Capítulo 3

Desarrollo de los sistemas AR

En el presente capítulo, se describe un proceso estocástico discreto así como uno continuo. Se muestran algunos de los modelos estocásticos discretos utilizados para modelar estructuras dinámicas lineales, se comentan algunos de los tipos de convergencias, el modelo estocástico discreto, se plantea el problema de la caja negra y el por qué de la estimación de parámetros.

3.1. Proceso estocástico: continuo y discreto

El término *proceso estocástico* o *proceso aleatorio* es utilizado para describir la evolución temporal de un fenómeno estadístico de acuerdo con las leyes probabilísticas. La evolución temporal del fenómeno significa que el proceso estocástico es una función del tiempo, definida en algún intervalo de observación. La naturaleza estadística del fenómeno significa que, a priori, no se puede definir exactamente cómo evoluciona en el tiempo. Algunos ejemplos de procesos estocásticos son: las señales de voz, de televisión, señales de radar, datos digitales en una computadora, la salida de un canal de comunicaciones, datos sismológicos, y ruido, entre otras.

La forma del proceso estocástico que es de interés para este trabajo, es aquel que está definido como *discreto* y con *muestras de instantes de tiempo uniformemente espaciadas*. Alternativamente, el proceso estocástico puede ser definido para un rango continuo de valores reales de tiempo, pero antes del procesamiento, éste, se muestrea uniformemente en el tiempo, con la razón de muestreo mayor que el doble de la componente con la frecuencia más alta contenida en el proceso.

Un proceso estocástico *no* es sólo una función del tiempo; sino, representa, en teoría, un número infinito de *diferentes* realizaciones del proceso. Una realización particular de un proceso estocástico discreto en tiempo es llamada *serie de tiempo*.

En [11], un proceso estocástico escalar (vectorial) $\{x(t), t \in T\} \in (\Omega, \mathcal{F}, \mathrm{P})$ es una familia de variables (vectores) aleatorias indexadas por el conjunto de parámetros T. Donde el parámetro t es el tiempo. Si las variables (vectores) aleatorias $x(t)$ son discretas, se dice que el proceso estocástico tiene un *espacio de estados discreto*. Si son continuos, se dice que el proceso tiene un *espacio de estados continuo*. El conjunto de parámetros T también podrían ser discretos ($T = \{1, 2, \ldots, n\}, T = \{1, 2, \ldots\}$) o continuos ($T = [0, 1], T = \{t : t \geq 0\}$). Si el conjunto de parámetros es discreto, el proceso estocástico es un *proceso de parámetros discretos*; si éste es continuo, se dice que el proceso estocástico es un *proceso de parámetros continuo*. El diagrama correspondiente aparece en la Tabla 3.1.

Tabla 3.1: Clasificación de procesos estocásticos.

		Conjunto de parámetros: Discreto	Conjunto de parámetros: Continuo
Espacio de estados	Continuo	Secuencia aleatoria	Proceso (estocástico) Función aleatoria
	Discreto	Cadena de parámetros discretos	Cadena de parámetros continuos

3.2. Modelos estocásticos discretos

La clase más popular de modelos de series en tiempo lineales, como el mostrado en la Figura 3.1, consiste de modelos autorregresivos de promedios móviles (ARMA), incluyendo como casos especiales a los puramente autorregresivos (AR). Los modelos ARMA son frecuentemente utilizados para modelar las estructuras dinámicas lineales, para describir las relaciones entre las variables de retraso, y para servir como vehículos de predicción lineal. Una clase particular de modelos útiles, contiene los llamados modelos autorregresivos integrados de promedios móviles (ARIMA), los cuales incluyen los procesos ARMA del tipo estacionario como una subclase.

Los modelos estocásticos discretos que se describirán son: AR, MA, ARMA y ARIMA.

Figura 3.1: Modelo estocástico.

3.2.1. Modelos Autorregresivos (AR)

Se dice que la serie en el tiempo $\{\tilde{w}(k-i), i \in \mathbb{Z}_+\}$, representa la realización de un *proceso autorregresivo (AR) de orden M* si satisface la ecuación en diferencias (3.1), [4, 8].

$$\tilde{w}(k) + a_1^* \tilde{w}(k-1) + \cdots + a_M^* \tilde{w}(k-M) = v(k), \tag{3.1}$$

donde $\{a_i^*\}$ son constantes llamadas *parámetros* del modelo AR y $v\,(k)$ es el ruido blanco. Los términos $a_i^* \tilde{w}\,(k-i)$ son la versión escalar del *producto interno* de a_i y $\tilde{w}\,(k-i)$, donde $i = 1, \ldots, M$.

Para explicar la razón del término *autorregresivo* reescribase (3.1) en (3.2):

$$v\,(k) = \sum_{i=0}^{M} a_i^* \tilde{w}\,(k-i)\,, \text{con } a_0^* = 1, \tag{3.2}$$

donde $a_i^* := -a_i, \forall i \in \mathbb{Z}_+$. Así se ve que los valores presentes del proceso – esto es, $\tilde{w}(k)$ – son iguales a una *combinación finita de valores pasados* del proceso, $\tilde{w}(k-1), \ldots, \tilde{w}(k-M)$, más un *término de error* $v(k)$. Ahora veamos la razón del término *autorregresivo*. El modelo lineal de la forma (3.3), considera $a_0^* = 1$,

$$v\,(k) = \sum_{i=1}^{M} a_i^* \tilde{w}\,(k-i) + \tilde{w}\,(k)\,, \tag{3.3}$$

y relaciona la variable *dependiente* $v\,(k)$ con el conjunto de variables *independientes* $\tilde{w}(1), \tilde{w}(2), \ldots, \tilde{w}(M)$ más un término de error $\tilde{w}(k)$. Esto es frecuentemente referido como un *modelo de regresión*, y se dice que $y(k)$ es *regresada* sobre $\tilde{w}(1), \tilde{w}(2), \ldots, \tilde{w}(M)$. En la ecuación (3.2), la variable $u(k)$ es *regresada* sobre sus valores previos, de ahí el término *autorregresivo* y es ilustrado en las Figuras 3.2 y 3.3.

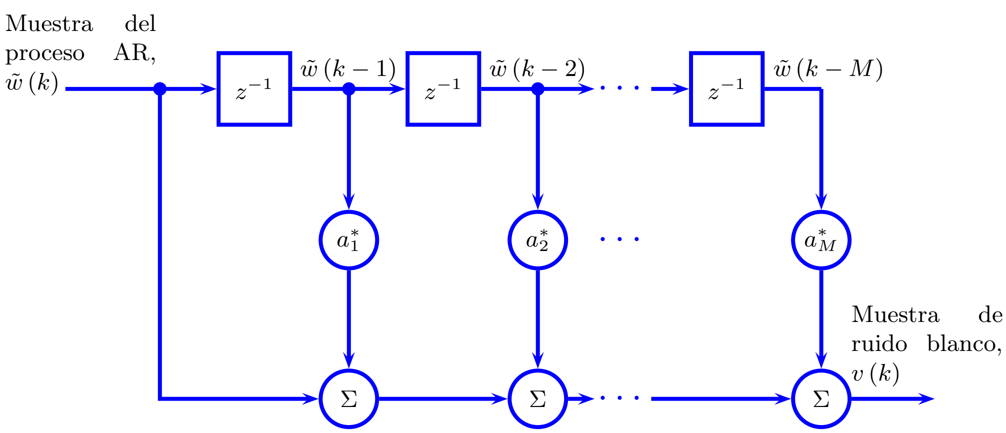

Figura 3.2: AR: Analizador del proceso.

Ahora la respuesta del modelo descrito en (3.3), es posible regresarla y contar con la excitación considerando la configuración escalera de la Figura 3.3.

Y es expresada en (3.4) como,

$$\tilde{w}\,(k) = v\,(k) - \sum_{i=1}^{M} a_i^* v\,(k-i) \tag{3.4}$$

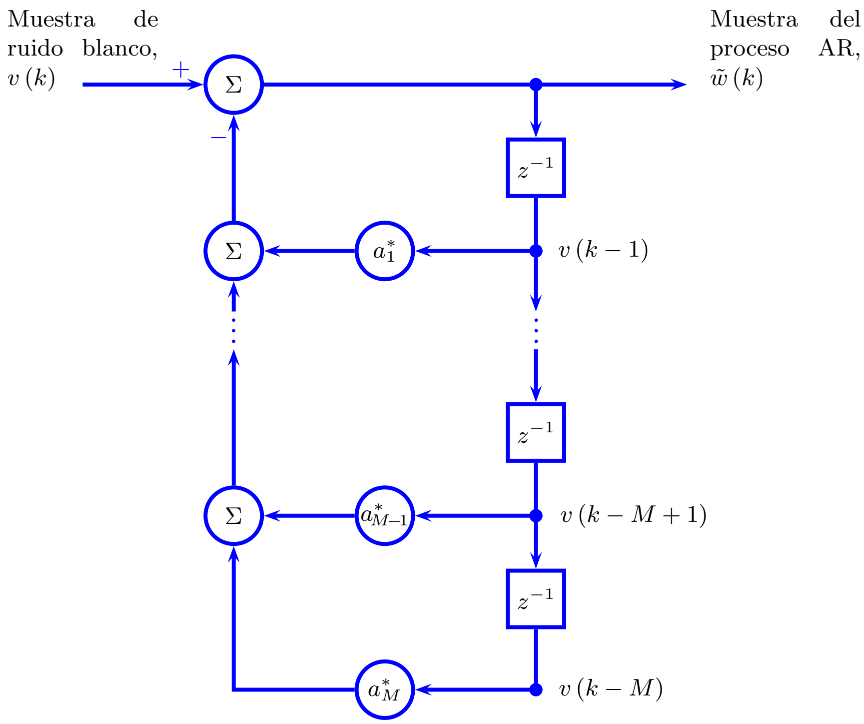

Figura 3.3: AR: Generador del proceso.

3.2.2. Modelos de Promedios Móviles (MA)

En un *modelo de promedios móviles (MA)*, el filtro lineal discreto (Figura 3.4) consiste de un *filtro todos-ceros* manejado por ruido blanco. El proceso resultante $y(k)$, producido a la salida del filtro, es descrito por la ecuación en diferencias (3.5), [4, 8].

$$y(k) = v(k) + b_1^* v(k-1) + \cdots + b_K^* v(k-K) \tag{3.5}$$

donde $b_1, \ldots, b_K$ son constantes llamadas *parámetros* del modelo MA y $v(k)$ es ruido blanco con media cero y varianza $\sigma_v^2 < \infty$. Excepto por $v(k)$, cada término del lado derecho de (3.5) representa la versión escalar de un producto interno. El *orden* del proceso MA es igual a K, como puede verse en la Figura 3.4.

$$y(k) = \sum_{i=0}^{K} b_i^* v(k-i) + v(k). \tag{3.6}$$

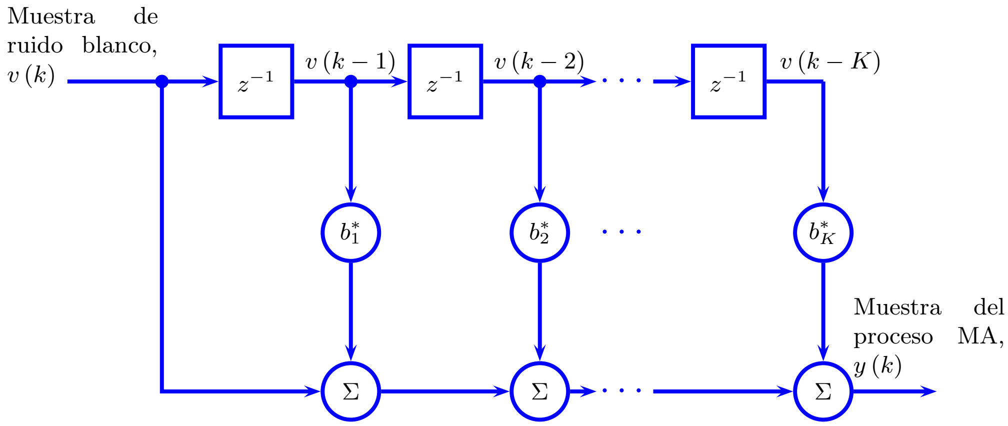

Figura 3.4: MA: Generador del proceso.

3.2.3. Modelos Autorregresivos de Promedios Móviles (ARMA)

Para generar un *proceso autorregresivo de promedios móviles (ARMA)*, se considera a la función de transferencia que relaciona $y(k)$ con $\tilde{w}(k)$ o sea la interrelación entre la Figura 3.2 y la Figura 3.4. La función de transferencia consiste de *polos* y *ceros*. Dado el ruido blanco $\tilde{w}(k)$ como la entrada del filtro, el proceso ARMA $y(k)$ producido a la salida del filtro, es descrito por la ecuación en diferencias (3.7), [4, 8].

$$y(k) + a_1^* y(k-1) + \cdots + a_M^* y(k-M) = \tilde{w}(k) + b_1^* \tilde{w}(k-1) + \cdots + b_K^* \tilde{w}(k-K) \tag{3.7}$$

donde $a_1, \ldots, a_M$ y $b_1, \ldots, b_K$ son llamados los *parámetros* del modelo ARMA. Excepto por $y(k)$ en el lado izquierdo de (3.7) y $\tilde{w}(k)$ en el lado derecho, todos los demás términos representan el orden del sistema considerado. El *orden* del proceso ARMA es igual a (M, K), como puede verse en la Figura 3.5.

3.2.4. Modelos Autorregresivos Integrados de Promedios Móviles (ARIMA)

Una subclase útil de modelos ARMA consiste de los llamados modelos *estacionarios*. La estacionariedad refleja ciertamente invariancia de acuerdo a algún parámetro del proceso, propiedades de las series y es en algún modo una condición necesaria para hacer una inferencia estadística. Sin embargo, las series en tiempo real frecuentemente exhiben tendencias en el tiempo (tal como lentos incrementos) y características estocásticas que están más allá de la capacidad de los modelos estacionarios ARMA. Una práctica común es preprocesar la información para remover esas componentes *inestables*. Tomando la diferencia (más de una vez si es necesario) entre estados adyacentes es conveniente y efectivo para minimizar la tendencia y destemporizar. Después de remover las tendencias temporales, se puede modelar la nueva serie por un modelo estacionario ARMA. Debido a que la serie original es la integración de series diferenciadas, la llamamos, proceso *autorregresivo integrado de promedios móviles* (ARIMA), [4].

Una serie en tiempo $\{Y_t\}$ es llamada un *proceso autorregresivo integrado de promedios móviles* (ARIMA) con orden p, d, y q, denotado como $\{Y_t\} \sim ARIMA(p, d, q)$, si su diferencia $X_t = (1-B)^d Y_t$ de orden d es un proceso estacionario $ARMA(p, q)$, donde $d \geq 1$ es un entero, esto es,

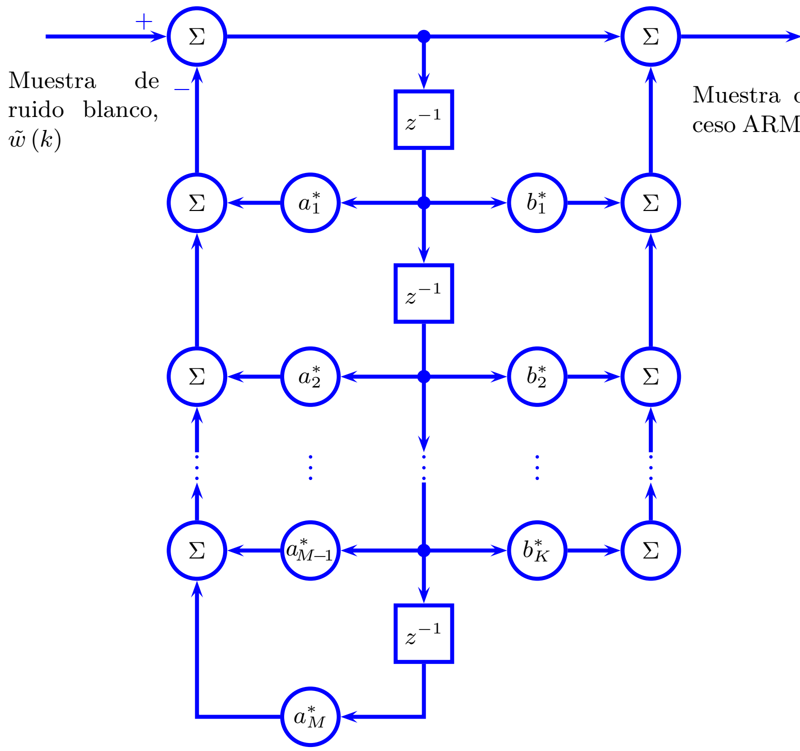

Figura 3.5: ARMA: Generador del proceso.

$b(B)(1-B)^d Y_t = a(B)\,\epsilon_t$.

Es fácil ver que un modelo $ARIMA(p,d,q)$, es un modelo especial de los modelos $ARMA(p+d,q)$ que es típicamente no estacionario dado que $b(B)(1-B)^d$ es un polinomio de orden $p+d$.

3.3. Tipos de convergencias

La siguiente definición introduce los seis principales conceptos de convergencia en la teoría de probabilidad, [15, 16].

Definición 1 *La secuencia de variables aleatorias* $\{\xi_n\}_{n\geq 1}$*, definida en el espacio de probabilidad* $(\Omega,\mathcal{F},\mathrm{P})$*, converge a una variable aleatoria* $\xi \in (\Omega,\mathcal{F},\mathrm{P})$*):*

1. ***En distribución****, esto es denotado por* (3.8)*,*

$$\xi_n \xrightarrow[n\to\infty]{d} \xi \tag{3.8}$$

si y sólo si, se cumple (3.9)*,*

$$\mathrm{E}\{f(\xi_n)\} \xrightarrow[n\to\infty]{d} \mathrm{E}\{f(\xi)\} \tag{3.9}$$

para cualquier función continua acotada $f(x)$*, o equivalentemente, si y sólo si se tiene la convergencia de las funciones de distribución, de acuerdo con* (3.10)*,*

$$F_{\xi_n}(x) \xrightarrow[n\to\infty]{} F_{\xi}(x) \tag{3.10}$$

en cualquier punto x *donde* $F_{\xi}(x)$ *es continua.*

2. ***En probabilidad****, esto es denotado por* (3.11)*,*

$$\xi_n \xrightarrow[n\to\infty]{\mathrm{P}} \xi \tag{3.11}$$

si y sólo si para cualquier $\epsilon > 0$*, se cumple* (3.12)*,*

$$\mathrm{P}\{\omega : |\xi_n - \xi| > \epsilon\} \xrightarrow[n\to\infty]{} 0. \tag{3.12}$$

3. ***Con probabilidad uno*** *o* ***casi seguro*** *(*almost surely (a.s.)*), esto es denotado por* (3.13)*,*

$$\xi_n \xrightarrow[n\to\infty]{a.s.} \xi \tag{3.13}$$

si y sólo si, se cumple (3.14)*,*

$$\mathrm{P}\{\omega : \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\} = 1, \tag{3.14}$$

i.e., el conjunto de puntos de muestra $\omega \in \Omega$ *para los cuales* $\xi_n(\omega)$ *hace que no converja a* $\xi(\omega)$ *tiene probabilidad cero.*

4. ***En promedio de orden*** p*, esto es denotado por* (3.15)*,*

$$\xi_n \xrightarrow[n\to\infty]{L^p} \xi \tag{3.15}$$

si y sólo si para algún $p > 0$*, se cumple* (3.16)*,*

$$\mathrm{E}\{|\xi_n - \xi|^p\} \xrightarrow[n\to\infty]{} 0 \tag{3.16}$$

para $p = 2$ *se tiene* ***la convergencia de la media al cuadrado****, que es denotada por* (3.17)*,*

$$\lim_{n\to\infty} \xi_n = \xi. \tag{3.17}$$

5. ***Puntualmente***, *esto es denotado por* (3.18),

$$\xi_n \xrightarrow[n\to\infty]{\forall\omega\in\Omega} \xi \tag{3.18}$$

si y sólo si, se cumple (3.19),

$$\text{lím}_{n\to\infty}\, \xi_n(\omega) = \xi(\omega) \;\text{ para todo } \omega \in \Omega. \tag{3.19}$$

6. ***Completamente***, *esto es denotado por* (3.20),

$$\xi_n \xrightarrow[n\to\infty]{c.c.} \xi \tag{3.20}$$

si y sólo si, para cualquier $\epsilon > 0$, *se cumple* (3.21),

$$\sum_{n=1}^{\infty} P\left\{|\xi_n - \xi| > \epsilon\right\} < \infty \tag{3.21}$$

o equivalentemente, por el Lema de Borel-Cantelli, si y sólo si $|\xi_n - \xi| \leq \epsilon$ *(a.s.) después de un número (quizás aleatorio)* $n_0(\omega)$ *finito de pasos.*

3.4. Modelo estocástico discreto

El término *modelo* es utilizado para cualquier hipótesis que pueda ser aplicada para explicar o describir las leyes desconocidas que se supone gobiernan o restringen la generación de información física de interés. La representación de un proceso estocástico por un modelo data desde la idea propuesta por [Yule, 1927]. La idea es que una serie en tiempo $u(k)$ consistente de observaciones altamente correlacionadas pueda ser generada aplicando una serie de *pulsos* estadísticamente independientes a un filtro lineal.

Los pulsos son variables aleatorias seleccionados de una distribución fija que es usualmente supuesta como Gaussiana con media cero y varianza acotada. Tales series de variables aleatorias constituyen un proceso puramente aleatorio, comúnmente referido como *ruido Gaussiano*. Específicamente, se puede describir la entrada $\tilde{w}(k)$ en la Figura 3.1, en términos estadísticos con (3.22) y (3.23).

$$\mathrm{E}\left\{\tilde{w}(k)\right\} = 0 \text{ para todo } k, \tag{3.22}$$

y

$$\mathrm{E}\left\{\tilde{w}(k)\,\tilde{w}^*(n)\right\} = \begin{cases} \sigma_{\tilde{w}}^2 < \infty, & n = k, \\ 0, & \text{otro caso.} \end{cases} \tag{3.23}$$

donde $\sigma_{\tilde{w}}^2$, es la varianza del ruido. La ecuación (3.22) resulta de la suposición de la media cero, y la ecuación (3.23) resulta de la suposición de ruido blanco.

En general, la descripción en el dominio del tiempo de la relación entrada-salida para el modelo estocástico de la Figura 3.5 puede ser descrito como sigue:

(valor presente de la salida del modelo) + (combinación lineal de los valores pasados de la salida del modelo) = (combinación lineal de los valores presentes y pasados de la entrada del modelo.)

Un proceso estocástico descrito así, es referenciado como *proceso estocástico lineal.*

La estructura del modelo estocástico utilizado para el desarrollo de este trabajo tiene la forma descrita en (3.24),

$$\begin{aligned} x(k+1) &= ax(k) + bw(k) \\ y(k) &= cx(k) + dv(k) \end{aligned} \tag{3.24}$$

donde a, b, c, d son parámetros constantes del modelo contenidos en $\mathbb{Z}_+$; $w(k)$ y $v(k)$ son señales estocásticas de ruido acotadas por su respectiva función de distribución; $x(k)$ es la variable de estado; $y(k)$ es la variable de estado de salida, acotadas por su respectiva función de distribución.

3.5. El problema de la caja negra

Al no conocer la dinámica existente que se genera en el sistema de referencia descrito por medio de bloques, en donde no se conoce dentro de él al número de estados, cuántas ganancias tiene, la relación entre ganancias y estados, etc., es llamado caja negra.

La caja negra como lo indica su nombre, consiste de un bloque al cual llega una entrada y que de la cual sale una respuesta, como se ve en la Figura 3.6.

Figura 3.6: Sistema tipo caja negra.

Del bloque como caja negra no es posible de forma directa saber qué es lo que ocurre con el comportamiento del sistema de referencia, tan sólo saber que la relación entre la entrada y la salida se encuentra acotada y que entre ellas no se tiene una relación mayor a la unidad.

Al ser propuesto cualquiera de los anteriores modelos se observa una relación de ganancias: Al aplicar una señal de excitación a la entrada, ésta, es procesada en el interior de la caja negra, y finalmente después de un grupo de operaciones en muchos casos desconocidas, se obtiene una señal de respuesta conocida como la señal observable.

¿Cuál es el resultado de haber descrito al sistema de referencia por medio de cualquiera de los modelos AR, MA, ARMA y ARIMA? ¿Qué orden tienen? En su forma más simple cualquiera de los modelos propuestos cuenta con un retardo y no permite saber qué valor de parámetro o parámetros requiere el sistema para cada evolución.

3.6. El por qué de la estimación de parámetros

Considerando los modelos estocásticos descritos como AR, MA, ARMA y ARIMA, cuentan con relaciones de retardos respecto de la señal de excitación, y que es en su forma simple tan sólo multiplicada por un grupo de ganancias, con respecto al producto punto se vería como un vector o matriz de ganancias y un vector formado por la señal de entrada retardada de acuerdo al orden del sistema (si es de cuarto orden corresponde a cuatro retardos). En cualquiera de los casos, la estimación es un elemento primordial para establecer de manera dinámica la relación entre la entrada y la salida del sistema de referencia en cada estado de evolución.

Con la estimación se logrará encontrar dada una excitación al sistema de referencia, un conjunto de valores que de acuerdo al modelo seleccionado permitirá aproximarse con un nivel de convergencia a la respuesta deseada.

3.7. Ganancia proporcional o por modos deslizantes

El estimador, se convierte en un selector de ganancias para lograr establecer la mejor relación con el sistema de referencia. La ganancia es del tipo proporcional, en la que se busca que la respuesta del filtro en su conjunto se aproxime a una señal de referencia, que es la del sistema considerado. Claro que a diferencia de los sistemas que utilizan una ganancia proporcional seleccionada de forma heurística, en el caso presente y de acuerdo a los modelos AR, MA, ARMA, ARIMA, así como a las condiciones de estabilidad, sus valores propios requieren estar acotados dentro de $[0, 1)$. En relación a la velocidad de cambio de la ganancia, estará en función de la relación entre la covarianza y la varianza de la señal de salida de referencia y su retardo; lo que nos permite considerar una condición acotada de evolución de ganancia. Como elemento que relaciona la entrada con la salida del sistema, queda acotada y se busca que sea tendiente a 1 para el mejor de los casos; es decir, en el que la transferencia energética entre la entrada y la salida fue sin pérdidas.

El estimador en su proceso de llevar la respuesta del modelo a la señal de referencia requiere de un proceso de adaptación que permita medir el error de convergencia y usarlo para ajustar su dinámica. La forma de lograr llevarlo a una convergencia global a una condición deseada se logra por medio de modos deslizantes. La cual es una función que de manera brusca permite el cambio de signo de una ganancia específica (el error de identificación) para atenuar la respuesta del estimador, garantizando la disminución del error a lo largo del tiempo. De su forma más simple queda representada por una función tangente con cambio de estado por cada intervalo de evolución.