



UNIVERSITAT^{DE}
BARCELONA

Treball final de grau

GRAU DE MATEMÀTIQUES

Facultat de Matemàtiques i Informàtica
Universitat de Barcelona

Filtro de Kalman y sus
aplicaciones

Maria Cristina Munuera Raga

Director: Dr. José Manuel Corcuera Valverde
Realitzat a: Departament
Probabilitat i Estadística

Barcelona, 27 de junio de 2018

Abstract

Filtering problems concern the estimation of noise-interfered variables. The Kalman filter provides the linear Mean Least Squares estimation, combining the evolution of the variable with the observations obtained. Moreover, it is a recursive process, and consequently the storage of information is needed at all times. The main applications are found in a wide range of fields such as finance, wireless communications or GPS systems.

Resumen

Los problemas de filtrado se ocupan de la estimación de una variable interferida por el ruido. El filtro de Kalman proporciona el estimador lineal con menor error cuadrático medio, combinando la evolución de la variable con observaciones recogidas. Es además un proceso recursivo, por lo que no es necesario almacenar la información del estado en todos los instantes. Las principales aplicaciones las encontramos en el campo de las finanzas, en el mundo de las comunicaciones inalámbricas o en la localización GPS.

Agradecimientos

Mi primer agradecimiento a mi familia y amigos por ser el mayor apoyo en todos mis años de estudiante.

Agradezco sinceramente al Dr. José Manuel Corcuera Valverde su dedicación y disponibilidad ante todas mis dudas, además de su paciencia y proactividad.

Agradecer también a todos los profesores que me han acompañado en cada año de mi vida académica y me han formado profesional y personalmente.

Índice

1. Introducción	1
2. El Filtro de Kalman: el algoritmo	3
2.1. Introducción al algoritmo de Kalman	3
2.2. El algoritmo	4
2.2.1. Matriz de varianzas-covarianzas del error	5
2.2.2. Ganancia de Kalman	6
2.2.3. Ecuaciones del algoritmo	8
2.3. Filtro de Kalman: el estimador lineal óptimo	8
2.4. Ecuaciones de Kalman: recursividad	9
3. Aplicaciones Filtro de Kalman	13
3.1. Tracking	13
3.2. Estimación del canal en sistemas de comunicación inalámbricos	20
3.3. Filtro adaptativo. RLS: algoritmo recursivo de error cuadrático mínimo	23
3.4. Estimación del PIB mensual	24
3.4.1. Rasgos generales del modelo	25
3.4.2. Estimación máxima verosimilitud	26
3.4.3. Modelo autorregresivo sin series relacionadas	27
3.4.4. Modelo autorregresivo con series relacionadas	29
4. Extensiones del filtro de Kalman	31
4.1. Filtro de Kalman extendido	32
4.2. Filtro de Kalman Unscented	33
4.2.1. Transformación unscented	33
4.2.2. Filtro de Kalman Unscented	35
4.3. Filtro de partículas	36
4.3.1. Método de Montecarlo	36
4.3.2. Filtro de partículas	36
5. ANEXO	39

1. Introducción

El filtro de Kalman se sitúa dentro de los problemas de filtrado, que quedan englobados en la teoría de procesos estocásticos. A qué nos referimos cuando decimos que nos adentramos en los modelos estocásticos? La necesidad de ir más allá de los sistemas deterministas -aquellos en los que el azar no interviene- se hace patente ante sus deficiencias, las cuales vamos a intentar evidenciar brevemente.

Dado un fenómeno físico, lo intentamos modelizar a partir de sus comportamientos, y lo haremos utilizando leyes físicas conocidas u observaciones recogidas, las cuales relacionaremos entre sí para obtener una salida del sistema. Sin embargo, un sistema determinista no es suficiente para llevar a cabo este análisis. Las razones son diversas. Por un lado, un modelo matemático acaba por recoger solamente aquellas características dominantes, por lo que muchos efectos quedan fuera del modelo. Por otro lado, los sistemas dinámicos no quedan definidos simplemente por los efectos que recogemos, sino que también existen ruidos que no podemos modelar de forma determinista. Por ejemplo, en el lanzamiento de una pelota hay factores como la velocidad del viento que no somos capaces de controlar. Esto formaría parte de la aleatoriedad que un sistema determinista no recoge. Por último, otra de las deficiencias es que los aparatos de medida no proporcionan datos exactos sobre el sistema, sino que incorporan también sus propias dinámicas y distorsiones. En resumen, no podemos asumir un conocimiento perfecto de todas las cantidades necesarias para describir un sistema completo, y es precisamente eso lo que necesitaríamos para poder conformarnos con un modelo determinista.

Ante estos nos surgen dudas: como recoger en un modelo las incertidumbres descritas? Como podemos optimizar un modelo a pesar de que las observaciones recogidas sean ruidosas o insuficientes? Para responder a esto nos adentramos en los modelos estocásticos, y más concretamente en los problemas de filtrado, de los cuales nos ocuparemos en nuestro trabajo.

El problema de filtrado desarrolla un algoritmo para estimar una cierta variable (o vector de variables) que se ve interferida por ruidos y errores, los cuales queremos aislar para obtener un estimador lo más limpio posible. El gran avance en este ámbito llegó en 1960 de la mano de Rudolf Kalman en su publicación «A New Approach to Linear Filtering and Prediction Problems». El filtro de Kalman es un algoritmo que estima una variable a partir de datos medidos. Lo hace siguiendo dos pasos: por un lado la predicción del estado del sistema y por otro la incorporación de las observaciones recogidas una vez corregidas. Así pues, el objetivo es obtener un estimador óptimo, -en términos del Error Cuadrático Medio (ECM)-, en base a la dinámica del sistema y a las observaciones ruidosas. En definitiva, el filtro de Kalman procesa todas las medidas disponibles, independientemente de su precisión. Lo hará con el objetivo de estimar el valor de las variables de interés, basándose en el conocimiento del sistema y de las observaciones, junto con la descripción de su ruido y errores.

Una de sus mayores ventajas es que se trata de un proceso recursivo. Esto nos permite ir incorporando las nuevas observaciones sin tener que reformular todo el algoritmo.

El filtro de Kalman tiene un amplio abanico de posibilidades para su aplicación. Es ampliamente utilizado en econometría: en la evaluación de los riesgos, en la previsión del crecimiento del Producto Interior Bruto (PIB) de un país, en el estudio de los efectos de la inflación, en el estudio de la evolución del precio de la vivienda, o en la variabilidad del precio del petróleo. Pero es también utilizado en muchos otros campos, como en la biología, para el estudio de las especies en extinción, o en el mundo de las comunicaciones inalámbricas, como veremos más adelante en un par de ejemplos. Pero una de las principales aplicaciones, y que también estudiaremos en mayor detalle, está en el tracking: el rastreo de la localización de algún elemento; por ello son amplias sus aplicaciones en la localización por señal GPS.

Nuestro trabajo quedará estructurado de la siguiente manera. En la sección 2 deduciremos el algoritmo de Kalman: definiremos los conceptos básicos y vemos como se deducen las ecuaciones de actualización del comportamiento del modelo y del filtrado del ruido. Veremos también como la propia construcción del algoritmo nos garantiza el carácter óptimo del estimador. En la sección 3 aplicaremos todos los conceptos ya presentados a tres de los campos principales: el tracking, el sistema de comunicaciones y el conocimiento del PIB de un país. En la sección 4 nos iniciaremos en una extensión del filtro de Kalman, la cual nos permitirá trabajar con sistemas no lineales.

2. El Filtro de Kalman: el algoritmo

2.1. Introducción al algoritmo de Kalman

Introduzcamos los fundamentos del algoritmo del filtro de Kalman. El punto de inicio es una ecuación de transición del estado, que describirá la evolución de la variable a lo largo del tiempo. Juntamente con esto, tendremos una ecuación de observaciones, que nos describirá la relación entre las observaciones medidas con la variable de estado. El filtro de Kalman busca minimizar el Error Cuadrático Medio (ECM), y para ello el objetivo será encontrar el valor de la ganancia que nos garantiza esta propiedad. Así pues, el filtro de Kalman acabará por darnos, en el instante k , un estimador lineal que será combinación de los estimadores hasta el instante $k - 1$ y de la nueva observación del instante k .

Como decimos, el filtro de Kalman nos dará un estimador que minimizará el error cuadrático medio a partir de unas observaciones. Consideramos un vector aleatorio X de dimensión 1. Siendo $\hat{X}_{k|j}$ el estimador de X en el instante k a partir de las observaciones tomadas hasta el instante j , Z_1, \dots, Z_j , con $j \leq k$, tenemos:

$$\hat{X}_{k|j} = \min_{\hat{X}_{k|j} \in R^n} \text{traza} \{ E \left[(X_k - \hat{X}_{k|j})(X_k - \hat{X}_{k|j})^t \right] \}$$

donde, recordemos, que Z^j se refiere al conjunto de observaciones y t indica transposición. Como veremos más adelante, el estimador que obtenemos no es más que el valor esperado del estado en el tiempo k , condicionado por las observaciones tomadas hasta el tiempo j . De manera que:

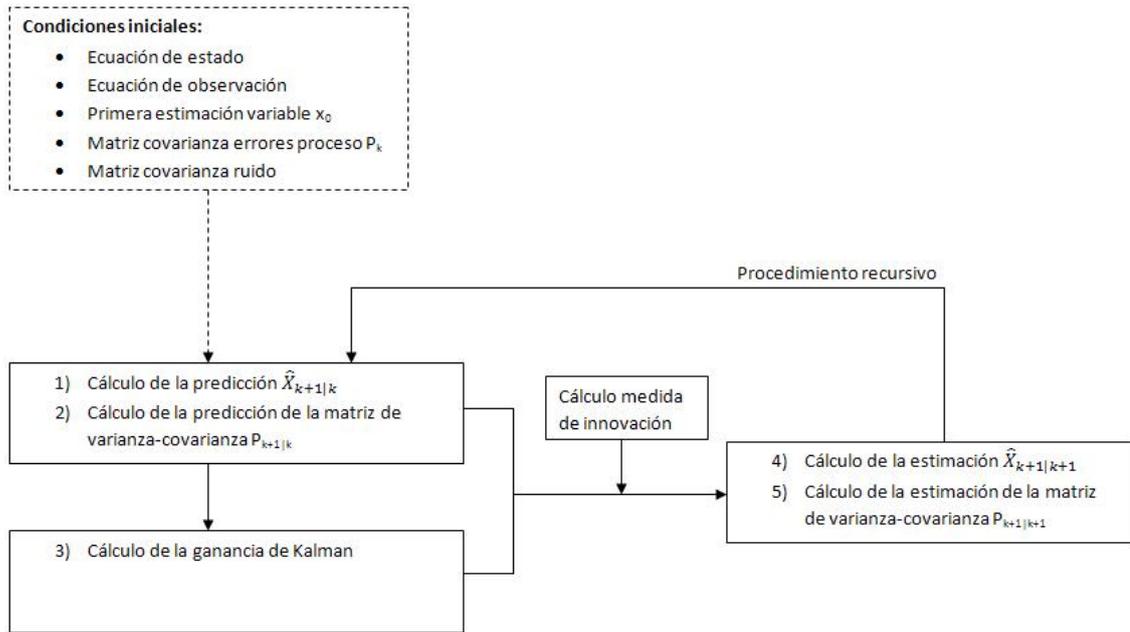
$$\hat{X}_{k|j} = E [X_k | Z^j]$$

Definimos la matriz de varianzas-covarianzas del error:

$$P_{k|j} = E \left[(X_k - \hat{X}_{k|j})(X_k - \hat{X}_{k|j})^t \right]$$

A partir de ahora el estimador $\hat{X}_{k|k-1}$, que hace referencia a la estimación de la variable X en el instante k a partir de las observaciones hasta el instante $k-1$, será lo que conoceremos como predicción, pues aún no hemos usado la observación en el instante k .

El algoritmo de Kalman queda resumido en este diagrama:



2.2. El algoritmo

Profundicemos en el desarrollo del algoritmo de Kalman y sus formulaciones. Recordemos que si escribimos, por ejemplo, $\hat{X}_{k|k-1}$ estaremos considerando la predicción de la variable o vector de variables X en el instante k y utilizando las variables observadas hasta el instante $k-1$. Lo mismo sucederá con la notación de la matriz de varianzas-covarianzas $P_{k|k-1}$.

Definimos la ecuación de transición del modelo, que recoge la evolución del vector de variables X :

$$\mathbb{R}^l \ni X_k = F_k X_{k-1} + B_k U_k + W_k$$

donde $F \in \mathbb{M}_{l \times l}$ es la matriz de transición y $U \in \mathbb{M}_{l \times l}$ es un control basado en la información anterior, es decir, $U_k = h(Z_1, \dots, Z_{k-1})$, donde Z_i son las observaciones.

La ecuación de observación será:

$$\mathbb{R}^m \ni Z_k = H_k X_k + V_k$$

donde la matriz $H_k \in \mathbb{M}_{m \times l}$ relaciona la observación y el vector de estado sin tener en cuenta el ruido, que viene determinado por $V_k \in \mathbb{R}^m$. A partir de ahora supondremos $l = m$.

Establezcamos condiciones sobre los vectores aleatorios. Consideramos que los ruidos $V_k, W_k \in \mathbb{R}^l$ son conjuntamente gaussianos, sin correlación entre ellos y de media cero. Es decir:

$$E(W_k) = E(V_k) = 0$$

y la matriz de varianzas-covarianzas correspondiente es:

$$E(W_i W_j^t) = \delta_{ij} Q_i, \quad E(V_i V_j^t) = \delta_{ij} R_i,$$

donde R_i y Q_i son matrices $l \times l$ no singulares.

También asumimos la no correlación de los errores, esto es:

$$E(W_i V_j^t) \equiv 0 \quad \forall i, j$$

Definamos formalmente dos conceptos. Por un lado definimos la matriz de varianzas-covarianzas del *error a posteriori* -se refiere al error de estimación- $P_{k|k} \in \mathbb{M}_{l \times l}$, esto es, $cov(\hat{X}_{k|k} - X_k, \hat{X}_{k|k} - X_k)$. Este concepto es fundamental en todo el desarrollo, pues Kalman tiene la propiedad de que se construye minimizando la traza de esta matriz. Recordamos que minimizar el ECM: $E(\|X_k - \hat{X}_{k|k}\|^2)$ es equivalente a minimizar la traza de la matriz P .

Por otro lado definimos *el error de predicción o medida de innovación*, esto es $\mathbb{R}^l \ni \delta_k = Z_k - H \hat{X}_{k|k-1}$. Más adelante este concepto será desarrollado con mayor detalle.

2.2.1. Matriz de varianzas-covarianzas del error

Recordamos que el fundamento del filtro de Kalman es la combinación entre la predicción por el modelo teórico y las observaciones. Así pues, lo que buscamos es una nueva estimación \hat{X}'_k que sea combinación lineal de la predicción $\hat{X}_{k|k-1}$ y de la observación Z_k . Esto es:

$$\hat{X}'_k = \hat{X}_{k|k-1} + K_k (Z_k - H_k \hat{X}_{k|k-1}) = \hat{X}_{k|k} \quad (2.1)$$

donde $Z_k - H_k \cdot \hat{X}_{k|k-1}$ es la medida de innovación, es decir, la discrepancia entre la observación y *lo que se esperaba*. Por otro lado, el coeficiente K será lo que conoceremos como ganancia de Kalman. Será el responsable de la corrección, y minimizará la traza de la covarianza del error a posteriori. En consecuencia será el mejor estimador entre los que dependen linealmente de las observaciones. Por ello, usaremos técnicas de derivación para encontrar el valor de K .

Añadimos a la ecuación (2.1) la relación entre observación (Z_k) y estado interno (X_k):

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k (H_k X_k + V_k - H_k \hat{X}_{k|k-1}) \quad (2.2)$$

Dado que K minimizará la covarianza de la traza del error a posteriori, debemos estudiar la matriz $P_{k|k}$, que habíamos definido ya previamente. Empezaremos, pues, estudiando el error $X_k - \hat{X}_{k|k}$. Por lo tanto, usando (2.2)

$$X_k - \hat{X}_{k|k} = X_k - \hat{X}_{k|k-1} - K_k H_k X_k - K_k V_k + K_k H_k \hat{X}_{k|k-1}$$

$$X_k - \hat{X}_{k|k} = X_k (1 - K_k H_k) - \hat{X}_{k|k-1} (1 - K_k H_k) - K_k (V_k)$$

$$X_k - \hat{X}_{k|k} = \left(X_k - \hat{X}_{k|k-1} \right) (1 - K_k H_k) - K_k V_k$$

Dado que no hay relación entre las variables y el vector del ruido, podemos continuar:

$$\text{cov} \left(X_k - \hat{X}_{k|k} \right) = \text{cov} \left(\left(X_k - \hat{X}_{k|k-1} \right) (1 - K_k H_k) \right) + \text{cov} (K_k V_k)$$

Aplicando la propiedad de la covarianza: $\text{cov} (AX + a) = A \text{cov} (X) A^t$:

$$\begin{cases} \text{cov} \left(\left(X_k - \hat{X}_{k|k-1} \right) (1 - KH) \right) = (1 - KH) \text{cov} \left(X_k - \hat{X}_{k|k-1} \right) (1 - KH)^t \\ \text{cov} (K \cdot V_k) = KRK^t \end{cases}$$

donde $R_k = \text{cov}(V_k)$

En definitiva, con esto ya hemos formulado, de una forma recursiva, la matriz de covarianzas:

$$P_{k|k} = (1 - K_k H_k) P_{k|k-1} (1 - K_k H_k)^t + K_k R_k K_k^t \quad (2.3)$$

Por otro lado, a partir de la definición de la matriz de varianzas-covarianzas como ECM, podemos deducir la ecuación de predicción de esta matriz. Veamos:

$$P_{k|k-1} = E[(X_k - \hat{X}_{k|k-1})(X_k - \hat{X}_{k|k-1})^t]$$

A partir de la ecuación de la predicción tenemos que:

$$P_{k|k-1} = E[(X_k - F_k \hat{X}_{k-1|k-1} - B_k U_k)(X_k - F_k \hat{X}_{k-1|k-1} - B_k U_k)^t]$$

Incorporando la definición del estado interno X_k :

$$P_{k|k-1} = E[(F_k X_{k-1} + w_k - F_k \hat{X}_{k-1|k-1})(F_k X_{k-1} + w_k - F_k \hat{X}_{k-1|k-1})^t]$$

Por lo tanto:

$$P_{k|k-1} = F_k E[(X_{k-1} - \hat{X}_{k-1|k-1})(X_{k-1} - \hat{X}_{k-1|k-1})^t] F_k^T + E[w_k w_k^t]$$

En definitiva, simplificando:

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^t + Q_k$$

2.2.2. Ganancia de Kalman

Recordemos que nuestro objetivo era encontrar la expresión de la ganancia de Kalman K . Queremos un algoritmo óptimo, entendiendo por óptimo aquel que minimiza el error cuadrático medio. Tal y como hemos dicho, el ECM nos mide la diferencia entre el valor real y el valor estimado, y lo hace considerando el promedio

del error al cuadrado, es decir, $E(e_k^2)$, donde e_k es el vector error. Así pues, queremos un coeficiente K que nos minimice la traza de la matriz de covarianzas. Si tenemos el vector $e_k := X_k - \hat{X}_{k|k} = \left((X_k - \hat{X}_{k|k})^1, \dots, (X_k - \hat{X}_{k|k})^l \right)$, entonces:

$$P_{k|k} = \begin{bmatrix} cov((X_k - \hat{X}_{k|k})^1, (X_k - \hat{X}_{k|k})^1) & \dots & cov((X_k - \hat{X}_{k|k})^1, (X_k - \hat{X}_{k|k})^l) \\ \vdots & \vdots & \vdots \\ cov((X_k - \hat{X}_{k|k})^i, (X_k - \hat{X}_{k|k})^1) & cov((X_k - \hat{X}_{k|k})^i, (X_k - \hat{X}_{k|k})^i) & \dots \\ \vdots & \vdots & \vdots \\ cov((X_k - \hat{X}_{k|k})^l, (X_k - \hat{X}_{k|k})^1) & \dots & cov((X_k - \hat{X}_{k|k})^l, (X_k - \hat{X}_{k|k})^l) \end{bmatrix}$$

Recordamos que a partir de la derivación lograremos encontrar el valor K que nos minimizará la traza: igualaremos la derivada a zero y encontraremos la fórmula de K .

Ahora utilizaremos que la traza de la suma de matrices es la suma de trazas y que la traza de una matriz es invariante por transposición. Así, a partir de la definición de $P_{k|k}$ tendremos:

$$\begin{aligned} tr P_{k|k} &= tr((1 - K_k H_k) P_{k|k-1} (1 - K_k H_k)^t + K_k R_k K_k^t) \\ tr P_{k|k} &= tr((P_{k|k-1} - K_k H_k P_{k|k-1})(1 - H_k^t K_k^t) + K_k R_k K_k^t) \\ tr P_{k|k} &= tr(P_{k|k-1} - K_k H_k P_{k|k-1} - P_{k|k-1} H_k^t K_k^t + K_k H_k P_{k|k-1} H_k^t K_k^t + K_k R_k K_k^t) \\ tr P_{k|k} &= tr(P_{k|k-1}) - tr(K_k H_k P_{k|k-1}) - tr(P_{k|k-1} H_k^t K_k^t) + tr(K_k (H_k P_{k|k-1} H_k^t + R_k) K_k^t) \end{aligned}$$

Ahora derivamos la traza respecto K :

$$\frac{\partial tr(P_{k|k})}{\partial K_k} = -2(H_k P_{k|k-1})^t + 2K_k (H_k P_{k|k-1} H_k^t + R_k)$$

Igualamos a zero:

$$\begin{aligned} -2(H_k P_{k|k-1})^t + 2K_k (H_k P_{k|k-1} H_k^t + R_k) &= 0 \\ (H_k P_{k|k-1})^t &= K_k H_k P_{k|k-1} H_k^t + R_k \\ K_k &= (H_k P_{k|k-1})^t (H_k P_{k|k-1} H_k^t + R_k)^{-1} \end{aligned}$$

Y así hemos llegado a la fórmula de la ganancia de Kalman.

Podemos simplificar aún más está fórmula definiendo la matriz $S_k = H_k P_{k|k-1} H_k^t + R_k$. Con esta notación y la simetría de una matriz de covarianzas, nos queda la ganancia de Kalman como:

$$K_k = P_{k|k-1} H_k^t S_k^{-1} \quad (2.4)$$

Cabe destacar que la ganancia de Kalman hará que nos pondere más la observación o el estimador según sea la variable con menor error. Debido a esto la ganancia de Kalman será un valor entre 0 y 1 que multiplicará la diferencia entre la observación y el estimador. De hecho, tal y como hemos visto, si EST_k se refiere al estimador en el tiempo k , OBS se refiere a la observación y $PRED_k$ a la predicción en tiempo k con observaciones hasta $k-1$, tenemos:

$$EST_k = PRED_k + K_k [OBS - H_k \cdot PRED_k]$$

Anteriormente, en (2.3) habíamos encontrado una expresión recursiva para la actualización de la matriz de covarianzas. Ahora, con la fórmula de K , podemos simplificar esta expresión.

$$\begin{aligned}
P_{k|k} &= (I - K_k H_k) P_{k|k-1} (I - K_k H_k)^t + K_k R_k K_k^t \\
&= P_{k|k-1} - K_k H_k P_{k|k-1} - P_{k|k-1} H_k^t K_k^t + K_k (H_k P_{k|k-1} H_k^t + R_k) K_k^t \\
&= P_{k|k-1} - K_k H_k P_{k|k-1} - P_{k|k-1} H_k^t K_k^t + K_k S_k K_k^t
\end{aligned}$$

A partir de la definición de K_k en (2.4) tenemos: $K_k S_k K_k^t = P_{k|k-1} H_k^t K_k^t$. Y por tanto:

$$\begin{aligned}
P_{k|k} &= P_{k|k-1} - K_k H_k P_{k|k-1} \\
&= (I - K_k H_k) P_{k|k-1}
\end{aligned} \tag{2.5}$$

2.2.3. Ecuaciones del algoritmo

En definitiva, las ecuaciones del filtro de Kalman las dividiremos entre predicción (tiempo) y corrección (observaciones). Estas son:

Ecuaciones de predicción:

$$\begin{cases} \hat{X}_{k|k-1} = F_k \hat{X}_{k-1|k-1} + B_k U_k \\ P_{k|k-1} = F_k P_{k-1|k-1} F_k^t + Q_k \end{cases}$$

Ecuaciones de corrección (actualización por observación):

$$\begin{cases} \hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k \cdot (Z_k - H_k \hat{X}_{k|k-1}) \\ P_{k|k} = (I - K_k H_k) \cdot P_{k|k-1} \end{cases}$$

donde K_k es la ganancia de Kalman definida anteriormente

2.3. Filtro de Kalman: el estimador lineal óptimo

Definamos un estimador lineal. Un estimador \hat{X} de X es lineal cuando se puede expresar como combinación lineal de las observaciones, es decir, $\hat{X} = \sum A(i) Z_i$, donde Z_i son las observaciones.

Entre los estimadores lineales, el filtro de Kalman es el que nos garantiza un menor error de covarianza, independientemente de la hipótesis de modelo gaussiano. Sin embargo, mientras que en el caso gaussiano podemos afirmar que el filtro de Kalman es el óptimo entre toda clase de estimadores, esto no es así cuando no trabajamos bajo hipótesis gaussianas. En ese caso solo podemos afirmar que se trata del mejor de los estimadores lineales, pues puede existir un mejor estimador no lineal -entendiendo por mejor que tenga menor varianza-.

En general, como veremos a continuación, se puede demostrar que la esperanza condicionada constituye el estimador de menor varianza, siendo, además, un estimador insesgado. Precisamente por ello, en el caso no gaussiano, no podemos asegurar el carácter óptimo de Kalman entre todos los estimadores, pues este no coincide con la esperanza condicionada. Esto es porque, tal y como se explica en el anexo (ver sección Esperanza condicionada), en el caso gaussiano la esperanza condicionada es lineal, y en cambio esta linealidad no es siempre cierta cuando estamos en el caso no gaussiano. Demostremos ahora que, efectivamente, la esperanza condicionada es el estimador que minimiza el error cuadrático medio.

Teorema. Sea X un vector de variables. El estimador \hat{X} que minimiza el ECM $E[(X - \hat{X})^2]$ es $\hat{X} = E[X|Z]$

Demostración. Para esta demostración usaremos propiedades de la esperanza condicionada que están demostradas en el anexo.

Consideramos el estimador \hat{X} función de Z : $\hat{X} = g(Z)$. Entonces:

$$\begin{aligned} E[(X - \hat{X})^2] &= E[(X - E[X|Z] + E[X|Z] - \hat{X})^2] = \\ &= E \left[(X - E[X|Z])^2 + 2(X - E[X|Z])(E[X|Z] - \hat{X}) + (E[X|Z] - \hat{X})^2 \right] = \\ &= E[(X - E[X|Z])^2] + 2E[(X - E[X|Z])(E[X|Z] - \hat{X})] + E[(E[X|Z] - \hat{X})^2] = \\ &= E[(X - E[X|Z])^2] + E[(E[X|Z] - \hat{X})^2] + 2E[E[(X - E[X|Z])(E[X|Z] - \hat{X})|Z]] = \end{aligned}$$

Se cumple que:

$$E[(X - E[X|Z])(E[X|Z] - \hat{X})|Z] = E[(E[X|Z] - \hat{X})E[(X - E[X|Z])|Z]]$$

y como

$$E[(X - E[X|Z])|Z] = E[X|Z] - E[X|Z] = 0$$

resulta que:

$$E[(X - \hat{X})^2] = E[(X - E[X|Z])^2] + E[(E[X|Z] - \hat{X})^2]$$

y dado que $E[(E[X|Z] - \hat{X})^2] \geq 0$ tenemos

$$E[(X - \hat{X})^2] \geq E[(X - E[X|Z])^2]$$

□

2.4. Ecuaciones de Kalman: recursividad

Otra forma de obtener las ecuaciones de Kalman es con un procedimiento recursivo. Para ello necesitamos asentar un concepto definido previamente: la medida de innovación.

Sean Z^k el conjunto de observaciones. Así, queremos prever la observación Z_k a partir de las $k-1$ anteriores. Esto es, queremos el estimador $\hat{Z}_{k|k-1}$.

$$\begin{aligned}\hat{Z}_{k|k-1} &= E[Z_k|Z^{k-1}] = \\ &= E[H_k X_k + V_k|Z^{k-1}] = H_k E[X_k|Z^{k-1}] + E[V_k|Z^{k-1}] = H_k \hat{X}_{k|k-1}\end{aligned}$$

La innovación será la diferencia entre la observación real Z_k y esta estimación. Así pues:

$$\delta_k = Z_k - H_k \hat{X}_{k|k-1} \in \mathbb{R}^l$$

Y por lo tanto la innovación nos servirá también para darnos una idea de la eficacia del estimador.

Queremos ver que la medida de innovación es un ruido blanco. Para ello primero veremos que se trata de un residuo de media cero.

$$E[\delta_k|Z^{k-1}] = E[Z_k - H_k \hat{X}_{k|k-1}|Z^{k-1}]$$

Por otro lado demostraremos que es un residuo no autocorrelacionado. Esto será equivalente a ver que $E[\delta_i^t \delta_j^t] = N_i \delta_{ij}$ para una cierta matriz $N \in \mathbb{M}_{l \times l}$ y donde δ_{ij} es la delta de Kronecker. Calculamos la matriz de varianzas-covarianzas de la innovación.

$$\begin{aligned}N_k &= E[\delta_k \delta_k^t] = E[(Z_k - H_k \hat{X}_{k|k-1})(Z_k - H_k \hat{X}_{k|k-1})^t] \\ &= E[(H_k X_k + v_k - H_k \hat{X}_{k|k-1})(H_k X_k + v_k - H_k \hat{X}_{k|k-1})^t] = \\ &= E[(H_k(X_k - \hat{X}_{k|k-1}) + v_k)(H_k(X_k - \hat{X}_{k|k-1}) + v_k)^t] = \\ &= R_k + H_k P_{k|k-1} H_k^t\end{aligned}$$

Ahora queda ver que si $i \neq j$, entonces $E[\delta_i \delta_j^t] = 0$. Lo haremos utilizando que su media es 0.

$$E[\delta_i \delta_j^t] = E[E[\delta_i \delta_j^t|Z^{k-1}]] = E[E[\delta_i|Z^{k-1}] \delta_j^t] = 0$$

Con todo esto hemos probado ya la propiedad más importante de la medida de innovación.

Demostremos algunos lemas que nos servirán para terminar la sección deduciendo las ecuaciones de Kalman de forma recursiva. No obstante, antes de empezar definiremos la probabilidad condicionada como proyección. Veamos.

Consideraremos el mismo modelo con el que hemos trabajado hasta ahora -salvo que ahora prescindimos del término de control- :

$$\begin{cases} X_{k+1} = F_k X_k + W_{k+1} \\ Z_k = H_k X_k + V_k \end{cases}$$

con R y Q matriz de varianzas-covarianzas de V y W respectivamente.

Hasta ahora entendemos $P(X|Z_0, \dots, Z_k)$ como la probabilidad de X sabiendo que son ciertas primeras k -ésimas observaciones. Si consideramos S el haz de observaciones Z_i con $i = 1, \dots, k$, entonces la componente j -ésima de $P(X|Z_0, \dots, Z_k)$, que notamos por $P(X|S)^j$ coincidirá con la proyección de la componente j -ésima de X , X^j , sobre S . Así, entenderemos $P(X|S)$ como una proyección. Para ello en esta sección notaremos:

$$P_k(X) := P(X|Z_0, \dots, Z_k), k = 0, 1, 2, \dots$$

Vayamos a los lemas

Lema 2.1. *Siguiendo la notación expuesta, el vector $P_k(X)$ cumple que $[X - P_k(X)] \perp Z_i$, con $i = 0, \dots, k$*

Demostración. Es consecuencia de la propia definición de $P_k(X)$ □

Lema 2.2. *Consideremos Y, X variables con momentos de segundo orden finitos y de dimensión q y r respectivamente. Entonces: $P(X|Y) = MY$, donde M es una matriz $q \times r$ con $M = E(XY^t)(E(YY^t))^{-1}$.*

Lema 2.3. *Los vectores Z_0, \dots, Z_k y $Z_0, \dots, Z_{k-1}, \delta_k$ generan el mismo espacio, es decir, contienen la misma información. De aquí se deduce que:*

$$P_k(X) = P(X|Z_0, \dots, Z_{k-1}, \delta_k) = P(X|Z_0, \dots, \delta_{k-1}) + P(X|\delta_k) = P_{k-1}(X) + P(X|\delta_k)$$

Consideremos X un vector de variables de dimensión m bajo el modelo definido en la página anterior. Las ecuaciones recursivas de predicción pero el estimador $\hat{X}_{k+1|k}$ se definen como

$$\hat{X}_{k+1|k} = F_k \hat{X}_{k|k-1} + \Theta_k \Delta_k^{-1} \delta_k$$

donde

$$\begin{aligned} \Theta_k &= F_k P_{k|k-1} H_k^T \\ \Delta_k &= H_k P_{k|k-1} H_k^T + R_k \end{aligned}$$

Demostración. por el lema 2.2 tenemos

$$\begin{aligned} \Theta_k &= E[X_{k+1} \delta_k^t] = E[(F_k X_k + W_{k+1})((X_k - X_{k|k-1})^t H_k^t + V_k^t)] = F_k P_{k|k-1} H_k^t \square \\ \Delta_k &= E[\delta_k \delta_k^t] = H_k P_{k|k-1} H_k^t + R_k \end{aligned}$$

Consideramos X un vector de variables de dimensión r . Los estimadores $X_{k|k}$ y las matrices de covarianza $P_{k|k}$ se pueden definir de forma recursiva y las ecuaciones correspondientes son:

$$\begin{cases} \hat{X}_{k|k} = X_{k|k-1} + P_{k|k-1} H_k^t \Delta_k^{-1} (Z_k - H_k \hat{X}_{k|k-1}) \\ P_{k|k} = P_{k|k-1} - P_{k|k-1} H_k^t \Delta_k^{-1} H_k P_{k|k-1} \end{cases}$$

Demostración. Por el lema 2.3 tenemos que $P_k(X_k) = P_{k-1}(X_k) + P(X_k|Z_k)$. Podemos aplicar el lema 2.2 y tendremos que

$$P_k(X_k) = P_{k-1}(X_k) + M\delta_k$$

donde $M = E(X_k\delta_k^t)(E(\delta_k\delta_k^t))^{-1}$

Incorporando a M la definición de δ_k tenemos:

$$M = E(X_k\delta_k^t)(E(\delta_k\delta_k^t))^{-1} = E(X_k(H_k(X_k - \hat{X}_{k-1}) + V_k)^t)\Delta_k^{-1}$$

Y finalmente, con la definición de $P_{k|k-1} = E[(X_k - \hat{X}_{k|k-1})(X_k - \hat{X}_{k|k-1})^t]$:

$$M = E(X_k(H_k(X_k - \hat{X}_{k|k-1}) + V_k)^t)\Delta_k^{-1} = P_{k|k-1}H_k\Delta_k^{-1}$$

Ya hemos llegado a la ecuación que queríamos

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + P_{k|k-1}H_k\Delta_k^{-1}\delta_k$$

Ahora vemos la ecuación con la matriz de varianzas-covarianzas

Por otro lado, usando que $P_k(X_k) - P_{k-1}(X_k) = M\delta_k$, que es lo deducido del lema 2.2, tenemos:

$$X_k - P_{k-1}(X_k) = X_k - P_k(X_k) + P_k(X_k) - P_{k-1}(X_k) = X_k - P_k(X_k) + M\delta_k$$

Finalmente usando el lema 2.1 y 2.3, por el cual $P_k(X) = P(X|Z_0, \dots, Z_k) = P(X|Z_0, \dots, Z_{k-1}, \delta_k)$, llegamos a lo que queríamos:

$$P_{k|k-1} = P_{k|k} + P_{k|k-1}H_k^t\Delta_k^{-1}H_kP_{k|k-1}^t$$

□

3. Aplicaciones Filtro de Kalman

3.1. Tracking

Una de las principales aplicaciones del filtro de Kalman es el tracking: el seguimiento de cierto objeto. Veamos un ejemplo de aplicación en esta línea. Nuestro objetivo será determinar la posición, velocidad y aceleración de un objeto.

Sabemos que lo primero a determinar son las ecuaciones del modelo, es decir, queremos definir las matrices F_k y H_k de forma que:

$$\hat{X}_{k|k-1} = F_k \hat{X}_{k-1|k-1} + w_k$$

$$Z_k = H_k X_k + v_k$$

donde, tal y como hemos definido hasta ahora, X_i es la variable de estado -en la ecuación suponemos que no hay término de control- y Z_i son las observaciones. Trabajamos con las hipótesis de que los vectores aleatorios w_k y v_k siguen una distribución gaussiana de media cero y además no están correlacionados, es decir, $E[v_i w_j^T] = 0 \forall i, j$

El objetivo es estimar la posición, velocidad y aceleración. En el caso que nos ocupa, el movimiento será en línea recta manteniendo la aceleración constante. A partir de las conocidas ecuaciones del movimiento, siendo v_i la velocidad, p_i la posición y a_i la aceleración:

$$v_{k+1} = v_k + T a_k$$

$$x_{k+1} = x_k + v_k T + \frac{1}{2} a_k T^2$$

Expresándolo en lenguaje matricial, tenemos que el vector de estado -la variable que queremos estimar- es:

$$X_k = \begin{bmatrix} p_k \\ v_k \\ a_k \end{bmatrix}$$

Así pues, con las ecuaciones que acabamos de definir, la matriz F será:

$$F = \begin{bmatrix} 1 & T & \frac{1}{2} T^2 \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix}$$

Y por lo tanto es invariante en el tiempo.

Y el vector w_k lo consideraremos nulo. Suponemos que medimos la posición, velocidad y aceleración, las cuales almacenaremos en el vector Z_i . Entonces, ya podemos escribir la ecuación que nos queda:

$$Z_k = H_k X_k + v_k$$

donde

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

que permanece también invariante en el tiempo.

Ahora aplicaremos las ecuaciones de Kalman para lograr la estimación que buscamos. Sin embargo, primero necesitamos establecer unas condiciones iniciales en la posición, velocidad, aceleración y también establecer los errores en el proceso y en las observaciones. Supongamos que partimos con: $p_0 = 100m$, $v_0 = 20m/s$ y $a_0 = 3m/s^2$. Consideraremos un intervalo de tiempo de $\Delta T = 1s$. Además consideraremos unos errores propios del proceso de 10 m en la posición, $2m/s$ en la velocidad y $0,1m/s^2$ en la aceleración. Estos serán necesarios para definir la matriz de covarianza del error.

Así, aplicamos la primera ecuación del filtro de Kalman:

$$X_{1|0} = \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 100 \\ 20 \\ 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 121,5 \\ 23 \\ 3 \end{bmatrix} = \begin{bmatrix} p_1 \\ v_1 \\ a_1 \end{bmatrix}$$

Ahora actualizaremos la matriz de covarianza. Sabemos que, siendo P la matriz de covarianza de los errores del proceso, entonces:

$$P = \begin{bmatrix} e_p^2 & e_p e_v & e_p e_a \\ e_p e_v & e_v^2 & e_v e_a \\ e_p e_a & e_v e_a & e_a^2 \end{bmatrix}$$

donde e_p es el error en la posición, e_v es el error en la velocidad y e_a el error en la aceleración. Con los errores que hemos establecido anteriormente, la matriz, en las condiciones iniciales, nos queda:

$$P = \begin{bmatrix} 100 & 20 & 1 \\ 20 & 4 & 0,2 \\ 1 & 0,2 & 0,01 \end{bmatrix}$$

En nuestro caso la ecuación de actualización de la matriz de covarianza es:

$$P_{k|k-1} = F P_{k-1|k-1} F^T$$

Esto supone que:

$$\begin{aligned} P_{1|0} &= \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 100 & 20 & 1 \\ 20 & 4 & 0,2 \\ 1 & 0,2 & 0,01 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \frac{1}{2} & 1 & 1 \end{bmatrix} = \\ &= \begin{bmatrix} 145,2025 & 25,305 & 1,205 \\ 25,3015 & 4,41 & 0,21 \\ 1,205 & 0,21 & 0,01 \end{bmatrix} \end{aligned}$$

Cuando nos ocupamos de problemas de rastreo, se suelen considerar simplemente los errores de la diagonal de la matriz. Así, la matriz de varianzas-covarianzas que usaremos será:

$$P_{1|0} = \begin{bmatrix} 145,2025 & 0 & 0 \\ 0 & 4,41 & 0 \\ 0 & 0 & 0,01 \end{bmatrix}$$

Ahora calculamos el valor de la ganancia de Kalman. Recordemos la fórmula para calcularlo:

$$K_k = P_{k|k-1} H^T S_k^{-1}$$

donde

$$S_k = H P_{k|k-1} H^T + R$$

y R es la matriz de varianzas-covarianzas del error en las observaciones v_k . Primero, pues, establecemos los errores en las observaciones -los cuales consideraremos constantes para todas las observaciones- y definimos, a partir de ellos, la matriz R. Consideraremos $e'_p = 15m$, $e'_v = 4m/s$ y $e'_a = 0,2m/s^2$. Entonces:

$$R = \begin{bmatrix} 225 & 60 & 3 \\ 60 & 16 & 0,8 \\ 3 & 0,8 & 0,04 \end{bmatrix}$$

De nuevo podemos desestimar los elementos que quedan fuera de la diagonal y trabajaremos con la matriz:

$$R = \begin{bmatrix} 225 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 0,04 \end{bmatrix}$$

Calculamos ahora la matriz S:

$$\begin{aligned} S_1 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 145,2025 & 0 & 0 \\ 0 & 4,41 & 0 \\ 0 & 0 & 0,01 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 225 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 0,04 \end{bmatrix} = \\ &= \begin{bmatrix} 370,2025 & 0 & 0 \\ 0 & 20,41 & 0 \\ 0 & 0 & 0,05 \end{bmatrix} \end{aligned}$$

Para calcular K necesitaremos calcular primero la matriz inversa de S:

$$S_1^{-1} = \begin{bmatrix} 0,00270122 & 0 & 0 \\ 0 & 0,0489956 & 0 \\ 0 & 0 & 20 \end{bmatrix}$$

Así pues:

$$K_1 = \begin{bmatrix} 145,2025 & 0 & 0 \\ 0 & 4,41 & 0 \\ 0 & 0 & 0,01 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0,00270122 & 0 & 0 \\ 0 & 0,0489956 & 0 \\ 0 & 0 & 20 \end{bmatrix} =$$

$$= \begin{bmatrix} 0,392224 & 0 & 0 \\ 0 & 0,2160706 & 0 \\ 0 & 0 & 0,2 \end{bmatrix}$$

Ahora empezamos a considerar las nuevas observaciones que medimos. Suponemos una nueva observación $p_1 = 130m$, $v_1 = 23m/s$, $a_1 = 3m/s^2$. Podemos aplicar las ecuaciones definidas anteriormente como ecuaciones de corrección. Por lo tanto, tenemos $\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k \cdot \delta_k$. Calculemos, con nuestros datos, la medida de innovación $\delta_k = Z_k - H\hat{X}_{k|k-1}$.

$$\delta_k = \begin{bmatrix} 130 \\ 23 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 121,5 \\ 23 \\ 3 \end{bmatrix} = \begin{bmatrix} 8,5 \\ 0 \\ 0 \end{bmatrix}$$

Ya podemos calcular la actualización de la variable:

$$\begin{aligned} X_{1|1} &= \begin{bmatrix} 121,5 \\ 23 \\ 3 \end{bmatrix} + \begin{bmatrix} 0,392224 & 0 & 0 \\ 0 & 0,2160706 & 0 \\ 0 & 0 & 0,2 \end{bmatrix} \cdot \begin{bmatrix} 8,5 \\ 0 \\ 0 \end{bmatrix} = \\ &= \begin{bmatrix} 121,5 \\ 23 \\ 3 \end{bmatrix} + \begin{bmatrix} 3,3339 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 124,8339 \\ 23 \\ 3 \end{bmatrix} \end{aligned}$$

Ahora actualizaremos la matriz de varianzas-covarianzas. La ecuación para hacerlo es: $P_{k|k} = (I - K_k H)P_{k|k-1}$. Por lo tanto:

$$\begin{aligned} P_{1|1} &= \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0,392224 & 0 & 0 \\ 0 & 0,215878 & 0 \\ 0 & 0 & 0,2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \times \\ &\quad \times \begin{bmatrix} 145,2025 & 0 & 0 \\ 0 & 4,41 & 0 \\ 0 & 0 & 0,01 \end{bmatrix} = \\ &= \begin{bmatrix} 0,607776 & 0 & 0 \\ 0 & 0,784122 & 0 \\ 0 & 0 & 0,8 \end{bmatrix} \cdot \begin{bmatrix} 145,2025 & 0 & 0 \\ 0 & 4,41 & 0 \\ 0 & 0 & 0,01 \end{bmatrix} = \\ &= \begin{bmatrix} 88,25 & 0 & 0 \\ 0 & 3,457 & 0 \\ 0 & 0 & 0,008 \end{bmatrix} \end{aligned}$$

Ahora empezaremos con la segunda iteración. Actualizamos la variable de estado a partir de la fórmula $X_{2|1} = FX_{1|1} + w_2$. Sabemos, porque así lo hemos establecido anteriormente, que tomamos un intervalo de tiempo T de 1 seg. Siendo así:

$$X_{2|1} = \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 124,8339 \\ 23 \\ 3 \end{bmatrix} =$$

$$= [149, 3339 \quad 26 \quad 3]$$

Hacemos lo correspondiente con la matriz de varianzas-covarianzas a partir de $P_{2|1} = FP_{1|1}F^T$

$$\begin{aligned} P_{2|1} &= \begin{bmatrix} 1 & 1 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 88,25 & 0 & 0 \\ 0 & 3,457 & 0 \\ 0 & 0 & 0,008 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \frac{1}{2} & 1 & 1 \end{bmatrix} = \\ &= \begin{bmatrix} 91,71 & 3,461 & 0,004 \\ 3,461 & 3,465 & 0,008 \\ 0,004 & 0,008 & 0,008 \end{bmatrix} \end{aligned}$$

Considerando, de nuevo, solo los términos de la diagonal, nos queda:

$$P_{2|1} = \begin{bmatrix} 91,71 & 0 & 0 \\ 0 & 3,465 & 0 \\ 0 & 0 & 0,008 \end{bmatrix}$$

Suponemos que tomamos una segunda observación: $p_2 = 150m$, $v_2 = 25m/s$, $a_2 = 2,5m/s^2$. Para calcular la nueva estimación, calcularemos primero la nueva ganancia de Kalman. La matriz $S_k = HP_{k|k-1}H^T + R$ será:

$$\begin{aligned} S_2 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 91,71 & 0 & 0 \\ 0 & 3,465 & 0 \\ 0 & 0 & 0,008 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 225 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 0,04 \end{bmatrix} = \\ &= \begin{bmatrix} 316,71 & 0 & 0 \\ 0 & 19,465 & 0 \\ 0 & 0 & 0,048 \end{bmatrix} \end{aligned}$$

Su inversa es:

$$S_2^{-1} = \begin{bmatrix} 0,0031575 & 0 & 0 \\ 0 & 0,0513743 & 0 \\ 0 & 0 & 20,8333 \end{bmatrix}$$

Así, la ganancia de Kalman $K_2 = P_{2|1}H^T S_2^{-1}$ queda como:

$$\begin{aligned} K_2 &= \begin{bmatrix} 91,71 & 0 & 0 \\ 0 & 3,465 & 0 \\ 0 & 0 & 0,008 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0,0031575 & 0 & 0 \\ 0 & 0,0513743 & 0 \\ 0 & 0 & 20,8333 \end{bmatrix} = \\ &= \begin{bmatrix} 0,28957 & 0 & 0 \\ 0 & 0,178 & 0 \\ 0 & 0 & 0,167 \end{bmatrix} \end{aligned}$$

A partir de la nueva observación que hemos definido, podemos calcular la medida de innovación:

$$\delta_2 = \begin{bmatrix} 150 \\ 25 \\ 2,5 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 149,3339 \\ 26 \\ 3 \end{bmatrix} = \begin{bmatrix} 0,6661 \\ -1 \\ -0,5 \end{bmatrix}$$

Y con ello ya tenemos la nueva estimación $x_{2|2} = x_{2|1} + K_2 \cdot \delta_2$

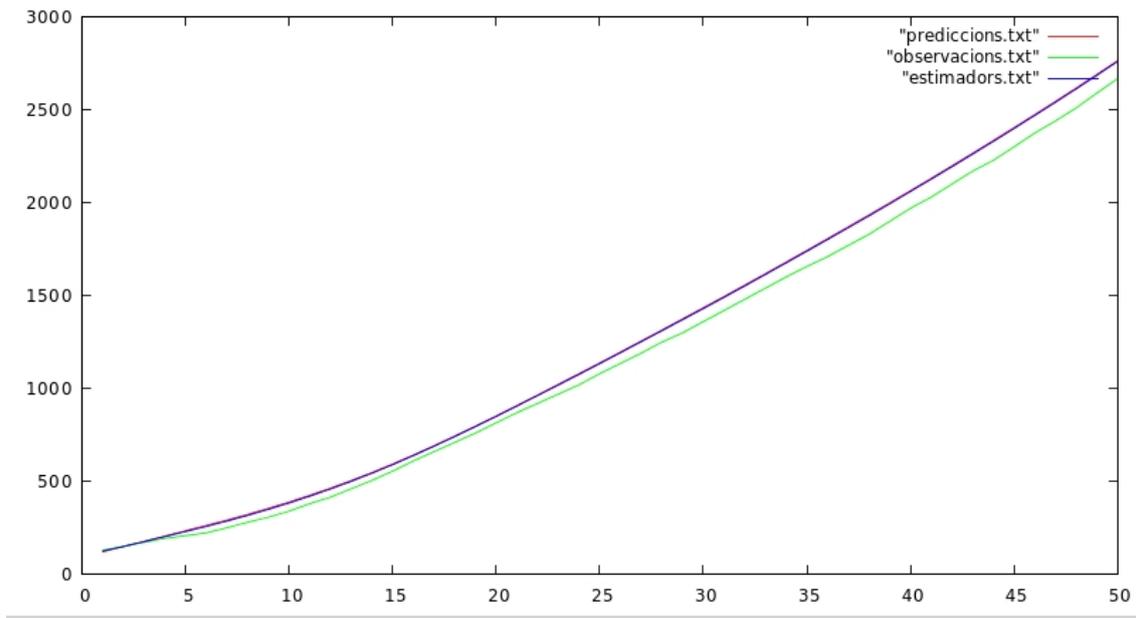
$$\begin{aligned} X_{2|2} &= \begin{bmatrix} 149,3339 \\ 26 \\ 3 \end{bmatrix} + \begin{bmatrix} 0,28957 & 0 & 0 \\ 0 & 0,178 & 0 \\ 0 & 0 & 0,167 \end{bmatrix} \cdot \begin{bmatrix} -0,6661 \\ -1 \\ -0,5 \end{bmatrix} = \\ &= \begin{bmatrix} 149,3339 \\ 26 \\ 3 \end{bmatrix} + \begin{bmatrix} 0,1929 \\ -0,178 \\ -0,0835 \end{bmatrix} = \begin{bmatrix} 149,5268 \\ 25,822 \\ 2,9165 \end{bmatrix} \end{aligned}$$

Y la nueva matriz de varianzas-covarianzas $P_{2|2} = (I - K_k H)P_{2|1}$:

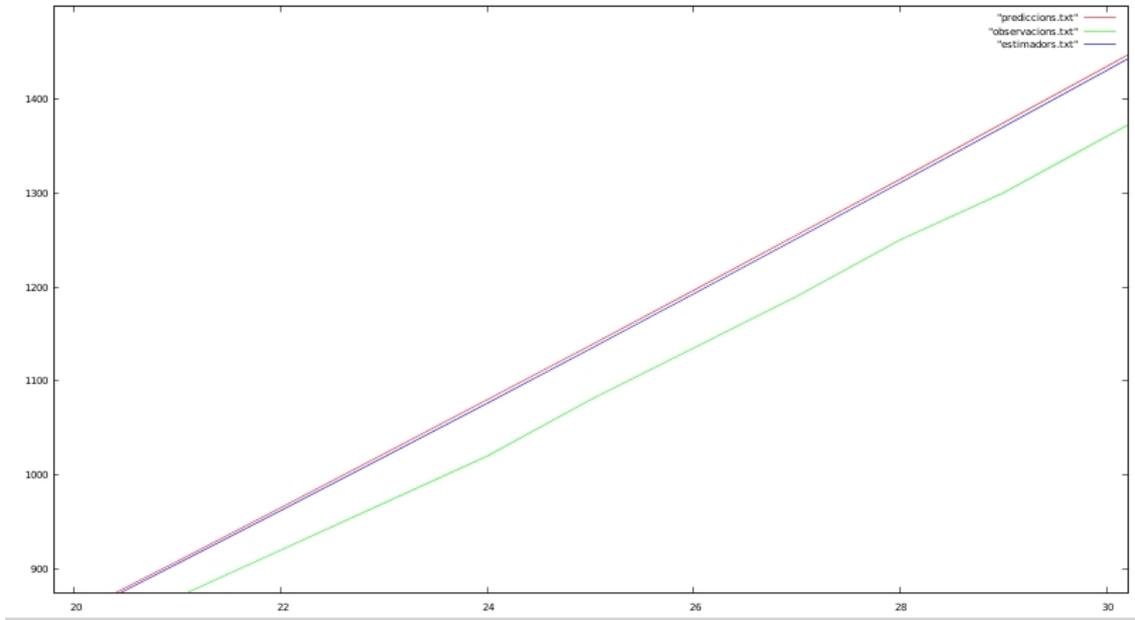
$$\begin{aligned} P_{2|2} &= \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0,28957 & 0 & 0 \\ 0 & 0,178 & 0 \\ 0 & 0 & 0,167 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right) \times \\ &\times \begin{bmatrix} 91,71 & 0 & 0 \\ 0 & 3,465 & 0 \\ 0 & 0 & 0,008 \end{bmatrix} = \begin{bmatrix} 65,15 & 0 & 0 \\ 0 & 2,8482 & 0 \\ 0 & 0 & 0,0067 \end{bmatrix} \end{aligned}$$

Queremos ver que pasa si repetimos sucesivamente estas iteraciones, añadiendo cada vez nuevas observaciones. Para hacerlo, escribiremos un código que nos automatice este proceso. Queremos como salida del programa las estimaciones $\hat{x}_{i|i} \forall i = 1 \dots n$ donde n son el número total de iteraciones que consideraremos. Una vez hecho esto, dibujaremos las gráficas con las predicciones, observaciones y estimaciones. En el anexo incorporaremos las observaciones utilizadas, así como las predicciones y estimaciones resultantes.

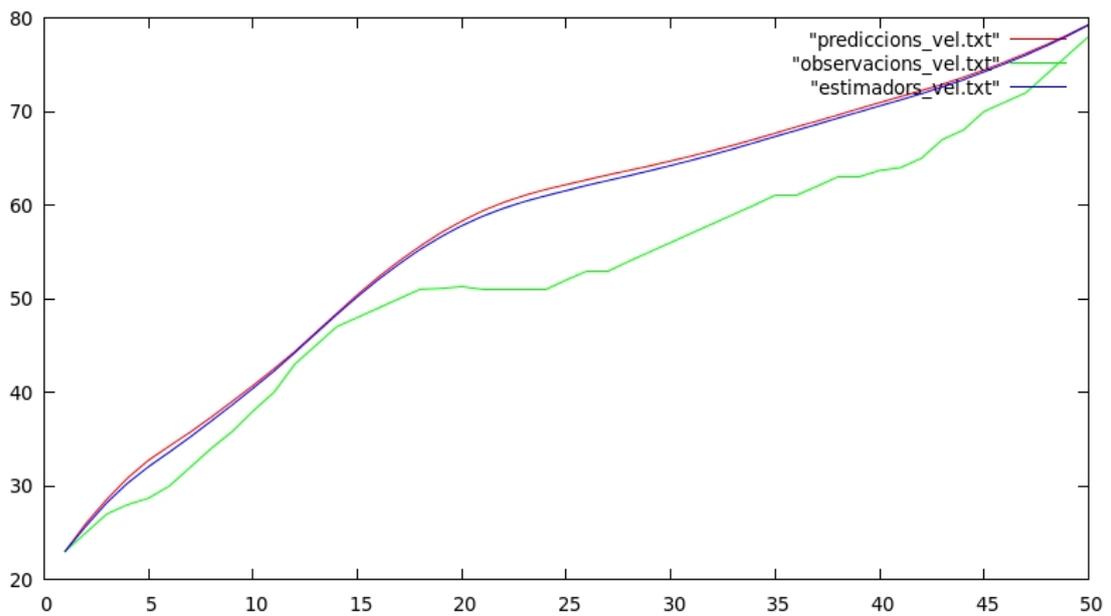
Respecto a la variable posición, el resultado hasta $T = 50$ es el siguiente:



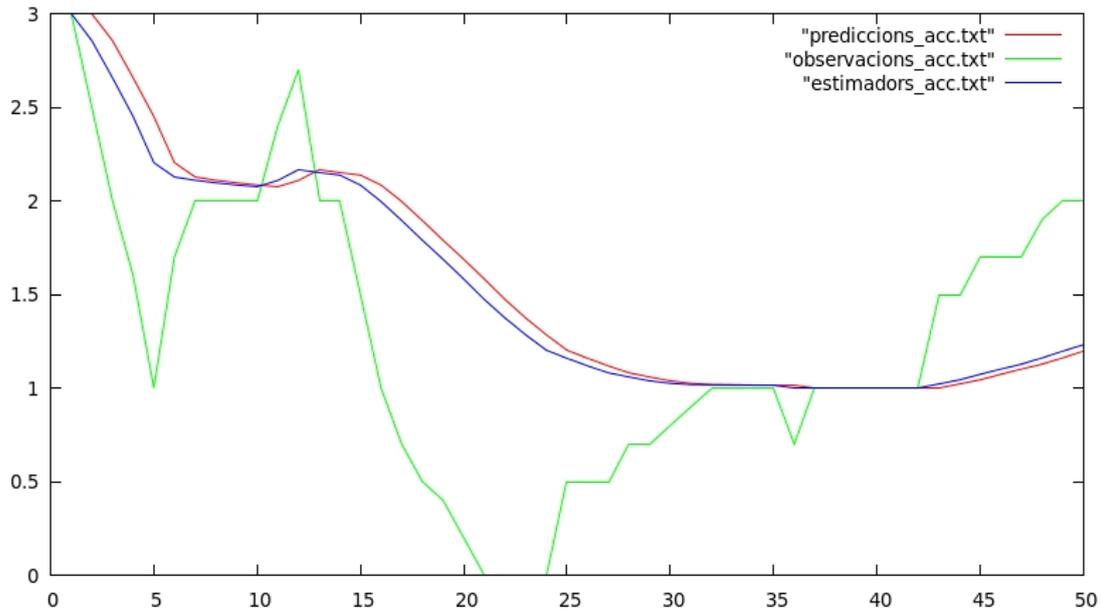
Efectivamente, las estimaciones están entre las predicciones y las observaciones. Su cercanía a una u otra viene determinada por la ganancia de Kalman, que lo que hará, tal y como hemos explicado teóricamente, es ponderar el peso de cada error. Si aumentamos el zoom podremos verlo con mejor detalle:



Hacemos lo mismo con la variable velocidad



Y con la variable aceleración



3.2. Estimación del canal en sistemas de comunicación inalámbricos

Otro ámbito en el que se suele aplicar el algoritmo del filtro de Kalman es el de las comunicaciones inalámbricas. Consideremos el proceso de transmisión y recepción de una señal. Sabemos que, en general, las señales emitidas por el transmisor llegan al receptor con interferencias y ruido: se superponen las que llegan directamente, con rebote, con retraso. . . Nuestro objetivo será estudiar la salida de un filtro digital. En síntesis, sabemos que los filtros tienen como entrada una señal analógica o digital y en su salida emiten otra señal con modificaciones.

Para este ejemplo consideraremos los filtros digitales cuya respuesta es de dimensión finita, en concreto de dimensión 3. Lo que haremos es buscar los coeficientes respuesta del filtro, que cambian en el tiempo. Lo haremos observando la transmisión y recepción de la señal. Recordemos que consideramos la respuesta del filtro como 3-dimensional y lo que buscamos son sus coeficientes.

Podemos considerar que las ecuaciones del modelo son:

$$\begin{bmatrix} x_k \\ x'_k \\ x''_k \end{bmatrix} = \begin{bmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ x'_{k-1} \\ x''_{k-1} \end{bmatrix} + \begin{bmatrix} w_k \\ w'_k \\ w''_k \end{bmatrix}$$

donde $W_k = [w_k, w'_k, w''_k]$ es el vector error de distribución gaussiana, con media 0 y matriz de covarianza Q .

Por otro lado, la ecuación de las observaciones será:

$$z_k = [h_k \ h_{k-1} \ h_{k-2}] \begin{bmatrix} x_k \\ x'_k \\ x''_k \end{bmatrix} + v_k$$

Donde v_k es el error en las observaciones y es de distribución gaussiana, de media 0 y con matriz de covarianza R.

Ahora fijemos los datos con los que trabajaremos. Tomaremos $\alpha_1 = 0,85$, $\alpha_2 = 1,001$ y $\alpha_3 = -0,95$. La matriz de varianzas-covarianzas del vector v será:

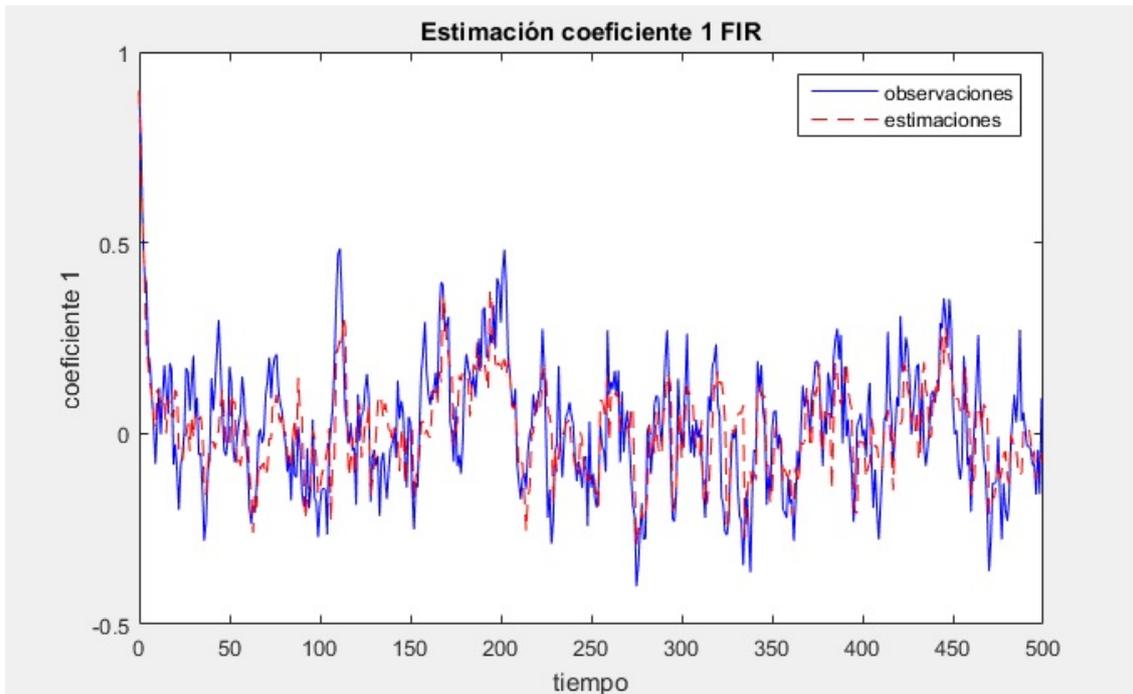
$$Q = \begin{bmatrix} 0,1 & 0 & 0 \\ 0 & 0,1 & 0 \\ 0 & 0 & 0,1 \end{bmatrix}$$

Para conocer la ecuación de observación necesitamos $[h_k, h_{k-1}, h_{k-2}]$. La única condición que imponemos sobre estos coeficientes es que sean de distribución gaussiana. Tomaremos $h_k \sim N(0, 1)$. En consecuencia, la matriz de varianzas-covarianzas del error v_k en las observaciones será:

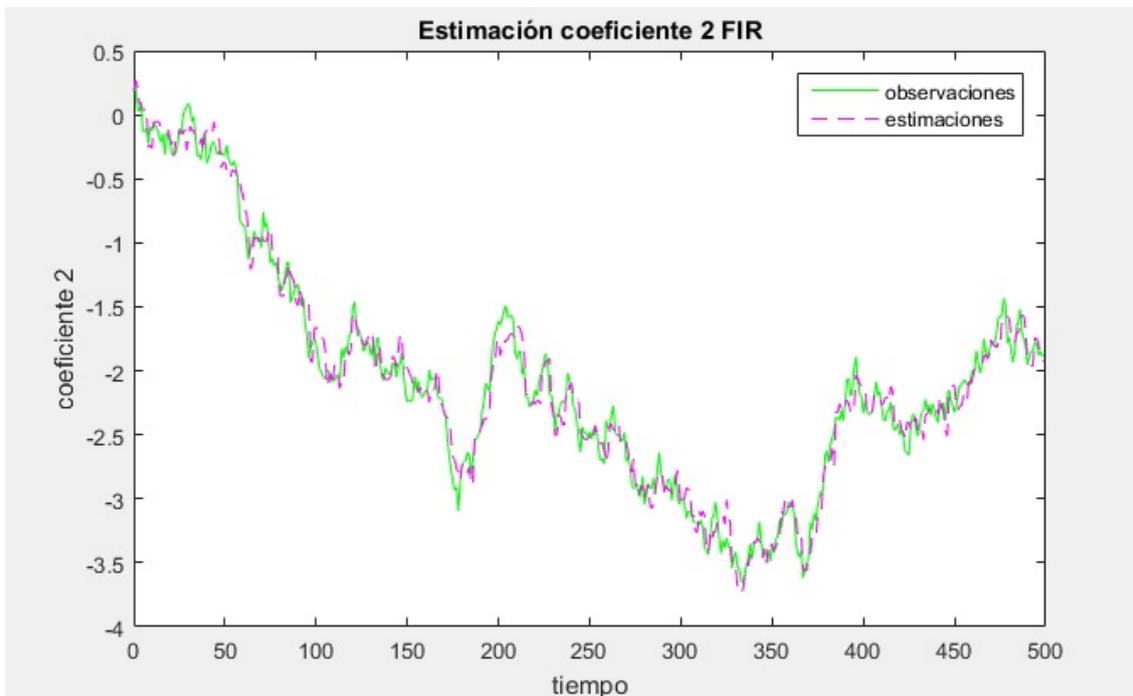
$$R = \begin{bmatrix} 0,1 & 0 & 0 \\ 0 & 0,1 & 0 \\ 0 & 0 & 0,1 \end{bmatrix}$$

Por último, las condiciones iniciales para el filtro de Kalman serán $P_{0|0} = Id$ y la variable X de los coeficientes la inicializaremos como un vector de dimensión 3 de distribución normal. Esta vez recurriremos a Matlab para programar el algoritmo del filtro de Kalman que nos dará como salida los coeficientes X de la señal respuesta. Veamos los resultados.

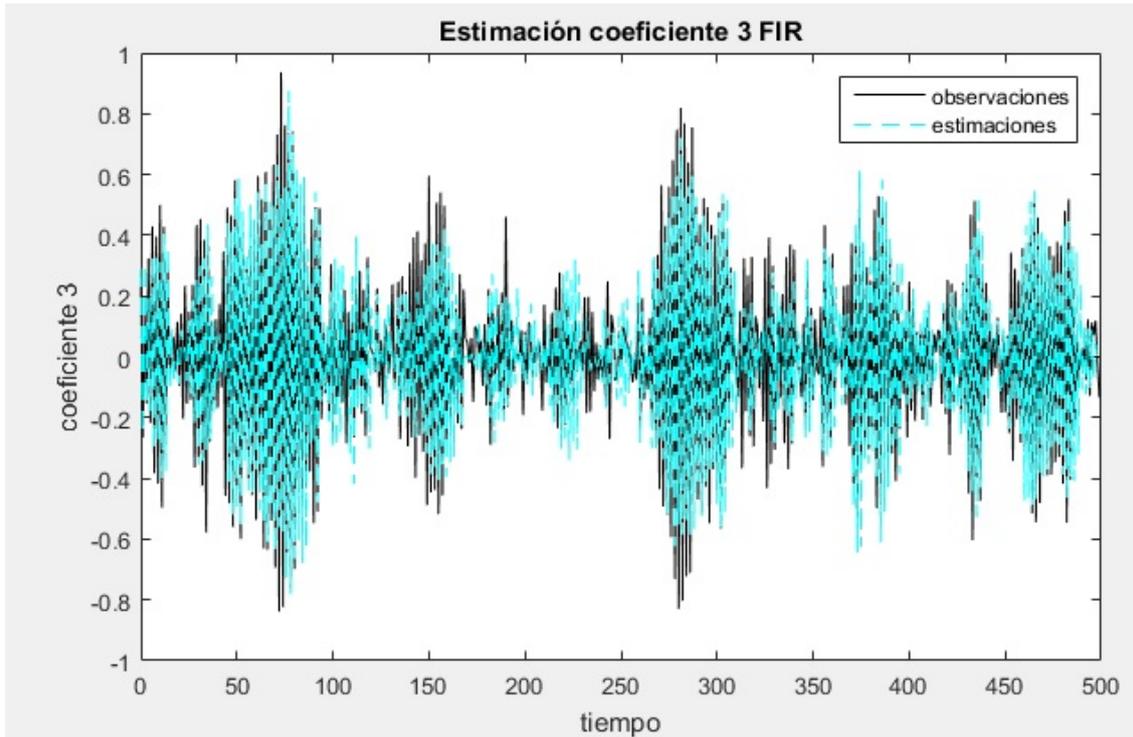
Si dibujamos en un gráfico los valores observados frente a los estimados del primer coeficiente respuesta tenemos:



Por lo que respecta al coeficiente 2 tenemos como resultado:



Por último, con el coeficiente 3 tenemos:



3.3. Filtro adaptativo. RLS: algoritmo recursivo de error cuadrático mínimo

Siguiendo con el procesamiento de una señal, nos volvemos a ocupar de los filtros. Nuestro objetivo ahora es encontrar los coeficientes de salida que nos proporcionan el mínimo error cuadrático, entendiendo por error la diferencia entre la señal que se esperaba y la que nos acaba proporcionando el filtro.

El algoritmo que consideramos es conocido como RLS (Recursive Least Squares algorithm), pues para optimizar el proceso opera de forma recursiva. A diferencia de antes, el tamaño de la señal será finito pero D-dimensional.

Ahora nuestra ecuación de estado será:

$$\begin{bmatrix} x_k^1 \\ x_k^2 \\ \dots \\ x_k^D \end{bmatrix} = \begin{bmatrix} \lambda^{\frac{-1}{2}} & 0 & \dots & 0 \\ 0 & \lambda^{\frac{-1}{2}} & \dots & 0 \\ 0 & 0 & \lambda^{\frac{-1}{2}} & \dots \\ 0 & 0 & \dots & \lambda^{\frac{-1}{2}} \end{bmatrix} \begin{bmatrix} x_{k-1}^1 \\ x_{k-1}^2 \\ \dots \\ x_{k-1}^D \end{bmatrix}$$

Como vemos, ahora el error del modelo es nulo.

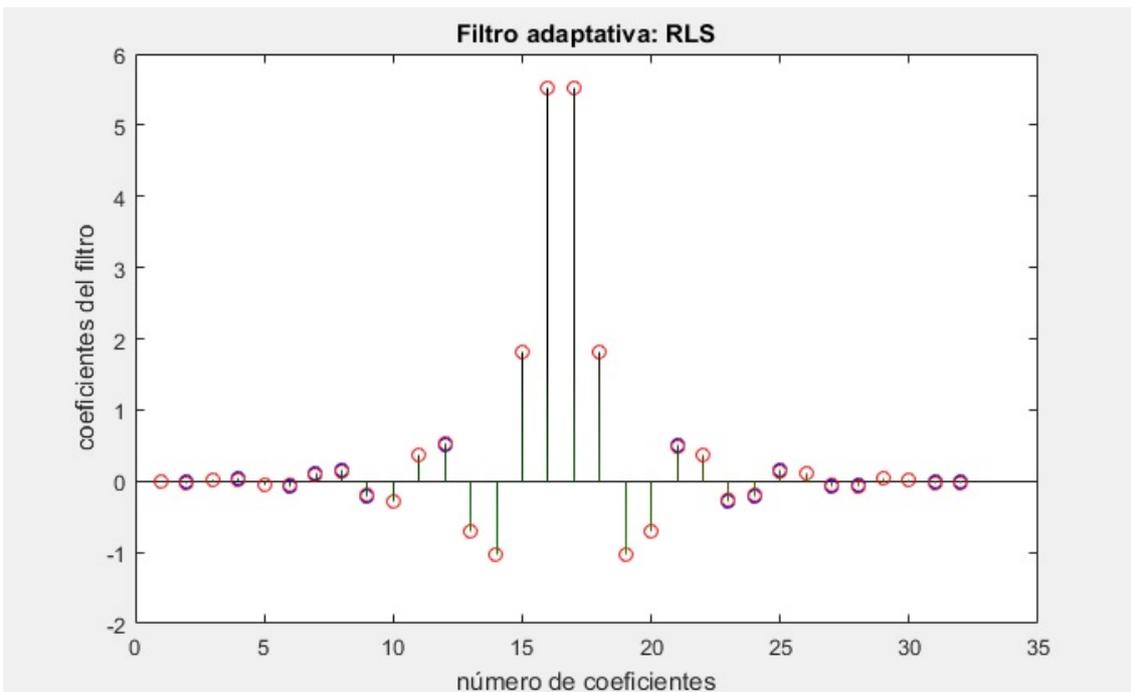
Ahora escribimos la ecuación de observación. Sean Z_k las observaciones en el tiempo k , entonces la ecuación que las relaciona con el vector X es:

$$Z_k = [h_t \quad h_{t-1} \quad \dots \quad h_{t-D+1}] \begin{bmatrix} x_k^1 \\ x_k^2 \\ \dots \\ x_k^D \end{bmatrix} + v_k$$

En este caso el vector error es v_k , y es gaussiano de media zero y con matriz de varianzas-covarianzas la identidad.

Como hemos hecho hasta ahora, utilizaremos el filtro de Kalman para estimar el vector de variables X que es cambiante con el tiempo. Establezcamos los datos de nuestro problema. Consideraremos $D = 32$, $\lambda = 0,99$. De igual forma que en el problema anterior, la única condición que imponemos a los coeficientes c_i es que sigan una distribución gaussiana. Por último inicializaremos la matriz de covarianza $P_{0|0}$ como la matriz identidad.

Para programar este algoritmo usaremos, de nuevo, Matlab. Los resultados que nos muestra son los siguientes:



3.4. Estimación del PIB mensual

Uno de los indicadores económicos más importantes en la descripción de la economía de un país es el PIB. El PIB se define como la suma de los bienes y servicios producidos por un país en un cierto espacio de tiempo, que puede ser mensual, trimestral, anual...

La primera publicación de los datos oficiales del PIB es trimestral, es decir, los primeros datos que se publican son del total de la producción en tres meses. Sin embargo, hay que tener en cuenta que la frecuencia de un modelo será, como mínimo, igual que la del indicador menos frecuente. Es por esto que la publicación de los datos del PIB trimestralmente puede ser un inconveniente para un estudio económico, pues el PIB supone una gran fuente de información de la que no quisiéramos prescindir en un análisis cuantitativo de menor frecuencia a la trimestral. Así pues, el objetivo de nuestro ejemplo será interpolar los datos trimestrales del PIB para lograr una estimación del PIB mensual. La ventaja del filtro de Kalman en esta interpolación es que nos permite añadir observaciones conforme estas van llegando.

Para elaborar nuestros modelos deberemos tener en cuenta distintos factores: la estacionariedad e integración de las variables -una variable es integrada si se puede expresar mediante un modelo autorregresivo- o la necesidad de incluir series relacionadas -esto es la necesidad de incorporar indicadores externos cuya correlación con el PIB sea muy elevada y que a su vez sean publicados con una frecuencia superior a 3 meses-.

Antes de continuar aclaremos que entendemos por modelo autorregresivo. Un modelo es autorregresivo si la variable de salida establece una relación lineal con sus propios valores anteriores. Formalizando, diremos que un modelo autorregresivo se define como:

$$X_k = a + \sum_{i=1}^p \beta X_{k-i}$$

El valor p será el orden del modelo.

En los modelos que expondremos consideraremos que las variables son integradas de orden igual o superior a 1 -es decir, que se pueden escribir siguiendo un modelo autorregresivo-. Si no fuera así no aplicaríamos el filtro de Kalman y haríamos las estimaciones a partir de deducciones y cálculos. Siempre bajo la hipótesis de modelo autorregresivo, desarrollaremos dos modelos: en uno de ellos incluiremos series relacionadas y en el otro no.

3.4.1. Rasgos generales del modelo

Antes de entrar en la formulación de cada modelo, veamos los rasgos más generales. Recordemos que en el filtro de Kalman necesitamos una ecuación de transición, que formula la evolución de la variable de estado, y la ecuación de observaciones, que relaciona los datos de las observaciones con la variable de estado.

$$\begin{aligned} X_{k+1} &= F_k X_k + C_k U_k + w_{k+1} \\ Z_k &= H_k X_k + A_k U'_k + v_k \end{aligned}$$

La primera es la ecuación de transición y la segunda la ecuación de observación. Tenemos el vector de estados $X_k \in \mathbb{R}^3$, dado que se refiere a los datos del PIB mensual, que agruparemos de 3 en 3 ya que los datos conocidos son los del PIB trimestral. El hecho de definir este vector de variables como dimensión 3, nos permitirá imponer restricciones como que la suma de todos los componentes del vector sea igual al PIB trimestral publicado. Con esta restricción ya tendremos definida la ecuación de observaciones y dado que la suma de los 3 meses debe ser exactamente el dato trimestral del PIB nos desaparecerá el término del error v_k . Por otro lado, las variables U_k y U'_k se refieren a las series relacionadas. Por último tenemos los términos de error w_k y v_k , que siguen una distribución normal de media 0 y con matriz de covarianzas Q y R respectivamente. Los términos de error no están correlacionados entre ellos. Así pues, necesitaremos estimar los coeficientes de las matrices F_k , C_k , H_k y A_k . Las estimaciones que elegiremos serán aquellas que maximicen la función de máxima verosimilitud, es decir, aquellos coeficientes que hagan máxima la densidad. A partir de ahora consideraremos que estas matrices son invariantes en el tiempo.

Con todo lo descrito hasta ahora ya tenemos el modelo general para la estimación del PIB mensual, en caso de incorporar series relacionadas este es:

$$X_{k+1} = F_k X_k + C_k U_k + w_{k+1}$$

$$Z_k = H_k X_k + A_k U'_k$$

Y en caso de no incluir series relacionadas, la forma general del modelo es:

$$X_{k+1} = F_k X_k + w_{k+1}$$

$$Z_k = H_k X_k$$

3.4.2. Estimación máxima verosimilitud

Recordemos brevemente como maximizar la función de máxima verosimilitud de una variable $X = (x_1, \dots, x_n)$ en función de parámetros a_i . Sea $f(x_i; a_1, \dots, a_k)$ la función de densidad de la variable x_i , y al ser ésta independiente de las n variables restantes, la función de densidad de la variable X será $\prod_{i=1}^n f(x_i; a_1, \dots, a_k)$ y queremos los a_i que hagan máximo el valor de esta función. Para mayor simplicidad en los cálculos, podemos maximizar $\ln(\prod_{i=1}^n f(x_i; a_1, \dots, a_k)) = \sum_{i=1}^n \ln f(x_i; a_1, \dots, a_k)$.

Encontremos la fórmula general para la estimación de las matrices. Supondremos que el PIB trimestral sigue una distribución normal: $Z_k \sim N(H_k X_{k|k-1} + A_k U'_k, H_k P_{k|k-1} H_k)$, donde la matriz P es la matriz del error del filtro de Kalman. Recordemos que, si tenemos una variable $Y \sim N(\mu, \sigma^2)$, entonces su función de densidad es

$$f_Y(y) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

Antes ya hemos dicho que, para simplicidad de los cálculos, maximizaríamos el logaritmo de la función de densidad de $Z = (Z_1, \dots, Z_{3r})$. Si consideramos n el total de meses que estimaremos, dado que Z solo contiene los datos trimestrales, el vector Z será de tamaño $n/3$. Por lo tanto, al calcular la densidad conjunta del vector Z tendremos:

$$\ln \left(\prod_{i=1}^{n/3} f_{Z_i}(z) \right) = \sum_{i=1}^{n/3} \ln (f_{Z_i}(z)) = \sum_{i=1}^{n/3} \left(-\ln(\sqrt{2\pi}) - \ln(\sqrt{H_k P_{k|k-1} H_k}) - \frac{(z - H_k X_{k|k-1} - A_k U'_k)^2}{2H_k P_{k|k-1} H_k} \right)$$

Ahora, dado que la matriz P hace referencia a la matriz de errores del filtro, aplicamos la ecuación de proyección del filtro de Kalman para esta matriz, y añadimos también la ecuación de transición para la variable $X_{k|k-1}$. Suponemos que las matrices permanecen invariantes con el tiempo.

$$\begin{aligned} &= \sum_{i=1}^{n/3} \left(\frac{-1}{2} \ln(2\pi) - \frac{1}{2} \ln (H(FP_{k-1|k-1}F^t + Q)H^t) - \right. \\ &\quad \left. - \frac{(z - H(FX_{k-1|k-1} + CU) - AU')^2}{2H(FP_{k-1|k-1}F^t + Q)H} \right) = \\ &= -\frac{n}{6} \ln(2\pi) - \frac{1}{2} \sum_{i=1}^{n/3} (\ln(H(FP_{k-1|k-1}F^t + Q)H)) - \\ &\quad - \frac{1}{2} \sum_{i=1}^{n/3} \left(\frac{(z - H(FX_{k-1|k-1} + CU) - AU')^2}{H(FP_{k-1|k-1}F^t + Q)H} \right) \end{aligned}$$

Así pues, para cada modelo deberemos maximizar esta función y así hallar los coeficientes de las matrices desconocidas F , C , Q , H , A .

3.4.3. Modelo autorregresivo sin series relacionadas

Ahora entremos a detallar cada modelo. El primer modelo que consideraremos será integrado de orden 1 y sin incorporar series relacionadas. Dado que, en este caso, suponemos que el PIB mensual es una variable integrada de orden 1, y a partir de lo expuesto en el anexo acerca la diferenciación, podemos escribir la ecuación de transición como:

$$\Delta x_k = a\Delta x_{k-1} + w_k$$

donde w_k es el término del error que sigue una distribución normal y con matriz de covarianzas

$$Q = \begin{pmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix}$$

Y por la definición de Δ podemos reescribir esta ecuación con orden 2:

$$x_k = (1 + a)x_{k-1} - ax_{k-2} + w_k$$

Anteriormente hemos dicho que consideraríamos un vector que incluyera los tres meses que forman un trimestre, y por lo tanto, tomaremos el vector $X = (x_k, x_{k-1}, x_{k-2})$. Escribimos la ecuación anterior en forma matricial y obtenemos:

$$\begin{pmatrix} x_{k+1} \\ x_k \\ x_{k-1} \end{pmatrix} = \begin{pmatrix} 1+a & -a & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_k \\ x_{k-1} \\ x_{k-2} \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} w_{k+1} \\ w_k \\ w_{k-1} \end{pmatrix}$$

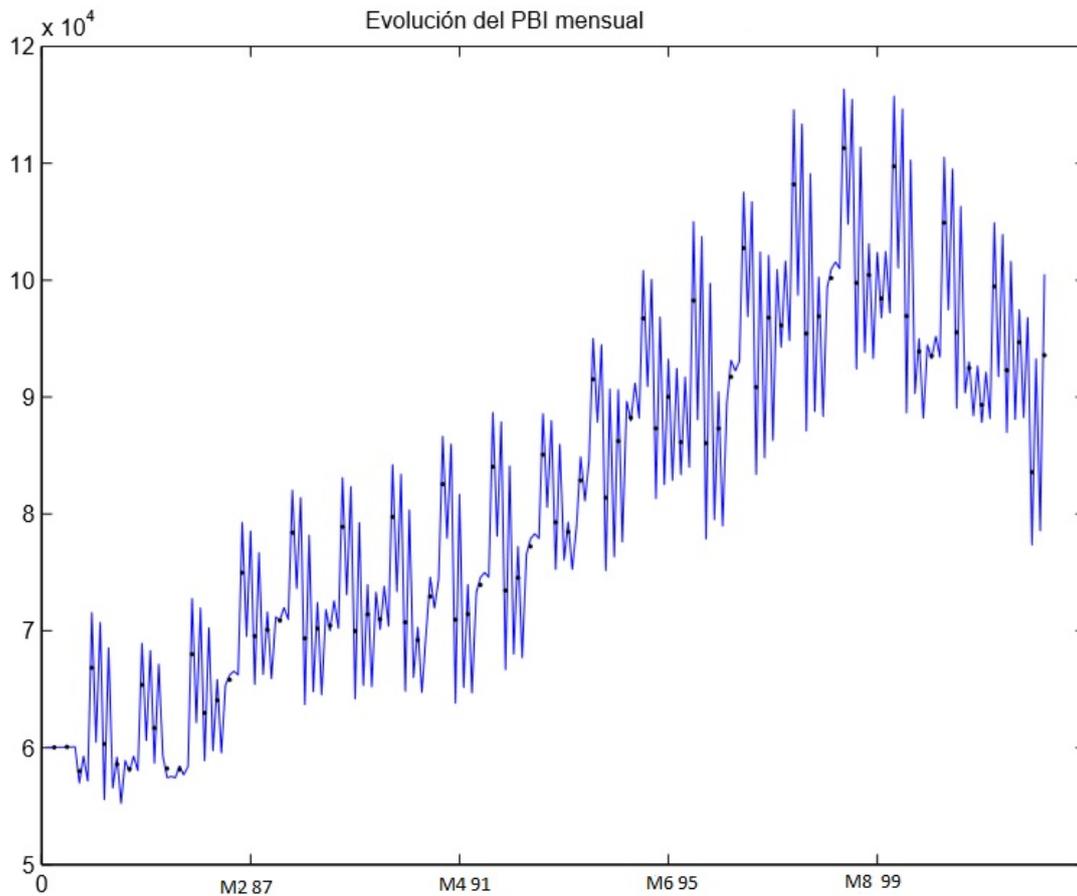
Ahora consideremos el vector $z = (0, 0, z_3, \dots, 0, 0, z_{3i}, \dots)$ que contiene las observaciones del PIB trimestral. Por ello solo tendrán valor las componentes múltiples de 3, y estas coincidirán con la suma de las componentes del vector X. Así pues, tendremos:

$$z_k = \begin{cases} [1 & 1 & 1] \begin{bmatrix} x_k \\ x_{k-1} \\ x_{k-2} \end{bmatrix} & \text{cuando } 3|k \\ 0 & \text{en caso contrario} \end{cases}$$

Con esto ya tenemos formulada la ecuación de las observaciones. Por lo tanto, los únicos coeficientes que tendremos que estimar por máxima verosimilitud serán a y σ . Una vez calculados estos valores aplicaremos filtro de Kalman para encontrar finalmente la estimación del PIB mensual. Tal y como hemos visto hasta ahora, las ecuaciones del modelo a las que aplicaremos Kalman serán:

$$\begin{cases} \begin{bmatrix} x_{k+1} \\ x_k \\ x_{k-1} \end{bmatrix} = \begin{bmatrix} 1+a & -a & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_k \\ x_{k-1} \\ x_{k-2} \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} w_{k+1} \\ w_k \\ w_{k-1} \end{bmatrix} \\ z_k = [1 & 1 & 1] \begin{bmatrix} x_k \\ x_{k-1} \\ x_{k-2} \end{bmatrix} \end{cases}$$

Por la complejidad y laboriosidad al programar todo este proceso -lo hemos hecho en los anteriores ejemplos más sencillos- cogemos los resultados obtenidos por otras investigaciones y así veremos que conclusiones podemos obtener aplicando el filtro de Kalman a estos modelos. En el anexo incorporamos los resultados numéricos. Veamos aquí los gráficos que nos resultan.



El estudio comprende el periodo que va entre enero de 1983 y diciembre de 2002 en Uruguay. Los puntos se refieren a cada trimestre. Es evidente el carácter oscilatorio de estos datos. Pero, en definitiva, esta técnica nos ha permitido interpolar los datos trimestrales del PIB obteniendo así datos mensuales, que era nuestro objetivo inicial.

3.4.4. Modelo autorregresivo con series relacionadas

Centrémonos ahora en el segundo modelo, el cual se diferenciará respecto al primero en que este sí incorporará series relacionadas. Dado que volvemos a suponer que se puede corregir, mediante la diferenciación, la no estacionariedad de los datos, de forma que logramos escribir el PIB mensual de forma autorregresiva, las ecuaciones de este segundo modelo serán iguales que las del primero, pero con la diferencia de que esta vez se incorpora el término $C_k U_k$ por la presencia de series relacionadas. Así pues, como incorporamos esta información externa? Podemos suponer que la serie relacionada mantiene una relación lineal con el PIB mensual, es decir, que podemos escribir $X = \beta U + s$, donde X es el PIB mensual, U se refiere a la serie relacionada y s es el término del error.

Como consecuencia, también el PIB trimestral admite esta relación lineal: solo

debemos multiplicar esta ecuación por la matriz

$$D = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & \dots & & & \\ \dots & & & & & & & & & \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Por lo tanto, la ecuación nos queda como:

$$DX = DU\beta + s$$

Renombrando las variables:

$$\bar{X} = \bar{U}\beta + \bar{s}$$

donde \bar{X} hace referencia al PIB trimestral. Algunos de los indicadores externos que nos aportarían información altamente relacionada con el PIB son: la producción industrial, el total de exportaciones, los índices de empleo...

4. Extensiones del filtro de Kalman

Ahora planteamos el filtro de Kalman desde un enfoque bayesiano, es decir, valorando que las observaciones Z_i afectan al estado de nuestra variable X . Esto es equivalente a decir que profundizaremos en el filtro de Kalman como una herramienta de propagación de la densidad de X condicionada a las observaciones Z_i . Veamos qué queremos decir con esto.

En una aproximación bayesiana del filtro de Kalman, el objetivo es construir el estado de una variable en base a toda la información previa conocida. Tal y como hemos estado viendo, si notamos X_k como el estado de la variable X en el instante k , entonces, este quedará supeditado a las observaciones tomadas hasta el instante k . Es decir, la probabilidad de X_k dependerá de las mediciones $Z_i, i = 1, \dots, k$. Esto se traduce en una densidad de probabilidad condicional: $f_{X_k|Z_1, \dots, Z_k}(x|z_1, \dots, z_k)$. En esta densidad queda recogida toda la información disponible acerca de X_k , e indica la probabilidad de X de alcanzar un determinado estado en el instante k en función de las observaciones tomadas.

El filtro de Kalman se encarga de la propagación de esta función de densidad, y lo hace de forma recursiva, pues a medida que avanzamos en el tiempo vamos conociendo nuevas observaciones que podemos incorporar sin reformular el modelo. Como ya habíamos comentado, esta es una de las grandes ventajas del filtro de Kalman, pues no es necesario almacenar toda la información que nos llega ya que no reprocesamos el modelo con cada dato nuevo que nos llega.

Bajo este enfoque bayesiano, y aplicando el teorema del mismo nombre, tenemos:

$$P(X_k|Z^k) = \frac{P(Z^k|X_k)P(X_k|Z^{k-1})}{P(Z^k|Z^{k-1})}$$

Como aclaración de notación decir que $P(X_k|Z_k)$ se refiere a la probabilidad de X_k conociendo el conjunto de observaciones Z^k . Esta es la base de la solución bayesiana óptima.

Hasta ahora hemos trabajado bajo la hipótesis de linealidad del modelo: la ecuación de transición es lineal ($X_{k+1} = F_k X_k + w_k$), así como la relación entre las observaciones y el estimador ($Z_k = H_k X_k + v_k$). Estos sistemas son muy deseables, ya que la teoría que disponemos para trabajar con ellos es más amplia y su complejidad menor. Ahora consideraremos el caso en que estas ecuaciones no son lineales, es decir, trabajamos con:

$$\begin{cases} X_{k+1} = f(X_k, w_k) \\ Z_k = h(X_k, v_k) \end{cases}$$

donde f o h (o ambas) no son lineales. Cuando estas funciones no pueden ser escritas de forma lineal, tal y como hemos venido haciendo hasta ahora, una solución pasa por la linealización local mediante la derivación. En esto se basa el filtro de Kalman

extendido: ahora haremos que $P_{X_k|Z_k}$ siga una distribución gaussiana -hipótesis imprescindible a lo largo de todo el trabajo- que dependa de funciones linealizadas. Estas linealizaciones que fundamentaran el filtro de Kalman extendido serán de orden 1.

Una mejora a este filtro de Kalman extendido llega con el filtro de Kalman Unscented, pues no nos basamos en la linealización, sino en la *transformación unscented*, que propaga la media y la covarianza con mayor precisión (de orden 3, a diferencia del filtro de Kalman extendido que lo hacía a orden 1), es decir, aproxima mejor la no linealidad. Esto es importante porque una distribución gaussiana queda determinada solo con los primeros dos momentos de la distribución. Si tenemos una transformación no lineal $Y = g(X)$, lo que haremos en este filtro es seleccionar cuidadosamente un conjunto de puntos del vector X cuya media y covarianza serán conocidas. A estos le aplicaremos directamente la función no lineal. La media y covarianza resultantes de la transformación serán un buen estimador de la media y covarianza de Y . En definitiva, con esto lo que conseguimos es calcular los primeros momentos de la densidad de probabilidad de una variable gaussiana Y que resulta de aplicar una transformación no lineal g .

Sin embargo, hasta ahora hemos visto como el filtro de Kalman se construye alrededor de la hipótesis fundamental de que el estado sigue una distribución gaussiana. Pero qué pasa si no es así? El filtro de partículas proporciona una buena solución. El filtro de partículas es otro filtro bayesiano recursivo basado en el método de Monte Carlo. Lo que hará es recoger un conjunto de muestras (partículas) a las que asociará unos pesos en función de como estas se ajustan a la función de densidad. Ahora entramos en mayor detalle de cada una de estas extensiones.

4.1. Filtro de Kalman extendido

Hasta ahora hemos trabajado bajo la hipótesis de linealidad del modelo: la ecuación de transición es lineal ($X_{k+1} = F_k X_k + w_k$), así como la relación entre las observaciones y el estimador ($Z_k = H_k X_k + v_k$). De nuevo consideramos el modelo:

$$\begin{cases} X_{k+1} = f(X_k, w_k) \\ Z_k = h(X_k, v_k) \end{cases}$$

con las hipótesis típicas de gaussianidad y ruido blanco, y donde f o h (o ambas) no son lineales. En este punto trabajamos con el filtro de Kalman extendido, que no hará más que utilizar una aproximación lineal de la función. Así pues, la única condición sobre f y h es que sean funciones diferenciables.

En esta hipótesis no lineal, como actualizamos en observación la matriz de covarianza del error? Es decir, cuál es la relación entre la matriz $P_{k|k}$ y $P_{k|k-1}$? Al no tratarse de un sistema lineal no podemos multiplicar por la matriz de f y/o g . Por

eso consideraremos su matriz jacobiana, es decir, definimos:

$$F_k = \frac{\partial f}{\partial x_k}, \quad H_k = \frac{\partial h}{\partial x_k}$$

Entonces, las ecuaciones de Kalman para la matriz de covarianza serán:

- Actualización por predicción

$$P_{k+1|k} = F_{k+1}P_{k|k}F_{k+1}^T + Q_k$$

donde Q_k es la matriz de covarianza del vector de aleatoriedad w_k .

- Actualización por observación

$$P_{k|k} = (I - K_k H_k)P_{k|k-1}$$

donde la ganancia de Kalman se definirá análogamente a los sistemas lineales pero con las jacobianas. Esto es:

$$K_k = P_{k|k-1}H_k^T (H_k P_{k|k-1}H_k^T + R_k)^{-1}$$

donde R_k es la matriz de covarianza del vector de aleatoriedad v_k . Por lo que respecta a la actualización de la variable X , el procedimiento será el mismo que en el caso lineal, pero usando las funciones f y h :

- Actualización por predicción:

$$\hat{X}_{k+1|k} = f(\hat{X}_{k|k}, w_k)$$

- Actualización por observación: Esto será igual que en el caso lineal pero considerando la aplicación h no lineal. Así pues nos queda:

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k (Z_k - h(\hat{X}_{k|k-1}, v_k))$$

donde K_k es la ganancia de Kalman, ya definida en la actualización por predicción.

$$K_k = P_{k|k-1}H_k^T (H_k P_{k|k-1}H_k^T + V_k R_k V_k^T)^{-1}$$

Con esto ya hemos definido todas las iteraciones cuando el modelo viene definido por un sistema no lineal al que aplicamos FKE (Filtro de Kalman extendido).

4.2. Filtro de Kalman Unscented

4.2.1. Transformación unscented

Consideramos el mismo modelo que en el caso anterior

$$\begin{cases} X_{k+1} = f(X_k, w_k) \\ Z_k = h(X_k, v_k) \end{cases}$$

bajo las hipótesis típicas de gaussianidad y ruido blanco. De nuevo suponemos que f o h (o ambas) no son lineales.

Como hemos dicho este filtro constituye una mejor aproximación de la no linealidad. Si consideramos una transformación no lineal $Y = g(X)$ seleccionaremos un conjunto de puntos de X del que conoceremos toda la información acerca la variancia y la media. Una vez transformados estos puntos directamente por el sistema no lineal, obtendremos la variancia y la media de Y con una precisión de hasta tercer orden. Veamos como debemos elegir estos puntos y cuál será el algoritmo.

Asumimos que $X \in \mathbb{R}^n$ es de media \bar{X} y con matriz de covarianzas P_X . Para seleccionar el conjunto de puntos construiremos una matriz $M \in \mathbb{M}_{n \times n}$ donde:

$$M_i = \bar{X} + \left(\sqrt{(n + \lambda)P_X} \right)_i, \quad i = 1, \dots, n$$

$$M_i = \bar{X} - \left(\sqrt{(n + \lambda)P_X} \right)_{i-n}, \quad i = n + 1, \dots, 2n$$

con $\lambda = \alpha^2(n + \kappa) - n$:

- α se refiere a la desviación de los puntos alrededor de la media \bar{X} . Normalmente adopta un valor pequeño (por ejemplo 1×10^{-3}), y nunca superior a 1.
- κ es un parámetro al que solemos otorgar el valor 0.
- $\left(\sqrt{(n + \lambda)P_X} \right)_i$ se refiere a la fila i -ésima de la matriz de covarianzas

A cada uno de estos puntos se le asigna un peso A_i . Estos se encuentran mediante el siguiente algoritmo:

$$A_i = 1/(2(n + \lambda))$$

Ahora transformamos cada uno de estos puntos mediante la función g . Sea N la matriz homóloga a M pero con la variable Y , entonces: $N_i = g(M_i)$ $i = 1, \dots, 2n$.

Y ahora ya podemos encontrar la media y la covarianza de esta variable Y .

$$\bar{Y} = \sum_{i=1}^{2n} A_i N_i$$

$$P_Y = \sum_{i=1}^{2n} A_i (N_i - \bar{Y})(N_i - \bar{Y})^t$$

Ahora veamos porque el Filtro de Kalman Unscented se fundamenta en esta transformación.

4.2.2. Filtro de Kalman Unscented

Ahora generalizamos la transformación unscented para definir el Filtro de Kalman Unscented (FKU).

Tomamos como condiciones iniciales

$$\hat{X}_0 = E[X_0]$$

y

$$P_0 = E[(X_0 - \hat{X}_0)(X_0 - \hat{X}_0)^t]$$

Ahora extendemos el vector de estado de forma que incorpore también los ruidos. Es decir, tomamos $X_k^a = [X_k^t w_k^t v_k^t]$. Ahora aplicaremos la transformación que acabamos de explicar a este vector, de forma que su correspondiente matriz es M_k^a . Siguiendo con las condiciones iniciales, establiremos:

$$\hat{X}_0^a = [\hat{X}_0^t 00]^t$$

$$P_0^a = E[(X_0^a - \hat{X}_0^a)(X_0^a - \hat{X}_0^a)^t]$$

Calculamos el conjunto de puntos propios de la transformación. Serán:

$$M_{k-1}^a = [\hat{X}_{k-1}^a \hat{X}_{k-1}^a \sqrt{(n + \lambda)P_{k-1}^a}]$$

Las ecuaciones de actualización por predicción serán:

$$M_{k|k-1}^X = f(M_{k-1}^X, M_{k-1}^w)$$

$$\hat{X}_{k|k-1} = \sum_{i=1}^{2n} A_i M_{i,k|k-1}^X$$

$$P_{k|k-1} = \sum_{i=1}^{2n} A_i [M_{i,k|k-1}^X - \hat{X}_{k|k-1}] [M_{i,k|k-1}^X - \hat{X}_{k|k-1}]^t$$

$$N_{k|k-1} = h(X_{k|k-1}, M_{k-1}^v)$$

$$\hat{Y}_{k|k-1} = \sum_{i=1}^{2n} A_i N_{i,k|k-1}$$

Las ecuaciones de actualización por observación serán:
Primero definimos:

$$P_{\hat{Y}_k, \hat{Y}_k} = \sum_{i=1}^{2n} A_i [N_{i,k|k-1} - \hat{Y}_{k|k-1}] [N_{i,k|k-1} - \hat{Y}_{k|k-1}]^t$$

$$P_{\hat{X}_k, \hat{Y}_k} = \sum_{i=1}^{2n} A_i [M_{i,k|k-1}^X - \hat{X}_{k|k-1}] [N_{i,k|k-1} - \hat{Y}_{k|k-1}]^t$$

Entonces, las fórmulas de actualización son:

$$K = P_{\hat{X}_k, \hat{Y}_k} P_{\hat{Y}_k, \hat{Y}_k}^{-1}$$

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K(Y_k - \hat{Y}_{k|k-1})$$

$$P_{k|k} = P_{k|k-1} - K P_{\hat{Y}_k, \hat{Y}_k} K^t$$

donde $X^a = [X^t \ w^t \ v^t]$, $M^a = [(M^x)^t \ (M^v)^t \ (M^n)^t]^t$

4.3. Filtro de partículas

Hasta ahora hemos trabajado bajo la hipótesis de estado gaussiano. Cómo podemos resolverlo cuando esta hipótesis no se cumple? En estos casos la mejor solución la proporciona el filtro de partículas, la cual se basa en el método de Monte Carlo.

4.3.1. Método de Montecarlo

El método de Montecarlo se encarga de resolver problemas que no tienen una solución estrictamente numérica o analítica, sino que tienen un fuerte componente aleatorio. Este método resolverá el problema mediante la simulación de la variable aleatoria y recogiendo las muestras obtenidas.

4.3.2. Filtro de partículas

En el filtro de partículas recogemos un conjunto de muestras (partículas) a las cuales asignaremos unos valores (pesos). Las partículas representan muestras del espacio de estado mientras que los pesos serán indicadores de la importancia que cada una de estas partículas adquiere. El objetivo es aproximar una función de densidad a partir de las partículas y con el uso de ecuaciones de recurrencia bayesana, que se derivan del teorema de Bayes. Más formalmente, podemos decir que el objetivo es estimar la función de densidad de probabilidad $P(X_k|Z^k)$, es decir, el estado en el instante k del vector X conociendo el conjunto de observaciones Z^k . Diferenciamos cuatro etapas: inicialización, actualización, estimación y predicción.

En la inicialización recogemos un conjunto de puntos que son muestras aleatorias de la variable de estado. En la actualización asignamos el peso de cada partícula, en función de la similitud entre el estado de la partícula y el modelo teórico. En la estimación creamos un nuevo conjunto de partículas que reemplazará al anterior. Para elegir estos nuevos puntos, tendremos en cuenta los pesos que habíamos atribuido a cada partícula, por lo que las partículas con mayor peso asociado tendrán una mayor probabilidad de encontrarse en este nuevo conjunto. Por último, en la etapa de predicción, modificamos levemente el estado de cada partícula con tal de estimar el estado del objeto en la siguiente unidad de tiempo. Volvemos a la etapa de actualización e iremos siguiendo este proceso recursivamente hasta terminar la

secuencia de datos. Entremos ahora en mayor detalle.

Con el teorema de Bayes lograremos la estimación de $P(X_k|Z_1, \dots, Z_k)$ a partir de la probabilidad condicionada $P(Z_1, \dots, Z_k|X_k)$. La fórmula nos dice que:

$$P(X_k|Z_1, \dots, Z_k) = \frac{P(Z_1, \dots, Z_k)P(X_k)}{P(Z_1, \dots, Z_k)}$$

Por tanto, lo que haremos es aproximar la densidad a posteriori, y a partir de esta aproximación lograremos tener información del estado del vector X en el instante k .

Para menor complejidad del algoritmo nos valdremos de la recursividad, es decir, queremos lograr una aproximación de $P(X_k|Z_1, \dots, Z_k)$ basándonos solo en la nueva observación z_k y en la anterior densidad a posteriori $P(X_{k-1}|Z_1, \dots, Z_{k-1})$. Por esto consideraremos al filtro de partículas un filtro bayesiano recursivo.

Tal y como hemos estado viendo en los algoritmos desarrollados hasta ahora, podemos diferenciar entre ecuaciones de predicción y de actualización. La ecuación de predicción $P(X_k|Z_{k-1})$ la deduciremos a partir de la ecuación de Chapman-Kolmogorov, por la cual tenemos:

$$P(X_k|Z_{k-1}) = \int_{R^n} P(X_k|X_{k-1})P(X_{k-1}|Z_{k-1})dX_{k-1}$$

En concreto, la predicción depende de la última densidad a posteriori estimada $P(X_{k-1}|Z_{k-1})$ y del modelo, más concretamente de la ecuación de transición.

Ahora nos fijamos en la ecuación de actualización. Para formularla nos volvemos a fijar en el teorema de Bayes, aplicándolo de forma que, por recursividad, logremos la dependencia solo respecto la observación inmediatamente anterior. Así pues, nos queda:

$$P(X_k|Z_1, \dots, Z_k) = \frac{(P(Z_k|X_k)P(X_k|Z_{k-1}))}{P(Z_k|Z_{k-1})}$$

Hemos dicho que la aproximación de la función de densidad dependerá de las muestras aleatorias tomadas y de sus respectivos pesos. El único inconveniente a esta técnica es la gran cantidad de muestras que pueden llegar a ser necesarias para obtener una estimación fiable. Este problema de recogida de muestras, las cuales conoceremos como partículas, se solucionará con el método de Montecarlo.

La técnica de Montecarlo nos permite obtener las muestras aleatorias para aproximar la función de densidad de probabilidad, ante la imposibilidad de hacerlo analíticamente. Sin embargo, hay veces que no es posible recoger muestras de la función de probabilidad $P(x)$ que queremos. Para solucionarlo tendremos que definir una nueva función de probabilidad $Q(x)$ de forma que esta no se anule en

los puntos en que $P(x)$ no lo hace. Esta es la única condición que impondremos a la nueva función $Q(x)$. Una vez recogidas las muestras aleatorias de Q deberemos corregir el error introducido al no usar la función P que debíamos. Aquí es donde intervendrán los pesos. Es evidente que las muestras de $Q(x)$ estarán concentradas en la región de Q con mayor probabilidad. Para ello definiremos

$$A_x = \frac{P(x)}{Q(x)}$$

Con esto lograremos compensar la cantidad de veces que una muestra aparece en $Q(x)$ por la cantidad de veces que debería aparecer en $P(x)$. Esto es lo que conoceremos como muestreo por importancia.

Finalmente, la función de probabilidad quedará aproximada como:

$$P(x) = \sum_{i=1}^N A^{(i)} \delta_{x^{(i)}}(x)$$

donde $\langle x^{(i)}, w^{(i)} \rangle$ con $i = 1, \dots, N$, siendo N el número de muestras. Y donde $\delta_{x^{(i)}}(x)$ se refiere a la delta de Dirac centrada en $x^{(i)}$ -es una función con valor infinito para $x = x^{(i)}$, y con valor 0 para cualquier otro valor de x . Su integral de $-\infty$ a $+\infty$ es 1.

5. ANEXO

A

Código tracking en C

```
    include <stdio.h>
include <stdlib.h>
include <math.h>

    double** prodMatriu (double **A, double **B, int filA, int colA, int colB)
double **res
; double suma = 0;
int i, j, k;
res = (double**)calloc(filA, sizeof(double*));
for(i = 0; i <filA; ++i) res[i] = (double*)calloc(colB, sizeof(double));
for(i = 0; i <filA; ++i)
for(j = 0; j <colB; ++j)
for(k = 0; k <colA; ++k)
suma = suma + A[i][k] * B[k][j];

res[i][j] = suma;
suma = 0;

return res;

    int main(void)
int n=50; //n son las iteraciones
int i, j, t;
double **nouP, **aux, **Kalmangain, **P, **F, **H, **Ftrans;
double *nouX, *Z, *X;
FILE *entrada, *sortida1, *sortida2, *sortida3, *sortida4, *sortida5, *sortida6, *sor-
tida7, *sortida8, *sortida9, *sortida10;
char fiEnt[30];
printf(Çom es diu el fitxer d'entrada?");
scanf(çchar Pred [] = "prediccions.txt";
charvPred [] = "prediccions_v.el.txt";
charaPred [] = "prediccions_a.cc.txt";
charEst [] = "estimadors.txt";
charvEst [] = "estimadors_v.el.txt";
charaEst [] = "estimadors_a.cc.txt";
charObs [] = "observacions.txt";
```

```

charvObs [] = "observacions_v.el.txt";
charaObs [] = "observacions_acc.txt";
charVar [] = "matriu_v.ariances.txt";
entrada = fopen(fiEnt, "r");
sortida1 = fopen(Pred, "w");
sortida2 = fopen(Est, "w");
sortida3 = fopen(Obs, "w");
sortida4 = fopen(Var, "w");
sortida5 = fopen(vPred, "w");
sortida6 = fopen(aPred, "w");
sortida7 = fopen(vEst, "w");
sortida8 = fopen(aEst, "w");
sortida9 = fopen(vObs, "w");
sortida10 = fopen(aObs, "w");
P = (double **)calloc(3, sizeof(double*));
F = (double **)calloc(3, sizeof(double*));
H = (double **)calloc(3, sizeof(double*));
Ftrans = (double **)calloc(3, sizeof(double*));
aux = (double **)calloc(3, sizeof(double*));
nouP = (double **)calloc(3, sizeof(double*));
Z = (double *)calloc(3, sizeof(double));
Kalmangain = (double **)calloc(3, sizeof(double*));
X = (double *)calloc(3, sizeof(double));
nouX = (double *)calloc(3, sizeof(double));
for(i = 0; i < 3; ++i) F[i] = (double *)calloc(3, sizeof(double)); Ftrans[i] = (double *)calloc(3, sizeof(double));
P[0][0] = 100;
P[1][0] = P[0][1] = 20;
P[2][0] = P[0][2] = 1;
P[1][1] = 4;
P[1][2] = P[2][1] = 0,2;
P[2][2] = 0,01;
F[0][0] = F[0][1] = F[1][1] = F[1][2] = F[2][2] = 1;
F[0][2] = 0,5;
F[1][0] = F[2][0] = F[2][1] = 0;
H[0][0] = H[1][1] = H[2][2] = 1;
H[0][1] = H[1][0] = H[1][2] = H[2][1] = H[2][0] = H[0][2] = 0;
Ftrans[0][1] = Ftrans[0][2] = Ftrans[1][2] = 0;
Ftrans[0][0] = Ftrans[1][0] = Ftrans[1][1] = Ftrans[2][1] = Ftrans[2][2] = 1;
Ftrans[2][0] = 0,5;
X[0] = 100;
X[1] = 20;
X[2] = 3;
for(t = 1; t <= n; ++t) aux = prodMatriu(F, P, 3, 3, 3); nouP = prodMatriu(aux, Ftrans, 3, 3, 3);
return 0;

```

A.2. Entrada observaciones en tracking

130 23 3
150 25 2.5
172 27 2
193 28 1.6
208 28.7 1
223 30 1.7
250 32 2
280 34 2
307 35.8 2
340 38 2
380 40 2.4
415 43 2.7
460 45 2
505 47 2
555 48 1.5
610 49 1
660 50 0.7
710 51 0.5
760 51.1 0.4
815 51.3 0.2
870 51 0
920 51 0
970 51 0
1020 51 0
1080 52 0.5
1135 53 0.5
1190 53 0.5
1250 54 0.7
1300 55 0.7
1360 56 0.8
1420 57 0.9
1480 58 1
1540 59 1
1600 60 1
1657 61 1
1710 61 0.7
1770 62 1
1830 63 1
1900 63 1
1970 63.7 1
2030 64 1
2100 65 1
2170 67 1.5
2228 68 1.5

2300 70 1.7
2375 71 1.7
2440 72 1.7
2510 74 1.9
2590 76 2
2668 78 2
2735 80 2

A.3. Predicción de la posición

1 1.215000e+02
2 1.493339e+02
3 1.766648e+02
4 2.051128e+02
5 2.342713e+02
6 2.629698e+02
7 2.916064e+02
8 3.221956e+02
9 3.548641e+02
10 3.889750e+02
11 4.251119e+02
12 4.637941e+02
13 5.043723e+02
14 5.476201e+02
15 5.932262e+02
16 6.412655e+02
17 6.917660e+02
18 7.439487e+02
19 7.975437e+02
20 8.522600e+02
21 9.082164e+02
22 9.651307e+02
23 1.022452e+03
24 1.080046e+03
25 1.137802e+03
26 1.196332e+03
27 1.255274e+03
28 1.314550e+03
29 1.374452e+03
30 1.434382e+03
31 1.494940e+03
32 1.556109e+03
33 1.617883e+03
34 1.680254e+03
35 1.743216e+03
36 1.806609e+03
37 1.870176e+03
38 1.934287e+03
39 1.998941e+03
40 2.064571e+03
41 2.131128e+03
42 2.198075e+03
43 2.265887e+03
44 2.334597e+03

45 2.403649e+03
46 2.473751e+03
47 2.545021e+03
48 2.616995e+03
49 2.689941e+03
50 2.764327e+03

A.4. Predicción en la velocidad

1 2.300000e+01
2 2.600000e+01
3 2.856564e+01
4 3.085900e+01
5 3.275057e+01
6 3.426581e+01
7 3.574789e+01
8 3.734446e+01
9 3.902031e+01
10 4.073010e+01
11 4.250852e+01
12 4.436040e+01
13 4.639474e+01
14 4.841703e+01
15 5.042986e+01
16 5.230854e+01
17 5.403591e+01
18 5.561472e+01
19 5.705510e+01
20 5.830689e+01
21 5.939156e+01
22 6.028580e+01
23 6.103680e+01
24 6.166481e+01
25 6.218683e+01
26 6.271070e+01
27 6.323756e+01
28 6.370874e+01
29 6.420098e+01
30 6.471285e+01
31 6.524868e+01
32 6.581210e+01
33 6.640616e+01
34 6.702808e+01
35 6.767542e+01
36 6.834597e+01
37 6.897177e+01
38 6.962189e+01
39 7.029436e+01
40 7.093876e+01
41 7.159071e+01
42 7.223066e+01
43 7.289209e+01
44 7.364259e+01

45 7.442954e+01
46 7.530368e+01
47 7.621361e+01
48 7.715600e+01
49 7.818002e+01
50 7.928360e+01

A.5. Predicción en la aceleración

1 3.000000e+00
2 3.000000e+00
3 2.855215e+00
4 2.656728e+00
5 2.450236e+00
6 2.203732e+00
7 2.127516e+00
8 2.110046e+00
9 2.096214e+00
10 2.085004e+00
11 2.075746e+00
12 2.108980e+00
13 2.166314e+00
14 2.150971e+00
15 2.137670e+00
16 2.083832e+00
17 1.995872e+00
18 1.894507e+00
19 1.789123e+00
20 1.687489e+00
21 1.581927e+00
22 1.472851e+00
23 1.374030e+00
24 1.284198e+00
25 1.202285e+00
26 1.158531e+00
27 1.118415e+00
28 1.081544e+00
29 1.059261e+00
30 1.038691e+00
31 1.025282e+00
32 1.018372e+00
33 1.017377e+00
34 1.016451e+00
35 1.015589e+00
36 1.014785e+00
37 9.988015e-01
38 9.988614e-01
39 9.989176e-01
40 9.989702e-01
41 9.990196e-01
42 9.990659e-01
43 9.991096e-01
44 1.022208e+00

45 1.043971e+00
46 1.073497e+00
47 1.101365e+00
48 1.127691e+00
49 1.161278e+00
50 1.197357e+00

A.6. Estimación en la posición

1 1.248339e+02
2 1.495268e+02
3 1.755822e+02
4 2.027459e+02
5 2.298059e+02
6 2.569223e+02
7 2.859062e+02
8 3.168919e+02
9 3.492874e+02
10 3.836413e+02
11 4.204882e+02
12 4.590607e+02
13 5.002785e+02
14 5.438652e+02
15 5.899988e+02
16 6.387281e+02
17 6.892813e+02
18 7.413832e+02
19 7.947969e+02
20 8.496158e+02
21 9.055813e+02
22 9.621027e+02
23 1.019023e+03
24 1.076216e+03
25 1.134200e+03
26 1.192596e+03
27 1.251382e+03
28 1.310780e+03
29 1.370189e+03
30 1.430204e+03
31 1.490807e+03
32 1.551985e+03
33 1.613734e+03
34 1.676049e+03
35 1.738770e+03
36 1.801703e+03
37 1.865165e+03
38 1.929146e+03
39 1.994132e+03
40 2.060037e+03
41 2.126343e+03
42 2.193494e+03
43 2.261465e+03
44 2.329741e+03

45 2.398984e+03
46 2.469358e+03
47 2.540403e+03
48 2.612342e+03
49 2.685642e+03
50 2.760227e+03

A.7. Estimación en la velocidad

1 2.300000e+01
2 2.571043e+01
3 2.820227e+01
4 3.030033e+01
5 3.206207e+01
6 3.362038e+01
7 3.523442e+01
8 3.692409e+01
9 3.864510e+01
10 4.043278e+01
11 4.225142e+01
12 4.422843e+01
13 4.626606e+01
14 4.829219e+01
15 5.022471e+01
16 5.204003e+01
17 5.372021e+01
18 5.526598e+01
19 5.661940e+01
20 5.780964e+01
21 5.881295e+01
22 5.966277e+01
23 6.038061e+01
24 6.098455e+01
25 6.155217e+01
26 6.211915e+01
27 6.262719e+01
28 6.314172e+01
29 6.367416e+01
30 6.422340e+01
31 6.479373e+01
32 6.538878e+01
33 6.601163e+01
34 6.665983e+01
35 6.733118e+01
36 6.797296e+01
37 6.862303e+01
38 6.929544e+01
39 6.993979e+01
40 7.059169e+01
41 7.123159e+01
42 7.189298e+01
43 7.262038e+01
44 7.338557e+01

45 7.423018e+01
46 7.511224e+01
47 7.602830e+01
48 7.701874e+01
49 7.808624e+01
50 7.922897e+01

A.8. Estimación en la aceleración

1 3.000000e+00
2 2.855215e+00
3 2.656728e+00
4 2.450236e+00
5 2.203732e+00
6 2.127516e+00
7 2.110046e+00
8 2.096214e+00
9 2.085004e+00
10 2.075746e+00
11 2.108980e+00
12 2.166314e+00
13 2.150971e+00
14 2.137670e+00
15 2.083832e+00
16 1.995872e+00
17 1.894507e+00
18 1.789123e+00
19 1.687489e+00
20 1.581927e+00
21 1.472851e+00
22 1.374030e+00
23 1.284198e+00
24 1.202285e+00
25 1.158531e+00
26 1.118415e+00
27 1.081544e+00
28 1.059261e+00
29 1.038691e+00
30 1.025282e+00
31 1.018372e+00
32 1.017377e+00
33 1.016451e+00
34 1.015589e+00
35 1.014785e+00
36 9.988015e-01
37 9.988614e-01
38 9.989176e-01
39 9.989702e-01
40 9.990196e-01
41 9.990659e-01
42 9.991096e-01
43 1.022208e+00
44 1.043971e+00

45 1.073497e+00
46 1.101365e+00
47 1.127691e+00
48 1.161278e+00
49 1.197357e+00
50 1.231519e+00

B

Kalman en Matlab (para estimación del canal y RLS)

```
function [ m_n, P_n ] = kalmanfilter(m_n1, P_n1, y_n, F_n, Q_n, H_n, R_n)
m_n n1 = F_n * m_n1;
P_n n1 = F_n * P_n1 * F_n' + Q_n;
S_n = H_n * P_n n1 * H_n' + R_n;
K_n = P_n n1 * H_n' / S_n;
m_n = m_n n1 + K_n * (y_n - H_n * m_n n1);
P_n = P_n n1 - K_n * H_n * P_n n1;
end
```

C

Código para aplicación 3.2. en Matlab

```
alpha = [0.85; 1.001; -0.95]
F_n = diag(alpha);
sigma_x = 0,1;
Q_n = sigma_x^2 * eye(3);
h = randn(501, 1);
sigma_y = 0,1;
R_n = sigma_y^2;

x (1:3,1) = randn (3,1) ;
m (1:3,1)= x (1:3,1) ;
P (1:3,1:3,1) = eye (3);

for n = 2 : 500 ,

    v_n sigma_x * randn(3, 1);
x(1 : 3, n) = F_n * x(1 : 3, n - 1) + v_n;

    H_n = h(n + 1 : 1 : n - 1, 1)';
w_n = sigma_y * randn(1, 1);
y_n = H_n + x(1 : 3, n) + w_n;

    [m (1:3,n) , P (1:3,1:3,n)] = kalman_filter(m(1 : 3, n - 1),
P(1 : 3, 1 : 3, n - 1), y_n, F_n, Q_n, H_n, R_n);
End
```

D

Código para aplicación 3.3. en Matlab

```
D = 32;
lambda = 0,99 ;
Fn = 1/sqrt(lambda);
Qn = zeros(D, D);
h = 10 * randn(500 + D - 2, 1);
sigmay = 1;
Rn = sigmay2;

x(1:D,1) = fir1(D-1, 0,5);
m(1:D,1) = zeros(D,1);
P(1:D,1:D,1) = eye(D) ;

for n = 2 : 500 ,

    x(1:D,n) = Fn * x(1:D,n-1);

    Hn = h(n + D - 2 : -1 : n - 1, 1)';
    w - n = sigma - y * randn(1, 1);
    yn = Hn * x(1:D,n) + wn;

    [m(1:D,n), P(1:D,1:D,n)] = kalman_filter(m(1:D,n-1),
    P(1:D,1:D,n-1), yn, Fn, Qn, Hn, -Rn);
End
```

E

Salida estimación mensual del PIB Uruguay

Mes	PBI	Mes	PBI	Mes	PBI	Mes	PBI
M1 83	60022.191	M1 88	73618.642	M1 93	80554.868	M1 98	98710.894
M2 83	60022.191	M2 88	81399.078	M2 93	87976.57	M2 98	113374.51
M3 83	60022.191	M3 88	63686.115	M3 93	75246.454	M3 98	87087.676
M4 83	60022.191	M4 88	78186.758	M4 93	85968.622	M4 98	109119.5
M5 83	60022.191	M5 88	64784.438	M5 93	76058.584	M5 98	88756.434
M6 83	60079.87	M6 88	72431.155	M6 93	79296.522	M6 98	100263.02
M7 83	60044.417	M7 88	64520.812	M7 93	75251.051	M7 98	88327.596
M8 83	60077.185	M8 88	71832.001	M8 93	78990.106	M8 98	99358.994
M9 83	56965.137	M9 88	70007.606	M9 93	84896.229	M9 98	100940.51
M10 83	59274.039	M10 88	72570.861	M10 93	81101.924	M10 98	101572.69
M11 83	57140.021	M11 88	70201.755	M11 93	84608.836	M11 98	100988.39
M12 83	71553.886	M12 88	83105.21	M12 93	95032.743	M12 98	116356.82
M1 84	60435.399	M1 89	73083.915	M1 94	87827.919	M1 99	104788.79
M2 84	70711.737	M2 89	82346.167	M2 94	94487.027	M2 99	115480.62
M3 84	55554.377	M3 89	64169.829	M3 94	75157.889	M3 99	92383.512
M4 84	68558.183	M4 89	79259.865	M4 94	90680.692	M4 99	111381.65
M5 84	56539.325	M5 89	65312.794	M5 94	76333.633	M5 99	93822.49
M6 84	59195.51	M6 89	73954.961	M6 94	90630.511	M6 99	103128.66
M7 84	55237.81	M7 89	65212.033	M7 94	77600.804	M7 99	93296.521
M8 84	58895.741	M8 89	73292.744	M8 94	89643.601	M8 99	102383.94

M8 84	58895.741	M8 89	73292.744	M8 94	89643.601	M8 99	102383.94
M9 84	57944.067	M9 89	70124.352	M9 94	87944.768	M9 99	96762
M10 84	59255.548	M10 89	73817.307	M10 94	91191.813	M10 99	102451.87
M11 84	58043.402	M11 89	70404.068	M11 94	88190.709	M11 99	97192.968
M12 84	68941.974	M12 89	84232.386	M12 94	100847.96	M12 99	115759.64
M1 85	60607.374	M1 90	73349.135	M1 95	90906.594	M1 00	101036.08
M2 85	68310.686	M2 90	83408.056	M2 95	100094.97	M2 00	114644.43
M3 85	58669.901	M3 90	64825.969	M3 95	81325.304	M3 00	88653.013
M4 85	67132.263	M4 90	80349.795	M4 95	96846.098	M4 00	110290.91
M5 85	59310.867	M5 90	66001.791	M5 95	82500.897	M5 00	90291.934
M6 85	57418.621	M6 90	70320.332	M6 95	93249.361	M6 00	95000.664
M7 85	57545.873	M7 90	64738.962	M7 95	82868.73	M7 00	88199.438
M8 85	57428.259	M8 90	69897.582	M8 95	92463.1	M8 00	94485.518
M9 85	58450.526	M9 90	74592.899	M9 95	83357.191	M9 00	93274.979
M10 85	57668.111	M10 90	71935.69	M10 95	91731.033	M10 00	95211.561
M11 85	58391.264	M11 90	74391.634	M11 95	83991.451	M11 00	93421.661
M12 85	72764.052	M12 90	86646.523	M12 95	105026.56	M12 00	110517.59
M1 86	62154.082	M1 91	77902.233	M1 96	88052.76	M1 01	97461.932
M2 86	71960.42	M2 91	85984.204	M2 96	103740.91	M2 01	109528.72
M3 86	58867.127	M3 91	63804.553	M3 96	77857.795	M3 01	89037.436
M4 86	70252.248	M4 91	81688.978	M4 96	99756.614	M4 01	106316.35
M5 86	59729.471	M5 91	65159.174	M5 96	79516.479	M5 01	90346.194

M5 86	59729.471	M5 91	65159.174	M5 96	79516.479	M5 01	90346.194
M6 86	65840.905	M6 91	73962.704	M6 96	90456.288	M6 01	93015.27
M7 86	59549.778	M7 91	64674.918	M7 96	78964.144	M7 01	88398.61
M8 86	65364.395	M8 91	73259.219	M8 96	89585.838	M8 01	92665.59
M9 86	66219.285	M9 91	74553.94	M9 96	93147.093	M9 01	87824.426
M10 86	66536.621	M10 91	74998.083	M10 96	92234.138	M10 01	92139.36
M11 86	66243.321	M11 91	74587.581	M11 96	93077.943	M11 01	88151.252
M12 86	79277.165	M12 91	88679.227	M12 96	107534.06	M12 01	104918.01
M1 87	69499.644	M1 92	78092.833	M1 97	96881.713	M1 02	91746.848
M2 87	78536.586	M2 92	87877.381	M2 97	106727.22	M2 02	103920.39
M3 87	65413.809	M3 92	66676.199	M3 97	83357.225	M3 02	86977.147
M4 87	76694.508	M4 92	84110.005	M4 97	102419.99	M4 02	101625.11
M5 87	66268.244	M5 92	67996.689	M5 97	84801.098	M5 02	88086.629
M6 87	71615.368	M6 92	77235.028	M6 97	102123.88	M6 02	97495.122
M7 87	65910.632	M7 92	67691.111	M7 97	86297.796	M7 02	88247.298
M8 87	71183.274	M8 92	76512.143	M8 97	100925.16	M8 02	96794.664
M9 87	70897.407	M9 92	77882.07	M9 97	94264.34	M9 02	77342.134
M10 87	71977.22	M10 92	78308.975	M10 97	101640.05	M10 02	93267.331
M11 87	70979.195	M11 92	77914.405	M11 97	94822.999	M11 02	78548.357
M12 87	82036.686	M12 92	88584.78	M12 97	114576.2	M12 02	100516.32

F

Esperanza condicionada

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad y consideremos dos variables aleatorias discretas X, Y entonces,

$$\mathbb{E}(X|Y = y) := \sum_x xP(X = x|Y = y) \quad (\text{suponemos } P(Y = y) > 0)$$

y podemos definir, a partir de aquí, la variable

$$\mathbb{E}(X|Y)(\omega) = \mathbb{E}(X|Y = y), \text{ si } Y(\omega) = y$$

o lo que es lo mismo

$$\mathbb{E}(X|Y) = \sum_y \mathbb{E}(X|Y = y)\mathbf{1}_{\{Y=y\}}$$

Sea ahora $A \in \mathcal{F}$ (suponemos $P(A) > 0$)

$$\mathbb{E}(X|A) := \sum xP(X = x|A),$$

consideremos la σ -álgebra generada por un partición, (A_i) , de Ω , llamémosle \mathcal{A} , esto es $\mathcal{A} = \{\emptyset, \Omega, \cup_j A_{i_j}\}$, podemos ahora definir la variable aleatoria

$$\mathbb{E}(X|\mathcal{A})(\omega) = \mathbb{E}(X|A_i), \text{ si } \omega \in A_i,$$

o equivalentemente

$$\mathbb{E}(X|\mathcal{A}) = \sum_i \mathbb{E}(X|A_i)\mathbf{1}_{A_i}.$$

Notemos que $\mathbb{E}(X|\mathcal{A})$ es una variable aleatoria \mathcal{A} -medible:

$$\mathbb{E}(X|\mathcal{A})^{-1}(z_i) = A_i \in \mathcal{A} \text{ si } z_i = \mathbb{E}(X|A_i),$$

además para todo $A \in \mathcal{A}$

$$\mathbb{E}(X\mathbf{1}_A) = \mathbb{E}(\mathbb{E}(X|\mathcal{A})\mathbf{1}_A).$$

En efecto, si tomamos $A = A_i$

$$\mathbb{E}(X|\mathcal{A})\mathbf{1}_A = \mathbb{E}(X|A_i)\mathbf{1}_{A_i},$$

de manera que

$$\begin{aligned} \mathbb{E}(\mathbb{E}(X|\mathcal{A})\mathbf{1}_A) &= \mathbb{E}(\mathbb{E}(X|A_i)\mathbf{1}_{A_i}) = \mathbb{E}(X|A_i)P(A_i) \\ &= \sum_x xP(X = x|A_i)P(A_i) \\ &= \sum_x xP(X = x, A_i) = \mathbb{E}(X\mathbf{1}_{A_i}). \end{aligned}$$

Esto sugiere la siguiente definición general. Consideremos una variable aleatoria integrable X y sea \mathcal{A} una σ -álgebra contenida en \mathcal{F} , definimos la esperanza condicionada de X dado \mathcal{A} a una variable aleatoria, digamos Z , con las siguientes propiedades

- i)* Z es \mathcal{A} -medible
- ii)* Para todo $A \in \mathcal{A}$, $\mathbb{E}(Z\mathbf{1}_A) = \mathbb{E}(X\mathbf{1}_A)$.

Se puede demostrar que tal variable siempre existe y es única casi seguramente.

Propiedades

1. Linealidad: $\mathbb{E}(aX + bY|\mathcal{A}) = a\mathbb{E}(X|\mathcal{A}) + b\mathbb{E}(Y|\mathcal{A})$, $a, b \in \mathbb{R}$.
2. $\mathbb{E}(\mathbb{E}(X|\mathcal{A})) = \mathbb{E}(X)$. Basta tomar $A = \Omega$ en *ii*).
3. Si X y \mathcal{A} son independientes $\mathbb{E}(X|\mathcal{A}) = \mathbb{E}(X)$. De hecho la constante $\mathbb{E}(X)$ es trivialmente \mathcal{A} -medible, y para todo $A \in \mathcal{A}$, tenemos $\mathbb{E}(X\mathbf{1}_A) = \mathbb{E}(X)\mathbb{E}(\mathbf{1}_A) = \mathbb{E}(\mathbb{E}(X)\mathbf{1}_A)$.
4. Si X es \mathcal{A} -medible, entonces $\mathbb{E}(X|\mathcal{A}) = X$. Trivial a partir de la definición.
5. Si Y está acotada y es \mathcal{A} -medible, entonces $\mathbb{E}(YX|\mathcal{A}) = Y\mathbb{E}(X|\mathcal{A})$. De hecho la variable aleatoria $Y\mathbb{E}(X|\mathcal{A})$ es integrable any \mathcal{A} -medible, y para todo $A \in \mathcal{A}$ tenemos

$$\mathbb{E}(\mathbb{E}(X|\mathcal{A})Y\mathbf{1}_A) = \mathbb{E}(XY\mathbf{1}_A),$$

donde la igualdad se sigue de *ii*) y del teorema de convergencia dominada. Esta propiedad significa que las variables \mathcal{A} -medibles se comportan como constantes y pueden ser factorizadas fuera de la esperanza condicionada con respecto a \mathcal{A} . Esta propiedad es cierta siempre que $\mathbb{E}(|XY|) < \infty$. En particular si $X, Y \in L^2(\Omega)$.

6. En particular si $X, Y \in L^2(\Omega)$, por la propiedad anterior con $A = \Omega$, tenemos que si Y es \mathcal{A} -medible $X - \mathbb{E}(X|\mathcal{A})$ es ortogonal a Y :

$$\mathbb{E}((X - \mathbb{E}(X|\mathcal{A}))Y) = \mathbb{E}(XY) - (\mathbb{E}(X|\mathcal{A})Y) = 0,$$

7. (Propiedad de torre) Dadas dos σ -álgebras $\mathcal{B} \subset \mathcal{A}$, entonces $\mathbb{E}(\mathbb{E}(X|\mathcal{A})|\mathcal{B}) = \mathbb{E}(X|\mathcal{B})$. En efecto, sea

$$B \in \mathcal{B},$$

tenemos que ver que

$$\mathbb{E}(\mathbb{E}(X|\mathcal{A})\mathbf{1}_B) = \mathbb{E}(X\mathbf{1}_B),$$

pero como $\mathcal{B} \subset \mathcal{A}$, $B \in \mathcal{A}$ y la igualdad es cierta por definición de $\mathbb{E}(X|\mathcal{A})$.

Referencias

- [1] HIRSA, A., *Computational Methods in Finance*. Chapman Hall, 2012.
- [2] ANDERSON, B.; MOORE, J. *Optimal filtering*. Prentice-Hall, 1979.
- [3] KOVVALI, N.; BANAVAR, M.; SPANIAS, A. *An Introduction to Kalman Filtering with MATLAB examples*. Morgan Claypool Publishers, 2013.
- [4] BROCKWELL, P.; DAVIS, R. *State-Space Models and the Kalman Recursions*. Springer, 1991.
- [5] DURRANT-WHYTE, H. *Introduction to Estimation and the Kalman Filter*. The University of Sidney, 2001.
- [6] GÓMEZ BRUQUE, J.L. *Filtros de Kalman Extendido y Unscented.^{en} Sistemas Estocásticos no Lineales con Observaciones Inciertas*. Universidad de Granada, Trabajo fin de máster, 2011.
- [7] CASTAÑEDA, J.A.; NIETO, M.A.; ORTIZ, V.A. *Análisis y aplicación del filtro de Kalman a una señal con ruido aleatorio* Universidad Tecnológica de Pereira, 2013.
- [8] CAMACHO, M; PÉREZ-QUIROS, G *Ñ-sting: España short term indicator of growth* Banco de España, 2009.
- [9] CAMACHO, M.; PÉREZ-QUIROS, G *Introducing the euro-sting: short term indicator of euro area growth* Banco de España, 2008.
- [10] VÁZQUEZ, A. *Estimaciones mensuales del PBI mediante el filtro de Kalman: Evidencia de Uruguay* Universidad de la República, 2003.
- [11] CORCUERA, J.M. *Modelización: apuntes* Universitat de Barcelona
- [12] MAYBECK, P. *Stochastic models, estimation, and control. Volume 1* Academic Press, 1979.
- [13] RODRÍGUEZ, A. *Estudio del filtro de partículas aplicado al seguimiento de objetos en secuencias de imágenes* Universidad Carlos III de Madrid, 2009.
- [14] WAN, E.; VAN DER MERWE, R. *The Unscented Kalman Filter for Nonlinear Estimation* Oregon Graduate Institute of Science Technology, 2003.