

MONITORIZACIÓN E INTERVENCIÓN CONJUNTA EN UN MODELO LINEAL DINÁMICO COMO UN PROBLEMA DE DECISIÓN

Pilar Gargallo Valero¹, Manuel Salvador Figueras².

¹Dpto. de Métodos Estadísticos
EUEEZ.Universidad de Zaragoza
pigarga@posta.unizar.es

²Dpto. de Métodos Estadísticos
Facultad de Económicas. Universidad de Zaragoza
salvador@posta.unizar.es

RESUMEN

En este trabajo se presenta un procedimiento que monitoriza de forma conjunta outliers, cambios en media, varianza y autocorrelaciones e indica cómo intervenir en el modelo de una manera selectiva, es decir, cómo actuar en función del tipo de cambio. Se parte de la función de pérdida implícita en el algoritmo propuesto por West y Harrison (1997) para construir un algoritmo de monitorización conjunta de los errores de predicción a un paso, basado en un esquema decisionista bayesiano, que permite identificar el tipo de deterioro observado e intervenir de acuerdo al mismo. Finalmente, la metodología se ilustra con un conjunto de series reales procedentes de campos de estudio muy dispares tales como Medio-Ambiente, Comercio, Sociología o Finanzas.

Palabras y frases clave: Monitorización, MLD, Función de Pérdida.

Area temática: 7. Métodos Cuantitativos

1. INTRODUCCIÓN

Una serie temporal contiene, a menudo, observaciones atípicas o discordantes debido a múltiples factores como, por ejemplo, errores de transcripción en la recopilación de datos, ocurrencia de sucesos atípicos como huelgas, desastres naturales, guerras, etc. y puntos de cambio estructural debido a cambios de régimen político, irrupción de nuevos competidores en el mercado, etc. Este tipo de observaciones pueden causar muchos problemas en la construcción de modelos dinámicos que describan la evolución de las series analizadas.

En este trabajo se presenta un procedimiento que monitoriza de forma conjunta outliers, cambios en media, varianza y autocorrelaciones residuales e indica cómo intervenir en el modelo de una manera selectiva, es decir, como actuar en función del tipo de cambio. Se parte de la función de pérdida implícita en el algoritmo propuesto por West y Harrison (1997) para construir un algoritmo de monitorización conjunta de

los errores de predicción a un paso, basado en un esquema decisionista bayesiano, que permite identificar el tipo de deterioro observado e intervenir de acuerdo al mismo.

Finalmente, el algoritmo propuesto se aplica a diversas series temporales analizadas previamente en la literatura y procedentes de campos de estudio tan dispares como Medio-Ambiente, Comercio, Sociología o Finanzas. Estas series son ejemplos reales que contienen algunos outliers, diversas clases de tendencias o patrones estacionales, así como, la presencia de heterocedasticidades y, por tanto, se utilizan para comprobar el buen o mal funcionamiento del algoritmo.

El trabajo se estructura como sigue: en la sección 2 se plantea el problema y el modelo a analizar, en la sección 3 se expresa el procedimiento de West y Harrison (1997) como un problema de decisión indicando cuál es la función de pérdida implícita. En la sección 4 se describe el proceso de monitorización conjunta de varios tipos de cambio como un problema de decisión que generaliza el procedimiento de West y Harrison detallando cada una de las componentes de dicho esquema. En la sección 5 se proponen las intervenciones que se suponen adecuadas para cada tipo de cambio y en la sección 6 se especifican los pasos a seguir en el algoritmo de monitorización conjunta e intervención automática propuesto. En la sección 7 se ilustra la metodología con un conjunto de series reales y finalmente, en la sección 8 se concluye.

2. EL MODELO LINEAL DINÁMICO

Los modelos lineales dinámicos son muy utilizados en la modelización de series temporales univariantes. Ello se debe, fundamentalmente, a tres razones: la primera es que se puede deducir una gran cantidad de modelos realistas a partir de ellos; la segunda es que muchos modelos no lineales se pueden aproximar por un modelo de este tipo y, por último, la tercera, de carácter más matemático, es la existencia de un algoritmo, el filtro de Kalman, que permite su estimación de una manera rápida y sencilla. Un *Modelo Lineal Dinámico (MLD) univariante* general viene dado por las ecuaciones:

$$\begin{aligned} Y_t &= F_t' \theta_t + v_t & \text{con } v_t &\sim N(0, V_t) & \text{siendo } \phi_t &= V_t^{-1} \\ \theta_t &= G_t \theta_{t-1} + w_t & \text{con } w_t &\sim T_{n_{t-1}}(0, W_t) \end{aligned}$$

con distribuciones a priori:

$$(\theta_0 / D_0) \sim T_{n_0} (m_0, C_0) \quad y \quad (\phi_0 / D_0) \sim G\left(\frac{n_0}{2}, \frac{n_0 S_0}{2}\right)$$

con $\{v_t, w_t; t=1, \dots, T\}$ independientes entre sí e independientes de (θ_0 / D_0) y (ϕ_0 / D_0) . Suponemos, además, que la matriz de varianzas y covarianzas de evolución se construye mediante factores de descuento (ver West y Harrison (1997), por ejemplo). Si tenemos un MLD construido como superposición de varios MLD tomaremos un factor de descuento diferente para cada una de sus h componentes, es decir, $\delta_{m1}, \dots, \delta_{mh}$. Suponemos, también, que la varianza observacional es desconocida y variable en el tiempo.

El procesamiento de la información se realiza de forma secuencial mediante una metodología bayesiana que utiliza el filtro de Kalman (1960) para obtener, de manera recursiva, los parámetros de las distribuciones *a posteriori* del vector de estados, θ_t , en un instante determinado con la información disponible hasta ese momento.

Si denotamos $D_t; t = 0, \dots, T$ al conjunto de información del analista en el periodo t donde D_0 es el conjunto de información inicial del analista previo a la observación de la serie y partimos de:

$$(\theta_{t-1} / D_{t-1}) \sim T_{n_{t-1}} (m_{t-1}, C_{t-1})$$

$$(\phi_{t-1} / D_{t-1}) \sim G\left(\frac{n_{t-1}}{2}, \frac{n_{t-1} S_{t-1}}{2}\right)$$

en este caso, las ecuaciones de actualización vienen dadas por:

$$(\theta_t / D_{t-1}) \sim T_{n_{t-1}} (a_t, R_t) \quad \text{donde } a_t = G_t m_{t-1} \text{ y } R_t = G_t C_{t-1} G_t' + W_t \text{ donde } W_t \text{ está}$$

construida utilizando los h factores de descuento $\delta_{m1}, \dots, \delta_{mh}$.

$$(\phi_t / D_{t-1}) \sim G\left(\frac{\delta_v n_{t-1}}{2}, \frac{\delta_v n_{t-1} S_{t-1}}{2}\right)$$

$$(Y_t / D_{t-1}) \sim T_{\delta_v n_{t-1}} (f_t, Q_t) \quad \text{donde } f_t = F_t' a_t \text{ y } Q_t = F_t' R_t F_t + S_{t-1}$$

$$(\theta_t / D_t) \sim T_{n_t} (m_t, C_t) \quad \text{donde } m_t = a_t + A_t u_t \text{ y } C_t = \frac{S_t}{S_{t-1}} (R_t - A_t A_t' Q_t)$$

$$\text{siendo } u_t = Y_t - f_t \text{ y } A_t = R_t F_t Q_t^{-1}$$

$$(\phi_t/D_t) \sim G\left(\frac{n_t}{2}, \frac{n_t S_t}{2}\right) \quad \text{con } n_t = \delta_v n_{t-1} + 1 \text{ y } S_t = \delta_v S_{t-1} + \frac{S_{t-1}}{n_t} \left(\frac{u_t^2}{Q_t} - 1\right)$$

3. PROCEDIMIENTO DE WEST Y HARRISON COMO UN PROBLEMA DE DECISIÓN

El algoritmo de monitorización propuesto por West y Harrison (1997) utiliza como material de trabajo los errores predictivos a un paso estandarizados, $\{e_t; t=1, \dots, T\}$ y como herramienta matemática básica el Factor Bayes Acumulado dado por la expresión:

$$H_1(r) = \frac{P_1(e_1, e_2, \dots, e_r / D_0)}{P_0(e_1, e_2, \dots, e_r / D_0)} = \frac{\prod_{t=1}^r P_1(e_t / D_{t-1})}{\prod_{t=1}^r P_0(e_t / D_{t-1})}$$

donde M_0 es el modelo que se somete a evaluación continua y M_1 es un modelo alternativo que se construye para examinar el comportamiento de M_0 por comparación. La forma que adopta M_1 depende de los aspectos de M_0 que el analista considere necesario vigilar.

Con esta notación, el esquema de monitorización propuesto por West y Harrison queda, en la etapa r -ésima, como sigue:

$$\begin{cases} \text{Si } H_1(r) \geq \frac{1}{\tau} & \Rightarrow \text{ se acepta } M_1 \\ \text{Si } H_1(r) \leq 1 & \Rightarrow \text{ se acepta } M_0 \\ \text{Si } 1 < H_1(r) < \frac{1}{\tau} & \Rightarrow \text{ se continua} \end{cases}$$

A continuación vamos a plantear este procedimiento como un problema de decisión en el que los modelos M_0 y M_1 representan los dos *estados de la naturaleza* y las tres decisiones o *alternativas* que pueden tomarse son:

d_0 : El monitor indica que el proceso funciona correctamente y se acepta el modelo estándar M_0 .

- d_c : La información de que se dispone no es concluyente para decidir entre M_0 y M_1 y, por tanto, se toma una nueva observación.
- d_1 : El monitor indica que el proceso no funciona bien y se acepta el modelo alternativo M_1 .

Si denotamos $P(M_0 / D_0)$ y $P(M_1 / D_0)$ a las probabilidades a priori de los estados M_0 y M_1 , respectivamente, se escoge como solución óptima aquella cuyo riesgo o pérdida esperada a posteriori sea menor. El siguiente teorema proporciona la forma de la función de pérdida implícita en el proceso de monitorización de West y Harrison cuando suponemos, además, que $P(M_0 / D_0) = P(M_1 / D_0)$.

Teorema 1.- En las condiciones que se acaban de describir la *función de pérdida* implícita en el procedimiento de West y Harrison, $\ell(M_i, d_j)$, es proporcional a:

$\ell(M_i, d_j)$	M_0	M_1
d_0	0	1
d_1	$1/\alpha$	0
d_c	$1/2$	$1/2$

siendo

$$\alpha = \frac{2\tau}{1 + \tau}$$

Observación 1- Este teorema proporciona un medio para conectar la forma de la función de pérdida con el umbral τ y sirve como punto de partida para el algoritmo de monitorización conjunta que proponemos en este capítulo.

4. EL PROCESO DE MONITORIZACIÓN CONJUNTA DE OUTLIERS Y CAMBIOS ESTRUCTURALES COMO UN PROBLEMA DE DECISIÓN

En esta sección presentamos un procedimiento que analiza conjuntamente la existencia de outliers, cambios de nivel, varianza y desarrollo de autocorrelaciones locales de carácter residual planteándolo como un problema de decisión cuya función de pérdida generaliza a la función implícita en West y Harrison.

Es obvio que con una única observación sólo podemos estudiar la presencia de outliers y que necesitamos como mínimo dos observaciones para poder monitorizar también cambios de nivel, varianza y presencia de autocorrelaciones residuales. Por esta razón, el problema planteado es un problema de decisión secuencial en el que el conjunto de *estados de la naturaleza*, el conjunto de *alternativas* o decisiones posibles y la *función de pérdida* dependen de si estamos analizando una sola observación o varias. A continuación se exponen con detalle cada una de las componentes fundamentales de este problema de decisión.

4.1. EL CONJUNTO Θ DE LOS ESTADOS DE LA NATURALEZA

Se comienza determinando el conjunto Θ de los *estados de la naturaleza* que está formado por el modelo estándar M_0 y todas las alternativas que modelizan las posibles desviaciones respecto de M_0 que se desean detectar. Dichas alternativas dependen, como se ha comentado con anterioridad, del número de observaciones tomadas, es decir:

$$\Theta = \begin{cases} M_0, M_{out+}, M_{out-} & \text{si } r = 1 \\ M_0, M_{out+}, M_{out-}, M_{\mu,h+}, M_{\mu,h-}, M_{\sigma,k+}, M_{\sigma,k-}, M_{\rho,g+}, M_{\rho,g-} & \text{si } r \geq 2 \end{cases}$$

donde:

M_0 : representa el modelo estándar y, por tanto, la distribución condicional de los errores de predicción a un paso estandarizados es:

$$(e_t / D_{t-1}) \sim N(0, 1) \quad t = 1, \dots, r$$

M_{out+} y M_{out-} : representan modelos en los que se supone la existencia de un outlier de signo positivo o negativo, respectivamente, en el instante r -ésimo, es decir:

$$(e_t / D_{t-1}) \sim \begin{cases} N(0, 1) & \text{si } t = 1, \dots, r-1 \\ N(A, 1) & \text{si } t = r \end{cases} \quad A > 0 \quad \text{si } M_{out+}$$

$$(e_t / D_{t-1}) \sim \begin{cases} N(0, 1) & \text{si } t = 1, \dots, r-1 \\ N(-A, 1) & \text{si } t = r \end{cases} \quad A > 0 \quad \text{si } M_{\text{out-}}$$

$M_{\mu, h+}$ y $M_{\mu, h-}$: representan modelos con cambios de nivel positivo y negativo, respectivamente. En este caso, la distribución condicional de los errores de predicción a un paso estandarizados tiene media no nula, es decir:

$$\begin{aligned} (e_t / D_{t-1}) &\sim N(h, 1) & t = 1, \dots, r & \quad h > 0 & \text{si } M_{\mu, h+} \\ (e_t / D_{t-1}) &\sim N(-h, 1) & t = 1, \dots, r & \quad h > 0 & \text{si } M_{\mu, h-} \end{aligned}$$

$M_{\sigma, k+}$ y $M_{\sigma, k-}$: representan modelos con un aumento y disminución en varianza, respectivamente. En este caso, la distribución condicional de los errores de predicción a un paso estandarizados tiene varianza distinta de uno, es decir:

$$\begin{aligned} (e_t / D_{t-1}) &\sim N(0, k) & t = 0, 1, \dots, r & \quad k > 1 & \text{si } M_{\sigma, k+} \\ (e_t / D_{t-1}) &\sim N(0, 1/k) & t = 0, 1, \dots, r & \quad k > 1 & \text{si } M_{\sigma, k-} \end{aligned}$$

$M_{\rho, g+}$ y $M_{\rho, g-}$: representan modelos con errores de predicción a un paso estandarizados dependientes y con correlación positiva y negativa, respectivamente. Es decir:

$$\begin{aligned} e_t &= g e_{t-1} + a_t & \text{con } a_t \sim N(0, 1) \text{ independientes } t = 1, \dots, r & \text{si } M_{\rho, g+} \\ e_t &= -g e_{t-1} + a_t & \text{con } a_t \sim N(0, 1) \text{ independientes } t = 1, \dots, r & \text{si } M_{\rho, g-} \end{aligned}$$

donde para ambos casos $0 < g < 1$ y $e_0 = 0$.

Para resolver un problema de decisión en ambiente de riesgo se necesita conocer la ley de probabilidad correspondiente a cada uno de los estados de la naturaleza $M_i \in \Theta$. Por tanto, si se supone que las probabilidades a priori de todos los estados, $P(M_i/D_0)$, son idénticas, en cada etapa r del proceso la probabilidad a posteriori de $M_i \in \Theta$, $P(M_i/D_r)$, viene dada por:

$$P(M_i / D_r) = \frac{\prod_{t=1}^r P_i(e_t / D_{t-1})}{\sum_{M_k \in \Theta} \prod_{t=1}^r P_k(e_t / D_{t-1})}$$

4.2. EL CONJUNTO DE ALTERNATIVAS

La notación utilizada para las distintas *alternativas* o decisiones es:

d_0 : El monitor indica que el proceso funciona correctamente.

d_c : La información de que se dispone no es concluyente para decidir si hay o no outlier y, por lo tanto, se toma una nueva observación.

$d_{c,i} (i \neq 0, out+, out-)$: Se sospecha un deterioro en la forma descrita por $M_i \in \Theta$ pero la información de que se dispone no es concluyente para decidir y, por tanto, se toma una nueva observación.

$d_i (i \neq 0)$: El monitor indica que el proceso no funciona bien debido a que se ha producido un deterioro en los errores de predicción, descrito por el modelo $M_i \in \Theta$.

Con esta notación el conjunto de alternativas o posibles decisiones viene dado por:

$$D = \begin{cases} d_c, d_{out+}, d_{out-} & \text{si } r = 1 \\ d_0, d_{out+}, d_{out-}, d_{\mu,h+}, d_{\mu,h-}, d_{\sigma,k+}, d_{\sigma,k-}, d_{\rho,g+}, d_{\rho,g-}, d_{c,\mu,h+}, d_{c,\mu,h-}, d_{c,\sigma,k+}, d_{c,\sigma,k-}, d_{c,\rho,g+}, d_{c,\rho,g-} & \text{si } r \geq 2 \end{cases}$$

La razón de tomar así el conjunto D estriba en el interés por detectar outliers unipuntuales aislados. Notar que para el resto de los casos tomamos al menos dos observaciones para poder distinguir el cambio producido siendo, además, este requisito imprescindible en la detección de autocorrelaciones locales.

4.3. LA FUNCIÓN DE PÉRDIDA

La función de pérdida adoptada depende también del número de observaciones tomadas, r , puesto que los estados de la naturaleza y las alternativas varían. Así, en el caso $r = 1$ asignamos una pérdida proporcional a $1/\alpha$ con $0 < \alpha < 1$ al hecho de decidir

que existe un outlier cuando no es así con el fin de evitar intervenciones innecesarias y garantizar que sólo intervenimos cuando hay evidencia suficiente. Además, penalizamos no detectar atípicos o detectarlos con signo cambiado con una pérdida proporcional a 1. Es decir, la función de pérdida $\ell(\mathbf{M}_i, \mathbf{d}_j)$ viene dada por una matriz proporcional a:

$\ell(\mathbf{M}_i, \mathbf{d}_j)$	M_0	$M_{\text{out}+}$	$M_{\text{out}-}$
d_c	0	1	1
$d_{\text{out}+}$	$1/\alpha$	0	1
$d_{\text{out}-}$	$1/\alpha$	1	0

En el caso de que $r \geq 2$, la función de pérdida la construimos generalizando la función implícita en el procedimiento de West y Harrison, pero teniendo en cuenta además la presencia de varios modelos alternativos. Por lo tanto, dicha función viene dada por una matriz proporcional a:

$\ell(\mathbf{M}_i, \mathbf{d}_j)$	M_0	$M_{\text{out}+}$	$M_{\text{out}-}$	$M_{\mu, h+}$	$M_{\mu, h-}$	$M_{\sigma, k+}$	$M_{\sigma, k-}$	$M_{\rho, g+}$	$M_{\rho, g-}$
d_0	0	1	1	1	1	1	1	1	1
$d_{\text{out}+}$	$1/\alpha$	0	1	1	1	1	1	1	1
$d_{\text{out}-}$	$1/\alpha$	1	0	1	1	1	1	1	1
$d_{\mu, h+}$	$1/\alpha$	1	1	0	1	1	1	1	1
$d_{\mu, h-}$	$1/\alpha$	1	1	1	0	1	1	1	1
$d_{\sigma, k+}$	$1/\alpha$	1	1	1	1	0	1	1	1
$d_{\sigma, k-}$	$1/\alpha$	1	1	1	1	1	0	1	1
$d_{\rho, g+}$	$1/\alpha$	1	1	1	1	1	1	0	1
$d_{\rho, g-}$	$1/\alpha$	1	1	1	1	1	1	1	0
$d_{c, \mu, h+}$	$1/2$	1	1	$1/2$	1	1	1	1	1
$d_{c, \mu, h-}$	$1/2$	1	1	1	$1/2$	1	1	1	1
$d_{c, \sigma, k+}$	$1/2$	1	1	1	1	$1/2$	1	1	1
$d_{c, \sigma, k-}$	$1/2$	1	1	1	1	1	$1/2$	1	1
$d_{c, \rho, g+}$	$1/2$	1	1	1	1	1	1	$1/2$	1
$d_{c, \rho, g-}$	$1/2$	1	1	1	1	1	1	1	$1/2$

donde:

$$\alpha = \frac{2\tau}{1+\tau} \text{ con } 0 < \tau < 1$$

siendo τ el umbral del algoritmo de monitorización como ocurría en West y Harrison.

Una vez fijado el problema de decisión procedemos, a continuación, a describir la forma del procedimiento de monitorización conjunta. Los teoremas 2 y 3 deducen la regla de decisión cuando el número de observaciones tomadas es $r = 1$ y $r \geq 2$, respectivamente. La notación que vamos a utilizar es la siguiente:

$$H_j(r) = \frac{\prod_{t=1}^r P_j(e_t / D_{t-1})}{\prod_{t=1}^r P_0(e_t / D_{t-1})}$$

Teorema 2.- En las condiciones anteriores con $r = 1$ la regla de decisión óptima consiste en:

$$\begin{cases} \text{Si } H_{\text{out-}}(1) \geq \frac{1}{\alpha} & \Rightarrow \text{decisión } d_{\text{out-}} \\ \text{Si } \max \{H_{\text{out-}}(1), H_{\text{out+}}(1)\} < \frac{1}{\alpha} & \Rightarrow \text{decisión } d_c \\ \text{Si } H_{\text{out+}}(1) \geq \frac{1}{\alpha} & \Rightarrow \text{decisión } d_{\text{out+}} \end{cases}$$

El siguiente corolario describe la regla de decisión óptima dada en el teorema 2, en términos del error de predicción a un paso e_1 , simplificando así el tratamiento de los outliers.

Corolario 1.- La regla de decisión óptima anterior es equivalente a:

$$\begin{cases} \text{Si } e_1 \leq -0.5A + \frac{\ln \alpha}{A} & \Rightarrow \text{decisión } d_{\text{out-}} \\ \text{Si } -0.5A + \frac{\ln \alpha}{A} < e_1 < 0.5A - \frac{\ln \alpha}{A} & \Rightarrow \text{decisión } d_c \\ \text{Si } e_1 \geq 0.5A - \frac{\ln \alpha}{A} & \Rightarrow \text{decisión } d_{\text{out+}} \end{cases}$$

Observación 2. Para especificar el valor de A se impone que, dado un error máximo, e_{\max} , y un valor fijado α , se verifique la siguiente relación:

$$\frac{\exp \left\{ -0.5(e_{\max} - A)^2 \right\}}{\exp \left\{ -0.5e_{\max}^2 \right\}} = \frac{1}{\alpha}$$

El siguiente teorema describe el algoritmo cuando el número de observaciones tomadas es $r \geq 2$.

Teorema 3.- En las condiciones anteriores con $r \geq 2$ la regla de decisión óptima consiste en:

Se decide d_0 si:

$$\max_{\ell \in L} \{H_{\ell}(r)\} \leq 1 \quad \text{y} \quad \max_{k \in K} \{H_k(r)\} \leq \frac{1}{\alpha}$$

Se decide d_k ($k \in K$) si:

$$H_k(r) \geq \max \left\{ \frac{1}{\alpha}, \{H_i(r)\}_{\forall i \neq k}, \left\{ \frac{1}{2\tau} + \frac{1}{2} H_{\ell}(r) \right\}_{\forall \ell \in L} \right\}$$

Se decide d_{ℓ} ($\ell \in L$) si:

$$H_{\ell}(r) \geq \max \left\{ \frac{1}{\tau}, \{H_i(r)\}_{\forall i \neq \ell}, \left\{ \frac{1}{2\tau} + \frac{1}{2} H_{\ell'}(r) \right\}_{\forall \ell' \neq \ell \in L} \right\}$$

Se decide $d_{c,\ell}$ ($\ell \in L$) si:

$$\max \left\{ 1, \left\{ 2H_i(r) - \frac{1}{\tau} \right\}_{\forall i \neq \ell}, \{H_{\ell'}(r)\}_{\forall \ell' \neq \ell \in L} \right\} < H_{\ell}(r) < \frac{1}{\tau}$$

donde $K = \{\text{out}^+; \text{out}^-\}$ y $L = \{\mu, h^+; \mu, h^-; \sigma, k^+; \sigma, k^-; \rho, g^+; \rho, g^-\}$.

5. INTERVENCIONES ADECUADAS AL TIPO DE CAMBIO

Uno de los atractivos del algoritmo de monitorización propuesto es que ofrece la posibilidad de intervenir retrospectivamente de acuerdo con el tipo de cambio detectado utilizando la información proporcionada por las observaciones más recientes de la serie analizada. En esta sección proponemos las intervenciones que consideramos adecuadas para los distintos tipos de cambios estudiados.

5.1. OUTLIERS

Si se detecta un outlier la intervención que proponemos consiste en *ignorar la observación* Y_t para predicciones futuras, o equivalentemente, hacer $D_t = D_{t-1}$. Por tanto, la distribución a posteriori del vector de estados en el instante t coincide con la distribución a priori de dicho vector, es decir, $(\theta_t / D_t) \sim T_{n_t}(m_t, C_t)$ donde $m_t = a_t$ y $C_t = R_t$; y la distribución a posteriori en el instante t de la precisión coincide con la del instante $t-1$, es decir, $S_t = S_{t-1}$ y $n_t = n_{t-1}$. Además de ignorar la observación Y_t , y puesto que no estamos seguros del tipo de cambio que realmente se está produciendo, se disminuyen todos los factores de descuento, es decir, $\delta_{m1}, \dots, \delta_{mh}$, y el factor de descuento de la varianza δ_v consiguiendo así que el modelo se adapte mejor a los nuevos datos.

Para cambios en media y varianza se utiliza el modo de *intervención subjetiva* que consiste simplemente en cambiar los momentos de la distribución a priori de θ_t , a_t y R_t , por los valores a_t^* y R_t^* proporcionados por el analista. En nuestro caso concreto, si se detecta un cambio en el instante t , proponemos intervenir en el instante $t-r+1$, ya que el Factor Bayes Acumulado $H_j(r)$ indica que ha sido r instantes atrás donde se ha iniciado el deterioro. A partir de ese instante los datos se vuelven a actualizar con la intervención ya incorporada, utilizando el filtro de Kalman de la forma usual.

5.2. CAMBIOS DE NIVEL

Si se detecta un ***cambio en nivel*** obtenemos una estimación de dicho cambio calculando la media de los últimos r errores estandarizados a un paso que han intervenido en el factor Bayes acumulado multiplicada por la desviación típica de la distribución predictiva a un paso con el fin de deshacer la estandarización. Por tanto, si denotamos:

$$\hat{h} = \frac{\sum_{j=t-r+1}^t e_j}{r}$$

con e_t los errores de predicción a un paso estandarizados, y utilizamos la notación del epígrafe anterior, la modificación que proponemos es:

$$a_{t-r+1}^* = a_{t-r+1} + \hat{h} \sqrt{Q_{t-r+1}} E_1$$

siendo E_1 el vector coordenado de dimensión n que contiene todas sus componentes nulas salvo la primera que vale uno, es decir, modificamos solamente la primera componente de la media de la distribución a priori del vector de estados correspondiente al nivel de la serie.

Además de la modificación anterior disminuimos todos los factores de descuento $\delta_{m1}, \dots, \delta_{mh}$ con el fin de hacer el modelo más adaptativo a los nuevos datos.

5.3. CAMBIOS EN VARIANZA

Si se detecta un ***cambio en varianza*** la modificación que proponemos está basada en el siguiente razonamiento. Denotando $Q_{t-r+1}^* = \frac{Q_{t-r+1}}{S_{t-r}}$ (ver capítulo 4 de West y Harrison (1997)) se verifica que:

$$\frac{y_{t-r+1} - f_{t-r+1}}{(Q_{t-r+1}^*)^{1/2}} \sim N(0, V_{t-r+1}) \quad \Leftrightarrow \quad \frac{y_{t-r+1} - f_{t-r+1}}{(Q_{t-r+1}^*)^{1/2} V_{t-r+1}^{1/2}} \sim N(0, 1)$$

Además, dado que:

$$\phi_{t-r} / D_{t-r} \sim G \left\{ \frac{n_{t-r}}{2}, \frac{n_{t-r} S_{t-r}}{2} \right\} \text{ y } \phi_{t-r+1} / D_{t-r+1} \sim G \left\{ \frac{n_{t-r+1}}{2}, \frac{n_{t-r+1} S_{t-r+1}}{2} \right\}$$

las matrices S_{t-r} y S_{t-r+1} son estimaciones de la matriz de varianzas de observación en los instantes $t-r$ y $t-r+1$, respectivamente.

De manera que si suponemos que en el instante $t-r+1$ se ha producido un cambio en varianza, se tiene que:

$$\frac{y_{t-r+1} - f_{t-r+1}}{(Q_{t-r+1}^*)^{1/2}} \sim N(0, k V_{t-r+1}) \quad \Leftrightarrow \quad \frac{y_{t-r+1} - f_{t-r+1}}{(Q_{t-r+1}^*)^{1/2} V_{t-r+1}^{1/2} \sqrt{k}} \sim N(0, 1)$$

Por tanto, si utilizamos como estimación de V_{t-r} a S_{t-r} , es decir, un valor anterior al cambio, la estimación de V_{t-r+1} en el instante $t-r+1$ vendrá dada por $S_{t-r+1} \approx S_{t-r} k$.

Así, si estimamos el cambio en varianza, k , con la media de los cuadrados de los últimos r errores estandarizados a un paso que han intervenido en el cálculo del factor Bayes acumulado, es decir:

$$\hat{k} = \frac{\sum_{j=t-r+1}^t e_j^2}{r}$$

y utilizamos además la notación del epígrafe anterior, la modificación que proponemos es:

$$n_{t-r+1}^* = r$$

$$S_{t-r+1}^* = S_{t-r} \hat{k}$$

$$R_{t-r+1}^* = \frac{S_{t-r+1}^*}{S_{t-r}} R_{t-r+1}$$

Dado que los cambios en varianza pueden provenir de deterioros en la evolución de las distintas componentes del vector de parámetros θ_t , además de las modificaciones anteriores, disminuimos todos los factores de descuento $\delta_{m1}, \dots, \delta_{mh}$ y el factor de descuento de la varianza δ_v con el fin de hacer el modelo más adaptativo a los nuevos datos.

5.4. AUTOCORRELACIONES RESIDUALES

Finalmente, si se detecta *la presencia de autocorrelación residual*, proponemos disminuir los factores de descuento $\delta_{m1}, \dots, \delta_{mh}$ si la correlación es positiva y aumentarlos si es negativa. La disminución o aumento de dichos factores de descuento la realizamos paulatinamente desde el instante $t-r+1$ hasta el instante t volviendo a actualizar los datos a partir del instante $t-r+1$ con los nuevos factores de descuento, utilizando el filtro de Kalman de la forma usual. En el caso de correlaciones negativas,

si el aumento de los factores de descuento no consigue arreglar el problema, se procede a disminuir dichos factores como en el caso de correlaciones positivas.

6. ALGORITMO DE MONITORIZACIÓN E INTERVENCIÓN AUTOMÁTICA

En las secciones 4 y 5 hemos proporcionado la regla de decisión y las intervenciones que debemos aplicar para los diferentes deterioros analizados. Por tanto, combinando esta información, podemos construir nuestro algoritmo de monitorización e intervención automática cuyo esquema puede verse en el organigrama de la figura 1.

El proceso consiste en obtener, como se ha visto en los teoremas 2 y 3, la decisión óptima, es decir, la decisión de menor riesgo o pérdida esperada a posteriori, y con ella, la forma de actuar es muy simple. Si la decisión óptima es:

- d_0 : se actualiza el modelo normalmente y se reinicializa el monitor.
- d_{out+} ó d_{out-} : se interviene según se ha indicado en el epígrafe 5.1 y se reinicializa el monitor.
- $d_{\mu,h+}$ ó $d_{\mu,h-}$: se interviene según se ha indicado en el epígrafe 5.2 y se reinicializa el monitor.
- $d_{\sigma,k+}$ ó $d_{\sigma,k-}$: se interviene según se ha indicado en el epígrafe 5.3 y se reinicializa el monitor.
- $d_{p,g+}$ ó $d_{p,g-}$: se interviene según se ha indicado en el epígrafe 5.4 y se reinicializa el monitor.
- En cualquier otro caso se continua y se toma una nueva observación.

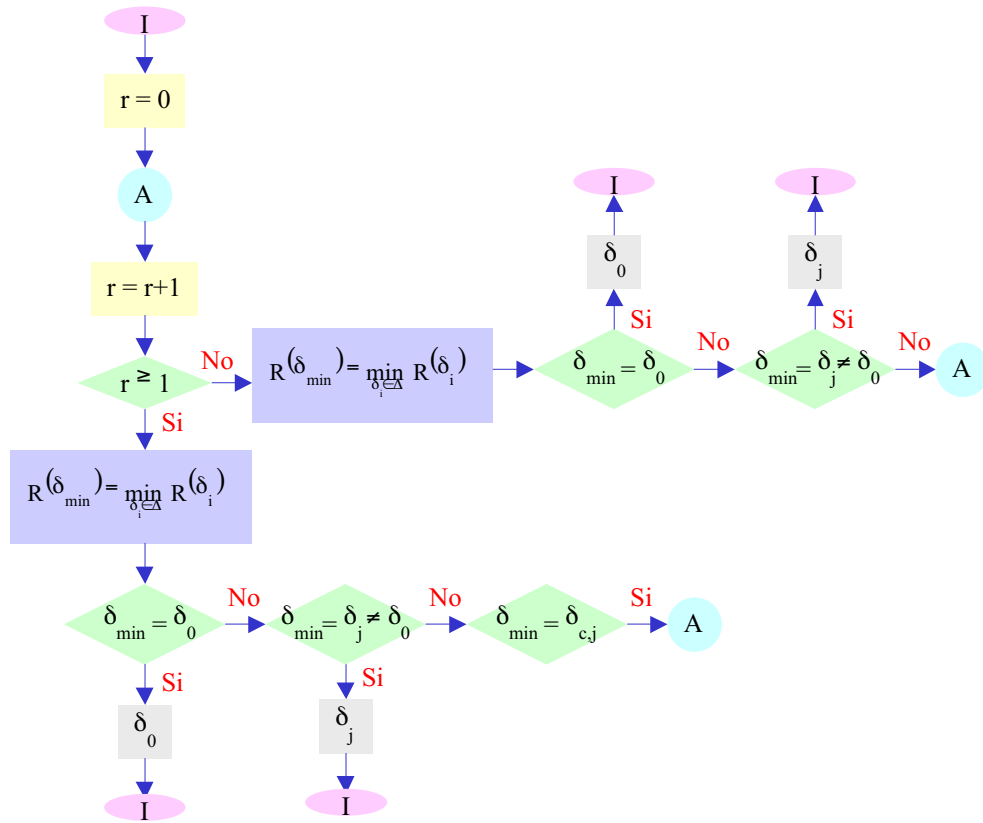


Figura 1. Organigrama del esquema de monitorización conjunta

7. APLICACIONES

En este apartado, el algoritmo propuesto en la sección anterior se aplica a diversas series temporales analizadas previamente en la literatura y procedentes de campos de estudio tan dispares como Medio-Ambiente, Comercio, Sociología o Finanzas. Por brevedad, simplemente se describen las series con las que se ha trabajado omitiéndose los gráficos y resultados obtenidos.

Así, en primer lugar, dicho algoritmo se ha aplicado sobre la serie temporal que recoge el volumen anual del río Nilo durante los años 1871-1970 tomada de Cobb (1978) detectando un atípico negativo en la observación 1877 y la presencia de correlación positiva en el año 1892. Esta serie ha sido examinada por gran cantidad de autores que habían detectado un cambio de nivel en 1899, sin embargo, con nuestro algoritmo detectamos la presencia de correlación positiva en un instante anterior lo cual

puede ser indicativo de que el cambio de nivel no ocurrió de una forma brusca sino que ya venía anunciado varios instantes atrás.

El siguiente ejemplo que se ha analizado es la serie CP6 presentada por West y Harrison (1997) y que recoge las ventas mensuales de tabaco de una importante compañía en el Reino Unido desde Enero de 1955 hasta Diciembre de 1959. El algoritmo detecta un aumento de nivel en Junio de 1955, dos aumentos en varianza en Julio de 1955 y en Diciembre de 1956, una disminución en varianza en Enero de 1957 y una disminución en nivel en Agosto de 1958.

Otra serie que se ha analizado es la serie mensual que recoge el número de conductores muertos y seriamente heridos en Gran Bretaña desde Enero de 1969 hasta Diciembre de 1984 tomada de Harvey y Durbin (1986). El algoritmo detecta un atípico negativo en Enero de 1974 y la presencia de correlación positiva en Febrero de 1974 y en Octubre de 1982. Estos cambios parecen estar de acuerdo con el aumento de los precios de la gasolina como respuesta de la crisis del petróleo y con la introducción de la ley de obligatoriedad del cinturón de seguridad en 1983.

La siguiente serie analizada ha sido tomada de West y Harrison (1997) y recoge (en miles) el número de nuevos matrimonios registrados trimestralmente en el Reino Unido entre los años 1965 a 1970 ambos inclusive. Esta serie presenta un aumento en estacionalidad en 1967 y un cambio de patrón estacional en 1969. Es interesante destacar que este algoritmo no está preparado para captar cambios en los patrones estacionales y, por tanto, los confunde con un aumento en varianza en el tercer trimestre de 1967, sin embargo, esta intervención proporciona unos resultados aceptables.

Otra serie sobre la que se aplica el algoritmo es el ejemplo clásico de las líneas aéreas que se puede encontrar en el libro de Box-Jenkins (1976) y que contiene desde Enero de 1949 hasta Diciembre de 1960 el número de pasajeros de líneas aéreas internacionales. Dicho algoritmo detecta únicamente una disminución en nivel en Agosto de 1957.

También presentamos un ejemplo de regresión dinámica analizado por Pole et al. (1994) y que estudia la participación en el mercado de un producto de consumo en

términos de un nivel subyacente, el precio del producto y los esfuerzos de promoción tanto por parte de la propia compañía como por parte de los competidores. Este conjunto de datos contiene cuatro series semanales durante los años 1990 y 1991. El algoritmo propuesto detecta dos disminuciones en varianza una en la semana 21 y otra en la 33 de 1990 y un aumento en varianza en la semana 32 de ese mismo año.

Finalmente, se aplica el algoritmo de monitorización sobre la serie heterocedástica del Ibex 35 que recoge desde Enero de 1990 hasta Diciembre de 2000, las cotizaciones diarias de las 35 empresas más líquidas del mercado, ponderadas por capitalización y como es lógico, el monitor, en este caso, detecta una gran cantidad de cambios en varianza.

8. CONCLUSIONES

En este trabajo se ha planteado el proceso de monitorización como un problema de decisión bayesiana construyendo un procedimiento general que analiza conjuntamente la existencia de outliers, cambios en media, varianza y permitiendo también detectar la presencia de autocorrelaciones locales en los errores de predicción. Además, se indica como intervenir en el modelo de una manera selectiva, es decir, cómo actuar para cada tipo de cambio, a diferencia del algoritmo de monitorización e intervención propuesto por West y Harrison (1997), en el que las intervenciones eran las mismas con independencia del monitor que daba la señal de deterioro.

9. REFERENCIAS

- Box, G. E. P. y Jenkins, G. M. (1976): *Time Series Analysis: Forecasting and Control*, (2ª Ed.). Holden-Day, San Francisco.
- Cobb, G. W. (1978): "The Problem of the Nile: Conditional Solution to a Change-Point Problem". *Biometrika*. 65, 243-251.
- Harvey, A. C. y Durbin, J. (1986): "The effects of seat belt legislation on British road casualties: a case study in structural time series modelling (with discussion)". *Journal of the Royal Statistical Society Series A*, 149, 187-227.
- Pole, A.; West, M. y Harrison, P. J. (1994): *Applied Bayesian Forecasting and Time Series Analysis*. Chapman & Hall, New York.
- West, M. y Harrison, P. J. (1997): *Bayesian Forecasting and Dynamic Models*. Springer, 2ª Ed.