STATE-SPACE MODELS

AND

THE KALMAN RECURSIONS

State-space models

Replacing the deterministic trend component

$$m_t=a+bt$$

in the **linear trend model**

$$Y_t=m_t+U_t,\ t=1,2,3,\ldots$$

by a cumulative sum of i.i.d. random variables (**random walk**)

$$M_t=V_1+\ldots+V_t,$$

we obtain the **stochastic trend model**

$$Y_t=M_t+U_t,\ t=1,2,3,\ldots,$$

which can be represented in state-space form as

$$Y_t=M_t+U_t,\ t=1,2,3,\ldots,$$

$$M_t=M_{t-1}+V_t,\ t=1,2,3,\ldots\ .$$

A **state-space model** consists of two equations. The **observation equation** (or **measurement equation**) expresses the observation Y_t as a linear function of the state M_t plus noise. The **state equation** (or **transition equation**) expresses the state M_t as a linear function of the previous state M_{t-1} plus noise.

All noise terms in the two equations are taken to be uncorrelated with each other and also with the initial state M_0.

Replacing the deterministic trend component

$$m_t=a+bt$$

in the **linear trend model**

$$Y_t=m_t+U_t,\ t=1,2,3,\ldots$$

first by a **random walk with drift**

$$M_t=(b+V_1)+\ldots+(b+V_t)$$

and then the deterministic drift term b by another random walk

$$B_t=W_1+\ldots+W_t,$$

we obtain the **local linear trend model**, which can be written as

$$Y_t=M_t+U_t,\ t=1,2,3,\ldots,$$

$$M_t= M_{t-1}+B_t+V_t,\ t=1,2,3,\ldots,$$

$$B_t=B_{t-1}+W_t,\ t=1,2,3,\ldots,$$

or in state-space form with a 2-dimensional state vector as

$$Y_t=\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}M_t\\ B_t^*\end{pmatrix}+U_t,\ t=1,2,3,\ldots,$$

$$\begin{pmatrix}M_t\\ B_t^*\end{pmatrix}=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\begin{pmatrix}M_{t-1}\\ B_{t-1}^*\end{pmatrix}+\begin{pmatrix}V_t\\ W_t^*\end{pmatrix},\ t=1,2,3,\ldots,$$

where $B_t^*=B_{t+1}$ and $W_t^*=W_{t+1}$.

The Kalman recursions

Consider the state-space model

$$Y_t = HX_t + U_t,\ X_{t+1} = FX_t + V_t$$

for the possibly multivariate observations $Y_1,\dots,Y_n$.

Assuming that all covariance matrices are finite and choosing $Y_0=(1,\dots,1)^T$, we define the one-step predictor

$$\hat{X}_t = P_{t-1}(X_t) = P(X_t \mid Y_{t-1},\dots,Y_0)$$

of X_t in terms of $Y_{t-1},\dots,Y_0$ as the (componentwise) projection of X_t onto the closed subspace spanned by the components of $Y_{t-1},\dots,Y_0$.

Using

$$J_t = Y_t - P_{t-1}(Y_t) = HX_t + U_t - P_{t-1}(HX_t + U_t) = H(X_t - \hat{X}_t) + U_t,$$

$$E(J_t J_t^T) = H\underbrace{E(X_t - \hat{X}_t)(X_t - \hat{X})^T}_{\Omega_t} H^T + \underbrace{EU_t U_t^T}_{\Sigma_U},$$

and

$$E(X_{t+1} J_t^T) = E(FX_t + V_t)((X_t - \hat{X}_t)^T H^T + U_t^T) = F\Omega_t H^T$$

we obtain

$$\begin{aligned}\hat{X}_{t+1} &= P(X_{t+1} \mid Y_t,\dots,Y_1)\\ &= P(X_{t+1} \mid J_t) + P(X_{t+1} \mid Y_{t-1},\dots,Y_1)\\ &= E(X_{t+1} J_t^T)(E(J_t J_t^T))^{-1} J_t + P_{t-1}(FX_t + V_t)\\ &= \underbrace{F\Omega_t H^T (H\Omega_t H^T + \Sigma_U)^{-1}}_{G(\Omega_t)} J_t + F\hat{X}_t.\end{aligned}$$

The covariance matrices

$$\Omega_t=E(X_t-\hat{X}_t)(X_t-\hat{X}_t)^T$$

can also be obtained recursively, because Ω_{t+1} can be written as

$$\begin{aligned}
&E(X_{t+1}X_{t+1}^T)-E(\hat{X}_{t+1}\hat{X}_{t+1}^T)\\
&=E(FX_t+V_t)(FX_t+V_t)^T-E(G(\Omega_t)J_t+F\hat{X}_t)(G(\Omega_t)J_t+F\hat{X}_t)^T\\
&=F(E(X_tX_t^T)-E(\hat{X}_t\hat{X}_t^T))F^T+EV_tV_t^T-G(\Omega_t)E(J_tJ_t^T)G^T(\Omega_t)\\
&=F\Omega_tF^T+\Sigma_V-F\Omega_tH^T(H\Omega_tH^T+\Sigma_U)^{-1}(H\Omega_tH^T+\Sigma_U)G^T(\Omega_t)\\
&=F\Omega_tF^T+\Sigma_V-F\Omega_tH^T(H\Omega_tH^T+\Sigma_U)^{-1}(F\Omega_tH^T)^T.
\end{aligned}$$

<u>Exercise:</u> Suppose that the inverse of $E(J_tJ_t^T)$ exists. Show that

$$E(X_{t+1}-MJ_t)J_t^T=0$$

implies that

$$M=E(X_{t+1}J_t^T)(E(J_tJ_t^T))^{-1}.$$

Estimation of state-space models

Assuming that all variables are jointly normally distributed, we can write the joint density of the m-dimensional observations $Y_1,\ldots,Y_n$ (conditional on Y_0) as

$$f(Y_1,\ldots,Y_n \mid Y_0) = \prod_{t=1}^{n} f(Y_t \mid Y_{t-1},\ldots,Y_0)$$

$$= \prod_{t=1}^{n} (2\pi)^{-\frac{m}{2}} (\det \Sigma_t)^{-\frac{1}{2}} \exp\left(-\tfrac{1}{2} J_t^T \Sigma_t^{-1} J_t\right),$$

where

$$J_t = Y_t - P_{t-1}(Y_t) = Y_t - H\hat{X}_t$$

and

$$\Sigma_t = E(Y_t - P_{t-1}(Y_t))(Y_t - P_{t-1}(Y_t))^T = E(J_t J_t^T) = H\Omega_t H^T + \Sigma_U.$$

For given parameter matrices H, F, Σ_U, Σ_V, the joint density can be evaluated with the Kalman recursions. To find the parameter matrices that maximize the joint density (maximum likelihood estimates) we must use some nonlinear optimization algorithm.

State-space representations of ARMA models

A univariate AR(p) process represented by

$$y_t=\phi_1 y_{t-1}+\ldots+\phi_p y_{t-p}+u_t$$

can be written in state-space form as

$$y_t=(0,0,\ldots,0,1)X_t,$$

$$X_{t+1}=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \phi_p & \phi_{p-1} & \phi_{p-2} & \cdots & \phi_1 \end{pmatrix} X_t+\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} u_{t+1},$$

where

$$X_t=(y_{t-p+1},y_{t-p+2},\ldots,y_{t-1},y_t)^T.$$

State-space representations are not unique. For example, the MA(1) process

$$y_t=\theta_1 u_{t-1}+u_t$$

can be written as

$$y_t=(1,\theta_1)X_t,$$

$$X_{t+1}=\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}X_t+\begin{pmatrix} 1 \\ 0 \end{pmatrix}u_{t+1},$$

where

$$X_t=\begin{pmatrix} u_t \\ u_{t-1} \end{pmatrix},$$

and also as

$$y_t=(1,0)X_t,$$

$$X_{t+1}=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}X_t+\begin{pmatrix} 1 \\ \theta_1 \end{pmatrix}u_{t+1},$$

where

$$X_t=\begin{pmatrix} u_t+\theta_1 u_{t-1} \\ \theta_1 u_t \end{pmatrix}.$$