 5: Retain only the highest frequency iterations from step 4. Then in these iterations, keep only the highest frequency identical combination iterations.
- Step 6: Now use the iterations picked out in Step 5 to:
  1. Calculate the means of the same distinct realization in the iterations, these means are estimates $\boldsymbol{\mu}^*$ for $\boldsymbol{\mu}$.
  2. Calculate the mode of the numbers of duplications for every single distinct

realization in these iterations and divide this mode by p_0 , we get estimates $\boldsymbol{\pi}^*$ for $\boldsymbol{\pi}$.

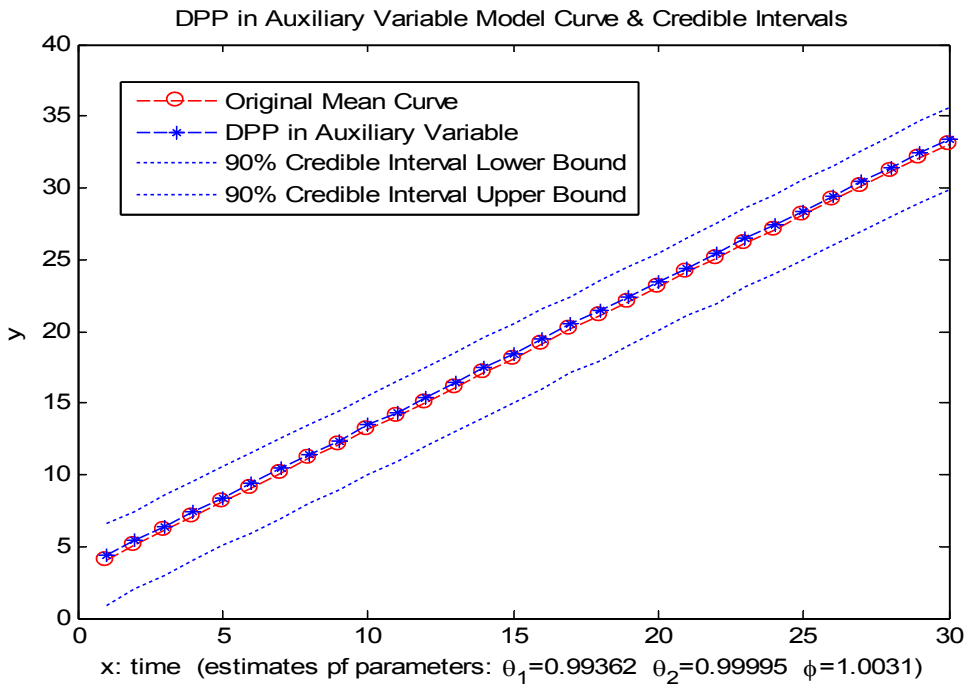
3. Thus we obtain estimates p^* (in step 3), $\boldsymbol{\mu}^*$ and $\boldsymbol{\pi}^*$ for p , $\boldsymbol{\mu}$ and $\boldsymbol{\pi}$, respectively.

8.4 Simulation

For simplicity, we consider Case A and run a simulation of 3000 iterations (the first 2000 burn-ins dropped) by assuming $T=30$, $\theta=(\theta_1 \ \theta_2)=(1 \ 1)$, $\alpha=0.5$, $\phi=1$ and that the auxiliary variable data come from a population of a mixture of two normal distributions, with probability $\pi_1=0.3$ from $N(0,1)$ and probability $\pi_2=0.7$ from $N(3,1)$. Using the aforementioned algorithm, we get these estimates :

1. $p^*=2$ (the same as the number of components in the original data).
2. $\mu^*=(0.55017, 2.71725)$, $\pi_1^*=0.2$ and $\pi_2^*=0.8$ (very close to the values above).
3. $(\theta_1^* \ \theta_2^* \ \phi_1^*) = (0.99362 \ 0.99995 \ 1.00312)$, almost the same as the above values. Standard deviations are $(0.17921 \ 4.38442 \ 0.66974)$.
4. Adjusted R^2 for model using DPP is 0.9546 (better than 0.95295 for not using DPP model).

Figure 8.1



8.5 Conclusion

In this chapter we have presented growth curve models with auxiliary variables containing uncertain data distributions based on mixtures of standard components and using normal distributions in our simulation example. We have also developed a useful algorithm in estimating the model parameters. The results from this algorithm show that our model can successfully estimate the number of components in the mixture normals, the probabilities from which the auxiliary variables arise and the means of the normal distributions in the components of the mixture normals. The estimates and goodness of fit statistics (adjusted R^2), show that models with DPP can outperform those models without DPP.

References

Bhattacharya, S. and SenGupta, A. (2009). Bayesian Analysis of Semiparametric Linear-Circular Models. *Journal of Agricultural Biological Environmental Statistics*, **14**, No. 1, pp. 33-65.

Dellaportas, P. and Stephens, S. A. (1995). Bayesian Analysis of Errors-in-Variables Regression Models. *Biometrics*, **51**, pp.1085-1095.

Escobar, M. D. And West M. (1995). Bayesian Density Estimation and Inference Using Mixtures. *Journal of American Statistical Association*, **90**, No. 430, pp. 577-588.

Ferguson, T. S. (1983). Bayesian Density Estimation by Mixtures of Normal Distributions. *In Recent Advances in Statistics*, eds. H. Rizvi and J. Rustagi, New York: Academic Press, pp. 287-302.

MacEachern, S. N. And Muller P. (2000). Efficient MCMC Schemes for Robust Model Extensions Using Encompassing Dirichlet Process Mixture Models. *In Robust Bayesian Analysis*. Lecture Notes in Statistics, **152**, pp.295-316. Springer, New York.

Muller, P. and Quintana F. A. (2004). Nonparametric Bayesian Data Analysis. *Statistical Science.*,**19**, N0. 1, pp. 95-110.

West, M. (1990). Bayesian Kernel Density Estimation. *Discussion Paper* 90-A02. Duke University, Institute of Statistics and Decision Sciences.