UNA NOTA ACERCA DE LA DIFERENCIA ENTRE LOS MODELOS DE ESTADOS GAUSSIANOS CONVENCIONAL Y CONDICIONAL

por

Fabio H. Nieto*

Resumen

Nieto, F. H.: Una nota acerca de la diferencia entre los modelos de estados gaussianos convencional y condicional. Rev. Acad Colomb. Cienc. **19** (72): 117 - 119. 1994 ISSN 0370-3908.

En ocasiones las hipótesis estadísticas convencionales de un modelo de estados para una serie cronológica se deben reemplazar por hipótesis condicionales sobre cierta información conocida. En este artículo se destaca la diferencia básica entre los dos conjuntos de hipótesis y se demuestra que las ecuaciones del Filtro de Kalman son válidas en ambos casos.

Abstract

Sometimes the conventional statistical hypotheses of a state space model for a time series must be replaced for conditional hypotheses on some known information. In this paper, the basic distintion betwen the two sets of hypotheses is emphasized and the fact that the Kalman Filter equations are valid in both cases is shown.

1. Introducción.

En algunos libros o artículos sobre modelos de estados o el Filtro de Kalman (FK) se considera solo el modelo en el cual las matrices de covarianzas de los términos de error son no singulares y la distribución de probabilidad de ellos es multinormal (no singular), ver por ejemplo Harvey (1989) o **West y Harrison** (1989), quizás los más completos y clásicos sobre el tema. En otros, estas hipótesis son debilitadas en el sentido de permitir matrices de covarianzas singulares, tal es el caso por ejemplo de **Anderson y Moore** (1979), **Kohn y Ansley** (1983), **Catlin** (1989) y **Aoki** (1990), pero en ellos no se utiliza la teoría de la distribución multinormal singular para deducir el filtro. Recientemente **Nieto** (1993) considera el modelo en el cual se supone que los términos de error tienen distribución multinormal *singular* y se hace la deducción del FK bajo esta hipótesis fundamental. Este modelo es una extensión del de **Harvey** (1989) o del de **West y Harrison** (1989) y salvo por distribuciones, cubre también el caso del modelo de **Anderson y Moore, Kohn y Ansley, Catlin y Aoki.** Para futura referencia, el modelo tratado por Nieto (1993) será llamado *Gaussiano convencional* o simplemente *convencional.*

En ocasiones, es necesario considerar que la distribución de un término de error (o ambos) en un tiempo t en un modelo de estados está *condicionada* por la información observada hasta el tiempo t-1 tal como es tratado por **Harvey** (1989) y **West y Harrison** (1989). En esta forma surge el modelo *condicional.* **Harvey** afirma que las ecuaciones del FK del caso convencional son todavía válidas pero no se demuestra, ni se indica una justificación sobre la afirmación. Tampoco se indican referencias al respecto. Más aún, **Harvey** considera solo un modelo con matrices de covarianzas definidas positivas. De otro lado **West y Harrison** si presenta la deducción del FK pero en el contexto bayesiano y suponiendo matrices de covarianzas también definidas positivas.

* Profesor Asociado, Departamento de Matemáticas y Estadística, Universidad Nacional de Colombia, Santafé de Bogotá, D. C.

En este trabajo, se destaca la *diferencia* entre las hipótesis del modelo convencional y el modelo condicional y se demuestra que el FK es válido también en el último caso.

2. Modelos de estados Gaussianos condicionales.

Usando la notación de **Harvey** (1989), pp. 156, se define aquí también el modelo Gaussiano *condicional* o simplemente condicional por medio de las ecuaciones

$$\left.\begin{aligned} Y_t &= Z_t(\mathbf{Y}_{t-1})\,\alpha_t + \varepsilon_t \\ \alpha_t &= T_t(\mathbf{Y}_{t-1})\,\alpha_{t-1} + \eta_t \end{aligned}\right\} \quad t = 1, 2, \ldots \qquad (1)$$

donde α_t, Y_t, ε_t, y η_t son vectores aleatorios, $Z_t(\mathbf{Y}_{t-1})$, $T_t(\mathbf{Y}_{t-1})$ son matrices que dependen determinísticamente del vector $\mathbf{Y}_{t-1} = (Y_1, \ldots, Y_{t-1})'$ donde " ' " denota transposición. Las hipótesis probabilísticas básicas sobre los vectores aleatorios ε_t, y η_t son las siguientes:

(i) $\varepsilon_t \mid \mathbf{Y}_{t-1} \sim N_s(0, H_t(\mathbf{Y}_{t-1}))$ y $\eta_t \sim N_s(0, Q_t(\mathbf{Y}_{t-1}))$ para cada t = 1, 2, ..., donde N_s significa *multinormal singular* (Ver **Nieto** (1993) o **Anderson** (1984)).

(ii) $\alpha_t \mid \mathbf{Y}_{t-1}$ y $\varepsilon_t \mid \mathbf{Y}_{t-1}$, y, $\alpha_{t-1} \mid \mathbf{Y}_{t-1}$ y $\eta_t \mid \mathbf{Y}_{t-1}$ son, cada par, independientes estocásticamente para cada t = 1, 2, ...

(iii) $\alpha_0 \sim N_s(a_0, P_0)$, a_0 y P_0 conocidos.

Observación: aunque no se dice explícitamente, se supone que $Z_t(\mathbf{Y}_{t-1})$, $T_t(\mathbf{Y}_{t-1})$, $H_t(\mathbf{Y}_{t-1})$ y $Q_t(\mathbf{Y}_{t-1})$ son conocidas para cada t, t = 1, 2, ...

En el modelo convencional no se asume que Z_t, T_t, H_t y Q_t dependen de la información hasta el tiempo t-1, $\mathbf{Y}_{t-1}$, y las hipótesis básicas son:

(i) $\varepsilon_t \sim N_s(0, H_t)$ y $\eta_t \sim N_s(0, Q_t)$ para cada t = 1, 2, ...

(ii) Las sucesiones $\{\varepsilon_t\}$ y $\{\eta_t\}$ son, cada una, independientes mutuamente estocásticamente y cualquier conjunto finito de la forma $\{\varepsilon_t, \eta_{s_1}, \eta_{s_k}\}$ es mutuamente independiente.

(iii) $\alpha_0 \sim N_s(a_0, P_0)$, y α_0, $\{\varepsilon_t\}$ y $\{\eta_t\}$ son independientes mutuamente estocásticamente

Es claro que las diferencias esenciales están en las hipótesis (ii) y (iii) de ambos modelos. Más aún, al condicionar por $\mathbf{Y}_{t-1}$, las hipótesis del modelo convencional son reducidas notablemente, ya que la independencia mútua exigida en la hipótesis (ii) y la independencia mútua entre α_0, $\{\varepsilon_t\}$ y $\{\eta_t\}$ de la hipótesis (iii) desaparecen.

Ahora bien, un hecho importante en el modelo convencional es que las hipótesis (ii) y (iii) implican que α_t y ε_t, y, α_{t-1} y η_t, son, cada par, independientes estocásticamente y esta proposición es análoga a la hipótesis (ii) del modelo condicional . En virtud de esto se podría decir que la hipótesis (ii) del modelo condicional *absorbe* las hipótesis (ii) y (iii) del convencional, salvo por el supuesto $\alpha_0 \sim N_s(a_0, P_0)$. Este hecho permite la deducción del FK en el caso condicional como se mostrará a continuación.

3. Deducción del Filtro de Kalman en el caso condicional.

En lo que sigue se hará referencia al modelo (1) y la deducción se hace sin seguir el esquema formal *Teorema - Demostración.*

El vector de estado en t = 1 está dado por

$$\alpha_1 = T_1(\mathbf{Y}_0)\,\alpha_0 + \eta_1.$$

Al condicionar sobre $\mathbf{Y}_0$ (que representa solo el *conocimiento* de la distribución de α_0) se obtiene que (ver **Nieto** (1993))

$\alpha_1 \mid \mathbf{Y}_0 \sim N_s(T_1 a_0, T_1 P_0 T'_1 + Q_1)$ $(T_1 \equiv T_1(\mathbf{Y}_0), Q_1 \equiv Q_1(\mathbf{Y}_0))$

Sean $a_{1|0} = T_1 a_0$

$$P_{1|0} = T_1 P_0 T'_1 + Q_1.$$

Ahora, sean $V_1 = (\alpha'_1\ Y'_1)'$ y $X = (X'_1\ X'_2)'$ tales que X_1 y X_2 son compatibles con α_1 y Y_1, respectivamente, para realizar producto interno. Entonces

$$X'V_1 = (X'_1 + X'_2 Z_1)\,\alpha_1 + X'_2\,\varepsilon_1 \qquad (Z_1 \equiv Z_1(\mathbf{Y}_0)))$$

Nuevamente condicionando por $\mathbf{Y}_0$ se obtiene que V_1 tiene distribución multinormal singular (**Nieto** (1993)) con vector de medias

$$\begin{pmatrix} a'_{1|0} & a'_{1|0} Z'_1 \end{pmatrix}'$$

y matriz de covarianzas

$$\begin{pmatrix} P_{1|0} & P_{1|0} Z'_1 \\ Z_1 P_{1|0} & Z_1 P_{1|0} Z'_1 + H_1 \end{pmatrix}.$$

con $(H_1 \equiv H_1(\mathbf{Y}_0)$.

Aplicando los resultados de **Marsaglia** (1992) se obtiene que la distribución de α_1 condicional sobre un valor particular de Y_1 (dado $\mathbf{Y}_0$) es multinormal singular con media

$$a_1 = a_{1|0} + P_{1|0} Z'_1 F_1^{-} (Y_1 - Z_1 a_{1|0})$$

y matriz de covarianza

$$P_1 = P_{1|0} - P_{1|0} Z_1' F_1^- Z_1 P_{1|0}$$

donde

$$F_1 = Z_1 P_{1|0} Z_1' + H_1 (\mathbf{Y}_0)$$

y F_1^- indica la seudo - inversa de F_1.

Si se supone que se conoce la distribución de α_{t-1} dado $\mathbf{Y}_{t-1}$, es decir,

$$\alpha_{t-1} \sim N_s(a_{t-1}, P_{t-1}),$$

se utilizan las hipótesis del modelo, los resultados de **Nieto** (1993) sobre distribución multinormal singular y se procede como en el caso anterior t = 1, entonces se obtiene que

$$\alpha_t \mid Y_t, \mathbf{Y}_{t-1} \sim N_s(a_t, P_t),$$

donde

$$a_t = a_{t|t-1} + P_{t|t-1} Z_t' (\mathbf{Y}_{t-1})\ F_t^- \left(Y_t - Z_t (\mathbf{Y}_{t-1})\ a_{t|t-1}\right) \quad (2)$$

$$P_t = P_{t|t-1} - P_{t|t-1} Z_t' (\mathbf{Y}_{t-1})\ F_t^- Z_t (\mathbf{Y}_{t-1})\ P_{t|t-1} \quad (3)$$

con

$$a_{t|t-1} = T_t (\mathbf{Y}_{t-1})\ a_{t-1}$$

$$P_{t|t-1} = T_t (\mathbf{Y}_{t-1})\ P_{t-1}\ T_t (\mathbf{Y}_{t-1})' + Q_t(\mathbf{Y}_{t-1}) \quad \text{y}$$

$$F_t = Z_t(\mathbf{Y}_{t-1})\ P_{t|t-1}\ Z_t (\mathbf{Y}_{t-1})' + H_t(\mathbf{Y}_{t-1})$$

Las ecuaciones (2) y (3) conforman el Filtro de Kalman para el modelo condicional y son análogas al caso convencional (**Nieto** (1992)).

En realidad el modelo de estados condicional presentado acá podría extenderse al caso en el cual el conjunto informativo sobre el que se condiciona, además de $\mathbf{Y}_{t-1}$, contiene información *externa* al proceso $\{Y_t\}$, inclusive hasta el tiempo t. Si I_t denota este nuevo conjunto y las hipótesis (i) - (ii) se reescriben condicionando sobre I_t, entonces las ecuaciones (2) - (3) siguen válidas. Esto extiende los resultados de **West** y **Harrison** (1989) sobre el mismo tópico.

Referencias

Anderson, T. W. (1984), *An Introduction to Multivariante Statistical Analysis*, John Wiley and Sons, Inc, New York.

Anderson, D. O. and Moore, J. B. (1979), *Optimal Filtering*, Prentice -Hall, Inc., New Jersey (U. S. A.).

Aoki, M. (1990), *State Space Modeling of Time Series*, Springer Verlag, Berlín.

Catlin, D. (1989), *Estimation, Control and the Discrete Kalman Filter*, Springer-Verlag, Berlín.

Harvey, A. C. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge.

Kohn, R. and Ansley, C. F. (1983), Fixed interval estimation in state sapace models when some of the data are missing or aggregated, *Biometrika*, Vol. 70, 3, pp 683 - 8.

Marsaglia, G. (1964), Conditional Means and Covariances of Normal Variables with Singular Covariance Matrix, *Journal of the American Statistical Association*, 59, pp. 1203 - 1204.

Nieto, F. H., (1993), Deducción del Filtro de Kalman en el caso de modelos de estados gaussianos singulares, *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales*, Vol. XVIII, 71, pp 539-543.

West, M and Harrison, P. J. (1989), *Bayesian Forecasting and Dynamic Models*, Springer-Verlag, Berlín.