RECONSTRUCCIÓN DE ESTADO

Reconstrucción Óptima Lineal de Estado

Todas las versiones de problemas analizados y resueltos por medio del *Modelo de estado* hasta ahora, parten de una misma suposición y es que el vector de estado se puede determinar con exactitud y en forma completa.

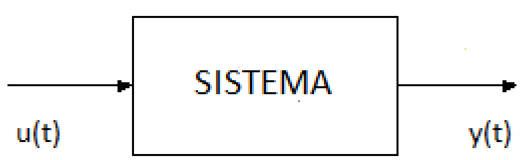
Sin embargo, esta suposición no siempre es válida. Una situación muy frecuente se produce cuando no puede accederse a determinar los parámetros del sistema para conocer el vector de estado.

Esto significa que en la ecuación diferencial de estado, dada por $\dot{\mathbf{x}}= [\mathbf{A}]\{\mathbf{x}\} + [\mathbf{b}]\, \mathbf{u}$; es imposible determinar el estado {x} por no poder acceder a los parámetros del sistema, contenidos en la matriz [A].

Se trata de una situación típica para los ingenieros de mantenimiento, cuando se encuentran frente a la necesidad de reparar o ajustar equipos muy viejos, de los cuales no se dispone de los esquemas de circuito ni los diagramas de conexión o equipos nuevos cuyos módulos se encuentran encapsulados y embebidos en resina y no pueden realizarse mediciones sobre los componentes para buscar una falla.

Aun así, en todos los casos es visible y accesible la salida del sistema, que será un cable por el que circula una señal eléctrica, un tubo que conduce un fluido a presión, un mecanismo que transmite un movimiento, un radiante que propaga calor o cualquier otra magnitud física; según se trate de un sistema eléctrico, hidrodinámico, mecánico, térmico o cualquier otro.

En general, aunque en alguna etapa o en todo el conjunto se encuentre un gabinete hermético o un módulo sellado, siempre será visible y accesible su entrada y su salida:



Sobre esta magnitud de salida pueden realizarse mediciones y obtener de ellas algunas combinaciones lineales de estado, siendo que la salida es:

$y(t)= [C]\, \{x(t)\}$

La magnitud de la salida {y} es la de un vector L-dimensional, donde (L) es usualmente menor que la dimensión (n) del vector de estado {x}; donde, además, {y} se designa como la *variable observada*.

El objetivo de este estudio es encontrar métodos para reconstruir el vector de estado desconocido o encontrar aproximaciones al vector de estado {x} desde el conocimiento de la variable observada {y}.

En particular, se desea encontrar un vector funcional F, tal que sea igual a una aproximación del vector de estado:

$\hat{x}(t)=F[y(\tau),\tau]$

donde:

$t_o \leq \tau \leq t$; para todo $t_o \leq t$; tal que $\hat{x}(t) \cong x(t)$

Este funcional **$F[y(\tau),\tau]$** produce el vector de estado reconstruido que es $\hat{x}(t)$, y es función de observaciones pasadas $y(\tau)$ y no depende, en absoluto, de observaciones futuras.

Una vez que el estado ha sido reconstruido se debería recurrir a las leyes del control de estado, para permitir un conocimiento completo del vector de estado, reemplazando el estado *real* {x(t)} por el estado *reconstruido* {$\hat{x}(t)$}.

Un método de trabajo consiste en introducir el *observador*, que es un sistema dinámico cuya salida aproxima, en tiempo creciente, al estado que se busca reconstruir.

Aunque esta aproximación no toma en cuenta, explícitamente, las dificultades que surgen como consecuencia del ruido que surge al observar, la misma busca formas de reconstruir el estado que implican un cierto grado de filtrado implícito del ruido.

Finalmente, por medio de un análisis estocástico, se puede encontrar un *observador óptimo*.

OBSERVADORES

Para reconstruir el estado {x} del sistema dado por la expresión $\dot{\mathbf{x}}\mathbf{=Ax+Bu}$; (1) donde la variable observada {y} está dada por y = C x; se propone un sistema diferencial lineal cuya salida será una aproximación conveniente del estado {x}.

Lo que debe investigarse es qué estructura debería tener este sistema y cómo debería comportarse.

Observador completo:

Si se designa ahora como $\hat{x}(t)$ a la variable de estado reconstruida, la ecuación diferencial de estado del sistema n-dimensional observador es:

$\dot{\hat{x}}$ **(t) = F $\hat{x}$ (t) + G y(t) + H u(t)** ②

Que es un observador completo para diferenciarlo de los observadores de *orden reducido* que sólo aproximan una parte del vector de estado, entregando una reconstrucción de orden $m < n$.

Este observador de orden completo entrega a su salida un vector de estado reconstruido $\{\hat{x}\}$ que es de la misma dimensión (n) que el estado {x} del sistema original.

Si se investigan las condiciones que deben satisfacer las matrices [F], [G] y [H] para que el sistema ② califique como observador para el sistema original ①; tal que sea capaz de reconstruir su vector de estado, se puede establecer que estos requisitos son los siguientes:

[F] = [A] – [K] [C] ③
[G] = [K]
[H] = [B]

Donde [K] es una matriz arbitraria, variante en el tiempo, que se denomina *ganancia del observador* y ya se verá que cumple un papel fundamental en el desempeño del sistema observador. Las matrices [A], [B] y [C] pertenecen al sistema original y son, por lo tanto, las que se desconocen por la inaccesibilidad del sistema.

Mientras que [F], [G] y [H] son matrices del observador y por lo tanto parte del diseño de este sistema y resultan conocidas o evaluables.

¿De dónde surgen estas relaciones entre las matrices?

De la diferencia entre las ecuaciones del vector de estado original ① y el vector de estado reconstruido ②; a saber:

$\dot{x}(t) - \dot{\hat{x}}(t) = [Ax+Bu] - [F\,\hat{x} + Gy + Hu]$

donde y=Cx; entonces:

$\dot{x}(t) - \dot{\hat{x}}(t) = [Ax+Bu] - [F\,\hat{x} + GCx + Hu]$ por lo tanto

$\boldsymbol{\dot{x}(t) - \dot{\hat{x}}(t) = [A - GC]\, x - [F]\, \hat{x} + [B - H]u}$ ④

En esta última ecuación se observa que, si el vector de estado reconstruido es igual al vector de estado original, $x(t) = \hat{x}(t)$; la diferencia entre sus derivadas será nula:

$\dot{x}(t) - \dot{\hat{x}}(t) = 0$

por lo tanto

$0 = [A - GC]\, x - [F]\, \hat{x} + [B - H]u$

y, operando:

$[F]\, \hat{x} + [H]\, u = [A - GC]\, x + [B]\, u$

donde, obviamente, para que $x = \hat{x}$

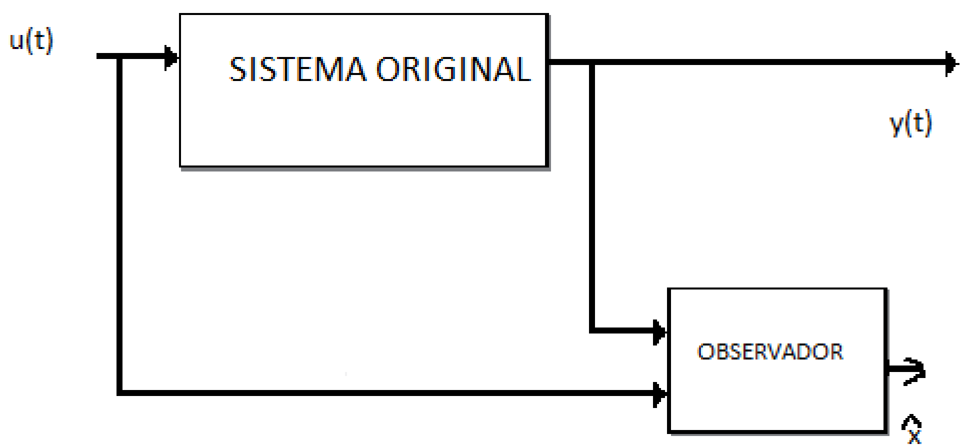
deberá ser

$[F]=[A - GC]$

$[H]=[B]$

que era lo que establecían las ecuaciones 3.

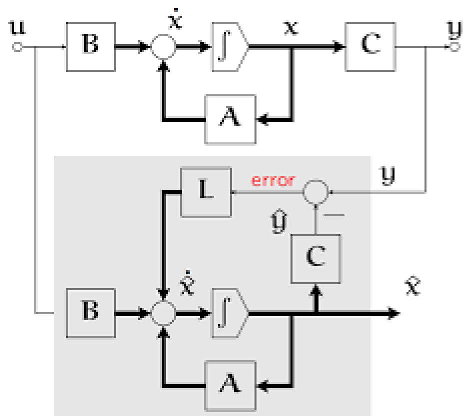
Esquema del Observador:



Esquemáticamente, el Observador es un sistema que se alimenta de la entrada y la salida del sistema original y entrega a su salida el vector de estado reconstruido $\{\hat{x}\}$

El Diagrama de Flujo de la implementación del Observador, donde se ha esquematizado su arquitectura interna, es el siguiente.

Donde los bloques sombreados simbolizan el módulo que corresponde al Observador propiamente dicho:



ERROR DE RECONSTRUCCIÓN

Dado que el vector de estado reconstruido es una aproximación del vector de estado del sistema original inaccesible, esto significa que habrá una diferencia entre ambos, que se conoce como *error de reconstrucción*.

Aunque la aproximación del sistema observador sea muy buena, siempre existirá un pequeño error en algún rango de tiempo $t > t_o$; particularmente para instantes cercanos a la observación inicial.

La evaluación de este error de reconstrucción es muy importante para conocer la aptitud del vector de estado reconstruido de representar al vector de estado original en aplicaciones del sistema observado.

En general, se define el error como:

$e(t) = x(t) - \hat{x}(t)$

Si se restan las dos ecuaciones diferenciales de estado dadas por las fórmulas ① y ②; se obtiene la ecuación de diferencia de las derivadas ④; pero como deben cumplirse las relaciones ③; resulta:

$$\dot{e}(t) = \dot{x}(t) - \dot{\hat{x}}(t) = [A - KC]\,[x(t) - \hat{x}(t)]$$

por lo tanto:

$$\mathbf{\dot{e}(t) = [A - KC]\, e(t)} \qquad ⑤$$

Esta última es la ecuación diferencial del error. Si se resuelve se podrá encontrar la función error de reconstrucción e(t); que es la solución de la ED ⑤.

Como ya hemos visto, se puede resolver por Transformada de Laplace, pero también se podría resolver por integración directa ya que se trata de una ED simple de primer orden. Vamos a utilizar las dos formas:

1) Resolución por Transformada de Laplace:

En la ED ⑤ se debe transformar miembro a miembro y se obtiene una nueva igualdad, a saber:

$\mathfrak{L}[\dot{e}(t)] = \mathfrak{L}\{[A - KC]\, e(t)\}$

Luego, operando:

$s\,E(s) - e(0) = [A - KC]\, E(s)$

para un sistema lineal dinámico, invariante en el tiempo, el factor matricial [A – KC] se comporta como si fuera una constante en términos escalares desde el punto de vista de la Transformada.

reagrupando:

$s\,E(s) - [A - KC]\, E(s) = e(0)$

donde se puede sacar factor común E(s) en el primer miembro. Y en el segundo se puede llamar E_o= e(0) *error inicial*; entonces:

$\{s[I] - [A - KC]\}\, E(s) = E_o$

y despejando E(s) por medio de premultiplicar miembro a miembro por la inversa de la matriz {s[I] – [A–KC]}, resulta:

$E(s) = \{s[I] - [A - KC]\}^{-1}\, E_o$ ⑥

Aunque no sería correcto, pero para visualizar la estructura de la antitransformada, esto es equivalente, en términos escalares, a la siguiente expresión:

$E(s) = \frac{E_o}{s[I] - [A - KC]}$ (*)

(*) Nota: Esta expresión es inválida, la correcta es la utilizada en ⑥, en un renglón anterior.

Expresión que tiene la forma análoga a la transformada de un exponencial, aunque con argumento matricial en lugar de escalar, siendo

$F(s) = \frac{a}{s-b}$

cuya solución resulta:

$f(t)=\mathfrak{L}^{-1}\left[\frac{a}{s-b}\right] = a\, e^{bt}$

Si se antitransforma esta expresión 6, por lo tanto, se obtendrá el error de reconstrucción buscado e(t):

$e(t) = \mathfrak{L}^{-1}\{E(s)\} = \mathfrak{L}^{-1}\{s[I] - [A-KC]\}^{-1} E_o$

por analogía:

$e(t) = E_o\, e^{[A-KC]t}$

donde todavía conviene separar los exponenciales para que se vea mejor la dependencia del error:

$$\boxed{e(t) = E_o\, e^{[A]t}\, e^{-[KC]t}} \qquad 7$$

Función que adoptamos como expresión final del error de reconstrucción. Como se puede observar, las matrices [A] y [C] corresponden al sistema original y son las que hemos supuesto inaccesibles, por lo tanto la regulación del error dependerá solamente de la matriz de ganancia del observador [K].

2) Resolución como ED de Primer Orden:

Todavía resta intentar la resolución de la ED del error 5 por el método tradicional de ecuaciones, sin utilizar la Transformada de Laplace y, por supuesto, obtendremos el mismo resultado.

El procedimiento es muy simple porque se trata de una ED Ordinaria Lineal de 1° Orden, luego:

$\dot{e}(t) = [A - KC]\, e(t)$

donde:

$\dot{e}(t) = \frac{de}{dt}$

entonces

$\frac{de}{dt} = [A-KC]\, e$

luego

$\left(\frac{1}{e}\right) \frac{de}{dt} = [A-KC]$

este factor matricial [A – KC] es como si fuera una constante para un sistema invariante en el tiempo, por lo tanto, integrando miembro a miembro resulta:

$\int (\frac{1}{e}) \frac{de}{dt} dt = \int [A\text{-}KC]\, dt = [A - KC] \int dt$

de donde:

$\ln e = [A\text{-}KC]\, t + C_1$

donde C_1 es una constante arbitraria de integración, entonces, despejando el error:

$e(t) = e^{[A\text{-}KC]t+C_1}$

y desagregando los exponenciales:

$e(t) = e^{[A]t}\, e^{-[KC]t}\, e^{C_1}$

donde el exponencial con la constante arbitraria de integración será, a su vez, una constante, que llamamos E_o igual que antes

$e^{C_1} = E_o$

finalmente:

$$\mathbf{e(t) = E_o\, e^{[A]t}\, e^{-[KC]t}}$$

Que es el mismo error de reconstrucción 7 que ya se había calculado antes mediante el procedimiento de la Transformada de Laplace.

Ganancia Óptima

En una primera mirada se aprecia en los factores exponenciales con argumento matricial de la fórmula del error que si la ganancia del observador K es grande, el error será pequeño. Tanto que, en el límite, si $K \to \infty$ entonces $e(t) \to 0$

Esto estaría sugiriendo que basta con aumentar la ganancia K del observador para que el error de reconstrucción disminuya, lo cual es matemáticamente cierto.

Sin embargo, aparece otro problema que debe ser tenido en cuenta. Cuando la matriz K aumenta, el sistema observador se hace más sensible al ruido.

Por lo tanto, el valor óptimo de la ganancia K deberá ser un compromiso entre la disminución del error de reconstrucción y la sensibilidad al ruido de observación.

A continuación se estudiará el método para encontrar una Matriz de Ganancia Óptima (K_O) que cumpla con las dos condiciones: producir el mínimo error de reconstrucción y minimizar el efecto perjudicial del ruido.

Para ello se debe realizar el estudio completo de las perturbaciones, tanto las propias del sistema original observado como las que introduce la observación.

Para contemplar estos efectos no deseados se deben modificar las ecuaciones diferenciales de estado del sistema, incluyendo las perturbaciones o ruidos.

$$\begin{cases} \dot{\mathrm{x}}(\boldsymbol{t}) = \boldsymbol{Ax} + \boldsymbol{Bu} + \boldsymbol{w}_1(\boldsymbol{t}) \\ \boldsymbol{y}(\boldsymbol{t}) = \boldsymbol{Cx} + \boldsymbol{w}_2(\boldsymbol{t}) \end{cases}$$

donde $w_1(t)$ y $w_2(t)$ son las componentes del vector de ruido $w(t)=\begin{Bmatrix} w_1 \\ w_2 \end{Bmatrix}$; que se denomina *ruido blanco*, w(t): White noise; porque la Transformada de Fourier de w(t) muestra una distribución espectral en la cual el ruido tiene la misma densidad de potencia para todas las frecuencias del espectro, por eso, en una analogía física con el proceso de la luz blanca y su distribución espectral se lo llama ruido blanco.

Cada una de las componentes del vector de ruido tiene su propio significado y su propio peso en las ecuaciones de estado, a saber:

w_1(t): ruido de excitación de estado. También llamado ruido térmico, es intrínseco.

w_2(t): ruido de observación o medición. Es externo y se “introduce” al observar.

En este punto es importante aclarar que en el estudio de errores por ruido de observación se trabaja con el Error Cuadrático Medio, MSE (Medium Squared Error); porque en cada observación el error viene con su signo, que puede ser negativo (error por defecto) o positivo (error por exceso).

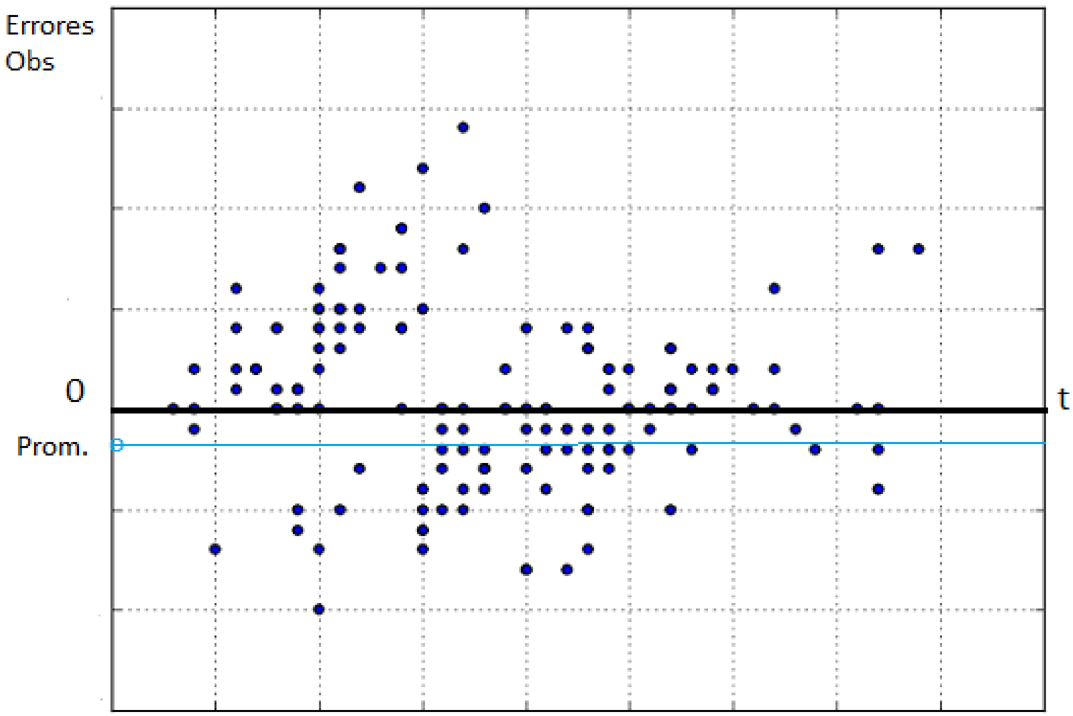
Si en los errores puros, con su signo, se toma el promedio, podría suceder que se compensen negativos con positivos, incluso podrían cancelarse y la media estaría arrojando un error muy bajo o nulo absurdamente y sin correspondencia con la magnitud promediada de errores cometidos.

Por eso, se elevan todos los valores al cuadrado, para independizarse del signo y luego se toma el promedio de éstos, que resultará una ponderación más veraz y confiable sobre la media real de errores existente a lo largo de todas las múltiples observaciones realizadas.

ERROR MEDIO CUADRÁTICO

En el siguiente cuadro se muestran observaciones aleatorias de errores por exceso y por defecto y, por lo tanto, con su signo.

Su promedio aritmético no reflejará la media de errores sino una compensación de los mismos entre sí.



En la práctica, el estudio de los fenómenos relacionados con los ruidos y perturbaciones que afectan a un sistema observado se realiza utilizando procesos estocásticos como modelo matemático de análisis.

El proceso estocástico se puede pensar como una familia de funciones del tiempo, cada una de las cuales se denomina realización. El interés radica en encontrar la media y la matriz de covarianza del proceso estocástico.

Desde este punto de vista estocástico el error de reconstrucción se define como la esperanza matemática del producto entre el error y el vector de ruido:

$E\{e(t)w(t)\}$

En definitiva, la optimización se trata de un proceso de prueba y error, calculando el error de reconstrucción $e(t)$ para un determinado valor de la matriz K y observando paralelamente el comportamiento del ruido $w(t)$.

Luego se repite el cálculo para otro juego de valores de K y w; hasta encontrar el valor óptimo de K, que produce el menor error de reconstrucción con la mínima influencia del ruido w(t).

Todo esto no se hace en forma manual sino que se utiliza la Ecuación matricial de Ricatti para el cálculo y se corre en forma algorítmica esta repetición recursiva de valores.

Finalmente se puede llegar hasta K_O que es la Matriz de Ganancia Óptima. La ecuación de Ricatti, utilizada en estadística, es la que permite obtener la derivada de la matriz de variancia, necesaria para el cálculo y el procedimiento algorítmico.

Siendo la derivada del error de reconstrucción del sistema con la ponderación del ruido incluido:

$$\dot{e}(t) = [A - KC]\, e(t) + [I - K] \begin{Bmatrix} w_1 \\ w_2 \end{Bmatrix}$$

si se designa a la media del error de reconstrucción como $\tilde{e}(t)$, resulta:

$$E\{e(t)\} = \tilde{e}(t)$$

de manera que

$E\{e(t) - \tilde{e}(t)\} = \breve{Q}(t)$ es la matriz de varianza del error

Finalmente, se puede verificar que la matriz de varianza del error $\breve{Q}(t)$ se minimiza para $K(\tau) = K_O(\tau)$ donde la matriz de ganancia óptima es:

$$\mathbf{K_O(\tau) = Q(\tau)\, C^T(\tau)\, [W_0(\tau)]^{-1}}$$

Esta ganancia óptima minimiza el error de reconstrucción cuadrático medio para todo ($t \geq t_o$) y si este resultado se implementa por medio de un filtro, este filtro será el mínimo estimador lineal cuadrático medio. Es decir que no se podrá encontrar otro funcional lineal de las observaciones $y(\tau)$ y de la entrada $u(\tau)$ que produzca una estimación de estado {x(t)} con un error de reconstrucción cuadrático medio más pequeño.

H. B. 2020