

01169
2ej. 2

SOBRE APLICACIONES DE LA TEORIA DE INFORMACION
AL PROCESAMIENTO DIGITAL DE SEÑALES

Créditos asignados a la tesis DOCE (12)

APROBADO POR EL JURADO:

Presidente: DR. FEDERICO KUHLMANN RODRIGUEZ

Vocal: DR. ANDRES BUZO DE LA PEÑA

Secretario: DR. FRANCISCO GARCIA UGALDE

Suplente: DR. RAFAEL KELLY MARTINEZ

Suplente: M en C LUIS MARCIAL HERNANDEZ ORTEGA

Federico Kuhlmann
Andrés Buzo
Francisco García Ugalde
Rafael Kelly Martínez
Luis Marcial Hernández Ortega

TESIS CON
PAELLA DE ORIGEN



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

RESUMEN

"Sobre Aplicaciones de la Entropía Cruzada al Procesamiento Digital de Señales"

Se discuten las diferencias entre una medida y una distorsión indicando las características convenientes de un criterio de similitud para aplicarlo a problemas de simulación, clasificación y filtrado de procesos estocásticos discretos. Luego se resumen las definiciones y propiedades de algunos conceptos de teoría de información, según se aplican a los procesos estocásticos discretos. Se muestran 7 proposiciones que permiten bajo ciertas condiciones resolver los problemas de simulación y clasificación de procesos discretos estacionarios, ilustrando con el diseño de un vocoder de canal. Posteriormente se plantea el problema de filtrado en forma análoga al filtrado de Wiener, pero con la razón de cambio de entropía cruzada entre el proceso de error y un ruido blanco gaussiano, analizando primero el filtrado no causal y posteriormente el causal. Se muestra como la solución de Wiener es en general, una solución particular para este criterio, y como se relaciona con las soluciones aquí presentadas. Finalmente se estudian los procesos discretos periódicos.

CAPITULO. 1

INTRODUCCION

Frecuentemente se presenta la necesidad de clasificar observaciones o resultados de un experimento, de acuerdo a ciertos lineamientos característicos para cada aplicación. Todos los problemas de clasificación de señales, sin embargo, tienen algunos rasgos comunes [1]: Generalmente, las señales observadas a clasificar pertenecen a un espacio continuo, mientras que los patrones o señales patrón en que serán clasificadas forman un espacio discreto finito. Dependiendo del problema específico es necesario establecer un método para seleccionar las señales patrón que formarán el espacio de patrones, y establecer un criterio para asociar a cada señal observada una de las señales patrón del espacio discreto. Una vez que se cuenta con el espacio de patrones, esta asociación puede llevarse a cabo al determinar el valor numérico de una medida de distancia o de similitud entre la señal observada y cada una de las señales patrón, encontrando a cual de las señales patrón la señal observada es "más cercana" o "más parecida" o, en general, con respecto a cual de los patrones tiene un mayor número de características comunes. A este criterio de asociación se le conoce con el nombre de "método del vecino más cercano de acuerdo a la medida utilizada" [1].

Entonces, el diseñar un sistema para clasificar señales usando el método del vecino más cercano, presenta los problemas de seleccionar el criterio de similitud que será usado para el reconocimiento y elegir el conjunto de patrones.

En este trabajo se estudian algunas características y aplicaciones de un cri-

terio de similitud que ha sido empleado exitosamente en problemas de clasificación. Además, este criterio de similitud puede ser usado para construir el espacio discreto de los patrones a través de secuencias de entrenamiento que consisten en realizaciones típicas de la señal (proceso estocástico) a clasificar.

Las señales que serán clasificadas son realizaciones de procesos estocásticos discretos en el tiempo, o en su defecto, en el caso de ser analógicas, se supone que han sido muestreadas lo suficientemente rápido en forma tal que su versión discreta en el tiempo contenga la mayor parte de la información contenida en la señal analógica.

El criterio de similitud entre señal y patrón debe ser una función de las muestras de la señal y de una representación paramétrica del patrón. Generalmente, en problemas en que se clasifica con el método del vecino más cercano, el criterio de similitud es en realidad una medida de distancia; ésto es, satisface las propiedades de una métrica: Sean x , y , y z tres señales y sea $d(x,y)$ la similitud entre x y y .

Entonces

1. $d(x,y) \geq 0$, con igualdad si $x=y$
2. $d(x,y) = d(y,x)$, simetría
3. $d(x,y) \leq d(x,z) + d(z,y)$, desigualdad del triángulo

Sin que la intención sea restar mérito a los criterios de similitud que a su vez son métricas, cabe mencionar que existen otras características igualmente importantes que debe poseer el criterio de similitud y que la simetría y la

desigualdad del triángulo no siempre son indispensables para que un criterio de similitud sea adecuado para algunas aplicaciones específicas.

En lo sucesivo, se llamará "distorsión" a la medida de similitud, para incluir aquellas que no son métricas.

Los atributos que debe tener la distorsión para que pueda ser empleada en el método del vecino más cercano son:

1. A cada pareja de señal observada y patrón a reconocer, le debe asignar un número mayor o igual a cero. Este número representa el costo de interpretar la señal observada como alguno de los patrones.
2. Debe ser subjetivamente sensible, lo cual quiere decir que valores pequeños y grandes de la distorsión deben corresponder a cercanía o lejanía perceptible entre observación y patrón.
3. Debe ser analíticamente tratable, en el sentido de que permita analizar y diseñar técnicas de clasificación que sean prácticamente realizables.
4. Debe ser computable, lo cual indica que el valor numérico de la distorsión entre la señal observada y cada uno de los patrones debe ser calculable.

La intención de los siguientes capítulos es estudiar la entropía cruzada como una medida de distorsión que cumple los atributos 1, 3, y 4, y que para algunas aplicaciones ha mostrado satisfacer el 2, además de tener características

interesantes. Estas son: tiene una interpretación que considera a las señales como eventos en un espacio de probabilidades, tiene una interpretación en los dominios del tiempo y la frecuencia, y puede ser evaluada en ambos dominios.

Además de la clasificación, otras aplicaciones de interés para una medida de distorsión aparecen en el área de simulación de procesos y en los problemas de filtrado (predicción pura de orden finito e infinito, filtrado puro de orden finito e infinito, interpolación ("smoothing"), etc.)

En el área de simulación uno de los problemas fundamentales es el de generar una señal sintética que según algún criterio prefijado sea lo más parecida a las señales o familia de señales observadas. En este problema, la distorsión es una medida de fidelidad, es decir, de que tan similar es el proceso producido por vía de la simulación al proceso real que se está simulando; se desea que la reproducción sea lo más fiel posible, ésto es, que la distorsión sea mínima.

Los atributos que debe tener la distorsión para que pueda ser empleada como criterio de fidelidad son análogos a los requeridos para su uso en el método del vecino más cercano para reconocimiento de patrones:

1. A cada par de proceso sintético y proceso observado, le debe asignar un número mayor ó igual a cero.
2. La cercanía lograda debe ser perceptible por el humano.
3. Debe ser analíticamente tratable, en el sentido de que permita diseñar el sistema simulador.

4. Debe ser calculable, de tal forma que su valor numérico sirva para describir la calidad de la simulación.

En el caso de filtrado la distorsión es una medida de desempeño. Definiendo el error como la diferencia entre la señal ya filtrada, o señal estimada, y la señal original, o señal a estimar, el desempeño del filtro estimador se mide como la distorsión presente entre el error y una señal de referencia. Es evidente que la selección de la señal de referencia juega un papel crucial, y que en función de ella y de la aplicación se desea minimizar ó maximizar la medida de desempeño. Tomando en cuenta estas consideraciones, los atributos correspondientes a la medida distorsión son:

1. A la pareja formada por el error obtenido y el proceso de referencia debe asignarse un número mayor o igual a cero.
2. Debe proporcionar alguna ventaja perceptible del proceso filtrado sobre el proceso perturbado a partir del cual se efectúa la estimación.
3. Debe ser analíticamente tratable, en el sentido que permita diseñar el filtro.
4. Debe ser calculable, para que su valor numérico sirva su función como índice de desempeño.

En el capítulo 2 de este trabajo, se presentan algunas definiciones y resultados de la teoría de información con objeto de establecer un marco de referencia para su uso en los capítulos posteriores.

En el capítulo 3, con base en conceptos de teoría de la información, se derivan las propiedades más importantes de la cantidad de información promedio para discriminar entre procesos estacionarios; con el enfoque que aquí se presenta, es posible obtener, en forma simple, las propiedades más importantes de este criterio de discriminación, permitiendo además derivar interpretaciones nuevas para resultados conocidos, de interés tanto teórico como práctico. Este criterio ha sido empleado exitosamente en algunos problemas de clasificación [1,2] y síntesis [3,4] de señales discretas en el tiempo.

En el capítulo 4 se aplican los resultados del capítulo 3 a los problemas de clasificación y simulación de procesos y se ilustra la forma en que esos resultados se aplican al diseño del vocoder de canal.

En el capítulo 5 se muestra que la cantidad de información de discriminación promedio puede ser usada como una medida de desempeño para el diseño de filtros óptimos, evaluando la distorsión entre el proceso de error y un proceso de referencia. La clase de problemas de Filtrado de Wiener se reformula basándose en esta distorsión, tanto para predicción como para filtrado. La obtención de un punto crítico para este criterio bajo hipótesis de Gaussianidad conduce en primera instancia a una ecuación de Wiener-Hopf ampliada y descrita en el dominio de la frecuencia. En particular, para el caso no-causal, se muestra como el filtro de Wiener satisface su ecuación correspondiente. Para el caso causal, se establecen condiciones (necesarias y suficientes) para que el filtro de Wiener sea una solución de la ecuación planteada; particularizando un poco más, en el caso de ruido aditivo blanco ortogonal, se establece una condición necesaria para que exista un punto crítico. Se muestra luego que el filtro de Wiener constituye un punto silla para el caso no-causal, y tam-

bién para el caso causal bajo las condiciones establecidas, y cómo, a partir de estas soluciones, pueden diseñarse filtros globalmente óptimos según el criterio de información discriminatoria, para diferentes casos particulares importantes.

Finalmente, en el capítulo 6, se estudian los procesos periódicos y se concluye que existe una forma especial para ellos de la distorsión empleada.

Con la intención de hacer lo más legible el trabajo, a la vez de autocontenido, se incluyen varios apéndices donde se efectúan las demostraciones más elaboradas.

Suposiciones Generales:

Se supondrá que las señales observadas son procesos estocásticos reales, estrictamente estacionarios, ya sea regulares (Capítulos 2, 3, 4 y 5) o predecibles (Capítulo 6) [5]; las señales sintéticas y las señales patrón son modelos paramétricos de estos procesos estacionarios y el criterio de similitud es la cantidad de información promedio que se tiene para discriminar entre el modelo paramétrico del proceso y el proceso observado, dado que se observa cierta estadística del segundo. Este criterio se traduce a la razón de cambio de la entropía cruzada (RCEC) [6] para procesos regulares y a la entropía cruzada normalizada para procesos predecibles, y ha recibido mucha atención en los últimos años [7,8]. En general, el índice de definición de los procesos será el tiempo discreto, de menos infinito a más infinito, con la excepción de los problemas de filtrado con entradas estacionarias aplicadas desde un tiempo

inicial $n_0=0$, en los que las respuestas de los filtros son asintóticamente estacionarias y el índice de definición de dichas respuestas varía de cero a infinito.

CAPITULO 2

LA ENTROPIA CRUZADA COMO MEDIDA DE DISTORSION ENTRE PROCESOS

Entropia Cruzada

Una clase de medidas de distorsión entre variables o secuencias de variables aleatorias que ha sido usada con frecuencia en diversos contextos, consiste en la clase de Ali-Silvey [10], la cual se define de la siguiente forma:

Sean $P_u^N(x)$ y $P_v^N(x)$ las funciones de densidad de probabilidad conjunta correspondientes a las variables aleatorias u y v (donde u y v son vectores aleatorios de dimensión N , y $x = (x_1, x_2, \dots, x_N)^T$ es un vector mudo con N componentes); sea $\delta x^N = dx_1 \cdot dx_2 \dots \cdot dx_N$ el elemento diferencial en R^N .

Defínase

$$d(P_u^N, P_v^N) = F\left(\int c\left(\frac{P_u^N(x)}{P_v^N(x)}\right) P_u^N(x) \delta x^N\right)$$

(donde la integral es sobre el soporte de $P_u^N(x)$), como la distorsión o costo por interpretar la variable u por la variable v , donde c es una función convexa y continua en $(0, \infty)$ y F es una función no decreciente. A esta clase de medidas pertenecen las medidas de

1. Kolmogorov: $F(a) = a$, $c(a) = (a-1)^2$ ó $F(a) = \frac{a}{2}$, $c(a) = a-1$
2. Matusita: $F(a) = \frac{a}{2}$, $c(a) = (a-1)^2$
3. Chernoff: $F(a) = -\lambda \ln a$, $c(a) = a^{1-\alpha}$; $0 \leq \alpha \leq 1$
4. Entropía cruzada: $F(a) = a$, $c(a) = \ln a$; [1]

A la entropía cruzada también se le conoce como número de Kullback y Leibler [2] o cantidad de información discriminatoria. Nótese que la definición presupone la existencia de las funciones de densidad de probabilidad, o lo que es lo mismo, que las funciones de distribución son absolutamente continuas, y además, que el soporte de $P_U^N(x)$ contiene al soporte de $P_V^N(x)$. En lo que sigue, se supone que todas las funciones de densidad de probabilidad, que formen parte de un integrando, tienen el mismo soporte, y que la integral se efectúa sobre todo el soporte de estas funciones de densidad de probabilidad.

La entropía cruzada entre las variables u y v se denota por $H(P_U^N, P_V^N)$, puede expresarse como

$$H(P_U^N, P_V^N) = \int \delta x P_U^N(x) \ln(P_U^N(x)/P_V^N(x)) \quad (7.2.1)$$

y tiene las siguientes propiedades:

1. $H(P_U^N, P_V^N) \geq 0$, con igualdad si, y sólo si, $P_U^N(x) = P_V^N(x)$ (Para mostrarlo, aplíquese la desigualdad de Jensen a $-H(P_U^N, P_V^N)$).
2. Es invariante a cambios del sistema de referencia, ésto es:

$$H(P_{g(u)}^N, P_{g(v)}^N) = H(P_U^N, P_V^N)$$

donde $g(x) \in \mathbb{R}^N$ y es una función invertible. Esto es, cuando en lugar de observar las variables originales se observan las transformaciones $g(u)$ y $g(v)$, no se altera la similitud entre las variables (Para mostrarlo, considere las funciones de densidad de probabilidad

resultantes de la transformación [5, pág. 143], la definición de entropía cruzada, y el jacobiano de la transformación).

3. Si u y $v \in \mathbb{R}^N$ y las componentes de v son estadísticamente independientes, entonces

$$H(P_U^N, P_V^N) = -H(P_U^N) + \sum_{i=1}^N \int dx_i P_{u_i}(x_i) \ln(1/P_{v_i}(x_i)) \quad (2.3.1)$$

$$\geq \sum_{i=1}^N H(P_{u_i}, P_{v_i}) \quad (2.3.2)$$

$$= N H(P_{u_1}, P_{v_1}) \quad (2.3.3)$$

donde $H(P_U^N)$ es la entropía diferencial del vector de variables aleatorias $(u_1, u_2, \dots, u_N)^T$, definida por

$$H(P_U^N) = \int \delta x P_x^N(x) \ln(1/P_U^N(x)) \quad (2.3.4)$$

y $P_{u_i}(x)$, $P_{v_i}(x)$ son las funciones de densidad de probabilidad marginal de las variables componentes u_i y v_i respectivamente. (2.3.2)

se cumple con igualdad si, y sólo si, las componentes de $u \in \mathbb{R}^N$ son también independientes (Para mostrar (2.3.2) recuérdese la desigualdad $H(P_U^N) \leq \sum_{i=1}^N H(P_{u_i})$; (2.3.3) por la suposición de estacionariedad).

4. Según la definición de información propuesta en la teoría de la información, la entropía cruzada mide la cantidad promedio de información para discriminar la variable aleatoria v de la variable aleatoria u , dado que se observa a la variable aleatoria u , o sea,

$$H(P_U^N, P_V^N) = \int \delta x P_U^N(x) (I_V(x) - I_U(x))$$

donde $I_U(x) = -\ln P_U^N(x)$ es la información que proporciona el evento $u=x$ sobre la variable aleatoria u . Debido a esta propiedad la entropía cruzada recibe el nombre de número de información discriminatoria promedio.

5. Sean Q_+^N y P^N f.d.p. (funciones de densidad de probabilidad) con el mismo soporte $D \subset \mathbb{R}^N$. Sean ciertas funciones conocidas $f_k(x)$ (ó $g_k(x)$), $k=1,2,\dots,M$, con dominio en \mathbb{R}^N y rango en \mathbb{R} , las funciones de restricción. Sean \bar{f}_k (ó \bar{g}_k), $k=1,2,\dots,M$, constantes definidas por

$$\int_D f_k(x) Q_+^N(x) \delta x = \bar{f}_k \quad (2.4.1)$$

ó

$$\int_D g_k(x) Q_+^N(x) \delta x \geq \bar{g}_k \quad (2.4.2)$$

Sea $Q_{\#}^N$ cualquier f.d.p. con soporte D que pueda ser usada en lugar de Q_+^N en las ecuaciones (2.4.1) (ó (2.4.2)); es decir, $Q_{\#}^N$ es cualquier f.d.p. con soporte D que satisface las restricciones.

Sea Q^N tal que

$$H(Q^N, P^N) = \min_{Q_{\#}^N} H(Q_{\#}^N, P^N) \quad (2.4.3)$$

Entonces:

$$H(Q_+^N, P^N) \geq H(Q_+^N, Q^N) + H(Q^N, P^N), \quad (\text{relación de triángulo}) \quad (2.4.4)$$

Si la minimización se realiza bajo un conjunto finito de restricciones de igualdad (2.4.1), siempre que exista solución a éste problema, se satisface la igualdad en (2.4.4), y además, Q^N es de la forma

$$Q^N(x) = P^N(x) \exp \left\{ -\lambda_0 - \sum_{k=1}^M \lambda_k f_k(x) \right\} \quad (2.5.1)$$

(con la posible excepción de un conjunto de puntos donde $Q^N(x) = 0$; sin embargo, esta situación no será considerada aquí); las λ_i son los multiplicadores de Lagrange que deben escogerse en tal forma que $Q^N(x)$ satisfaga las restricciones.

Q_+^N es la f.d.p. a estimar, P^N es un estimado a priori de Q_+^N determinado antes de conocer las restricciones, y Q_-^N es el estimado a posteriori de Q_+^N determinado tomando en cuenta la información proporcionada por las restricciones.

6. Sea Q_*^N cualquier f.d.p. de la forma

$$Q_*^N(x) = P^N(x) \exp \left\{ -\lambda_0 - \sum_{k=1}^M \lambda_k f_k(x) \right\} \quad (2.5.2)$$

Entonces $H(Q_+^N, Q_*^N)$ es mínima si Q_*^N satisface las restricciones (2.4.1).

7. De las propiedades 5. y 6., se sigue que el término $H(Q_+^N, Q^N)$ en (2.4.4) es mínimo, si la minimización (2.4.3) se realiza bajo restricciones de igualdad, ya que Q^N es de la forma (2.5.2) y satisface las restricciones (2.4.1).

Para la demostración de 5. y 6., y de otras propiedades interesantes,

véase [8].

Cabe aclarar que la entropía cruzada no satisface ni la propiedad de simetría ni la desigualdad del triángulo de las métricas. La relación de triángulo (2.4.4) debe entenderse bajo el contexto de la minimización (2.4.3) con restricciones (2.4.1) ó (2.4.2).

Razón de Cambio de la Entropía Cruzada

Cuando en lugar de medir la similitud entre vectores de variables aleatorias, se desea medir la similitud entre procesos aleatorios estacionarios discretos en el tiempo, con funciones de densidad de probabilidad conjunta definidas para cualquier orden finito, es posible extender la definición de entropía cruzada, dando lugar a la razón de cambio de la entropía cruzada (RCEC), la cual está definida por:

$$H(U,V) = \lim_{N \rightarrow \infty} \frac{1}{N} \int \dots \int \delta x \, P_U^N(x_1, \dots, x_N) \ln \left(\frac{P_U^N(x_1, \dots, x_N)}{P_V^N(x_1, \dots, x_N)} \right) \quad (2.6.1)$$

donde U y V representan a los procesos $\{u_n\}_{n=-\infty}^{\infty}$ y $\{v_n\}_{n=-\infty}^{\infty}$.

El cálculo de la RCEC presenta la dificultad de requerir las funciones de densidad de probabilidad conjunta de todo orden finito de los procesos entre los cuales se desea establecer la similitud, lo que, salvo en casos especiales, resulta imposible.

La RCEC posee las mismas propiedades que la entropía cruzada, con modificaciones obvias; por ejemplo, cuando tanto U como V son sucesiones de variaia

bles aleatorias independientes e idénticamente distribuidas (v.a.i.i.d.), por la propiedad 3,

$$\bar{H}(U,V) = H(P_{u_i}, P_{v_i}) \quad (2.7.1)$$

y la ecuación (2.6.1) se reduce a la entropía cruzada entre las variables aleatorias u_i y v_i , las cuales tienen la misma f.d.p. que la f.d.p. marginal de primer orden del proceso U y V respectivamente. En este caso particular el cálculo de la RCEC se simplifica considerablemente.

Al aplicar la propiedad 2 de la RCEC deberá ejercerse cierto cuidado al considerar que la transformación debe ser invertible (véase apéndice A).

Otro caso en el cual la expresión para la RCEC se reduce a una más sencilla es cuando los dos procesos son gaussianos con funciones de densidad espectral $f_u(\omega)$ y $f_v(\omega)$, respectivamente, en cuyo caso la RCEC se reduce a [6; apéndice D]

$$\bar{H}(U,V) = \frac{1}{2} \int_{-\pi}^{\pi} \left(\frac{f_u(\omega)}{f_v(\omega)} - \ln \frac{f_u(\omega)}{f_v(\omega)} - 1 \right) \frac{d\omega}{2\pi} \quad (2.7.2)$$

El doble del miembro derecho de esta última ecuación recibe el nombre de distorsión de Itakura-Saito [10] entre dos procesos estacionarios y se emplea para medir la similitud entre cualesquier par de procesos estacionarios regulares considerando únicamente sus estadísticas de segundo orden (ó funciones de densidad espectral); la igualdad en (2.7.2) sólo se cumple si los procesos en cuestión son gaussianos y estacionarios.

Cuando los procesos no son gaussianos no es posible escribir la RCEC en fun-

ción de su estadística de segundo orden. Sin embargo, cuando la RCEC se usa en situaciones donde se pretende medir la similitud entre un proceso estacionario arbitrario y un modelo paramétrico de un proceso estacionario, pueden resultar algunas simplificaciones, dependiendo del modelo que se use para representar al proceso con respecto al cual se pretende medir su similitud. En particular, si los procesos patrón o procesos sintéticos son procesos autorregresivos (AR), de promedio móvil (MA), o procesos ARMA, la expresión de la RCEC se simplifica considerablemente, lo cual será derivado en el capítulo 3, donde además se trata el caso en el cual la señal observada puede ser modelada como un proceso estocástico predecible.

Entropía Diferencial.

La entropía diferencial de la f.d.p. conjunta de orden N , $P_U^N(x)$, asociada al vector aleatorio u , está definida por:

$$H(P_U^N) = \int \dots \int \delta x P_U^N(x) \ln(1/P_U^N(x)) \quad (2.8.1)$$

Razón de Cambio de Entropía Diferencial

La razón de cambio de la entropía diferencial (RCED) del proceso estocástico $U = \{u_n\}_{n=-\infty}^{\infty}$, está definida por:

$$\bar{H}_U = \lim_{N \rightarrow \infty} \frac{1}{N} \int \dots \int \delta x P_U^N(x_1, \dots, x_N) \ln(1/P_U^N(x_1, \dots, x_N)) \quad (2.8.2)$$

La RCED tiene la siguiente propiedad:

Sea $U = \{u_n\}_{n=-\infty}^{\infty}$ un proceso estocástico estacionario definido en el tiempo discreto, la entrada desde $n=0$ a un sistema causal y estable lineal invariante en el tiempo con función de transferencia $H(e^{j\omega})$. Entonces, la salida $V = \{v_n\}_{n=0}^{\infty}$ es un proceso estocástico asintóticamente estacionario con RCED dada por

$$\bar{H}_V = \bar{H}_U + \lambda_n |h[0]| \quad (2.9.1)$$

donde $h[n]$ es la secuencia de respuesta impulso del sistema; si además el sistema es de fase mínima, entonces:

$$\lambda_n |h[0]| = \int_{-\pi}^{\pi} \lambda_n |H(e^{j\omega})| \frac{d\omega}{2\pi} \quad (2.9.2)$$

[5, pág. 534; 11, pág. 299; apéndice A.6; apéndice B.7]

Información Mutua.

La información mutua promedio entre los vectores aleatorios u y v , con f.d.p. conjunta respectivas $P_U^N(x)$ y $P_V^N(y)$, está definida por:

$$I(u;v) = \int \dots \int \delta x^N \delta y^N P_{UV}^{2N}(x,y) \lambda_n \left(\frac{P_{UV}^{2N}(x,y)}{P_U^N(x)P_V^N(y)} \right) \quad (2.9.3)$$

donde $P_{UV}^{2N}(x,y)$ es la f.d.p. conjunta del vector aleatorio de dimensión $2N$ formada por la concatenación de u y v .

Razón de Cambio de Información Mútua.

La razón de cambio de información mútua (RCIM) entre los procesos estocásticos conjuntamente estacionarios $U = \{u_n\}_{n=-\infty}^{\infty}$ y $V = \{v_n\}_{n=-\infty}^{\infty}$ está definida por [6, pág. 76]:

$$\bar{I}(U;V) = \lim_{N \rightarrow \infty} \frac{1}{2N} \int \dots \int_{\delta x}^N \int_{\delta y}^N P_{uv}^{2N}(x_1, \dots, x_N, y_1, \dots, y_N) \ln \left(\frac{P_{uv}^{2N}(x_1, \dots, x_N, y_1, \dots, y_N)}{P_u^N(x_1, \dots, x_N) P_v^N(y_1, \dots, y_N)} \right) \ln \quad (2.10.1)$$

1. Se puede demostrar [6, pág 175], que si los procesos U y V son conjuntamente estacionarios gaussianos y regulares, la RCIM entre ellos está dada por:

$$\bar{I}(U; V) = - \frac{1}{4\pi} \int_{-\pi}^{\pi} \ln(1 - \gamma_{uv}^2(\omega)) d\omega \quad (2.10.2)$$

donde $\gamma_{uv}^2(\omega)$ es la coherencia entre U y V , definida por:

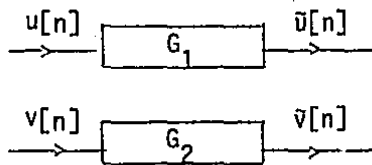
$$\gamma_{uv}^2(\omega) = \frac{f_{uv}(e^{j\omega}) f_{vu}(e^{j\omega})}{f_u(e^{j\omega}) f_v(e^{j\omega})} \quad (2.10.3)$$

donde $f_{uv}(e^{j\omega})$ es la función de densidad espectral cruzada entre u y v , $f_{vu}(e^{j\omega}) = f_{uv}^*(e^{j\omega}) = f_{uv}(e^{-j\omega})$ (los procesos en este trabajo se consideran reales), y $f_u(e^{j\omega})$, $f_v(e^{j\omega})$ son las funciones de densidad espectral de U y V respectivamente. Cuando $\gamma_{uv}^2(\omega) \equiv 1$, se dice que U y V son coherentes, y entonces $\bar{I}(U; V) = \infty$; ésto sucede si un proceso puede obtenerse a partir del otro por medio de un filtro.

La RCIM tiene además las siguientes propiedades:

2. Sean U y V procesos conjuntamente estacionarios gaussianos y regulares que excitan a los sistemas lineales invariantes en el tiempo y estables con funciones de transferencia $G_1(e^{j\omega})$ y $G_2(e^{j\omega})$ respectivamente. Entonces $\bar{I}(\tilde{U}; \tilde{V}) = \bar{I}(U; V)$, donde \tilde{U} y \tilde{V} son los procesos a la salida de cada uno de los sistemas. (Nótese que no es necesario que los sistemas sean causales, pero el papel de los procesos como entradas y salidas no puede intercambiarse para aplicar el resultado a los sistemas inversos a menos que los sistemas inversos sean estables).

Demostración:



[5, pág. 282]:

$$f_{\tilde{U}}(e^{j\omega}) = G_1(e^{j\omega}) G_1^*(e^{j\omega}) f_U(e^{j\omega})$$

$$f_{\tilde{V}}(e^{j\omega}) = G_2(e^{j\omega}) G_2^*(e^{j\omega}) f_V(e^{j\omega})$$

$$f_{\tilde{V}\tilde{U}}(e^{j\omega}) = G_2(e^{j\omega}) G_1^*(e^{j\omega}) f_{VU}(e^{j\omega})$$

$$f_{\tilde{U}\tilde{V}}(e^{j\omega}) = G_1(e^{j\omega}) G_2^*(e^{j\omega}) f_{UV}(e^{j\omega})$$

$$\therefore \gamma_{\tilde{U}\tilde{V}}^2 = \frac{f_{\tilde{U}\tilde{V}}(e^{j\omega}) f_{\tilde{V}\tilde{U}}(e^{j\omega})}{f_{\tilde{U}}(e^{j\omega}) f_{\tilde{V}}(e^{j\omega})} = \frac{f_{UV}(e^{j\omega}) f_{VU}(e^{j\omega})}{f_U(e^{j\omega}) f_V(e^{j\omega})} = \gamma_{UV}^2$$

y usando (2.10.2), $\bar{I}(\tilde{U}; \tilde{V}) = \bar{I}(U; V)$

3. Sean U y V procesos conjuntamente estacionarios que excitan a los sistemas lineales e invariantes en el tiempo (L.I.T.), causales y estables, con funciones de transferencia $H^{(1)}(z)$ y $H^{(2)}(z)$ respectivamente. Entonces $\bar{I}(\tilde{U}; \tilde{V}) = \bar{I}(U; V)$, donde \tilde{U} y \tilde{V} son los procesos a la salida de cada uno de los sistemas (La causalidad de los sistemas es una condición suficiente dado que $h_0^{(1)} \neq 0$ y $h_0^{(2)} \neq 0$). Nótese que no es necesario que los procesos sean gaussianos. La igualdad $\bar{I}(\tilde{U}; \tilde{V}) = \bar{I}(U; V)$ se sigue cumpliendo cuando \tilde{U} y \tilde{V} son las entradas a los sistemas inversos y U y V son las salidas, si los sistemas son de fase mínima), si U y V se aplican desde $n_0 = 0$.

Para una demostración, véase la sección A.7 del apéndice A.

Razón de Cambio de Entropía Diferencial Condicionada:

La razón de cambio de entropía diferencial condicionada (RCEDC) se define por:

$$\bar{H}(U|Z) = \lim_{N \rightarrow \infty} \frac{-1}{N} \int \dots \int_{U|Z} P_{UZ}^{2N}(x_1, \dots, x_N, t_1, \dots, t_N) \ln P_{U|Z}^N(x_1, \dots, x_N | t_1, \dots, t_N) \delta x \delta t$$

(2.12.1)

donde los procesos U y Z son conjuntamente estacionarios, $P_{uz}^{2N}(x,t)$ es la f.d.p. conjunta del vector aleatorio de dimensión $2N$ formado por la concatenación de u y z (u y z son vectores aleatorios de dimensión N) y $P_{u|z}^N(x|t)$ es la función de densidad de probabilidad conjunta condicional de u dado z . ($x, t \in \mathbb{R}^N$).

Algunas Relaciones Entre las Cantidades Definidas:

En forma análoga a la demostración de las relaciones fundamentales correspondientes de la teoría de información, se puede demostrar:

1. $\bar{H}_U = \bar{I}(U; U)$ (si las definiciones son las dadas en [6])
2. $\bar{H}(U|Z) = \bar{H}_U - \bar{I}(U; Z)$
3. Sean U, V , y $Z = U + V$ procesos discretos para los cuales están definidas las RCED de cada uno de ellos y las RCIM entre cualquier par de ellos. Entonces:

$$\bar{H}_Z = \bar{I}(U; Z) + \bar{H}_V - \bar{I}(U; V) \quad (2.13.1)$$

$$\bar{H}_Z = \bar{I}(V; Z) + \bar{H}_U - \bar{I}(U; V) \quad (2.13.2)$$

y por tanto:

$$\bar{I}(U; Z) - \bar{I}(V; Z) = \bar{H}_U - \bar{H}_V \quad (2.13.3)$$

Demostración:

$$\bar{I}(U; U+V) = \bar{H}_Z - \bar{H}(U+V|U), \text{ por 2.}$$

$$= \bar{H}_Z - \bar{H}(V|U), \quad [5, \text{pág. 566}]$$

$$= \bar{H}_Z - (\bar{H}_V - \bar{I}(U; V)), \quad \text{por 2.}$$

$$\therefore \bar{I}(U; Z) = \bar{H}_Z - \bar{H}_V + \bar{I}(U; V)$$

Lo que da (2.13.1). La demostración de (2.13.2) es análoga.

CAPITULO 3

DISCRIMINACION ENTRE PROCESOS ESTACIONARIOS Y SINTETICOS:

En este capítulo se derivan las propiedades más importantes de la cantidad de información discriminatoria promedio entre procesos discretos estacionarios y regulares, con media cero, y algunos de los modelos paramétricos de procesos estacionarios más usados en la ingeniería, aquí llamados procesos sintéticos.

Es este capítulo el primer argumento de la RCEC es un proceso estocástico fijo, y se busca un segundo argumento que la minimice. En el capítulo 5, se tendrá fijo el segundo argumento (el proceso de referencia), y se buscará un primer argumento que minimice la RCEC.

En el apéndice G se resumen algunas características de los procesos regulares. La discriminación entre procesos predecibles se discute en el capítulo 6.

Procesos Sintéticos: Una clase importante de modelos de procesos estacionarios, son los generados a la salida de un sistema S dinámico, discreto, invariante en el tiempo e invertible, tal que S y su inverso S^{-1} son estables y causales, cuando es excitado por una sucesión de v.a.i.i.d. (variables aleatorias independientes e idénticamente distribuidas) $Z = \{z_n\}_{n=-\infty}^{\infty}$, con media cero, variancia unitaria (a ésta clase de procesos se les conoce con el nombre de ruido blanco en el sentido estricto con potencia unitaria) y distribución de probabilidad marginal absolutamente continua con densidad de probabilidad marginal $P_Z(x)$, ésto es:

$$p_z^N(x_1, \dots, x_N) = \prod_{i=1}^N p_z(x_i); \quad N \geq 1$$

$$E\{z_n\} = 0 \quad (3.2.1)$$

$$E\{z_n^2\} = 1$$

donde $E\{\cdot\}$ representa al operador esperanza matemática.

A esta clase de procesos se les llamará procesos sintéticos. De esta forma, si $V = \{v_n\}_{n=-\infty}^{\infty}$ es un proceso sintético, entonces V está determinado por la densidad de probabilidad marginal del proceso estacionario de excitación $p_z(x)$, y por el sistema S con los atributos antes mencionados.

Ejemplos de éstos modelos de procesos estacionarios regulares son: los generadores de números aleatorios, donde al excitar con v.a.i.i.d. con distribución uniforme en el intervalo $[0, 1]$ a un sistema sin memoria, por lo general no lineal, se generan v.a.i.i.d. con una función de densidad de probabilidad prefijada; los procesos AR, MA y ARMA, donde al excitar por medio de v.a.i.i.d. a un sistema S lineal, se generan procesos estacionarios con estadísticas de segundo orden (funciones de autocorrelación) preestablecidas.

Cuando los procesos patrón o procesos a generar son procesos sintéticos como los aquí definidos, RCEC se simplifica, dando lugar a algunos resultados interesantes.

Notación: Sea U un proceso discreto en el tiempo, estacionario y regular; sea V un proceso sintético caracterizado por el sistema S y por la fun-

ción de densidad de probabilidad marginal $P_Z(x)$ que caracteriza a la excitación de v.a.i.i.d. del sistema S . Entonces, después de aplicar la generalización de la propiedad de invariancia a cambios del sistema de referencia de la entropía cruzada (apéndice A.5), la RCEC se reduce a

$$\bar{H}(U, V) = \bar{H}(R, Z) \quad (3.3.1)$$

donde $R = \{r_n\}_{n=-\infty}^{\infty}$ es el proceso que se obtiene al excitar el sistema inverso al sistema S , literalmente S^{-1} , por el proceso observado U , y Z es la secuencia de v.a.i.i.d. que genera al proceso V . Al proceso R se le denomina proceso residual. La ecuación (3.3.1) significa que la RCEC entre el proceso observado U y el proceso sintético V , es igual a la RCEC entre el proceso residual y el proceso de excitación que genera al proceso sintético. De la ecuación (3.3.1) se obtiene la siguiente proposición:

Proposición 1:

La RCEC entre un proceso estacionario regular U y un proceso sintético V es igual a:

$$\bar{H}(U, V) = I(r_n; (r_{n-1}, r_{n-2}, \dots)) + H(P_r, P_z) \quad (3.3.2)$$

donde $I(r_n; (r_{n-1}, r_{n-2}, \dots))$ es la información mutua promedio entre r_n y $(r_{n-1}, r_{n-2}, \dots)$, o sea, el promedio de la cantidad de información que se obtiene sobre el valor del residuo r_n por medio de sus valores anteriores.

De esta proposición se concluye que para generar un proceso sintético lo más

cercano posible, según la RCEC, a un proceso dado, se requiere un sistema S , cuyo inverso S^{-1} al ser usado como filtro del proceso a observar, genere una secuencia de v.a.i.i.d., y que el proceso de excitación del sistema S tenga la misma función de densidad de probabilidad que el proceso residual. Vale la pena mencionar que el encontrar un filtro que transforme un proceso estacionario regular cualquiera, no gaussiano, en ruido blanco en el sentido estricto, es un problema no resuelto. Además, resulta interesante notar la relación tan estrecha que tiene la generación de procesos sintéticos óptimos (lo más cercano posible, según la RCEC) con el problema de filtrado óptimo.

Demostración de (3.3.2):

Usando la ecuación (2.3.1), la ecuación (3.3.1) puede ser reescrita como:

$$\bar{H}(U, V) = -\bar{H}_r + \int P_r(x) \ln \left[\frac{1}{P_z(x)} \right] dx \quad (3.4.1)$$

(ya que Z es una suceción de v.a.i.i.d.), donde \bar{H}_r es la razón de cambio de la entropía diferencial (RCEC) del proceso estacionario R , definida por (2.8.2) y $P_r(x)$ es la función de densidad de probabilidad marginal de orden 1 de dicho proceso.

Sumando y restando al miembro derecho de la ecuación (3.4.1) la entropía diferencial de la función de densidad de probabilidad marginal de orden uno del proceso de residuos $P_r(x)$, definida por la ecuación (2.8.1), se obtiene:

$$\bar{H}(U, V) = -\bar{H}_r + H(P_r) + H(P_r, P_z) ;$$

ya que $\bar{H}_r = \lim_{m \rightarrow \infty} H(r_n | (r_{n-1}, \dots, r_{n-m}))$, [5, pág. 533], entonces [5, pág.521]:

$$I(r_n; (r_{n-1}, r_{n-2}, \dots)) = H(P_r) - \bar{H}_r \quad (3.5.1)$$

y la proposición 1 queda probada.

Proposición 2:

Sea U un proceso estacionario regular, S un sistema dado tal que $E\{r_n^2\} = 1$, y los procesos sintéticos V y U^* , tales que U^* es el proceso sintético generado por una secuencia de v.a.i.i.d. con densidad marginal igual a la densidad marginal de orden uno $P_r(x)$ del proceso de residuos de U y V un proceso sintético generado por el mismo sistema S y una función de probabilidad $P_z(x)$ cualquiera. Entonces, la siguiente igualdad del triángulo se cumple:

$$\bar{H}(U, V) = \bar{H}(U, U^*) + \bar{H}(U^*, V) \quad (3.5.2)$$

$$= \bar{H}(U, U^*) + H(P_r, P_z) \quad (3.5.3)$$

Por lo que, si el sistema que produce al proceso sintético está fijo y cumple con la condición de que la potencia promedio del proceso residual sea unitaria, el proceso sintético más cercano, según la RCEC, al proceso U , es aquel que se produce al excitar el sistema S por la secuencia de v.a.i.i.d. con función de densidad de probabilidad marginal igual a la del proceso residual.

Demostración de (3.5.2) y (3.5.3)

La demostración se logra directamente al observar que el primero y segundo término de estas ecuaciones coinciden con el primero y segundo término de (3.3.2) (téngase en mente 3.3.1)

Una suposición más realista que el conocer la función de densidad de probabilidad de primer orden del proceso residual, es el suponer conocidas algunas estadísticas marginales de este proceso residual. Por ejemplo, los valores esperados de $f_k(r_n)$ para M funciones distintas (con $f_1(x) = x^2$). Entonces, el estimado de la función de densidad de probabilidad de primer orden del proceso residual, que satisface el principio de máxima entropía [5] o mínima entropía cruzada con una distribución a-priori gaussiana con media cero y variancia unitaria [7], es una función de densidad de probabilidad de la forma (de 2.5.1):

$$\begin{aligned} Q(r_n) &= P(r_n) \exp \left\{ -\lambda_0 - \sum_{k=1}^M \lambda_k f_k(r_n) \right\} \\ &= \frac{1}{\sqrt{2\pi}} \exp \left\{ -\lambda_0 - \frac{f_1(r_n)}{2} - \sum_{k=1}^M \lambda_k f_k(r_n) \right\} \\ &= \exp \left\{ -a_0 - \sum_{k=1}^M a_k f_k(x) \right\} \end{aligned}$$

Por ésto, si sólo se conocen algunas estadísticas marginales del proceso residual, conviene restringir las funciones de densidad de probabilidad de los procesos de excitación del sistema S para construir los procesos sintéticos, en la forma en que se hace en la siguiente proposición:

Proposición 3:

Sean: U un proceso estacionario regular, un sistema S , y M funciones pre-establecidas $f_k(x)$ con $f_1(x) = x^2$. Entonces, si del proceso residual se conocen los valores esperados de $f_k(r_n)$, esto es,

$$E_{p_r} \{f_1(r_n)\} = 1$$

$$E_{P_r} \{f_k(r_n)\} = \bar{f}_k, \quad 1 < k \leq M$$

y la función de densidad de probabilidad del proceso de excitación del sistema S que genera al proceso sintético se restringe a tener la forma paramétrica

$$P_Z(x) = \exp\{-a_0 - \sum_{k=1}^M a_k f_k(x)\},$$

entonces, la RCEC entre U y un proceso sintético V con este tipo de excitación se reduce a

$$\bar{H}(U, V) = \bar{H}(U, V^*) + H(P_Z^*, P_Z) \quad (3.7.1)$$

$$= \bar{H}(U, V^*) + \bar{H}(V^*, V), \quad (3.7.2)$$

donde V^* es el proceso sintético que se genera con el sistema S y con el proceso de v.a.i.i.d. Z^* que tiene función de densidad de probabilidad de primer orden $P_Z^*(x)$, definida por

$$P_Z^*(x) = \exp\{-a_0^* - \sum_{k=1}^M a_k^* f_k(x)\}$$

con $a_0^*, a_1^*, \dots, a_M^*$ tales que

$$E_{P_Z^*} \{f_k(x)\} = \bar{f}_k, \quad 0 \leq k \leq M$$

con $f_0(x) = 1$ y $\bar{f}_0 = \bar{f}_1 = 1$

y

$$\bar{H}(U, V^*) = -\bar{H}_r + \sum_{k=0}^M a_k^* \bar{f}_k \quad (3.7.3)$$

Nótese que el primer término de la ecuación (3.7.3) es una medida del grado de dependencia estadística entre r_n y sus valores anteriores y el segundo término es función de las restricciones.

Demostración:

Por la ecuación (2.4.4),

$$H(P_r, P_z) = H(P_r, P_z^*) + H(P_z^*, P_z).$$

Por lo que sustituyendo esta igualdad en la ecuación (3.3.2)

$$\bar{H}(U, V) = I(r_n; (r_{n-1}, r_{n-2}, \dots)) + H(P_r, P_z^*) + H(P_z^*, P_z).$$

Por la proposición 1, $\bar{H}(U, V^*)$ es igual a la suma de los dos primeros términos de esta última ecuación, de donde se obtiene la ecuación (3.7.1). La ecuación (3.7.2) se sigue inmediatamente de la ecuación (3.7.1) por ser Z^* y Z secuencias de v.a.i.i.d. transformadas por el mismo sistema S .

Observando que usando (3.5.1) $\bar{H}(U, V^*)$ se puede escribir como

$$\bar{H}(U, V^*) = -\bar{H}_r + H(P_r) + H(P_r, P_z^*)$$

y expresando la entropía cruzada entre $P_r(x)$ y $P_z^*(x)$ como

$$H(P_r, P_z^*) = -H(P_r) + \int P_r(x) \ln \left\{ \frac{1}{P_z^*(x)} \right\} dx,$$

de las dos últimas ecuaciones, después de cancelar $H(P_r)$ y sustituir $P_z^*(x)$

para efectuar la integral se obtiene la ecuación (3.7.3).

Corolario:

Sean U^* como en la proposición 2, V y V^* como en la proposición 3; entonces

$$\bar{H}(U, V) = \bar{H}(U, U^*) + \bar{H}(U^*, V)$$

$$\bar{H}(U^*, V) = \bar{H}(U^*, V^*) + \bar{H}(V^*, V)$$

y de estas dos ecuaciones

$$\bar{H}(U, V) = \bar{H}(U, U^*) + \bar{H}(U^*, V^*) + \bar{H}(V^*, V) \quad (3.9.1)$$

De la ecuación (3.9.1) se concluye que la RCEC entre un proceso observado U y un proceso sintético V (producido por una excitación con $P_z(x)$ como en la proposición 3), está compuesta aditivamente por un término debido a que la excitación es estrictamente blanca ($\bar{H}(U, U^*)$); un segundo término debido a que la función de densidad de probabilidad de la excitación no es la residual, $P_z^*(x) \neq P_r(x)$, ($\bar{H}(U^*, V^*)$); y un tercer término debido a que $P_z(x)$ no satisface las restricciones.

Además de los procesos sintéticos producidos por excitaciones como la de la proposición 3, hay otra clase de procesos sintéticos, para los cuales se obtienen resultados de interés teórico y práctico, y es la clase ARMA, cuyos procesos resultan al excitar con una sucesión de v.a.i.i.d., como las definidas en (3.2.1), a un filtro digital, lineal, causal, invariante en el tiempo y de fase mínima, con función de transferencia $G(z)$ dada por:

$$G(z) = \frac{\sigma B(z)}{A(z)}$$

donde

$$B(z) = \sum_{k=0}^M b_k z^{-k}, \quad b_0 = 1$$

$$A(z) = \sum_{k=0}^N a_k z^{-k}, \quad a_0 = 1 \neq$$

Proposición 4:

La RCEC entre un proceso estacionario regular U , con función de densidad espectral $f_u(\omega)$, y un proceso ARMA, se reduce a:

$$\bar{H}(U, V) = -\bar{H}_e - \int_{-\pi}^{\pi} \ln \left\{ \frac{\sigma_u}{\sigma} \right\} + \int P_r(x) \ln \left\{ \frac{1}{P_z(x)} \right\} dx \quad (3.10.1)$$

donde \bar{H}_e es la RCED del proceso de innovaciones $E = \{e_n\}_{n=-\infty}^{\infty}$ asociado al proceso U , el cual está definido por [5]:

$$e_n = (u_n - \hat{u}_n) / \sigma_u$$

con \hat{u}_n igual al mejor predictor lineal (en el sentido del error cuadrático medio) de u_n , dados todos sus valores anteriores, y σ_u^2 es el error cuadrático medio del mejor predictor lineal, de un paso, dado todo su pasado.

Esto es [5, pág. 424]:

$$\sigma_u^2 = E\{(u_n - \hat{u}_n)^2\} = \exp \int_{-\pi}^{\pi} \ln(f_u(\omega)) \frac{d\omega}{2\pi} \quad (3.10.2)$$

≠ Nótese que por la forma de $B(z)$ y $A(z)$, $\lim_{z \rightarrow \infty} G(z) = \sigma$, la cual es una constante finita y diferente de cero, como debe de ser para sistemas de fase mínima [14, pág. 347].

Demostración :

Por ser U la respuesta del filtro de innovación excitado por el proceso de innovaciones [5, pág. 534]

$$\bar{H}_U = \bar{H}_e + \frac{1}{2} \ln \sigma_u^2$$

Como $G(z)$ tiene todos sus polos y ceros dentro del círculo unitario por ser de fase mínima, es invertible. Por ser U la respuesta del filtro $G(z)$ excitado por el proceso residual

$$\bar{H}_U = \bar{H}_r + \frac{1}{2} \ln \sigma^2$$

Combinando las dos últimas ecuaciones con (3.4.1), se obtiene (3.10.1)

Proposición 5 :

Sea un proceso regular U con función de densidad de potencia espectral $f_U(\omega)$ y un proceso ARMA V^* producido por una excitación con densidad de probabilidad marginal igual a la del proceso residual y por un sistema S lineal con función de transferencia $G(z)$ con ganancia σ tal que

$$\sigma^2 = \alpha = \int_{-\pi}^{\pi} \left| \frac{A(e^{j\omega})}{B(e^{j\omega})} \right|^2 f_U(e^{j\omega}) \frac{d\omega}{2\pi} \quad (3.11.1)$$

o sea, un proceso ARMA, tal que la potencia promedio del proceso residual sea unitaria. Entonces se aplica la siguiente cota superior:

$$\bar{H}(U, V^*) \leq \bar{H}(E, W) + 1/2 \ln \left(\frac{\alpha}{\sigma_u^2} \right) \quad (3.11.2)$$

$$\leq 1/2 \ln \left(\frac{\alpha}{\sigma_u^2} \right) \quad (3.12.1)$$

con igualdad en (3.11.2) y (3.12.1) si y sólo si el proceso U es gaussiano y donde E y W representan el proceso de innovaciones de U y un proceso de ruido blanco gaussiano con potencia unitaria, respectivamente.

Para caracterizar un proceso sintético que minimice esta cota superior es suficiente encontrar los coeficientes de $A(z)$ y $B(z)$ que minimizan (3.12.1), lo que equivale a minimizar α definido en (3.11.1). El doble del miembro derecho de la desigualdad (3.12.1) recibe el nombre de distorsión de Itakura-Saito con ganancia optimizada [4], y es una medida de similitud entre los espectros de potencia normalizados del proceso observado y el proceso patrón o sintético. Este tipo de distorsión ha sido usado en situaciones en las cuales la forma de onda del espectro, y no la ganancia, debe ser clasificada o reproducida.

Demostración:

Si la potencia promedio de los residuos es unitaria, entonces la ganancia del filtro ARMA está definida por (3.11.1), y si el proceso de excitación del filtro ARMA tiene la misma densidad de probabilidad que la marginal de primer orden del proceso residual, entonces la ecuación (3.10.1) puede ser reescrita como:

$$\bar{H}(U, V^*) = -\bar{H}_e + \frac{1}{2} \ln \left(\alpha / \sigma_u^2 \right) + H(P_r) \quad (3.12.2)$$

y debido a que $E\{r_n^2\} = 1$, es posible acotar superiormente a la entropía dife

rencial $H(P_r)$, por la entropía diferencial de la densidad gaussiana con variancia unitaria, ésto es:

$$H(P_r) \leq 1/2 \ln(2\pi e)$$

con igualdad, si y sólo si $P_r(x)$ es la densidad de probabilidad gaussiana. Sustituyendo esta expresión en (3.12.2) y observando que si W representa a un proceso de ruido blanco gaussiano con potencia unitaria, entonces $\bar{H}(E, W)$ es igual a:

$$\bar{H}(E, W) = -\bar{H}_e + 1/2 \ln(2\pi e)$$

con lo que la proposición 5 queda demostrada.

Proposición 6:

La RCEC entre un proceso regular U , con función de densidad de potencia espectral $f_u(\omega)$ y un proceso sintético gaussiano ARMA V , está dada por

$$\bar{H}(U, V) = \bar{H}(E, W) + \frac{1}{2} d_{IS}(f_u(e^{j\omega}), f_v(e^{j\omega})), \quad (3.13.1)$$

donde $d_{IS}(f_u(e^{j\omega}), f_v(e^{j\omega}))$ es la distorsión de Itakura-Saito definida como el doble del miembro derecho de la ecuación (2.7.2) y $f_v(e^{j\omega})$ es la función de densidad de potencia espectral del proceso sintético ARMA definida por

$$f_v(e^{j\omega}) = \sigma^2 \left| \frac{B(e^{j\omega})}{A(e^{j\omega})} \right|^2$$

Demostración:

De la ecuación (3.10.1) con $P_z(x)$ igual a la densidad gaussiana, y después de sumar y restar $(1/2)$ al miembro derecho de (3.10.1):

$$\bar{H}(U,V) = -\bar{H}_e - \frac{1}{2} \ln\left(\frac{\sigma_u^2}{\sigma^2}\right) + \frac{1}{2} \ln(2\pi e) - \frac{1}{2} + \frac{E\{r_n^2\}}{2}$$

y reescribiendo $E\{r_n^2\} = \int_{-\pi}^{\pi} \frac{f_u(e^{j\omega})}{f_v(e^{j\omega})} \frac{d\omega}{2\pi}$

se obtiene la ecuación (3.13.1) (téngase en mente 2.9.2).

Nótese que si U es gaussiano, E coincide con W y la RCEC entre procesos gaussianos se reduce, como se dijo antes (ecuación (2.7.2)), a la mitad de la distorsión de Itakura-Saito.

La importancia de estas dos últimas proporciones, radica en el hecho de que en un gran número de situaciones, del proceso U a clasificar o a simular, tan sólo se conocen estadísticas de segundo orden, y quizás distribuciones marginales de primer orden, por lo que si se conoce la estadística de segundo orden del proceso U , es posible seleccionar un proceso sintético que minimice la cota superior de la proposición 5 (suponiendo conocida la función de densidad de probabilidad marginal del proceso residual), o el proceso gaussiano más cercano, según la RCEC, al proceso observado.

Esto es posible debido a que en los dos casos el primer término del miembro derecho de (3.11.2), y de (3.13.1), es independiente de la estadística de segundo orden de los procesos involucrados, mientras que los segundos términos

en dichas ecuaciones, son función exclusivamente de la estadística de segundo orden de los procesos involucrados. (Véase el apéndice B.6)

Finalmente, en situaciones en las cuales los procesos sintéticos son generados con un modelo (filtro) ARMA, y bajo la restricción de que sólo es posible observar algunas estadísticas marginales del proceso residual, se obtiene la siguiente proposición:

Proposición 7 :

Para un proceso sintético ARMA y bajo las mismas suposiciones que en la proposición 3, se obtienen los mismos resultados que en la proposición 3 y la RCEC entre U y el proceso sintético V^* (ecuación (3.7.3)), se reduce a:

$$\bar{H}(U, V^*) = -\bar{H}_e + \frac{1}{2} \ln\left(\frac{\alpha}{\sigma_u^2}\right) + \sum_{k=0}^M a_k^* \bar{f}_k$$

con $\bar{f}_0 = \bar{f}_1 = 1$ y α definida por (3.11.1).

Demostración:

La demostración de esta proposición es inmediata, usando la ecuación (3.7.3) y $\bar{H}_r = \bar{H}_e + \ln\left(\frac{\sigma_u}{\sigma}\right)$.

CAPITULO 4

APLICACIONES A SIMULACION, CLASIFICACION Y DISEÑO DEL VOCODER.

Simulación de Procesos: Se analiza a continuación el problema de generar un proceso sintético que esté lo más cercano posible a un proceso observado U , según el criterio de fidelidad empleado en éste trabajo, ésto es, según la información discriminatoria promedio normalizada.

De acuerdo a la proposición 1, para que exista un proceso sintético V' , tal que $\bar{H}(U, V') = 0$, se requiere que $E\{r_n^2\} = 1$, ya que al definir los procesos sintéticos se restringió la variancia de la excitación a ser unitaria. Por tanto, el sistema S^{-1} debe incluir una ganancia γ de tal modo que se cumpla esta condición. Pero [16, pág. 31; 17 pág. 145].

$$I(\gamma r_n; (\gamma r_{n-1}, \gamma r_{n-2}, \dots)) = I(r_n; (r_{n-1}, r_{n-2}, \dots))$$

y por ende, dicho escalamiento no afecta en ninguna forma los resultados de la proposición 1 del capítulo 3.

Sin embargo, caracterizar el sistema S que produzca el proceso sintético óptimo V' es un problema abierto, para el cual se conoce la solución única_{mente} en contadas excepciones; por ejemplo, si el proceso U es gaussiano. Sin embargo, si se restringen los procesos sintéticos V a pertenecer a la clase de procesos ARMA gaussianos, es posible encontrar el proceso V^+ tal que

$$\bar{H}(U, V^+) = \min \bar{H}(U, V)$$

donde la minimización es sobre la clase mencionada. Para ello, basta minimizar la distorsión de Itakura-Saito en (3.13.1), lo que es equivalente a minimizar α definida en (3.11.1) [4].

Si no se limitan los procesos sintéticos a ser gaussianos, pero se conoce la función de densidad de primer orden del proceso residual, se procede a minimizar la cota (3.11.2), lo que requiere nuevamente minimizar α .

La minimización de α es en general un problema de optimización no lineal. Sin embargo, para procesos AR, se reduce a un problema de optimización lineal.

De la definición (3.11.1) :

$$\alpha = \int_{-\pi}^{\pi} |A(e^{j\omega})|^2 f_u(e^{j\omega}) \frac{d\omega}{2\pi} = \sum_{k=0}^N \sum_{\ell=0}^N a_k a_{\ell} R_u[k-\ell] \quad (4.2.1)$$

$$= R_u[0] R_a[0] + 2 \sum_{n=1}^N R_u[n] R_a[n] \quad (4.2.2)$$

donde

$$R_a[n] = \sum_{k=0}^{N-n} a_k a_{k+n}, \quad \text{para } n = 0, 1, \dots, N \quad (4.2.3)$$

$$R_u[k] = \int_{-\pi}^{\pi} e^{j\omega k} f_u(e^{j\omega}) \frac{d\omega}{2\pi}, \quad k = 0, 1, \dots, N \quad (4.2.4)$$

son la correlación de los coeficientes del filtro y los coeficientes de auto-

correlación del proceso observado U , respectivamente.

En (4.2.1), se observa que α es el error cuadrático medio de un predictor lineal de un paso hacia adelante de orden N [5, pág. 430] .

Se saben que, para procesos AR, la α mínima ocurre cuando los coeficientes del filtro son la solución del sistema de ecuaciones normales (ó ecuaciones de Yule-Walker) [4] :

$$\begin{bmatrix} R_u[0] & R_u[1] & \dots & R_u[N-1] \\ R_u[1] & R_u[0] & \dots & R_u[N-2] \\ \vdots & \vdots & \ddots & \vdots \\ R_u[N-1] & & & R_u[0] \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} R_u[1] \\ R_u[2] \\ \vdots \\ R_u[N] \end{bmatrix} \quad (4.3.1)$$

A continuación, denotando con σ_N a la α mínima de un proceso AR de orden N y con $A_N(z)$ al polinomio $A(z)$ formado con los coeficientes a_1, \dots, a_N que satisfacen las ecuaciones de Yule-Walker, se resumen 4 propiedades de los procesos AR :

- 1] El proceso residual que resulta al filtrar el proceso U con el sistema de función de transferencia $\frac{A_N(z)}{\sigma_N}$, en general, no es una sucesión de v.a.i.i.d. Obsérvese, sin embargo, que $A_N(z)$ es la función de transferencia del filtro de error y por tanto [5, pág. 431], r_n es ortogonal a r_{n-1}, \dots, r_{n-N} . Se concluye que si $N \rightarrow \infty$, el proceso residual será una sucesión de variables aleatorias ortogonales e idénticamente distri

buídas [5, pág. 439], con distribución marginal de primer orden igual a la del proceso de innovaciones del proceso U. (apéndices B.4, B.5 y B.6).

- 2] Cualquier proceso sintético generado con un filtro de función de transferencia $\frac{\sigma_N}{A_N(z)}$, tiene la propiedad de que sus primeros N+1 coeficientes de autocorrelación son idénticos a los primeros N+1 coeficientes de autocorrelación del proceso observado, ésto es [5, pág. 437] :

$$R_V [k] = R_U [k] , \quad k = 0, 1, \dots, N$$

- 3] La densidad espectral del proceso sintético generado por $\frac{\sigma_N}{A_N(z)}$, la cual es independiente del proceso de excitación, es la estimación de la densidad espectral del proceso observado U que maximiza la entropía, dado que sólo se conocen los primeros N+1 momentos de segundo orden del proceso U [5, pág. 448].

- 4] Se sabe que [5, págs. 430, 424, 436] :

$$\sigma_N^2 = \sum_{i=0}^N a_i R_U [i] \quad ; \quad a_0 = 1$$

$$\sigma_U^2 = \sum_{i=0}^{\infty} a_i R_U [i] \quad ; \quad a_0 = 1$$

$$\sigma_M^2 \geq \sigma_{M+1}^2 \geq \dots \geq \sigma_N^2 ; \quad \lim_{N \rightarrow \infty} \sigma_N^2 = \sigma_U^2 .$$

Por lo que, incrementar el orden del filtro puede únicamente disminuir el término $\frac{1}{2} \ln \left(\frac{\sigma_N^2}{\sigma_U^2} \right)$ en (3.11.2), y en el caso límite, cuando N tiende a ∞ , dicho término es cero.

CLASIFICACION DE SEÑALES

Sea U un proceso observado y $C = \{ V_i \}_{i=1}^L$ la colección de procesos sintéticos que representan los patrones a clasificar. La clasificación del proceso U en el patrón V_j (la cual será representada por $U \rightarrow V_j$), por el método del vecino más cercano usando la RCEC como criterio de fidelidad, se lleva a cabo cuando

$$\bar{H}(U, V_j) \leq \bar{H}(U, V_i) \quad \forall V_i \in C \quad (4.5.1)$$

Si los procesos sintéticos V_i son procesos ARMA generados por sistemas con función de transferencia

$$G(z)_i = \sigma_i \left[\frac{B(z)}{A(z)} \right]_i, \quad i = 1, 2, \dots, L$$

excitados con procesos caracterizados por la función de densidad de probabilidad de primer orden $P_{Z_i}(x)$, $i = 1, 2, \dots, L$, la condición (4.5.1), después de simplificar ambos miembros de la desigualdad, se reduce a (de 3.10.1):

$$\begin{aligned} U \rightarrow V_j, \text{ si: } & \ln \sigma_j - \int P_{R_j}(x) \ln P_{Z_j}(x) dx \\ & \leq \ln \sigma_i - \int P_{R_i}(x) \ln P_{Z_i}(x) dx, \quad i = 1, 2, \dots, L \end{aligned} \quad (4.5.2)$$

donde $P_{R_i}(x)$ son las funciones de densidad de probabilidad marginal de primer orden de los procesos residuales que resultan al pasar el proceso U por los filtros con función de transferencia $G^{-1}(z)_i$.

De la condición (4.5.2) se puede observar que la decisión para clasificar la

señal observada es función de las densidades de probabilidad $P_{r_j}(x)$, las cuales a su vez son funciones de toda la estadística de los procesos residuales que resultan al pasar el proceso U por cada uno de los sistemas $G^{-1}(z)_i$, o de toda la estadística del proceso U y de los coeficientes que definen las funciones de transferencia $G(z)_i$. Debido a que las propiedades de segundo orden de los procesos aleatorios tienen interpretaciones físicas interesantes, resulta conveniente hacer esta decisión exclusivamente en base a la estadística de segundo orden del proceso observado. Esto se logra si los patrones de referencia son procesos gaussianos, o equivalentemente, si $P_{z_i}(x) = P_w(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad \forall i$, lo cual simplifica el criterio de decisión a:

$$U \rightarrow V_j \quad \text{si} \quad E\{r_{j_n}^2\} + \lambda n \sigma_j^2 \leq E\{r_{i_n}^2\} + \lambda n \sigma_i^2, \quad i=1, \dots, L \quad (4.6.1)$$

donde $E\{r_{i_n}^2\}$ es la potencia promedio del proceso residual que resulta al pasar el proceso observado U por los sistemas $G^{-1}(z)_i$, ésto es,

$$E\{r_{i_n}^2\} = \int_{-\pi}^{\pi} |G^{-1}(e^{j\omega})_i|^2 f_U(e^{j\omega}) \frac{d\omega}{2\pi}$$

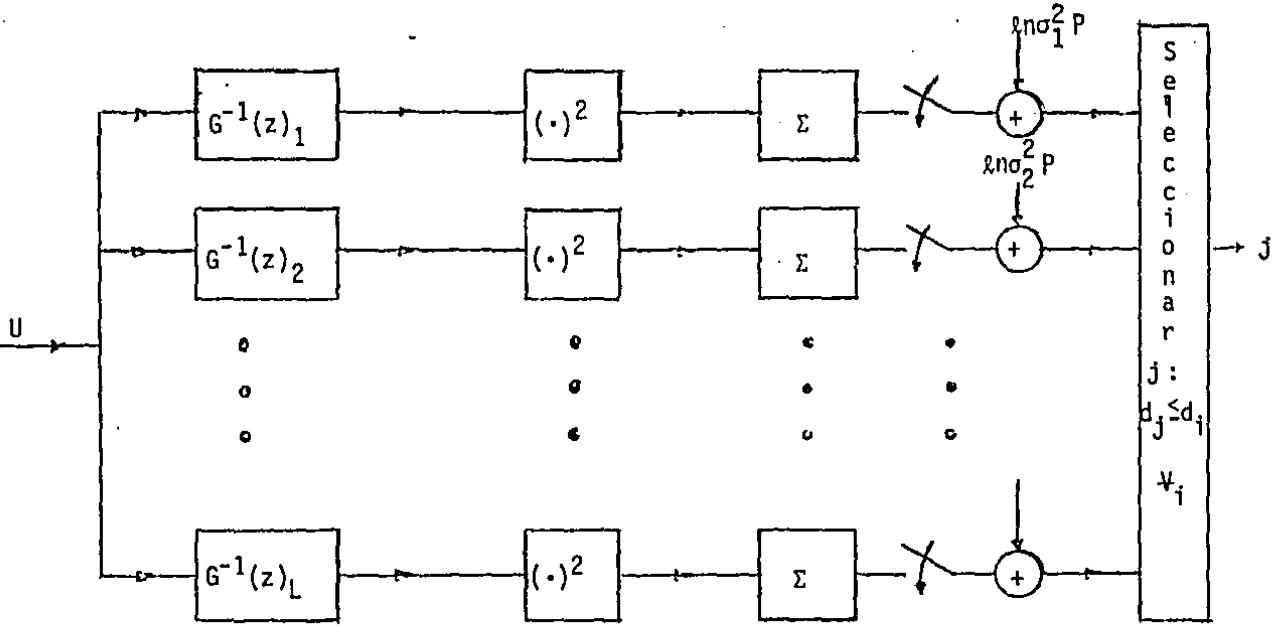
Es importante notar que en el criterio de decisión dado en (4.6.1), no hay ninguna restricción sobre la estadística del proceso observado U , aparte de la estacionariedad del mismo. En otras palabras, la condición (4.6.1) dice la forma de seleccionar dentro de una colección de procesos gaussianos ARMA, aquél que esté lo más cercano al proceso observado U según la RCEC.

Para llevar a cabo la clasificación según la regla (4.6.1), los únicos parámetros que son necesarios calcular son los del conjunto $E\{r_{i_n}^2\}$, $i=1, \dots, L$.

Estos, en un caso práctico, donde sólo se conoce un conjunto de muestras $\{u_\ell\}_{\ell=1}^P$ de una realización del proceso U , pueden ser aproximados suponiendo ergodicidad, por medio de

$$E\{r_{i_n}^2\} = \int x^2 P_{r_i}(x) dx \approx \frac{1}{P} \sum_{\ell=1}^P r_{i_\ell}^2 \quad (4.7.1)$$

donde las secuencias $\{r_{i_\ell}\}_{\ell=1}^P$, $i=1, \dots, L$, son las muestras que resultan al pasar la secuencia $\{u_\ell\}_{\ell=1}^P$ por los filtros con función de transferencia $G^{-1}(z)_i$, por lo que esquemáticamente la identificación se realiza según lo indica la siguiente figura



La selección de los filtros $\left[\frac{B(z)}{A(z)} \right]_i$ y el valor de sus ganancias σ_i , que son usados en el sistema clasificador, se lleva a cabo con el siguiente criterio: Supóngase que para seleccionar cada uno de los procesos patrón, se cuenta con una colección de K realizaciones del proceso a caracterizar, las cua

les se representarán por u^k , donde el superíndice k denota la k -ésima realización del mismo proceso. Entonces, una opción para seleccionar el patrón, es encontrar el proceso sintético que en promedio esté lo más cercano, según la RCEC, a la colección de realizaciones $\{u^k\}$. Restringiéndonos a procesos sintéticos gaussianos, el problema se reduce a encontrar los coeficientes $(a_1, a_2, \dots, a_N; b_1, b_2, \dots, b_M)$ del filtro $\frac{B(z)}{A(z)}$, con $b_0 = a_0 = 1$, que minimicen la expresión [4]

$$\beta = \frac{1}{K} \sum_{k=1}^K \int_{-\pi}^{\pi} \left| \frac{A(e^{j\omega})}{B(e^{j\omega})} \right|^2 f_{u^k}(e^{j\omega}) \frac{d\omega}{2\pi} = \int_{-\pi}^{\pi} \left| \frac{A(e^{j\omega})}{B(e^{j\omega})} \right|^2 \left\{ \frac{1}{K} \sum_{k=1}^K f_{u^k}(e^{j\omega}) \right\} \frac{d\omega}{2\pi}$$

(donde $f_{u^k}(e^{j\omega})$ es la densidad espectral estimada con cada una de las realizaciones), y asignar a σ^2 el valor mínimo de β así obtenido.

$$\sigma^2 = \beta_{\text{mín}}$$

Entonces el problema es similar a la minimización de α en el caso de simulación, con la diferencia de que en lugar de usar la densidad del espectro de potencia $f_U(e^{j\omega})$ de un proceso en particular, se debe usar el promedio de las densidades espectrales estimadas con cada una de las realizaciones. En el caso particular de que $M = 0$ y $N \neq 0$, el problema se reduce a uno de optimización lineal que consiste en resolver un sistema de ecuaciones lineales idéntico al sistema mostrado en (4.3.1) con la diferencia de que en lugar de trabajar con las autocorrelaciones de una realización en particular, se usa el promedio aritmético de las autocorrelaciones de las K realizaciones del proceso U .

DISEÑO DE VOCODER PARA SIMULACION DE PROCESOS POR MEDIO DE LA RCEC:

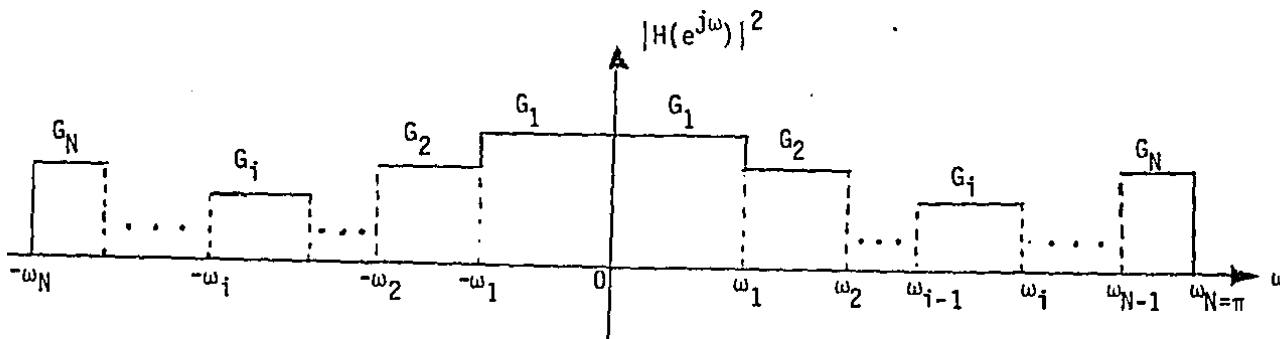
Las técnicas anteriores pueden aplicarse al diseño del vocoder de canal ("codificador de canal de voz"), cuando la amplitud de la función de transferencia $H(e^{j\omega})$ del vocoder no es cero en ningún punto del intervalo $[-\pi, \pi]$. Esto garantiza que se satisfaga la condición de Paley-Wiener [18, pág. 229]:

$$\int_{-\pi}^{\pi} \ln |H(e^{j\omega})| d\omega < \infty$$

y por tanto, que existe al menos una fase para la cual el vocoder es causal.

Se requerirá que el vocoder también sea estable, y que el sistema inverso al vocoder también sea estable y causal. Esto es, se pedirá que el vocoder sea de fase mínima. Entonces la fase está determinada en forma única y está dada por [14, pág. 346]:

$$\arg[H(e^{j\omega})] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln |H(e^{j\theta})| \cot\left(\frac{\theta - \omega}{2}\right) d\theta$$



"El vocoder de canal". (Cada filtro G_i no es causal)

Ya que la discontinuidad de $|H(e^{j\omega})|$ en un número finito de puntos no es un problema para el cálculo de éste argumento [19, Ej. 10-9, pág.208], para un vocoder similar al ilustrado en la figura se encuentra que la fase asociada para que sea de fase mínima está dada por:

$$\angle H(e^{j\omega}) = \frac{1}{2\pi} \ln \left\{ \prod_{i=1}^N \left(\frac{\tan \frac{\omega}{2} + \tan \frac{\omega_{i-1}}{2}}{\tan \frac{\omega}{2} - \tan \frac{\omega_{i-1}}{2}} \right) \left(\frac{\tan \frac{\omega}{2} - \tan \frac{\omega_i}{2}}{\tan \frac{\omega}{2} + \tan \frac{\omega_i}{2}} \right) \right\}$$

(Nótese que esta fase es una función impar).

Ahora, para un sistema de fase mínima [5, pág. 474]

$$\angle n|h[0]| = \frac{1}{2\pi} \int_{-\pi}^{\pi} \angle n|H(e^{j\omega})| d\omega = \frac{1}{2\pi} \sum_{i=1}^N (\angle n G_i)(\omega_i - \omega_{i-1}) \quad (4.10.1)$$

ya que $|H(e^{j\omega})| = \sqrt{G_i}$ para $\omega_{i-1} < \omega < \omega_i$. $h[0]$ representa la respuesta al pulso unitario del filtro $H(e^{j\omega})$ en $n=0$. Se puede pensar en el vocoder como un modelo ARMA de dimensión infinita; por esto, $h[0]$ es igual a σ en la ecuación (3.10.1), la cual se reescribe a continuación:

$$\bar{H}(U, V) = -\bar{H}_e - \ln \left\{ \frac{\sigma_u}{h_0} \right\} + \int P_r(x) \ln \left\{ \frac{1}{P_z(x)} \right\} dx \quad (4.10.2)$$

donde se supone que el vocoder está siendo excitado con un proceso estrictamente blanco.

Si $P_z(x) = P_r(x)$, (3.11.2) y (3.11.1) pueden ser reescritas, respectivamente:

$$\bar{H}(U, V^*) \leq \bar{H}(E, W) + \frac{1}{2} \ln \left\{ \frac{h_0^2}{\sigma_u^2} \right\} \quad (4.10.3)$$

y

$$i = \int_{-\pi}^{\pi} \frac{f_u(e^{j\omega})}{|H(e^{j\omega})|^2} \frac{d\omega}{2\pi} \quad (4.11.1)$$

Si $P_z(x) = P_w(x)$, (3.13.1) puede ser reescrita como

$$\bar{H}(U, V) = \bar{H}(E, W) + \frac{1}{2} d_{IS}(f_u(e^{j\omega}), |H(e^{j\omega})|^2) \quad (4.11.2)$$

Nótese que el vocoder es función de las constantes G_i , $i=1, \dots, N$, y ω_i , $i=1, \dots, N$. Se procede a continuación a minimizar (4.10.3) sujeto a (4.11.1), y posteriormente a minimizar (4.11.2), ambas con respecto a las constantes G_i , para lo cual se suponen dadas las constantes ω_i .

Minimización de la cota (4.10.3), sujeta a (4.11.1):

Se requiere minimizar (4.10.1) sujeta a (4.11.1) suponiendo fijos los intervalos de frecuencia (ω_{i-1}, ω_i) , $i=1, \dots, N$.

Aplicando el método de multiplicadores de Lagrange:

$$\begin{aligned} L &= \frac{1}{2\pi} \sum_{i=1}^N (\ln G_i)(\omega_i - \omega_{i-1}) + \lambda \left(1 - \int_{-\pi}^{\pi} \frac{f_u(e^{j\omega})}{|H(e^{j\omega})|^2} \frac{d\omega}{2\pi} \right) \\ &= \frac{1}{2\pi} \sum_{i=1}^N (\ln G_i)(\omega_i - \omega_{i-1}) + \lambda \left(1 - \sum_{i=1}^N \int_{\omega_{i-1}}^{\omega_i} \frac{f_u(e^{j\omega})}{G_i} \frac{d\omega}{\pi} \right); \end{aligned}$$

tomando parciales con respecto a las G_i e igualando a cero

$$\frac{\partial L}{\partial G_i} = \frac{1}{2\pi} \left(\frac{\omega_i - \omega_{i-1}}{G_i} \right) + \frac{\lambda}{\pi} \int_{\omega_{i-1}}^{\omega_i} \frac{f_u(e^{j\omega})}{G_i^2} d\omega = 0$$

de donde

$$G_i = -2\lambda \int_{\omega_{i-1}}^{\omega_i} \frac{f_u(e^{j\omega})}{\omega_i - \omega_{i-1}} d\omega \quad (4.12.1)$$

Sustituyendo las G_i dadas por (4.12.1) en la restricción, se obtiene el valor de $\lambda = -1/2$. Usando este valor en (4.12.1):

$$G_i = \int_{\omega_{i-1}}^{\omega_i} \frac{f_u(e^{j\omega})}{\omega_i - \omega_{i-1}} d\omega \quad (4.12.2)$$

Minimización de (4.11.2) :

Suponiendo fijos los intervalos de frecuencia (ω_{i-1}, ω_i) ,

$$\begin{aligned} d_{IS}(f_u(e^{j\omega}), |H(e^{j\omega})|^2) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \frac{f_u(e^{j\omega})}{|H(e^{j\omega})|^2} - \ln \frac{f_u(e^{j\omega})}{|H(e^{j\omega})|^2} - 1 \right\} d\omega \\ &= \frac{1}{\pi} \sum_{i=1}^N \int_{\omega_{i-1}}^{\omega_i} \left\{ \frac{f_u(e^{j\omega})}{G_i} - \ln \frac{f_u(e^{j\omega})}{G_i} - 1 \right\} d\omega \end{aligned}$$

Derivando la expresión anterior con respecto a las G_i e igualando a cero

$$\int_{\omega_{i-1}}^{\omega_i} \left\{ -\frac{f_u(e^{j\omega})}{G_i^2} + \frac{1}{G_i} \right\} d\omega = 0, \text{ de donde}$$

se obtiene nuevamente (4.12.2). Obsérvese que este valor es el que se asigna comúnmente a las G_i y que es función de las ω_i .

Se concluye que para minimizar (4.10.3) sujeto a (4.11.1), o minimizar (4.11.2), dadas las frecuencias ω_i , $i=1, \dots, N$, se debe usar (4.12.2). Además, (4.12.2) implica que

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |H(e^{j\omega})|^2 d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_U(e^{j\omega}) d\omega \quad (4.13.1)$$

Obsérvese también que con estas G_i , (4.11.2) se reduce al miembro derecho de (4.10.3), ésto es, la d_{IS} se reduce a la "distorsión de Itakura-Saito con ganancia optimizada".

Entonces, para minimizar ambas expresiones, dadas las G_i óptimas por (4.12.2), con respecto a las ω_i , basta minimizar $h[0]$.

Minimización de $h[0]$ (con respecto a $\omega_1, \dots, \omega_N$):

Se desea minimizar $h[0]$, o lo que es lo mismo, minimizar $\ln|h[0]|^{2\pi}$.

La función objetivo es

$$\ln|h[0]|^{2\pi} = \sum_{i=1}^N (\ln(G_i)) (\omega_i - \omega_{i-1}) \quad (4.13.2)$$

y las restricciones son:

$$\omega_0 = 0, \omega_N = \pi, G_i = \int_{\omega_{i-1}}^{\omega_i} \frac{f_U(e^{j\omega})}{\omega_i - \omega_{i-1}} d\omega$$

$$y \quad -(\omega_i - \omega_{i-1}) \leq 0.$$

Puede mostrarse que la solución está dentro de la frontera, y el problema se reduce a:

$$\frac{\partial \ln|h[0]|^{2\pi}}{\partial \omega_i} = 0, \quad i = 1, \dots, N-1 \quad (4.14.1)$$

donde deberá verificarse que las ω_i resulten tales que $0 < \omega_1 < \omega_2 < \dots < \omega_N = \pi$

(4.14.1) se traduce a

$$\frac{\partial}{\partial \omega_i} \left\{ \ln \left[\frac{\int_{\omega_{i-1}}^{\omega_i} f_U(e^{j\omega}) d\omega}{\omega_i - \omega_{i-1}} \right] (\omega_i - \omega_{i-1}) + \ln \left[\frac{\int_{\omega_i}^{\omega_{i+1}} f_U(e^{j\omega}) d\omega}{\omega_{i+1} - \omega_i} \right] (\omega_{i+1} - \omega_i) \right\} = 0 \quad (4.14.2)$$

Después de derivar y simplificar se obtiene

$$\frac{f_U(e^{j\omega_i})}{G_i} + \ln G_i = \frac{f_U(e^{j\omega_i})}{G_{i+1}} + \ln G_{i+1} \quad (4.14.3)$$

lo cual puede escribirse en la forma

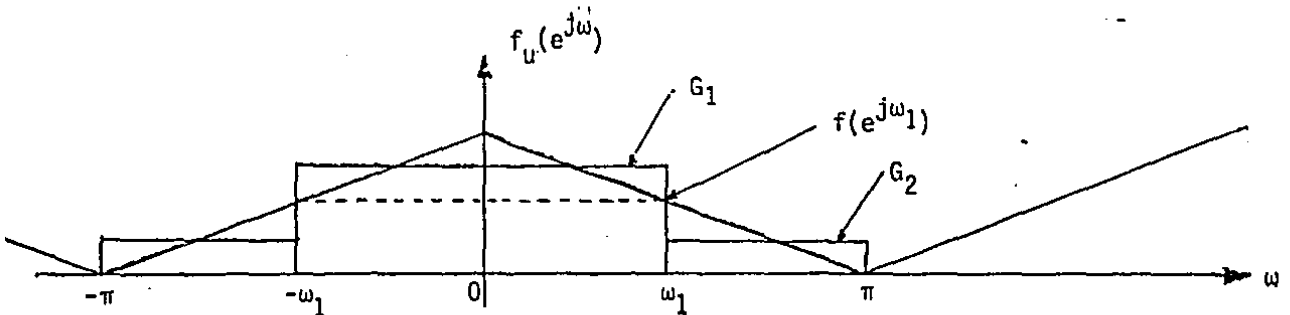
$$\frac{f_U(e^{j\omega_i})}{G_i} - \ln \frac{f_U(e^{j\omega_i})}{G_i} - 1 = \frac{f_U(e^{j\omega_i})}{G_{i+1}} - \ln \frac{f_U(e^{j\omega_i})}{G_{i+1}} - 1 \quad (4.14.4)$$

de donde puede darse una interpretación en términos de la distorsión de Itakura-Saito que se ilustra en el siguiente ejemplo, así como la idea general para resolver (4.14.3) numéricamente. Cabe aclarar que (4.14.3) es sólo una condición necesaria para el mínimo y no se ha probado aquí que sea suficiente:

Ejemplo:

$$\text{Sea } f_u(e^{j\omega}) = \begin{cases} -\omega/\pi + 1, & \omega \in [0, \pi] \\ \omega/\pi + 1, & \omega \in [-\pi, 0] \\ f(e^{j(\omega + n2\pi)}) \end{cases}$$

la cual tiene la siguiente representación gráfica:



y que será aproximada por un vocoder de dos intervalos, ésto es, $\omega_0 = 0$,

$$\omega_1 = x, \omega_2 = \pi.$$

La interpretación sugerida por (4.14.4) es que debe cumplirse $d_{IS}(f(e^{jx}), G_1) = d_{IS}(f(e^{jx}), G_2)$.

Debido a (4.12.2):

$$G_1 = \frac{1+y}{2}; \quad G_2 = \frac{y}{2}; \quad \text{donde } y \stackrel{\Delta}{=} f(e^{jx}).$$

estas relaciones implican que $d_{IS}(f(e^{j\omega_1}), G_2)$ es constante y está dada por

$$d_{IS}(y, y/2) = 2 - \ln 2 - 1 = 0.306853 = c$$

así que debe cumplirse:

$$d_{IS}\left(y, \frac{1+y}{2}\right) = c$$

de donde

$$\frac{2y}{1+y} - \ln\left(\frac{2y}{1+y}\right) = 1 + c;$$

ésta última ecuación se resuelve numéricamente y se encuentra que la raíz es $y = 0.255$.

Como $x = \pi(y-1)$, se obtuvo $\omega_1 = 0.745\pi$

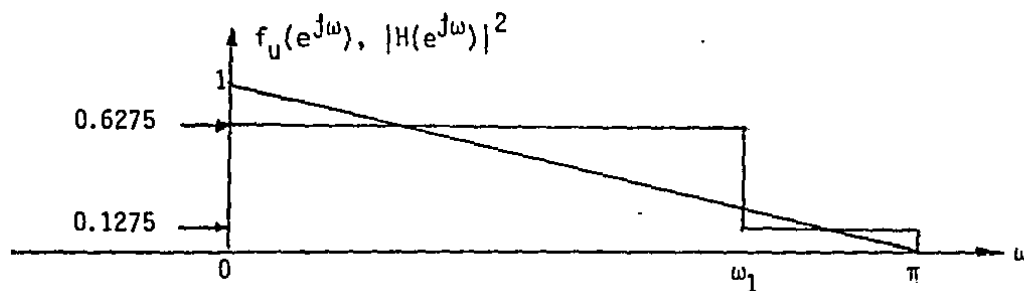


Fig. "El vocoder óptimo de 2 intervalos para una línea recta, en el sentido de minimizar la cota superior de la RCEC (4.10.3) sujeta a (4.11.1), ó minimizar (4.11.2)".

Resulta interesante comparar el vocoder óptimo en el sentido de la RCEC, con aquel que hubiera resultado si el criterio de diseño fuera minimizar el error cuadrático entre la densidad espectral de potencia $f_u(e^{j\omega})$ y su aproximación por el vocoder; para este ejemplo, ésto es minimizar

$$E.C. = \frac{1}{\pi} \int_0^{a\pi} \left(1 - \frac{\omega}{\pi} - G_1 \right)^2 d\omega + \frac{1}{\pi} \int_{a\pi}^{\pi} \left(1 - \frac{\omega}{\pi} - G_2 \right)^2 d\omega$$

$$\text{con } G_1 = \frac{1 + (1-a)}{2}, \quad G_2 = \frac{(1-a)}{2}, \quad (4.17.1)$$

lo cual resulta en $a = 1/2$.

Se procede ahora a comparar las distorsiones; para ello,

$$\begin{aligned} d_{IS} &= \frac{1}{\pi} \int_0^{a\pi} \left\{ \frac{(1-\omega/\pi)}{G_1} - \ln \frac{1-\omega/\pi}{G_1} - 1 \right\} d\omega + \frac{1}{\pi} \int_{a\pi}^{\pi} \left\{ \frac{1-\omega/\pi}{G_2} - \ln \frac{(1-\omega/\pi)}{G_2} - 1 \right\} d\omega \\ &= \frac{1}{\pi} \int_0^{a\pi} \left\{ \frac{(1-\omega/\pi)}{G_1} + \ln G_1 \right\} d\omega + \frac{1}{\pi} \int_{a\pi}^{\pi} \left\{ \frac{(1-\omega/\pi)}{G_2} + \ln G_2 \right\} d\omega \\ &\quad - \frac{1}{\pi} \int_0^{\pi} \ln(1-\omega/\pi) d\omega - 1 \end{aligned}$$

haciendo el cambio $x = +\omega/\pi$,

$$d_{IS} = \int_0^a \left\{ \frac{1-x}{G_1} + \ln G_1 \right\} dx + \int_a^1 \left\{ \frac{(1-x)}{G_2} + \ln G_2 \right\} dx - \int_0^1 \ln(1-x) dx - 1$$

los últimos dos términos suman cero y así:

$$d_{IS} = a \ln G_1 + (1-a) \ln G_2 - \left\{ \frac{(1-x)^2}{2G_1} \Big|_0^a + \frac{(1-x)^2}{2G_2} \Big|_a^1 \right\}$$

$$d_{IS} = a \ln G_1 + (1-a) \ln G_2 - \left\{ \frac{(1-a)^2}{2-a} - \frac{1}{(2-a)} - \frac{(1-a)^2}{(1-a)} \right\}$$

donde se ha usado (4.17.1). De aquí,

$$d_{IS} = a \ln G_1 + (1-a) \ln G_2 + 1$$

$$d_{IS} = a \ln (2-a) - a \ln 2 + (1-a) \ln (1-a) - (1-a) \ln 2 + 1$$

$$d_{IS} = a \ln (2-a) + (1-a) \ln (1-a) + 1 - \ln 2 \quad (4.18.1)$$

para $a = 1/2$, $d_{IS} = 0.1630$

para $a = 0.745$, $d_{IS} = 0.1276$

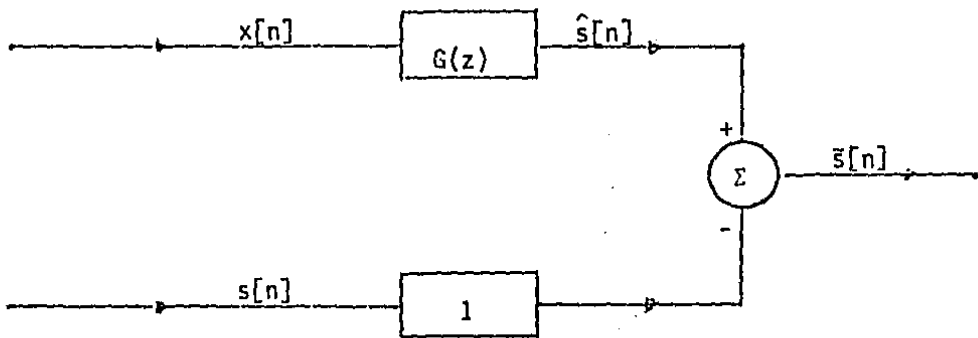
lo que indica un error del 0.277 ó 27.7% al usar mínimo E.C., en relación con el óptimo aquí propuesto.

Usando la ecuación (4.18.1) puede verificarse que se ha encontrado un mínimo.

CAPITULO 5

FILTRADO OPTIMO BAJO EL CRITERIO DE RCEC:

Hasta aquí, se ha tenido como herramienta fundamental en este trabajo la entropía cruzada entre un proceso observado U y un proceso sintético V , $\bar{H}(U, V)$, donde el proceso U está fijo, y se busca la manera de producir un proceso V que minimice la distorsión. Básicamente, se estudió el efecto de sistemas en paralelo pero no conectados entre sí. En este capítulo se seguirá trabajando con procesamiento en paralelo, pero el proceso estocástico a la entrada de los sistemas, en algunos casos (predictores), será común a ambos sistemas, y en todos los casos la respuesta de interés será la diferencia resultante entre las salidas de cada uno de los sistemas, diferencia la cual se denominará el error:



En la figura, obsérvese que uno de los sistemas está fijo y es simplemente un 1. El proceso \tilde{S} es el error. Se llamará a S el proceso original y a X el proceso perturbado; a \hat{S} se le llamará el proceso predicho o estimado. El objetivo de diseñar un filtro (sistema lineal e invariable en el

tiempo) $G(z)$ tal que haga que \hat{S} sea "una buena estimación" de S , bajo un criterio matemático de similitud. Por ejemplo, N. Wiener utilizó el criterio de minimizar la potencia del error. Recuérdese aquí que el filtro de Wiener causal para el caso en que no hay predicción [5, pág. 451] está dado por

$$G_w(z) = \frac{1}{L_x(z)} \left[\frac{f_{sz}(z)}{L_x(1/z)} \right]^+$$

donde $L_x(z)$ es de fase mínima y $f_x(z) = L_x(z)L_x(1/z)$; $[\cdot]^+$ denota la parte del desarrollo en potencias de z con exponente cero ó negativo, y $f_{sx}(z)$ es la transformada z de $R_{sx}[k] = E\{s[n]x[n-k]\}$.

Aquí se propone otro criterio para satisfacer el mismo objetivo. Obsérvese primero que las RCEC $H(S, \hat{S})$ y $H(\hat{S}, S)$ no son funciones de la correlación cruzada entre X y S , pero intuitivamente dos procesos son más parecidos conforme estén más correlacionados entre sí; por esta razón, no se escogen como criterio. Además, $\bar{I}(\hat{S}; S) = \bar{I}(X; S)$ si los procesos son gaussianos y $G(z)$ es estable, o si $G(z)$ es causal y estable, como se vió en la pág.(2.11), de donde el filtrado por un sistema L.I.T. no recupera información sobre el proceso original, sino sólo la dispone en una forma más útil.

Se propone pues como criterio la RCEC $H(\tilde{S}, \sigma^2 W)$, donde W representa a un proceso de ruido blanco gaussiano con potencia promedio unitaria, σ^2 es la potencia del proceso de referencia y funcionará como un parámetro en función del cual quedará definida la solución.

Nótese que $\sigma^2 W$ corresponde en posición al anterior proceso sintético V , pero está fijo, y \tilde{S} corresponde en posición al proceso observado U , pero sobre

él se efectúa la minimización.

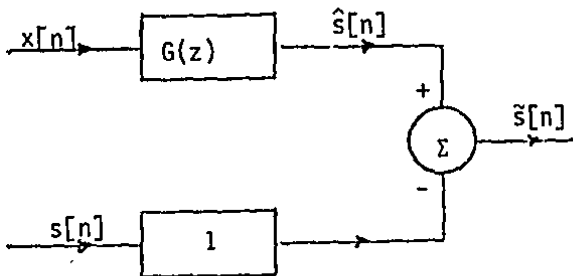
Se procede así en consecuencia con la suposición de diseño de que el proceso de error está disponible para su observación, mientras que $\sigma^2 W$ es sintético en el sentido de no estar relacionado con X ni con S , sino que sólo es una referencia.

Debido a que para una potencia dada σ^2 , el proceso blanco gaussiano tiene la máxima razón de cambio de entropía diferencial, el criterio aquí propuesto es equivalente a buscar el filtro que haga a \tilde{S} lo más parecido a un proceso de máxima entropía; el proceso $\sigma^2 W$ es lo menos predecible, ya que sus muestras son independientes; por tanto, también \tilde{S} deberá hacerse lo menos predecible posible. Dada la optimización del filtro, $S = \hat{S} + \tilde{S} \approx \hat{S} + \sigma^2 W$ lo que indica que se ha dejado fuera de la predicción (estimación) una "componente" de S impredecible (no-estimable). Debido a que la descomposición anterior no es única, los resultados serán desarrollados como función de σ^2 .

Bajo la suposición de gaussianidad, si \tilde{S} es blanco y de potencia σ^2 , \tilde{S} es una componente totalmente impredecible, e intuitivamente resulta que el valor más interesante de σ^2 es aquel mínimo para que ésto sea posible. Es decir, interesa la descomposición de $S = \hat{S} + \tilde{S}$ en una parte predecible y otra impredecible tal que la parte impredecible tenga la potencia mínima. Se demostrará posteriormente cual es esta σ^2 mínima, y que para valores menores \tilde{S} no puede ser totalmente impredecible (blanco), pero si lo menos predecible posible (bajo filtrado L.I.T.).

Densidad de Potencia Espectral del Proceso de Error:

Ya sea que $\tilde{s}[n] = \hat{s}[n] - s[n]$ sea estacionario o asintóticamente estacionario, la densidad de potencia espectral de \tilde{S} , $f_{\tilde{S}}(e^{j\omega})$, en términos del filtro $G(e^{j\omega})$, de las densidades de potencia espectral de los procesos estacionarios X y S , $f_X(e^{j\omega})$ y $f_S(e^{j\omega})$, y de las densidades de potencia cruzadas $f_{XS}(e^{j\omega})$ y $f_{SX}(e^{j\omega})$, puede encontrarse como sigue (se supone que la distribución de potencia espectral $F_{\tilde{S}}(e^{j\omega})$ es absolutamente continua y regular) [5, pág. 282]:



$$f_{\tilde{S}}(e^{j\omega}) = f_{\hat{S}}(e^{j\omega}) - f_{\hat{S}S}(e^{j\omega})$$

$$- f_{S\hat{S}}(e^{j\omega}) + f_S(e^{j\omega})$$

$$f_{\hat{S}\hat{S}}(e^{j\omega}) = G(e^{j\omega}) f_{XS}(e^{j\omega})$$

$$f_{S\hat{S}}(e^{j\omega}) = f_{SS}^*(e^{j\omega})$$

$$f_{\tilde{S}}(e^{j\omega}) = |G(e^{j\omega})|^2 f_X(e^{j\omega})$$

Con lo cual

$$f_{\tilde{S}}(e^{j\omega}) = G(e^{j\omega})G^*(e^{j\omega})f_X(e^{j\omega}) - G(e^{j\omega})f_{XS}(e^{j\omega}) - G^*(e^{j\omega})f_{SX}(e^{j\omega}) + f_S(e^{j\omega}) \quad (5.4.1)$$

ó bien

$$f_{\tilde{S}}(e^{j\omega}) = \left| G(e^{j\omega})L_X(e^{j\omega}) - \frac{f_{SX}(e^{j\omega})}{L_X^*(e^{j\omega})} \right|^2 - \frac{|f_{SX}(e^{j\omega})|^2}{f_X(e^{j\omega})} + f_S(e^{j\omega}) \quad (5.4.2)$$

Una cota superior para $\bar{H}(\tilde{S}, \sigma^2 W)$ de validez "general":

De la definición de la RCEC:

$$\bar{H}(\tilde{S}, \sigma^2 W) = -\bar{H}_{\tilde{S}} + \lambda n \sqrt{2\pi\sigma^2} + \frac{1}{2\sigma^2} \left(\lim_{N \rightarrow \infty} E\{\tilde{s}^2[N]\} \right)$$

Aplicando la propiedad 2 de la RCIM (considerando que los procesos son gaussianos) o la propiedad 3 de la RCIM (si el filtro $G(z)$ es causal y los procesos X y S se aplican a partir de $n=0$):

$$\bar{I}(\hat{S}; S) = \bar{I}(X; S)$$

Además, por la ecuación (2.13.1):

$$\bar{H}_{\tilde{S}} = \bar{I}(\hat{S}; \tilde{S}) + \bar{H}_S - \bar{I}(\hat{S}; S) = \bar{I}(\hat{S}; \tilde{S}) + \bar{H}_S - \bar{I}(X; S)$$

Por lo que

$$\bar{H}(\tilde{S}, \sigma^2 W) = -\bar{H}_S + \bar{I}(X; S) + \lambda n \sqrt{2\pi\sigma^2} + \frac{1}{2\sigma^2} \left(\lim_{N \rightarrow \infty} E\{s^2[N]\} \right) - \bar{I}(\hat{S}; \tilde{S}) \quad (5.5.1)$$

donde los tres primeros términos son independientes del filtro, y en consecuencia, para minimizar la RCEC se debe minimizar por medio del filtro la cantidad

$$\frac{1}{2\sigma^2} \int_{-\pi}^{\pi} f_{\tilde{S}}(e^{j\omega}) \frac{d\omega}{2\pi} - \bar{I}(\hat{S}; \tilde{S})$$

o aplicando nuevamente la propiedad dos o la tres de la RCIM (págs.2.11 y 2.12):

$$\frac{1}{2\sigma^2} \int_{-\pi}^{\pi} f_{\tilde{S}}(e^{j\omega}) \frac{d\omega}{2\pi} = \bar{I}(X; \tilde{S}).$$

Ya que $\bar{I}(X; \tilde{S}) = \bar{I}(\hat{S}; \tilde{S}) \geq 0$, de (5.5.1) se obtiene:

$$\bar{H}(\tilde{S}, \sigma^2 W) \leq -\bar{H}_{\tilde{S}} + \bar{I}(X; S) + \ln \sqrt{2\pi\sigma^2} + \frac{1}{2\sigma^2} \int_{-\pi}^{\pi} f_{\tilde{S}}(e^{j\omega}) \frac{d\omega}{2\pi} \quad (5.6.1)$$

De (5.6.1) se concluye que la cota superior obtenida se minimiza con el filtro que minimiza el error cuadrático medio.

Se verá en la siguiente sección que en el caso de predictores con procesos gaussianos, $\bar{I}(X; \tilde{S})$ queda indeterminada. Por tanto, la cota obtenida es válida para

- a) procesos gaussianos aplicados desde $n_0 = -\infty$;
 $G(z)$ estable; filtrado
- b) procesos aplicados desde $n_0 = 0$; $G(z)$ causal y estable;
 para predicción con tiempo inicial $n_0 = 0$, la cota 5.6.1 no será necesaria, y no es válida si además los procesos son gaussianos.

UNA EXPRESION GENERAL PARA LA RCEC CUANDO LOS PROCESOS SON GAUSSIANOS Y

$n_0 = 0$. CONDICIONES INICIALES NULAS

Se sabe que $H(\tilde{S}, \sigma^2 W) = \frac{1}{2} d_{IS}(f_{\tilde{S}}(e^{j\omega}), \sigma^2)$ cuando los procesos X y S son gaussianos y se aplican desde $n_0 = -\infty$, porque \tilde{S} a partir de cualquier

tiempo finito es estacionario. Sin embargo, la expresión anterior no se cumple si los procesos se aplican desde $n_0 = 0$, ya que entonces \tilde{S} es tan sólo asintóticamente estacionario. A continuación se deriva una expresión que es válida para procesos X y S gaussianos estacionarios que se apliquen desde $n_0 = 0$.

Sean los vectores aleatorios $\underline{X}_n^T = (x_0, x_1, \dots, x_{n-1})$ y $\underline{\tilde{S}}_n^T = (\tilde{s}_0, \tilde{s}_1, \dots, \tilde{s}_{n-1})$. Sea $C_T = E \{ \text{col}(\underline{X}_n, \underline{\tilde{S}}_n) \text{col}^T(\underline{X}_n, \underline{\tilde{S}}_n) \}$ y $\Delta_T = \det(C_T)$; sea $C_X = E\{\underline{X}_n \underline{X}_n^T\}$ y $\Delta_X = \det(C_X)$, sea $C_S = E\{\underline{\tilde{S}}_n \underline{\tilde{S}}_n^T\}$ y $\Delta_S = \det(C_S)$. Entonces bajo la suposición de gaussianidad:

$$I(x_0, x_1, \dots, x_{n-1}; (x_0, x_1, \dots, x_{n-1})) = \frac{1}{2} \ln \left(\frac{\Delta_X \Delta_S}{\Delta_T} \right) \quad (5.7.1)$$

Demostración: Apéndice D.1

Cuando los procesos X y S se aplican desde $n_0 = 0$, se satisface la siguiente ecuación:

$$\underline{\tilde{S}}_n = G_n \underline{X}_n - \underline{S}_n$$

donde $\underline{S}_n^T = (s_0, s_1, \dots, s_{n-1})$ y $G_n = \begin{bmatrix} g_0 & & & 0 \\ g_1 & g_0 & & \\ \vdots & & & \\ g_{n-1} & \dots & \dots & g_0 \end{bmatrix}$

con la secuencia $\{g_k\}_{k=0}^{\infty}$ igual a la respuesta al pulso unitario del filtro causal $G(z)$.

Sea $C_{xs} = E\{\underline{x}_n \underline{s}_n^T\}$, $C_{sx} = E\{\underline{s}_n \underline{x}_n^T\}$, y $C_s = E\{\underline{s}_n \underline{s}_n^T\}$. Entonces Δ_T es independiente del filtro y

$$\Delta_T = \det \begin{bmatrix} C_x & C_{xs} \\ C_{sx} & C_s \end{bmatrix} \quad (5.8.1)$$

Demostración: Apéndice D.2

De (5.7.1),

$$I(\underline{x}_n; \tilde{\underline{s}}_n) = -\frac{1}{2} \ln \Delta_s + \frac{1}{2} \ln \left(\frac{\Delta_x \Delta_s}{\Delta_T} \right) + \frac{1}{2} \ln \Delta_{\tilde{s}} \quad , \quad \text{con } \Delta_s = \det(C_s);$$

de (5.8.1),

$\Delta_T = \det E\{ \text{col}(\underline{x}_n, \underline{s}_n) \text{col}^T(\underline{x}_n, \underline{s}_n) \}$, por lo que, aplicando nuevamente (5.7.1)

$$I(\underline{x}_n; \tilde{\underline{s}}_n) = -\frac{1}{2} \ln \Delta_s + I(\underline{x}_n, \underline{s}_n) + \frac{1}{2} \ln \Delta_{\tilde{s}} ;$$

dividiendo entre N , tomando el límite cuando N tiende a infinito y aplicando el teorema de Szego:

$$\bar{I}(X; \tilde{S}) = -\frac{1}{2} \int_{-\pi}^{\pi} \ln f_s(e^{j\omega}) \frac{d\omega}{2\pi} + \bar{I}(X; S) + \lim_{N \rightarrow \infty} \left(\frac{1}{2N} \ln \Delta_{\tilde{s}} \right) \quad (5.8.2)$$

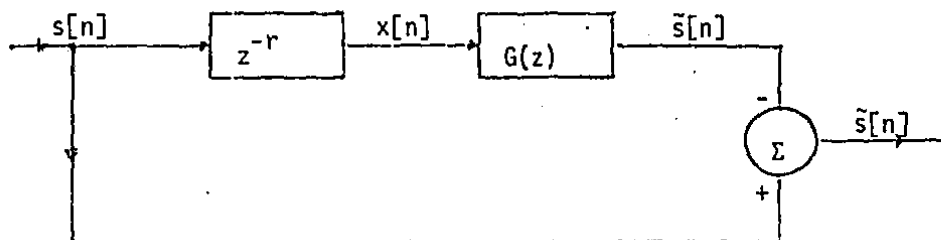
donde el último término no puede simplificarse igual que el primero por ser \tilde{S} tan sólo asintóticamente estacionario.

Usando (5.8.2) en (5.5.1) (ya que $\bar{I}(\tilde{S}; \tilde{S}) = \bar{I}(X; \tilde{S})$):

$$\bar{H}(\tilde{S}, \sigma^2 W) = -\frac{1}{2} \ln(2\pi e) + \ln \sqrt{\frac{2\pi\sigma^2}{2\pi\sigma^2}} + \frac{1}{2\sigma^2} \int_{-\pi}^{\pi} f_{\tilde{S}}(e^{j\omega}) \frac{d\omega}{2\pi} - \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \ln \Delta_{\tilde{S}}$$

$$\bar{H}(\tilde{S}, \sigma^2 W) = \frac{1}{2} \left[\int_{-\pi}^{\pi} \frac{f_{\tilde{S}}(e^{j\omega})}{\sigma^2} \frac{d\omega}{2\pi} - \left(\lim_{N \rightarrow \infty} \frac{\ln \Delta_{\tilde{S}} / \sigma^2}{N} \right) - 1 \right] \quad (5.9.1)$$

Obsérvese que (5.9.1) se reduciría a $1/2 \, d_{IS}(f_{\tilde{S}}(e^{j\omega}), \sigma^2)$ si el proceso \tilde{S} fuera estacionario. Esto se demuestra en el apéndice D.3.



si el atrazo z^{-r} se incorpora al filtro $G(z)$, $C_{SX} = C_S = C_X$, y la matriz C_T es singular.

RELACION CON LA PLANITUD ESPECTRAL :

La medida de planitud espectral ("spectral flatness measure") está definida por [24]:

$$\gamma_X^2 \triangleq \frac{\exp\left\{ \int_{-\pi}^{\pi} \ln f_X(e^{j\omega}) \frac{d\omega}{2\pi} \right\}}{\frac{1}{2\pi} \int_{-\pi}^{\pi} f_X(e^{j\omega}) d\omega}$$

Por otro lado, para procesos gaussianos estacionarios ($n_0 = -\infty$):

$$\bar{H}(\bar{S}, \sigma^2 W) = \frac{1}{2} \int_{-\pi}^{\pi} \left\{ \frac{f_{\bar{S}}(e^{j\omega})}{\sigma^2} - \ln \frac{f_{\bar{S}}(e^{j\omega})}{\sigma^2} - 1 \right\} \frac{d\omega}{2\pi} ;$$

sea la potencia del proceso de referencia

$$\sigma^2 = \int_{-\pi}^{\pi} f_{\bar{S}}(e^{j\omega}) \frac{d\omega}{2\pi} ,$$

entonces

$$\bar{H}(\bar{S}, \sigma^2 W) = \frac{1}{2} \left[\ln \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f_{\bar{S}}(e^{j\omega}) d\omega \right) - \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln f_{\bar{S}}(e^{j\omega}) d\omega \right]$$

$$\bar{H}(\bar{S}, \sigma^2 W) = \frac{1}{2} \ln \frac{1}{\gamma_{\bar{S}}^2}$$

Recuérdese que a) γ_X^2 es la razón de la media geométrica a la media aritmética

b) $0 \leq \gamma_X^2 \leq 1$

c) γ_X^{-2} es una medida de la predictibilidad de la forma de onda. Si $\gamma_X^{-2} = \infty$ el proceso no es regular.

FILTRO NO-CAUSAL

Considerando un filtro $G(z)$ no necesariamente causal, lineal, estable e inva-

riante en el tiempo, y que los procesos X y S son gaussianos estacionarios aplicados desde $n_0 = -\infty$, de tal forma que el proceso de error \tilde{S} a partir de cualquier tiempo finito es también estacionario, se tiene (Apéndice D.3-B):

$$\bar{H}(\tilde{S}, \sigma^2_W) = \frac{1}{2} d_{IS}(f_{\tilde{S}}(e^{j\omega}), \sigma^2)$$

donde $f_{\tilde{S}}(e^{j\omega})$ está dado por (5.4.1) ó (5.4.2). Escribiendo $G(e^{j\omega}) = A(\omega) + jB(\omega)$, la nulidad de la primera variación de d_{IS} con respecto a $G(e^{j\omega})$, se reduce a las condiciones (Apéndice E-1)

$$\frac{\partial d_{IS}}{\partial A} = 0, \quad \frac{\partial d_{IS}}{\partial B} = 0,$$

necesarias para un extremo local. La solución de estas condiciones es

$$G(e^{j\omega}) = \frac{f_{SX}(e^{j\omega})}{f_X(e^{j\omega})} \triangleq G_W^{NC}(e^{j\omega}),$$

la cual es el filtro de Wiener no-causal. Sin embargo, esta solución es un "punto silla" para el criterio d_{IS} , es decir, la planitud de la densidad espectral del error puede aumentarse o disminuirse en torno a esta solución.

De hecho, dada una densidad espectral cualquiera $|K(e^{j\omega})|^2$ que satisfaga la condición $|K(e^{j\omega})|^2 \geq f_{\tilde{S}_W}^{NC}(e^{j\omega}) \forall \omega$ (donde $f_{\tilde{S}_W}^{NC}(e^{j\omega})$ representa la densidad espectral del error cuando se usa el filtro de Wiener no-causal, esto es:

$$f_{\tilde{S}_W}^{NC}(e^{j\omega}) = - \frac{|f_{SX}(e^{j\omega})|^2}{f_X(e^{j\omega})} + f_S(e^{j\omega}),$$

puede encontrarse un filtro $G_W^{NC}(e^{j\omega}) + \Delta(e^{j\omega})$ tal que el espectro del error evaluado en este filtro sea

$$f_{\bar{s}_\Delta}(e^{j\omega}) = |K(e^{j\omega})|^2.$$

$\Delta(e^{j\omega})$ puede encontrarse de (5.4.2):

$$|L_x(e^{j\omega}) \Delta(e^{j\omega})|^2 = |K(e^{j\omega})|^2 - f_{\bar{s}_W}^{NC}(e^{j\omega}).$$

Obsérvese que el filtro de Wiener no-causal no solamente minimiza la potencia del error, sino que también, para cada valor de ω minimiza la densidad espectral del error (ya que hace nulo el único término de (5.4.2) que depende del filtro). Se concluye que el problema de minimizar la d_{IS} es uno sujeto a la restricción $f_{\bar{s}}(e^{j\omega}) \geq f_{\bar{s}_W}^{NC}(e^{j\omega})$.

La solución es función de σ^2 y se da a continuación para cada uno de tres rangos de este parámetro:

a) $\sigma^2 \geq \max_{\omega} \{f_{\bar{s}_W}^{NC}(e^{j\omega})\} :$

Obsérvese aquí que si se escoge $|K(e^{j\omega})|^2 = \sigma^2$, siempre se cumple la restricción y existe $\Delta(e^{j\omega})$ tal que $f_{\bar{s}}(e^{j\omega}) = |K(e^{j\omega})|^2 = \sigma^2 \geq f_{\bar{s}_W}^{NC}(e^{j\omega})$. Entonces $d_{IS}(|K(e^{j\omega})|^2, \sigma^2) = d_{IS}(\sigma^2, \sigma^2) = 0$ y evidentemente se tiene el mínimo deseado.

b) $\min_{\omega} f_{\bar{s}_W}^{NC}(e^{j\omega}) < \sigma^2 < \max_{\omega} f_{\bar{s}_W}^{NC}(e^{j\omega})$

Sean los conjuntos $A = \{ \omega : f_{\bar{s}_W}^{NC}(e^{j\omega}) \geq \sigma^2 \}$

$$B = \{ \omega: f_{\tilde{s}_w}^{NC}(e^{j\omega}) < \sigma^2 \}$$

En el conjunto B, si $f_{\tilde{s}}(e^{j\omega}) = \sigma^2$, se satisface la restricción y además

$$\int_B \left\{ \frac{f_{\tilde{s}}(e^{j\omega})}{\sigma^2} - \ln \frac{f_{\tilde{s}}(e^{j\omega})}{\sigma^2} - 1 \right\} \frac{d\omega}{2\pi} = 0, \quad (5.13.1)$$

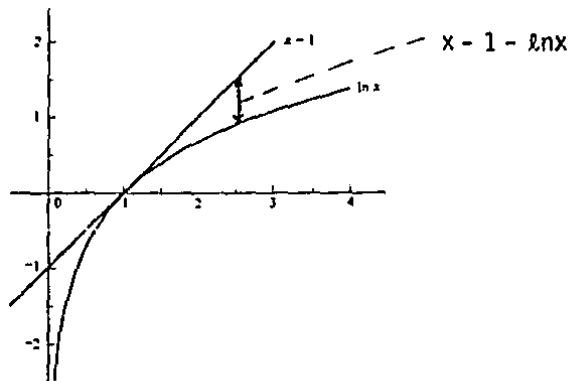
por lo que este término es mínimo. Por otro lado, en el conjunto A:

$$\int_A \left\{ \frac{f_{\tilde{s}}(e^{j\omega})}{\sigma^2} - \ln \frac{f_{\tilde{s}}(e^{j\omega})}{\sigma^2} - 1 \right\} \frac{d\omega}{2\pi} \geq \int_A \left\{ \frac{f_{\tilde{s}_w}^{NC}(e^{j\omega})}{\sigma^2} - \ln \frac{f_{\tilde{s}_w}^{NC}(e^{j\omega})}{\sigma^2} - 1 \right\} \frac{d\omega}{2\pi} \quad (5.13.2)$$

ya que

$$x - \ln x - 1 \geq y - \ln y - 1 \geq 0$$

si $x \geq y \geq 1$. (véase la figura)



Debido a la restricción, la igualdad ocurre en (5.13.2) si y sólo si $f_{\tilde{s}}(e^{j\omega}) = f_{\tilde{s}_w}^{NC}(e^{j\omega})$ para toda $\omega \in A$, y por tanto, el miembro derecho de (5.13.2) es el mínimo del miembro izquierdo.

$$\text{Sea } |K(\omega)|^2 = \begin{cases} \sigma^2 & \omega \in B \\ f_{\tilde{s}_w}^{NC}(e^{j\omega}) & \omega \in A \end{cases},$$

entonces, existe $\Delta(e^{j\omega})$ tal que

$$f_{\tilde{s}_\Delta}(e^{j\omega}) = |K(e^{j\omega})|^2 \quad \text{y además}$$

$$d_{IS}(f_{\tilde{s}_\Delta}(e^{j\omega}), \sigma^2) = \int_A \left\{ \frac{f_{\tilde{s}_\Delta}(e^{j\omega})}{\sigma^2} - \ln \left(\frac{f_{\tilde{s}_\Delta}(e^{j\omega})}{\sigma^2} \right) - 1 \right\} \frac{d\omega}{2\pi}$$

$$+ \int_B \left\{ \frac{f_{\tilde{s}_\Delta}(e^{j\omega})}{\sigma^2} - \ln \left(\frac{f_{\tilde{s}_\Delta}(e^{j\omega})}{\sigma^2} \right) - 1 \right\} \frac{d\omega}{2\pi}$$

es mínima sujeta a la restricción, ya que ambos términos son positivos y mínimos.

$$c) \min_{\omega} f_{\tilde{s}_w}^{NC}(e^{j\omega}) \geq \sigma^2 :$$

En este caso, el conjunto A del inciso b) ocupa todo el intervalo de $-\pi$ a π , y la distorsión mínima, sujeta a la restricción $f_{\tilde{s}}(e^{j\omega}) \geq$

$\geq f_{\tilde{s}_w}^{NC}(e^{j\omega})$, ocurre de acuerdo a (5.13.2) precisamente cuando

$f_{\tilde{s}_\Delta}(e^{j\omega}) = f_{\tilde{s}_w}^{NC}(e^{j\omega})$ ésto es, cuando $\Delta(e^{j\omega}) = 0$ y el filtro es el filtro de Wiener.

FILTRO CAUSAL, $n_0 = -\infty$. CONDICIONES NECESARIAS PARA EL OPTIMO :

Considerando un filtro $G(z)$ causal, lineal, estable e invariante en el tiempo, y que los procesos X y S son gaussianos, estacionarios aplicados desde

$n_0 = -\infty$, de tal forma que el proceso de error \tilde{S} a partir de cualquier tiempo finito es estacionario, se tiene (Apéndice D.3-B)

$$\bar{H}(\tilde{S}, \sigma^2 W) = \frac{1}{2} d_{IS}(f_{\tilde{S}}(e^{j\omega}), \sigma^2)$$

donde $f_{\tilde{S}}(e^{j\omega})$ está dado por (5.4.1) o (5.4.2). Se introduce ahí la restricción de que el filtro sea causal escribiéndolo en la forma $G(z) = \sum_{k=0}^{\infty} g_k z^{-k}$. Las condiciones necesarias para un extremo local son

$$\frac{\partial}{\partial g_k} d_{IS}(f_{\tilde{S}}(e^{j\omega}), \sigma^2) = 0, \quad k=0, 1, 2, \dots \quad (5.15.1)$$

de donde se obtiene :

$$\int_{-\pi}^{\pi} \operatorname{Re}\{[G(e^{j\omega}) f_x(e^{j\omega}) - f_{sx}(e^{j\omega})] e^{j\omega k}\} \left[1 - \frac{\sigma^2}{f_{\tilde{S}}(e^{j\omega})}\right] \frac{d\omega}{2\pi} = 0 \quad k=0, 1, 2, \dots$$

Definiendo $G(e^{j\omega}) \triangleq A(\omega) + jB(\omega)$, $f_{sx}(e^{j\omega}) = R(\omega) + jI(\omega)$ y sustituyendo en la última ecuación, se sigue que

$$\int_{-\pi}^{\pi} (A(\omega) f_x(e^{j\omega}) - R(\omega)) \left[1 - \frac{\sigma^2}{f_{\tilde{S}}(e^{j\omega})}\right] \cos(\omega k) \frac{d\omega}{2\pi} = \int_{-\pi}^{\pi} \left\{ B(\omega) f_x(e^{j\omega}) - I(\omega) \right\} \left\{ 1 - \frac{\sigma^2}{f_{\tilde{S}}(e^{j\omega})} \right\} \sin(\omega k) \frac{d\omega}{2\pi}, \quad k=0, 1, 2, \dots \quad (5.15.2)$$

Ambos miembros de (5.15.2) serán iguales para $k=1, 2, 3, \dots$, si y sólo si la transformada de Fourier inversa de la función $F(e^{j\omega})$, (Apéndice C.1),

$$F(e^{j\omega}) = [(A(\omega) f_x(e^{j\omega}) - R(\omega)) - j(B(\omega) f_x(e^{j\omega}) - I(\omega))] \left[1 - \frac{\sigma^2}{f_{\tilde{S}}(e^{j\omega})}\right]$$

es una función causal. Equivalentemente

$$F(1/z) = [G(z) f_x(z) - f_{sx}(z)] \left[1 - \frac{\sigma^2}{f_s(z)} \right] \quad (5.16.1)$$

debe ser una función anticausal. De la ecuación (5.16.1) pueden deducirse va rias cosas:

a) Si $\sigma^2 \rightarrow 0$, en el límite, la solución es el filtro de Wiener causal, ésto es:

$$F(1/z) = G(z)f_x(z) - f_{sx}(z) \text{ debe ser anticausal}$$

$$\Rightarrow \frac{F(1/z)}{L_x(1/z)} = G(z)L_x(z) - \frac{f_{sx}(z)}{L_x(1/z)} \text{ debe ser anticausal}$$

$$\Rightarrow G(z) = \frac{1}{L_x(z)} \left[\frac{f_{sx}(z)}{L_x(1/z)} \right]^+ \triangleq G_w^c(e^{j\omega}) \quad (5.16.2)$$

por tanto (5.16.2) es solución de (5.15.2) para $k=1, 2, 3, \dots$

Para $k=0$, se verifica que (5.16.2) también es solución: ya que

$$\int_{-\pi}^{\pi} \{B(\omega) f_x(e^{j\omega}) - I(\omega)\} \frac{d\omega}{2\pi} = 0, \text{ para } k=0, (5.15.2) \text{ es lo mismo que}$$

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \{G(e^{j\omega})f_x(e^{j\omega}) - f_{sx}(e^{j\omega})\} d\omega = 0$$

donde

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} L_x(e^{-j\omega}) \left[\frac{f_{sx}(e^{j\omega})}{L_x(e^{-j\omega})} \right]^- d\omega = 0$$

se verifica ya que el integrando contiene sólo potencias positivas de $e^{j\omega}$.

- b) Si se permite que σ^2 no sea una constante, sino una función de $e^{j\omega}$, que represente la densidad espectral del proceso de comparación, (5.16.1) toma la forma

$$F(1/z) = [G(z)f_x(z) - f_{sx}(z)] \left[\frac{1}{\sigma^2(z)} - \frac{1}{f_{\xi}(z)} \right]$$

de donde, si $\sigma^2(z) = \frac{f_{\xi_w}^C(z)}{1 + af_{\xi_w}^C(z)}$, la función

$F(1/z) = [G(z)f_x(z) - f_{sx}(z)] a$ es anticausal cuando $G(z)$ es el filtro de Wiener. Nótese que si $a \rightarrow \infty$, $\sigma^2(e^{j\omega}) \approx 1/a$, lo cual es blanco y de una potencia que tiende a cero. ($f_{\xi_w}^C(e^{j\omega})$ se obtiene con (5.16.2) en (5.4.2))

- c) Recuérdese que el filtro de Wiener no-causal hace cero la primera variación de la d_{IS} con respecto al filtro $G(z)$ para todo valor de σ^2 . Si se pide que $F(1/z)$ en (5.16.1) sea anticausal para todo valor de σ^2 , escribiendo

$$F(1/z) = [G(z)f_x(z) - f_{sx}(z)] - \sigma^2 \left\{ \frac{G(z)f_x(z) - f_{sx}(z)}{f_{\xi}(z)} \right\}$$

o bien

$$F(1/z) = F_1(z) - \sigma^2 F_2(z)$$

y si pueden existir $F_1(z)$ y $F_2(z)$ tales que, para algún valor particular σ^2 , $F(1/z)$ sea anticausal, la cancelación de polos y ceros nece

saría para ésto no ocurrirá para otros valores de σ^2 . Por tanto, se requiere que tanto $F_1(z)$ como $F_2(z)$ sean anticausales. En el inciso a) se vió que $F_1(z)$ es anticausal si el filtro es el filtro de Wiener causal. De la condición

$$\frac{F_2(z)L_{\tilde{S}}(1/z)}{L_X(1/z)} = \frac{G(z)L_X(z)}{L_{\tilde{S}}(z)} - \frac{f_{sX}(z)}{L_{\tilde{S}}(z)L_X(1/z)},$$

donde el miembro izquierdo es anticausal y el primer término del miembro derecho es causal, se llega a

$$G(z) = \frac{L_{\tilde{S}}(z)}{L_X(z)} \left\{ \left[\frac{f_{sX}(z)}{L_X(1/z)L_{\tilde{S}}(z)} \right]^+ \right\};$$

entonces la condición para que $F(1/z)$ sea anticausal para todo valor de σ^2 es

$$G(z) = \frac{1}{L_X(z)} \left[\frac{f_{sX}(z)}{L_X(1/z)} \right]^+ = \frac{L_{\tilde{S}}(z)}{L_X(z)} \left[\frac{f_{sX}(z)}{L_{\tilde{S}}(z)L_X(1/z)} \right]^+ \quad (5.18.1)$$

de tal forma que si (5.18.1) se cumple, (5.15.2) se cumple para todo σ^2 y para $k=1, 2, 3, \dots$; (5.15.2) también se cumple para $k=0$, ya que en (5.18.1) se está restringiendo el filtro de Wiener a cierto tipo de datos, y como se vió en el inciso a), el filtro de Wiener satisface (5.15.2) para $k=0$.

La condición (5.18.1) es pues necesaria y suficiente para que (5.15.1) se cumpla para toda σ^2 , estabilizando la d_{IS} .

Hay dos formas al menos en que (5.18.1) puede cumplirse:

i) $L_S(z) = \text{constante}$. Esta es la situación en que el filtro de Wiener produce un proceso de error de densidad espectral constante, lo cual ciertamente es un caso redundante.

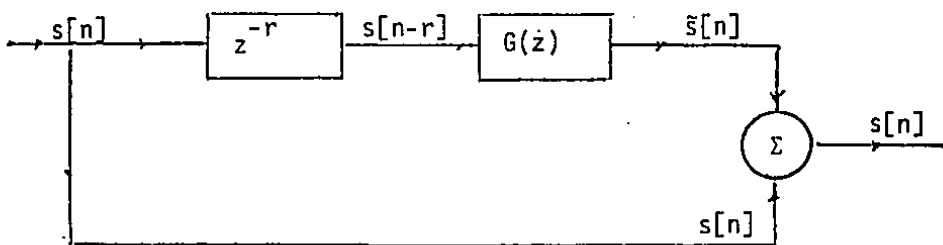
$$\text{ii) } \frac{f_{s_x}(z)}{L_x(z)} = f_{se_x}(z) = \sum_{k=0}^{\infty} R_{se_x}[k]z^{-k}, \quad |z| < 1, \quad (5.19.1)$$

donde e_x representa al proceso de innovaciones de x . Cuando (5.19.1) se cumple, el filtro de Wiener no-causal, resulta causal:

$$\begin{aligned} G_{W \text{ no-causal}}(z) &= \frac{f_{s_x}(z)}{f_x(z)} = \frac{1}{L_x(z)} \left[\frac{f_{s_x}(z)}{L_x(1/z)} \right] = \frac{1}{L_x(z)} \left[\frac{f_{s_x}(z)}{L_x(1/z)} \right]^* \\ &= G_W^C = G_W^{NC}(z) \end{aligned}$$

En el apéndice F se dan unos modelos de canal y ruido para los causales (5.19.1) se cumple.

d) Aplicación al predictor:



Aquí, $x[n] = s[n-r]$, donde r es el número de pasos del predictor. Se tiene que $f_x(z) = f_s(z) = L_s(z)L_s(1/z)$, donde $L_s(z)$ es de fase mínima. Sea $L_s(z) = \sum_{k=0}^{\infty} l_k z^{-k}$. De la figura, con $G(z)$ igual al predictor de

Wiener, ésto es,

$$G(z) = G_W^c(z) = \frac{z^r}{L_S(z)} \left\{ \sum_{k=r}^{\infty} \ell_k z^{-k} \right\},$$

se sigue

$$f_{\tilde{S}_W}(z) = \sum_{k=0}^{r-1} (\ell_k z^{-k}) \left(\sum_{k=0}^{r-1} \ell_k z^k \right) = L_{\tilde{S}_W}(z) L_{\tilde{S}_W}(1/z),$$

donde $L_{\tilde{S}_W}(z) = \sum_{k=0}^{r-1} \ell_k z^{-k}$ si y sólo si el miembro derecho es de fase mínima. La ecuación (5.16.1) toma ahora la forma

$$\begin{aligned} F(1/z) &= L_z(1/z) \left[-z^r \sum_{k=0}^{r-1} \ell_k z^{-k} \right] \left[1 - \frac{\sigma^2}{\left(\sum_{k=0}^{r-1} \ell_k z^{-k} \right) \left(\sum_{k=0}^{r-1} \ell_k z^k \right)} \right] \\ &= L_S(1/z) \left[-z^r \sum_{k=0}^{r-1} \ell_k z^{-k} + \frac{\sigma^2 z^r}{\sum_{k=0}^{r-1} \ell_k z^k} \right] \end{aligned}$$

donde el primer término es obviamente anticausal, y la región de convergencia del segundo término es la región de convergencia del factor (apéndice A.3):

$$\frac{1}{\sum_{k=0}^{r-1} \ell_k z^k},$$

la cual es el interior de un círculo. Si el polinomio $\sum_{k=0}^{r-1} \ell_k z^k$ tiene ceros dentro del círculo unitario, dicho círculo de convergencia es interior al círculo unitario, y si bien el término sigue siendo anticausal, la transformada de Fourier no converge. Por tanto, para que el predictor

de Wiener sea solución, es necesario que el filtro de error, dado por

$$\frac{\sum_{k=0}^{r-1} \ell_k z^{-k}}{L_S(z)}$$

sea de fase mínima. Posteriormente se mostrará que ésto es una condición suficiente para la minimización de la d_{IS} .

- e) A partir de (5.16.1) se puede obtener una solución general, escribiendo $G_{\Delta}(z) = G_W^C(z) + \frac{\Delta(z)}{L_X(z)}$, donde $G_W^C(z)$ es el filtro de Wiener causal y $\Delta(z)$ es causal.

Sea $f(z) \triangleq \left[\frac{f_{SX}(z)}{L_X(1/z)} \right]^-$, donde $[]^-$ indica la parte del desarrollo en potencias de z con exponente positivo. Entonces, el espectro del error usando el filtro de Wiener causal es

$$f_{\tilde{S}_W}^C(z) = f(z)f(1/z) + f_{\tilde{S}_W}^{NC}(z) \quad (5.21.1)$$

donde el espectro del error usando el filtro de Wiener no-causal es

$$f_{\tilde{S}_W}^{NC}(z) = - \frac{f_{SX}(z)f_{SX}(1/z)}{f_X(z)} + f_S(z) \quad (5.21.2)$$

Además, el espectro del error utilizando el filtro $G_{\Delta}(z)$ es

$$f_{\tilde{S}_{\Delta}}(z) = [\Delta(z) - f(z)][\Delta(1/z) - f(1/z)] + f_{\tilde{S}_W}^{NC}(z) \quad (5.21.3)$$

Ahora (5.16.1), usando $G(z) = G_{\Delta}(z)$, toma la forma:

$$F(1/z) = L_x(1/z) [\Delta(z) - f(z)] \left[1 - \frac{\sigma^2}{f_{S_\Delta}(z)} \right]$$

o bien ,

$$f(z) + \frac{F(1/z)}{L_x(1/z)} = \Delta(z) - \frac{\sigma^2(\Delta(z) - f(z))}{f_{S_\Delta}(z)}$$

y como el miembro izquierdo es anticausal, las partes causales del miembro derecho deben cancelarse:

$$\Delta(z) = \left[\frac{\sigma^2 (\Delta(z) - f(z))}{(\Delta(z) - f(z))(\Delta(1/z) - f(1/z)) + f_{S_W}^{NC}(z)} \right]^+ \quad (5.22.1)$$

Nótese que

- i) en el caso de predicción, $f_{S_W}^{NC}(z) = 0$ y $f(z) = z^r \sum_{k=0}^{r-1} \ell_k z^{-k}$,
y (5.22.1) toma la forma

$$\Delta(z) = \left[\frac{\sigma^2}{\Delta(1/z) - z^{-r} \left(\sum_{k=0}^{r-1} \ell_k z^{+k} \right)} \right]^+ ; \text{ haciendo}$$

$$\Delta(z) = 0,$$

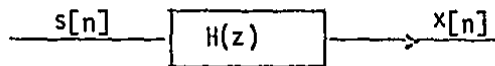
$$0 = \left[\frac{\sigma^2 z^r}{\sum_{k=0}^{r-1} \ell_k z^k} \right]^+, \text{ lo cual se verifica.}$$

- ii) Si existe $\Delta(z)$ tal que $f_{S_\Delta}(z) = \sigma^2$, (5.22.1) se reduce a

$$\Delta(z) = \Delta(z) .$$

y por tanto, tal $\Delta(z)$ es solución, como era de esperarse.

- iii) Sea $x[n]$ obtenido a partir de $s[n]$ por medio de un filtro de fase mínima $H(z)$:



El filtro de Wiener causal es entonces $G_w^C(z) = \frac{1}{H(z)}$; Además,

$$f_{s_x}(z) = H(1/z)f_s(z) ; f_x(z) = H(z)H(1/z)f_s(z)$$

$$f_{S_w}^{NC}(z) = - \frac{H(z)f_s(z)f_s(z)H(1/z)}{H(z)H(1/z)f_s(z)} + f_s(z) = 0$$

$$f(z) = \left[\frac{f_{s_x}(z)}{L_x(1/z)} \right]^- = \left[\frac{H(1/z)f_s(z)}{H(1/z)L_s(1/z)} \right]^- = [L_s(z)]^- = 0$$

La ec. (5.22.1) es entonces

$$\Delta(z) = \left[\frac{\sigma^2 \Delta(z)}{\Delta(z) \Delta(1/z)} \right]^+ = \left[\frac{\sigma^2}{\Delta(1/z)} \right]^+ ;$$

la solución es $\Delta(z) = \pm \sigma$.

Se observa que

$$G_{\Delta}(z) = \frac{1}{H(z)} + \frac{\sigma}{L_X(z)} \quad \text{y} \quad G_W^C(z) = \frac{1}{H(z)}; \text{ usando (5.4.1):}$$

$$\begin{aligned} f_{S_W}^C &= f_X(z) \frac{1}{H(z)H(1/z)} - f_{S_X}(z) \frac{1}{H(1/z)} - f_{S_X}(1/z) \frac{1}{H(z)} + f_S(z) = \\ &= \frac{H(z)H(1/z)f_S(z)}{H(z)H(1/z)} - H(1/z) \frac{f_S(z)}{H(1/z)} - H(z)f_S(z)/H(z) + f(z) = 0 \end{aligned}$$

$$\text{y} \quad d_{IS}(f_{S_W}^C, \sigma^2) = - \int_{-\pi}^{\pi} \ln \frac{f_{S_W}^C(e^{j\omega})}{\sigma^2} \frac{d\omega}{2\pi} - 1 = \infty. \quad (\sigma^2 \neq 0)$$

sin embargo

$$f_{S_{\Delta}}(e^{j\omega}) = \sigma^2$$

y

$$d_{IS}(f_{S_{\Delta}}, \sigma^2) = 0$$

Se concluye:

$$G_W^C(z) = 1/H(z); \quad f_{S_W}^C(z) = 0; \quad d_{IS}(f_{S_W}^C, \sigma^2) = \infty$$

$$G_{\Delta}(z) = \frac{1}{H(z)} \pm \frac{\sigma}{L_X(z)}; \quad f_{S_{\Delta}}(z) = \sigma^2; \quad d_{IS}(f_{S_{\Delta}}, \sigma^2) = 0.$$

de lo cual, $G_{\Delta}(z) \rightarrow G_W^C(z)$ si $\sigma \rightarrow 0$. Este ejemplo trivial muestra que la solución de Wiener es mejor aún cuando la distorsión pueda ser infinita. Sin embargo, la solución de Wiener es un caso particular.

- iv) Sea $f_s(z) = T$, y sea $f_n(z) = N$, $x[n] = s[n] + n[n]$, donde $n[n]$ es ortogonal a $s[n]$; entonces $f_x(z) = N + T$, $L_X(z) = \sqrt{N + T}$,

$$f_{SX}(z) = f_S(z) = T$$

El filtro de Wiener causal es:

$$G_W^C(z) = \frac{1}{L_X(z)} \left[\frac{f_{SX}(z)}{L_X(1/z)} \right]^+ = \frac{1}{\sqrt{N+T}} \left[\frac{T}{\sqrt{N+T}} \right]^+ = \frac{T}{N+T}$$

La densidad espectral de potencia utilizando el filtro de Wiener es:

$$\begin{aligned} f_{S_W}^C(z) &= G_W^C(z) G_W^C(1/z) f_X(z) - f_{SX}(1/z) G_W^C(z) - f_{SX}(z) G_W^C(1/z) + f_S(z) = \\ &= \frac{T^2}{N+T} - \frac{T^2}{N+T} - \frac{T^2}{N+T} + T = \frac{NT}{N+T} = f_{S_W}^{NC}(z) \end{aligned}$$

Ahora, notando que $f(z) = 0$, (5.22.1) luce así:

$$\Delta(z) = \left[\frac{\sigma^2 \Delta(z)}{\Delta(z)\Delta(1/z) + \frac{NT}{N+T}} \right]^+ ; \text{ esta ecuación}$$

puede ser satisfecha por una constante. Sea $\Delta(z) = \delta_0$; entonces

$$\delta_0 = \frac{\sigma^2 \delta_0}{\delta_0^2 + \frac{NT}{N+T}} \Rightarrow \delta_0^2 = \left(\sigma^2 - \frac{NT}{N+T} \right)$$

$$\Rightarrow \delta_0 = \pm \sqrt{\sigma^2 - \frac{NT}{N+T}}; \quad G_\Delta(z) = \frac{T}{N+T} \pm \frac{\sqrt{\sigma^2 - \frac{NT}{N+T}}}{\sqrt{N+T}}$$

$$G_\Delta(z) = \frac{T}{N+T} \pm \frac{\sqrt{(N+T)\sigma^2 - NT}}{N+T}$$

Obsérvese que la solución sólo es real si $\sigma^2 \geq \frac{NT}{N+T}$. La densidad espectral de potencia utilizando el filtro $G_\Delta(z)$, es

$$f_{\tilde{s}_\Delta}^c(z) = G_\Delta(z)G_\Delta(1/z)f_x(z) - f_{s_x}(z)G_\Delta(1/z) - f_{s_x}(1/z)G_\Delta(z) + f_s(z) = \sigma^2$$

Concluyendo:

$$d_{IS}(f_{\tilde{s}_w}^c, \sigma^2) = d_{IS}\left(\frac{NT}{N+T}, \sigma^2\right)$$

$$d_{IS}(f_{\tilde{s}_\Delta}^c, \sigma^2) = d_{IS}(\sigma^2, \sigma^2) = 0, \text{ en tanto } \sigma^2 \geq \frac{NT}{N+T}.$$

Se observa además, que $G_\Delta(z) \rightarrow G_w(z)$ en tanto que $\sigma^2 \rightarrow \frac{NT}{N+T}$.

- v) En el inciso a), pág (5.16), se vió que si $\sigma^2 \rightarrow 0$, en el límite, la solución es el filtro de Wiener. A partir de (5.22.1) puede verse que en general, si $\sigma^2 \neq 0$, la solución no es el filtro de Wiener. Haciendo $\Delta(z) = 0$ en (5.22.1):

$$0 = \left[\frac{-\sigma^2 f(z)}{f_{\tilde{s}_w}^c(z)} \right]^+ \quad \text{debiera cumplirse.}$$

Pero en general no es cierto, ya que $f_{\tilde{s}_w}^c(z)$ contiene tanto potencias positivas como negativas de z . Esto concuerda con los ejemplos en iii) y iv) y es congruente con el resultado en b).

FILTRO CAUSAL, $n_0 = -\infty$. CONDICIONES SUFICIENTES PARA EL OPTIMO .

- a) Si el proceso de referencia es blanco, de potencia σ^2 , la $d_{IS}(f_s(e^{j\omega}), \sigma^2)$ es mínima en el filtro de Wiener, si $\sigma^2 \rightarrow 0$.

Prueba : Sea $G_\Delta(z) = G_W^C(z) + \frac{\Delta(z)}{L_X(z)}$ y sea $f_{\tilde{s}_\Delta}(e^{j\omega})$ la densidad espectral del error usando éste filtro. Sea $f_{\tilde{s}_W}^C(e^{j\omega})$ la densidad espectral del error cuando $\Delta(z) = 0$. Se puede escribir entonces:

$$d_{SI}(f_{\tilde{s}_\Delta}(e^{j\omega}), \sigma^2) = d_{IS}(f_{\tilde{s}_\Delta}(e^{j\omega}), f_{\tilde{s}_W}^C(e^{j\omega})) + d_{IS}(f_{\tilde{s}_W}^C(e^{j\omega}), \sigma^2) + 1 + \int_{-\pi}^{\pi} \left\{ \frac{f_{\tilde{s}_\Delta}(e^{j\omega})}{\sigma^2} - \frac{f_{\tilde{s}_\Delta}(e^{j\omega})}{f_{\tilde{s}_W}^C(e^{j\omega})} - \frac{f_{\tilde{s}_W}^C(e^{j\omega})}{\sigma^2} \right\} \frac{d\omega}{2\pi}$$

usando ahora $1 = \int_{-\pi}^{\pi} \frac{f_{\tilde{s}_W}^C(e^{j\omega})}{f_{\tilde{s}_W}^C(e^{j\omega})} \frac{d\omega}{2\pi}$, se obtiene :

$$d_{IS}(f_{\tilde{s}_\Delta}(e^{j\omega}), \sigma^2) - d_{IS}(f_{\tilde{s}_W}^C(e^{j\omega}), \sigma^2) = d_{IS}(f_{\tilde{s}_\Delta}(e^{j\omega}), f_{\tilde{s}_W}^C(e^{j\omega})) +$$

$$\int_{-\pi}^{\pi} \{ |\Delta(e^{j\omega}) - f(e^{j\omega})|^2 - |f(e^{j\omega})|^2 \} \left\{ \frac{1}{\sigma^2} - \frac{1}{f_{\tilde{s}_W}^C(e^{j\omega})} \right\} \frac{d\omega}{2\pi}$$

≥ 0 si $\sigma^2 \rightarrow 0$ y $\Delta(e^{j\omega})$ es finito.

- b) Si en lugar de usar una densidad espectral de potencia constante para el proceso de referencia, se considera que éste tiene una densidad espectral

de potencia de la familia paramétrica $\sigma^2(\omega) = f_{S_W}^C(e^{j\omega}) / (1 + a f_{S_W}^C(e^{j\omega}))$, entonces la distorsión de Itakura-Saito $d_{IS}(f_{S_\Delta}^C(e^{j\omega}), \sigma^2(\omega))$ es mínima cuando $\Delta(z) \equiv 0$, ésto es, cuando se usa el filtro de Wiener causal, independientemente del valor del parámetro $a \geq 0$.

Prueba : En forma análoga a como se hizo en a), se puede escribir:

$$d_{IS}(f_{S_\Delta}^C(e^{j\omega}), \sigma^2(\omega)) - d_{IS}(f_{S_W}^C(e^{j\omega}), \sigma^2(\omega)) = d_{IS}(f_{S_\Delta}^C(e^{j\omega}), f_{S_W}^C(e^{j\omega})) +$$

$$a \int_{-\pi}^{\pi} \{ |\Delta(e^{j\omega}) - f(e^{j\omega})|^2 - |f(e^{j\omega})|^2 \} \frac{d\omega}{2\pi} = d_{IS}(f_{S_\Delta}^C(e^{j\omega}), f_{S_W}^C(e^{j\omega})) +$$

$$a \int_{-\pi}^{\pi} |\Delta(e^{j\omega})|^2 \frac{d\omega}{2\pi} \geq 0$$

donde en la última igualdad se ha aplicado ortogonalidad.

Discusión : El espectro del error utilizando el filtro de Wiener causal resulta ser lo más cercano posible a la familia paramétrica de densidades espectrales de potencia del proceso de referencia

$$\sigma^2(\omega) = \frac{f_{S_W}^C(e^{j\omega})}{1 + a f_{S_W}^C(e^{j\omega})}, \text{ el cual es función de los datos del proble-}$$

ma. Sin embargo, cuando "a" crece indefinidamente, $\sigma^2(\omega) \approx \frac{1}{a} + 0$, y esto es lo mismo que se ha establecido en a): El filtro de Wiener causal es el que "más blanquea" si la potencia del proceso de referencia tiende a cero.

- c) Cuando los datos del problema son tales que se satisface (5.19.1), (5.21.1) toma la forma $f_{S_W}^C(e^{j\omega}) = f_{S_W}^{NC}(e^{j\omega})$ y (5.21.3) la forma

$$f_{\tilde{S}_\Delta}^C(z) = \Delta(z)\Delta(1/z) + f_{\tilde{S}_W}^{NC}(z)$$

Entonces, dada una densidad de potencia espectral deseada para el error $|K(\omega)|^2$ tal que

$$|K(\omega)|^2 > f_{\tilde{S}_W}^{NC}(e^{j\omega}) = f_{\tilde{S}_W}^C(e^{j\omega}),$$

existe $\Delta(z)$, de fase mínima, tal que

$$|\Delta(e^{j\omega})|^2 = |K(\omega)|^2 - f_{\tilde{S}_W}^{NC}(e^{j\omega}).$$

No obstante, si $|K(\omega)|^2 = f_{\tilde{S}_W}^{NC}(e^{j\omega})$ en cierto intervalo, $\Delta(z)$ no es causal, excepto si la igualdad se cumple para toda ω , en cuyo caso $\Delta(z) = 0$. Se concluye que la solución obtenida para el caso no-causal cuando

$$\sigma^2 \leq \min_{\omega} f_{\tilde{S}_W}^{NC}(e^{j\omega})$$

es ahora la misma.

- d) Aplicación al predictor: Sea $L_{\tilde{S}_W}(z)$ de fase mínima dado por $L_{\tilde{S}_W}(z) = \sum_{k=0}^{r-1} \ell_k z^{-k}$, y recuérdese que $L_S(z) = \sum_{k=0}^{\infty} \ell_k z^{-k}$, de tal forma que la densidad espectral de potencia del error usando el predictor de Wiener es

$$f_{\tilde{S}_W}^C(z) = L_{\tilde{S}_W}(z) L_{\tilde{S}_W}(1/z).$$

$$\text{Sea } G_{\Delta}(z) = \frac{z^r}{L_S(z)} \left\{ \sum_{k=r}^{\infty} (\ell_k + \delta_k) z^{-k} \right\} = G_W^C(z) + \frac{\Delta(z)}{L_X(z)}$$

otro filtro causal tal que

$$L_{\tilde{s}_\Delta}(z) \triangleq \sum_{k=0}^{r-1} \ell_k z^{-k} - \sum_{k=r}^{\infty} \delta_k z^{-k}, \quad f_{\tilde{s}_\Delta}(z) = L_{\tilde{s}_\Delta}(z) L_{\tilde{s}_\Delta}(1/z)$$

también es de fase mínima.

Entonces :

$$\begin{aligned} d_{IS}(f_{\tilde{s}_\Delta}, \sigma^2) &= \int_{-\pi}^{\pi} \left\{ \frac{f_{\tilde{s}_\Delta}(e^{j\omega})}{\sigma^2} - \ln \frac{f_{\tilde{s}_\Delta}(e^{j\omega})}{\sigma^2} - 1 \right\} \frac{d\omega}{2\pi} \\ &= \frac{1}{\sigma^2} \left\{ \sum_{k=0}^{r-1} \ell_k^2 + \sum_{k=r}^{\infty} \delta_k^2 \right\} - \ln \frac{\ell_0^2}{\sigma^2} - 1 \\ &\geq \frac{1}{\sigma^2} \sum_{k=0}^{r-1} \ell_k^2 - \ln \frac{\ell_0^2}{\sigma^2} - 1 = d_{IS}(f_{\tilde{s}_w}, \sigma^2) \end{aligned}$$

con igualdad si y sólo si $\delta_k = 0$, $k = r, r+1, \dots$, ésto es, si y sólo si $\Delta(z) = 0$.

En resumen, dentro de la clase de predictores que producen un filtro de error de fase mínima, el predictor de Wiener es óptimo en el sentido de minimizar $d_{IS}(f_{\tilde{s}}(e^{j\omega}), \sigma^2)$, cuando los procesos X y S se aplican desde el pasado infinitamente remoto.

Nota : Sea el predictor de un paso adelante de orden N , [5], definido por

$$H_N(z) = - \sum_{k=1}^N a_k^N z^{-k};$$

el filtro de error está definido por

$$E_N(z) = \sum_{k=0}^N a_k^N z^{-k}, \quad a_0^N = 1.$$

Se sabe que el filtro de error para este predictor según se determina por el algoritmo de Levinson tiene todos sus ceros dentro del círculo unitario [5]. (Recuérdese que se están considerando procesos regulares).

Por tanto y por lo visto en ésta sección, el predictor obtenido bajo el criterio de error cuadrático medio mínimo es óptimo en el sentido ya descrito.

- e) El ejemplo ii) de la ecuación (5.22.1) merece especial atención. Repitiendo (5.21.3) :

$$f_{\tilde{S}_\Delta}(z) = [\Delta(z) - f(z)][\Delta(1/z) - f(1/z)] + f_{\tilde{S}_W}^{NC}(z) \quad (5.31.1)$$

y excepto cuando $f(z) \equiv 0$ (situación en que el filtro de Wiener no causal es igual al filtro de Wiener causal), $f_{\tilde{S}_\Delta}(e^{j\omega}) > f_{\tilde{S}_W}^{NC}(e^{j\omega})$ a.e. (ésto es, para algunas frecuencias aisladas $|\Delta(e^{j\omega}) - f(e^{j\omega})|^2$ puede ser igual a cero, lo cual puede verse construyendo algunos ejemplos). Por tanto, aplicando el teorema de Paley-Wiener

$$f_{\tilde{S}_\Delta}(z) - f_{\tilde{S}_W}^{NC}(z) = L(z) L(1/z) \quad (5.31.2)$$

donde $L(z)$ es de fase mínima.

Sea $H_{0f}(z)$ un filtro pasa todas (estable), con sus ceros iguales a los

polos de $f(z)$, de tal forma que el producto $f(z)H_{of}(z)$ es causal y estable. De (5.31.1) y (5.31.2) se llega a

$$|\Delta(e^{j\omega})H_{of}(e^{j\omega}) - f(e^{j\omega})H_{of}(e^{j\omega})|^2 = |L(e^{j\omega})|^2$$

y por tanto existe un filtro pasa todas (estable) $H_{oL}(z)$ tal que

$$\Delta(z)H_{of}(z) - f(z)H_{of}(z) = H_{oL}(z)L(z) \quad (5.32.1)$$

ésto es :

$$\Delta(z) = \frac{H_{oL}(z)L(z)}{H_{of}(z)} + f(z) \quad (5.32.2)$$

Como $\Delta(z)$ es causal, debe de cumplirse que

$$\left[\frac{H_{oL}(z)L(z)}{H_{of}(z)} \right]^- = -f(z) \quad (5.32.3)$$

Entonces, para encontrar $\Delta(z)$ se puede primero resolver (5.32.3) para $H_{oL}(z)$ y después tomar

$$\Delta(z) = \left[\frac{H_{oL}(z)L(z)}{H_{of}(z)} \right]^+ \quad (5.32.4)$$

Si (5.32.3) no tiene solución $H_{oL}(z)$ pasa todas, entonces el espectro $f_{S_{\Delta}}(z)$ no puede alcanzarse con un filtro causal. En particular interesa que $f_{S_{\Delta}}(z) = \sigma^2$, y es evidente que deben cumplirse, por lo menos :

$$i) \quad \sigma^2 > \max_{\omega} f_{S_{\omega}}^{NC}(e^{j\omega})$$

$$ii) \sigma^2 > \int_{-\pi}^{\pi} f_{S_W}^C(e^{j\omega}) \frac{d\omega}{2\pi}$$

A continuación, se ilustra este procedimiento con un ejemplo.

Ejemplo : Se sabe que $f_{S_W}^C(e^{j\omega}) = |f(e^{j\omega})|^2 + f_{S_W}^{NC}(e^{j\omega})$

Sea $f(z) = \frac{0.5z}{z-2}$, $|z| < 1$. Sea el espectro deseado para el error

$$f_{S_{\Delta}}^C(z) = \frac{(z-1/4)(z^{-1}-1/4)}{(z-1/3)(z^{-1}-1/3)} + f_{S_W}^C(z)$$

Entonces

$$L(z)L(1/z) = f_{S_{\Delta}}^C(z) - f_{S_W}^{NC}(z) = \frac{(z-1/4)(z^{-1}-1/4)}{(z-1/3)(z^{-1}-1/3)} + \frac{1/4}{(z-2)(z^{-1}-2)}$$

y después de un poco de álgebra:

$$L(z)L(1/z) = \frac{72z^4 - 498z^3 + 949z^2 - 498z + 72}{(12z-4)(12-4z)(z-2)(1-2z)} ;$$

se puede mostrar que una de las raíces de la ecuación resolvente de 3er. grado del polinomio del numerador es 2, y luego se encuentra que las raíces del numerador son las raíces de las ecuaciones:

$$24z^2 - (83 + \sqrt{449})z + 24 = 0$$

$$y \quad 24z^2 + (-83 + \sqrt{449})z + 24 = 0 ; \text{ ésto es:}$$

$$z_1 = 4.09716284, \quad z_2 = 0.24407133, \quad z_3 = 2.09901960, \quad z_4 = 0.47641289$$

$$\text{De aquí que } L(z) = \frac{\sqrt{72}(z-z_2)(z-z_4)}{(12z-4)(1-2z)}$$

De los polos de $f(z)$, se obtiene $H_{of}(z) = \frac{0.5z-1}{z-0.5} = \frac{z-2}{2z-1}$;

como $f(z)$ tiene tan sólo un polo, se propone

$$H_{oL}(z) = \frac{\alpha z - 1}{z - \alpha},$$

y debe encontrarse α tal que $\left[\frac{H_{oL}(z)L(z)}{H_{of}(z)} \right]^- = -f(z)$:

$$\left[\frac{H_{oL}(z)L(z)}{H_{of}(z)} \right]^- = \left[\left(\frac{\alpha z - 1}{z - \alpha} \right) \left(\frac{\sqrt{72} (z - z_2)(z - z_4)}{(12z - 4)(1 - 2z)} \right) \left(\frac{2z - 1}{z - 2} \right) \right]^- = \frac{Dz}{z - 2},$$

$$\text{donde } D = \lim_{z \rightarrow 2} \left\{ \left(\frac{\alpha z - 1}{z - \alpha} \right) \left(\frac{\sqrt{72} (z - z_2)(z - z_4)}{(12z - 4)(1 - 2z)} \right) \left(\frac{2z - 1}{z} \right) \right\}$$

$$D = \left(\frac{2\alpha - 1}{2 - \alpha} \right) (-0.56751901) = -0.5$$

y de aquí, $\alpha = 0.95870493$; el valor encontrado implica que $H_{oL}(z)$ es estable y finalmente,

$$\Delta(z) = \left[\frac{H_{oL}(z)L(z)}{H_{of}(z)} \right]^+.$$

Se concluye que siempre que pueda encontrarse $H_{oL}(z)$ (estable) tal que $f_{\bar{s}_\Delta}(z) = \sigma^2$, se ha minimizado $d_{IS}(f_{\bar{s}_\Delta}(e^{j\omega}), \sigma^2) = d_{IS}(\sigma^2, \sigma^2) = 0$.

FILTRO CAUSAL, $n_0 = -\infty$. UNA COTA SUPERIOR PARA LA DISTORSION

Se ha visto que el filtro de Wiener causal no minimiza, en general, la $d_{IS}(f_{\tilde{s}}(e^{j\omega}), \sigma^2)$ a menos que σ^2 tienda a cero. Sin embargo, como se muestra a continuación, minimiza una cota superior de la d_{IS} . El resultado puede enunciarse así:

La distorsión de Itakura-Saito entre la densidad espectral de potencia del proceso de error evaluada en el filtro $G_{\Delta}(e^{j\omega}) = G_W^C(e^{j\omega}) + \frac{\Delta(e^{j\omega})}{L_X(e^{j\omega})}$, literalmente $f_{\tilde{s}_{\Delta}}(e^{j\omega})$, y una densidad espectral de potencia constante σ^2 , está acotada superiormente por:

$$d_{IS}(f_{\tilde{s}_{\Delta}}(e^{j\omega}), \sigma^2) \leq \int_{-\pi}^{\pi} \frac{|\Delta(e^{j\omega})|^2}{\sigma^2} \frac{d\omega}{2\pi} + \int_{-\pi}^{\pi} \frac{|f(e^{j\omega})|^2}{\sigma^2} \frac{d\omega}{2\pi} + d_{IS}(f_{\tilde{s}_W^{NC}}(e^{j\omega}), \sigma^2).$$

Esta cota se minimiza con $\Delta(z) \equiv 0$, ésto es, $G_{\Delta}(z) = G_W^C(z)$, independientemente del valor de σ^2 .

Prueba Con $f_{\tilde{s}_{\Delta}}(e^{j\omega})$ dada por (5.21.3),

$$d_{IS}(f_{\tilde{s}_{\Delta}}(e^{j\omega}), \sigma^2) = \int_{-\pi}^{\pi} \frac{|\Delta(e^{j\omega}) - f(e^{j\omega})|^2 + f_{\tilde{s}_W^{NC}}(e^{j\omega})}{\sigma^2} \frac{d\omega}{2\pi} - \int_{\pi}^{\pi} \ln \left\{ \frac{|\Delta(e^{j\omega}) - f(e^{j\omega})|^2 + f_{\tilde{s}_W^{NC}}(e^{j\omega})}{\sigma^2} \right\} \frac{d\omega}{2\pi} - 1 =$$

$$= \int_{-\pi}^{\pi} \frac{|\Delta(e^{j\omega})|^2 + |f(e^{j\omega})|^2 + f_{S_w}^{NC}(e^{j\omega})}{\sigma^2} \frac{d\omega}{2\pi} - \int_{-\pi}^{\pi} \ln \left(\left\{ \frac{|\Delta(e^{j\omega}) - f(e^{j\omega})|^2}{f_{S_w}^{NC}(e^{j\omega})} + 1 \right\} \frac{d\omega}{2\pi} - 1 \right)$$

donde se ha aplicado la ortogonalidad entre $\Delta(e^{j\omega})$ y $f(e^{j\omega})$. Observando que $1 + \frac{|\Delta(e^{j\omega}) - f(e^{j\omega})|^2}{f_{S_w}^{NC}(e^{j\omega})} \geq 1$ para toda ω , se concluye que

$$\int_{-\pi}^{\pi} \ln \left(\frac{|\Delta(e^{j\omega}) - f(e^{j\omega})|^2}{f_{S_w}^{NC}(e^{j\omega})} + 1 \right) \frac{d\omega}{2\pi} \geq 0 \quad \text{y de aquí}$$

se obtiene la cota propuesta. Obviamente, el mínimo de la cota se obtiene cuando $\Delta(e^{j\omega}) \equiv 0$; ésto termina la prueba.

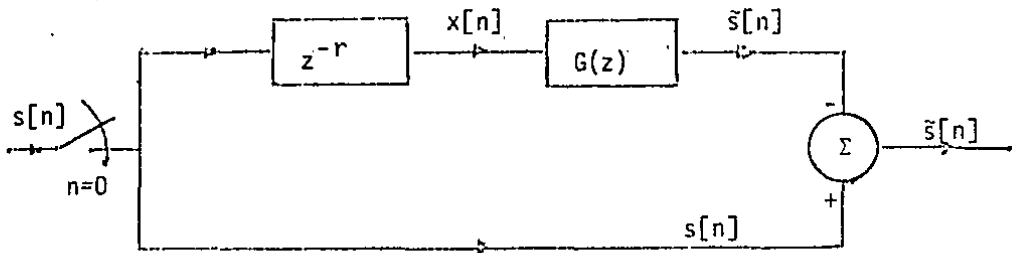
Nótese que

$$d_{IS}(f_{S_w}^C(e^{j\omega}), \sigma^2) - d_{IS}(f_{S_w}^{NC}(e^{j\omega}), \sigma^2) = \int_{-\pi}^{\pi} \frac{|f(e^{j\omega})|^2}{\sigma^2} \frac{d\omega}{2\pi} - \int_{-\pi}^{\pi} \ln \left(1 + \frac{|f(e^{j\omega})|^2}{f_{S_w}^{NC}(e^{j\omega})} \right) \frac{d\omega}{2\pi} \quad (5.36.1)$$

lo que indica que, conforme σ^2 crece, la diferencia entre las distorsiones del filtro causal y del no causal se hace menor. Puede verificarse fácilmente que en tanto $\sigma^2 \leq \min_{\omega} f_{S_w}^{NC}(e^{j\omega})$, la diferencia en (5.36.1) es positiva.

PREDICCIÓN PURA. $n_0 = 0$, CONDICIONES INICIALES NULAS

Para el filtro predictor, cuando la entrada es un proceso estocástico estrictamente estacionario la cual se aplica a partir de un tiempo finito, por ejemplo, desde $n_0 = 0$, se demuestra que la RCEC $H(\tilde{S}, \sigma^2 W)$ se minimiza con el filtro que minimiza la potencia del proceso de error, es decir, con el predictor de Wiener. No tiene importancia considerar el caso de un predictor no-causal, ya que éste es simplemente un atraso. El resultado es pues para un predictor causal y es independiente de la función de densidad de probabilidad del proceso S y de si el filtro es de fase mínima o no. El esquema que corresponde a este problema se ilustra en la figura.



De la definición de la RCEC:

$$H(\tilde{S}, \sigma^2 W) = -\bar{H}_{\tilde{S}} + \lambda n \sqrt{2\pi\sigma^2} + \frac{1}{2\sigma^2} \int_{-\pi}^{\pi} f_{\tilde{S}}(e^{j\omega}) \frac{d\omega}{2\pi} \quad (5.37.1)$$

El proceso de error es la respuesta asintóticamente estacionaria del filtro de error $H(z) = 1 - z^{-r}G(z)$ a la entrada S . Si la respuesta al pulso unitario de $H(z)$ es $h[k]$, como $G(z)$ es causal y $r \geq 1$, se tiene que

$$h[0] = \lim_{z \rightarrow \infty} H(z) = 1 \quad (5.37.2)$$

y por la ecuación (2.9.1):

$$\bar{H}_S = \bar{H}_S + \lambda n |h[0]|$$

y por (5.37.2)

$$\bar{H}_S = \bar{H}_S$$

por lo que (5.37.1) se puede reescribir como:

$$\bar{H}(\bar{S}, \sigma^2 W) = -\bar{H}_S + \lambda n \sqrt{2\pi\sigma^2} + \frac{1}{2\sigma^2} \int_{-\pi}^{\pi} f_{\bar{S}}(e^{j\omega}) \frac{d\omega}{2\pi}$$

de donde se concluye, que para minimizar $\bar{H}(\bar{S}, \sigma^2 W)$, es necesario y suficiente minimizar la potencia del proceso de error (ya que \bar{S} es asintóticamente estacionario la densidad espectral de potencia $f_{\bar{S}}(e^{j\omega})$ está definida solamente en sentido asintótico). Nótese que el resultado es independiente del valor de σ^2 .

Hay que recalcar que se tomaron condiciones iniciales nulas y que esto no es estrictamente necesario; basta con que sean conocidas, pero el resultado no se sigue si las condiciones iniciales son aleatorias.

CAPITULO 6

PROCESOS PERIODICOS Y ENTROPIA CRUZADA NORMALIZADA

Los procesos periódicos constituyen un caso particular para la determinación de la entropía cruzada, debido a ciertas características especiales en la función de distribución conjunta del proceso. Considerando en este capítulo procesos estocásticos con índice de definición el tiempo discreto, estacionarios en sentido amplio, se hará primero la distinción entre dos tipos diferentes de periodicidad; como se verá, esta diferencia es sutil para procesos de tiempo discreto.

1) Procesos periódicos en el sentido cuadrático medio o periódicos C.M.

Un proceso de tiempo discreto se denomina periódico C.M. de período N , si [5, pág. 230]:

$$E \{ |X[n+N] - X[n]|^2 \} = 0, \text{ para todo } N.$$

Esto implica, por la desigualdad de Tchebycheff

$$x(n, \omega) = x(n + N, \omega) \text{ a.s.}$$

(Esto es, para n fija y casi toda $\omega \in \Omega$, siendo Ω el espacio muestral)

Esta definición es equivalente al decir que los procesos periódicos C.M. son aquellos

para los cuales la correlación es periódica [5, pág. 230]:

$$R_x[l] = R_x[l + mN], \quad \forall l, m \text{ enteros.}$$

2) Procesos estrictamente periódicos ó de realizaciones periódicas, ó procesos circulares:

Un proceso de tiempo discreto estacionario en sentido amplio es estrictamente periódico, de período N , si [20, pág. 76; 21, pág. 461]:

$$x[n + mN] = x[n], \quad \forall m, n \text{ enteros.}$$

Se demuestra en el apéndice H.1 que dado un proceso discreto periódico C.M. $X[n]$ puede definirse un proceso estrictamente periódico $\hat{X}[n]$ tal que

a) $R_{\hat{X}}[l] = R_X[l]$

b) $E\{|X[n] - \hat{X}[n]|^2\} = 0, \quad \forall n.$

c) $P\{x(n, \omega) = \hat{x}(n, \omega)\} = 1 \quad \forall n \text{ fija.}$

Adicionalmente, se discute en el apéndice H el hecho de que, para procesos discretos, $P\{x(n, \omega) = \hat{x}(n, \omega) \quad \forall n\} = 1$, lo que se representará aquí por $P\{x(\cdot) = \hat{x}(\cdot)\} = 1$. En otras palabras, para procesos discretos periódicos C.M., las realizaciones no periódicas ocurren con probabilidad cero.

En general, éste no es el caso para procesos de tiempo continuo, porque

el índice de definición no es contable.

La relación $P\{x(n,\omega) = \hat{x}(n,\omega)\} = 1, \forall n$ fija, se conoce como equivalencia estocástica [15, pág. 210; 22, pág. 189].

Ya que los procesos estocásticamente equivalentes tienen distribuciones de dimensión finita iguales, se justifica usar, para propósitos de entropía cruzada, el proceso $\hat{X}[n]$ en lugar de $X[n]$, siempre que se defina la entropía cruzada entre procesos periódicos C.M. en términos de distribuciones de dimensión finita. Esto tiene la ventaja, como se verá más adelante, de poder representar la entropía cruzada entre procesos periódicos C.M. en términos de la distribución de los coeficientes C_k (véase apéndice H).

RAZON DE CAMBIO DE ENTROPIA CRUZADA ENTRE DOS PROCESOS PERIODICOS :

Recuérdese que por función de densidad se entiende una función $P_{\underline{x}}^k(\underline{x})$ medible según Lebesgue, la cual satisface ($\underline{x} \in R^k$) :

$$a) \quad P_{\underline{x}}^k(\underline{x}) \geq 0$$

$$b) \quad \int_{R^k} P_{\underline{x}}^k(\underline{x}) m(d\underline{x}) = 1$$

donde $m(A)$ es la medida de Lebesgue del conjunto A . [9, pág. 39].

Considérese ahora el vector de variables aleatorias, dado por un segmento de

longitud k de un proceso de período N , ($k > N$)

$$\underline{X} = (X_1, X_2, \dots, X_N, X_{N+1}, \dots, X_{(k-1)N}, \dots, X_k) \in R^k$$

el cual contiene $(k-1)$ períodos completos (a menos que $k=KN$, en cuyo caso contiene K períodos completos).

Sea $F_{\underline{X}}^k(\underline{x})$ la función de distribución de probabilidad conjunta de estas k variables, y sea $F_{\underline{X}}^N(x_1, \dots, x_N)$ la función de distribución conjunta de las primeras N variables.

La condición de periodicidad, dada por

$$X_1 = X_{N+1} = \dots = X_{(K-1)N+1} \tag{6.4.1}$$

$$X_2 = X_{N+2} = \dots = X_{(K-1)N+2}$$

$$X_N = X_{2N} = \dots = X_{(K-1)N} (= X_{kN} \text{ si } k = KN)$$

define en R^k una región A^k que coincide con R^N . Además, se sabe que si (X_1, X_2, \dots, X_k) en (Ω, f, P) son variables aleatorias que inducen el espacio de probabilidad (R^k, β^k, μ_k) , entonces se tiene [23, pág. 45]

$$\begin{aligned} \int_{\Delta} P(d\omega) &= \int_{B^k \in \beta^k} \dots \int \mu_k(dx_1, \dots, dx_k) = \int_{B^k \in \beta^k} \dots \int \mu_k(\underline{dx}) = \\ &= P\{\omega \in \Delta\} = P\{\underline{X} \in B^k\}; \quad \chi(\Delta) = B^k. \end{aligned}$$

Por tanto, en $(A^k)^c$, $\mu_k(\underline{dx}) = 0$, lo que es una forma de decir que las realizaciones no periódicas tienen probabilidad nula; se está suponiendo que μ_k es una medida de probabilidad regular, o distribución, esto es $\mu_k(\underline{dx}) = dF_{\underline{x}}^k(\underline{x})$.

En A^k , $dF_{\underline{x}}^k(\underline{x})$ puede ser diferente de cero, pero dicha región A^k tiene medida de Lebesgue cero en R^k .

Por tanto, $\int_{R^k} p_{\underline{x}}^k(\underline{x})m(\underline{dx}) = 0$; en otras palabras, no existe función de densidad. Esto último implica que la distribución μ_k no es absolutamente continua (con respecto a la medida de Lebesgue [9, pág. 68]). Por esto, la definición de entropía cruzada que hasta aquí se había estado usando no es aplicable y se recurre a la definición más general [3]:

$$H(\gamma_k, \mu_k) \triangleq \begin{cases} \int \left[\frac{d\gamma_k}{d\mu_k} \right] \ln \left[\frac{d\gamma_k}{d\mu_k} \right] \mu_k(\underline{dx}) & \text{si } \gamma_k \ll \mu_k \\ +\infty & \text{de otra forma} \end{cases} \quad (6.5.1)$$

donde $\underline{x} \in R^k$; $\gamma_k, \mu_k \in P^k =$ distribuciones de dimensión k ;

$\gamma_k \ll \mu_k$ significa γ_k absolutamente continua con respecto a μ_k ;

$\left[\frac{d\gamma_k}{d\mu_k} \right]$ es la derivada de Radon-Nicodym de γ_k con respecto a μ_k .

En correspondencia, hay una definición más general para la RCEC :

$$\bar{H}(X, Y) = H(\gamma, \mu) = \lim_{k \rightarrow \infty} \begin{cases} \frac{1}{k} \int \left[\frac{d\gamma_k}{d\mu_k} \right] \ln \left[\frac{d\gamma_k}{d\mu_k} \right] \mu_k(\underline{dx}), & \gamma_k \ll \mu_k \\ +\infty & \text{de otra forma.} \end{cases} \quad (6.5.2)$$

siempre y cuando existan $\gamma, \mu \in P^\infty =$ distribuciones de orden infinito.

Proposición 1 : Sea $\gamma_N \ll \mu_N$, y sean γ_N y μ_N absolutamente continuas con respecto a la medida de Lebesgue, es decir, existen $\left[\frac{d\gamma_N}{d\mu_N} \right]$, $P_{\underline{x}}^N(x_1, \dots, x_N)$ y $P_{\underline{y}}^N(x_1, \dots, x_N)$, que son la derivada de Radon-Nicodym y las funciones de densidad de los vectores formados por un período de los procesos X y Y , los cuales tienen el mismo período.

Entonces, la entropía cruzada entre los vectores $\underline{x}, \underline{y} \in \mathbb{R}^k$, está dada por un sólo período ($k > N$) :

$$H(\underline{x}, \underline{y}) = \int \cdots \int P_{\underline{x}}^N(\alpha_1, \dots, \alpha_N) \ln \frac{P_{\underline{x}}^N(\alpha_1, \dots, \alpha_N)}{P_{\underline{y}}^N(\alpha_1, \dots, \alpha_N)} d\alpha_1, \dots, d\alpha_N$$

y la RCEC es :

$$\bar{H}(X, Y) = 0$$

Demostración :

En A^k , tomando en cuenta (6.4.1) :

$$F_{\underline{x}}^k(x) = F_{\underline{x}}^N(x_1, \dots, x_N) P\{X_{N+1} \leq x_1, \dots, X_k \leq x_i \mid x_1, \dots, x_N\}$$

donde $i \in \{1, 2, \dots, N\}$, y esto es

$$F_{\underline{x}}^k(\underline{x}) = F_{\underline{x}}^N(x_1, \dots, x_N) .$$

Por tanto, para el proceso X

$$\gamma_k(\underline{dx}) = \begin{cases} dF_{\underline{x}}^N(x_1, \dots, x_N), & \underline{dx} \in A^k, \\ 0 & \underline{dx} \notin A^k \end{cases} \quad (6.7.1)$$

donde $F_{\underline{x}}^N(x_1, \dots, x_N)$ es la función de distribución conjunta de las variables aleatorias X_1, \dots, X_N , y para el proceso Y

$$\mu_k(\underline{dx}) = \begin{cases} dG_Y^N(x_1, \dots, x_N), & \underline{dx} \in A^k \\ 0 & \underline{dx} \notin A^k \end{cases} \quad (6.7.2)$$

donde $G_Y^N(x_1, \dots, x_N)$ es la función de distribución conjunta de las variables aleatorias Y_1, \dots, Y_N ; ésto es:

$$G_Y^N(x_1, \dots, x_N) = P\{Y_1 \leq x_1, \dots, Y_N \leq x_N\}$$

Se procede ahora a mostrar que $\gamma_k \ll \mu_k$: Sea cualquier conjunto $B^k \in \mathcal{B}^k$

$$\gamma_k(B^k) = \int_{B^k \in \mathcal{B}^k} \dots \int \gamma_k(\underline{dx}), \quad \underline{dx} \in R^k;$$

por (6.7.1) :

$$\gamma_k(B^k) = \int_{B^k \cap A^k \in \mathcal{B}^N} \dots \int dF_{\underline{x}}^N(x_1, \dots, x_N); \quad (6.7.3)$$

por ser γ_N absolutamente continua (con respecto a la medida de Lebesgue "m" en R^N) :

$$\gamma_k(B^k) = \int \cdot \cdot \cdot \int_{B^N} P_{\underline{x}}^N(\underline{x}) m(d\underline{x}), \quad \underline{x}, \quad d\underline{x} \in R^N$$

por tanto: $\gamma_k(B^k) = \gamma_N(B^N)$, donde $B^N = B^k \cap A^k$. Similarmente:

$$\mu_k(B^k) = \mu_N(B^N).$$

Y por ser $\gamma_N \ll \mu_N$ por hipótesis, se sigue que $\gamma_k \ll \mu_k$.

Por el teorema de Radon-Nicodym, existe entonces la derivada de Radon-Nicodym

$\left[\frac{d\gamma_k}{d\mu_k} \right]$ tal que

$$\gamma_k(B^k) = \int \cdot \cdot \cdot \int_{B^k} \left[\frac{d\gamma_k}{d\mu_k} \right] \mu_k(d\underline{x}), \quad d\underline{x} \in R^k;$$

por (6.7.2)

$$\gamma_N(B^N) = \gamma_k(B^k) = \int \cdot \cdot \cdot \int_{B^N} \left[\frac{d\gamma_k}{d\mu_k} \right] dG_{\underline{y}}^N(x_1, \dots, x_N) \quad (6.8.1)$$

Como B^N es arbitrario y la derivada es única, de (6.8.1) y (6.7.3) se sigue que:

$$\left[\frac{d\gamma_k}{d\mu_k} \right] = \left[\frac{d\gamma_N}{d\mu_N} \right] = \frac{\gamma_N(d\underline{x})}{\mu_N(d\underline{x})} = \frac{P_{\underline{x}}^N(x_1, \dots, x_N)}{P_{\underline{y}}^N(x_1, \dots, x_N)} \quad (6.8.2)$$

teniendo presente (6.7.1), (6.7.2) y $B^N = B^k \cap A^k$, $d\underline{x} \in R^N$.

Usando (6.8.2) en (6.5.1) y (6.5.2) se obtiene el resultado de la proposición.

Nótese que se ha considerado que ambos procesos X y Y tienen el mismo período N . El caso en que tienen diferente período carece de interés, pues uno se distingue del otro simplemente por el período.

Con el fin de tener para procesos periódicos una medida de similitud normalizada, semejante a la RCEC de los procesos aperiódicos, se usará para procesos periódicos C.M. la entropía cruzada normalizada.

$$H^*(X^N, Y^N) = \frac{1}{N} H(P_{\underline{X}}^N, P_{\underline{Y}}^N)$$

donde X^N, Y^N son procesos de período N , $P_{\underline{X}}^N$ y $P_{\underline{Y}}^N$ son funciones de densidad de probabilidad conjunta correspondientes a las variables aleatorias en un período de X^N y Y^N , y $H(P_{\underline{X}}^N, P_{\underline{Y}}^N)$ está definida en (2.2.1).

RELACION DE $H^*(X^N, Y^N)$ CON LA DISTORSION DE ITAKURA-SAITO :

Sean $X^N = \{x[n]\}_{n=-\infty}^{\infty}$, $Y^N = \{y[n]\}_{n=-\infty}^{\infty}$, $E^N = \{e[n]\}_{n=-\infty}^{\infty}$ procesos periódicos C.M., con media cero, ésto es, $E\{x[n]\} = E\{y[n]\} = E\{e[n]\} = 0$ para toda n .

Del apéndice H.1, se sabe que

$$x[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} C_k^x W_N^{-kn} \quad \text{a.s.} \quad (6.10.1)$$

y que Y^N y E^N tiene desarrollos similares. Haciendo $C_k^x = a_k^x - jb_k^x$, $0 \leq k < N$; $b_0 = 0$, y usando el hecho de que $(C_{N-k}^x)^* = C_k^x$, (6.10.1) puede escribirse en la forma:

$$x[n] = \frac{1}{\sqrt{N}} \left\{ a_0^x + 2 \sum_{k=1}^M \left[a_k^x \cos\left(\frac{2\pi kn}{N}\right) + b_k^x \sin\left(\frac{2\pi kn}{N}\right) \right] \right\} \quad \text{a.s.} \quad (6.10.2)$$

donde $M = \begin{cases} (N-1)/2 & \text{si } N-1 \text{ es par} \\ (N-2)/2 + 1 & \text{si } N \text{ es par.} \end{cases}$

con desarrollos similares para Y^N y E^N .

De la definición de los coeficientes C_k^x (apéndice H.1) :

$$C_k^x = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] W_N^{kn}$$

se obtiene :

$$a_k^x = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] \cos\left(\frac{2\pi kn}{N}\right), \quad 0 \leq k \leq M \quad (6.10.3)$$

$$b_k^x = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] \sin\left(\frac{2\pi kn}{N}\right), \quad 1 \leq k \leq M \quad (6.10.4)$$

Por otro lado, usando la relación (véase apéndice H.1)

$$a) \quad E\{(C_k^X)(C_j^X)^*\} = T_k \delta[k-j], \quad 0 \leq k \leq M, \quad 0 \leq j \leq M$$

junto con la igualdad $(C_{N-j}^X)^* = C_j^X$, se concluye que

$$b) \quad E\{C_k^X C_j^X\} = 0, \quad 0 \leq k \leq M, \quad 1 \leq j \leq M;$$

sustituyendo en a) y b) $C_k^X = a_k^X - jb_k^X$ se obtienen sistemas de ecuaciones cuya solución establece las siguientes relaciones:

$$\left. \begin{aligned} E\{(a_k^X)^2\} &= E\{(b_k^X)^2\} = \frac{1}{2} T_k \\ E\{a_k^X b_k^X\} &= 0 \end{aligned} \right\} \quad 1 \leq k \leq M$$

$$E\{a_k^X a_\ell^X\} = E\{b_k^X b_\ell^X\} = E\{a_k^X b_\ell^X\} = E\{b_k^X a_\ell^X\} = 0, \quad 1 \leq k \leq M, \quad 1 \leq \ell \leq M, \quad k \neq \ell.$$

$$E\{(a_0^X)^2\} = T_k \quad (6.11.1)$$

(obviamente, las ecuaciones análogas correspondientes a Y^N y E^N son válidas)

Se harán ahora las siguientes suposiciones:

c) N es un número non.

d) las secuencias de autocorrelación de los procesos X^N , Y^N y E^N es-

tán dadas por :

$$R_x[n] = \frac{1}{N} \sum_{k=0}^{N-1} T_k w_N^{-nk}$$

$$R_y[n] = \frac{1}{N} \sum_{k=0}^{N-1} S_k w_N^{-nk}$$

$$R_e[n] = R_y[n]$$

e) x^N y e^N son gaussianos.

Como a partir de las combinaciones lineales (6.10.3) y (6.10.4), por la suposición e), se obtienen dos vectores de $2M+1 = N$ variables aleatorias gaussianas independientes y variancias (por (6.11.1)) :

$$\begin{aligned} E\{(a_0^x)^2\} &= T_0 \\ E\{(a_k^x)^2\} &= E\{(b_k^x)^2\} = \frac{1}{2} T_k, \quad 1 \leq k \leq M \end{aligned} \quad (6.12.1)$$

$$\begin{aligned} E\{(a_0^e)^2\} &= S_0 \\ E\{(a_k^e)^2\} &= E\{(b_k^e)^2\} = \frac{1}{2} S_k, \quad 1 \leq k \leq M \end{aligned} \quad (6.12.2)$$

Nótese que (6.12.2) se cumple también para a_k^y y b_k^y , independientemente de que y^N sea gaussiano o no.

Definiendo los vectores de dimensión N

$$z_x \triangleq (a_0^x, a_1^x, b_1^x, \dots, a_M^x, b_M^x)$$

$$Z_e = (a_0^e, a_1^e, b_1^e, \dots, a_M^e, b_M^e)$$

Sus elementos tienen las funciones de densidad de probabilidad conjunta:

$$P_{Z_x}^N(\alpha_0, \alpha_1, \beta_1, \dots, \alpha_M, \beta_M) = \frac{e^{-\alpha_0^2/2T_0}}{\sqrt{2\pi T_0}} \prod_{k=1}^M \left\{ \frac{1}{\pi T_k} e^{-(\alpha_k^2 + \beta_k^2)/T_k} \right\}$$

$$P_{Z_e}^N(\alpha_0, \alpha_1, \beta_1, \dots, \alpha_M, \beta_M) = \frac{e^{-\alpha_0^2/2S_0}}{\sqrt{2\pi S_0}} \prod_{k=1}^M \left\{ \frac{1}{\pi S_k} e^{-(\alpha_k^2 + \beta_k^2)/S_k} \right\}$$

También se define el vector de dimensión N

$$Z_y \stackrel{\Delta}{=} (a_0^y, a_1^y, b_1^y, \dots, a_M^y, b_M^y)$$

Ya que las ecuaciones (6.10.3) y (6.10.4), aplicadas en un período, definen una transformación invertible, por la invariancia de la entropía cruzada bajo este tipo de transformación:

$$H^*(Y^N, X^N) = \frac{1}{N} H(P_{\underline{Y}}^N, P_{\underline{X}}^N) = \frac{1}{N} H(P_{Z_y}^N, P_{Z_x}^N)$$

Sustituyendo en (2.2.1), se obtiene :

$$H^*(Y^N, X^N) = -\frac{1}{N} H(P_{Z_y}^N) + \frac{1}{N} \left\{ \ln \sqrt{2\pi T_0} + \sum_{k=1}^M \ln \pi T_k + \frac{E\{(a_0^y)^2\}}{2T_0} + \sum_{k=1}^M E \left\{ \frac{(a_k^y)^2 + (b_k^y)^2}{T_k} \right\} \right\} \quad (6.13.1)$$

donde $H(P_{z_y}^N)$ está definida en (2.8.1).

Pero $T_k = T_{-k}$ y $S_k = S_{-k}$ (véase apéndice H.1), por lo que pueden escribirse las identidades :

$$\ln \sqrt{2\pi T_0} + \sum_{k=1}^M \ln \pi T_k = -\frac{1}{2} \left[\sum_{k=-M}^M \ln \left(\frac{S_k}{T_k} \right) \right] + \sum_{k=1}^M \ln \pi S_k + \ln \sqrt{2\pi S_0} \quad (6.14.1)$$

y

$$\frac{S_0}{2T_0} + \sum_{k=1}^M \frac{S_k}{T_k} = \frac{1}{2} \sum_{k=-M}^M \frac{S_k}{T_k} \quad (6.14.2)$$

Usando (6.14.1) y (6.14.2) en (6.13.1), se obtiene :

$$H^*(Y^N, X^N) = -\frac{1}{N} H(P_{z_y}^N) + \frac{1}{N} \left\{ \frac{1}{2} \sum_{k=-M}^M \left(\frac{S_k}{T_k} - \ln \frac{S_k}{T_k} - 1 \right) \right\} + \frac{1}{N} \left\{ \ln \sqrt{2\pi S_0} + \sum_{k=1}^M \ln \pi S_k + \frac{N}{2} \right\}$$

y como $H^*(Y^N, E^N)$ está dado por (6.13.1) con S_k en lugar de T_k y $M+1/2 = N/2$:

$$H^*(Y^N, X^N) = H^*(Y^N, E^N) + \frac{1}{2N} \sum_{k=-M}^M \left(\frac{S_k}{T_k} - \ln \frac{S_k}{T_k} - 1 \right) \quad (6.14.3)$$

Nótese que si Y^N es también gaussiano, $P_{z_y}^N(\underline{\alpha}) = P_{z_e}^N(\underline{\alpha})$, y se obtiene que

la entropía cruzada entre dos procesos gaussianos es (N impar) :

$$H^*(Y^N, X^N) = \frac{1}{2N} \sum_{k=-M}^M \left\{ \frac{S_k}{T_k} - \ln \frac{S_k}{T_k} - 1 \right\} \quad (6.15.1)$$

El doble del miembro derecho de (6.15.1) es la versión discreta de la distorsión de Itakura-Saito en (2.7.2) [1].

La ecuación (6.14.3) corresponden entonces a (3.13.1).

A.1 INVERSIÓN DE MATRICES DE RESPUESTA IMPULSO :a) Sistemas causales :

Un sistema discreto lineal e invariante en el tiempo (L.I.T.) causal, puede verse como una transformación lineal, representada por una matriz de dimensión infinita A , de una secuencia $\{u_i\}_{i=1}^n$ a otra secuencia $\{v_i\}_{i=1}^n$; simbólicamente

$$\{v_i\}_{i=1}^n = A\{u_i\}_{i=1}^n \quad (1)$$

Cada v_i esta dada por una combinación lineal de los elementos de $\{u_n\}_{n=1}^i$, de acuerdo a la sumatoria de convolución

$$v_i = \sum_{k=0}^{i-1} h_k u_{i-k} \quad (2)$$

donde $\{h_k\}_{k=0}^{\infty}$ es la respuesta impulso del sistema. Usando este hecho, la ecuación (1) puede desplegarse como sigue :

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} h_0 & 0 & 0 & 0 & \dots & \dots \\ h_1 & h_0 & 0 & 0 & \dots & \dots \\ h_2 & h_1 & h_0 & 0 & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \quad (3)$$

La inversa de A puede obtenerse a través de la respuesta impulso del sistema inverso: Sea $H(z)$ la transformada Z de $\{h_k\}_{k=0}^{\infty}$, ésto es, la función de transferencia; sea $G(z) = 1/H(z)$. En general, $\{g_k\}_{k=-m}^{\infty}$, la transformada Z inversa de $G(z)$, no es una secuencia causal; u_i puede escribirse como

$$u_i = \sum_{k=-m}^{i-1} g_k v_i - k$$

donde n puede ser infinito; en consecuencia, la transformación inversa se despliega (ya que las primeras n v_i son cero) :

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} g_0 & g_{-1} & g_{-2} & \dots & g_{-m} & 0 & 0 & 0 & \dots & v_1 \\ g_1 & g_0 & g_{-1} & & g_{-m+1} & g_{-m} & 0 & 0 & \dots & v_2 \\ g_2 & g_1 & g_0 & & g_{-m+2} & g_{-m+1} & 0 & 0 & \dots & v_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

o, abreviando

$$\{u_i\}_{i=1} = B\{v_i\}_{i=1}^{n+m} \quad (4)$$

Comparando (1) y (4), puede verse que si el dominio de B coincide con la imagen de A , la imagen de B no es el dominio de A . En otras palabras, para que B sea la transformación inversa de A , es necesario que $m=0$. Esto es: para que la transformación A que representa a un

sistema LIT discreto causal sea invertible, es necesario que el sistema inverso sea causal.

b) Sistemas no causales :

Correspondiente a la sumatoria de convolución

$$v_i = \sum_{k=-N}^{i-1} h_k u_{i-k} \quad (5)$$

la cual relaciona la entrada y la salida de un sistema discreto LIT no causal, con N pasos de anticipación (donde N puede ser infinito), se define la transformación lineal A tal que

$$\{v_i\}_{i=1-N}^n = A \{u_i\}_{i=1}^N \quad (6)$$

o, desplegando

$$\begin{bmatrix} v_{1-N} \\ v_{2-N} \\ \vdots \\ v_0 \\ v_1 \\ \vdots \\ v_n \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} h_{-N} & 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ h_{-N+1} & h_{-N} & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ h_{-1} & h_{-2} & \dots & h_{-N} & 0 & 0 & \dots & \dots & \dots & \dots \\ h_0 & h_{-1} & \dots & \dots & h_{-N} & 0 & \dots & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{n-1} & h_{n-2} & \dots & \dots & \dots & h_{-N} & 0 & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \\ u_{N+1} \\ \vdots \\ u_{N+n} \\ \vdots \\ \vdots \end{bmatrix} \quad (7)$$

Se busca ahora la función de transferencia $G^1_z = 1/H(z)$; como

$$H(z) = \sum_{k=-N}^{\infty} h_k z^{-k} = z^N \sum_{k=-N}^{\infty} h_k z^{-(k+N)} = z^N \sum_{\ell=0}^{\infty} h_{\ell-N} z^{-\ell}$$

$$H(z) = z^N H_1(z)$$

donde $H_1(z)$ es causal; supóngase que $H_1(z)$ tiene una inversa, que es causal, $G(z) = 1/H_1(z)$. La transformación lineal B^1 asociada al sistema inverso $G^1(z) = 1/H(z) = z^{-N} G(z)$ puede desplegarse (observando que z^{-N} es un atraso de N pasos) :

$$\begin{bmatrix} u_1^1 \\ u_2^1 \\ \vdots \\ u_N^1 \\ u_{N+1}^1 \\ u_{N+2}^1 \\ \vdots \\ u_{N+n}^1 \\ \vdots \\ u_{2N+n}^1 \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ g_0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ g_1 & g_0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{n-1} & g_{n-2} & \dots & g_0 & \dots & \dots & \dots & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{n-1+N} & g_{n-2+N} & \dots & \dots & \dots & \dots & g_0 & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} v_{1-N} \\ v_{2-N} \\ \vdots \\ \vdots \\ v_0 \\ v_1 \\ v_2 \\ \vdots \\ \vdots \\ v_n \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \quad (8)$$

o, abreviando

$$\{u_i^1\}_{i=1}^{n+2N} = B^1 \{v_i\}_{i=1-N}^n$$

donde $u_1^i = u_2^i = \dots = u_N^i = 0$ y $u_{i-N}^i = u_i$ para $i > N$.

Por tanto, se puede escribir

$$\{u_i\}_{i=1}^{n+N} = B\{v_i\}_{i=1-N}^n \quad (9)$$

donde

$$B = \begin{bmatrix} g_0 & 0 & 0 & \dots & \dots \\ g_1 & g_0 & 0 & \dots & \dots \\ g_2 & g_1 & g_0 & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

y $\{g_k\}_{k=0}^{\infty}$ es la transformada z_1 inversa de $G(z) = 1/H_1(z)$.

Comparando (6) y (9), se concluye que B es la inversa de A , siempre que $G(z) = 1/H_1(z)$ sea causal. Aunque la matriz inversa existe, no está directamente asociada al sistema inverso $G^1(z)$.

A.2 ESTABILIDAD

En las definiciones de RCEC, RCED, RCIM, y RCEDC, se considera uno o dos procesos estocásticos estacionarios que comienzan en un tiempo finito arbitrario y se continúan hasta el infinito. Con este tipo de procesos como entradas a sistemas discretos LIT estables, los procesos a la salida son tan sólo asintó

ticamente estacionarios [5, pág. 238]. Este problema ha sido tratado en [6]. En éste trabajo no se harán distinciones entre procesos "asintóticamente estacionarios" y "estacionarios". Se ha visto que para hablar de una transformación lineal uno a uno es necesario que si $H(z)$ es causal, su inverso $G(z) = 1/H(z)$ sea causal. Con el fin de garantizar la convergencia de la transformada de Fourier de la función de transferencia $H(e^{j\omega})$, de un sistema causal, se requerirá que los sistemas causales sean estables. Además, con el fin de que los procesos de salida tengan variancia finita y sean asintóticamente estacionarios, se pedirá que los sistemas, causales o no, sean estables. Para este último objetivo no se requerirá, sin embargo, que el sistema inverso $G(z) = 1/H(z)$ sea estable, si éste se conecta en cascada con el sistema $H(z)$, ya que, de acuerdo al siguiente esquema

si $u[n]$ a la entrada es estacionario, $u[n]$ a la salida es estacionario, y la transformación lineal inversa asociada a $G(z)$ es convergente para las entradas $\{v_i\}_{i=1}^{\infty}$ asintóticamente estacionarias a la salida de $H(z)$; en otras palabras, la imagen de la transformación inversa B asociada a $G(z)$ es estacionaria, y en particular, de variancia finita, si su dominio se restringe al rango de la transformación lineal A (aquí se presupuso que $H(z)$ y $G(z)$ son causales).

A.3 EJEMPLOS DE INVERSION DE MATRICES DE RESPUESTA IMPULSO PARA SISTEMAS CAU-

SALES :

a) Ejemplo 1.: El sistema inverso es estable pero no causal :

Sea $H(z) = z^{-1}$, $|z| > 0$, el cual es causal y estable; entonces, la respuesta impulso es

$$h_k = \begin{cases} 1, & k = 1 \\ 0, & k \neq 1 \end{cases}$$

El sistema inverso está dado por $G(z) = z$, $0 < |z| < \infty$. Obsérvese que la región de convergencia de $G(z)$ está incluida en la región de convergencia de $H(z)$; la respuesta impulso del sistema estable pero no causal $G(z)$ es

$$g_k = \begin{cases} 1, & k = -1 \\ 0, & k \neq -1 \end{cases}$$

A partir de la respuesta impulso de cada sistema se escriben las matrices de respuesta impulso :

A partir de h_k (usando (3)) :

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \cdot \\ \cdot \end{bmatrix} = A \{u_i\}_{i=1}^n$$

A partir de g_k (usando la forma desplegada de (4)) :

$$\begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ \vdots \end{bmatrix} = B \{v_i\}_{i=1}^n$$

Multiplicando A por B :

$$AB = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \neq I \quad \therefore B \text{ no es la inversa de } A$$

b) Ejemplo 2: El sistema inverso es causal pero no estable :

Sea $H(z) = \frac{z+1}{z} = 1 + z^{-1}$, $|z| \geq 1$, el cual es causal y estable; su respuesta impulso es

$$h_k = \begin{cases} 1, & k = 0, 1 \\ 0, & k = 2, 3, 4, \dots \end{cases}$$

El sistema inverso está dado por división larga :

$$z+1 \left| \begin{array}{r} 1-z^{-1}+z^{-2}-z^{-3} \dots\dots \\ \hline z+1 \\ \hline -1 \\ \hline -1-z^{-1} \\ \hline +z^{-1} \\ \hline z^{-1}+z^{-2} \\ \hline -z^{-2} \end{array} \right. \quad \text{ó} \quad 1+z \left| \begin{array}{r} z-z^2+z^3 \dots\dots \\ \hline z+z^2 \\ \hline -z^2 \\ \hline -z^2-z^3 \\ \hline +z^3 \end{array} \right.$$

y se escoge $G(z) = 1 - z^{-1} + z^{-2} - z^{-3} \dots$, ya que su región de convergencia ($|z| > 1$) está incluida en la región de convergencia de $H(z)$. La respuesta impulso del sistema inverso, el cual es causal pero no estable, está dada por :

$$g_k = \begin{cases} (-1)^k, & k = 0, 1, 2, \dots \\ 0, & k < 0 \end{cases}$$

Las respectivas matrices de respuesta impulso son

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots \\ 1 & 1 & 0 & 0 & \dots \\ 0 & 1 & 1 & 0 & \dots \\ 0 & 0 & 1 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} ; \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots \\ -1 & 1 & 0 & 0 & \dots \\ 1 & -1 & 1 & 0 & \dots \\ -1 & 1 & -1 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Efectuando el producto de estas matrices

$$AB = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \end{bmatrix} = I$$

de donde se concluye que B es la inversa de A

A.4 INVERTIBILIDAD DE TRANSFORMACIONES LINEALES ASOCIADAS A SISTEMAS CAUSALES EN LA TOMA DE UN LIMITE :

En las definiciones de RCEC, RCED, RCIM y RCEDC, aparece la toma de un límite cuando N tiende a infinito. Por esto se entiende que cada una de estas cantidades es tan cercana como se desee a su valor límite cuando N es suficientemente grande, pero aún finita. Aún más, las f.d.p. conjunta $P_u^N(x)$ de un proceso estocástico están definidas para todo N finito. En consecuencia, cuando se habla de una transformación invertible de los procesos involucrados en dichas definiciones, el punto de interés es que la transformación sea invertible para todo N finito.

En la sección A.1 se vió que para N infinita un sistema causal tiene asociada una transformación lineal invertible sólo si el sistema inverso es causal. Sin embargo, ésta no es la situación cuando N es finita, en cuyo caso la ecuación (3) se simplifica a

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_N \end{bmatrix} = \begin{bmatrix} h_0 & 0 & 0 & 0 & \dots & 0 \\ h_1 & h_0 & 0 & 0 & \dots & 0 \\ h_2 & h_1 & h_0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{N-1} & h_{N-2} & h_{N-3} & \vdots & \vdots & h_0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_N \end{bmatrix} \quad (10)$$

o, abreviadamente

$$u = A_N u$$

El jacobiano de la transformación es el determinante de A_N , $J = h_0^N$, y la transformación lineal de orden finito es invertible (uno a uno) si y sólo si $h_0 \neq 0$.

En las siguientes secciones de este apéndice se supone que los sistemas son causales y estables y que $h_0 \neq 0$. La estacionariedad asintótica de los procesos de salida es necesaria para la existencia del límite.

A.5 INVARIANCIA DE LA RCEC BAJO LA TRANSFORMACION POR SISTEMAS DISCRETOS LIT CAUSALES Y ESTABLES :

Se demuestra la invariancia de la RCEC entre dos procesos discretos estacionarios cuando ambos son transformados por el mismo sistema discreto LIT causal y estable.

Sean $u, v, \tilde{u}, \tilde{v}$ vectores aleatorios en \mathbb{R}^N

$$\tilde{u} = A_N u ; \tilde{v} = A_N v , A_N \text{ dada en (10)}$$

Se sabe que

$$P_u^N(x) = \frac{P_u^N(y)}{|h_0|^N} , P_v^N(x) = \frac{P_v^N(y)}{|h_0|^N} , \text{ donde } x = A_N y ; x, y \in \mathbb{R}^N$$

$$\text{y que } \delta_x^N = |h_0|^N \delta_y^N .$$

Utilizando éstas relaciones en la ecuación (11.1), se obtiene

$$\bar{H}(\tilde{U}, \tilde{V}) = \bar{H}(U, V)$$

Nótese que si el sistema $H(z)$ es de fase mínima, el papel de los procesos puede intercambiarse, en el sentido de aplicar \tilde{U} y \tilde{V} como entradas al sistema inverso.

A.6 RELACION DE RCED ENTRE LA ENTRADA Y LA SALIDA DE UN SISTEMA DISCRETO LIT CAUSAL Y ESTABLE :

Sean u, v vectores aleatorios en \mathbb{R}^N

$$v = A_N u , A_N \text{ dada en (10)}$$

$$\text{y } P_v^N(x) = \frac{P_u^N(y)}{|h_0|^N} , \text{ donde } x = A_N y ; x, y \in \mathbb{R}^N$$

$$\text{Adem\u00e1s, } \delta_x^N = |h_0|^N \delta_y^N ;$$

sustituyendo estas relaciones en (13.1), se obtiene (13.2) [12] .

A.7 INVARIANCIA DE LA RCIM BAJO LA TRANSFORMACION POR SISTEMAS DISCRETOS LIT CAUSALES Y ESTABLES :

Se demuestra la propiedad 3 de la RCIM .

Sean $u, v, \tilde{u}, \tilde{v}$, vectores aleatorios en R^N

$$\tilde{u} = A_N^{(1)} u, \quad \tilde{v} = A_N^{(2)} v; \quad A_N^{(1)} \quad \text{y} \quad A_N^{(2)} \quad \text{dadas en (10) .}$$

Se sabe que

$$P_{\tilde{u}}^N(\tilde{x}) = \frac{P_u^N(x)}{h_0^{(1)N}}, \quad P_{\tilde{v}}^N(\tilde{y}) = \frac{P_v^N(y)}{h_0^{(2)N}} \quad (11)$$

donde $\tilde{x} = A_N^{(1)} x, \quad \tilde{y} = A_N^{(2)} y ; \quad \tilde{x}, x, \tilde{y}, y \in R^N .$

$$\text{Tambi\u00e9n, } \delta_{\tilde{x}}^N = |h_0^{(1)}|^N \delta_x^N, \quad \delta_{\tilde{y}}^N = |h_0^{(2)}|^N \delta_y^N \quad (12)$$

Por otro lado, la concatenación de los vectores \tilde{u} y \tilde{v} está relacionada con la concatenación de u y v por

$$\begin{bmatrix} \tilde{u} \\ \tilde{v} \end{bmatrix} = \begin{bmatrix} A_N^{(1)} & 0_N \\ 0_N & A_N^{(2)} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} ;$$

el jacobiano de esta transformación es $(h_0^{(1)} h_0^{(2)})^N$, por lo cual

$$P_{\tilde{u}\tilde{v}}^{2N}(\tilde{x}, \tilde{y}) = \frac{P_{uv}^{2N}(x, y)}{|h_0^{(1)} h_0^{(2)}|^N} \quad (13)$$

Sustituyendo (11), (12) y (13) en (14.1) y tomando el límite cuando N tiende a infinito, se obtiene $\bar{I}(\tilde{U}, \tilde{V}) = \bar{I}(U, V)$

NOTA : Las demostraciones en A.5, A.6 y A.7 pueden generalizarse a cualquier transformación que sea invertible para todo N finito.

B.1 COTA SUPERIOR DE LA RCED DEL PROCESO DE INNOVACIONES :

Sea $E = \{e_n\}_{n=-\infty}^{\infty}$ un proceso estocástico discreto estacionario de densidad espectral de potencia constante σ^2 , ésto es, blanco en sentido amplio con potencia σ^2 .

De [5, pág. 532],

$$\bar{H}_e = \lim_{m \rightarrow \infty} H(e_n | (e_{n-1}, e_{n-2}, \dots, e_{n-m})) \leq H(e_n) \stackrel{\Delta}{=} H(P_e) \quad (\text{B.1.1})$$

donde $P_e(x)$ es la función de densidad de probabilidad marginal de primer orden del proceso E . Pero la entropía diferencial está acotada superiormente por la entropía diferencial de una variable aleatoria gaussiana de la misma variancia, por lo cual

$$\bar{H}_e \leq \ln \sqrt{2\pi e \sigma^2} = \bar{H}_{\sigma W} \quad (\text{B.1.2})$$

Cuando el proceso E es gaussiano, (B.1.1) y (B.1.2) se cumplen con igualdad, por lo qué, el proceso blanco (en sentido amplio) con máxima RCED es el proceso blanco gaussiano, para una potencia dada σ^2 .

B.2 UNA COTA INFERIOR (O SUPERIOR) PARA $H(\tilde{s}, \sigma^2 W)$ SEGUN $\bar{H}_{e_{\tilde{s}}}$ SEA NEGATIVA O POSITIVA :

De la definición de la RCED :

$$\bar{H}(\tilde{s}, \sigma^2 W) = -\bar{H}_{\tilde{s}} + \ln \sqrt{2\pi \sigma^2} + \frac{1}{2\sigma^2} \int_{-\pi}^{\pi} f_{\tilde{s}}(e^{j\omega}) \frac{d\omega}{2\pi} \quad (\text{B.1.3})$$

Sea $E_{\tilde{S}} = \{e_{s_n}\}_{n=-\infty}^{\infty}$ el proceso de innovaciones de \tilde{S} , con RCED $\bar{H}_{e_{\tilde{S}}}$. Entonces

$$H_{\tilde{S}} = H_{e_{\tilde{S}}} + \frac{1}{2} \int_{-\pi}^{\pi} \ln f_{\tilde{S}}(e^{j\omega}) \frac{d\omega}{2\pi} \quad (\text{B.2.1})$$

y con

$$\bar{H}_W = \frac{1}{2} \ln(2\pi e) \quad (\text{B.2.2})$$

se tiene de (B.1.3), (B.2.1) y (B.2.3),

$$\bar{H}(\tilde{S}, \sigma^2 W) = \bar{H}_W - \bar{H}_{e_{\tilde{S}}} + \frac{1}{2} d_{IS}(f_{\tilde{S}}(e^{j\omega}), \sigma^2) \quad (\text{B.2.3})$$

Por apéndice B.1, $\bar{H}_{e_{\tilde{S}}} < \bar{H}_W$. Se concluye que, si $\bar{H}_{e_{\tilde{S}}} \geq 0$

$$\bar{H}(\tilde{S}, \sigma^2 W) < \bar{H}_W + \frac{1}{2} d_{IS}(f_{\tilde{S}}(e^{j\omega}), \sigma^2)$$

e independientemente del signo de $\bar{H}_{e_{\tilde{S}}}$

$$H(\tilde{S}, \sigma^2 W) \geq \frac{1}{2} d_{IS}(f_{\tilde{S}}(e^{j\omega}), \sigma^2)$$

En realidad, de (B.1.3) es fácil ver que

$$\bar{H}(E_{\tilde{S}}, W) = -\bar{H}_{e_{\tilde{S}}} + \bar{H}_W,$$

ésto es, (B.2.3) corresponde a (3.) con un filtro ARMA $G(z) = \sigma$.

Sea $P_{e_{\tilde{S}}}(x)$ la función de densidad de probabilidad marginal de primer orden

del proceso de innovaciones de \tilde{S} ; entonces

$$\bar{H}(E_{\tilde{S}}, W) = [\bar{H}_W - H(P_{e_{\tilde{S}}})] + [H(P_{e_{\tilde{S}}}) - \bar{H}_{e_{\tilde{S}}}]$$

$$\bar{H}(E_{\tilde{S}}, W) = [\bar{H}_W - H(P_{e_{\tilde{S}}})] + I(e_{\tilde{S}_n} | (e_{\tilde{S}_{n-1}}, e_{\tilde{S}_{n-2}}, \dots))$$

El primer término entre paréntesis del miembro derecho de la ecuación anterior es una "medida" de la gaussianidad de las observaciones, y el segundo término es una "medida" de la independencia de las innovaciones; (B.2.3) puede reescribirse ahora

$$\bar{H}(\tilde{S}, \sigma W) = \bar{H}_W - H(P_{e_{\tilde{S}}}) + I(e_{\tilde{S}_n} | (e_{\tilde{S}_{n-1}}, e_{\tilde{S}_{n-2}}, \dots)) + \frac{1}{2} d_{IS}(f_{\tilde{S}}(e^{j\omega})\sigma^2) \quad (B.3.1)$$

En la proposición 6 la constante σ es arbitraria. Escogiendo $\sigma^2 =$

$$= \int_{-\pi}^{\pi} f_{\tilde{S}}(e^{j\omega}) \frac{d\omega}{2\pi},$$

$$H(S, W) = H_W - H(P_{e_S}) + I(e_{S_n} | (e_{S_{n-1}}, e_{S_{n-2}}, \dots)) - \frac{1}{2} \ln \gamma_S^2 \quad (B.3.2)$$

B.4 LA PROPOSICIÓN 6 CON G(z) IGUAL AL FILTRO DE INNOVACIONES DE U :

En éste caso el proceso de innovaciones de U coincide con su proceso residual y $f_U(e^{j\omega}) = f_V(e^{j\omega})$, por lo qué, de (3.) y del apéndice (B.2.):

$$\bar{H}(U, V) = H(E, W) = \bar{H}_W - H(P_e) + I(e_n; (e_{n-1}, e_{n-2}, \dots)) \quad (B.3.3)$$

B.5 LA PROPOSICION 5 CON $G(z)$ IGUAL AL FILTRO DE INNOVACIONES DE U :

En éste caso $\sigma^2 = \sigma_U^2 = \alpha$ y de (3.18.), con $R = E$,

$$\bar{H}(U, V^*) = I(e_n; (e_{n-1}, e_{n-2}, \dots)) \quad (\text{B.4.1})$$

B.6 ERROR POR EL USO DE UN PROCESO SINTETICO GAUSSIANO :

De (B.3.3) y (B.4.1), donde $G(z)$ es el filtro de innovaciones de U :

$$\bar{H}(U, V) - \bar{H}(U, V^*) = \bar{H}_W - H(P_e) \quad (\text{B.4.2})$$

por lo cual, al excitar al filtro de innovaciones con un proceso blanco gaussiano de potencia unitaria en lugar de excitarlo con el proceso residual, la distorsión se incrementa en función exclusiva de que tan gaussiana es la función de densidad de probabilidad marginal de primer orden del proceso de innovaciones $P_e(x)$.

B.7 RCED PARA PROCESOS GAUSSIANS CUANDO LA ENTRADA SE APLICA EN $n = 0$ Y CUANDO SE APLICA EN $n = -\infty$:

Sea U un proceso estocástico estacionario regular y gaussiano (la generalización de lo que aquí se presente a procesos no gaussianos es inmediata); sea la densidad espectral de potencia de U , $f_U(e^{j\omega}) = L_U(e^{j\omega})L_U(e^{-j\omega})$, donde

$L_u(e^{j\omega})$ es de fase mínima y representa el filtro de innovaciones de U .

Sea W un proceso de ruido blanco gaussiano con potencia unitaria. Sea $H_0(z)$ un filtro pasa todas (estable), ésto es :

$$H_0(z) = \frac{Q}{\pi} \prod_{i=1}^Q \frac{z z_i^* - 1}{z - z_i}, \quad |z_i| < 1.$$

$$|H_0(e^{j\omega})| = 1, \quad \forall \omega.$$

Sea $H(z) = H_0(z) L_u(z)$. Entonces :

- a) si el proceso de innovaciones W se aplica a la entrada del filtro $L_u(z)$ desde $n = -\infty$, y simultáneamente

al filtro $H(z)$ (refiérase a la figura adjunta); entonces [5, pág.533]:

$$\bar{H}_u = \bar{H}_W + \frac{1}{2} \int_{-\pi}^{\pi} \ln(|L_u(e^{j\omega})|^2) \frac{d\omega}{2\pi}$$

y como

$$|H(e^{j\omega})|^2 = |L_u(e^{j\omega})|^2,$$

se concluye

$$\bar{H}_V = \bar{H}_U \quad (\text{B.6.1})$$

- b) si el proceso de innovaciones W se aplica a la entrada del filtro $L_U(z)$ a partir de $n = 0$ (suponiendo condiciones iniciales nulas), y simultáneamente al filtro $H(z)$ (refiérase a la figura adjunta), entonces [(2.) y (2.)] :

las respuestas de los filtros son \hat{U} y \hat{V} , las cuales son procesos asintóticamente estacionarios y convergentes hacia U y V , respectivamente, y además, se cumplen :

$$H_{\hat{V}} = H_W + \varepsilon_n |h[0]| \quad (\text{B.6.2})$$

$$\bar{H}_{\hat{V}} = \bar{H}_W + \varepsilon_n |\varepsilon_U[0]| = H_W + \frac{1}{2} \int_{-\pi}^{\pi} \varepsilon_n (|L_U(e^{j\omega})|^2) \frac{d\omega}{2\pi} \quad (\text{B.6.3})$$

donde $h[0]$ y $\varepsilon_U[0]$ son la respuesta al pulso unitario en $n = 0$ de los filtros $H(z)$ y $L_U(z)$ respectivamente. Es inmediato que $\bar{H}_{\hat{V}} = \bar{H}_U$.

Sea $h_0[0]$ la respuesta al pulso unitario en $n = 0$ del filtro $H_0(z)$; cómo

$$\frac{z z_i - 1}{z - z_i} = z_i + (z_i^2 - 1) z^{-1} + \dots$$

entonces $h_0[0] = \prod_{i=1}^Q z_i, |z_i| < 1$

y

$$|h_0[0]| < 1$$

También

$$h[0] = \lambda_u[0] h_0[0] \quad (\text{B.7.1})$$

y se concluye, de (B.7.1), (B.6.3) y (B.6.2) :

$$\bar{H}_V^{\wedge} = \bar{H}_U + \lambda_n |h_0[0]|, |h_0[0]| < 1 \quad (\text{B.7.2})$$

$$\bar{H}_U^{\wedge} = \bar{H}_U \quad (\text{B.7.3})$$

- c) conclusiones: i) cuando la entrada se aplica desde $n = 0$, la RCED de la respuesta depende de la característica de fase del filtro
- ii) cuando la entrada se aplica desde $n = -\infty$, la RCED de la respuesta de un filtro no depende de la característica de fase del filtro y siempre puede usarse para efectos del cálculo de la RCED el filtro de fase mínima que tiene la misma característica de amplitud que el filtro original.

- iii) Si el filtro es $L_u(e^{j\omega}) \equiv 1, \forall \omega$, la RCED de la respuesta es igual a la RCED de la entrada, si ésta se aplica desde $n = 0$. Pero si se dispone en cascada con $L_u(e^{j\omega})$ un filtro pasatodas, la RCED de la respuesta es menor que la de la entrada. (B.7.2).
- iv) La información adicional sobre la característica de fase del filtro contenida en la respuesta, se encuentra en el transitorio de la respuesta. La RCED de la respuesta de un filtro cuando la entrada se aplica en $n = -\infty$, es igual a la RCED del filtro de fase mínima con la misma característica de amplitud cuando la entrada se aplica desde $n = 0$. Esto puede entenderse como que el transitorio de la respuesta de un filtro de fase mínima no contiene información adicional sobre la característica de fase del filtro a aquella contenida en el estado estacionario de la respuesta. Recíprocamente, el estado estacionario de la respuesta sólo proporciona información sobre la característica del filtro de fase mínima que la produce. Esto explica porqué, basándose tan sólo en el estado permanente (entrada aplicada desde $n = -\infty$) sólo puede diseñarse un predictor dentro de la clase de predictores de fase mínima.

C.1. TRANSFORMADAS SENO Y COSENO DE SECUENCIAS CAUSALES :

Sea $f[n]$ una secuencia de números reales tal que $\sum_{n=-\infty}^{\infty} |f[n]| < \infty$. Entonces está definida la transformada de Fourier de la secuencia por la relación

$$F(\omega) = \sum_{n=-\infty}^{\infty} f[n] e^{-j\omega n}$$

la cual tiene, entre otras, las siguientes propiedades :

a) Si $f[n]$ es par, $F(\omega) = f[0] + 2 \sum_{k=1}^{\infty} f[k] \cos \omega k$.

b) Si $f[n]$ es impar, $F(\omega) = -2j \sum_{k=1}^{\infty} f[k] \sin \omega k$.

c) En general, si $f[n]$ es real, $f[-n] = F(-\omega) = F^*(\omega)$

Sea $F(\omega) = F_R(\omega) + F_I(\omega)$, donde $F_R(\omega)$ es real y $F_I(\omega)$ es puramente imaginaria; por ser $F(\omega)$ par-conjugada,

$$F_R(\omega) = F_R(-\omega)$$

$$F_I(\omega) = -F_I(-\omega)$$

d) Sea $f[n]$ real, $f[n] = f_e[n] + f_o[n]$, donde $f_e[n]$ es la parte par y $f_o[n]$ la parte impar de $f[n]$. Sea $f[n]$ causal, esto es, $f[n] = 0, \forall n < 0$. Entonces

$$f[n] = \begin{cases} f_e[0], & n = 0 \\ 2f_e[n], & n \geq 1 \\ 2f_o[n], & n \geq 1. \end{cases}$$

y además

$$f_e[n] \xleftrightarrow{F} F_R(\omega)$$

$$f_o[n] \xleftrightarrow{F} F_I(\omega)$$

c) de a), $F_R(\omega) = f[0] + 2 \sum_{k=1}^{\infty} f[k] \cos \omega k$

de donde :

$$f[0] = \frac{1}{2\pi} \int_{-\pi}^{\pi} F_R(\omega) d\omega$$

$$f[n] = 2f_e[n] = \frac{1}{\pi} \int_{-\pi}^{\pi} F_R(\omega) \cos \omega n d\omega, \quad n \geq 1 \quad 1$$

de b), $F_I(\omega) = -2j \sum_{k=1}^{\infty} f[k] \sin \omega k$.

de donde

$$f[n] = 2f_o[n] = \frac{1}{\pi} \int_{-\pi}^{\pi} jF_I(\omega) \sin \omega n d\omega, \quad n \geq 1 \quad 2$$

igualando 1 y 2 :

$$\frac{1}{\pi} \int_{-\pi}^{\pi} F_R(\omega) \cos \omega n d\omega = \frac{1}{\pi} \int_{-\pi}^{\pi} j F_I(\omega) \sin \omega n d\omega, \quad n \geq 1$$

Esta ecuación se cumple si $f[n]$ es real, causal y estable. ($F_I(\omega) = j2n(F(\omega))$).

C.2 PARTES CAUSALES Y ANTICAUSALES :

Sea $f_{se_x}(z)$ una función compleja con dominio en el plano complejo z , cuya región de convergencia incluye el círculo unitario de tal forma que su transformada de Fourier está definida. Se entiende por parte anticausal de $f_{se_x}(z)$ la serie formada por los términos del desarrollo de $f_{se_x}(z)$ en serie formada por los términos del desarrollo de $f_{se_x}(z)$ en serie de Laurent en torno al origen, tales que su región de convergencia es el interior de un círculo (exterior al círculo unitario para garantizar la existencia de la transformada de Fourier); es decir, todos los polos de la parte anticausal están fuera de dicho círculo. A su vez, la parte causal corresponde a los términos que convergen fuera de un círculo interior al círculo unitario.

Dada la serie de Laurent, $f_{se_x}(z) = \sum_{m=-\infty}^{\infty} R_{se_x}[m] z^{-m}$, es claro que los términos con potencia de z negativos corresponden a la parte causal y aquéllos con potencias positivas a la parte anticausal. Sin embargo, el término con potencia de z nula, puede ser separado en $R_{se_x}[0] = k + R_{se_x}[0] - k$, donde k es una constante arbitraria, y así la parte causal y anticausal quedan definidas hasta la constante k y $R_{se_x}[0] - k$ respectivamente.

El mismo problema ocurre si se toma el desarrollo en fracciones parciales como base de la definición de partes causales y anticausales (Véase la nota en la pág. 449 de Papoulis).

En este trabajo se usa la siguiente notación :

a) $f_{se_x}^+(z)$ es la parte del desarrollo en serie de $f_{se_x}(z)$ que contiene po-

tencias de z positivas o cero:

$$f_{se_x}^+(z) = \sum_{m=0}^{\infty} R_{se_x}[m]z^{-m}$$

y por tanto

$$f_{se_x}^+(1/z) = \sum_{m=0}^{\infty} R_{se_x}[m]z^m$$

- b) $f_{se_x}^-(z)$ es la parte del desarrollo en serie de $f_{se_x}(z)$ que contiene potencias de z negativas :

$$f_{se_x}^-(z) = \sum_{m=-\infty}^{-1} R_{se_x}[m]z^{-m}$$

y por tanto

$$f_{se_x}^-(1/z) = \sum_{m=-\infty}^{-1} R_{se_x}[m]z^m = \sum_{m=1}^{\infty} R_{se_x}[-m]z^{-m}$$

- c) Debido a que este "truncamiento" no conmuta con el cambio de argumento de z a $1/z$, se usan paréntesis rectangulares para indicar que primero se efectúa el cambio del argumento y posteriormente el truncamiento :

$$f_{se_x}(1/z) = \sum_{m=-\infty}^{\infty} R[m]z^m = \sum_{m=-\infty}^{\infty} R[-m]z^{-m}$$

$$[f_{se_x}(1/z)]^+ = \sum_{m=0}^{\infty} R_{se_x}[-m]z^{-m}$$

$$[f_{se_x}(1/z)]^- = \sum_{m=-\infty}^{-1} R_{se_x}[-m]z^{-m} = \sum_{m=1}^{\infty} R_{se_x}[m]z^m$$

d) Se concluye que se satisfacen las siguientes relaciones :

$$[f_{se_x}(1/z)]^+ = f_{se_x}^-(1/z) + R_{se_x}[0]$$

$$[f_{se_x}(1/z)]^- = f_{se_x}^+(1/z) - R_{se_x}[0] .$$

D.1

$$I(\underline{X}_n; \underline{S}_n) = E \left\{ \ln \left[\frac{\frac{1}{(2\pi)^n \Delta_T^{1/2}} \exp\left\{-\frac{1}{2} \text{col}^T(\underline{X}_n, \underline{S}_n) \cdot C_T^{-1} \text{col}(\underline{X}_n, \underline{S}_n)\right\}}{\frac{1}{(2\pi)^{n/2} \Delta_x^{1/2}} \exp\left\{-\frac{1}{2} \underline{X}_n^T C_x^{-1} \underline{X}_n\right\} \cdot \frac{1}{(2\pi)^{n/2} \Delta_s^{1/2}} \exp\left\{-\frac{1}{2} \underline{\tilde{S}}_n^T C_s^{-1} \underline{\tilde{S}}_n\right\}} \right] \right\}$$

$$= \ln \frac{\Delta_x^{1/2} \Delta_s^{1/2}}{\Delta_T^{1/2}} + E \left\{ \frac{1}{2} (\underline{X}_n^T C_x^{-1} \underline{X}_n + \underline{\tilde{S}}_n^T C_s^{-1} \underline{\tilde{S}}_n - \text{col}^T(\underline{X}_n, \underline{\tilde{S}}_n) C_T^{-1} \text{col}(\underline{X}_n, \underline{\tilde{S}}_n)) \right\}$$

Observando que los términos del valor esperado son escalares y que

$t_r(AB) = t_r(BA)$, se tiene :

$$\begin{aligned} E\{t_r \text{col}^T(\underline{X}_n, \underline{\tilde{S}}_n) C_T^{-1} \text{col}(\underline{X}_n, \underline{\tilde{S}}_n)\} &= t_r E\{C_T^{-1} \text{col}(\underline{X}_n, \underline{\tilde{S}}_n) \text{col}^T(\underline{X}_n, \underline{\tilde{S}}_n)\} = \\ &= t_r C_T^{-1} C_T = 2n \end{aligned}$$

y en forma similar :

$$E\{\underline{X}_n^T C_x^{-1} \underline{X}_n\} = n ; \quad E\{\underline{S}_n^T C_s^{-1} \underline{\tilde{S}}_n\} = n$$

$$\therefore I(\underline{X}_n; \underline{\tilde{S}}_n) = \frac{1}{2} \ln \left[\frac{\Delta_x \Delta_s}{\Delta_T} \right]$$

D.2

$$C_{x\tilde{s}} = E\{\underline{X}_n \underline{S}_n^T\} = E\{\underline{X}_n (\underline{X}_n^T G_n^T - \underline{S}_n^T)\} = C_x G_n^T - C_{xs}$$

$$C_{\tilde{S}X} = E\{\tilde{S}_n X_n^T\} = E\{(G_n X_n - S_n) X_n^T\} = G_n C_X - C_{SX} = C_{XS}^T$$

$$C_{\tilde{S}} = E\{\tilde{S}_n \tilde{S}_n^T\} = E\{(G_n X_n - S_n) (X_n^T G_n^T - S_n^T)\}$$

$$C_{\tilde{S}} = G_n C_X G_n^T - G_n C_{XS} - C_{SX} G_n^T + C_S$$

$$C_T = \begin{bmatrix} C_X & C_{X\tilde{S}} \\ C_{\tilde{S}X} & C_{\tilde{S}} \end{bmatrix}$$

Ahora bien, recordando la fórmula

$$\det \begin{pmatrix} A & D \\ C & B \end{pmatrix} = \det(A) \det(B - CA^{-1}D)$$

resulta

$$\Delta_T = \det \begin{pmatrix} C_X & C_{X\tilde{S}} \\ C_{\tilde{S}X} & C_{\tilde{S}} \end{pmatrix} = \det(C_X) \det(C_S - C_{\tilde{S}X} C_X^{-1} C_{X\tilde{S}}) =$$

$$= \det(C_X) \det(G_n C_X G_n^T - G_n C_{XS} - C_{SX} G_n^T + C_S - (G_n C_X - C_{SX}) C_X^{-1} (C_X G_n^T - C_{XS}))$$

$$= \det(C_X) \det(C_S - C_{SX} C_X^{-1} C_{XS}) = \det \begin{pmatrix} C_X & C_{XS} \\ C_{SX} & C_S \end{pmatrix}$$

D.3-A. RELACION ENTRE RCEC Y DISTORSION DE ITAKURA-SAITO. 1a. DEMOSTRACION:

Sean $U = \{u_k\}_{k=0}^{\infty}$ y $V = \{v_k\}_{k=0}^{\infty}$ procesos discretos estacionarios gaussianos

nos de media cero. Las funciones de densidad de probabilidad de orden N correspondientes son:

$$P_U^N(\underline{x}) = \frac{1}{\sqrt{(2\pi)^N \Delta_U}} \exp \left\{ -\frac{1}{2} (\underline{x}^T C_U^{-1} \underline{x}) \right\}$$

$$P_V^N(\underline{x}) = \frac{1}{\sqrt{(2\pi)^N \Delta_V}} \exp \left\{ -\frac{1}{2} (\underline{x}^T C_V^{-1} \underline{x}) \right\}$$

y como la razón de cambio de entropía cruzada está dada por

$$\bar{H}(U;V) = \lim_{N \rightarrow \infty} \frac{1}{N} \int \dots \int P_U^N(\underline{x}) \ln \frac{P_U^N(\underline{x})}{P_V^N(\underline{x})} \delta_{\underline{x}}^N$$

donde $\underline{x} = (x_0, x_1, \dots, x_{N-1}) \in \mathbb{R}^N$ y $\delta_{\underline{x}}^N = dx_0 dx_1 \dots dx_{N-1}$, se tiene que :

$$\begin{aligned} \bar{H}(U;V) &= \lim_{N \rightarrow \infty} \frac{1}{N} \int \dots \int \frac{1}{\sqrt{(2\pi)^N \Delta_U}} \exp \left\{ -\frac{1}{2} \underline{x}^T C_U^{-1} \underline{x} \right\} \left\{ \ln \left[\frac{\Delta_V^{1/2}}{\Delta_U^{1/2}} - \frac{\underline{x}^T C_U^{-1} \underline{x}}{2} + \right. \right. \\ &\quad \left. \left. + \frac{\underline{x}^T C_V^{-1} \underline{x}}{2} \right\} = \lim_{N \rightarrow \infty} \left\{ \frac{1}{N} \left(\frac{1}{2} \ln \left[\frac{\Delta_V}{\Delta_U} \right] - \frac{1}{2} (t_r C_U^{-1} C_U) + \frac{1}{2} E_U \{ \underline{x}^T C_V^{-1} \underline{x} \} \right) \right\} \\ &= \lim_{N \rightarrow \infty} \frac{E_U \{ \underline{x}^T C_V^{-1} \underline{x} \}}{2N} + \frac{1}{2} \left[\int_{-\pi}^{\pi} \ln \frac{f_V(e^{j\omega})}{f_U(e^{j\omega})} \frac{d\omega}{2\pi} \right] - \frac{1}{2} \end{aligned}$$

donde se ha utilizado la relación

$$\lim_{N \rightarrow \infty} \frac{\ln \Delta_{\underline{x}}}{N} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln f_{\underline{x}}(e^{j\omega}) d\omega, \quad [5, \text{pág 478}]$$

pero $\lim_{N \rightarrow \infty} \frac{1}{2} t_r C_v^{-1} C_u =$

D.3-B RELACION ENTRE RCEC Y DISTORSION DE ITAKURA-SAITO. 2a. DEMOSTRACION :

De la ecuación :

$$\bar{H}(\tilde{S}, \sigma^2 W) = -\bar{H}_{\tilde{S}} + \ln \sqrt{2\pi\sigma^2} + \frac{1}{2} \left\{ \lim_{N \rightarrow \infty} E\{S^2[N]\} \right\}$$

y como $\bar{H}_{\tilde{S}}$, siendo \tilde{S} gaussiano y estrictamente estacionario, está dado por [5, pág. 533] :

$$\bar{H}_{\tilde{S}} = \frac{1}{2} \ln 2\pi e + \frac{1}{2} \int_{-\pi}^{\pi} \ln f_{\tilde{S}}(e^{j\omega}) \frac{d\omega}{2\pi}$$

se sigue :

$$\bar{H}(\tilde{S}, \sigma^2 W) = - \left[\frac{1}{2} \ln 2\pi e + \frac{1}{2} \int_{-\pi}^{\pi} \ln f_{\tilde{S}}(e^{j\omega}) \frac{d\omega}{2\pi} \right] + \ln \sqrt{2\pi e} - \frac{1}{2} + \frac{1}{2} \ln \sigma^2 +$$

$$+ \frac{1}{2\sigma^2} \left\{ \lim_{N \rightarrow \infty} E\{S^2[N]\} \right\}$$

$$H(\tilde{S}, \sigma^2 W) = \frac{1}{2} \int_{-\pi}^{\pi} \left\{ \frac{f_{\tilde{S}}(e^{j\omega})}{\sigma^2} - \ln \frac{f_{\tilde{S}}(e^{j\omega})}{\sigma^2} - 1 \right\} \frac{d\omega}{2\pi} = \frac{1}{2} d_{IS}(f_{\tilde{S}}(e^{j\omega}), \sigma^2)$$

La diferencia entre ésta segunda demostración y la primera, es que aquí se considera un proceso de referencia blanco, lo que simplifica la deducción.

También es posible deducir este resultado a partir de (5.), por medio de la aplicación de (14.2). Nótese que en el caso de predicción, $\gamma_{XS}^2 = 1$ y $\bar{I}(X; S)$ queda indeterminado, lo que corresponde a Δ_T singular en el apéndice D.2 .

E.1 ESTABILIZACION DE LA DISTORSION DE ITAKURA-SAITO CON UN FILTRO NO-CAUSAL:

Repetiendo (5.),

$$f_{\xi}(e^{j\omega}) = G(e^{j\omega})G^*(e^{j\omega})f_x(e^{j\omega}) - g(e^{j\omega})f_{sx}^*(e^{j\omega}) - G^*(e^{j\omega})f_{sx}(e^{j\omega}) + f_s(e^{j\omega}) \quad (\text{E.1.1})$$

Sea $G(e^{j\omega}) \triangleq A(\omega) + jB(\omega)$ y $f_{sx}(e^{j\omega}) \triangleq R(\omega) + jI(\omega)$ donde $A(\omega)$, $B(\omega)$, $R(\omega)$, $I(\omega)$ son funciones reales, con lo que la densidad espectral del error puede escribirse:

$$f_{\xi}(e^{j\omega}) = [A^2(\omega) + B^2(\omega)]f_x(e^{j\omega}) - 2[A(\omega)R(\omega) + B(\omega)I(\omega)] + f_s(e^{j\omega}) \quad (\text{E.1.2})$$

Se usará el cálculo de variaciones para obtener la condición de extremo local, suponiendo que el filtro está dado en forma única por los espectros. Posteriormente habrá de verificarse si se trata de un mínimo y que la solución es única para los espectros dados .

Se tiene : $d_{IS}(f_{\xi}(e^{j\omega}) \sigma^2) =$

$$\int_{-\pi}^{\pi} \left\{ \frac{(A^2(\omega) + B^2(\omega))f_x(e^{j\omega}) - 2(A(\omega)R(\omega) + B(\omega)I(\omega)) + f_s(e^{j\omega}))}{\sigma^2} \right\} \\ - \ln \left[\frac{[A^2(\omega) + B^2(\omega)]f_x(e^{j\omega}) - 2(A(\omega)R(\omega) + B(\omega)I(\omega)) + f_s(e^{j\omega})}{\sigma^2} \right] - 1 \right\} \frac{d\omega}{2\pi}$$

(E.1.3)

Sea ahora $A(\omega) = A_0(\omega) + \epsilon \eta_1(\omega)$

$$B(\omega) = B_0(\omega) + \epsilon \eta_2(\omega)$$

donde ϵ es una variable escalar real; sustituyendo estas dos definiciones en la d_{IS} y derivando con respecto a ϵ , se obtiene la primera variación de la funcional d_{IS} :

$$\left. \frac{d d_{IS}}{d\epsilon} \right|_{\epsilon} = \int_{-\pi}^{\pi} \frac{\omega}{\pi} \left\{ (2A_0(\omega) f_X(e^{j\omega}) - 2R(\omega)) \eta_1(\omega) \left[\frac{1}{\sigma^2} - \frac{1}{f_S(e^{j\omega})} \right] + (2B_0(\omega) f_X(e^{j\omega}) - 2I(\omega)) \eta_2(\omega) \left[\frac{1}{\sigma^2} - \frac{1}{f_S(e^{j\omega})} \right] \right\} = 0$$

Al igualar a cero, se establece la condición necesaria de extremo local. Como $\eta_1(\omega)$ y $\eta_2(\omega)$ son arbitrarias, escogiendo sucesivamente $\eta_1(\omega) \equiv 0$ y $\eta_2(\omega) \equiv 0$, se obtiene el par de condiciones suficientes:

$$2A_0(\omega) f_X(e^{j\omega}) - 2R(\omega) = 0$$

$$2B_0(\omega) f_X(e^{j\omega}) - 2I(\omega) = 0$$

las cuales se pueden combinar en

$$[A_0(\omega) + jB_0(\omega)] f_X(e^{j\omega}) - [R(\omega) + jI(\omega)] = 0$$

o sea

$$G_0(e^{j\omega}) = \frac{f_{SX}(e^{j\omega})}{f_X(e^{j\omega})}$$

Este filtro es el filtro de Wiener no-causal y es un resultado determinado en forma única por el espectro cruzado $f_{sx}(e^{j\omega})$ y la densidad espectral del proceso perturbado $f_x(e^{j\omega})$. Se muestra en el texto que no es, en general, un mínimo.

F. MODELOS DE CANAL Y RUIDO QUE RESULTAN EN IGUALDAD ENTRE EL FUTURO DE WIENER NO-CAUSAL Y EL FILTRO DE WIENER CAUSAL .

F.1 ANTECEDENTES : MODELOS EQUIVALENTES :

Supóngase que son dados $f_s(z)$, $f_{sx}(z)$ y $f_x(z)$. Se puede entonces encontrar un modelo de la forma : [11, pág. 382]

$$x[n] = h[n] * s[n] + v[n] ;$$

esquemáticamente :

donde

$$H(z) = \sum_{k=-\infty}^{\infty} h[k] z^{-k} .$$

Se denominará como "el canal" al sistema LIT $H(z)$, y al proceso $v[n]$ el ruido. La relación entre $f_v(z)$, $f_{sv}(z)$ y $H(z)$ con los datos está dada por las ecuaciones deducidas a continuación :

de la figura : $y[n] = h[n] * s[n]$

$$x[n] = y[n] + v[n]$$

$$R_x[k] = E\{(y[n] + \gamma[n])(y[n-k] + \gamma[n-k])\} = R_y[k] + R_{\gamma y}[k] + R_{y\gamma}[k] + R_\gamma[k]$$

de donde

$$f_x(z) = f_y(z) + f_{\gamma y}(z) + f_{y\gamma}(z) + f_\gamma(z)$$

$$\Rightarrow f_x(z) = H(z)H(1/z)f_s(z) + H(1/z)f_{s\gamma}(z) + H(z)f_{\gamma s}(z) + f_\gamma(z)$$

$$f_x(z) = H(z)H(1/z)f_s(z) + H(1/z)f_{s\gamma}(1/z) + H(z)f_{\gamma s}(z) + f_\gamma(z) \quad (\text{F.2.1})$$

Por otro lado :

$$R_{sx}[k] = E\{(y[n] + \gamma[n]) S[n-k]\}$$

$$R_{sx}[k] = R_{ys}[k] + R_{\gamma s}[k]$$

de donde

$$f_{sx}(z) = H(1/z) f_s(z) + f_{s\gamma}(z) \quad (\text{F.2.2})$$

Estas dos relaciones (F.2.1 y F.2.2), dan lugar a una serie de situaciones diferentes; se tiene especial interés aquí en que $f_{sx}(z) = \frac{f_{sx}(z)}{L_x(1/z)}$ es una función causal, y en adelante se supone que tal es el caso.

F.2 RUIDO ORTOGONAL : $f_{s\gamma}(z) = 0$

de (F.2.2) :
$$H(z) = \frac{f_{sx}(1/z)}{f_s(z)} \quad (\text{F.3.1})$$

de (F.2.1) :
$$f_Y(z) = f_X(z) - \frac{f_{sx}(z)f_{sx}(1/z)}{f_s(z)} \quad (\text{F.3.2})$$

a) Canal no-causal : Se tiene, de (F.3.1) :

$$H(z) = \frac{f_{sex}(1/z)L_X(z)}{L_S(z)L_S(1/z)} = \left[\frac{L_X(z)}{L_S(z)} \right] \cdot \left[\frac{f_{sex}(1/z)}{L_S(1/z)} \right]$$

Como $f_{sex}(z)$ es causal, se concluye que el canal debe ser de la forma:

$$H(z) = \left[\frac{L_X(z)}{L_S(z)} \right] \phi(z) , \quad (\text{F.3.3})$$

donde $\phi(z)$ tiene, en su desarrollo en potencias de z , exponentes cero y positivos.

La densidad espectral del ruido debe ser, con $f_{sY}(z) = 0$ en (F.2.1) :

$$f_Y(z) = f_X(z) - [H(z)f_s(z)H(1/z)]$$

$$f_Y(z) = f_X(z) - f_X(z)\phi(z)\phi(1/z)$$

$$f_Y(z) = f_X(z)[1 - \phi(z)\phi(1/z)] , \quad (\text{F.3.4})$$

Nota: En la práctica puede ser que el ruido sea especificado de antemano, ya que sobre él no se tiene control, de tal forma que lo que se puede di-

señal es el canal; de (F.3.4) :

$$|\phi(e^{j\omega})|^2 = 1 - \frac{f_v(e^{j\omega})}{f_x(e^{j\omega})} \geq 0 \quad (\text{F.4.1})$$

se escoge luego $f_x(e^{j\omega}) > f_v(e^{j\omega})$ y se encuentra $\phi(z)$ a partir de (F.4.1) por factorización espectral, lo cual fija el canal con (F.3.3). ($f_s(z)$ está dado).

b) Canal causal : La función $\phi(z)$ en (F.3.3) se limita a ser una constante

$$H(z) = \frac{L_x(z)}{L_s(z)} K \quad (\text{F.4.2})$$

y a partir de (F.3.4)

$$f_x(z) = \frac{f_v(z)}{1 - K^2}, \quad K < 1.$$

si $f_v(z) = L_v(z) L_v(1/z)$, con $L_v(z)$ de la fase mínima ,

$$L_x(z) = \frac{L_v(z)}{\sqrt{1 - K^2}} \quad (\text{F.4.3})$$

Se concluye que si se da un ruido con densidad espectral de potencia $f_v(z)$, para que el filtro de Wiener no-causal sea igual al causal, se requiere que el canal sea

$$H(z) = \frac{L(z) K}{L_s(z) [\sqrt{1 - K^2}]} \quad (\text{F.4.4})$$

F.3 RUIDO CORRELACIONADO CON LA SEÑAL. CANAL CAUSAL :

Recuérdese que se está tomando a $f_{sex}(z)$ como causal; dividiendo ambos miembros de (F.2.2) por $L_x(1/z)$:

$$\frac{f_{sx}(z)}{L_x(1/z)} = f_{sex}(z) = H(1/z)f_s(z) + \frac{f_{sv}(z)}{L_x(1/z)}$$

o equivalentemente :

$$\frac{f_{sex}(1/z)}{L_x(1/z)} = \frac{H(z)L_s(z)}{L_x(z)} + \frac{f_{sv}(1/z)}{L_x(z)L_s(1/z)} \quad (F.5.1)$$

$$\text{defínase ahora } \phi(z) \triangleq \frac{f_{sv}(1/z)}{L_x(z)L_s(1/z)} \quad (F.5.2)$$

ya que el miembro izquierdo de (F.5.1) es anticausal, excepto por una constante (véase apéndice C.2), las partes causales del miembro derecho deben cancelarse :

$$\frac{H(z)L_s(z)}{L_x(z)} = -\phi^+(z) + K$$

o bien

$$H(z) = \frac{L_x(z)}{L_s(z)} \left\{ -\phi^+(z) + K \right\} \quad (F.5.3)$$

Dada la densidad espectral de la señal y la densidad espectral cruzada entre

el ruido y la señal, debe todavía determinarse la densidad espectral de x para poder utilizar la fórmula (F.5.3).

El canal resultante dependerá de $f_v(z)$ a través de $f_x(z)$. Sustituyendo (F.5.3) en (F.2.1)

$$\begin{aligned}
 f_x(z) &= f_x(z) \{ (-\phi^+(z)+k)(-\phi^+(1/z)+k) + (-\phi^+(1/z)+k)\phi(z) + (-\phi^+(z)+k)\phi(1/z) \} + \\
 &+ f_v(z) = f_x(z) \{ \phi^+(z)\phi^+(1/z) - k\phi^+(1/z) - k\phi^+(z) + k^2 + (-\phi^+(1/z)+k)(\phi^+(z) + \\
 &+ \phi^-(z)) + (-\phi^+(z)+k)(\phi^+(1/z) + \phi^-(1/z)) + f_v(z) = \\
 &= f_x(z) \{ \phi^+(z)\phi^+(1/z) - k\phi^+(1/z) - k\phi^+(z) + k^2 - \phi^+(1/z)\phi^+(z) + k\phi^+(z) - \\
 &- \phi^+(1/z)\phi^-(z) + k\phi^-(z) - \phi^+(z)\phi^+(1/z) + k\phi^+(1/z) - \phi^+(z)\phi^-(1/z) + k\phi^-(1/z) \} + \\
 &+ f_v(z) = f_x(z) \{ -\phi^+(z)\phi^+(1/z) - \phi^+(1/z)\phi^-(z) - \phi^+(z)\phi^-(1/z) + k(\phi^{-1}(z) + \\
 &+ \phi^-(1/z) + k^2) \} + f_v(z) \\
 &= f_x(z) \{ -(\phi^+(z) + \phi^-(z))(\phi^+(1/z) + \phi^-(1/z)) + \phi^-(z)\phi^-(1/z) + k(\phi^-(z) + \phi^-(1/z) \\
 &+ k^2) \} f_v(z) = f_x(z) \{ -\phi(z)\phi(1/z) + (\phi^-(z) + k)(\phi^-(1/z) + k) \} + f_v(z)
 \end{aligned}$$

usando (F.5.3) :

$$f_x(z) = f_x(z) \left\{ \frac{f_{sv}(1/z)f_{sv}(z)}{f_x(z)f_s(z)} \right\} + f_x(z) \{ (\phi^-(z) + k)(\phi^-(1/z) + k) \} + f(z)$$

$$\rightarrow f_x(z) = - \frac{f_{sv}(1/z)f_{sv}(z)}{f_s(z)} + f_x(z)[(\phi^-(z)+k)(\phi^-(1/z)+k)] + f_v(z)$$

$$f_x(z)[1 - (\phi^-(z)+k)(\phi^-(1/z)+k)] = f_v(z) - \frac{f_{sv}(z)f_{sv}(1/z)}{f_s(z)}$$

y finalmente

$$f_x(z) = \frac{f_v(z) - \frac{f_{sv}(z)f_{sv}(1/z)}{f_s(z)}}{[1 - (\phi^-(z)+k)(\phi^-(1/z)+k)]} \quad (\text{F.7.1})$$

Entonces, primero se encuentra $\phi(z)$ usando (F.5.2), después $f_x(z)$ usando (F.7.1); con $f_x(z)$ se puede determinar $L_x(z)$ para sustituir en (F.5.3). La constante k está dada por $k = \frac{f_{se_x}(\infty)}{L_s(\infty)}$, y como $f_{se_x}(\infty) = R_{se_x}[0]$ es arbitraria, puede escogerse libremente con tal que el denominador sea positivo para toda ω , $z = e^{j\omega}$.

F.4 RUIDO SEMICORRELACIONADO CON LA SEÑAL CANAL CAUSAL :

Si en la sección F.3 se pide además de que $\phi(z)$ sea causal, esto es, que $\phi^-(z) = 0$, la ecuación (F.7.1) se transforma en

$$f_x(z) = \frac{f_v(z) - \frac{f_{sv}(z)f_{sv}(1/z)}{f_s(z)}}{1 - k^2}; \quad k^2 < 1 \quad (\text{F.7.2})$$

Además, de (F.5.2) :

$$\phi(z) = \frac{f_{sv}(1/z)}{L_x(z)L_s(1/z)} = \frac{f_{vs}(z)}{L_s(1/z)L_x(z)} = \frac{f_{ve_s}(z)}{L_x(z)}$$

Como $\phi^+(z) = \phi(z)$, se requiere que $f_{ve_s}(z)$ sea causal, o bien, $f_{e_s v}(1/z)$ es "anticausal". En términos de la correlación cruzada es

$$R_{e_s v}[k] = E\{e_s[n]v[n-k]\} = \begin{cases} 0, & k > 0 \\ \text{arbitraria}, & k \leq 0 \end{cases}$$

o bien :

$$E\{e_s[n]v[m]\} = \begin{cases} 0, & n > m \\ \text{arbitraria}, & n \leq m \end{cases} \quad (\text{F.8.1})$$

Comparando con (4), pág. 148 de T. Kailath. "A view of three decades of linear filtering theory", IEEE Trans. on Information Theory, Vol. IT-20, No.2, March 1974, se concluye que (F.8.2) es la situación opuesta, pero aquí no hay realimentación.

De (F.5.3), se obtiene el canal ($\phi^+(z) = \phi(z)$) :

$$H(z) = -\frac{f_{ve_s}(z)}{L_s(z)} + \frac{k L_x(z)}{L_x(z)} \quad (\text{F.8.2})$$

finalmente, si en (F.7.2) y (F.8.2) se toma $k=0$, se obtiene

$$f_x(z) = f_v(z) \left[1 - \frac{f_{sv}(z)f_{sv}(1/z)}{f_v(z)f_s(z)} \right] \quad (\text{F.8.3})$$

y

$$H(z) = -f_{ve_s}(z) / L_s(z) = -\frac{f_{vs}(z)}{f_s(z)} \quad (\text{F.9.1})$$

Comparando (F.8.3) con (F.3.2) y (F.9.1) con (F.3.1) se observa cierta simetría, en la cual los papeles de "x" y "v" se han intercambiado.

F.5 EJEMPLO: RUIDO BLANCO ORTOGONAL

Considérese el siguiente sistema:

donde A: ganancia del transmisor

$L_s(z)$: filtro de innovaciones del proceso $S = \{s_n\}_{n=-\infty}^{\infty}$

E_s : proceso de innovaciones del proceso $S = \{e_{s_n}\}_{n=-\infty}^{\infty}$

v: ruido blanco ortogonal $v = \{v_n\}_{n=-\infty}^{\infty}$,

$R_v[m] = N\delta[m]$; $R_{ve_s}[m] \equiv R_{vs}[m] \equiv 0$.

X: señal recibida

\hat{S} : estimado de S

\tilde{S} : error.

Se muestra que en éste ejemplo (5.) se satisface. Por simplicidad, se supone que todos los procesos son gaussianos, así que el ruido es independiente de S y E_s , y además, $H(\tilde{S}, \sigma^2) = 1/2 d_{IS}(f_{\tilde{S}}(e^{j\omega}), \sigma^2)$.

Como

$$x[n] = A e_s[n] + v[n]$$

se sigue que

$$f_x(z) = A^2 + N,$$

y por tanto

$$L_x(z) = \sqrt{A^2 + N} = \ell_0$$

Ahora bien

$$e_x[n] = \frac{A e_s[n] + v[n]}{\ell_0}$$

y

$$R_{se_x}[m] = E \left\{ S[n] \left[\frac{A e_s[n-m] + v[n-m]}{\ell_0} \right] \right\} = \frac{A}{\ell_0} R_{se_s}[m]$$

o bien

$$f_{se_x}(z) = \frac{A}{\ell_0} L_s(z)$$

lo cual muestra que se satisface (5.), y el filtro de Wiener causal es igual al no-causal. Nótese que (F.4.2), (F.4.3) y (F.4.4) se satisfacen con

$$k = A/\ell_0 .$$

La densidad de potencia espectral es de (5.) :

$$f_{\xi}(z) = [G_1(z)\ell_0 - \frac{A}{\ell_0} L_S(z)] [G_1(1/z)\ell_0 - \frac{A}{\ell_0} L_S(1/z)] - \frac{A^2}{\ell_0^2} f_S(z) + f_S(z)$$

y el filtro de Wiener es $G_W(z) = \frac{1}{\ell_0} [\frac{A}{\ell_0} L_S(z)]^+ = \frac{A}{\ell_0^2} L_S(z) .$

Por tanto:

- a) Si se usa el filtro de Wiener, $G_1(z) = G_W(z)$, la potencia del error es mínima y la densidad espectral de potencia es:

$$f_{\xi}(z) = f_S(z) [1 - \frac{A^2}{\ell_0^2}]$$

- b) Se puede aumentar la potencia del error, pero incrementar la planitud de la densidad espectral de potencia del error; si se propone $G_1(z) = \frac{A L_S(z)}{\ell_0^2} + \frac{\Delta(z)}{\ell_0}$, entonces

$$f_{\xi}(z) = \Delta(z)\Delta(1/z) + f_S(z) [1 - \frac{A^2}{\ell_0^2}] ;$$

Sea $\Delta(z)$ tal que

$$\Delta(z)\Delta(1/z) = k - f_S(z) [1 - \frac{A^2}{\ell_0^2}] , \text{ con } k$$

tal que $|\Delta(e^{j\omega})|^2 \geq 0, \forall \omega$. Entonces, existe $\Delta(z)$ causal, de fase mínima, tal que

$$f_{\xi}(z) = k$$

G.1 PROCESOS REGULARES :

Se cita a continuación un teorema que define los procesos regulares [15, pág. 252-264]

Teorema : Para que un proceso escalar $\xi[n]$ con la distribución espectral $F(\omega)$ constituya la respuesta de un filtro causal a cuya entrada llega un proceso discreto blanco con media cero y variancia unitaria ($\zeta_0[n]$), ésto es, para que $\xi[n]$ sea regular, es necesario y suficiente que su distribución espectral $F(\omega)$ sea absolutamente continua (y por tanto existe la densidad de potencia espectral $f(\omega) = \frac{d}{d\omega} F(\omega)$), y la densidad espectral $f(\omega)$ satisfaga la condición

$$\int_{-\pi}^{\pi} \lambda_n f(\omega) \frac{d\omega}{2\pi} > -\infty \quad (\text{G.1.1})$$

Cumplidas éstas condiciones, $\xi[n]$ se puede representar en la forma :

$$\xi[n] = \sum_{k=-\infty}^n h[n-k] \zeta_0[k] ; \sum_{n=0}^{\infty} |h[n]|^2 < \infty$$

Nota : La condición (G.1.1) sólo se satisface si $f(\omega) = 0$ a lo más en un conjunto con medida de Lebesgue cero.

El concepto de proceso regular tiene su complemento en el concepto de proceso singular:

Procesos Singulares : Si la distribución espectral no tiene componente abso

lutamente continua, o si la tiene pero ésta no satisface la condición (G.1.1), por ejemplo, si $f(\omega) = \frac{d}{d\omega} F_1(\omega)$ se anula en el conjunto de la medida positiva de Lebesgue, el proceso se llama singular.

Nota : $F_1(\omega)$ es la componente absolutamente continua de la distribución espectral.

Nota : En estos conceptos se sobreentiende que los procesos son discretos y estacionarios, y que el índice de definición de los mismos va desde $n = -\infty$ hasta $n = \infty$.

G.2 FACTORIZACION ESPECTRAL :

De acuerdo a [13, pág. 501], una función de densidad espectral de la forma:

$$f(\omega) = c' e^{j\omega a} \frac{\prod_{j=1}^{\alpha'} (e^{j\omega} - \omega_j^i)}{\prod_{j=1}^{\beta'} (e^{j\omega} - z_j^i)}, \quad \begin{array}{l} c', \omega_j^i, z_j^i \neq 0 \\ \omega_j^i \neq z_j^i, \neq i, j \\ |z_j^i| \neq 1 \end{array}$$

puede escribirse en la forma

$$f(\omega) = \left| \frac{\sum_{i=0}^{\alpha} A_i e^{ji\omega}}{\sum_{i=0}^{\beta} B_i e^{ji\omega}} \right|^2,$$

donde las raíces de $\sum_{j=0}^{\beta} A_j z^j$, ω_j , son las ω_j^i de módulo ≤ 1 , y las raíces de

$\sum_{j=0}^{\alpha} B_j z^j$, z^j , z_j , son las z_j de módulo < 1 .

Esto es : $f(z) = H(z)H(z^{-1})$, donde $H(z)$ tiene todos sus polos dentro del círculo unitario, y sus ceros dentro y sobre el círculo unitario. Esto no concuerda exactamente con [5, pág. 297]. Esto es, $H(z)$ no tiene que ser de fase mínima, porque su inverso no tiene que ser estable. Sin embargo, si $\alpha = \beta$, $H(z)$ es causal estable, y su inverso es causal, y de acuerdo al Apéndice A.1, $H(z)$ tiene asociada una transformación lineal (infinita) invertible.

H.1 REPRESENTACION ESPECTRAL DE PROCESOS PERIODICOS C.M.

El siguiente teorema es la versión correspondiente a procesos de tiempo discreto del teorema dado para procesos periódicos C.M. de tiempo continuo en [5, pág. 301].

Teorema : Sea $x[n]$ un proceso de tiempo discreto, estacionario en sentido amplio, periódico cuadrático medio con período N , de tal forma que su correlación es periódica y puede expandirse en una serie discreta de Fourier, éste es:

$$R_x[l] = \frac{1}{N} \cdot \sum_{i=0}^{N-1} T_i W_N^{-il} ;$$

$$T_i = \sum_{l=0}^{N-1} R_x[l] W_N^{il} ; W_N = e^{-j\frac{2\pi}{N}} , \quad (\text{H.1.1})$$

Sea $\hat{x}[n]$ un proceso de tiempo discreto, estacionario en sentido amplio, estrictamente periódico (ésto es, con realizaciones periódicas), definido de la siguiente forma :

$$\hat{x}(n, \omega) = \frac{1}{\sqrt{N}} \cdot \sum_{k=0}^{N-1} C_k(\omega) W_N^{-kn} \quad (\text{H.1.2})$$

donde los coeficientes C_k son variables aleatorias definidas por :

$$C_k(\omega) = \frac{1}{\sqrt{N}} \cdot \sum_{n=0}^{N-1} x(n, \omega) W_N^{kn} \quad (\text{H.1.3})$$

Entonces se cumple lo siguiente :

a) Los coeficientes C_k no están correlacionados, éste es :

$$E\{C_k C_j^*\} = T_k \delta_N[k-j] \quad F$$

b) $x[n]$ y $\hat{x}[n]$ tienen la misma correlación :

$$R_x[l] = R_{\hat{x}}[l]$$

c) $x[n]$ y $\hat{x}[n]$ son iguales en el sentido cuadrático medio

$$E\{x[n] - \hat{x}[n]\}^2 = 0, \text{ para toda } n.$$

d) $x[n]$ y $\hat{x}[n]$ son estocásticamente equivalentes

e) $P\{x(n, \omega) = \hat{x}(n, \omega), \forall n\} = 1$, ésto es, las realizaciones no periódicas de $x[n]$ tienen probabilidad de ocurrencia cero.

$$F \quad \delta_N[i] \triangleq \begin{cases} 1 & i = kN, 0, \pm 1, \pm 2, \dots \\ 0 & i \neq kN \end{cases}$$

Pruebas :

$$a) \quad E\{C_k C_j^*\} = \frac{1}{N} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} E\{x[n]x^*[m]\} W_N^{kn-jm}$$

$$= \frac{1}{N} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} R[n-m] W_N^{k(n-m)} W_N^{m(k-j)}$$

$$= \frac{1}{N} \sum_{m=0}^{N-1} W_N^{m(k-j)} \sum_{n=0}^{N-1} R[n-m] W_N^{k(n-m)}$$

Notando que $R[n]$ y W_N^{kn} son de período N y usando (H.1.1) :

$$\sum_{n=0}^{N-1} R[n-m] W_N^{k(n-m)} = \sum_{n=-m}^{N-1-m} R[n] W_N^{kn} = T_k$$

por lo cual

$$E\{C_k C_j^*\} = \frac{T_k}{N} \sum_{m=0}^{N-1} W_N^{m(k-j)} = T_k \delta_N[k-j]$$

b) La autocorrelación de $x[n]$ está dada por

$$\begin{aligned} R_{\hat{x}}[\ell] &= E\{\hat{x}[n+\ell] x^*[n]\} = \frac{1}{N} \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} E\{C_k C_j^*\} W_N^{-kn-k\ell+jn} = \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} T_k \delta_N[k-j] W_N^{-n(k-j)-k\ell} = \\ &= \frac{1}{N} \sum_{k=0}^{N-1} T_k W_N^{-k\ell} = R_x[\ell] \end{aligned}$$

donde se usó (H.1.1).

c) $x[n]$ y $\hat{x}[n]$ son iguales en el sentido cuadrático medio, ya que

$$\begin{aligned} E\{(x[n] - \hat{x}[n])(x^*[n] - \hat{x}^*[n])\} &= \\ &= R_x[0] + R_{\hat{x}}[0] - E\{x[n] \hat{x}^*[n] + \hat{x}[n] x^*[n]\} \end{aligned}$$

Además

$$E\{x[n] \hat{x}^*[n]\} = \frac{1}{\sqrt{N}} \sum_{k=1}^{N-1} E\{x[n] C_k^*\} W_N^{kn} \quad (\text{H.3.1})$$

El valor esperado requerido en la expresión anterior está dado por:

$$E\{x[n]C_k^*\} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} R_x[n-i] W_N^{k(n-i)} W_N^{-kn}$$

y notando que:

$$\begin{aligned} \sum_{i=0}^{N-1} R_x[n-i] W_N^{k(n-i)} &= \sum_{i=0}^{N-1} R_x[i-n] W_N^{-k(i-n)} = \sum_{i=-n}^{N-1-n} R_x[i] W_N^{-ki} = \\ &= \sum_{i=-(N-1)}^0 R_x[i] W_N^{ki} = \sum_{i=0}^{N-1} R_x[i] W_N^{ki} = T_k = T_{-k} \end{aligned} \quad (\text{H.4.1})$$

donde se ha usado la periodicidad de $R_x[\ell]$ y $W_N^{k\ell}$, al que $R_x[\ell]$ es par y el que la sumatoria sea sobre un período; entonces

$$E\{x[n]C_k^*\} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} R_x[i] W_N^{ki} W_N^{-kn} \quad (\text{H.4.2})$$

sustituyendo (H.4.2) en (H.3.1) :

$$E\{x[n]x^*[n]\} = \frac{1}{N} \sum_{k=1}^N T_k = R_x[0]$$

Similarmente, se demuestra que

$$E\{\hat{x}[n] x^*[n]\} = R_x[0]$$

y como por b), $R_x[0] = R_{\hat{x}}[0]$, se concluye

$$E\{(X[n] - \hat{x}[n])^2\} = 0, \quad \forall n.$$

d) $P\{|x[n] - \hat{x}[n]| < \epsilon\} \geq$

$$1 - \frac{E\{(x[n] - \hat{x}[n] - E\{x[n] - x[n]\})^2\}}{\epsilon^2}$$

para cualquier $\epsilon > 0$, por el teorema de Tchebycheff. Pero :

$$E\{\hat{x}[n]\} = \frac{1}{N} \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} E\{x[n]\} W_N^{jk} W_N^{-kn}$$

de (H.1.2) y (H.1.3), por lo cual

$$E\{\hat{x}[n]\} = E\{x[n]\} \cdot \sum_{j=0}^{N-1} \delta_N[j-n]$$

por ser $x[n]$ estacionario en sentido amplio.

Además ,

$$\sum_{j=0}^{N-1} \delta_N[j-n] = 1, \quad \forall n.$$

por lo que $E\{\hat{x}[n]\} = E\{x[n]\}$, y aplicando el resultado c) de éste apéndice :

$$P\{|x[n] - \hat{x}[n]| < \epsilon\} = 1, \quad \forall \epsilon > 0.$$

o sea

$$P\{x(n, \omega) = \hat{x}(n, \omega)\} = 1, \quad \text{para toda } n \text{ fija.}$$

Esta última es la condición de equivalencia estocástica [15] .

e) Sea $u(n, \omega) = x(n, \omega) - \hat{x}(n, \omega)$

Entonces, por d) ,

$P\{u(n, \omega) = 0\} = 1$ para toda n fija, y se sabe que [13, pág. 51 y 53] :

$$P\{u(n, \omega) = 0, \forall n\} = 1 \quad (\text{H.6.1})$$

o bien

$$P\{x(n, \omega) = \hat{x}(n, \omega), \forall n\} = 1$$

ya que para cada valor de n se tiene un conjunto Λ_n de puntos ω con $P\{\Lambda_n\} = 0$ tal que $x(n, \omega) \neq \hat{x}(n, \omega)$ si $\omega \in \Lambda_n$. La unión de estos conjuntos es enumerable por ser el índice n enumerable, y aplicando la desigualdad de Boole

$$P\left\{\bigcup_{n=-\infty}^{\infty} \Lambda_n\right\} \leq \sum_{n=-\infty}^{\infty} P\{\Lambda_n\} = 0, \text{ y se sigue (H.6.1) .}$$

Nota : En el caso de procesos de tiempo continuo, a partir de la equivalencia estocástica no se sigue que $P\{x(t, \omega) = \hat{x}(t, \omega), \forall t\} = 1$. Una condición suficiente, pero no necesaria, para efectuar ésta deducción, es pedir que el proceso sea separable [13, pág. 53]. Sin embargo, los procesos de tiempo discreto son implícitamente separables [13, pág. 56] .

REFERENCIAS

- [1] J.E. Shore y R.M. Gray, "Minimum cross-entropy pattern classification and cluster analysis", IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI-4, pág. 11-41.

- [2] S. Kullback, Information Theory and Statistics, New York: Dover, 1969.

- [3] R.M. Gray, A.H. Gray, Jr., G. Rebolledo y J.E. Shore, "Rate-distortion speech coding with a minimum discrimination information distortion measure", IEEE Trans. Inform. Theory, Vol.IT-27, pág.708-721 Nov.1969.

- [4] R.M. Gray, A. Buzo, A.H. Gray, Jr., y Y. Matsuyama, "Distortion measures for speech processing", IEEE Trans. Acoust., Speech, Signal Processing, Vol.ASSP-28, pág.367-376, Ago. 1980.

- [5] A.Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, 1984.

- [6] M.Pinsker, Information and Information Stability of Random Variables and Processes, Ize. Akad. Nauk. SSSR, Moscow, 1960. (Traducción al inglés: San Francisco, CA: Holden-Day, 1964).

- [7] J.E. Shore y R.W. Johnson, "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy", IEEE Trans. Inform. Theory, Vol.IT-26, pág. 26-37, Enero 1980.

[8]

[9] K. Itô, "Introduction to Probability Theory", Cambridge University Press, 1984.

[10] F. Itakura y S. Saito, "Analysis synthesis telephony based on the maximum likelihood method", Rep. 6th Int. Cong. Acoust., Y. Yonasi. Ed., Tokyo, 1968, pág. C12-C20.

[11] W.A. Gardner, "Introduction to Random Processes with Applications to Signals and Systems", MacMillan Publishing Co, 1986.

[12] A. Papoulis, "Maximum entropy and spectral estimation: a review". IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-29, pág 1176-1186, Dic. 1981.

[13] J.L. Doob, "Stochastic Processes", John Wiley and Sons Inc., 1953

[14] Alan V. Oppenheim, Ronald W. Schaffer, "Digital Signal Processing", Prentice-Hall, 1975.

[15] V.S. Koroliuk, "Manual de la Teoría de Probabilidades y Estadística Matemática", Editorial Mir, 1981.

[16] R.G. Gallager, "Information Theory and Reliable Communication", John

- Wiley and Sons Inc., 1968.
- [17] D.S. Jones, Elementary Information Theory, Oxford University Press, 1974.
- [18] A. Papoulis, Signal Analysis
- [19] A. Papoulis, The Fourier Integral and its Applications
- [20] L.H. Koopmans, The Spectral Analysis of Time Series, Academic Press, 1974.
- [21] G.M. Jenkins, D.G. Watts, Spectral Analysis and its Applications, Holden-Day Inc. 1968.
- [22] J. Lamperti, Stochastic Processes. A Survey of the Mathematical Theory, Springer-Verlag, New York Inc. 1977.
- [23] K.L. Chung, A Course in Probability Theory, Academic Press Inc, 1974.
- [24] N.S. Jayant, P. Noll, Digital Coding of Waveforms. Principles and Applications to Speech and Video, Prentice-Hall, 1984.