EL MODELO LOCAL LEVEL

"Una Descripción Detallada"

PRÁCTICA INVESTIGATIVA IV

Ana Catalina Ruiz

Asesor: Freddy H. Marín.

Ingeniería Matemática

DEPARTAMENTO DE CIENCIAS BÁSICAS
UNIVERSIDAD EAFIT
Medellín

2008

El Modelo Local Level

Introducción

En este trabajo se pretende presentar formalmente uno de los modelos básicos asociados al análisis de series de tiempo por métodos de estado espacio Gaussianos, el modelo Local Level.
Este modelo tiene algunos elementos algebraicos un poco complejos y por ello se quiere describir los procedimientos paso a paso para su mejor comprensión; entre ellos se encuentran algunas técnicas básicas del análisis del Espacio Estado como filtrado, suavización, inicialización y estimación de los parámetros. Se presentará además, un ejemplo ilustrativo con una serie simulada en la que se aplican las técnicas y procedimientos descritos en este trabajo.

El Modelo Local Level

El modelo local level se puede ser descrito por los 2 siguientes procesos

$$\begin{cases} y_t = \alpha_t + \varepsilon_t, & \varepsilon_t \sim \mathrm{N}(0,\sigma_\varepsilon^2) \\ \alpha_{t+1} = \alpha_t + \eta_t, & \eta_t \sim \mathrm{N}(0,\sigma_\eta^2) \end{cases} \quad , \qquad (1)$$

donde y_t se conoce como la ecuación de medida y α_t como la ecuación de transición.

$$\begin{aligned} y_t &= \alpha_t + \varepsilon_t \\ &= \alpha_{t-1} + \eta_{t-1} + \varepsilon_t \\ &= \alpha_{t-2} + \eta_{t-2} + \eta_{t-1} + \varepsilon_t \\ &\vdots \\ &= \alpha_1 + \eta_1 + \eta_2 + \cdots + \eta_{t-2} + \eta_{t-1} \end{aligned}$$

Por tanto

$$y_t = \alpha_1 + \sum_{j=1}^{t-1} \eta_j + \varepsilon_t, \qquad t = 1, \ldots, n$$

Las observaciones y_t generadas por el modelo local level se representan como un vector y de $n \times 1$, tal que

$$y \sim \mathrm{N}(\mathbf{1}a_1, \mathbf{\Omega}), \quad \text{con } y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{1} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad \text{y} \quad \mathbf{\Omega} = \mathbf{1}\mathbf{1}' P_1 + \Sigma$$

donde el elemento (i,j) de la matriz Σ de $n \times n$ está dado por

$$\Sigma_{ij} = \begin{cases} (i-1)\sigma_\eta^2, & i<j \\ \sigma_\varepsilon^2 + (i+1)\sigma_\eta^2, & i=j \qquad i,j=1,\dots,n \\ (j-1)\sigma_\eta^2, & i>j \end{cases}$$

Para $i = j$

$$\begin{aligned}
\Sigma_{ij} &= \text{cov}(y_i, y_i) \\
&= E(y_i^2) - E(y_i)^2 \\
&= E\left(\alpha_1^2 + \sum_{t=1}^{i-1}\eta_t^2 + \varepsilon_i^2 + 2\alpha_1\sum_{t=1}^{i-1}\eta_t + 2\alpha_1\varepsilon_i + 2\varepsilon_i\sum_{t=1}^{i-1}\eta_t + 2\sum_{t=1}^{i-1}\sum_{t_1=t+1}^{i-1}\eta_t\eta_{t_1}\right) - a_1^2 \\
&= E(\alpha_1^2) + \sum_{t=1}^{i-1}E(\eta_t^2) + E(\varepsilon_i^2) + 2E(\alpha_1\sum_{t=1}^{i-1}\eta_t) + 2E(\alpha_1\varepsilon_i) + 2E(\varepsilon_1\sum_{t=1}^{i-1}\eta_t) + \\
&\quad + 2\sum_{t=1}^{i-1}\sum_{t_1=t+1}^{i-1}E(\eta_t\eta_{t_1}) - a_1^2
\end{aligned}$$

Como las ε_t y las η_t son mutuamente independientes y son independientes de α_1, entonces

$$\begin{aligned}
\Sigma_{ij} &= E(\alpha_1^2) + \sum_{t=1}^{i-1}E(\eta_t^2) + E(\varepsilon_i^2) + 2E(\alpha_1)\sum_{t=1}^{i-1}E(\eta_t) + 2E(\alpha_1)E(\varepsilon_i) + 2E(\varepsilon_1)\sum_{t=1}^{i-1}E(\eta_t) + \\
&\quad + 2\sum_{t=1}^{i-1}\sum_{t_1=t+1}^{i-1}E(\eta_t)E(\eta_{t_1}) - a_1^2 \\
&= E(\alpha_1^2) + \sum_{t=1}^{i-1}\sigma_\eta^2 + \sigma_\varepsilon^2 - a_1^2 \\
&= E(\alpha_1^2) - a_1^2 + \sum_{t=1}^{i-1}\sigma_\eta^2 + \sigma_\varepsilon^2 \\
&= \text{var}(\alpha_1^2) + (i-1)\sigma_\eta^2 + \sigma_\varepsilon^2 \\
&= P_1 + (i-1)\sigma_\eta^2 + \sigma_\varepsilon^2
\end{aligned}$$

Para $i < j$

$$\begin{aligned}
\Sigma_{ij} &= \text{cov}(y_i, y_j) \\
&= E(y_i y_j) - E(y_i)E(y_j) \\
&= E(y_i y_j) - a_1^2
\end{aligned}$$

Como $y_i = \alpha_1 + \sum_{t=1}^{i-1}\eta_t + \varepsilon_i$ y $y_j = \alpha_1 + \sum_{t=1}^{j-1}\eta_t + \varepsilon_j$, entonces

$$y_j = y_i + \sum_{t=i}^{j-1}\eta_t - \varepsilon_i + \varepsilon_j$$

$$\Sigma_{ij} = E\left[y_i \left(y_i + \sum_{t=i}^{j-1} \eta_t - \varepsilon_i + \varepsilon_j \right)\right] - a_1^2$$
$$= E\left(y_i^2 + \sum_{t=i}^{j-1} y_i \eta_t - y_i \varepsilon_i + y_i \varepsilon_j \right) - a_1^2$$
$$= E(y_i^2) + \sum_{t=i}^{j-1} E(y_i \eta_t) - E(y_i \varepsilon_i) + E(y_i \varepsilon_j)$$

Como y_i es independiente de ε_j y de η_t para $t \geq i$

$$\Sigma_{ij} = E(y_i^2) + \sum_{t=i}^{j-1} E(y_i)E(\eta_t) - E(y_i\varepsilon_i) + E(y_i)E(\varepsilon_j)$$
$$= E(y_i^2) - E(y_i\varepsilon_i) - a_1^2$$
$$= E(y_i^2) - E\big((\alpha_i + \varepsilon_i)\varepsilon_i\big) - a_1^2$$
$$= E(y_i^2) - E\big(\alpha_i\varepsilon_i + \varepsilon_i^2\big) - a_1^2$$
$$= E(y_i^2) - E(\alpha_i\varepsilon_i) - E(\varepsilon_i^2) - a_1^2$$

Como α_i y ε_i son independientes

$$\Sigma_{ij} = E(y_i^2) - E(\alpha_i)E(\varepsilon_i) - E(\varepsilon_i^2) - a_1^2$$
$$= E(y_i^2) - E(\varepsilon_i^2) - a_1^2$$
$$= E(y_i^2) - a_1^2 - E(\varepsilon_i^2)$$
$$= \operatorname{cov}(y_i, y_i) - \sigma_\varepsilon^2$$
$$= P_1 + (i-1)\sigma_\eta^2 + \sigma_\varepsilon^2 - \sigma_\varepsilon^2$$
$$= P_1 + (i-1)\sigma_\eta^2$$

De la misma manera se puede probar que para $i > j$, $\Sigma_{ij} = P_1 + (j-1)\sigma_\eta^2$

El Filtro de Kalman

El objetivo del filtro de Kalman es actualizar el conocimiento del sistema cada vez que es recogida una nueva observación y_t.

Sea Y_{t-1} el conjunto de observaciones pasadas $\{y_1,\ldots,y_{t-1}\}$ y asúmase que la distribución condicional de α_t dado Y_{t-1} es $N(a_t, P_t)$ donde a_t y P_t serán determinadas. Dado que a_t y P_t son conocidas, el objetivo es calcular a_{t+1} y P_{t+1} cuando y_t es recogida.

Dado que $a_{t+1} = E(\alpha_{t+1} \mid Y_t) = E(\alpha_t + \eta_t \mid Y_t)$ y $P_{t+1} = \operatorname{var}(\alpha_{t+1} \mid Y_t) = \operatorname{var}(\alpha_t + \eta_t \mid Y_t)$, de (1) se tiene

$$a_{t+1} = E(\alpha_t \mid Y_t), \quad P_{t+1} = \operatorname{var}(\alpha_t \mid Y_t) + \sigma_\eta^2$$

Sea $v_t = y_t - a_t$ y $F_t = \text{var}(v_t)$, entonces

$$
\begin{aligned}
E(v_t \mid Y_{t-1}) = E(y_t - a_t \mid Y_{t-1}) = E(\alpha_t + \varepsilon_t - a_t \mid Y_{t-1}) \\
&= E(\alpha_t \mid Y_{t-1}) + E(\varepsilon_t \mid Y_{t-1}) - E(a_t \mid Y_{t-1}) \\
&= a_t + 0 - a_t \\
&= 0
\end{aligned}
$$

Así $E(v_t) = E\left[E(v_t \mid Y_{t-1})\right] = 0$ y

$$
\begin{aligned}
\text{cov}(v_t, y_j) &= E(v_t y_j) - E(v_t)E(y_j) \\
&= E(v_t y_j) \\
&= E\left[E(v_t \mid Y_{t-1}) y_j\right] \\
&= E\left[0(y_j)\right] \\
&= 0
\end{aligned}
$$

Así, v_t y y_j son independientes para $j = 1,\ldots, t\text{-}1$. Cuando Y_t es fija $Y_{t\text{-}1}$ y y_t son fijas, así $Y_{t\text{-}1}$ y v_t son fijas y viceversa. Por tanto, $E(\alpha_t \mid Y_t) = E(\alpha_t \mid Y_{t-1}, y_t) = E(\alpha_t \mid Y_{t-1}, v_t + a_t) = E(\alpha_t \mid Y_{t-1}, v_t)$ y de la misma forma $\text{var}(\alpha_t \mid Y_t) = \text{var}(\alpha_t \mid Y_{t-1}, v_t)$.

$$
E(\alpha_t \mid Y_t) = E(\alpha_t \mid Y_{t-1}, v_t) = E(\alpha_t \mid Y_{t-1}) + \text{cov}(\alpha_t, v_t)\,\text{var}(v_t)^{-1} v_t \qquad (2)
$$

Donde

$$
\begin{aligned}
\text{cov}(\alpha_t, v_t) &= E(\alpha_t v_t) - E(\alpha_t)E(v_t) \\
&= E(\alpha_t (y_t - a_t)) \\
&= E(\alpha_t (\alpha_t + \varepsilon_t - a_t)) \\
&= E(\alpha_t (\alpha_t - a_t)) \\
&= E(E(\alpha_t (\alpha_t - a_t) \mid Y_{t-1})) \\
&= E(E(\alpha_t^2 \mid Y_{t-1}) - E(\alpha_t a_t \mid Y_{t-1})) \\
&= E(E(\alpha_t^2 \mid Y_{t-1}) - a_t E(\alpha_t \mid Y_{t-1})) \\
&= E(E(\alpha_t^2 \mid Y_{t-1}) - a_t^2) \\
&= E(E(\alpha_t^2 \mid Y_{t-1}) - E(\alpha_t \mid Y_{t-1})^2) \\
&= E(\text{var}(\alpha_t \mid Y_{t-1})) \\
&= E(P_t) \\
&= P_t
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{var}(v_t) = F_t &= \operatorname{var}(\alpha_t + \varepsilon_t - a_t) \\
&= E\left[(\alpha_t + \varepsilon_t - a_t)^2\right] - E(\alpha_t + \varepsilon_t - a_t)^2 \\
&= E(\alpha_t^2 + \varepsilon_t^2 + a_t^2 + 2\alpha_t\varepsilon_t - 2\alpha_t a_t - 2\varepsilon_t a_t) - E(\alpha_t + \varepsilon_t - a_t)^2 \\
&= E(\alpha_t^2) + E(\varepsilon_t^2) + E(a_t^2) + 2E(\alpha_t\varepsilon_t) - 2E(\alpha_t a_t) - 2E(\varepsilon_t a_t) - \left[E(\alpha_t) + E(\varepsilon_t) - E(a_t)\right]^2
\end{aligned}
$$

Dado que ε_t es independiente de α_t y de a_t,

$$
\begin{aligned}
\operatorname{var}(v_t) &= E(\alpha_t^2) + \operatorname{var}(\varepsilon_t) + E(a_t^2) + 2E(\alpha_t)E(\varepsilon_t) - 2E(\alpha_t a_t) - 2E(\varepsilon_t)E(a_t) - \left[E(\alpha_t) - E(a_t)\right]^2 \\
&= E(\alpha_t^2) + \operatorname{var}(\varepsilon_t) + E(a_t^2) - 2E(\alpha_t a_t) - \left[E\left[E(\alpha_t \mid Y_{t-1})\right] - E(a_t)\right]^2 \\
&= E(\alpha_t^2) + \operatorname{var}(\varepsilon_t) + E(a_t^2) - 2E\left[E(\alpha_t a_t \mid Y_{t-1})\right] - \left[E(a_t) - E(a_t)\right]^2 \\
&= E(\alpha_t^2) + \operatorname{var}(\varepsilon_t) + E(a_t^2) - 2E\left[a_t E(\alpha_t \mid Y_{t-1})\right] - 0 \\
&= E(\alpha_t^2) + \operatorname{var}(\varepsilon_t) + E(a_t^2) - 2E\left[a_t^2\right] \\
&= E\left[E(\alpha_t^2 \mid Y_{t-1})\right] + \operatorname{var}(\varepsilon_t) - E\left[a_t^2\right] \\
&= E\left[E(\alpha_t^2 \mid Y_{t-1}) - a_t^2\right] + \operatorname{var}(\varepsilon_t) \\
&= E\left[E(\alpha_t^2 \mid Y_{t-1}) - E(\alpha_t \mid Y_{t-1})^2\right] + \operatorname{var}(\varepsilon_t) \\
&= E\left[\operatorname{var}(\alpha_t \mid Y_{t-1})\right] + \operatorname{var}(\varepsilon_t) \\
&= E\left[P_t\right] + \sigma_\varepsilon^2 \\
&= P_t + \sigma_\varepsilon^2
\end{aligned}
$$

Puesto que $a_t = E(\alpha_t \mid Y_{t-1})$, de (2) se tiene que

$$E(\alpha_t \mid Y_t) = a_t + K_t v_t$$

Donde $K_t = P_t / F_t$ es el coeficiente de regresión de α_t en v_t.

$$
\begin{aligned}
\operatorname{var}(\alpha_t \mid Y_t) &= \operatorname{var}(\alpha_t \mid Y_{t-1}, v_t) \\
&= \operatorname{var}(\alpha_t \mid Y_{t-1}) - \operatorname{cov}(\alpha_t, v_t)^2 \operatorname{var}(v_t)^{-1} \\
&= P_t - P_t^2 / F_t \\
&= P_t(1 - K_t)
\end{aligned}
$$

Las relaciones de actualización del tiempo t al tiempo $t+1$, son

$$
\begin{aligned}
v_t &= y_t - a_t, \qquad F_t = P_t + \sigma_\varepsilon^2, \qquad K_t = P_t / F_t \\
a_{t+1} &= a_t + K_t v_t, \qquad P_{t+1} = P_t(1 - K_t) + \sigma_\eta^2
\end{aligned}
\qquad (3)
$$

Descomposición de Cholesky

La densidad conjunta de $y_1,\dots,y_n$ es

$$p(y_1,\dots,y_n) = p(y_1)\prod_{t=2}^{n} p(y_t \mid Y_{t-1}) \qquad (3)$$

$$v_1 = y_1 - a_1$$

$$\begin{aligned} v_2 &= y_2 - a_2 \\ &= y_2 - a_1 - K_1 v_1 \\ &= y_2 - a_1 - K_1(y_1 - a_1) \end{aligned}$$

$$\begin{aligned} v_3 &= y_3 - a_3 \\ &= y_3 - a_2 - K_2 v_2 \\ &= y_3 - a_1 - K_1(y_1 - a_1) - K_2(y_2 - a_2) \\ &= y_3 - a_1 - K_1(y_1 - a_1) - K_2(y_2 - a_1 - K_1 v_1) \\ &= y_3 - a_1 - K_1(y_1 - a_1) - K_2(y_2 - a_1) + K_2 K_1(y_1 - a_1) \\ &= y_3 - a_1 - K_2(y_2 - a_1) - K_1(1 - K_2)(y_1 - a_1) \end{aligned}$$

$$\begin{aligned} v_4 &= y_4 - a_4 \\ &= y_4 - a_3 - K_3 v_3 \\ &= y_4 - a_1 - K_2(y_2 - a_1) - K_1(1 - K_2)(y_1 - a_1) - K_3(y_3 - a_3) \\ &= y_4 - a_1 - K_2(y_2 - a_1) - K_1(1 - K_2)(y_1 - a_1) - K_3(y_3 - a_1 - K_2(y_2 - a_1) - \\ &\quad - K_1(1 - K_2)(y_1 - a_1)) \\ &= y_4 - a_1 - K_2(y_2 - a_1) - K_1(1 - K_2)(y_1 - a_1) - K_3(y_3 - a_1) + K_3 K_2(y_2 - a_1) + \\ &\quad + K_3 K_1(1 - K_2)(y_1 - a_1)) \\ &= y_4 - a_1 - K_2(y_2 - a_1)(1 - K_3) - K_1(1 - K_2)(y_1 - a_1)(1 - K_3) - K_3(y_3 - a_1) \\ &= y_4 - a_1 - K_3(y_3 - a_1) - K_2(1 - K_3)(y_2 - a_1) - K_1(1 - K_2)(1 - K_3)(y_1 - a_1) \end{aligned}$$

y en general,

$$\begin{aligned} v_j = y_j &- a_1 - K_{j-1}(y_{j-1} - a_1) - K_{j-2}(1 - K_{j-1})(y_{j-2} - a_1) - \cdots - \\ &- K_{j-i}(1 - K_{j-i+1})\dots(1 - K_{j-2})(1 - K_{j-1})(y_{j-i} - a_1) - \cdots - \\ &- K_2(1 - K_3)\dots(1 - K_{j-2})(1 - K_{j-1})(y_2 - a_1) - \\ &- K_1(1 - K_2)\dots(1 - K_{j-2})(1 - K_{j-1})(y_1 - a_1) \end{aligned}$$

Se puede escribir *v* matricialmente como

$$v=\begin{bmatrix}1&0&0&&0\\c_{21}&1&0&&0\\c_{31}&c_{32}&1&&0\\&&&\ddots&\vdots\\c_{n1}&c_{n2}&c_{n3}&\cdots&1\end{bmatrix}\begin{bmatrix}y_1-a_1\\y_2-a_1\\y_3-a_1\\\\y_n-a_1\end{bmatrix}=C(y-\mathbf{1}a_1),$$

donde C es la matriz superior

$$C=\begin{bmatrix}1&0&0&&0\\c_{21}&1&0&&0\\c_{31}&c_{32}&1&&0\\&&&\ddots&\vdots\\c_{n1}&c_{n2}&c_{n3}&\cdots&1\end{bmatrix};$$

Con $\begin{aligned}c_{i,i-1}&=-K_{i-1},\\ c_{ij}&=-K_j(1-K_{j+1})\cdots(1-K_{i-2})(1-K_{i-1})\end{aligned}$

$$p(y_1,\dots,y_n)=p(v_1,\dots,v_n)\cdot|J|$$

Donde $J=\begin{vmatrix}\frac{\partial v_1}{\partial y_1}&\frac{\partial v_1}{\partial y_2}&\frac{\partial v_1}{\partial y_3}&&\frac{\partial v_1}{\partial y_n}\\\frac{\partial v_2}{\partial y_1}&\frac{\partial v_2}{\partial y_2}&\frac{\partial v_2}{\partial y_3}&&\frac{\partial v_2}{\partial y_n}\\\frac{\partial v_3}{\partial y_1}&\frac{\partial v_3}{\partial y_2}&\frac{\partial v_3}{\partial y_3}&&\frac{\partial v_3}{\partial y_n}\\&&&\ddots&\vdots\\\frac{\partial v_n}{\partial y_1}&\frac{\partial v_n}{\partial y_2}&\frac{\partial v_n}{\partial y_3}&\cdots&\frac{\partial v_n}{\partial y_n}\end{vmatrix}$

Así $J=\begin{vmatrix}1&0&0&&0\\-K_1&1&0&&0\\-K_1(1-K_2)&-K_2&1&&0\\&&&\ddots&\vdots\\-K_1(1-K_2)\cdots(1-K_{n-1})a&-K_2\cdots(1-K_{n-1})&-K_3\cdots(1-K_{n-1})a&\cdots&1\end{vmatrix}$

Como la anterior es una matriz triangular inferior, su determinante es el producto de los elementos de la diagonal y por tanto $|J|=1$ y $p(y_1,\dots,y_n)=p(v_1,\dots,v_n)$.

Sustituyendo $p(v_1) = p(y_1)$ y $p(v_t) = p(y_t|y_{t-1})$ en (3)

$$p(v_1,\ldots,v_n) = \prod_{t=1}^{n} p(v_t)$$

$$\begin{aligned} E(v) = E\left[C(y-\mathbf{1}a_1)\right] &= CE\left[(y-\mathbf{1}a_1)\right] \\ &= C\left[E(y)-E(\mathbf{1}a_1)\right] \\ &= C\left[\mathbf{1}a_1-\mathbf{1}a_1\right] \\ &= \mathbf{0} \end{aligned}$$

$$\begin{aligned} \mathrm{var}(v) = \mathrm{var}\left[C(y-\mathbf{1}a_1)\right] &= E\left[C(y-\mathbf{1}a_1)\left[C(y-\mathbf{1}a_1)\right]'\right] \\ &= E\left[C(y-\mathbf{1}a_1)(y-\mathbf{1}a_1)'C'\right] \\ &= CE\left[(y-\mathbf{1}a_1)(y-\mathbf{1}a_1)'\right]C' \\ &= CE\left[(y-E(y))(y-E(y))'\right]C' \\ &= C\Omega C' \end{aligned}$$

En consecuencia, las v_t son independientemente distribuidas.

Se define el error de estimación del estado como

$$x_t = \alpha_t - a_t \qquad \text{con} \qquad \mathrm{Var}(x_t)=P_t$$

$$\begin{aligned} \mathrm{var}(x_t) &= \mathrm{var}(\alpha_t - a_t) \\ &= E\left[(\alpha_t - a_t)^2\right] - E(\alpha_t - a_t)^2 \\ &= E\left[\alpha_t^2 - 2\alpha_t a_t + a_t^2\right] - \left[E(\alpha_t) - E(a_t)\right]^2 \\ &= E(\alpha_t^2) - 2E(\alpha_t a_t) + E(a_t^2) - \left[E(\alpha_t) - E(a_t)\right]^2 \\ &= E\left[E(\alpha_t^2 \mid Y_{t-1})\right] - 2E\left[E(\alpha_t a_t \mid Y_{t-1})\right] + E(a_t^2) - \left[E\left[E(\alpha_t \mid Y_{t-1})\right] - E(a_t)\right]^2 \\ &= E\left[E(\alpha_t^2 \mid Y_{t-1})\right] - 2E\left[a_t E(\alpha_t \mid Y_{t-1})\right] + E(a_t^2) - \left[E(a_t) - E(a_t)\right]^2 \\ &= E\left[E(\alpha_t^2 \mid Y_{t-1})\right] - 2E\left[a_t^2\right] + E(a_t^2) \\ &= E\left[E(\alpha_t^2 \mid Y_{t-1})\right] - E(a_t^2) \\ &= E\left[E(\alpha_t^2 \mid Y_{t-1})\right] - E(E(\alpha_t \mid Y_{t-1})^2) \\ &= E\left[E(\alpha_t^2 \mid Y_{t-1}) - E(\alpha_t \mid Y_{t-1})^2\right] \\ &= E\left[\mathrm{var}(\alpha_t \mid Y_{t-1})\right] \\ &= E\left[P_t\right] \\ &= P_t \end{aligned}$$

$$
\begin{aligned}
v_t &= y_t - a_t \\
&= \alpha_t + \varepsilon_t - a_t \\
&= \alpha_t - a_t + \varepsilon_t \\
&= x_t + \varepsilon_t
\end{aligned}
$$

$$
\begin{aligned}
x_{t+1} &= \alpha_{t+1} - a_{t+1} \\
&= \alpha_t + \eta_t - a_t - K_t v_t \\
&= \alpha_t - a_t + \eta_t - K_t v_t \\
&= x_t + \eta_t - K_t (x_t + \varepsilon_t) \\
&= (1 - K_t) x_t + \eta_t - K_t \varepsilon_t \\
&= L_t x_t + \eta_t - K_t \varepsilon_t
\end{aligned}
$$

donde $L_t = 1 - K_t = \sigma_\varepsilon^2 / F_t$.

Así, análogamente a las relaciones

$$y_t = \alpha_t + \varepsilon_t, \qquad \alpha_{t+1} = \alpha_t + \eta_t,$$

Se tienen las relaciones de error

$$v_t = x_t + \varepsilon_t, \qquad x_{t+1} = L_t x_t + \eta_t - K_t \varepsilon_t, \qquad t = 1, \dots, n,$$

Con $x_1 = \alpha_1 - a_1$.

Como

$$
\begin{aligned}
P_{t+1} = \operatorname{cov}(\alpha_{t+1}, v_{t+1}) &= \operatorname{cov}(\alpha_{t+1}, x_{t+1} + \varepsilon_{t+1}) \\
&= \operatorname{cov}(\alpha_{t+1}, x_{t+1}) - \operatorname{cov}(\alpha_{t+1}, \varepsilon_{t+1}) \\
&= \operatorname{cov}(\alpha_{t+1}, x_{t+1})
\end{aligned}
$$

Entonces,

$$
\begin{aligned}
P_{t+1} &= \operatorname{var}(x_{t+1}) = \operatorname{cov}(x_{t+1}, \alpha_{t+1}) = \operatorname{cov}(x_{t+1}, \alpha_t + \eta_t) \\
&= \operatorname{cov}(L_t x_t + \eta_t - K_t \varepsilon_t, \alpha_t + \eta_t) \\
&= \operatorname{cov}(L_t x_t, \alpha_t + \eta_t) + \operatorname{cov}(\eta_t, \alpha_t + \eta_t) - \operatorname{cov}(K_t \varepsilon_t, \alpha_t + \eta_t) \\
&= \operatorname{cov}(L_t x_t, \alpha_t) + \operatorname{cov}(L_t x_t, \eta_t) + \operatorname{cov}(\eta_t, \alpha_t) + \operatorname{cov}(\eta_t, \eta_t) - \operatorname{cov}(K_t \varepsilon_t, \alpha_t) - \operatorname{cov}(K_t \varepsilon_t, \eta_t) \\
&= L_t \operatorname{cov}(x_t, \alpha_t) + L_t \operatorname{cov}(x_t, \eta_t) + \operatorname{cov}(\eta_t, \alpha_t) + \operatorname{cov}(\eta_t, \eta_t) - K_t \operatorname{cov}(\varepsilon_t, \alpha_t) - K_t \operatorname{cov}(\varepsilon_t, \eta_t) \\
&= L_t P_t + \sigma_\eta^2
\end{aligned}
$$

Suavización del estado

Ahora se considerará la estimación de $\alpha_1,\dots,\alpha_n$ dada la muestra completa Y_n.
Dado que todas las distribuciones son normales, la densidad condicional de α_t dado y es $N(\hat{\alpha}_t, V_t)$.

Los errores de pronóstico $v_1,\dots,v_n$ son mutuamente independientes y son una transformación lineal de $y_1,\dots,y_n$, además, $v_1,\dots,v_n$ son independientes de $y_1,\dots,y_{t-1}$ con medias cero.

$$
\begin{aligned}
\hat{\alpha}_t &= E(\alpha_t \mid y) = E(\alpha_t \mid Y_{t-1}, v_t, \dots, v_n) \\
&= E(\alpha_t \mid Y_{t-1}) + \operatorname{cov}\left[\alpha_t, (v_t,\dots,v_n)'\right]\operatorname{var}\left[(v_t,\dots,v_n)'\right]^{-1}(v_t,\dots,v_n)' \\
&= a_t + \begin{pmatrix} \operatorname{cov}(\alpha_t, v_j) \\ \vdots \\ \operatorname{cov}(\alpha_t, v_n) \end{pmatrix}' \begin{bmatrix} F_t & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & F_n \end{bmatrix}^{-1} \begin{pmatrix} v_t \\ \vdots \\ v_n \end{pmatrix} \\
&= a_t + \left[\operatorname{cov}(\alpha_t, v_j) \quad \cdots \quad \operatorname{cov}(\alpha_t, v_n)\right] \begin{bmatrix} \frac{1}{F_t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{F_n} \end{bmatrix} \begin{pmatrix} v_t \\ \vdots \\ v_n \end{pmatrix} \\
&= a_t + \left[\operatorname{cov}(\alpha_t, v_j) \quad \cdots \quad \operatorname{cov}(\alpha_t, v_n)\right] \begin{pmatrix} \frac{v_t}{F_t} \\ \vdots \\ \frac{v_n}{F_n} \end{pmatrix} \\
&= a_t + \sum_{j=t}^{n} \operatorname{cov}(\alpha_t, v_j) F_j^{-1} v_j \qquad (4)
\end{aligned}
$$

Como a_t depende de v_j solo para $j = 1,\dots,t\text{-}1$, entonces $\operatorname{cov}(a_t, v_j) = 0$ para $j = t,\dots,n$ y

$$\operatorname{cov}(x_t, v_j) = \operatorname{cov}(\alpha_t - a_t, v_j) = \operatorname{cov}(\alpha_t, v_j) - \operatorname{cov}(a_t, v_j) = \operatorname{cov}(\alpha_t, v_j) \text{ para } j = t,\dots,n$$

Como $E(v_t)=0$ para todo t, entonces

$$
\begin{aligned}
\operatorname{cov}(x_t, v_t) &= E(x_t v_t) - E(x_t)E(v_t) = E(x_t v_t) = E\left[x_t(x_t + \varepsilon_t)\right] \\
&= E(x_t^2) + E(x_t \varepsilon_t) = E(x_t^2) = \operatorname{var}(x_t) = P_t
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{cov}(x_t, v_{t+1}) &= E(x_t v_{t+1}) = E\left[x_t(x_{t+1} + \varepsilon_{t+1})\right] = E(x_t x_{t+1}) = E\left[x_t(L_t x_t + \eta_t - K_t \varepsilon_t)\right] \\
&= L_t E(x_t^2) + E(x_t \eta_t) - K_t E(x_t \varepsilon_t) = L_t E(x_t^2) = L_t \operatorname{var}(x_t) = L_t P_t
\end{aligned}
$$

$$\begin{aligned}\operatorname{cov}(x_t, v_{t+2}) &= E(x_t v_{t+2}) = E\left[x_t (x_{t+2} + \varepsilon_{t+2})\right] = E(x_t x_{t+2}) = E\left[x_t (L_{t+1} x_{t+1} + \eta_{t+1} - K_{t+1}\varepsilon_{t+1})\right] \\ &= L_{t+1} E(x_t x_{t+1}) = L_{t+1} L_t P_t\end{aligned}$$

$$\begin{aligned}\operatorname{cov}(x_t, v_{t+3}) &= E(x_t v_{t+3}) = E\left[x_t (x_{t+3} + \varepsilon_{t+3})\right] = E(x_t x_{t+3}) = E\left[x_t (L_{t+2} x_{t+2} + \eta_{t+2} - K_{t+2}\varepsilon_{t+2})\right] \\ &= L_{t+2} E(x_t x_{t+2}) = L_{t+2} L_{t+1} L_t P_t\end{aligned}$$

y en general, $\operatorname{cov}(x_t, v_n) = P_t L_t L_{t+1} \cdots L_{n-1}$

Sustituyendo en (4)

$$\begin{aligned}\hat{\alpha}_t &= a_t + P_t \frac{v_t}{F_t} + P_t L_t \frac{v_{t+1}}{F_{t+1}} + P_t L_t L_{t+1} \frac{v_{t+2}}{F_{t+2}} + \cdots \\ &= a_t + P_t r_{t-1}\end{aligned}$$

donde

$$r_{t-1} = \frac{v_t}{F_t} + L_t \frac{v_{t+1}}{F_{t+1}} + L_t L_{t+1} \frac{v_{t+2}}{F_{t+2}} + L_t L_{t+1} L_{t+2} \frac{v_{t+3}}{F_{t+3}} + \cdots + L_t L_{t+1} \ldots L_{n-1} \frac{v_n}{F_n}$$

$$r_t = \frac{v_{t+1}}{F_{t+1}} + L_{t+1} \frac{v_{t+2}}{F_{t+2}} + L_{t+1} L_{t+2} \frac{v_{t+3}}{F_{t+3}} + \cdots + L_{t+1} L_{t+2} \ldots L_{n-1} \frac{v_n}{F_n} \tag{5}$$

$$r_{t-1} = \frac{v_t}{F_t} + L_t r_t \tag{6}$$

$$r_{t-1} = F_t^{-1} v_t + L_t r_t, \qquad \hat{\alpha}_t = a_t + P_t r_{t-1}, \qquad t = n, \ldots, 1,$$

$$\begin{aligned}V_t &= \operatorname{var}(\alpha_t \mid y) = \operatorname{var}(\alpha_t \mid Y_{t-1}, v_t, \ldots, v_n) \\ &= \operatorname{var}(\alpha_t \mid Y_{t-1}) - \operatorname{cov}\left[\alpha_t, (v_t, \ldots, v_n)'\right] \operatorname{var}\left[(v_t, \ldots, v_n)'\right]^{-1} \operatorname{cov}\left[\alpha_t, (v_t, \ldots, v_n)'\right]'\end{aligned}$$

$$= P_t - \begin{pmatrix}\operatorname{cov}(\alpha_t, v_j) \\ \vdots \\ \operatorname{cov}(\alpha_t, v_n)\end{pmatrix}' \begin{bmatrix} F_t & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & F_n \end{bmatrix}^{-1} \begin{pmatrix}\operatorname{cov}(\alpha_t, v_j) \\ \vdots \\ \operatorname{cov}(\alpha_t, v_n)\end{pmatrix}$$

$$= P_t - \left[\mathrm{cov}(\alpha_t, v_j) \quad \cdots \quad \mathrm{cov}(\alpha_t, v_n)\right] \begin{bmatrix} \frac{1}{F_t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{F_n} \end{bmatrix} \begin{pmatrix} \mathrm{cov}(\alpha_t, v_t) \\ \vdots \\ \mathrm{cov}(\alpha_t, v_n) \end{pmatrix}$$

$$= P_t + \left[\mathrm{cov}(\alpha_t, v_j) \quad \cdots \quad \mathrm{cov}(\alpha_t, v_n)\right] \begin{pmatrix} \mathrm{cov}(\alpha_t, v_t) F_t^{-1} \\ \vdots \\ \mathrm{cov}(\alpha_t, v_n) F_n^{-1} \end{pmatrix}$$

$$= P_t + \sum_{j=t}^{n} \left[\mathrm{cov}(\alpha_t, v_j)\right]^2 F_j^{-1}$$

$$V_t = P_t - P_t^2 \frac{1}{F_t} - P_t^2 L_t^2 \frac{1}{F_{t+1}} - P_t^2 L_t^2 L_{t+1}^2 \frac{1}{F_{t+2}} - \cdots - P_t^2 L_t^2 L_{t+1}^2 \cdots L_{n-1}^2 \frac{1}{F_n}$$

$$= P_t + P_t^2 N_{t-1}$$

donde

$$N_{t-1} = \frac{1}{F_t} + L_t^2 \frac{1}{F_{t+1}} + L_t^2 L_{t+1}^2 \frac{1}{F_{t+2}} + L_t^2 L_{t+1}^2 L_{t+2}^2 \frac{1}{F_{t+3}} + \cdots + L_t^2 L_{t+1}^2 \cdots L_{n-1}^2 \frac{1}{F_n}$$

$$N_t = \frac{1}{F_{t+1}} + L_{t+1}^2 \frac{1}{F_{t+2}} + L_{t+1}^2 L_{t+2}^2 \frac{1}{F_{t+3}} + \cdots + L_{t+1}^2 L_{t+2}^2 \cdots L_{n-1}^2 \frac{1}{F_n} \qquad (7)$$

$$N_{t-1} = \frac{1}{F_t} + L_t^2 N_t \qquad (8)$$

$$N_{t-1} = F_t^{-1} + L_t^2 N_t, \qquad \mathrm{V}_t = P_t - P_t^2 N_{t-1}, \qquad t = n, \ldots, 1,$$

$$\mathrm{var}(r_t) = \frac{1}{F_{t+1}} \mathrm{var}(v_{t+1}) + \frac{L_{t+1}^2}{F_{t+2}^2} \mathrm{var}(v_{t+2}) + \frac{L_{t+1}^2 L_{t+2}^2}{F_{t+3}^2} \mathrm{var}(v_{t+3}) + \cdots + \frac{L_{t+1}^2 L_{t+2}^2 \ldots L_{n-1}^2}{F_n^2} \mathrm{var}(v_n)$$

$$= \frac{1}{F_{t+1}^2} F_{t+1} + \frac{L_{t+1}^2}{F_{t+2}^2} F_{t+2} + \frac{L_{t+1}^2 L_{t+2}^2}{F_{t+3}^2} F_{t+3} + \cdots + \frac{L_{t+1}^2 L_{t+2}^2 \ldots L_{n-1}^2}{F_n^2} F_n$$

$$= \frac{1}{F_{t+1}} + L_{t+1}^2 \frac{1}{F_{t+2}} + L_{t+1}^2 L_{t+2}^2 \frac{1}{F_{t+3}} + \cdots + L_{t+1}^2 L_{t+2}^2 \cdots L_{n-1}^2 \frac{1}{F_n}$$

$$= N_t$$

Suavización de la perturbación del error $\hat{\varepsilon}_t = E(\varepsilon_t \mid y)$

$$\hat{\varepsilon}_t = \sigma_\varepsilon^2 u_{t,} \qquad t = n, \ldots, 1. \qquad (9)$$

Donde, $u_t = F_t^{-1} v_t - K_t r_t$, y donde la recursión para r_t está dada por (6).

El escalar u_t es llamado error de suavización.

La varianza suavizada $\operatorname{var}(\varepsilon_t \mid y)$ es obtenido por

$\operatorname{var}(\varepsilon_t \mid y) = \sigma_\varepsilon^2 - \sigma_\varepsilon^4 D_{t,} \qquad t = n, \ldots, 1.$, donde,

$D_t = F_t^{-1} + K_t^2 N_t$, y donde la recursión para N_t está dada por (8).

De (5), v_t es independiente de r_t, y $\operatorname{Var}(r_t) = N_t$, tenemos

$$\begin{aligned}\operatorname{var}(u_t) &= \operatorname{var}(F_t^{-1} v_t - K_t r_t) = F_t^{-2} \operatorname{var}(v_t) + K_t^2 \operatorname{var}(r_t) = F_t^{-2} F_t + K_t^2 N_t \\ &= F_t^{-1} + K_t^2 N_t = D_t\end{aligned}$$

en consecuencia, de (9) se obtiene $\operatorname{var}(\hat{\varepsilon}_t) = \sigma_\varepsilon^4 \operatorname{var}(u_t) = \sigma_\varepsilon^4 D_t$

Los métodos para calcular $\hat{\alpha}_t$ y $\hat{\varepsilon}_t$ son consistentes, desde que

$$1 - K_t = 1 - \frac{P_t}{F_t} = \frac{F_t - P_t}{F_t} = \frac{P_t + \sigma_\varepsilon^2 - P_t}{F_t} = \sigma_\varepsilon^2 F_t^{-1}$$

$$\begin{aligned}\hat{\varepsilon}_t &= y_t - \hat{\alpha}_t \\ &= y_t - a_t - P_t r_{t-1} \\ &= v_t - P_t (F_t^{-1} v_t + L_t r_t) \\ &= v_t - P_t \left[F_t^{-1} v_t + (1 - K_t) r_t\right] \\ &= v_t - P_t (F_t^{-1} v_t + \sigma_\varepsilon^2 F_t^{-1} r_t) \\ &= v_t - P_t F_t^{-1} v_t - \sigma_\varepsilon^2 P_t F_t^{-1} r_t \\ &= F_t^{-1} v_t (F_t - P_t) - \sigma_\varepsilon^2 P_t F_t^{-1} r_t \\ &= F_t^{-1} v_t (P_t + \sigma_\varepsilon^2 - P_t) - \sigma_\varepsilon^2 K_t r_t \\ &= F_t^{-1} v_t \sigma_\varepsilon^2 - \sigma_\varepsilon^2 K_t r_t \\ &= \sigma_\varepsilon^2 (F_t^{-1} v_t - K_t r_t) \qquad t = n, \ldots, 1.\end{aligned}$$

La media suavizada de la perturbación $\hat{\eta}_t = E(\eta_t \mid y)$ es calculada por

$\hat{\eta}_t = \sigma_\varepsilon^2 r_{t,}$ $\quad t = n, \dots, 1.$ donde la recursión para r_t es dada por (6).

La varianza suavizada del error $\mathrm{var}(\eta_t \mid y)$ está dada por

$\mathrm{var}(\eta_t \mid y) = \sigma_\eta^2 - \sigma_\eta^4 N_{t,}$ $\quad t = n, \dots, 1.$ donde la recursión para N_t está dada por (8).

Dado que Var(r_t) = N_t, tenemos que $\mathrm{var}(\hat{\eta}_t) = \sigma_\eta^4 \ \mathrm{var}(r_t) = \sigma_\eta^4 N_t$.

El método para calcular $\hat{\eta}_t$ es consistente con la definición $\eta_t = \alpha_{t+1} - \alpha_t$, pues

$$
\begin{aligned}
\hat{\eta}_t &= \hat{\alpha}_{t+1} - \hat{\alpha}_t \\
&= a_{t+1} + P_{t+1} r_t - a_t + P_t r_{t-1} \\
&= a_t + K_t v_t - a_t + P_{t+1} r_t + P_t r_{t-1} \\
&= a_t + K_t v_t - a_t + \left[P_t(1-K_t) + \sigma_\eta^2\right] r_t + P_t(F_t^{-1} v_t + L_t r_t) \\
&= a_t + K_t v_t - a_t + (P_t L_t + \sigma_\eta^2) r_t + P_t(F_t^{-1} v_t + L_t r_t) \\
&= a_t + K_t v_t - a_t + P_t L_t r_t + \sigma_\eta^2 r_t + P_t F_t^{-1} v_t + P_t L_t r_t \\
&= a_t + K_t v_t - a_t + P_t L_t r_t + \sigma_\eta^2 r_t + K_t v_t + P_t L_t r_t \\
&= \sigma_\eta^2 r_t
\end{aligned}
$$

Se considerará ahora el cálculo de $\hat{\varepsilon}_t = E(\varepsilon_t \mid y)$ por regresión directa de $\varepsilon_t = (\varepsilon_1, \dots, \varepsilon_n)'$ en el vector y para obtener $\hat{\varepsilon}_t = (\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_n)'$

$$
\begin{aligned}
\hat{\varepsilon} = E(\varepsilon \mid y) &= E(\varepsilon) + \mathrm{cov}(\varepsilon, y)\, \mathrm{var}(y)^{-1}\left[y - E(y)\right] \\
&= \mathrm{cov}(\varepsilon, y)\Omega^{-1}\left[y - \mathbf{1} a_1\right]
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{cov}(\varepsilon_i, y_j) &= E(\varepsilon_i y_j) - E(\varepsilon_i)E(y_j) = E(\varepsilon_i y_j) \\
&= E\left[\varepsilon_i\left(\alpha_1 + \sum_{t=1}^{j-1}\eta_t + \varepsilon_j\right)\right] \\
&= E(\varepsilon_i \alpha_1) + E\left(\varepsilon_i \sum_{t=1}^{j-1}\eta_t\right) + E(\varepsilon_i \varepsilon_j) \\
&= E(\varepsilon_i)E(\alpha_1) + E(\varepsilon_i)\sum_{t=1}^{j-1} E(\eta_t) + E(\varepsilon_i \varepsilon_j) \\
&= E(\varepsilon_i \varepsilon_j)
\end{aligned}
$$

Si $i \neq j$, $\mathrm{cov}(\varepsilon_i, y_j) = E(\varepsilon_i \varepsilon_j) = E(\varepsilon_i)E(\varepsilon_j) = 0$.

Si $i = j$, $\mathrm{cov}(\varepsilon_i, y_j) = E(\varepsilon_i^2) = \sigma_\varepsilon^2$

Por tanto $\mathrm{cov}(\varepsilon_i, y_j) = \begin{bmatrix} \sigma_\varepsilon^2 & 0 & & 0 \\ 0 & \sigma_\varepsilon^2 & & 0 \\ & & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_\varepsilon^2 \end{bmatrix} = \sigma_\varepsilon^2 \begin{bmatrix} 1 & 0 & & 0 \\ 0 & 1 & & 0 \\ & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \sigma_\varepsilon^2 \mathbf{I_n}$

De la descomposición de Cholesky considerada anteriormente, se tiene $\Omega^{-1} = C'F^{-1}C$ y $C(y - \mathbf{1}a_1) = v$. Por consiguiente, $\hat{\varepsilon} = \sigma_\varepsilon^2 I_n C'F^{-1}C(y - \mathbf{1}a_1) = \sigma_\varepsilon^2 C'F^{-1}v$

$$C'F^{-1}v = \begin{bmatrix} 1 & c_{21} & c_{31} & & c_{n1} \\ 0 & 1 & c_{32} & & c_{n2} \\ 0 & 0 & 1 & & c_{n3} \\ & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{F_1} & 0 & & 0 \\ 0 & \frac{1}{F_2} & & 0 \\ & & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{F_n} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$= \begin{bmatrix} 1 & c_{21} & c_{31} & & c_{n1} \\ 0 & 1 & c_{32} & & c_{n2} \\ 0 & 0 & 1 & & c_{n3} \\ & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \frac{v_1}{F_1} \\ \frac{v_2}{F_2} \\ \vdots \\ \frac{v_n}{F_n} \end{bmatrix}$$

El término t de el vector $C'F^{1}v$ es:

$$\left[C'F^{-1}v\right]_t = \frac{v_t}{F_t} - K_t \frac{v_{t+1}}{F_{t+1}} - (1-K_{t+1})K_t \frac{v_{t+2}}{F_{t+2}} - (1-K_{t+2})(1-K_{t+1})K_t \frac{v_{t+3}}{F_{t+3}} - \cdots -$$

$$-(1-K_{n-2})\cdots(1-K_{t+2})(1-K_{t+1})K_t \frac{v_{n-1}}{F_{n-1}} - (1-K_{n-1})\cdots(1-K_{t+2})(1-K_{t+1})K_t \frac{v_n}{F_n}$$

$$= \frac{v_t}{F_t} - K_t \frac{v_{t+1}}{F_{t+1}} - L_{t+1}K_t \frac{v_{t+2}}{F_{t+2}} - L_{t+2}L_{t+1})K_t \frac{v_{t+3}}{F_{t+3}} - \cdots -$$

$$-L_{n-2}\cdots L_{t+2}L_{t+1}K_t \frac{v_{n-1}}{F_{n-1}} - L_{n-1}\cdots L_{t+2}L_{t+1}K_t \frac{v_n}{F_n}$$

$$= \frac{v_t}{F_t} - K_t \left[\frac{v_{t+1}}{F_{t+1}} + L_{t+1} \frac{v_{t+2}}{F_{t+2}} + L_{t+1}L_{t+2} \frac{v_{t+3}}{F_{t+3}} + \cdots + L_{t+1}L_{t+2}\cdots L_{n-2} \frac{v_{n-1}}{F_{n-1}} + L_{t+1}L_{t+2}\cdots L_{n-1} \frac{v_n}{F_n} \right]$$

$$= v_t F_t^{-1} - K_t r_t$$

$$= u_t$$

Así,

$$\hat{\varepsilon} = \sigma_{\varepsilon}^{2} u , \qquad u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

Donde

$$u = C'F^{-1}v \quad \text{con} \quad v = C(y - \mathbf{1}a_1)$$

Esto lleva a que

$u = C'F^{-1}C(y - \mathbf{1}a_1) = \Omega^{-1}(y - \mathbf{1}a_1)$
Donde $\Omega = \text{var}(y)$ y $F = C\Omega C'$

Observaciones Faltantes

Una ventaja considerable de el enfoque espacio estado es la facilidad con la que pueden ser tratadas las observaciones faltantes.

Supóngase que se tieneun modelo local level en el que no se tienen las observaciones y_j, con $j = \tau,\ldots,\tau^*-1$, para $1 < \tau < \tau^* \leq n$. La forma más obvia de tratar este problema es definir nuevas series y_t^* donde $y_t^* = y_t$ para t = 1,…,τ-1 y $y_t^* = y_{t+\tau^*-\tau}$ para $t = t,\ldots,n^*$ con $n^* = n-(\tau^*-\tau)$.

Como $y_\tau^* = y_{\tau^*} = \alpha_{\tau^*} + \varepsilon_{\tau^*}$, y

$$\begin{aligned} \alpha_{\tau^*} &= \alpha_{\tau^*-1} + \eta_{\tau^*-1} \\ &= \alpha_{\tau^*-2} + \eta_{\tau^*-2} + \eta_{\tau^*-1} \\ &\vdots \\ &= \alpha_{\tau} + \eta_{\tau} + \cdots + \eta_{\tau^*-2} + \eta_{\tau^*-1} \\ &= \alpha_{\tau-1} + \eta_{\tau-1} + \eta_{\tau} + \cdots + \eta_{\tau^*-2} + \eta_{\tau^*-1} \\ &= \alpha_{\tau-1} + \eta_{\tau-1}^* \end{aligned}$$

Donde

$$\begin{aligned} E(\eta_{\tau-1}^*) &= E(\eta_{\tau-1} + \eta_{\tau} + \cdots + \eta_{\tau^*-2} + \eta_{\tau^*-1}) \\ &= E(\eta_{\tau-1}) + E(\eta_{\tau}) + \cdots + E(\eta_{\tau^*-2}) + E(\eta_{\tau^*-1}) \\ &= 0 \end{aligned}$$

Y

$$\begin{aligned}\operatorname{var}(\eta^*_{\tau-1}) &= \operatorname{var}(\eta_{\tau-1} + \eta_\tau + \cdots + \eta_{\tau^*-2} + \eta_{\tau^*-1}) \\ &= \operatorname{var}(\eta_{\tau-1}) + \operatorname{var}(\eta_\tau) + \cdots + \operatorname{var}(\eta_{\tau^*-2}) + \operatorname{var}(\eta_{\tau^*-1}) \\ &= \sigma^2_\eta + \sigma^2_\eta + \cdots + \sigma^2_\eta + \sigma^2_\eta \\ &= (\tau^* - \tau)\sigma^2_\eta\end{aligned}$$

Entonces, el modelo para y^*_t la escala de tiempo $t = t, \ldots, n^*$ es el mismo que en (1), excepto que $\alpha_\tau = \alpha_{\tau-1} + \eta_{\tau-1}$ donde $\eta^*_{\tau-1} \sim N\left[0, (\tau^* - \tau)\sigma^2_\eta\right]$.

Sin embargo es más fácil y más transparente proceder usando el dominio de tiempo original, de la siguiente manera.

En el tiempo t = $\tau, \ldots, \tau^*$-1, se tiene

$$\begin{aligned}\alpha_t &= \alpha_{t-1} + \eta_{t-1} \\ &= \alpha_{t-2} + \eta_{t-2} + \eta_{t-1} \\ &\vdots \\ &= \alpha_\tau + \eta_\tau + \cdots + \eta_{t-2} + \eta_{t-1} \\ &= \alpha_\tau + \sum_{j=\tau}^{t-1} \eta_j\end{aligned}$$

Entonces

$$\begin{aligned}a_t = E(\alpha_t \mid Y_{t-1}) = E(\alpha_t \mid Y_{\tau-1}) = E\left(\alpha_\tau + \sum_{j=\tau}^{t-1} \eta_j \middle| Y_{\tau-1}\right) &\\ = E(\alpha_\tau \mid Y_{\tau-1}) + \sum_{j=\tau}^{t-1} E(\eta_j \mid Y_{\tau-1}) &\\ = E(\alpha_\tau \mid Y_{\tau-1}) &\\ = a_\tau &\end{aligned}$$

Y

$$\begin{aligned}P_t = \operatorname{var}(\alpha_t \mid Y_{t-1}) = \operatorname{var}(\alpha_t \mid Y_{\tau-1}) = \operatorname{var}\left(\alpha_\tau + \sum_{j=\tau}^{t-1} \eta_j \middle| Y_{\tau-1}\right) &\\ = \operatorname{var}(\alpha_\tau \mid Y_{\tau-1}) + \sum_{j=\tau}^{t-1} \operatorname{var}(\eta_j \mid Y_{\tau-1}) &\\ = P_\tau + \sum_{j=\tau}^{t-1} \sigma^2_\eta &\\ = P_\tau + (t - \tau)\sigma^2_\eta &\end{aligned}$$

Pero $a_{t+1} = a_\tau = a_t$

y

$$
\begin{aligned}
P_{t+1} &= P_\tau + (t+1-\tau)\sigma_\eta^2 \\
&= P_\tau + (t-\tau)\sigma_\eta^2 + \sigma_\eta^2 \\
&= P_t + \sigma_\eta^2
\end{aligned}
$$

Obteniendo las siguientes ecuaciones del filtro de Kalman

$$a_{t+1} = a_t, \qquad P_{t+1} = P_t + \sigma_\eta^2, \qquad t = \tau, \ldots, \tau^* - 1 \qquad (10)$$

Los demás valores de a_t y P_t son dados por (3) para $t = 1,\ldots,\tau$ y $t = t^*,\ldots,n$. Se puede entonces usar el filtro original (3) para todo t tomando $v_t = 0$ y $K_t = 0$ en los puntos de tiempo faltantes. El mismo procedimiento puede ser usado cuando hay mas de un grupo de observaciones faltantes.

Las recursiones de error en los puntos de tiempo faltantes son

$$v_t = x_t + \varepsilon_t, \qquad x_{t+1} = x_t + \eta_t, \qquad t = \tau, \ldots, \tau^* - 1$$

Dado que $K_t = 0$ y por lo tanto $L_t = 1$. Las covarianzas entre el estado en los puntos faltantes y las innovaciones después del periodo faltante están dadas por

$$
\begin{aligned}
&\operatorname{cov}(\alpha_t, v_{\tau^*}) = P_t L_t L_{t+1} \cdots L_{\tau^*-1} = P_t \cdot 1 \cdot 1 \cdot \cdots \cdot 1 = P_t \\
&\operatorname{cov}(\alpha_t, v_j) = P_t L_t L_{t+1} \cdots L_{j-1} = P_t \cdot 1 \cdot \cdots \cdot 1 \cdot L_{\tau^*} L_{\tau^*+1} \cdots L_{j-1} = P_t L_{\tau^*} L_{\tau^*+1} \cdots L_{j-1},
\end{aligned}
$$

para $j = \tau^* + 1, \ldots, n, \qquad t = \tau, \ldots, \tau^* - 1$

Como $v_t = 0$ para t = $\tau,\ldots,\tau^*$-1, entonces, de (4)

$$\hat{\alpha}_t = a_t + \sum_{j=\tau^*}^{n} \operatorname{cov}(\alpha_t, v_j) F_j^{-1} v_j \qquad \text{para t} = \tau,\ldots,\tau^*\text{-1}$$

$$r_t = r_{t-1}, \qquad \hat{\alpha}_t = a_t + P_t r_{t-1}, \qquad t = \tau, \ldots, \tau^* - 1$$

Pronóstico

Sea $\bar{y}_{n+j}$ el pronóstico de y_{n+j} que da el error cuadrático medio mínimo, dada la serie de tiempo $y_1,\ldots,y_n$ para $j = 1,2,\ldots,J$ con J algún entero positivo pre-definido. El pronóstico del mínimo error cuadrático medio es la función $\bar{y}_{n+j}$ de $y_1,\ldots,y_n$ que minimiza $E\left[(y_{n+j} - \bar{y}_{n+j})^2 \mid Y_n\right]$. Entonces $\bar{y}_{n+j} = E(y_{n+j} \mid Y_n)$. Esto sigue del resultado de que si x es una variable aleatoria con media μ el valor de λ que minimiza $E(x-\lambda)^2$ es $\lambda = \mu$. La varianza del error de pronóstico es denotada por $\bar{F}_{n+j} = \operatorname{var}(y_{n+j} \mid Y_n)$.

El pronóstico para el modelo local level consiste en filtrar las observaciones $y_1,\ldots,y_n,y_{n+1},\ldots,y_{n+J}$ usando la recursión (3) y tratando las últimas J observaciones $y_{n+1},\ldots,y_{n+J}$ como faltantes.

Tomando $\bar{a}_{n+j} = E(\alpha_{n+j} \mid Y_n)$ y $\bar{P}_{n+j} = \text{var}(\alpha_{n+j} \mid Y_n)$, se sigue de la ecuación (10)

$$\bar{y}_{n+j} = \bar{a}_{n+j}, \qquad \bar{F}_{n+j} = \bar{P}_{n+j} + \sigma_\varepsilon^2$$

$$\bar{a}_{n+j+1} = \bar{a}_{n+j}, \qquad \bar{P}_{n+j+1} = \bar{P}_{n+j} + \sigma_\eta^2, \qquad j = 1,\ldots,J-1$$

$$P_1 \to \infty$$

Implementación.

opsdjhósdjfhjgfnh

Bibliografía.

- J. Durbin and S. Koopman, Time series análisis by state space methods. Oxford Statistical Science Series. 2005.