El filtro de Kalman

José Antonio Camarena Ibarrola

Que es un Filtro de Kalman?

- Es un algoritmo recursivo óptimo de procesamiento de datos
- Hay muchas formas de definir "óptimo"
- El filtro de Kalman es óptimo respecto a casi cualquier criterio que tenga sentido
- El filtro de Kalman toma en cuenta toda la información que se tenga.
- Procesa todas las mediciones sin importar su precisión para estimar el valor de las variables de interés
- Hace uso de:

1. Conocimiento del sistema y de la dinámica de los instrumentos de medición,
2. La descripción estadística de los errores de medición y de la incertidumbre de los modelos dinámicos.
3. Cualquier información disponible de las condiciones iniciales de las variables de interés

Ejemplo

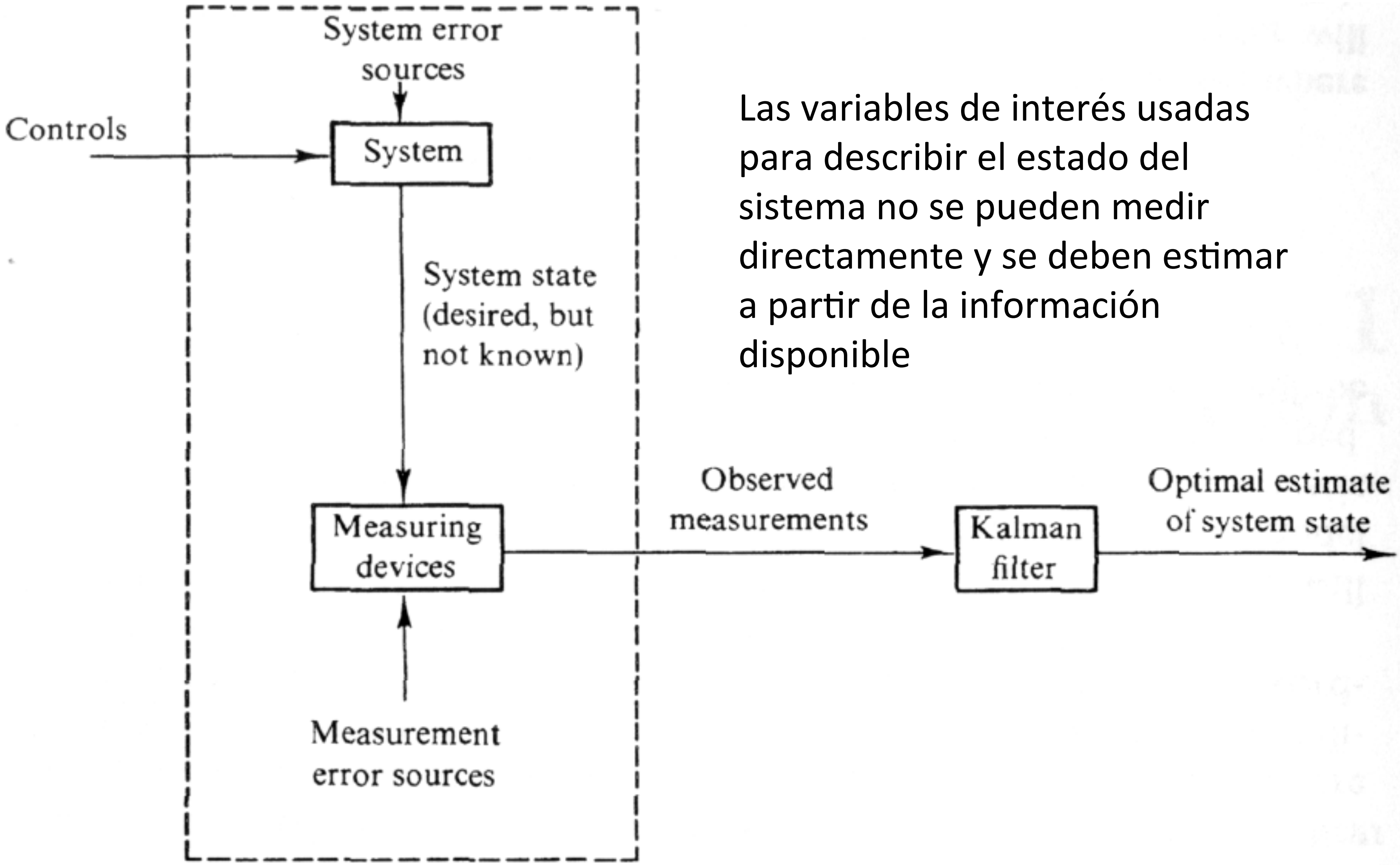
- Para determinar la velocidad de una aeronave podríamos diseñar un filtro de Kalman que tome en cuenta:

1. Un radar de efecto Doppler
2. Velocímetro de navegación inercial
3. La presión estática
4. Viento relativo
5. Lo que diga el piloto

Filtro recursivo?

- Es recursivo porque no requiere almacenar todos los datos previos y reprocesarlos cada vez que llega una nueva información
- Es un filtro en el sentido de que es un programa que trabaja con muestras tomadas en tiempo discreto en lugar de entradas en tiempo continuo

Uso del filtro de Kalman



Las variables de interés usadas para describir el estado del sistema no se pueden medir directamente y se deben estimar a partir de la información disponible

Por qué necesitamos al filtro de Kalman?

- Los sistemas están normalmente controlados por mas variables que las que el diseñador de los sistemas proveen como “variables de control”
- La relación entre las variables de estado y las variables que se miden se conoce solo con cierto grado de incertidumbre
- Las mediciones son corrompidas en cierto grado con ruido, desviaciones e imprecisiones de dispositivos
- Las diferentes dinámicas de los diferentes instrumentos de medición utilizados se pueden confundir con la dinámica del sistema
- El filtro de Kalman produce una estimación óptima estadísticamente
- Toma en cuenta todos las mediciones mas la información a priori que se tenga del sistema y de los dispositivos de medición

Enfoque Bayesiano

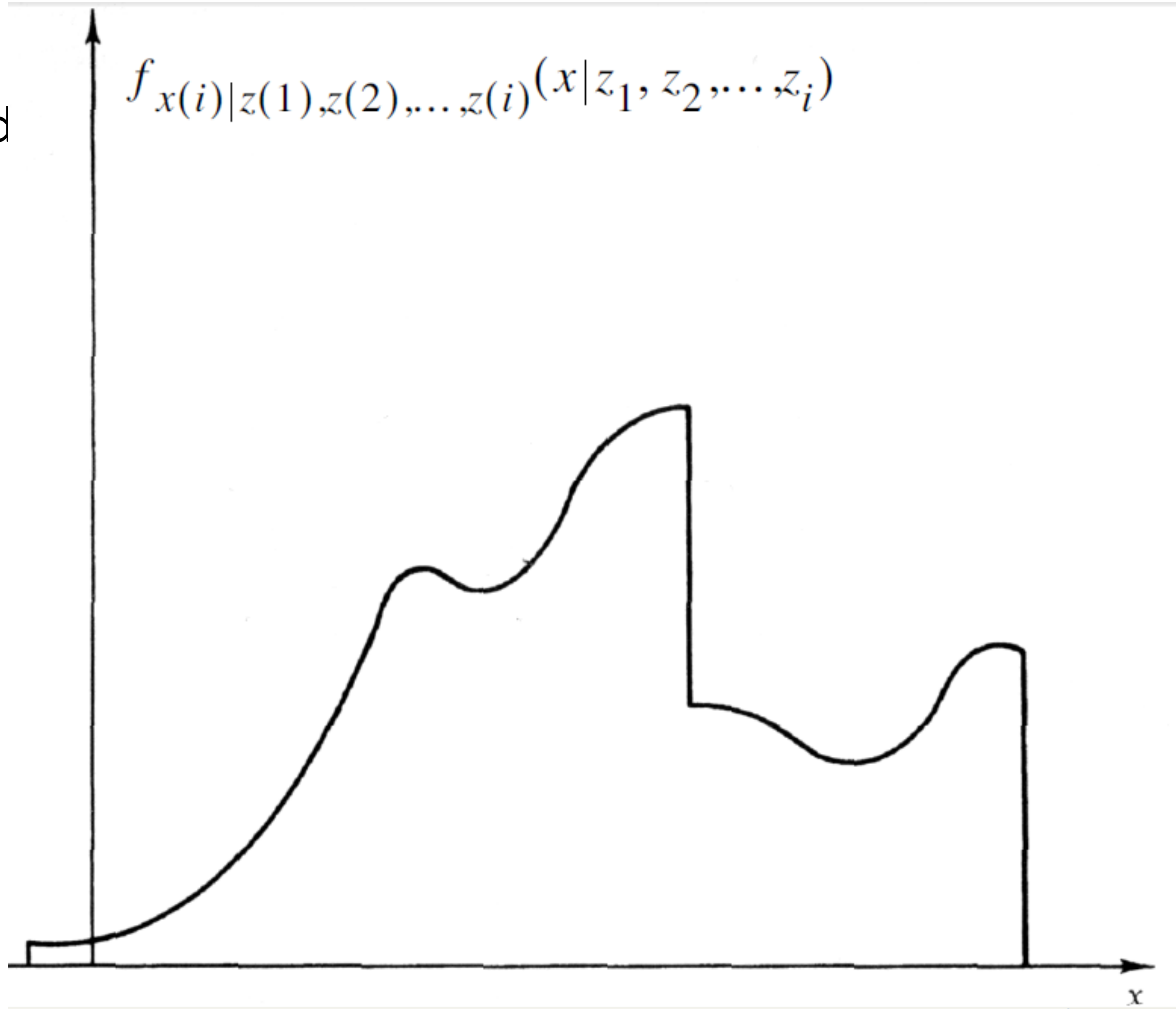
- El filtro debe propagar la densidad de probabilidad condicional de las cantidades deseadas condicionado al conocimiento de los datos que provienen de los dispositivos de medición.
- Considere la densidad de probabilidad condicional del valor de una variable *x* en un instante *i (x(i)),* condicionada al conocimiento de que una variable medida en el instante 1 (*z(1)*) tomó el valor *z1* y al medirla en el instante 2 (*z(2)*) tomó el valor *z2*, etc.

$$f_{x(i)|z(1),z(2),\ldots,z(i)}(x|z_1, z_2,\ldots,z_i)$$

- Por ejemplo *x(i)* pudiera ser la posición de un vehículo en una sola dimensión en el instante *i* y *z(j)* pudiera ser la posición del vehículo en dos dimensiones reportada por dos radares separados en el instante *j*

La densidad condicional

- La densidad condicional contiene toda la información disponible acerca de *x(i).*
- Indica la probabilidad de *x(i)* de asumir cualquier valor específico , para el valor dado de todas las mediciones tomadas hasta el instante *i* .
- Es condicional porque el valor reportado depende de todas las mediciones tomadas
- Si fuera un pico tendríamos certeza del valor de x(i)

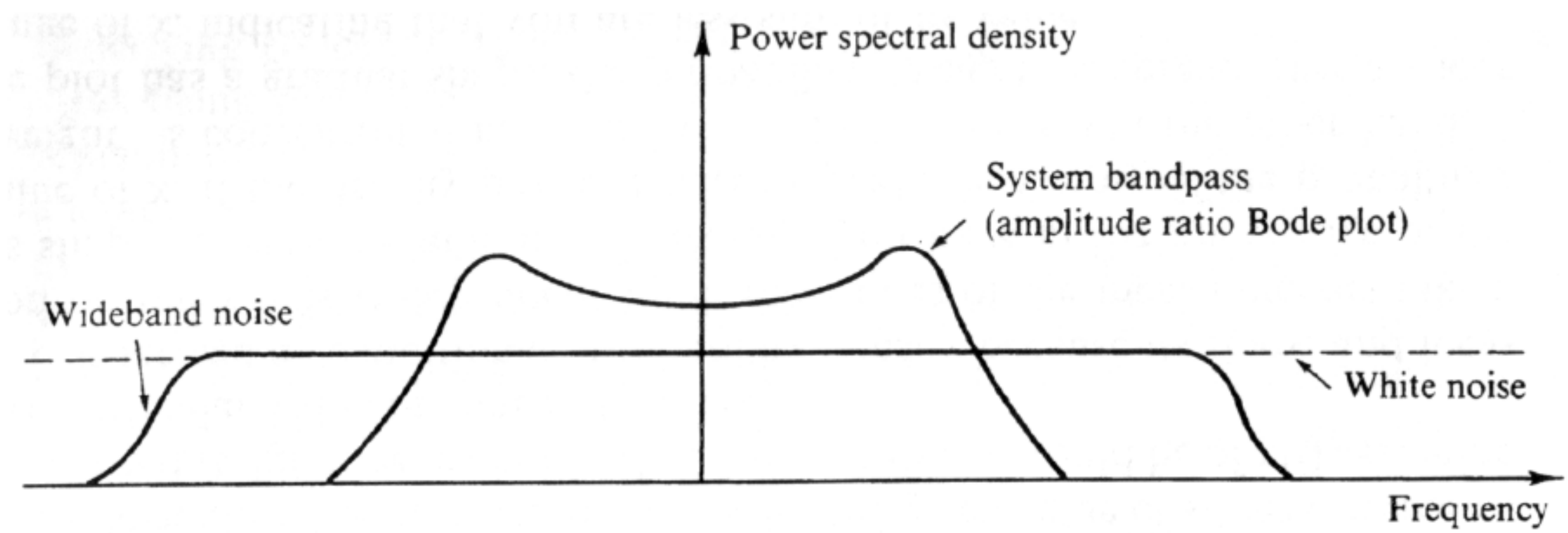


Restricciones

- Un filtro de Kalman realiza la propagación de la densidad de probabilidad para problemas en las que el sistema puede ser descrito mediante un modelo lineal y los ruidos en las mediciones son gaussianos
- Bajo estas restricciones, el filtro de Kalman es óptimo
- Al hacer la propagación podemos utilizar la media, la mediana o la moda
- En la distribución gaussiana estas 3 medidas coinciden
- Si se elimina la presunción de gaussianidad, el filtro de Kalman es un filtro con el menor error de la clase de filtros no-sesgados
- Para muchas aplicaciones estas restricciones se cumplen

Presunciones básicas

- Un modelo lineal. Se puede ajustar a lineal cuando hay no-linealidades.
- Los sistemas lineales son muy deseables pues hay muchas herramientas y teoría mas completa y práctica
- Ruido blanco significa que no tiene correlación con el tiempo (en nada ayuda conocer el tiempo en la predicción del valor del ruido)
- Al considerar que el ruido es blanco se vuelve tratable el problema, de todos modos al sistema le da lo mismo el ruido blanco que el ruido de ancho de banda amplia para el ancho de banda al que responde el sistema

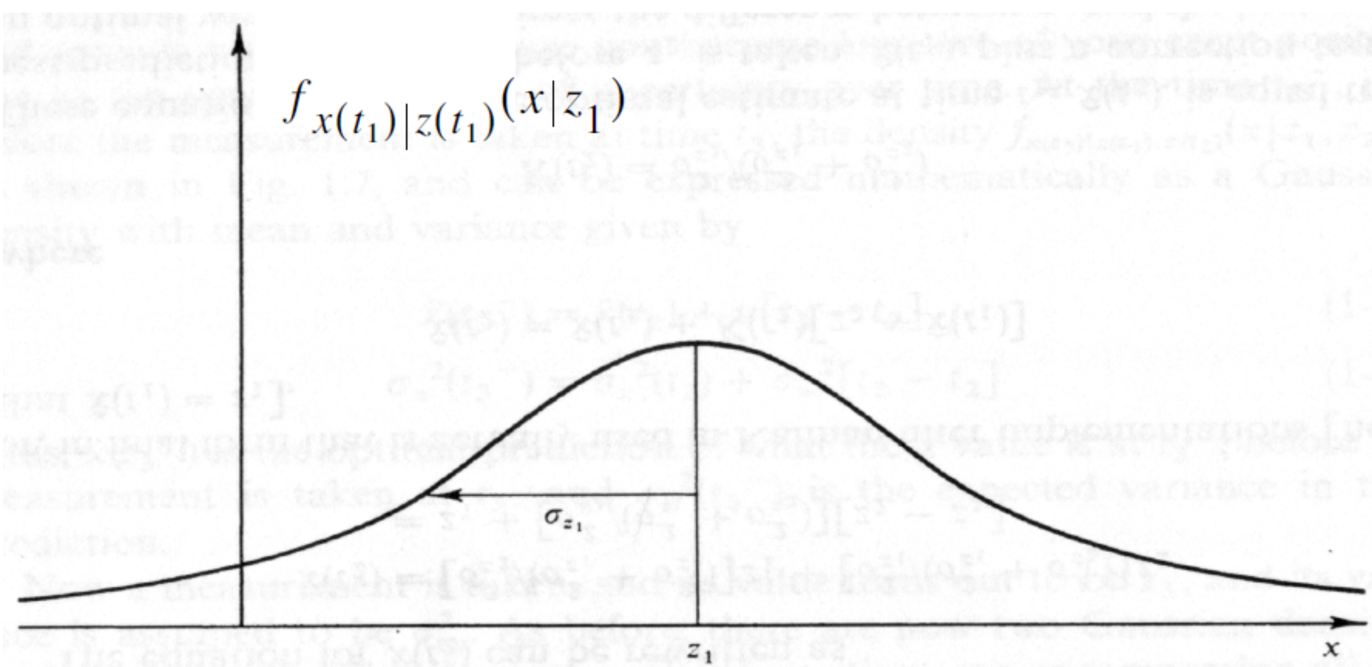


Presunciones básicas

- Una distribución gaussiana queda determinada solo con los primeros dos momentos de la distribución los cuales son fáciles de averiguar normalmente
- Las tres presunciones juntas convierten el problema en tratable

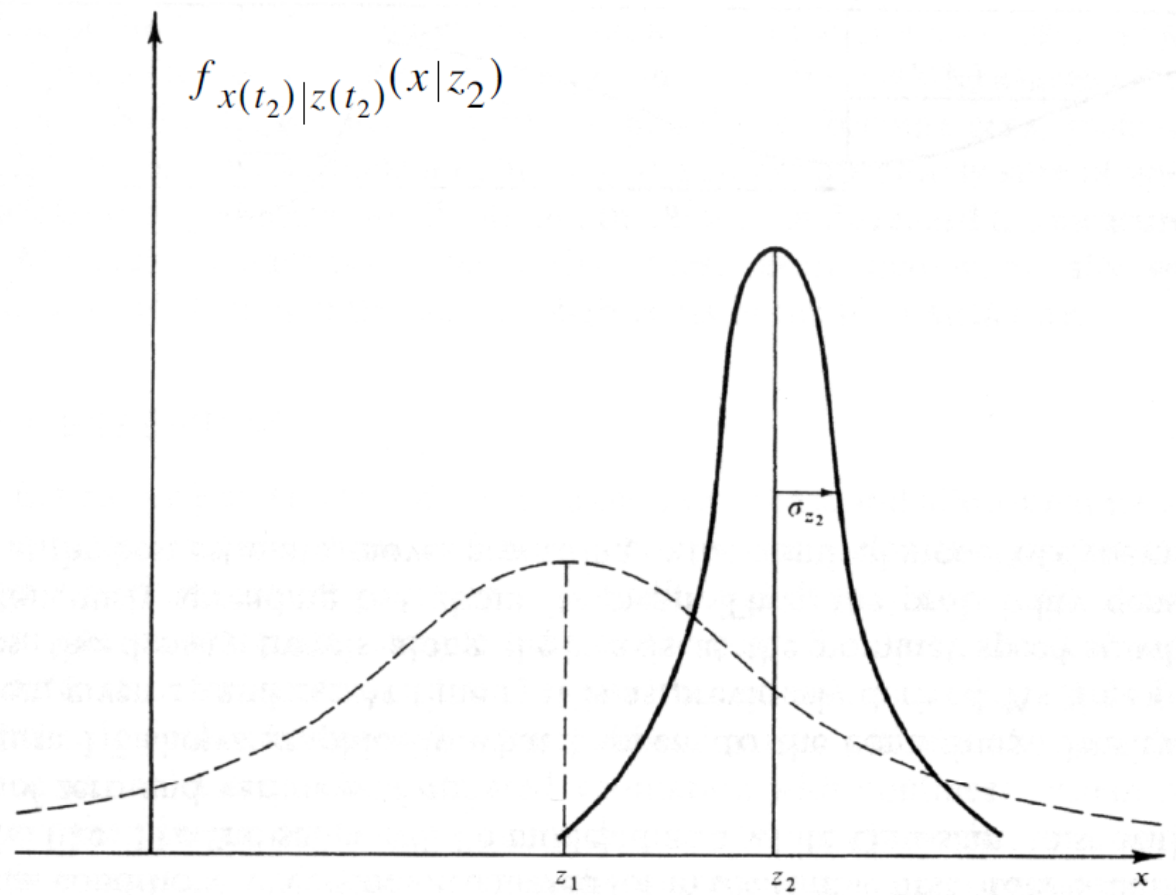
Ejemplo

- Suponga que no tiene idea de su posición en el mar
- Localiza una estrella en el cielo para tratar de ubicarse
- Suponga por simplicidad que el sistema es de una sola dimensión
- En el instante t1, determinas que tu posición es z1
- Debido a las imprecisiones inherentes la posición es incierta
- Supongamos que decides que tu precisión es tal que la desviación estándar es σ_{z_1}
- Entonces podemos establecer que la probabilidad condicional de $x(t_1)$, condicionada al valor observado de la medición es z1



Ejemplo (cont)

- Hasta este momento la mejor estimación de la posición es $\hat{x}(t_1) = z_1$ y la varianza es $\sigma_x^2(t_1) = \sigma_{z_1}^2$
- Enseguida, un marino con mas experiencia toma otra medición en el instante t2 muy poco después de t1 de manera que la posición real no ha cambiado y obtiene z2 con varianza σ_{z_2} menor puesto que es mas experimentado



Actualizando la distribución

$$\mu = [\sigma_{z_2}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]z_1 + [\sigma_{z_1}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]z_2$$

$$1/\sigma^2 = (1/\sigma_{z_1}^2) + (1/\sigma_{z_2}^2)$$

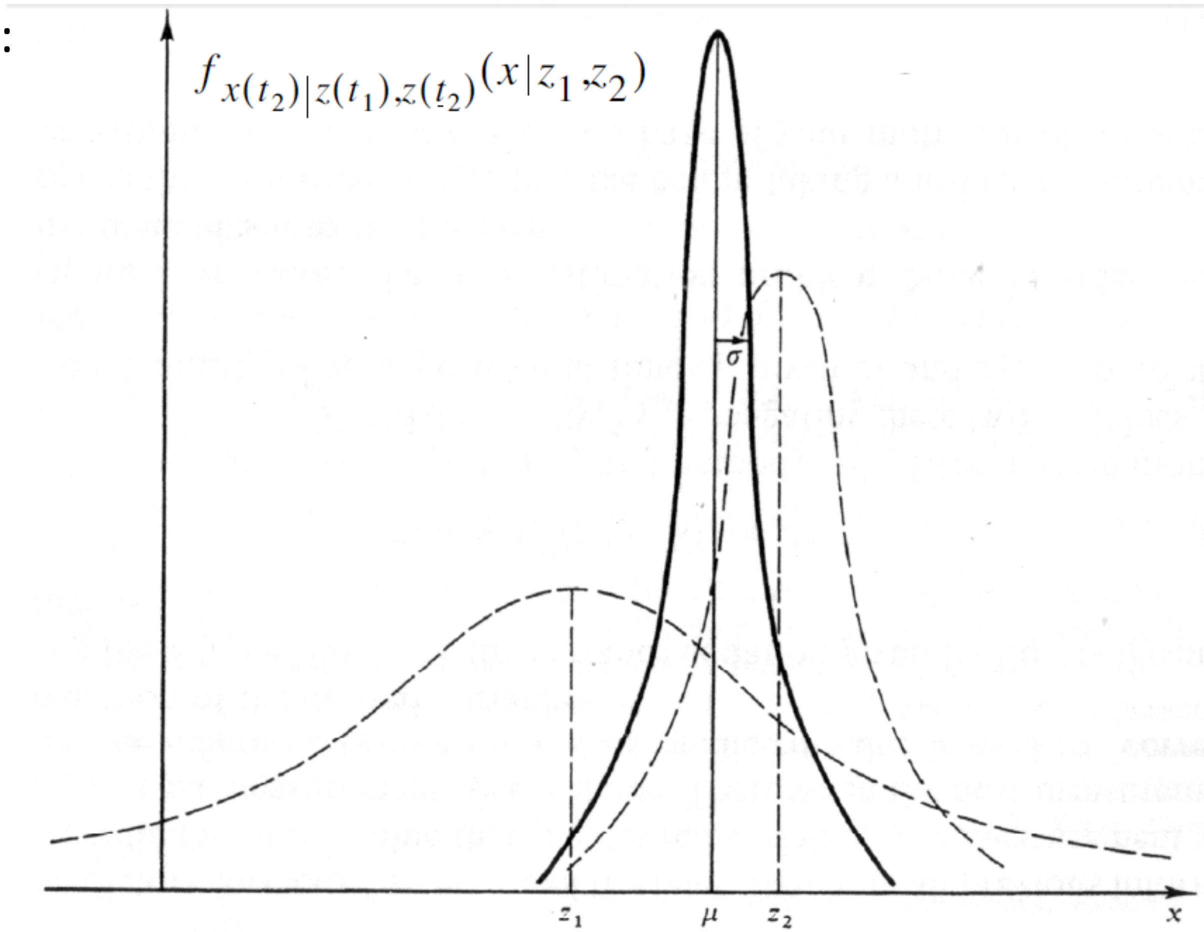
Y la mejor estimación es:

$$\hat{x}(t_2) = \mu$$

Observe que

$$\sigma < \sigma_{z_1}$$

$$\sigma < \sigma_{z_2}$$



Estructura predictor-corrector del filtro de Kalman

$$\hat{x}(t_2) = [\sigma_{z_2}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]z_1 + [\sigma_{z_1}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)]z_2$$
$$= z_1 + [\sigma_{z_1}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)][z_2 - z_1]$$

Puede reescribirse como: $$\hat{x}(t_2) = \hat{x}(t_1) + K(t_2)[z_2 - \hat{x}(t_1)]$$

donde: $$K(t_2) = \sigma_{z_1}^2/(\sigma_{z_1}^2 + \sigma_{z_2}^2)$$

La mejor estimación en el instante t2 es igual a la mejor predicción antes de que llegara z2 mas un término de corrección formado por un factor de peso optimizado que pondera la diferencia entre z2 y la mejor predicción que había antes de la llegada de z2

Estructura predictor-corrector del filtro de Kalman

- Basada en toda la información previa hacemos una predicción del valor que las variables deseadas tendrán la próxima vez que se haga la medición
- Cuando la siguiente medición se hace, la diferencia entre esta y la predicción se usa para corregir la predicción de las variables deseadas
- La varianza también puede reescribirse como:

$$\sigma_x^2(t_2) = \sigma_x^2(t_1) - K(t_2)\sigma_x^2(t_1)$$

- Observe que $\hat{x}(t_2)$ y $\sigma_x^2(t_2)$ encierran toda la información de $f_{x(t_2)|z(t_1),z(t_2)}(x \mid z_1, z_2)$
- Propagando estas dos variables, la densidad condicional de tu posición en el instante t2 dados z1 y z2 queda completamente especificada

Incorporando dinámica

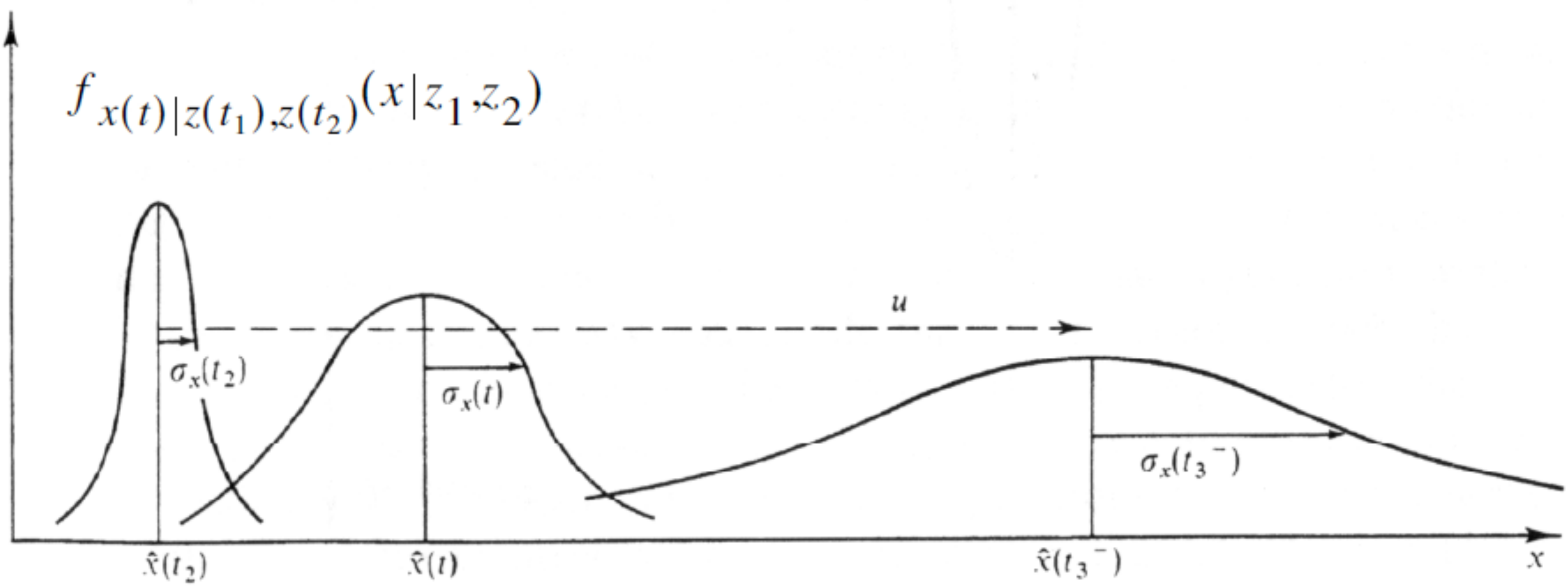
- Suponga que viaja por algún tiempo antes de tomar otra medición
- Asuma que el modelo del movimiento tiene la forma $dx/dt=u+w$

donde *u* es la velocidad y *w* es ruido que representa incertidumbre sobre la velocidad real debida a disturbios y condiciones no controlables

- Asuma que *w* se modela mediante una gaussiana con media cero y varianza σ_w^2

Propagación de la densidad condicional

A medida que el tiempo avanza, la densidad condicional viaja a lo largo del eje x a la velocidad u mientras que se ensancha alrededor de su media debido a la constante adición de incertidumbre



Primero la predicción

En el instante t_3^- , justo antes de que se tome la medición en t_3
la densidad $f_{x(t_3)|z(t_1),z(t_2)}(x \mid z_1, z_2)$

es una gaussiana con media y varianza dadas por:

$$\hat{x}(t_3^-) = \hat{x}(t_2) + u[t_3 - t_2]$$

$$\sigma_x^2(t_3^-) = \sigma_x^2(t_2) + \sigma_w^2[t_3 - t_2]$$

$\hat{x}(t_3^-)$ es la predicción óptima del valor que tendrá x en t_3^-

$\sigma_x^2(t_3^-)$ es la incertidumbre de la predicción

Y luego la corrección

Ahora se toma una medición y es z_3 y su varianza se asume es $\sigma_{z_3}^2$

Como antes, tenemos dos gaussianas que contienen información acerca de la posición, una con la información antes de la medición y la otra es la que con la información que trae consigo la medición

Combinamos la densidad con media $\hat{x}(t_3^-)$ y varianza $\sigma_x^2(t_3^-)$ con la densidad con media z_3 y varianza $\sigma_{z_3}^2$

$$\hat{x}(t_3) = \hat{x}(t_3^-) + K(t_3)[z_3 - \hat{x}(t_3^-)]$$

$$\sigma_x^2(t_3) = \sigma_x^2(t_3^-) - K(t_3)\sigma_x^2(t_3^-)$$

donde $$K(t_3) = \sigma_x^2(t_3^-)/[\sigma_x^2(t_3^-) + \sigma_{z_3}^2]$$

Sin confianza en la medición

Si σ_{z_3} es grande, entonces $K(t_3)$ será pequeño

Si $\sigma_{z_3}^2 \rightarrow \infty$ entonces $K(t_3)$ será cero y entonces $\hat{x}(t_3) = \hat{x}(t_3^-)$

Y la medición nueva infinitamente ruidosa es completamente ignorada

Sin confianza en la dinámica del sistema

Si σ_w^2 es grande, entonces también lo será $K(t_3)$

Esto significa que no tienes mucha confianza en el modelo lineal y se le dará importancia a la medición, en un caso extremo si

$$\sigma_w^2 \to \infty\,,\ \sigma_x^2(t_3^-) \to \infty \ \text{ y } \ K(t_3) \to 1$$

$$\hat{x}(t_3) = \hat{x}(t_3^-) + 1 \cdot [z_3 - \hat{x}(t_3^-)] \ = \ z_3$$

Entonces en el caso extremo de no tener ninguna confianza en el modelo, la solución óptima es ignorar la predicción y tomar la medición como la estimación óptima

Certeza en la predicción

Si $\sigma_x^2(t_3^-) = 0$ entonces $K(t_3) = 0$

Significa que se tiene absoluta seguridad en la predicción y se descartará la próxima medición

Ejemplo

- Un tanque de agua llenándose a velocidad Cte $L_t = L_{t-1} + f$
- El estado del sistema se define mediante el vector

$$x = \begin{pmatrix} x_l \\ x_f \end{pmatrix}$$

- El nivel del líquido es x_l
- La rapidez con la que se llena el tanque es $x_f = dx_l / dt$
- Observe que $\ddot{x}_l = \dot{x}_f = 0$ Puesto que el tanque se llena a Velocidad constante

$$\dot{x} = Ax$$

$$\begin{pmatrix} \dot{x}_l \\ \dot{x}_f \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_l \\ x_f \end{pmatrix}$$

Transformando A en F

- Necesitamos un proceso discreto, entonces la matriz A debe ser tiempo-discretizada
- $$F = e^{A\Delta t}$$

Sabemos que: $e^{x} = 1 + \sum_{i=1}^{\infty} \frac{x^{i}}{i!}$

Entonces: $F = I + \sum_{i=1}^{\infty} \frac{A^{i}\Delta t^{i}}{i!}$

$$F = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \Delta t \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \frac{\Delta t^{2}}{2} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} + \ldots$$

$$F = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix}$$

Si $\Delta t = 1$ Entonces: $F = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

Modelando el ruido de la medición

- Asumiendo que el ruido ocurre solo en la parte de medición asociada al llenado

$$\mathbf{Q} = \begin{pmatrix} 0 & 0 \\ 0 & q_f \end{pmatrix}$$

Donde q_f es el ruido del llenado

Modelando el ruido de la medición

- El proceso continuo se puede aproximar mediante un proceso discreto en tiempo mediante:

$$\mathbf{Q}(\Delta t) = \int_0^{\Delta t} e^{\mathbf{A}\tau}\mathbf{Q}e^{\mathbf{A}^\top \tau}\,d\tau$$

$$e^{A\tau} = I + \sum_{i=1}^{\infty} \frac{A^i \tau^i}{i!}$$

$$e^{A\tau} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \tau \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \frac{\tau^2}{2} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} + \ldots = \begin{pmatrix} 1 & \tau \\ 0 & 1 \end{pmatrix}$$

Modelando el ruido de la medición

- De manera similar

$$e^{A^T\tau} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \tau \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \frac{\tau^2}{2} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} + \ldots = \begin{pmatrix} 1 & 0 \\ \tau & 1 \end{pmatrix}$$

$$e^{A\tau} Q e^{A^T\tau} = \begin{pmatrix} 1 & \tau \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & q_f \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \tau & 1 \end{pmatrix} = \begin{pmatrix} 0 & \tau q_f \\ 0 & q_f \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \tau & 1 \end{pmatrix} = \begin{pmatrix} \tau^2 q_f & \tau q_f \\ \tau q_f & q_f \end{pmatrix}$$

$$\mathbf{Q}(\Delta t) = \int_0^{\Delta t} e^{\mathbf{A}\tau} \mathbf{Q} e^{\mathbf{A}^\top \tau} d\tau = \int_0^{\Delta t} \begin{pmatrix} \tau^2 q_f & \tau q_f \\ \tau q_f & q_f \end{pmatrix} d\tau$$

$$= \begin{pmatrix} \frac{\Delta t^3}{3} q_f & \frac{\Delta t^2}{2} q_f \\ \frac{\Delta t^2}{2} q_f & q_f \end{pmatrix}$$

Si $\Delta t = 1$ $\mathbf{Q} = \begin{pmatrix} q_f/3 & q_f/2 \\ q_f/2 & q_f \end{pmatrix}$

Modelando el proceso de medición

- Asumiendo que solo medimos el nivel y no el régimen de llenado

$$\mathbf{H} = (1\ ,\ 0)$$

$$\mathbf{y} = (y\ ,\ 0)^{\top}$$

Ejemplo

- Estimación inicial
- Predicción

$$x_{0|0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad P_{0|0} = \begin{pmatrix} 1000 & 0 \\ 0 & 1000 \end{pmatrix}$$

$$q_f = 0.00001$$

$$\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \hat{\mathbf{x}}_{t-1|t-1}$$

$$\mathbf{P}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^T + \mathbf{Q}_t$$

$$\hat{x}_{t|t-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$P_{t|t-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1000 & 0 \\ 0 & 1000 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} \frac{0.0001}{3} & \frac{0.0001}{2} \\ \frac{0.0001}{2} & 0.0001 \end{pmatrix} = \begin{pmatrix} 2000.00000333 & 1000.000005 \\ 1000.000005 & 1000.00001 \end{pmatrix}$$

Corrección

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t \left(\mathbf{y}_t - \mathbf{H}_t \hat{\mathbf{x}}_{t|t-1}\right)$$

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{H}_t^T \left(\mathbf{H}_t \mathbf{P}_{t|t-1}\mathbf{H}_t^T + \mathbf{R}_t\right)^{-1}$$

$$\mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t)\mathbf{P}_{t|t-1}$$

$$K_t = \begin{pmatrix} 2000.000003 & 1000.000005 \\ 1000.000005 & 1000.00001 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\left(\begin{pmatrix} 1 & 0 \end{pmatrix}\begin{pmatrix} 2000.000003 & 1000.000005 \\ 1000.000005 & 1000.00001 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1\right)^{-1}$$

$$K_t = \begin{pmatrix} 0.9995 \\ 0.49975 \end{pmatrix}$$

Supongamos una medición de $y_t = 2.462$

(Ruido real de 5)

$$\hat{x}_{t|t} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.9995 \\ 0.49975 \end{pmatrix}\left(2.462 - \begin{pmatrix} 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 2.46079 \\ 1.2304 \end{pmatrix}$$

$$p_{t|t} = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.9995 \\ 0.49975 \end{pmatrix}\begin{pmatrix} 1 & 0 \end{pmatrix}\right)\begin{pmatrix} 2000.000003 & 1000.000005 \\ 1000.000005 & 1000.00001 \end{pmatrix} = \begin{pmatrix} 0.9995 & 0.49975 \\ 0.49975 & 500.24988 \end{pmatrix}$$

En Matlab

```
F=[1 1;0 1];
H=[1 0]; %Matriz que relaciona mediciones con el estado del sistema
q=0.00001; %ruido de las mediciones;
Q=[q/3 q/2;q/2 q]; %Matriz de covarianzas de ruido de mediciones;
r=1; %ruido de las mediciones
I=[1 0;0 1] %Matriz identidad
x(:,1)=[0 0]'; %Estado inicial (nivel cero y regimen de llenado de cero)
P=[1000 0;0 1000]; % Matriz de covarianzas de distribucion de edo inicial
for i=1:1:40
   z(i)=i*0.1;
end
y=random('norm',0,1,1,40); %Ruido gaussiano con media cero y varianza 1
y=z+y;
for t=1:1:39
%Prediccion
x(:,t+1)=F*x(:,t)   %observe que no hay entradas (B=0; u=0)
P=F*P*F'+Q
%Correccion
K=P*H'/[H*P*H'+r] % K=P*H'*inv([H*P*H'+r]) Para este ejemplo es lo mismo
x(:,t+1)=x(:,t+1)+K*(y(t+1)-H*x(:,t+1))
P=(I-K*H)*P
t=t+1;
end
```