

Tarea 1 - Filtros en series de tiempo

MDS7204 - Aprendizaje de Máquinas Avanzado

Camilo Carvajal Reyes

26 de septiembre de 2022

1. Introducción

El presente informe aborda las implementaciones de dos métodos de predicción en series de tiempo: el filtro de Kalman y el filtro de partículas. Ambos integran un conjunto de métodos predictivos secuenciales cuya naturaleza general abordaremos a continuación para el caso unidimensional.

Supongamos que queremos encontrar una sucesión (indexada en los naturales) X definida en los reales. Cada elemento X_n de la sucesión describe algún fenómeno en el instante de tiempo n . Diremos que su comportamiento dinámico estará dado por $X_n = f_n(X_{n-1}, V_{n-1})$. Acá cada f_n es una función $\mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ que no necesariamente será lineal, y $(V_n)_{n \in \mathbb{N}}$ son realizaciones independientes de una v.a. que modela ruido.

Normalmente no tendremos acceso a los valores reales de X , pero si podremos hacer observaciones, lo cual nos genera una sucesión Y . La naturaleza de tal observaciones estarán dadas por la expresión $Y_n = h_n(X_n, W_n)$. Nuevamente los h_n serán funciones $\mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ (posiblemente) no lineales y W será una sucesión aleatoria i.i.d. que modela el ruido.

1. **Predicción:** Primero computamos la probabilidad del estado siguiente dadas las observaciones hasta el tiempo anterior usando probabilidades totales.

$$p(X_n | Y_{1:n-1}) = \int p(X_n | X_{n-1}) p(X_{n-1} | Y_{1:n-1}) dX_{n-1}$$

2. **Actualización:** En este paso incorporamos la información de la medición en el tiempo actual, con lo cual se puede usar el teorema de Bayes.

$$p(X_n | Y_{1:n}) = \frac{p(Y_n | X_n) p(X_n | Y_{1:n-1})}{p(Y_n | Y_{1:n-1})}$$

2. Filtro de Kalman

Primero consideremos f como $f(X_n) = AX_{n-1} + BV_n$ y h como $h_n(X_n) = CX_n + DW_n$, lo cual corresponde a un modelo lineal, pues A, B, C, D son constantes. Además tanto V como W son sucesiones de ruido Gaussiano $\mathcal{N}(0, 1)$. Consideramos estos datos como input para nuestro modelo, que evolucionará en dos etapas:

1. **Predicción:** primeramente se computan las predicciones a priori tanto de X_n como de la varianza, que denotamos P_n .

$$\hat{X}_{n|n-1} = A\hat{X}_{n-1}, \quad P_{n|n-1} = A^2 P_{n-1|n-1} + Q$$

donde Q es la varianza del ruido en la ecuación de estado. Notemos que en este caso, esto es igual a B , pues hemos definido ruidos normales estándar.

2. **Actualización:** una vez que la medición se vuelve disponible, podemos incorporarla en nuestra estimación. Esto se hace usando un valor que llamaremos la ganancia de Kalman. Este parámetro nos dice cuanto priorizar la medición versus nuestro estimado usando la confianza en las mediciones, y estará dada por:

$$K_n = \frac{P_{n|n-1}}{(P_{n|n-1} + \sigma^2)}, \quad \hat{X}_{n|n} = \hat{X}_{n|n-1} + K_n(Y_n - A\hat{X}_{n|n-1}), \quad P_{n|n} = (1 - K_n)P_{n|n-1}$$

Consideremos un modelo lineal. Sampleamos los valores de las constantes incluyendo el estado inicial X_0 obteniendo los siguientes valores y series como muestra la figura 1:

parámetro	A	B	C	D	Posición inicial
valor	0.5488	1.0860	0.6027	0.5179	1.4940

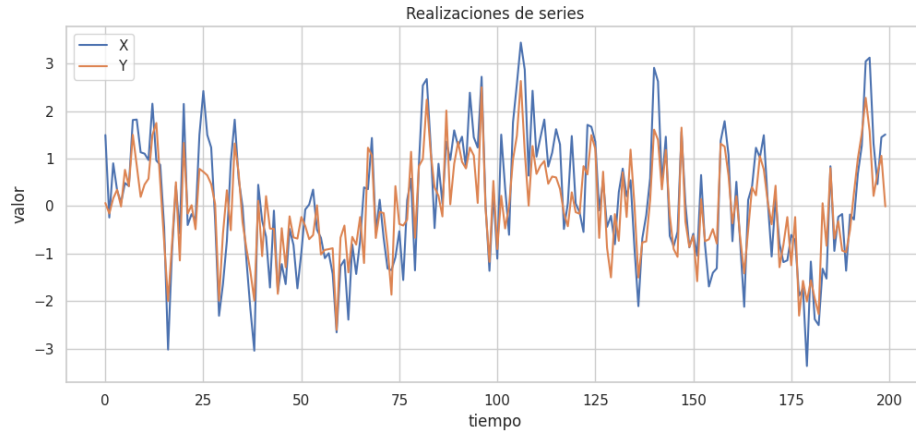


Figura 1: Serie original X y mediciones Y .

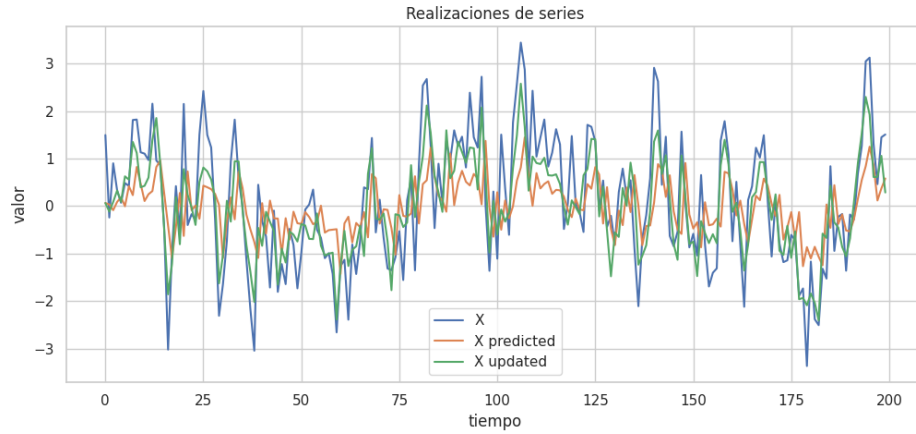


Figura 2: Serie X junto con las predicciones del filtro pre y post actualización

La implementación del filtro y su uso sobre la serie Y se grafican en la figura 2. En ella, la serie original X a aproximar se muestra con líneas azules, mientras que los valores que resultan de la aplicación del filtro de Kalman están en verde. Se vislumbra que hay una fuerte correlación entre ambas series. La aproximación se vuelve menos precisa en aquellos puntos de tiempo en los cuales hay saltos repentinos en X .

Por otro lado, incluimos también la serie $X_{predicted}$ (en **naranja**), la cual representa los valores de las predicciones a priori $X_{n:n-1}$ para $n \in \{0, \dots, 200\}$. Esto nos permite ver que tan alejada era la aproximación antes de la observación correspondiente. Es claro que la serie luego de la actualización de Kalman (**verde**) aproxima de mejor modo X (**azul**), con lo cual podemos confirmar que ponderar usando la ganancia de Kalman efectivamente acerca la predicción final a la serie original. Para poder cuantificar la calidad de las predicciones usamos el error cuadrático medio (MSE), con lo cual obtenemos 0.457. A modo de referencia, el error de la serie pre-actualización es de 1.329.

Finalmente, en la figura 3 incluimos un mapa de calor que nos ayuda a entender el efecto de las varianzas del problema. Mientras más oscura la celda es mejor pues la predicción está más cerca de la serie original. Esto nos dice que es preferible tener baja varianza en las mediciones, lo cual es intuitivo. Matemáticamente, este valor influye en la confianza que podemos tener de las mediciones (mientras más bajo V , más alta la ganancia de Kalman). Por otro lado, una baja varianza en el proceso nos empeora los resultados. Esto se explica puesto a que tener una peor varianza del medidor que del proceso mismo resultará en mediciones consistentemente malas, con las cuales ningún filtro logrará predicciones adecuadas.

3. Filtro de Partículas

Para esta pregunta nos centraremos en un modelo de volatilidad estocástica. Simulamos 500 pasos de un tal proceso X , que se muestran en la figura 4 junto con las mediciones Y . Notamos en él la naturaleza multiplicativa del ruido de las mediciones.

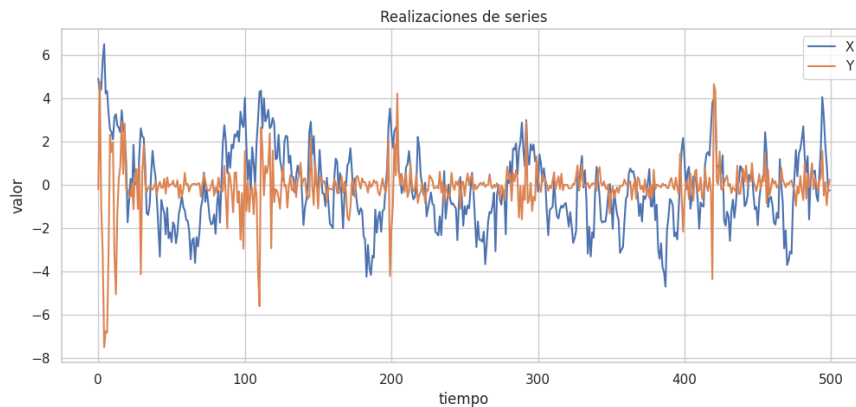


Figura 4: Serie original X y mediciones Y para el modelo Volatilidad Estocástica.

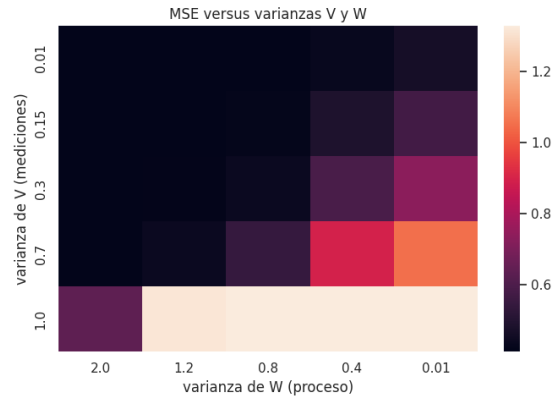


Figura 3: MSE versus varianzas del modelo

Para llevar a cabo el sampleo de nuestras partículas escogemos la importance distribution siguiente:

$$q(x_n|x_{n-1}^i, y_n) := p(x_n|x_{n-1}^i) \sim \mathcal{N}(\alpha x_{n-1}^i, \sigma^2)$$

Esta distribución cumple la naturaleza Markoviana que buscamos, además de ser una adivinanza razonable de un paso siguiente dado el modelo. Más aún, tal distribución nos facilita el cálculo de los pesos, pues luego sólo basta multiplicar el peso anterior correspondiente por la probabilidad de la nueva observación dada una nueva partícula (previa normalización). Luego de la implementación de un filtro de partículas basado en Sequential Importance Sampling obtenemos el gráfico de la figura 5.

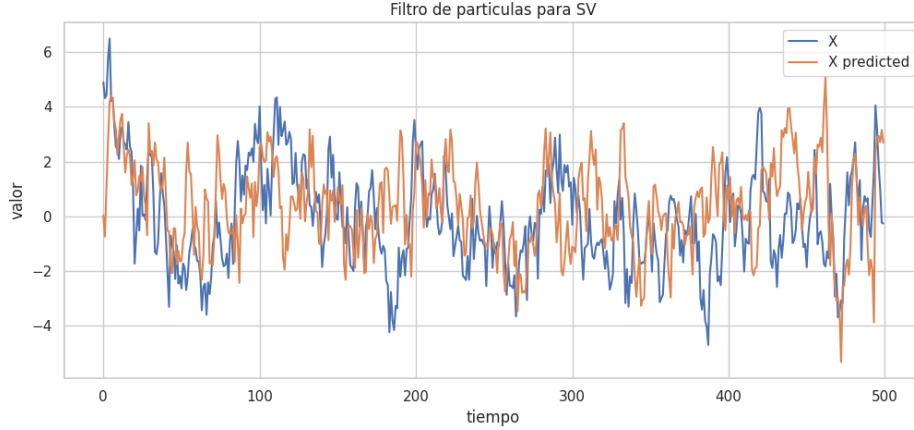


Figura 5: Serie X junto con las predicciones del filtro sin resampling

El filtro es poco efectivo para aproximar X . Esto se debe a que las trayectorias de las partículas se empiezan a alejar de los valores reales, lo cual causa una inestabilidad que se puede observar desde las primeras 100 iteraciones. Este resultado es esperable para un filtro basado únicamente en SIS, y su MSE es de 4.539 para 1000 partículas (no se observaron cambios significativos con un N más grande). Por otro lado, las aproximaciones mejoran sustancialmente cuando aplicamos resampling, como se vislumbra en la figura 6.

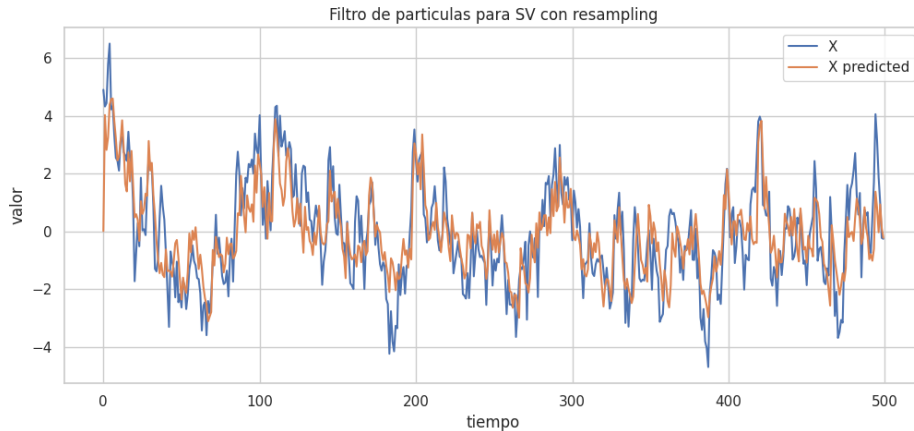


Figura 6: Serie X junto con las predicciones del filtro con resampling

Este filtro sí muestra un comportamiento estable, ya que muchas de las trayectorias que tienen baja probabilidad de explicar el proceso (dadas las observaciones) serán descartadas en el proceso de resampling. En cada paso obtendremos entonces nuevas partículas que se ajustan de mejor manera a las observaciones.