Una nueva visión del estimador de Kalman

*Por: Darío Valencia Restrepo**

1. INTRODUCCION

El estimador introducido por Kalman, conocido como el filtro de Kalman, ha recibido considerable atención en la literatura técnica y particularmente en las aplicaciones relacionadas con el control de sistemas estocásticos. Es lógico que un resultado tan importante haya sido deducido por varios caminos, aunque éstos con frecuencia no son muy útiles desde el punto de vista didáctico en razón de su complejidad. Aquí se presentará una visión simple del filtro, que no exige comportamientos gaussianos, basada en un modelo lineal que preserva propiedades de primeros y segundos momentos, con lo cual se demuestra que el estimador proviene de un análisis de segundo orden.

Sin el ánimo de agotar la bibliografía, algunos caminos de deducción son los siguientes:

- Kumar y Varaiya (referencia 1, pág. 94) y Whittle (2, 147) suponen que los vectores aleatorios en juego se comportan conjuntamente según la distribución multinormal (comportamiento gaussiano).

- Davis y Vinter (3, 118) encuentran el mejor estimador lineal con ayuda de las proyecciones ortogonales en un espacio euclídeo de dimensión finita.

- Bras y Rodríguez-Iturbe (4, 437) siguen un enfoque bayesiano basado en una idea sugerida por Schweppe (5).

- Bryson y Ho (6, 349) utilizan un estimador mínimo cuadrático ponderado.

- Bertsekas (7, 158) utiliza el estimador mínimo - cuadrático.

- Gelb (8, 105) supone un estimador que es una combinación lineal de una predicción y una medición, para luego hallar los coeficientes de la combinación mediante la aplicación de un cierto criterio de optimalidad.

Antes de proceder a la deducción del filtro de Kalman, se mostrará su importancia en el contexto de la teoría de control y se probará algunos teoremas relacionados con un modelo lineal auxiliar, los cuales son la base para la mencionada deducción. Ello será materia de las dos secciones siguientes.

2. EL FILTRO DE KALMAN EN LA TEORIA DE CONTROL

Muchos sistemas dinámicos de la realidad pueden representarse por una ecuación diferencial ordinaria, de orden n, lineal y no homogénea, en donde el tiempo es la variable independiente y un cierto atributo del sistema es la variable dependiente. Mediante un tratamiento vectorial que aumente hasta "n" el número de atributos del sistema, es fácil llegar a una ecuación diferencial equivalente de primer orden, que en términos discretos correspondería a una expresión de la forma.

$$x_{k+1} = A_k x_k + B_k u_k \tag{1}$$

En donde x_k es el llamado vector de estado (conjunto de atributos del sistema) en el tiempo k; u_k es un vector de control (impulso o estímulo que se da al sistema) en k; y A_k y B_k son matrices de dimensiones apropiadas, que en general varían con el tiempo k. En lo que sigue, k = 0,1, ...

Es de anotar que a una ecuación del tipo (1), llamada ecuación de estado, es posible llegar partiendo de otra más general:

$$x_{k+1} = f_k (x_k, u_k) \tag{2}$$

por medio de una expansión en serie de Taylor (siempre que ella sea posible) en el entorno de un cierto pun-

* Profesor Titular de la Universidad Nacional de Colombia

to (x_k^*, u_k^*), y luego despreciando términos de orden mayor que uno. Se tendría,

$$\delta x_{k+1} = A_k\, \delta x_k + B_k\, \delta u_k \qquad (3)$$

siendo

$$\delta x_k = x_k - x_k^*;\ \delta u_k = u_k - u_k^*$$

δx_k y δu_k se conocen como las perturbaciones de estado y de control, en tanto que x_k^* y u_k^* son el estado y el control "nominales", todo en el tiempo k.

Es bien posible que un modelo dado por la ecuación (3) contenga errores de modelación (su carácter lineal puede ser una aproximación) o provenientes de la incertidumbre en la estimación de las matrices paramétricas A_k y B_k. Además, puede ser que el sistema esté sujeto a impulsos aleatorios, inherentes a él y distintos del control. Por estas razones, suele sumarse una componente aleatoria α_k, de dimensión apropiada, que con frecuencia se supone distribuida multinormalmente, con media cero y no correlacionada con x_k ni con u_k. Así

$$\delta x_{k+1} = A_k\, \delta x_k + B_k\, \delta u_k + \alpha_k \qquad (4)$$

De otro lado, es común que el vector de estado no sea directamente medible, pero que en el tiempo K se posea una medición y_k relacionada con x_k y afectada de un error aleatorio (el contenido de agua en varios embalses al comienzo de un cierto período se obtiene a partir de la lectura de los respectivos niveles del agua, por ejemplo, y dicha lectura está afectada de error). Se tendría una ecuación de medición lineal o eventualmente linealizable, dada en forma perturbacional por

$$\delta y_k = C_k\, \delta x_k + \beta_k \qquad (5)$$

en donde β_k suele suponerse distribuida multinormalmente, con media cero y no correlacionada con x_k.

Un problema importante es definir una política de control, dada por $u_0, u_1, \ldots$, que a partir de un estado dado x_0 obligue al sistema a evolucionar de una manera que sea de interés. La teoría de control demuestra que si las ecuaciones de estado y medición son del tipo (4) y (5), si los comportamientos son gaussianos y si la evolución que interesa cumple un criterio de optimalidad de tipo cuadrático, el control en k vendrá dado por

$$\delta u_k = -M_k\, \delta x_k \qquad (6)$$

Por lo tanto, si δx_k no se conoce directamente es necesario estimarlo de alguna manera, empleando toda la información disponible proporcionada por el modelo y las mediciones, incluyendo la última, y_k. El filtro de Kalman cumple esa finalidad, entregando un estimado de δx_k con agradables propiedades.

3. PROPIEDADES DEL MODELO LINEAL AUXILIAR

Sean x, y dos vectores aleatorios de dimensiones respectivas m y n, cuya distribución conjunta se desconoce pero con respecto a los cuales se conocen los primeros y segundos momentos (vectores medios y matrices de covarianza):

$$E[x] = \bar{x} \qquad E[y] = \bar{y}$$

$$E[(x-\bar{x})(x-\bar{x})^T] = S_{xx} \qquad E[(y-\bar{y})(y-\bar{y})^T] = S_{yy}$$

$$E[(x-\bar{x})(y-\bar{y})^T] = S_{xy}$$

También a veces se denota

$$S_{xx} = \mathrm{Cov}\,[x] \qquad S_{xy} = \mathrm{Cov}\,[x; y]$$

En la práctica, es frecuente conocer el valor de y, de manera que es importante estimar a x a partir de esa información y tener idea de la bondad de esa estimación. Se propone el siguiente modelo lineal como una aproximación al comportamiento condicional de x dado y, introducido por Valencia y Schaake (9):

$$x = ay + b + w \qquad (7)$$

en donde las matrices de parámetros a y b, así como las propiedades del vector aleatorio w (que trata de representar el comportamiento de x no descrito por la función lineal de y), deben determinarse. En un análisis de segundo orden, la determinación se hará con base en la preservación de primeros y segundos momentos.

Si E [w] = 0 y se toma valor esperado en (7):

$$\bar{x} = a\bar{y} + b$$

Entonces el modelo resulta

$$x - \bar{x} = a(y-\bar{y}) + w \qquad (8)$$

Si la ecuación (8) se posmultiplica por $(y-\bar{y})^T$ y se toma valor esperado en ambos lados

$S_{xy} = a\, S_{yy}$ siempre que w, y sean no correlacionados.

$$a = S_{xy}\, S_{yy} \text{ si } S_{yy} \text{ es no singular}$$

Finalmente, si ambos lados de (8) se posmultiplican por ellos mismos traspuestos

$$S_{xx} = a\, S_{yy} a^T + S_{ww} \qquad (9)$$

Reemplazando el valor de a antes obtenido, se obtiene la siguiente expresión para la matriz de covarianza de w:

$$S_{ww} = S_{xx} - S_{xy} S_{yy}^{-1} S_{yx} \qquad (10)$$

Todo lo anterior permite señalar que se encontraron expresiones para a y b, así como para los primeros y segundos momentos de w. El modelo resulta ser

$$x = \bar{x} + S_{xy} S_{yy}^{-1} (y-\bar{y}) + w \qquad (11)$$

Teorema uno. Si el comportamiento condicional de un vector aleatorio x dado otro y viene dado por un modelo como el de la ecuación (11), la media condicional de x dado y es

$$E\,[x/y] = \bar{x} + S_{xy} S_{yy}^{-1} (y-\bar{y}) \qquad (12)$$

Para lo cual basta tomar valor esperado condicional en (11).

En lo que sigue (12) se usará como estimado de x y se denotará por $\hat{x}$:

$$\hat{x} = E\,[x/y] = \bar{x} + S_{xy} S_{yy}^{-1} (y-\bar{y}) \qquad (13)$$

El error de dicho estimado es

$$\tilde{x} = x - \hat{x}$$

Una medida de la bondad del estimador viene dada por matriz de covarianza del error, la cual se hallará en el siguiente teorema.

Teorema dos. El error $\tilde{x}$ del estimado $\hat{x}$ tiene como matriz de covarianza.

$$S_{\tilde{x}\tilde{x}} = S_{xx} - S_{xy} S_{yy}^{-1} S_{yx} \qquad (14)$$

En efecto, con ayuda de las ecuaciones (11) y (13) se ve que

$$\tilde{x} = x - \hat{x} = w$$

De modo que las matrices $S_{\tilde{x}\tilde{x}}$ y S_{ww} son iguales, y la ecuación (10) prueba el teorema.

Teorema tres. El error $\tilde{x}$ del estimado $\hat{x}$ no está correlacionado con y ni tampoco con $\hat{x}$.

En efecto, $\tilde{x} = w$ y ya se sabe que los vectores w, y no están correlacionados. Además como $\hat{x}$ sólo depende de y, $\tilde{x}$ y $\hat{x}$ tampoco están correlacionados.

Antes de continuar con los tres últimos teoremas, se destaca enseguida ciertas propiedades del estimador $\hat{x}$.

- Es fácil ver en la ecuación (13)

$$E\,[\hat{x}] = \bar{x} \qquad (15)$$

Ello quiere decir que el estimador es insesgado, o, en otras palabras, que el valor esperado del error x es nulo, pues

$$E\,[\tilde{x}] = E\,[x - \hat{x}] = \bar{x} - \bar{x} = 0 \qquad (16)$$

- La ecuación (14) muestra la bondad del estimado $\hat{x}$ frente a otro estimado que fuese igual a la media $\bar{x}$, pues este último tendría un error de estimación cuya matriz de covarianza es S_{xx}, en tanto que aquél tiene un error cuya matriz de covarianza viene dada por (14) (téngase en cuenta que S_{xx} y S_{yy} son matrices definidas positivamente y que $S_{xy} S_{yy} S_{xy}$ es al menos definida no negativamente). En el peor de los casos, cuando los vectores x, y no están correlacionados ($S_{xy} = 0$), $S_{\tilde{x}\tilde{x}} = S_{xx}$. Muy ilustrativo es el caso cuando x, y son variables aleatorias escalares, de manera que las ecuaciones (13) y (14) se reducen a

$$\hat{x} = \mu_x + \rho_{xy} \frac{\sigma_x}{\sigma_y} (y - \mu_y)$$

$$\sigma_{\tilde{x}}{}^2 = \sigma_x{}^2 \, (1 - \rho_{xy}{}^2)$$

en donde μ_x y μ_y son las medias, $\sigma_x{}^2$ y $\sigma_y{}^2$ las varianzas y ρ_{xy} el coeficiente de correlación entre las variables x, y. El valor $(1 - \rho_{xy}{}^2)$ señala la proporción en que la varianza marginal, $\sigma_x{}^2$, es disminuida cuando se usa el estimado condicional.

- La teoría de la distribución multinormal demuestra que si los vectores x, y se distribuyen conjuntamente según la multinormal, entonces

$$E\,[x/y] = \bar{x} + S_{xy} S_{yy}^{-1} (y - \bar{y})$$

$$\mathrm{Cov}\,[x/y] = S_{xx} - S_{xy} S_{yy}^{-1} S_{yx}$$

Como en el caso gaussiano el vector medio condicional y la matriz de covarianza condicional definen completamente la distribución condicional de x dado y (que también es gaussiana), y como las expresiones de aquéllos coinciden con las ecuaciones (13) y (14), en ese caso el modelo lineal dado por (11) es una descripción completa del comportamiento condicional de x dado y. En caso contrario, el modelo de todas maneras preserva las propiedades de primero y segundo orden de x. Basta operar sobre la ecuación (11) para obtener

$$E[x] = \bar{x}$$

$$E[(x - \bar{x})(x - \bar{x})^T] = S_{xx}$$

$$E[(x - \bar{x})(y - \bar{y})^T] = S_{xy}$$

Se ha puesto así de presente la estrecha correspondencia entre los modelos lineales y los comportamientos gaussianos.

Es preciso mencionar que sólo en el caso gaussiano la media y la varianza condicionales proporcionadas por el modelo coinciden con la media y la varianza condicionales de la variable. Se presenta a continuación otra interpretación de la media condicional dada por el modelo.

- Supóngase que se tiene dos vectores aleatorios x, y. Disponiéndose del conocimiento de y, y sabiéndose que los vectores están correlacionados, se desea encontrar un estimador lineal de x que minimice el valor esperado del error cuadrático del estimado. Sin pérdida de generalidad, se supondrá que los vectores tienen media nula. El problema es, pues, encontrar una matriz a tal que minimice la expresión:

$$E[(x - \hat{x})(x - \hat{x})^T] = E[(x - ay)(x - ay)^T]$$

Desarrollando la expresión

$$E[(x - ay)(x - ay)^T] = \\ = S_{xx} - S_{xy}\, a^T - a\, S_{yx} + a\, S_{yy}\, a^T$$

Por simple manipulación algebráica, si S_{yy} es no singular, se llega a

$$E[(x - ay)(x - ay)^T] = \\ = (a - S_{xy} S_{yy}^{-1})\, S_{yy}\, (a - S_{xy} S_{yy}^{-1})^T + S_{xx} - S_{xy} S_{yy}^{-1} S_{yx}$$

Puesto que S_{yy} es definida positivamente, dicho valor esperado alcanza un mínimo cuando

$$a = S_{xy} S_{yy}^{-1}$$

Se concluye que la matriz así definida proporciona el error mínimo - cuadrático y que la expresión para a coincide con la del parámetro matricial del modelo lineal (11). Pero obsérvese que a diferencia del estimador lineal mínimo - cuadrático (que es la base para efectuar la regresión lineal múltiple multivariante, o ajuste mínimo - cuadrático, de una muestra de y sobre una muestra de x), el modelo lineal no es de la forma $\hat{x} = ay$ sino más bien de la forma $x = ay + w$, en el cual el valor de a y las propiedades de primero y segundo momentos de w se encontraron por medio de un análisis de segundo orden.

Continuando con el tema de esta sección, se va a generalizar los teoremas uno, dos y tres para cuando se desee estimar el comportamiento condicional de x dados los vectores y, z, ..., siendo éstos mutuamente no correlacionados. Por el interés inmediato que tiene para la sección 4, se considerará el caso de sólo dos vectores y, z. Ahora se propone el modelo

$$x = ay + bz + c + w \qquad (17)$$

Tomando valor esperado en ambos lados de (17) se encuentra el valor de c, y si $E[w] = 0$, se tendrá

$$x - \bar{x} = a(y - \bar{y}) + b(z - \bar{z}) + w \qquad (18)$$

Si w no está correlacionado con y ni con z, es fácil llegar a

$$a = S_{xy} S_{yy}^{-1} \qquad b = S_{xz} S_{zz}^{-1} \qquad (19)$$

Siempre que S_{yy} y S_{zz} sean no singulares. El nuevo modelo es

$$x = \bar{x} + S_{xy} S_{yy}^{-1} (y - \bar{y}) + S_{xz} S_{zz}^{-1} (z - \bar{z}) + w \qquad (20)$$

Teorema cuatro. Si el comportamiento condicional de x dados y, z viene dado por (20), entonces

$$E[x/y, z] = \bar{x} + S_{xy} S_{yy}^{-1} (y - \bar{y}) + S_{xz} S_{zz} (z - \bar{z}) \qquad (21)$$

Lo anterior es inmediato a partir de la misma ecuación (20).

Teorema cinco. Si el comportamiento condicional de x dados y, z viene dado por la ecuación (20), el error de un estimado que use el valor esperado condicional definido en el teorema cuatro, tendrá como matriz de covarianza.

$$S_{\tilde{x}\tilde{x}} = S_{xx} - S_{xy} S_{yy}^{-1} S_{yx} - S_{xz} S_{zz}^{-1} S_{zx} \qquad (22)$$

En efecto, las ecuaciones (20) y (21) conducen a

$$\tilde{x} = x - \hat{x} = w$$

Además, si en la ecuación (20) se resta $\tilde{x}$ de ambos lados, lo obtenido se posmultiplica en ambos lados por estos mismos traspuestos, y se toma valor esperado

$$S_{xx} = S_{xy}\, S_{yy}^{-1}\, S_{yy}\, S_{yy}^{-1}\, S_{yx} + S_{xz}\, S_{zz}^{-1}\, S_{zz}\, S_{zz}^{-1}\, S_{zx} + S_{ww}$$

siempre que w no esté correlacionado con y ni con z. Simplificando,

$$S_{ww} = S_{xx} - S_{xy}\, S_{yy}^{-1}\, S_{yx} - S_{xz}\, S_{zz}^{-1}\, S_{zx}$$

Q. E. D.

Teorema seis. $\tilde{x}$ no está correlacionado con y ni con z, y, por lo tanto, tampoco con $\hat{x}$.

En efecto, como $\tilde{x} = w$, $\tilde{x}$ no estará correlacionada con y ni con z. Además, como $\hat{x} = E[x/y, z]$ sólo depende de los vectores y, z según la ecuación (21), entonces $\tilde{x}$ y $\hat{x}$ no están correlacionados.

La sección termina mostrando que el estimado $\hat{x}$ dado por (21) es el estimador lineal mínimo - cuadrático. Con medias nulas, esa ecuación se reduce a

$$\hat{x} = ay + bz = [a \,\vdots\, b]\begin{bmatrix} y \\ \cdots \\ z \end{bmatrix} = c\, r$$

Bien se sabe que el valor de c en el caso mínimo cuadrático es

$$c = S_{xr}\, S_{rr}^{-1} = [S_{xy} \,\vdots\, S_{xz}] \begin{bmatrix} S_{yy} & 0 \\ 0 & S_{zz} \end{bmatrix}^{-1}$$

$$= [S_{xy}\, S_{yy}^{-1} \,\vdots\, S_{xz}\, S_{zz}^{-1}] = [a \,\vdots\, b]$$

Y estos valores de a y b coinciden con los de la ecuación (19).

4. DEDUCCION DEL FILTRO DE KALMAN

Volviendo a las ecuaciones básicas de estado y de medición vistas en la sección 2:

$$\delta x_{k+1} = A_k\, \delta x_k + B_k\, \delta u_k + \alpha_k \qquad (23)$$

$$\delta y_k = C_k\, \delta x_k + \beta_k \qquad (24)$$

Como se dijo en esa sección, el objetivo es, a partir de la dos ecuaciones y unas condiciones iniciales, hallar para cada k un estimado del estado y una medida de la bondad de dicho estimado dada por la matriz de covarianza del error del estimado. Para ello se tratará de usar toda la información suministrada por las ecuaciones (23) y (24), así como por las mediciones, incluyendo la última medición y_k.

En la misma sección 2 se señaló que si los comportamientos son gaussianos y se emplea un criterio de optimalidad de tipo cuadrático, la teoría del regulador lineal - cuadrático - gaussiano demuestra que la política de control es del tipo.

$$\delta u_k = - M_k\, \delta x_k$$

Existe, además, un teorema de dicha teoría, conocido como el teorema de la separabilidad, que permite resolver independientemente dos problemas:

- El problema de hallar la matriz M_k, que puede resolverse en forma determinística, sin tener en cuenta los aspectos estocásticos.

- El problema de hallar un estimado de δx_k, que puede resolverse mirando el control en forma determinística.

Por lo tanto, en lo que sigue se considerará, sin pérdida de generalidad, $u_k = 0$ para todo k. Además, no siendo necesaria, se levantará la hipótesis de comportamiento gaussiano.

Es conveniente definir la siguiente nomenclatura:

R_k : información hasta el tiempo k. incluyendo la medición y_k.

$\delta x_{k+1/k} = E\,[\delta x_{k+1}/R_k]$: estimado de δx_{k+1} condicionado a la información R_k.

$\delta \tilde{x}_{k+1/k} = \delta x_{k+1} - \delta x_{k+1/k}$: error del estimado $\delta x_{k+1/k}$

$\Sigma_{k+1/k} = \mathrm{Cov}\,[\delta \tilde{x}_{k+1/k}/R_k]$: matriz de covarianza del error del estimado, condicionada a R_k.

$\delta x_{k+1/k+1} = E\,[\delta x_{k+1}/R_{k+1}]$: estimado de δx_{k+1} condicionado a la información R_{k+1}

$\delta x_{k+1/k+1} = \delta x_{k+1} - \delta x_{k+1/k+1}$: error del estimado $\delta x_{k+1/k+1}$

$\Sigma_{k+1/k+1} = \mathrm{Cov}\,[\delta \tilde{x}_{k+1/k+1}/R_{k+1}]$: matriz de covarianza del error del estimado, condicionada a R_{k+1}

Para empezar, se supone que se conoce $\delta x_{0/0}$ y $\Sigma_{0/0}$, producto de las condiciones iniciales del sistema. Así, el producto de las condiciones iniciales del sistema. Así, el problema es escontrar recursivamente $\delta x_{k+1/k+1}$ y $\Sigma_{k+1/k+1}$ a partir de $\delta x_{k/k}$ y $\Sigma_{k/k}$.

Se procederá en tres etapas: predicción, cálculos auxiliares y actualización.

4.1 PREDICCION: Condicionado a la información R_k se hallará $\delta x_{k+1/k}$ y $\Sigma_{k+1/k}$ mediante una predicción basa en el modelo (23) o ecuación de estado con u_k = 0. Tomando valor esperado en ambos lados condicionado a R_k:

$$\delta x_{k+1/k} = E\,[\delta x_{k+1}/R_k] =$$

$$= Ak\, E\,[\delta x_k/R_k] = Ak\, \delta x_{k/k} \tag{25}$$

si α_k no está correlacionado con R_k (con los vectores incluidos en esta información) y si $E\,[\alpha_k] = 0$.

Si de la ecuación de estado se resta la (25)

$$\delta x_{k+1} - \delta x_{k+1/k} = A_k\,(\delta x_k - \delta x_{k/k}) + \alpha_k$$

$$\delta \tilde{x}_{k+1/k} = A_k\, \delta \tilde{x}_{k/k} + \alpha_k \tag{26}$$

Como el modelo dado por (23) es de tipo lineal, es aplicable la primera propiedad vista después del teorema tres, o sea la dada por la ecuación (16), y se tiene:

$$E\,[\delta \tilde{x}_{k+1/k}] = 0$$

Si se toma valor esperado en (26) :

$$0 = - Ak\, E\,[\delta \tilde{x}_{k/k}] + 0 \quad \therefore \quad E\,[\delta \tilde{x}_{k/k}] = 0$$

Si ambos lados de la ecuación (26) se posmultiplican por ellos mismos traspuestos y se toma valor esperado:

$$\Sigma_{k+1/k} = A_k\, \Sigma_{k/k}\, A_k^T + SA_k \tag{27}$$

siempre que α_k no esté correlacionado con R_k y que $E\,[\alpha_k\, \alpha_k^T] = SA_k$

Las ecuaciones (25) y (27) dan la predicción buscada así como la matriz de covarianza del error de la predicción, todo condicionado a R_k.

4.2 CALCULOS AUXILIARES: Ahora se encontrarán dos expresiones relacionadas con la medición δy_{k+1} y el valor esperado condicional de esa medición dado R_k, las cuales son necesarias para el numeral 4.3. Si se parte de la ecuación de medición (24) en el tiempo k+1 y definiendo

$$\delta y_{k+1/k} = E\,[\delta y_{k+1}/R_k]$$

puede llegarse a la siguiente expresión si se procede en forma análoga a como se obtuvo la ecuación (26)

$$\delta \tilde{y}_{k+1/k} = C_{k+1}\, \delta \tilde{x}_{k+1/k} + \beta_{k+1} \tag{28}$$

siempre que β_{k+1} no esté correlacionado con R_k y que $E\,[\beta_{k+1}] = 0$.

Como antes se llegó a la ecuación (27) también ahora se llega a una ecuación análoga para la matriz de covarianza de $\delta \tilde{y}_{k+1/k}$:

$$\Sigma^y_{k+1/k} = C_{k+1}\, \Sigma_{k+1/k}\, C_{k+1}{}^T + SB_{k+1} \tag{29}$$

siendo $SB_{k+1} = E\,[\beta_{k+1}\, \beta_{k+1}{}^T]$ y β_{k+1} no correlacionado con R_k.

Premultiplicando, en ambos lados de la ecuación (28) traspuesta, por δx_{k+1} y tomando valor esperado

$$E\,[\delta x_{k+1}\, \delta \tilde{y}_{k+1/k}{}^T] = E\,[\delta x_{k+1}\, \delta \tilde{x}_{k+1/k}{}^T]\, C_{k+1}{}^T \tag{30}$$

siempre que β_{k+1} no esté correlacionado con δx_{k+1}.

Pero recordando que $\delta x_{k+1} = \delta \tilde{x}_{k+1/k} + \delta x_{k+1/k}$ y que en virtud del teorema tres de la sección 3, $\delta x_{k+1/k}$ y $\delta \tilde{x}_{k+1/k}$ no están correlacionados:

$$E\,[\delta x_{k+1}\, \delta \tilde{y}_{k+1/k}{}^T] =$$

$$= E\,[\,(\delta \tilde{x}_{k+1/k} + \delta x_{k+1/k})\, \delta x_{k+1/k}{}^T]\, C^T_{k+1} =$$

$$= E\,[\delta \tilde{x}_{k+1/k}\, \delta \tilde{x}_{k+1/k}{}^T]\, C^T_{k+1}$$

$$= \Sigma_{k+1/k}\, C^T_{k+1} \tag{31}$$

4.3 ACTUALIZACION: En esta última etapa se hallará el estimado del estado incluyendo la última observación y_{k+1}, y se hallará también la correspondiente matriz de covarianza. Se tiene

$$\delta x_{k+1/k+1} = E\,[\delta x_{k+1}/R_{k+1}] \tag{32}$$

$$\delta y_{\tilde{k}+1/k} = \delta y_{k+1} - \delta \bar{y}_{k+1/k} \tag{33}$$

La última ecuación dice que R_{k+1} contiene la misma información que R_k y $\delta y_{k+1/k}$. Entonces la ecuación (32) es equivalente a

$$\delta x_{k+1/k+1} = E\,[\delta x_{k+1}/R_k, \delta y_{\tilde{k}+1/k}] \tag{34}$$

Como según el teorema tres de la sección 3, $\delta y_{\tilde{k}+1/k}$ no está correlacionado con R_k (evento condicionante), es aplicable el resultado del teorema cuatro de esa misma sección, dado por la ecuación (21):

$$\delta x_{k+1/k+1} = E\,[\delta x_{k+1}/R_k] + \mathrm{Cov}\,[\delta x_{k+1}; \delta y_{\tilde{k}+1/k}]$$
$$(\Sigma^{y}_{k+1/k})^{-1}\, \delta y_{\tilde{k}+1/k} \tag{35}$$

Obsérvese que $E\,[\delta x_{k+1}/R_k]$ corresponde a los dos primeros términos de (21). Pero además

$$\mathrm{Cov}\,[\delta x_{k+1};\delta y_{\tilde{k}+1/k}] = E\,[\delta x_{k+1}\, \delta y_{\tilde{k}+1/k}{}^{T}] =$$
$$= \Sigma_{k+1/k}\, C^{T}_{k+1} \tag{36}$$

en virtud de la ecuación (31).

Si (29) y (36) se llevan a (35):

$$\delta x_{k+1/k+1} = \delta x_{k+1/k} + \Sigma_{k+1/k}\, C^{T}_{k+1}$$
$$(C_{k+1}\, \Sigma_{k+1/k}\, C^{T}_{k+1} + SB_{k+1})^{-1}\, \delta y_{\tilde{k}+1/k} \tag{37}$$

Aplicando el teorema cinco de la sección 3, se obtiene la matriz de covarianza

$$\Sigma_{k+1/k+1} = \Sigma_{k+1/k} - E\,[\delta x_{k+1}\, \delta y_{\tilde{k}+1/k}{}^{T}]$$
$$(\Sigma^{y}_{k+1/k})^{-1}\, E[\delta y_{\tilde{k}+1/k}\, \delta x^{T}_{k+1}] \tag{38}$$

Obsérvese que $\Sigma_{k+1/k}$ corresponde a los dos primeros términos de (22). Usando los resultados (29) y (31) del numeral pasado, la expresión (38) se convierte en:

$$\Sigma_{k+1/k+1} = \Sigma_{k+1/k} - \Sigma_{k+1/k}\, C^{T}_{k+1}$$
$$(C_{k+1}\, \Sigma_{k+1/k}\, C^{T}_{k+1} + SB_{k+1})^{-1}\, C_{k+1}\, \Sigma_{k+1/k} \tag{39}$$

Téngase en cuenta que $\Sigma_{k+1/k}$ es una matriz simétrica.

4.4 COMENTARIOS FINALES: Las ecuaciones (37) y (39) definen el filtro de Kalman y sobre ellas conviene destacar:

- La ecuación (37) puede verse también como

$$\delta x_{k+1/k+1} = \delta x_{k+1/k} + H_{k+1}\,(\delta y_{k+1} - \delta y_{k+1/k}) \tag{40}$$

De modo que el estimado es igual a la predicción más una corrección tanto mayor cuanto más difiera la medición predicha de la medición real.

También (37) puede presentarse así:

$$\delta x_{k+1/k+1} = \delta x_{k+1/k} + H_{k+1}\,(\delta y_{k+1} - C_{k+1}\, \delta x_{k+1/k})$$

$$= (I - H_{k+1}\, C_{k+1})\, \delta x_{k+1/k} + H_{k+1}\, \delta y_{k+1}$$

O sea, el estimado es una combinación lineal de la predicción y la medición.

- Recuérdese que los estimados y las covarianzas se obtuvieron mediante valores esperados condicionales del tipo

E [x/y]	E [x/y, z]
Cov [x/y]	Cov [x/y, z]

calculados con base en modelos lineales.

Si las distribuciones subyacentes son multinormales, los resultados obtenidos son realmente las medias y covarianzas condicionales. En caso contrario, así no respondan a medias y covarianzas condicionales, los resultados obtenidos provienen de preservar primeros y segundos momentos (marginales y conjuntos), es decir, provienen de un análisis de segundo orden. También, según se vio al final de la sección 3, el filtro es el estimador lineal mínimo cuadrático.

- Si los comportamientos son gaussianos, puede demostrarse que el estimador de Kalman es también el de mínima varianza y el de máxima verosimilitud.

- La matriz de covarianza dada por la ecuación (39) no depende de las mediciones. Por lo tanto, es posible calcular su valor para todo k de interés antes que se conozca la evolución real del sistema en el tiempo. En otras palabras, la bondad del filtro puede establecerse con anticipación. No es éste el caso del estimado $\delta x_{k+1/k+1}$, que exige la llegada de la nueva medición δy_{k+1}.

BIBLIOGRAFIA

1. KUMAR, P. R. y P. VARAIYA. Stochastic Systems - Estimation, identification and adaptative control, Prentice Hall, Inc., Englewood Cliffs, New Jersey, Estados Unidos, 1986.

2. WHITTLE, P. Prediction and Regulation by Linear Least-Square Methods, University of Minnesota Press, Minneapolis, Minnesota, Estados Unidos, 1983.

3. DAVIS, M. H. A., y R. B. VINTER. Stochastic Modelling and Control, Chapman and Hall, Londres, Inglaterra, 1985.

4. BRAS, R. L. e I. RODRIGUEZ - ITURBE. Random Functions and Hydrology, Addison Wesley, Reading, Massachusetts, Estados Unidos, 1985.

5. SCHWEPPE, F. C. Uncertain Dynamic Systems, Prentice - Hall, New York, Estados Unidos, 1973.

6. BRYSON, A. E. y YU-CHI HO. Applied Optimal Control, Hemisphere Publishing Corporation, New York, Estados Unidos, 1975.

7. BERTSEKAS, D. P. Dynamic Programming and Stochastic Control, Academic Press, New York, Estados Unidos, 1976.

8. GELB, A. (Editor). Applied Optimal Estimation, The M. I. T. Press, Cambridge, Massachusetts. Estados Unidos, 1974.

9. Valencia, D. y J. C. SCHAAKE. Disaggregation processes in stochastic Hydrology, Water Resources Research, vol. 9, No. 3, pág. 580-585, American Geophysical Union, Estados Unidos, 1973.