



UNIVERSITAT DE
BARCELONA

Facultat de Matemàtiques
i Informàtica

GRADO DE MATEMÁTICAS

Trabajo final de grado

EL FILTRO DE KALMAN

Autor: Víctor Pelegrí Àlvarez

Director: Dr. José Manuel Corcuera Valverde

Realitzat a: Departamento
de Matemáticas y Informática

Barcelona, 20 de junio de 2019

Abstract

In this work we will study the Kalman filter, highlighting its main equations and developing how they are. We will also deal with Extended Kalman Filter case analogously and Particle Filter case and a brief example for each one. Then we will see some more general considerations about its implementation and how we measure its performance applied to an example.

Resumen

En este trabajo estudiaremos el Filtro de Kalman, destacando sus ecuaciones principales y desarrollando como se encuentran. Trataremos también el caso del Filtro de Kalman Extendido de forma análoga y el caso del Filtro de la Partícula y un breve ejemplo para cada uno. Luego veremos algunas consideraciones más generales sobre la implementación de este y como medir su rendimiento aplicado a un ejemplo.

Agradecimientos

A todos los profesores que he tenido a lo largo de estos años y a mi tutor José Manuel Corcuera por la dedicación y el tiempo que me han ofrecido. A mi familia, amigas y compañeras por hacer del apoyo mutuo y la solidaridad, no solo un eslogan, sino nuestra mejor arma en todo momento.

Índice general

1. Introducción	3
1.1. Objetivos del Filtro de Kalman	4
1.2. Modelos de estado	5
2. Resultados previos	7
3. Algoritmo del Filtro de Kalman	9
3.1. Estimación del estado y covarianza del error en la estimación del estado .	9
3.2. Innovación	10
3.3. La matriz de ganancia	11
3.3.1. Actualización del estimador del estado	11
3.3.2. Cálculo de la matriz de ganancia	12
3.4. Pasos del algoritmo del Filtro de Kalman	13
3.4.1. Estimación	13
3.4.2. Actualización	13
4. Filtros no lineales	15
4.1. Filtro de Kalman Extendido	15
4.1.1. Predicción del estado	15
4.1.2. La predicción de la observación y la innovación	17
4.1.3. Actualización de ecuaciones	18
4.2. El Filtro de la Partícula	18
4.2.1. Filtro Recursivo Bayesiano	18
4.2.2. Muestreo por Importancia	19
4.2.3. El Filtro de la Partícula	21
4.2.4. Aplicación: Modelos de Volatilidad Estocástica	23
4.2.5. Estimación de los parámetros	25
5. Implementación del Filtro de Kalman	27

<i>ÍNDICE GENERAL</i>	1
5.1. Planteamiento general	28
5.2. Inicialización	29
5.3. Estabilidad del rendimiento	29
5.4. Consistencia del filtro	30
5.4.1. Test sobre el sesgo de la innovación	30
5.4.2. Test de incorrelación de la innovación	31
6. Ejemplo del Filtro de Kalman Extendido	33
7. Conclusiones	37
8. Anexo	39
8.1. Esperanza condicionada	39
8.1.1. Definiciones previas	39
8.1.2. La esperanza condicionada, definiciones y propiedades	40
8.1.3. Cálculo de la esperanza condicionada	41
9. Bibliografía	45

Capítulo 1

Introducción

El **Filtro de Kalman** es un algoritmo recursivo desarrollado por Rudolf E. Kalman en 1960 que sirve para poder calcular la estimación de un estado que evoluciona en el tiempo y del cual tenemos información precisa. En este modelo supondremos que tenemos una ecuación para el estado que queremos estudiar, y otra para las observaciones de este estado, que son combinación lineal del estado y el **ruido**, es decir, una variable aleatoria que perturba las observaciones. Este algoritmo garantiza que la estimación obtenida minimiza el error cuadrático medio (la diferencia entre el valor de la estimación del estado y el del estado). Veremos que el estimador del estado corresponde a la esperanza condicionada y que este es el mejor estimador posible, además en el caso gaussiano es lineal. El resultado que nos proporciona el filtro es una estimación del estado que es una suma ponderada de la estimación anterior y la nueva observación, esta es una suposición que hacemos y a partir de la cual encontramos la matriz de ganancia que, como hemos dicho antes, minimiza el error cuadrático medio. La importancia del Filtro de Kalman es consecuencia de que no requiere almacenar toda la información sobre las observaciones pasadas y que al tratar con estimaciones lineales es un algoritmo de uso relativamente sencillo y por lo tanto es ideal para resolver problemas complejos de estimación.

Las ideas de Rudolf E. Kalman sobre el filtraje (consideradas la contribución más importante en la teoría de estimación de los últimos 50 años) fueron recibidas al principio con gran escepticismo, tanto es así que se vio obligado a presentar sus resultados en publicaciones de ingeniería mecánica en vez de ingeniería eléctrica o de sistemas (la publicación se titulaba "A new approach to linear filtering and prediction problems" en el Journal of Basic Engineering 82 (1): 35–45). Posteriormente tuvo más éxito durante su visita al NASA Ames Research Center. Esto implicó el uso del algoritmo en el programa Apolo, en el transbordador espacial de la NASA, en submarinos de guerra de la Marina, en armas y vehículos no tripulados y en el desarrollo del sistema de posicionamiento GPS. Por todo eso el Filtro de Kalman tiene numerosas aplicaciones en tecnología. Además, se utiliza ampliamente en campos como el procesamiento de señales y la econometría, como por ejemplo en el cálculo del PIB de un país.

1.1. Objetivos del Filtro de Kalman

Las aplicaciones del Filtro de Kalman abarcan muchos campos, pero su uso como herramienta tiene normalmente dos objetivos:

1. **Estimación de estados de sistemas dinámicos:** casi todos los sistemas físicos son dinámicos, no son constantes. Si queremos estimaciones precisas de sus características en el tiempo, tenemos que tener en cuenta sus dinámicas. Pero las dinámicas no se conocen normalmente de forma precisa. Por eso se expresan en probabilidades. El filtro nos permite estimar el estado de sistemas dinámicos con ciertos tipos de comportamientos aleatorios a partir de información estadística. Estos son algunos ejemplos de estimación de sistemas dinámicos

Aplicación	Sistema dinámico	Tipos de sensor
Control de procesos	Planta química	Presión Temperatura Tasa de flujo Analizador de gas
Predicción de inundaciones	Sistema del río	Nivel del agua Pluviómetro Radar del tiempo
Rastreo	Nave espacial	Radar Sistema de imágenes
Navegación	Barco	Sextante Giroscopio Acelerómetro

2. **Análisis del rendimiento de sistemas de estimación:** el objetivo del análisis es determinar si se ajustan bien los diferentes tipos de sensores para un conjunto de criterios de diseño dados. Estos criterios están normalmente relacionados con la precisión de la estimación y el coste del sistema. El filtro de Kalman usa una caracterización paramétrica de la distribución de probabilidad del error de estimación para determinar las matrices de ganancia óptimas, y su distribución de probabilidad se usa evaluando su rendimiento como una función de "diseño de parámetros" de un sistema de estimación, como por ejemplo:

- a) Los tipos de sensores que se usan.
- b) Las posiciones y orientaciones usando diferentes tipos de sensores con respecto al sistema que se quiere estimar.
- c) Las características del ruido admitido por los sensores.
- d) El prefiltrado de métodos para suavizar el ruido.
- e) Las tasas de muestreo de datos para varios tipos de sensores.
- f) El nivel de simplificación del modelo para reducir la implementación de requerimientos.

1.2. Modelos de estado

Describimos a continuación algunos modelos de series temporales que pueden ser representados como estados lineales. Indicamos que la serie $\{Z_n, n = 1, 2, \dots\}$ satisface una ecuación de la forma

$$Z_n = H_n X_n + V_n \quad n = 1, 2, \dots$$

donde

$$X_n = F_n X_{n-1} + W_n \quad n = 1, 2, \dots$$

Esta última ecuación puede ser interpretada como la evolución de un **estado** X_n (vector de dimensión d) de un sistema en el instante n en términos de una secuencia conocida de matrices $F_1, F_2, \dots$ de dimensión $d \times d$. Entonces se define la secuencia de **observaciones**, Z_n , que se obtiene aplicando una transformación lineal a X_n y añadiendo el vector aleatorio $V_n, t = 1, 2, \dots$. La ecuación de X_n se generaliza en la teoría de control incluyendo un término adicional $B_n U_n$, que representa el efecto de aplicar un control U_n en un instante de tiempo n .

Veamos algunas hipótesis necesarias para el uso y análisis de la ecuación de estado y la de observación.

1. $F_1, F_2, \dots$ es una secuencia de matrices deterministas $d \times d$.
2. $H_1, H_2, \dots$ es una secuencia de matrices deterministas $r \times d$.
3. $\{X_n, (V_n^T, W_n^T)^T, n = 1, 2, \dots\}$ es una secuencia ortogonal de vectores aleatorios con momentos de segundo orden finitos.
4. $\mathbb{E}[V_n] = 0$ y $\mathbb{E}[W_n] = 0 \quad n = 1, 2, \dots$
5. $\mathbb{E}[V_n V_n^T] = R_n, \mathbb{E}[W_n W_n^T] = Q_n$ donde las matrices R_n, Q_n son de dimensión $r \times r$ y $d \times d$ respectivamente.

Observación 1.2.1. En algunos casos es posible que las matrices F_n, H_n, R_n, Q_n sean independientes de n .

Observación 1.2.2. Por las características que presentan, X_n y Z_n tienen forma de función para $n = 1, 2, \dots$

$$X_n = f_n(X_1, W_1, \dots, W_n) \quad Z_n = h_n(X_1, W_1, \dots, W_n, V_n)$$

Observación 1.2.3. Por la Observación 1.1.2 y la hipótesis número 3 tenemos las relaciones de ortogonalidad siguientes

$$W_n \perp X_m, \quad W_n \perp Z_m, \quad V_n \perp X_m, \quad V_n \perp Z_m, \quad 1 \leq m < n$$

Ejemplo 1.2.4. (Una dirección aleatoria variante con ruido añadido). Si β es una constante, $W_n \sim N(0, \sigma^2)$ y Y_1 es una variable aleatoria incorrelacionada con $\{W_n, n = 1, 2, \dots\}$, entonces el proceso estocástico $\{Y_n, n = 1, 2, \dots\}$ definido por

$$Y_{n+1} = Y_n + \beta + W_n = Y_1 + n\beta + W_1 + \dots + W_n, \quad n = 1, 2, \dots$$

puede generar un modelo de estado en que las trayectorias son aproximadamente lineales.

Para construir una representación del estado para Y_n , introducimos el vector $X_n = [Y_n \ \beta]^T$, que se puede describir de forma equivalente como

$$X_{n+1} = \begin{bmatrix} Y_{n+1} \\ \beta \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} X_n + W_n, \quad n = 1, 2, \dots$$

Donde $W_n = [W_n^1 \ 0]^T$. El proceso Y_n viene determinado por $Y_n = [1 \ 0] X_n$. Se puede añadir el ruido V_n a Y_n de manera que nos queda la secuencia

$$Z_n = [1 \ 0] X_n + V_n$$

donde $V_n \sim N(0, v^2)$. Si $\{X_1, V_1, W_1, V_2, W_2, \dots\}$ es una sucesión ortogonal, entonces las ecuaciones de X_n y Z_n constituyen un modelo de estado para el proceso Y_n como quieramos demostrar. Para este modelo tenemos

$$F_n = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, H_n = [1 \ 0], Q_n = \begin{bmatrix} \sigma^2 & 0 \\ 0 & 0 \end{bmatrix}, R = v^2$$

Capítulo 2

Resultados previos

Veamos algunos resultados importantes de la esperanza condicionada en la estimación para entender la motivación del Filtro de Kalman.

Proposición 2.0.1. *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, $X := (X_k)_{1 \leq k \leq n}$ un proceso estocástico y $Z := (Z_k)_{1 \leq k \leq n}$ otro proceso estocástico relacionado con X , entonces el mejor estimador de X bajo el error cuadrático medio, llamado $\hat{X}_n$, es la esperanza condicionada, es decir, $\hat{X}_n = \mathbb{E}[X_n | Z]$.*

Demostración. Sea $F(Z_1, \dots, Z_n) = \hat{X}_n$ el estimador, el error cuadrático medio entre el estado y la estimación viene dado por,

$$\begin{aligned} \mathbb{E}[(X_n - \hat{X}_n)^2] &= \mathbb{E}[(X_n - \mathbb{E}[X_n | Z] + \mathbb{E}[X_n | Z] - \hat{X}_n)^2] \\ &= \mathbb{E}[(X_n - \mathbb{E}[X_n | Z])^2 + (\mathbb{E}[X_n | Z] - \hat{X}_n)^2 + 2(\mathbb{E}[X_n | Z] - \hat{X}_n)(X_n - \mathbb{E}[X_n | Z])] = \\ &= \mathbb{E}[(X_n - \mathbb{E}[X_n | Z])^2] + \mathbb{E}[(\mathbb{E}[X_n | Z] - \hat{X}_n)^2] + 2\mathbb{E}[(\mathbb{E}[X_n | Z] - \hat{X}_n)(X_n - \mathbb{E}[X_n | Z])]. \end{aligned}$$

Notemos que

$$\mathbb{E}[(\mathbb{E}[X_n | Z] - \hat{X}_n)(X_n - \mathbb{E}[X_n | Z])] = 0,$$

ya que por la propiedad 8 de la esperanza condicionada (véase anexo) se puede transformar en

$$\mathbb{E}[(\mathbb{E}[X_n | Z] - \hat{X}_n)(X_n - \mathbb{E}[X_n | Z])] = \mathbb{E}\left[\mathbb{E}[(\mathbb{E}[X_n | Z] - \hat{X}_n)(X_n - \mathbb{E}[X_n | Z]) | Z]\right]$$

como $\mathbb{E}[X_n | Z] - \hat{X}_n$ es Z -medible

$$\mathbb{E}\left[\mathbb{E}[(\mathbb{E}[X_n | Z] - \hat{X}_n)(X_n - \mathbb{E}[X_n | Z]) | Z]\right] = \mathbb{E}[(\mathbb{E}[X_n | Z] - \hat{X}_n)\mathbb{E}[(X_n - \mathbb{E}[X_n | Z]) | Z]]$$

entonces el término de dentro

$$\mathbb{E}[(X_n - \mathbb{E}[X_n | Z]) | Z] = \mathbb{E}[X_n | Z] - \mathbb{E}[X_n | Z] = 0$$

Tenemos ahora que

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

y como $\mathbb{E}[(\mathbb{E}[X_n|Z] - \hat{X}_n)^2] \geq 0$ nos queda lo que queríamos:

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

□

Proposición 2.0.2. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, sea $X, Z_1, \dots, Z_n$ variables aleatorias tales que $(X, Z_1, \dots, Z_n) \sim N_{n+1}(\mu, \Sigma)$ demostremos que $\mathbb{E}[X|Z_1, \dots, Z_n] = AZ + B$ donde $Z = (Z_1, \dots, Z_n)$, A matriz $1 \times n$ y $B \in \mathbb{R}$. Esto es, que la esperanza condicionada, en el caso que las observaciones y el estado sean conjuntamente gaussianas, es una función lineal de las observaciones.

Demostración. Sea

$$\Sigma = \begin{pmatrix} \text{Cov}(X, X) & \text{Cov}(X, Z_1) & \text{Cov}(X, Z_2) & \dots & \text{Cov}(X, Z_n) \\ \text{Cov}(Z_1, X) & \text{Cov}(Z_1, Z_1) & \dots & \dots & \text{Cov}(Z_1, Z_n) \\ \dots & \dots & \dots & \dots & \dots \\ \text{Cov}(Z_n, X) & \text{Cov}(Z_n, Z_1) & \text{Cov}(Z_n, Z_2) & \dots & \text{Cov}(Z_n, Z_n) \end{pmatrix}$$

Sea $Z^T = (Z_1, \dots, Z_n)$ y $Y = X - \Sigma_{21}\Sigma_{11}^{-1}Z$. Entonces $\text{Cov}(Y, Z) = 0$ y por la propiedad 10 de la esperanza condicionada (véase anexo),

$$\mathbb{E}[Y|Z] = \mathbb{E}[X|Z] - \Sigma_{21}\Sigma_{11}^{-1}Z = \mathbb{E}[Y]$$

por tanto queda,

$$\mathbb{E}[X|Z] = \mathbb{E}[Y] + \Sigma_{21}\Sigma_{11}^{-1}Z = \mathbb{E}[X] + \Sigma_{21}\Sigma_{11}^{-1}(Z - \mathbb{E}[Z])$$

□

Capítulo 3

Algoritmo del Filtro de Kalman

En el filtro de Kalman suponemos que tanto para la evolución del estado como para las observaciones tenemos un sistema lineal. Para la evolución del vector de variables aleatorias X_n de dimensión d , llamado **estado** para el instante t_n (denotemos $X_{t_n} \equiv X_n$) tenemos

$$X_n = F_n X_{n-1} + B_n U_n + W_n$$

donde F_n es la matriz $d \times d$ de transición (determinista), B_n es la matriz $d \times d$ de modelo de control de entrada (determinista) aplicada al vector de variables aleatorias predictivo de control U_n de dimensión d , es decir, $U_n \in \sigma(Z_1, \dots, Z_{n-1})$ (Z_n se definirá a continuación). Y $W_n \sim N_d(0, Q_n)$ es un proceso estocástico donde Q_n , para cada n , es la matriz de covarianzas de W_n . La componente W_n es la que hace que X_n sea aleatorio.

Suponemos también que el proceso estocástico de las **observaciones** del estado Z_n de dimensión r en el mismo instante sigue la ecuación

$$Z_n = H_n X_n + V_n$$

donde H_n es la matriz de observaciones $r \times d$ (determinista) y $V_n \sim N_r(0, R_n)$ es el proceso estocástico llamado **ruido** donde R_n , para cada n son las columnas de la matriz de covarianzas de V_n . Supondremos que $W_1, \dots, W_d$ y $V_1, \dots, V_r$ son mutuamente independientes.

3.1. Estimación del estado y covarianza del error en la estimación del estado

Proposición 3.1.1. *La estimación del estado en el instante n , dadas las observaciones hasta el instante $n - 1$, sigue la ecuación*

$$\hat{X}_{n|n-1} = F_n \hat{X}_{n-1|n-1} + B_{n-1} U_{n-1}$$

Señalamos que la notación $n|n-1$ se refiere a la esperanza condicionada y que $Z^{n-1} := (Z_1, \dots, Z_{n-1})$.

Demostración. Entonces,

$$\begin{aligned}\hat{X}_{n|n-1} &= \mathbb{E}[X_n|Z^{n-1}] \\ &= \mathbb{E}[F_n X_{n-1} + B_n U_n + W_n|Z^{n-1}] = F_{n-1} \mathbb{E}[X_{n-1}|Z^{n-1}] + B_n U_n + \mathbb{E}[W_n|Z^{n-1}]\end{aligned}$$

Donde hemos usado que U_n es Z^{n-1} -medible y que por lo tanto cumple que $\mathbb{E}[U_n|Z^{n-1}] = U_n$. Como $W_n \sim N(0, Q_n)$ y W_n es independiente de $Z^{n-1} \implies \mathbb{E}[W_n|Z^{n-1}] = \mathbb{E}[W_n] = 0$, y nos queda

$$\hat{X}_{n|n-1} = F_n \hat{X}_{n-1|n-1} + B_{n-1} U_{n-1}$$

□

Proposición 3.1.2. *La covarianza del error producido en la estimación del estado en el instante n dadas las observaciones hasta $n - 1$ sigue la ecuación*

$$P_{n|n-1} = F_n P_{n-1|n-1} F_n^T + Q_n$$

Demostración. Por definición,

$$P_{n|n-1} := \mathbb{E}[(X_n - \hat{X}_{n|n-1})(X_n - \hat{X}_{n|n-1})^T | Z^{n-1}] = *$$

Por un lado tenemos que

$$- \hat{X}_{n|n-1} = -F_n \hat{X}_{n-1|n-1} - B_n U_n$$

$$X_n = F_n X_{n-1} + B_n U_n + W_n$$

por otro lado notemos que $\mathbb{E}[(X_{n-1} - \hat{X}_{n-1|n-1})|Z^{n-1}] = 0$ entonces tenemos que,

$$* = F_n \mathbb{E}[(X_{n-1} - \hat{X}_{n-1|n-1})(X_{n-1} - \hat{X}_{n-1|n-1})^T | Z^{n-1}] F_n^T + \mathbb{E}[W_n W_n^T | Z^{n-1}]$$

Por la definición de covarianza y como $W_n \sim N(0, Q_n) \implies Q_n = \mathbb{E}[W_n W_n^T] - \mathbb{E}[W_n] \mathbb{E}[W_n^T] = \mathbb{E}[W_n W_n^T]$ y acaba quedando,

$$P_{n|n-1} = F_n P_{n-1|n-1} F_n^T + Q_n$$

□

3.2. Innovación

Definición 3.2.1. *En el Filtro de Kalman llamamos innovación del estado, $\hat{\delta}_n$, al error en la estimación de las observaciones,*

$$\hat{\delta}_n := Z_n - \hat{Z}_{n|n-1}$$

Proposición 3.2.2. *La innovación satisface que:*

$$\hat{\delta}_n = Z_n - H_n \hat{X}_{n|n-1}$$

Demostración. Por un lado por definición y desarrollando,

$$\hat{Z}_{n|n-1} := \mathbb{E}[Z_n | Z^{n-1}] = \mathbb{E}[H_n X_n + V_n | Z^{n-1}] = H_n \hat{X}_{n|n-1}$$

donde hemos usado que $V_1, \dots, V_n$ son independientes de $V_n \implies \mathbb{E}[V_n | Z^{n-1}] = \mathbb{E}[V_n] = 0$. Por otro lado por definición de innovación y sustituyendo lo anterior,

$$\hat{\delta}_n := Z_n - \hat{Z}_{n|n-1} = Z_n - H_n \hat{X}_{n|n-1}$$

□

Como no conocemos el estado real, no es posible compararlo con el estimador del estado. Por lo tanto la innovación es una herramienta útil a la hora de decidir si un estimador se adapta bien o no.

Proposición 3.2.3. *La covarianza de la innovación, cumple que:*

$$S_n = H_n P_{n|n-1} H_n^T + R_n$$

Demostración. Por definición y como la esperanza de la innovación es zero,

$$S_n := E[\hat{\delta}_n \hat{\delta}_n^T | Z^{n-1}] = E[(Z_n - H_n \hat{X}_{n|n-1})(Z_n - H_n \hat{X}_{n|n-1})^T | Z^{n-1}]$$

Usando que $Z_n = H_n X_n + V_n$

$$= E[(H_n(X_n - \hat{X}_{n|n-1}) + V_n)((H_n(X_n - \hat{X}_{n|n-1}) + V_n)^T | Z^{n-1}]$$

de forma análoga a la última demostración y como $V_n \sim N(0, R_n)$

$$S_n = H_n P_{n|n-1} H_n^T + R_n$$

□

3.3. La matriz de ganancia

3.3.1. Actualización del estimador del estado

Definición 3.3.1. *En el Filtro de Kalman definimos el error en la estimación del estado como*

$$\tilde{X}_{n|n} := \hat{X}_{n|n} - X_n$$

El objetivo del Filtro de Kalman es encontrar un estimador que sea combinación lineal de la predicción del estado y de la innovación, por eso empezaremos suponiendo que

$$\hat{X}_{n|n} = K_n' \hat{X}_{n|n-1} + K_n Z_n$$

donde K_n' y K_n son las matrices de ganancia que se eligen de forma que minimizen la expresión

$$\mathbb{E}[\tilde{X}_{n|n}^T \tilde{X}_{n|n} | Z^n] = \text{traza} \left[\mathbb{E}[\tilde{X}_{n|n} \tilde{X}_{n|n}^T | Z^n] \right] = \text{traza}[P_{n|n}]$$

La igualdad es consecuencia de que $\tilde{X}_{n|n}^T \tilde{X}_{n|n}$ resulta una matriz cuya suma de los términos de la diagonal es el producto escalar de estos dos vectores.

Proposición 3.3.2. : K'_n y K_n se relacionan a través de la expresión

$$K'_n = 1 - K_n H_n$$

Demostración. El error en la estimación del estado en términos de las matrices de ganancia se expresa como

$$\begin{aligned}\tilde{X}_{n|n} &:= \hat{X}_{n|n} - X_n = K'_n \hat{X}_{n|n-1} + K_n Z_n - X_n \\ &= K'_n \hat{X}_{n|n-1} + K_n (H_n X_n + V_n) - X_n\end{aligned}$$

Usando que $\hat{X}_{n|n-1} = X_n + \tilde{X}_{n|n-1}$,

$$= (K'_n + K_n H_n - 1)X_n + K'_n \tilde{X}_{n|n-1} + K_n V_n$$

Tomando esperanzas y sabiendo que $\mathbb{E}[V_n|Z^k] = 0$ y $\mathbb{E}[\tilde{X}_n|Z^k] = 0$

$$\mathbb{E}[\tilde{X}_n|Z^k] = (K'_n + K_n H_n - 1)\mathbb{E}[X_n|Z^k]$$

y como $\mathbb{E}[\tilde{X}_n|Z^k] = 0$ tiene que darse $K'_n = 1 - K_n H_n$ □

Corolario 3.3.3. La actualización del estimador del estado sigue la ecuación

$$\hat{X}_{n|n} = \hat{X}_{n|n-1} + K_n \hat{\delta}_n$$

Demostración.

$$\hat{X}_{n|n} = K'_n \hat{X}_{n|n-1} + K_n Z_n = (1 - K_n H_n) \hat{X}_{n|n-1} + K_n Z_n = \hat{X}_{n|n-1} + K_n \hat{\delta}_n$$

□

3.3.2. Cálculo de la matriz de ganancia

Seguimos calculando el error en la estimación sustituyendo, por un lado $K'_n = 1 - K_n H_n$ en $\hat{X}_{n|n}$

$$\hat{X}_{n|n} = (1 - K_n H_n) \hat{X}_{n|n-1} + K_n Z_n$$

y por otro lado sabemos que $\hat{X}_{n|n-1} = \tilde{X}_{n|n-1} + X_n$ y $Z_n = H_n X_n + V_n$,

$$= (1 - K_n H_n) \tilde{X}_{n|n-1} + X_n - \cancel{X_n K_n H_n} + \cancel{K_n H_n X_n} + K_n V_n$$

el error queda

$$\tilde{X}_n = ((1 - K_n H_n) \tilde{X}_{n|n-1} + \cancel{X_n} + K_n V_n) - \cancel{X_n}$$

Sigamos ahora calculando $P_{n|n}$ para encontrar K_n usando lo que tenemos hasta el momento.

$$\begin{aligned}P_{n|n} &= (1 - K_n H_n) \mathbb{E}[\tilde{X}_{n|n-1} \tilde{X}_{n|n-1}^T | Z^{n-1}] (1 - K_n H_n)^T + K_n \mathbb{E}[V_n V_n^T] K_n^T \\ &\quad + 2(1 - K_n H_n) \mathbb{E}[\tilde{X}_{n|n-1} V_n^T] K_n^T\end{aligned}$$

Señalamos que $\mathbb{E}[\tilde{X}_{n|n-1}V_n^T] = 0$, $P_{n|n-1} = \mathbb{E}[\tilde{X}_{n|n-1}\tilde{X}_{n|n-1}^T|Z^n]$, $R_n = \mathbb{E}[V_nV_n]$ y obtenemos

$$\begin{aligned} P_{n|n} &= (I - K_n H_n)P_{n|n-1}(I - K_n H_n)^T + K_n R_n K_n^T \\ &= P_{n|n-1} - K_n H_n P_{n|n-1} - P_{n|n-1} H_n^T K_n^T + K_n (H_n P_{n|n-1} H_n^T + R_n) K_n^T \\ &\quad P_{n|n-1} - K_n H_n P_{n|n-1} - P_{n|n-1} H_n^T K_n^T + K_n S_n K_n^T \end{aligned}$$

Donde hemos usado al final que $S_n = H_n P_{n|n-1} H_n^T + R_n$. Ahora pasemos a minimizar la expresión $\text{traza}(P_{n|n})$.

$$\begin{aligned} \frac{\partial \text{traza}(P_{n|n})}{\partial K_n} &= 0 - \frac{\partial \text{traza}(K_n H_n P_{n|n-1})}{\partial K_n} - \frac{\partial \text{traza}(P_{n|n-1} H_n^T K_n^T)}{\partial K_n} + \frac{\partial \text{traza}(K_n S_n K_n^T)}{\partial K_n} \\ &= 0 \end{aligned}$$

Aplicamos el hecho de que para cualquier matriz A y una matriz simétrica B

$$\begin{aligned} \frac{\partial \text{traza}(ABA^T)}{\partial K_n} &= 2AB \\ \frac{\partial \text{traza}(P_{n|n})}{\partial K_n} &= -2(H_n P_{n|n-1})^T + 2K_n S_n = 0 \\ K_n &= P_{n|n-1} H_n^T S_n^{-1} \end{aligned}$$

3.4. Pasos del algoritmo del Filtro de Kalman

Para esquematizar todo lo demostrado anteriormente, exponemos a continuación las ecuaciones y pasos del algoritmo:

3.4.1. Estimación

Estimación del estado y la covarianza del error en la estimación:

$$\begin{aligned} \hat{X}_{n|n-1} &= F_n \hat{X}_{n-1|n-1} + B_{n-1} U_{n-1} \\ P_{n|n-1} &= F_n P_{n-1|n-1} F_n^T + Q_n \end{aligned}$$

3.4.2. Actualización

Actualización del estimador del estado y la covarianza del error en la estimación en el instante n :

$$\begin{aligned} \hat{X}_{n|n} &= \hat{X}_{n|n-1} + K_n \hat{\delta}_n \\ P_{n|n} &= P_{n|n-1} - K_n H_n P_{n|n-1} - P_{n|n-1} H_n^T K_n^T + K_n S_n K_n^T \end{aligned}$$

donde la matriz de ganancia viene dada por la ecuación

$$K_n = P_{n|n-1} H_n^T S_n^{-1}$$

Capítulo 4

Filtros no lineales

El algoritmo del Filtro de Kalman visto anteriormente se limita al caso de sistemas lineales. Si las observaciones y/o los estados siguen una dinámica no lineal usamos otras versiones del Filtro de Kalman.

4.1. Filtro de Kalman Extendido

En este caso supondremos que:

$$X_n = F(X_{n-1}, U_n) + W_n, Z_n = H(X_n) + V_n$$

Donde F y H son funciones no lineales, U_n el vector de control y $V_n \sim N_r(0, R_n)$ y $W_n \sim N_d(0, Q_n)$, de forma análoga al filtro de Kalman.

Las funciones F, H pueden ser usadas para calcular la predicción de la observación y el estado, pero para calcular la covarianza hay que usar las Jacobianas de la funciones con el objetivo de linearizarlas.

4.1.1. Predicción del estado

Proposición 4.1.1. *La predicción del estado en el Filtro Extendido de Kalman es de la forma*

$$\hat{X}_{n|n-1} \approx F(\hat{X}_{n-1|n-1}, U_n) + O([X_{n-1} - \hat{X}_{n-1|n-1}]^2)$$

Demostración. Empezamos suponiendo que $\hat{X}_{n-1|n-1} \approx \mathbb{E}[X_{n-1}|Z^{n-1}]$. Para encontrar la predicción de X_n , desarrollamos Taylor de X_n en $\hat{X}_{n-1|n-1}$

$$X_n = F(\hat{X}_{n-1|n-1}, U_n) + \nabla F_{X_n}(X_{n-1} - \hat{X}_{n-1|n-1}) + O([X_{n-1} - \hat{X}_{n-1|n-1}]^2) + W_n$$

y lo sustituimos en,

$$\begin{aligned}\hat{X}_{n|n-1} &\approx \mathbb{E}[X_n|Z^{n-1}] \\ &\approx \mathbb{E}[F(\hat{X}_{n-1|n-1}, U_n) + \nabla F_{X_n}(X_{n-1} - \hat{X}_{n-1|n-1}) + O([X_{n-1} - \hat{X}_{n-1|n-1}]^2) + W_n|Z^{n-1}] \\ &= F(\hat{X}_{n-1|n-1}, U_n) + O([X_{n-1} - \hat{X}_{n-1|n-1}]^2)\end{aligned}$$

ya que $\nabla F_{X_n} \mathbb{E}[(X_{n-1} - \hat{X}_{n-1|n-1})|Z^{n-1}] \approx 0$ y $\mathbb{E}[W_n|Z^{n-1}] = 0$ □

Con esto podemos encontrar también el error de la predicción.

Corolario 4.1.2. *El error en la estimación del estado en el Filtro Extendido de Kalman sigue la ecuación:*

$$\tilde{X}_{n|n-1} \approx \nabla F_{X_n} \tilde{X}_{n-1|n-1} + W_n + O([X_{n-1} - \hat{X}_{n-1|n-1}]^2)$$

Demostración.

$$\begin{aligned}\tilde{X}_{n|n-1} &:= X_n - \hat{X}_{n|n-1} \\ &\approx \cancel{F(\hat{X}_{n-1|n-1}, U_n)} + \nabla F_{X_n}(X_{n-1} - \hat{X}_{n-1|n-1}) + O([X_{n-1} - \hat{X}_{n-1|n-1}]^2) + W_n \\ &\quad - \cancel{F(\hat{X}_{n-1|n-1}, U_n)} = \nabla F_{X_n}(\tilde{X}_{n-1|n-1}) + W_n + O([X_{n-1} - \hat{X}_{n-1|n-1}]^2)\end{aligned}$$

□

Se sigue que también podemos demostrar la forma de la covarianza del error en la estimación del estado.

Proposición 4.1.3. *La covarianza del error en la estimación del estado en el Filtro Extendido de Kalman sigue la ecuación:*

$$P_{n|n-1} \approx \nabla F_{X_n} P_{n-1|n-1} \nabla F_{X_n}^T + Q_n + O([X_{n-1} - \hat{X}_{n-1|n-1}]^4)$$

Demostración.

$$\begin{aligned}P_{n|n-1} &:= \mathbb{E}[\tilde{X}_{n|n-1} \tilde{X}_{n|n-1}^T | Z^{n-1}] \\ &\approx \mathbb{E}[(\nabla F_{X_n} \tilde{X}_{n-1|n-1} + W_n + O([X_{n-1} - \hat{X}_{n-1|n-1}]^2))(\nabla F_{X_n} \tilde{X}_{n-1|n-1} + W_n \\ &\quad + O([X_{n-1} - \hat{X}_{n-1|n-1}]^2))^T | Z^{n-1}] \\ &= \nabla F_{X_n} \mathbb{E}[\tilde{X}_{n-1|n-1} \tilde{X}_{n-1|n-1}^T | Z^{n-1}] \nabla F_{X_n}^T + \mathbb{E}[W_n W_n^T | Z^{n-1}] + \mathbb{E}[\nabla F_{X_n} \tilde{X}_{n-1|n-1} W_n^T | Z^{n-1}] \\ &\quad + \mathbb{E}[W_n \tilde{X}_{n-1|n-1}^T \nabla F_{X_n}^T | Z^{n-1}] + O([X_{n-1} - \hat{X}_{n-1|n-1}]^4)\end{aligned}$$

Y como por hipótesis W_n y $\tilde{X}_n$ están incorrelacionados (ya que W_n con $\hat{X}_n$ y X_n lo están) $\implies \mathbb{E}[W_n \tilde{X}_{n-1|n-1}^T \nabla F_{X_n}^T | Z^{n-1}] = 0$ y $\mathbb{E}[\nabla F_{X_n} \tilde{X}_{n-1|n-1} W_n^T | Z^{n-1}] = 0$

$$P_{n|n-1} \approx \nabla F_{X_n} P_{n-1|n-1} \nabla F_{X_n}^T + Q_n + O([X_{n-1} - \hat{X}_{n-1|n-1}]^4)$$

□

4.1.2. La predicción de la observación y la innovación

Es necesario calcularlas para actualizar el estimador del estado.

Proposición 4.1.4. *La estimación de la observación para el Filtro Extendido de Kalman sigue la ecuación $\hat{Z}_{n|n-1} \approx H(\hat{X}_{n|n-1}) + O([\hat{X}_{n|n-1} - X_n]^2)$.*

Demostración. Análogamente a los procedimientos anteriores hacemos el desarrollo de Taylor de Z_n en $\hat{X}_{n|n-1}$

$$Z_n = H(X_n) + V_n \approx H(\hat{X}_{n|n-1}) + \nabla H_{X_n}(\hat{X}_{n|n-1} - X_n) + O([\hat{X}_{n|n-1} - X_n]^2) + V_n$$

Sutituyéndolo en $\hat{Z}_{n|n-1} \approx \mathbb{E}[Z_n|Z^{n-1}]$

$$\approx \mathbb{E}[H(\hat{X}_{n|n-1}) + \nabla H_{X_n}(\hat{X}_{n|n-1} - X_n) + O([\hat{X}_{n|n-1} - X_n]^2) + V_n|Z^{n-1}]$$

Y usando que $\mathbb{E}[\tilde{X}_n|Z^{n-1}] = 0$ y $\mathbb{E}[V_n|Z^{n-1}] = 0$

$$\hat{Z}_{n|n-1} \approx H(\hat{X}_{n|n-1}) + O([\hat{X}_{n|n-1} - X_n]^2)$$

□

Señalamos que en el Filtro Extendido de Kalman hay una distinción entre la innovación y el error en la estimación de la observación. La innovación se calcula como $\hat{\delta}_n = Z_n - H(\hat{X}_{n|n-1})$ y el error en la estimación de la observación

$$\begin{aligned} \tilde{Z}_{n|n-1} &:= Z_n - \hat{Z}_{n|n-1} \\ &\approx \cancel{H(\hat{X}_{n|n-1})} + \nabla H_{X_n}(\hat{X}_{n|n-1} - X_n) + O([\hat{X}_{n|n-1} - X_n]^2) + V_n - \cancel{H(\hat{X}_{n|n-1})} \\ \tilde{Z}_{n|n-1} &\approx \nabla H_{X_{n|n-1}}(\hat{X}_{n|n-1} - X_n) + V_n + O([\hat{X}_{n|n-1} - X_n]^2) \end{aligned}$$

Finalmente podemos ver como se comporta la covarianza de la innovación.

Proposición 4.1.5. *La covarianza de la innovación en el Filtro Extendido de Kalman es de la forma*

$$S_n \approx \nabla H_{X_n} P_{n|n-1} \nabla H_{X_n}^T + R_n + O([\hat{X}_{n|n-1} - X_n]^2)$$

Demostración.

$$\begin{aligned} S_n &:= \mathbb{E}[\tilde{Z}_{n|n-1} \tilde{Z}_{n|n-1}^T | Z^{n-1}] \\ &\approx \mathbb{E}[(\nabla H_{X_{n|n-1}}(\hat{X}_{n|n-1} - X_n) + V_n + O([\hat{X}_{n|n-1} - X_n]^2)) \\ &\quad (\nabla H_{X_{n|n-1}}(\hat{X}_{n|n-1} - X_n) + V_n + O([\hat{X}_{n|n-1} - X_n]^2))^T | Z^{n-1}] \end{aligned}$$

Y de forma análoga al cálculo de la covarianza del estado

$$= \nabla H_{X_n} P_{n|n-1} \nabla H_{X_n}^T + R_n + O([\hat{X}_{n|n-1} - X_n]^4)$$

□

4.1.3. Actualización de ecuaciones

El siguiente paso es encontrar el estimador lineal recursivo del estado $\hat{X}_{n|n}$. Para ello suponemos $\hat{X}_{n|n} = \hat{X}_{n|n-1} + K_n(Z_n - H(\hat{X}_{n|n-1}))$, análogo al caso lineal. Como sabemos por el capítulo 3, la matriz de ganancia minimiza

$$\text{traza}(P_{n|n})$$

Para tener $P_{n|n}$ es necesario calcular $\tilde{X}_{n|n}$

$$\tilde{X}_{n|n} = \hat{X}_{n|n} - X_n = \hat{X}_{n|n-1} + K_n(Z_n - H(\hat{X}_{n|n-1})) - X_n$$

Usando el desarrollo de Taylor de Z_n en $\hat{X}_{n|n-1}$,

$$\begin{aligned} &\approx \hat{X}_{n|n-1} + K_n(\underbrace{(H(\hat{X}_{n|n-1}))}_{-H(\hat{X}_{n|n-1})} + \nabla H_{X_n}(\hat{X}_{n|n-1} - X_n) + V_n + O([\hat{X}_{n|n-1} - X_n]^2)) \\ &\quad - \underbrace{H(\hat{X}_{n|n-1})}_{-H(\hat{X}_{n|n-1})} - X_n = (I - K_n \nabla H_{X_n}) \tilde{X}_{n|n-1} + K_n V_n + O([\hat{X}_{n|n-1} - X_n]^2) \end{aligned}$$

Luego

$$\begin{aligned} P_{n|n} &:= \mathbb{E}[\tilde{X}_{n|n} \tilde{X}_{n|n}^T | Z^{n-1}] \\ &\approx (I - K_n \nabla H_{X_n}) \mathbb{E}[\tilde{X}_{n|n-1} \tilde{X}_{n|n-1}^T | Z^{n-1}] (I - K_n \nabla H_{X_n})^T + K_n \mathbb{E}[V_n V_n^T | Z^{n-1}] K_n^T \\ &\quad + O([\hat{X}_{n|n-1} - X_n]^4) \approx (I - K_n \nabla H_{X_n}) P_{n|n-1} (I - K_n \nabla H_{X_n})^T + K_n R_n K_n^T \\ &\quad + O([\hat{X}_{n|n-1} - X_n]^4) \end{aligned}$$

Derivamos respecto K_n e igualamos a 0.

$$\frac{\delta \text{traza}(P_{n|n})}{K_n} = -2(I - K_n \nabla H_{X_n}) P_{n|n-1} \nabla H_{X_n}^T + 2K_n R_n = 0$$

Podemos sumar las correspondientes derivadas porque son vectores. Finalmente

$$K_n = P_{n|n-1} \nabla H_{X_n}^T (\nabla H_{X_n} P_{n|n-1} \nabla H_{X_n}^T + R_n)^{-1} = P_{n|n-1} \nabla H_{X_n}^T S_n^{-1}$$

4.2. El Filtro de la Partícula

4.2.1. Filtro Recursivo Bayesiano

Veamos este método que usaremos en el Filtro de la Partícula. Suponemos que tenemos dos ecuaciones, para el estado y las observaciones, que no son lineales,

$$x_n = f_n(x_{n-1}, u_n) \quad Z_n = g_n(x_n, v_n)$$

Donde f_n es la función de transición entre los instantes $n-1$ y n y g_n la función de observación. Este método usa que toda la información necesaria para aproximar $x_0, \dots, x_n$ está contenida en la función de filtraje, es decir en $f(x_0, \dots, x_n | z^n)$. Por el Teorema de

Bayes y usando que $f(x_0, \dots, x_n | z^{n-1}) = f(x_n | x_0, \dots, x_n, z^{n-1})f(x_0, \dots, x_{n-1} | z^{n-1})$ tenemos que sigue la ecuación

$$f(x_0, \dots, x_n | z^n) = \frac{f(z_n | z^{n-1}, x_0, \dots, x_n) f(x_n | x_0, \dots, x_n, z^{n-1}) f(x_0, \dots, x_{n-1} | z^{n-1})}{\int_{x_0, \dots, x_n} f(z_n | z^{n-1}, x_0, \dots, x_n) f(x_0, \dots, x_n | z^{n-1}) dx_0, \dots, dx_n}$$

Sabemos que $f(z_n | z^{n-1}, x_0, \dots, x_n) = f(z_n | x_n)$ ya que las observaciones en diferentes instantes son independientes entre ellas dados los estados y no dependen del estado en el instante actual. Y $f(x_n | x_0, \dots, x_{n-1}, z^{n-1}) = f(x_n | x_{n-1})$ ya que el estado actual solo depende del estado anterior, entonces queda

$$f(x_0, \dots, x_n | z^n) = \frac{f(z_n | z^{n-1}) f(x_n | x_{n-1}) f(x_0, \dots, x_{n-1} | z^{n-1})}{\int_{x_0, \dots, x_n} f(z_n | x_n) f(x_0, \dots, x_n | z^{n-1}) dx_0, \dots, dx_n}$$

Esta ecuación es recursiva y se descompone en dos pasos:

1. El primer paso es aproximar $f(x_0, \dots, x_n | z^{n-1})$ con las probabilidades $f(x_n | x_{n-1})$ y $f(x_0, \dots, x_{n-1} | z^{n-1})$.
2. El segundo es obtener $f(x_0, \dots, x_n | z^n)$ usando $f(z^n | z_n)$.

Para obtener estimación de x_n , usamos la esperanza condicionada (demostrado anteriormente que es el mejor estimador)

$$\mathbb{E}[F(x_0, \dots, x_n)] = \int_{x_0, \dots, x_n} F(x_0, \dots, x_n) f(x_0, \dots, x_n | z_n) dx_0 \cdots dx_n$$

donde F es una función ponderada. Si las densidades son normales entonces la estimación se calcula directamente con el Filtro de Kalman (visto anteriormente). Sino, tenemos dos tipos de métodos:

1. Analíticos: como el Filtro Extendido de Kalman.
2. Numéricos: en que se calcula la aproximación de forma empírica y dividiendo por partes la función de densidad, como es el caso del **Filtro de la Partícula**.

4.2.2. Muestreo por Importancia

Antes de detallar el Filtro de la Partícula debemos introducir el Muestreo por Importancia. El Muestreo por Importancia es uno de los Métodos de Monte Carlo. Estos métodos tienen el propósito de generar muestras $x_0, \dots, x_n$ a partir de una distribución de probabilidad dada o bien estimar esperanzas de funciones bajo una distribución determinada. Se asume que dicha distribución es lo suficientemente compleja para que estas esperanzas no puedan evaluarse por métodos exactos. Una de las propiedades clave de Monte Carlo es que la exactitud de la estimación no depende de la dimensión del espacio muestreado.

El Muestreo por Importancia, en concreto, es un método para estimar esperanzas de funciones bajo una distribución determinada f . Para ello usamos una función de densidad más simple q a partir de la cual podemos generar muestras $x_0^{(i)}, \dots, x_n^{(i)}$. Si estos puntos fueran muestras de f , entonces podríamos hacer la estimación que queremos. Pero las muestras que generamos con q son valores que evaluados en q son mayores o menores que f . Es por ello que se introduce el concepto de peso, para ajustar la importancia de cada punto. Lo aplicamos a continuación a nuestro caso.

Una vez que hemos aproximado $f(x_0, \dots, x_n | z^n)$, hay que evaluar el estimador

$$\int_{x_0, \dots, x_n} F(x_0, \dots, x_n) f(x_0, \dots, x_n | z^n) dx_0 \cdots dx_n$$

. El método de Monte Carlo nos permite aproximar esta integral mediante la construcción de una variable aleatoria acorde con $f(x_0, \dots, x_n | z^n)$. Para ello se usa una **función propuesta**, $q(x_0, \dots, x_n | z^n)$, que contiene a $f(x_0, \dots, x_n | z^n)$ y se puede muestrear, de manera que

$$\begin{aligned} \mathbb{E}_f[F(x_0, \dots, x_n)] &= \int_{x_0, \dots, x_n} f(x_0, \dots, x_n) \frac{f(x_0, \dots, x_n | z^n)}{q(x_0, \dots, x_n | z^n)} q(x_0, \dots, x_n | z^n) dx_0 \cdots dx_n \\ &= \mathbb{E}_Q \left[F(x_0, \dots, x_n) \frac{f(x_0, \dots, x_n | z^n)}{q(x_0, \dots, x_n | z^n)} \right] \end{aligned}$$

Para N iteraciones suponemos que $x_0^{(i)}, \dots, x_n^{(i)} \sim q(x_0, \dots, x_n | z_1, \dots, z_n)$ $i = 1, \dots, N$, y podemos aproximar el estimador anterior por

$$\widehat{\mathbb{E}}_f[F(x_0, \dots, x_n)] = \frac{1}{N} \sum_{i=1}^N F(x_0^{(i)}, \dots, x_n^{(i)}) \frac{f(x_0^{(i)}, \dots, x_n^{(i)} | z^n)}{q(x_0^{(i)}, \dots, x_n^{(i)} | z^n)}$$

Por la Ley de los Grandes Números podemos asegurar que este estimador converge a $\mathbb{E}_f[F(x_0, \dots, x_n)]$ cuando N tiende a infinito. Definimos las **ponderaciones por importancia** como

$$w_n^{(i)} = \frac{f(x_0^{(i)}, \dots, x_n^{(i)} | z^n)}{q(x_0^{(i)}, \dots, x_n^{(i)} | z^n)} = \frac{f(z^n | x_0^{(i)}, \dots, x_n^{(i)}) f(x_0^{(i)}, \dots, x_n^{(i)})}{f(z^n) q(x_0^{(i)}, \dots, x_n^{(i)} | z^n)}$$

donde $f(z_1, \dots, z_n) = \int_{x_0, \dots, x_n} f(z^n | x_0^{(i)}, \dots, x_n^{(i)}) f(x_0, \dots, x_n) dx_0 \cdots dx_n$, cuya expresión es difícil de calcular, pero de todas formas se puede usar que

$$\widehat{\mathbb{E}}_f[F(x_0, \dots, x_n)] = \frac{1}{N} \sum_{i=1}^N F(x_0^{(i)}, \dots, x_n^{(i)}) \frac{w_n^{(i)}}{\sum_{j=1}^N w_n^{(j)}}$$

con

$$\lim_{n \rightarrow \infty} w_n^{(i)} \approx \frac{f(z^n | x_0^{(i)}, \dots, x_n^{(i)}) f(x_0, \dots, x_n)}{q(x_0^{(i)}, \dots, x_n^{(i)} | z^n)}$$

Ahora se puede hacer el Muestreo por Importancia recursivo, para obtener así el algoritmo del Filtro de la Partícula.

4.2.3. El Filtro de la Partícula

La idea por lo tanto es representar y aproximar de forma empírica la función de probabilidad $f(x_0, \dots, x_n | z^n)$ a partir de una muestra ponderada de tamaño N de la forma $\{x_0^{(i)}, \dots, x_n^{(i)}, w_n^{(i)}\}$

$$f(x_0, \dots, x_n | z^n) \approx \frac{1}{N} \sum_{i=1}^N w_n^{(i)} \delta_{x_0^{(i)}, \dots, x_n^{(i)}}(x_0, \dots, x_n)$$

Donde $x_0^{(i)}, \dots, x_n^{(i)}$ se llaman **partículas** y δ es la **delta de Dirac**, una función de distribución que nos sirve para aproximar la distribución en cuestión. Cada partícula es una posible solución de la aproximación del estado y su ponderación asociada representa la calidad de éste acorde con las observaciones que disponemos.

La muestra $S_n = \{x_0^{(i)}, \dots, x_n^{(i)}, w_n^{(i)}\}$ se calcula a partir de la muestra anterior, al igual que $f(x_0, \dots, x_n | z_1, \dots, z_n)$. Para ello es necesario ejecutar los siguientes pasos:

1. Una exploración del estado, durante la cual propagamos las partículas a través de la función propuesta.
2. Corrección de la calidad de la partícula, que consiste en calcular la nueva ponderación.
3. **Remuestreo**: una selección de las partículas.

El esquema general del Filtro de la Partícula entre los instantes $n - 1$ y n se puede resumir en el siguiente algoritmo:

1. Representación de $f(x_0, \dots, x_{n-1} | z_1, \dots, z_{n-1})$ a través del conjunto de partículas

$$\{x_0^{(i)}, \dots, x_{n-1}^{(i)}, w_{n-1}^{(i)}\}, i = 1, \dots, N$$

2. Propagación, o exploración del estado con una función propuesta

$$x_n^{(i)} \approx q(x_n | x_0^{(i)}, \dots, x_{n-1}^{(i)}, z^n)$$

3. Corrección o evaluación de la calidad de la partícula con las observaciones, calculando las ponderaciones:

$$w_n^{(i)} = \frac{f(x_0^{(i)}, \dots, x_n^{(i)} | z^n)}{q(x_0^{(i)}, \dots, x_n^{(i)} | z^n)}$$

$$\lim_{n \rightarrow \infty} w_n^{(i)} \approx \frac{f(z_n | x_n^{(i)}) f(x_n^{(i)} | x_{n-1}^{(i)} | z^{n-1}) f(x_0^{(i)}, \dots, x_{n-1}^{(i)} | z^{(n-1)})}{q(x_0^{(i)}, \dots, x_n^{(i)} | z^n)}$$

Suponiendo que $q(x_n^{(i)} | z^n) = q(x_n^{(i)} | x_0^{(i)}, \dots, x_{n-1}^{(i)}, z^{(n)}) q(x_0^{(i)}, \dots, x_{n-1}^{(i)} | z^{(n-1)})$, tenemos que

$$\lim_{n \rightarrow \infty} w_n^{(i)} \approx \frac{f(z_n | x_n^{(i)}) f(x_n^{(i)} | x_{n-1}^{(i)} | z^{n-1}) f(x_0^{(i)}, \dots, x_{n-1}^{(i)} | z^{(n-1)})}{q(x_n^{(i)} | x_0^{(i)}, \dots, x_{n-1}^{(i)}, z^{(n)}) q(x_0^{(i)}, \dots, x_{n-1}^{(i)} | z^{(n-1)})}$$

$$\lim_{n \rightarrow \infty} w_n^{(i)} \approx w_{n-1}^{(i)} \frac{f(z_n | x_n^{(i)}) f(x_n | x_{n-1}^{(i)})}{q(X_n^{(i)} | x_0^{(i)}, \dots, x_{n-1}^{(i)})}$$

Y las ponderaciones se normalizan $\tilde{w}_n^{(i)} = \frac{w_n^{(i)}}{\sum_{j=1}^N w_n^{(j)}}$

4. Aproximación de la función de distribución esperada del filtraje mediante

$$\mathbb{E}(F(x_0, \dots, x_n)) \approx \frac{1}{N} \tilde{w}_n^{(i)} F(x_0^{(i)}, \dots, x_n^{(i)})$$

5. Remuestreo.

Filtro Extendido de la Partícula

El algoritmo que describimos a continuación es un caso particular del Filtro de la Partícula que usa el Filtro Extendido de Kalman:

1. Para $n = 0$ escogemos x_0 y $p_0 > 0$. Para i tal que $1 \leq i \leq N$ cogemos $x_0^{(i)} = x_0 + \sqrt{p_0} u^{(i)}$ donde $u^{(i)}$ es un número simulado distribuido de forma Normal. También cogemos $p_0^{(i)} = p_0$ y $w_0^{(i)} = 1/N$.

Mientras $1 \leq k \leq N$

2. Para cada simulación de i se calcula $\hat{x}_n^{(i)}$ con el Filtro Extendido de Kalman y su asociada $p_n^{(i)}$ (covarianza)

3. Para cada i entre 1 y N

$$\tilde{x}_n^{(i)} = \hat{x}_n^{(i)} + \sqrt{p_n^{(i)}} u^{(i)}$$

4. Calculamos las ponderaciones asociadas

$$w_n^{(i)} = w_{n-1}^{(i)} \frac{f(z_n | \tilde{x}_n^{(i)}) f(\tilde{x}_n^{(i)} | x_{n-1}^{(i)})}{q(\tilde{x}_n^{(i)} | x_{n-1}^{(i)}, z^n)}$$

5. Normalizamos las ponderaciones

$$\tilde{w}_n^{(i)} = \frac{w_{n-1}^{(i)}}{w_{n-1}^{\sum_{i=1}^N (i)}}$$

6. Remuestreamos $\tilde{x}_n^{(i)}$ y escogemos $x_n^{(i)}$ y restablecemos $w_n^{(i)} = \tilde{w}_n^{(i)} = 1/N$

7. Incrementamos n , volvemos al paso 2. hasta el final del ciclo.

Estimación de los parámetros

Para estimar el conjunto de parámetros de un sistema, debemos maximizar la función de verosimilitud. La verosimilitud en el instante n sería

$$l_n = f(z_n|z^n) = \int f(z_n|x_n)f(x_n|z^n)$$

Entonces la función de verosimilitud total es el producto de todas las l_n y por tanto la función logarítmica de verosimilitud a maximizar será

$$\ln(L_1, \dots, L_n) = \sum_{n=1}^N \ln(l_n)$$

donde l_n puede ser descrita como

$$l_n = \int f(z_n|x_n) \frac{f(x_n|z^{n-1})}{q(x_n|x_{n-1}, z^n)} q(x_n|x_{n-1}, z^n) dx_n$$

$\tilde{x}_n^{(i)}$ se distribuye acorde a Q , considerando el restablecimiento de $w_n^{(i)}$ a una constante $1/N$ durante el paso de remuestreo, podemos aproximar l_n por

$$\tilde{l}_n = \sum_{i=1}^N w_n^{(i)}$$

que proporciona una interpretación de la verosimilitud como una ponderación total.

4.2.4. Aplicación: Modelos de Volatilidad Estocástica

El Filtro de la Partícula se usa en numerosas areas como por ejemplo en la comunicación, la biología, economía, geociencia, ciencias sociales,... La volatilidad es la inestabilidad de los precios en los mercados financieros. Es decir, volatilidad es lo que hace variar la rentabilidad de un activo financiero respecto de su media a lo largo de un periodo de tiempo determinado.

En esta sección estudiaremos el Modelo de Heston, utilizado por los profesionales financieros para valorar opciones de precios. Este utiliza métodos estadísticos para calcular y pronosticar el precio de las opciones con el supuesto de que la volatilidad es arbitraria. La suposición de que la volatilidad es arbitraria en lugar de constante es el factor clave que hace que los modelos estocásticos de volatilidad sean únicos. Otros tipos de modelos de volatilidad estocástica incluyen el modelo SABR, el modelo Chen y el modelo GARCH.

Estudiemos el modelo de Volatilidad estocástico de Heston Euler-discretizado

$$\ln(S_n) = \ln(S_{n-1}) + (r_n - 0,5v_n)T + \sqrt{v_n}\sqrt{T}V_{n-1}$$

$$v_n = v_{n-1} + (w - \theta v_{n-1})T + \xi\sqrt{v_{n-1}}\sqrt{T}W_{n-1}$$

donde S_{n-1} es el precio de mercado en el instante n , T es el intervalo de tiempo entre instantes, r_n la tasa libre de riesgo, v_n la varianza del mercado y V_n, W_n dos secuencias incorrelacionadas temporalmente (no lo están en cambio V_n, W_n) de variables aleatorias Normales de correlación ρ . El conjunto de parámetros del modelo es $\Psi = (\omega, \theta, \xi, \rho)$.

Vamos ahora a eliminar la correlación entre el sistema y el ruido de las observaciones. Consideramos v_n como el estado oculto X_n y $\ln(S_{n+1})$ como la ecuación de observaciones Z_n . Pasando $\ln(S_n)$ al otro lado de la igualdad de la primera ecuación, tenemos que

$$v_n = v_{n-1} + (\omega - \theta v_{n-1})T + \xi \sqrt{v_{n-1}} \sqrt{T} W_{n-1} - \rho \xi [\ln(S_n) - \ln(S_{n-1}) + (r_n - 0,5v_n)T + \sqrt{v_n} \sqrt{T} V_{n-1}]$$

Definiendo

$$\widetilde{W}_n = \frac{1}{\sqrt{1 - \rho^2}} (W_n - \rho V_n)$$

tendremos como queríamos $\widetilde{W}_n$ incorrelacionado con V_n y

$$X_n = v_n = v_{n-1} + [\omega - \rho \xi r_n - (\theta - 0,5\rho \xi) v_{n-1}]T + \rho \xi \ln\left(\frac{S_n}{S_{n-1}}\right) + \xi \sqrt{1 - \rho^2} \sqrt{v_{n-1}} \sqrt{T} \widetilde{W}_{n-1}$$

y $Z_n = \ln(S_{n+1})$

Filtro de la partícula

Podemos aplicar el algoritmo del Filtro de la Partícula a este problema. Estableciendo la notación

$$N(X, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

entonces definimos

$$Q(\widetilde{X}_n^{(i)} | X_{n-1}^{(i)}, Z^n) = N(\widetilde{X}_n^{(i)}, \widehat{X}_n^{(i)}, \sqrt{P_n^{(i)}})$$

$$\mathbb{P}(Z_n | \widetilde{X}_n^{(i)}) = N(Z_n, Z_{n-1} + (r_n - 0,5\widetilde{X}_n^{(i)})T, \sqrt{\widetilde{X}_n^{(i)}}\sqrt{T})$$

y

$$\mathbb{P}(\widetilde{X}_n^{(i)} | X_{n-1}^{(i)}) = N(\widetilde{X}_n^{(i)}, \mu_X, \xi \sqrt{1 - \rho^2} (X_{n-1}^{(i)})^p \sqrt{T})$$

donde

$$\begin{aligned} \mu_X = & X_{n-1}^{(i)} + [\omega - \rho \xi r_n (X_{n-1}^{(i)})^{p-0,5} - (\theta - 0,5\rho \xi (X_{n-1}^{(i)})^{p-0,5}) X_{n-1}^{(i)}]T \\ & + \rho \xi (X_{n-1}^{(i)})^{p-0,5} (Z_{n-1} - Z_{n-2}) \end{aligned}$$

De esta manera la ecuación recursiva

$$w_n^{(i)} = w_{n-1}^{(i)} \frac{\mathbb{P}(Z_n | \widetilde{X}_n^{(i)}) \mathbb{P}(\widetilde{X}_n^{(i)} | X_{n-1}^{(i)})}{Q(\widetilde{X}_n^{(i)} | X_{n-1}^{(i)}, Z^n)}$$

proporciona la implementación del filtro.

4.2.5. Estimación de los parámetros

Por el Filtro de la Partícula, debemeos maximizar

$$\sum_{n=1}^N \ln\left(\sum_{i=1}^N w_n^{(i)}\right)$$

para encontrar el conjunto de parámetros $\hat{\Psi} = (\hat{\omega}, \hat{\theta}, \hat{\xi}, \hat{\rho})$. Podemos escoger $\hat{\Psi}^* = (0,02, 0,5, 0,05, -0,5)$. Tomando $S_0 = 1000$, $v_0 = 0,04$, $r_n = 0,027$ y $T = 1$, donde $\nu_n = Z_n - \hat{Z}_n^-$. La serie temporal se genera a través de las primeras ecuaciones de v_n y $\ln(S_n)$. Entonces los parámetros óptimos de estimación serian: $\hat{\Psi}_{EPF} = (0,019357, 0,500021, 0,033354, -0,581794)$

Capítulo 5

Implementación del Filtro de Kalman

Normalmente se invierte más esfuerzo en ajustar un filtro que en usar su algoritmo. La dificultad en la implementación es consecuencia de no saber cuando el rendimiento del filtro está limitado por la adecuación del sensor y los procesos del modelo.

Habitualmente para saber si podemos mejorar nuestros modelos o estamos limitados por el rendimiento del algoritmo, tendremos que comprobarlo usando otros posibles filtros con respecto a las mismas observaciones que se han hecho.

Aun y así existen algunas reglas para enfocar mejor el problema de la implementación:

1. **Entender el sensor:** familiarizarse con las características que trata este e intentar obtener el máximo de información posible en diferentes situaciones con el sensor que estemos usando.
2. **Entender el proceso:** otra vez también familiarizarse con el proceso es fundamental: observar el proceso en diferentes modos operativos, con instrumentos adicionales si hace falta...
3. **Modelar el sensor:** en el Filtro de Kalman esto se reduce a la construcción de un apropiado H_n y W_n .
4. **Modelar el proceso:** el primer paso de este punto es construir un "modelo verdadero" lo más preciso posible describiendo todos las componentes del proceso para ser estimado. Este modelo será demasiado largo pero servirá para entender qué estados son importantes y cuales no. El segundo paso si que será reducir el modelo a aquellos estados que tengan más impacto en el rendimiento del filtro y construir apropiadamente F_n y V_n .
5. **Codificar el filtro:** es la parte más sencilla: calcular las diversas componentes del algoritmo, estimaciones de los estados, covarianzas...

6. **Inicialización:** como el filtro de Kalman es un algoritmo recursivo debemos aportar una "aproximación" para $\hat{X}_{0|0}$ y $P_{0|0}$. Aunque a largo plazo el efecto que tenga en el rendimiento disminuye, la inicialización es un paso importante.
7. **Análisis de la secuencia de la innovación:** es el primer método y el más efectivo para analizar el rendimiento del filtro.
8. **Análisis del rendimiento:** sabemos que la innovación se ve afectada por las observaciones y los modelos del proceso. Por lo tanto se debe separar estos dos factores y analizar simultáneamente las secuencia de observaciones y la estabilidad de las componentes del filtro; el estimador del estado y su covarianza.
9. **Análisis de las condiciones de error:** eliminar errores o información falsa en el proceso de filtraje.

En las siguientes secciones veremos un caso práctico de implementación del Filtro de Kalman.

5.1. Planteamiento general

El ejemplo que trataremos es el problema de la velocidad constante unidimensional de una partícula en movimiento. Las ecuaciones del estado del modelo vienen dadas por

$$\begin{bmatrix} X_n \\ \dot{X}_n \end{bmatrix} = \begin{bmatrix} 1 & \Delta T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_{n-1} \\ \dot{X}_{n-1} \end{bmatrix} + \begin{bmatrix} \Delta T^2/2 \\ \Delta T \end{bmatrix} W_n$$

Donde ΔT es el intervalo de tiempo que pasa entre el instante $n - 1$ y n , y suponemos que $W_n \sim N_2(0, Q_n)$ donde

$$Q_n = \begin{bmatrix} \Delta T^3/3 & \Delta T^2/2 \\ \Delta T^2/2 & \Delta T \end{bmatrix} \sigma_q^2$$

Vamos a suponer que las observaciones de la localización de la partícula se hacen en el instante n y que siguen la ecuación

$$Z_n = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} X_n \\ \dot{X}_n \end{bmatrix} + V_n$$

Donde $V_n \sim N(0, R_n)$ y $R_n = \sigma_r^2$. Los parámetros de entrada del algoritmo son

$$F_n = \begin{bmatrix} 1 & \Delta T \\ 0 & 1 \end{bmatrix}, H_n = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

V_n y W_n son calculados usando un generador de números pseudo-aleatorios para proporcionar una secuencia de números distribuidos como una normal de media 0 y covarianzas σ_q^2 y σ_r^2 respectivamente.

5.2. Inicialización

El primer paso es escoger $\hat{X}_{0|0}$ y su $P_{0|0}$. Como hemos dicho antes, el efecto de estas estimaciones iniciales disminuye con el tiempo. En un filtro lineal simple, este convergerá a la misma trayectoria estable independientemente de estos valores iniciales. Aún y así la elección de un buen estimador inicial mejora la convergencia y es esencial en el caso del Filtro Extendido de Kalman. Inicializaremos el filtro con $\sigma_q^2 = 0,01$ y $\sigma_r^2 = 0,1$ y usando las primeras observaciones para obtener una estimación de la posición y la velocidad

$$\hat{X}_{1|1} = Z_1, \hat{\dot{X}}_{1|1} = \frac{Z_2 - Z_1}{\Delta T}$$

El método más simple para inicializar $P_{0|0}$ es escoger una constante α y establecer $P_{0|0} = \alpha^2 Q_n$ (normalmente $\alpha = 10$).

5.3. Estabilidad del rendimiento

Vamos a calcular ahora la actualización de la estimación de la posición y la velocidad. Recordemos que

$$\hat{X}_{n|n} = (1 - K_n H_n) \hat{X}_{n|n-1} + K_n Z_n$$

Donde $K_n = P_{n|n-1} H_n^T S_n^{-1}$. Calculamos entonces

$$K_n = P_{n|n-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} S_n^{-1}$$

Sabemos que $S_n = H_n P_{n|n-1} H_n^T + R_n$. Por lo tanto

$$S_n = \begin{bmatrix} 1 & 0 \end{bmatrix} P_{n|n-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \sigma_r^2 = P_{n|n-1}^{xx} + \sigma_r^2$$

Donde $P_{n|n-1}^{xx}$ indica la componente de la fila 1 y la columna 1 de la matriz $P_{n|n-1}$. Usando esto nos queda que

$$K_n = \begin{bmatrix} \frac{P_{n|n-1}^{xx}}{P_{n|n-1}^{xx} + \sigma_r^2} \\ \frac{P_{n|n-1}^{x\dot{x}}}{P_{n|n-1}^{xx} + \sigma_r^2} \end{bmatrix}$$

Finalmente sustituyendo K_n y calculando $1 - K_n H_n$, encontramos que la actualización de la estimación de la posición de la partícula sigue la ecuación

$$\hat{X}_{n|n} = \frac{\sigma_r^2}{P_{n|n-1}^{xx} + \sigma_r^2} \hat{X}_{n|n-1} + \frac{P_{n|n-1}^{xx}}{P_{n|n-1}^{xx} + \sigma_r^2} Z_n$$

Y usando que $\hat{X}_{n|n} = \hat{X}_{n|n-1} + K_n \hat{\delta}_n$, podemos hallar la actualización de la estimación de la velocidad

$$\hat{\dot{X}}_{n|n} = \hat{\dot{X}}_{n|n-1} + \frac{P_{n|n-1}^{x\dot{x}}}{P_{n|n-1}^{xx} + \sigma_r^2} (Z_n - \hat{X}_{n|n-1})$$

(Cálculo de cada componente). Donde hemos usado que $\hat{\delta}_n = Z_n - H_n \hat{X}_{n|n-1} = Z_n -$

$$\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{X}_{n|n-1} \\ \hat{\dot{X}}_{n|n-1} \end{bmatrix} = Z_n - \hat{X}_{n|n-1}$$

5.4. Consistencia del filtro

Definición 5.4.1. un estimador se llama **consistente** si $\hat{X}_{n|n} = \mathbb{E}[X_n|Z^n]$ y $\mathbb{E}[(X_n - \hat{X}_{n|n})(X_n - \hat{X}_{n|n})^T|Z^n] = P_{n|n}$.

En principio para comprobar la consistencia del filtro, deberíamos construir el error cuadrático medio normalizado entre el estado y el estimador

$$p_n = (X_n - \hat{X}_{n|n})^T P_{n|n}^{-1} (X_n - \hat{X}_{n|n})$$

Recordemos que estamos en el caso gaussiano y por lo tanto bajo la hipótesis de consistencia (que se adapte bien el filtro) $p_n \sim \chi_n^2$, por ser el producto de dos variables normales independientes, donde $n = \dim(X_n)$ y $\mathbb{E}[p_n] = n$. Esto es lo que habría que comprobar, pero como hemos dicho anteriormente nunca tenemos el valor de X_n . Para comprobar la consistencia entonces habrá que ver si la media de $\hat{\delta}_n$ es 0 y con covarianza S_n . A veces observando la secuencia de la innovación es suficiente, pero para obtener mayor precisión a la hora de decidir la consistencia del filtro debemos comprobar que la innovación no es sesgada y se trata de ruido blanco.

5.4.1. Test sobre el sesgo de la innovación

Por un lado para contrastar que no sea sesgada, hay que calcular la secuencia

$$q_n^{(i)} = \hat{\delta}_n^T (S_n^{(i)})^{-1} \hat{\delta}_n$$

Para $i = 1, \dots, N$ donde (i) es un superíndice. Notemos que la secuencia $q_n^{(i)}$ se obtiene corriendo una única muestra. Si las hipótesis del filtro son correctas $q_n^{(i)} \sim \chi_m^2$ y por lo tanto debemos ver que $\mathbb{E}[q_n] = m$.

Para aproximar $\mathbb{E}[q_n]$ deberíamos hacer N iteraciones del Filtro de Kalman y calcular

$$\bar{q}_n = \frac{1}{N} \sum_{i=1}^N q_n^{(i)}$$

Por las propiedades de suma de variable de Ji-cuadrado, $N\bar{q}_n \sim \chi_{Nm}^2$ y con la tabla de χ^2 podemos construir una región de confianza donde $\mathbb{E}[q_n^i] = m$.

Pero hay otra manera de aproximar $\bar{q}_n$. Como hemos supuesto que la innovación está incorrelacionada para cada instante, entonces es ergódica, y por el Teorema Ergódico de Birkhoff, podemos aproximar $\bar{q}_n$ por

$$\bar{q} = \frac{1}{N} \sum_{n=1}^N \bar{q}_n$$

y comprobar que $N\bar{q} \sim \chi_{Nm}^2$.

Para el ejemplo que estabamos tratando, vamos a suponer que tenemos $N = 100$ iteraciones. Para comprobar lo dicho anterior debemos ahora construir un intervalo de confianza donde $\bar{q} \in [b_1, b_2]$ con probabilidad $1 - \alpha$

$$\mathbb{P}(\bar{q} \in [b_1, b_2] | H_0) = 1 - \alpha$$

Donde H_0 se refiere a la hipótesis nula de que $N\bar{q} \sim \chi_{Nm}^2$. Fijamos $1 - \alpha = 0,95$ y por las tablas χ_{100}^2

$$[b_1 \quad b_2] = \left[\frac{\chi_{100}^2(0,025)}{100} \quad \frac{\chi_{100}^2(0,975)}{100} \right] = [0,74 \quad 1,3]$$

Este intervalo proporciona una región de aceptación de nivel de significación 0,05. Y calculando el gráfico veríamos si $\bar{q} \in [0,74 \quad 1,3]$.

5.4.2. Test de incorrelación de la innovación

Para ver que la innovación está incorrelacionada, el test consiste en comprobar que

$$\mathbb{E}[\widehat{\delta}_i^T \widehat{\delta}_j] = S_i \epsilon_{ij}$$

Donde ϵ_{ij} es la Delta de Dirac, que se define como

$$\epsilon_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Notemos que la Delta de Dirac se suele escribir como una δ pero no lo hacemos para no confundir con el símbolo de la innovación.

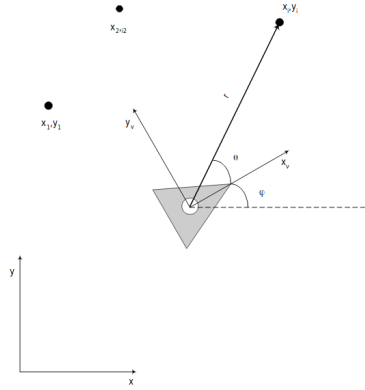
Definimos

$$r(\tau) = \frac{1}{N - \tau} \sum_{i=0}^{N-\tau-1} \widehat{\delta}_i^T \widehat{\delta}_{i+\tau}$$

normalizada de forma que $r(0) = 1$. Para N suficientemente grande, bajo la suposición de incorrelación, si $\tau \neq 0$ entonces $r(\tau) \sim N(0, \frac{1}{\sqrt{N}})$, dado que $\mathbb{E}[\widehat{\delta}_i^T \widehat{\delta}_{i+\tau}] = 0$ y $Var(\widehat{\delta}_i^T \widehat{\delta}_{i+\tau}) = 1$. Por lo tanto se define una región de aceptación con un nivel de significación de 0,05 tal que $\mathbb{P}\left(r(\tau) \in \left[\frac{-2}{\sqrt{N}} \quad \frac{2}{\sqrt{N}}\right]\right) = 0,95$. Hipótesis que aceptaríamos o no según el gráfico de $r(\tau)$.

Capítulo 6

Ejemplo del Filtro de Kalman Extendido



Consideramos el movimiento de un vehículo en el plano. El estado X_n del vehículo en un instante n viene determinado por su localización (x_n, y_n) y su orientación $(\phi, \text{ángulo con la componente } x \text{ de su localización})$. El vector de control U_n viene determinado por la velocidad demandada V_n y la dirección del recorrido ψ_n . El movimiento del vehículo sigue la ecuación no lineal

$$X_n = \begin{bmatrix} x_n \\ y_n \\ \phi_n \end{bmatrix} = \begin{bmatrix} x_{n-1} + TV_n \cos(\phi_{n-1} + \psi_n) \\ y_{n-1} + TV_n \sin(\phi_{n-1} + \psi_n) \\ \phi_{n-1} + T \frac{V_n}{B} \sin(\psi_n) \end{bmatrix} + \begin{bmatrix} q_n^x \\ q_n^y \\ q_n^\phi \end{bmatrix}$$

Donde T es el intervalo de tiempo entre cada instante n , B el punto de partida de las ruedas y $q_n = \begin{bmatrix} q_n^x & q_n^y & q_n^\phi \end{bmatrix}^T$ un vector aleatorio.

Supondremos que el vehículo tiene equipado un sensor que mide la distancia y el rango respecto a diferentes puntos de referencia del plano $B(i) = [s(i) \ t(i)]^T, i = 1, \dots, N$. Entonces las observaciones para cada punto de referencia vienen dadas por el cálculo de

estas dos magnitudes

$$\begin{bmatrix} Z_n^r(i) \\ Z_n^\theta(i) \end{bmatrix} = \begin{bmatrix} \sqrt{(s(i) - x_n)^2 + (t_i - y_n)^2} \\ \arctan(\frac{t(i) - y_n}{s(i) - x_n}) - \phi_n \end{bmatrix} + \begin{bmatrix} r_n^r(i) \\ r_n^\theta(i) \end{bmatrix}$$

Donde $r_n(i) = [r_n^r(i) \ r_n^\theta(i)]^T$ es el ruido de la observación.

Aplicando que la estimación del estado es $\hat{X}_{n|n-1} = F(\hat{X}_{n|n-1}, U_n)$ tenemos que

$$\begin{bmatrix} \hat{x}_n \\ \hat{y}_n \\ \hat{\phi}_n \end{bmatrix} = \begin{bmatrix} \hat{x}_{n-1|n-1} + TV_n \cos(\hat{\phi}_{n-1|n-1} + \psi_n) \\ \hat{y}_{n-1|n-1} + TV_n \sin(\hat{\phi}_{n-1|n-1} + \psi_n) \\ \hat{\phi}_{n-1|n-1} + T \frac{V_n}{B} \sin(\psi_n) \end{bmatrix}$$

Para calcular la covarianza del error en la estimación del estado, sabemos que sigue la ecuación

$$P_{n|n-1} = \nabla F_{X_n} P_{n-1|n-1} \nabla F_{X_n}^T + Q_n$$

Calculamos primero ∇F_{X_n} , es decir

$$\begin{aligned} \nabla F_{X_n} &= \begin{bmatrix} \frac{\delta}{\delta \hat{x}_{n-1|n-1}} \hat{x}_{n-1|n-1} & \frac{\delta}{\delta \hat{y}_{n-1|n-1}} \hat{x}_{n-1|n-1} & \frac{\delta}{\delta \hat{\phi}_{n-1|n-1}} \hat{x}_{n-1|n-1} \\ \frac{\delta}{\delta \hat{x}_{n-1|n-1}} \hat{y}_{n-1|n-1} & \frac{\delta}{\delta \hat{y}_{n-1|n-1}} \hat{y}_{n-1|n-1} & \frac{\delta}{\delta \hat{\phi}_{n-1|n-1}} \hat{y}_{n-1|n-1} \\ \frac{\delta}{\delta \hat{x}_{n-1|n-1}} \hat{\phi}_{n-1|n-1} & \frac{\delta}{\delta \hat{y}_{n-1|n-1}} \hat{\phi}_{n-1|n-1} & \frac{\delta}{\delta \hat{\phi}_{n-1|n-1}} \hat{\phi}_{n-1|n-1} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & -TV_n \sin(\hat{\phi}_{n-1|n-1} + \psi_n) \\ 0 & 1 & TV_n \cos(\hat{\phi}_{n-1|n-1} + \psi_n) \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

Podemos suponer que $P_{n-1|n-1} = \text{diag}[(\sigma^x)^2 \ (\sigma^y)^2 \ (\sigma^\psi)^2]$ donde cada componente son las covarianzas del error de la estimación del estado y $Q_n = \text{diag}[(q^x)^2 \ (q^y)^2 \ (q^\psi)^2]$ dado que Q_n es la matriz de covarizanzas de q_n y sus componentes q^x, q^y, q^ψ son mutuamente independientes. Entonces multiplicando

$$\begin{aligned} P_{n|n-1} &= \begin{bmatrix} (\sigma^x)^2 + T^2 V_n^2 \sin^2(\hat{\phi}_{n-1|n-1} + \psi_n)(\sigma^\phi)^2 + (q^x)^2 \\ -T^2 V_n^2 \sin(\hat{\phi}_{n-1|n-1} + \psi_n) \cos(\hat{\phi}_{n-1|n-1} + \psi_n)(\sigma^\phi)^2 \\ TV_n \sin(\hat{\phi}_{n-1|n-1} + \psi_n)(\sigma^\phi)^2 \\ -T^2 V_n^2 \sin(\hat{\phi}_{n-1|n-1} + \psi_n) \cos(\hat{\phi}_{n-1|n-1} + \psi_n)(\sigma^\phi)^2 \\ (\sigma^y)^2 + T^2 V_n^2 \cos^2(\hat{\phi}_{n-1|n-1} + \psi_n) + (q^y)^2 \\ TV_n \cos(\hat{\phi}_{n-1|n-1} + \psi_n)(\sigma^\phi)^2 \\ -TV_n \sin(\hat{\phi}_{n-1|n-1} + \psi_n)(\sigma^\phi)^2 \\ TV_n \cos(\hat{\phi}_{n-1|n-1} + \psi_n)(\sigma^\phi)^2 \\ (\sigma^\phi)^2 + (q^\phi)^2 \end{bmatrix} \end{aligned}$$

Ahora calculamos la estimación de la observación que sabemos que sigue la fórmula $\hat{Z}_{n|n-1} = H(\hat{X}_{n|n-1})$

$$\begin{bmatrix} \hat{Z}_{n|n-1}^r(i) \\ \hat{Z}_{n|n-1}^\theta(i) \end{bmatrix} = \begin{bmatrix} \sqrt{(s(i) - \hat{x}_{n|n-1})^2 + (t(i) - \hat{y}_{n|n-1})^2} \\ \arctan\left(\frac{t(i) - \hat{y}_{n|n-1}}{s(i) - \hat{x}_{n|n-1}}\right) - \hat{\phi}_{n|n-1} \end{bmatrix}$$

Vamos a calcular la covarianza de la innovación, que se expresa de la forma

$$S_n = \nabla H_{X_n} P_{n|n-1} \nabla H_{X_n}^T + R_n$$

Calculamos primero la jacobiana ∇H_{X_n}

$$\begin{aligned} \nabla H_n^x &= \begin{bmatrix} \frac{\delta}{\delta \hat{x}_{n|n-1}} Z_n^r(i) & \frac{\delta}{\delta \hat{y}_{n|n-1}} Z_n^r(i) & \frac{\delta}{\delta \hat{\phi}_{n|n-1}} Z_n^r(i) \\ \frac{\delta}{\delta \hat{x}_{n|n-1}} Z_n^\theta(i) & \frac{\delta}{\delta \hat{y}_{n|n-1}} Z_n^\theta(i) & \frac{\delta}{\delta \hat{\phi}_{n|n-1}} Z_n^\theta(i) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\hat{x}_{n|n-1} - s(i)}{\sqrt{(s(i) - \hat{x}_{n|n-1})^2 + (t(i) - \hat{y}_{n|n-1})^2}} & \frac{\hat{y}_{n|n-1} - t(i)}{\sqrt{(s(i) - \hat{x}_{n|n-1})^2 + (t(i) - \hat{y}_{n|n-1})^2}} & 0 \\ -\frac{\hat{y}_{n|n-1} - t(i)}{\sqrt{(s(i) - \hat{x}_{n|n-1})^2 + (t(i) - \hat{y}_{n|n-1})^2}} & \frac{\hat{x}_{n|n-1} - s(i)}{\sqrt{(s(i) - \hat{x}_{n|n-1})^2 + (t(i) - \hat{y}_{n|n-1})^2}} & -1 \end{bmatrix} \end{aligned}$$

Análogamente a como hemos hecho para obtener $P_{n|n-1}$, suponemos que $P_{n-1|n-1} = \text{diag}[(\sigma^x)^2 \ (\sigma^y)^2 \ (\sigma^\psi)^2]$ y $R_n = \text{diag}[(r^r)^2 \ (r^\theta)^2]$ dado que R_n es la matriz de covarianzas de r_n y sus componentes r^r, r^θ son independientes.

Definimos $d = \text{sqrt}(s(i) - x_n)^2 + (t(i) - y_n)^2$

$$\begin{aligned} S_n &= \frac{1}{(s(i) - x_n)^2 + (t(i) - y_n)^2} \cdot \\ &\begin{bmatrix} (s(i) - \hat{x}_{n|n-1})^2 (\sigma^x)^2 + (t(i) - \hat{y}_{n|n-1})^2 (\sigma^y)^2 + (r^r)^2 \\ (s(i) - \hat{x}_{n|n-1}) + (t(i) - \hat{y}_{n|n-1})^2 ((\sigma^x)^2 - (\sigma^y)^2) / d \\ (s(i) - \hat{x}_{n|n-1})(t(i) - \hat{y}_{n|n-1})^2 ((\sigma^x)^2 - (\sigma^y)^2) / d \\ ((s(i) - \hat{x}_{n|n-1})^2 (\sigma^x)^2 + (t(i) - \hat{y}_{n|n-1})^2 (\sigma^y)^2) / d^2 + ((r^\theta)^2 + (\sigma^\theta)^2) d^2 \end{bmatrix} \end{aligned}$$

Capítulo 7

Conclusiones

En este trabajo he conseguido entender y desarrollar sobre todo de forma teórica un algoritmo recursivo, que sin conocerlo antes, he descubierto que tiene muchas aplicaciones en otras áreas como la electrónica y la economía (que de hecho se trata como un tema en asignaturas de otros Grados o Másters) y que ha pasado por épocas de mayor y menor uso. Un gran número de avances científicos no hubieran sido posibles sin éste. Podemos afirmar que el Filtro de Kalman fue una de las piezas que permitió la era espacial, ya que la precisión y eficiencia que se obtuvo a la hora de resolver el problema de la navegación de las naves espaciales a través del sistema solar no habría sido posibles sin éste.

Aunque parezca a simple vista un algoritmo sencillo de implementar, la comprensión de su origen y objetivo requiere de un buen conocimiento en la teoría de la medida y procesos estocásticos, así como el dominio de diversas herramientas de la teoría de Probabilidad y Estadística en general. Dado que durante el Grado no se transmiten demasiados conceptos sobre estimación, al principio me ha resultado difícil asimilar y entender el Filtro de Kalman, pero a la vez me encontraba dispuesto y motivado para llenar el vacío de conocimiento que tenía sobre estas áreas de las Matemáticas. También me ha resultado difícil de entender el Filtro de la Partícula, dado que a pesar de haber muchas referencias con información sobre éste, no todas eran muy claras ni lo desarrollaban con profundidad y de forma detallada.

El Trabajo Final de Grado me ha servido para obtener una gran agilidad y familiarización con el lenguaje Latex que no tenía hasta el momento y que me permitirá digitalizar cualquier expresión matemática, realizar otros trabajos con mayor calidad... También he mejorado a la hora de formalizar definiciones y procedimientos matemáticos esclareciendo todas las hipótesis así como a usar la notación correspondiente y adecuada. El trabajo me ha ayudado a aplicar y seguir ampliando lo aprendido en otras asignaturas y me puede servir en el futuro para tratar de realizar alguna implementación en lenguaje C del Filtro de Kalman y usarlo para algún caso concreto.

Capítulo 8

Anexo

8.1. Esperanza condicionada

8.1.1. Definiciones previas

Veamos algunas definiciones elementales antes de entrar en la temática de la esperanza condicionada.

Definición 8.1.1. *Dada una colección Σ_0 de subconjuntos de un conjunto S se llama σ -álgebra en S si:*

1. $S \in \Sigma_0$
2. $F \in \Sigma_0 \implies F^c := S - F \in \Sigma_0$
3. $F, G \in \Sigma_0 \implies F \cup G \in \Sigma_0$
4. $F_n \in \Sigma \ (n \in \mathbb{N}) \implies \bigcup_n F_n \in \Sigma$

Señalamos que $\emptyset = S^c \in \Sigma_0$ y $F, G \in \Sigma_0 \implies F \cap G = (F^c \cup G^c)^c \in \Sigma_0$ y que si Σ es un σ -álgebra donde S y $F_n \in \Sigma \ (n \in \mathbb{N})$, entonces

$$\bigcap_n F_n = \left(\bigcup_n F_n^c \right)^c \in \Sigma$$

Definición 8.1.2. *Un espacio de probabilidad es una terna $(\Omega, \mathcal{F}, \mathbb{P})$ donde:*

1. Ω es un conjunto formado por elementos ω que son los posibles fenómenos aleatorios.
2. $\mathcal{F}$ es una familia de subconjuntos de Ω con estructura de σ -álgebra. Sus elementos se llaman eventos.

3. $\mathbb{P}$ es una aplicación

$$\mathbb{P} : A \in \mathcal{F} \longrightarrow \mathbb{P}(A) \in [0, 1]$$

tal que $\mathbb{P}(\Omega) = 1$ y para toda sucesión numerable $\{A_i, i \geq 1\}$ de conjuntos disjuntos dos a dos de $\mathcal{F}$ se tiene

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

Definición 8.1.3. Sea $\mathcal{F}$ un conjunto y $\mathcal{G}, \mathcal{H}$ σ -álgebras de $\mathcal{F}$ entonces $\mathcal{G}$ es una **sub- σ -álgebra** de $\mathcal{H}$ si $\mathcal{G} \subseteq \mathcal{H}$.

Definición 8.1.4. Un **espacio medible** es un par (S, Σ) , donde S es un conjunto y Σ es una σ -álgebra. Un elemento de Σ subconjunto de S se llama **Σ -medible**.

Definición 8.1.5. Sea (S, Σ) un espacio medible tal que Σ sea una σ -álgebra en S . Suponemos que $h : S \longrightarrow \mathbb{R}$ y para $A \subseteq \mathbb{R}$, se define $h^{-1}(A) := \{s \in S : h(s) \in A\}$. Luego h se llama **Σ -medible** si $h^{-1} : \mathcal{B} \longrightarrow \Sigma$, es decir, si $h^{-1}(A) \in \Sigma, \forall A \in \mathcal{B}$. Esta definición de medibilidad sirve también para variables aleatorias.

8.1.2. La esperanza condicionada, definiciones y propiedades

Definición 8.1.6. Para una variable aleatoria $X \in \mathcal{L}^1 = \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, definimos la **esperanza de X** , $\mathbb{E}[X]$, como

$$\mathbb{E}[X] := \int_{\Omega} X d\mathbb{P} = \int_{\Omega} X(w) \mathbb{P}(dw)$$

Para $X \in \mathcal{L}^1$ y $F \in \mathcal{F}$, se define $\mathbb{E}[X] := \int_F X(w) \mathbb{P}(dw) := \mathbb{E}[X \mathbb{I}_F]$

Teorema 8.1.7. Sea una terna $(\Omega, \mathcal{F}, \mathbb{P})$ y X una variable aleatoria con $\mathbb{E}[|X|] < \infty$. Sea $\mathcal{G}$ una sub- σ -álgebra de $\mathcal{F}$. Entonces existe una variable aleatoria Y tal que:

1. Y es $\mathcal{G}$ -medible.
2. $\mathbb{E}[|Y|] < \infty$.
3. Para cada conjunto G de $\mathcal{G}$ tenemos que $\int_G Y d\mathbb{P} = \int_G X d\mathbb{P}, \forall G \in \mathcal{G}$.

Además si $\tilde{Y}$ es una variable aleatoria con estas propiedades entonces $\tilde{Y} = Y$ y $\mathbb{P}[\tilde{Y} = Y] = 1$. Una variable aleatoria Y con estas propiedades se conoce como una **versión de la esperanza condicionada** $\mathbb{E}[X|\mathcal{G}]$ de X dado $\mathcal{G}$ y escribimos $Y = \mathbb{E}[X|\mathcal{G}]$.

Normalmente escribiremos $\mathbb{E}[X|Z]$ para $\mathbb{E}[X|\sigma(Z)]$ y $\mathbb{E}[X|Z_1, Z_2, \dots, Z_n]$ para

$$\mathbb{E}[X|\sigma(Z_1), \sigma(Z_2), \dots, \sigma(Z_n)].$$

8.1.3. Cálculo de la esperanza condicionada

Definición 8.1.8. Suponemos X, Z variables aleatorias con función de densidad conjunta $f_{X,Z}(x, z)$. Entonces $f_Z(z) = \int_{\mathbb{R}} f_{X,Z}(x, z) dx$ actúa como función de densidad de Z . Definimos la **función de densidad de X condicionada por Z** como $f_{X|Z}(x|z) = \frac{f_{X,Z}(x,z)}{f_Z(z)}$ si $f_Z(z) \neq 0$ o $f_{X|Z}(x|z) = 0$ si $f_Z(z) = 0$.

Definición 8.1.9. Sea h una función boreliana de $\mathbb{R}$ tal que $\mathbb{E}[|h(X)|] = \int_{\mathbb{R}} |h(x)| f_X(x) dx$. Definimos $g(z) := \int_{\mathbb{R}} h(x) f_{X|Z}(x|z) dx$. Entonces $Y := g(Z)$ es una **versión de la esperanza condicionada** de $h(X)$ dada $\sigma(Z)$.

Propiedades de la esperanza condicionada

Suponemos que todas las X 's satisfacen $\mathbb{E}[|X|] < \infty$ y que $\mathcal{G}, \mathcal{H}$ son sub- σ -álgebras de $\mathcal{F}$.

1. Si Y es una versión de $\mathbb{E}[X|\mathcal{G}] \implies \mathbb{E}[X] = \mathbb{E}[Y]$
2. Si X es $\mathcal{G}$ -medible $\implies \mathbb{E}[X|\mathcal{G}] = X$ c.s.
3. $\mathbb{E}[a_1 X_1 + a_2 X_2 | \mathcal{G}] = a_1 \mathbb{E}[X_1 | \mathcal{G}] + a_2 \mathbb{E}[X_2 | \mathcal{G}]$ c.s.
4. Si $X \geq 0 \implies \mathbb{E}[X|\mathcal{G}] \geq 0$ c.s.
5. Si $X_n \geq 0 \implies \mathbb{E}[\liminf X_n | \mathcal{G}] \leq \liminf (\mathbb{E}[X_n | \mathcal{G}])$ c.s.
6. Si $|X_n(w)| \leq V(w), \forall n, w$ y $\mathbb{E}[V] < \infty$ y $X_n \rightarrow X$ c.s. $\implies \mathbb{E}[X_n | \mathcal{G}] \rightarrow \mathbb{E}[X | \mathcal{G}]$ c.s.
7. Si $c : \mathbb{R} \rightarrow \mathbb{R}$ es convexa y $\mathbb{E}[c(X)] < \infty \implies \mathbb{E}[c(X) | \mathcal{G}] \geq c\mathbb{E}[X | \mathcal{G}]$ c.s.
8. Si $\mathcal{H}$ es una sub- σ -álgebra de $\mathcal{G} \implies \mathbb{E}[\mathbb{E}[X | \mathcal{G}] | \mathcal{H}] = \mathbb{E}[X | \mathcal{H}]$ c.s.
9. Si Z es $\mathcal{G}$ -medible i acotada $\implies \mathbb{E}[ZX | \mathcal{G}] = Z\mathbb{E}[X | \mathcal{G}]$ c.s.
10. Si $\mathcal{H}$ es independiente de $\sigma(\sigma(X), \mathcal{G}) \implies \mathbb{E}[X | \sigma(\mathcal{G}, \mathcal{H})] = \mathbb{E}[X | \mathcal{G}]$. En particular si X es independiente de $\mathcal{H} \implies \mathbb{E}[X | \mathcal{H}] = \mathbb{E}[X]$ c.s.
(c.s. se refiere a las siglas de casi seguramente)

Demostración. :

1. Si Y es una versión de $\mathbb{E}[X|\mathcal{G}]$, por definición $\implies \int_G Y d\mathbb{P} = \int_G X d\mathbb{P}, \forall G \in \mathcal{G} \implies \mathbb{E}[Y] = \mathbb{E}[X]$.
2. Si X es $\mathcal{G}$ -medible, entonces X cumple todas las propiedades para ser una versión de la esperanza condicionada $\mathbb{E}[X|\mathcal{G}] \implies \mathbb{E}[X|\mathcal{G}] = X$.

3. Suponemos que Y_1 es una versión de $\mathbb{E}[X_1|\mathcal{G}]$ y Y_2 una versión de $\mathbb{E}[X_2|\mathcal{G}]$ entonces como $\int_G Y_1 d\mathbb{P} = \int_G X_1 d\mathbb{P}$ y $\int_G Y_2 d\mathbb{P} = \int_G X_2 d\mathbb{P}$, $\forall G \in \mathcal{G}$ se sigue que

$$\int_G (a_1 Y_1 + a_2 Y_2) d\mathbb{P} = \int_G (a_1 X_1 + a_2 X_2) d\mathbb{P}$$

Por otro lado tenemos que $a_1 Y_1 + a_2 Y_2$ es $\mathcal{G}$ -medible y por lo tanto $a_1 Y_1 + a_2 Y_2$ es versión de $\mathbb{E}[a_1 X_1 + a_2 X_2|\mathcal{G}]$ luego $\mathbb{E}[a_1 X_1 + a_2 X_2|\mathcal{G}] = a_1 Y_1 + a_2 Y_2 = a_1 \mathbb{E}[X_1|\mathcal{G}] + a_2 \mathbb{E}[X_2|\mathcal{G}]$.

4. Sea Y una versión de $\mathbb{E}[X|\mathcal{G}]$. Suponemos que $\mathbb{P}(Y < 0) > 0$ (reducción al absurdo), entonces para algún n el conjunto $G := \{Y < -n^{-1}\}$ en $\mathcal{G}$ tiene probabilidad positiva. Entonces

$$0 \leq \int_G X d\mathbb{P} = \int_G Y d\mathbb{P} = -n^{-1} \mathbb{P}(G) < 0$$

La cual cosa contradice lo dicho anteriormente y por lo tanto la hipótesis inicial.

5. Se demuestra gracias al lemma de Fatou.
 6. Se demuestra gracias al Teorema de Convergencia dominada.
 7. Sabemos que para una función $f : \mathbb{R} \rightarrow \mathbb{R}$, $\exists m \in \mathbb{R}$, $\forall (x_0, f(x_0))$ tal que

$$f(x) \geq m(x - x_0) + f(x_0)$$

Entonces si $\mathbb{E}[x] = x_0$, tomando esperanzas

$$\mathbb{E}[f(x)] \geq \mathbb{E}[m(x - x_0)] + f(\mathbb{E}[x_0])$$

y como $\mathbb{E}[m(x - x_0)] = m(\mathbb{E}[x] - x_0) = 0$ entonces

$$\mathbb{E}[f(x)] \geq f(\mathbb{E}[x_0])$$

8. Si Y es una versión de $\mathbb{E}[X|\mathcal{G}]$ y Z es una versión de $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}]$ se cumple

$$\int_H Z d\mathbb{P} = \int_H \mathbb{E}[X|\mathcal{G}] d\mathbb{P} = \int_G \mathbb{E}[X|\mathcal{G}] d\mathbb{P} = \int_G Y d\mathbb{P}$$

Ya que $\forall H \in \mathcal{H} \implies H \in G \in \mathcal{G}$. Y como Y es $\mathcal{G}$ -medible $\implies Z$ es $\mathcal{H}$ -medible, por tanto $\mathbb{E}[X|\mathcal{G}] = Y = Z = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}]$.

9. Por linealidad podemos suponer que $X \geq 0$. Fijamos Y como versión de $\mathbb{E}[X|\mathcal{G}]$ y $G \in \mathcal{G}$. Si Z es $\mathcal{G}$ -medible, el indicador de G y apropiadamente integrable entonces $\int_G ZX d\mathbb{P} = \int_G ZY d\mathbb{P}$ dado que $\int_G X d\mathbb{P} = \int_G Y d\mathbb{P}$. Además $\mathbb{E}[|ZX|] < \infty$ dado que Z es acotada y $X \in \mathcal{L}^1$. Entonces $\int_G ZX d\mathbb{P} = Z \int_G X d\mathbb{P}$ y se cumple $\mathbb{E}[ZX|\mathcal{G}] = ZE[X|\mathcal{G}]$
 10. Suponemos que $X \geq 0$ y $\mathbb{E}[X] < \infty$. Sea $G \in \mathcal{G}$ y $H \in \mathcal{H}$, $X\mathbb{I}_G$ y $\mathbb{I}_H$ son independientes, por lo tanto

$$\mathbb{E}[X; G \cap H] = \mathbb{E}[(X\mathbb{I}_G)(\mathbb{I}_H)] = \mathbb{E}[X\mathbb{I}_G]\mathbb{E}[\mathbb{I}_H] = \mathbb{E}[X\mathbb{I}_G]\mathbb{P}(H)$$

Ahora si $Y = \mathbb{E}[X|\mathcal{G}]$ es una versión de la esperanza condicionada, como Y es $\mathcal{G}$ -medible y $Y\mathbb{I}_G$ es independiente de $\mathbb{I}_H$ entonces $\mathbb{E}[(X\mathbb{I}_G)(\mathbb{I}_H)] = \mathbb{E}[Y\mathbb{I}_G]\mathbb{P}(H)$ y también tenemos que $\mathbb{E}[X; G \cap H] = \mathbb{E}[Y; G \cap H]$

Así que las medidas $F \mapsto \mathbb{E}[X; F]$, $F \mapsto \mathbb{E}[Y; F]$ en $\sigma(\mathcal{G}, \mathcal{H})$ de la misma masa total finita concuerdan con el π -sistema de conjuntos de la forma $G \cap H$ y en todo $\sigma(\mathcal{G}, \mathcal{H})$, como queríamos demostrar.

□

Probabilidad condicionada

Para $F \in \mathcal{F}$, tenemos que $\mathbb{P}(F) = \mathbb{E}[\mathbb{I}_F]$.

Definición 8.1.10. Para $F \in \mathcal{F}$ y $\mathcal{G}$ una sub- σ -álgebra de $\mathcal{F}$ definimos $\mathcal{P}(F|\mathcal{G})$ como una versión de $E[\mathbb{I}_F|\mathcal{G}]$.

Esta probabilidad cumple que $\mathbb{P}(\bigcup F_n|\mathcal{G}) = \sum \mathbb{P}(F_n|\mathcal{G})$ c.s.

Definición 8.1.11. Una función $\mathbb{P}$ se llama **probabilidad regular condicionada** dado $\mathcal{G}$ si esta función

$$\mathbb{P}(,) : \Omega \times \mathcal{F} \longrightarrow [0, 1)$$

tal que

1. Para $F \in \mathcal{F}$, la función $w \mapsto \mathbb{P}(w, F)$ es una versión de P
2. Para cada casi todo w , la función

$$F \mapsto \mathbb{P}(w, F)$$

es una probabilidad de medida en $\mathcal{F}$

La existencia de esta función no se puede garantizar dado que hay muchas sucesiones de conjuntos disjuntos no numerables.

Capítulo 9

Bibliografia

Hirsa, A. (2016). Computational methods in finance. CRC Press.

Durrant-Whyte, H. (2001). Introduction to Estimation and the Kalman Filter.

Brockwell, P. J., Davis, R. A., & Fienberg, S. E. (1991). Time Series: Theory and Methods: Theory and Methods. Springer Science Business Media.

Williams, D. (1991). Probability with martingales. Cambridge university press.

Mohinder S.. Grewal, & Angus P.. Andrews. (2001). Kalman filtering: theory and practice using matlab. Wiley.

Dubuisson, S. (2015). Tracking with particle filter for high-dimensional observation and state spaces. John Wiley Sons.

Javaheri, A., Lautier, D., & Galli, A. (2003). Filtering in finance. Wilmott, 3, 67-83.