Linear Gaussian State Space Models

- Structural Time Series Models
  - Level and Trend Models
  - Basic Structural Model (BSM)
- Dynamic Linear Models
  - State Space Model Representation
    - Level, Trend, and Seasonal Models
    - Time Varying Regression Model
  - Extensions
    - Multivariate Time Series Analysis
    - Bayesian Time Series Analysis

Structural Time Series Models

- Local level Model

$$y_t = \alpha_{t-1} + v_t, \quad v_t \sim N(0, \sigma_v^2)$$

$$\alpha_t = \alpha_{t-1} + w_t, \quad w_t \sim N(0, \sigma_w^2)$$

- Local Trend Model

$$y_t = \alpha_{t-1} + v_t, \quad v_t \sim N(0, \sigma_v^2)$$

$$\alpha_t = \alpha_{t-1} + \beta_{t-1} + w_{\alpha,t}, \quad w_{\alpha,t} \sim N(0, \sigma_\alpha^2)$$

$$\beta_t = \beta_{t-1} + w_{\beta,t}, \quad w_{\beta,t} \sim N(0, \sigma_\beta^2)$$

Structural Time Series Models

- Basic Structural Model (BSM)

$$y_t = \alpha_{t-1} + \gamma_{1,t-1} + v_t, \quad v_t \sim N(0, \sigma_v^2)$$

$$\alpha_t = \alpha_{t-1} + \beta_{t-1} + w_{\alpha,t}, \quad w_{\alpha,t} \sim N(0, \sigma_\alpha^2)$$

$$\beta_t = \beta_{t-1} + w_{\beta,t}, \quad w_{\beta,t} \sim N(0, \sigma_\beta^2)$$

$$\gamma_{1,t} = -\sum_{j=2}^{p-1} \gamma_{j,t-1} + w_{\gamma,t}, \quad w_{\gamma,t} \sim N(0, \sigma_\gamma^2)$$

$$\gamma_{j,t} = \gamma_{j,t-1}, \quad j = 2, \ldots, p-1$$

- Forecasting

$$\hat{y}_{t+1|t} = a_t + b_t + s_{t+1-p}, \quad t = 1, 2, \ldots, n$$

$$\hat{y}_{n+h|n} = a_n + hb_n + s_{n+h-p}, \quad h = 1, 2, \ldots \le p$$

Dynamic Linear Models

- Observation Equation

$$\underset{m\times 1}{y_t} = \underset{m\times p}{F_t}\ \underset{p\times 1}{\theta_t} + \underset{m\times 1}{v_t}, \quad \underset{m\times 1}{v_t} \sim iid\ N(\underset{m\times 1}{0}, \underset{m\times m}{V_t})$$

- State Equation

$$\underset{p\times 1}{\theta_t} = \underset{p\times p}{G_t}\ \underset{p\times 1}{\theta_{t-1}} + \underset{p\times 1}{w_t}, \quad \underset{p\times 1}{w_t} \sim iid\ N(\underset{p\times 1}{0}, \underset{p\times p}{W_t})$$

- Initial State Distribution

$$\theta_0 \sim N(m_0, C_0)$$

- $$E(v_t\theta_0^{'}) = E(w_t\theta_0^{'}) = 0 \quad \forall t$$
  $$E(v_t w_s^{'}) = \underset{m\times p}{0} \quad \forall t, s$$

Dynamic Linear Models

- Local Level Model

$$y_t = \alpha_t + v_t, \quad v_t \sim N(0, \sigma_v^2)$$
$$\alpha_t = \alpha_{t-1} + w_t, \quad w_t \sim N(0, \sigma_w^2)$$

  - State Space Model Representation

$$y_t = F_t \theta_t + v_t$$
$$\theta_t = G_t \theta_{t-1} + w_t \quad (m = 1,\ p = 1)$$
$$\theta_t = \alpha_t, F_t = 1, G_t = 1, V_t = \sigma_v^2, W_t = \sigma_w^2$$

Dynamic Linear Models

- Local Trend Model

$$y_t = \alpha_t + v_t, \quad v_t \sim N(0, \sigma_v^2)$$

$$\alpha_t = \alpha_{t-1} + \beta_{t-1} + w_{\alpha,t}, \quad w_{\alpha,t} \sim N(0, \sigma_\alpha^2)$$

$$\beta_t = \beta_{t-1} + w_{\beta,t}, \quad w_{\beta,t} \sim N(0, \sigma_\beta^2)$$

$$\Rightarrow \begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_{t-1} \\ \beta_{t-1} \end{bmatrix} + \begin{bmatrix} w_{\alpha,t} \\ w_{\beta,t} \end{bmatrix}$$

  - State Space Model Representation

$$y_t = F_t \theta_t + v_t$$

$$\theta_t = G_t \theta_{t-1} + w_t \quad (m = 1,\ p = 2)$$

$$\theta_t = \begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix}, F_t = \begin{bmatrix} 1 & 0 \end{bmatrix}, G_t = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, V_t = \sigma_v^2, W_t = \begin{bmatrix} \sigma_\alpha^2 & 0 \\ 0 & \sigma_\beta^2 \end{bmatrix}$$

Dynamic Linear Models

- Time Varying Regression Parameters

$$y_t = \alpha_t + \beta_t x_t + v_t, \quad v_t \sim N(0, \sigma_v^2)$$

$$\alpha_t = \alpha_{t-1} + w_{\alpha,t}, \quad w_{\alpha,t} \sim N(0, \sigma_\alpha^2)$$

$$\beta_t = \beta_{t-1} + w_{\beta,t}, \quad w_{\beta,t} \sim N(0, \sigma_\beta^2)$$

  - State Space Model Representation

$$y_t = \begin{bmatrix} 1 & x_t \end{bmatrix} \begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix} + v_t \quad \Leftarrow F_t = \begin{bmatrix} 1 & x_t \end{bmatrix}, G_t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, V_t = \sigma_v^2$$

$$\begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_{t-1} \\ \beta_{t-1} \end{bmatrix} + \begin{bmatrix} w_{\alpha,t} \\ w_{\beta,t} \end{bmatrix} \Leftarrow \quad \theta_t = \begin{bmatrix} \alpha_t \\ \beta_t \end{bmatrix}, w_t = \begin{bmatrix} w_{\alpha,t} \\ w_{\beta,t} \end{bmatrix}, W_t = \begin{bmatrix} \sigma_\alpha^2 & 0 \\ 0 & \sigma_\beta^2 \end{bmatrix}$$

Dynamic Linear Models

- Model Estimation

$$\{y_t\} \Rightarrow \begin{cases} y_t = F_t\theta_t + v_t \\ \theta_t = G_t\theta_{t-1} + w_t \end{cases}$$

  - Filtering (filtered estimate of θ)

    $$E(\theta_t \mid I_t = \{y_1, \ldots, y_t\})$$

  - Smoothing (smoothed estimate of θ)

    $$E(\theta_t \mid I_T = \{y_1, \ldots, y_T\})$$

Model Estimation

- The Kalman Filter is a set of recursion equations for determining the optimal estimates of θ_t given I_t. The filter consists of two sets of equations:
  - Prediction Equation
  - Update Equation
- Using the following notations

$$m_t = E(\theta_t \mid I_t) = optimal\ estimator\ of\ \theta_t\ based\ on\ I_t$$

$$C_t = E[(\theta_t - m_t)(\theta_t - m_t)' \mid I_t) = MSE\ matrix\ of\ m_t$$

Model Estimation

- Prediction Equations
  - Given m_{t-1} and C_{t-1} at t-1, the optimal predictor of θ_t and its MSE matrix are

$$m_{t|t-1} = E(\theta_t \mid I_{t-1}) = G_t m_{t-1}$$

$$C_{t|t-1} = E[(\theta_t - m_{t-1})(\theta_t - m_{t-1})' \mid I_{t-1}) = G_t C_{t-1} G_t' + W_t$$

  - The corresponding optimal predictor of y_t at t-1 is

$$y_{t|t-1} = E[y_t \mid I_{t-1}] = F_t m_{t|t-1}$$

  - The predictive error and its MSE matrix are

$$e_t = y_t - y_{t|t-1} = y_t - F_t m_{t|t-1} = F_t(\theta_t - m_{t|t-1}) + v_t$$

$$E(e_t e_t') = Q_t = F_t C_{t|t-1} F_t' + V_t$$

Model Estimation

- Update Equations
  - When new observation y_t become available, the optimal predictor $m_{t|t-1}$ and its MSE matrix are updated using

$$m_t = m_{t|t-1} + C_{t|t-1} F_t^{'} Q_t^{-1} (y_t - F_t m_{t|t-1})$$

$$= m_{t|t-1} + C_{t|t-1} F_t^{'} Q_t^{-1} e_t$$

$$C_t = C_{t|t-1} - C_{t|t-1} F_t^{'} Q_t^{-1} F_t C_{t|t-1}$$

$$Note : K_t = C_{t|t-1} F_t^{'} Q_t^{-1} = Kalman\ Gain\ Matrix$$

  - Kalman Gain Matrix gives the weight on new information e_t in the update equation for m_t.

Model Estimation

- Kalman Smoother
  - Once all data I_T is observed, the optimal estimators $E(\theta_t | I_T)$ can be computed using the backwards Kalman smoothing recursions

$$E(\theta_t \mid I_T) = m_{t|T} = m_t + C_t^*(m_{t+1|T} - G_{t+1} m_t)$$

$$E[(\theta_t - m_{t|T})(\theta_t - m_{t|T})' \mid I_T) = C_{t|T} = C_t^*(C_{t+1|T} - C_{t=1|t})C_t^{*'}$$

$$C_t^* = C_t G_{t+1}' C_{t+1|t}^{-1}$$

  - The algorithm starts by setting $m_{T|T} = m_T$ and $C_{T|T} = C_T$ and then proceed backwards for t = T-1, …,1.

Maximum Likelihood Estimation

- For a linear Gaussian state space model, let ψ denote the parameters of the model (embedded in the system matrices F_t, G_t, V_t, and W_t). The prediction error decomposition of the Gaussian log-likelihood function is

$$y_{t|t-1} \sim N(F_t(\psi)m_{t|t-1}(\psi), Q_t(\psi))$$

$$e_t(\psi) = y_t - y_{t|t-1} = y_t - F_t(\psi)m_{t|t-1}(\psi) \sim N(0, Q_t(\psi))$$

$$f(y_{t|t-1};\psi) = (2\pi Q_t(\psi))^{-1/2} \exp\left\{-\frac{1}{2} e_t(\psi)' Q_t^{-1}(\psi) e_t(\psi)\right\}$$

$$\ln L(\psi \mid y) = \sum_{t=1}^{T} \ln f(y_t \mid I_{t-1};\psi)$$

$$= -\frac{NT}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{N} \ln|Q_t(\psi)| - \frac{1}{2}\sum_{t=1}^{N} e_t(\psi)' Q_t^{-1}(\psi) e_t(\psi)$$

$$\Rightarrow \hat{\psi}_{MLE} = \arg\max_{\psi} \ln L(\psi \mid y)$$

Forecasting

- The Kalman filter prediction equations produces in-sample 1-step ahead forecasts and MSE matrices.
- Out-of-sample h-step ahead predictions and MSE matrices can be computed from the prediction equations by extending the data set $y_1, \ldots, y_T$ with a set of h missing values.
- When y_τ is missing the Kalman filter reduces to the prediction step so a sequence of h missing values at the end of the sample will produce a set of h-step ahead forecasts

Forecasting

- One-Step Ahead Forecast at t = p, p+1,…

$$\hat{y}_{t+1|t} = a_t + b_t + s_{t+1-p}$$

$$\hat{y}_{1|0} = a_0 + b_0 + s_{1-p}$$

$$\hat{y}_{2|1} = a_1 + b_1 + s_{2-p}$$

$$\ldots$$

$$\hat{y}_{p|p-1} = a_{p-1} + b_{p-1} + s_0$$

$$\hat{\varepsilon}_{t+1|t} = y_{t+1} - \hat{y}_{t+1|t}$$

$$MAE = mean(|\,\hat{\varepsilon}_{t+1|t}\,|)$$

$$MAPE = 100\; mean(|\,\hat{\varepsilon}_{t+1|t} \,/\, y_{t+1}\,|)$$

$$MSE = mean(\hat{\varepsilon}^2_{t+1|t})$$

$$RMSE = \sqrt{mean(\hat{\varepsilon}^2_{t+1|t})}$$

$$RMSPE = 100\sqrt{mean[(\hat{\varepsilon}_{t+1|t} \,/\, y_{t+1})^2]}$$

Example 1 (Continued)

- China Shanghai Common Stock
  - High Frequency Daily Index
  - Monthly Index Time Series
    - Trend, Seasonality
  - Dynamic Linear Model
    - Correlation with Exchange Rate?

Example 2

- Chinese Yuan vs. U.S. Dollar
  - Exchange Rate Time Series
    - Trend
    - Intervention
  - Dynamic Linear Model
    - Correlation with Stock Market?
    - ARMA