



5 Filtros Kalman

El filtro Kalman completa el conjunto de **filtros óptimos lineales**. Algunas características distintivas son las siguientes:

- Se estudiará en términos de la **descripción por variables de estado**.
- La solución obtenida con el filtro Kalman se obtiene **recursivamente**.
- Puede utilizarse en la estimación de **estados y parámetros** de un sistema, aunque se utilizará esencialmente en la primera aplicación, a partir de la utilización de filtros FIR.
- El concepto de **innovación** es particularmente importante en su caracterización.
- Servirá esencialmente para un tratamiento unificado de una familia de algoritmos de filtrado adaptivo: **Cuadrados Mínimos Recursivo** (RLS).
- Se estudiará primero el **problema de estimación recursiva de mínima media cuadrática** (MMS) para el caso escalar y luego se extenderán los resultados al caso vectorial.

5.1 Estimación recursiva de MMS, caso escalar

- Asumiendo que los datos en (y antes de) $n = 0$ son cero y en base a un conjunto de datos $y(1), y(2), \dots, y(n-1)$ (Y_{n-1}), se ha estudiado el estimador de **mínima media cuadrática (MMS)** de una variable $x(n)$, $\hat{x}(n-1|Y_{n-1})$.
- Con $y(n)$, en el instante n es posible obtener una estimación $\hat{x}(n|Y_n)$ en base ahora a los datos almacenados $y(1), y(2), \dots, y(n-1)$ e $y(n)$.
- Computacionalmente es eficiente obtener esa estimación en forma **recursiva**. Existen varios procedimientos para realizarlo. Definiendo el error de predicción forward como

$$f_{n-1}(n) = y(n) - \hat{y}(n|Y_{n-1}), \quad n = 1, 2, \dots$$

donde $\hat{y}(n|Y_{n-1})$ es la **predicción de un paso** de $y(n)$ utilizando los datos pasados hasta $n-1$. $f_{n-1}(n)$ es la salida de un filtro de predicción forward de orden $n-1$, cuyo orden aumenta linealmente con n .

- Por el principio de ortogonalidad, $f_{n-1}(n)$ es ortogonal a $y(1), y(2), \dots, y(n-1)$ y puede interpretarse como una medida de la información **nueva** que aporta $y(n)$, por ello se introduce el nombre de **innovación**. El concepto está relacionado con que $y(n)$ no aporta información completamente nueva, sino que existe una parte predecible $\hat{y}(n|Y_{n-1})$.
- La parte realmente nueva asociada a $y(n)$ está dada por $f_{n-1}(n)$, que definimos como la innovación dada por

$$\alpha(n) = f_{n-1}(n), \quad n = 1, 2, \dots$$

- Algunas propiedades del proceso de innovación $\alpha(n)$:

1. $\alpha(n)$, asociado con $y(n)$ es ortogonal a los datos pasados $y(1), y(2), \dots, y(n-1)$, ya que

$$E[\alpha(n)y^*(k)] = 0, \quad 1 \leq k \leq n-1$$

Esto es simplemente otra forma de escribir el principio de ortogonalidad.

2. $\alpha(1), \alpha(2), \dots, \alpha(n)$ son ortogonales entre si, de forma que

$$E[\alpha(n)\alpha^*(k)] = 0, \quad 1 \leq k \leq n-1$$

Esto es la reformulación de

$$E[f_{n-1}(n)f_{n-1}^*(k)] = 0, \quad 1 \leq k \leq n-1$$

Esto implica que $\alpha(n)$ es un proceso **blanco**.

3. Existe una correspondencia uno a uno entre $y(1), y(2), \dots, y(n)$ y $\alpha(1), \alpha(2), \dots, \alpha(n)$, en el sentido que una secuencia puede obtenerse de la otra por medio de un filtro causal (y causalmente inversible) sin ninguna pérdida de información, o sea

$$[y(1), y(2), \dots, y(n)] \Leftrightarrow [\alpha(1), \alpha(2), \dots, \alpha(n)]$$

Para probar esta propiedad se utilizará una forma del **procedimiento de ortogonalización de Gram-Schmidt**.

- Suponiendo los datos $y(1), y(2), \dots, y(n)$ algebraicamente independientes, es posible hacer $\alpha(1) = y(1)$ (asumiendo $\hat{y}(1|Y_0)$).

- Luego hacemos $\alpha_2 = y(2) + a_{1,1}y(1)$. El coeficiente $a_{1,1}$ se elige de forma que $\alpha(1)$ y $\alpha(2)$ sean ortogonales, o sea $E[\alpha(2)\alpha(1)] = 0$. Esto se cumple para

$$a_{1,1} = -\frac{E[y(2)y^*(1)]}{E[y(1)y(1)]}$$

- Haciendo luego $\alpha(3) = y(3) + a_{2,1}y(2) + a_{2,2}y(1)$, donde $a_{2,1}$ y $a_{2,2}$ se eligen tal que $\alpha(3)$ sea ortogonal a $\alpha(1)$ y $\alpha(2)$.
- Así en general

$$\begin{bmatrix} \alpha(1) \\ \alpha(2) \\ \vdots \\ \alpha(n) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_{1,1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ a_{n-1,n-1} & a_{n-1,n-2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(n) \end{bmatrix}$$

- Los elementos de la fila k de esta matriz triangular inferior de transformación se denotan $a_{k-1,k-1}, a_{k-1,k-2}, \cdots, 1$, para $k = 1, 2, \cdots, n$. Estos elementos representan los coeficientes de un **filtro de error de predicción** de orden $k - 1$.

- En base a la propiedad 3 es posible escribir

$$\begin{aligned}\hat{x}(n|Y_n) &= \text{estimación de MMS de } x(n) \text{ dados } y(1), y(2), \dots, y(n) \\ &= \text{estimación de MMS de } x(n) \text{ dados } \alpha(1), \alpha(2), \dots, \alpha(n)\end{aligned}$$

- Luego, definiendo $\hat{x}(n|Y_n)$ como una combinación lineal de las innovaciones $\alpha(1), \alpha(2), \dots, \alpha(n)$, o sea

$$\hat{x}(n|Y_n) = \sum_{k=1}^n b_k \alpha(k)$$

donde los coeficientes b_k deben ser determinados.

- Utilizando la ortogonalidad de las innovaciones $\alpha(k)$ y eligiendo los b_k para minimizar la media cuadrática de $x(n) - \hat{x}(n|Y_n)$ (usando un procedimiento de ortogonalización como el anterior), hallamos que

$$b_k = \frac{E[x(n)\alpha^*(k)]}{E[\alpha(k)\alpha^*(k)]}, \quad 1 \leq k \leq n$$

- Notar que $\hat{x}(n|Y_n) = \sum_{k=1}^{n-1} b_k \alpha(k) + b_n \alpha(n)$, donde $b_n = \frac{E[x(n)\alpha^*(n)]}{E[\alpha(n)\alpha^*(n)]}$. Pero, por definición, esto se puede escribir en forma recursiva como

$$\hat{x}(n|Y_n) = \hat{x}(n-1|Y_{n-1}) + b_n \alpha(n) \quad (43)$$

5.2 El problema de filtrado Kalman

- Un sistema dinámico lineal discreto puede describirse por

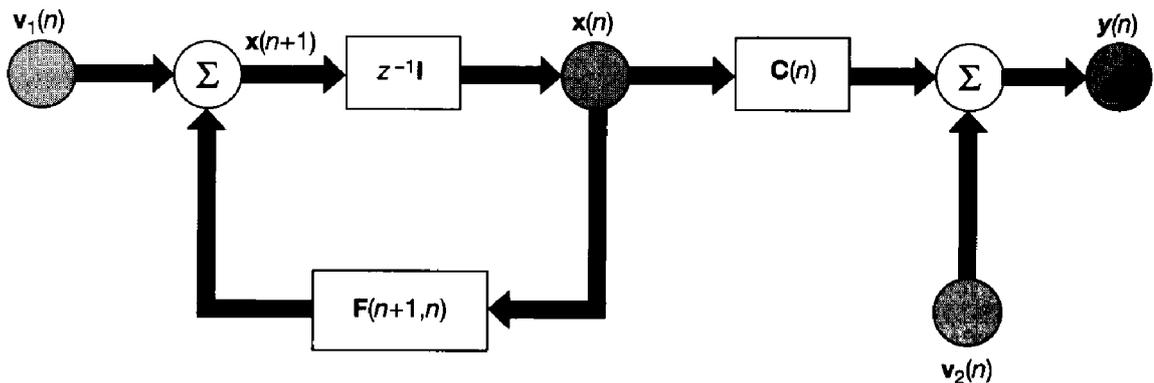


Figura 14: Representación de un sistema dinámico lineal discreto.

- $\mathbf{x}(n)$ (de dimensión M , desconocido) es suficiente para describir unívocamente el comportamiento dinámico del sistema no forzado. Para estimarlo se utiliza el conjunto de datos observados $\mathbf{y}(n)$, de dimensión N .
- Las ecuaciones relacionadas son las siguientes:

1. Una **ecuación del proceso**, dada por

$$\mathbf{x}(n+1) = \mathbf{F}(n+1, n)\mathbf{x}(n) + \mathbf{v}_1(n) \quad (44)$$

donde $\mathbf{F}(n+1, n)$ ($M \times M$) es la **matriz de transición de estados**, conocida. $\mathbf{v}_1(n)$ ($M \times 1$) representa el **ruido de proceso**, modelado por un proceso de ruido blanco, media cero y matriz correlación

$$E[\mathbf{v}_1(n)\mathbf{v}_1^H(n)] = \begin{cases} \mathbf{Q}_1(n), & n = k \\ \mathbf{0} & n \neq k \end{cases}$$

2. Una ecuación de medición, descripta por

$$\mathbf{y}(n) = \mathbf{C}(n)\mathbf{x}(n) + \mathbf{v}_2(n) \quad (45)$$

donde $\mathbf{C}(n)$ ($N \times M$) es la **matriz de medición**, conocida. $\mathbf{v}_2(n)$ ($N \times 1$) representa el **ruido de medición**, modelado por un proceso de ruido blanco, media cero y matriz correlación

$$E[\mathbf{v}_2(n)\mathbf{v}_2^H(n)] = \begin{cases} \mathbf{Q}_2(n), & n = k \\ \mathbf{0} & n \neq k \end{cases}$$

Se supone que el valor inicial del estado $\mathbf{x}(0)$ no está correlacionado con $\mathbf{v}_1(n)$ y $\mathbf{v}_2(n)$ para $n \geq 0$ y además que $\mathbf{v}_1(n)$ y $\mathbf{v}_2(n)$ son estadísticamente independientes, o sea $E[\mathbf{v}_1(n)\mathbf{v}_2^H(k)] = \mathbf{0}$ para todo n y k .

El problema de filtrado Kalman puede formalizarse como sigue:

Utilizar los datos completos, consistentes en $\mathbf{y}(1)$, $\mathbf{y}(2)$, \dots , $\mathbf{y}(n)$ para obtener, para cada $n \geq 1$, la estimación de mínima media cuadrática de (MMS) de los componentes del estado $\mathbf{x}(i)$.

El problema se denomina de **filtrado** si $i = n$, de **predicción** si $i > n$ y de **suavizado** si $1 \leq i < n$.

5.3 El proceso de Innovación $\alpha(n)$

- Sea $\hat{\mathbf{y}}(n|\mathbf{Y}_{n-1})$ la estimación MMS de los datos $\mathbf{y}(n)$ en el instante n , dados $\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(n-1)$ que forman el espacio vectorial \mathbf{Y}_{n-1} .
- El proceso de **innovación** asociado a $\mathbf{y}(n)$ se define como

$$\alpha(n) = \mathbf{y}(n) - \hat{\mathbf{y}}(n|\mathbf{Y}_{n-1}), \quad n = 1, 2, \dots$$

- Generalizando los resultados del caso escalar, las propiedades del proceso de innovación $\alpha(n)$ ($M \times 1$) son las siguientes
 1. $\alpha(n)$ asociado a $\mathbf{y}(n)$ es ortogonal a $\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(n-1)$, o sea

$$E[\alpha(n)\mathbf{y}^H(k)] = \mathbf{0}, \quad 1 \leq k \leq n-1$$

2. La secuencia de procesos de innovación es ortogonal, o sea

$$E[\alpha(n)\alpha^H(k)] = \mathbf{0}, \quad 1 \leq k \leq n-1$$

3. Existe una correspondencia uno a uno entre la secuencia de datos y las secuencias de innovaciones, en el sentido que una puede obtenerse de la otra por medio de operadores lineales estables sin pérdida de información. Esto puede escribirse como

$$[\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(n-1)] \rightleftharpoons [\alpha(1), \alpha(2), \dots, \alpha(n-1)]$$

para obtener este mapeo puede utilizarse un procedimiento de ortogonalización como el analizado anteriormente.

5.3.1 Matriz correlación del proceso de innovación $\mathbf{R}(n)$

- Para obtener la matriz de correlación del proceso de innovación $\boldsymbol{\alpha}(n)$ tendremos en cuenta la solución de la ecuación (44), dada por

$$\mathbf{x}(k) = \mathbf{F}(k, 0)\mathbf{x}(0) + \sum_{i=1}^{k-1} \mathbf{F}(k, i+1)\mathbf{v}_1(i)$$

donde se observa que $\mathbf{x}(k)$ es una combinación lineal de $\mathbf{x}(0)$ y $\mathbf{v}_1(1), \mathbf{v}_1(2), \dots, \mathbf{v}_1(k-1)$ y además se utilizaron las siguientes definiciones y suposiciones

1. $\mathbf{x}(0)$ es el valor inicial del estado.
2. Los datos (y $\mathbf{v}_1(0)$) son 0 para $n \leq 0$.
3. La matriz de transición de estados satisface

$$\mathbf{F}(k, k-1)\mathbf{F}(k-1, k-2) \cdots \mathbf{F}(i+1, i) = \mathbf{F}(k, i)$$

y $\mathbf{F}(k, k) = \mathbf{I}$. Para un sistema invariante en el tiempo

$$\mathbf{F}(n+1, n) = \mathbf{F}(n+1-n) = \mathbf{F}(1) = \mathbf{I}$$

- Por hipótesis, es fácil verificar que: ($\mathbf{v}_2(n)$ no correlacionado con $\mathbf{x}(0)$ y $\mathbf{v}_1(n)$)

$$E[\mathbf{x}(k)\mathbf{v}_2^H(n)] = \mathbf{0}, \quad k, n \leq 0$$

- Luego de (45) es posible concluir que

$$E[\mathbf{y}(k)\mathbf{v}_2^H(n)] = \mathbf{0}, \quad 0 \leq k \leq n-1$$

- Además es posible escribir que

$$E[\mathbf{y}(k)\mathbf{v}_1^H(n)] = \mathbf{0}, \quad 0 \leq k \leq n$$

- Dados los datos pasados $\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(n-1)$ y usando la ecuación (45) es posible hallar que la estimación MMS de $\mathbf{y}(n)$ es: $\hat{\mathbf{y}}(n|\mathbf{Y}_{n-1}) = \mathbf{C}(n)\hat{\mathbf{x}}(n|\mathbf{Y}_{n-1}) + \hat{\mathbf{v}}_2(n|\mathbf{Y}_{n-1})$.
- Pero la estimación $\hat{\mathbf{v}}_2(n|\mathbf{Y}_{n-1})$ es cero ya que $\mathbf{v}_2(n)$ es ortogonal a \mathbf{Y}_{n-1} . De esto puede obtenerse

$$\hat{\mathbf{y}}(n|\mathbf{Y}_{n-1}) = \mathbf{C}(n)\hat{\mathbf{x}}(n|\mathbf{Y}_{n-1})$$

de donde el proceso de innovación puede escribirse

$$\boldsymbol{\alpha}(n) = \mathbf{y}(n) - \mathbf{C}(n)\hat{\mathbf{x}}(n|\mathbf{Y}_{n-1}) \quad (46)$$

- Sustituyendo la ecuación de medición (45) en la anterior se obtiene

$$\boldsymbol{\alpha}(n) = \mathbf{C}(n)\boldsymbol{\epsilon}(n, n-1) + \mathbf{v}_2(n)$$

donde $\boldsymbol{\epsilon}(n, n-1)$ es el **error de predicción** del vector de estados en el instante n utilizando datos hasta el instante $n-1$, o sea

$$\boldsymbol{\epsilon}(n, n-1) = \mathbf{x}(n) - \hat{\mathbf{x}}(n|\mathbf{Y}_{n-1}) \quad (47)$$

Notar que $\boldsymbol{\epsilon}(n, n-1)$ es ortogonal a $\mathbf{v}_1(n)$ y $\mathbf{v}_2(n)$.

- Finalmente la matriz de correlación del proceso de innovación se define por

$$\mathbf{R}(n) = E[\boldsymbol{\alpha}(n)\boldsymbol{\alpha}^H(n)]$$

Luego, sustituyendo (47) en la anterior, expandiendo términos, y usando la ortogonalidad entre $\boldsymbol{\epsilon}(n, n - 1)$ y $\mathbf{v}_2(n)$, obtenemos

$$\mathbf{R}(n) = \mathbf{C}(n)\mathbf{K}(n, n - 1)\mathbf{C}^H(n) + \mathbf{Q}_2(n) \quad (48)$$

donde $\mathbf{K}(n, n - 1)$ es la matriz de correlación del error de predicción de estados, definida por

$$\mathbf{K}(n, n - 1) = E[\boldsymbol{\epsilon}(n, n - 1)\boldsymbol{\epsilon}^H(n, n - 1)]$$

5.4 Predicción usando $\alpha(n)$

- La estimación de MMS de $\mathbf{x}(i)$ puede obtenerse como una combinación lineal de la secuencia de procesos de innovación $\alpha(1)$, $\alpha(2)$, \dots , $\alpha(n)$, de la misma forma que en el caso escalar, o sea

$$\hat{\mathbf{x}}(i|\mathbf{Y}_n) = \sum_{k=1}^n \mathbf{B}_i(k)\alpha(k) \quad (49)$$

donde $\mathbf{B}_i(k)$, $k = 1, 2, \dots, n$ es un conjunto de matrices de $(M \times N)$ a ser determinado.

- De acuerdo al principio de ortogonalidad, el error de predicción del vector de estados es ortogonal al proceso de innovación, como mostrado por

$$E[\epsilon(i, n)\alpha^H(m)] = E\{[\mathbf{x}(i) - \hat{\mathbf{x}}(i|\mathbf{Y}_n)]\alpha^H(m)\} = \mathbf{0}, \quad m = 1, 2, \dots, n$$

- Sustituyendo la primera ecuación en la segunda y utilizando la ortogonalidad del proceso de innovación se obtiene

$$E[\mathbf{x}(i)\alpha^H(m)] = \mathbf{B}_i(m)E[\alpha(m)\alpha^H(m)] = \mathbf{B}_i(m)\mathbf{R}(m)$$

de donde

$$\mathbf{B}_i(m) = E[\mathbf{x}(i)\alpha^H(m)]\mathbf{R}^{-1}(m)$$

- Finalmente sustituyendo esta ecuación en (49) la estimación de MMS puede escribirse

$$\begin{aligned} \hat{\mathbf{x}}(i|\mathbf{Y}_n) &= \sum_{k=1}^n E[\mathbf{x}(i)\alpha^H(k)]\mathbf{R}^{-1}(k)\alpha(k) \\ &= \sum_{k=1}^{n-1} E[\mathbf{x}(i)\alpha^H(k)]\mathbf{R}^{-1}(k)\alpha(k) + E[\mathbf{x}(i)\alpha^H(n)]\mathbf{R}^{-1}(n)\alpha(n) \end{aligned}$$

y para $i = n + 1$ es posible escribir

$$\begin{aligned}\hat{\mathbf{x}}(n+1|\mathbf{Y}_n) &= \sum_{k=1}^n E[\mathbf{x}(n+1)\boldsymbol{\alpha}^H(k)]\mathbf{R}^{-1}(k)\boldsymbol{\alpha}(k) \\ &= \sum_{k=1}^{n-1} E[\mathbf{x}(n+1)\boldsymbol{\alpha}^H(k)]\mathbf{R}^{-1}(k)\boldsymbol{\alpha}(k) + E[\mathbf{x}(n+1)\boldsymbol{\alpha}^H(n)]\mathbf{R}^{-1}(n)\boldsymbol{\alpha}(n)\end{aligned}$$

- Para reducir esta ecuación, en particular, usando la matriz de transición de estados que relaciona $\mathbf{x}(n+1)$ con $\mathbf{x}(n)$, es posible escribir para $0 \leq k \leq n$

$$E[\mathbf{x}(n+1)\boldsymbol{\alpha}^H(k)] = E\{\mathbf{F}(n+1, n)\mathbf{x}(n) + \mathbf{v}_1(n)\}\boldsymbol{\alpha}^H(k) = \mathbf{F}(n+1, n)E[\mathbf{x}(n)\boldsymbol{\alpha}^H(k)] \quad (50)$$

Luego

$$\begin{aligned}\sum_{k=1}^{n-1} E[\mathbf{x}(n+1)\boldsymbol{\alpha}^H(k)]\mathbf{R}^{-1}(k)\boldsymbol{\alpha}(k) &= \mathbf{F}(n+1, n) \sum_{k=1}^{n-1} E[\mathbf{x}(n)\boldsymbol{\alpha}^H(k)]\mathbf{R}^{-1}(k)\boldsymbol{\alpha}(k) \\ &= \mathbf{F}(n+1, n)\hat{\mathbf{x}}(n|\mathbf{Y}_{n-1})\end{aligned}$$

Para obtener una forma más simple del estimador se discutirán e introducirán a continuación algunas definiciones.

5.4.1 Ganancia de Kalman $\mathbf{G}(n)$

- Definamos la matriz de $(M \times N)$, $\mathbf{G}(n) = E[\mathbf{x}(n+1)\boldsymbol{\alpha}^H(n)]\mathbf{R}^{-1}(n)$.
- Luego usando el resultado de la sección anterior es posible escribir la estimación MMS de estados como

$$\hat{\mathbf{x}}(n+1|\mathbf{Y}_n) = \mathbf{F}(n+1, n)\hat{\mathbf{x}}(n|\mathbf{Y}_{n-1}) + \mathbf{G}(n)\boldsymbol{\alpha}(n) \quad (51)$$

- Esta es una versión vectorial de (43) y es especialmente importante porque muestra que para obtener la estimación de MMS de $\hat{\mathbf{x}}(n+1|\mathbf{Y}_n)$ de un sistema dinámico es suficiente sumar a la estimación previa $\hat{\mathbf{x}}(n|\mathbf{Y}_{n-1})$ un término de corrección.
- Esencialmente la matriz $\mathbf{G}(n)$ juega el papel de una **ganancia** en relación al proceso de innovación $\boldsymbol{\alpha}(n)$.
- Para obtener una forma computacionalmente factible de $\mathbf{G}(n)$ es posible tener en cuenta que

$$\begin{aligned} E[\mathbf{x}(n+1)\boldsymbol{\alpha}^H(n)] &= \mathbf{F}(n+1, n)E[\mathbf{x}(n)\boldsymbol{\alpha}^H(n)] \\ &= \mathbf{F}(n+1, n)E[\mathbf{x}(n)(\mathbf{C}(n)\boldsymbol{\epsilon}(n, n-1) + \mathbf{v}_2(n))^H] \\ &= \mathbf{F}(n+1, n)E[\mathbf{x}(n)\boldsymbol{\epsilon}^H(n, n-1)]\mathbf{C}^H(n) \end{aligned}$$

- Como $\epsilon(n, n-1)$ es ortogonal a $\hat{\mathbf{x}}(n | \mathbf{Y}_{n-1})$ es posible reemplazar

$$\begin{aligned} E[\mathbf{x}(n+1)\alpha^H(n)] &= \mathbf{F}(n+1, n)E[\epsilon(n, n-1)\epsilon^H(n, n-1)]\mathbf{C}^H(n) \\ &= \mathbf{F}(n+1, n)\mathbf{K}(n, n-1)\mathbf{C}^H(n) \end{aligned}$$

de forma que sustituyendo esta ecuación en la definición de $\mathbf{G}(n)$ se obtiene

$$\mathbf{G}(n) = \mathbf{F}(n+1, n)\mathbf{K}(n, n-1)\mathbf{C}^H(n)\mathbf{R}^{-1}(n) \quad (52)$$

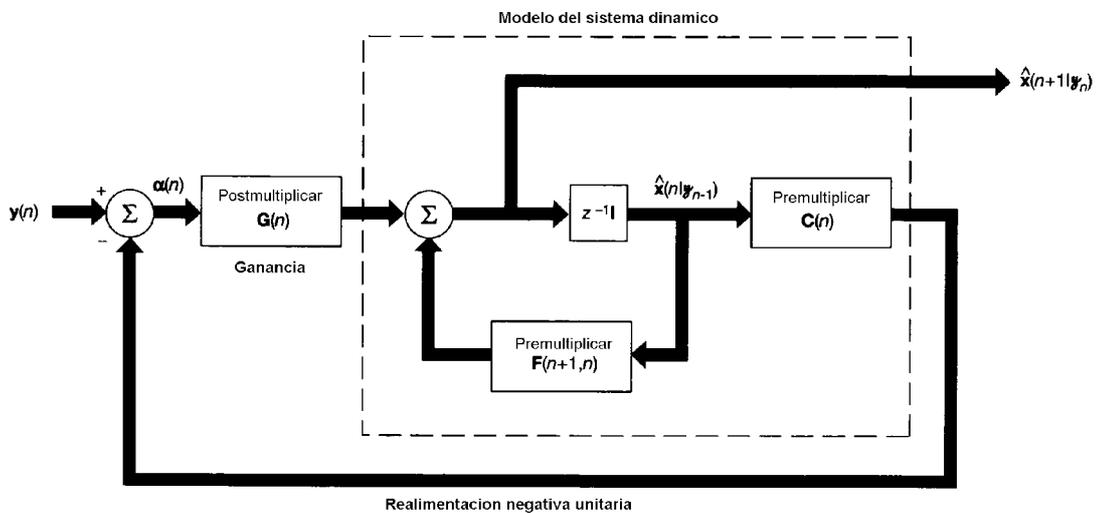


Figura 15: Predictor de un paso.

5.4.2 Ecuación de Riccati

- La ecuación de $\mathbf{G}(n)$ aún no es particularmente útil dado que se requiere $\mathbf{K}(n, n-1)$. El objetivo entonces es obtener recursivamente $\mathbf{K}(n, n-1) = E[\boldsymbol{\epsilon}(n+1, n)\boldsymbol{\epsilon}^H(n+1, n)]$.
- Luego como

$$\begin{aligned}\boldsymbol{\epsilon}(n+1, n) &= \mathbf{F}(n+1, n)[\mathbf{x}(n) - \hat{\mathbf{x}}(n|Y_{n-1})] - \mathbf{G}(n)[\mathbf{y}(n) - \mathbf{C}(n)\hat{\mathbf{x}}(n|Y_{n-1})] + \mathbf{v}_1(n) \\ &= [\mathbf{F}(n+1, n) - \mathbf{G}(n)\mathbf{C}(n)]\boldsymbol{\epsilon}(n, n-1) + \mathbf{v}_1(n) - \mathbf{G}(n)\mathbf{v}_2(n)\end{aligned}$$

se obtiene que

$$\begin{aligned}\mathbf{K}(n+1, n) &= [\mathbf{F}(n+1, n) - \mathbf{G}(n)\mathbf{C}(n)]\mathbf{K}(n, n-1)[\mathbf{F}(n+1, n) - \mathbf{G}(n)\mathbf{C}(n)]^H \\ &\quad + \mathbf{Q}_1(n) - \mathbf{G}(n)\mathbf{Q}_2(n)\mathbf{G}(n)\end{aligned}$$

- Expandiendo el lado derecho de esta ecuación y utilizando (52) y (48) se obtiene la **ecuación a diferencias de Riccati** para obtener en forma recursiva la matriz de correlación del error de predicción de los estados, dada por

$$\mathbf{K}(n+1, n) = \mathbf{F}(n+1, n)\mathbf{K}(n)\mathbf{F}^H(n+1, n) + \mathbf{Q}_1(n) \quad (53)$$

donde, en particular usando $\mathbf{F}(n+1, n)\mathbf{F}(n, n+1) = \mathbf{I}$, se obtiene la recursividad para $\mathbf{K}(n)$ ($M \times M$), dada por

$$\mathbf{K}(n) = \mathbf{K}(n, n-1) - \mathbf{F}(n, n+1)\mathbf{G}(n)\mathbf{C}(n)\mathbf{K}(n, n-1) \quad (54)$$

esta matriz tiene una interpretación directa en el contexto de filtrado que se discutirá en la sección siguiente.



Comentarios

- La entrada al filtro Kalman son los datos $\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(n)$ y la salida es la predicción del estado $\hat{\mathbf{x}}(n+1|\mathbf{Y}_n)$.
- Como $\mathbf{F}(n+1, n), \mathbf{C}(n), \mathbf{Q}_1(n)$ y $\mathbf{Q}_2(n)$ se suponen conocidas, de las ecuaciones (51) y (53) es posible verificar que $\mathbf{K}(n+1, n)$ es independiente de $\mathbf{y}(n)$, como era de esperar. La ganancia $\mathbf{G}(n)$ es también independiente de $\mathbf{y}(n)$. Luego ambas pueden ser obtenidas **antes** de poner al filtro en operación.
- $\mathbf{K}(n+1, n)$ permite un análisis estadístico del error en la predicción del vector de estados, de forma de conocer a priori si la solución que se obtiene con el filtro de Kalman es satisfactoria.
- Si las matrices $\mathbf{F}(n+1, n), \mathbf{C}(n), \mathbf{Q}_1(n)$ y $\mathbf{Q}_2(n)$ no son totalmente conocidas a priori, puede generalizarse los resultados del filtro de Kalman cuando estas dependen de $\mathbf{y}(n)$. En ese caso $\hat{\mathbf{x}}(n+1|\mathbf{Y}_n)$ y $\mathbf{K}(n+1, n)$ aún se obtienen por (51) y (53), respectivamente, pero ahora dependerán de $\mathbf{y}(n)$. En particular $\mathbf{K}(n+1, n)$ es una matriz de error de predicción **condicional** sobre la entrada $\mathbf{y}(n)$.

5.5 Filtrado

- En particular se desea obtener la estimación $\hat{\mathbf{x}}(n|\mathbf{Y}_n)$ usando el algoritmo de predicción de un paso descrito anteriormente.
- Para hallar la estimación de MMS de $\mathbf{x}(n+1)$ a partir de la ecuación del proceso (44) (teniendo en cuenta que $\mathbf{v}_1(n)$ es independiente de $\mathbf{x}(n)$ y de las entradas)

$$\hat{\mathbf{x}}(n+1|\mathbf{Y}_n) = \mathbf{F}(n+1, n)\hat{\mathbf{x}}(n|\mathbf{Y}_n)$$

- Luego

$$\hat{\mathbf{x}}(n|\mathbf{Y}_n) = \mathbf{F}^{-1}(n+1, n)\hat{\mathbf{x}}(n+1|\mathbf{Y}_n) \quad (55)$$

$$= \mathbf{F}(n, n+1)\hat{\mathbf{x}}(n+1|\mathbf{Y}_n) \quad (56)$$

5.5.1 Error de estimación y factor de conversión en filtrado

- En un contexto de filtrado es natural definir un **error de estimación de filtrado**, dado por

$$\mathbf{e}(n) = \mathbf{y}(n) - \mathbf{C}(n)\hat{\mathbf{x}}(n|\mathbf{Y}_n)$$

- Esta definición es similar a la definición de $\boldsymbol{\alpha}(n)$, ecuación (46), excepto que se ha sustituido la predicción $\hat{\mathbf{x}}(n|\mathbf{Y}_{n-1})$ por la estimación de filtrado $\hat{\mathbf{x}}(n|\mathbf{Y}_n)$. Usando (51) en la anterior se obtiene

$$\begin{aligned} \mathbf{e}(n) &= \mathbf{y}(n) - \mathbf{C}(n)\hat{\mathbf{x}}(n|\mathbf{Y}_{n-1}) - \mathbf{C}(n)\mathbf{F}(n, n+1)\mathbf{G}(n)\boldsymbol{\alpha}(n) \\ &= \boldsymbol{\alpha}(n) - \mathbf{C}(n)\mathbf{F}(n, n+1)\mathbf{G}(n)\boldsymbol{\alpha}(n) \\ &= [\mathbf{I} - \mathbf{C}(n)\mathbf{F}(n, n+1)\mathbf{G}(n)]\boldsymbol{\alpha}(n) \end{aligned}$$

- La matriz $[\mathbf{I} - \mathbf{C}(n)\mathbf{F}(n, n+1)\mathbf{G}(n)]$ se denomina **factor de conversión**, y convierte el vector de innovación $\boldsymbol{\alpha}(n)$ en el vector de error de estimación de filtrado $\mathbf{e}(n)$. Usando (52) es posible eliminar $\mathbf{G}(n)$ de esta ecuación para obtener

$$\mathbf{e}(n) = \mathbf{Q}_2(n)\mathbf{R}^{-1}(n)\boldsymbol{\alpha}(n)$$

- Así, excepto por $\mathbf{Q}_2(n)$, la matriz $\mathbf{R}(n)$ juega el papel de factor de conversión en la teoría de filtrado Kalman. Además en el caso en que $\mathbf{Q}_2(n) = \mathbf{I}$, \mathbf{R}^{-1} es el factor de conversión exacto.

5.5.2 Matriz de correlación del error de estados en filtrado $\mathbf{K}(n)$

- Analizaremos una interpretación de la matriz $\mathbf{K}(n)$ introducida en la formulación de la ecuación de Riccati, a partir del cual será posible concluir que esta matriz es igual a la matriz de correlación del error de estimación en filtrado.
- Definimos el vector de error de estimación en filtrado $\boldsymbol{\epsilon}(n)$ como

$$\boldsymbol{\epsilon}(n) = \mathbf{x}(n) - \hat{\mathbf{x}}(n|\mathbf{Y}_n)$$

- Usando (51) y (55) en la anterior se obtiene

$$\begin{aligned} \boldsymbol{\epsilon}(n) &= \mathbf{x}(n) - \hat{\mathbf{x}}(n|\mathbf{Y}_{n-1}) - \mathbf{F}(n, n+1)\mathbf{G}(n)\boldsymbol{\alpha}(n) \\ &= \boldsymbol{\epsilon}(n, n-1) - \hat{\mathbf{x}}(n|\mathbf{Y}_{n-1}) - \mathbf{F}(n, n+1)\mathbf{G}(n)\boldsymbol{\alpha}(n) \end{aligned}$$

- Luego la matriz correlación del error de estimación en filtrado será

$$\begin{aligned} E[\boldsymbol{\epsilon}(n)\boldsymbol{\epsilon}^H(n)] &= E[\boldsymbol{\epsilon}(n, n-1)\boldsymbol{\epsilon}^H(n, n-1)] \\ &\quad + \mathbf{F}(n, n+1)\mathbf{G}(n)E[\boldsymbol{\alpha}(n)\boldsymbol{\alpha}^H(n)]\mathbf{G}^H(n)\mathbf{F}^H(n, n+1) \\ &\quad - 2E[\boldsymbol{\epsilon}(n, n-1)\boldsymbol{\alpha}^H(n)]\mathbf{G}^H(n)\mathbf{F}^H(n, n+1) \end{aligned}$$

- Para reducir esta expresión es posible tener en cuenta que las esperanzas del lado derecho pueden interpretarse como sigue:
 1. El primer término es la matriz correlación del error de predicción de los estados $\mathbf{K}(n, n - 1)$.
 2. El segundo término es la matriz de correlación de la innovación $\mathbf{R}(n)$.
 3. Para interpretar el último término tenemos en cuenta que

$$\begin{aligned} E[\boldsymbol{\epsilon}(n, n - 1)\boldsymbol{\alpha}^H(n)] &= E[(\mathbf{x}(n) - \hat{\mathbf{x}}(n|\mathbf{Y}_{n-1}))\boldsymbol{\alpha}^H(n)] \\ &= E[\mathbf{x}(n)\boldsymbol{\alpha}^H(n)] \end{aligned}$$

Luego usando (50) para $k = n$ y premultiplicando por $\mathbf{F}^{-1}(n + 1, n) = \mathbf{F}(n, n + 1)$, se obtiene

$$\begin{aligned} E[\mathbf{x}(n)\boldsymbol{\alpha}^H(n)] &= \mathbf{F}(n, n + 1)E[\mathbf{x}(n + 1)\boldsymbol{\alpha}^H(n)] \\ &= \mathbf{F}(n, n + 1)\mathbf{G}(n)\mathbf{R}(n) \end{aligned}$$

de donde $E[\boldsymbol{\epsilon}(n, n - 1)\boldsymbol{\alpha}^H(n)] = \mathbf{F}(n, n + 1)\mathbf{G}(n)\mathbf{R}(n)$

- Además teniendo en cuenta que

$$\mathbf{G}(n)\mathbf{R}(n) = \mathbf{F}(n + 1, n)\mathbf{K}(n, n - 1)\mathbf{C}^H(n)$$

es posible finalmente obtener

$$E[\boldsymbol{\epsilon}(n)\boldsymbol{\epsilon}^H(n)] = \mathbf{K}(n, n - 1) - \mathbf{K}(n, n - 1)\mathbf{C}^H(n)\mathbf{G}^H(n)\mathbf{F}^H(n, n + 1)$$

o en forma equivalente

$$E[\boldsymbol{\epsilon}(n)\boldsymbol{\epsilon}^H(n)] = \mathbf{K}(n, n - 1) - \mathbf{F}(n, n + 1)\mathbf{C}(n)\mathbf{G}(n)\mathbf{K}(n, n - 1)$$

de donde se verifica, comparando con (54), que $E[\boldsymbol{\epsilon}(n)\boldsymbol{\epsilon}^H(n)] = \mathbf{K}(n)$.

5.6 Algoritmo completo

- Para operar adecuadamente los algoritmos de predicción de un paso y de filtrado descritos en las secciones anteriores es necesario obviamente especificar **condiciones iniciales**.
- El estado inicial de la ecuación del proceso no es conocido en general en forma exacta, de forma que es posible caracterizarlo por su media y correlación como

$$\hat{\mathbf{x}}(1|\mathbf{Y}_0) = E[\mathbf{x}(1)]$$

$$\mathbf{K}(1,0) = E[(\mathbf{x}(1) - E[\mathbf{x}(1)])(\mathbf{x}(1) - E[\mathbf{x}(1)])^H]$$

- La tabla siguiente presenta un resumen del filtro de Kalman para la predicción de un paso

Vector de entrada:

$$\text{Observaciones} = \{\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(n)\}$$

Parámetros conocidos:

$$\text{Matriz de transición de estados} = \mathbf{F}(n+1, n)$$

$$\text{Matriz de medición} = \mathbf{C}(n)$$

$$\text{Matriz de correlación del ruido de proceso} = \mathbf{Q}_1(n)$$

$$\text{Matriz de correlación del ruido de medición} = \mathbf{Q}_2(n)$$

Condiciones iniciales:

$$\hat{\mathbf{x}}(1|\mathbf{Y}_0) = E[\mathbf{x}(1)]$$

$$\mathbf{K}(1,0) = E[(\mathbf{x}(1) - E[\mathbf{x}(1)])(\mathbf{x}(1) - E[\mathbf{x}(1)])^H] = \Pi_0$$

Calcular: Para $n = 1, 2, 3, \dots$

$$\mathbf{G}(n) = \mathbf{F}(n+1, n)\mathbf{K}(n, n-1)\mathbf{C}^H(n)[\mathbf{C}(n)\mathbf{K}(n, n-1)\mathbf{C}^H(n) + \mathbf{Q}_2(n)]^{-1}$$

$$\boldsymbol{\alpha}(n) = \mathbf{y}(n) - \mathbf{C}(n)\hat{\mathbf{x}}(n|\mathbf{Y}_{n-1})$$

$$\hat{\mathbf{x}}(n+1|\mathbf{Y}_n) = \mathbf{F}(n+1, n)\hat{\mathbf{x}}(n|\mathbf{Y}_{n-1}) + \mathbf{G}(n)\boldsymbol{\alpha}(n)$$

$$\mathbf{K}(n) = \mathbf{K}(n, n-1) - \mathbf{F}(n, n+1)\mathbf{G}(n)\mathbf{C}(n)\mathbf{K}(n, n-1)$$

$$\mathbf{K}(n+1, n) = \mathbf{F}(n+1, n)\mathbf{K}(n)\mathbf{F}^H(n+1, n) + \mathbf{Q}_1(n)$$

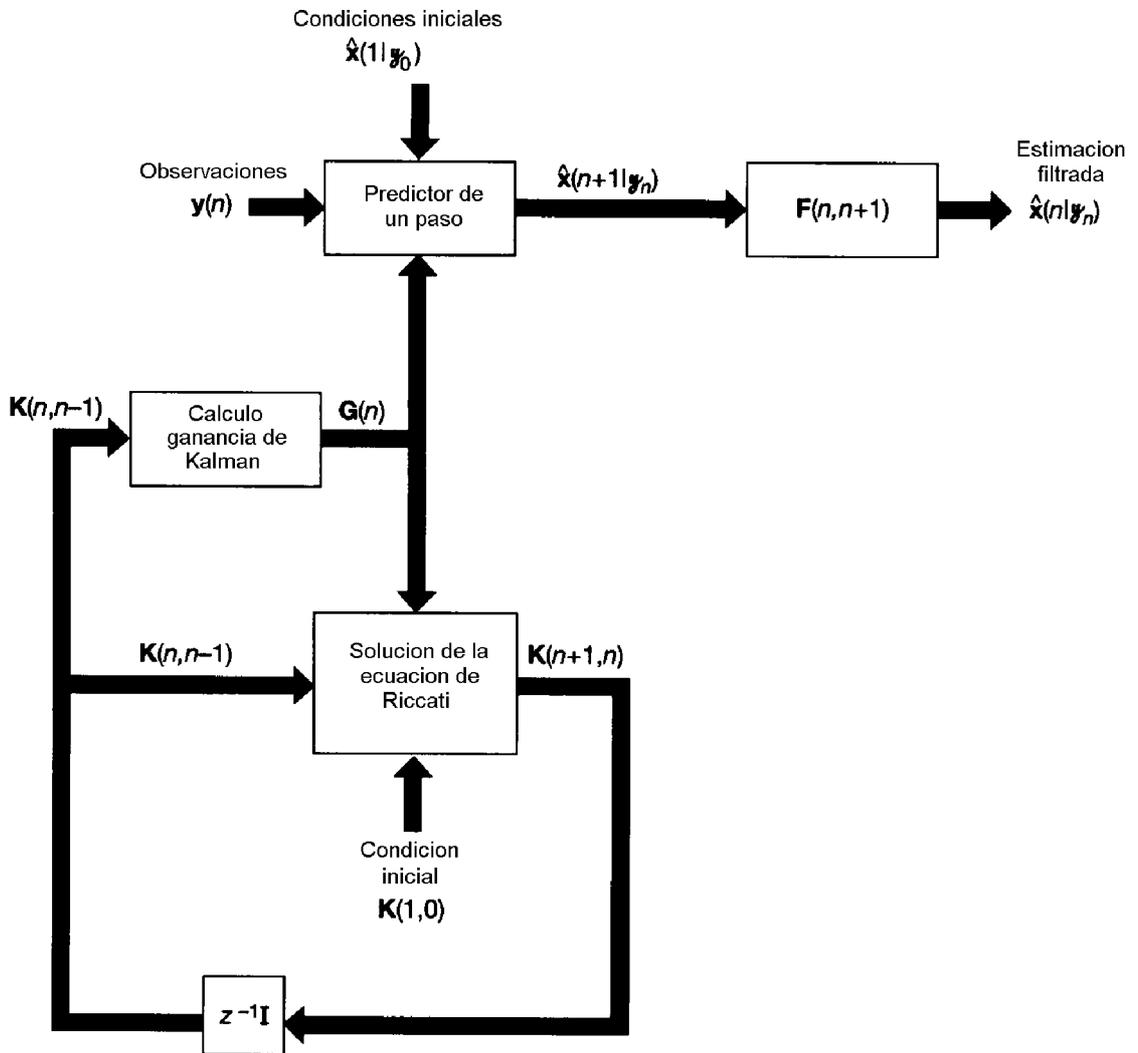


Figura 16: Diagrama en bloques del filtro de Kalman en base a predicción de un paso.

5.7 Variantes del filtro de Kalman

La principal razón para introducir el filtro de Kalman es la de proveer un marco general para el estudio de ciertos algoritmos de filtrado adaptivo conocidos colectivamente como **algoritmos de Cuadrados Mínimos Recursivos** (RLS).

5.7.1 Caso especial: sin ruido de proceso

- Consideremos un sistema dinámico lineal

$$\begin{aligned}\mathbf{x}(n+1) &= \lambda^{-1/2} \mathbf{x}(n) \\ y(n) &= \mathbf{u}^H(n) \mathbf{x}(n) + \nu(n)\end{aligned}$$

donde λ es un escalar positivo, el ruido de proceso ($\mathbf{v}_2(n)$) es cero, y el ruido de medición $\nu(n)$ es un proceso escalar, de media cero y varianza unitaria, o sea

$$E[\nu^2(n)] = \begin{cases} 1, & n = k \\ 0, & n \neq k \end{cases}$$

- Comparandolo con el modelo general

$$\begin{aligned}\mathbf{F}(n+1, n) &= \lambda^{-1/2} \mathbf{I} & \mathbf{Q}_1(n) &= \mathbf{0} \\ \mathbf{C}(n) &= \mathbf{u}^H(n) & \mathbf{Q}_1(n) &= \mathbf{0}\end{aligned}$$

- El sistema así descrito se denomina **modelo dinámico no forzado** porque la ecuación del proceso no tiene una excitación externa.
- Este modelo es la clave para analizar la familia de algoritmos RLS. En ese caso la constante λ tiene un papel importante en el funcionamiento de esos algoritmos.

5.7.2 Algoritmo de filtrado tipo covarianza

- El algoritmo general de la tabla anterior está diseñado para utilizar recursivamente la matriz de correlación $\mathbf{K}(n+1, n)$ asociada al error en la estimación del estado $\hat{\mathbf{x}}(n+1|\mathbf{Y}_n)$.
- Este algoritmo se denomina **de filtrado Kalman tipo covarianza**.
- En la tabla siguiente se resume el algoritmo de filtrado Kalman tipo covarianza para el modelo dinámico no forzado.

Entrada escalar:

Observaciones = $y(1), y(2), \dots, y(n)$

Parámetros conocidos:

Matriz de transición de estados = $\lambda^{-1/2}\mathbf{I}$

Matriz de medición = $\mathbf{u}^H(n)$

Matriz de correlación del ruido de medición $\nu(n) = 1$

Condiciones iniciales:

$\hat{\mathbf{x}}(1|\mathbf{Y}_0) = E[\mathbf{x}(1)]$

$\mathbf{K}(1, 0) = E[(\mathbf{x}(1) - E[\mathbf{x}(1)])(\mathbf{x}(1) - E[\mathbf{x}(1)])^H] = \mathbf{\Pi}_0$

Calcular: Para $n = 1, 2, 3, \dots$

$\mathbf{g}(n) = \frac{\lambda^{-1/2}\mathbf{K}(n-1)\mathbf{u}(n)}{\mathbf{u}^H(n)\mathbf{K}(n-1)\mathbf{u}(n)+1}$

$\alpha(n) = y(n) - \mathbf{u}^H(n)\hat{\mathbf{x}}(n|Y_{n-1})$

$\hat{\mathbf{x}}(n+1|Y_n) = \lambda^{-1/2}\hat{\mathbf{x}}(n|Y_{n-1}) + \mathbf{g}(n)\alpha(n)$

$\mathbf{K}(n) = \lambda^{-1}\mathbf{K}(n-1) - \lambda^{-1/2}\mathbf{g}(n)\mathbf{u}^H(n)\mathbf{K}(n-1)$

5.7.3 Algoritmo de filtrado tipo información

- El filtro Kalman también puede implementarse utilizando recursivamente la matriz $\mathbf{K}^{-1}(n)$. Esta matriz está relacionada con la **matriz de Información de Fisher**, la cual permite una interpretación del desempeño del filtro en términos de Teoría de Estimación (en particular asociada a la varianza mínima de un estimador o Límite de Cramer-Rao).
- Por ello este tipo de filtro Kalman se denomina de **tipo información**.

- Para obtener el algoritmo tipo información para el modelo dinámico no forzado seguiremos los siguientes pasos

1. A partir de la ecuación de Riccati (última línea del algoritmo anterior) y de la ecuación de ganancia $\mathbf{g}(n)$ se obtiene

$$\mathbf{K}^{-1}(n) = \lambda \mathbf{K}^{-1}(n-1) + \lambda \mathbf{u}(n) \mathbf{u}^H(n)$$

2. De la primera, segunda y tercera líneas del algoritmo anterior se obtiene

$$\mathbf{K}^{-1}(n) \hat{\mathbf{x}}(n+1|Y_n) = \lambda^{1/2} [\mathbf{K}^{-1}(n-1) \hat{\mathbf{x}}(n|Y_{n-1}) + \mathbf{u}(n)y(n)]$$

3. Combinando el resultado de los pasos anteriores

$$\begin{aligned} \hat{\mathbf{x}}(n+1|Y_n) &= \mathbf{K}(n) (\mathbf{K}^{-1}(n) \hat{\mathbf{x}}(n+1|Y_n)) \\ &= [\mathbf{K}^{-1}(n)]^{-1} (\mathbf{K}^{-1}(n) \hat{\mathbf{x}}(n+1|Y_n)) \end{aligned}$$

- A pesar que los algoritmos tipo covarianza e información son algebraicamente equivalentes (tienen igual complejidad computacional $O(M^2)$), sus propiedades numéricas difieren sustancialmente. La tabla siguiente describe el algoritmo de filtrado Kalman para el modelo dinámico no forzado

Entrada escalar:

$$\text{Observaciones} = y(1), y(2), \dots, y(n)$$

Parámetros conocidos:

$$\text{Matriz de transición de estados} = \lambda^{-1/2} \mathbf{I}$$

$$\text{Matriz de medición} = \mathbf{u}^H(n)$$

$$\text{Matriz de correlación del ruido de medición } \nu(n) = 1$$

Condiciones iniciales:

$$\hat{\mathbf{x}}(1|Y_0) = E[\mathbf{x}(1)]$$

$$\mathbf{K}(1,0) = E[(\mathbf{x}(1) - E[\mathbf{x}(1)])(\mathbf{x}(1) - E[\mathbf{x}(1)])^H] = \mathbf{\Pi}_0$$

Calcular: Para $n = 1, 2, 3, \dots$

$$\mathbf{K}^{-1}(n) = \lambda \mathbf{K}^{-1}(n-1) + \lambda \mathbf{u}(n) \mathbf{u}^H(n)$$

$$\mathbf{K}^{-1}(n) \hat{\mathbf{x}}(n+1|Y_n) = \lambda^{1/2} [\mathbf{K}^{-1}(n-1) \hat{\mathbf{x}}(n|Y_{n-1}) + \mathbf{u}(n)y(n)]$$

$$\hat{\mathbf{x}}(n+1|Y_n) = [\mathbf{K}^{-1}(n)]^{-1} (\mathbf{K}^{-1}(n) \hat{\mathbf{x}}(n+1|Y_n))$$

5.7.4 Factorización de $\mathbf{K}(n)$

- El algoritmo de filtrado Kalman tipo covarianza sufre de problemas numéricos graves. En particular, dado que $\mathbf{K}(n)$ está definida por la diferencia entre dos matrices no negativas solo una gran precisión numérica evitará que en alguna iteración $\mathbf{K}(n)$ sea indefinida. Esto es lógicamente inaceptable porque esta matriz representa una matriz correlación.
- Estos problemas de condicionamiento numérico con aritmética de longitud finita se denominan **fenómeno de divergencia**.
- Estos problemas pueden evitarse trabajando con una factorización adecuada de la matriz $\mathbf{K}(n)$. En particular utilizando recursivamente la matriz $\mathbf{K}(n)$ a partir de la **factorización de Cholesky**

$$\mathbf{K}(n) = \mathbf{K}^{1/2}(n)\mathbf{K}^{T/2}(n)$$

donde $\mathbf{K}^{1/2}(n)$ es una matriz triangular inferior. Esta factorización se denomina de **raíz cuadrada** y es mucho menos propensa a problemas numéricos.

- Para el caso del algoritmo tipo información puede utilizarse recursivamente una factorización de $\mathbf{K}^{-1}(n)$ de la forma

$$\mathbf{K}^{-1}(n) = \mathbf{K}^{-1/2}(n)\mathbf{K}^{-T/2}(n)$$