



# UNIVERSIDAD DE GUANAJUATO

---

---

Facultad de Ingeniería Mecánica Eléctrica y Electrónica

*“Efectos de longitud finita de palabra en filtros digitales adaptivos”*

TESIS PROFESIONAL

QUE PARA OBTENER EL TÍTULO DE:  
INGENIERO EN COMUNICACIONES Y ELECTRÓNICA

PRESENTA:  
ANDRÉS FRÍAS VELÁZQUEZ

ASESOR:  
DR. RENÉ DE J. ROMERO TRONCOSO

SALAMANCA, GTO.

SEPTIEMBRE DE 2005

“La ciencia no puede resolver el misterio último que guarda la naturaleza debido a que nosotros mismos formamos parte de ese misterio”

MAX PLANCK

# *Agradecimientos*

*A Dios, por concederme su amor y su gracia*

*A mis padres, por su cariño, esfuerzo y enseñanzas de vida*

*A mi mamá María y papá Chucho, por ser mis ángeles de la guarda*

*A mis hermanos Irene, Avelina y Juan, por su eterna amistad y cariño*

*A mis abuelos Petra y Antonio, por su respaldo y bendiciones*

*A mis tíos Mario e Isabel, por su apoyo y ejemplo*

*A mis tíos Martha e Israel, por ayudarme a superarme*

*A Tere, porque me enseñaste la humildad y la caridad*

*A Chiquis, porque me enseñaste la rectitud y me brindas tu cariño*

*A mis tíos Chano y Luzma, por su apoyo y afecto*

*A mi asesor René de Jesús, por sus sabios consejos y forjarme  
íntegramente como estudiante*

*A los doctores Oscar y Amparo, por su amistad y apoyo incondicional*

*A mis amigos, Alejandro, Gustavo, Sergio, Alberto R., Anayansi, Mary Helena, Chucho, Ordaz, Vite, Cabal, Diana, Ale U., Aurora, Maribel, Chabe, Carolina, Almendra, Agustín, Mata, Luis, Alberto, Carlos Santiago, etc.,etc., así como a todos mis amigos de CEM-UFEC, mi más profundo respeto y agradecimiento.*

---

# CONTENIDO

Contenido . . . . .	i
Lista de Figuras . . . . .	iii
Lista de Tablas . . . . .	iv
Nomenclatura . . . . .	v
<b>1 Introducción al Filtrado Adaptivo</b>	<b>1</b>
<i>1.1 Introducción.</i> . . . . .	1
1.1.1 Objetivos . . . . .	2
1.1.2 Justificación . . . . .	3
<i>1.2 Introducción al Filtrado Adaptivo</i> . . . . .	5
1.2.1 Algoritmos Adaptivos . . . . .	7
1.2.2 Aplicaciones . . . . .	10
<i>1.3 Fundamentos de Filtrado Adaptivo</i> . . . . .	15
1.3.1 Clasificación de Señales . . . . .	15
1.3.2 Procesos Estocásticos . . . . .	16
1.3.3 Parámetros Estadísticos . . . . .	19
1.3.4 Ergodicidad . . . . .	21
1.3.5 Densidad de Potencia Espectral . . . . .	21
1.3.6 Matriz de Correlación . . . . .	22
<i>1.4 Filtro Wiener</i> . . . . .	23
<i>1.5 Filtro LMS</i> . . . . .	26
<i>1.6 Efectos de Palabra Finita</i> . . . . .	28
1.6.1 Modelo de Ruido de Cuantización . . . . .	30

---

<b>2</b>	<b>Modelo Matemático</b>	<b>34</b>
2.1	<i>Filtro LMS implementado con aritmética de punto fijo</i>	34
2.2	<i>Modelo de comportamiento transitorio y estado estable</i>	36
2.3	<i>Análisis de Estabilidad</i>	46
<b>3</b>	<b>Simulaciones</b>	<b>56</b>
3.1	<i>Procedimiento de Simulación</i>	56
3.2	<i>Simulaciones</i>	57
<b>4</b>	<b>Conclusiones</b>	<b>62</b>
4.1	<i>Conclusiones</i>	62
4.2	<i>Perspectivas de investigación futura</i>	63
4.3	<i>Aprendizaje</i>	63
	<b>Referencias</b>	<b>64</b>

---

# LISTA DE FIGURAS

Fig. 1.1	Configuración general de un filtro adaptivo. . . . .	6
Fig. 1.2	(a) Estructura transversal IIR de 4to. orden . . . . .	6
Fig. 1.2	(b) Estructura canónica directa IIR de 3er. orden . . . . .	7
Fig. 1.3	Configuración de modelado . . . . .	10
Fig. 1.4	Configuración de modelado inverso. . . . .	11
Fig. 1.5	Configuración de predicción lineal . . . . .	12
Fig. 1.6	Configuración de cancelación de interferencia . . . . .	14
Fig. 1.7	Modelo probabilístico de ruido blanco. . . . .	16
Fig. 1.8	Realización del proceso $N(m)$ . . . . .	17
Fig. 1.9	Estructura de filtro LMS . . . . .	27
Fig. 1.10	Formato numérico de punto fijo. . . . .	28
Fig. 1.11	Modelo de proceso de cuantización . . . . .	28
Fig. 1.12	Proceso de truncamiento . . . . .	29
Fig. 1.13	Modelo estadístico de cuantización . . . . .	29
Fig. 1.14	Distribución de error para truncamiento complemento a dos . . . . .	30
Fig. 1.15	Distribución de error para redondeo . . . . .	31
Fig. 1.16	Distribución resultante de la diferencia de dos dist. uniformes . . . . .	32
Fig. 2.1	Algoritmo LMS de precisión finita para configuración de identificación de sistemas. . . . .	33
Fig. 2.2	(a) Gráfica de sensibilidad $S_{\sigma_x^2}^E$ . . . . .	48
Fig. 2.2	(b) Comparación de (2.87) con varias aproximaciones de (2.89) . . . . .	48
Fig. 2.3	Variación del RMSE al aumentar $\nu$ . . . . .	50
Fig. 3.1	C.A. comparativa utilizando (a) redondeo (b) truncamiento . . . . .	56
Fig. 3.2	C.A. comparativa con $b_d = 7$ y $b_c = 11$ para (a) redondeo (b) truncamiento . . . . .	58
Fig. 3.3	C.A. comparativa cuando $\gamma_2 = -6/5$ para (a) redondeo (b) truncamiento . . . . .	59

---

---

# LISTA DE TABLAS

Tabla 2.1	Primer y segundo momento de distribuciones cuantizadas . . . . .	34
Tabla 2.2	Límites lineales con $\nu$ grados de libertad . . . . .	50
Tabla 2.3	Rangos de operación propuestos por varios autores . . . . .	53
Tabla 3.1	Comparación de error de estado estable para ambos casos de cuantización . . . . .	57
Tabla 3.2	Exceso de kurtosis para varias distribuciones . . . . .	59

---

# NOMENCLATURA

$b_c$	Número de bits de cuantizador de coeficientes
$b_d$	Número de bits de cuantizador de datos
$d(k)$	Señal deseada
$diag[\mathbf{x}]$	Matriz diagonal
$E$	Exceso de error de estado estable
$e(k)$	Señal error
$e_o(k)$	Señal de error óptimo
$E[\cdot]$	Función valor esperado
$\gamma_2$	Exceso de kurtosis
$\eta_d(k)$	Señal de ruido de cuantización de la señal deseada
$\eta_e(k)$	Señal de ruido de cuantización de la señal error
$\boldsymbol{\eta}_u(k)$	Ruido de cuantización del término de actualización
$\eta_y(k)$	Señal de ruido de cuantización de la señal de salida
$\mathbf{I}$	Matriz identidad
$\mathbf{K}(k)$	Matriz de correlación del vector de error ponderado
$\boldsymbol{\Lambda}$	Matriz de eigenvalores
$\lambda$	Eigenvalor
$\mu$	Factor de convergencia
$\boldsymbol{\mu}_u$	Media estadística del vector de ruido de actualización
$\boldsymbol{\mu}_v(k)$	Valor esperado del vector de error ponderado
$N$	Orden de filtro
$n(k)$	Ruido de medición

---

---

$n_\nu$	Límite lineal con $\nu$ grados de libertad
$\nu$	Grados de libertad
$\mathbf{p}$	Vector de correlación cruzada
$\mathbf{Q}$	Matriz de eigenvectores
$\mathbf{R}$	Matriz de correlación
$\sigma_e^2$	Varianza del ruido de la señal error
$\sigma_n^2$	Varianza del ruido de medición
$\sigma_x^2$	Varianza de la señal de entrada
$S_{\sigma_x^2}^E$	Sensitividad
$\text{tr}[\cdot]$	Función traza
$u(k)$	Función escalón
$\mathbf{v}(k)$	Vector de error ponderado
$\mathbf{w}(k)$	Vector ponderado
$\mathbf{w}_0$	Vector de coeficientes
$x(k)$	Señal de entrada
$\xi_{\min}$	Error mínimo cuadrático
$\xi(k)$	Error medio cuadrático
$y(k)$	Señal de salida
$\mathcal{Z}\{\cdot\}$	Transformada z
$\mathcal{Z}^{-1}\{\cdot\}$	Transformada z inversa

---

# Capítulo 1

## INTRODUCCIÓN AL FILTRADO ADAPTIVO

*"El hombre encuentra a Dios detrás de cada puerta que la ciencia logra abrir"*

**-Albert Einstein-**

---

### *1.1 Introducción*

En un mundo en el que la tecnología cada vez más forma parte del estilo de vida de una persona y crea una sociedad más sofisticada, es poco probable no haber escuchado o leído acerca de las señales que están presentes en los dispositivos que utilizamos diariamente, como las señales de música en un reproductor de disco compacto, voz en un teléfono celular, video en una televisión, por mencionar algunos ejemplos. La mayoría de estas señales son generadas por medios naturales donde las variables que determinan dichas señales suelen ser tiempo, distancia, posición, temperatura, presión etc., y que generalmente son transformadas a señales eléctricas para su procesamiento. Sin embargo, algunas de estas señales pueden también ser sintetizadas por otros medios como la simulación por computadora.

Una señal inherentemente contiene información acerca de la entidad o fenómeno que generó dicha señal, por lo que el procesamiento de señales tiene como objetivo la extracción de esta información. Como resultado de lo anterior, podemos definir en forma general que el procesamiento de señales se ocupa de la interpretación matemática de las señales y del desarrollo de algoritmos para la extracción de la información contenida en ellas.

De acuerdo a la forma de implementarse, el procesamiento de señales puede ser analógico ó digital, el primero está basado fundamentalmente en resistores, capacitores y

amplificadores operacionales, los cuales manipulan la señal en el dominio natural en la que ésta es generada (analógico); mientras que en el procesamiento digital las señales requieren ser muestreadas para transformarse a tiempo discreto, y mediante convertidores A/D obtener la señal digital, posteriormente se procesa la información mediante microprocesadores, los cuales enviarán los resultados hacia otro convertidor encargado de que la señal resultante sea transformada a tiempo continuo nuevamente.

Debido principalmente a que el procesamiento analógico es muy susceptible al ruido, temperatura, y robusto para aplicaciones serias, es por lo que el procesamiento digital ha ganado terreno. En los últimos 30 años, el procesamiento digital de señales ha sido uno de los campos que más ha prosperado a la par con el desarrollo de circuitos digitales, cuestión que ha permitido la implementación de algoritmos complejos en un sinnúmero de sistemas aplicables que abarcan desde electrodomésticos hasta instrumentos aeroespaciales. La sinergia resultante de estos dos campos ha dado pie a que se produzcan sistemas más confiables, precisos, flexibles y de tamaño reducido.

Actualmente existe una gran variedad de dispositivos mediante los cuales se puede efectuar el procesamiento digital de señales, el de uso más común es el llamado DSP (*Digital Signal Processor*), el cual es un tipo de microprocesador que tiene la capacidad de realizar operaciones compuestas como la multiplicación y suma en un solo ciclo de máquina además de poder simularse y desempeñarse en tiempo real con un bajo consumo de potencia. Por otro lado, existen otros sistemas que realizan funciones similares a un DSP, pero que varían en la flexibilidad, confiabilidad, costo, consumo de potencia y desempeño. Un ejemplo de ello son los FPGAs (*Field Programmable Gate Arrays*) que tienen como principal arma la capacidad de ser reconfigurables y ofrecer un mejor desempeño en funciones específicas. Sin embargo, los FPGAs son significativamente más caros y típicamente disipan mayor potencia que los DSPs. Otra opción a considerar son los ASICs (*Application-Specific ICs*) los cuales su desempeño es superior en comparación a los casos anteriores, además tienen una alta eficiencia de potencia, su principal desventaja es que su flexibilidad es nula pues no son programables, y el tiempo de salida al mercado del chip es mayor. En contraste con los ASICs, los microprocesadores de propósito general no son utilizados en aplicaciones serias pues no tienen la capacidad de procesar en tiempo real, además de que el consumo de potencia es mucho mayor; a pesar de estas desventajas pueden utilizarse en aplicaciones simples.

Una de las principales operaciones que se realiza en el procesamiento de señales es el filtrado, el cual tiene como objetivo manipular las señales para obtener la información contenida en éstas; tradicionalmente, los filtros son utilizados para dejar pasar ciertas frecuencias y suprimir lo más posible aquellas que sean indeseables.

Un filtro digital es aquél que puede procesar señales en tiempo discreto representadas en un formato digital. Los filtros digitales más comunes son los que sus parámetros y estructura del filtro son fijos, esto significa que son invariantes con el tiempo. Cuando los parámetros fijos de un filtro no son conocidos o las especificaciones de un filtro no son satisfechas por los filtros invariantes en el tiempo se utiliza el llamado filtrado adaptivo. Los filtros adaptivos son variantes con el tiempo porque sus parámetros están cambiando continuamente hasta obtener el desempeño deseado.

El hecho de que los parámetros de un filtro estén representados en un formato digital, implica que se generen errores con respecto al diseño teórico de un filtro, los cuales pueden afectar desde una forma mínima en el desempeño, cambiar drásticamente la respuesta en frecuencia del filtro diseñado, o hasta convertirlo en un oscilador. De esta manera, es indispensable tomar en cuenta estos errores en la implementación de cualquier filtro, llamados también como efectos de palabra finita.

### 1.1.1 Objetivos

Un modelo matemático para el análisis de los efectos de palabra finita de la curva de aprendizaje del algoritmo LMS (*Least Mean Squares*) es presentado en este trabajo. Tomar en consideración los efectos de palabra finita en el diseño de un filtro adaptivo LMS es un aspecto indudablemente importante al momento de la implementación, ya que dichos efectos pueden repercutir seriamente en las variables que miden el desempeño transitorio y de estado estable del filtro, tal como la rapidez de convergencia y el desajuste respectivamente. Existen diversos trabajos relacionados a este tema que han permitido analizar los errores de precisión finita de acuerdo a distintos criterios como el modelo de cuantización del algoritmo, entorno estacionario o no estacionario, análisis transitorio o de estado estable, tipo de entrada, tipo de cuantización etc.; sin embargo, al momento de diseñar es indispensable tener un panorama completo del sistema bajo la unificación de condiciones y criterios de todo el análisis. El modelo matemático propuesto en este trabajo permite analizar la curva de aprendizaje (C.A.) con los efectos de precisión finita asumiendo que la señal de entrada es estacionaria en el sentido amplio, real, media cero, y los elementos del vector de entrada son mutuamente independientes, así mismo el análisis contempla la cuantización por redondeo y truncamiento complemento a dos. Simulaciones realizadas en Matlab confirman la validez del modelo matemático propuesto, y sus ecuaciones permiten realizar un diseño confiable del filtro considerando los efectos de palabra finita.

## 1.1.2 Justificación

Actualmente las necesidades comerciales demandan aplicaciones tecnológicas cada vez más rápidas y precisas, llevando al límite la capacidad de los sistemas; un ejemplo de ello son los sistemas de posicionamiento global, en los cuales el procesamiento en tiempo real y la alta precisión son vitales para un buen desempeño del sistema. En el caso del algoritmo LMS, éste es implementado generalmente con aritmética de punto fijo, cuestión que genera inherentemente efectos que ponen en detrimento la rapidez de convergencia y estabilidad del filtro, así mismo el exceso de error medio cuadrático puede resultar mucho mayor que el estimado por el análisis de precisión infinita.

Existen diversos trabajos relacionados a este tema que han permitido analizar los efectos de palabra finita de forma satisfactoria en relación a los resultados prácticos obtenidos; sin embargo, la complejidad del análisis obliga a limitar el problema a ciertas condiciones tales como, tipo de señal de entrada, cuantización, aplicación, entre otras. El sinnúmero de variables involucradas hace difícil formular un modelo genérico para el análisis de los efectos finitos de palabra del algoritmo LMS; a pesar de esto, un intento de lo anterior acotado para el análisis de estado estable fue propuesto por Yousef y Sayed [1]. Por otro lado, en la mayoría de los análisis propuestos se considera un buen referente hacer uso del típico problema de identificación de sistemas como base para integrar un modelo estadístico de los errores de cuantización. Andrews y Fitch [2] se enfocan en analizar los efectos de convergencia generados por la cuantización sufrida en la ecuación de actualización del filtro para el caso de redondeo y truncamiento complemento a dos. Diniz [3] presenta un modelo estadístico práctico para el cálculo del error de estado estable para entradas gaussianas estacionarias; mientras que Seara et al. [4] proponen un modelo estadístico mejorado que enfatiza en el problema referente al fenómeno conocido como “stopping”, así mismo Gupta y Hero III [5] analizan el comportamiento transitorio del mismo fenómeno, pero bajo otra concepción referida como “frenado”(slowdown) .

En este trabajo se analiza el comportamiento transitorio y de estado estable de forma conjunta incorporando los efectos de cuantización por truncamiento y redondeo, se asume que la señal de entrada es estacionaria en el sentido amplio, real, media cero, y los elementos del vector de entrada son mutuamente independientes. Una de las principales aportaciones de este trabajo es que, no solo se analizan los efectos cuando la entrada es gaussiana sino también cuando la distribución no es mesokurtica, esto significa que el exceso kurtosis de la señal de entrada es diferente de cero, de esta manera se puede predecir el comportamiento del algoritmo cuando no se cuenta con un filtro para el preprocesamiento de la señal de entrada. Otro resultado importante es el análisis de estabilidad del filtro el cual mediante un análisis de sensibilidad del exceso de MSE permite un mejor entendimiento de las restricciones del factor de convergencia.

## 1.2 Introducción al Filtrado Adaptivo

En la actualidad el filtrado adaptivo se ha convertido en una herramienta muy útil en aplicaciones, principalmente en áreas como control, instrumentación, procesamiento de imágenes y telecomunicaciones. La característica que fundamenta el filtrado adaptivo es la capacidad de autoajustar su desempeño, esto significa que si se aplica a la entrada del filtro una señal que está variando constantemente, el sistema se adaptará en relación a los cambios que presenta la misma. Como introducción a este tema se había mencionado que los filtros adaptivos eran variantes en el tiempo, y que a diferencia de los filtros fijos, los coeficientes de un filtro adaptivo están variando o autodiseñándose constantemente para ajustarse al desempeño deseado, al inferir lo anterior es obvio que el análisis de este tipo de filtros es más complejo que los fijos.

La configuración básica de un filtro adaptivo es mostrada en la Fig. 1.1, en donde  $k$  es el número de iteración,  $x(k)$  la señal de entrada,  $y(k)$  la señal de salida del filtro adaptivo,  $d(k)$  define la señal deseada y la señal error  $e(k)$  es calculada como  $d(k) - y(k)$ . La función error es la que determina el desempeño del filtro adaptivo, y que sirve como base en el algoritmo adaptivo para actualizar el filtro hasta que cumpla un criterio de minimización. La especificación completa de un sistema adaptivo, está formada por tres partes [3]:

- 1) **Estructura del Filtro Adaptivo:** El filtro adaptivo puede ser implementado en un número diferente de estructuras o realizaciones. La elección de un tipo de estructura puede influir en la complejidad computacional (cantidad de operaciones aritméticas por iteración) del proceso y del número de iteraciones necesarias para lograr el nivel deseado de desempeño. Básicamente, hay dos principales clases de realizaciones para filtros digitales adaptivos, filtros de respuesta finita al impulso (FIR, *Finite Impulse Response*) y filtros de respuesta infinita al impulso (IIR, *Infinite Impulse Response*). La realización FIR utilizada más comúnmente, es la estructura de filtro transversal (Fig. 1.2(a)), la cual es una realización directa de la función de transferencia, que tiene como principal característica el que no tiene lazo de retroalimentación. En el caso de las realizaciones de los filtros IIR, la estructura más común es la forma canónica directa (Fig. 1.2(b)), debido a su simple implementación y análisis. Sin embargo, existen problemas inherentes a los filtros recursivos que tienen origen en su estructura, como lo pueden ser la inestabilidad de los polos del filtro y la rapidez de convergencia. Estos problemas han sido resueltos mediante otras estructuras que intentan aminorar o eliminar las limitaciones de la realización de forma directa, tal como la estructura en cascada, de red, y la paralela.

- 2) **Algoritmo:** El algoritmo es el procedimiento usado para ajustar los coeficientes del filtro mediante un criterio establecido. El algoritmo es definido en función del método de búsqueda utilizado para lograr la minimización de la señal error.
  
- 3) **Aplicación:** El tipo de aplicación es definido por la clase de señales que son adquiridas, y por la configuración típica del sistema adaptable. Algunos ejemplos de estas configuraciones son la cancelación de eco, igualación de canal dispersivo, identificación de sistemas, cancelación de ruido, y control.

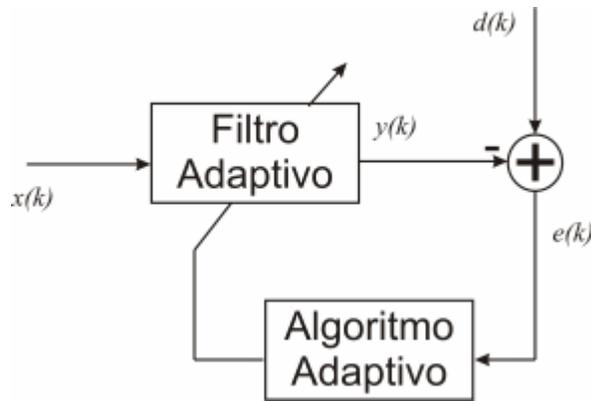
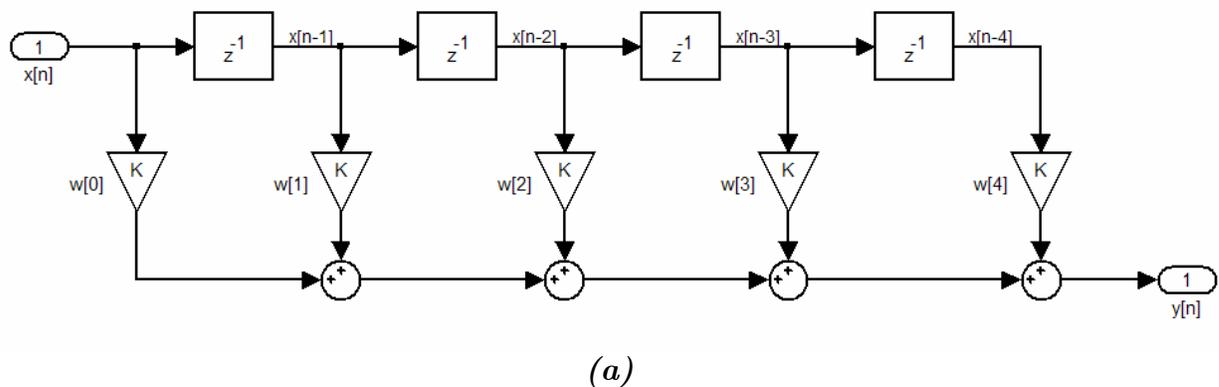


Fig. 1.1 Configuración General de un Filtro Adaptivo.



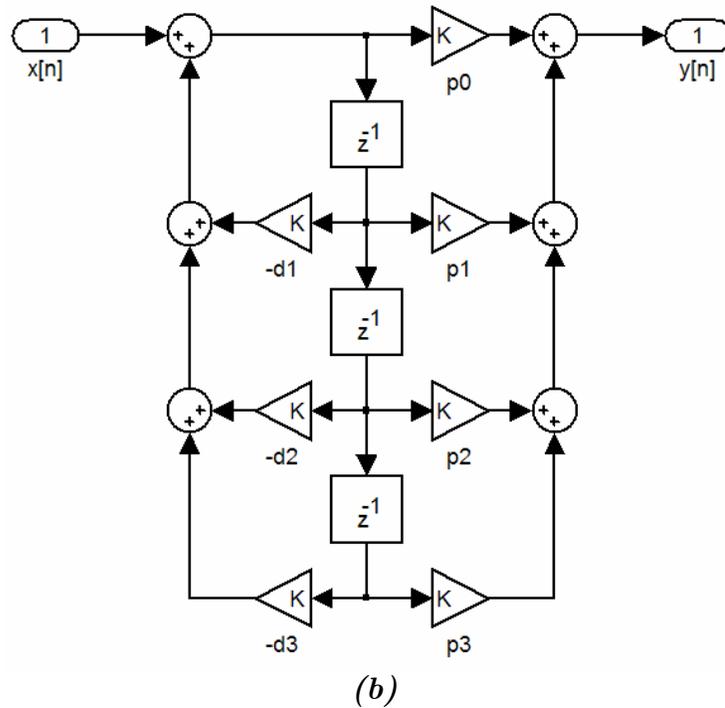


Fig. 1.2 (a) Estructura transversal FIR de 4to. orden.  
(b) Estructura canónica directa IIR de 3er. orden.

## 1.2.1 Algoritmos Adaptivos

La palabra *adaptación* tiende a ser un término muy común en la naturaleza, así como en el quehacer diario del hombre, pues constantemente estamos sometidos a variables que hacen ajustarnos a las necesidades de nuestro entorno; la forma en realizar ese ajuste es lo que podríamos llamar un algoritmo adaptivo.

Para desarrollar un algoritmo adaptivo es necesario generar una función objetivo, la cual nos indique la variable de interés que desea ser minimizada, así como las variables que intervienen en el proceso; en forma general la función objetivo se denota de la siguiente manera  $F = F[e(k)] = F[x(k), y(k), d(k)]$ . Dicha función tiene que cumplir con ciertas propiedades [3]:

- 1) No negatividad:  $F[e(k)] \geq 0$
- 2) Optimación:  $F[e(k)] = 0$

Existen diversas funciones que satisfacen dichas propiedades, sin embargo, utilizar funciones muy complejas inevitablemente hace que el análisis sea más difícil así como su implementación; debido a lo anterior es que generalmente se utilizan funciones elevadas al

cuadrado en conjunto con algún promedio estadístico. Las funciones que se utilizan con mayor frecuencia se describen a continuación [3]:

- Error Medio Cuadrático (MSE, *Mean Square Error*):  $F[e(k)] = E[|e(k)|^2]$ ;
- Mínimos Cuadrados (LS, *Least Squares*):  $\frac{1}{k+1} \sum_{i=0}^k |e(k-i)|^2$ ;
- Mínimos Cuadrados Ponderados (WLS, *Weighted Least Squares*):  
 $F[e(k)] = \sum_{i=0}^k \lambda^i |e(k-i)|^2$ , donde  $\lambda$  es una constante menor a 1.
- Valor Cuadrático Instantáneo (ISV, *Instantaneous Squared Values*):  $F[e(k)] = |e(k)|^2$ .

Además de las diferencias en cuanto a complejidad, cada una de las funciones anteriores difiere en convergencia, susceptibilidad al ruido, rapidez de cálculo etc., que hace que uno u otro sea más adecuado para cierta aplicación.

La variable de mayor interés dentro de un proceso adaptivo es la señal de error, esto se hace evidente ya que está presente en todas las funciones descritas arriba. Como se mencionó anteriormente, el error se genera de la diferencia de una señal de referencia y la salida del filtro adaptivo.

Hasta ahora se han definido en forma general las propiedades de un algoritmo adaptivo, así como los criterios de minimización más comunes; sin embargo, una tarea importante en el filtrado adaptivo es llevar al sistema a un estado óptimo, a dicho estado se le conoce como solución óptima de Wiener, la cual es derivada de las ecuaciones *Wiener-Hopf* que se detallarán más adelante. Para obtener la solución óptima de Wiener se requiere la inversión de una matriz, cuestión que hace poco probable su implementación práctica mediante esta concepción; sin embargo, existen métodos de optimización que recursivamente minimizan la función objetivo y que pueden implementarse sin ningún problema, también llamados comúnmente como métodos de búsqueda.

Los métodos de optimización más comunes son los siguientes:

- Método steepest-descent: En este método se asume que la función objetivo es convexa, el algoritmo inicia en un punto arbitrario y realiza pequeños cambios en la dirección en que la función objetivo disminuye más rápidamente, de tal manera que en cada iteración se acerca más al punto mínimo de la superficie convexa, en ese punto se habrá minimizado la función objetivo y obtenido la solución óptima de Wiener. A continuación se muestra la ecuación de actualización de filtro utilizando este método:

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu \nabla_k [e(k)] \tag{1.1}$$

donde  $\mu$  es el factor que controla la convergencia del algoritmo y  $\nabla$  es el operador gradiente definido por el siguiente vector:

$$\nabla = \left[ \frac{\partial}{\partial w_0} \quad \frac{\partial}{\partial w_1} \quad \dots \quad \frac{\partial}{\partial w_{N-1}} \right]^T \quad (1.2)$$

- Método de Newton: Este método es similar al anterior solo que en lugar que el gradiente apunte hacia la dirección en que se minimiza la función objetivo, el gradiente es rotado para que apunte hacia la dirección a la que se encuentra la solución óptima. La ecuación de actualización este método es la siguiente:

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu \mathbf{R}^{-1} \nabla_k [e(k)] \quad (1.3)$$

La ecuación es muy similar a la del método steepest-descent, solo que el gradiente está multiplicado por la inversa de una matriz llamada *matriz de correlación* la cual será explicada a detalle más adelante. El producto  $\mathbf{R}^{-1} \nabla_k [e(k)]$  hace que el gradiente sea rotado hacia la dirección de la solución óptima, y por consecuencia la convergencia hacia ese punto es más rápida que para el caso anterior.

- Métodos *Quasi-Newton*: Estos algoritmos son una versión simplificada del método Newton, ya que intentan estimar por diversos medios la inversa de la matriz de correlación. El compromiso de estos algoritmos es de que sean prácticos en la implementación, así como rápidos para converger; sin embargo, existen problemas de inestabilidad asociados al estimar la matriz de correlación, debido a que dichas estimaciones son obtenidas por medios iterativos.

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu \mathbf{S}(k) \nabla_k [e(k)] \quad (1.4)$$

donde  $\mathbf{S}(k) = \hat{\mathbf{R}}^{-1}$ , de tal manera que  $\lim_{k \rightarrow \infty} \mathbf{S}(k) = \mathbf{R}^{-1}$ .

Una forma más didáctica de entender todo lo que se ha explicado en esta sección sería cuando tratamos de mantener una escoba en equilibrio sostenida por la palma de una mano, nuestra función objetivo se basaría en el criterio de mantener la escoba en una posición vertical, la función error se produciría en relación al ángulo generado con la vertical, mientras quien sostiene la escoba trata de que la ésta llegue al punto de equilibrio (solución de Wiener) mediante pequeños desplazamientos de la palma.

## 1.2.2 Aplicaciones

En diversos textos relacionados sobre filtrado adaptivo se exponen un sinnúmero de aplicaciones donde puede ser utilizado este tipo de filtrado, sin embargo, cada aplicación puede ser clasificada dentro de cuatro clases básicas como es sugerido en [6]: *Modelado*, *Modelado Inverso*, *Predicción Lineal* y *Cancelación de Interferencia*.

### MODELADO

El *modelado* es uno de los casos típicos en el que se utiliza el filtrado adaptivo, debido a que es una forma básica en la que se puede comprender el análisis y las implicaciones de este tipo de filtrado. La configuración de este sistema se muestra en la Fig. 1.3. Como el nombre lo sugiere, el objetivo de las aplicaciones que están dentro de esta clasificación, es estimar un modelo de la función de transferencia de un sistema el cual se tiene muy poca información de forma a priori. De acuerdo a la figura,  $H(z)$  representa la función de transferencia del sistema desconocido,  $W(z)$  representa el filtro que constantemente se esta actualizando hasta que alcanza un valor óptimo. La diferencia de las señales filtradas genera la señal error, la cual es utilizada para actualizar  $W(z)$ .

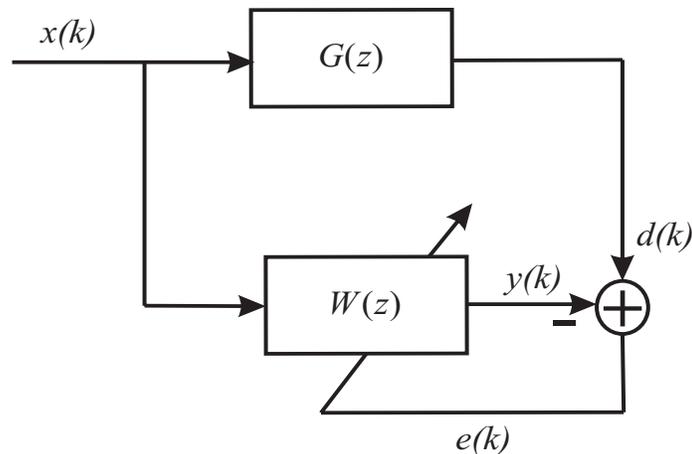


Fig. 1.3 Configuración de Modelado.

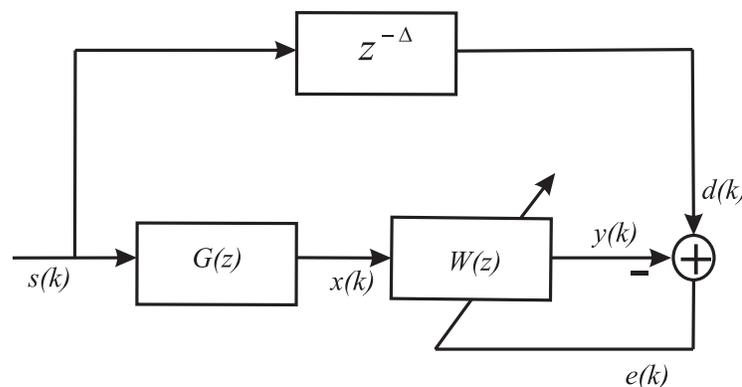
Una de las principales aplicaciones en que se utiliza este modelo, es el llamado identificación de sistemas, que hace exactamente lo que se describió arriba. Esta aplicación es muy utilizada dentro del área de control, pues a veces los sistemas que se desean controlar están compuestos por muchas variables determinísticas y aleatorias, que hacen casi una tarea imposible modelar matemáticamente dicho sistema. Gracias al filtrado adaptivo se puede estimar dicha caja negra y se puede diseñar un esquema de control de lazo cerrado.

Uno de los temas de investigación aún vigentes dentro del campo de comunicaciones es la igualación de canal, el objeto de esta investigación se centra en eliminar los efectos producidos por un canal (cable, fibra óptica, enlace inalámbrico) al transmitir datos por este medio. Un canal puede ser modelado en forma práctica como una caja negra, debido a que las condiciones del canal pueden intervenir un sinnúmero de factores que no solo dependen del material que sirve como medio físico, sino también de las condiciones externas a las que está sometido. Al poder estimarse la función de transferencia del canal mediante el filtrado adaptivo, se puede obtener la inversa de dicha función (igualación de canal) y aplicarse a los datos recibidos. Cabe aclarar que estrictamente hablando, a esta aplicación se le denomina identificación de canal, pues en realidad lo que se logra estimar es la función de transferencia del canal y no su inversa, ya que la inversa se obtiene en forma independiente del filtro adaptivo.

Otra aplicación importante dentro de las comunicaciones es la cancelación de eco. Dicho fenómeno es producido básicamente por un desacoplamiento de impedancias en la línea telefónica, esto produce que la calidad del enlace sea pobre y sea difícil entablar comunicaciones a larga distancia. El filtrado adaptivo es utilizado para estimar la señal de eco, la cual posteriormente es utilizada para eliminarla de la señal recibida.

### MODELADO INVERSO

El funcionamiento de esta clase de filtrado adaptivo es muy similar al de modelado, solo que las señales de entrada del filtro son conceptualizadas de manera diferente, esto con el objetivo de que las salidas del filtro puedan ser utilizadas activamente dentro de un proceso. En comparación con el caso anterior donde el principal objetivo era obtener la función de transferencia de una planta desconocida, en este caso se busca encontrar la inversa de dicha función. En la Fig. 1.4 se muestra el modelo básico de esta clase de filtrado, en el cual se puede ver que la señal de entrada del filtro  $W(z)$  es precisamente la salida generada por la planta desconocida  $G(z)$ , mientras que la señal deseada es una versión retrasada de la señal de entrada de la planta desconocida.



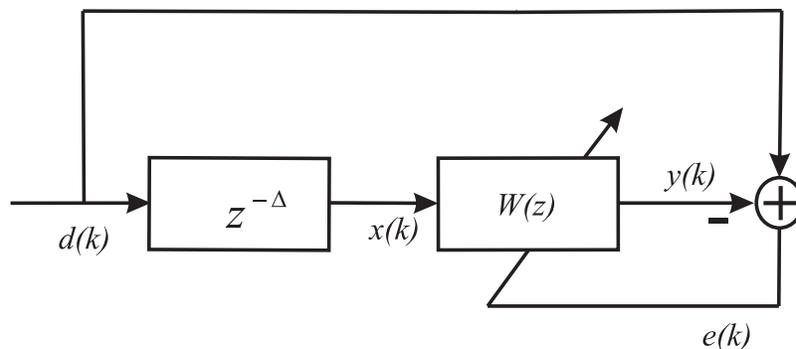
**Fig. 1.4 Configuración de Modelado Inverso.**

Este tipo de modelo es referido frecuentemente como igualador de canal, pues en comparación con el sistema de identificación de canal descrito anteriormente, este filtro obtiene directamente la inversa del canal, esto se debe a que en la señal deseada se encuentra una secuencia de símbolos de entrenamiento los cuales igualmente son enviados por el emisor a través del canal, la señal recibida por el receptor se introduce al filtro adaptivo, y su salida es comparada con los símbolos de entrenamiento, de tal forma que el error generado es utilizado para actualizar los coeficientes del filtro hasta que el error sea mínimo, al llegar a este punto, el lazo de la señal de error se abre terminando así el proceso de entrenamiento y se empiezan a enviar datos los cuales son recuperados a la salida del filtro adaptivo y procesados con algún criterio de decisión. Este tipo de filtrado es utilizado comúnmente en módems y sistemas inalámbricos de comunicación.

Otra aplicación de este modelo es utilizado en el área de control, pues es común encontrar que un planta tenga ciertas características no ideales, las cuales pueden afectar en el tiempo de respuesta y estabilidad del sistema de control. Este problema puede ser resuelto haciendo uso de este tipo de filtrado para autodiseñar un compensador que aminore dichas características no ideales.

## PREDICCIÓN LINEAL

Para realizar cualquier tipo de predicción ya sea climática, económica, tecnológica, social etc., es necesario tener información previa del fenómeno que lo origina, de tal manera que evaluando dicha información podamos predecir con cierto grado de confiabilidad un evento. Este mismo criterio se utiliza en diversas aplicaciones que utilizan el filtrado adaptivo como modelo de predicción lineal. La estructura del filtro se muestra en la Fig. 1.5. A diferencia de los casos anteriores, aquí la señal deseada es tomada directamente de la entrada del proceso sin ser manipulada, mientras que la señal que entra al filtro  $W(z)$  son valores anteriores de la misma entrada del proceso; dichos valores anteriores representan la información previa y mediante el error generado se puede sintonizar el filtro para que actúe como predictor.



**Fig. 1.5 Configuración de Predicción Lineal.**

Una aplicación muy importante dentro del procesamiento y transmisión de voz e imagen, es la codificación de predicción lineal, la cual permite codificar dichas señales a su representación digital con el menor número de bits posibles, de esta manera se puede transmitir de forma más rápida la información esencial de una señal de voz e imagen. Para recuperar la información transmitida, se envían también los coeficientes del filtro adaptivo, los cuales servirán para que el receptor pueda decodificar la información.

Una de las características importantes que son necesarias para realizar la predicción lineal, es que los valores generados en una señal estén correlacionados con ellos mismos. Esta propiedad permite que cuando una señal está compuesta por dos señales y una de ellas sus valores presentes y pasados están correlacionados y en la otra no, se pueda obtener la separación de cada una de ellas mediante un esquema de predicción lineal; a este tipo de aplicación se le conoce como incrementador de línea adaptivo. Se le conoce así porque la mayoría de las señales que son de espectro estrecho sus valores están correlacionados, y por lo tanto las señales pueden ser separadas por el predictor lineal, mientras que las señales de espectro amplio precisamente debido a su dispersión en frecuencia es poco probable que sus valores estén correlacionados, y en consecuencia son suprimidos por el predictor. Un ejemplo claro de esta aplicación es cuando se desea eliminar el ruido de 60Hz en un electrocardiograma, donde la señal de 60Hz es la señal de espectro estrecho y la del electrocardiograma la de espectro amplio, de tal manera que la señal error contiene precisamente la señal del electrocardiograma sin el ruido del 60Hz.

El esquema de predicción lineal es también utilizado ampliamente para estimar la densidad de potencia espectral, la cual refleja el contenido espectral de un proceso estocástico en función de la frecuencia, la estimación se realiza utilizando el modelo de autoregresión el cual puede verse más a detalle en [6].

## CANCELACIÓN DE INTERFERENCIA

En la mayoría de los sistemas electrónicos se producen distintos tipos de ruidos así como interferencias, algunos de estos ruidos resultan difíciles poder eliminarse por medio de técnicas de filtrado fijo, es por eso que se recurre al filtrado adaptivo. El principio que se utiliza para realizar la cancelación de interferencia se basa en que la señal de interferencia (obtenida de forma independiente de la señal deseada) es utilizada como referencia, la cual sirve de entrada para el filtro adaptivo, posteriormente este último intentará estimar la interferencia contenida en la señal deseada. La señal deseada se le conoce también como entrada primaria, la cual esta compuesta por la señal útil y la interferencia, de tal suerte que cuando se resta la señal deseada de la interferencia estimada por el filtro adaptivo, se obtendrá la señal útil sin la interferencia. El modelo de este esquema de filtrado se muestra en la Fig. 1.6.

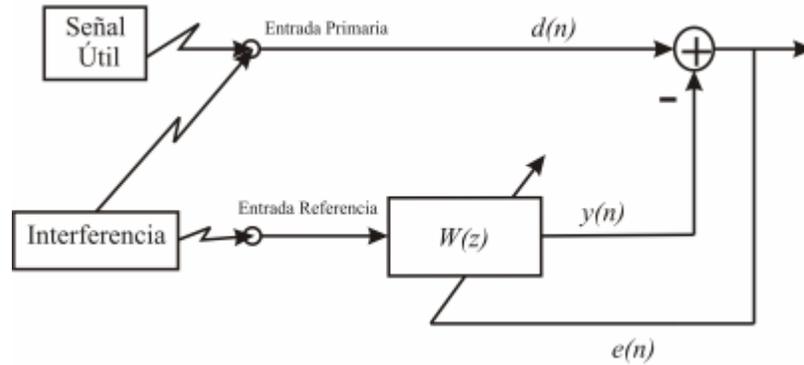


Fig. 1.6 Configuración de Cancelación de Interferencia.

Una forma práctica de ver este modelo es cuando en la entrada primaria tenemos un micrófono en el cual evidentemente la voz se mezcla con ruido debido al dispositivo, al ambiente y el eco generado por la bocina. La señal de referencia se utiliza para captar todas estas interferencias y estimar por medio del filtro adaptivo la interferencia contenida en la señal deseada, posteriormente cuando el filtro converge a una solución óptima, la diferencia de ambas señales interfiere destructivamente del tal manera que solo sobrevive la señal útil.

Una aplicación muy específica de este modelo de filtrado, es control activo de ruido, el cual es utilizado principalmente en sistemas de ventilación y ductos. El sistema está compuesto por un micrófono el cual capta la señal de ruido, la cual es utilizada como señal de referencia, también se instala una bocina cerca de la fuente de ruido la cual está conectada a la salida del filtro adaptivo y obviamente de un DAC (Digital-Analog Converter); delante de donde fue instalada la bocina se coloca otro micrófono el cual hace la función de receptor de la señal error del sistema. El sistema funciona de la siguiente manera, primero se capta la señal de ruido mediante el micrófono de referencia, después el filtro adaptivo manda una señal a la bocina tratando de que el ruido sea interferido destructivamente y posteriormente el resultado es registrado por el micrófono de error, el filtro tratará de ajustarse hasta que el ruido haya sido cancelado completamente.

Durante esta sección se pudo conocer que existe una gran variedad de aplicaciones del filtrado adaptivo dentro de los diversos campos de la ciencia, algunos de ellos no han sido incluidos dentro de esta tesis debido a que el objetivo es presentar la clasificación general de los modelos existentes, así como crear un panorama breve acerca de estas aplicaciones, si el lector requiere más información acerca de este tema puede consultar en [2,3].

## 1.3 Fundamentos de Filtrado Adaptivo

Como se ha podido constatar en temas tratados anteriormente, el filtrado adaptivo ha permitido resolver diversos problemas que con el filtrado tradicional hubiese resultado poco práctico o hasta imposible de resolver; el motivo por el cual el filtrado adaptivo ha ganado terreno ante el filtrado fijo, es debido a que la mayoría de las señales con las que el hombre está en contacto como voz, ruido, etc., son de tipo aleatorio o pseudoaleatorio; de tal manera que el filtrado adaptivo tiene la capacidad de tratar dichas señales de mejor forma que el filtrado fijo.

En las siguientes secciones se estudiará con mayor detalle la naturaleza de las señales utilizadas en el filtrado adaptivo, el análisis del filtro Wiener, así como el filtro LMS.

### 1.3.1 Clasificación de Señales

Dentro del campo del procesamiento de señales existen diversas clasificaciones para las señales, todo depende de la característica o características que está siendo analizadas en cierto grupo de señales. De acuerdo a las características más fundamentales de una señal, las señales pueden dividirse en determinísticas o aleatorias, estas últimas conocidas también como señales estocásticas. Dentro de cada clasificación puede ocurrir que las señales sean continuas, discretas, cuantizadas o digitales. Debido a que nuestro principal interés es el procesamiento digital y a que su teoría es desarrollada para señales discretas por razones de generalidad, de aquí en adelante se manejarán definiciones y conceptos para señales de tiempo discreto a menos que se especifique lo contrario.

#### SEÑALES DETERMINÍSTICAS

Una señal determinística en tiempo discreto es aquella que se puede calcular en forma exacta los valores de una señal en cualquier parte del espacio discreto. Un ejemplo de lo anterior puede ser la siguiente señal:

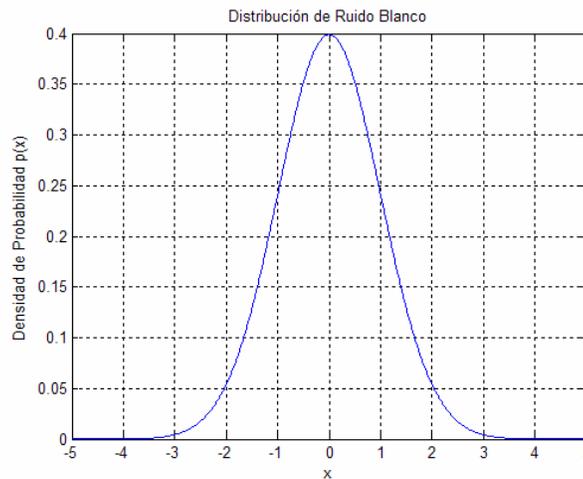
$$x(k) = e^{-\alpha k} \cos(\omega k) \quad (1.5)$$

donde  $k = 0, \pm 1, \pm 2, \pm 3, \dots$  representa el espacio discreto. Como se muestra en la ecuación (1.5), el comportamiento de la señal puede predecirse de manera analítica, así como su respuesta en frecuencia.

## SEÑALES ALEATORIAS

En comparación a las señales determinísticas, las señales aleatorias no pueden ser modeladas por una ecuación, ya que no se puede predecir cuales serán los valores próximos generados por la fuente de dicha señal. Algunas de estas señales en tiempo discreto pueden obtenerse al muestrear ruido o voz, las cuales se consideran señales mayormente aleatorias. El concepto de aleatoriedad está cercanamente asociado a los términos de información y ruido, debido precisamente a que en los sistemas de comunicación la mayoría de las señales que se manejan son aleatorias o pseudoaleatorias. Aunque no es posible obtener una función matemática que pueda describir las señales aleatorias con la exactitud de una señal determinística, generalmente la mayoría de estas señales exhiben características bien definidas que permiten ser descritas de acuerdo a dichas características tales como media, mediana, varianza, densidad de potencia espectral etc., así como de un modelo probabilístico.

A continuación se muestra en la Fig. 1.7 el modelo probabilístico de ruido blanco.



**Fig. 1.7** *Modelo probabilístico de ruido blanco.*

### 1.3.2 Procesos Estocásticos

El término de proceso estocástico es usado comúnmente para describir un proceso aleatorio. En un sentido estricto, un proceso estocástico representa el modelo probabilístico de una clase de señales aleatorias, tal como los procesos Gaussianos, procesos de Markov, procesos de Poisson etc. Los procesos aleatorios en tiempo discreto pueden ocurrir en forma natural, obtenerse mediante el muestreo de procesos aleatorios de tiempo continuo o limitados en banda. El término de “*proceso estocástico en tiempo discreto*” hace referencia a la clase de señales aleatorias en tiempo discreto  $X(m)$ , caracterizado por un modelo

probabilístico [7]. Cada realización del proceso estocástico discreto  $X(m)$ , puede ser indexada en tiempo y en espacio como  $x(m,s)$ , donde  $m$  es el índice de tiempo discreto, y  $s$  la variable entera que designa el índice de realización del proceso.

## ENSAMBLADO DE UN PROCESO ESTOCÁSTICO

El conjunto de realizaciones de un proceso aleatorio se le conoce como *ensamblado* o *espacio* del proceso [7]. Es claro que al obtener muestras independientes de ruido blanco de un canal de comunicación, las señales obtenidas serán a primera vista completamente diferentes unas de otras, sin embargo al obtener el histograma de cada una de las muestras se observará que todas tienden a parecerse al modelo probabilístico gaussiano, de esta manera, cada una de las muestras representan una realización específica del proceso, formando así parte del *ensamblado* o *espacio* de un proceso. Cada una de las realizaciones del proceso se denota indicando el índice de tiempo discreto y el correspondiente al espacio respectivo  $n(m,s)$ , y el proceso estaría indicado como  $N(m)$ . La estadística “verdadera” de un proceso estocástico se obtiene al ensamblar los promedios estadísticos de todas las realizaciones existentes en el proceso. Sin embargo, en muchos casos prácticos solo se puede obtener una realización del proceso, en estos casos se recurre a la propiedad de ergodicidad de un proceso, en el cual se considera que el promedio en tiempo en una realización del proceso es igual al promedio de ensamblado. Este concepto será revisado a detalle más adelante.

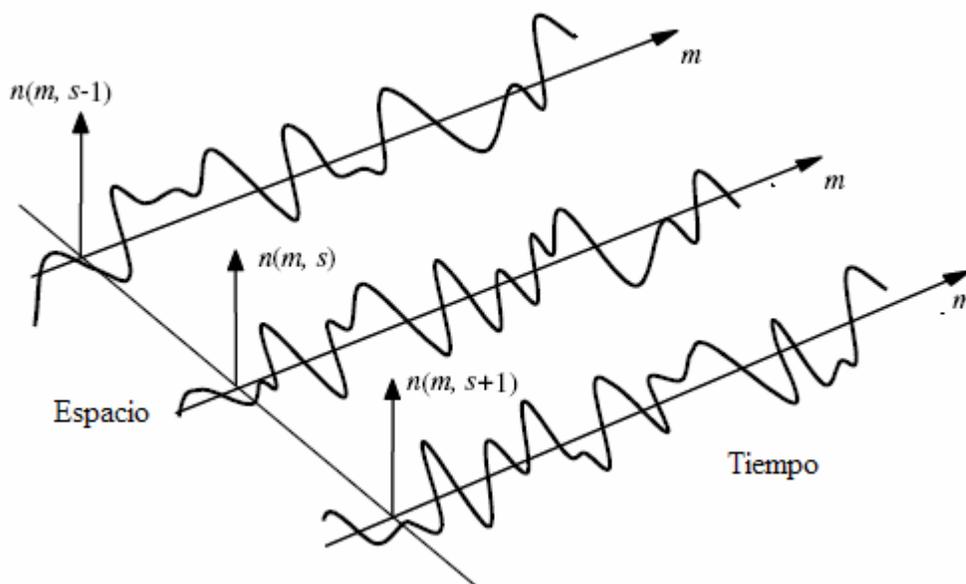


Fig. 1.8 Realizaciones del proceso  $N(m)$ .

## MODELOS PROBABILÍSTICOS

En forma general podemos definir que los modelos probabilísticos proveen la descripción matemática exacta de un proceso aleatorio. Un modelo probabilístico está compuesto esencialmente por variables aleatorias, las cuales se definen como el conjunto de muestras de un instante de tiempo fijo  $m$  obtenidas de diversas realizaciones en el espacio  $s$  de un proceso  $X(m) = \{x(m, s)\}$ . Los modelos probabilísticos pueden ser aplicados no solo a variables aleatorias, sino también a procesos estocásticos. Una variable aleatoria está compuesta de todos los elementos probables que pueden ocurrir, este espacio de elementos puede ser dividido en subespacios de acuerdo a cierto criterio, a cada uno de los subespacios se les conoce como eventos. La probabilidad de que un evento ocurra está definido como la razón del número de elementos que se observaron en cierto evento entre el número total de observaciones. Obteniendo las probabilidades de cada uno de los eventos se puede obtener el modelo probabilístico de la variable aleatoria formando así la función de probabilidad ponderada, en la cual se puede determinar la probabilidad de que ocurra cierto evento de una variable aleatoria. La notación de lo anterior se muestra de la siguiente manera  $P(X = x_i)$ .

## PROCESOS ALEATORIOS ESTACIONARIOS Y NO ESTACIONARIOS

Cuando en un proceso aleatorio, la señal que se genera de éste es invariante en el tiempo se dice que es estacionario, esto significa que los parámetros estadísticos del modelo probabilístico que caracteriza a un proceso no varían con el tiempo, si dichos parámetros cambian con el tiempo entonces se dice que el proceso es no estacionario. Los parámetros estadísticos que determinan si un proceso es estacionario o no, son la media, la varianza, la densidad de potencia espectral, y los momentos de orden superior. Sin embargo, existen grados en que una señal puede ser estacionaria, debido a que algunos de los parámetros estadísticos puede que sean estacionarios y otros no.

La mayoría de las señales con las que el hombre tiene relación diariamente son no estacionarias, sin embargo se utiliza mucho el concepto de señales estacionarias para desarrollar métodos de procesamiento de señales. La voz por ejemplo, que se considera una señal no estacionaria, pero puede ser considerada estacionaria para periodos cortos de tiempo. La mayoría de los procesos estocásticos tales como las señales de video, audio, información financiera, meteorológica, biomédica etc., no son estacionarias debido a que dichos datos son generados por sistemas que dependen de variables externas que cambian con el tiempo.

En el campo del procesamiento de señales se definen dos clases de procesos estacionarios:

- 1) **Procesos Estacionarios en el sentido estricto:** Se dice que un proceso pertenece a esta clase si los parámetros estadísticos de media, autocorrelación y densidad de potencia espectral son invariantes en el tiempo.
- 2) **Procesos Estacionarios en el sentido amplio:** Esta clase de procesos es menos restrictivo que el caso anterior, debido a que solo se requiere que la media y la autocorrelación sean invariantes en el tiempo.

### 1.3.3 Parámetros Estadísticos

Los parámetros estadísticos además de determinar si un proceso es estacionario o no, también definen la clase de modelo probabilístico del proceso, tal como Gaussiano, de Poisson etc.

Uno de los parámetros estadísticos más comunes es la esperanza, la cual también es conocida como promedio de ensamblado, debido a que el promedio se realiza sobre el eje de tiempo del proceso estocástico, por ejemplo, si  $X(n) = \{x(n,s)\}$  es un proceso estocástico, la media del elemento  $n$  del proceso será:

$$m_x(n) = E[X(n)] \tag{1.6}$$

donde  $E[\cdot]$  denota el operador de esperanza. Cabe aclarar que no es posible obtener, la esperanza de un proceso estocástico promediando en tiempo una realización del mismo, a menos que el proceso tenga cierta propiedad especial que se detallara en la sección 1.3.4. Otros promedios estadísticos se derivan del operador descrito arriba y se han utilizado a través de este capítulo bajo el sustento de que el lector tiene la idea básica de ellos; a continuación se definirán formalmente dichos promedios.

- 1) *Función de autocorrelación:* Para un proceso estocástico definido como  $X(n)$ , la autocorrelación se define como

$$\phi_{xx}(n, m) = E[X(n)X^*(m)] \tag{1.7}$$

donde el asterisco representa el conjugado de la señal. Las señales que se manejan en los procesos pueden corresponder a señales reales o complejas, es por eso que se define formalmente de esa manera.

- 2) *Función de correlación cruzada*: Esta definida para dos procesos estocásticos  $X(n)$  y  $Y(m)$  como

$$\phi_{xy}(n, m) = E[X(n)Y^*(m)] \quad (1.8)$$

De igual manera que en el caso continuo, la correlación es una forma de estimar el grado de relación entre una señal y otra.

Las definiciones descritas en los incisos anteriores son válidas para cualquier tipo de proceso, estacionario o no estacionario, sin embargo, debido a que mucho del análisis se realiza bajo procesos estacionarios en sentido amplio, definiremos de forma específica los parámetros estadísticos bajo este criterio. En concordancia a lo definido, un proceso es estacionario en el sentido amplio si su media y la correlación son invariantes en el tiempo, por lo tanto se tienen que cumplir estas dos condiciones

$$m_x(n) = m_x(n+k) \quad (1.9)$$

$$\phi_{xx}(n, m) = \phi_{xx}(n+k, m+k) \quad (1.10)$$

donde  $k$  representa solo un desplazamiento temporal. Estas dos condiciones implican que  $m_x(n)$  es constante para cualquier valor de  $n$  y que  $\phi_{xx}(n, m)$  solo depende de la diferencia  $n - m$ , por lo tanto podríamos definir la función de autocorrelación de forma más apropiada como

$$\phi_{xx}(k) = E[X(n)X^*(n-k)] \quad (1.11)$$

Si en el caso de la correlación cruzada los promedios de los procesos  $X(n)$  y  $Y(m)$  son independientes de  $n$  y dicha correlación solo depende de la diferencia  $n - m$  entonces se dice que son procesos conjuntamente estacionarios en el sentido amplio, y por lo tanto puede ser definida de la siguiente manera

$$\phi_{xy}(k) = E[X(n)Y^*(n-k)] \quad (1.12)$$

### 1.3.4 Ergodicidad

En muchos problemas de procesamiento de señales, solo es posible obtener una realización del proceso aleatorio, del cual se pueda obtener la media, correlación, y densidad espectral. En tales casos, se utilizan los promedios estadísticos obtenidos en tiempo como promedios de ensamblado, los cuales como ya se mencionó son calculados en base a diferentes realizaciones del proceso. Por lo tanto, se dice que un proceso estocástico es ergódico si exhibe las mismas características estadísticas a lo largo del eje de tiempo de una sola realización como a través del espacio o ensamble de diferentes realizaciones del proceso [7]. De esta manera una sola realización del proceso que es muestreada para un número grande de elementos, puede poseer las mismas características estadísticas en todo el espacio del proceso.

$$\underset{n \rightarrow \infty}{\text{promedios estadísti}} \cos[x(n, s)] = \underset{\text{a través del espacio } s}{\text{promedios estadísti}} \cos[x(n, s)]$$

### 1.3.5 Densidad de Potencia Espectral

La densidad de potencia espectral (PSD, *Power Spectral Density*) refleja el contenido espectral de un determinado proceso en función de la frecuencia. El estudio de este promedio estadístico de segundo orden es debido a que la convergencia de un filtro adaptivo depende directamente de la potencia espectral de la señal de entrada. La densidad de potencia espectral está definida como

$$P_{XX}(f) = E[X(f)X^*(f)] \tag{1.13}$$

$$= \sum_{n=-\infty}^{\infty} \phi_{xx}(k) e^{-j2\pi fn} \tag{1.14}$$

donde  $\phi_{xx}(k)$  y  $P_{XX}(f)$  representan la autocorrelación del proceso  $X(n)$  y la densidad de potencia espectral respectivamente,  $f$  representa la variable de frecuencia.

La densidad de potencia espectral es una función real no negativa, expresada en unidades de Watts/Hz. Para el caso en que se manejan procesos con valores reales y estacionarios, la autocorrelación es simétrica y por consecuencia la densidad de potencia espectral puede ser obtenida mediante la siguiente ecuación

$$P_{XX}(f) = \phi_{xx}(0) + \sum_{n=1}^{\infty} 2\phi_{xx}(n) \cos(2\pi fn) \tag{1.15}$$

### 1.3.6 Matriz de Correlación

La matriz de correlación forma parte importante dentro del comportamiento de un filtro adaptivo pues en muchos casos es quien determina la convergencia de un algoritmo, así como ciertas propiedades en estado estable. Si los elementos de la señal de entrada están representados en forma vectorial de la forma

$$\mathbf{x}(k) = [x_0(k) \quad x_1(k) \dots x_N(k)]^T \quad (1.16)$$

la matriz de correlación está definida como  $\mathbf{R} = E[\mathbf{x}(k)\mathbf{x}^H(k)]$  donde el superíndice  $H$  representa la transposición hermitiana, el cual es un operador matricial compuesto, que indica la transposición de la matriz, seguida de aplicar el conjugado de la misma. La matriz de correlación puede ser expresada en forma matricial de la siguiente manera

$$\mathbf{R} = \begin{bmatrix} E[|x_0(k)|^2] & E[x_0(k)x_1^*(k)] & \dots & E[x_0(k)x_N^*(k)] \\ E[x_1(k)x_0^*(k)] & E[|x_1(k)|^2] & \dots & E[x_1(k)x_N^*(k)] \\ \vdots & \vdots & \ddots & \vdots \\ E[x_N(k)x_0^*(k)] & E[x_N(k)x_1^*(k)] & \dots & E[|x_N(k)|^2] \end{bmatrix} \quad (1.17)$$

Las principales propiedades de la matriz de correlación son las siguientes [3]:

- 1) La matriz de correlación es positiva semidefinida.

Dado un vector arbitrario complejo  $\mathbf{w}$ , se puede obtener la señal dada por

$$y(k) = \mathbf{w}^H \mathbf{x}(k) \quad (1.18)$$

La magnitud al cuadrado de  $y(k)$  es

$$y(k)y^*(k) = |y(k)|^2 = \mathbf{w}^H \mathbf{x}(k)\mathbf{x}^H(k)\mathbf{w} \geq 0 \quad (1.19)$$

El valor medio cuadrático de  $y(k)$  está dado por

$$E[|y(k)|^2] = \mathbf{w}^H E[\mathbf{x}(k)\mathbf{x}^H(k)]\mathbf{w} = \mathbf{w}^H \mathbf{R}\mathbf{w} \geq 0 \quad (1.20)$$

Por lo tanto  $\mathbf{R}$  es positiva semidefinida.

- 2) El hermitiano de la matriz de correlación es la misma matriz.

$$\mathbf{R} = \mathbf{R}^H \quad (1.21)$$

$$\mathbf{R}^H = E\{[\mathbf{x}(k)\mathbf{x}^H(k)]^H\} = E[\mathbf{x}(k)\mathbf{x}^H(k)] = \mathbf{R}$$

- 3) Una matriz es Toeplitz si los elementos de la diagonal principal y de alguna diagonal secundaria son iguales. Cuando el vector de la señal de entrada está compuesta de versiones retrasadas de la misma señal tomadas de un proceso estacionario en el sentido amplio, la matriz  $\mathbf{R}$  es Toeplitz.

Si  $\mathbf{R}$  se define como se describió anteriormente entonces

$$\mathbf{R} = \begin{bmatrix} \phi_x(0) & \phi_x(1) & \cdots & \phi_x(N) \\ \phi_x(-1) & \phi_x(0) & \cdots & \phi_x(N-1) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_x(-N) & \phi_x(-N+1) & \cdots & \phi_x(0) \end{bmatrix} \quad (1.22)$$

Por simple inspección, y de acuerdo a la definición dada podemos concluir que  $\mathbf{R}$  es Toeplitz.

## 1.4 Filtro Wiener

El estudio de los filtros Wiener es la base para entender la mayoría de las implicaciones del filtrado adaptivo, a esta clase de filtros también se les conoce como filtros lineales óptimos debido a que involucran un proceso de estimación lineal de una señal deseada de otra relacionada con ésta. El estudio de este filtro se enfocará para el caso de señales de tiempo discreto y de valores reales, las señales utilizadas se considerarán como procesos aleatorios.

Como se mencionó en la sección 1.2.1, uno de los algoritmos adaptivos más importantes y más utilizados es el de error medio cuadrático (MSE), del cual se definió de la siguiente manera

$$F[e(k)] = \xi(k) = E[|e(k)|^2] \quad (1.23)$$

y de acuerdo a la Fig. 1.1, el error es generado como  $e(k) = d(k) - y(k)$ , por lo tanto substituyendo en la ecuación anterior obtenemos que

$$\xi(k) = E[|e(k)|^2] = E[d^2(k) - 2d(k)y(k) + y^2(k)] \quad (1.24)$$

Si se considera que la estructura del filtro debe consistir de una combinación lineal, entonces se opta por la utilización de una estructura FIR como la mostrada en la Fig. 1.2(a) la cual su salida en forma genérica esta dada por la siguiente ecuación

$$y(k) = \sum_{i=0}^N w_i(k)x(k-i) = \mathbf{w}^T(k)\mathbf{x}(k) \quad (1.25)$$

donde  $\mathbf{x}(k)=[x(k) \ x(k-1) \dots x(k-N)]^T$  y  $\mathbf{w}(k)=[w_0(k) \ w_1(k) \dots w_N(k)]^T$  son la señal de entrada y los coeficientes del filtro adaptivo respectivamente. Por lo tanto podemos reescribir la ecuación (1.24) de la siguiente manera

$$E[e^2(k)] = \xi(k) \tag{1.26}$$

$$\begin{aligned} &= E[d^2(k) - 2d(k)\mathbf{w}^T(k)\mathbf{x}(k) + \mathbf{w}^T(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{w}(k)] \\ &= E[d^2(k)] - 2E[d(k)\mathbf{w}^T(k)\mathbf{x}(k)] + E[\mathbf{w}^T(k)\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{w}(k)] \end{aligned} \tag{1.27}$$

Para un filtro con coeficientes fijos invariantes en el tiempo, el error medio cuadrático está dado como

$$\begin{aligned} \xi &= E[d^2(k)] - 2\mathbf{w}^T E[d(k)\mathbf{x}(k)] + \mathbf{w}^T E[\mathbf{x}(k)\mathbf{x}^T(k)]\mathbf{w} \\ &= E[d^2(k)] - 2\mathbf{w}^T \mathbf{p} + \mathbf{w}^T \mathbf{R} \mathbf{w} \end{aligned} \tag{1.28}$$

donde  $\mathbf{p} = E[d(k)\mathbf{x}(k)]$  es el vector de correlación cruzada, entre la señal deseada y la señal de entrada, y  $\mathbf{R} = E[\mathbf{x}(k)\mathbf{x}^T(k)]$  es la matriz de correlación de la señal de entrada. Como puede notarse, el error medio cuadrático es una función que varía dependiendo de los coeficientes del filtro, pues tanto como  $\mathbf{p}$  y  $\mathbf{R}$  son conocidos de acuerdo a los procesos aleatorios que les dan origen. Por lo tanto, para obtener el error mínimo posible se aplica la función gradiente a la función objetivo, para obtener los coeficientes que minimizan el error medio cuadrático del sistema.

$$\begin{aligned} \nabla_{\mathbf{w}} &= \frac{\partial \xi}{\partial \mathbf{w}^T} = \begin{bmatrix} \frac{\partial \xi}{\partial w_0} & \frac{\partial \xi}{\partial w_1} & \dots & \frac{\partial \xi}{\partial w_N} \end{bmatrix} \\ &= -2\mathbf{p} + 2\mathbf{R} \mathbf{w} \end{aligned} \tag{1.29}$$

Igualando el vector gradiente a cero y asumiendo que  $\mathbf{R}$  no es singular, los valores óptimos de los coeficientes del filtro pueden ser obtenidos mediante la siguiente ecuación

$$\mathbf{w}_o = \mathbf{R}^{-1} \mathbf{p} \tag{1.30}$$

A esta solución se le conoce como *solución óptima de la ecuación Wiener-Hopf*, la cual fue mencionada en secciones anteriores. De aquí se deriva el lema conocido como inversión de matriz, debido precisamente al problema de obtener la inversión de la matriz de correlación para obtener la solución óptima. Por otro lado, debido a que no se pueden obtener estimaciones muy precisas acerca de la matriz de correlación y correlación cruzada, se utiliza el principio de ergodicidad para obtener los promedios estadísticos de ensamblaje, esta característica generalmente va implícita cuando se desarrollan la mayoría de los filtros adaptivos.

Para obtener el error mínimo cuadrático se sustituye la solución de Wiener en la ecuación (1.28) obteniendo finalmente la siguiente ecuación

$$\begin{aligned}\xi_{\min} &= E[d^2(k)] - 2\mathbf{w}_o^T \mathbf{p} + \mathbf{w}_o^T \mathbf{R} \mathbf{R}^{-1} \mathbf{p} \\ &= E[d^2(k)] - \mathbf{w}_o^T \mathbf{p}\end{aligned}\tag{1.31}$$

Aquí se puede ver claramente que el error depende directamente de la señal deseada y la señal de entrada, si ocurriese que en una aplicación la señal de entrada y la señal deseada no estuvieran correlacionadas, el error solo dependería de la potencia de la señal deseada, sin embargo en la mayoría de los casos existe cierta relación entre ambas variables.

Una importante propiedad del filtro Wiener puede ser deducida al analizar el gradiente de error en la solución óptima.

$$\nabla_{\mathbf{w}} = \frac{\partial E[e^2(k)]}{\partial \mathbf{w}^T} = E\left[2e(k) \frac{\partial e(k)}{\partial \mathbf{w}^T}\right] = -E[2e(k)\mathbf{x}(k)]\tag{1.32}$$

Al igualar el gradiente a cero para minimizar el error obtenemos que

$$E[e_o(k)\mathbf{x}(k)] = \mathbf{0}\tag{1.33}$$

Lo anterior se le conoce como *principio de ortogonalidad* el cual infiere que la señal de error óptima está rotada  $90^\circ$  de la señal de entrada, por lo cual se demuestra que ambas señales no están correlacionadas. Esto se origina del hecho de que todas las variables aleatorias que poseen momentos finitos de segundo orden constituyen un espacio lineal, y que el producto punto de ambas determina el grado de correlación entre una variable y otra. En el caso en que la señal de error o la señal de entrada tengan una media de cero, el principio de ortogonalidad implica que ambas señales no están correlacionadas.

Una derivación importante del principio de ortogonalidad, es la que demuestra que la señal de error en su valor óptimo y la señal de salida del filtro también son ortogonales, como se muestra a continuación

$$\begin{aligned}E[e_o(k)y(k)] &= E[y(k)e_o(k)] \\ &= E[\mathbf{w}_o^T \mathbf{x}(k)e_o(k)] \\ &= \mathbf{w}_o^T E[\mathbf{x}(k)e_o(k)] \\ &= 0\end{aligned}\tag{1.34}$$

## 1.5 Filtro LMS

Como se pudo constatar en la sección anterior, los principales problemas de implementar el filtro Wiener surgen debido a que no es posible obtener mediciones precisas de la matriz de correlación  $\mathbf{R}$  y del vector de correlación cruzada  $\mathbf{p}$ , por lo que se estiman por medio de promedios temporales acudiendo a la propiedad de ergodicidad. Sin embargo, el problema real radica en resolver el lema conocido como inversión de matriz, el cual puede ser resuelto de manera apropiada con métodos numéricos capaces de ser implementados de manera simple en hardware. El algoritmo adaptivo más utilizado para lograr optimizar al filtro es el steepest-descent debido a su simplicidad de implementación y análisis. Al utilizar una estructura FIR, el criterio de error medio cuadrático, así como el algoritmo steepest-descent para minimizar la función objetivo, se dice que se utiliza un filtro LMS (*Least Mean Square*).

Como se había establecido anteriormente, la solución óptima de Wiener se obtenía mediante la siguiente ecuación

$$\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p}$$

donde  $\mathbf{R} = E[\mathbf{x}(k)\mathbf{x}^H(k)]$  y  $\mathbf{p} = E[d(k)\mathbf{x}(k)]$ , asumiendo que  $d(k)$  y  $\mathbf{x}(k)$  son conjuntamente estacionarias en sentido amplio. Si la matriz de correlación y el vector de correlación cruzada son expresados en forma estricta como una estimación, y se utiliza el método de búsqueda steepest-descent entonces, la solución óptima de Wiener puede ser obtenida como

$$\begin{aligned} \mathbf{w}(k+1) &= \mathbf{w}(k) - \mu \hat{\nabla}_{\mathbf{w}}(k) \\ &= \mathbf{w}(k) + 2\mu(\hat{\mathbf{p}}(k) - \hat{\mathbf{R}}(k)\mathbf{w}(k)) \end{aligned} \quad (1.35)$$

para  $k=0,1,2,\dots$ , donde  $\hat{\nabla}_{\mathbf{w}}(k)$  representa el gradiente estimado de la función objetivo con respecto a los coeficientes del filtro.

Las estimaciones de la matriz de correlación y el vector de correlación cruzada pueden ser obtenidas mediante estimaciones instantáneas de cada una de ellas

$$\hat{\mathbf{R}}(k) = \mathbf{x}(k)\mathbf{x}^T(k) \quad (1.36)$$

$$\hat{\mathbf{p}}(k) = d(k)\mathbf{x}(k) \quad (1.37)$$

Al sustituir las ecuaciones (1.36) y (1.37) en (1.35) obtenemos que la ecuación de actualización del filtro está dada como

$$\begin{aligned}
 \mathbf{w}(k+1) &= \mathbf{w}(k) + 2\mu(d(k)\mathbf{x}(k) - \mathbf{x}(k)\mathbf{x}^T(k)\mathbf{w}(k)) \\
 &= \mathbf{w}(k) + 2\mu\mathbf{x}(k)(d(k) - y(k)) \\
 &= \mathbf{w}(k) + 2\mu e(k)\mathbf{x}(k)
 \end{aligned}
 \tag{1.38}$$

Parte de la rapidez de convergencia de este filtro depende que  $\mu$ , así como del rango que garantiza la convergencia. Finalmente, podemos describir en forma completa el filtro LMS mediante las siguientes ecuaciones.

Inicialización  $\mathbf{x}(0) = \mathbf{w}(0) = [0 \ 0 \ \dots \ 0]^T$  (1.39)

Para  $k \geq 0$

$$e(k) = d(k) - \mathbf{x}^T(k)\mathbf{w}(k) \tag{1.40}$$

$$\mathbf{w}(k+1) = \mathbf{w}(k) + 2\mu e(k)\mathbf{x}(k) \tag{1.41}$$

Es preciso aclarar que la inicialización de los coeficientes del filtro no necesariamente debe empezar en ceros, pues en semejanza a cuando se utilizan los métodos de búsqueda para encontrar las raíces de un polinomio, la solución puede obtenerse más rápido si se puede estimar la vecindad en la que se encuentra la solución.

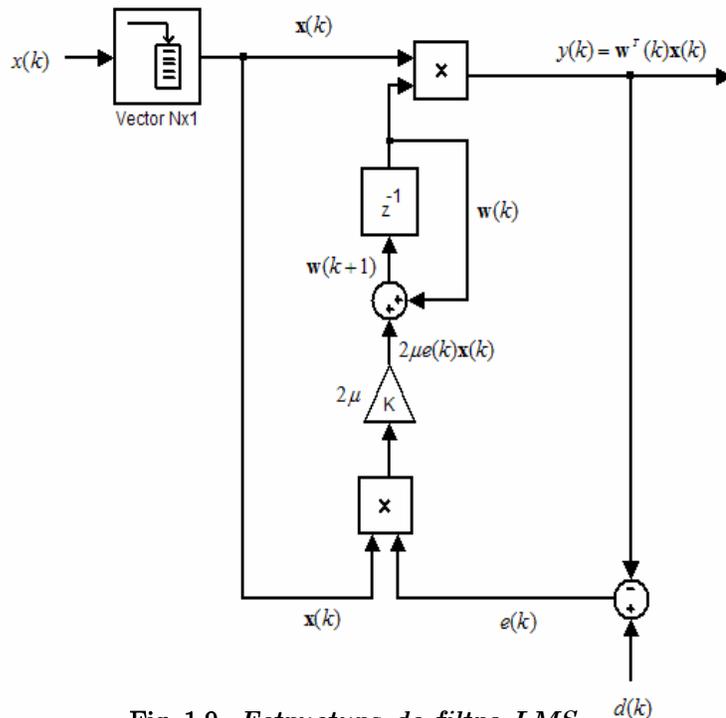


Fig. 1.9 Estructura de filtro LMS.

## 1.6 Efectos de palabra finita

Todo sistema desarrollado en tiempo discreto y que desea implementarse en hardware requiere que todas sus señales de entrada, coeficientes de filtros, operaciones etc., se realicen con una precisión finita determinada, esto implica que el sistema discreto (*precisión infinita*) sea convertido a un sistema estrictamente digital. De tal manera que al diseñar un filtro en tiempo discreto, es necesario que los coeficientes del mismo sean cuantizados. Esto último quiere decir que dichos valores deben ser transformados a un formato binario, tal como el de punto fijo o el de punto flotante. En nuestro caso nos enfocaremos al tratado del formato de punto fijo, el cual asume que el punto binario está fijo en una posición, y en la implementación en hardware se toma en cuenta dicha posición para realizar operaciones aritméticas [8]. Los números en punto fijo son siempre representados como fracciones debido precisamente a que la multiplicación o suma de dos números de punto fijo resulta en un valor que mantiene la posición del punto binario.

La representación en punto fijo se puede ver en la Fig. 1.10 en la que se asume que la longitud de palabra disponible es de  $(b+1)$  bits, donde el bit más significativo representa el signo y cualquier número representado con este formato está cuantizado en pasos de  $2^{-b}$ , llamado paso o ancho de cuantización.

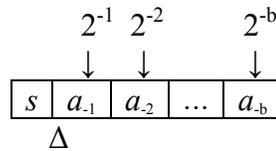


Fig. 1.10 Formato numérico de punto fijo.

Antes de la cuantización, la longitud de palabra es mucho mayor que la indicada al cuantizar. Si se asume que un dato  $x$  es representado con una precisión de  $(\beta+1)$  bits cuando  $\beta \gg b$ , entonces para convertir dicho valor a una fracción de  $(b+1)$  bits, se le aplica un modelo de cuantización como se observa en la Fig. 1.11. Existen dos tipos de cuantización: *truncamiento* y *redondeo*, el truncamiento puede ser de *signo-magnitud*, *complemento a uno* o *complemento a dos*, dependiendo de la representación numérica que se utilice; el tipo más común es el de truncamiento complemento a dos.

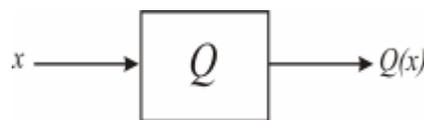
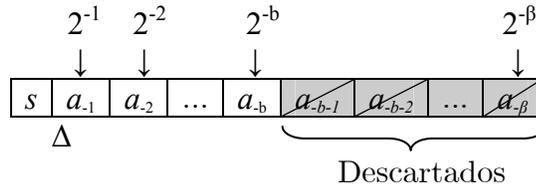


Fig. 1.11 Modelo del proceso de cuantización.

Para realizar el truncamiento de un número de punto fijo de precisión  $(\beta + 1)$  bits a  $(b + 1)$ , simplemente se descartan los bits menos significativos  $(\beta - b)$  como se muestra en la Fig. 1.12, y el error generado por este truncamiento está definido como

$$\varepsilon_t = Q(x) - x \tag{1.42}$$



**Fig. 1.12** *Proceso de truncamiento*

Para el caso de truncamiento complemento a dos, el error siempre es negativo y tiene un rango de

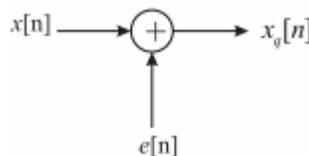
$$-(2^{-b} - 2^{-\beta}) \leq \varepsilon_t \leq 0 \tag{1.43}$$

En el caso de redondeo, el número es cuantizado al nivel de cuantización más próximo, si un número se encuentra exactamente a la mitad de los dos niveles de cuantización, dicho valor es redondeado al nivel superior próximo. Es de aclarar que el error de redondeo, no depende la representación numérica que se está utilizando pues únicamente está basado en la magnitud del número, por lo tanto el rango de redondeo está dado por

$$-\frac{1}{2}(2^{-b} - 2^{-\beta}) < \varepsilon_r \leq \frac{1}{2}(2^{-b} - 2^{-\beta}) \tag{1.44}$$

### 1.6.1 Modelo de Ruido de Cuantización

Para modelar los efectos de cuantización en un sistema, lo más común es utilizar un modelo estadístico tal como se muestra en la Fig. 1.13, este modelo tiene origen precisamente en la ecuación (1.42), la cual permite estimar el error como una perturbación aditiva. Debido a que este es un modelo estadístico, es necesario obtener las distribuciones de ruido generadas por cada uno de los tipos de truncamiento.



**Fig. 1.13** *Modelo Estadístico de cuantización.*

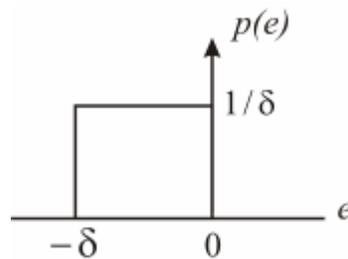
Las distribuciones de cada uno de los tipos de truncamiento puede ser obtenidas mediante las ecuaciones (1.43) y (1.44), considerando que  $\beta \gg b$  y que la distribución es uniforme. Para el caso de truncamiento complemento a dos tenemos que

$$\int_{-2^{-b}}^0 P_{\varepsilon_t} d\varepsilon_t = 1 \quad (1.45)$$

Si  $P_{\varepsilon_t}$  es la densidad de probabilidad de error, la cual es constante debido a que la distribución es uniforme entonces obtenemos que

$$P_{\varepsilon_t} = 2^b \quad -2^{-b} \leq \varepsilon_t \leq 0 \quad (1.46)$$

En la Fig. 1.14 se muestra la función de densidad de probabilidad de error para el caso de truncamiento complemento a dos, donde  $\delta = 2^{-b}$ .



**Fig. 1.14** *Distribución de error para truncamiento complemento a dos.*

Debido a que estas funciones de densidad de probabilidad de error son utilizadas para modelar sistemas cuantizados, es necesario conocer los promedios estadísticos de estas distribuciones ya que son muy útiles durante el análisis. Por lo tanto, la media y la varianza para truncamiento se obtienen de la siguiente manera

*Media*

$$\begin{aligned} \mu_t &= E[2^b] \\ &= \int_{-2^{-b}}^0 2^b x dx = 2^b \frac{x^2}{2} \Big|_{-2^{-b}}^0 \\ &= -2^{-b-1} \end{aligned} \quad (1.47)$$

*Varianza*

$$\begin{aligned} \sigma_t^2 &= E[(2^b - \mu)^2] \\ &= \int_{-2^{-b}}^0 (x + 2^{-b-1})^2 2^b dx \\ &= \frac{2^{-2b}}{12} \end{aligned} \quad (1.48)$$

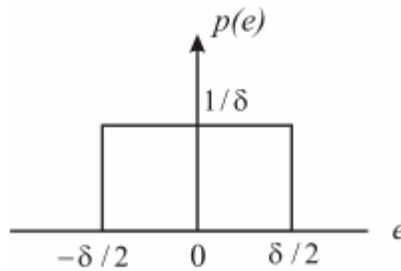
Para el caso de redondeo tenemos que

$$\int_{-2^{-b-1}}^{2^{-b-1}} P_{\varepsilon_t} d\varepsilon_t = 1 \quad (1.49)$$

Si la densidad de nuevo se considera constante, desarrollando la ecuación obtenemos

$$P_{\varepsilon_t} = 2^b \quad -2^{-b-1} \leq \varepsilon_t \leq 2^{-b-1} \quad (1.50)$$

En la Fig. 1.15 se muestra la densidad de probabilidad de error para el redondeo



**Fig. 1.15** *Distribución de error para redondeo.*

Los promedios estadísticos para este caso son:

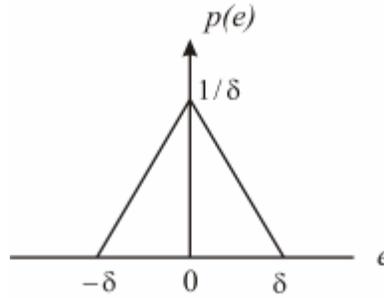
*Media*

$$\begin{aligned} \mu_t &= E[2^b] \\ &= \int_{-2^{-b-1}}^{2^{-b-1}} 2^b x dx = 2^b \left. \frac{x^2}{2} \right|_{-2^{-b-1}}^{2^{-b-1}} \\ &= 0 \end{aligned} \quad (1.51)$$

*Varianza*

$$\begin{aligned} \sigma_r^2 &= E[(2^b)^2] \\ &= \int_{-2^{-b-1}}^{2^{-b-1}} x^2 2^b dx \\ &= \frac{2^{-2b}}{12} \end{aligned} \quad (1.52)$$

Un caso especial que será muy utilizado en el siguiente capítulo es la distribución de la diferencia de dos densidades iguales de probabilidad de error uniforme, pero que cada una de ellas es generada independientemente. Por experimentación se descubrió que la función de densidad resultante es una función triangular para el caso en que las distribuciones sean generadas por truncamiento o por redondeo. La función de distribución puede ser vista en la Fig. 1.14 y sus promedios estadísticos pueden ser obtenidos de los promedios obtenidos anteriormente.



**Fig. 1.16** *Distribución resultante de la diferencia de dos distribuciones uniformes.*

*Media.-* Considerando que  $\mu_{\varepsilon_1} = \mu_{\varepsilon_2}$ , para ambos casos, truncamiento o redondeo.

$$\begin{aligned} E[(P_{\varepsilon_1} - P_{\varepsilon_2})] &= E[P_{\varepsilon_1}] - E[P_{\varepsilon_2}] \\ &= \mu_{\varepsilon_1} - \mu_{\varepsilon_2} \\ &= 0 \end{aligned} \tag{1.53}$$

*Varianza.-* Considerando que  $\sigma_{\varepsilon_1}^2 = \sigma_{\varepsilon_2}^2$  y  $\mu_{\varepsilon_1} = \mu_{\varepsilon_2}$  para ambos casos, truncamiento o redondeo.

$$\begin{aligned} E[(P_{\varepsilon_1} - P_{\varepsilon_2})^2] &= E[(P_{\varepsilon_1})^2] - 2E[P_{\varepsilon_1}P_{\varepsilon_2}] + E[(P_{\varepsilon_2})^2] \\ &= (\sigma_{\varepsilon_1}^2 + \mu_{\varepsilon_1}^2) - 2\mu_{\varepsilon_1}\mu_{\varepsilon_2} + (\sigma_{\varepsilon_2}^2 + \mu_{\varepsilon_2}^2) \\ &= 2\sigma_{\varepsilon}^2 \\ &= \frac{2^{-2b}}{6} \end{aligned} \tag{1.54}$$

Los efectos que generan estas perturbaciones en los sistemas pueden ser diversos, por ejemplo, al cuantizar los coeficientes de un filtro puede que la respuesta en frecuencia se modifique con respecto a su diseño original, de igual manera los filtros que cuentan con una estructura IIR deben tomar en cuenta estos errores ya que el sistema en lugar de filtrar oscile debido a su propiedad recursiva. Otro problema que puede surgir es cuando el error de cuantización es comparable a los valores de la señal cuantizada, ya que la relación señal a ruido disminuye, afectando de esta manera el desempeño de cualquier sistema electrónico.

---

## Capítulo 2

# MODELO MATEMÁTICO

"Si la gente no cree que las matemáticas son simples, es solo porque no se han dado cuenta cuan complicada es la vida "

-John von Neumann-

---

### 2.1 Filtro LMS implementado con aritmética de punto fijo

El diagrama de bloques detallado en Fig. 2.1 muestra el modelo típico de identificación de sistemas que incluye los errores de precisión finita modelados como ruido aditivo.

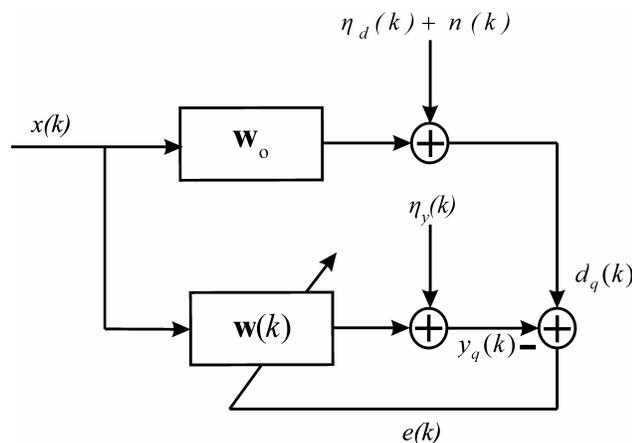


Fig. 2.1 Algoritmo LMS de precisión finita para configuración identificación de sistemas.

La señal de entrada es una de las principales variables que determina el comportamiento del algoritmo LMS, así como una fuente importante de información para establecer las condiciones del análisis matemático. En la Fig. 2.1,  $x(k)$  representa la señal de entrada, la cual se asume no sufre cuantización y como se puntualizó anteriormente dichos valores de entrada son mutuamente independientes, de media cero y generados por

un proceso estacionario. El siguiente análisis puede ser utilizado ampliamente con señales de ruido de banda ancha tal como el ruido blanco, comúnmente utilizado en esta configuración debido a que con dicha entrada el algoritmo converge más rápido; por otro lado, se aprovecha que las muestras sucesivas del ruido de banda ancha no están correlacionadas entre ellas [6] lo cual simplifica el análisis matemático. El vector  $\mathbf{w}_o = [w_{o1} w_{o2} \dots w_{oN-1}]^T$  representa la respuesta al impulso del sistema desconocido,  $\mathbf{w}(k) = [w_0(k) w_1(k) \dots w_{N-1}(k)]^T$  representa el vector ponderado de longitud  $N$  de la  $k$ va iteración obtenido de la ecuación de actualización, la cual esta dada como

$$\mathbf{w}(k+1) = \mathbf{w}(k) + 2\mu_q e(k)\mathbf{x}(k) + \boldsymbol{\eta}_u(k) \quad (2.1)$$

en donde  $\mu_q$  representa el factor de convergencia cuantizado,  $e(k)$  es la señal error,  $\mathbf{x}(k)$  representa el vector de corrimiento de la señal de entrada, y  $\boldsymbol{\eta}_u(k)$  representa el ruido producto de la cuantización de la multiplicación resultante de  $2\mu_q e(k)\mathbf{x}(k)$ , cabe precisar que todas las variables con subíndice  $q$  indican que han sufrido cuantización, así mismo se supone que la señal de entrada está propiamente escalada por lo que no ocurre error debido a sobreflujo. Las perturbaciones  $\eta_d(k)$  y  $\eta_y(k)$  son debidas a la cuantización realizada después de la suma de los productos internos entre el vector de entrada, y el vector ponderado correspondiente. La señal  $n(k)$  se le conoce comúnmente como ruido de medición (measurement noise).

Se asume que las señales de ruido  $\eta_d(k)$ ,  $\eta_y(k)$ , y  $\boldsymbol{\eta}_u(k)$  no están correlacionadas entre ellas y tienen una distribución uniforme, y se definen formalmente de la siguiente manera

$$\eta_d(k) = d_q(k) - d(k) \quad (2.2)$$

$$\eta_y(k) = y_q(k) - y(k) \quad (2.3)$$

$$\boldsymbol{\eta}_u = [2\mu_q e(k)\mathbf{x}(k)]_q - 2\mu_q e(k)\mathbf{x}(k) \quad (2.4)$$

El primer y segundo momento central de las distribuciones de ruido debidas al redondeo y truncamiento se muestran en la siguiente tabla.

**Tabla 2.1**  
**Primer y segundo momento de distribuciones cuantizadas.**

Distribuciones de error	Truncamiento		Redondeo	
	Media	Varianza	Media	Varianza
$\eta_d(k), \eta_y(k)$	$-2^{-b_d-1}$	$\frac{2^{-2b_d}}{12}$	0	$\frac{2^{-2b_d}}{12}$
$\boldsymbol{\eta}_u(k)$	$-2^{-b_c-1}$	$\frac{2^{-2b_c}}{12}$	0	$\frac{2^{-2b_c}}{12}$

El proceso de cuantización del algoritmo involucra dos cuantizadores de diferente resolución, uno para los coeficientes el cual es de  $b_c$  bits más el de signo y otro para los datos de  $b_d$  bits más el de signo.

## 2.2 *Modelo de comportamiento transitorio y de estado estable*

El análisis conjunto del comportamiento transitorio y de estado estable de la curva de aprendizaje es sin duda una tarea difícil; sin embargo, revela información importante del sistema, tal como la dependencia de ambos comportamientos y permite analizar distintos parámetros de desempeño bajo las mismas condiciones y criterios establecidos para todo el análisis.

De la Fig. 2.1 podemos deducir las siguientes ecuaciones que nos permitirán desarrollar el modelo matemático.

$$e(k) = d_q(k) - y_q(k) \quad (2.5)$$

$$d_q(k) = \mathbf{w}_o^T \mathbf{x}(k) + n(k) + \eta_d(k) \quad (2.6)$$

$$y_q(k) = \mathbf{w}^T(k) \mathbf{x}(k) + \eta_y(k) \quad (2.7)$$

Dado que nuestro objetivo es obtener la curva de aprendizaje, ésta última es obtenida utilizando el criterio de optimación de error medio cuadrático, el cual está definido como

$$E[|e(k)|^2] \quad (2.8)$$

de esta manera procedemos a sustituir las ecuaciones (2.6) y (2.7) en (2.5), y obtenemos la siguiente ecuación

$$e(k) = n(k) + \eta_e(k) - \mathbf{v}^T(k) \mathbf{x}(k) \quad (2.9)$$

donde  $\mathbf{v}(k) = \mathbf{w}(k) - \mathbf{w}_o$ , el cual se le conoce como el vector de error ponderado, así mismo  $\eta_e(k) = \eta_d(k) - \eta_y(k)$  permite conjuntar ambos ruidos de cuantización en uno solo y al analizar el primer y segundo momento central de esta distribución nos permitirá ahorrar trabajo algebraico extra; de acuerdo a lo obtenido en el capítulo anterior dichos momentos centrales de  $\eta_e(k)$  corresponden precisamente a las ecuaciones (1.53) y (1.54).

De la ecuación (2.9) podemos deducir fácilmente que el algoritmo alcanza el error óptimo cuando  $\mathbf{v}(k) = \mathbf{0}$ , por lo que se define de la siguiente manera

$$e_o(k) = n(k) + \eta_e(k) \quad (2.10)$$

donde  $n(k)$  se considera como ruido blanco de media cero como es comúnmente asumido, y dado que  $\eta_e(k)$  es también una ruido de media cero, la distribución de error óptimo conservará este parámetro estadístico.

Habiendo detallado en los elementos de la ecuación (2.9), procedemos a aplicar el criterio de optimación en dicha ecuación, resultando (2.11). Después de manipular algebraicamente (2.11) se asume que  $E[n(k)\eta_e(k)] = 0$ , así mismo se asume que  $\mathbf{x}(k)$  y  $\mathbf{v}(k)$  son vectores independientes lo que permite manipular el término  $E[e_o(k)\mathbf{v}^T(k)\mathbf{x}(k)]$  como se muestra en (2.12).

$$E[|e(k)|^2] = E[(e_o(k) - \mathbf{v}^T(k)\mathbf{x}(k))^2] = E[n^2(k)] + E[\eta_e^2(k)] + E[(\mathbf{v}^T(k)\mathbf{x}(k))^2] \quad (2.11)$$

$$E[e_o(k)\mathbf{v}^T(k)\mathbf{x}(k)] = E[E[e_o(k)\mathbf{x}^T(k)]\mathbf{v}(k)] \quad (2.12)$$

Citando el principio de ortogonalidad [6] el cual establece (2.13), podemos concluir (2.14).

$$E[e_o(k)\mathbf{x}^T(k)] = \mathbf{0}^T \quad (2.13)$$

$$E[e_o(k)\mathbf{v}^T(k)\mathbf{x}(k)] = 0 \quad (2.14)$$

Por otro lado, de acuerdo a las características previamente establecidas para  $n(k)$ , se obtiene que  $E[n^2(k)] = \sigma_n^2$ , así mismo podemos recordar que el término  $E[\eta_e^2(k)]$  es obtenido por medio de la ecuación (1.54), mientras que el último término puede ser expresado como  $\text{tr}[\mathbf{R}\mathbf{K}(k)]$  de acuerdo a lo propuesto en [6], donde  $\mathbf{R} = E[\mathbf{x}(k)\mathbf{x}^T(k)]$ ,  $\mathbf{K}(k) = E[\mathbf{v}(k)\mathbf{v}^T(k)]$  y  $\text{tr}[\cdot]$  representa la función traza. Considerando lo anterior podemos rescribir la ecuación (2.11) como (2.15).

$$\xi(k) = E[|e(k)|^2] = \sigma_n^2 + \sigma_e^2 + \text{tr}[\mathbf{R}\mathbf{K}(k)] \quad (2.15)$$

Hasta ahora se ha definido una ecuación que nos permite visualizar parte de la solución del problema, y nos indica la dirección hacia lo que debemos enfocarnos para continuar el análisis. Es claro que los dos primeros términos de la ecuación (2.15) pertenecen al comportamiento de estado estable, y sin duda el tercer término contiene el comportamiento transitorio, es por eso que nos enfocaremos primeramente al análisis de la matriz  $\mathbf{K}(k)$ .

La matriz de correlación del vector de error ponderado  $\mathbf{K}(k)$  está definida como

$$\mathbf{K}(k) = E[\mathbf{v}(k)\mathbf{v}^T(k)] \quad (2.16)$$

Si sustituimos la ecuación (2.10) en (2.9) obtenemos que la señal error esta dada por (2.17) y sustituyendo en la ecuación (2.1) podemos redefinir el vector de error ponderado como (2.18).

$$e(k) = e_o(k) - \mathbf{v}^T(k)\mathbf{x}(k) \quad (2.17)$$

$$\mathbf{v}(k+1) = (\mathbf{I} - 2\mu_q\mathbf{x}(k)\mathbf{x}^T(k))\mathbf{v}(k) + 2\mu_q e_o(k)\mathbf{x}(k) + \boldsymbol{\eta}_u(k) \quad (2.18)$$

Ahora procederemos a encontrar  $\mathbf{K}(k+1)$  utilizando la ecuación (2.18) asumiendo la independencia entre  $e_o(k)$ ,  $\boldsymbol{\eta}_u(k)$ ,  $\mathbf{v}(k)$  y  $\mathbf{x}(k)$ .

$$\mathbf{K}(k+1) = E[\mathbf{v}(k+1)\mathbf{v}^T(k+1)] \quad (2.19)$$

$$\begin{aligned} \mathbf{K}(k+1) &= E\left[ (\mathbf{I} - 2\mu_q\mathbf{x}(k)\mathbf{x}^T(k))\mathbf{v}(k)\mathbf{v}^T(k)(\mathbf{I} - 2\mu_q\mathbf{x}(k)\mathbf{x}^T(k)) \right] \\ &\quad + 4\mu_q^2 E[e_o^2(k)\mathbf{x}(k)\mathbf{x}^T(k)] + E[\boldsymbol{\eta}_u(k)\boldsymbol{\eta}_u^T(k)] \\ &\quad + E[\boldsymbol{\eta}_u(k)\mathbf{v}^T(k)(\mathbf{I} - 2\mu_q\mathbf{x}(k)\mathbf{x}^T(k))] + E\left[ (\mathbf{I} - 2\mu_q\mathbf{x}(k)\mathbf{x}^T(k))\mathbf{v}(k)\boldsymbol{\eta}_u^T(k) \right] \end{aligned} \quad (2.20)$$

Reduciendo (2.20) resulta (2.21) donde  $\mu_u$  y  $\sigma_u^2$  son el primer y segundo momento central respectivamente de la distribución  $\eta_u(k)$ , la matriz  $\mathbf{O}$  representa una matriz de unos, y  $\boldsymbol{\mu}_v(k) = E[\mathbf{v}(k)]$ .

$$\begin{aligned} \mathbf{K}(k+1) &= \mathbf{K}(k) - 2\mu_q (\mathbf{K}(k)\mathbf{R} + \mathbf{R}\mathbf{K}(k)) + 4\mu_q^2 E\left[ \mathbf{x}(k)\mathbf{x}^T(k)\mathbf{K}(k)\mathbf{x}(k)\mathbf{x}^T(k) \right] \\ &\quad + 4\mu_q^2 \sigma_o^2 \mathbf{R} + \sigma_u^2 \mathbf{I} + \mu_u^2 \mathbf{O} + \boldsymbol{\mu}_u \boldsymbol{\mu}_v^T(k) (\mathbf{I} - 2\mu_q \mathbf{R}) + (\mathbf{I} - 2\mu_q \mathbf{R}) \boldsymbol{\mu}_v(k) \boldsymbol{\mu}_u^T \end{aligned} \quad (2.21)$$

Analizando la ecuación (2.21) podemos observar que se sigue comportando como una ecuación recursiva, en donde la matriz de correlación del vector de ponderado no es lo único que está variando en cada iteración sino también la esperanza del vector de error ponderado; cabe precisar que los últimos tres términos de la ecuación es la aportación debida a la cuantización por truncamiento por la presencia de  $\mu_u$  en dichos términos.

Antes de proseguir estudiando la ecuación (2.21), nos enfocaremos a obtener la esperanza del vector de error ponderado la cual será utilizada posteriormente; por lo tanto, aplicando la esperanza a la ecuación (2.18) obtenemos (2.22), la cual que es una clásica ecuación de diferencias, por lo que procedemos a aplicar la transformada  $\mathcal{Z}$  en ambos lados de la igualdad para obtener (2.23) donde  $\boldsymbol{\mu}_v(0) = -\mathbf{w}_0$  y después de agrupar se obtiene (2.24).

$$\boldsymbol{\mu}_v(k+1) = (\mathbf{I} - 2\mu_q \mathbf{R}) \boldsymbol{\mu}_v(k) + \boldsymbol{\mu}_u \quad (2.22)$$

$$z \boldsymbol{\mu}_v(z) - z \boldsymbol{\mu}_v(0) = (\mathbf{I} - 2\mu_q \mathbf{R}) \boldsymbol{\mu}_v(z) + \frac{\boldsymbol{\mu}_u}{1-z^{-1}} \quad (2.23)$$

$$\boldsymbol{\mu}_v(z) = [\mathbf{I} - (\mathbf{I} - 2\mu_q \mathbf{R}) z^{-1}]^{-1} \frac{z^{-1}}{1-z^{-1}} \boldsymbol{\mu}_u + [\mathbf{I} - (\mathbf{I} - 2\mu_q \mathbf{R}) z^{-1}]^{-1} \boldsymbol{\mu}_v(0) \quad (2.24)$$

Si utilizamos la transformación de similaridad unitaria para descomponer la matriz de correlación definida por (2.25) podemos reescribir (2.24) como (2.26).

$$\mathbf{R} = \mathbf{Q} \boldsymbol{\Lambda} \mathbf{Q}^{-1} \quad (2.25)$$

$$\boldsymbol{\mu}_v(z) = \mathbf{Q} [\mathbf{I} - (\mathbf{I} - 2\mu_q \boldsymbol{\Lambda}) z^{-1}]^{-1} \frac{z^{-1}}{1-z^{-1}} \mathbf{Q}^{-1} \boldsymbol{\mu}_u + \mathbf{Q} [\mathbf{I} - (\mathbf{I} - 2\mu_q \boldsymbol{\Lambda}) z^{-1}]^{-1} \mathbf{Q}^{-1} \boldsymbol{\mu}_v(0) \quad (2.26)$$

Dado que  $\boldsymbol{\Lambda}$  es una matriz diagonal, la inversión de matriz se obtiene de forma directa, cuestión que permite aplicar la transformada  $\mathcal{Z}$  inversa de forma más simple como se establece en (2.27).

$$\begin{aligned} \boldsymbol{\mu}_v(k) = & \mathbf{Q} \mathcal{Z}^{-1} \left\{ [\mathbf{I} - (\mathbf{I} - 2\mu_q \boldsymbol{\Lambda}) z^{-1}]^{-1} \frac{z^{-1}}{1-z^{-1}} \right\} \mathbf{Q}^{-1} \boldsymbol{\mu}_u \\ & + \mathbf{Q} \mathcal{Z}^{-1} \left\{ [\mathbf{I} - (\mathbf{I} - 2\mu_q \boldsymbol{\Lambda}) z^{-1}]^{-1} \right\} \mathbf{Q}^{-1} \boldsymbol{\mu}_v(0) \end{aligned} \quad (2.27)$$

Desarrollando el argumento de la primera transformada  $\mathcal{Z}$  inversa en (2.27) obtenemos (2.28) y al resolver un elemento de la matriz, se puede generalizar el resultado como se muestra en (2.29). Siguiendo el mismo procedimiento para el segundo término en (2.27) obtenemos (2.30).

$$[\mathbf{I} - (\mathbf{I} - 2\mu_q \boldsymbol{\Lambda}) z^{-1}]^{-1} \frac{z^{-1}}{1-z^{-1}} = \begin{bmatrix} \frac{1}{1-(1-2\mu_q \lambda_0) z^{-1}} \frac{z^{-1}}{1-z^{-1}} & 0 & \cdots & 0 \\ 0 & \frac{1}{1-(1-2\mu_q \lambda_1) z^{-1}} \frac{z^{-1}}{1-z^{-1}} & & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & \frac{1}{1-(1-2\mu_q \lambda_{N-1}) z^{-1}} \frac{z^{-1}}{1-z^{-1}} \end{bmatrix} \quad (2.28)$$

$$\mathcal{Z}^{-1}\left\{\left[\mathbf{I} - (\mathbf{I} - 2\mu_q\Lambda)z^{-1}\right]^{-1} \frac{z^{-1}}{1-z^{-1}}\right\} = \frac{1}{2\mu_q}\Lambda^{-1}\left[\mathbf{I} - (\mathbf{I} - 2\mu_q\Lambda)^k\right] \quad (2.29)$$

$$\mathcal{Z}^{-1}\left\{\left[\mathbf{I} - (\mathbf{I} - 2\mu_q\Lambda)z^{-1}\right]^{-1}\right\} = (\mathbf{I} - 2\mu_q\Lambda)^k \quad (2.30)$$

Si sustituimos las ecuaciones (2.29) y (2.30) en (2.27) y  $\mathbf{Q} = \mathbf{Q}^{-1} = \mathbf{I}$  debido a que la entrada es mutuamente independiente y de media cero podemos concluir (2.31).

$$\boldsymbol{\mu}_v(k) = (\mathbf{I} - 2\mu_q\Lambda)^k \left( \boldsymbol{\mu}_v(0) - \frac{1}{2\mu_q}\Lambda^{-1}\boldsymbol{\mu}_u \right) + \frac{1}{2\mu_q}\Lambda^{-1}\boldsymbol{\mu}_u \quad (2.31)$$

Volviendo a la ecuación (2.21) procedemos a multiplicar dicha ecuación por  $\mathbf{R}$  en ambos lados de la igualdad, así mismo se aplica la función traza y se obtiene (2.32).

$$\begin{aligned} \text{tr}[\mathbf{R}\mathbf{K}(k+1)] &= \text{tr}[\mathbf{R}\mathbf{K}(k) - 2\mu_q\mathbf{R}\mathbf{K}(k)\mathbf{R} - 2\mu_q\mathbf{R}^2\mathbf{K}(k) \\ &\quad + 4\mu_q^2\mathbf{R}E[\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{K}(k)\mathbf{x}(k)\mathbf{x}^T(k)] + 4\mu_q^2\sigma_o^2\mathbf{R}^2 + \sigma_u^2\mathbf{R} \\ &\quad + \mu_u^2\mathbf{R}\mathbf{O} + \mathbf{R}\boldsymbol{\mu}_u\boldsymbol{\mu}_v^T(k)(\mathbf{I} - 2\mu_q\mathbf{R}) + \mathbf{R}(\mathbf{I} - 2\mu_q\mathbf{R})\boldsymbol{\mu}_v(k)\boldsymbol{\mu}_u^T] \end{aligned} \quad (2.32)$$

Por cuestiones prácticas los términos independientes del tiempo se agrupan en una sola matriz llamada  $\mathbf{S}$  como se muestra en (2.33) y dado que la función traza es una función lineal utilizamos dicha propiedad para obtener la siguiente ecuación (2.34).

$$\mathbf{S} = 4\mu_q^2\sigma_o^2\mathbf{R}^2 + \sigma_u^2\mathbf{R} + \mu_u^2\mathbf{R}\mathbf{O} \quad (2.33)$$

$$\begin{aligned} \text{tr}[\mathbf{R}\mathbf{K}(k+1)] &= \text{tr}[\mathbf{R}\mathbf{K}(k)] - 2\mu_q\text{tr}[\mathbf{R}\mathbf{K}(k)\mathbf{R}] - 2\mu_q\text{tr}[\mathbf{R}^2\mathbf{K}(k)] \\ &\quad + 4\mu_q^2\text{tr}\left[\mathbf{R}E[\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{K}(k)\mathbf{x}(k)\mathbf{x}^T(k)]\right] \\ &\quad + \text{tr}\left[\mathbf{R}\boldsymbol{\mu}_u\boldsymbol{\mu}_v^T(k)(\mathbf{I} - 2\mu_q\mathbf{R})\right] + \text{tr}\left[\mathbf{R}(\mathbf{I} - 2\mu_q\mathbf{R})\boldsymbol{\mu}_v(k)\boldsymbol{\mu}_u^T\right] + \text{tr}[\mathbf{S}] \end{aligned} \quad (2.34)$$

Aunque sabemos que la multiplicación de matrices no es conmutativa, la función traza permite la permutación cíclica como se establece en (2.35). Otra propiedad importante se detalla en (2.36)

$$\text{tr}[\mathbf{ABC}] = \text{tr}[\mathbf{BCA}] = \text{tr}[\mathbf{CAB}] \quad (2.35)$$

$$\text{tr}[E[\mathbf{A}(k)]] = E[\text{tr}[\mathbf{A}(k)]] \quad (2.36)$$

Utilizando las dos propiedades (2.35) y (2.36) podemos reducir algunos elementos de la ecuación (2.34) para obtener (2.37) y (2.38).

$$\text{tr}[\mathbf{R}\mathbf{K}(k)\mathbf{R}] = \text{tr}[\mathbf{R}^2\mathbf{K}(k)] \quad (2.37)$$

$$\begin{aligned} \text{tr}[\mathbf{R}E[\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{K}(k)\mathbf{x}(k)\mathbf{x}^T(k)]] &= \text{tr}[E[\mathbf{R}\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{K}(k)\mathbf{x}(k)\mathbf{x}^T(k)]] \\ &= E[\text{tr}[\mathbf{R}\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{K}(k)\mathbf{x}(k)\mathbf{x}^T(k)]] \\ &= E[\text{tr}[\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{R}\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{K}(k)]] \\ &= \text{tr}[E[\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{R}\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{K}(k)]] \end{aligned} \quad (2.38)$$

Analizando el término  $E[\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{R}\mathbf{x}(k)\mathbf{x}^T(k)]$  en forma separada y considerando que las muestras de son mutuamente independientes podemos establecer (2.39) y (2.40) donde  $x^2(k-\ell)$  y  $\lambda_\ell = E[x^2(k-\ell)]$  representan elementos diagonales de las matrices  $\mathbf{P} = \mathbf{x}(k)\mathbf{x}^T(k)$  y  $\mathbf{R}$  respectivamente.

$$\mathbf{x}^T(k)\mathbf{R}\mathbf{x}(k) = \text{tr}[\mathbf{R}\mathbf{P}] = \sum_{\ell=0}^{N-1} x^2(k-\ell)\lambda_\ell \quad (2.39)$$

$$E[\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{R}\mathbf{x}(k)\mathbf{x}^T(k)] = E[\text{tr}[\mathbf{R}\mathbf{P}]\mathbf{P}] \quad (2.40)$$

Por otro lado, podemos descomponer la matriz  $\mathbf{P}$  como se muestra en (2.41), donde  $\text{diag}[x^2(k), x^2(k-1), \dots, x^2(k-N+1)]$  representa la matriz compuesta por los elementos diagonales de  $\mathbf{P}$ , y  $\mathbf{U}$  representa la matriz con los elementos restantes.

$$\mathbf{P} = \text{diag}[x^2(k), x^2(k-1), \dots, x^2(k-N+1)] + \mathbf{U} \quad (2.41)$$

Sustituyendo la ecuación (2.41) en (2.40) obtenemos (2.42), y evaluando los elementos de la matriz  $\mathbf{T} = E[\text{tr}[\mathbf{R}\mathbf{P}]\mathbf{U}]$  podemos obtener (2.43) y en consecuencia (2.44).

$$E[\text{tr}[\mathbf{R}\mathbf{P}]\mathbf{P}] = E[\text{tr}[\mathbf{R}\mathbf{P}]\text{diag}[x^2(k), x^2(k-1), \dots, x^2(k-N+1)] + \text{tr}[\mathbf{R}\mathbf{P}]\mathbf{U}] \quad (2.42)$$

$$\begin{aligned} t_{ii} &= 0 \quad i = 0, 1, \dots, N-1 \\ t_{ij} &= E[x(k-i)x(k-j)\text{tr}[\mathbf{R}\mathbf{P}]] = 0 \quad i \neq j \end{aligned} \quad (2.43)$$

$$\mathbf{T} = E[\text{tr}[\mathbf{R}\mathbf{P}]\mathbf{U}] = \mathbf{0} \quad (2.44)$$

Evaluando los elementos de la matriz  $\mathbf{F} = E[\text{tr}[\mathbf{R}\mathbf{P}]\text{diag}[x^2(k), x^2(k-1), \dots, x^2(k-N+1)]]$ , y utilizando (2.39) obtenemos (2.45), donde  $E[x^4(k-i)] = (\gamma_2 + 3)\lambda_i$  y  $u(\ell \neq i) = u(\ell) - \delta(\ell - i)$  es la función escalón que es cero únicamente cuando  $\ell = i$  y  $\ell < 0$ . De esta manera del resultado de la ecuación (2.45) podemos concluir (2.46), donde  $\gamma_2$  es conocido como el exceso de kurtosis el cual es un valor fijo que depende del tipo de distribución de la señal de entrada; y finalmente podemos reescribir (2.38) como (2.47).

$$\begin{aligned}
 f_{ii} &= E\left[x^2(k-i)\sum_{\ell=0}^{N-1}x^2(k-\ell)\lambda_{\ell}\right] \\
 &= E\left[x^4(k-i)\lambda_i + x^2(k-i)\sum_{\ell=0}^{N-1}x^2(k-\ell)\lambda_{\ell}u(\ell \neq i)\right] \\
 &= (\gamma_2 + 3)\lambda_i^3 + \lambda_i\sum_{\ell=0}^{N-1}\lambda_{\ell}^2u(\ell \neq i) \\
 &= (\gamma_2 + 2)\lambda_i^3 + \lambda_i\sum_{\ell=0}^{N-1}\lambda_{\ell}^2 \quad i = 0, 1, \dots, N-1
 \end{aligned} \tag{2.45}$$

$$E[\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{R}\mathbf{x}(k)\mathbf{x}^T(k)] = (\gamma_2 + 2)\mathbf{R}^3 + \mathbf{R}\text{tr}[\mathbf{R}^2] \tag{2.46}$$

$$\text{tr}[\mathbf{R}E[\mathbf{x}(k)\mathbf{x}^T(k)\mathbf{K}(k)\mathbf{x}(k)\mathbf{x}^T(k)]] = \text{tr}[(\gamma_2 + 2)\mathbf{R}^2 + \text{tr}[\mathbf{R}^2]\mathbf{I}]\mathbf{R}\mathbf{K}(k) \tag{2.47}$$

Usando propiedades de simetría de la matriz de correlación, así como la transpuesta y permutación cíclica podemos manipular el quinto término de (2.34) y obtener (2.48).

$$\begin{aligned}
 \text{tr}[\mathbf{R}\boldsymbol{\mu}_u\boldsymbol{\mu}_v^T(k)(\mathbf{I} - 2\mu_q\mathbf{R})] &= \text{tr}[(\mathbf{R}\boldsymbol{\mu}_u\boldsymbol{\mu}_v^T(k)(\mathbf{I} - 2\mu_q\mathbf{R}))^T] \\
 &= \text{tr}[(\mathbf{I} - 2\mu_q\mathbf{R})^T(\boldsymbol{\mu}_u\boldsymbol{\mu}_v^T(k))^T\mathbf{R}^T] \\
 &= \text{tr}[\mathbf{R}(\mathbf{I} - 2\mu_q\mathbf{R})\boldsymbol{\mu}_v(k)\boldsymbol{\mu}_u^T]
 \end{aligned} \tag{2.48}$$

Habiendo simplificado los términos de las ecuaciones (2.37), (2.47) y (2.48) los sustituimos en la ecuación (2.34) para obtener (2.49), la cual su transformada  $\mathcal{Z}$  se muestra en (2.50) donde  $\mathbf{K}(0) = \mathbf{w}_o\mathbf{w}_o^T$ .

$$\begin{aligned}
 \text{tr}[\mathbf{R}\mathbf{K}(k+1)] &= \text{tr}[\mathbf{R}\mathbf{K}(k)] - 4\mu_q\text{tr}[\mathbf{R}^2\mathbf{K}(k)] + 4\mu_q^2\text{tr}[(\gamma_2 + 2)\mathbf{R}^2 + \text{tr}[\mathbf{R}^2]\mathbf{I}]\mathbf{R}\mathbf{K}(k) \\
 &\quad + 2\text{tr}[\mathbf{R}(\mathbf{I} - 2\mu_q\mathbf{R})\boldsymbol{\mu}_v(k)\boldsymbol{\mu}_u^T] + \text{tr}[\mathbf{S}]
 \end{aligned} \tag{2.49}$$

$$\begin{aligned} \text{tr}[\mathbf{R}(z\mathbf{K}(z) - z\mathbf{K}(0))] &= \text{tr}[\mathbf{R}\mathbf{K}(z)] - 4\mu_q \text{tr}[\mathbf{R}^2\mathbf{K}(z)] \\ &+ 4\mu_q^2 \text{tr}\left[\left((\gamma_2 + 2)\mathbf{R}^2 + \text{tr}[\mathbf{R}^2]\mathbf{I}\right)\mathbf{R}\mathbf{K}(z)\right] \\ &+ 2 \text{tr}\left[\mathbf{R}(\mathbf{I} - 2\mu_q\mathbf{R})\boldsymbol{\mu}_v(z)\boldsymbol{\mu}_u^T\right] + \frac{\text{tr}[\mathbf{S}]}{1 - z^{-1}} \end{aligned} \quad (2.50)$$

Ahora (2.51) se obtiene agrupando términos comunes y considerando que  $\mathbf{R} = \sigma_x^2\mathbf{I}$  podemos además reducir a (2.52).

$$\begin{aligned} \text{tr}\left\{\left[z\mathbf{I} - \mathbf{I} + 4\mu_q\mathbf{R} - 4\mu_q^2\left((\gamma_2 + 2)\mathbf{R}^2 + \text{tr}[\mathbf{R}^2]\mathbf{I}\right)\right]\mathbf{R}\mathbf{K}(z)\right\} &= \text{tr}[z\mathbf{R}\mathbf{K}(0)] \\ &+ 2 \text{tr}\left[\mathbf{R}(\mathbf{I} - 2\mu_q\mathbf{R})\boldsymbol{\mu}_v(z)\boldsymbol{\mu}_u^T\right] + \frac{\text{tr}[\mathbf{S}]}{1 - z^{-1}} \end{aligned} \quad (2.51)$$

$$\begin{aligned} \left[z - (1 - 4\mu_q\sigma_x^2 + 4\mu_q^2\sigma_x^4(N + \gamma_2 + 2))\right]\text{tr}[\mathbf{R}\mathbf{K}(z)] &= \text{tr}[z\mathbf{R}\mathbf{K}(0)] \\ &+ 2\sigma_x^2(1 - 2\mu_q\sigma_x^2)\text{tr}\left[\boldsymbol{\mu}_v(z)\boldsymbol{\mu}_u^T\right] + \frac{\text{tr}[\mathbf{S}]}{1 - z^{-1}} \end{aligned} \quad (2.52)$$

Despejando  $\text{tr}[\mathbf{R}\mathbf{K}(z)]$  y sustituyendo  $\alpha = 1 - 4\mu_q\sigma_x^2 + 4\mu_q^2\sigma_x^4(N + \gamma_2 + 2)$  obtenemos (2.53). La transformada  $\mathcal{Z}$  inversa de (2.53) es (2.54).

$$\text{tr}[\mathbf{R}\mathbf{K}(z)] = \frac{z}{z - \alpha} \text{tr}[\mathbf{R}\mathbf{K}(0)] + \frac{2\sigma_x^2(1 - 2\mu_q\sigma_x^2)}{z - \alpha} \text{tr}[\boldsymbol{\mu}_v(z)\boldsymbol{\mu}_u^T] + \frac{z \text{tr}[\mathbf{S}]}{(z - \alpha)(z - 1)} \quad (2.53)$$

$$\begin{aligned} \text{tr}[\mathbf{R}\mathbf{K}(k)] &= \text{tr}[\mathbf{R}\mathbf{K}(0)] \mathcal{Z}^{-1}\left\{\frac{z}{z - \alpha}\right\} + 2\sigma_x^2(1 - 2\mu_q\sigma_x^2) \mathcal{Z}^{-1}\left\{\frac{\text{tr}[\boldsymbol{\mu}_v(z)\boldsymbol{\mu}_u^T]}{z - \alpha}\right\} \\ &+ \text{tr}[\mathbf{S}] \mathcal{Z}^{-1}\left\{\frac{z}{(z - \alpha)(z - 1)}\right\} \end{aligned} \quad (2.54)$$

La primera y tercer transformada del lado derecho de (2.54) pueden ser obtenidas de forma simple como se muestra en (2.55) y (2.56), respectivamente.

$$\mathcal{Z}^{-1}\left\{\frac{z}{z - \alpha}\right\} = \alpha^k u(k) \quad (2.55)$$

$$\mathcal{Z}^{-1}\left\{\frac{z}{(z - \alpha)(z - 1)}\right\} = \frac{1 - \alpha^k}{1 - \alpha} u(k) \quad (2.56)$$

Para obtener la transformada inversa del segundo término en (2.54) recurrimos a la ecuación (2.26) actualizándola a las condiciones presentes obtenemos (2.57).

$$\mathbf{\mu}_v(z) = \frac{1}{[1 - (1 - 2\mu_q\sigma_x^2)z^{-1}]} \frac{z^{-1}}{1 - z^{-1}} \mathbf{\mu}_u + \frac{1}{1 - (1 - 2\mu_q\sigma_x^2)z^{-1}} \mathbf{\mu}_v(0) \quad (2.57)$$

Si:  $\beta = 1 - 2\mu_q\sigma_x^2$ ,  $a^{-1} = (\alpha - \beta)(\alpha - 1)$ ,  $b^{-1} = (\beta - \alpha)(\beta - 1)$  y  $c^{-1} = (1 - \alpha)(1 - \beta)$ ; entonces se establece (2.58) a (2.60).

$$\text{tr}[\mathbf{\mu}_v(z)\mathbf{\mu}_u^T] = \frac{z}{(z - \beta)(z - 1)} N\mu_u^2 + \frac{z}{z - \beta} \text{tr}[\mathbf{\mu}_v(0)\mathbf{\mu}_u^T] \quad (2.58)$$

$$\mathcal{Z}^{-1}\left\{\frac{\text{tr}[\mathbf{\mu}_v(z)\mathbf{\mu}_u^T]}{z - \alpha}\right\} = \mathcal{Z}^{-1}\left\{\frac{z}{(z - \alpha)(z - \beta)(z - 1)} N\mu_u^2 + \frac{z}{(z - \alpha)(z - \beta)} \text{tr}[\mathbf{\mu}_v(0)\mathbf{\mu}_u^T]\right\} \quad (2.59)$$

$$\mathcal{Z}^{-1}\left\{\frac{z}{(z - \alpha)(z - \beta)(z - 1)}\right\} = (a\alpha^k + b\beta^k + c)u(k) \quad (2.60)$$

Ahora procedemos a obtener la transformada  $\mathcal{Z}$  inversa como se establece en (2.61) donde  $d^{-1} = (\alpha - \beta)$  y  $e^{-1} = (\beta - \alpha)$  y sustituimos en (2.54) para obtener (2.62) y agrupando términos comunes concluimos (2.63).

$$\mathcal{Z}^{-1}\left\{\frac{z}{(z - \alpha)(z - \beta)}\right\} = (d\alpha^k + e\beta^k)u(k) \quad (2.61)$$

$$\begin{aligned} \text{tr}[\mathbf{RK}(k)] &= \alpha^k \text{tr}[\mathbf{RK}(0)] + 2\sigma_x^2\beta \left\{ N\mu_u^2(a\alpha^k + b\beta^k + c) + \text{tr}[\mathbf{\mu}_v(0)\mathbf{\mu}_u^T](d\alpha^k + e\beta^k) \right\} \\ &\quad + \frac{(1 - \alpha^k)\text{tr}[\mathbf{S}]}{(1 - \alpha)} \end{aligned} \quad (2.62)$$

$$\begin{aligned} \text{tr}[\mathbf{RK}(k)] &= \left[ \text{tr}[\mathbf{RK}(0)] + 2N\mu_u^2\sigma_x^2\beta a + 2\sigma_x^2\beta d \text{tr}[\mathbf{\mu}_v(0)\mathbf{\mu}_u^T] - \frac{\text{tr}[\mathbf{S}]}{1 - \alpha} \right] \alpha^k \\ &\quad + \left[ 2N\mu_u^2\sigma_x^2\beta b + 2\sigma_x^2\beta e \text{tr}[\mathbf{\mu}_v(0)\mathbf{\mu}_u^T] \right] \beta^k + \frac{\text{tr}[\mathbf{S}]}{1 - \alpha} + 2N\sigma_x^2\mu_u^2\beta c \end{aligned} \quad (2.63)$$

Simplificando (2.63) podemos finalmente obtener (2.64), donde de (2.65) a (2.68) se describen sus componentes.

$$\text{tr}[\mathbf{RK}(k)] = (J_C + J_T)\alpha^k + J'_T\beta^k + E \quad (2.64)$$

$$J_C = \sigma_x^2 \sum_{i=0}^{N-1} k_{ii}(0) + \frac{N\sigma_x^2 (4\mu_q^2 \sigma_o^2 \sigma_x^2 + \sigma_u^2)}{\alpha - 1} \quad (2.65)$$

$$J_T = \frac{\left[ N(\alpha + \beta)\mu_u + 2\beta(\alpha - 1) \sum_{i=0}^{N-1} \mu_{v_i}(0) \right] \sigma_x^2 \mu_u}{\alpha^2 - (\beta + 1)\alpha + \beta} \quad (2.66)$$

$$J'_T = \frac{2 \left[ N\mu_u + (\beta - 1) \sum_{i=0}^{N-1} \mu_{v_i}(0) \right] \sigma_x^2 \beta \mu_u}{\beta^2 - (\alpha + 1)\beta + \alpha} \quad (2.67)$$

$$E = \frac{\mu_q (\sigma_n^2 + \sigma_e^2) N \sigma_x^2}{1 - \mu_q \sigma_x^2 (\gamma_2 + 2 + N)} + \frac{N \sigma_u^2}{4\mu_q (1 - \mu_q \sigma_x^2 (\gamma_2 + 2 + N))} \quad (2.68)$$

$$+ \frac{N \mu_u^2 (1 - \mu_q \sigma_x^2)}{4\mu_q^2 \sigma_x^2 (1 - \mu_q \sigma_x^2 (\gamma_2 + 2 + N))}$$

Hasta ahora se analizado el problema asumiendo que la matriz de correlación es constante durante todo el análisis; sin embargo, es común asumir que el vector de inicio es puesto a ceros y por lo tanto se requiere  $N$  iteraciones para que este lleno de datos y considerar la matriz de correlación constate, para este caso la condiciones iniciales cambian, así mismo la ecuación solo es valida para  $k \geq N-1$  entonces la ecuación (2.64) puede ser rescrita como (2.69) y las condiciones iniciales se aproximan mediante (2.70) y (2.71).

$$\text{tr}[\mathbf{RK}(k)] = \left[ (J_C + J_T)\alpha^{k-N+1} + J'_T\beta^{k-N+1} + E \right] u(k - N + 1) \quad (2.69)$$

$$\mathbf{\mu}_v(N-1) \approx \left[ -(1 - 2\mu_q \sigma_x^2)^{N-1} w_{o0} \quad -(1 - 2\mu_q \sigma_x^2)^{N-2} w_{o1} \quad \dots \quad -(1 - 2\mu_q \sigma_x^2) w_{oN-1} \right]^T \quad (2.70)$$

$$\sum_{i=0}^{N-1} k_{ii}(N-1) = \mathbf{\mu}_v^T(N-1) \mathbf{\mu}_v(N-1) \quad (2.71)$$

Finalmente sustituyendo (2.64) en (2.15) podemos concluir que la curva de aprendizaje puede ser modelado como se establece en (2.72).

$$\xi(k) = (J_C + J_T)\alpha^k + J'_T \beta^k + E + \sigma_o^2 \quad (2.72)$$

Para el caso en que el vector de la señal de entrada se ponga a ceros solo se sustituye la ecuación (2.69) en la ecuación (2.15).

### 2.3 *Análisis de estabilidad*

La estabilidad del filtro es sin duda un elemento importante en cualquier proceso recursivo, así mismo es un criterio de diseño imprescindible dentro de los filtros adaptivos pues la rapidez de convergencia está directamente asociada a la estabilidad. A continuación se mostrará el análisis de estabilidad del modelo propuesto en la sección anterior.

Revisando la ecuación (2.72) podemos afirmar que el algoritmo será estable cuando se cumpla lo (2.73) y (2.74).

$$-1 < \alpha < 1 \quad (2.73)$$

$$-1 < \beta < 1 \quad (2.74)$$

Si  $\alpha = 1 - 4\mu_q\sigma_x^2 + 4\mu_q^2\sigma_x^4(N + \gamma_2 + 2)$  y  $\beta = 1 - 2\mu_q\sigma_x^2$  entonces podemos establecer (2.75) de (2.73). Dado que  $\alpha$  es una función cuadrática en la cual la variable de interés es  $\mu_q$  procedemos a evaluar en el límite superior como se establece en (2.76).

$$-1 < 1 - 4\mu_q\sigma_x^2 + 4\mu_q^2\sigma_x^4(N + \gamma_2 + 2) < 1 \quad (2.75)$$

$$\begin{aligned} 1 - 4\mu_q\sigma_x^2 + 4\mu_q^2\sigma_x^4(N + \gamma_2 + 2) &= 1 \\ 4\sigma_x^2\mu_q[(N + \gamma_2 + 2)\sigma_x^2\mu_q - 1] &= 0 \end{aligned} \quad (2.76)$$

De (2.76) obtenemos dos soluciones (2.77) y (2.78). El rango esta dado por (2.79).

$$\mu_{q1} = 0 \quad (2.77)$$

$$\mu_{q2} = \frac{1}{(N + \gamma_2 + 2)\sigma_x^2} \quad (2.78)$$

$$0 < \mu_q < \frac{1}{(N + \gamma_2 + 2)\sigma_x^2} \quad (2.79)$$

Si se evalúa en el límite inferior las soluciones son complejas, por lo que se descartan debido a que solo estamos considerando el caso de valores reales.

Dado que  $\alpha$  no es una función lineal, entonces es importante conocer el valor que minimiza la función la cual está dada por (2.80).

$$\begin{aligned}\frac{d\alpha}{d\mu_q} &= -4\sigma_x^2 + 8\mu_q\sigma_x^4(N + \gamma_2 + 2) \\ 0 &= -4\sigma_x^2 + 8\mu_q\sigma_x^4(N + \gamma_2 + 2) \\ \mu_q &= \frac{1}{2\sigma_x^2(N + \gamma_2 + 2)}\end{aligned}\tag{2.80}$$

Para la desigualdad en (2.74) podemos fácilmente obtener (2.81).

$$\begin{aligned}-1 &< 1 - 2\mu_q\sigma_x^2 < 1 \\ -2 &< -2\mu_q\sigma_x^2 < 0 \\ 0 &< \mu_q < \frac{1}{\sigma_x^2}\end{aligned}\tag{2.81}$$

Teóricamente los límites de estabilidad del algoritmo serían restringidos por las desigualdades en las ecuaciones (2.79) y (2.81); sin embargo, en la realidad existen algunos efectos que hacen que el comportamiento real del algoritmo y el obtenido en la ecuación (2.72) no concuerden, una de las causas es debida al fenómeno conocido como frenado [5] en el cual durante el comportamiento transitorio el término de actualización llega a cero debido a que este término es muy pequeño para ser cuantizado apropiadamente, por lo que el algoritmo se frena de tal manera que el ruido actualiza los valores del filtro, por lo tanto para prevenir dicho problema entonces establecemos que

$$|2\mu_q e_o(k)x(k)| > 2^{-B}\tag{2.82}$$

donde  $B = b_c$  cuando se utiliza truncamiento y  $B = b_c + 1$  para redondeo; al elevar al cuadrado y aplicar la esperanza a (2.82) obtenemos (2.83) y (2.84).

$$4\mu_q^2\sigma_o^2\sigma_x^2 > 2^{-2B}\tag{2.83}$$

$$\mu_q > \frac{2^{-B}}{2\sigma_x\sqrt{\sigma_n^2 + \sigma_e^2}}\tag{2.84}$$

En (2.84) se establece el factor de convergencia mínimo necesario para evitar el fenómeno de frenado.

Otro aspecto que debe considerarse, es cuando los valores de  $\mu_q$  son grandes, en estos casos es común experimentar cambios bruscos en el comportamiento de estado estable, así como variaciones pequeñas o considerables en el comportamiento transitorio con respecto al teórico. Estas diferencias se atribuyen a que las suposiciones de independencia utilizadas en el desarrollo teórico no son válidas para valores grandes de  $\mu_q$  [6], así mismo se puede ver en la ecuación (2.68) que el exceso de error estable no es una función lineal y que es extremadamente sensible a los parámetros de  $\mu_q$  y  $\sigma_x^2$ . Una forma de ver claro lo anterior es definiendo la sensibilidad  $S_{\sigma_x^2}^E$  como el cambio en el exceso de error de estado estable a causa de un cambio en la varianza de la señal de entrada, el cual es definido en (2.85) y cuando es aplicado a (2.68) podemos obtener (2.86).

$$S_{\sigma_x^2}^E = \frac{\sigma_x^2}{E} \frac{\partial E}{\partial \sigma_x^2} \quad (2.85)$$

$$S_{\sigma_x^2}^E = \frac{1}{1 + \left[ \frac{\mu_q \sigma_x^2 (\sigma_u^2 - \mu_u^2)}{4 \mu_q^2 \sigma_o^2 \sigma_x^4 + \mu_u^2} \right]} + \frac{\mu_q (N + \gamma_2 + 2) \sigma_x^2}{1 - \mu_q \sigma_x^2 (N + \gamma_2 + 2)} \quad (2.86)$$

Si los valores de  $\sigma_u^2$  y  $\mu_u^2$  son valores muy pequeños en la mayoría de las aplicaciones podemos normalizar (2.86) como se muestra en (2.87), donde  $n$  es definida en (2.87),  $0 < n < 1$ , y  $\sigma_x^2$  es constante.

$$S_{\sigma_x^2}^E = 1 + \frac{n}{1 - n} \quad (2.87)$$

$$\mu_q = \frac{n}{\sigma_x^2 (N + \gamma_2 + 2)} \quad (2.88)$$

La grafica de la ecuación (2.87) se muestra en Fig. 2.2a y se comprueba que para valores grandes de  $\mu_q$  la sensibilidad crece rápidamente y es la causa de la inestabilidad del filtro en estado estable pues pequeños cambios en la estadística de la señal de entrada producen cambios grandes en el exceso de error de estado estable, por lo tanto, generalmente se descarta utilizar valores grandes  $\mu_q$ , lo anterior pudiera interpretarse como algo que afecta drásticamente a la rapidez de convergencia del algoritmo; sin embargo, dado que  $\alpha$  es una función cuadrática y por lo tanto simétrica alrededor al valor encontrado en (2.80) en donde se tiene la máxima rapidez de convergencia, se puede obtener convergencias similares utilizando valores  $n \leq 0.5$ , otra ventaja de utilizar dichos valores es que el error de estado estable es menor ya que el desajuste es menor al 100%.

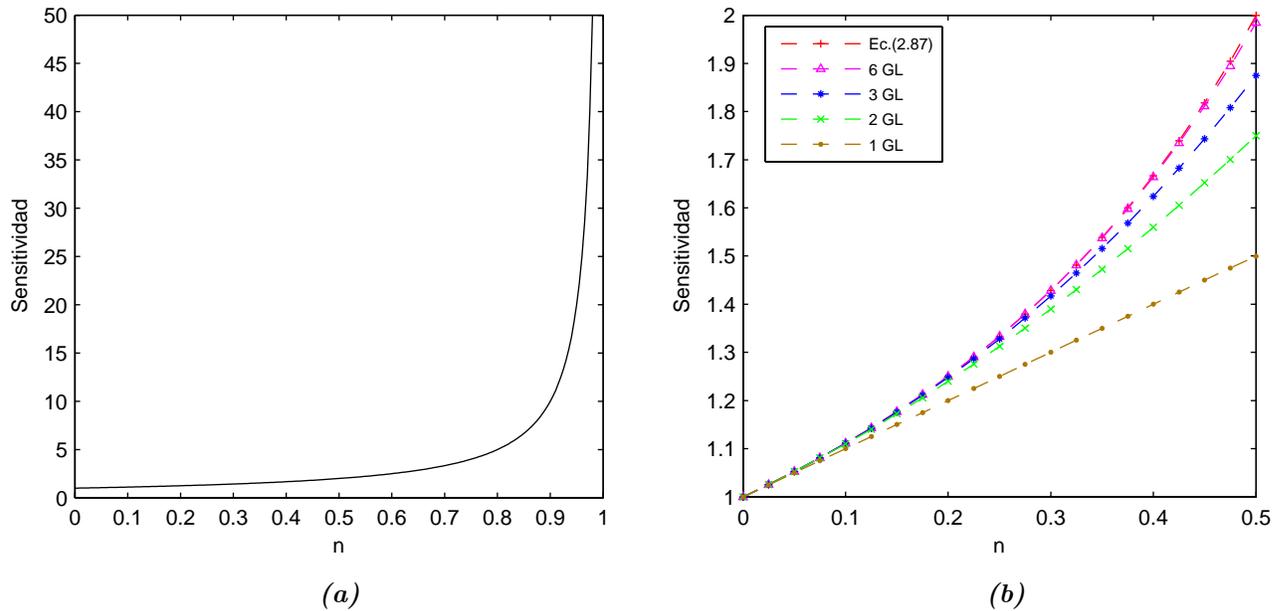


Fig. 2.2 (a) Gráfica de sensibilidad  $S_{\sigma_x^2}^E$ , (b) Comparación de (2.87) con varias aproximaciones de (2.89).

Hasta ahora hemos visto como el rango de operación del algoritmo ha sido cada vez más restrictivo y encontrar un límite superior a este rango resulta una tarea difícil, pues aunque para  $n \leq 0.5$  la sensibilidad es mucho menor, la variabilidad estadística de la señal de entrada puede causar picos en valores cercanos a  $n = 0.5$ , los cuales pueden ser indeseables en muchas aplicaciones. En consecuencia, es imperante establecer criterios más restrictivos, uno de ellos sería mediante el máximo desajuste permitido o encontrando una región de seguridad. Si ahora analizamos la ecuación (2.86) pero bajo la condición de que lo que varía es  $\sigma_x^2$  y lo que se mantiene constante es  $\mu_q$  podemos llegar a una ecuación similar a (2.87), donde  $n$  ahora representa la variabilidad de  $\sigma_x^2$  y no de  $\mu_q$ , por lo que podemos remitirnos a la misma Fig. 2.2a en la cual podemos interpretarla de la siguiente manera: si se está trabajando con valores grandes de  $n$  al ocurrir un cambio positivo en la estadística, el exceso de error aumentará y a la vez el lugar de trabajo por lo que la sensibilidad aumentará en forma no lineal; ante esta situación, sin duda trabajar en la región lo más lineal posible es muy importante para asegurar la estabilidad pues ante los cambios de la estadística estos tendrán un efecto menor al cambiar el punto de trabajo. Para encontrar dicha zona lineal podemos aproximar la ecuación (2.87) como una serie de potencias (2.89) donde  $\nu$  representa el grado de libertad.

$$S_{\sigma_x^2}^E \approx 1 + n + n^2 + n^3 \dots + n^\nu \quad (2.89)$$

En la Figura 2.2b se muestra la ecuación (2.87) con varias curvas de diferente grado de libertad. Como se puede ver, al aumentar el grado de libertad la curva de la ecuación (2.89) se aproxima a la curva real en el rango  $0 < n < 0.5$  que es ahora nuestra región de trabajo, así mismo el rango que se puede considerar lineal se hace menos exigente. Una forma muy común para medir el grado de ajuste de una curva es mediante el *RMSE* (*Root Mean Square Error*), mediante esto podemos estimar cuantos grados de libertad son necesarios para ajustar la curva en el rango  $0 < n \leq 0.5$ . En nuestro caso particular podemos definir el *RMSE* como se establece en (2.90).

$$RMSE = \sqrt{MSE} = \sqrt{2 \int_0^{0.5} \varepsilon^2(n) dn} \quad (2.90)$$

donde

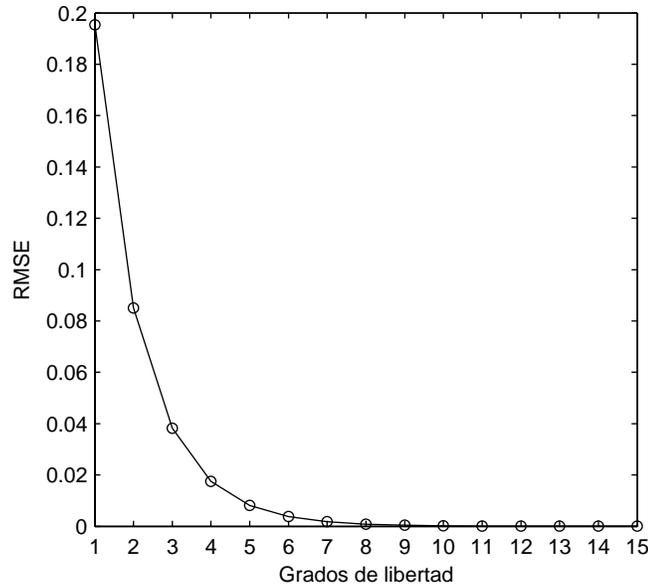
$$\varepsilon(n) = \frac{n^{\nu+1}}{(1-n)} \quad (2.91)$$

el cual es el valor residuo de la expansión realizada en la ecuación (2.89) y que además representa el error entre la curva original y la expansión de series.

Resolviendo la integral en (2.90) obtenemos (2.92) podemos obtener el *RMSE* de cualquier valor de  $\nu$  y así la Fig. 2.3 en la cual es evidente que con al menos  $\nu = 10$  se obtiene un ajuste óptimo de la curva, así mismo representa que existen al menos 10 grados de linealidad en el intervalo  $0 < n \leq 0.5$ , y que podemos encontrar dichos puntos utilizando la ecuación (2.91), ya que podemos encontrar el error óptimo en el punto máximo (2.93) y utilizar este mismo error para encontrar los límites restantes.

$$RMSE = \sqrt{2 \sum_{i=0}^{2\nu} \frac{(i+1)(0.5)^{2\nu-i+1}}{(2\nu-i+1)} - 4 \ln(2)(\nu+1) + 2} \quad (2.92)$$

$$\varepsilon_{opt} = \frac{0.5^{11}}{0.5} = 9.765625 \times 10^{-4} \quad (2.93)$$



**Fig. 2.3.** Variación del RMSE al aumentar  $\nu$ .

Ahora la ecuación (2.91) se convierte en una ecuación de orden  $\nu+1$  como se muestra en (2.94). Al resolver para los distintos valores de  $\nu$  obtenemos la tabla 2.2.

$$n^{\nu+1} + \varepsilon_{opt}n - \varepsilon_{opt} = 0 \tag{2.94}$$

**Tabla 2.2**  
*Límites lineales con  $\nu$  grados de libertad.*

	$\nu$									
	1	2	3	4	5	6	7	8	9	10
$n$	0.0308	0.0959	0.1688	0.2368	0.2970	0.3494	0.3949	0.4345	0.4693	0.5
$n_\nu = 1/n$	32.5	10.424	5.9244	4.2222	3.3669	2.8623	2.5325	2.3014	2.1308	2

Algunos de los límites mostrados en la tabla 2.2 son fácilmente verificables comparando con la Fig. 2.2b, así mismo se puede ver como al trabajar en una región estrictamente lineal el rango de operación se vuelve extremadamente reducido; sin embargo, en la mayoría de la veces se puede tolerar cierto grado de no linealidad aumentando así el rango de operación. En relación al rango propuesto en [6,9] podemos interpretar diciendo que se sugiere trabajar con aproximadamente 6 grados de libertad. Sin embargo, si se considera el límite superior propuesto inicialmente por Horowitz y Senne [10], y después utilizado como base del desarrollo de [9], bajo las condiciones establecidas en este trabajo, el rango de operación en realidad coincide con (2.78), lo cual es substancialmente diferente del resultado final obtenido en [9].

Finalmente, utilizando (2.84) y (2.88) podemos concluir que el rango de trabajo esta dado por (2.95) donde  $n_\nu$  representa el límite lineal con  $\nu$  grados de libertad.

$$\frac{2^{-B}}{2\sigma_x\sqrt{\sigma_n^2 + \sigma_e^2}} < \mu_q \leq \frac{1}{n_v(N + \gamma_2 + 2)\sigma_x^2} \quad (2.95)$$

Las variaciones estadísticas de la señal de entrada no solo afectan al comportamiento de estado estable, sino también al transitorio; sin embargo, al trabajar en dentro del rango propuesto en (2.95), la convergencia del comportamiento transitorio está garantizada, solo que pueden existir variaciones con respecto a lo estimado teóricamente en la ecuación (2.72) en donde las simulaciones muestran que la rapidez de convergencia puede ser mayor en ciertas regiones del rango de operación con respecto a lo estimado teóricamente.

Finalmente podemos resumir que el modelo matemático propuesto para el análisis de los efectos de la palabra finita en la curva de aprendizaje utilizando el modelo estadístico de la Fig. 2.1, es descrito básicamente por las ecuaciones (2.72) y (2.95). En el presente modelo se considera la cuantización por redondeo así como por truncamiento, así mismo se asume que entrada estacionaria en sentido amplio, real, media cero, y los elementos del vector de la señal de entrada son mutuamente independientes.

Tal como se mencionó anteriormente, la caracterización de la curva de aprendizaje puede ser obtenida de acuerdo a dos casos típicos de inicialización del vector de entrada:

1. Si el vector de inicio de la señal de entrada esta formado por elementos de ruido

$$\begin{aligned} \xi(k) &= (J_C + J_T)\alpha^k + J_T' \beta^k + E + \sigma_o^2 \\ \mu_v(0) &= -\mathbf{w}_0 \end{aligned} \quad (2.96)$$

2. Si el vector de inicio de la señal de entrada es puesto a ceros

$$\xi(k) = \left[ (J_C + J_T)\alpha^{k-N+1} + J_T' \beta^{k-N+1} + E + \sigma_o^2 \right] u(k-N+1) \quad (2.97)$$

$$\mu_v(N-1) \approx \left[ -(1-2\mu_q\sigma_x^2)^{N-1} w_{o0} \quad -(1-2\mu_q\sigma_x^2)^{N-2} w_{o1} \quad \dots \quad -(1-2\mu_q\sigma_x^2) w_{oN-1} \right]^T$$

La curva de aprendizaje es válida dentro del rango variable mostrado a continuación

$$\frac{2^{-B}}{2\sigma_x\sqrt{\sigma_n^2 + \sigma_e^2}} < \mu_q \leq \frac{1}{n_v(N + \gamma_2 + 2)\sigma_x^2} \quad (2.98)$$

donde  $n_v$  es obtenido a través de la tabla 2.

Como se muestra, el modelo contempla una referencia completa de curva de aprendizaje que incluye diversos aspectos prácticos del filtro adaptivo LMS, los cuales proporcionan información importante al diseñador tal como las variables que intervienen en el comportamiento transitorio y de estado estable, así como la interdependencia de ambos comportamientos; por otro lado se puede identificar claramente las aportaciones debidas al error por truncamiento y redondeo en el comportamiento de la curva y su comparación con la de precisión infinita.

En la mayoría de los trabajos previos el análisis se limita básicamente al estudio de comportamiento transitorio, estado estable o rango de convergencia; esto sin duda no permite la uniformidad de resultados, pues los modelos propuestos generalmente no son los mismos, así mismo ciertas condiciones en un modelo pueden resultar no válidas en otro, de tal suerte que no se puede tener certidumbre de que la caracterización completa del sistema es confiable al momento de diseñar.

Uno de los aspectos poco tratados en la literatura es el análisis de los efectos finitos de palabra es el comportamiento transitorio, pues como es mostrado en [3,6,8] este se aproxima generalmente mediante curvas de decaimiento geométrico (2.99) y (2.100) las cuales son derivadas del comportamiento transitorio del algoritmo steepest-descent.

$$r = e^{-k/\tau} \quad (2.99)$$

$$\tau = \frac{1}{4\mu\lambda} \quad (2.100)$$

Este método solo es válido cuando se desea aproximar el comportamiento transitorio con precisión infinita; sin embargo, comparando con (2.64) dependiendo del tipo de cuantización utilizada el comportamiento transitorio puede variar en forma considerable en relación al estimado de precisión infinita.

Por otro lado la interdependencia entre los comportamientos de la curva de aprendizaje es sin duda vital para el entendimiento general de la curva de aprendizaje, así como establecer criterios de convergencia más confiables, ya que es una de las partes más sensitivas en el diseño de un filtro adaptivo. La mayoría de las veces se acude al conocimiento empírico del sistema así como a la simulación extensiva para establecer rangos de operación que estimen de mejor manera dicho intervalo en comparación a las propuestas teóricas. Algunas propuestas para estimar teóricamente el rango de convergencia bajo las condiciones establecidas para este trabajo se muestra en la tabla 2.3.

**Tabla 2.3**  
**Rangos de operación propuestos por diversos autores**

Referencia	Rango
Widrow et al. [11]	$0 < \mu < \frac{1}{\lambda_{\max}}$
Haykin [12]	$0 < \mu < \frac{2}{\text{tr}[\mathbf{R}]}$
Diniz [3]	$\frac{2^{-b_c}}{4\sigma_x \sqrt{\sigma_e^2 + \sigma_n^2}} < \mu < \frac{1}{\text{tr}[\mathbf{R}]}$
Feuer y Weinstein [9]	$0 < \mu < \frac{1}{3 \text{tr}[\mathbf{R}]}$
Frías y Romero	$\frac{2^{-B}}{2\sigma_x \sqrt{\sigma_n^2 + \sigma_e^2}} < \mu_q \leq \frac{1}{n_v(\text{tr}[\mathbf{R}] + (\gamma_2 + 2)\lambda)}$

**Nota:**  $\text{tr}[\mathbf{R}] = N\sigma_x^2$  y  $\lambda = \sigma_x^2$

Establecer el límite superior teórico del rango de operación es sin duda una tarea difícil pues diversas variables y consideraciones están involucradas en el análisis, de tal forma que cualquier rango teórico propuesto puede llegar a fallar si su derivación es solo una consecuencia matemática del análisis transitorio y de estado estable y no un análisis completo que evalúe otros aspectos importantes que causan la inestabilidad del sistema tal como la variabilidad de los parámetros estadísticos. En el presente trabajo se obtuvieron condiciones de convergencia y estabilidad derivadas del modelo matemático, pero además se realizó un análisis de sensibilidad del exceso de error de estado estable con respecto al factor de convergencia y la estadística de la señal de entrada, lo anterior permite entender de mejor manera los efectos asociados a la inestabilidad, así como la necesidad de establecer regiones de seguridad. Sin duda la solución de este problema no consiste en establecer rangos fijos más restrictivos que los encontrados actualmente, sino enfocarnos al estudio de dos puntos fundamentales: suposiciones de independencia de los modelos y el efecto de la variación de los parámetros estadísticos de la señal de entrada. Como se mencionó al principio de este trabajo, el conocimiento de las características de la señal de entrada es un factor que determina el análisis del filtro adaptivo, así mismo para el análisis de estabilidad y convergencia si se tiene más información sobre el comportamiento estadístico de la señal de entrada se puede obtener un rango operativo fijo y confiable, que corresponda con alguno de los rangos de seguridad establecidos en este trabajo.

---

# Capítulo 3

## SIMULACIONES

*“Quien ama la práctica sin la teoría  
es como un marinero que sube a un barco y no sabe a dónde va”*

**-Leonardo da Vinci-**

---

### *3.1 Procedimiento de Simulación*

A continuación se presenta el procedimiento estándar de simulación formulado en base a la Fig. 2.1.

1. Definir constantes del filtro adaptivo como orden del filtro, la planta a reconocer, número de iteraciones, número de corridas independientes, varianza de la señal de entrada, exceso de kurtosis de la señal de entrada, varianza del ruido de medición, precisión del cuantizador de datos y el de coeficientes, y tipo de cuantización utilizada.
2. Establecer región y factor de convergencia de acuerdo al paso 1 y a la tabla 2.3.
3. Establecer las condiciones de inicio para vector de entrada  $\mathbf{x}(k)$ , generación de las muestras del ruido de medición así como el de la señal de entrada para el número de iteraciones requeridas y cuantización del factor de convergencia.
4. Introducir muestra de la señal de entrada mediante el corrimiento del vector de entrada.
5. Obtener la señal deseada mediante la realización el producto interno entre el vector de entrada y el de la planta desconocida, y sumar una muestra del ruido de medición,  $d(k) = \mathbf{w}_0^T \mathbf{x}(k) + n(k)$ .
6. Cuantizar la señal deseada.

7. Obtener la señal de salida  $y(k) = \mathbf{w}^T(k)\mathbf{x}(k)$ .
8. Cuantizar la señal de salida.
9. Generar la señal error mediante  $e(k) = d_q(k) - y_q(k)$ .
10. Obtener el término de actualización de filtro  $2\mu_q e(k)\mathbf{x}(k)$ .
11. Cuantizar término de actualización de filtro.
12. Realizar actualización del filtro  $\mathbf{w}(k+1) = \mathbf{w}(k) + [2\mu_q e(k)\mathbf{x}(k)]_q$ .
13. Se eleva al cuadrado el error y se acumula en una posición específica del vector MSE, la cual corresponde a la iteración realizada.
14. Se repite el paso 4 hasta que el número de iteraciones definidas se cumpla.
15. Se repite el paso 3 hasta que el número de corridas independientes definidas se cumpla.
16. Se divide el vector MSE entre el número de iteraciones totales.

### 3.2 Simulaciones

Diversas simulaciones en Matlab fueron realizadas para validar el modelo propuesto. La simulación en la Fig. 3.1 muestra la curva de aprendizaje para un filtro de orden  $N=15$ , entrada de ruido blanco de varianza  $\sigma_x^2=1$ , ruido blanco aditivo  $\sigma_n^2=0.001$  con una precisión de  $b_c=b_d=12$  bits, 200 corridas independientes y con aproximadamente 3 grados de libertad  $\mu=0.01$ .

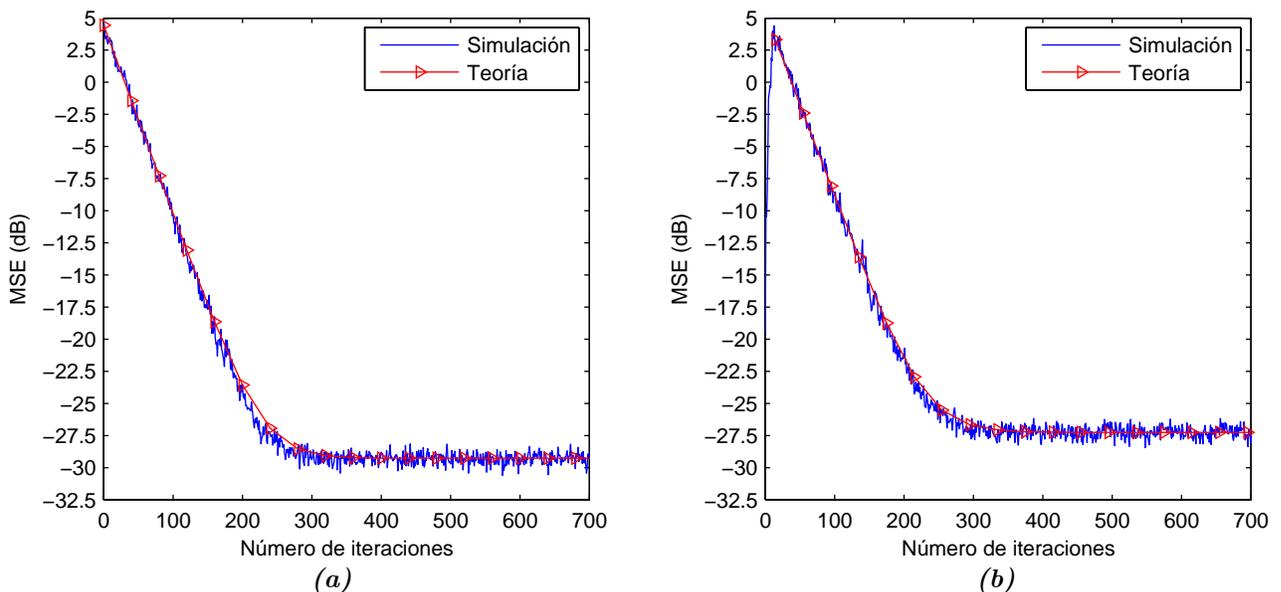


Fig. 3.1. Curva de aprendizaje comparativa utilizando: (a) redondeo, (b) truncamiento.

En la Fig. 3.1 se muestra como el desempeño de las curvas de aprendizaje teóricas se comportan de manera aceptable en relación a lo simulado, así mismo se puede confirmar que el error de estado estable es mayor para el truncamiento que para el redondeo; en el caso de la Fig. 3.1a el vector de inicio está formado por muestras de ruido blanco, mientras en la Fig. 3.1b el vector de inicio es puesto a ceros por lo tanto requiere de 15 iteraciones para que la matriz de correlación se considere constante, de esta manera se comprueba la validez de la ecuación (2.70) en la cual se estima  $\mu_v(N-1)$ .

En la Tabla 3.1 se analiza el error de estado estable para diferentes números de bits, bajo las mismas condiciones establecidas inicialmente, así como para los dos tipos de cuantización.

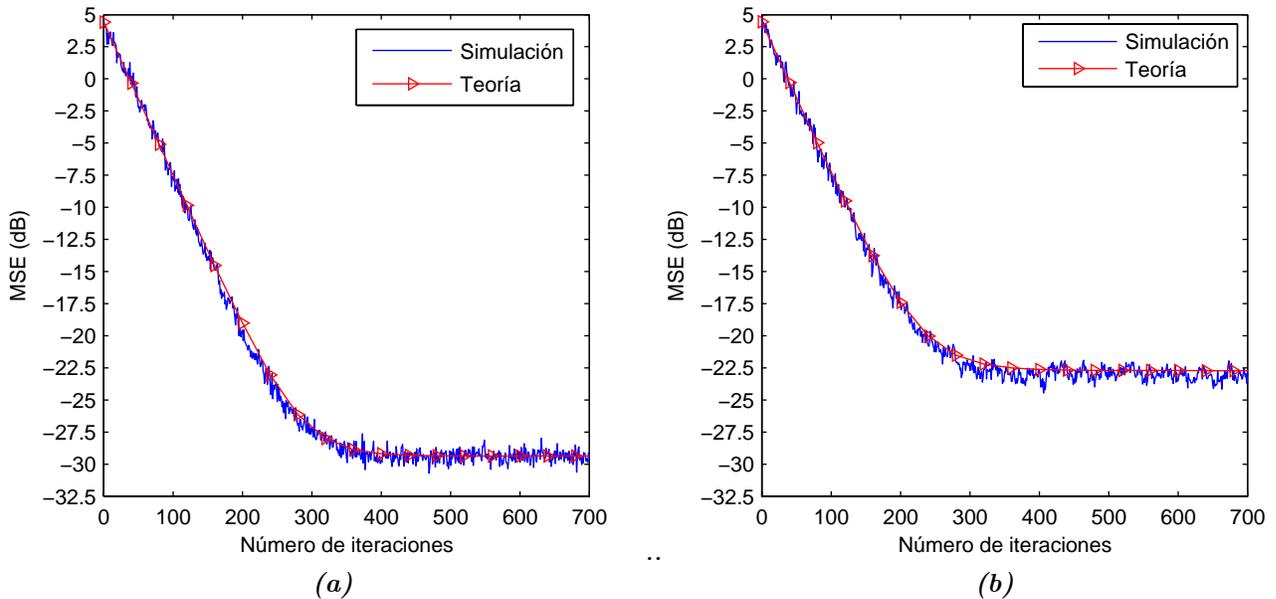
**Tabla 3.1**  
*Error de estado estable para ambos casos de cuantización.*

$b = b_c = b_d$	Error de estado estable			
	Redondeo		Truncamiento	
	Simulación	Teoría	Simulación	Teoría
10	$1.1994 \times 10^{-3}$	$1.2124 \times 10^{-3}$	$1.2486 \times 10^{-2}$	$1.2342 \times 10^{-2}$
12	$1.1824 \times 10^{-3}$	$1.1832 \times 10^{-3}$	$1.8839 \times 10^{-3}$	$1.8735 \times 10^{-3}$
16	$1.1799 \times 10^{-3}$	$1.1806 \times 10^{-3}$	$1.1840 \times 10^{-3}$	$1.1832 \times 10^{-3}$
20	$1.1812 \times 10^{-3}$	$1.1807 \times 10^{-3}$	$1.1838 \times 10^{-3}$	$1.1807 \times 10^{-3}$
24	$1.1804 \times 10^{-3}$	$1.1807 \times 10^{-3}$	$1.1807 \times 10^{-3}$	$1.1807 \times 10^{-3}$

Las diferencias de utilizar redondeo y truncamiento son evidentes para cuando el número de bits es limitado (p.e.  $b \leq 16$ ), así mismo se puede verificar que las predicciones teóricas para ambos casos de cuantización concuerdan en forma general con la simulación. En este caso en particular el error de estado estable no se vio afectado por las variaciones de la estadística de la señal de entrada debido a que la región en que se encuentra el factor de convergencia es una zona de trabajo segura como lo muestra la Fig. 2.2b ya que con 3 grados de libertad se conserva en una región lineal; sin embargo cuando el orden del filtro aumenta indudablemente la región de convergencia se hace más estrecha y debido a la cuantización del factor de convergencia, este último puede pasar de una región estrictamente lineal a una región con 10 grados de libertad, por lo que hay que cuidar que la precisión que estemos manejando sea la adecuada así como maximizar el rango de operación variando la estadística de la señal de entrada.

En las Fig. 3.2a y Fig. 3.2b se muestra las C.A. cuando se trabaja con precisión muy limitada, en la que es necesario que la precisión de los coeficientes sea mayor que en la de datos para evitar el efecto de frenado, en las dos graficas se muestra que el modelo matemático sigue siendo válido, en el caso de la gráfica de redondeo la convergencia es alcanzada a un error de estado estable mucho menor que en el truncamiento así mismo se

puede ver como la rapidez de convergencia para redondeo es mayor comparando las iteraciones que son necesarias para que la C.A. de redondeo alcance el error de convergencia de la de truncamiento; este comportamiento es debido principalmente a que la distribución  $\eta_u$  no es de media cero lo cual como se vio, este término está presente en la ecuaciones (2.66)-(2.68) lo cual afecta en forma considerable el desempeño transitorio y de estado estable.



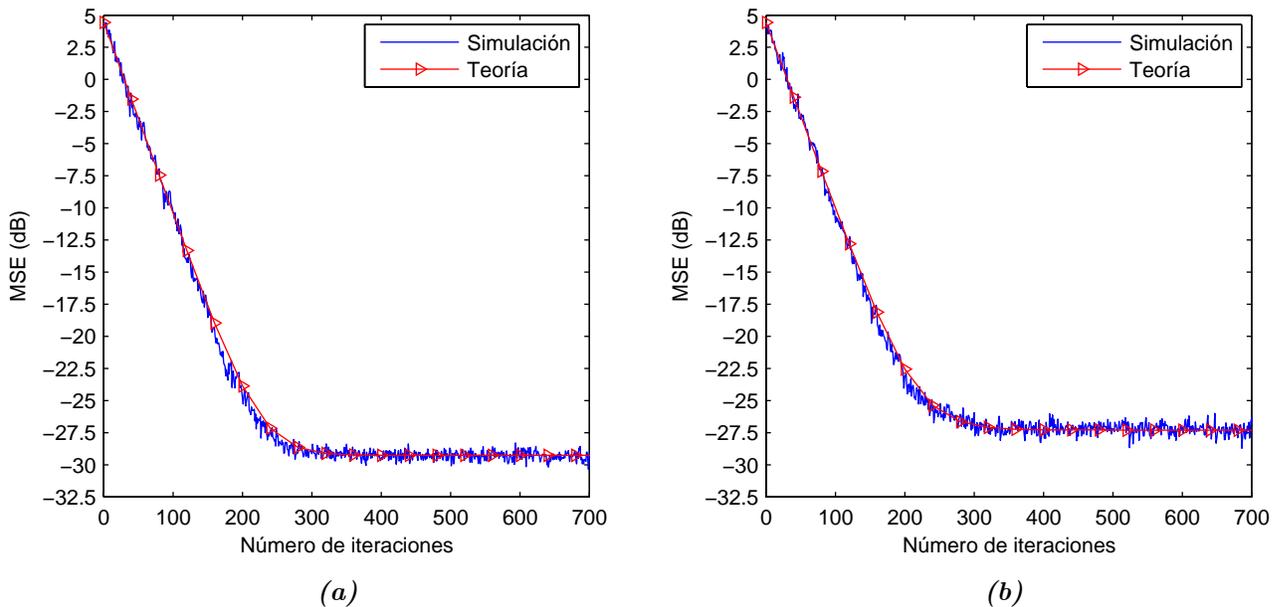
**Fig. 3.2.** Curva de aprendizaje con  $b_d = 7$  y  $b_c = 11$  para (a) redondeo, (b) truncamiento.

Como se mencionó anteriormente, un resultado importante de este trabajo es la posibilidad de analizar el desempeño de la curva de aprendizaje no solo cuando la señal de estrada es ruido blanco, sino también cuando la distribución no es perfectamente gaussiana, así como cuando se utiliza otro tipo de ruido de banda ancha. En la tabla 3.2 se muestra los diversos valores de  $\gamma_2$  para distintas distribuciones.

**Tabla 3.2**  
**Exceso de kurtosis para varias distribuciones.**

Distribución	Exceso de kurtosis ( $\gamma_2$ )
Bernoulli	$1/(1-p) + 1/p - 6$
Binomial	$(6p^2 - 6p + 1)/(np(1-p))$
$\chi^2$	$12/r$
Laplace	3
Lognormal	$e^{4s^2} + 2e^{3s^2} + 3e^{2s^2} - 6$
Maxwell	$-4/3$
Normal	0
Poisson	$1/\nu$
Uniforme	$-6/5$

En Fig. 3.3a y Fig. 3.3b se muestra las curvas de aprendizaje bajo las mismas condiciones establecidas inicialmente en esta sección solo que ahora se considera que la señal de entrada es ruido uniforme, y de acuerdo a la tabla 3.2  $\gamma_2 = -6/5$ ; por otro lado, es evidente que el modelo propuesto en este trabajo sigue siendo válido ante el cambio mencionado, así mismo se puede verificar como la rapidez de convergencia es ligeramente mayor para este caso en comparación con la Fig 3.1 dado que la variable  $\alpha$  es sensitiva a  $\gamma_2$  y específicamente ésta última es negativa, de esta manera se puede confirmar que la rapidez de convergencia es mayor.



**Fig. 3.3.** Curva de aprendizaje cuando  $\gamma_2 = -6/5$  para (a) redondeo, (b) truncamiento.

---

## Capítulo 4

# CONCLUSIONES

*"La diferencia entre una persona exitosa y otros, no es la falta de fortaleza,  
ni de conocimiento, sino la falta de voluntad"*

**-Vince Lombardi-**

---

### *4.1 Conclusiones*

Un análisis completo de los efectos de palabra finita para el filtro LMS fue desarrollado en este trabajo. La realización de un modelo analítico de la C.A. del algoritmo LMS permitió analizar en forma integral un modelo estadístico bajo la unificación de criterios y condiciones en todo el desarrollo matemático, y por consecuencia la uniformidad de los resultados. En comparación a diversos trabajos previos enfocados al análisis de estado estable con redondeo, en este trabajo se puede tener una referencia completa de la curva de aprendizaje con los efectos de cuantización por redondeo y truncamiento. Por otro lado, aunque la mayoría de los procesadores digital de señales utilizan cuantizadores por redondeo, actualmente el creciente uso de los procesadores de señales embebidos implementados en ASICs así como en FPGAs pueden hacer uso de cuantizadores por redondeo o por truncamiento debido a su arquitectura abierta, lo cual permite evaluar el desempeño del diseño de un filtro adaptivo LMS bajo las dos aproximaciones citadas en este trabajo.

Finalmente podemos concluir que la predicción analítica de la C.A. resultó ser satisfactoria dentro de los límites de estabilidad encontrados, así mismo se puede concluir que el límite superior del factor de convergencia se encuentra en relación directa a las variaciones estadísticas de la señal de entrada, así como del rango lineal en que se trabaja.

## 4.2 *Perspectivas de investigación futura*

En el presente trabajo se abarcó de manera satisfactoria la mayoría de los aspectos más importantes relacionados al análisis y diseño de un filtro adaptivo LMS, así mismo, durante el desarrollo del modelo matemático se ha podido vislumbrar algunos temas interesantes de estudio que podrían ser desarrollados posteriormente y que complementarían el presente trabajo. Algunos de esos temas serían los siguientes:

1. Análisis de la curva de aprendizaje del filtro LMS cuando la media estadística de la señal de entrada es diferente de cero, así como la dispersión de los eigenvalores de la matriz de correlación.
2. Estudios exhaustivos sobre propiedades estocásticas de ruido pseudoaleatorio, específicamente sobre suposiciones de independencia y variabilidad de parámetros estadísticos durante la realización de un proceso.
3. Desarrollo y verificación del modelo propuesto bajo descripción VHDL.

## 4.3 *Aprendizaje*

Durante el transcurso de la carrera como en el desarrollo de esta tesis diversos retos se me plantearon, y personalmente me siento completamente satisfecho de haber superado cada uno de ellos, pues no solo me han permitido crecer académicamente sino también como persona, así mismo me han ayudado a formarme un criterio más amplio y sentido de responsabilidad y ética. Creo firmemente que ser llamado *profesionista* no solo involucra obtener un papel que lo sustente sino realmente actuar como tal en todo momento.

Sin duda el realizar esta tesis me ha permitido no solo adquirir mayor conocimiento y experiencia en las áreas de procesamiento digital de señales, álgebra lineal, y procesos estocásticos sino que además me ha alentado a confirmar mi vocación por la investigación.

---

# REFERENCIAS

- [1] N.R. Yousef , A.H. Sayed, Fixed-point steady-state analysis of adaptive filters, Wiley Internat. J. Adapt. Control Signal Process. 17 (3) (2003) 237-258.
- [2] M. Andrews, R. Fitch, Finite word length arithmetic computational error effects on the LMS adaptive weights, Proc. IEEE Internat. Conf. Acoust. Speech Signal Process., Washington, DC, April 1977, pp. 628–631.
- [3] P.S.R. Diniz, Adaptive Filtering: Algorithms and Practical Implementation, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.
- [4] R. Seara, J.C.M. Bermudez, W.P. Carpes Jr., An improved quantization model for the finite precision LMS adaptive algorithm, IEEE Internat. Symp. Circuits Systems, Chicago, IL, May 1993, pp. 858-861.
- [5] R. Gupta, A.O. Hero III, Transient behavior of fixed point LMS adaptation, Proc. IEEE Internat. Conf. Acoust. Speech Signal Process., Istanbul, Turkey, June 2000, pp. 376-379.
- [6] B. Farhang-Boroujeny, Adaptive Filters: Theory and Applications, John Wiley & Sons, New York, 1998.
- [7] Vaseghi, S.V., Advanced Digital Signal Processing and Noise Reduction, John Wiley & Sons, New York, 2000.
- [8] Mitra, S.K., Digital Signal Processing: A Computer-Based Approach, New York, 2001.
- [9] A. Feuer, E. Weinstein, Convergence analysis of LMS filters with uncorrelated gaussian data, IEEE Trans. Acoust. Speech Signal Process. 33 (1) (1985) 222-230.

- 
- [10] L. I. Horowitz, K. D. Senne, Performance advantage of complex LMS for controlling narrow-bank adaptive arrays, *IEEE Trans. Acoust. Speech Signal Process.* 29 (3) (1981) 722-736.
- [11] B. Widrow et al., Stationary and nonstationary learning characteristics of the LMS adaptive filter, *Proc. IEEE* 64 (8) (1976) 1151-1162.
- [12] Haykin, S., *Adaptive Filters Theory*, Prentice Hall, New Jersey, 1996.

