Chapter 6
State Space Methods for Latent Trajectory and Parameter Estimation by Maximum Likelihood

Jacques J. F. Commandeur, Siem Jan Koopman, and Kees van Montfort

Abstract We review Kalman filter and related smoothing methods for the latent trajectory in multivariate time series. The latent effects in the model are modelled as vector unobserved components for which we assume particular dynamic stochastic processes. The parameters in the resulting multivariate unobserved components time series models will be estimated by maximum likelihood methods. Some essential details of the state space methodology are discussed in this chapter. An application in the modelling of traffic safety data is presented to illustrate the methodology in practice.

6.1 Introduction

This chapter concerns multivariate state space analysis and discusses some particular issues of interest, see Durbin and Koopman (2001) and Commandeur and Koopman (2007).

Multivariate state space analysis is applicable to situations where two or more time series need to be analysed simultaneously. However, the material in this chapter also provides a unified treatment for univariate time series. In classical regression analysis a linear relationship is assumed between the dependent variable y_i and an independent variable x_i. The standard regression model for n realizations or

Jacques J. F. Commandeur
Department of Econometrics, VU University Amsterdam
e-mail: jcommandeur@feweb.vu.nl

Siem Jan Koopman
Department of Econometrics, VU University Amsterdam
e-mail: s.j.koopman@feweb.vu.nl

Kees van Montfort
Department of Econometrics, VU University Amsterdam
e-mail: kvmontfort@feweb.vu.nl

K. van Montfort et al., *Longitudinal Research with Latent Variables*,
DOI 10.1007/978-3-642-11760-2_6,

observations of y_i and covariate x_i for $i = 1, \ldots, n$ can be represented by

$$y_i = a + bx_i + \varepsilon_i,$$

where the disturbances or errors $\varepsilon_1, \ldots, \varepsilon_n$ are normally and independently distributed with mean zero and variance σ_ε^2. The coefficients a and b are unknown and fixed and are usually estimated by employing the regression method. It is implied in a classical regression analysis that the observations y_i, after the corrections for intercept and for independent variable x_i, are assumed to be independent of each other. In a time series context, it is not realistic to assume that the observations are conditionally independent because they are expected to be interrelated through time. When statistical inference is carried out when the observations are known to be subject to serial correlations (time dependencies), various problems can arise when it is based on classical regression analysis. For instance, the well-known F-test and t-test statistics do not have proper F- and t-distributions, respectively, under the null hypothesis. Time series analysis has the primary task to uncover the dynamic development of observations measured over time. By using state space methodology it is assumed that the dynamic properties cannot be observed directly from the data. The unobserved dynamic process at time t can be measured indirectly and is referred to as the state of the time series. The state of the time series may consist of several unobserved components and can be estimated by the Kalman filter.

State space methods originated in the field of control engineering, starting with the ground-breaking paper of Kalman (1960). They were initially (and still are) deployed for the purpose of accurately tracking the position and velocity of moving objects such as ships, airplanes, missiles, and rockets.

Around the eighties of the last century it was recognized by scientists involved in other fields than control engineering that these ideas could well be applied to time series analysis generally as well. Since then state space methods have been applied in a wide range of subjects, including economics, finance, political science, environmental science, the social sciences, road safety and medicine.

The outline of this chapter is as follows. In Section 6.2 we formulate the general multivariate state space model and we discuss several well-known sub models. Section 6.3 deals with the Kalman filter and the estimation of the unobserved states and the unknown model parameters. In Section 6.4 we discuss some tests to check the model assumptions such as normality, independency and homoscedasticity. Finally, we will present an empirical example.

6.2 Linear Gaussian State Space Models

A time series is a set of observations which are sequentially ordered over time. In a state space analysis the time series observations are assumed to depend linearly on a *state vector* that is unobserved and is generated by a stochastically time-varying process (a dynamic system). The observations are further assumed to be subject to measurement error that is independent of the state vector. The state vector can be estimated or identified once a sufficient set of observations becomes available. In this section we concentrate on the state space model and its special cases. In Section 6.3 we discuss methods for estimation, residual analysis and forecasting on the basis of state space models. The expositions rely mostly on the introductory textbook by Commandeur and Koopman (2007) and on the more advanced textbook by Durbin and Koopman (2001).

The general linear Gaussian state space model for the n-dimensional observation sequence $y_1,\dots,y_n$ is given by

$$y_t = Z_t\alpha_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathrm{NID}(0,H_t), \tag{6.1}$$

$$\alpha_{t+1} = T_t\alpha_t + R_t\eta_t, \qquad \eta_t \sim \mathrm{NID}(0,Q_t), \qquad t = 1,\dots,n, \tag{6.2}$$

where α_t is the state vector, ε_t and η_t are disturbance vectors and the system matrices Z_t, T_t, R_t, H_t and Q_t are fixed and known but a selection of elements may depend on an unknown parameter vector. Equation (6.1) is called the *observation* or *measurement equation*, while (6.2) is called the *state* or *transition equation*. The $p \times 1$ observation vector y_t contains the p observations at time t and the $m \times 1$ state vector α_t is unobserved. The $p \times 1$ irregular vector ε_t has zero mean and $p \times p$ variance matrix H_t.

The $p \times m$ matrix Z_t links the observation vector y_t with the unobservable state vector α_t and may consist of regression variables. The $m \times m$ transition matrix T_t in (6.2) determines the dynamic evolution of the state vector. The $r \times 1$ disturbance vector η_t for the state vector update has zero mean and $r \times r$ variance matrix Q_t. The observation and state disturbances ε_t and η_t are assumed to be serially independent and independent of each other at all time points. In many standard cases, $r = m$ and matrix R_t is the identity matrix I_m. In other cases, matrix R_t is an $m \times r$ selection matrix with $r < m$. Although matrix R_t can be specified freely, it is often composed of a selection from the first r columns of the identity matrix I_m.

The initial state vector α_1 is assumed to be generated as

$$\alpha_1 \sim \mathrm{NID}(a_1,P_1),$$

independently of the observation and state disturbances ε_t and η_t. Mean a_1 and variance P_1 can be treated as given and known in almost all stationary processes for the state vector. For nonstationary processes and regression effects in the state vector, the associated elements in the initial mean a_1 can be treated as unknown and need to be estimated. For an extensive discussion of *initialisation* in state space analysis, we refer to Durbin and Koopman (2001, Chapter 5).

6.2.1 Local Level Model and Other Unobserved Component Models

By appropriate choices of the vectors α_t, ε_t and η_t, and of the matrices Z_t, T_t, H_t, R_t and Q_t, a wide range of different time series models can be derived from (6.1) and (6.2). Here we discuss the class of *unobserved components time series models*. A number of special cases will be discussed in some detail. Special attention is given to the univariate *local level model*.

Letting

$$\alpha_t = \mu_t, \quad \eta_t = \xi_t, \quad Z_t = T_t = R_t = 1, \quad H_t = \sigma_\varepsilon^2, \quad Q_t = \sigma_\xi^2,$$

(all of order 1×1) for $t = 1, \ldots, n$, model (6.1)-(6.2) reduces to the local level model as given by

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, & \varepsilon_t &\sim \text{NID}(0, \sigma_\varepsilon^2), \\ \mu_{t+1} &= \mu_t + \xi_t, & \xi_t &\sim \text{NID}(0, \sigma_\xi^2), \end{aligned} \tag{6.3}$$

for $t = 1, \ldots, n$. The level component μ_t can be conceived of as the equivalent of the intercept in the classical linear regression model $y_t = \mu + \varepsilon_t$ which is obtained by setting all the level disturbances ξ_t in (6.3) equal to zero and with $\mu = \mu_1$. The key difference is that the intercept μ in a regression model is fixed whereas the level component μ_t in (6.3) is allowed to change from time point to time point.

Since the second equation in (6.3) defines a random walk, the local level model is also referred to as the random walk plus noise model (where the noise refers to the irregular component ε_t). It can be shown that the dynamic process for $x_t = y_{t+1} - y_t = \eta_t + \varepsilon_{t+1} - \varepsilon_t$, for $t = 1, \ldots, n$, reduces to the moving average process $x_t = \varepsilon_t + \theta \varepsilon_{t-1}$ where θ relates to the *signal-to-noise ratio* $q = \sigma_\xi^2 / \sigma_\varepsilon^2$ via a quadratic function. Furthermore, the forecasting function of observations generated by the local level model is equivalent to the *exponentially weighted moving average* scheme or *exponential smoothing*.

By defining

$$\alpha_t = \begin{pmatrix} \mu_t \\ \nu_t \end{pmatrix}, \quad \eta_t = \begin{pmatrix} \xi_t \\ \zeta_t \end{pmatrix}, \quad T_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad Z_t = \begin{pmatrix} 1 & 0 \end{pmatrix},$$

$$H_t = \sigma_\varepsilon^2, \quad Q_t = \begin{bmatrix} \sigma_\xi^2 & 0 \\ 0 & \sigma_\zeta^2 \end{bmatrix}, \quad \text{and} \quad R_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

the scalar notation of (6.1) and (6.2) leads to

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, & \varepsilon_t &\sim \text{NID}(0, \sigma_\varepsilon^2), \\ \mu_{t+1} &= \mu_t + \nu_t + \xi_t, & \xi_t &\sim \text{NID}(0, \sigma_\xi^2), \\ \nu_{t+1} &= \nu_t + \zeta_t, & \zeta_t &\sim \text{NID}(0, \sigma_\zeta^2), \end{aligned} \tag{6.4}$$

for $t = 1, \ldots, n$, and we obtain the *local linear trend model*.

The local linear trend model requires a 2×1 state vector α_t: one element for the level component μ_t and one element for the slope component ν_t. The slope component can be conceived of as the equivalent of the regression coefficient in the classical regression model where the observed time series y_t is regressed on the independent variable time t: $y_t = \mu + \nu t + \varepsilon_t$ with $\mu = \mu_1$ and $\nu = \nu_1$. Again, the important difference is that the regression coefficient or weight ν is fixed in classical linear regression, whereas the slope ν_t in the local linear trend model is allowed to change over time.

In the situation that the observed time series consists of quarterly or monthly data, for example, the local level and the local linear trend model can be extended with a stochastic seasonal dummy component denoted here by γ_t. In the case of a quarterly time series (the seasonal length is 4), by defining

$$\alpha_t = \begin{pmatrix} \mu_t \\ \gamma_{1,t} \\ \gamma_{2,t} \\ \gamma_{3,t} \end{pmatrix}, \quad \eta_t = \begin{pmatrix} \xi_t \\ \omega_t \end{pmatrix}, \quad T_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad Z_t = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix},$$

$$H_t = \sigma_\varepsilon^2, \qquad Q_t = \begin{bmatrix} \sigma_\xi^2 & 0 & 0 & 0 \\ 0 & \sigma_\omega^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad R_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

and expanding (6.1) and (6.2) in scalar notation, we obtain

$$\begin{aligned} y_t &= \mu_t + \gamma_{1,t} + \varepsilon_t, & \varepsilon_t &\sim \mathrm{NID}(0, \sigma_\varepsilon^2), \\ \mu_{t+1} &= \mu_t + \xi_t, & \xi_t &\sim \mathrm{NID}(0, \sigma_\xi^2), \\ \gamma_{1,t+1} &= -\gamma_{1,t} - \gamma_{2,t} - \gamma_{3,t} + \omega_t, & \omega_t &\sim \mathrm{NID}(0, \sigma_\omega^2), \\ \gamma_{2,t+1} &= \gamma_{1,t}, \\ \gamma_{3,t+1} &= \gamma_{2,t}, \end{aligned} \tag{6.5}$$

for $t = 1, \ldots, n$, which is a local level and dummy seasonal model for a quarterly time series where the seasonal component is allowed to change over time. The seasonal dummy model is not the only approach to incorporate time-varying seasonal effects in unobserved components time series model. For example, the trigonometric seasonal can also be considered. For details about such alternative specifications of the seasonal we refer to Harvey (1989) and Durbin and Koopman (2001).

The textbook of Harvey (1989) was instrumental in the dissemination of state space models outside the field of control engineering. When a slope component is included in (6.5) as well, Harvey calls this model the *basic structural time series model*. A typical application of this model is for the *seasonal adjustment* of time series. A seasonally adjusted time series is defined in this context simply by $\hat{y}_t - \gamma_t$ for $t = 1, \ldots, n$.

Another extension is to include one or more cycles to any of the special models within the class of unobserved components time series models. By defining

$$\alpha_t = \begin{pmatrix} \mu_t \\ c_t \\ c_t^* \end{pmatrix}, \quad \eta_t = \begin{pmatrix} \xi_t \\ \kappa_t \\ \kappa_t^* \end{pmatrix}, \quad T_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \rho\cos(\lambda_c) & \rho\sin(\lambda_c) \\ 0 & -\rho\sin(\lambda_c) & \rho\cos(\lambda_c) \end{bmatrix}, \quad Z_t = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix},$$

$$H_t = \sigma_\varepsilon^2, \quad Q_t = \begin{bmatrix} \sigma_\xi^2 & 0 & 0 \\ 0 & \sigma_c^2(1-\rho^2) & 0 \\ 0 & 0 & \sigma_c^2(1-\rho^2) \end{bmatrix}, \quad \text{and} \quad R_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

in (6.1) and (6.2), we obtain the following local level plus cycle model as given by

$$\begin{aligned} y_t &= \mu_t + c_t + \varepsilon_t, & \varepsilon_t &\sim \text{NID}(0, \sigma_\varepsilon^2), \\ \mu_{t+1} &= \mu_t + \xi_t, & \xi_t &\sim \text{NID}(0, \sigma_\xi^2), \\ c_{t+1} &= \rho[\cos(\lambda_c)c_t + \sin(\lambda_c)c_t^*] + \kappa_t, & \kappa_t &\sim \text{NID}(0, \sigma_c^2(1-\rho^2)), \\ c_{t+1}^* &= \rho[-\sin(\lambda_c)c_t + \cos(\lambda_c)c_t^*] + \kappa_t^*, & \kappa_t^* &\sim \text{NID}(0, \sigma_c^2(1-\rho^2)), \end{aligned} \tag{6.6}$$

for $t = 1, \ldots, n$, where $0 < \rho \leq 1$ is the *damping factor* and λ_c is the frequency of the cycle measured in radians so that $2\pi / \lambda_c$ is the *period* of the cycle. In case $\rho = 1$, the cycle reduces to a fixed sine-cosine wave but the component is still stochastic since the initial values c_1 and c_1^* are stochastic variables with mean zero and variance σ_c^2. A typical application of this model is for the signal extraction of *business cycles* from macro-economic time series.

6.2.2 Regression and Intervention Effects

Another extension of the local level and local linear trend models concerns the incorporation of fixed explanatory and intervention variables. In the case of one regression variable x_t and one intervention variable w_t, for example, we have $y_t = \mu_t + \beta x_t + \lambda w_t + \varepsilon_t$ for the local level model and a state vector of three elements is required: one for the level component μ_t, one for the regression coefficient β, and one for the regression coefficient λ. The substitution of

$$\alpha_t = \begin{pmatrix} \mu_t \\ \beta_t \\ \lambda_t \end{pmatrix}, \quad \eta_t = \xi_t, \quad T_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad Z_t = \begin{pmatrix} 1 & x_t & w_t \end{pmatrix},$$

$$H_t = \sigma_\varepsilon^2, \qquad Q_t = \begin{bmatrix} \sigma_\xi^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad R_t = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

in (6.1) and (6.2) results in

$$\begin{aligned} y_t &= \mu_t + \beta_t x_t + \lambda_t w_t + \varepsilon_t, & \varepsilon_t &\sim \text{NID}(0, \sigma_\varepsilon^2), \\ \mu_{t+1} &= \mu_t + \xi_t, & \xi_t &\sim \text{NID}(0, \sigma_\xi^2), \\ \beta_{t+1} &= \beta_t, \\ \lambda_{t+1} &= \lambda_t, \end{aligned} \tag{6.7}$$

where $\beta = \beta_1 = \beta_t$ and $\lambda = \lambda_1 = \lambda_t$ for $t = 1, \ldots, n$. This is the local level model with one continuous explanatory variable x_t and one intervention variable w_t. By adding disturbance terms to the state equation for β_t in (6.7), this regression weight is effectively subjected to a random walk, thus allowing for the estimation of time-varying regression effects.

Letting τ denote the time point at which an intervention effect occurred, variable w_t can either be coded as a pulse intervention:

$$w_t = \begin{cases} 0, & t < \tau, \quad t > \tau \\ 1, & t = \tau \end{cases}$$

(to model an outlier observation), or as a level intervention:

$$w_t = \begin{cases} 0, & t < \tau, \\ 1, & t \geq \tau, \end{cases}$$

(to model a structural break in the level of the series), or as a slope intervention:

$$w_t = \begin{cases} 0, & t < \tau, \\ 1 + t - \tau, & t \geq \tau, \end{cases}$$

(to model a structural break in the slope of the series). Other types of intervention effects can be modelled as well, see Box and Tiao (1975).

6.2.3 *Structural Time Series Model*

What emerges – and this a key advantage of state space methods – is their *structural* approach to time series analysis: the different unobserved components or building blocks responsible for the dynamics of the series such as trend, seasonal, cycle, and the effects of explanatory and intervention variables are identified separately before being put together in a state space model. It is the responsibility of the researcher to decide what components are required in a specific situation, and then to consider whether they apply to the time series under consideration. This explains why state space models are also known as *structural time series models*.

6.2.4 Multivariate Models

All this is easily extended to multivariate time series. For example, letting y_t denote a $p \times 1$ vector of observations, a multivariate local level model can be applied to the p time series simultaneously:

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, & \varepsilon_t &\sim \text{NID}(0, \Sigma_\varepsilon), \\ \mu_{t+1} &= \mu_t + \xi_t, & \xi_t &\sim \text{NID}(0, \Sigma_\xi), \end{aligned} \tag{6.8}$$

for $t = 1, \ldots, n$, where μ_t, ε_t, and ξ_t are $p \times 1$ vectors and Σ_ε and Σ_ξ are $p \times p$ variance matrices. In what is known as the *seemingly unrelated time series equations model* (6.8), the series are modelled as in the univariate situation, but the disturbances driving the level components are allowed to be instantaneously correlated across the p series. When slope, seasonal, or cycle components are involved, each of these three components also has an associated $p \times p$ variance matrix allowing for correlated disturbances across series.

If it is found that the rank r of Σ_ξ in (6.8) is smaller than p, then this indicates that the p series have *r common levels*. Such common factors may not only have a nice interpretation, but may also result in more efficient inferences and forecasts.

6.3 State Space Analysis

For given values of all system matrices – and for given initial conditions a_1 and P_1 – the state vector can be estimated in three different ways, yielding what are known as the *filtered*, the *predicted*, and the *smoothed* state vector. Depending on the types of state estimates required in the analysis, the estimates of the state vector can be obtained by performing one or two passes through the observed time series:

1. a *forward* pass, from $t = 1, \ldots, n$, using a recursive algorithm known as the *Kalman filter* enables the computation of filtered and predicted states and prediction errors;
2. a *backward* pass, from $t = n, \ldots, 1$, using output of the Kalman filter and using recursive algorithms known as *state and disturbance smoothers* enables the computation of smoothed estimates of states and disturbances.

6.3.1 Kalman Filter for Prediction, Filtering and Forecasting

The forward pass through the data with the well-known Kalman (1960) filter provides all estimates that are relevant for the filtered and the predicted state. The main purpose of the Kalman filter is to obtain optimal estimates of the state at time point t, *only* considering the observations $\{y_1, y_2, \ldots, y_{t-1}\}$. A key property of the predicted state and its related estimates is therefore that they are only based on *past values* of the observed time series. The recursive formulas for the Kalman filter are

$$\begin{aligned} v_t &= y_t - Z_t a_t, & F_t &= Z_t P_t Z_t' + H_t, \\ K_t &= T_t P_t Z_t' F_t^{-1}, & L_t &= T_t - K_t Z_t, \\ a_{t+1} &= T_t a_t + K_t v_t, & P_{t+1} &= T_t P_t L_t' + R_t Q_t R_t', \end{aligned} \tag{6.9}$$

for $t = 1,\dots,n$. The values of a_t in (6.9) represent the predicted state, while the values of P_t quantify the estimation error variance matrix of the predicted state a_t. Under the assumption of normality, the latter variances are useful for the construction of *confidence intervals* for the predicted state, which – assuming that we are interested in their 90% confidence limits for example – can be calculated as

$$a_t \pm 1.64\sqrt{P_t},$$

for $t = 1,\dots,n$. A modification of the Kalman filter also allows the computation of the filtered estimate of the state vectors, that is

$$a_{t|t} = a_t + P_t Z_t' F_t^{-1} v_t, \qquad P_{t|t} = P_t - P_t Z_t' F_t^{-1} Z_t P_t, \qquad t = 1,\dots,n,$$

where $a_{t|t}$ is the optimal estimate of the state at time point t by considering the observations $\{y_1, y_2, \dots, y_t\}$ while $P_{t|t}$ is the state filtered estimation error variance matrix. The values of v_t in (6.9) are called the *one-step ahead prediction* or *forecast* errors, since they quantify the lack of accuracy of a_t in predicting the observed value of y_t at time point t; the values of F_t are the variances of these one-step ahead prediction errors v_t.

One of the convenient features of state space methods is the ease with which they deal with two important aspects of time series analysis – forecasting and missing observations: by treating them in exactly the same way. Missing observations are handled by setting K_t and v_t in (6.9) equal to 0. Forecasts for $y_{n+1},\dots,y_{n+k}$ given $y_1,\dots,y_n$ are simply obtained by applying the Kalman filter for $t = 1,\dots,n,n+1,\dots,n+k$ and by treating $y_{n+1},\dots,y_{n+k}$ as missing observations.

6.3.2 State and Disturbance Smoothing

The backward pass through the data is only required for smoothing that leads to estimates such as the smoothed states and smoothed disturbances. The main purpose of state and disturbance smoothing is to obtain estimated values of the state and disturbance vectors at time point t, considering *all* available observations $\{y_1, y_2, \dots, y_n\}$. The recursive formulas for state smoothing are

$$r_{t-1} = Z_t' F_t^{-1} v_t + Z_t' r_t, \qquad N_{t-1} = Z_t' F_t^{-1} Z_t + L_t' N_t L_t, \tag{6.10}$$

$$\hat{\alpha}_t = a_t + P_t r_{t-1}, \qquad V_t = P_t - P_t N_{t-1} P_t, \tag{6.11}$$

for $t = n, \ldots, 1$. The recursive formulas for smoothing (6.10) are initialised with $r_n = 0$ and $N_n = 0$. The state smoothing equations (6.11) yield the smoothed state estimate $\hat{\alpha}_t$ and is defined as the optimal estimate of α_t using the full set of observations $\{y_1, y_2, \ldots, y_n\}$; the state smoothing equations also yield the corresponding smoothed state estimation error variance matrix V_t.

Analogous to the predicted state, under the assumption of normality the smoothed state estimation error variance matrix V_t is useful for the construction of confidence intervals for the smoothed state components, which – should we happen to be interested in their 90% confidence limits for example – can be calculated as

$$\hat{\alpha}_t \pm 1.64\sqrt{V_t},$$

for $t = 1, \ldots, n$.

The recursions for r_{t-1} and N_{t-1} in (6.10) also enable the computation of the smoothed estimates of the disturbances ε_t and η_t in the following way,

$$\hat{\varepsilon}_t = H_t \left(F_t^{-1} v_t - K_t' r_t\right), \qquad \mathrm{Var}(\hat{\varepsilon}_t) = H_t \left(F_t^{-1} + K_t' N_t K_t\right) H_t, \tag{6.12}$$

$$\hat{\eta}_t = Q_t R_t' r_t, \qquad \mathrm{Var}(\hat{\eta}_t) = Q_t R_t' N_t R_t Q_t, \tag{6.13}$$

for $t = n, \ldots, 1$. The equations (6.12) and (6.13) compute the smoothed observation disturbances $\hat{\varepsilon}_t$, the smoothed state disturbances $\hat{\eta}_t$, and their corresponding smoothed estimation error variance matrices $\mathrm{Var}(\hat{\varepsilon}_t)$ and $\mathrm{Var}(\hat{\eta}_t)$.

6.3.3 Diagnostic Checking

All significance tests in linear Gaussian state space models – and the construction of confidence intervals – are based on three assumptions concerning the residuals of the analysis. The residuals should satisfy independence, homoscedasticity, and normality, in this order of importance. Whether the residuals satisfy these three assumptions can be established by diagnosing what are known as the *standardised prediction errors*. They are defined as

$$\frac{v_t}{\sqrt{F_t}}, \tag{6.14}$$

for $t = 1, \ldots, n$. For the computations of the one step-ahead prediction errors v_t and their variances F_t in (6.14), we refer to the recursive formulas for the Kalman filter given in (6.9). The assumptions of independence and normality of the residuals can be diagnosed using the Box-Ljung test statistic and the Bowman and Shenton test statistic, respectively. The assumption of homoscedasticity can be checked by testing whether the variance of the standardised prediction errors in the first third part of the series is equal to the variance of the errors corresponding to the last third part of the series. For further details concerning these diagnostic tests, we refer to Harvey (1989), Durbin and Koopman (2001) and Commandeur and Koopman (2007).

A second diagnostic tool for determining the appropriateness of a model is provided by inspection of what are known as the *auxiliary residuals*. As already mentioned above, the disturbance smoothing filters applied in the backward pass through the data yield, amongst others, estimates of the smoothed observation and state disturbances, and of their variances. The auxiliary residuals are obtained by dividing the smoothed observation and state disturbances with the square root of their corresponding variances, as follows:

$$\frac{\hat{\varepsilon}_t}{\sqrt{\mathrm{Var}(\hat{\varepsilon}_t)}}, \text{ and } \frac{\hat{\eta}_t}{\sqrt{\mathrm{Var}(\hat{\eta}_t)}}, \tag{6.15}$$

for $t = 1, \ldots, n$, resulting in *standardised* smoothed disturbances. Inspection of the standardised smoothed observation disturbances (shown at the left of (6.15)) allows for the detection of possible *outlier* observations in a time series, while inspection of the standardised smoothed state disturbances (shown at the right of (6.15)) makes it possible to detect *structural breaks* in the underlying development of a time series.

Each auxiliary residuals can be considered as a t-test for the null hypothesis that there was no outlier observation (when inspecting the auxiliary residuals at the left of (6.15)) or as a t-test for the null hypothesis that there was no structural break in the corresponding unobserved component of the observed time series (when inspecting the auxiliary residuals at the right of (6.15)). Applying the usual 95% confidence limits of ± 1.96 corresponding to a two-tailed t-test, possible outlier observations or structural breaks in the unobserved components making up the state vector are thus easily detected.

6.3.4 Parameter Estimation

So far, we have presented all of the results that can be obtained with state space methods as if the disturbance variances, the fixed regression effects, the parameters ρ and λ_c associated with cycles, etcetera, are given and known. In practice, of course, these parameters are unknown, and have to be estimated.

It can be shown that the Kalman filter presented in (6.9) also provides the necessary ingredients required for evaluating the log-likelihood function, which is given by

$$\log L(y|\psi) = -\frac{np}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\left(\log|F_t| + v_t' F_t^{-1} v_t\right), \tag{6.16}$$

where the v_t are the one-step ahead prediction errors, the F_t are their variances for $t = 1, \ldots, n$ defined in (6.9), and ψ denotes the vector of unknown parameters. The log-likelihood (6.16) is maximised with respect to ψ numerically using the score vector or the EM algorithm.

Numerical quasi-Newton methods for likelihood maximization such as the one of Broyden-Fletcher-Goldfarb-Shanno (BFGS) are generally regarded as computationally efficient in terms of convergence speed and numerical stability, see also the book of Nocedal and Wright (1999). The BFGS iterative optimization method is based on information from the gradient and terminated when some pre-chosen convergence criterion is satisfied. The convergence criterion is usually based on the gradient evaluated at the current estimate, the parameter change compared to the previous estimate or the likelihood value change compared to the previous estimate. The number of iterations required to satisfy these criteria depends on the choice of the initial parameter values, the tightness of the chosen criterion and the shape of the likelihood surface.

Several problems may arise when maximizing the likelihood function with respect to the parameter vector of a high dimension. For example, the number of required iterations may be too large for a feasible procedure, different initial parameter values and different convergence criteria may lead to different estimates. Also, flat likelihood surfaces may not allow the optimization procedure to converge.

An alternative method for computing ML estimates is the use of the EM-algorithm. The EM-algorithm is not an alternative to ML, but it is an alternative way to obtain the ML estimates. We may compare the different estimation methods in terms of required calculation time. The EM algorithm in the setting of a state space model was developed by Shumway and Stoffer (1982) and Watson and Engle (1983). The basic EM procedure works roughly as follows. Consider the joint density $p(y_1,\ldots,y_n,\alpha_1,\ldots,\alpha_n)$. The Expectation (E) step takes the expectation of the components of the joint density conditional on $y_1,\ldots,y_n$ and maximizes the resulting expression with respect to ψ. The E step mainly consists of evaluating the estimated state vector using state space smoothing algorithms. The next step is the Maximization (M) step which usually can be done analytically and is simpler than maximizing the full likelihood function directly. Given the "new" estimate from the M step, we can go back to the E step and evaluate the smoothed estimates based on the new estimate. This iterative procedure converges to the ML estimate of ψ. Under fairly weak conditions it can be proven that each iteration of the EM algorithm only increases the value of the likelihood, and that the EM estimate converges to a maximum of the likelihood. The algorithm has similar properties as a well chosen numerical ML algorithm.

6.4 An Illustration of Multivariate State Space Analysis

In this section we present the practical implications of a multivariate state space analysis. Various results of a simultaneous analysis of two time series will be discussed in some detail. In February 1983 a law was introduced in the United Kingdom (UK) obligating front seat passengers in cars (including the driver) to wear a seat belt. In Durbin and Koopman (2001) and Commandeur and Koopman (2007) the effect of this law was investigated by applying a bivariate local level with seasonal model to the log of the monthly numbers of *front* seat passengers killed or seriously injured (KSI) in cars and to the log of the monthly numbers of *rear* car seat passengers KSI, but only for the period 1969 – 1984 (thus yielding a total of $12 \times 16 = 192$ observations per series). These series were analysed previously with univariate state space models in Harvey and Durbin (1986).

In these studies the numbers of UK front seat passengers KSI in cars were treated as a *treatment* series, while the UK rear seat passengers KSI in cars were used as a *control* series, based on the assumption that the rear seat passengers KSI in cars were not affected by the introduction of this seat belt law. It was indeed found that the seat belt law resulted in a significant 28.4% to 30.5% decrease in the number of front seat passengers KSI in cars, but did not affect the number of UK rear car seat passengers KSI.

In this section we re-investigate the effect of the introduction of this law, but now applied to the same two series supplemented with monthly observations for the years 1985 – 2007, resulting in a total of $12 \times 39 = 468$ observations per series. The logs of the two series are displayed in Figure 6.1.

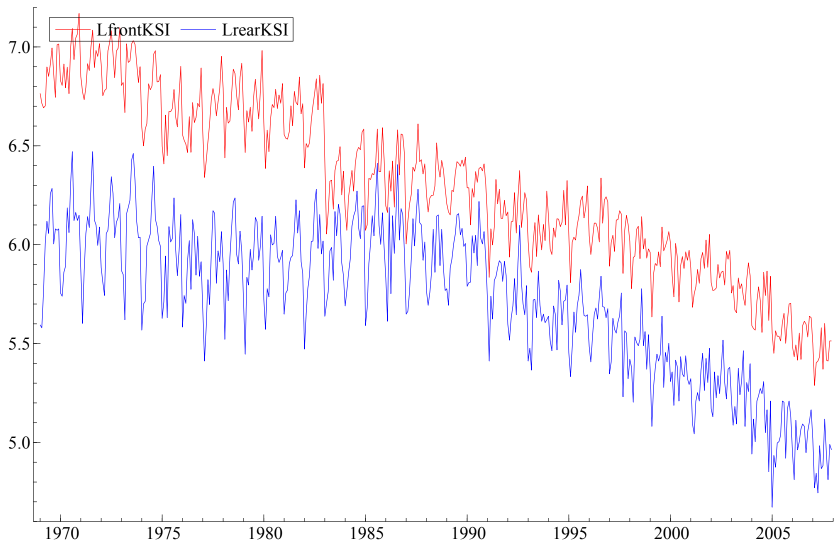

Fig. 6.1 Log of monthly numbers of front seat passengers (top) and rear seat passengers (bottom) in cars killed or seriously injured in the UK in the period 1969–2007.

These extended series not only make it possible to confirm or falsify the value and significance of the effect of the February 1983 seat belt law on front seat passengers in cars previously found in Durbin and Koopman (2001) and Commandeur and Koopman (2007) for the monthly 1969 – 1984 series, but also to investigate the effects of the introduction of two other seat belt laws in the UK: the obligation for children in the rear seat of cars to wear a seat belt in September 1989, and for adults in the rear seat of cars to wear a seat belt in July 1991. In the evaluation of the effects of the latter two laws it is typically the monthly number of rear seat passengers KSI that act as a treatment series while the monthly number of front seat passengers KSI can now be used as a control series.

All the analyses discussed in this chapter were performed in `STAMP 8` of Koopman, Harvey, Doornik, and Shephard (2007). `STAMP 8` is an easy-to-use package designed to model and forecast time series, based on uni- and multivariate structural time series models. No coding is required because all the models are simply formulated by clicking options in dialog windows. Other software packages that currently have functions for analysing time series with state space methods (but with a programmatic interface) include `SsfPack`, `R`, `Matlab`, `Eviews`, `Gauss`, `Stata`, `SAS`, `RATS`, and `Gretl`.

We start by adding three intervention variables to a bivariate local linear trend with monthly seasonal model applied to both series (in logs). These intervention variables are: the introduction of the seat belt law for car drivers and front seat car passengers in February 1983, the introduction of the seat belt law for children in the rear seat of cars in September 1989, and the introduction of the seat belt law for adults in the rear seat of cars in July 1991, all applied to both series simultaneously.

The bivariate time series analysis aims to assess the effects of the introduction of these three seat belt laws in the UK. The intervention of February 1983 is expected to affect the car drivers and front seat car passengers only, and not the rear seat car passengers. In contrast, the interventions of September 1989 and July 1991 are expected to affect the rear seat car passengers only, and not the car drivers and front seat car passengers. As we already mentioned, the car drivers and front seat car passengers series can be considered as a treatment series for the evaluation of the February 1983 intervention, while the rear seat car passengers series can be used as a control series in this case. For the evaluation of the seat belt laws implemented in September 1989 and July 1991, on the other hand, the reverse holds true: in that case it is the car drivers and front seat car passengers series that takes on the role of a control series, while the rear seat car passengers series can be used as a treatment series in these two cases.

The residual and fit diagnostics of this analysis are as follows:

```
Summary statistics
                 LfrontKSI     LrearKSI
 T                  468.00       468.00
 p                  3.0000       3.0000
 std.error        0.084885      0.10540
 Normality          1.9352       10.135
 H(150)            0.82104      0.91225
 DW                 1.9910       2.1078
 r(1)            0.0013231    -0.058876
 q                  24.000       24.000
 r(q)            -0.029281    -0.078653
 Q(q,q-p)           43.697       37.213
 Rs^2              0.39786      0.43512
```

The Box-Ljung diagnostic tests for the independence of residuals for the front and rear seat passengers KSI series are $Q(21) = 43.697$ and $Q(21) = 37.213$, respectively. Since these should be tested against $\chi^2_{(21;0.05)} = 32.6705$, the residuals of both series are somewhat serially correlated. The tests for homoscedasticity of the residuals for the front and rear seat passengers KSI series are equal to $H(150) = 0.82104$ and $H(150) = 0.91225$, respectively. Since $F_{(150,150;0.025)} \approx 1.43$, and $1/H(150) = 1.22$ and $1/H(150) = 1.10$, the assumption of homoscedasticity is satisfied for both series. The Bowman-Shenton diagnostic tests for normality of the residuals are $N = 1.9117$ and $N = 13.679$, respectively, implying that the assumption of normality is only satisfied for the front seat passengers KSI series. This is not something to worry about very much since we are dealing with 468 observations. The values of the Akaike Information Criterion (AIC) for the two series are -4.8603 and -4.4273, respectively.

The estimates of the variance matrices (where the upper off-diagonal elements denote correlations) for this bivariate state space model are:

```
Level disturbance variance matrix:       Slope disturbance variance matrix:
            LfrontKSI    LrearKSI                    LfrontKSI    LrearKSI
LfrontKSI   0.0002752      0.8798        LfrontKSI  2.249e-008       1.000
LrearKSI    0.0002047   0.0001967        LrearKSI   3.329e-008  4.927e-008

Seasonal disturbance variance matrix:    Irregular disturbance variance matrix:
            LfrontKSI    LrearKSI                    LfrontKSI    LrearKSI
LfrontKSI  7.080e-007      0.8030        LfrontKSI    0.005460      0.5935
LrearKSI   1.186e-006  3.082e-006        LrearKSI     0.004033    0.008459
```

The t-tests for the regression weights of the three level shift intervention variables are:

```
Equation LfrontKSI: regression effects in final state at time 2007(12)

                    Coefficient         RMSE      t-value       Prob
Level break 1983(2)    -0.33634      0.05107     -6.58646 [0.00000]
Level break 1989(9)     0.04346      0.05108      0.85077 [0.39535]
Level break 1991(7)    -0.03793      0.05108     -0.74260 [0.45811]

Equation LrearKSI: regression effects in final state at time 2007(12)

                    Coefficient         RMSE      t-value       Prob
Level break 1983(2)     0.02321      0.05208      0.44564 [0.65607]
Level break 1989(9)     0.05752      0.05208      1.10445 [0.26999]
Level break 1991(7)    -0.06484      0.05206     -1.24556 [0.21357]
```

These t-tests indicate that the regression coefficient for the February 1983 level shift intervention variable applied to the front seat passengers KSI series is very significant, unlike any of the other five intervention variables. The estimated regression coefficient for the February 1983 level shift intervention variable on front seat passengers KSI is -0.33634, implying a $100 \times (\exp(-0.33634) - 1) = -28.56\%$ change in the number of front seat passengers KSI due to the introduction of this seat belt law in the UK.

Although the disturbance variances of the two slope components for both series are quite small, we decide to keep the slope components in all further multivariate analyses of these two series because the values of these components in December 2007 are found to significantly deviate from zero:

```
Equation LfrontKSI
                              Value      Prob
Slope                      -0.00395 [0.00486]

Equation LrearKSI
Slope                      -0.00476 [0.01114]
```

We now present the results of the same analysis after removing the five non-significant level shift intervention variables from the previous model. The residual and fit diagnostics are:

```
Summary statistics
                LfrontKSI    LrearKSI
 T                 468.00      467.00
 p                 3.0000      3.0000
 std.error       0.085015     0.10567
 Normality         1.8309      9.5254
 H(151)           0.81288     0.91048
 DW                1.9886      2.1155
 r(1)           0.0021785   -0.063392
 q                 24.000      24.000
 r(q)           -0.032486   -0.072441
 Q(q,q-p)          43.508      34.964
 Rs^2             0.39334     0.42977
```

The Box-Ljung diagnostic tests for the independence of the residuals for the front and rear seat passengers KSI series in this analysis are $Q(21) = 43.508$ and $Q(21) = 34.964$, respectively. The residuals of both series are therefore still serially correlated, although to a somewhat lesser extent than in the previous analysis. The tests for homoscedasticity of the residuals for the front and rear seat passengers KSI series for this analysis are equal to $H(151) = 0.81288$ and $H(151) = 0.91048$, respectively. Since $F_{(151,151;0.025)} \approx 1.43$, and $1/H(151) = 1.23$ and $1/H(151) = 1.10$, the assumption of homoscedasticity is still satisfied for both series. The Bowman-Shenton diagnostic tests for normality of the residuals are now $N = 1.8309$ and $N = 9.5254$, respectively, meaning that the assumption of normality is still only satisfied for the front seat passengers KSI series. Again, this is not something to worry about very much due to the large amount of observations in this data set. The AIC for the two series are now -4.8658 and -4.4351, respectively, indicating a better fit than in the previous analysis.

The estimates of the variance matrices (where the upper off-diagonal elements again denote correlations) for this analysis are:

```
Level disturbance variance matrix:      Slope disturbance variance matrix:
            LfrontKSI    LrearKSI                   LfrontKSI    LrearKSI
LfrontKSI   0.0002708      0.8734       LfrontKSI  2.408e-008       1.000
LrearKSI    0.0002110   0.0002155       LrearKSI   3.581e-008  5.325e-008

Seasonal disturbance variance matrix:   Irregular disturbance variance matrix:
            LfrontKSI    LrearKSI                   LfrontKSI    LrearKSI
LfrontKSI  7.038e-007      0.8051       LfrontKSI    0.005457      0.5927
LrearKSI   1.190e-006  3.105e-006       LrearKSI     0.004005    0.008368
```

With a value of -10.55 for the t-test, the estimated regression coefficient for the February 1983 level shift intervention variable applied to the front seat passengers KSI series is now -0.35120, implying a significant $100 \times (\exp(-0.35120) - 1) = -29.62\%$ level change in the number of front seat passengers KSI.

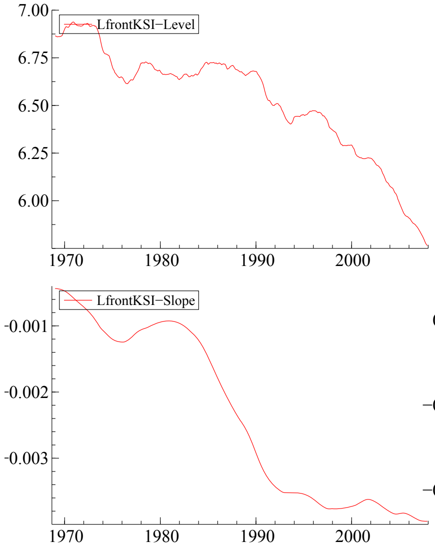

Fig. 6.2 Levels and slope components of full rank model for monthly numbers of front seat passengers (left) and rear seat passengers (right) killed or seriously injured in the UK in the period 1969–2007.

There is a perfect correlation between the slope disturbances of the two series, probably due to their very small variances. However, the just mentioned variance matrix of the level disturbances indicates that the level disturbances are also quite highly correlated. This is confirmed by the following eigenvalue decompositions of the level and slope disturbance variance matrices:

```
Analysis of variance matrices
Level disturbance variance matrix is 2 x 2 with imposed rank 2 and actual rank 2
Variance/correlation matrix
            LfrontKSI    LrearKSI
LfrontKSI   0.0002708      0.8734
LrearKSI    0.0002110   0.0002155
Cholesky decomposition LDL' with L and D
            LfrontKSI    LrearKSI
LfrontKSI       1.000      0.0000
LrearKSI       0.7791       1.000
diag(D)     0.0002708   5.112e-005
Eigenvectors and eigenvalues
              LfrontKSI    LrearKSI
LfrontKSI        0.7517      0.6596
LrearKSI         0.6596     -0.7517
eigenvalues   0.0004560  3.036e-005
percentage        93.76       6.243

Slope disturbance variance matrix is 2 x 2 with imposed rank 2 and actual rank 1
Variance/correlation matrix
            LfrontKSI    LrearKSI
LfrontKSI  2.408e-008       1.000
LrearKSI   3.581e-008  5.325e-008
```

```
Eigenvectors and eigenvalues
              LfrontKSI    LrearKSI
LfrontKSI       -0.5580      0.8298
LrearKSI        -0.8298     -0.5580
eigenvalues   7.732e-008  4.850e-020
percentage        100.0   6.272e-011
```

The first eigenvalue of the level disturbance variance matrix explains almost 94% of the variance in this matrix. This indicates that the model for the analysis of these two series could be simplified by imposing *rank one restrictions* on both these matrices, thus treating the level and slope components as *common to both series*.

We therefore repeat the analysis only applying a level shift intervention variable in February 1983 on the front seat passengers KSI series, *and* restricting the level and slope disturbance matrices to be of rank *one*. The residual and fit diagnostics of this final model are:

```
Summary statistics
                LfrontKSI    LrearKSI
 T                 468.00      467.00
 p                 3.0000      3.0000
 std.error       0.085009     0.10604
 Normality         1.7758      10.106
 H(151)           0.82200     0.94560
 DW                1.9739      2.0539
 r(1)            0.010066   -0.029112
 q                 24.000      24.000
 r(q)           -0.031552   -0.075277
 Q(q,q-p)          43.087      34.490
 Rs^2             0.39342     0.42570
```

The Box-Ljung diagnostic tests for the independence of the residuals for the front and rear seat passengers KSI series are now $Q(21) = 43.087$ and $Q(21) = 34.490$, respectively. The residuals of both series are therefore still serially correlated, although again to a somewhat lesser extent than in the previous analysis. The tests for homoscedasticity of the residuals for the front and rear seat passengers KSI series for this analysis equal $H(151) = 0.82200$ and $H(151) = 0.94560$, respectively. Since $F_{(151,151;0.025)} \approx 1.43$, and $1/H(151) = 1.22$ and $1/H(151) = 1.06$, the assumption of homoscedasticity is again satisfied for both series. The Bowman-Shenton diagnostic tests for normality of the residuals are $N = 1.7758$ and $N = 10.106$, respectively, implying that the assumption of normality is still only satisfied for the front seat passengers KSI series. The AIC for the two series are now -4.8659 and -4.428, respectively, indicating that the previous analysis results in a marginally better fit than the present one.

The estimates of the variance matrices for this last analysis are:

```
Level disturbance variance/correlation matrix:
            LfrontKSI    LrearKSI
LfrontKSI   0.0002639       1.000
LrearKSI    0.0001810   0.0001241
Level disturbance factor variance for LfrontKSI: 0.000263925
Level disturbance factor loading  for LrearKSI: 0.685819
            LfrontKSI    LrearKSI
Constant       0.0000      0.9298
```

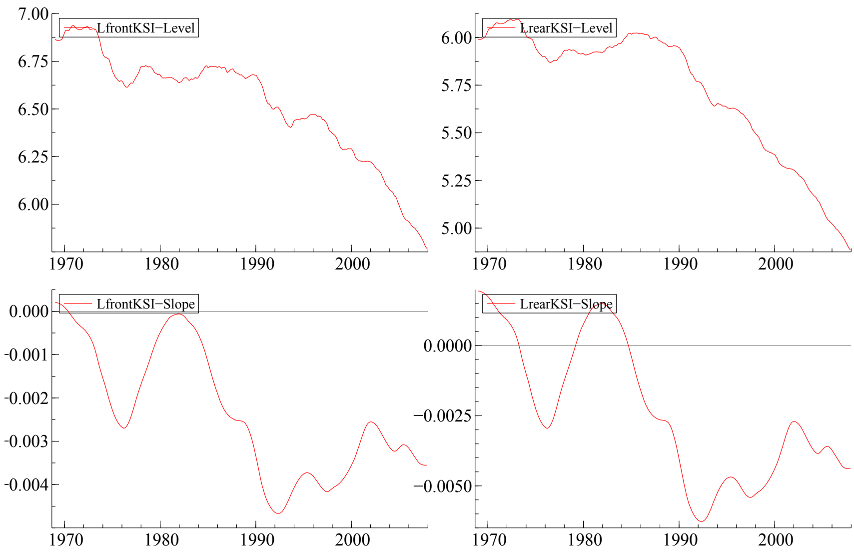

Fig. 6.3 Levels and slope components of rank one model for monthly numbers of front seat passengers (left) and rear seat passengers (right) killed or seriously injured in the UK in the period 1969–2007.

```
Slope disturbance variance/correlation matrix:
            LfrontKSI     LrearKSI
LfrontKSI  7.084e-008        1.000
LrearKSI   1.195e-007   2.014e-007
Slope disturbance factor variance for LfrontKSI: 7.08383e-008
Slope disturbance factor loading  for LrearKSI: 1.68624
            LfrontKSI     LrearKSI
Constant       0.0000     0.001602

Seasonal disturbance variance/correlation matrix:
            LfrontKSI     LrearKSI
LfrontKSI  7.048e-007       0.8072
LrearKSI   1.187e-006   3.067e-006

Irregular disturbance variance/correlation matrix:
            LfrontKSI     LrearKSI
LfrontKSI    0.005467       0.5899
LrearKSI     0.004066     0.008690
```

The t-test for the regression weight of the only level shift intervention variable is:

```
Equation LfrontKSI: regression effects in final state at time 2007(12)

                    Coefficient         RMSE     t-value       Prob
Level break 1983(2)    -0.35111      0.03124   -11.23930 [0.00000]
```

With a t-value of -11.24, the estimated regression coefficient for the February 1983 level shift intervention variable in this final analysis equals -0.35111, indicating a significant $100 \times (\exp(-0.35111) - 1) = -29.61\%$ level change in the number of front seat passengers KSI.

The most important graphical results of this final analysis are presented in Figures 6.3, 6.4, and 6.5. Figure 6.4 displays the estimated trends for the front and rear passengers series KSI series (first row in Figure 6.4), the estimated trigonometric seasonals (second row in Figure 6.4), and the corresponding irregular components (third row in Figure 6.4), while Figure 6.5 contains the correlograms of the residuals of the two series. In Figure 6.3 the common level and slope components of the two series are shown.

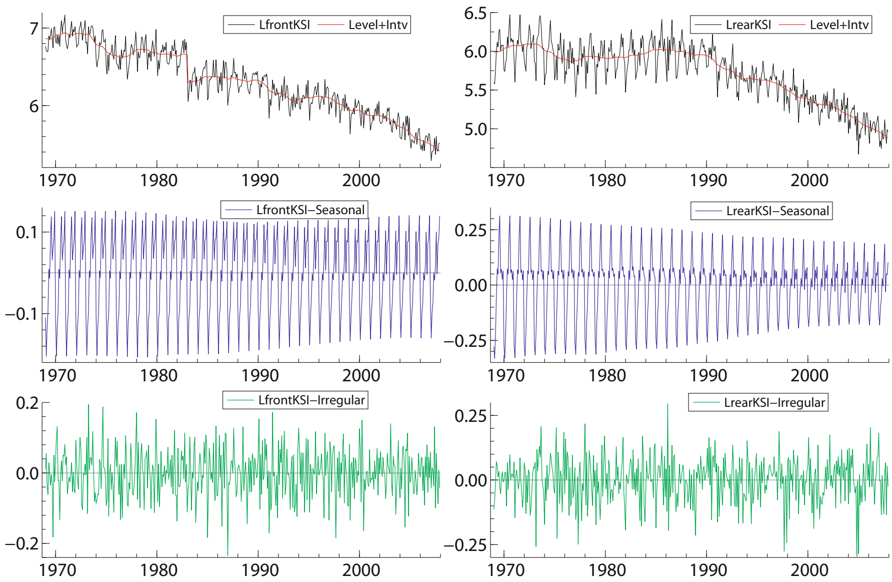

Fig. 6.4 Trends, seasonals and irregular components of rank one model for monthly numbers of front seat passengers (top) and rear seat passengers (bottom) killed or seriously injured in the UK in the period 1969–2007.

The correct implementation of the rank one restrictions is confirmed by the output of the `STAMP 8` program of Koopman, Harvey, Doornik, and Shephard (2007):

```
Level disturbance variance/correlation matrix:
            LfrontKSI    LrearKSI
LfrontKSI   0.0002639       1.000
LrearKSI    0.0001810   0.0001241
Level disturbance factor variance for LfrontKSI: 0.000263925
Level disturbance factor loading  for LrearKSI: 0.685819

Slope disturbance variance/correlation matrix:
            LfrontKSI    LrearKSI
LfrontKSI  7.084e-008       1.000
LrearKSI   1.195e-007  2.014e-007
Slope disturbance factor variance for LfrontKSI: 7.08383e-008
Slope disturbance factor loading  for LrearKSI: 1.68624
```

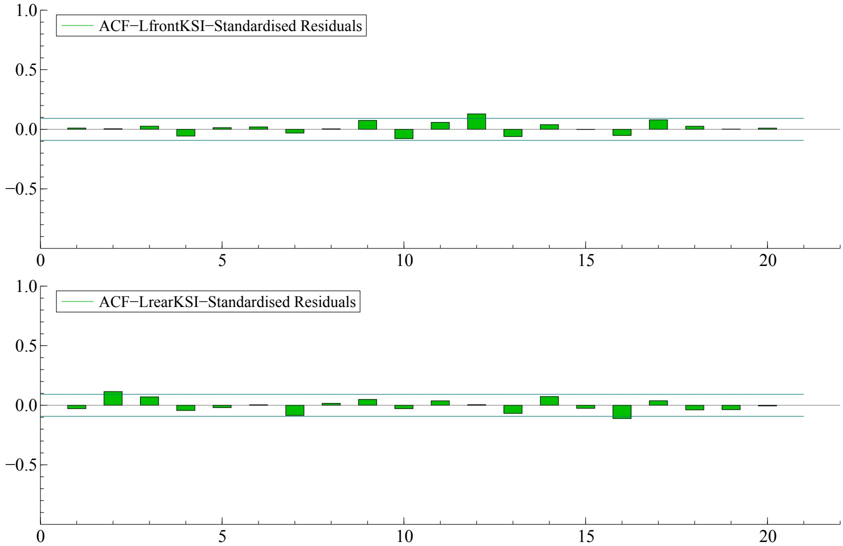

Fig. 6.5 Correlograms of rank one model for monthly numbers of front seat passengers (top) and rear seat passengers (bottom) killed or seriously injured in the UK in the period 1969–2007.

and

```
Analysis of variance matrices
Level disturbance variance matrix is 2 x 2 with imposed rank 1 and actual rank 1
Factors are determined by series LfrontKSI
Variance/correlation matrix
            LfrontKSI    LrearKSI
LfrontKSI   0.0002639       1.000
LrearKSI    0.0001810   0.0001241
Eigenvectors and eigenvalues
              LfrontKSI    LrearKSI
LfrontKSI        0.8247      0.5656
LrearKSI         0.5656     -0.8247
eigenvalues   0.0003881 -5.559e-021
percentage        100.0 -1.432e-015

Slope disturbance variance matrix is 2 x 2 with imposed rank 1 and actual rank 1
Factors are determined by series LfrontKSI
Variance/correlation matrix
            LfrontKSI    LrearKSI
LfrontKSI  7.084e-008       1.000
LrearKSI   1.195e-007  2.014e-007
Eigenvectors and eigenvalues
              LfrontKSI    LrearKSI
LfrontKSI       -0.5101      0.8601
LrearKSI        -0.8601     -0.5101
eigenvalues  2.723e-007  1.016e-023
percentage        100.0  3.731e-015
```

showing that *all* of the variation in the level and slope disturbance matrices is now explained by the first dimension, as expected. It follows that the state equations of the two level and slope components can be written as

$$\begin{aligned}
\mu_{t+1}^{(1)} &= \mu_t^{(1)} + \nu_t^{(1)} + \xi_t^{(1)},\\
\mu_t^{(2)} &= \mu_t^{(1)} + \nu_t^{(2)},\\
\nu_{t+1}^{(1)} &= \nu_t^{(1)} + \zeta_t^{(1)},\\
\nu_t^{(2)} &= 1.68624\nu_t^{(1)} + 0.001602.
\end{aligned}$$

Notwithstanding the fact that the residual diagnostic tests of the analyses presented in this section do not satisfy all of the model assumptions of independency and normality perfectly, we conclude that the impressive reduction in the UK number of front seat passengers KSI of 28.4% to 30.5% found in Durbin and Koopman (2001) and Commandeur and Koopman (2007) as a result of the introduction of the seat belt law in February 1983 is confirmed in the present analyses, even after adding 24 years of monthly observations to these time series data. However, the introduction of the UK seat belt laws for children and adults in the rear seat of cars in September 1989 and July 1991 apparently failed to have any significant impact on these types of road users.

6.5 Conclusions

We have presented an overview of uni- and multivariate state space time series analysis. An illustration of how the methodology based on state space can be implemented is given for the simultaneous analysis of two time series of traffic safety data. This account is far from complete and more details – such as how to deal with nonlinear models and non Gaussian error distributions – can be found in the references given.

References

Box, G. E. P., & Tiao, G. C. (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association, 70,* 70-79.

Commandeur, J. J. F., & Koopman, S. J. (2007). *An Introduction to State Space Time Series Analysis.* Oxford: Oxford University Press.

Durbin, J., & Koopman, S. J. (2001). *Time Series Analysis by State Space Methods.* Oxford: Oxford University Press.

Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter.* Cambridge: Cambridge University Press.

Harvey, A. C., & J. Durbin (1986). The effects of seat belt legislation on British road casualties: A case study in structural time series modelling. *Journal of the Royal Statistical Society A, 149 (3),* 187-227.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal Basic Engineering, Transactions ASMA Series D, 82,* 35-45.

Koopman, S. J., Harvey, A. C., Doornik, J. A., & Shephard, N. (2007). *Stamp 8.0: Structural Time Series Analyser, Modeller and Predictor.* London: Timberlake Consultants.

Nocedal, J., & Wright, S. J. (1999). *Numerical Optimization.* New York: Springer Verlag.
Shumway, R. H., & Stoffer, D. S. (1982). An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis, 3,* 253-264.
Watson, M. W., & Engle, R. F. (1983). Alternative algorithms for the estimation of dynamic factor, MIMIC and varying coefficient regression. *Journal of Econometrics, 23,* 385-400.