

**Técnicas Recursivas Para Estimación Dinámica
Una Introducción Matemática al Filtro Kalman**

Guillermo Prado Obando
gprado@etb.net.co

**Trabajo Para Optar por el Título de
Matemático**

Director: Pervys Rengifo
Ingeniero Civil Universidad Nacional de Colombia.

Fundación Universitaria Konrad Lorenz
Facultad de Matemáticas
Decanatura
Bogotá D.C 02 de Junio de 2005

Resumen

El filtro de Kalman es un conjunto de ecuaciones matemáticas que proveen una solución recursiva eficiente del método de mínimos cuadrados. Esta solución permite calcular un estimador lineal, insesgado y óptimo del estado de un proceso en cada momento del tiempo con base en la información disponible en el momento $t-1$, y actualizar, con la información adicional disponible en el momento t , dichas estimaciones. Este filtro es el principal algoritmo para estimar sistemas dinámicos especificados en la forma de estado-espacio

Abstract

The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) solution of the least-squares method. The goal is to find unbiased minimum variance lineal estimator of the state at time t with base in available information at time $t-1$ and update with the additional available information at time t that estimator. This filter is the principal algorithm to estimate dynamic systems specified in state-space form.

Indice General.

Introducción

1. Marco Conceptual.

1.1 Probabilidad

1.1.1 Variables Aleatorias

1.1.2 Media y varianza.

1.1.3 Distribución Normal o Gaussina.

1.1.4 Continuidad, independencia y probabilidad condicional.

1.1.5 Señales espaciales Vs. Señales espectrales

1.2 Estimación Dinámica.

1.3 Técnicas Recursivas Para la Estimación Dinámica.

1.4 Estimación Estocástica.

1.4.1 Modelos Estado Espacio

2. El Filtro Kalman.

2.1 Filtro Kalman Discreto

2.2 Ejemplo

2.5 Aplicaciones.

Introducción

Este trabajo se realiza en el área de la estadística, dentro de los estudios sobre estimación dinámica los cuales han tomado mucha fuerza en los últimos años, gracias al avance de la computación como herramienta de soporte para guardar información y realizar cálculos, y a la necesidad creciente de métodos de estimación para realizar predicciones en tiempo real. Dentro de las herramientas matemáticas que pueden ser usadas para estimación estocástica de medidas de ruido en sensores, se encuentra una muy conocida denominada el filtro Kalman. La importancia de estudiar el algoritmo de Kalman, radica en que se constituye en el principal procedimiento para estimar sistemas dinámicos representados en la forma de estado - espacio, los cuales tienen muchas aplicaciones de interés.

El filtro Kalman fué presentado hace aproximadamente 30 años, pero su uso se ha propagado recientemente en una variedad de aplicaciones. Estas aplicaciones van desde la simulación de instrumentos musicales en realidad virtual, hasta a la extracción de secuencias de movimiento de los labios de oradores en vídeo conferencias. El filtro Kalman es el mejor estimador posible para una amplia clase de problemas y un muy efectivo estimador para una clase aún mayor. El filtro Kalman es muy fácil de entender y de usar con pocas herramientas conceptuales. El filtro de Kalman es un algoritmo recursivo que estima el estado no observable de un sistema dinámico (con expectativas racionales en este caso) dado un conjunto de observaciones que proporcionan información acerca de dicho estado en cada instante.

El presente trabajo de grado pretende realizar una introducción de los principios matemáticos en los que se fundamenta el filtro Kalman, ilustrando esta técnica mediante un ejemplo, así como también identificar las áreas de interés en las que se aplica. Se espera que esta investigación sea el inicio de nuevos proyectos que amplíen y profundicen en esta técnica, para aplicarla a la solución de problemas reales.

Este trabajo se realiza en dos partes. En la primera parte se presentan los fundamentos matemáticos para entender el funcionamiento del filtro Kalman, entre otros se tratan: conceptos fundamentales de probabilidad, estimación dinámica, los sistemas lineales de ecuaciones diferenciales y los sistemas de ecuaciones en diferencias.

La segunda parte se enfoca en la explicación del Filtro Kalman, empezando con una reseña histórica y el desarrollo del modelo matemático. En la parte final se presenta la parte algorítmica para su posible uso en el desarrollo de programas de cómputo, finalizando con la presentación de un ejemplo en donde el uso del filtro Kalman es aplicado.

1. Marco Conceptual.

1.1 Probabilidad de un Evento.

La noción de lo que significa que un evento tenga una ocurrencia al azar es muy concisa, o la probabilidad de que ocurra cierto evento en un espacio de muestra. Formalmente, la probabilidad que el resultado de un evento discreto (por ejemplo el lanzamiento de una moneda) favorecerá un evento particular está definida por;

$$p(A) = \frac{\text{Posibles resultados favorables del evento } A}{\text{Total de posibles resultados}}$$

En el caso de eventos independiente la probabilidad que ocurra A o B está dada por:

$$p(A \cup B) = p(A) + p(B), \quad 1.1.1$$

si la probabilidad de dos resultados es independiente, es decir una no afecta a otra, entonces la probabilidad de que ambos ocurran está dada por:

$$p(A \cap B) = p(A) \cdot p(B) \quad 1.1.2$$

Si A y B son dos sucesos, la probabilidad de que ocurra B dado A se denota por $p(B \mid A)$. y es llamada probabilidad condicional y está dada por:

$$p(B \mid A) = \frac{p(A \cap B)}{p(A)} \quad 1.1.3$$

1.1.1 Variables Aleatorias.

Una variable aleatoria es básicamente una función que mapea todos los puntos en el espacio de muestra a los números reales. Las variables aleatorias pueden ser discretas, cuando se pueden contar el conjunto de sus resultados posibles, o continuas, cuando toma valores en una escala continua.

Por ejemplo la variable aleatoria continua $X(t)$ puede mapear tiempo a posición. En cualquier punto en el tiempo, $X(t)$ nos dirá la posición esperada. En el caso de variables aleatorias continuas, la probabilidad de cualquier evento simple discreto A es de hecho 0. Esto es, $P(A) = 0$. Una función común que representa la probabilidad de una variable aleatoria, es definida como la función de distribución acumulativa. Esta función representa la probabilidad acumulativa de las variables aleatorias continuas X para todos (los no contables) eventos hasta x .

$$F_X(x) = P(-\infty, x]. \quad 1.1.1.1$$

Algunas propiedades importantes de las funciones de distribución acumulativa son:

$$F_X(x) \rightarrow 0 \text{ cuando } x \rightarrow -\infty$$

$$F_X(x) \rightarrow 1 \text{ cuando } x \rightarrow +\infty$$

$F_X(x)$ es una función no decreciente de x

la ecuación 1.1.1.1 es usada más comúnmente en forma de su derivada y se conoce como la función de densidad de probabilidad

$$f_X(x) = \frac{d}{dx} F_X(x), \quad 1.1.1.2$$

la función de densidad también tiene las siguientes propiedades:

$f_X(x)$ es una función no negativa

$$\int_{-\infty}^{+\infty} f_X(x) dx = 1$$

$$\begin{aligned} P(a < X < b) &= \int_a^b f_X(x) dx \\ &= P(X < b) - P(X < a) \end{aligned}$$

1.1.2 Media y varianza.

Es muy familiar el concepto de promedio de una secuencia de números para algún espacio muestral N (*Se define espacio muestral como todos los resultados posibles de un experimento estadístico*) de una variable aleatoria discreta X . En lo sucesivo se hará referencia a este valor como la media de la variable aleatoria X , y se expresa como μ_X o simplemente μ cuando se sabe a qué variable aleatoria se refiere. Es común referirse a esta media como el valor esperado de la variable aleatoria X , y se expresa como $E(X)$. El promedio o la media está dada por:

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_N}{N}$$

si X es una variable aleatoria discreta con distribución de probabilidad $f(x)$ la media o valor esperado de X es:

$$\mu = E(X) = \sum_x xf(x) \quad 1.1.2.1$$

el el caso de X continua es:

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x).dx \quad 1.1.2.2$$

si X es una variable aleatoria discreta con distribución de probabilidad $f(x)$ y media μ , la varianza de X se expresa como σ^2 y se define de la siguiente forma:

$$\sigma^2 = E[(X - \mu)^2] = \sum_x (x - \mu)^2 f(x), \quad 1.1.2.3$$

cuando la variable aleatoria X es continua:

$$\sigma^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \quad 1.1.2.4$$

La varianza es una propiedad estadística muy usada en las señales aleatorias, porque si se supiera la varianza de una señal que de otro modo fuera supuesta para ser *constante* al rededor de un valor, la magnitud de la varianza daría una ponderación de cuanto *ruido* hay en la señal.

Si X es continua la raíz cuadrada de la varianza $\sqrt{\sigma^2} = \sigma$, se llama desviación estándar de X .

1.1.3 Distribución Normal o Gaussina.

Esta distribución de probabilidad ha tenido un uso muy extendido en modelamiento de sistemas aleatorios por muchas razones. Muchos procesos de la naturaleza parecen tener una distribución normal o muy cercana a esta. En efecto, bajo ciertas condiciones, se puede probar que una suma de variables aleatorias con cualquier distribución tiende hacia una distribución normal. El teorema que define esta propiedad se llama el teorema del límite central. La distribución normal tiene algunas propiedades que la hacen matematicamente manejable y hasta interesante.

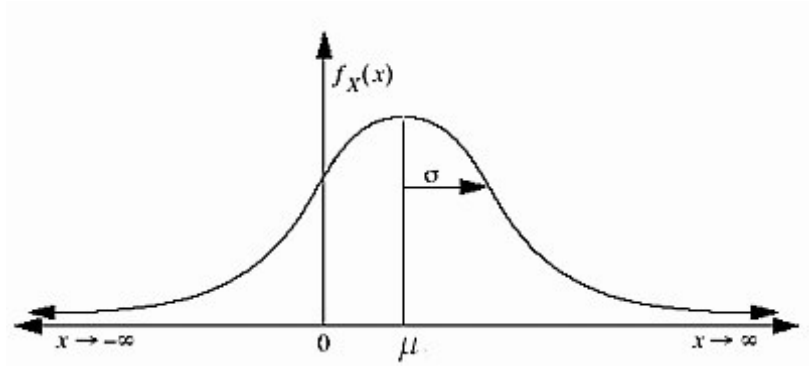


Figura 1.1.3.1: Función de distribución de Probabilidad Normal o Gaussiana

Una variable aleatoria continua X que tiene la distribución en forma de campana de la figura (1.1.3.1) se llama variable aleatoria normal. La ecuación matemática para la distribución de probabilidad de la variable normal depende de los parámetros μ y σ . Por lo tanto se representan los valores de densidad de X por $n(x; \mu, \sigma)$.

la función de densidad de probabilidad para una variable aleatoria normal X , con media μ y varianza σ^2 está dada por:

$$n(x; \mu, \sigma) = f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}} \text{ para } -\infty < x < \infty \quad 1.1.3.1$$

Donde $\pi = 3.14159..$ y $e = 2.71828$

Cualquier función lineal de un proceso aleatorio normalmente distribuido también es un proceso aleatorio normalmente distribuido. En particular si $X \sim N(\mu, \sigma^2)$ y $Y = aX + b$, entonces

$$Y \sim N(a\mu + b, a^2\sigma^2) \quad 1.1.3.2$$

la función de densidad de probabilidad para Y es dada por:

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

Finalmente si X_1 y X_2 son independientes, $X_1 \sim N(\mu_1, \sigma_1^2)$ y $X_2 \sim N(\mu_2, \sigma_2^2)$ entonces:

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2) \quad 1.1.3.3$$

$$f_X(X_1 + X_2) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} e^{-\frac{1}{2} \frac{(x - (\mu_1 + \mu_2))^2}{(\sigma_1^2 + \sigma_2^2)}} \quad 1.1.3.4$$

1.1.4 Continuidad, Independencia y Probabilidad Condicional.

Se puede ver que en las ecuaciones (1.1.2) y (1.1.3), la probabilidad está definida para variables aleatorias continuas. Dos variables aleatorias continuas X y Y son llamadas estadísticamente independientes si su probabilidad conjunta $f_{XY}(x,y)$ es igual a la suma de sus probabilidades marginales, es decir, se considera independiente si $f_{XY}(x,y) = f_X(x) f_Y(y)$.

Definition 1.1.4.1 *Adicionalmente la regla de Bayes tomando la ecuación (1.1.3) ofrece una manera de especificar la densidad de probabilidad condicional de una variable aleatoria X dada (en la presencia de) Y , Variable aleatoria. En este caso la Regla de Bayes está dada por:*

$$f_{X|Y}(x) = \frac{f_{Y|X}(y) f_X(x)}{f_Y(y)}$$

Definition 1.1.4.2 *Dado un proceso discreto X y un proceso continuo Y , la función discreta de masa de probabilidad para X condicionado sobre $Y = y$ y está dada por*

$$P_X(x | Y = y) = \frac{f_Y(y | X = x) P_X(x)}{\sum_z f_Y(y | X = z) P_X(z)} \quad 1.1.4.1$$

1.1.5 Señales Espaciales Vs. Señales Espectrales.

Se puede decir que la magnitud de la varianza de una señal indicar cuanto "ruido" tiene la señal. Sin embargo, la varianza de una señal no dice nada acerca del espaciamiento o la tasa de cambio con el paso del tiempo.

Una característica útil tiempo-relación de una señal aleatoria es su autocorrelación R_X , su relación con sí misma a través del tiempo. Formalmente se puede definir la autocorrelación de una señal aleatoria $X(t)$ de la siguiente forma:

$$R_X(t_1, t_2) = E[X(t_1)X(t_2)] \quad 1.1.5.1$$

Para los tiempos de muestra t_1 y t_2 . Si el proceso es estacionario, es decir que la función de densidad no varía con el tiempo, la ecuación (1.1.5.1) sólo depende de la diferencia $\tau = t_1 - t_2$.

En este caso, la autocorrelación se puede escribir como.

$$R_X(\tau) = E[X(t)X(t + \tau)] \quad 1.1.5.2$$

Dos funciones de autocorrelación se muestran en la figura (1.1.5.1), se observa que la señal aleatoria X_2 respecto a la señal X_1 es relativamente más pequeña y más ancha. Mientras $|\tau|$ aumenta (cuando se mueve lejos de $\tau = 0$ en el centro de la curva) la señal de autocorrelación hará X_2 baja relativamente rápido. Esto indica que X_2 está menos correlacionada con si misma que X_1 .

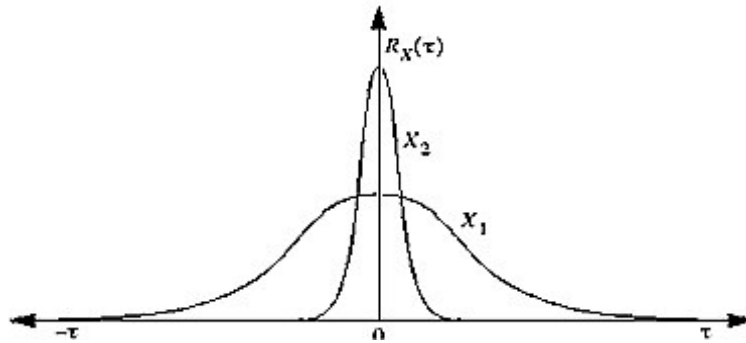


Figura 1.1.5.1: Ejemplos de funciones de autocorrelación X_1 y X_2

Se puede ver que la autocorrelación es una función del tiempo. Esto significa que tiene una interpretación espectral también en el dominio de la frecuencia. Nuevamente para un proceso estacionario, existe una importante relación temporal-espectral conocida como la relación o el teorema Wiener-Khinchine:

$$S_X(jw) = \mathfrak{F}[R_X(\tau)] = \int_{-\infty}^{\infty} R_X(\tau) e^{-jw\tau} d\tau$$

Donde $\mathfrak{F}[\bullet]$ indica la transformada de Fourier, y w indica 2π ciclos por segundo. La función $S_X(jw)$ es llamada la energía de la densidad espectral de la señal aleatoria. Como se puede ver, esta relación enlaza ambas representaciones del espectro de tiempo y frecuencia de la misma señal.

El ruido Blanco.

Un importante caso de una señal aleatoria es cuando la función de autocorrelación es una función *Dirac delta* $\delta(\tau)$ que tiene valor cero en todas partes excepto cuando $\tau = 0$. En otras palabras donde:

$$R_X(\tau) = \left\{ \begin{array}{l} A, \text{ si } \tau = 0 \\ 0, \text{ en otro caso} \end{array} \right\},$$

para alguna magnitud constante A . En este caso especial donde la autocorrelación es un "pulso", la transformada de Fourier resulta en un espectro de frecuencia *constante*, es decir, todas las frecuencias son igualmente importantes o tienen la misma potencia espectral, y están completamente no auto correlacionadas, excepto en el tiempo actual cuando $\tau = 0$ como se muestra en la figura 1.1.5.2. Esta es una descripción de **ruido blanco**. Cualquier muestra de la señal en el tiempo es completamente independiente (sin correlación) de una muestra en cualquier otro tiempo.

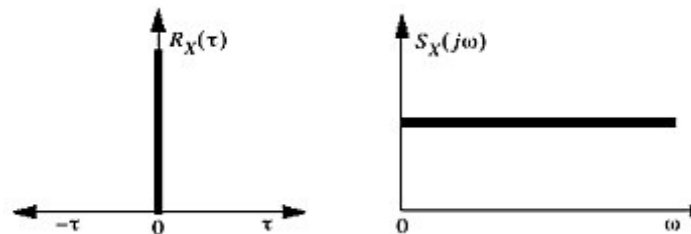


Figura 1.1.5.2: Ruido Blanco mostrado en el tiempo (izquierda) y en el dominio de la frecuencia

Aunque no es posible ver en la práctica (ningún sistema puede exhibir energía infinita a través de un espectro infinito), el **ruido blanco** es una parte importante en el diseño y el análisis. Las señales aleatorias pueden ser modeladas como **ruido blanco** filtrado o **formado**. Literalmente esto quiere decir que se podría filtrar la salida de una (hipotética) fuente de **ruido blanco** para lograr una fuente de ruido no blanca o **colorida** que es delimitada en el dominio de la frecuencia, y más correlacionada con el dominio de tiempo.

1.4 Estimación Estocástica.

Mientras que existen muchos avances para calcular (estimar) un estado desconocido desde un conjunto de medidas de un proceso, muchos de estos métodos, intrínsecamente, no toman en consideración la naturaleza típicamente **ruidosa** de las medidas. Por ejemplo, concidere un trabajo en análisis gráficos interactivos de computadora. Mientras que los requerimientos para el analisis de la información varían con la aplicación, la fuente principal de información es la misma: Las estimaciones son derivadas de medidas eléctricas **ruidosas** de sensores mecánicos, de inercia, ópticos, acústicos, o magnéticos. Este **ruido** es típicamente estadístico por naturaleza(o puede ser modelado efectivamente como tal), lo cual nos conduce hacia métodos estocásticos para tratar los problemas. Aquí se proporciona una introducción muy básica al tema, dirigido sobre todo a preparar al lector para la sección 2.

1.4.1 Modelos Estado-Espacio.

Los modelos *estado-espacio* son esencialmente una notación conveniente para los problemas

de estimación y control, desarrollados para hacer manejable lo que de otra manera sería un análisis notacionalmente intratable. Considere un proceso dinámico descrito por una ecuación en diferencias de orden *n-esimo* (similar a una ecuación diferencial) de la forma

$$y_{i+1} = a_{0,i} y_i + \dots + a_{n-1,i} y_{i-n+1} + u_i, i \geq 0,$$

Donde $\{u_i\}$ es una media-cero (estadísticamente) blanco

$$E(u_i, u_j) = R_u = Q_i \delta_{ij}$$

y los valores iniciales $\{y_0, y_{-1}, \dots, y_{-n+1}\}$ son variables aleatorias con media-cero, con una matrix de covarianza conocida $n \times n$

$$P_0 = E(y_{-j}, y_{-k}), j, k \in \{0, n-1\}$$

También se asume que

$$E(u_i, y_j) = 0 \text{ para } -n+1 \leq j \leq 0 \text{ y para } i \geq 0$$

esto asegura que

$$E(u_i, y_j) = 0, i \geq j \geq 0$$

En otras palabras, el ruido es estadísticamente independiente del proceso a ser estimado.

$$\vec{x}_{i+1} \equiv \begin{bmatrix} y_{i+1} \\ y_i \\ y_{i-1} \\ \vdots \\ y_{i-n+2} \end{bmatrix} = \underbrace{\begin{bmatrix} a_0 & a_1 & \dots & a_{n-2} & a_{n-1} \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}}_A \underbrace{\begin{bmatrix} y_i \\ y_{i-1} \\ y_{i-2} \\ \vdots \\ y_{i-n+1} \end{bmatrix}}_{\vec{x}_i} + \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_G u_i$$

Que conduce al modelo de *estado-espacio*

$$\vec{x}_{i+1} = A \vec{x}_i + G u_i$$

$$\vec{y}_i = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} \vec{x}_i$$

de manera más general

$$\vec{x}_{i+1} = A \vec{x}_i + Gu_i \quad 1.4.1.1$$

$$\vec{y}_i = H_i x_i \quad 1.4.1.2$$

La ecuación (1.4.1.1) representa la forma en que un nuevo estado \vec{x}_{i+1} es modelado como una combinación lineal del estado \vec{x}_i , y algún u_i proceso de ruido. La ecuación (1.4.1.2) describe la manera en que las observaciones o mediciones del proceso \vec{y}_i son derivadas a partir del estado interno \vec{x}_i . Estas dos ecuaciones son llamadas respectivamente **modelo de proceso** y **modelo de medida**, y sirven como base para casi todos los métodos de estimación lineales, como es el caso del filtro Kalman.

1.4.2 El Problema de Diseño del Observador.

Existe un problema general relacionado en el área de la teoría de los sistemas lineales generalmente llamado el problema de diseño del observador. El problema básico es determinar (estimar) los estados internos de un sistema lineal, teniendo acceso solamente a las salidas del sistema. Esto está relacionado con la idea de una caja negra en donde se tiene acceso a algunas señales que vienen de la caja pero no se puede observar directamente que pasa dentro de la caja. Las aproximaciones a este problema se basan típicamente en los modelos **espacio-estado** (1.4.1). Existe típicamente un **modelo** de proceso que modela la transformación del estado del proceso. Esto se puede presentar generalmente como una ecuación en diferencias lineal estocástica, similar a la ecuación (1.4.1.1)

$$x_k = A x_{k-1} + Bu_k + w_{k-1}. \quad 1.4.2.1$$

Además, hay una cierta forma de modelo de la medición, que describe la relación entre el estado de proceso y las mediciones. Esta puede ser representada como una expresión lineal similar a la ecuación (1.4.1.2)

$$z_k = Hx_k + v_k, \quad 1.4.2.2$$

los términos w_k y v_k son variables aleatorias que representan el ruido del proceso y de la medición respectivamente. Se puede ver que en la ecuación (1.4.2.2) se cambia la variable dependiente por z_k en lugar de y_k como en la ecuación (1.4.1.2). La justificación es reforzar la noción de que más medidas no tienen que ser específicamente elementos del estado, sino que pueden ser cualquier combinación de elementos de estado.

Ruido en la Medición y en el Proceso

Se debe tener en cuenta el caso de las mediciones del ruido en sensores. Existen muchas fuentes de ruido en estas mediciones. Por ejemplo, cada tipo de sensor tiene limitaciones fundamentales relacionadas con el medio físico asociado, y al sobrepasar estas limitaciones las señales se degradan. Además, el sensor y los circuitos eléctricos agregan alguna cantidad de ruido eléctrico aleatorio. La variación del ruido eléctrico en las señales afecta la cantidad y la calidad de la información. El resultado es que la información obtenida de cualesquier sensor debe ser calificada como parte de una secuencia global de estimaciones, y los modelos analíticos de medición, típicamente incorporan alguna noción de medida de ruido aleatorio o incertidumbre.

Existe el problema adicional de que el estado actual del modelo es completamente desconocido. Aunque se pueden hacer las predicciones sobre pequeños intervalos usando modelos basados en la transformación del estado reciente, estas predicciones asumen que las transformaciones son predecibles, lo cual no es siempre el caso.

2.El Filtro Kalman.

El filtro Kalman es un conjunto de ecuaciones matemáticas que proveen una eficiente manera computacional para estimar el estado de un proceso, de un modo que minimiza la varianza estimada del error utilizando mínimos cuadrados. El filtro es muy poderoso en varios aspectos: da soporte de estimaciones pasadas, presentes y de futuros estados y puede hacer esto aun cuando la naturaleza del sistema modelado sea desconocida. El filtro Kalman es el mejor estimador posible para una amplia clase de problemas y un muy efectivo estimador para una clase aun mayor.

El filtro de Kalman proporciona un buen marco para la estimación incremental de una cantidad en una situación en la cual las mediciones relacionadas con la misma están disponibles a lo largo del tiempo. Concretamente, se trata de una técnica de estimación Bayesiana empleada para seguir sistemas estocásticos dinámicos observados mediante sensores ruidosos.

El filtro es un procedimiento matemático que opera por medio de un mecanismo de predicción y corrección. En esencia este algoritmo pronostica el nuevo estado a partir de su estimación previa añadiendo un término de corrección proporcional al error de predicción, de tal forma que este último es minimizado estadísticamente

2.1 El Filtro Kalman Discreto.

2.1.1 El Proceso a Ser Estimado.

En el proceso llamado filtrado de Kalman se definen tres modelos o partes fundamentales a las cuales se les suele llamar modelo del proceso a ser estimado, y dos fases que constituyen el filtrado de Kalman propiamente dicho. Las partes (*modelos*) del proceso y las *fases* son las siguientes:

2.1.1.1 Modelo del Sistema.

Describe la evolución en el tiempo de la cantidad que se desea estimar, esta cantidad es expresada mediante un vector de estado $x_k \in \mathbb{R}^n$. La transición entre estados ($x_k \rightarrow x_{k+1}$), se caracteriza por la matriz de transición A_k , y la adición de un ruido gaussiano w_k que tiene media cero y una matriz de covarianza Q . Todo esto se representa mediante una ecuación en diferencias lineal estocástica de la siguiente forma:

$$\begin{aligned}x_k &= A_{k-1}x_{k-1} + B_k u_k + w_{k-1} \\x_{k+1} &= A_k x_k + B_k u_k + w_k,\end{aligned}\tag{2.1.1.1}$$

La matriz A $n \times n$ de la ecuación en diferencias(2.1.1.1) relaciona el estado del paso previo $k - 1$ con el estado en el paso actual k en ausencia de una función controlante o una perturbación de proceso. Note que en la práctica la matriz A podría cambiar con cada paso de tiempo, pero aquí se asume como constante.

También se puede observar a u_k , que se define como la entrada del sistema. La medición de esta entrada es z_k , que se relaciona con el vector de estado x_k , por medio de la matriz de medición H_k , esto se muestra en el siguiente paso.

2.1.1.2 Modelo de la Medición.

Relaciona el vector de medida $z_k \in \mathbb{R}^m$ con el estado del sistema x_k a través de la matriz de medición H_k y la adición de un ruido gaussiano v_k con matriz de covarianza R .

$$z_k = H_k x_k + v_k\tag{2.1.1.2}$$

La matrix B $n \times l$ relaciona el control opcional de entrada $u \in \mathbb{R}^l$ para el estado x . La matrix H $n \times m$ en la ecuación de medición (2.1.1.1.2) relaciona el estado con la medición z_k . En la práctica H podría cambiar con cada paso de tiempo, pero se asume como constante.

2.1.1.3 Modelo a Priori.

Aquí se describe el conocimiento previo del vector de estado en el instante inicial $x(0)$, en cuanto al valor esperado y a su matriz de covarianza $P(0)$. Con respecto a las variables aleatorias w_k y v_k representan el ruido del proceso y de la medición respectivamente. Se asume que son independientes entre ellas, que son ruido blanco y que tienen distribución de probabilidad normal, y por lo tanto cumplen:

$$E\{w_k\} = E\{v_k\} = 0$$

$$E\{w_k v_k^T\} = E\{v_k w_k^T\} = 0$$

$$E\{w_k w_k^T\} = Q$$

$$E\{v_k v_k^T\} = R$$

$$E\{w_k w_j^T\} = 0, \forall k \neq j \text{ y } E\{v_k v_j^T\} = 0, \forall k \neq j$$

En la práctica las matrices de covarianza de los ruidos del proceso Q y de la medición R , podrían cambiar con el paso del tiempo o la medición, de cualquier forma se asumen que son constantes. El ruido del sistema puede considerarse que se genera en su interior o bien que se introduce a la entrada del sistema, y el ruido de la medición es el error que se comete al medir la salida, es decir, será el error que cometen los sensores al medir. La figura siguiente aclara estos dos conceptos:

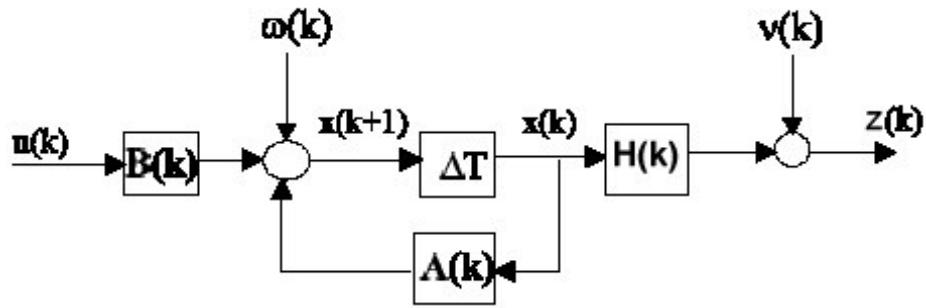


Figura 2.1.1.3:Diagrama sobre la relación de los ruidos

Ahora el problema consiste en estimar el valor óptimo del vector de estado x_k , basándose en las medidas ruidosas z_1, z_2, \dots, z_k ; además, se debe tener en cuenta que el vector de estado estará contaminado con el ruido del sistema. Para esto se consideran las mediciones más la predicción del estado a partir del conocimiento de su dinámica

2.1.2 Origen Computacional del Filtro Kalman.

El filtro de Kalman estima el proceso anterior utilizando una especie de control de retroalimentación, esto es, estima el proceso a algún momento en el tiempo y luego obtiene la retroalimentación por medio de los datos observados

Se define \hat{x}_k^- (note el "-" superior) como un estado *a priori*, estado declarado en el paso k a partir del conocimiento del proceso antes del paso k , y se define $\hat{x}_k \in \mathbb{R}^n$ como el estado *a posteriori*, estimado en el paso k a partir de la medida z_k . Luego se puede definir los errores *a priori* y *a posteriori*.

$$e_k^- \equiv x_k - \hat{x}_k^-$$

$$e_k \equiv x_k - \hat{x}_k$$

La covarianza *a priori* estimada del error es:

$$P_k^- = E[e_k^- e_k^{-T}], \quad 2.1.2.1$$

la covarianza *a posteriori* estimada del error es

$$P_k = E[e_k e_k^T]. \quad 2.1.2.2$$

derivando las ecuaciones para el filtro Kalman, comenzamos con el propósito de encontrar una ecuación que calcule un estado estimado *a posteriori* \hat{x} como una combinación lineal de un estado estimado *a priori* \hat{x}_k^- y una diferencia ponderada entre una medida real z_k y una medida de predicción $H\hat{x}_k^-$ de la siguiente forma

$$\hat{x} = \hat{x}_k^- + K(Z_k - H \hat{x}_k^-) \quad 2.1.2.3$$

La diferencia $(Z_k - H \hat{x}_k^-)$ en la ecuación (2.1.2.3) es llamada la medida de *innovación* o *residual*. El residuo refleja la diferencia entre la medida prevista $H \hat{x}_k^-$ y la medida real Z_k . Un residuo igual a cero significa que las dos están en completo acuerdo.

La matriz K de $n \times m$ en (2.1.2.3) es seleccionada para ser el factor de ganancia o el factor de mezcla que minimiza la ecuación de covarianza del error *a posteriori* (2.1.2.2). Esta minimización se puede lograr primero sustituyendo (2.1.2.3) en la definición anterior para e_k , luego sustituyendo esto en (2.1.2.2), calculando los valores esperados, tomando la derivada de la traza del resultado con respecto a K igualando el resultado a cero, y luego resolviendo para K .

$$\begin{aligned} K_k &= P_k^- H^T (H P_k^- H^T + R)^{-1} \\ &= \frac{P_k^- H^T}{(H P_k^- H^T + R)} \end{aligned} \quad 2.1.2.4$$

Mirando (2.1.2.4) observamos que la covarianza del error en la medición R aproxima a cero, la ganancia K pondera el residuo con más peso en la ecuación (2.1.2.3). Específicamente,

$$\lim_{R_k \rightarrow 0} K_k = H^{-1}$$

por otra parte, conforme la covarianza estimada del error *a priori* P_k^- se aproxima a cero, la ganancia K pondera el residuo menos duramente. Específicamente,

$$\lim_{P_k^- \rightarrow 0} K_k = 0$$

Otra manera de ver la carga para K es que conforme la covarianza del error de la medición R se acerca a cero, la medida actual Z_k es más y más *confiable*, mientras que la medida predicha $H \hat{x}_k^-$ es menos y menos *confiable*. De otra parte conforme la covarianza estimada del error *a priori* P_k^- se acerca a cero la medida actual Z_k es menos y menos *confiable* mientras que la medida predicha $H \hat{x}_k^-$ es más y más *confiable*.

2.1.3 Origen Probabilístico del Filtro Kalman.

La justificación para la ecuación (2.1.2.3) está basada en la probabilidad de la estimación *a priori* \hat{x}_k^- condicionado en todas las medidas anteriores z_k (regla de Bayes). Por ahora esto es suficiente para señalar que el filtro Kalman preserva los dos primeros momentos de la distribución de estado.

$$E[x_k] = \hat{x}_k$$

$$E\left[\left(x_k - \hat{x}_k\right)\left(x_k - \hat{x}_k\right)^T\right] = P_k.$$

La ecuación de estado estimado *a posteriori* (2.1.2.3) refleja la media (el primer momento) de la distribución del estado, esta está normalmente distribuida si las condiciones de las ecuaciones (2.1.1.3) y (2.1.1.4). La ecuación de la covarianza estimada del error *a posteriori* (2.1.2.2) refleja la varianza de la distribución del estado (el segundo momento no central), esto es,

$$\begin{aligned} p(x_k | z_k) &\sim N\left(E[x_k], E\left[\left(x_k - \hat{x}_k\right)\left(x_k - \hat{x}_k\right)^T\right]\right) \\ &= N\left(\hat{x}_k, P_k\right). \end{aligned}$$

2.1.3 El Algoritmo del Filtro Kalman Discreto.

Se pretende presentar la parte central de las ecuaciones y cómo usarlas con la versión discreta del filtro.

El filtro Kalman estima un proceso utilizando un control de retroalimentación, esto significa que estima el proceso en algún momento en el tiempo y luego obtiene la retroalimentación por medio de los datos observados, esta retroalimentación tiene forma de mediciones *ruidosas*. Desde este punto de vista las ecuaciones que se utilizan para derivar el filtro de Kalman se pueden dividir en dos grupos: las que *actualizan el tiempo* o ecuaciones de predicción y las que actualizan los datos observados o ecuaciones de *actualización de medida*. Las del primer grupo son responsables de la proyección hacia adelante (en el tiempo) del estado al momento k tomando como referencia el estado en el momento $k - 1$ y de la actualización intermedia de la

matriz de covarianza del estado (error de covarianza estimado). Para obtener las estimaciones *a priori* para el siguiente paso en el tiempo $k + 1$. Las ecuaciones de actualización de medida son responsables de la retroalimentación, es decir, incorporan una nueva medición dentro de la estimación *a priori* con lo cual se llega a una estimación mejorada del estado *a posteriori*.

Las ecuaciones que actualizan el tiempo pueden también ser pensadas como ecuaciones de pronóstico, mientras que las ecuaciones que incorporan nueva información pueden considerarse como ecuaciones de corrección. Efectivamente, el algoritmo de estimación final se asemeja a un algoritmo predictor-corrector de los utilizados para resolver problemas numéricos. Así, el filtro de Kalman funciona por medio de un mecanismo de proyección y corrección al pronosticar el nuevo estado y su incertidumbre y corregir la proyección con la nueva medida. Como se muestra en la siguiente figura.

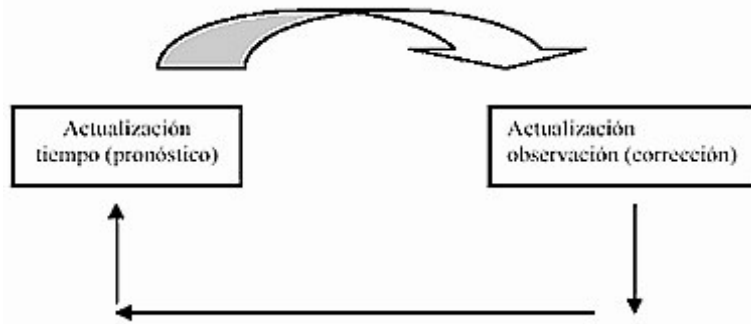


Figura 2.1.3.1: Ciclo del Filtro Kalman.

El primer paso consiste en generar un pronóstico del estado hacia adelante en el tiempo tomando en cuenta toda la información disponible en ese momento y en un segundo paso, se genera un pronóstico mejorado del estado, de tal manera que el error es minimizado estadísticamente.

Las ecuaciones específicas para la actualización del tiempo y la medida son detalladas a continuación.

Las ecuaciones de actualización de tiempo para el filtro Kalman discreto son:

$$\hat{x}_k^- = A \hat{x}_k + Bu_k \quad 2.1.3.1$$

$$P_k^- = AP_{k-1}A^T + Q \quad 2.1.3.2$$

Note cómo las ecuaciones anteriores pronostican las estimaciones del estado y la covarianza desde $k - 1$ a k . La matriz A relaciona el estado en el momento previo $k - 1$ con el estado al momento actual k , esta matriz podría cambiar para los diferentes momentos en el tiempo (k). Q representa la covarianza de la perturbación aleatoria del proceso que trata de estimar el estado.

las ecuaciones de corrección o de actualización de la medida para el filtro Kalman discreto son:

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} \quad 2.1.3.3$$

$$\hat{x}_k = \hat{x}_k^- + K_k \left(z_k - H \hat{x}_k^- \right) \quad 2.1.3.4$$

$$P_k = (I - K_k H) P_k^- \quad 2.1.3.5$$

La primera tarea durante la corrección de la proyección del estado es el cálculo de la ganancia de Kalman, K_k (2.1.3.3). Note que la ecuación dada aquí como (2.1.3.3) es la misma que en (2.1.2.4) El siguiente paso es medir el proceso para obtener z_k y luego generar una nueva estimación del estado que incorpora la nueva observación, como en la ecuación (2.1.3.4). El paso final es obtener una nueva estimación de la covarianza del error mediante la ecuación (2.1.3.5).

Después de cada par de actualizaciones, tanto del tiempo como de la medida, el proceso es repetido tomando como punto de partida las nuevas estimaciones del estado y de la covarianza del error. Esta naturaleza recursiva es una de las características llamativas del filtro Kalman. Esto hace posible implementaciones prácticas. El filtro Kalman condiciona recurrentemente la estimación actual en todas las últimas medidas La siguiente figura ofrece una imagen completa del funcionamiento del filtro, combinando el diagrama (2.1.3.1) con las ecuaciones (2.1.3.1, 2.1.3.2, 2.1.3.3, 2.1.3.4, 2.1.3.5)

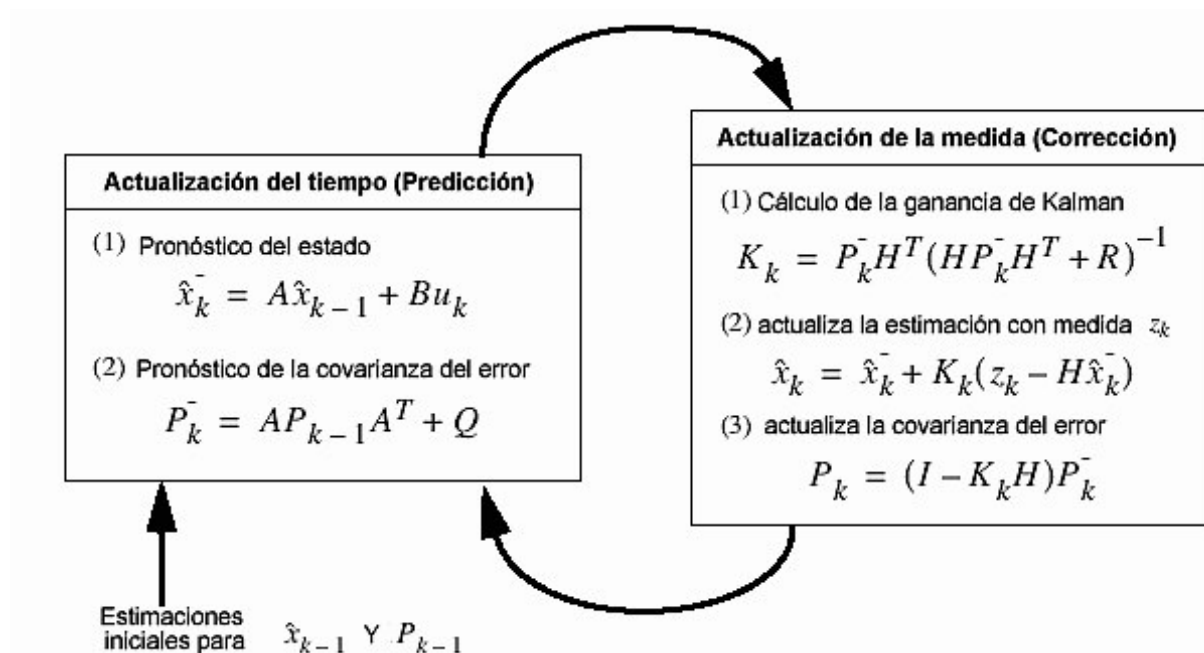


Figura 2.1.3.2: Imagen completa de la operación del filtro Kalman

2.2 El Filtro Kalman Extendido (FKE).

2.2.1 El Proceso a ser Estimado.

Según lo explicado en la sección 2.1.1, el filtro Kalman se ocupa del problema general de intentar estimar el estado $x \in \mathfrak{R}^n$ de un proceso controlado en tiempo discreto que es gobernado por una ecuación estocástica lineal el diferencia. Pero que sucede cuando el proceso a ser estimado y/o la relación de medida para el proceso es no lineal?. Algunos de las más interesantes y exitosas aplicaciones de la filtración Kalman están dentro de esta pregunta. Un filtro Kalman que lineariza la media y la covarianza actuales se denomina *filtro kalman extendido* o FKE

Mediante algo semejante a una serie de Taylor se puede linearizar la estimación al rededor del estimador actual usando las derivadas parciales de las funciones de proceso y medida para calcular estimaciones aun en el lado de las relaciones no lineales. De esta forma se modifican las ecuaciones presentadas en la sección 2.1. Se asume nuevamente que el proceso tiene un vector de estado $x \in \mathfrak{R}^n$, pero que este proceso es ahora gobernado por una ecuación en diferencias estocástica no lineal

$$x_k = f(x_{k-1}, u_k, w_{k-1}),$$

2.2.1.1

Con una medida $z \in \mathbb{R}^m$ esto es

$$z_k = h(x_k, v_k), \quad 2.2.1.2$$

Las variables aleatorias w_k y v_k representan el ruido del proceso y de la medida respectivamente como en las ecuaciones (2.1.1.3) y (2.1.1.4). En este caso la función *no lineal* f en la ecuación diferencial (2.2.1.1) relaciona el estado en el paso previo $k - 1$ con el estado en el paso actual k . Incluye como parámetros cualquier función controlante u_k y el ruido del proceso con media cero w_k . La función *no lineal* h en la ecuación de la medida (2.2.1.2) relaciona el estado x_k con la medida z_k .

Naturalmente en la práctica se desconocen los valores individuales del ruido w_k y v_k para cada paso de tiempo. Sin embargo se puede aproximar el vector de estado y de la medida de la siguiente forma

$$\tilde{x}_k = \left(f^{\wedge} x_{k-1}, u_k, 0 \right) \quad 2.2.1.3$$

$$\tilde{z}_k = h(\tilde{x}_k, 0), \quad 2.2.1.4$$

Donde \tilde{x}_k es alguna estimación *a posteriori* del estado (para un paso previo k)

Es importante mencionar que un defecto fundamental del FKE es que las distribuciones (o las densidades en el caso continuo) de las variables aleatorias no son normales después de sufrir la transformación no lineal. El FKE es simplemente un estimador de estado que solo aproxima la optimalidad de la regla de Bayes por linealización.

2.2.2 El Origen Computacional del Filtro Kalman Extendido.

Para estimar un proceso con diferencia no lineal y relaciones de medida, se re escriben las ecuaciones que linealizan un estimador sobre la ecuación (2.2.1.3) y la ecuación (2.2.1.4),

$$x_k \approx \tilde{x}_k + A \begin{pmatrix} x_{k-1} \\ \hat{x}_{k-1} \end{pmatrix} + Ww_{k-1}, \quad 2.2.1.5$$

$$z_k \approx \tilde{z}_k + H \begin{pmatrix} x_k \\ \tilde{x}_k \end{pmatrix} + Vv_k. \quad 2.2.1.6$$

Donde

- - * x_k y z_k son los vectores actuales del estado y la medida,
- * \tilde{x}_k y \tilde{z}_k son los vectores del estado de aproximación y de medida de la ecuación (2.2.1.3) y la ecuación (2.2.1.4)
- * \hat{x}_k es el estimado *a posteriori* del estado en el paso k .
- * las variables aleatorias w_k y v_k representan el ruido del proceso y de la medida, respectivamente, como en la ecuación (2.1.1.3) y (2.1.1.4).
- * A es la matriz Jacobiana de derivadas parciales de f con respecto a x esto es

$$A_{[i,j]} = \frac{\partial f_{[i]}}{\partial x_{[j]}} \begin{pmatrix} \hat{x}_{k-1} \\ u_k, 0 \end{pmatrix},$$

- - * W es la matriz Jacobiana de derivadas parciales de f con respecto a w ,

$$W_{[i,j]} = \frac{\partial f_{[i]}}{\partial w_{[j]}} \begin{pmatrix} \hat{x}_{k-1} \\ u_k, 0 \end{pmatrix},$$

- - * •
- * H es la matriz Jacobiana de derivadas parciales de h con respecto a x ,

- - * •

$$H_{[i,j]} = \frac{\partial h_{[i]}}{\partial x_{[j]}} (\tilde{x}_k, 0),$$

- * V es la matriz Jacobiana de derivadas parciales de h con respecto a v ,
-

$$V_{[i,j]} = \frac{\partial h_{[i]}}{\partial v_{[j]}} (\tilde{x}_k, 0),$$

Observe que en las notaciones de las ecuaciones anteriores no se utiliza el subíndice del paso en el tiempo k con los jacobianos A, W, H, V , aunque ellos, en efecto, son diferentes en cada paso del tiempo

Ahora se define una nueva notación para la predicción del error,

$$\tilde{e}_{x_k} \equiv x_k - \tilde{x}_k, \quad 2.2.1.7$$

y la medición residual,

$$\tilde{e}_{z_k} \equiv z_k - \tilde{z}_k. \quad 2.2.1.8$$

Es necesario recordar que en la práctica no se tiene acceso a x_k en la ecuación (2.2.1.7), el cual es el *actual* vector de estado, es decir la cantidad que se intenta estimar. Por otra parte, se tiene acceso a z_k en la ecuación (2.2.1.8), esta es la medida actual que se usa para estimar x_k . Usando la ecuación (2.2.1.7) y la ecuación (2.2.1.8) se pueden escribir las ecuaciones que gobiernan los *procesos de error* como:

$$\tilde{e}_{x_k} \approx A \begin{pmatrix} x_{k-1} & -\hat{x}_{k-1} \end{pmatrix} + \varepsilon_k, \quad 2.2.1.9$$

$$\tilde{e}_{z_k} \approx H \tilde{e}_{x_k} + \eta_k, \quad 2.2.1.10$$

donde ε_k y η_k representan nuevas variables aleatorias independientes con media cero y matrices de covarianza WQW^T y VRV^T , con Q y R como en (2.1.1.3) y (2.1.1.4) respectivamente.

se nota que las ecuaciones (2.2.1.9) y (2.2.1.10) son lineales y esto se asemeja a las ecuaciones del proceso y la medición, ecuaciones (2.1.1.1) y (2.1.1.2) del filtro Kalman discreto. Esto motiva a utilizar el residuo de la medida actual \tilde{e}_{z_k} en la ecuación (2.2.1.8) y un segundo (hipotético) Filtro Kalman para estimar el error de la predicción \tilde{e}_{x_k} dado por la ecuación (2.2.1.9). esta estimación llamada $\tilde{\tilde{e}}_k$, luego se puede usar con la ecuación (2.2.1.7) para obtener el estado estimado *a posteriori* estimado para el proceso no lineal original. De la siguiente forma:

$$\tilde{x}_k = \hat{x}_k + \tilde{e}_k \quad 2.2.1.11$$

las variables aleatorias de las ecuaciones (2.2.1.9) y (2.2.1.10) tienen aproximadamente las siguientes distribuciones de probabilidad

$$p(\tilde{e}_{x_k}) \sim N(0, E[\tilde{e}_{x_k} \tilde{e}_{x_k}^T])$$

$$p(\varepsilon_{x_k}) \sim N(0, WQ_k W^T)$$

$$p(\eta_{x_k}) \sim N(0, VR_k V^T)$$

Dadas estas aproximaciones y dejando el valor previsto de \hat{e}_k a cero, la ecuación de filtro de Kalman usada para estimar \hat{e}_k es

$$\hat{e}_k = K_k \tilde{e}_{z_k} . \quad 2.2.1.12$$

Sustituyendo (2.2.1.12) en (2.2.1.11) y haciendo uso de la ecuación (2.2.1.8) se ve que no es necesario el segundo (hipotético) filtro Kalman.

$$\hat{x}_k = \tilde{x}_k + K_k \tilde{e}_{z_k}$$

$$= \tilde{x}_k + K_k(z_k - \tilde{z}_k) \quad 2.2.1.13$$

La ecuación (2.2.1.13) puede ser actualizada para la actualización de la medida en el filtro

Kalman extendido, con \hat{x}_k^- y \hat{x}_k viniendo de las ecuaciones (2.2.1.1) y (2.2.1.2), y la ganancia Kalman de la ecuación (2.1.3.3) con la sustitución apropiada para la covarianza del error de la medición

El conjunto completo de ecuaciones para el FKE se muestran a continuación. Note que se ha sustituido \hat{x}^- por \hat{x}_k^- para ser consistentes con la notación anterior del super índice "-", para el estado *a priori* y que ahora se ha colocado un sub índice k a los Jacobianos A, W, H y V .

Las ecuaciones de actualización de tiempo son:

$$\hat{x}_k^- = \left(f \hat{x}_{k-1}, u_k, 0 \right), \quad 2.2.1.14$$

$$P_k^- = A_k P_{k-1} A_k^T + W_k Q_{k-1} W_k^T \quad 2.2.1.15$$

Como en el filtro Kalman discreto, las dos ecuaciones anteriores (2.2.1.14) y (2.2.1.15) proyectan las estimaciones del estado y de la covarianza del paso de tiempo anterior $k - 1$ al paso del tiempo actual k . Nuevamente f en la ecuación (2.2.1.14) viene de la ecuación (2.2.1.3), A_k y W_k son los Jacobianos del proceso en el paso k , y Q_k es la covarianza del ruido del proceso(2.1.1.3) en el paso k .

Las ecuaciones de actualización de la medida son:

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + V_k R_k V_k^T)^{-1} \quad 2.2.1.16$$

$$\hat{x}_k = \hat{x}_k^- + K_k \left(Z_k - h \left(\hat{x}_k^-, 0 \right) \right) \quad 2.2.1.17$$

$$P_k = (I - K_k H_k) P_k^- \quad 2.2.1.18$$

Igual que con el filtro Kalman discreto, las ecuaciones de actualización de la (2.2.1.16), (2.2.1.17) y (2.2.1.18) corrigen la estimación del estado y de la covarianza con la medida z_k . Nuevamente h en la ecuación (2.2.1.17) viene de la ecuación (2.1.2.4), H_k y V son los Jacobianos de la medición en el paso k , y R_k es la covarianza del ruido de la medición(2.1.1.4)

en el paso k (note el subíndice R que cambia con cada medida).

La operación básica del FKE es la misma que la del filtro Kalman discreto como se muestra en la siguiente figura se muestra una completa imagen combinando el diagrama (2.1.3.1) con las ecuaciones (2.2.1.14, 2.2.1.15, 2.2.1.16, 2.2.1.17, 2.2.1.18)

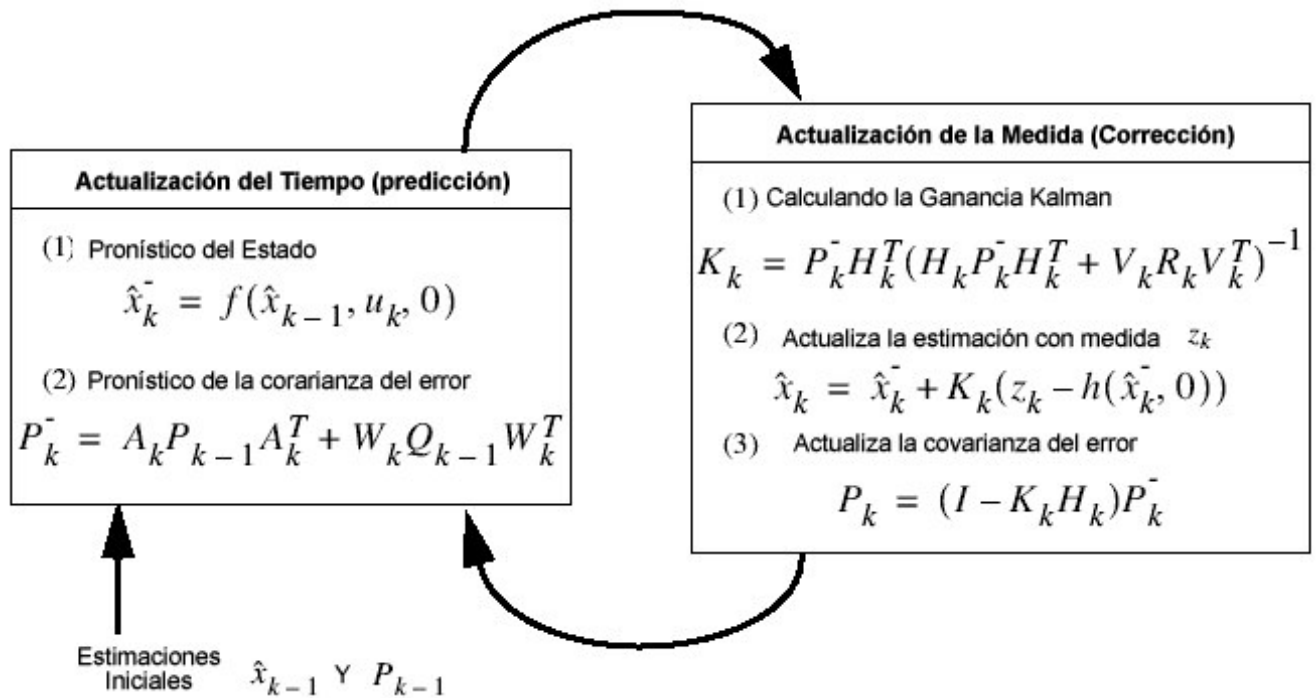


Figura 2.2.1.1: Imagen completa de la operación del filtro Kalman extendido

Una característica importante del FKE es que el Jacobiano H_k en la ecuación para la ganancia K_k sirve para propagar o "magnificar" sólo el componente relevante de la información de la medida. Por ejemplo, si no existe un mapeo uno a uno entre la medida z_k y el estado a través de h , el Jacobiano H_k afecta la ganancia de Kalman de modo que sólo magnifica la porción del residuo $z_k - h(\hat{x}_k, 0)$ que afecta al estado.

Por supuesto que si sobre todas las medidas no existe un mapeo uno a uno entre la medida z_k y el estado a través de h , entonces el filtro diverge rápidamente. En este caso el proceso es *no observable*.

2.3 Ejemplo

2.3.1 Estimando una Constante Aleatoria.

En las secciones (2.1) y (2.2) se presentaron las consideraciones matemáticas para el filtro Kalman discreto y el filtro Kalman extendido, para entender mejor el concepto se presenta el siguiente ejemplo

2.3.1.1 El Modelo del Proceso.

En este ejemplo se desea estimar un escalar aleatorio constante, por ejemplo un voltaje. Se asume que se tiene la capacidad de medir la constante pero las medidas están perturbadas por un *ruido blanco* de medida 0.1 Voltios (el convertidor que se utiliza de análogo a digital no es muy exacto). En este ejemplo el proceso es gobernado por una ecuación lineal el diferencia

$$x_k = Ax_{k-1} + Bu_k + w_k,$$

$$= x_{k-1} + w_k$$

Con una medida $z \in \mathbb{R}^1$ esto es

$$z_k = Hx_k + v_k$$

$$x_k + v_k$$

El estado no se altera de paso a paso, por lo tanto $A = 1$. No hay control de entrada por lo tanto $u = 0$. La medida del ruido se toma del estado directamente, entonces $H = 1$. Note que se cambió el subíndice k en varios lugares porque los parámetros respectivos siguen siendo constantes en el modelo

2.3.1.2 Los Parámetros y las Ecuaciones del Filtro.

Las ecuaciones de actualización del tiempo son:

$$\hat{x}_k^- = \hat{x}_{k-1} ,$$

$$P_k^- = P_{k-1} + Q ,$$

Las ecuaciones para la actualización de la medida son:

$$K_k = P_k^- (P_k^- + R)^{-1}$$

$$= \frac{P_k^-}{P_k^- + R} \quad 2.3.1.2.1$$

$$\hat{x}_k = \hat{x}_k^- + K_k \left(Z_k - \hat{x}_k^- \right) ,$$

$$P_k = (1 - K_k) P_k^- .$$

Presumiendo un muy pequeña varianza del proceso, se toma $Q = 1e - 5$. (se podría hacer $Q = 0$ pero se asume un valor pequeño diferente a cero esto da mayor flexibilidad para "afinar" el filtro, como se muestra más adelante). También se asume por experiencia que el valor verdadero de la constante aleatoria tiene una distribución de probabilidad normal estándar, por esta razón se fijará el filtro con la suposición de que la constante es 0. En otras palabras, antes de iniciar tenemos $\hat{x}_{k-1} = 0$

De la misma forma se necesita fijar un valor inicial para P_{k-1} , denominado P_0 . Si se está absolutamente seguro que el estado inicial estimado $\hat{x}_0 = 0$ es correcto, se dejaría $P_0 = 0$. Sin embargo, dada la incertidumbre en la estimación inicial \hat{x}_0 , al elegir $P_0 = 0$, causaría que el filtro cree inicialmente y para siempre $\hat{x}_k = 0$. Como resulta la elección alternativa no es crítica. Se podría elegir casi cualquier $P_0 \neq 0$ y el filtro eventualmente converge. Se comienza el filtro con $P_0 = 1$. para este ejemplo.

2.3.1.3 Las Simulaciones.

Para empezar, se realiza la selección aleatoria de una constante escalar $z = 0.37727$ (no hay "sombrero" sobre la z porque esta representa la "verdadera"). Se simulan 50 medidas distintas z_k que tienen una distribución normal del error al rededor del cero con una desviación estandar de 0.1 (se debe recordar que se asume que las medidas son corruptas por un *ruido blanco* de medida 0.1 Voltios. Se pueden generar las medidas individuales dentro del ciclo del filtro, pero pre generar el conjunto de 50 medidas permite correr varias simulaciones con las mismas medidas exactas (es decir la misma medida de ruido) de modo que las comparaciones entre las simulaciones con diversos parámetros resultan más significativas.

En la primera simulación se fija la varianza de la medida en $R = (0.1)^2 = 0.01$. Debido a que esta es la "verdadera" medida de la varianza del error, se espera el "mejor" funcionamiento en términos de balanceo de sensibilidad y varianza estimada. Esto se hará más evidente en la segunda y tercera simulación. La siguiente figura muestra los resultados de esta simulación inicial. El valor verdadero de la constante aleatoria $x = -0.37727$ es dado por la línea horizontal, las medidas del ruido por las marcas en cruz, y la estimación del filtro por la curva sobrante.

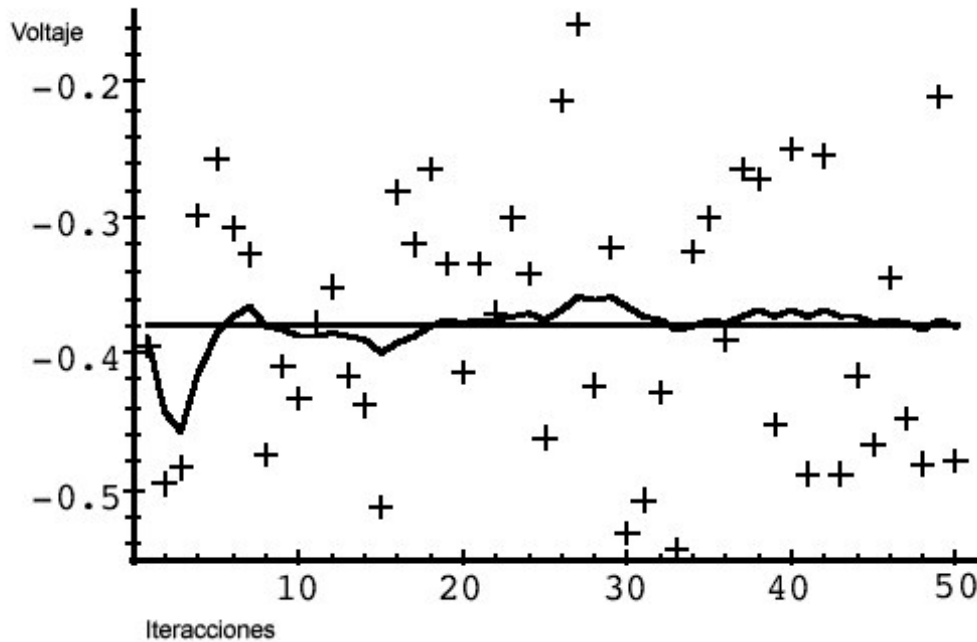


Figura 2.3.1.31.: La primera simulación con $R = (0.1)^2 = 0.01$. el Valor real de la constante aleatoria $x = -0.37727$ este es dado por la línea sólida horizontal, la medida del ruido por las marcas en cruz, y la estimación del filtro por la otra curva.

Al condicionar la elección para para P_0 en la parte de arriba, se menciona que la elección no es crítica, mientras $P_0 \neq 0$ porque el filtro eventualmente converge. En la siguiente figura se puede ver la interacción entre el valor de P_k versus la interacción. Para la interacción número 50,

se ha fijado desde el valor inicial seleccionado en 1 hasta aproximadamente 0.0002 (voltios²).

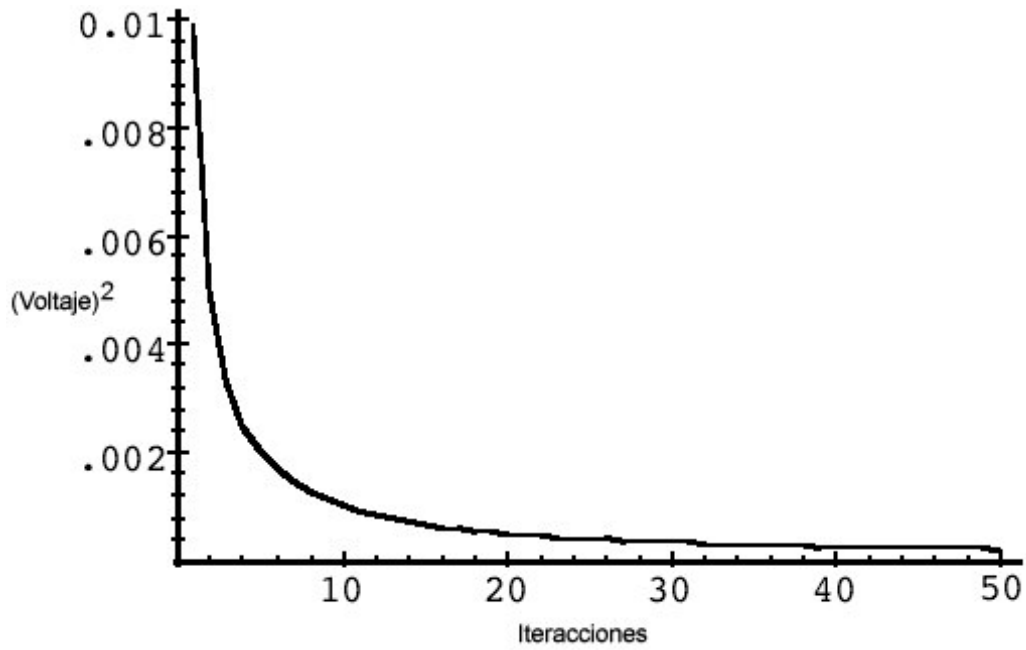


Figura 2.3.1.3.2: Después de 50 iteraciones el valor P_k de la covarianza de error fijado en 1 se ha fijado aproximadamente en 0.0002 (voltios²)

En las siguientes figuras se puede ver que pasa cuando R se incrementa o decrementa en un factor de 100. En la siguiente figura $R = 1$ el filtro responde lentamente a la medida, resultando en una estimación reducida de la varianza.

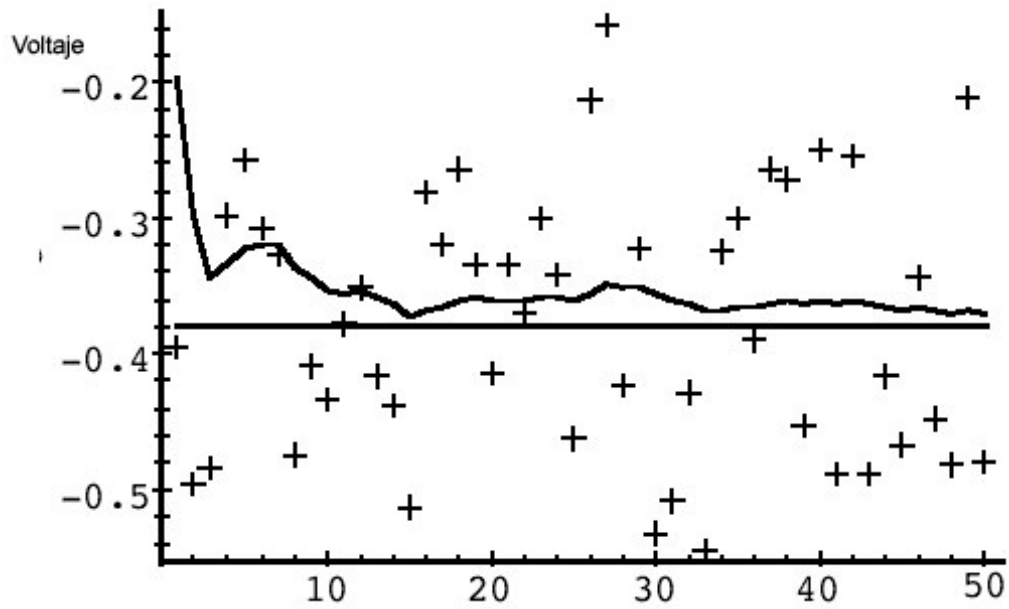


Figura 2.3.1.3.3.: La segunda simulación con $R = 1$. El filtro responde lentamente a la medida, resultando en una estimación reducida de la varianza.

En la siguiente figura la medida de la varianza es 100 veces más pequeña $R = 0.0001$

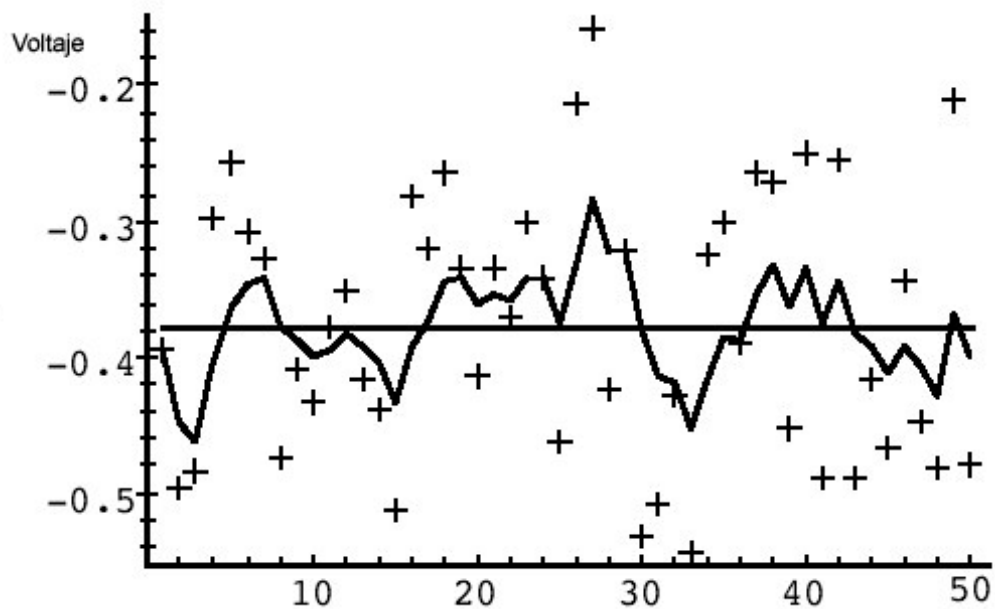


Figura 2.3.1.3.4.: La tercera simulación con $R = 0.001$. el filtro responde rápidamente a la medida, incrementando la estimación de la varianza.

Aunque la estimación de una constante es relativamente directa, esto sirve para demostrar claramente el trabajo del filtro Kalman.

En la figura 2.3.1.3.3 en particular en la filtración Kalman es evidente que la estimación aparece considerablemente más suave que las medidas del ruido..

3. Conclusiones.

Después de estudiar el Filtro Kalman podemos decir que.

*El filtro de Kalman utiliza el método de mínimos cuadrados para generar recursivamente un estimador del estado al momento k , que es lineal, insesgado y de varianza mínima. De aquí su enorme poder, para resolver un amplio rango de problemas en inferencia estadística.

*El filtro se distingue por su habilidad para predecir el estado de un modelo en el pasado, presente y futuro, aún cuando la naturaleza precisa del sistema modelado es desconocida. La modelación dinámica de un sistema es una de las características claves que distingue el método de Kalman. Los modelos lineales dinámicos son modelos con una transición lineal desde un periodo al próximo, los cuales pueden describir la mayoría de los modelos comúnmente utilizados en trabajos de series de tiempo.

*El desarrollo del filtro de Kalman, tal como se encuentra en el documento original, supone un conocimiento amplio en teoría de probabilidades, específicamente con el tema de la condicionalidad gaussiana en las variables aleatorias, lo cual puede originar una limitante para su estudio y aplicación.

*Cuando se desarrolla para modelos autorregresivos los resultados están condicionados a la información pasada de la variable en cuestión. En este sentido el pronóstico con series de tiempo representa la fuerza o inercia que actualmente presenta el sistema y son eficientes únicamente en el corto plazo.

*Existen muchas herramientas para realizar estimaciones dinámicas pero es casi nulo nuestro conocimiento de ellas. La importancia de este conocimiento radica en el hecho de poder relacionarlas con los conceptos teóricos aprendidos en los diferentes cursos y ver su funcionalidad en la solución de problemas prácticos.

4. Bibliografía

- [Bla02]Blake, Andrew (2002) "*State-Space Models and the Kalman Filter: Application, Formulation and Estimation*", Bank of England.
- [KAL60]Kalman, R.E., (1960) "*A New Approach to Linear Filtering and Prediction Problems*", Trans. ASME, J. Basic Engineering, vol 82, March 1960, pp 94-35.
- [MAY79]Maybech, Peter S. (1979) "*Stochastic Models, Estimation and Control*", Volumen 1, Academic Press, INC.
- [WEL02]Welch, Greg and Gary Bishop (2002). "*An Introduction to the Kalman Filter*", TR 95-041, Department of Computer Science, University of North Carolina at Chapel Hill, 2002.
- [MIS02]Misas A. y Diego Vásquez, (2002) "*Expectativas de Inflación en Colombia: Un Ejercicio Econométrico*", Subgerencia de Estudios Económicos, Banco de la República, Colombia.
- [KIK03]Ana Cecilia Kikut V. (2003). "*Técnicas recursivas de estimación de los coeficientes de regresión*", División económica departamento de investigaciones económicas informe técnico
- [RAM03]Álvaro Solera Ramírez.(2003). "*EL FILTRO DE KALMAN*", Banco central de costa rica división económica departamento de investigaciones económicas julio del 2003.
- [HTTP]The Kalman Filter. "*Some tutorials, references, and research on the Kalman filter*" <http://www.cs.unc.edu/~welch/kalman/>