



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE INGENIERIA

DIVISION DE ESTUDIOS DE POSGRADO

TESIS

EXTENSION A DOS DIMENSIONES DE LOS FILTROS
ANTICIPATORIOS PARA ESTIMACION, DE ACUERDO
AL CRITERIO DE INFORMACION DE DISCRIMINACION
NORMALIZADA MINIMA DEBIDO A LEPE-CASILLAS, Y
SU IMPLANTACION PARA RECONSTRUIR IMAGENES.

PRESENTADA POR:

DAMIAN FEDERICO VARGAS SANDOVAL

PARA OBTENER EL GRADO DE:

MAESTRO EN INGENIERIA

(COMUNICACIONES)

DIRIGIDA POR:

M. en I. FERNANDO LEPE CASILLAS

Ciudad Universitaria, mayo de 1998.



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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 esta 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.



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 esta 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.

Extensión a Dos Dimensiones de los Filtros
Anticipatorios para Estimación, de Acuerdo al
Criterio de Información de Discriminación
Normalizada Mínima debido a Lepe-Casillas, y su
Implantación para Reconstruir Imágenes

Damián Federico Vargas Sandoval

Mayo 1998

Resumen

Se presenta en este trabajo la extensión a dos dimensiones de un filtro para estimación basado en la Teoría de Información desarrollado por Lepe-Casillas, y su aplicación a la restauración o reconstrucción de imágenes. En el Capítulo 1 se presentan los requerimientos mínimos para este trabajo de la Teoría de Información: definiciones y algunas relaciones importantes. En el Capítulo 2 se presenta un breve resumen de los resultados obtenidos en una dimensión, de la Teoría de Información aplicada a filtrado, de los cuales se obtienen 3 Criterios de filtrado diferentes. Considerando el caso no causal, del Primer Criterio se obtiene el **Filtro de Wiener No Causal**; con el Segundo Criterio se obtiene el **Filtro de Igualación de Espectro con Fase Optimizada**, y con el Tercer Criterio se obtiene el **Filtro de Mínima Razón de Información de Discriminación de Propósito Dual** y motivo principal de este trabajo, el cual tiene un comportamiento intermedio entre ambos; es decir, mientras el Filtro de Wiener minimiza la potencia del error y el de Igualación de Espectro iguala el espectro de la señal filtrada con el de la señal original, el Filtro de Propósito Dual trata de hacer ambas cosas a la vez, de donde se deriva el nombre. Las definiciones y relaciones básicas requeridas de la Teoría de Información en dos dimensiones, se presentan en el Capítulo 3, mientras que la extensión a dos dimensiones de estos criterios, se presentan en el Capítulo 4. Finalmente, en el Capítulo 5, se presentan los resultados de la aplicación de estos criterios al procesamiento digital de imágenes, específicamente en reconstrucción o restauración, mediante simulación por computadora. Se incluyen 2 Apéndices en los que se fundamentan las consideraciones teóricas empleadas en muchas partes del desarrollo de este trabajo.

A MI ESPOSA MARGARITA
(Antes mi novia, y, en todo momento, mi amor).

A MI PAPÁ GABRIEL
A MI MAMÁ LIDIA
(ahora ausente pero siempre presente en mi corazón).

A MIS HERMANOS Y SOBRINOS,
por compartir conmigo momentos inolvidables.

AGRADEZCO SINCERAMENTE

Al Profesor Fernando Lepe Casillas por sus invaluable consejos, ayuda incondicional y la guía que significó durante mi paso por la DEPFI.

A mis amigos y compañeros Héctor y Alberto.

Al Dr. Francisco García Ugalde por la ayuda que me brindó.

A todo el personal académico de la Sección de Ingeniería Eléctrica de la DEPFI, en particular al Dr. Boris Escalante Ramírez, por todo su apoyo.

G(2) 503790

Contenido

1	Requerimientos de la Teoría de Información	5
1.1	Introducción.	5
1.2	Definiciones Fundamentales.	5
1.2.1	Entropía Diferencial Normalizada por Muestra.	5
1.2.2	Información de Discriminación Normalizada por Muestra.	6
1.2.3	Información Mutua Normalizada por Muestra.	6
1.3	Algunas Relaciones Importantes.	6
1.4	Observaciones Finales.	7
2	Aplicaciones de la Teoría de Información a Filtrado	8
2.1	Introducción.	8
2.2	Primer Criterio de Filtrado basado en la Teoría de Información.	8
2.3	Segundo Criterio de Filtrado Basado en la Teoría de Información.	10
2.4	Tercer Criterio de Filtrado Basado en la Teoría de Información.	10
2.5	Observaciones Finales.	13
3	Definiciones de Teoría de Información en dos Dimensiones.	14
3.1	Introducción.	14
3.2	Campos Aleatorios.	14
3.2.1	Elementos Probabilísticos [6].	14
3.2.2	Relaciones Entrada-Salida para un SLIC.	15
3.3	Definiciones.	16
3.3.1	Entropía Diferencial Normalizada.	16
3.3.2	Información de Discriminación Normalizada.	17
3.3.3	Información Mutua Normalizada.	17
3.4	Algunos Resultados Importantes.	17
4	Extensión a Dos Dimensiones del Criterio de Lepe-Casillas para Filtrado.	19
4.1	Introducción.	19
4.2	Resultados de la Extensión.	19
4.3	Resultados de la optimización.	21
4.4	Comentarios Finales.	23

5	Aplicaciones a Procesamiento de Imágenes	24
5.1	Introducción	24
5.2	Aspectos Numéricos	24
5.2.1	Imágenes y Matrices	25
5.2.2	La expresiones de los filtros	25
5.2.3	Observaciones Adicionales.	28
5.3	Programación	28
5.4	Resultados	30
5.5	Conclusiones	31
A		33
A.1	Introducción.	33
A.2	Información de Discriminación Normalizada entre los procesos \hat{S} y S	33
A.3	Procesos Asintóticamente Regulares y Asintóticamente Estacionarios.	34
A.3.1	Equivalencia Asintótica entre Secuencias de Matrices [22].	34
A.3.2	Definición de Procesos Asintóticamente Estacionarios y Asintóticamente Regulares.	36
A.3.3	Conclusiones.	39
A.4	Algunos resultados de la Teoría de Predicción Lineal	40
A.4.1	Predicción Lineal de un Paso de Orden Infinito.	40
A.4.2	Predicción Lineal de Un Paso de Orden constante Finito.	41
A.4.3	Predicción Lineal de Un Paso de Orden Variable Finito.	44
A.5	Aplicaciones.	46
A.5.1	Caso 1: \hat{S} y S Regulares y Estacionarios.	46
A.5.2	Caso 2: \hat{S} Asintóticamente Regular y Asintóticamente Estacionario y S Regular y Estacionario.	48
A.5.3	Conclusiones.	48
B		49
B.1	Introducción.	49
B.2	Conceptos Preliminares.	49
B.2.1	Regiones Causales.	49
B.2.2	Campos Regulares	49
B.2.3	Matrices de Correlaciones de Campos Homogéneos	51
B.2.4	Campos Asintóticamente Regulares y Asintóticamente Homogéneos.	55
B.3	Resultados de Predicción Lineal en Dos Dimensiones.	59
B.4	Reducción de $\bar{D}(\hat{S}, S)$	62

Glosario

- $\equiv \{\mathbf{x}_n\}_{n=-\infty}^{\infty}$ Proceso estocástico discreto real.
- \bar{H}_x Entropía Diferencial Normalizada por Muestra del Proceso \mathbf{X} .
- $\bar{D}(\mathbf{X}, \mathbf{Y})$ Información de Discriminación Normalizada por Muestra del proceso \mathbf{X} con respecto al proceso \mathbf{Y}
- $\bar{I}(\mathbf{X}, \mathbf{Y})$ Información Mutua Normalizada por Muestra entre los procesos conjuntamente estacionarios \mathbf{X} y \mathbf{Y}
- $f_x(e^{j\omega})$ Función de Densidad Espectral del proceso \mathbf{X} .
- \mathbf{E} Proceso de Error.
- $\hat{\mathbf{S}}$ Proceso Estimador de \mathbf{S} .
- \mathbf{S} Proceso original.
- $\mathbf{G}(e^{j\omega})$ Filtro estimador.
- $\mathbf{G}_{nc}^W(e^{j\omega})$ Filtro de Wiener No Causal.
- $\mathbf{G}^{SE}(e^{j\omega})$ Filtro de Igualación de espectro con fase optimizada.
- $\gamma_{xs}(e^{j\omega})$ Coherencia entre los procesos \mathbf{X} y \mathbf{S} .
- $\mathbf{G}^o(e^{j\omega})$ Filtro de Mínima Razón de Información de discriminación de Propósito Dual.
- $X(\mathbf{t}) = X$ Campo Aleatorio Estocástico Real
- $R_X(\mathbf{t}_1, \mathbf{t}_2)$ Función de Correlación del campo X .
- \bar{H}_X Entropía Diferencial Normalizada del campo X
- $\bar{D}(X, Y)$ Información de Discriminación Normalizada del campo X con respecto al campo Y
- $\bar{I}(X, Y)$ Información Mutua Normalizada entre los campos X y Y
- $f_X(e^{j\omega_1}, e^{j\omega_2})$ Función de Densidad Espectral de Potencia del campo X
- $f_Y(e^{j\omega_1}, e^{j\omega_2})$ Función de Densidad Espectral de Potencia del campo Y
- $G^o(e^{j\omega_1}, e^{j\omega_2})$ Filtro de Mínima Razón de Información de discriminación de Propósito Dual Bidimensional
- $G^W(e^{j\omega_1}, e^{j\omega_2})$ Filtro de Wiener Bidimensional
- $G^{SE}(e^{j\omega_1}, e^{j\omega_2})$ Filtro de Igualación de Espectro Bidimensional con Fase Optimizada
- $\mathbf{R}_{\mathbf{X}(N)}$ Matriz de correlaciones de orden N del proceso \mathbf{X}

$\Delta_{s(N)}$ Determinante de $\underline{\mathbf{R}}_{x(N)}$

$\underline{\mathbf{A}}_N$ Secuencia de matrices $\underline{\mathbf{A}}$ de orden $N \times N$

$\underline{\mathbf{A}}_N \sim \underline{\mathbf{B}}_N$ Equivalencia asintótica entre $\underline{\mathbf{A}}_N$ y $\underline{\mathbf{B}}_N$

$\underline{\mathbf{L}}_{(N)}$ Matriz de orden N asociada a la transformación lineal efectuada por el sistema con función de transferencia $\mathbf{L}(e^{j\omega})$

$R_{X(N)}$ Matriz de Correlaciones de orden N del campo X

Capítulo 1

Requerimientos de la Teoría de Información

1.1 Introducción.

Se presentan en este capítulo algunas definiciones y relaciones importantes de la Teoría de Información tomadas de [1]. Estas definiciones son de la mayor relevancia para este trabajo. Su antecedente directo se encuentra en [4] sobre Teoría de Información aplicada a Procesos Estocásticos empleando profusamente Teoría de la Medida. Se supone que los procesos estocásticos involucrados son discretos, reales, cuyas funciones de densidad de probabilidad conjuntas de orden N existen para todo valor de N . Se suponen además estacionarios y regulares, y si se aplican a Sistemas Lineales Invariantes con el Tiempo (SLIT) desde $-\infty$, los procesos de salida también son estacionarios y regulares. Cuando se aplican desde cero, se da lugar a otro tipo de procesos, los *asintóticamente estacionarios* y *asintóticamente regulares*, de los que se habla más adelante.

1.2 Definiciones Fundamentales.

Se presentan a continuación las definiciones fundamentales para este trabajo.

1.2.1 Entropía Diferencial Normalizada por Muestra.

La Entropía Diferencial Normalizada por Muestra del proceso $\mathbf{X} \equiv \{\mathbf{x}_n\}_{n=-\infty}^{\infty}$ se define como

$$\bar{H}_x = \lim_{N \rightarrow \infty} \frac{1}{N} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} P_x^N(\mathbf{u}^N) \ln \left\{ \frac{1}{P_x^N(\mathbf{u}^N)} \right\} \delta_u^N; \quad (1.1)$$

donde $\delta_u^N = du_1 du_2 \dots du_N$ es un elemento diferencial en R^N , $\mathbf{u}^N = (u_1, u_2, \dots, u_N)^T$ es un vector mudo de N elementos también en R^N , y $P_x^N(\mathbf{u}^N)$ es la función de densidad de probabilidad (f.d.p.) conjunta de orden N del proceso \mathbf{X} . También se le conoce como Razón de Cambio de Entropía Relativa.

1.2.2 Información de Discriminación Normalizada por Muestra.

La Información de Discriminación Normalizada por Muestra del proceso \mathbf{X} con respecto al proceso \mathbf{Y} está dado por

$$\bar{D}(\mathbf{X}, \mathbf{Y}) = \lim_{N \rightarrow \infty} \frac{1}{N} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} P_x^N(\underline{\mathbf{u}}^N) \ln \left\{ \frac{P_x^N(\underline{\mathbf{u}}^N)}{P_y^N(\underline{\mathbf{u}}^N)} \right\} \delta_u^N ; \quad (1.2)$$

donde $P_y^N(\underline{\mathbf{u}}^N)$ es la f.d.p. conjunta de orden N del proceso \mathbf{Y} . Nótese que no es simétrica. Se le conoce también como Razón de Cambio de Entropía Cruzada.

1.2.3 Información Mutua Normalizada por Muestra.

La Información Mutua Normalizada por Muestra entre los procesos conjuntamente estacionarios \mathbf{X} y \mathbf{Y} está definida por

$$\bar{I}(\mathbf{X}, \mathbf{Y}) = \lim_{N \rightarrow \infty} \frac{1}{N} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} P_{xy}^{2N}(\underline{\mathbf{u}}^N, \underline{\mathbf{v}}^N) \ln \left\{ \frac{P_{xy}^{2N}(\underline{\mathbf{u}}^N, \underline{\mathbf{v}}^N)}{P_x^N(\underline{\mathbf{u}}^N) P_y^N(\underline{\mathbf{v}}^N)} \right\} \delta_u^N \delta_v^N ; \quad (1.3)$$

siendo $P_{xy}^{2N}(\underline{\mathbf{u}}^N, \underline{\mathbf{v}}^N)$ la f.d.p. conjunta de orden $2N$ de los procesos \mathbf{X} y \mathbf{Y} . Es un caso particular de la Información de Discriminación Normalizada cuando se consideran los procesos bivariados $\mathbf{Z} = (\mathbf{X}, \mathbf{Y})$ y $\mathbf{Z}^\times = (\mathbf{X} \times \mathbf{Y})$ cuyas f.d.p. de orden $2N$ son ¹

$$P_z^{2N}(\underline{\mathbf{u}}^N, \underline{\mathbf{v}}^N) = P_{xy}^{2N}(\underline{\mathbf{u}}^N, \underline{\mathbf{v}}^N) \text{ y } P_{z^\times}^{2N}(\underline{\mathbf{u}}^N, \underline{\mathbf{v}}^N) = P_x^N(\underline{\mathbf{u}}^N) P_y^N(\underline{\mathbf{v}}^N),$$

respectivamente. Así, se tiene que

$$\bar{D}(\mathbf{Z}, \mathbf{Z}^\times) = \lim_{N \rightarrow \infty} \frac{1}{N} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} P_z^{2N}(\underline{\mathbf{u}}^N, \underline{\mathbf{v}}^N) \ln \left\{ \frac{P_z^{2N}(\underline{\mathbf{u}}^N, \underline{\mathbf{v}}^N)}{P_{z^\times}^{2N}(\underline{\mathbf{u}}^N, \underline{\mathbf{v}}^N)} \right\} \delta_u^N \delta_v^N ,$$

$$\bar{D}(\mathbf{Z}, \mathbf{Z}^\times) = \bar{I}(\mathbf{X}, \mathbf{Y}) = \bar{I}(\mathbf{Y}, \mathbf{X}) . \quad (1.4)$$

Obsérvese que 1.4 es la definición de la Información de Discriminación Normalizada para *Procesos Bivariados*.

1.3 Algunas Relaciones Importantes.

1) Si $\mathbf{X} \equiv \{\mathbf{x}_n\}_{n=-\infty}^{\infty}$ es un proceso aplicado desde $n = 0$ a la entrada de un SLIT estable, con función de transferencia $\mathbf{H}(e^{j\omega})$, condiciones iniciales nulas y respuesta al impulso $h[n] \equiv h_n$, $n = 0, 1, 2, \dots$; entonces la salida \mathbf{Y} es un proceso estocástico *asintóticamente estacionario* (ver Apéndice A), cuya razón de cambio de entropía \bar{H}_y es [17, pag. 299]:

$$\bar{H}_y = \bar{H}_x + \frac{1}{2} \ln |h_0|^2 , \quad h_0 \neq 0 . \quad (1.5)$$

¹Notación Introducida por Lepe en [2]

Si la misma entrada se aplica al mismo sistema, pero ahora desde $n = -\infty$, el proceso de salida \mathbf{Y} es estacionario y su razón de cambio de entropía es ahora [14, pag. 590]:

$$\bar{\mathbf{H}}_y = \bar{\mathbf{H}}_x + \frac{1}{2} \int_{-\pi}^{\pi} \ln |\mathbf{H}(e^{j\omega})|^2 \frac{d\omega}{2\pi}; \quad (1.6)$$

ambos resultados coinciden si el sistema es de fase mínima. Lepe-Casillas hace especial énfasis en la diferencia entre 1.5 y 1.6: la primera la emplea en el diseño de filtros y predictores causales, mientras que emplea la segunda en los filtros no causales. Finalmente, si \mathbf{X} es un proceso gaussiano, de la definición se tiene que

$$\bar{\mathbf{H}}_x = \frac{1}{2} \ln(2\pi e) + \frac{1}{2} \int_{-\pi}^{\pi} \ln f_x(e^{j\omega}) \frac{d\omega}{2\pi}; \quad (1.7)$$

donde $f_x(e^{j\omega})$ es la *función de densidad espectral de potencia* (f.d.e.p.) del proceso \mathbf{X} .

2) Sean $\mathbf{X} \equiv \{\mathbf{x}_n\}_{n=-\infty}^{\infty}$ y $\mathbf{Y} \equiv \{\mathbf{y}_n\}_{n=-\infty}^{\infty}$ procesos gaussianos con espectros de densidad de potencia $f_x(e^{j\omega})$ y $f_y(e^{j\omega})$, y matrices de correlación de orden N $\mathbf{R}_x = E \{\underline{\mathbf{x}}^N \underline{\mathbf{x}}^{NT}\}$ y $\mathbf{R}_y = E \{\underline{\mathbf{y}}^N \underline{\mathbf{y}}^{NT}\}$, respectivamente, se cumple entonces que

$$\begin{aligned} \bar{\mathbf{D}}(\mathbf{X}, \mathbf{Y}) &= \frac{1}{2} \int_{-\pi}^{\pi} \left(\frac{f_x(e^{j\omega})}{f_y(e^{j\omega})} - \ln \left\{ \frac{f_x(e^{j\omega})}{f_y(e^{j\omega})} \right\} - 1 \right) \frac{d\omega}{2\pi} \\ \bar{\mathbf{D}}(\mathbf{X}, \mathbf{Y}) &= \frac{1}{2} d_{IS}(f_x(e^{j\omega}), f_y(e^{j\omega})); \end{aligned} \quad (1.8)$$

es decir, la razón de cambio de entropía cruzada del proceso \mathbf{X} con respecto al proceso \mathbf{Y} , cuando ambos son gaussianos, es igual a un medio de la *Distorsión de Itakura Saito*.

3) Sean $\mathbf{X} \equiv \{\mathbf{x}_n\}_{n=-\infty}^{\infty}$, $\mathbf{Y} \equiv \{\mathbf{y}_n\}_{n=-\infty}^{\infty}$ y $\mathbf{Z} = \mathbf{X} \pm \mathbf{Y}$ procesos con razones de cambio de entropía definidas individualmente y para cualquier par de ellos. Se cumple que [1]:

$$\begin{aligned} \bar{\mathbf{H}}_z &= \bar{\mathbf{I}}(\mathbf{X}, \mathbf{Z}) + \bar{\mathbf{H}}_y - \bar{\mathbf{I}}(\mathbf{X}, \mathbf{Y}); \\ \bar{\mathbf{H}}_z &= \bar{\mathbf{I}}(\mathbf{Y}, \mathbf{Z}) + \bar{\mathbf{H}}_x - \bar{\mathbf{I}}(\mathbf{X}, \mathbf{Y}). \end{aligned} \quad (1.9)$$

1.4 Observaciones Finales.

En las definiciones aquí introducidas, los procesos involucrados se suponen regulares y estacionarios. Sin embargo, es posible considerar otro tipo de procesos de gran importancia en este trabajo: los *asintóticamente regulares y asintóticamente estacionarios* [5], que se tratan en el Apéndice A. Estos incluyen el transitorio y son de especial interés en el trabajo de Lepe Casillas sobre filtros y predictores causales. Por otra parte, es importante observar la estructura de las definiciones, ya que servirán de base para las correspondientes en dos dimensiones.

Capítulo 2

Aplicaciones de la Teoría de Información a Filtrado

2.1 Introducción.

En esta parte se presenta brevemente la teoría de Estimación Lineal, con aplicaciones a filtrado, en una dimensión, con mínima información de discriminación normalizada, desarrollada y detallada por Lepe-Casillas en [3]. En dicho trabajo se desarrollan tres criterios de filtrado empleando conceptos de Teoría de Información: Entropía Diferencial Mínima, Igualación de Espectro con fase según la potencia mínima del error, y el Filtro de Propósito Dual. En la sección 2.1 se presenta el primer criterio, el cual corresponde a la relación entre la minimización de la potencia del error y la minimización de la información de discriminación. En la sección 2.3 se expone el segundo criterio de estimación que trata la igualación del espectro, y en la 2.4 se presenta el tercer criterio, el cual da lugar al Filtro de Propósito Dual (FPD), y motivo principal de este trabajo. Las expresiones resultantes finales se comparan entre sí y con las del filtro de Wiener.

2.2 Primer Criterio de Filtrado basado en la Teoría de Información.

Considérese el sistema de la Fig.2.1, donde \mathbf{S} y \mathbf{X} son procesos reales, discretos, de media cero, regulares y con funciones de densidad espectral $f_s(e^{j\omega})$ y $f_x(e^{j\omega})$, respectivamente; $\mathbf{G}(e^{j\omega})$ es un SLIT causal y estable. El tiempo de aplicación n_0 de las entradas a los filtros determina el tipo de proceso a la salida: estacionario si $n_0 = -\infty$, o asintóticamente estacionario si $n_0 = 0$. Debido a que se toma el límite en las definiciones introducidas en el capítulo anterior, los resultados que se obtienen al emplear cualquiera de ambos valores, son equivalentes (ver Apéndice A) cuando ha cesado el transitorio. En adelante, por simplicidad, se supondrá que $n_0 = -\infty$, salvo en las ocasiones en que explícitamente se diga otra cosa.

Weidemann y Stear [13] tomaron como criterio que en el caso ideal, la información mutua normalizada entre los procesos \mathbf{E} y $\hat{\mathbf{S}}$ debe ser cero, $\bar{I}(\mathbf{E}, \hat{\mathbf{S}}) = 0$; esto significa que debe minimizarse de modo tal que \mathbf{E} ya no pueda proporcionar información para determinar el proceso

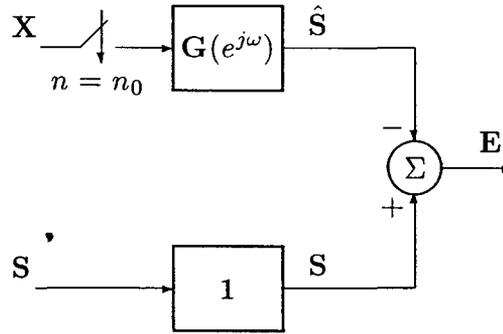


Figura 2.1: Sistema de Filtrado Lineal.

estimado \hat{S} . De la definición se tiene que

$$\bar{I}(\mathbf{E}, \hat{S}) = \bar{D}((\mathbf{E}, \hat{S}), (\mathbf{E} \times \hat{S})) ; \quad (2.1)$$

con el fin de tener una cota superior para esta cantidad en términos de estadísticas de segundo orden, se cambia $(\mathbf{E} \times \hat{S})$ por $(\mathbf{W} \times \hat{S})$, donde \mathbf{W} es un proceso gaussiano con f.d.p. $f_w(e^{j\omega})$ e independiente de \mathbf{X} y \mathbf{S} , llamado por Lepe Casillas *proceso de ponderación*. Así, al minimizar $\bar{D}((\mathbf{E}, \hat{S}), (\mathbf{W} \times \hat{S}))$ con respecto al filtro, se obliga a \mathbf{E} a ser tan gaussiano como sea posible al mismo tiempo que independiente de \mathbf{S} , es decir

$$\bar{D}((\mathbf{E}, \hat{S}), (\mathbf{W} \times \hat{S})) = \bar{D}(\mathbf{E}, \mathbf{W}) + \bar{I}(\mathbf{E}, \hat{S}) \geq \bar{I}(\mathbf{E}, \hat{S}) . \quad (2.2)$$

Puesto que el filtro $G(e^{j\omega})$ no genera información, o bien puede destruirla, se cumple que

$$\bar{I}(\mathbf{E}, \mathbf{X}) \geq \bar{I}(\mathbf{E}, \hat{S}) ; \quad (2.3)$$

con lo que se obtiene

$$\bar{D}((\mathbf{E}, \hat{S}), (\mathbf{W} \times \hat{S})) \leq \bar{D}((\mathbf{E}, \mathbf{X}), (\mathbf{W} \times \mathbf{X})) = \bar{D}(\mathbf{E}, \mathbf{W}) + \bar{I}(\mathbf{E}, \mathbf{X}) ; \quad (2.4)$$

con igualdad si \mathbf{X} y \mathbf{S} son conjuntamente gaussianos. Entonces

$$\begin{aligned} \bar{D}((\mathbf{E}, \mathbf{X}), (\mathbf{W} \times \mathbf{X})) &= \bar{D}(\mathbf{E}, \mathbf{W}) + \bar{I}(\mathbf{E}, \mathbf{X}) \\ &= -\bar{H}_s + \bar{I}(\mathbf{S}, \mathbf{X}) + \frac{1}{2} \ln(2\pi) + \frac{1}{2} \int_{-\pi}^{\pi} \ln f_w(e^{j\omega}) \frac{d\omega}{2\pi} + \frac{1}{2} \int_{-\pi}^{\pi} \frac{f_e(e^{j\omega})}{f_w(e^{j\omega})} \frac{d\omega}{2\pi} ; \end{aligned} \quad (2.5)$$

donde $f_e(e^{j\omega})$ es la función de densidad espectral de potencia del proceso de error \mathbf{E} ¹. Sólo el último término de la ecuación anterior depende del filtro, y al minimizar con respecto a éste, considerando **un filtro no causal**, se obtiene el filtro de Wiener no causal:

$$G_{nc}^W(e^{j\omega}) = \frac{f_{sx}(e^{j\omega})}{f_x(e^{j\omega})} , \quad (2.6)$$

¹Se denomina también espectro de densidad de potencia o simplemente espectro.

Por esto, la minimización de la expresión 2.5 es equivalente, en el caso no causal, a la minimización del valor medio del cuadrado del error empleada en el filtrado convencional; es decir, ambos enfoques de minimización proporcionan el mismo resultado. Según Lepe, este primer criterio es equivalente al criterio de mínimo valor medio del cuadrado del error, sólo si el filtro es acausal. Si el filtro se restringe a ser causal, el óptimo no es independiente de $f_w(e^{j\omega})$. Este problema se estudia en [7].

2.3 Segundo Criterio de Filtrado Basado en la Teoría de Información.

El filtro de igualación de espectro se ha estudiado y empleado con anterioridad [10, 11, 12]. Sin embargo, antes de [3], no se tiene conocimiento de que se le haya estudiado bajo la perspectiva de la Teoría de Información. En ella, se le ve entonces como la minimización de $d_{IS}(f_s(e^{j\omega}), f_x(e^{j\omega}))$ considerando filtrado no causal. El resultado, el cual no depende de la fase, se enriqueció con el empleo de la fase que minimiza el valor medio de la potencia del error ², y que abajo se indica. En resumen, la función de transferencia del filtro es

$$G^{SE}(e^{j\omega}) = \left(\frac{f_s(e^{j\omega})}{f_x(e^{j\omega})} \right)^{\frac{1}{2}} \exp(j a(e^{j\omega})) ; \quad (2.7)$$

donde $a(e^{j\omega})$ se propone sea

$$a(e^{j\omega}) = \tan^{-1} \left\{ \frac{\text{Im} [f_{sx}(e^{j\omega})]}{\text{Re} [f_{sx}(e^{j\omega})]} \right\} ; \quad (2.8)$$

así, el espectro de la salida del filtro queda

$$f_s^{SE}(e^{j\omega}) = f_s(e^{j\omega}) ; \quad (2.9)$$

y el espectro del proceso de error es

$$f_e^{SE}(e^{j\omega}) = f_s(e^{j\omega}) \{2(1 - |\gamma_{xs}(e^{j\omega})|)\} , \quad (2.10)$$

donde $\gamma_{xs}(e^{j\omega})$ es la *coherencia* entre los procesos X y S (se define posteriormente).

2.4 Tercer Criterio de Filtrado Basado en la Teoría de Información.

Lepe Casillas propuso adicionalmente dos modificaciones al planteamiento de las secciones anteriores. Por una parte, el proceso a estimar S se supone gaussiano; y por otra, se minimiza

²Cannon [11] no propone el empleo de una fase definida, sino por lo contrario, su propuesta es seleccionarla de un cierto grupo según convenga, y presenta algunos ejemplos y resultados. Por otra parte, la historia de este segundo criterio comienza, en palabras del propio Lepe Casillas, cuando L. A. Buzo, después de conocer los resultados del Tercer criterio, propuso escoger una fase tal que minimizara el valor medio de la potencia del error. Lepe resolvió el problema así planteado y encontró que dicha fase era la misma que la del filtro de Wiener, y que la del filtro obtenido con el Tercer Criterio.

ahora la razón de cambio de entropía cruzada del proceso bidimensional $(\mathbf{E}, \hat{\mathbf{S}})$ con respecto al proceso también bidimensional $(\mathbf{W} \times \mathbf{S})$ con el propósito de hacer a $\hat{\mathbf{S}}$ lo más parecido posible a \mathbf{S} . De las definiciones se tiene que

$$\bar{\mathbf{D}}((\mathbf{E}, \hat{\mathbf{S}}), (\mathbf{W} \times \mathbf{S})) = \bar{\mathbf{D}}(\mathbf{E}, \mathbf{W}) + \bar{\mathbf{I}}(\mathbf{E}, \hat{\mathbf{S}}) + \bar{\mathbf{D}}(\hat{\mathbf{S}}, \mathbf{S}). \quad (2.11)$$

Como puede apreciarse, esta expresión mide la razón de cambio de información mutua entre el proceso de error \mathbf{E} y el proceso estimado $\hat{\mathbf{S}}$, la razón de cambio de entropía cruzada de $\hat{\mathbf{S}}$ con respecto a \mathbf{S} y además, la razón de cambio entropía cruzada entre \mathbf{E} y el proceso de referencia \mathbf{W} . Considerando nuevamente que el filtro $\mathbf{G}(e^{j\omega})$ no genera información, entonces

$$\bar{\mathbf{D}}((\mathbf{E}, \hat{\mathbf{S}}), (\mathbf{W} \times \mathbf{S})) \leq \bar{\mathbf{D}}(\mathbf{E}, \mathbf{W}) + \bar{\mathbf{I}}(\mathbf{E}, \mathbf{X}) + \bar{\mathbf{D}}(\hat{\mathbf{S}}, \mathbf{S}); \quad (2.12)$$

$$\begin{aligned} \bar{\mathbf{D}}(\mathbf{E}, \mathbf{W}) + \bar{\mathbf{I}}(\mathbf{E}, \mathbf{X}) + \bar{\mathbf{D}}(\hat{\mathbf{S}}, \mathbf{S}) &= \bar{\mathbf{D}}((\mathbf{E}, \mathbf{X}), (\mathbf{W} \times \mathbf{X})) \\ &+ \frac{1}{2} d_{IS}(f_{\hat{s}}(e^{j\omega}), f_s(e^{j\omega})) - \bar{\mathbf{H}}_x + \frac{1}{2} \ln(2\pi) + \frac{1}{2} \int_{-\pi}^{\pi} \ln f_x(e^{j\omega}) \frac{d\omega}{2\pi}. \end{aligned} \quad (2.13)$$

Nuevamente, al minimizar 2.11 con respecto al filtro, se obliga simultáneamente al proceso de error \mathbf{E} a ser tan gaussiano como sea posible, al proceso estimado $\hat{\mathbf{S}}$ a ser lo más parecido posible a \mathbf{S} y al mismo tiempo a \mathbf{E} a ser lo más independiente de $\hat{\mathbf{S}}$. Además, si se minimiza con respecto al espectro de densidad de potencia del proceso de ponderación \mathbf{W} , $f_w(e^{j\omega})$, se llega a

$$f_w(e^{j\omega}) = f_e(e^{j\omega}). \quad (2.14)$$

Al sustituir 2.5 en 2.13 y considerando el resultado de 2.14, se observa que la minimización de 2.13 con respecto al filtro consiste simplemente en minimizar la suma del último término del lado derecho de 2.5 y de la distorsión de Itakura-Saito entre $f_{\hat{s}}(e^{j\omega})$ y $f_s(e^{j\omega})$, ya que son los únicos términos que dependen del filtro $\mathbf{G}(e^{j\omega})$. Así, al minimizar 2.13 con respecto al filtro, y al mismo tiempo con respecto al espectro del proceso de referencia $f_w(e^{j\omega})$ para tener un proceso de ponderación "óptimo", se obtienen los siguientes resultados [3] (suponiendo a \mathbf{X} gaussiano). El valor mínimo de 2.11 es

$$\bar{\mathbf{D}}((\mathbf{E}, \hat{\mathbf{S}}), (\mathbf{W} \times \mathbf{S})) = \frac{1}{2} \int_{-\pi}^{\pi} \left\{ \ln \left(\frac{\alpha(e^{j\omega}) - 1}{1 - \gamma_{xs}^2(e^{j\omega})} \right) + \frac{\alpha(e^{j\omega})}{2\alpha(e^{j\omega}) - 1} - \ln \left(\frac{\alpha(e^{j\omega})}{2\alpha(e^{j\omega}) - 1} \right) - 1 \right\} \frac{d\omega}{2\pi}, \quad (2.15)$$

donde

$$\gamma_{xs}^2(e^{j\omega}) = \frac{(\alpha^2(e^{j\omega}) - 3\alpha(e^{j\omega}) + 1)^2}{(2\alpha(e^{j\omega}) - 1) \alpha(e^{j\omega})}, \quad (2.16)$$

y

$$\alpha(e^{j\omega}) = 1 + \frac{f_w(e^{j\omega})}{f_s(e^{j\omega})}; \quad (2.17)$$

siempre que el filtro de propósito dual sea

$$\mathbf{G}^o(e^{j\omega}) = \left(\frac{f_s(e^{j\omega})}{f_x(e^{j\omega})} \right)^{\frac{1}{2}} \frac{f_{sx}(e^{j\omega})}{|f_{sx}(e^{j\omega})|} \left(\frac{\alpha(e^{j\omega})}{2\alpha(e^{j\omega}) - 1} \right)^{\frac{1}{2}}. \quad (2.18)$$

De esta ecuación se observa que el filtro de mínima razón de información de discriminación tiene la misma fase que el de Wiener. Empleando 2.17, este filtro toma la forma

$$\mathbf{G}^o(e^{j\omega}) = \frac{f_{sx}(e^{j\omega})}{2f_x(e^{j\omega})(1 + f_w^n(e^{j\omega}))} \left\{ 1 + \sqrt{1 + \frac{4f_x(e^{j\omega})f_s(e^{j\omega})}{|f_{sx}(e^{j\omega})|^2} f_w^n(e^{j\omega})(1 + f_w^n(e^{j\omega}))} \right\}, \quad (2.19)$$

donde

$$f_w^n(e^{j\omega}) = \alpha(e^{j\omega}) - 1 = \frac{f_w(e^{j\omega})}{f_s(e^{j\omega})}. \quad (2.20)$$

Con estos resultados, los valores "óptimos" de los espectros de los procesos de referencia, de error y estimado toman la forma

$$f_w^o(e^{j\omega}) = f_s(e^{j\omega}) \{ \alpha(e^{j\omega}) - 1 \}, \quad (2.21)$$

$$f_e^o(e^{j\omega}) = f_w^o(e^{j\omega}), \quad (2.22)$$

y

$$f_s^o(e^{j\omega}) = f_s(e^{j\omega}) \left\{ \frac{\alpha(e^{j\omega})}{2\alpha(e^{j\omega}) - 1} \right\}; \quad (2.23)$$

respectivamente. Para efectos de comparación, el filtro de Wiener no causal [3] tiene las siguientes características. La función de transferencia del filtro es

$$\mathbf{G}_{nc}^W(e^{j\omega}) = \frac{f_{sx}(e^{j\omega})}{f_x(e^{j\omega})}; \quad (2.24)$$

el espectro del proceso estimado tiene la forma

$$f_s^W(e^{j\omega}) = \gamma_{xs}^2(e^{j\omega}) f_s(e^{j\omega}), \quad (2.25)$$

donde $\gamma_{xs}^2(e^{j\omega})$ es la *coherencia* entre los procesos X y S dada por

$$\gamma_{xs}^2(e^{j\omega}) = \frac{|f_{sx}(e^{j\omega})|^2}{f_x(e^{j\omega})f_s(e^{j\omega})}; \quad (2.26)$$

y el espectro del proceso de error es

$$f_e^W(e^{j\omega}) = f_s(e^{j\omega}) \{ 1 - \gamma_{xs}^2(e^{j\omega}) \}. \quad (2.27)$$

Con estos resultados, puede verificarse que se cumple

$$\begin{aligned} \mathbf{G}^o(e^{j\omega}) &= \mathbf{G}^W(e^{j\omega}) \frac{1}{|\gamma_{xs}(e^{j\omega})|} \left(\frac{\alpha(e^{j\omega})}{2\alpha(e^{j\omega}) - 1} \right)^{\frac{1}{2}}, \\ &= \mathbf{G}^{SE}(e^{j\omega}) \left(\frac{\alpha(e^{j\omega})}{2\alpha(e^{j\omega}) - 1} \right)^{\frac{1}{2}}; \end{aligned} \quad (2.28)$$

$$\begin{aligned}
f_e^o(e^{j\omega}) &= f_e^W(e^{j\omega}) \frac{\alpha(e^{j\omega}) - 1}{1 - \gamma_{xs}^2(e^{j\omega})}, \\
&= f_e^{SE}(e^{j\omega}) \frac{\alpha(e^{j\omega}) - 1}{2(1 - |\gamma_{xs}(e^{j\omega})|)}; \tag{2.29}
\end{aligned}$$

$$\begin{aligned}
f_s^o(e^{j\omega}) &= f_s^W(e^{j\omega}) \frac{1}{\gamma_{xs}^2(e^{j\omega})} \left(\frac{\alpha(e^{j\omega})}{2\alpha(e^{j\omega}) - 1} \right), \\
&= f_s^{SE}(e^{j\omega}) \frac{\alpha(e^{j\omega})}{2\alpha(e^{j\omega}) - 1} \tag{2.30}
\end{aligned}$$

2.5 Observaciones Finales.

Se presenta en este capítulo la relación descubierta en [3] que existe entre la Teoría de Información y el filtrado clásico y las modificaciones propuestas en [3] para obtener un criterio diferente de filtrado. En este nuevo criterio se aprovecha la suposición de que la f.d.e.p. del proceso a estimar se conoce y es gaussiano. De las expresiones 2.28-2.30, Lepe [3] demuestra que el *filtro propósito dual* se comporta como el filtro de Wiener para valores grandes de la coherencia (cercaos a uno), mientras que para valores pequeños de ella (cercaos a cero) su comportamiento es más parecido al del filtro del método de igualación de espectro. Para valores intermedios, su desempeño se ubica entre ambos. Por este motivo, este filtro puede emplearse en lugar del de Wiener en problemas en que los valores de la coherencia sean pequeños, por ejemplo, en restauración de imágenes, por supuesto, habiendo obtenido primeramente las expresiones correspondientes en dos dimensiones.

Capítulo 3

Definiciones de Teoría de Información en dos Dimensiones.

3.1 Introducción.

En este capítulo se presentan las definiciones introducidas en el capítulo 2 para el caso bidimensional, empleando requerimientos básicos de campos aleatorios. En la sección 4.2 se definen algunos conceptos probabilísticos elementales de campos aleatorios reales. En la sección 4.3 se presentan las definiciones de razón de cambio de entropía cruzada, información mutua y entropía diferencial para dos dimensiones, mientras que en la sección 4.4 se derivan algunas relaciones entre ellas.

3.2 Campos Aleatorios.

Se presentan a continuación los elementos básicos de campos aleatorios.

3.2.1 Elementos Probabilísticos [6] .

Un *campo aleatorio o estocástico* real $X(\underline{t})$ puede verse como una extensión de un proceso estocástico en la cual a cada evento ζ del espacio de todos los posibles eventos, se le asigna una función $X(\underline{t}, \zeta)$ que depende de un vector de parámetros $\underline{t} = (t_1, t_2, \dots, t_n)$ de orden n y del evento ζ . Es esencialmente parecido a un proceso estocástico, con la diferencia de que el espacio de observación o localización es un vector n -dimensional \underline{t} en lugar de ser unidimensional como en los procesos estocásticos.

Puesto que el valor de un campo aleatorio en la localización o punto de observación constante \underline{t} es una variable aleatoria (v.a) $X = X(\underline{t})$, se define su función de distribución acumulada como

$$\mathcal{P}[X = X(\underline{t}) \leq x] = F(x, \underline{t}); \quad (3.1)$$

donde x es el valor de la realización $x(\underline{t})$ en la observación \underline{t} . Nótese que es una función que depende del conjunto de parámetros \underline{t} . La función de densidad de probabilidad de la v.a. X se

define como

$$p_x(x, \underline{t}) = \frac{\partial F(x, \underline{t})}{\partial x}; \quad (3.2)$$

de modo semejante se pueden definir las f.d.p conjunta para cualquier orden N . Generalmente, en el argumento se omite la dependencia explícita de $\underline{t}_1, \underline{t}_2, \dots, \underline{t}_m$, y se usa la forma simplificada ya conocida para procesos unidimensionales.

Se dice que un campo aleatorio es *homogéneo* si todas las f.d.p conjuntas de cualquier orden m no cambian cuando el conjunto de observaciones $\underline{t}_1, \underline{t}_2, \dots, \underline{t}_m$ experimenta un corrimiento del origen, es decir

$$p_x(x_1, \dots, x_m; \underline{t}_1, \dots, \underline{t}_m) = p_x(x_1, \dots, x_m; \underline{t}_1 + \underline{k}, \dots, \underline{t}_m + \underline{k}). \quad (3.3)$$

Cuando el campo es homogéneo, su media y su variancia son constantes y las funciones de correlación dependen sólo de la posición relativa de las observaciones \underline{t}_1 y \underline{t}_2 , con lo que se tiene que, considerando el campo con media cero

$$R(\underline{t}_1, \underline{t}_2) = E [X(\underline{t}_1)X(\underline{t}_2)] = R(\underline{t}_1 - \underline{t}_2) = R(\underline{\tau}); \quad (3.4)$$

donde $\underline{\tau} = \underline{t}_1 - \underline{t}_2$ es un vector de componentes $\tau_1 = t_{11} - t_{21}, \dots, \tau_n = t_{1n} - t_{2n}$. En particular, si el campo es bidimensional se cumple que

$$R(\underline{\tau}) = R(\tau_1, \tau_2) = R(-\tau_1, -\tau_2) \quad \tau_1, \tau_2 \geq 0; \quad (3.5)$$

y

$$R(-\tau_1, \tau_2) = R(\tau_1, -\tau_2) \quad \tau_1, \tau_2 \geq 0. \quad (3.6)$$

Para un campo gaussiano bidimensional, de media cero, su f.d.p. conjunta de orden N toma la forma simplificada (se omite la dependencia explícita de los parámetros de observación \underline{t})

$$p(\underline{x}) = \frac{1}{(2\pi)^{N/2} \Delta_R^{1/2}} \exp \left\{ -\frac{1}{2} (\underline{x}^T \underline{R}^{-1} \underline{x}) \right\}, \quad (3.7)$$

donde $\underline{x}^T = (X(\underline{t}_1), \dots, X(\underline{t}_N))$ es un vector de N variables aleatorias; \underline{R} es la matriz de correlaciones de orden N del campo $X(\underline{t})$ dada por

$$\underline{R} = E \{ \underline{x} \underline{x}^T \}, \quad (3.8)$$

y Δ_R es su determinante. En esta expresión, la estructura de \underline{R} depende de los puntos de observación o región de observación que conformen al vector \underline{x} , así como de los ordenamientos que se formen en dicha región (ver Apéndice B).

3.2.2 Relaciones Entrada-Salida para un SLIC.

Sea un Sistema Lineal Invariante al Corrimiento (SLIC) discreto, causal y estable (ver Apéndice B) con respuesta al impulso $h(\underline{t})$ y entrada un campo real discreto homogéneo $X(\underline{t})$. El campo aleatorio de salida está dado por

$$Y(\underline{t}) = \sum_{\forall \underline{\tau}} h(\underline{t}) X(\underline{t} - \underline{\tau}). \quad (3.9)$$

Al suponer homogéneo el campo de entrada, el campo de salida también es homogéneo con media constante dada por

$$m_y = m_x \sum_{\forall \tau} h(\tau); \quad (3.10)$$

y función de autocorrelación

$$R_y(\tau) = \sum_{\forall \alpha} \sum_{\forall \beta} h(\alpha)h(\beta)R_x(\tau + \alpha - \beta); \quad (3.11)$$

donde $R_x(\tau)$ es la función de autocorrelación del campo de entrada $X(\underline{t})$. Si la entrada es no correlacionada con variancia σ^2 (por ejemplo, ruido blanco), la ecuación anterior se reduce a

$$R_y(\tau) = \sigma^2 \sum_{\forall \alpha} h(\alpha)h(\tau + \alpha) = \sigma^2 h(\tau) * h(-\tau). \quad (3.12)$$

Para el análisis en el dominio de la frecuencia, el espectro de densidad de potencia del campo X , $S_x(\underline{e}^{j\omega})$, con $\underline{e}^{j\omega} = (e^{j\omega_1}, e^{j\omega_2}, \dots, e^{j\omega_n})$, se define como la Transformada de Fourier de su función de autocorrelación $R_x(\tau)$, es decir,

$$S_x(\underline{e}^{j\omega}) = \sum_{\forall \tau} R_x(\tau) e^{-j\omega \tau}. \quad (3.13)$$

En el caso bidimensional, la relación entre las funciones de densidad espectral de la entrada y la salida es

$$S_y(e^{j\omega_1}, e^{j\omega_2}) = |H(e^{j\omega_1}, e^{j\omega_2})|^2 S_x(e^{j\omega_1}, e^{j\omega_2}) \quad -\pi \leq \omega_1, \omega_2 \leq \pi, \quad (3.14)$$

donde $H(e^{j\omega_1}, e^{j\omega_2})$ es la Transformada de Fourier de $h(\underline{t})$.

3.3 Definiciones.

Con los elementos de la teoría de la información, se repiten las definiciones de entropía diferencial, información mutua e información de discriminación para campos bidimensionales (se cumplen también para cualquier dimensión), sean estructuralmente iguales a las introducidas en el Capítulo 1, de hecho son las mismas excepto por el argumento: ahora cada variable aleatoria depende de \underline{t} en lugar de t . Para evitar posibles confusiones, se ha cambiado ligeramente la notación, aunque lo mejor es tener presente que ahora el espacio de observación es n -dimensional. Se supone que los campos son reales, discretos, de media cero y cuyas funciones de densidad de probabilidad conjuntas existen para cualquier orden N . Por simplicidad, se omite en los argumentos de las funciones de densidad conjuntas, la dependencia explícita de los vectores $\underline{t}_1, \dots, \underline{t}_N$ y del orden N .

3.3.1 Entropía Diferencial Normalizada.

La Entropía Diferencial Normalizada del campo $X(\underline{t})$ se define como

$$\bar{H}_X = \lim_{N \rightarrow \infty} \frac{1}{N} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} p_X(\underline{u}) \ln \left\{ \frac{1}{p_X(\underline{u})} \right\} \delta_u^N; \quad (3.15)$$

donde $\delta_u^N = du_1 du_2 \cdots du_N$ es un elemento diferencial en R^N , $\mathbf{u} = (u_1, u_2, \dots, u_N)^T$ es un vector mudo de N elementos también en R^N , y $p_X(\mathbf{u})$ es la función de densidad de probabilidad (f.d.p.) conjunta de orden N del campo $X(\mathbf{t})$.

3.3.2 Información de Discriminación Normalizada.

La Información de Discriminación Normalizada del campo $X(\mathbf{t}) \equiv X$ con respecto al campo $Y(\mathbf{t}) \equiv Y$ está dada por

$$\bar{D}(X, Y) = \lim_{N \rightarrow \infty} \frac{1}{N} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_X(\mathbf{u}) \ln \left\{ \frac{p_X(\mathbf{u})}{p_Y(\mathbf{u})} \right\} \delta_u^N ; \quad (3.16)$$

donde $p_Y(\mathbf{u})$ es la f.d.p. conjunta de orden N del campo Y .

3.3.3 Información Mutua Normalizada.

La Información Mutua Normalizada entre los campos X y Y se define por

$$\bar{I}(X, Y) = \lim_{N \rightarrow \infty} \frac{1}{N} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_{XY}(\mathbf{u}, \mathbf{v}) \ln \left\{ \frac{p_{XY}(\mathbf{u}, \mathbf{v})}{p_X(\mathbf{u}) p_Y(\mathbf{v})} \right\} \delta_u^N \delta_v^N ; \quad (3.17)$$

donde $p_{XY}(\mathbf{u}, \mathbf{v})$ es la f.d.p. conjunta de orden $2N$ de los campos X y Y . Nuevamente se puede ver como un caso especial de la Información de Discriminación Normalizada si se consideran los campos bivariados $Z = (X, Y)$ y $Z^\times = (X \times Y)$ con f.d.p. de orden $2N$ dadas por $p_Z(\mathbf{u}, \mathbf{v}) = p_{XY}(\mathbf{u}, \mathbf{v})$ y $p_{Z^\times}(\mathbf{u}, \mathbf{v}) = p_X(\mathbf{u}) p_Y(\mathbf{v})$ respectivamente. Así, se cumple que

$$\bar{D}(Z, Z^\times) = \bar{I}(X, Y) = \bar{I}(Y, X) . \quad (3.18)$$

3.4 Algunos Resultados Importantes.

1) Sea un SLIC bidimensional causal y estable, con respuesta al impulso $h(m, n)$ e inversa única, definida en una región causal, condiciones iniciales nulas, con entrada un campo X homogéneo aplicado en la misma región; la salida es un campo asintóticamente homogéneo Y (ver Anexo B.1), cuya entropía diferencial normalizada es

$$\bar{H}_Y = \bar{H}_X + \frac{1}{2} \ln |h_{0,0}|^2 \quad h_{0,0} \neq 0 . \quad (3.19)$$

En general, el segundo término del lado derecho, es una constante que sólo depende de $h(m, n)$, y en particular tiene este valor sólo cuando X y $h(m, n)$ son causales. Si este mismo campo de entrada X se aplica desde $m = -\infty$, y $n = -\infty$, se tiene entonces que Y es un campo regular y homogéneo y se cumple que (ver parte final del Apéndice B)

$$\bar{H}_Y = \bar{H}_X + \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln |H(e^{j\omega_1}, e^{j\omega_2})|^2 \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} . \quad (3.20)$$

En una dimensión, estas 2 últimas expresiones son iguales si el sistema es de fase mínima. El autor en este trabajo, no afirma ni niega que 3.19 y 3.20 sean iguales cuando $h(m, n)$ sea de fase mínima, sin embargo, es probable que así ocurra si las características de las matrices de correlaciones para el unidimensional, se conservan para el caso bidimensional. Cuando Z es un campo gaussiano, con función de densidad espectral de potencia $f_Z(e^{j\omega_1}, e^{j\omega_2})$, entonces (ver Apéndice B)

$$\bar{H}_Z = \frac{1}{2} \ln(2\pi e) + \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln f_Z(e^{j\omega_1}, e^{j\omega_2}) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}. \quad (3.21)$$

2) Sean X y Y campos homogéneos y regulares gaussianos, de media cero; la información de discriminación normalizada de X con respecto a Y está dada por

$$\bar{D}(X, Y) = \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left(\frac{f_X(e^{j\omega_1}, e^{j\omega_2})}{f_Y(e^{j\omega_1}, e^{j\omega_2})} - \ln \frac{f_X(e^{j\omega_1}, e^{j\omega_2})}{f_Y(e^{j\omega_1}, e^{j\omega_2})} - 1 \right) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}; \quad (3.22)$$

es decir, es igual a 1/2 de la *Distorsión de Itakura-Saito* bidimensional [9].

3) Sean X , Y y $Z = X \pm Y$ campos con entropías diferenciales normalizadas definidas individualmente y para cualquier par de ellos. Se cumple que [1],[13]

$$\begin{aligned} \bar{H}_Z &= \bar{I}(X, Z) + \bar{H}_Y - \bar{I}(X, Y); \\ \bar{H}_Z &= \bar{I}(Y, Z) + \bar{H}_X - \bar{I}(X, Y). \end{aligned} \quad (3.23)$$

Las siguientes dos relaciones para el campo X se demuestran en el Apéndice B.

4) Sea $\underline{\mathbf{R}}_x$ la matriz de correlaciones de orden N del campo X y $\underline{\mathbf{P}}$ una *matriz de permutación* del mismo orden que $\underline{\mathbf{R}}_x$. Se cumple que

$$\Delta_{R_x} = \Delta_P \Delta_{R_x} \Delta_{P^T} \quad (3.24)$$

y

$$\text{tr}\{\underline{\mathbf{R}}_x\} = \text{tr}\{\underline{\mathbf{P}}\underline{\mathbf{R}}_x\underline{\mathbf{P}}^T\}; \quad (3.25)$$

donde Δ_{R_x} , Δ_P y Δ_{P^T} , son los determinantes de $\underline{\mathbf{R}}_x$, $\underline{\mathbf{P}}$ y $\underline{\mathbf{P}}^T$, respectivamente, y $\text{tr}\{\}$ indica la traza. Lo anterior significa que ni el determinante ni la traza de una matriz de correlaciones se afectan si se premultiplican y postmultiplican por una matriz de permutación y su transpuesta, respectivamente, o viceversa. -

5) La relación entre el determinante de $\underline{\mathbf{R}}_x$ y la función de densidad espectral de potencia de X , $f_X(e^{j\omega_1}, e^{j\omega_2})$ puede escribirse como

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln \Delta_{R_x} = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln f_X(e^{j\omega_1}, e^{j\omega_2}) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}. \quad (3.26)$$

Capítulo 4

Extensión a Dos Dimensiones del Criterio de Lepe-Casillas para Filtrado.

4.1 Introducción.

En este capítulo se presenta la extensión del criterio reseñado en el Capítulo 3 al caso bidimensional. En la sección 5.2 se derivan las expresiones más importantes y se obtiene la expresión final a minimizar considerando nuevamente dos casos: \hat{S} regular y homogéneo y cuando sólo lo es asintóticamente, considerando en ambas condiciones iniciales nulas. En la sección 5.3 se presenta la expresión para el filtro resultante de la minimización y se compara con las expresiones tanto del filtro de Wiener no causal como con la del filtro del Método de Igualación de Espectros. Finalmente, en la sección 5.4 se comparan los resultados con los de Wiener y los del método de igualación de espectro.

4.2 Resultados de la Extensión.

De acuerdo a lo presentado en el Capítulo 2, en este criterio de filtrado se requiere la minimización de la cota superior ¹

$$\bar{D}((E, \hat{S}), (W \times S)) \leq \bar{D}(E, W) + \bar{I}(E, X) + \bar{D}(\hat{S}, S); \quad (4.1)$$

donde W es un campo gaussiano regular y homogéneo de media cero llamado de referencia, y los demás campos se relacionan como se indica en la Fig.4.1. De [3] se tiene que

$$\bar{D}(E, W) + \bar{I}(E, X) = -\bar{H}_E + \lim_{N \rightarrow \infty} \frac{1}{N} \ln(2\pi^{N/2} \Delta_W^{1/2}) + \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \{ \underline{\mathbf{R}}_E \underline{\mathbf{R}}_W^{-1} \} + \bar{I}(E, X); \quad (4.2)$$

¹Recuérdese que en esta expresión se ha sustituido $\bar{I}(E, \hat{S})$ por $\bar{I}(E, X)$, y como $\bar{I}(E, X) \geq \bar{I}(E, \hat{S})$, se produce la cota 4.1.

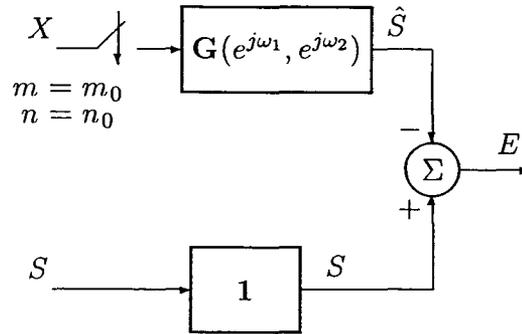


Figura 4.1: Sistema de Filtrado Bidimensional

donde Δ_W es el determinante de la matriz de correlaciones de orden N del campo W , y \mathbf{R}_E y \mathbf{R}_W son las matrices de correlaciones del mismo orden de los campos E y W , respectivamente. Esta expresión se reduce a [3]

$$\begin{aligned} \overline{\mathbf{D}}(E, W) + \overline{\mathbf{I}}(E, X) &= \overline{\mathbf{D}}((E, X), (W \times X)) \\ &= -\overline{\mathbf{H}}_S + \overline{\mathbf{I}}(S, X) + \frac{1}{2} \ln(2\pi) + \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln f_W(e^{j\omega_1}, e^{j\omega_2}) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \\ &\quad + \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{f_E(e^{j\omega_1}, e^{j\omega_2})}{f_W(e^{j\omega_1}, e^{j\omega_2})} \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}. \end{aligned} \quad (4.3)$$

Con este resultado, es posible obtener la expresión final a minimizar considerando las dos situaciones siguientes²:

a) \hat{S} regular y homogéneo. Se tiene en esta situación que [3]

$$\begin{aligned} \overline{\mathbf{D}}(E, W) + \overline{\mathbf{I}}(E, X) + \overline{\mathbf{D}}(\hat{S}, S) &= \overline{\mathbf{D}}((E, X), (W \times X)) - \overline{\mathbf{H}}_X + \frac{1}{2} \ln(2\pi e) \\ &\quad + \frac{1}{2} d_{IS}(f_{\hat{S}}(e^{j\omega_1}, e^{j\omega_2}), f_S(e^{j\omega_1}, e^{j\omega_2})) + \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln f_X(e^{j\omega_1}, e^{j\omega_2}) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}. \end{aligned} \quad (4.4)$$

a) \hat{S} asintóticamente regular y asintóticamente homogéneo. Considerando condiciones iniciales nulas, se tiene que

$$\begin{aligned} \overline{\mathbf{D}}(E, W) + \overline{\mathbf{I}}(E, X) + \overline{\mathbf{D}}(\hat{S}, S) &= \overline{\mathbf{D}}((E, X), (W \times X)) - \overline{\mathbf{H}}_X + \frac{1}{2} \ln(2\pi e) \\ &\quad + \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left(\frac{f_{\hat{S}}(e^{j\omega_1}, e^{j\omega_2})}{f_S(e^{j\omega_1}, e^{j\omega_2})} - \ln \frac{|g_{0,0}|^2 f_X(e^{j\omega_1}, e^{j\omega_2})}{f_S(e^{j\omega_1}, e^{j\omega_2})} - 1 \right) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} + \\ &\quad \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln f_X(e^{j\omega_1}, e^{j\omega_2}) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}, \end{aligned} \quad (4.5)$$

²Ver en el Apéndice B las demostraciones definiciones y conceptos relacionados.

donde $g_{0,0}$ es la respuesta al impulso evaluada en $(0,0)$ del filtro $G(e^{j\omega_1}, e^{j\omega_2})$. De la sección 3.4 del capítulo 3, si $g(m, n)$ es una transformación causal y estable con inversa única también causal y estable, ambas expresiones son iguales.

4.3 Resultados de la optimización.

La sección anterior muestra que la estructura de las ecuaciones obtenidas es igual a la de las ecuaciones unidimensionales del Capítulo 2, salvo por la dimensión del argumento de las funciones. De hecho, para el proceso de minimización, dicho argumento es irrelevante ya que la minimización es con respecto a las funciones. Por tanto, los resultados de [3] reseñados en el mismo Capítulo 2, se mantienen para dos dimensiones, es decir, el valor mínimo de 4.3 es

$$\begin{aligned} \bar{D}((E, X), (W \times X)) = & \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left\{ \ln \left(\frac{\alpha(e^{j\omega_1}, e^{j\omega_2}) - 1}{1 - \gamma_{XS}^2(e^{j\omega_1}, e^{j\omega_2})} \right) + \frac{\alpha(e^{j\omega_1}, e^{j\omega_2})}{2\alpha(e^{j\omega_1}, e^{j\omega_2}) - 1} \right. \\ & \left. - \ln \left(\frac{\alpha(e^{j\omega_1}, e^{j\omega_2})}{2\alpha(e^{j\omega_1}, e^{j\omega_2}) - 1} \right) - 1 \right\} \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}, \end{aligned} \quad (4.6)$$

donde la *coherencia* es ³

$$\gamma_{XS}^2(e^{j\omega_1}, e^{j\omega_2}) = \frac{(\alpha^2(e^{j\omega_1}, e^{j\omega_2}) - 3\alpha(e^{j\omega_1}, e^{j\omega_2}) + 1)^2}{(2\alpha(e^{j\omega_1}, e^{j\omega_2}) - 1)\alpha(e^{j\omega_1}, e^{j\omega_2})}, \quad (4.7)$$

y

$$\alpha(e^{j\omega_1}, e^{j\omega_2}) = 1 + \frac{f_W(e^{j\omega_1}, e^{j\omega_2})}{f_S(e^{j\omega_1}, e^{j\omega_2})}. \quad (4.8)$$

El filtro de *mínima razón de información de discriminación de propósito dual* es

$$G^o(e^{j\omega_1}, e^{j\omega_2}) = \left(\frac{J_S(e^{j\omega_1}, e^{j\omega_2})}{f_X(e^{j\omega_1}, e^{j\omega_2})} \right)^{\frac{1}{2}} \frac{J_{SX}(e^{j\omega_1}, e^{j\omega_2})}{|f_{SX}(e^{j\omega_1}, e^{j\omega_2})|} \left(\frac{\alpha(e^{j\omega_1}, e^{j\omega_2})}{2\alpha(e^{j\omega_1}, e^{j\omega_2}) - 1} \right)^{\frac{1}{2}}. \quad (4.9)$$

Nótese que, al igual que en una dimensión, este filtro tiene la misma fase que el de Wiener. Empleando 4.8 junto 4.7 con se tiene que

$$\begin{aligned} G^o(e^{j\omega_1}, e^{j\omega_2}) = & \frac{f_{SX}(e^{j\omega_1}, e^{j\omega_2})}{2f_X(e^{j\omega_1}, e^{j\omega_2})(1 + f_W^n(e^{j\omega_1}, e^{j\omega_2}))} \\ & \left\{ 1 + \sqrt{1 + \frac{4f_X(e^{j\omega_1}, e^{j\omega_2})f_S(e^{j\omega_1}, e^{j\omega_2})}{|f_{SX}(e^{j\omega_1}, e^{j\omega_2})|^2} f_W^n(e^{j\omega_1}, e^{j\omega_2})(1 + f_W^n(e^{j\omega_1}, e^{j\omega_2}))} \right\} \end{aligned} \quad (4.10)$$

donde

$$f_W^n(e^{j\omega_1}, e^{j\omega_2}) = \alpha(e^{j\omega_1}, e^{j\omega_2}) - 1 = \frac{f_W(e^{j\omega_1}, e^{j\omega_2})}{f_S(e^{j\omega_1}, e^{j\omega_2})}. \quad (4.11)$$

³Más adelante se da una definición formal de $\gamma_{XS}^2(e^{j\omega_1}, e^{j\omega_2})$.

Los valores de los espectros correspondientes al filtro óptimo (según el 3er. criterio) $G^o(e^{j\omega_1}, e^{j\omega_2})$, de los campos de referencia, de error y estimado toman la forma

$$f_W^o(e^{j\omega_1}, e^{j\omega_2}) = f_S(e^{j\omega_1}, e^{j\omega_2}) \{ \alpha(e^{j\omega_1}, e^{j\omega_2}) - 1 \} , \quad (4.12)$$

$$f_E^o(e^{j\omega_1}, e^{j\omega_2}) = f_W^o(e^{j\omega_1}, e^{j\omega_2}) , \quad (4.13)$$

y

$$f_{\hat{S}}^o(e^{j\omega_1}, e^{j\omega_2}) = f_S(e^{j\omega_1}, e^{j\omega_2}) \left\{ \frac{\alpha(e^{j\omega_1}, e^{j\omega_2})}{2\alpha(e^{j\omega_1}, e^{j\omega_2}) - 1} \right\} , \quad (4.14)$$

respectivamente. Procediendo de manera análoga a lo indicado en el capítulo 2 y para efectos de comparación, el filtro óptimo bidimensional no causal debido a Wiener es

$$G^W(e^{j\omega_1}, e^{j\omega_2}) = \frac{f_{SX}(e^{j\omega_1}, e^{j\omega_2})}{f_X(e^{j\omega_1}, e^{j\omega_2})} , \quad (4.15)$$

mientras que la f.d.e.p. del campo de salida cuando se usa este filtro tiene la forma

$$f_{\hat{S}}^W(e^{j\omega_1}, e^{j\omega_2}) = \gamma_{XS}^2(e^{j\omega_1}, e^{j\omega_2}) f_S(e^{j\omega_1}, e^{j\omega_2}) , \quad (4.16)$$

donde $\gamma_{XS}^2(e^{j\omega_1}, e^{j\omega_2})$ es la coherencia entre los campos X y S dada por

$$\gamma_{XS}^2(e^{j\omega_1}, e^{j\omega_2}) = \frac{|f_{SX}(e^{j\omega_1}, e^{j\omega_2})|^2}{f_S(e^{j\omega_1}, e^{j\omega_2}) f_X(e^{j\omega_1}, e^{j\omega_2})} , \quad (4.17)$$

y la f.d.e.p. del campo de error cuando se usa este mismo filtro es

$$f_E^W(e^{j\omega_1}, e^{j\omega_2}) = f_S(e^{j\omega_1}, e^{j\omega_2}) \{ 1 - \gamma_{XS}^2(e^{j\omega_1}, e^{j\omega_2}) \} . \quad (4.18)$$

Las relaciones correspondientes al método de igualación de espectro en dos dimensiones son: La expresión para el filtro es

$$G^{SE}(e^{j\omega_1}, e^{j\omega_2}) = \left(\frac{f_S(e^{j\omega_1}, e^{j\omega_2})}{f_X(e^{j\omega_1}, e^{j\omega_2})} \right)^{\frac{1}{2}} e^{ja(e^{j\omega_1}, e^{j\omega_2})} , \quad (4.19)$$

donde

$$a(e^{j\omega_1}, e^{j\omega_2}) = \tan^{-1} \left\{ \frac{Im[f_{SX}(e^{j\omega_1}, e^{j\omega_2})]}{Re[f_{SX}(e^{j\omega_1}, e^{j\omega_2})]} \right\} ; \quad (4.20)$$

la f.d.e.p. de la salida toma la forma

$$f_{\hat{S}}^{SE}(e^{j\omega_1}, e^{j\omega_2}) = f_S(e^{j\omega_1}, e^{j\omega_2}) , \quad (4.21)$$

y la f.d.e.p. del campo de error es

$$f_E^{SE}(e^{j\omega_1}, e^{j\omega_2}) = f_S(e^{j\omega_1}, e^{j\omega_2}) \{ 2(1 - |\gamma_{XS}(e^{j\omega_1}, e^{j\omega_2})|) \} . \quad (4.22)$$

Con lo anterior, se cumplen las relaciones siguientes,

$$G^o(e^{j\omega_1}, e^{j\omega_2}) = G^W(e^{j\omega_1}, e^{j\omega_2}) \frac{1}{|\gamma_{XS}(e^{j\omega_1}, e^{j\omega_2})|} \left(\frac{\alpha(e^{j\omega_1}, e^{j\omega_2})}{2\alpha(e^{j\omega_1}, e^{j\omega_2}) - 1} \right)^{\frac{1}{2}}, \quad (4.23)$$

$$= G^{SE}(e^{j\omega_1}, e^{j\omega_2}) \left(\frac{\alpha(e^{j\omega_1}, e^{j\omega_2})}{2\alpha(e^{j\omega_1}, e^{j\omega_2}) - 1} \right)^{\frac{1}{2}}, \quad (4.24)$$

$$f_E^o(e^{j\omega_1}, e^{j\omega_2}) = f_E^W(e^{j\omega_1}, e^{j\omega_2}) \frac{\alpha(e^{j\omega_1}, e^{j\omega_2}) - 1}{1 - \gamma_{XS}^2(e^{j\omega_1}, e^{j\omega_2})}, \quad (4.25)$$

$$= f_E^{SE}(e^{j\omega_1}, e^{j\omega_2}) \frac{\alpha(e^{j\omega_1}, e^{j\omega_2}) - 1}{2(1 - |\gamma_{XS}(e^{j\omega_1}, e^{j\omega_2})|)}, \quad (4.26)$$

$$f_S^o(e^{j\omega_1}, e^{j\omega_2}) = f_S^W(e^{j\omega_1}, e^{j\omega_2}) \frac{1}{\gamma_{XS}^2(e^{j\omega_1}, e^{j\omega_2})} \left(\frac{\alpha(e^{j\omega_1}, e^{j\omega_2})}{2\alpha(e^{j\omega_1}, e^{j\omega_2}) - 1} \right), \quad (4.27)$$

$$= f_S^{SE}(e^{j\omega_1}, e^{j\omega_2}) \frac{\alpha(e^{j\omega_1}, e^{j\omega_2})}{2\alpha(e^{j\omega_1}, e^{j\omega_2}) - 1} \quad (4.28)$$

4.4 Comentarios Finales.

Los resultados anteriores indican que la expresión para el *filtro de mínima razón de información de discriminación de propósito dual* en dos dimensiones, es la misma, salvo por el argumento de las funciones de densidad espectral de potencia, que la de una dimensión. Por ello, es de esperarse que ambos filtros se comporten en la misma forma. No obstante, se considera que la evaluación experimental del desempeño de estos 3 filtros es más ilustrativa en dos dimensiones, i.e. procesamiento de imágenes, en virtud de

- el ojo humano puede ser mejor discriminador que el oído,
- por tratarse de filtros no causales, es mejor aplicarlos a imágenes, donde la causalidad no es relevante, y comparar la calidad de las reconstrucciones;
- para una presentación escrita, son mejores las imágenes que las gráficas unidimensionales para evaluación práctica de los desempeños.

La conclusión más importante es que, salvo pequeñas modificaciones en los argumentos de las funciones, los resultados de [3] pueden aplicarse directamente a dos dimensiones.

Capítulo 5

Aplicaciones a Procesamiento de Imágenes

5.1 Introducción

Es este capítulo se presenta la aplicación del *filtro de mínima razón de información de discriminación* de Lepe-Casillas al procesamiento de imágenes, específicamente en restauración o reconstrucción. En la sección 5.2 se exponen algunos aspectos numéricos básicos sobre las expresiones del filtro y su implantación para simulación por computadora, así como otras expresiones relacionadas. En la sección 5.3 se resumen los algoritmos empleados en esta simulación y su codificación en lenguaje C. Por último, en la sección 5.4 se presentan los resultados obtenidos y se exponen las conclusiones relevantes.

5.2 Aspectos Numéricos

En este trabajo se tomó la decisión de construir todas las partes necesarias para la simulación por computadora de la aplicación del filtro en la restauración de imágenes, con el propósito de tener el mejor control posible de los aspectos numéricos y poder hacer la mejor evaluación posible. Las partes esenciales creadas para esta simulación son:

1. Un programa exhibidor de imágenes en pantalla de computadora.
2. Un programa para difuminar la imagen y agregarle ruido.
3. Un programa para restaurar o reconstruir la imagen empleando el filtro de Wiener, el del método de igualación de espectros con fase optimizada y el de mínima razón de información de discriminación de propósito dual.

El tipo de imagen a la que se aplicarán los métodos de restauración, debe tener una degradación producida por **difuminación** y por **ruido blanco normal aditivo**. Por cuestiones de simplicidad y rapidez en el cálculo numérico, todas las operaciones se realizan en el dominio de la frecuencia.

5.2.1 Imágenes y Matrices

Para este trabajo se emplearon imágenes monocromáticas de 256 por 256 píxeles, asignando a cada píxel 8 bits sin signo. Lo anterior hace que se tengan 256 niveles, en este caso de grises, del 0 (negro) a 255 (blanco). Cada imagen se maneja como matriz, es decir, a cada píxel se le asocia un elemento de la matriz de modo que al píxel que se encuentra en el extremo superior izquierdo de la imagen, le corresponde el primer elemento del arreglo matricial, mientras que al píxel que se encuentra en el extremo inferior derecho le corresponde el último. Debe mencionarse en esta parte que la **resolución** es un parámetro que se asocia el medio exhibidor y que no tiene que ver con el tamaño en píxeles de la imagen. Por supuesto, a mayor resolución en el medio exhibidor (puntos por pulgada cuadrada en impresiones y número de píxeles en la dirección horizontal y vertical en un monitor, por ejemplo), mejor claridad visual. Para este trabajo se escogió el tamaño mencionado antes de requerir ejecutar operaciones con operandos contenidos en la memoria expandida.

5.2.2 La expresiones de los filtros

Por conveniencia, se reproducen las ecuaciones presentadas en el capítulo anterior:

- **Filtro de Wiener.** La expresión para este filtro es

$$G^W(e^{j\omega_1}, e^{j\omega_2}) = \frac{f_{SX}(e^{j\omega_1}, e^{j\omega_2})}{f_X(e^{j\omega_1}, e^{j\omega_2})} \quad (5.1)$$

- **Filtro de mínima razón de de información de discriminación de propósito dual.** Este filtro toma la forma

$$G^o(e^{j\omega_1}, e^{j\omega_2}) = \frac{f_{SX}(e^{j\omega_1}, e^{j\omega_2})}{\left\{ 1 + \sqrt{1 + \frac{4}{\gamma^2(e^{j\omega_1}, e^{j\omega_2})} f_W^n(e^{j\omega_1}, e^{j\omega_2})(1 + f_W^n(e^{j\omega_1}, e^{j\omega_2}))} \right\}} \quad (5.2)$$

con

$$f_W^n(e^{j\omega_1}, e^{j\omega_2}) = \alpha(e^{j\omega_1}, e^{j\omega_2}) - 1 = \frac{f_W(e^{j\omega_1}, e^{j\omega_2})}{f_S(e^{j\omega_1}, e^{j\omega_2})}, \quad (5.3)$$

y

$$\gamma_{XS}^2(e^{j\omega_1}, e^{j\omega_2}) = \frac{|f_{SX}(e^{j\omega_1}, e^{j\omega_2})|^2}{f_S(e^{j\omega_1}, e^{j\omega_2}) f_X(e^{j\omega_1}, e^{j\omega_2})}, \quad (5.4)$$

que es la *coherencia normalizada entre los campos X y S*.

- **Filtro de Igualación de Espectro.** Este filtro se puede expresar como

$$G^{SE}(e^{j\omega_1}, e^{j\omega_2}) = \left(\frac{f_S(e^{j\omega_1}, e^{j\omega_2})}{f_X(e^{j\omega_1}, e^{j\omega_2})} \right)^{\frac{1}{2}} e^{ja(e^{j\omega_1}, e^{j\omega_2})}, \quad (5.5)$$

donde

$$a(e^{j\omega_1}, e^{j\omega_2}) = \tan^{-1} \left\{ \frac{\text{Im}[f_{SX}(e^{j\omega_1}, e^{j\omega_2})]}{\text{Re}[f_{SX}(e^{j\omega_1}, e^{j\omega_2})]} \right\}, \quad (5.6)$$

con esto, la fase de este filtro es la misma que la del filtro de Wiener.

Estas expresiones son las que se emplearon en los cálculos numéricos. Todos los espectros de los campos se calculan multiplicando la Transformada de Fourier del campo por su conjugado complejo, de acuerdo con la siguiente relación. Para los efectos numéricos de este trabajo, el espectro de densidad de potencia $f_X(e^{j\omega_1}, e^{j\omega_2})$ del campo X se calcula como:

$$f_X(e^{j\omega_1}, e^{j\omega_2}) = X(e^{j\omega_1}, e^{j\omega_2}) X^*(e^{j\omega_1}, e^{j\omega_2}) = |X(e^{j\omega_1}, e^{j\omega_2})|^2. \quad (5.7)$$

donde $X(e^{j\omega_1}, e^{j\omega_2})$ es la transformada de Fourier del campo X . A todos los campos se les resta el valor medio antes de efectuar las operaciones, y se restituye al finalizar estas. A partir del Filtro de Wiener se calculan los dos filtros restantes. Como se tratan imágenes (campos) degradadas por difuminación y ruido aditivo, 5.1 se expresa [8, pag. 278]

$$G^W(e^{j\omega_1}, e^{j\omega_2}) = \frac{H^*(e^{j\omega_1}, e^{j\omega_2}) f_S(e^{j\omega_1}, e^{j\omega_2})}{|H(e^{j\omega_1}, e^{j\omega_2})|^2 f_S(e^{j\omega_1}, e^{j\omega_2}) + f_N(e^{j\omega_1}, e^{j\omega_2})} = \frac{f_{SX}(e^{j\omega_1}, e^{j\omega_2})}{f_X}, \quad (5.8)$$

donde $H(e^{j\omega_1}, e^{j\omega_2})$ es el sistema difuminador y $f_N(e^{j\omega_1}, e^{j\omega_2})$ es el espectro del ruido aditivo blanco normal. Para efectos de cálculo, este puede manejarse como constante. La expresión 5.8 es compleja, y su fase se calcula simplemente como

$$\phi_W(e^{j\omega_1}, e^{j\omega_2}) = \tan^{-1} \left\{ \frac{\text{Im}[-H(e^{j\omega_1}, e^{j\omega_2})]}{\text{Re}[H(e^{j\omega_1}, e^{j\omega_2})]} \right\}. \quad (5.9)$$

Nótese que de este modo, la fase del filtro de Wiener y, por tanto, la del filtro de igualación de espectro, quedan determinadas por la fase del filtro difuminador $H(e^{j\omega_1}, e^{j\omega_2})$ ¹. A partir de 5.8 y de 5.9, se tiene que

$$\frac{f_S(e^{j\omega_1}, e^{j\omega_2})}{f_X(e^{j\omega_1}, e^{j\omega_2})} = \frac{f_S(e^{j\omega_1}, e^{j\omega_2})}{|H(e^{j\omega_1}, e^{j\omega_2})|^2 f_S(e^{j\omega_1}, e^{j\omega_2}) + f_N(e^{j\omega_1}, e^{j\omega_2})}, \quad (5.10)$$

con lo que 5.5 toma la forma

$$G^{SE}(e^{j\omega_1}, e^{j\omega_2}) = \left(\frac{f_S(e^{j\omega_1}, e^{j\omega_2})}{|H(e^{j\omega_1}, e^{j\omega_2})|^2 f_S(e^{j\omega_1}, e^{j\omega_2}) + f_N(e^{j\omega_1}, e^{j\omega_2})} \right)^{\frac{1}{2}} e^{j\phi_W(e^{j\omega_1}, e^{j\omega_2})}, \quad (5.11)$$

y la exponencial compleja se expresa como $\cos(\phi_W) + j \sin(\phi_W)$ para el cálculo numérico². Para el filtro de propósito dual, es necesario calcular dos valores adicionalmente, la coherencia entre

¹Recuérdese que la fase del filtro de mínima razón de información de discriminación de propósito dual, tiene la misma fase que el de Wiener.

²Por conveniencia se omite el argumento. En lo sucesivo, este argumento se omitirá de algunas de las funciones, exceptuando aquellas en las que se considere importante dentro del contexto.

S y X , $\gamma^2(e^{j\omega_1}, e^{j\omega_2})$ y el peso óptimo normalizado $f_W^n(e^{j\omega_1}, e^{j\omega_2})$. A partir de 5.4, empleando 5.8, se tiene que

$$\gamma_{XS}^2(e^{j\omega_1}, e^{j\omega_2}) = \frac{H^*(e^{j\omega_1}, e^{j\omega_2})f_S(e^{j\omega_1}, e^{j\omega_2}) (H^*(e^{j\omega_1}, e^{j\omega_2})f_S(e^{j\omega_1}, e^{j\omega_2}))^*}{f_S(e^{j\omega_1}, e^{j\omega_2})f_X(e^{j\omega_1}, e^{j\omega_2})}$$

$$\gamma_{XS}^2(e^{j\omega_1}, e^{j\omega_2}) = \frac{|H(e^{j\omega_1}, e^{j\omega_2})|^2 f_S(e^{j\omega_1}, e^{j\omega_2})}{|H(e^{j\omega_1}, e^{j\omega_2})|^2 f_S(e^{j\omega_1}, e^{j\omega_2}) + f_N(e^{j\omega_1}, e^{j\omega_2})}, \quad (5.12)$$

ya que $f_S(e^{j\omega_1}, e^{j\omega_2}) = f_S^*(e^{j\omega_1}, e^{j\omega_2})$. Por lo que respecta al peso óptimo $f_W^n(e^{j\omega_1}, e^{j\omega_2})$, se calculó empleando un método iterativo desarrollado por Lepe Casillas, cuya expresión fundamental, omitiendo los argumentos, es

$$f_W^{n(i+1)} = \left\{ \frac{1}{2[1 + f_W^{n(i)}]} \left[|\gamma| + \sqrt{\gamma^2 + 4f_W^{n(i)}[1 + f_W^{n(i)}]} - |\gamma| \right]^2 + 1 - \gamma^2 \right\}, \quad (5.13)$$

donde i es el índice de iteración. Ésta continúa hasta que se cumpla

$$\left| f_W^{n(i+1)}(e^{j\omega_1}, e^{j\omega_2}) - f_W^{n(i)}(e^{j\omega_1}, e^{j\omega_2}) \right| < \epsilon. \quad (5.14)$$

El valor de ϵ se escoge según la exactitud que se desee. Como primera aproximación se toma

$$f_W^{n(0)}(e^{j\omega_1}, e^{j\omega_2}) = 1 - \gamma^2(e^{j\omega_1}, e^{j\omega_2}).$$

Las expresiones anteriores suponen conocidos el espectro de la imagen original, el filtro de degradación, y el espectro del ruido. En la práctica, es difícil cumplir esto. Para la simulación realizada en este trabajo, se consideraron las siguientes situaciones

- *Espectro del Ruido:* Se toma como constante. Es uno de los parámetros de entrada en el programa de restauración que se describe más adelante.
- *Espectro de la imagen original:* Puesto que se cuenta con la imagen original, se cuenta con su espectro. Sin embargo, se presentan resultados empleando tanto este espectro como otros que se obtienen de imágenes que sirven de referencia. A esto se le conoce como *restauración a ciegas* ("Blind Restoration", en inglés) [12, pag. 392].
- *Filtro de degradación:* Se aplica lo mismo que se dijo con respecto al espectro de la imagen original: Se emplean el filtro original y otros distintos en la expresión 5.8.

Es muy importante mencionar que para efectuar los experimentos del modo en que se describen más adelante, es necesario suponer que las imágenes empleadas son realizaciones de campos ergódicos, de modo que el espectro de una de ellas es igual al espectro del campo que las genera como realizaciones. Además, la teoría presentada supone que estos procesos o campos son estacionarios, característica que no tienen las imágenes. El suponer una imagen como realización de un campo homogéneo y ergódico valida el procedimiento empleado para obtener los espectros empleados en los experimentos.

5.2.3 Observaciones Adicionales.

Las imágenes monocromáticas aquí empleadas ocupan 8 bits por pixel, y, como se mencionó al principio, cada pixel puede tener 256 valores posibles enteros positivos. La principal desventaja de mantener este formato, es que al realizar las operaciones aritméticas entre imágenes, puede perderse cierta precisión numérica, que a su vez puede proporcionar resultados que conduzcan a evaluaciones incorrectas. Lo anterior es particularmente notable cuando se manejan valores cercanos a los límites superior (255) e inferior(0). En este trabajo, para efectuar las operaciones entre imágenes, primero se cambió su formato numérico: cada valor de pixel se expresó como un valor numérico real. Una vez finalizado todo el proceso de operación numérica, los valores se regresan a su formato original, empleando una regla muy sencilla, si la parte decimal del número es mayor que 0.5, se toma el valor entero inmediato superior; si es menor, se toma el valor entero inmediato inferior. Con esto, se consideran las siguientes posibles situaciones:

1. Valor Máximo > 255 , Valor Mínimo < 0 y $|\text{Val. Max} - \text{Val. Min.}| > 256$;
2. Valor Máximo < 255 , Valor Mínimo > 0 y $|\text{Val. Max} - \text{Val. Min.}| < 256$;
3. Valor Máximo > 255 , Valor Mínimo > 0 y $|\text{Val. Max} - \text{Val. Min.}| < 256$;
4. Valor Máximo > 255 , Valor Mínimo > 0 y $|\text{Val. Max} - \text{Val. Min.}| > 256$;
5. Valor Máximo < 255 , Valor Mínimo < 0 y $|\text{Val. Max} - \text{Val. Min.}| < 256$;
6. Valor Máximo < 255 , Valor Mínimo < 0 y $|\text{Val. Max} - \text{Val. Min.}| > 256$.

En las 3 situaciones en que la diferencia es menor que 256, es posible representar fielmente los valores obtenidos con valores de grises cuantizados, aunque tenga que agregarse un cierto valor de desplazamiento al valor obtenido simplemente para colocarlo en el rango 0-255. En las situaciones en que la diferencia sea mayor que 256, la representación en valores de grises va no es directa y puede manejarse de dos modos básicos. Uno es escalar los valores al rango deseado; otro es sencillamente truncar los valores mayores que 255 a 255 y los menores que cero, a cero. Los resultados visuales dependerán también del número de valores que se encuentran fuera del rango deseado.

5.3 Programación

Los programas se desarrollaron en lenguaje C empleando el paquete *Borland C++* © 3.1 para DOS y se hizo uso de la memoria expandida (EMS) del sistema. Los valores de los pixeles se pasaron a un formato de punto flotante que en C ocupa 4 octetos (*bytes* en inglés) por valor. De este modo, por ejemplo, para una imagen de 256×256 pixeles, es necesario almacenar 65536 valores de 4 bytes cada uno, lo que da 262,144 bytes. Al realizar la Transformada Rápida de Fourier (*Fast Fourier Transform FFT*), se obtienen valores complejos, lo que duplica el valor anterior, es decir, para cada imagen es necesario disponer de 524,288 bytes. Se desarrollaron 4 programas fundamentales, cuyas características se describen a continuación.

- *Programa # 1:* Se emplea para degradar imágenes por difuminación y agregarles ruido de una variancia dada. Para lo anterior, se toman las transformadas de Fourier tanto de la imagen original como del filtro difuminador, y se realiza la multiplicación en la frecuencia. Hecho esto, al resultado se le toma la transformada inversa y a continuación se le agrega el ruido blanco normal. Los programas para calcular la transformada rápida de Fourier *radix 2*, N par e igual a 256, y para generar ruido, se tomaron, con algunas modificaciones convenientes, de [16]. La expresión para la transformada discreta de Fourier $\mathcal{F}(u, v)$ del campo $F(m, n)$ empleada es

$$\mathcal{F}(u, v) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F(m, n) e^{-j2\pi um/M - j2\pi vn/N}, \quad (5.15)$$

mientras que para la transformada inversa es

$$\bar{F}(m, n) = \frac{1}{\sqrt{MN}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \mathcal{F}(u, v) e^{j2\pi um/M + j2\pi vn/N}. \quad (5.16)$$

Para este trabajo, se cumple que $M = N = 256$.

- *Programa # 2:* En esta parte se programan las ecuaciones de la sección 6.2.2 para calcular los filtros de Wiener, de mínima información de discriminación normalizada de propósito dual, y de Igualación de Espectro. Con ellos, se restaura la imagen degradada en el programa anterior. Este segundo programa toma como parámetros de entrada la imagen degradada, el filtro empleado en la difuminación de la imagen (o bien, puede proponerse y emplearse otro distinto para evaluar los efectos en la restauración), una imagen de referencia y la variancia del ruido aditivo. Como se mencionó anteriormente, tanto el filtro como la imagen de referencia pueden ser los originales u otros propuestos. El espectro de la imagen de referencia se emplea en la fórmula de restauración.
- *Programa # 3:* Este programa se encarga de la normalización de los resultados obtenidos en el programa # 2, considerando las situaciones presentadas en la sección 6.2.3. Es decir, 1) normalizando cada resultado por separado de acuerdo a los niveles de sus valores extremos; 2) normalizando los 3 resultados de acuerdo a los valores extremos, máximo y mínimo obtenidos del conjunto de las 3 imágenes; o bien, 3) truncando los valores mayores que 255 o menores que 0. En esta última situación, no se hace ningún tipo de escalamiento. De hecho, en este trabajo, el escalar y el truncar son operaciones excluyentes.
- *Programa # 4:* Aquí se calculan los espectros de todas las imágenes involucradas. Para la visualización de los espectros, se toma el valor máximo y mínimo del conjunto de todos ellos y se hace un escalamiento para que queden dentro del rango de 0 a 256. Se hace lo mismo, por separado, para visualizar la coherencia y el filtro de difuminación. La versión numérica de 5.7 empleada para calcular el espectro del campo X es

$$P_X(u, v) = |\mathcal{X}(u, v)|^2, \quad (5.17)$$

donde $\mathcal{X}(u, v)$ es la transformada discreta de Fourier de X . La ecuación 5.17 proporciona lo que se conoce como el *periodograma* de X , el cual es un estimador sesgado de la f.d.e.p.

$f_{XX}(e^{j\omega_1}, e^{j\omega_2})$ de X . De hecho, el periodograma se puede ver como la Transformada de Fourier de la convolución del campo de autocorrelación real $R(\tau_1, \tau_2)$ y la ventana de Bartlett [26, pag. 65]. Cuando $m, n \rightarrow \infty$, el valor esperado del periodograma es igual a la f.d.e.p., sin embargo, la variancia nunca se aproxima a cero. Existen algunas alternativas para obtener mejores estimadores de espectros reduciendo la variancia, tales como la promediación de los periodogramas y el empleo sistemático de ventanas en él.

- *Programa # 5*: Es el programa que configura el monitor de una computadora como medio exhibidor de los resultados obtenidos en cada simulación. Este programa requiere de un monitor *Super VGA* y tarjeta de vídeo del mismo tipo con al menos 1Mbyte de memoria. Se puede exhibir en 3 resoluciones distintas: 640×400 , 800×600 y 1024×768 pixeles por pantalla. Las funciones básicas para estas resoluciones, se tomaron, con modificaciones convenientes, de [31]. Se incluyeron dos escalas para dar colores falsos, aparte de la de grises, para resaltar detalles. La ventaja de este programa es que no requiere el uso de ambiente *Windows* y sirve para cualquier marca de tarjeta de vídeo con las características ya mencionadas. La desventaja es que requiere el uso de memoria expandida y no es exportable a otras plataformas de C ya que emplea funciones exclusivas de *DOS*. Para la exhibición, se manejan 30 imágenes por separado de 256×256 *pixeles* = 64K*pixeles* cada una, almacenadas también cada una por separado en un archivo de formato simple (1 byte = 1 pixel o *raw format* en inglés), por lo cual se requiere de 64Kbytes por imagen y casi 2 Mbytes de memoria para el almacenamiento de todas. Finalmente, este programa exhibidor incluye una opción que permite guardar la pantalla deseada en un archivo de formato *TIFF*, la cual se ha empleado para la impresión de los resultados que se anexan.

Cada programa tiene subprogramas para la adecuación de formatos de datos y para la correcta comunicación de datos entre los programas principales. Los archivos de las imágenes generadas, pueden visualizarse con casi todos los visualizadores comerciales existentes ³.

5.4 Resultados

Se muestran los resultados de 6 experimentos realizados, empleando 6 imágenes distintas. En todos los casos, la imagen que se emplea como referencia es similar a la imagen original, pero con algunas zonas degradadas sintéticamente.

- **EXPERIMENTO 1**: La imagen original es una fotografía aérea y la de referencia es la misma pero girada 90° en sentido horario y con una región degradada. Las diferencias entre ambas pueden percibirse mejor en los espectros respectivos. En las imágenes restauradas se observa que la restauración de Wiener es de un modo "suave" que produce una imagen borrosa. Por lo contrario, el del Igualación de Espectro con fase mejorada, es el que quita menos ruido pero produce una imagen con mayor definición. El filtro de Lepe-Casillas está entre ambos: quita más ruido que el de igualación de espectros y produce una imagen menos borrosa que el de Wiener. En cierto modo, esto también se observa

³El autor no conoce uno que no pueda leer el formato sencillo.

en los espectros de la imágenes restauradas y en la imágenes de diferencias: el espectro de la imagen restaurada con el filtro de Wiener presenta características geométricas del filtro de difuminación, mientras que el espectro de los otros dos presenta características del espectro de la imagen de referencia. La imagen de diferencia entre la imagen original y la restaurada mediante el filtro de Wiener presenta los valores mayores justamente en los contornos de la imagen, mientras que en las otras imágenes diferencia, los contornos están menos delineados y el ruido distribuido de modo más uniforme. No obstante lo anterior, el filtro de Wiener proporciona la menor potencia de error de los tres. Por supuesto, con ese criterio, el Filtro de Wiener es "óptimo".

- **EXPERIMENTO 2:** La imagen original es sintética generada con un fondo cuadrulado y figuras geométricas sobrepuestas. Estas figuras son opacas y no permiten ver el fondo. La imagen de referencia es la misma pero con menos figuras. Se cumple lo expresado en el Experimento 1, pero aquí se presenta una particularidad: en las tres restauraciones se perciben contornos del fondo en el interior de las figuras que deben ser opacas. El efecto es mayor en la restauración de Wiener.
- **EXPERIMENTO 3:** La imagen original es LENA y la de referencia es la misma LENA con una región degradada. En este experimento se sigue cumpliendo en términos generales con los resultados anteriores, aunque el efecto de la degradación es grande.
- **EXPERIMENTO 4:** La imagen original es una toma de una secuencia de vídeo y la de referencia es la misma pero con regiones degradadas. En este experimento se observa un poco más claramente que la restauración de Wiener no es visualmente la mejor.
- **EXPERIMENTO 5:** La imagen original es parte del diapasón de un violín y la de referencia es otra parte del mismo. En este experimento son particularmente notables los resultados de las normalizaciones para la exhibición. Visualmente parece mejor truncar los valores mayores que 255 y menores que 0.
- **EXPERIMENTO 6:** La imagen original es una sección de una partitura y la de referencia es una sección distinta de la misma partitura. Nótese como la normalización de cada imagen por separado, parece perjudicar a la restauración por igualación de espectro y a la restauración del filtro de Lepe. Cuando se trunca, estas últimas parecen recuperarse un poco. Cuando se escala a los valores máximos considerando las tres juntas, se recuperan las características generales observadas en los primeros experimentos.

5.5 Conclusiones

Los resultados obtenidos confirman prácticamente lo demostrado en la teoría: El Filtro de Mínima Información de Discriminación Normalizada de Propósito Dual (Filtro de Lepe) tiene un comportamiento intermedio entre el Filtro de Wiener y el de Igualación de Espectros con fase mejorada. Sin embargo, como la potencia del error de restauración que produce este filtro no es mínima, como la de Wiener, puede verse afectado sobre todo en imágenes con valores

extremos muy cercanos a 0 o 255, ya que implica tener que truncar o escalar y esto, a su vez, puede producir efectos indeseados o no notables en otro tipo de restauraciones.

Otro factor muy importante a considerar, es que el estimador de espectro empleado, el periodograma, no es un buen estimador. Puede pensarse que en los experimentos afecta de igual modo a los tres filtros; no obstante, el Filtro de Lepe requiere además de la evaluación de la coherencia, la cual involucra estimación espectral. Indudablemente, mejores estimados de los espectros, tales como los modelos paramétricos o la segmentación de la imagen, que producen espectros más suaves, se traducirán en mejores resultados. En particular, el filtro de Lepe puede verse más favorecido por el hecho de que puede mejorarse la evaluación de la coherencia.

Finalmente, en los resultados se ha observado que los criterios de normalización para exhibición pueden afectar seriamente las evaluaciones visuales de los desempeños de los filtros. Para cada aplicación debe escogerse la normalización más adecuada.

Apéndice A

A.1 Introducción.

En este apéndice se presenta un método para obtener la *Distorsión de Itakura-Saito*, partiendo de la información de discriminación normalizada entre los procesos \hat{S} y S , $\bar{D}(\hat{S}, S)$, donde S es un proceso real, normal, discreto, de media cero, llamado proceso original o deseado; y \hat{S} , que es la respuesta de un filtro lineal, invariante con el tiempo, discreto, con función de transferencia $G(e^{j\omega})$, cuando como entrada tiene un cierto proceso X , llamado proceso observado, real, discreto de media cero y no necesariamente normal. Se define el proceso de error como $E = S - \hat{S}$. Se considerarán los casos en que \hat{S} sea estacionario y regular y asintóticamente estacionario y asintóticamente regular. En la sección A.2 se expone la estructura de la ecuación involucrada y se indican los términos relacionados con la *Distorsión de Itakura-Saito*. En la sección A.3 se definen los procesos asintóticamente estacionarios y asintóticamente regulares, y se explican sus relaciones con los procesos estacionarios y regulares. En la sección A.4 se presentan de forma sencilla algunos resultados de la teoría de predicción lineal y su aplicación en la factorización de matrices de correlaciones Toeplitz. Finalmente, en la sección A.5 se emplean todos estos resultados para la obtención de la *Distorsión de Itakura-Saito*.

A.2 Información de Discriminación Normalizada entre los procesos \hat{S} y S .

De las definiciones dadas en el capítulo 2, y tomadas de [30], se tiene que la *Información de Discriminación Normalizada* entre \hat{S} y S se desarrolla como

$$\begin{aligned}\bar{D}(\hat{S}, S) &= \lim_{N \rightarrow \infty} \frac{1}{N} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} P_{\hat{S}}^N(\underline{u}^N) \ln \left\{ \frac{P_{\hat{S}}^N(\underline{u}^N)}{P_S^N(\underline{u}^N)} \right\} \delta_u^N ; \\ &= -\bar{H}_{\hat{S}} + \lim_{N \rightarrow \infty} \frac{1}{N} E_{\hat{S}} \left\{ \ln \left\{ \frac{1}{P_S^N(\underline{u}^N)} \right\} \right\},\end{aligned}\tag{A.1}$$

donde $E_{\hat{S}}\{ \}$ indica el valor esperado con respecto a \hat{S} . Como el proceso S se supone normal, su función de densidad está dada por [14, pag. 197]:

$$P_S^N(\underline{\alpha}^N) = \frac{1}{(2\pi)^{\frac{N}{2}} (\Delta_{S(N)})^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} \underline{\alpha}^{NT} \mathbf{R}_{S(N)}^{-1} \underline{\alpha}^N \right\},$$

donde $\underline{\mathbf{R}}_{s(N)}$ es la matriz de correlaciones de orden N del proceso \mathbf{S} y $\Delta_{s(N)}$ es su determinante. Puesto que el producto matricial $\frac{1}{2}\alpha^{NT}\underline{\mathbf{R}}_{s(N)}^{-1}\alpha^N$ da un escalar, puede tomarse el valor esperado de la *traza* de este producto matricial, y por las propiedades de ésta:

$$\begin{aligned}\bar{\mathbf{D}}(\hat{\mathbf{S}}, \mathbf{S}) &= -\bar{\mathbf{H}}_{\hat{\mathbf{S}}} + \lim_{N \rightarrow \infty} \frac{1}{N} \ln \left\{ (2\pi)^{\frac{N}{2}} (\Delta_{s(N)})^{\frac{1}{2}} \right\} + \lim_{N \rightarrow \infty} \frac{1}{N} E_{\hat{\mathbf{S}}} \left\{ \frac{1}{2}\alpha^{NT}\underline{\mathbf{R}}_{s(N)}^{-1}\alpha^N \right\}, \\ \bar{\mathbf{D}}(\hat{\mathbf{S}}, \mathbf{S}) &= -\bar{\mathbf{H}}_{\hat{\mathbf{S}}} + \frac{1}{2} \ln(2\pi) + \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \ln \left\{ \Delta_{s(N)} \right\} + \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr}(\underline{\mathbf{R}}_{s(N)}^{-1}\underline{\mathbf{R}}_{\hat{\mathbf{S}}(N)}),\end{aligned}\quad (\text{A.2})$$

siendo $\underline{\mathbf{R}}_{\hat{\mathbf{S}}(N)}$ la matriz de correlaciones de orden N del proceso $\hat{\mathbf{S}}$ y $\bar{\mathbf{H}}_{\hat{\mathbf{S}}}$ la razón de cambio de entropía del proceso estimado $\hat{\mathbf{S}}$. Se demostrará que los dos últimos sumandos del lado derecho de la expresión A.2 están relacionados, respectivamente, con la integral del logaritmo de la función de densidad espectral de potencia de \mathbf{S} y la integral del cociente de las funciones de densidad espectral de potencia de $\hat{\mathbf{S}}$ y \mathbf{S} . Esta expresión A.2 se reducirá considerando los dos casos siguientes:

- *Caso 1:* $\hat{\mathbf{S}}$ y \mathbf{S} regulares y estacionarios.
- *Caso 2:* $\hat{\mathbf{S}}$ asintóticamente regular y estacionario y \mathbf{S} regular y estacionario.

A.3 Procesos Asintóticamente Regulares y Asintóticamente Estacionarios.

A.3.1 Equivalencia Asintótica entre Secuencias de Matrices [22].

En esta parte se reproducen algunos resultados de [22], dada la dificultad para su obtención. Estos resultados son fundamentales en las definiciones subsecuentes ¹.

Sean $\alpha_k, k = 0, 1, \dots, N-1$ los valores característicos de una matriz $\underline{\mathbf{A}}$ de $N \times N$ y $\lambda_k, k = 0, 1, \dots, N-1$ los valores característicos de la matriz hermitiana $\underline{\mathbf{A}}^*\underline{\mathbf{A}}$. La norma robusta de $\underline{\mathbf{A}}$ (conocida también como $\|\underline{\mathbf{A}}\|_2$, es decir la norma matricial $\|\underline{\mathbf{A}}\|_p$ cuando $p = 2$ [19, pag. 14]) se define como

$$\|\underline{\mathbf{A}}\|_R = \left(\max_{\underline{\mathbf{x}}^*\underline{\mathbf{x}}=1} \frac{\underline{\mathbf{x}}^*\underline{\mathbf{A}}^*\underline{\mathbf{A}}\underline{\mathbf{x}}}{\underline{\mathbf{x}}^*\underline{\mathbf{x}}} \right)^{\frac{1}{2}}, \quad \underline{\mathbf{x}} \in R^N; \quad (\text{A.3})$$

donde * indica conjugada transpuesta. Para esta norma se cumple que

$$\|\underline{\mathbf{A}}\|_R^2 \triangleq \lambda_M = \max_k \{ \lambda_k \}, \quad (\text{A.4})$$

siendo λ_M el máximo valor característico de $\underline{\mathbf{A}}^*\underline{\mathbf{A}}$. La norma débil o de *Hilbert-Schmidt* de $\underline{\mathbf{A}}$ (conocida también como la norma matricial de *Frobenius* [19, pag. 14]) se define como

$$\|\underline{\mathbf{A}}\|_F = \left(\frac{1}{N} \sum_{k=0}^{N-1} \sum_{j=0}^{N-1} |a_{k,j}|^2 \right)^{\frac{1}{2}} = \left(\frac{1}{N} \text{tr} [\underline{\mathbf{A}}^*\underline{\mathbf{A}}] \right)^{\frac{1}{2}} = \left(\frac{1}{N} \sum_{k=0}^{N-1} \lambda_k \right)^{\frac{1}{2}}, \quad (\text{A.5})$$

¹Se emplean los resultados de que para cualquier matriz cuadrada no singular, la traza es igual a la suma de sus valores característicos y el determinante es igual al producto de éstos [18, pag. 22].

donde los $a_{k,j}$ son los elementos de $\underline{\mathbf{A}}$. Para esta norma se cumple que

$$\|\underline{\mathbf{A}}\|_F^2 \geq \frac{1}{N} \sum_{k=0}^{N-1} |\alpha_k|^2; \quad (\text{A.6})$$

con igualdad si $\underline{\mathbf{A}}$ es normal. Dos *secuencias de matrices* $\underline{\mathbf{A}}_N$ y $\underline{\mathbf{B}}_N$ de orden $N \times N$ son *asintóticamente equivalentes* (y se abrevia $\underline{\mathbf{A}}_N \sim \underline{\mathbf{B}}_N$) si

(i) $\underline{\mathbf{A}}_N$ y $\underline{\mathbf{B}}_N$ son acotadas en la norma robusta

$$\|\underline{\mathbf{A}}_N\|_R, \|\underline{\mathbf{B}}_N\|_R \leq M < \infty; \quad (\text{A.7})$$

y

(ii) $\underline{\mathbf{A}}_N - \underline{\mathbf{B}}_N$ tiende a cero en la norma débil cuando $N \rightarrow \infty$,

$$\lim_{N \rightarrow \infty} \|\underline{\mathbf{A}}_N - \underline{\mathbf{B}}_N\|_F = 0. \quad (\text{A.8})$$

Así, se dice que si una de ellas es Toeplitz, entonces la otra es asintóticamente Toeplitz. Algunas propiedades de la equivalencia asintótica son:

a) Si $\underline{\mathbf{A}}_N \sim \underline{\mathbf{B}}_N$, entonces

$$\lim_{N \rightarrow \infty} \|\underline{\mathbf{A}}_N\|_F = \lim_{N \rightarrow \infty} \|\underline{\mathbf{B}}_N\|_F; \quad (\text{A.9})$$

b) Si $\underline{\mathbf{A}}_N \sim \underline{\mathbf{B}}_N$ y $\underline{\mathbf{B}}_N \sim \underline{\mathbf{C}}_N$, entonces

$$\underline{\mathbf{A}}_N \sim \underline{\mathbf{C}}_N; \quad (\text{A.10})$$

c) Si $\underline{\mathbf{A}}_N \sim \underline{\mathbf{B}}_N$ y $\underline{\mathbf{C}}_N \sim \underline{\mathbf{D}}_N$, entonces

$$\underline{\mathbf{A}}_N \underline{\mathbf{C}}_N \sim \underline{\mathbf{B}}_N \underline{\mathbf{D}}_N. \quad (\text{A.11})$$

d) Dadas dos matrices $\underline{\mathbf{A}}_N$ y $\underline{\mathbf{B}}_N$ asintóticamente equivalentes con valores característicos $\alpha_{N,k}$ y $\beta_{N,k}$, respectivamente, entonces

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} \alpha_{N,k} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} \beta_{N,k}; \quad (\text{A.12})$$

es decir, los valores promedio de las trazas son iguales cuando $N \rightarrow \infty$.

e) Si $\underline{\mathbf{A}}_N \sim \underline{\mathbf{B}}_N$ y

$$\|\underline{\mathbf{A}}_N^{-1}\|_R, \|\underline{\mathbf{B}}_N^{-1}\|_R \leq K < \infty, \quad (\text{A.13})$$

entonces

$$\underline{\mathbf{A}}_N^{-1} \sim \underline{\mathbf{B}}_N^{-1}. \quad (\text{A.14})$$

La equivalencia asintótica entre matrices implica que sus valores característicos, productos e inversas, se comportan similarmente. Es importante señalar que esto no significa la convergencia de cada elemento entre matrices, sino más bien se refiere a su comportamiento en promedio.

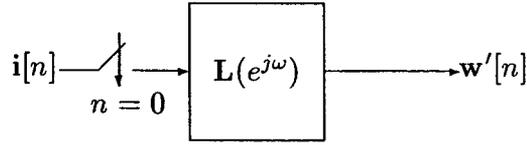


Figura A.1: Sistema Lineal I

A.3.2 Definición de Procesos Asintóticamente Estacionarios y Asintóticamente Regulares.

Para la definición de este tipo de procesos, se emplearán dos ejemplos sencillos ilustrativos. Primeramente, considérese la Fig. A.1 en la que se muestra un Sistema Lineal Invariante con el Tiempo, causal y estable, de fase mínima, con función de transferencia dada por

$$\mathbf{L}(e^{j\omega}) = \sum_{k=0}^{\infty} l[k](e^{j\omega})^{-k} = \sum_{k=0}^{\infty} l_k(e^{j\omega})^{-k}, \quad (\text{A.15})$$

$$\sum_{k=0}^{\infty} |l_k| < \infty, \quad (\text{A.16})$$

donde $l[n]$ es la respuesta al impulso del sistema; y como entrada, una secuencia de variables aleatorias independientes idénticamente distribuidas (v.a.i.i.d. o proceso de ruido blanco en sentido estricto), y de potencia unitaria $\mathbf{I} \equiv \mathbf{i}[n]$ (proceso de ruido blanco en sentido estricto), aplicada desde $n = 0$. El proceso de salida $\mathbf{W}' \equiv \mathbf{w}'[n]$ estará dado por

$$\mathbf{w}'[n] = \sum_{k=0}^{\infty} l[k]\mathbf{i}[n-k]. \quad (\text{A.17})$$

Por simplicidad, se supone inicialmente que este sistema tiene Respuesta al Impulso Finita de tamaño N . Entonces, las primeras N muestras de la salida pueden escribirse como:

$$w'_n = \sum_{k=0}^n l_k i_{n-k} \quad n = 0, 1, \dots, N-1,$$

$$\begin{pmatrix} w'_0 \\ w'_1 \\ w'_2 \\ \dots \\ w'_{N-1} \end{pmatrix} = \begin{pmatrix} l_0 & 0 & 0 & \dots & 0 \\ l_1 & l_0 & 0 & \dots & 0 \\ l_2 & l_1 & l_0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ l_{N-1} & l_{N-2} & l_{N-3} & \dots & l_0 \end{pmatrix} \begin{pmatrix} i_0 \\ i_1 \\ i_2 \\ \vdots \\ i_{N-1} \end{pmatrix};$$

$$\underline{\mathbf{w}}' = \underline{\mathbf{L}}_{(N)} \underline{\mathbf{i}}, \quad (\text{A.18})$$

donde $\underline{\mathbf{L}}_{(N)}$ se denominará **matriz asociada a la transformación lineal** efectuada por el sistema con función de transferencia $\mathbf{L}(e^{j\omega})$. De lo anterior, la matriz de correlaciones de orden

N de la salida queda

$$\begin{aligned}\mathbf{R}_{w'(N)} &= E \{ \mathbf{w}' \mathbf{w}'^T \} = E \{ \mathbf{L}_{(N)} \mathbf{i} (\mathbf{L}_{(N)} \mathbf{i})^T \} \\ \mathbf{R}_{w'(N)} &= (\mathbf{L}_{(N)} \mathbf{R}_{i(N)} \mathbf{L}_{(N)}^T); \end{aligned} \quad (\text{A.19})$$

y puesto que $E \{ \mathbf{i} \mathbf{i}^T \} = \mathbf{R}_{i(N)} = \mathbf{I}_{(N)}$, entonces

$$\mathbf{R}_{w'(N)} = \mathbf{L}_{(N)} \mathbf{L}_{(N)}^T. \quad (\text{A.20})$$

Nótese que a partir de esta ecuación se puede obtener el promedio de la potencia contenida en N muestras de $w'[n]$ como ²

$$P_{w'} = \frac{1}{N} \text{tr}(\mathbf{R}_{w'(N)}). \quad (\text{A.21})$$

Aún cuando los factores $\mathbf{L}_{(N)}$ y $\mathbf{L}_{(N)}^T$ son Toeplitz, $\mathbf{R}_{w'(N)}$ no lo es, por lo que el proceso de salida \mathbf{W}' no es, en términos estrictos, estacionario [21, pag. 710]. Por ejemplo, sea $N = 4$, la matriz $\mathbf{R}_{w'(N)}$ queda

$$\begin{aligned}\mathbf{R}_{w'(N)} &= \begin{pmatrix} l_0 & 0 & 0 & 0 \\ l_1 & l_0 & 0 & 0 \\ l_2 & l_1 & l_0 & 0 \\ l_3 & l_2 & l_1 & l_0 \end{pmatrix} \begin{pmatrix} l_0 & l_1 & l_2 & l_3 \\ 0 & l_0 & l_1 & l_2 \\ 0 & 0 & l_0 & l_1 \\ 0 & 0 & 0 & l_0 \end{pmatrix}, \\ \mathbf{R}_{w'(N)} &= \begin{pmatrix} l_0^2 & l_0 l_1 & l_0 l_2 & l_0 l_3 \\ l_0 l_1 & l_0^2 + l_1^2 & l_0 l_1 + l_1 l_2 & l_0 l_2 + l_1 l_3 \\ l_0 l_2 & l_0 l_1 + l_1 l_2 & l_0^2 + l_1^2 + l_2^2 & l_0 l_1 + l_1 l_2 + l_2 l_3 \\ l_0 l_3 & l_0 l_2 + l_1 l_3 & l_0 l_1 + l_1 l_2 + l_2 l_3 & l_0^2 + l_1^2 + l_2^2 + l_3^2 \end{pmatrix}, \end{aligned} \quad (\text{A.22})$$

o bien [22, pag. 58]:

$$r_{w'}[k, j] = \sum_{i=0}^{\min\{k, j\}} l_i l_{i+|k-j|}; \quad (\text{A.23})$$

donde se muestra la dependencia de dos variables ³.

Considérese ahora un sistema causal como el de la Fig. A.2. Aquí, el proceso de salida $\mathbf{W} \equiv w[n]$ es un *proceso regular*, y por supuesto, es estacionario con función de correlación dada por [23, pag. 65] (la entrada se aplica desde $n = -\infty$ y es una sucesión de v.a.i.i.d.):

$$r_w[m] = \sum_{i=-\infty}^{\infty} l_i l_{i+m}; \quad (\text{A.24})$$

y considerando el mismo valor de $N = 4$ del ejemplo anterior, se tiene que la matriz de correlaciones de la salida es

$$\mathbf{R}_{w(N)} = \begin{pmatrix} l_0^2 + l_1^2 + l_2^2 + l_3^2 & l_0 l_1 + l_1 l_2 + l_2 l_3 & l_0 l_2 + l_1 l_3 & l_0 l_3 \\ l_0 l_1 + l_1 l_2 + l_2 l_3 & l_0^2 + l_1^2 + l_2^2 + l_3^2 & l_0 l_1 + l_1 l_2 + l_2 l_3 & l_0 l_2 + l_1 l_3 \\ l_0 l_2 + l_1 l_3 & l_0 l_1 + l_1 l_2 + l_2 l_3 & l_0^2 + l_1^2 + l_2^2 + l_3^2 & l_0 l_1 + l_1 l_2 + l_2 l_3 \\ l_0 l_3 & l_0 l_2 + l_1 l_3 & l_0 l_1 + l_1 l_2 + l_2 l_3 & l_0^2 + l_1^2 + l_2^2 + l_3^2 \end{pmatrix}, \quad (\text{A.25})$$

²Si $w'[n]$ no es estacionario, cada muestra contiene en promedio, diferente potencia

³La expresión $r_{w'}[k, j]$ se refiere a la correlación como función de las variables k y j , mientras que la expresión $\mathbf{R}_{w'(N)}(l, m)$ se referirá al elemento l, m de la matriz $\mathbf{R}_{w'(N)}$.

la cual es una matriz Toeplitz. Comparando $\underline{\mathbf{R}}_{w'(N)}$ con $\underline{\mathbf{R}}_{w(N)}$ y A.23 con A.24, puede notarse la evolución del transitorio (condiciones iniciales del sistema nulas), y la convergencia de los valores de la primera a los de la segunda: el último renglón y la última columna en ambas matrices son iguales. Esto significa que

$$r_{w'}[k, j] \rightarrow r_w[k - j], \quad \text{si } k \rightarrow M \text{ o si } j \rightarrow M, \quad (\text{A.26})$$

donde M es el valor de la ordenada del punto más alejado del origen, y para este ejemplo, $M = N = 4$. Obsérvese que por la condición de estabilidad A.16, ambas matrices cumplen con A.7; y haciendo uso de una propiedad de secuencias convergentes (ver siguiente sección) se tiene que:

$$\lim_{N \rightarrow \infty} \|\underline{\mathbf{R}}_{w'(N)} - \underline{\mathbf{R}}_{w(N)}\|_F = \lim_{N \rightarrow \infty} \left(\frac{1}{N} \text{tr} [\underline{\mathbf{R}}_{w'(N)} - \underline{\mathbf{R}}_{w(N)}] \right)^{\frac{1}{2}} = 0; \quad (\text{A.27})$$

por lo que se cumple también con A.8. Entonces, las matrices $\underline{\mathbf{R}}_{w'(N)}$ y $\underline{\mathbf{R}}_{w(N)}$ son asintóticamente equivalentes; y puesto que $\underline{\mathbf{R}}_{w(N)}$ es Toeplitz, $\underline{\mathbf{R}}_{w'(N)}$ es asintóticamente Toeplitz. Por tanto, un proceso cuya matriz de correlaciones tenga una estructura como A.22, o bien, que pueda factorizarse como el producto de una matriz triangular inferior Toeplitz por su transpuesta, y que estos factores correspondan a un sistema causal, y estable, se define como un Proceso Asintóticamente Estacionario [15, pag. 225, 238]. En otras palabras, sus parámetros estadísticos variantes con el tiempo se aproximan a los valores de respuesta permanente o estacionaria invariantes con el tiempo cuando $n \rightarrow \infty$ [17, pag. 96].

Se define ahora la función de densidad espectral de potencia asintótica $f_{w'}(e^{j\omega})$ del proceso asintóticamente estacionario \mathbf{W}' como

$$f_{w'}(e^{j\omega}) = \sum_{n=-\infty}^{\infty} r_{w'}[n](e^{j\omega})^{-n}; \quad (\text{A.28})$$

donde

$$r_{w'}[n] \triangleq r_{w'}[k, j] \quad \forall k, j \text{ tales que } |k - j| = n; \quad \text{si } k \rightarrow M \text{ o si } j \rightarrow M, \quad (\text{A.29})$$

con M definido como arriba, por lo que

$$r_{w'}[n] = r_w[n]. \quad (\text{A.30})$$

Con esta definición se tiene que

$$\lim_{N \rightarrow \infty} f_{w'}(e^{j\omega}) = f_w(e^{j\omega}) = |\mathbf{L}(e^{j\omega})|^2, \quad (\text{A.31})$$

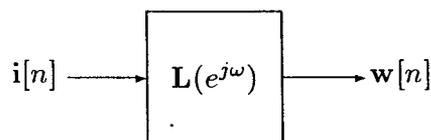


Figura A.2: Sistema Lineal II

donde $f_w(e^{j\omega})$ es la función de densidad espectral de potencia del proceso estacionario $w[n]$. Adicionalmente, se define: un Proceso es Asintóticamente Regular, y se denota como X^{as} , si se obtiene de un sistema como el de la Fig.A.1, con función de transferencia el filtro de innovaciones de un proceso Regular X , $L_x(e^{j\omega})$, de fase mínima, y entrada una sucesión de v.a.i.i.d. aplicada desde $n = 0$ [5]. Esta definición es una modificación a la presentada en [15, pag. 297].

Mediante las expresiones anteriores es posible relacionar las potencias promedio de los procesos de salida X^{as} , P_{xas} , y X , P_x , como sigue:

$$P_{xas} = \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr}(\underline{L}_{x(N)} \underline{L}_{x(N)}^T) = \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr}(\underline{R}_{x(N)}) = R_{x(N)}(0, 0) = r_x[0] = P_x; \quad (\text{A.32})$$

donde $\underline{L}_{x(N)}$ es la matriz de $N \times N$ asociada a $L_x(e^{j\omega})$, $R_{x(N)}(0, 0)$ es el primer elemento de $\underline{R}_{x(N)}$ y $r_x[0]$ es la función de autocorrelación de X evaluada en 0.

A.3.3 Conclusiones.

Los procesos asintóticamente estacionarios y asintóticamente regulares muestran un comportamiento similar al de los procesos estacionarios y regulares cuando el tiempo de observación es muy grande ya que el transitorio para entonces desaparece y el sistema alcanza una respuesta permanente. Por tanto, cuando la entrada se aplica desde $n = 0$, la salida de un SLIT causal y estable, para un tiempo de observación *tan grande como sea necesario* para que alcance su respuesta permanente, es igual a la que se observaría si la entrada se aplicase desde $n = -\infty$. Así, un proceso asintóticamente estacionario es equivalente a un proceso estacionario para un tiempo de observación infinitamente largo. Según Lepe [5], no todo proceso asintóticamente estacionario es asintóticamente regular. Nótese que cuando la función de transferencia del sistema, $L(e^{j\omega})$ es de fase mínima, su inverso $L^{-1}(e^{j\omega}) = \Gamma(e^{j\omega})$ es causal y estable, la matriz de transferencia $\underline{L}_{x(N)}$ es invertible para toda N (y su inversa se denota como $\underline{\Gamma}_{x(N)}$) y entonces es posible emplear sistemas inversos. La importancia de esto se verá más adelante. En general si X es un proceso regular, entonces

$$\mathbf{x}[n] = \sum_{k=0}^{\infty} l_x[k] \mathbf{i}[n - k], \quad (\text{A.33})$$

$$\underline{\mathbf{x}} = \underline{L}_{x(N)} \underline{\mathbf{i}}, \quad (\text{A.34})$$

$$\mathbf{i}[n] = \sum_{k=0}^{\infty} \gamma_x[k] \mathbf{x}[n - k], \quad (\text{A.35})$$

$$\underline{\mathbf{i}} = \underline{\Gamma}_{x(N)} \underline{\mathbf{x}}; \quad (\text{A.36})$$

donde $l_x[k] \equiv l_{x_k}$ y $\gamma_x[k] \equiv \gamma_{x_k}$ son las respuestas al impulso del filtro de innovaciones $L_x(e^{j\omega})$ y del filtro blanqueador $\Gamma_x(e^{j\omega})$, respectivamente, del proceso X , $\mathbf{i}[n]$ es una secuencia de v.a.i.i.d., y $\underline{L}_{x(N)}$ y $\underline{\Gamma}_{x(N)}$ son las matrices asociadas a $L_x(e^{j\omega})$ y $\Gamma_x(e^{j\omega})$, respectivamente, de orden N (donde $N \rightarrow \infty$).

A.4 Algunos resultados de la Teoría de Predicción Lineal .

Se reproducen a continuación algunos resultados importantes de [14, pag. 487] sobre Predicción Lineal, ⁴ los cuales permitirán mostrar la relación entre los coeficientes de los filtros predictores variantes e invariantes con el tiempo y su aplicación en la factorización de matrices Toeplitz.

A.4.1 Predicción Lineal de un Paso de Orden Infinito.

El proceso predictor de un paso $\hat{s}[n]$ de un proceso regular $S \triangleq s[n]$ en términos de todo su pasado $s[n-k]$, $k \geq 1$, se define como

$$\hat{s}[n] = \hat{E} \{s[n] \mid s[n-k], k \geq 1\} = \sum_{k=1}^{\infty} h[k]s[n-k], \quad (\text{A.37})$$

donde $h[k] \equiv h_k$ es la respuesta al impulso del filtro predictor $\mathbf{H}(e^{j\omega})$. En términos del filtro de innovaciones de S , $L_s(e^{j\omega})$, y de su respuesta al impulso $l_s[n] \equiv l_{s_n}$, se cumple que

$$\mathbf{H}(e^{j\omega}) = 1 - \frac{l_s[0]}{L_s(e^{j\omega})}; \quad (\text{A.38})$$

mientras que para el proceso de error de predicción $\varepsilon[n]$ se tiene que

$$\begin{aligned} \varepsilon[n] &= s[n] - \hat{s}[n] \\ &= \sum_{k=0}^{\infty} l_s[k]i[n-k] - \sum_{k=1}^{\infty} l_s[k]i[n-k] = l_s[0]i[n]; \end{aligned} \quad (\text{A.39})$$

donde $i[n]$ es un proceso de ruido blanco. El filtro del error de predicción $\mathbf{E}(e^{j\omega})$ queda entonces

$$\mathbf{E}(e^{j\omega}) = 1 - \mathbf{H}(e^{j\omega}) = \frac{l_s[0]}{L_s(e^{j\omega})} = l_s[0]\Gamma_s(e^{j\omega})$$

con lo que se llega a ⁵

$$\mathbf{E}(z) = 1 + \sqrt{P} \left(\gamma_{s_1} z^{-1} + \gamma_{s_2} z^{-2} + \dots + \gamma_{s_n} z^{-n} + \dots \right), \quad (\text{A.40})$$

donde $\Gamma_s(e^{j\omega}) = L_s^{-1}(e^{j\omega})$ es el filtro *blanqueador* de S , $\gamma_s[n] \equiv \gamma_{s_n}$ su respuesta al impulso, y P es la *potencia del error de predicción*, la cual, de A.39, es

$$P = E \{ \varepsilon^2[n] \} = l_s^2[0],$$

⁴Se trata de un breve resumen en el que se incluyen algunas notas aclarativas. Para mayor detalle, puede consultarse directamente la referencia

⁵Recuérdese que $\gamma_s[0] = \lim_{z \rightarrow \infty} \Gamma_s(z) = \frac{1}{l_s[0]}$

o bien, $l_s[0] = \sqrt{P}$. En términos de los coeficientes γ_{s_k} , el filtro predictor $\mathbf{H}(z)$ queda

$$\mathbf{H}(z) = 1 - \sqrt{P} \left(\gamma_{s_0} + \gamma_{s_1} z^{-1} + \gamma_{s_2} z^{-2} + \cdots + \gamma_{s_n} z^{-n} + \cdots \right),$$

y se obtiene la relación

$$h_k = -\sqrt{P} \gamma_{s_k}. \quad (\text{A.41})$$

Por otro lado, puesto que $\mathbf{L}_s(e^{j\omega})$ es de fase mínima,

$$\ln l^2[0] = \int_{-\pi}^{\pi} \ln |\mathbf{L}_s(e^{j\omega})|^2 \frac{d\omega}{2\pi} = \int_{-\pi}^{\pi} \ln f_s(e^{j\omega}) \frac{d\omega}{2\pi};$$

por lo que

$$P = \exp \left\{ \int_{-\pi}^{\pi} \ln f_s(e^{j\omega}) \frac{d\omega}{2\pi} \right\}, \quad (\text{A.42})$$

donde $f_s(e^{j\omega})$ es la función de densidad espectral de potencia de S. Finalmente, en términos de la respuesta al impulso del filtro blanqueador de S, la ecuación A.39 queda

$$\varepsilon[n] = l_0 i[n] = \sum_{k=0}^{\infty} \sqrt{P} \gamma_{s_k} s_{n-k}, \quad (\text{A.43})$$

o bien

$$\begin{pmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \varepsilon_2 \\ \cdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} \sqrt{P} \gamma_{s_0} & 0 & 0 & \cdots & 0 \\ \sqrt{P} \gamma_{s_1} & \sqrt{P} \gamma_{s_0} & 0 & \cdots & 0 \\ \sqrt{P} \gamma_{s_2} & \sqrt{P} \gamma_{s_1} & \sqrt{P} \gamma_{s_0} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sqrt{P} \gamma_{s_n} & \sqrt{P} \gamma_{s_{n-1}} & \sqrt{P} \gamma_{s_{n-2}} & \cdots & \sqrt{P} \gamma_{s_0} \end{pmatrix} \begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix}, \quad (\text{A.44})$$

donde $\underline{\Gamma}_{s(N)}$ es la matriz de asociada a la transformación $\underline{\Gamma}_s(e^{j\omega})$, de tamaño N (y se puede hacer que $N \rightarrow \infty$). Nótese que $\sqrt{P} \gamma_{s_0} = 1$ y que $\underline{\Gamma}_{s(N)}^{-1} = \underline{\mathbf{L}}_{s(N)}$.

A.4.2 Predicción Lineal de Un Paso de Orden constante Finito.

El proceso predictor de un paso $\hat{s}_N[n]$ del proceso regular $\mathbf{S} \triangleq s[n]$ en términos de los N valores pasados más recientes, toma la forma

$$\hat{s}_N[n] = \hat{E} \{ s[n] \mid s[n-k], 1 \leq k \leq N \} = \sum_{k=1}^N h_k^N s[n-k]; \quad (\text{A.46})$$

y la función de transferencia del filtro predictor se define como

$$\hat{\mathbf{H}}_N(e^{j\omega}) = \sum_{k=1}^N h_k^N (e^{j\omega})^{-k},$$

y el error de predicción de orden N , $\hat{\varepsilon}_N[n]$, como

$$\hat{\varepsilon}_N[n] = s[n] - \hat{s}_N[n].$$

Del principio de ortogonalidad y de la expresión para la potencia del error de predicción de orden N , $P_N = E \{ \hat{\varepsilon}_N^2[n] \}$, se obtiene el sistema de orden $N + 1$ por $N + 1$ conocido como *ecuaciones de Yule-Walker*. En particular, aplicando la regla de Cramer, se tiene que

$$P_N = \frac{\Delta_{s(N+1)}}{\Delta_{s(N)}}, \quad (\text{A.47})$$

donde $\Delta_{s(N+1)}$ y $\Delta_{s(N)}$ son los determinantes de las matrices de correlaciones $\mathbf{R}_{s(N+1)}$ de orden $N + 1$ y $\mathbf{R}_{s(N)}$ de orden N , respectivamente, del proceso S . Es evidente que si $N \rightarrow \infty$, entonces $\hat{s}_N[n] \rightarrow \hat{s}[n]$, o sea

$$\begin{aligned} \lim_{N \rightarrow \infty} \hat{\varepsilon}_N[n] &= \varepsilon[n], \\ \lim_{N \rightarrow \infty} h_k^N &= h_k = -\sqrt{P} \gamma_k, \\ \lim_{N \rightarrow \infty} P_N &= \lim_{N \rightarrow \infty} \frac{\Delta_{s(N+1)}}{\Delta_{s(N)}} = P. \end{aligned} \quad (\text{A.48})$$

Finalmente, en el Anexo A.1 se demuestra que

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln \Delta_{s(N)} = \ln P = \int_{-\pi}^{\pi} \ln f_s(e^{j\omega}) \frac{d\omega}{2\pi}. \quad (\text{A.49})$$

Anexo A.1 La demostración de A.49 es la respuesta al ejercicio 14.15 de [14, pag. 531].

§La secuencia P_n ($1 \leq n \leq \infty$) cumple que $P_1 \geq P_2 \geq \dots \geq P_i \dots \geq P_{N-1} \geq P_N \geq P_{N+1} \dots \geq P$, para $1 \leq i \leq N \leq \infty$, y además

$$\lim_{i \rightarrow N} P_i = P_N \quad y \quad \lim_{N \rightarrow \infty} P_N = P;$$

es decir, P_i es una secuencia convergente. La secuencia $\ln P_i$ también es convergente, por lo que se cumple que $\ln P_1 \geq \ln P_2 \geq \dots \geq \ln P_i \geq \dots \geq \ln P_{N-1} \geq \ln P_N \geq \ln P_{N+1} \geq \dots \geq \ln P$, y

$$\lim_{i \rightarrow N} \ln P_i = \ln P_N \quad y \quad \lim_{N \rightarrow \infty} \ln P_N = \ln P.$$

Por lo anterior, puede aplicarse la siguiente propiedad simple de las secuencias convergentes [14, pag. 567]

$$\lim_{k \rightarrow \infty} \left(\frac{1}{k} \sum_{m=1}^k x_m \right) = \lim_{k \rightarrow \infty} \left(\frac{1}{k} (x_1 + x_2 + x_3 + \dots + x_k) \right),$$

a $\ln P_N = \ln (\Delta_{s(N+1)} / \Delta_{s(N)})$, con lo que se llega a

$$\lim_{N \rightarrow \infty} \left(\frac{1}{N} (\ln P_1 + \ln P_2 + \ln P_3 + \dots + \ln P_N) \right) = \lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum_{i=1}^N \ln \frac{\Delta_{s(i+1)}}{\Delta_{s(i)}} \right) = S, \quad (\text{A.50})$$

donde S es el valor de la sumatoria. Se tiene entonces que

$$\lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum_{i=1}^N \ln \frac{\Delta_{s(i+1)}}{\Delta_{s(i)}} \right) = \lim_{N \rightarrow \infty} \left(\ln \frac{\Delta_{s(N+1)}}{\Delta_{s(N)}} \right) = S, \quad (\text{A.51})$$

o sea, de A.51 con A.48

$$\ln P = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \ln \frac{\Delta_{s(n+1)}}{\Delta_{s(n)}}, \quad (\text{A.52})$$

con lo que $\ln P$ se ha expresado en términos de una serie infinita convergente. De A.50 se tiene que

$$\lim_{N \rightarrow \infty} \ln \frac{\Delta_{s(N+1)}}{\Delta_{s(N)}} = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \Delta_{s(N+1)} - \lim_{N \rightarrow \infty} \frac{1}{N} \ln \Delta_{s(1)}, \quad (\text{A.53})$$

donde se observa que el segundo término del lado derecho se hace cero. Por otra parte, puesto que esta serie infinita es convergente, cumple

$$\lim_{N \rightarrow \infty} U_N = 0, \quad (\text{A.54})$$

donde

$$U_N = \frac{1}{N} \ln \frac{\Delta_{s(N+1)}}{\Delta_{s(N)}}$$

es el N -ésimo término de la serie. Por otro lado también se cumple que $U_N = S_N - S_{N-1}$, siendo S_N la suma de los primeros N términos de la serie y S_{N-1} la suma de los primeros $N-1$ términos de la misma. Empleando estos resultados se tiene que

$$\lim_{N \rightarrow \infty} U_N = \lim_{N \rightarrow \infty} (S_N - S_{N-1}) = 0 \quad (\text{A.55})$$

$$= \lim_{N \rightarrow \infty} S_N - \lim_{N \rightarrow \infty} S_{N-1} = 0, \quad (\text{A.56})$$

$$= \lim_{N \rightarrow \infty} \left(\frac{1}{N} \ln \Delta_{s(N+1)} - \frac{1}{N} \ln \Delta_{s(1)} \right) - \lim_{N \rightarrow \infty} \left(\frac{1}{N} \ln \Delta_{s(N)} - \frac{1}{N} \ln \Delta_{s(1)} \right) = 0, \quad (\text{A.57})$$

con lo que se llega a

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln \Delta_{s(N+1)} = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \Delta_{s(N)}. \quad (\text{A.58})$$

Entonces, sustituyendo en A.53 y de A.42 se tiene que

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln \Delta_{s(N)} = \ln P = \int_{-\pi}^{\pi} \ln f_s(e^{j\omega}) \frac{d\omega}{2\pi}. \quad (\text{A.59})$$

A.4.3 Predicción Lineal de Un Paso de Orden Variable Finito.

Se desea estimar el valor presente de un proceso regular $s[n]$ en términos de las n muestras más recientes de su pasado finito, empezando desde el origen. El proceso predictor tiene la forma

$$\hat{s}_n[n] = \sum_{k=1}^n h_k^n s[n-k]; \quad (\text{A.60})$$

y como puede apreciarse, la longitud n no es constante, a diferencia del predictor de la sección anterior de longitud fija N ; los coeficientes h_k^n también dependen de n . Por lo tanto, este predictor es un filtro lineal *variante con el tiempo*. El proceso de error se define como

$$\hat{\epsilon}_n[n] = s[n] - \hat{s}_n[n] = \sqrt{P_n} i[n], \quad (\text{A.61})$$

para el cual se verifica que

$$i[n] = \frac{1}{\sqrt{P_n}} \hat{\epsilon}_n[n] \quad E \{i^2[n]\} = 1 \quad \forall n. \quad (\text{A.62})$$

Los valores de P_n y de los coeficientes h_k^n pueden obtenerse con los mismos métodos de la predicción de orden finito. Por otra parte, considerando que $s[n]$ es regular, y denotando como $\gamma[n, k] \triangleq \gamma_k^n$ la respuesta, del filtro blanqueador de $s[k]$ variante con el tiempo, en el tiempo de observación n , al impulso $\delta[n-k]$ aplicado en el tiempo k , se tiene que [20, pag. 109]:

$$i[n] = \sum_{k=0}^n \gamma[n, k] s[k]; \quad (\text{A.63})$$

o en forma matricial

$$\begin{pmatrix} i_0 \\ i_1 \\ i_2 \\ \cdots \\ i_n \end{pmatrix} = \begin{pmatrix} \gamma_0^0 & 0 & 0 & \cdots & 0 \\ \gamma_0^1 & \gamma_1^1 & 0 & \cdots & 0 \\ \gamma_0^2 & \gamma_1^2 & \gamma_2^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma_0^n & \gamma_1^n & \gamma_2^n & \cdots & \gamma_n^n \end{pmatrix} \begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix};$$

$$\mathbf{i} = \hat{\Gamma}_{s(n+1)} \mathbf{s}, \quad (\text{A.64})$$

donde el subíndice $n+1$ indica el orden de la matriz. Se observa que si $s[n] = \delta[n-k]$, entonces $i[n] = \gamma_k^n$, $n \geq k$ ⁶. Al invertir la matriz $\hat{\Gamma}_{s(n+1)}$, se obtiene

$$\begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ \cdots \\ s_n \end{pmatrix} = \begin{pmatrix} l_0^0 & 0 & 0 & \cdots & 0 \\ l_0^1 & l_1^1 & 0 & \cdots & 0 \\ l_0^2 & l_1^2 & l_2^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ l_0^n & l_1^n & l_2^n & \cdots & l_n^n \end{pmatrix} \begin{pmatrix} i_0 \\ i_1 \\ i_2 \\ \vdots \\ i_n \end{pmatrix};$$

⁶Este proceso recibe el nombre de *Proceso de Innovaciones de Kalman*.

$$\underline{s} = \hat{\underline{L}}_{s(n+1)} \underline{i}; \quad (\text{A.65})$$

además, de esta ecuación se tiene que la matriz de correlaciones de orden $M = n + 1$, $\underline{\mathbf{R}}_{s(M)}$, del proceso regular S, está dada por

$$\underline{\mathbf{R}}_{s(M)} = \hat{\underline{L}}_{s(M)} \hat{\underline{L}}_{s(M)}^T, \quad (\text{A.66})$$

la cual se conoce como *factorización de Cholesky*. Haciendo $s[n] = \delta[n - k]$, de las expresiones A.60 y A.61 se tiene que

$$\begin{aligned} s[n] - \sum_{j=1}^n h_j^n s[n - k] &= \sqrt{P_n} i[n], \\ \delta[n - k] - \sum_{j=1}^n h_j^n \delta[n - k - j] &= \sqrt{P_n} \gamma_k^n; \end{aligned} \quad (\text{A.67})$$

de donde se obtienen las siguientes relaciones

$$\gamma_{n-k}^n = \frac{-1}{\sqrt{P_n}} h_k^n \quad ; \quad \gamma_n^n = \frac{1}{\sqrt{P_n}} \quad k \geq 1; \quad (\text{A.68})$$

con lo que finalmente la expresión A.61 toma la forma

$$\hat{\varepsilon}_n[n] = s[n] + \sum_{k=1}^n \sqrt{P_n} \gamma[n, n - k] s[n - k] = \sum_{k=0}^n \sqrt{P_n} \gamma_{n-k}^n s_{n-k}, \quad (\text{A.69})$$

con $\sqrt{P_i} \gamma_i^i = 1$, $i = 0, 1, \dots, n$; o bien

$$\begin{pmatrix} \hat{\varepsilon}_0 \\ \hat{\varepsilon}_1 \\ \hat{\varepsilon}_2 \\ \vdots \\ \hat{\varepsilon}_n \end{pmatrix} = \begin{pmatrix} \sqrt{P_0} \gamma_0^0 & 0 & 0 & \cdots & 0 \\ \sqrt{P_1} \gamma_0^1 & \sqrt{P_1} \gamma_1^1 & 0 & \vdots & 0 \\ \sqrt{P_2} \gamma_0^2 & \sqrt{P_2} \gamma_1^2 & \sqrt{P_2} \gamma_2^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sqrt{P_n} \gamma_0^n & \sqrt{P_n} \gamma_1^n & \sqrt{P_n} \gamma_2^n & \cdots & \sqrt{P_n} \gamma_n^n \end{pmatrix} \begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix}, \quad (\text{A.70})$$

Nótese la similitud de las ecuaciones A.43 y A.44 con A.69 y A.70, respectivamente. De hecho, cuando $n \rightarrow \infty$, los coeficientes en A.43 y A.69 son iguales y los procesos de error son iguales. Entonces se cumple que

$$\lim_{n \rightarrow \infty} \gamma_n^n = \gamma_0, \quad \lim_{n \rightarrow \infty} \gamma_{n-k}^n = \gamma_k; \quad (\text{A.71})$$

o sea, conforme n crece, los coeficientes *variantes* con el tiempo tienden a ser *invariantes* con el tiempo, es decir, sólo dependen de la diferencia entre el tiempo de aplicación de la entrada y el tiempo de observación de la salida. Considerando lo anterior, las secuencias de matrices $\underline{\mathbf{L}}_{s(M)}$, $\hat{\underline{\mathbf{L}}}_{s(M)}$, $\underline{\mathbf{\Gamma}}_{s(M)}$ y $\hat{\underline{\mathbf{\Gamma}}}_{s(M)}$, con $M = n + 1$, cumplen con A.7, y además

$$\lim_{M \rightarrow \infty} \left\| \underline{\mathbf{L}}_{s(M)} - \hat{\underline{\mathbf{L}}}_{s(M)} \right\|_F = 0,$$

y

$$\lim_{M \rightarrow \infty} \left\| \underline{\Gamma}_{s(M)} - \hat{\underline{\Gamma}}_{s(M)} \right\|_F = 0 ;$$

es decir

$$\underline{\mathbf{L}}_{s(M)} \sim \hat{\underline{\mathbf{L}}}_{s(M)} \quad y \quad \underline{\Gamma}_{s(M)} \sim \hat{\underline{\Gamma}}_{s(M)} . \quad (\text{A.72})$$

y por lo tanto, de A.13

$$\underline{\mathbf{L}}_{s(M)}^{-1} \sim \hat{\underline{\mathbf{L}}}_{s(M)}^{-1} \quad y \quad \underline{\Gamma}_{s(M)}^{-1} \sim \hat{\underline{\Gamma}}_{s(M)}^{-1} , \quad (\text{A.73})$$

resultados que se emplearán a continuación.

A.5 Aplicaciones.

En esta parte se aprovechan los resultados de las secciones anteriores para evaluar la expresión A.2. Se emplean las definiciones de Procesos Asintóticamente Regulares y Asintóticamente Estacionarios introducidas en la sección A.3.

A.5.1 Caso 1: $\hat{\mathbf{S}}$ y \mathbf{S} Regulares y Estacionarios.

De la ecuación A.49, se tiene que

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln \Delta_{s(N)} = \int_{-\pi}^{\pi} \ln f_s(e^{j\omega}) \frac{d\omega}{2\pi} .$$

Por otra parte, de la expresión A.66, se tiene que las matrices de correlaciones $\underline{\mathbf{R}}_{s(N)}$ y $\underline{\mathbf{R}}_{\hat{s}(N)}$ de los procesos \mathbf{S} y $\hat{\mathbf{S}}$, respectivamente, pueden escribirse como ⁷

$$\underline{\mathbf{R}}_{s(N)} = \hat{\underline{\mathbf{L}}}_{s(N)} \hat{\underline{\mathbf{L}}}_{s(N)}^T \quad \underline{\mathbf{R}}_{\hat{s}(N)} = \hat{\underline{\mathbf{L}}}_{\hat{s}(N)} \hat{\underline{\mathbf{L}}}_{\hat{s}(N)}^T ,$$

por lo que se tiene que

$$\begin{aligned} \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \underline{\mathbf{R}}_{s(N)}^{-1} \underline{\mathbf{R}}_{\hat{s}(N)} \right\} &= \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ (\hat{\underline{\mathbf{L}}}_{s(N)} \hat{\underline{\mathbf{L}}}_{s(N)}^T)^{-1} (\hat{\underline{\mathbf{L}}}_{\hat{s}(N)} \hat{\underline{\mathbf{L}}}_{\hat{s}(N)}^T) \right\} \\ &= \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \hat{\underline{\mathbf{L}}}_{s(N)}^{-1} \hat{\underline{\mathbf{L}}}_{\hat{s}(N)} \hat{\underline{\mathbf{L}}}_{\hat{s}(N)}^T \hat{\underline{\mathbf{L}}}_{s(N)}^{-T} \right\} . \end{aligned} \quad (\text{A.74})$$

Puesto que estas matrices son asintóticamente equivalentes a las asociadas a los filtros de innovaciones de los procesos $\hat{\mathbf{S}}$ y \mathbf{S} , o sea

$$\hat{\underline{\mathbf{L}}}_{s(N)}^{-1} \sim \underline{\mathbf{L}}_{s(N)}^{-1} , \quad \hat{\underline{\mathbf{L}}}_{s(N)}^{-T} \sim \underline{\mathbf{L}}_{s(N)}^{-T} \quad y \quad \hat{\underline{\mathbf{L}}}_{\hat{s}(N)} \sim \underline{\mathbf{L}}_{\hat{s}(N)} ,$$

y por la propiedad 5 de la equivalencia asintótica, se llega a

$$\frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \hat{\underline{\mathbf{L}}}_{s(N)}^{-1} \hat{\underline{\mathbf{L}}}_{\hat{s}(N)} \hat{\underline{\mathbf{L}}}_{\hat{s}(N)}^T \hat{\underline{\mathbf{L}}}_{s(N)}^{-T} \right\} = \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \underline{\mathbf{L}}_{s(N)}^{-1} \underline{\mathbf{L}}_{\hat{s}(N)} \underline{\mathbf{L}}_{\hat{s}(N)}^T \underline{\mathbf{L}}_{s(N)}^{-T} \right\} ;$$

⁷Los subíndices indican el proceso al cual pertenece la matriz

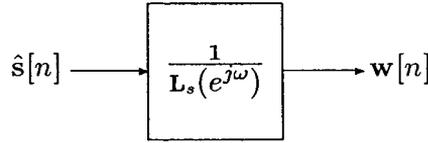


Figura A.3: Sistema Lineal III

$$\frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \mathbf{R}_{s(N)}^{-1} \mathbf{R}_{\hat{s}(N)} \right\} = \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \mathbf{L}_{s(N)}^{-1} \mathbf{R}_{\hat{s}(N)} \mathbf{L}_{s(N)}^{-T} \right\} ; \quad (\text{A.75})$$

el producto matricial involucrado en esta ecuación puede compararse con la ecuación A.20, aunque ésta corresponda a un proceso asintóticamente estacionario, y de esta comparación, se observa que este producto matricial puede asociarse a la matriz de correlaciones de un proceso de salida, \mathbf{W} por ejemplo, de un sistema como el que se muestra en la Fig.A.3, con función de transferencia el inverso del filtro de innovaciones $\mathbf{L}_s(e^{j\omega})$ del proceso regular \mathbf{S} , y entrada el proceso regular $\hat{\mathbf{S}}$. Por tanto, la expresión A.75 es la *potencia promedio por muestra* P_w [1, pag. D.4] del proceso de salida \mathbf{W} del sistema de la Fig.A.3, es decir

$$P_z = \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \mathbf{R}_{s(N)}^{-1} \mathbf{R}_{\hat{s}(N)} \right\} = \frac{1}{2} \int_{-\pi}^{\pi} \frac{f_{\hat{s}}(e^{j\omega})}{f_s(e^{j\omega})} \frac{d\omega}{2\pi} . \quad (\text{A.76})$$

Por otra parte, de la Fig.2.1 con $n_0 = -\infty$, y empleando los resultados de la sección 1.3 del capítulo 1, se tiene que

$$\bar{\mathbf{H}}_{\hat{s}} = \bar{\mathbf{H}}_x + \frac{1}{2} \int_{-\pi}^{\pi} \ln |\mathbf{G}(e^{j\omega})|^2 \frac{d\omega}{2\pi} ,$$

por lo que, con los resultados A.49 y A.76, y sumando y restando la expresión

$$\frac{1}{2} \int_{-\pi}^{\pi} \ln f_x(e^{j\omega}) \frac{d\omega}{2\pi} ,$$

donde $f_x(e^{j\omega})$ es la función de densidad espectral de potencia del proceso X , a la ecuación A.2, se tiene que

$$\begin{aligned} \bar{\mathbf{D}}(\hat{\mathbf{S}}, \mathbf{S}) &= \frac{1}{2} \int_{-\pi}^{\pi} \left(\frac{f_{\hat{s}}(e^{j\omega})}{f_s(e^{j\omega})} - \ln \left\{ \frac{|\mathbf{G}(e^{j\omega})|^2 f_x(e^{j\omega})}{f_s(e^{j\omega})} \right\} - 1 \right) \frac{d\omega}{2\pi} \\ &\quad - \bar{\mathbf{H}}_x + \frac{1}{2} \ln(2\pi e) + \frac{1}{2} \int_{-\pi}^{\pi} \ln f_x(e^{j\omega}) \frac{d\omega}{2\pi} , \end{aligned} \quad (\text{A.77})$$

donde el primer término de la ecuación anterior se reduce a un medio de la *Distorsión de Itakura-Saito*:

$$d_{IS}(\hat{\mathbf{S}}, \mathbf{S}) = \frac{1}{2} \int_{-\pi}^{\pi} \left(\frac{f_{\hat{s}}(e^{j\omega})}{f_s(e^{j\omega})} - \ln \left\{ \frac{f_{\hat{s}}(e^{j\omega})}{f_s(e^{j\omega})} \right\} - 1 \right) \frac{d\omega}{2\pi} . \quad (\text{A.78})$$

Debe tenerse en cuenta en este resultado, que el tiempo de aplicación de la entrada es $n_0 = -\infty$.

A.5.2 Caso 2: \hat{S} Asintóticamente Regular y Asintóticamente Estacionario y S Regular y Estacionario.

Ahora, el tiempo de aplicación es $n_0 = 0$, y la matriz de correlaciones de \hat{S} es $\mathbf{R}_{\hat{S}(N)}^a = \mathbf{L}_{\hat{S}(N)} \mathbf{L}_{\hat{S}(N)}^T$, y como $\mathbf{L}_{\hat{S}(N)} \sim \hat{\mathbf{L}}_{\hat{S}(N)}$, se llega directamente a las expresiones A.75. Sin embargo, de la misma figura 2.1 y de la sección 1.3 del capítulo 1 se tiene ahora que

$$\bar{\mathbf{H}}_{\hat{S}} = \bar{\mathbf{H}}_x + \frac{1}{2} \ln |g_0|^2,$$

y siguiendo un procedimiento análogo al de la sección anterior, se llega a

$$\begin{aligned} \bar{D}(\hat{S}, S) = & \frac{1}{2} \int_{-\pi}^{\pi} \left(\frac{f_{\hat{S}}(e^{j\omega})}{f_S(e^{j\omega})} - \ln \left\{ \frac{g_0^2 f_x(e^{j\omega})}{f_S(e^{j\omega})} \right\} - 1 \right) \frac{d\omega}{2\pi} \\ & - \bar{\mathbf{H}}_x + \frac{1}{2} \ln(2\pi e) + \frac{1}{2} \int_{-\pi}^{\pi} \ln f_x(e^{j\omega}) \frac{d\omega}{2\pi}, \end{aligned} \quad (\text{A.79})$$

donde g_0^2 es la respuesta al impulso del filtro $G(e^{j\omega})$ evaluada en cero. Si $G(e^{j\omega})$ es de fase mínima, entonces A.79 es igual a A.77. La ecuación A.79 aparece en [2, pag. 17] como el *Teorema 2*⁸.

A.5.3 Conclusiones.

Se ha visto que al considerar tiempos de observación infinitamente grandes, las trazas de las matrices de correlación de los procesos asintóticamente regulares y asintóticamente y estacionarios se comportan similarmente a las de los procesos regulares y estacionarios; es decir, los efectos transitorios y por la dependencia del tiempo desaparecen. Adicionalmente, al considerar filtros de fase mínima, en ambos tipos de procesos la Información de Discriminación Normalizada tiene la misma expresión. Esto es importante ya que permite construir modelos matemáticos de procesos que no son regulares, al menos hablando, a partir de resultados ya existentes en procesos regulares.

⁸Lepe también la ocupa en su Proyecto de Investigación de Doctorado.

Apéndice B

B.1 Introducción.

En esta parte, los conceptos introducidos en el Apéndice A se extienden a dos dimensiones. En la sección B.2 se presentan algunos conceptos básicos y resultados sobre Campos Aleatorios Regulares. En la sección B.3 se presentan algunos resultados de Predicción Lineal en Dos Dimensiones y en la sección B.4 se emplean estos resultados en la reducción de la Información de Discriminación Normalizada.

B.2 Conceptos Preliminares.

B.2.1 Regiones Causales.

El concepto de causalidad no se extiende de modo natural a dos dimensiones; sin embargo, para consideraciones teóricas, resulta conveniente definirlo. La región causal \mathcal{R} se define en [24, 10] como el conjunto de puntos (m, n) en el plano que satisfacen

- $(m, n) \in \mathcal{R}$ si y sólo si $(-m, -n) \notin \mathcal{R}$.
- Si $(m, n) \in \mathcal{R}$ y $(m', n') \in \mathcal{R}$ implica que $(m + m', n + n') \in \mathcal{R}$.
- $(0, 0) \in \mathcal{R}$.

Ejemplos de regiones causales se muestran en la Fig.B.1. Un *sistema causal* es aquel cuya respuesta al impulso tiene una región de soporte que toma cualquier forma dentro de la región causal \mathcal{R} [26, pag. 482]. Como caso particular, se tiene la región formada únicamente por el primer cuadrante del plano mn , la cual se muestra en la Fig.B.2. De acuerdo a la definición, ejemplos de posibles regiones *semicausales* y *no causales* [8, pag. 205], se muestran en la Fig.B.3. En Predicción Lineal en Dos Dimensiones se emplean el mismo tipo de regiones pero se excluye el origen ya que éste es el que se predice.

B.2.2 Campos Regulares

Sea un Sistema Lineal Invariante al Corrimiento (SLIC) Bidimensional $H(e^{j\omega_1}, e^{j\omega_2})$ causal y estable y con inverso único también causal y estable (en una dimensión se dice que es de fase

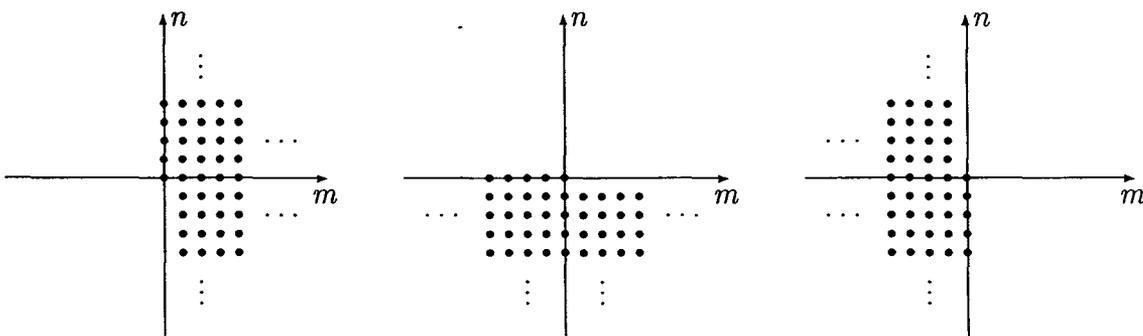


Figura B.1: Regiones Causales

mínima) con respuesta al impulso $h(m, n)$. Para este trabajo, se define a un campo homogéneo X como regular si cumple que

- sea el campo de salida del sistema $H(e^{j\omega_1}, e^{j\omega_2})$ cuando la entrada a éste sea una sucesión de v.a.i.i.d de media cero y variancia unitaria (campo de ruido blanco estricto $I(m, n)$) aplicada desde $m_0 = -\infty$, $n_0 = -\infty$ al mismo tiempo, como se indica en la Fig.B.4.

La consecuencia directa de esto es que la función de densidad espectral de potencia del campo X , $f_X(e^{j\omega_1}, e^{j\omega_2})$, puede factorizarse como

$$f_X(e^{j\omega_1}, e^{j\omega_2}) = |H(e^{j\omega_1}, e^{j\omega_2})|^2 = H(e^{j\omega_1}, e^{j\omega_2})H^*(e^{j\omega_1}, e^{j\omega_2}); \quad (\text{B.1})$$

donde $H(e^{j\omega_1}, e^{j\omega_2})$ es la transformada bidimensional de Fourier de la secuencia $h(m, n)$ de fase mínima [27, pag. 203]. Esto significa que la función de autocorrelación del campo regular X , $R_X(m, n)$, puede expresarse como la convolución de $h(m, n)$ y $h(-m, -n)$ (se supone que $h(m, n)$ es real), es decir, si

$$X(m, n) = \sum_{\forall k, l \in \hat{\mathcal{R}}} h(k, l) I(k - m, l - n) \quad \forall m, n \in \hat{\mathcal{R}}$$

y si

$$R_X(m, n) = E \{X(k, l) X(k + m, l + n)\}$$

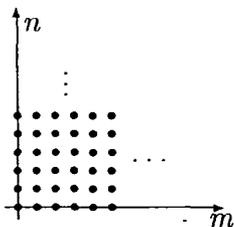


Figura B.2: Región en el cuadrante I.

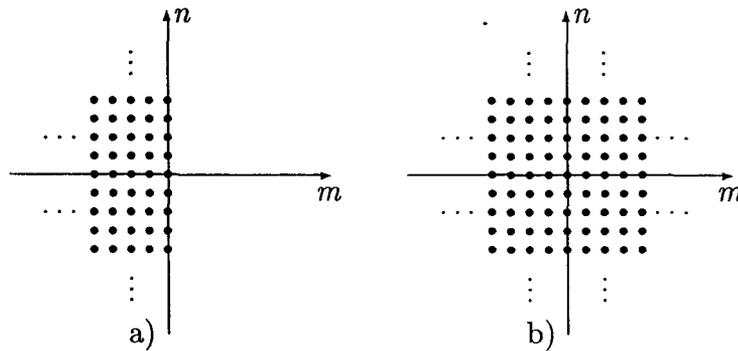


Figura B.3: a) Región Semicausal, b) Región Acausal

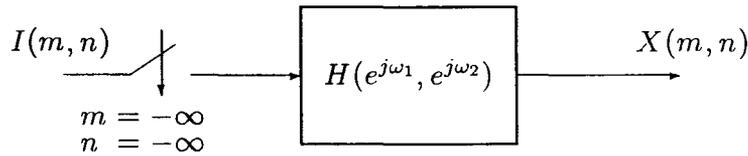


Figura B.4: Definición de Campo Regular

$$R_X(m, n) = E \left\{ \sum_{\forall i, j \in \hat{\mathcal{R}}} h(i, j) I(k - i, l - j) \sum_{\forall i', j' \in \hat{\mathcal{R}}} h(i', j') I(k + m - i', l + n - j') \right\}$$

se cumple entonces que

$$R_X(m, n) = h(m, n) * h(-m, -n) . \quad (\text{B.2})$$

Debe tenerse en cuenta que, por la condición de homogeneidad $R_X(m, n)$ sólo es función de la diferencia entre los puntos de observación, es decir, $m = m_1 - m_2$ y $n = n_1 - n_2$ si $P_1(m_1, n_1)$ y $P_2(m_2, n_2)$ son dos puntos de la región de soporte del campo. En general,

$$R_X(\underline{\mathbf{g}}) = R_X(\underline{\mathbf{j}} - \underline{\mathbf{k}}) \quad \forall \underline{\mathbf{j}}, \underline{\mathbf{k}} \in \hat{\mathcal{R}} \quad (\text{B.3})$$

B.2.3 Matrices de Correlaciones de Campos Homogéneos

Matrices de Permutación

Estas matrices se emplean para cambiar de posición elementos de otra matriz. Se construyen tomando un ordenamiento diferente de las columnas o renglones de la matriz identidad [18, pag. 73]. Por ejemplo, considérese la siguiente matriz de permutación \underline{P} de 4×4 ,

$$\underline{P} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

El primer renglón de \underline{P} corresponde al tercer renglón de la matriz identidad de orden 4, el segundo renglón al cuarto, mientras que el tercero y el cuarto corresponden al segundo y primero, respectivamente, de la misma. Entonces, el ordenamiento *por renglones* de \underline{P} , en términos de los renglones de la matriz identidad, es $\{3, 4, 2, 1\}$. Similarmente, la primera columna de \underline{P} corresponde a la cuarta columna de la matriz identidad, la segunda a la tercera, la tercera a la primera y la cuarta a la segunda, por lo que el ordenamiento *por columnas* de \underline{P} , en términos de las columnas de la matriz identidad, es $\{4, 3, 2, 1\}$. Sea \underline{A} una matriz de 4×4 escrita en términos de sus renglones $a_1^T, a_2^T, a_3^T, a_4^T$,

$$\underline{A} = \begin{pmatrix} a_1^T \\ a_2^T \\ a_3^T \\ a_4^T \end{pmatrix}$$

donde las a_i son los vectores que la conforman; se tiene entonces que

$$\underline{P}\underline{A} = \begin{pmatrix} a_3^T \\ a_4^T \\ a_2^T \\ a_1^T \end{pmatrix};$$

análogamente, expresando \underline{A} en términos de sus columnas como

$$\underline{A} = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 \end{pmatrix}$$

se cumple que

$$\underline{P}\underline{A} = \begin{pmatrix} a_4 & a_3 & a_1 & a_2 \end{pmatrix}.$$

Además, el producto de dos matrices de permutación es otra matriz de permutación y

$$\underline{P}\underline{P}^T = \underline{P}^T\underline{P} = \underline{I}, \quad (\text{D.4})$$

lo que significa que $\underline{P}^{-1} = \underline{P}^T$; es decir, las matrices de permutación son ortonormales, por lo cual

$$\det \underline{P} = \pm 1.$$

Finalmente, a partir de una matriz identidad de $N \times N$, se pueden construir $N!$ matrices de permutación. Las matrices de permutación son importantes en el estudio de los campos aleatorios, ya que permiten relacionar las diferentes matrices de correlaciones de un campo cuya respuesta al impulso tiene una cierta región de soporte, como se verá a continuación.

Sea $X(m,n)$ un campo regular y homogéneo¹, entonces su función de correlación dependerá sólo de la posición relativa entre las observaciones. Esto significa que la ubicación de la región

¹La definición de Campos Homogéneos se da en el Capítulo 3 y corresponde a lo que en procesos unidimensionales es estacionaridad.

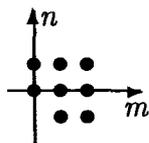


Figura B.5: Ejemplo de Región de Soporte de $h(m, n)$

de soporte de la respuesta al impulso $h(m, n)$ asociada a $X(m, n)$, no importa, lo que a su vez implica que la matriz de correlación debe tener ciertas características de simetría, como sucede en una dimensión. El número de puntos N que componen la región de soporte se pueden ordenar, en un vector \underline{x} , por columnas, renglones, diagonales o, además, en cualquier otro modo. Un cambio en el ordenamiento del vector \underline{x} producirá una matriz de correlaciones diferente; no obstante, un cierto ordenamiento a del vector \underline{x} , denotado \underline{x}_a , es igual al producto de una matriz de *permutación* \underline{P}_{ab} de orden $N \times N$ por otro vector \underline{x}_b , el cual resulta al ordenar al vector \underline{x} en la forma b , es decir

$$\underline{x}_a = \underline{P}_{ab} \underline{x}_b ;$$

por tanto, se cumple también que

$$\underline{x}_b = \underline{P}_{ab}^T \underline{x}_a = \underline{P}_{ab}^{-1} \underline{x}_a .$$

Considerando lo anterior, existen $N!$ formas de construir el vector \underline{x} y $N!$ matrices de permutación \underline{P} . Como éstas son ortonormales, la matriz de correlación de un campo homogéneo, $\underline{R}_{a(N)}$, asociada a un cierto ordenamiento a de su región de soporte, es igual a la matriz $\underline{R}_{b(N)}$ asociada a otro ordenamiento b premultiplicada por la correspondiente matriz de permutación \underline{P}_{ab} y posmultiplicada por \underline{P}_{ab}^{-1} , o sea

$$\underline{R}_{a(N)} = E \{ \underline{x}_a \underline{x}_a^T \} = E \{ \underline{P}_{ab} \underline{x}_b (\underline{P}_{ab} \underline{x}_b)^T \} = E \{ \underline{P}_{ab} \underline{x}_b \underline{x}_b^T \underline{P}_{ab}^T \} = \underline{P}_{ab} \underline{R}_{b(N)} \underline{P}_{ab}^T ; \quad (\text{B.5})$$

con lo que se obtiene que

$$\Delta_{Ra} = \Delta_{Pab} \Delta_{Rb} \Delta_{Pab}^T , \quad (\text{B.6})$$

donde Δ_{Ra} , Δ_{Pab} , Δ_{Rb} , Δ_{Pab}^T son los determinantes de las matrices $\underline{R}_{a(N)}$, \underline{P}_{ab} , $\underline{R}_{b(N)}$, \underline{P}_{ab}^T , respectivamente. Lo anterior significa que *los determinantes de todas las posibles matrices de correlación de un campo con una cierta región de soporte son iguales*. Como ejemplo, considérese la región mostrada en la Fig.B.5. Formando el vector \underline{x}^T con los elementos de la región, empezando por el extremo inferior izquierdo, y procediendo de abajo hacia arriba y luego de izquierda a derecha, de acuerdo a la figura, toma la forma

$$\underline{x}_a^T = \{ X(0, 0), X(0, 1), X(1, -1), X(1, 0), X(1, 1), X(2, -1), X(2, 0), X(2, 1) \} ;$$

y la matriz de correlaciones $\underline{\mathbf{R}}_{a(N)}$ asociada a este ordenamiento toma la forma (notación compacta donde la soberrraya indica que el número es negativo):

$$\underline{\mathbf{R}}_{a(N)} = \begin{pmatrix} \begin{pmatrix} R_{0,0} & R_{0,\bar{1}} \\ R_{0,1} & R_{0,0} \end{pmatrix} & \begin{pmatrix} R_{\bar{1},1} & R_{\bar{1},0} & R_{\bar{1},\bar{1}} \\ R_{\bar{1},2} & R_{\bar{1},1} & R_{\bar{1},0} \end{pmatrix} & \begin{pmatrix} R_{\bar{2},1} & R_{\bar{2},0} & R_{\bar{2},\bar{1}} \\ R_{\bar{2},2} & R_{\bar{2},1} & R_{\bar{2},0} \end{pmatrix} \\ \begin{pmatrix} R_{1,\bar{1}} & R_{1,\bar{2}} \\ R_{1,0} & R_{1,\bar{1}} \\ R_{1,1} & R_{1,0} \end{pmatrix} & \begin{pmatrix} R_{0,0} & R_{0,\bar{1}} & R_{0,\bar{2}} \\ R_{0,1} & R_{0,0} & R_{0,\bar{1}} \\ R_{0,2} & R_{0,1} & R_{0,0} \end{pmatrix} & \begin{pmatrix} R_{\bar{1},0} & R_{\bar{1},\bar{1}} & R_{\bar{1},\bar{2}} \\ R_{\bar{1},1} & R_{\bar{1},0} & R_{\bar{1},\bar{1}} \\ R_{\bar{1},2} & R_{\bar{1},1} & R_{\bar{1},0} \end{pmatrix} \\ \begin{pmatrix} R_{2,\bar{1}} & R_{2,\bar{2}} \\ R_{2,0} & R_{2,\bar{1}} \\ R_{2,1} & R_{2,0} \end{pmatrix} & \begin{pmatrix} R_{1,0} & R_{1,\bar{1}} & R_{1,\bar{2}} \\ R_{1,1} & R_{1,0} & R_{1,\bar{1}} \\ R_{1,2} & R_{1,1} & R_{1,0} \end{pmatrix} & \begin{pmatrix} R_{0,0} & R_{0,\bar{1}} & R_{0,\bar{2}} \\ R_{0,1} & R_{0,0} & R_{0,\bar{1}} \\ R_{0,2} & R_{0,1} & R_{0,0} \end{pmatrix} \end{pmatrix}. \quad (\text{B.7})$$

La estructura de B.7 corresponde, entonces, a la estructura general asociada a la matriz de correlaciones de un campo homogéneo. De acuerdo con B.2 y la Fig.B.5, se tiene que

$$\begin{aligned} R_{0,0} &= h_{0,0}^2 + h_{0,1}^2 + h_{1,\bar{1}}^2 + h_{1,0}^2 + h_{1,1}^2 + h_{2,\bar{1}}^2 + h_{2,0}^2 + h_{2,1}^2 \\ R_{0,1} &= h_{2,1}h_{2,0} + h_{2,0}h_{2,\bar{1}} + h_{1,1}h_{1,0} + h_{1,0}h_{1,\bar{1}} + h_{0,1}h_{0,0} = R_{0,\bar{1}} \\ R_{0,2} &= h_{2,1}h_{2,\bar{1}} + h_{1,1}h_{1,\bar{1}} = R_{0,\bar{2}} \\ R_{1,0} &= h_{2,1}h_{1,1} + h_{2,\bar{1}}h_{1,\bar{1}} + h_{2,0}h_{1,0} + h_{1,1}h_{0,1} + h_{1,0}h_{0,0} = R_{\bar{1},0} \\ R_{1,1} &= h_{2,1}h_{1,0} + h_{2,0}h_{1,\bar{1}} + h_{1,1}h_{0,0} = R_{\bar{1},\bar{1}} \\ R_{1,2} &= h_{2,1}h_{1,\bar{1}} = R_{\bar{1},\bar{2}} \\ R_{1,\bar{2}} &= h_{2,\bar{1}}h_{1,1} + h_{1,\bar{1}}h_{0,1} = R_{\bar{1},2} \\ R_{1,\bar{1}} &= h_{2,0}h_{1,1} + h_{2,\bar{1}}h_{1,0} + h_{1,0}h_{0,1} + h_{1,\bar{1}}h_{0,0} = R_{\bar{1},1} \\ R_{2,0} &= h_{2,1}h_{0,1} + h_{2,0}h_{0,0} = R_{\bar{2},0} \\ R_{2,1} &= h_{2,1}h_{0,0} = R_{\bar{2},\bar{1}} \\ R_{2,\bar{2}} &= h_{2,\bar{1}}h_{0,1} = R_{\bar{2},2} \\ R_{2,\bar{1}} &= h_{0,1}h_{2,0} + h_{2,\bar{1}}h_{0,0} = R_{\bar{2},1} \end{aligned} \quad (\text{B.8})$$

Si ahora se toma otro ordenamiento para la misma región por renglones, de izquierda a derecha y de arriba a abajo, se tiene que el vector toma la forma

$$\underline{\mathbf{x}}_b^T = \{X(0,1), X(1,1), X(2,1), X(0,0), X(1,0), X(2,0), X(1,-1), X(2,-1)\};$$

y la correspondiente matriz de correlaciones queda

$$\underline{\mathbf{R}}_{b(N)} = \begin{pmatrix} \begin{pmatrix} R_{0,0} & R_{\bar{1},0} & R_{\bar{2},0} \\ R_{1,0} & R_{0,0} & R_{\bar{1},0} \\ R_{2,0} & R_{1,0} & R_{0,0} \end{pmatrix} & \begin{pmatrix} R_{0,1} & R_{\bar{1},1} & R_{\bar{2},1} \\ R_{1,1} & R_{0,1} & R_{\bar{1},1} \\ R_{2,1} & R_{1,1} & R_{0,1} \end{pmatrix} & \begin{pmatrix} R_{\bar{1},2} & R_{\bar{2},2} \\ R_{0,2} & R_{\bar{1},2} \\ R_{1,2} & R_{0,2} \end{pmatrix} \\ \begin{pmatrix} R_{0,\bar{1}} & R_{\bar{1},\bar{1}} & R_{\bar{2},\bar{1}} \\ R_{1,\bar{1}} & R_{0,\bar{1}} & R_{\bar{1},\bar{1}} \\ R_{2,\bar{1}} & R_{1,\bar{1}} & R_{0,\bar{1}} \end{pmatrix} & \begin{pmatrix} R_{0,0} & R_{\bar{1},0} & R_{\bar{2},0} \\ R_{1,0} & R_{0,0} & R_{\bar{1},0} \\ R_{2,0} & R_{1,0} & R_{0,0} \end{pmatrix} & \begin{pmatrix} R_{\bar{1},1} & R_{\bar{2},1} \\ R_{0,1} & R_{\bar{1},1} \\ R_{1,1} & R_{0,1} \end{pmatrix} \\ \begin{pmatrix} R_{1,\bar{2}} & R_{0,\bar{2}} & R_{\bar{1},\bar{2}} \\ R_{2,\bar{2}} & R_{1,\bar{2}} & R_{0,\bar{2}} \end{pmatrix} & \begin{pmatrix} R_{1,\bar{1}} & R_{0,\bar{1}} & R_{\bar{1},\bar{1}} \\ R_{2,\bar{1}} & R_{1,\bar{1}} & R_{0,\bar{1}} \end{pmatrix} & \begin{pmatrix} R_{0,0} & R_{\bar{1},0} \\ R_{1,0} & R_{0,0} \end{pmatrix} \end{pmatrix}. \quad (\text{B.9})$$

Por supuesto, ambas matrices tienen los mismo elementos y se relacionan por

$$\mathbf{R}_{a(N)} = \mathbf{P}_{ab} \mathbf{R}_{b(N)} \mathbf{P}_{ab}^T ; \quad (\text{B.10})$$

donde \mathbf{P}_{ab} tiene la forma

$$\mathbf{P}_{ab} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} ;$$

además, $\mathbf{R}_{a(N)}$ y $\mathbf{R}_{b(N)}$ tienen la misma traza y el mismo determinante.

En el caso particular de que se trate de una región de soporte cuadrada, es decir, mismo número de observaciones en ambos ejes, la matriz de correlaciones resulta ser doblemente Toeplitz por bloques, es decir, la estructura por bloques (indicados por paréntesis interiores), es Toeplitz y cada bloque a su vez es Toeplitz. Para ejemplificar lo anterior, sea una región cuadrada de 3×3 , ordenada por columnas de abajo hacia arriba y luego de izquierda a derecha, empezando en el punto inferior izquierdo, la matriz de correlaciones toma la forma (en este caso, las submatrices se forman por la separación por columnas del vector que contiene todos los puntos, con el ordenamiento propuesto)

$$\mathbf{R}_{(N)} = \begin{pmatrix} \begin{pmatrix} R_{0,0} & R_{0,1} & R_{0,2} \\ R_{0,\bar{1}} & R_{0,0} & R_{0,1} \\ R_{0,\bar{2}} & R_{0,\bar{1}} & R_{0,0} \end{pmatrix} & \begin{pmatrix} R_{1,0} & R_{1,1} & R_{1,2} \\ R_{1,\bar{1}} & R_{1,0} & R_{1,1} \\ R_{1,\bar{2}} & R_{1,\bar{1}} & R_{1,0} \end{pmatrix} & \begin{pmatrix} R_{2,0} & R_{2,1} & R_{2,2} \\ R_{2,\bar{1}} & R_{2,0} & R_{2,1} \\ R_{2,\bar{2}} & R_{2,\bar{1}} & R_{2,0} \end{pmatrix} \\ \begin{pmatrix} R_{\bar{1},0} & R_{\bar{1},1} & R_{\bar{1},2} \\ P_{\bar{1},0} & P_{\bar{1},1} & P_{\bar{1},2} \end{pmatrix} & \begin{pmatrix} R_{0,0} & R_{0,1} & R_{0,2} \\ P_{0,0} & P_{0,1} & P_{0,2} \end{pmatrix} & \begin{pmatrix} R_{1,0} & R_{1,1} & R_{1,2} \\ P_{1,0} & P_{1,1} & P_{1,2} \end{pmatrix} \\ \begin{pmatrix} R_{\bar{1},\bar{2}} & R_{\bar{1},\bar{1}} & R_{\bar{1},0} \end{pmatrix} & \begin{pmatrix} R_{0,\bar{2}} & R_{0,\bar{1}} & R_{0,0} \end{pmatrix} & \begin{pmatrix} R_{1,\bar{2}} & R_{1,\bar{1}} & R_{1,0} \end{pmatrix} \\ \begin{pmatrix} R_{\bar{2},0} & R_{\bar{2},1} & R_{\bar{2},2} \\ R_{\bar{2},\bar{1}} & R_{\bar{2},0} & R_{\bar{2},1} \\ R_{\bar{2},\bar{2}} & R_{\bar{2},\bar{1}} & R_{\bar{2},0} \end{pmatrix} & \begin{pmatrix} R_{\bar{1},0} & R_{\bar{1},1} & R_{\bar{1},2} \\ R_{\bar{1},\bar{1}} & R_{\bar{1},0} & R_{\bar{1},1} \\ R_{\bar{1},\bar{2}} & R_{\bar{1},\bar{1}} & R_{\bar{1},0} \end{pmatrix} & \begin{pmatrix} R_{0,0} & R_{0,1} & R_{0,2} \\ R_{0,\bar{1}} & R_{0,0} & R_{0,1} \\ R_{0,\bar{2}} & R_{0,\bar{1}} & R_{0,0} \end{pmatrix} \end{pmatrix} \quad (\text{B.11})$$

En este caso, existen $9!$ posibles matrices de correlación para esta región cuadrada y otras tantas de permutación; sin embargo, puesto que, independientemente de cualquier ordenación que se tome, sobre la diagonal principal siempre se tienen puntos de observación iguales, *todas tienen la misma traza* y, de acuerdo a las propiedades de las matrices de permutación, *el mismo determinante*.

B.2.4 Campos Asintóticamente Regulares y Asintóticamente Homogéneos.

Sea un SLIC bidimensional de fase mínima como el mostrado en la Fig.B.4 cuya respuesta al impulso $h(m, n)$ tiene la región de soporte de $N = 8$ puntos mostrada en la Fig.B.5. Considerando

inicialmente como entrada un campo $X(\underline{t})$ con la misma región $\hat{\mathcal{R}}$ de soporte y observando la salida en la misma región, de la ecuación de convolución bidimensional introducida en la sección 3.2.2 del capítulo tres

$$Y(m, n) = \sum_{\forall k, l \in \hat{\mathcal{R}}} h(k, l) X(m - k, n - l) \quad \forall m, n \in \hat{\mathcal{R}},$$

se tiene que la salida Y está dada por

$$\begin{aligned} Y(0, 0) &= h(0, 0)X(0, 0) \\ Y(0, 1) &= h(0, 1)X(0, 0) + h(0, 0)X(0, 1) \\ Y(1, -1) &= h(1, -1)X(0, 0) + h(0, 0)X(1, -1) \\ Y(1, 0) &= h(1, -1)X(0, 1) + h(1, 0)X(0, 0) + h(0, 1)X(1, -1) + \\ &\quad h(0, 0)X(1, 0) \\ Y(1, 1) &= h(1, 1)X(0, 0) + h(1, 0)X(0, 1) + h(0, 1)X(1, 0) + \\ &\quad h(0, 0)X(1, 1) \\ Y(2, -1) &= h(2, -1)X(0, 0) + h(1, 0)X(1, -1) + h(1, -1)X(1, 0) + \\ &\quad h(0, 0)X(2, -1) \\ Y(2, 0) &= h(2, 0)X(0, 0) + h(2, -1)X(0, 1) + h(1, 1)X(1, -1) + \\ &\quad h(1, 0)X(1, 0) + h(1, -1)X(1, 1) + h(0, 1)X(2, -1) + \\ &\quad h(0, 0)X(2, 0) \\ Y(2, 1) &= h(2, 1)X(0, 0) + h(2, 0)X(0, 1) + h(1, 1)X(1, 0) + \\ &\quad h(1, 0)X(1, 1) + h(0, 1)X(2, 0) + h(0, 0)X(2, 1); \end{aligned} \quad (\text{B.12})$$

lo que matricialmente puede escribirse como (notación compacta):

$$\begin{pmatrix} Y_{0,0} \\ Y_{0,1} \\ Y_{1,\bar{1}} \\ Y_{1,0} \\ Y_{1,1} \\ Y_{2,\bar{1}} \\ Y_{2,0} \\ Y_{2,1} \end{pmatrix} = \begin{pmatrix} h_{0,0} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & h_{0,1} & 0 & 0 & 0 & 0 & 0 & 0 \\ h_{1,\bar{1}} & 0 & h_{0,0} & 0 & 0 & 0 & 0 & 0 \\ h_{1,0} & h_{1,\bar{1}} & h_{0,1} & h_{0,0} & 0 & 0 & 0 & 0 \\ h_{1,1} & h_{1,0} & 0 & h_{0,1} & h_{0,0} & 0 & 0 & 0 \\ h_{2,\bar{1}} & 0 & h_{1,0} & h_{1,\bar{1}} & 0 & h_{0,0} & 0 & 0 \\ h_{2,0} & h_{2,\bar{1}} & h_{1,1} & h_{1,0} & h_{1,\bar{1}} & h_{0,1} & h_{0,0} & 0 \\ h_{2,1} & h_{2,0} & 0 & h_{1,1} & h_{1,0} & 0 & h_{0,1} & h_{0,0} \end{pmatrix} \begin{pmatrix} X_{0,0} \\ X_{0,1} \\ X_{1,\bar{1}} \\ X_{1,0} \\ X_{1,1} \\ X_{2,\bar{1}} \\ X_{2,0} \\ X_{2,1} \end{pmatrix}, \quad (\text{B.13})$$

$$\underline{\mathbf{y}} = \mathcal{H} \underline{\mathbf{x}}. \quad (\text{B.14})$$

La matriz de correlaciones asociada a este proceso toma la forma, con N igual al número de puntos en la región de soporte,

$$\underline{\mathbf{R}}_{Y(N)} = E \{ \underline{\mathbf{y}} \underline{\mathbf{y}}^T \} = \mathcal{H} \underline{\mathbf{R}}_{X(N)} \mathcal{H}^T; \quad (\text{B.15})$$

con esta expresión, es posible calcular la potencia promedio del campo $Y(\underline{t})$ como

$$P_Y = \frac{1}{N} \text{tr} \{ \mathcal{H} \underline{\mathbf{R}}_{X(N)} \mathcal{H}^T \}. \quad (\text{B.16})$$

En el caso de que la entrada sea ruido blanco estricto de variancia unitaria, la ecuación B.15 se reduce a

$$\underline{\mathbf{R}}_{Y(N)} = \mathcal{H} \mathcal{H}^T, \quad (\text{B.17})$$

cuyo determinante $\Delta_{Y(N)}$ es

$$\Delta_{Y(N)} = \Delta_{\mathcal{H}} \Delta_{\mathcal{H}} = (h_{0,0})^{2N}, \quad (\text{B.18})$$

donde, de B.13, se tiene que el determinante $\Delta_{\mathcal{H}}$ de la matriz \mathcal{H} es $(h_{0,0})^N$, con N igual al número de observaciones en la región de soporte involucrada. La estructura de $\underline{\mathbf{R}}_{Y(N)}$ es

$$\underline{\mathbf{R}}_{Y(N)} = \begin{pmatrix} R_{(0,0),(0,0)} & R_{(0,1),(0,0)} & R_{(1,\bar{1}),(0,0)} & R_{(2,0),(0,0)} & R_{(1,1),(0,0)} & R_{(2,\bar{1}),(0,0)} & R_{(2,0),(0,0)} & R_{(2,1),(0,0)} \\ R_{(0,0),(0,1)} & R_{(0,1),(0,1)} & R_{(1,\bar{1}),(0,1)} & R_{(1,0),(0,1)} & R_{(1,1),(0,1)} & R_{(2,\bar{1}),(0,1)} & R_{(2,0),(0,1)} & R_{(2,1),(0,1)} \\ R_{(0,0),(1,\bar{1})} & R_{(0,1),(1,\bar{1})} & R_{(1,\bar{1}),(1,\bar{1})} & R_{(1,0),(1,\bar{1})} & R_{(1,1),(1,\bar{1})} & R_{(2,\bar{1}),(1,\bar{1})} & R_{(2,0),(1,\bar{1})} & R_{(2,1),(1,\bar{1})} \\ R_{(0,0),(1,0)} & R_{(0,1),(1,0)} & R_{(1,\bar{1}),(1,0)} & R_{(1,0),(1,0)} & R_{(1,1),(1,0)} & R_{(2,\bar{1}),(1,0)} & R_{(2,0),(1,0)} & R_{(2,1),(1,0)} \\ R_{(0,0),(1,1)} & R_{(0,1),(1,1)} & R_{(1,\bar{1}),(1,1)} & R_{(1,0),(1,1)} & R_{(1,1),(1,1)} & R_{(2,\bar{1}),(1,1)} & R_{(2,0),(1,1)} & R_{(2,1),(1,1)} \\ R_{(0,0),(2,\bar{1})} & R_{(0,1),(2,\bar{1})} & R_{(1,\bar{1}),(2,\bar{1})} & R_{(1,0),(2,\bar{1})} & R_{(1,1),(2,\bar{1})} & R_{(2,\bar{1}),(2,\bar{1})} & R_{(2,0),(2,\bar{1})} & R_{(2,1),(2,\bar{1})} \\ R_{(0,0),(2,0)} & R_{(0,1),(2,0)} & R_{(1,\bar{1}),(2,0)} & R_{(1,0),(2,0)} & R_{(1,1),(2,0)} & R_{(2,\bar{1}),(2,0)} & R_{(2,0),(2,0)} & R_{(2,1),(2,0)} \\ R_{(0,0),(2,1)} & R_{(0,1),(2,1)} & R_{(1,\bar{1}),(2,1)} & R_{(1,0),(2,1)} & R_{(1,1),(2,1)} & R_{(2,\bar{1}),(2,1)} & R_{(2,0),(2,1)} & R_{(2,1),(2,1)} \end{pmatrix} \quad (\text{B.19})$$

Nótese que los elementos de esta matriz son dependientes de los tiempos de observación. Los elementos diferentes (el número de elementos en la matriz es de $N \times N$, pero no todos son diferentes entre sí) cuyos argumentos están más alejados del origen son:

$$\begin{aligned} R_{(2,1),(2,1)} &= h_{0,0}^2 + h_{0,1}^2 + h_{1,0}^2 + h_{1,1}^2 + h_{2,0}^2 + h_{2,1}^2 \\ R_{(2,0),(2,1)} &= h_{2,1}h_{2,0} + h_{2,0}h_{2,\bar{1}} + h_{1,1}h_{1,0} + h_{1,0}h_{1,\bar{1}} + h_{0,1}h_{0,0} = R_{(2,1),(2,0)} \\ R_{(2,\bar{1}),(2,1)} &= h_{2,1}h_{2,\bar{1}} + h_{1,1}h_{1,\bar{1}} = R_{(2,1),(2,\bar{1})} \\ R_{(1,1),(2,1)} &= h_{2,1}h_{1,1} + h_{2,0}h_{1,0} + h_{1,1}h_{0,1} + h_{1,0}h_{0,0} = R_{(2,1),(1,1)} \\ R_{(1,0),(2,1)} &= h_{2,1}h_{1,0} + h_{2,0}h_{1,\bar{1}} + h_{1,1}h_{0,0} = R_{(2,1),(1,0)} \\ R_{(1,\bar{1}),(2,1)} &= h_{2,1}h_{1,\bar{1}} = R_{(2,1),(1,\bar{1})} \\ R_{(1,1),(2,\bar{1})} &= h_{2,\bar{1}}h_{1,1} + h_{1,\bar{1}}h_{0,1} = R_{(2,\bar{1}),(1,1)} \\ R_{(1,1),(2,0)} &= h_{2,0}h_{1,1} + h_{2,\bar{1}}h_{1,0} + h_{1,0}h_{0,1} + h_{1,\bar{1}}h_{0,0} = R_{(2,0),(1,1)} \\ R_{(0,1),(2,1)} &= h_{2,1}h_{0,1} + h_{2,0}h_{0,0} = R_{(2,1),(0,1)} \\ R_{(0,0),(2,1)} &= h_{2,1}h_{0,0} = R_{(2,1),(0,0)} \\ R_{(0,1),(2,\bar{1})} &= h_{2,\bar{1}}h_{0,1} = R_{(2,\bar{1}),(0,1)} \\ R_{(0,1),(2,0)} &= h_{0,1}h_{2,0} + h_{2,\bar{1}}h_{0,0} = R_{(2,0),(0,1)} \end{aligned} \quad (\text{B.20})$$

Comparando B.20 con B.8, se aprecia cómo las observaciones en puntos alejados del origen del campo de salida tienden a ser iguales a las de un campo homogéneo y regular²; esto es,

$$R_{Y(N)}(\underline{j}, \underline{k}) \rightarrow R_{Y(N)}(\underline{m}) \quad \forall \underline{j}, \underline{k} \text{ tales que } |\underline{j} - \underline{k}| = \underline{m} \text{ si } \underline{j} \rightarrow (X, Y) \text{ o si } \underline{k} \rightarrow (X, Y), \quad (\text{B.21})$$

²Por el tamaño y forma no simétrica de la región de soporte escogida para este ejemplo, aparecen algunos ceros por debajo de la diagonal principal de B.13 que producen algunas peculiaridades, por ejemplo $R_{(2,0),(2,0)}$ tiene mayor parecido a $R_{(0,0)}$ que $R_{(2,1),(2,1)}$.

donde (X, Y) son los valores de las coordenadas del punto de la región más alejado del origen. Por otra parte, puesto que el sistema $H(e^{j\omega_1}, e^{j\omega_2})$ es estable, se cumple que, siendo N el orden de la matriz,

$$\|\underline{\mathbf{R}}_{Y(N)}\|_R, \|\underline{\mathbf{R}}_{(N)}\|_R < \infty,$$

y de B.21, se tiene que

$$\lim_{N \rightarrow \infty} \|\underline{\mathbf{R}}_{Y(N)} - \underline{\mathbf{R}}_{(N)}\|_F = 0,$$

por lo que ambas son *asintóticamente equivalentes*:

$$\begin{aligned} \underline{\mathbf{R}}_{Y(N)} &\sim \underline{\mathbf{R}}_{(N)}, \\ \underline{\mathbf{R}}_{(N)} &\sim \mathcal{H} \mathcal{H}^T, \end{aligned} \quad (\text{B.22})$$

y se tiene que

$$\underline{\mathbf{R}}_{(N)}^{-1} \sim (\mathcal{H} \mathcal{H}^T)^{-1} \sim \mathcal{H}^{-T} \mathcal{H}^{-1}; \quad (\text{B.23})$$

resultado que se empleará más adelante. Al campo de salida de un SLIC de fase mínima observado en una región causal cuya entrada sea un campo de ruido blanco de variancia unitaria y media cero aplicado en la misma región, se define como *Campo Asintóticamente Homogéneo y Asintóticamente Regular (CAHAR)*. Con estos resultados, las observaciones hechas en el Apéndice A para procesos asintóticamente regulares y asintóticamente estacionarios se siguen cumpliendo para campos asintóticamente regulares y asintóticamente estacionarios. En el Anexo B.1 se extiende a dos dimensiones tanto el concepto como la solución del problema 34 de [17, pag. 299]. El resultado es análogo al que se obtiene en una dimensión, y se emplea más adelante.

Anexo B.1 Sea $\hat{\mathcal{R}}$ la región de soporte finita de N elementos de un sistema bidimensional invariante al corrimiento $g(m, n)$. Sea X un campo bidimensional con la misma región de soporte y que se aplica como entrada al sistema $g(m, n)$, y Y el campo de salida del sistema $g(m, n)$ con esta entrada y definido en la misma región. Sea $\underline{\mathbf{X}}_{NP}^T = \{X_{m_1, n_1}, X_{m_2, n_2}, \dots, X_{m_N, n_N}\}$ el vector de orden N correspondiente a la sustracción 1 de los N valores del campo X en los puntos de la región $\hat{\mathcal{R}}$. Se define la **entropía diferencial** $H(\underline{\mathbf{X}}_{NP})$ del vector $\underline{\mathbf{X}}_{NP}$ como

$$H(\underline{\mathbf{X}}_{NP}) \triangleq E \left\{ \ln \left[\frac{1}{p_{\underline{\mathbf{X}}_{NP}}(\underline{\mathbf{x}}_{NP})} \right] \right\}, \quad (\text{B.24})$$

donde $p_{\underline{\mathbf{X}}_{NP}}(\underline{\mathbf{x}}_{NP})$ es la función de densidad de probabilidad conjunta de orden N de $\underline{\mathbf{X}}_{NP}$. Se trata de obtener la entropía diferencial de

$$\underline{\mathbf{Y}}_{NP} = \underline{\mathbf{A}} \underline{\mathbf{X}}_{NP}, \quad (\text{B.25})$$

donde $\underline{\mathbf{A}}$ es la matriz asociada al sistema $g(m, n)$ (esta ecuación es equivalente a B.14).

La transformación lineal $\underline{\mathbf{Y}} = \underline{\mathbf{A}} \underline{\mathbf{X}} + \underline{\mathbf{b}}$, para la cual $\underline{\mathbf{A}}$ es una matriz cuadrada determinística y $\underline{\mathbf{b}}$ es un vector determinístico, tiene una f.d.p. $p_{\underline{\mathbf{Y}}}(\underline{\mathbf{y}})$ dada por [17, pag. 21]

$$p_{\underline{\mathbf{Y}}}(\underline{\mathbf{y}}) = \frac{p_{\underline{\mathbf{X}}}(\underline{\mathbf{A}}^{-1}[\underline{\mathbf{Y}} - \underline{\mathbf{b}}])}{|\Delta_{AN}|}, \quad (\text{B.26})$$

donde $\underline{\mathbf{A}}^{-1}$ es la inversa de $\underline{\mathbf{A}}$ y $|\Delta_{AN}|$ es el valor absoluto de su determinante. Puesto que $\underline{\mathbf{A}}^{-1}$ es única, se tiene que [14, pag. 565]

$$p_{\underline{\mathbf{Y}}}(\underline{\mathbf{y}}) = \frac{p_{\underline{\mathbf{X}}}(\underline{\mathbf{x}})}{|\Delta_{AN}|}, \quad (\text{B.27})$$

con lo cual

$$H(\underline{\mathbf{Y}}_{NP}) = E \left\{ \ln \left[\frac{1}{p_{\underline{\mathbf{Y}}_{NP}}(\underline{\mathbf{y}}_{NP})} \right] \right\} = E \left\{ \ln \left[\frac{1}{p_{\underline{\mathbf{X}}_{NP}}(\underline{\mathbf{x}}_{NP})} \right] \right\} + \ln |\Delta_{AN}|, \quad (\text{B.28})$$

es decir

$$H(\underline{\mathbf{Y}}_{NP}) = H(\underline{\mathbf{X}}_{NP}) + \ln |\Delta_{AN}|. \quad (\text{B.29})$$

La matriz $\underline{\mathbf{A}}$ puede tomar diversas formas según se tome el ordenamiento de la región de soporte $\hat{\mathcal{R}}$. Cada ordenamiento distinto produce distintos ordenamientos en $\underline{\mathbf{X}}_{NP}$, $\underline{\mathbf{A}}$ y $\underline{\mathbf{Y}}_{NP}$. De hecho, esta ecuación se cumple para cualesquiera regiones de soporte de X y $g(m, n)$, causales o no causales, siempre que la matriz asociada a la transformación sea cuadrada y con inversa única. Para el caso particular de que tanto X como $g(m, n)$ tengan una región de soporte causal, y que la salida se observe también una región causal, uno de esos ordenamientos produce una matriz diagonal inferior, como en B.13, cuyo determinante es simplemente $(g(0, 0))^N$, con N igual al número de elementos en la región. Empleando las matrices de permutación y sus propiedades, los determinantes de todas las posibles matrices $\underline{\mathbf{A}}$ son iguales, con lo cual

$$H(\underline{\mathbf{Y}}_{NP}) = H(\underline{\mathbf{X}}_{NP}) + N \ln |g(0, 0)|. \quad (\text{B.30})$$

Puesto que la entropía diferencial normalizada del campo Y puede escribirse como [14, pag. 587]

$$\bar{H}_Y = \lim_{N \rightarrow \infty} \frac{1}{N} H(\underline{\mathbf{Y}}_{NP}) = \lim_{N \rightarrow \infty} \frac{1}{N} H(Y_{m_1, n_1}, Y_{m_2, n_2}, \dots, Y_{m_N, n_N}), \quad (\text{B.31})$$

sustituyendo en B.30, se obtiene finalmente que

$$\bar{H}_Y = \bar{H}_X + \frac{1}{2} \ln |g(0, 0)|^2, \quad (\text{B.32})$$

donde \bar{H}_X es la entropía diferencial normalizada del campo X y el logaritmo se ha expresado en una forma que conviene en secciones posteriores.

B.3 Resultados de Predicción Lineal en Dos Dimensiones.

Sea un campo aleatorio bidimensional homogéneo $X(\underline{\mathbf{t}})$ con función de densidad espectral $f_X(e^{j\omega_1}, e^{j\omega_2})$; la variancia o potencia del error P de predecir $X(\underline{\mathbf{0}})$ en términos de todo su

“pasado” (todos los puntos fuera de la región causal \mathcal{R} , o complemento de \mathcal{R}), está dado,[24] y [28, pag. 224], por³

$$P = \exp \left\{ \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln f_X(e^{j\omega_1}, e^{j\omega_2}) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \right\}. \quad (\text{B.33})$$

Una región de soporte causal finita $\hat{\mathcal{R}}$ de N' elementos que excluye al origen, como se muestra en la figura B.6, se define como *región de predicción causal* [8, pag. 205] y es posible emplear una versión bidimensional de las ecuaciones de Yule-Walker [8, pag. 207] (algo equivalente a predecir en términos de los N valores pasados más recientes del caso unidimensional):

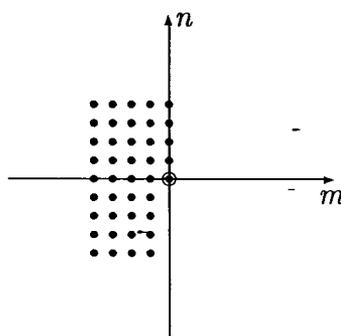


Figura B.6: Ejemplo de Región de Predicción Causal con $N' = 40$ elementos.

$$R(k, l) - \sum_{\forall i, j \in \hat{\mathcal{R}}} a(i, j) R(k - i, l - j) = 0, \quad (\text{B.34})$$

expresión que puede escribirse como un sistema de ecuaciones de $N' \times N'$ cuya solución proporciona el valor de los N' coeficientes $a(i, j)$ del estimado $\hat{Y}(t)$ de $Y(t)$. Por otro punto, la potencia del error de predicción de orden N' , $P_{N'}$ es [8, pag. 207]:

$$P_{N'} = R(0, 0) - \sum_{\forall i, j \in \hat{\mathcal{R}}} a(i, j) R(i, j). \quad (\text{B.35})$$

Las expresiones B.34 y B.35 forman un sistema aumentado de ecuaciones de orden $N' + 1 = N$ cuya solución proporciona los valores de los N' coeficientes $a(i, j)$ y el valor de $P_{N'}$. En particular, se tiene que [14, pag. 458]

$$P_{N'} = \frac{\Delta_{N'+1}}{\Delta_{N'}}, \quad \Delta_{N'} > 0; \quad (\text{B.36})$$

donde $\Delta_{N'}$ y $\Delta_{N'+1}$ son los determinantes de las matrices de los sistemas de orden N' y de orden $N' + 1$, respectivamente. Al aumentar el número de puntos o tamaño de la región de predicción $\hat{\mathcal{R}}$, de modo que tienda a ser la misma que el complemento de la región causal \mathcal{R} , el problema

³En una dimensión, Papoulis [15, pag. 424] le llama *Fórmula de Kolmogoroff*

tiende a ser igual al de predicción empleando todo el pasado del campo, es decir, por B.33 y B.36 se cumple que

$$P = \lim_{N' \rightarrow \infty} P_{N'} = \lim_{N' \rightarrow \infty} \frac{\Delta_{N'+1}}{\Delta_{N'}} = \exp \left\{ \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln f_X(e^{j\omega_1}, e^{j\omega_2}) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \right\}. \quad (\text{B.37})$$

Jain [29, pag. 291] presenta este resultado usando algoritmos que emplean recursividad en ambas dimensiones. Al tomar el logaritmo del error de predicción, se tiene que ⁴

$$\ln P = \lim_{N' \rightarrow \infty} \ln \frac{\Delta_{N'+1}}{\Delta_{N'}} = \lim_{N \rightarrow \infty} \frac{1}{N} \ln \Delta_N = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln f_X(e^{j\omega_1}, e^{j\omega_2}) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}; \quad (\text{B.38})$$

lo que es un resultado similar a [14, pag. 531]. La cualidad de B.38 es que relaciona determinantes con espectros en dos dimensiones sin emplear los teoremas de *Formas Asíntoticamente Toeplitz* de Grenander y Szegö, aunque para obtener la potencia del error de predicción P se modificaron estas ideas de Grenander y Szegö.

Como ejemplo, si se considera una región de predicción como la de la Fig.B.5 (sin incluir el origen), el sistema de $N = N' + 1$ queda

$$\begin{pmatrix} 1 & R_{0,\bar{1}} & R_{\bar{1},1} & R_{\bar{1},0} & R_{\bar{1},\bar{1}} & R_{2,\bar{1}} & R_{\bar{2},0} & R_{\bar{2},\bar{1}} \\ 0 & R_{0,0} & R_{\bar{1},2} & R_{\bar{1},1} & R_{\bar{1},0} & R_{\bar{2},2} & R_{\bar{2},1} & R_{\bar{2},0} \\ 0 & R_{1,\bar{2}} & R_{0,0} & R_{0,\bar{1}} & R_{0,\bar{2}} & R_{\bar{1},0} & R_{\bar{1},\bar{1}} & R_{\bar{1},\bar{2}} \\ 0 & R_{1,\bar{1}} & R_{0,1} & R_{0,0} & R_{0,\bar{1}} & R_{\bar{1},1} & R_{\bar{1},0} & R_{\bar{1},\bar{1}} \\ 0 & R_{1,0} & R_{0,2} & R_{0,1} & R_{0,0} & R_{\bar{1},2} & R_{\bar{1},1} & R_{\bar{1},0} \\ 0 & R_{2,\bar{2}} & R_{1,0} & R_{1,\bar{1}} & R_{1,\bar{2}} & R_{0,0} & R_{0,\bar{1}} & R_{0,\bar{2}} \\ 0 & R_{2,\bar{1}} & R_{1,1} & R_{1,0} & R_{1,\bar{1}} & R_{0,1} & R_{0,0} & R_{0,\bar{1}} \\ 0 & R_{2,0} & R_{1,2} & R_{1,1} & R_{1,0} & R_{0,2} & R_{0,1} & R_{0,0} \end{pmatrix} \begin{pmatrix} P_{N'} \\ a_{0,1} \\ a_{1,\bar{1}} \\ a_{1,0} \\ a_{1,1} \\ a_{2,\bar{1}} \\ a_{2,0} \\ a_{2,1} \end{pmatrix} = \begin{pmatrix} R_{0,0} \\ R_{0,1} \\ R_{1,\bar{1}} \\ R_{1,0} \\ R_{1,1} \\ R_{2,\bar{1}} \\ R_{2,0} \\ R_{2,1} \end{pmatrix}, \quad (\text{B.39})$$

y la matriz con determinante $\Delta_{N'+1}$ es

$$\mathbf{R} = \begin{pmatrix} R_{0,0} & R_{0,\bar{1}} & R_{\bar{1},1} & R_{\bar{1},0} & R_{\bar{1},\bar{1}} & R_{2,\bar{1}} & R_{\bar{2},0} & R_{\bar{2},\bar{1}} \\ R_{0,1} & R_{0,0} & R_{\bar{1},2} & R_{\bar{1},1} & R_{\bar{1},0} & R_{\bar{2},2} & R_{\bar{2},1} & R_{\bar{2},0} \\ R_{1,\bar{1}} & R_{1,\bar{2}} & R_{0,0} & R_{0,\bar{1}} & R_{0,\bar{2}} & R_{\bar{1},0} & R_{\bar{1},\bar{1}} & R_{\bar{1},\bar{2}} \\ R_{1,0} & R_{1,\bar{1}} & R_{0,1} & R_{0,0} & R_{0,\bar{1}} & R_{\bar{1},1} & R_{\bar{1},0} & R_{\bar{1},\bar{1}} \\ R_{1,1} & R_{1,0} & R_{0,2} & R_{0,1} & R_{0,0} & R_{\bar{1},2} & R_{\bar{1},1} & R_{\bar{1},0} \\ R_{2,\bar{1}} & R_{2,\bar{2}} & R_{1,0} & R_{1,\bar{1}} & R_{1,\bar{2}} & R_{0,0} & R_{0,\bar{1}} & R_{0,\bar{2}} \\ R_{2,0} & R_{2,\bar{1}} & R_{1,1} & R_{1,0} & R_{1,\bar{1}} & R_{0,1} & R_{0,0} & R_{0,\bar{1}} \\ R_{2,1} & R_{2,0} & R_{1,2} & R_{1,1} & R_{1,0} & R_{0,2} & R_{0,1} & R_{0,0} \end{pmatrix}, \quad (\text{B.40})$$

que es igual a la matriz de correlaciones B.7 asociada a la región causal de la Fig.B.5. Estos resultados se emplean en la siguiente sección.

⁴Ver sección A.4 del Apéndice A.

B.4 Reducción de $\bar{D}(\hat{S}, S)$.

Considerando que S es un campo regular, gaussiano y homogéneo, de la definición introducida en el Capítulo 4, se tiene que

$$\begin{aligned} \bar{D}(\hat{S}, S) &= -\bar{H}_{\hat{S}} + \frac{1}{2} \ln(2\pi) + \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \ln(\Delta_S) \\ &\quad + \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \mathbf{R}_{\hat{S}(N)} \mathbf{R}_{S(N)}^{-1} \right\}; \end{aligned} \quad (\text{B.41})$$

donde Δ_S es el determinante de la matriz de correlación de orden N del campo S y $\mathbf{R}_{\hat{S}(N)}$ y $\mathbf{R}_{S(N)}$ son las matrices de correlaciones de los campos \hat{S} y S respectivamente. Se consideran nuevamente dos casos: cuando \hat{S} es un campo asintóticamente regular y asintóticamente homogéneo y cuando es regular y homogéneo.

a) \hat{S} asintóticamente regular y asintóticamente homogéneo. Primeramente, se denota al campo \hat{S} como \hat{S}^a y a su matriz de correlaciones de orden $N \times N$ con $\mathbf{R}_{\hat{S}^a(N)}$ para diferenciarla de la matriz de correlaciones $\mathbf{R}_{\hat{S}(N)}$ del mismo campo cuando es regular y estacionario. Del Anexo B.1 se tiene

$$\bar{H}_{\hat{S}^a} = \bar{H}_X + \frac{1}{2} \ln |g(0, 0)|^2.$$

Por otra parte, de B.38 se tiene

$$\lim_{N \rightarrow \infty} \frac{1}{N} \ln \Delta_S = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln f_X(e^{j\omega_1}, e^{j\omega_2}) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}.$$

Puesto que $\mathbf{R}_{\hat{S}^a(N)} \sim \mathbf{R}_{\hat{S}(N)}$, por la propiedad c) de matrices asintóticamente equivalentes, se puede escribir

$$\frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \mathbf{R}_{\hat{S}^a(N)} \mathbf{R}_{S(N)}^{-1} \right\} = \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \mathbf{R}_{\hat{S}(N)} \mathbf{R}_{S(N)}^{-1} \right\}; \quad (\text{B.42})$$

y como se considera que S es regular y homogéneo, de B.23

$$\mathbf{R}_{S(N)}^{-1} \sim \mathcal{H}^{-T} \mathcal{H}^{-1};$$

donde \mathcal{H} es la matriz asociada a la respuesta al impulso de un sistema de fase mínima. Con lo anterior, se tiene

$$\begin{aligned} \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \mathbf{R}_{\hat{S}^a(N)} \mathbf{R}_{S(N)}^{-1} \right\} &= \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \mathbf{R}_{\hat{S}(N)} \mathbf{R}_{S(N)}^{-1} \right\} \\ &= \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \mathbf{R}_{\hat{S}(N)} (\mathcal{H}^{-T} \mathcal{H}^{-1}) \right\} \\ &= \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \mathcal{H}^{-1} \mathbf{R}_{\hat{S}(N)} \mathcal{H}^{-T} \right\}; \end{aligned} \quad (\text{B.43})$$

expresión que, comparando con B.16, proporciona la potencia de la señal de salida de un sistema que es el inverso de uno de fase mínima $H(e^{j\omega_1}, e^{j\omega_2})$, con entrada el campo regular \hat{S} , como se indica en la Fig.B.7. Se tiene entonces que

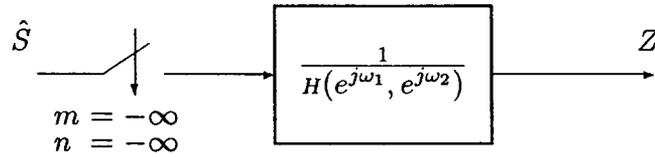


Figura B.7: Sistema Inverso

$$P_Z = \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \mathcal{H}^{-1} \mathbf{R}_{\hat{S}} \mathcal{H}^{-T} \right\} = \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{f_{\hat{S}}(e^{j\omega_1}, e^{j\omega_2})}{|H(e^{j\omega_1}, e^{j\omega_2})|^2} \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi};$$

es decir

$$\frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \text{tr} \left\{ \mathcal{H}^{-1} \mathbf{R}_{\hat{S}} \mathcal{H}^{-T} \right\} = \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{f_{\hat{S}}(e^{j\omega_1}, e^{j\omega_2})}{f_S(e^{j\omega_1}, e^{j\omega_2})} \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}. \quad (\text{B.44})$$

Sustituyendo B.38 y B.44 en B.41, y sumando y restando 1/2 y

$$\frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln f_X(e^{j\omega_1}, e^{j\omega_2}) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi},$$

se tiene

$$\begin{aligned} \overline{D}(\hat{S}, S) &= \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left(\frac{f_{\hat{S}}(e^{j\omega_1}, e^{j\omega_2})}{f_S(e^{j\omega_1}, e^{j\omega_2})} - \ln \frac{|g(0,0)|^2 f_X(e^{j\omega_1}, e^{j\omega_2})}{f_S(e^{j\omega_1}, e^{j\omega_2})} - 1 \right) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} + \\ &\quad - \overline{H}_X + \frac{1}{2} \ln(2\pi e) + \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln f_X(e^{j\omega_1}, e^{j\omega_2}) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}. \end{aligned} \quad (\text{B.45})$$

b) \hat{S} Regular y Homogéneo. Se demostrará primeramente la expresión 3.26. Supóngase, inicialmente, que X es un campo normal, con función de densidad de probabilidad conjunta de orden N dada por 3.7. Aplicando la definición 3.15 se tiene

$$\overline{H}_X = \lim_{N \rightarrow \infty} \frac{1}{N} E_X \left\{ \ln \left(\frac{1}{\frac{1}{\sqrt{(2\pi)^N \Delta_{X(N)}}} e^{\left\{ -\frac{1}{2} \mathbf{x}^T \mathbf{R}_{X(N)}^{-1} \mathbf{x} \right\}}} \right) \right\}, \quad (\text{B.46})$$

donde $E_X \{ \}$ indica el valor esperado con respecto a X , \mathbf{x} es el vector de N variables aleatorias del campo X , y $\Delta_{X(N)}$ es el determinante de la matriz de correlaciones $\mathbf{R}_{X(N)}$ del campo X . Entonces

$$\overline{H}_X = \frac{1}{2} \ln(2\pi e) + \frac{1}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \Delta_{X(N)}, \quad (\text{B.47})$$

ya que

$$E_X \left\{ -\frac{1}{2} \mathbf{x}^T \mathbf{R}_{X(N)}^{-1} \mathbf{x} \right\} = E_X \left\{ \frac{1}{2} \text{tr} (\mathbf{R}_{X(N)}^{-1} \mathbf{R}_{X(N)}) \right\} = \frac{1}{2} N.$$

Por B.38 se obtiene

$$\bar{H}_X = \frac{1}{2} \ln(2\pi e) + \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln f_X(e^{j\omega_1}, e^{j\omega_2}) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \quad (\text{B.48})$$

que es 3.21. Si ahora este proceso normal X se aplica a un SLIC con función de transferencia $G(e^{j\omega_1}, e^{j\omega_2})$, entonces el campo Y de salida será también normal con entropía diferencial normalizada

$$\bar{H}_Y = \frac{1}{2} \ln(2\pi e) + \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln \left\{ f_X(e^{j\omega_1}, e^{j\omega_2}) |G(e^{j\omega_1}, e^{j\omega_2})|^2 \right\} \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}, \quad (\text{B.49})$$

con lo que se obtiene

$$\bar{H}_Y = \bar{H}_X + \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln |G(e^{j\omega_1}, e^{j\omega_2})|^2 \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}. \quad (\text{B.50})$$

Suponga ahora que X no es normal. Del Anexo B.1 se obtiene el resultado general B.29.

$$H(\underline{Y}_{NP}) = H(\underline{X}_{NP}) + K;$$

donde $K = \ln |\Delta_{AN}|$ es una constante que sólo depende de los parámetros de la transformación lineal $G(e^{j\omega_1}, e^{j\omega_2})$ con respuesta al impulso $g(m, n)$ e inversa única, i.e, es independiente de X . Si ahora se considera una transformación con región de soporte que tienda a ser de tamaño infinito, se obtiene

$$\bar{H}_Y = \bar{H}_X + \lim_{N \rightarrow \infty} \ln |\Delta_{AN}|; \quad (\text{B.51})$$

el último término del lado derecho sigue siendo independiente de X , y además es una constante finita ya que se supone que $G(e^{j\omega_1}, e^{j\omega_2})$ es estable. Como se ha visto, si X es normal este término constante es igual a la integral en B.50, y puesto que es independiente de X , debe ser igual a esta integral para cualquier campo X [14, pag. 569], es decir, B.50 se cumple para cualquier campo X .

Con el resultado anterior, se tiene que

$$\bar{H}_{\hat{S}} = \bar{H}_X + \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln |G(e^{j\omega_1}, e^{j\omega_2})|^2 \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi},$$

y con un procedimiento similar al anterior, la ecuación B.41 se reduce a

$$\begin{aligned} \bar{D}(\hat{S}, S) &= \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left(\frac{f_{\hat{S}}(e^{j\omega_1}, e^{j\omega_2})}{f_S(e^{j\omega_1}, e^{j\omega_2})} - \ln \frac{|G(e^{j\omega_1}, e^{j\omega_2})|^2 f_X(e^{j\omega_1}, e^{j\omega_2})}{f_S(e^{j\omega_1}, e^{j\omega_2})} - 1 \right) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \\ &\quad - \bar{H}_X + \frac{1}{2} \ln(2\pi e) + \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln f_X(e^{j\omega_1}, e^{j\omega_2}) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \\ \bar{D}(\hat{S}, S) &= \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \left(\frac{f_{\hat{S}}(e^{j\omega_1}, e^{j\omega_2})}{f_S(e^{j\omega_1}, e^{j\omega_2})} - \ln \frac{f_{\hat{S}}(e^{j\omega_1}, e^{j\omega_2})}{f_S(e^{j\omega_1}, e^{j\omega_2})} - 1 \right) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \\ &\quad - \bar{H}_X + \frac{1}{2} \ln(2\pi e) + \frac{1}{2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln f_X(e^{j\omega_1}, e^{j\omega_2}) \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi}, \quad (\text{B.52}) \end{aligned}$$

es decir, sólo cambia el término que involucra al logaritmo en la integral. El primer término del lado derecho de B.52 se conoce como *Distorsión de Itakura-Saito Bidimensional* [9]. Nótese que los últimos tres términos de B.45 y B.52 no se cancelan ya que X no se considera gaussiano.

Referencias

- [1] Fernando Lepe Casillas. *Sobre Aplicaciones de la Teoría de Información al Procesamiento Digital de Señales*. Tesis de Maestría. DEPFI-UNAM, 1987. Dirigida por Luis Andrés Buzo de la Peña.
- [2] F. Lepe Casillas, Andrés Buzo. *Process Modeling and Filtering with Discrimination Information as the Fidelity Measure*. DEPFI-UNAM. Reporte Interno.
- [3] F. Lepe Casillas, A. Buzo. *A criterion founded on Information Theory for designing linear estimation filters*. Proceedings of EUSIPCO-90, september 18-21 1990, Barcelona, España. Vol. 1, pag. 513-516.
- [4] M. Pinsker. *Information and Information Stability of Random Variables and Processes*. Ize. Akad. Nauk, SSSR, Moscow, 1960.
- [5] Fernando Lepe Casillas. Sesiones del Proyecto de Tesis.
- [6] Erik Van Marcke. *Random Fields*. The MIT Press, Cambridge Massachusetts, 1983.
- [7] Fernando Lepe Casillas. *Filtro de Wiener causal con Ponderación Óptima según la Información de Discriminación Normalizada*. Memoria del XVII Congreso de la Academia de Ingeniería. Monterrey, Nuevo León. Septiembre 1991.
- [8] A. K. Jain. *Fundamentals of Digital Image Processing*. Ed. Prentice Hall, 1989.
- [9] H. Lev-Ari, S. R. Parker, T. Kailath. *Multidimensional Maximum Entropy Covariance Extension*. IEEE. Trans. Inf. Theory., Vol. 35, No. 3, may 1989.
- [10] B. R. Hunt. Capit. *Image Restoration. Digital Image Processing Techniques*. Editor M. P. Ekstrom. Academic Press, 1984.
- [11] T. M. Cannon. *Digital Image Deblurring by Nonlinear Homomorphic Filtering*. P.h. D. Thesis, Computer Science Department, University of Utah, Salt Lake City. 1974.
- [12] William K. Pratt. *Digital Image Processing*. 2nd. edition, John Wiley Interscience, 1991.
- [13] H. L. Weidemann, E. B. Stear. *Entropy Analysis of Estimating Systems*. IEEE Trans. Inf. Theory, Vol IT-16, No. 3, pp 264-270, May 1970.

- [14] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. 3rd Ed., McGraw-Hill, 1991.
- [15] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. 2nd. Ed., McGraw-Hill, 1984.
- [16] William H. Press *et al.* *Numerical Recipes in C*. Cambridge University Press, 1988.
- [17] W. A. Gardner. *INTRODUCTION TO RANDOM PROCESSES, With applications to Signals and Systems*. Macmillan Publishing Company, 1986.
- [18] Philip E. Gill, Walter Murray, Margaret H. Wright. *Numerical Linear Algebra and Optimization Volume 1*. Addison-Wesley Publishing Co., 1991.
- [19] Gene H. Golub, Charles F. Van Loan. *Matrix Computations*. The John Hopkins University Press, 1983.
- [20] R. A. Gabel, R. A. Roberts. *Señales y Sistemas Lineales*. Ed. LIMUSA, 1975.
- [21] C. Gueguen. *Course 12: An Introduction to Displacement Ranks and Related Fast Algorithms. Signal Processing Vol. II*. North-Holland, 1987.
- [22] R. M. Gray. *Toeplitz and Circulant Matrices II*; Stanford Electron. Lab., Stanford Univ., CA. Tech. Rep. 6504-1, April 1977.
- [23] A. V. Oppenheim, R. W. Schaffer. *Discrete-Time Signal Processing*. Prentice Hall, 1989.
- [24] H. Helson, D. Lowdenslager. *Prediction Theory and Fourier Series in Several Variables*, Acta Mathematica Vol 99, pp. 165-202, 1958.
- [25] M. P. Ekstrom, J. W. Woods. *Two Dimensional Spectral Factorization with Applications in Recursive Digital Filtering*, IEEE Trans. Acous. Speech and Signal Proc. Vol. ASSP 24, No. 2, pp. 115-128, april 1976.
- [26] S. M. Kay. *Modern Spectral Estimation, Theory and Application*. Prentice-Hall, 1988.
- [27] D. E. Dudgeon, R. M. Mersereau. *Multidimensional Digital Signal Processing*. Prentice-Hall, 1984.
- [28] M. Rosenblatt *Stationary Sequences and Random Fields*. Birkhäuser, 1985.
- [29] S. Ranganath, A. K. Jain. *Two Dimensional Linear Prediction Models, Part 1: Spectral Factorization and Realization*. IEEE Trans. Acous. Speech and Signal Proc. Vol. ASSP 33, No. 1, pp. 280-299, february 1985.
- [30] A. Buzo, F. Lepe, F. Kuhlmann. *Cantidad de Información de Discriminación entre Procesos Estacionarios*. Memoria del XI Congreso de la Academia Nacional de Ingeniería, San Luis Potosí, S.L.P. Septiembre 1985.
- [31] Spiro Gumas *et al.* *The Guru of Graphics Arrays*. McGraw-Hill, 1994.