The autoregressive and moving average models are combined into the ARMA class, thus an ARMA(1,1) model is given by

$$y_t - \theta y_{t-1} = \varepsilon_t - \phi - \varepsilon_{t-1}, \tag{2.9}$$

where y_t is the observation at time t, ϕ is the moving average parameter, θ the autoregressive parameter, the ε_t's are white noise, and y_0 is a known constant. Thus there are three parameters to the model. The ARMA model is parsimonious, that is, it is able to represent time series data with a few parameters, thus for a given set of observations {y(t), t = 1,2, …, n} an ARMA(1,1) model and an AR(8) (8th order autoregressive model) perhaps will give the same "fit" to the data. Parsimony was emphasized by Box and Jenkins (1970) in their development of the statistical analysis of time series which are generated by ARMA models. Indeed they say in practice one is not likely to encounter a case where either (or both) the moving average order a or (and) the autoregressive order p is in excess of two. The Box and Jenkins analysis of time series is a major contribution to statistical methodology and is today very popular. Compared to Box-Jenkins methodology, the Bayesian approach is in its infancy.

Nevertheless, the Bayesian posterior analysis and forecasting techniques of ARMA models are attempted in Chapter 5, where a complete treatment of autoregressive processes is given. The ARMA(1,1), MA(1) and MA(2) processes are also examined, but only partial success can be reported here for the MA(2) case. Monahan (1983) has recently developed a Bayesian analysis of ARMA models for low order processes.

Models for Structural Change

Two important references on structural change in linear models are Poirier (1976) and more recently the special issue of the Journal of Econometrics, which was edited by Broemeling (1982). The book by Poirier is a review of the statistical and econometric literature on structural change and presents new ideas about how one should model structural change with the aid of spline functions. The special issue contains ten articles on the subject and introduces the reader to a large variety of theory and methodology. This issue contains some material which was introduced before 1976, but for the most part includes ideas which appeared after Poirier's book.

The term "structural change" doesn't have a definite definition and is used by econometricians to denote a change in the parameters of a model which explains a relationship between economic variables. Also, the term is often employed to refer to a model which has been misspecified. Terms or phrases such as shift point, change point, transition function, separate regressions, switching regression, and two-phase regression, although not identical in meaning to structural change, are involved in some way with structural change.

When data are collected in an ordered sequence, often one is interested in the probabilistic structure of the measured variables from one subset of the time domain to the other. Suppose for example a major economic policy change is put into effect at some time, then the researcher is perhaps interested in assessing the effects of the change on the variables under study.

Two fundamental questions now occur:

(i) Has a change occurred among the variables under investigation during the observation period? This is called the detection problem because one wants to detect a change in the relationship.

(ii) Assuming a change has taken place, can one estimate the parameters of the model? For example, one may want to estimate the time where the change occurred as well as the other parameters of the model, namely those which explain the "before" and "after" relationship between the variables.

Problems of structural change are found in a wide variety of disciplines. Often the variables under study are economic where the policy change could be a new tax law (such as the windfall profits tax), a new government program (such as price supports), or a major disturbance to the economy (an oil embargo). In biology, one might focus on the life of an organism and the time when a change in growth pattern will occur. These examples and others are given by Holbert (1982), who reviews Bayesian developments in structural change from 1968 to the present.

An interesting example reported by Holbert is a two-phase regression problem illustrated with stock market sales volume.

On January 3, 1970, *Business Week* reported regional stock exchanges were hurt by abolition of give-ups (commission splitting) on December 5, 1968. McGee and Carleton (1970) analyzed the relationship between dollar volume sales of the Boston Stock Exchange and the combined New York and American Stock Exchanges. They did this to investigate the hypothesis that a structural change occurred (between dollar volume sales of the regional and national exchanges) after December 5, 1968, and their analysis was based on a piecewise regression methodology. Holbert (1982) reexamined the data, but used the Bayesian approach, and came to the same conclusion as McGee and Carleton, namely that abolition of split-ups did indeed hurt the regional exchanges.

The Bayesian approach was based on the two-phase regression model

$$y_i = \begin{cases} \alpha_1 + \beta_1 x_i + e_i, & i=1,2,\ldots,m, \\ & 2 \leqslant m \leqslant n-2 \\ \alpha_2 + \beta_2 x_i + e_i, & i=m+1,\ldots,n, \end{cases} \tag{2.10}$$

where the e_i's are n.i.d. (O,τ^{-1}), y_i. is the sales volume on the combined exchanges at time i, x_i. the corresponding total dollar volume on the Boston Stock Exchange at time i, m the unknown shift point, and α_i. and β_i. (i = 1,2) are unknown regression parameters. Thus the first m pairs $(x_i,y_i.)$ follow one relationship and the last n – m pairs $(x_i,y_i.)$ a regression relation with parameters α_2 and β_2. The main question here is what is m? Is m "close to" December 5, 1968? Does m have a large posterior probability corresponding to support the hypothesis the regional exchanges were damaged because of the abolition of give-ups?

Holbert's analysis adopted a vague improper prior for the parameters

$$\begin{aligned} &\xi(\alpha_1,\beta_1,\alpha_2,\beta_2) \propto \text{constant}, \quad \alpha_i \in R, \beta_i \in R \\ &\xi(m)=(n\text{-}3)^{-1}, \quad 2 \leqslant m \leqslant n-2 \\ &\xi(\tau) \propto 1/\tau, \quad \tau < 0 \end{aligned} \tag{2.11}$$

which when combined with the likelihood function gave

$$\xi(\mathrm{m}|\mathrm{data}) \propto [\mathrm{m}(\mathrm{n}-\mathrm{m}) \sum_{\mathrm{i}-1}^{\mathrm{m}} (\mathrm{x_i}-\bar{\mathrm{x}}_1)^2 \sum_{\mathrm{i}=\mathrm{m}+1}^{\mathrm{n}} (\mathrm{x_i}-\bar{\mathrm{x}}_2)^2]^{-1/2} \times [\sum_{\mathrm{i}=1}^{\mathrm{m}} (\mathrm{y_i}-\hat{\mathrm{y}}_{\mathrm{i,m}})^2 \sum_{\mathrm{i}=\mathrm{m}+1}^{\mathrm{n}} (\mathrm{y_i}-\hat{\mathrm{y}}_{\mathrm{i,\,m+1}})]^{-(\mathrm{n}-4)/2} \tag{2.12}$$

where $2 \leqslant m \leqslant n - 2$, as the marginal posterior mass function of m. In this formula $\bar{x}_{1,m}$ is the mean of the first m x-values, $\bar{x}_{2,m}$ the mean of the last n – m values, $\hat{y}_{i,m}$. is the predicted y_i. value based on the first m (x,y) pairs, and $\hat{y}_{i,m+1}$. the predicted value of y_i. based on the last n – m (x,y) pairs. Holbert computes the function (normalized) (2.12) for the months from February, 1967, to November, 1969, and obtains a "large" posterior probability of .2615 for November, 1968, the month just before the change went into effect.

The two-phase regression model is one of the many models of structural change which appear in Chapter 7. Other models which are examined are normal sequences, time series processes such as the autoregressive and regression models with autocorrelated errors. The multiple linear regression model with one change is investigated in detail and a thorough sensitivity study is done. In most of the models either a vague improper prior density or a quasiconjugate prior density is used.

Models for structural change usually incorporate only one change (or at most two changes). On the other hand, models which incorporate parameters which always change (from one time point to the next) are also very important and as we will see have been quite valuable to engineers and scientists in the communication and space industries.

Dynamic Linear Models

The Kalman filter is one of the major achievements of mathematical (statistical) modeling, however, it is not well-known to statisticians. The Kalman filter is an algorithm which recursively computes the mean of the posterior distribution of a parameter of a dynamic model. Models are called dynamic when the parameters of the model are actually always changing, thus the traditional models of regression and the design of experiments are static, since the parameters of these models are not thought to be always changing their values.

The dynamic linear model is for t = 1,2, …

$$\begin{aligned} \mathrm{y(t)} &= \mathrm{F(t)}\theta\mathrm{(t)} + \mathrm{U(t)}, \\ \theta\mathrm{(t)} &= \mathrm{G}\theta\mathrm{(t\text{-}1)} + \mathrm{V(t)}, \end{aligned} \tag{2.13}$$

where y (t) is a m × 1 observation vector, F(t) a m × p matrix, θ(t) a p × 1 state vector, U(t) a m × 1 unobservable observation error vector, G a p × p systems transition matrix, and V(t) a p × p unobservable systems error vector.

The primary parameters are the state vectors θ(0), θ(1),…, and consequently one must estimate them as the observations become available that is, given y(l), what is θ(1), given y(l) and y(2), what is θ(2), etc. The first equation of the model is the observation equation which relates the observation in a linear way to the state vector and by itself is the usual linear model of regression and the design of experiments. The second equation gives the model its dynamic flavor, and is called the systems equation and tells us how the parameters of the model change. The second equation is a first-order vector autoregressive model (examined in Chapter 8), which tells us the parameter (state vector) is always changing in such a way that the next value is a linear function of the present value apart from a random error.

One immediately notices that the dynamic model contains a large number of parameters when compared to their static counterparts, thus one usually assumes that F(t), t = 1,2, … and the covariance matrices of the error vectors U(t) and V(t) are known. Assuming these quantities are known as well as making other assumptions such as normality of the error vectors, Kalman (1960) found the mean of the marginal posterior normal distribution of θ(t) given y(l), y(2), …, y(t) for all t = 1,2, …, but more importantly, developed a recursive formula for its computation. It was very important to have a recursive formula since often the time between successive observations is very short, only a few seconds or less.

The Kalman filter is often used to compute the position of a satellite or a missile and tracking these craft requires the computation to be extremely fast, thus Kalman's filter was a major contribution.

Consider again the dynamic linear model, then θ(t) might represent the actual position of a satellite at time t, y(t) the measured position by radar or telemetry, and U(t) the error in the observed position. Knowing the relationship between the actual and observed positions allows one to know F(t) and knowing the noise characteristics of the radar or telemetry signals allows one to know the covariance matrix of the observation error vector U(t). Celestial mechanics assists one to find a law which relates the successive positions of the satellite from one time to the next, that is, one knows G, the transition matrix of the system equation. See Jazwinski (1970, page 324) for an example of orbit mechanics.

Tracking problems are easily handled by dynamic models (2.13) as are navigation problems where it is necessary to control a missile, aircraft, or submarine. A linear dynamic model for navigation is given by the observation equation of (2.13