**Dynamic Factor Models (DFM)**

Debemos tener en cuenta que la especificación del modelo responde a una especificación en el espacio de los estados. Dicha especificación del espacio de estados es una herramienta cada vez más habitual en el análisis de series temporales y es fundamental para los Modelos Factoriales Dinámicos (DFM).

La representación en espacio de estados consiste en dos ecuaciones principales:

1. En primer lugar una ecuación de medida (o de observación):

2. En segundo lugar, una ecuación de transición (o de estado):

Donde:

- es el vector de observaciones

- es el vector de estados (no observables)

- es la matriz de medida

- es la matriz de transición

- y son términos de error

La estimación de estos modelos, cuando hay frecuencia mixta (mensuales y trimestrales) se lleva a cabo mediante el algoritmo EM, y que usa el filtro de Kalman.

El algoritmo EM es una técnica iterativa utilizada para encontrar estimaciones de máxima verosimilitud de parámetros en modelos probabilísticos que involucran variables latentes o datos faltantes y que es el caso que presentamos aquí dado que las variables o indicadores trimestrales solo mostrarían un valor por cada uno de dichos trimestres.

Este método alterna entre dos pasos: el paso de Expectation (E) y el paso de Maximization (M). En el paso E, se calcula la expectativa de la función de log-verosimilitud, usando las estimaciones actuales de los parámetros y condicionada a los datos observados. En el paso M, se calculan nuevos parámetros maximizando la expectativa de la log-verosimilitud encontrada en el paso E. Estos dos pasos se repiten iterativamente hasta que se alcanza la convergencia, es decir, hasta que los cambios en los parámetros o en la log-verosimilitud sean menores que un umbral predefinido. El algoritmo EM es particularmente útil en situaciones donde la maximización directa de la función de verosimilitud es difícil, pero la función de verosimilitud puede ser simplificada asumiendo la existencia de valores para variables no observadas adicionales. Una de las ventajas clave del EM es su garantía de aumentar la verosimilitud en cada iteración, lo que asegura la convergencia a un máximo local. Sin embargo, no garantiza encontrar el máximo global, por lo que a menudo se utilizan múltiples inicios aleatorios para explorar diferentes regiones del espacio de parámetros. El algoritmo EM encuentra aplicaciones en una amplia gama de campos, incluyendo aprendizaje automático, visión por computadora, reconocimiento de patrones y bioinformática, siendo especialmente valioso en problemas como la estimación de mezclas de gaussianas, imputación de datos faltantes, y en nuestro caso, la estimación de parámetros en modelos factoriales dinámicos.

El uso del filtro de Kalman en la estimación de DFMs es útil por varias razones:

1. Estimación de estados latentes:

El filtro de Kalman permite estimar los valores de los factores no observables ( en nuestro modelo) en cada punto del tiempo, basándose en la información disponible hasta ese momento.

2. Manejo de datos faltantes:

Es especialmente útil cuando hay datos faltantes o cuando se trabaja con frecuencias mixtas (como en nuestro caso, que tiene datos mensuales y trimestrales).

3. Cálculo de la verosimilitud:

El filtro de Kalman proporciona una manera eficiente de calcular la función de verosimilitud del modelo, que es necesaria para la estimación de parámetros.

4. Predicción:

Permite realizar predicciones de las variables observables y de los factores latentes.

En el contexto del DFM que se usa en nuestra modelización de indicadores, el filtro de Kalman se aplica de la siguiente manera:

1. Inicialización:

Se inicia con una estimación inicial de los factores y su matriz de covarianza.

2. Predicción:

- Predice los factores para el siguiente período

- Calcula la matriz de covarianza de esta predicción

3. Actualización:

- Cuando se observan nuevos datos (), se actualiza la estimación de los factores

- Se calcula el error de predicción:

- Se actualiza la estimación de los factores y su matriz de covarianza

4. Suavizado:

Después de procesar todos los datos, se realiza un paso de suavizado para mejorar las estimaciones de los factores utilizando toda la información disponible.

El filtro de Kalman se integra en el algoritmo EM de la siguiente manera:

- Paso E: Se usa el filtro de Kalman para estimar los factores latentes y calcular las estadísticas suficientes necesarias.

- Paso M: Se actualizan los parámetros del modelo (, , etc.) basándose en estas estadísticas.

**Aplicación al DFMQ**

El Modelo Factorial Dinámico Trimestral (DFMQ) es, como se ha adelantado, una técnica estadística utilizada para analizar y predecir series temporales multivariadas con diferentes frecuencias temporales. Este modelo es particularmente útil en economía y finanzas, donde múltiples variables pueden estar influenciadas por factores comunes no observables.

El modelo básico, aplicado en este trabajo, se identifica con el explicado en la sección anterior, y se define de la siguiente manera:

donde:

• es un vector (n x 1) de datos observados en el tiempo t. Cada elemento de representa una serie temporal observada.

• es un vector (n x 1) de perturbaciones idiosincráticas en el tiempo t. Estas capturan la variabilidad específica de cada serie que no es explicada por los factores comunes.

• es un vector (r x 1) de factores no observados (latentes) en el tiempo t. Estos factores capturan las dinámicas comunes subyacentes en las series observadas.

• es un vector (r x 1) de perturbaciones del factor en el tiempo t, donde Q es la matriz de covarianza de estas perturbaciones.

• (Lambda) es una matriz (n x r) de cargas factoriales. Cada elemento representa la sensibilidad de la i-ésima variable observada al j-ésimo factor.

• son matrices (r x r) de coeficientes de autorregresión para i = 1, ..., p, donde p es el orden de la autorregresión. Estas matrices capturan la dinámica temporal de los factores.

Además, permitimos que las perturbaciones idiosincráticas estén serialmente correlacionadas, de modo que, donde . Aquí, es el coeficiente de autorregresión para la i-ésima perturbación idiosincrática, y es su varianza.

Así, según la salida que acompaña a los resultados del modelo obtenemos la información siguiente:

Texto

Descripción generada automáticamente

¿Cómo interpretar esta información?

El resultado muestra la estimación de un Modelo Factorial Dinámico (DFMQ) aplicado a una serie de indicadores económicos para la economía andaluza. El modelo se ha estimado utilizando datos desde el 31 de marzo de 2000 hasta el 30 de septiembre de 2023, con un total de 283 observaciones.

El modelo específico es un DFM con 2 factores en 2 bloques, lo que sugiere que se están capturando dos dinámicas latentes principales en los datos. Además, se ha utilizado una frecuencia mixta (M/Q), lo que indica que algunas variables son mensuales y otras trimestrales, y se ha aplicado un proceso AR(1) para modelar las perturbaciones idiosincráticas (.

La tabla de "Observation equation" muestra cómo cada variable observable se relaciona con los factores latentes:

1. Factor loadings (Λ): Las columnas 0 y 1 representan las cargas factoriales para cada uno de los dos factores. Por ejemplo, "Exportaciones de bienes" tiene cargas de -0.17 y -0.13 para los factores 0 y 1 respectivamente. Esto indica cómo cada variable se relaciona con los factores latentes no observables.

2. Idiosyncratic AR(1): Esta columna muestra los coeficientes de autorregresión () para cada variable. Por ejemplo, "Exportaciones de bienes" tiene un coeficiente de 0.60, lo que indica una persistencia moderada en su componente idiosincrático.

3. Var.: Muestra la varianza () del componente idiosincrático para cada variable.

Las secciones "Transition: Factor block 0" y "Transition: Factor block 1" muestran las matrices de transición para cada factor, que capturan la dinámica temporal de los factores latentes.

El modelo ha sido estimado utilizando el algoritmo EM, que convergió después de 68 iteraciones. Los criterios de información (AIC, BIC, HQIC) y el Log Likelihood proporcionan medidas de la bondad del ajuste del modelo.

Aplicando esto a nuestro DFMQ específico:

1. Ecuación de medida:

En nuestro caso, son las 9 variables económicas observadas (exportaciones, IPIAN, ICMA, etc.). La matriz corresponde a (las cargas factoriales que vemos en la tabla). son los dos factores latentes.

Por ejemplo, para "Exportaciones de bienes":

Donde sigue un proceso AR(1):

2. Ecuación de transición:

Esta describe cómo evolucionan los factores latentes. En nuestro modelo, vemos dos bloques de transición, uno para cada factor.

Para el Factor 0:

Para el Factor 1:

Intuición e interpretación:

1. Factores latentes: Representan las fuerzas subyacentes no observables que impulsan las variables económicas. En este caso, podríamos interpretar el Factor 0 como un indicador general de la actividad económica, dado que afecta negativamente a la mayoría de las variables.

2. Cargas factoriales: Muestran cómo cada variable observable responde a los factores latentes. Por ejemplo, el IPIAN y el ICMA son más sensibles al Factor 0 que las exportaciones.

3. Dinámica de los factores: Los coeficientes en las ecuaciones de transición (1.34, -0.76, etc.) indican cómo cada factor evoluciona en el tiempo. Los valores mayores que 1 sugieren que los factores tienen cierta persistencia y posiblemente comportamientos cíclicos.

4. Componentes idiosincráticos: Los coeficientes AR(1) en la columna "idiosyncratic" muestran cuánto de la variación en cada variable no se explica por los factores comunes y persiste en el tiempo.

**Inclusión del PIB en el sistema DFMQ**

En general, incluir el PIB en el sistema multivariante de indicadores cuando se quiere predecir esta misma variable presenta tanto ventajas como desventajas:

Ventajas de incluir el PIB:

* La inclusión del PIB en el modelo factorial dinámico ofrece varias ventajas significativas para el análisis económico y la predicción. En primer lugar, esta inclusión permite al modelo capturar de manera integral las complejas interrelaciones entre el PIB y los demás indicadores económicos. Esta captura de información completa no solo enriquece el modelo, sino que también puede conducir a una estimación más precisa y robusta de los factores latentes que impulsan la economía.
* Además, al incorporar el PIB, se logra una consistencia notable en el modelado económico. Este enfoque unificado proporciona una herramienta que sirve tanto para realizar análisis estructurales profundos como para generar predicciones confiables. Esta dualidad en la funcionalidad del modelo es particularmente valiosa en el campo de la economía, donde la comprensión de las estructuras subyacentes y la capacidad de pronóstico son igualmente importantes.
* Otro beneficio crucial es el aprovechamiento de la estructura temporal inherente al PIB. Al incluir esta variable en el modelo, se puede capturar su dinámica única y sus patrones de evolución a lo largo del tiempo. Esta característica es especialmente útil para realizar predicciones a corto plazo, ya que el modelo puede incorporar y proyectar las tendencias recientes y los ciclos propios del PIB. Esta capacidad de capturar y extrapolar la dinámica temporal del PIB mejora significativamente la precisión de las predicciones económicas a corto plazo, proporcionando información e intuición valiosas para la toma de decisiones en tiempo real.

Desventajas de incluir el PIB:

La incorporación del PIB en el modelo factorial dinámico, aunque beneficiosa en muchos aspectos, también presenta ciertos desafíos y potenciales inconvenientes que merecen una cuidadosa consideración.

* Uno de los riesgos más significativos es el del sobreajuste. Al incluir el PIB, existe la posibilidad de que el modelo se adapte excesivamente a los patrones históricos específicos de esta variable, lo que podría comprometer su capacidad para realizar predicciones precisas fuera de la muestra de entrenamiento. Este sobreajuste podría llevar a un modelo que funciona excepcionalmente bien con datos pasados, pero que pierde eficacia al enfrentarse a nuevas situaciones económicas. Esta desventaja puede minimizarse mediante la elección de los modelos y especificaciones que mejor ajusten las predicciones en un período de test a los datos reales observados.
* Otro aspecto problemático surge de la naturaleza de la publicación del PIB. Esta variable macroeconómica crucial generalmente se publica con un retraso significativo, lo que crea una dependencia de datos no inmediatamente disponibles. Esta característica puede complicar la realización de predicciones en tiempo real, ya que el modelo requeriría el valor más reciente del PIB para generar pronósticos actualizados. Esta limitación temporal podría reducir la utilidad del modelo en situaciones donde se necesitan estimaciones económicas rápidas y actuales.
* Además, la inclusión del PIB en un modelo cuyo objetivo principal es precisamente predecir esta variable puede introducir un elemento de circularidad en la lógica de predicción. Esta situación podría llevar a un razonamiento circular donde el modelo se basa en el PIB para predecir el mismo PIB, lo que potencialmente podría distorsionar los resultados o la interpretación de las predicciones. En todo caso, esta es la propia “filosofía” de los modelos ARIMA, lo que, dadas estas posibles peculiaridades, es de uso habitual en modelos de predicción. En este caso sería la misma idea que los modelos de vectores autorregresivos, dado que estamos ante sistemas multivariantes.
* Finalmente, es importante considerar que la adición del PIB al modelo incrementa inevitablemente su complejidad. Esta complejidad adicional no siempre se traduce en una mejora proporcional de la capacidad predictiva del modelo. En algunos casos, podría incluso dificultar la interpretación de los resultados o aumentar el riesgo de errores en la especificación del modelo, sin ofrecer beneficios sustanciales en términos de precisión predictiva.

Una razón de peso para incluir el PIB en el sistema de indicadores reside en el objetivo final de realizar un nowcasting del indicador ante la llegada de información novedosa.

Así, la incorporación del PIB en el DFMQ para el nowcasting ofrece una serie de ventajas significativas que enriquecen y potencian el análisis económico de este indicador en tiempo real. Esta inclusión permite al modelo capturar de manera integral las complejas interrelaciones entre el PIB y otros indicadores económicos, proporcionando una visión más completa y precisa del estado general de la economía. Al aprovechar la información histórica, el modelo puede aprender y utilizar eficazmente los patrones que relacionan los indicadores de alta frecuencia con el PIB, mejorando así la estimación del PIB actual.

Una de las fortalezas clave de los DFM es su capacidad para manejar datos de frecuencias mixtas, lo que resulta particularmente útil al combinar el PIB trimestral con indicadores mensuales o diarios. Esta característica permite una interpolación eficiente de los valores del PIB para períodos intermedios, esencial para un nowcasting preciso. Además, aunque el PIB se publica con cierto retraso, su inclusión en el modelo facilita actualizaciones más coherentes y robustas a medida que se dispone de nueva información, mejorando la precisión de las estimaciones a lo largo del tiempo.

La presencia del PIB en el modelo también contribuye a una mejor estimación de los factores latentes que impulsan la economía, haciéndolos más representativos del estado económico general. Esto, a su vez, asegura una mayor consistencia entre las proyecciones a corto y mediano plazo. Los DFMQ demuestran una notable capacidad para manejar revisiones de datos, permitiendo que el modelo se ajuste automáticamente a las actualizaciones del PIB, lo que refina continuamente la precisión del nowcasting.

La flexibilidad inherente a este enfoque no limita la realización de análisis específicos, ya que los factores estimados pueden extraerse y utilizarse en modelos complementarios si es necesario. Además, la inclusión del PIB permite una evaluación directa de cómo los shocks en diferentes indicadores afectan a esta variable clave, lo cual es muy útil para el análisis de escenarios en tiempo real. Finalmente, esta incorporación mejora la interpretabilidad de los factores latentes estimados, facilitando su comprensión en términos de su impacto en el crecimiento económico general.

Coeficiente de determinación (varianza explicada)

Imagen que contiene Texto

Descripción generada automáticamente

Esta salida muestra los resultados del análisis de factores dinámicos en términos de su capacidad explicativa de las diferentes variables incluidas en el modelo, específicamente los coeficientes de determinación (R^2) para dichas variables en relación con dos factores latentes. Veamos en detalle:

1. La tabla muestra las "Top ten variables explained" (las 10 principales variables explicadas) por dos factores, etiquetados como 0 y 1.
2. Para cada factor, se muestran las variables y su correspondiente R^2 (coeficiente de determinación).
3. Factor 0:
   1. La variable más explicada es "pib" con un R^2 de 0.89
   2. Seguida por "Cifra negocios del sector servicios - Andalucía" con 0.87
   3. Otras variables importantes incluyen ICMA General, matriculación de turismos, y afiliados SS Total en Andalucía
4. Factor 1:
5. La variable más explicada es "Matriculación de turismos - Andalucía" con un R^2 de 0.40
6. Seguida por "ICMA General sin estaciones de servicios" con 0.30
7. "Pernoctaciones - Andalucía" con 0.24
8. Observaciones generales:
9. El Factor 0 parece estar más relacionado con indicadores económicos generales (PIB, cifra de negocios, ICMA)
10. El Factor 1 parece estar más relacionado con turismo y servicios específicos (matriculación de turismos, pernoctaciones)
11. El Factor 0 explica una mayor proporción de la varianza de sus variables top (R^2 más altos) comparado con el Factor 1

Esta salida ayuda a identificar qué variables están más fuertemente asociadas con cada factor latente en el modelo, lo que puede ser útil para interpretar estos factores y entender las dinámicas subyacentes en los datos económicos de Andalucía.

**Nowcasting**

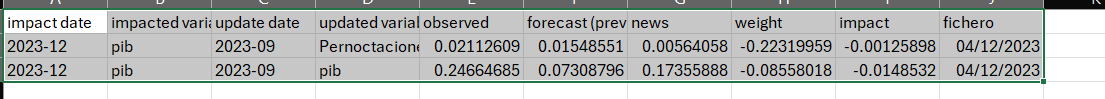
El nowcasting es una técnica de análisis económico que se ha vuelto cada vez más relevante en el contexto en el que encaja este encargo. El nowcasting, término que combina "now" (ahora) y "forecasting" (pronóstico), es una metodología que busca estimar el estado actual de la economía en tiempo real o casi real. A diferencia de las previsiones económicas tradicionales que se centran en el futuro, el nowcasting intenta capturar la situación económica presente utilizando datos de alta frecuencia y modelos estadísticos avanzados.

En el contexto de los indicadores económicos de Andalucía, el nowcasting sería particularmente útil por varias razones:

* 1. Retraso en los datos oficiales: Las estadísticas económicas oficiales, como el PIB trimestral, suelen publicarse con un retraso significativo. El nowcasting permite tener una estimación más inmediata del estado de la economía andaluza a través de la creación de un indicador sintético.
  2. Integración de múltiples fuentes de datos: Como se observa en el análisis de factores dinámicos, hay múltiples variables que influyen en la economía regional (PIB, cifra de negocios del sector servicios, matriculación de turismos, pernoctaciones, etc.). El nowcasting puede integrar estos diversos indicadores para proporcionar una visión más completa y actualizada.
  3. Manejo de datos de alta frecuencia: Algunos indicadores, como las cifras de turismo o las exportaciones, pueden tener una frecuencia mensual o incluso semanal. El nowcasting permite incorporar estos datos más recientes en las estimaciones económicas.
  4. Identificación de puntos de inflexión: En una economía dinámica como la de Andalucía, con sectores clave como el turismo y los servicios, el nowcasting puede ayudar a identificar rápidamente cambios en las tendencias económicas.
  5. Apoyo a la toma de decisiones: Para los responsables políticos y económicos de Andalucía, contar con estimaciones actualizadas del estado de la economía es crucial para la toma de decisiones informadas y oportunas.
  6. Manejo de la incertidumbre: En periodos de volatilidad económica, el nowcasting puede proporcionar estimaciones más robustas al incorporar datos en tiempo real.

El análisis de factores dinámicos es, de hecho, una herramienta que se utiliza a menudo en modelos de nowcasting. Estos modelos pueden ayudar a extraer señales comunes de un gran conjunto de variables económicas, lo que es esencial para obtener una imagen clara y actualizada de la economía andaluza.

Ejemplo de salida de fiuchero Nowcasting



En esta tabla se presenta, a modo de ejemplo, la información detallada sobre el impacto de actualizaciones de datos en las previsiones económicas, como parte de un ejercicio de nowcasting para la economía de Andalucía. Dicha tabla refleja el proceso de actualización y revisión de pronósticos económicos basados en nueva información. Cada fila representa una actualización específica y su impacto en las previsiones del PIB (Producto Interno Bruto).

Elementos clave de la tabla:

1. Impact date (Fecha de impacto): Ambas entradas muestran diciembre de 2023 (2023-12), indicando el último mes de la serie sobre la que queremos realizar el análisis de predicción e impacto de las variables (CONSEJO: indicar el último mes del trimestre sobre el que se está haciendo el análisis en tiempo real).
2. Impacted variable (Variable impactada): En ambos casos es el PIB, lo que sugiere que estamos observando cómo diferentes actualizaciones afectan la estimación del PIB.
3. Update date (Fecha de actualización): Septiembre de 2023 (2023-09) para ambas entradas, indicando cuándo se recibió la nueva información. Concretamente es el mes en el que se observa el nuevo dato para las variables o indicadores mostrados (updated variable).
4. Updated variable (Variable actualizada):
   1. Primera fila: Pernoctaciones.
   2. Segunda fila: PIB (una actualización directa de los datos del PIB).
5. Observed (Observado): Muestra el valor observado de la variable actualizada (en este modelo sería la tasa de crecimiento intertrimestral en %).
6. Pernoctaciones: 0.02112609
7. PIB: 0.24664685
8. Forecast (Pronóstico): El valor pronosticado antes de la actualización.
9. Pernoctaciones: 0.01548551
10. PIB: 0.07308796
11. Prev news (Noticias previas): Representa la diferencia entre el valor observado y el pronóstico anterior.
12. Weight (Peso): Indica la importancia relativa de cada variable en el modelo de pronóstico.
13. Pernoctaciones: -0.23199959
14. PIB: -0.08558018
15. Impact (Impacto): Muestra el efecto neto de la actualización en la previsión del PIB.
16. Actualización de Pernoctaciones: -0.00125898
17. Actualización directa del PIB: -0.0148532
18. Fichero (Archivo): Fecha del archivo de datos, 04/12/2023 para ambas entradas.

Interpretación en el contexto del nowcasting:

* 1. Actualización continua: La tabla muestra cómo el modelo de nowcasting incorpora nueva información a medida que está disponible. Las actualizaciones de septiembre están afectando las previsiones para diciembre (en realidad sería el IV trimestre).
  2. Múltiples variables: El modelo no solo usa actualizaciones directas del PIB, sino también indicadores relacionados como las pernoctaciones turísticas, lo que permite una visión más completa y actualizada de la economía.
  3. Cuantificación del impacto: El modelo asigna un impacto específico a cada actualización. Por ejemplo, la actualización de las pernoctaciones tiene un impacto menor (-0.00125898) en comparación con la actualización directa del PIB (-0.0148532).
  4. Revisión de pronósticos: La diferencia entre los valores observados y pronosticados muestra cómo el modelo ajusta sus previsiones basándose en nuevos datos.
  5. Ponderación de variables: Los pesos negativos sugieren que estas variables tienen una relación inversa con el pronóstico final del PIB en este período específico.
  6. Precisión temporal: La inclusión de fechas exactas para el impacto, actualización y archivo de datos demuestra la naturaleza en tiempo real del nowcasting.

Este tipo de análisis detallado permite tener una comprensión más precisa y actualizada del estado actual de la economía, facilitando decisiones más informadas y oportunas. La capacidad de cuantificar el impacto de diferentes variables y actualizaciones de datos es crucial para entender las dinámicas económicas en tiempo real y ajustar las previsiones económicas de manera más precisa.

¿Qué son los “weights”?

Los weights (pesos) en el contexto de un modelo de nowcasting típicamente provienen de la estimación econométrica. Matemáticamente, podemos explicar su origen de la siguiente manera:

1. Recordemos que el modelo tiene la forma:

Donde:

es el vector de variables observadas (como PIB, pernoctaciones, etc.)

es un vector de factores latentes

es la matriz de cargas factoriales (factor loadings)

es el término de error

Los weights estarían relacionados con los elementos de , que se estiman típicamente usando métodos como máxima verosimilitud o componentes principales.

2. Bridge equation:

Una ecuación puente podría tener la forma:

Donde son indicadores de alta frecuencia y son los coeficientes. En este caso, los weights estarían directamente relacionados con los estimados.

3. Cálculo del impacto:

El impacto de una actualización se calcularía como:

Impact = Weight \* (Observed - Forecast)

Donde el Weight es el coeficiente estimado que relaciona la variable actualizada con el PIB.

4. Estimación de los weights:

Los weights se obtienen generalmente minimizando una función de pérdida, como la suma de errores cuadráticos:

Usando métodos como mínimos cuadrados ordinarios o, en modelos más complejos, estimadores bayesianos o de máxima verosimilitud.

5. Interpretación:

Los weights negativos en la tabla (-0.23199959 para pernoctaciones, -0.08558018 para PIB) sugieren que en este período específico, un aumento en estas variables está asociado con una revisión a la baja del pronóstico del PIB. Esto podría deberse a factores estacionales, ajustes en el modelo, o características específicas del ciclo económico actual.

6. Actualización dinámica:

En un modelo de nowcasting, estos weights se actualizan frecuentemente a medida que llegan nuevos datos, mediante el uso de técnicas de filtrado como el filtro de Kalman:

Donde es la ganancia de Kalman, que determina cómo se actualizan las estimaciones con nueva información.