

MATRÍCULA DE HONOR

19/1x/2007

Apuntes de Tratamiento Digital de la Señal 2
4º ETSI Telecomunicación
Universidad de Málaga

Carlos García Argos (garcia.argos@gmail.com)

Curso 2006/2007

TRATAMIENTO DIGITAL DE LA SEÑAL 2 - ÍNDICE

TEMA 1: ANÁLISIS ESPECTRAL PARAMÉTRICO	1
1.- Señales aleatorias	1
2.- Introducción a la teoría de la estimación	6
3.- Estimación aplicada a las señales	12
4.- Filtros FIR en celosía	17
5.- Predicción lineal	25
5.1.- Algoritmo de Levinson-Durbin	31
5.2.- Propiedad de la señal de error de predicción	39
5.3.- Filtrado lineal óptimo (filtro de Wiener)	41
6.- Modelado de señales	43
6.1.- Ejemplos de modelos	56
7.- Análisis espectral clásico	59
7.1.- Periodograma	60
7.2.- Método de Bartlett	64
7.3.- Método de Welch	66
7.4.- Método de Blackman-Tukey	68
7.5.- Prestaciones	70
7.6.- Estimador de Daniell's	72
8.- Análisis espectral paramétrico	73
8.1.- Método de Yule-Walker	76
8.2.- Método de los coeficientes de reflexión	76
8.3.- Método de la covarianza	78
8.4.- Método de la covarianza modificada	79
8.5.- Elección del orden del modelo	81
8.6.- Aplicación del algoritmo de Levinson-Durbin	82
9.- Caracterización de los estimadores estudiados	84
TEMA 2: LA SEÑAL DE VOZ: MODELADO Y ANÁLISIS	87
1.- Introducción	87
2.- Métodos de análisis localizado (STP)	102
2.1.- Análisis localizado en el tiempo	104
2.2.- Análisis localizado en la frecuencia	114

TEMA 3: CODIFICACIÓN DE VOZ	123
1.- Introducción a la codificación de fuente analógica	123
2.- Cuantificación escalar	127
2.1.- Cuantificación uniforme	130
2.2.- Cuantificación óptima	134
2.3.- Cuantificación no uniforme	135
2.4.- Estándar PCM	140
2.5.- Calidad de la señal de voz	141
2.6.- Cuantificación adaptativa	142
2.7.- Cuantificación adaptativa realimentada (Jayant)	144
2.8.- Cuantificación diferencial	145
2.9.- Cuantificación diferencia adaptativa (ADPCM)	150
2.10.- Modulación delta	151
2.11.- Modulación delta adaptativa	153
3.- Vocoders mixtos o adaptativos	154

TEMA 1: ANÁLISIS ESPECTRAL PARÁMÉTRICO

1.- SEÑALES ALEATORIAS

señales / deterministas
aleatorias } \Rightarrow modelos de la realidad

- señal aleatoria: resultado de un experimento

espacio probabilístico = (Ω, \mathcal{P})

- Ω = conjunto de resultados del experimento
- \mathcal{P} = regla para aplicar una señal aleatoria a los resultados (probabilidades)

si para cada $\omega_i \in \Omega$, $\omega_i \xrightarrow{\mathcal{P}} x(t, \omega_i)$
 \uparrow
señal aleatoria

$x_d(u, \omega)$ = señal aleatoria de tiempo discreto

caracterizada por sus estadísticas

• caracterización \Rightarrow conocer su función de probabilidad en el instante, así como los conjuntos:

$$\cdot f_{X(u)}(x) \quad f_u$$

$$\cdot f_{X(u_1), X(u_2)}(x) \quad f_{u_1, u_2}$$

$$\cdot f_{X(u_1), X(u_2), X(u_3)}(x) \quad f_{u_1, u_2, u_3}$$

⋮

⋮

→ (hasta el infinito y más allá en el caso general)

Toda caracterización completa no es posible obtenerla, salvo para casos muy simples.

Por tanto, solamente contaremos con estadísticos de la señal de primer y segundo orden:

$$\cdot \underline{\text{media}}: \quad m_x(u) = E[x(u)]$$

$$\cdot \underline{\text{autocorrelación}}: \quad Y_x(u_1, u_2) = E[x(u_1)x^*(u_2)]$$

$$\cdot \underline{\text{covarianza}}: \quad Q_x(u_1, u_2) = E[(x(u_1) - m_x(u_1))(x(u_2) - m_x(u_2))^*] \\ = Y_x(u_1, u_2) - m_x(u_1)m_x^*(u_2)$$

$$\cdot \underline{\text{correlación cruzada}}: \quad Y_{xy}(u_1, u_2) = E[x(u_1)y^*(u_2)]$$

Ojo: considerar señales complejas

Señales estacionarias: (see)

- en sentido estricto: todas las estadísticas son insensibles a un desplazamiento del origen de tiempo

- en sentido amplio: la media no depende del instante de tiempo y la autocorrelación sólo depende del ancho del intervalo

$$m_x(u) = m_x$$

$$\gamma_x(u) = E[x(u)x^*(u-u)] \quad u=u_1-u_2$$

Señales estacionarias conjuntamente estacionarias:

$x(u)$, $y(u)$ estacionarias en sentido amplio, y

$$\gamma_{xy}(u_1, u_2) = \gamma_{xy}(u) \quad u=u_1-u_2$$

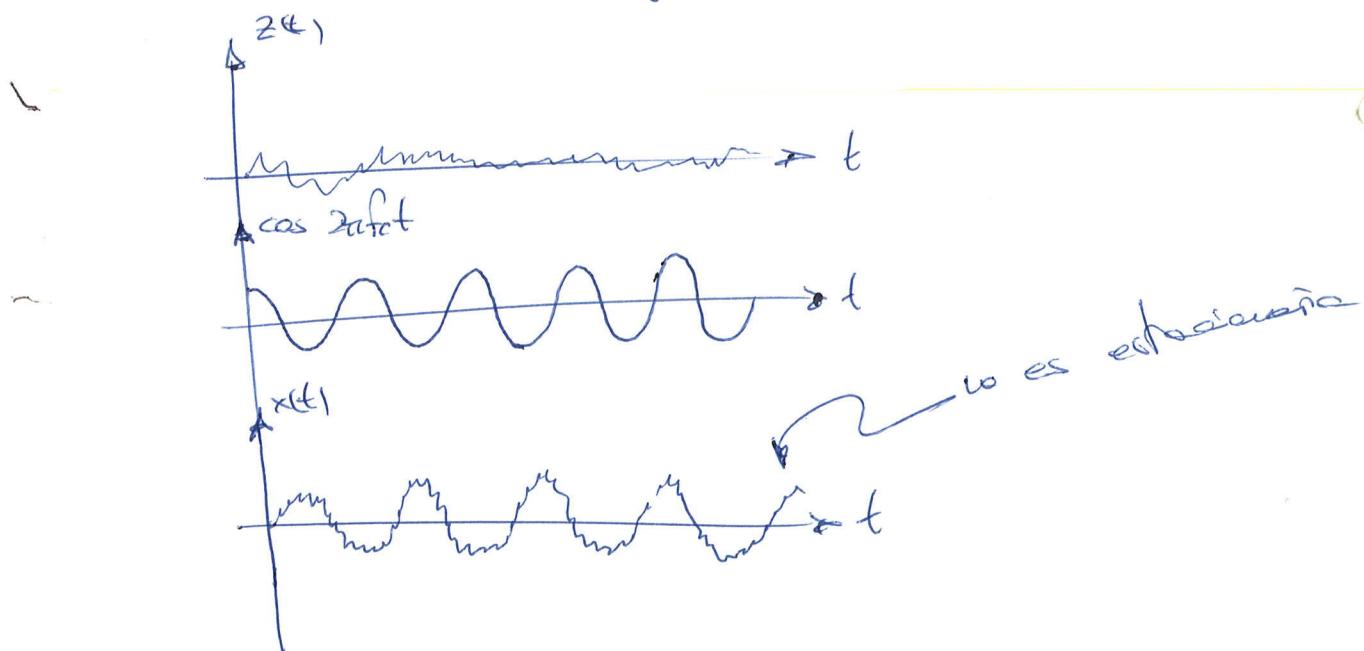
Sistemas ciclotrónicos:

La media y la autocorrelación dependen del origen de tiempo, pero de forma periódica

$$\left. \begin{array}{l} \mu_x(u) = \mu_x(u + kT) \\ Y_x(u) = Y_x(u + kT) \end{array} \right\}$$

Ejemplo: modelizar DBL $\Rightarrow x(t) = z(t) \cos 2\pi f_0 t$

$z(t)$ aleatoria y estacionaria (media nula)



$x(t)$ es ciclotrónico \Rightarrow se aprecia en receptor para decodificar

$$\text{Potencia } (x(t)) = \bar{t} \{ x^2(t) \} = Y_x(t, t)$$

\uparrow

depende del tiempo \Rightarrow cincelamiento

_____ o _____

Ejemplo: $x(t) = z(t) \cos(2\pi f_0 t + \phi)$

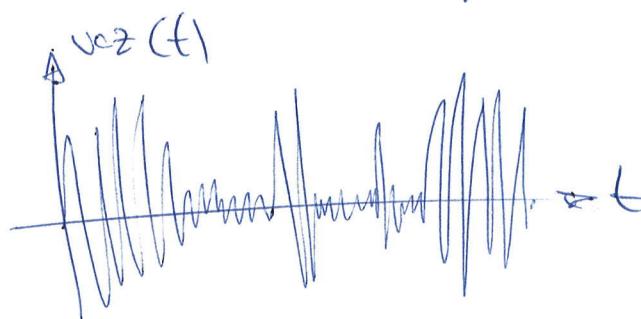
ϕ : v.c. uniforme en $[0, 2\pi]$

$\Rightarrow x(t)$ es estacionario

Método usado para calcular esp., dens. est. etc.

_____ o _____

Ejemplo: Señal de voz es muy compleja, con muchas
- b. freq. del tiempo



voz: constucción de
frecuencias

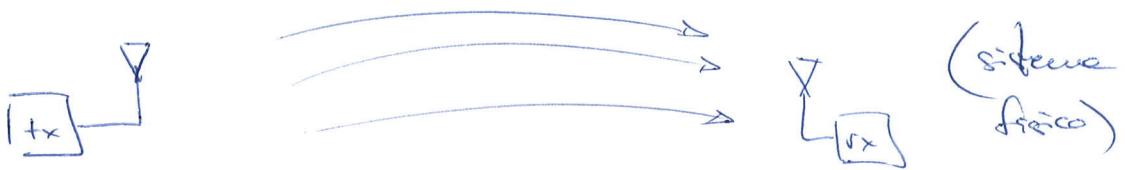
ondas vocales: energía ↑, periodidad:

(fase 2)

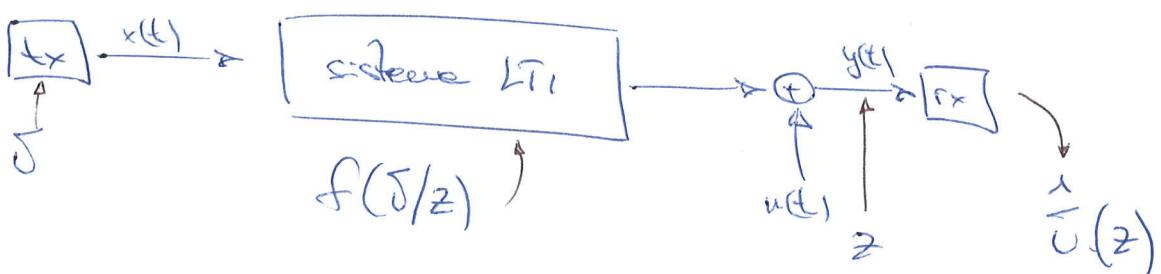
método de extracción: cuálisis en zonas localizadas

_____ o _____

2.- INTRODUCCIÓN A LA TEORÍA DE LA ESTIMACIÓN



modelado matemático



Para modelar bien el sistema de comunicación,
hay que usar variables estocásticas \Rightarrow estadísticas

$$\begin{cases} \bar{E}[x(t)] \\ \bar{E}[x(t) * (t-\tau)] \\ f_{\bar{E}(t)}(x) \end{cases}$$

El objetivo de la lección de estimación es
obtener estos parámetros estadísticos a partir de
medidas.

Se puede conocer:

- una realización en un intervalo: $x(t)$, $t \in [0, T]$
 - varias realizaciones en un intervalo: $x_1(t), x_2(t), \dots$, $t \in [0, T]$
 - una realización en varios intervalos
 - etc.
- señal ergódica: es aquella de la que se pueden extraer las estadísticas a partir de una sola realización.
- A veces se pueden extraer más pruebas y otros.
- vo. Ej.: "señal ergódica respecto a 6 medias".

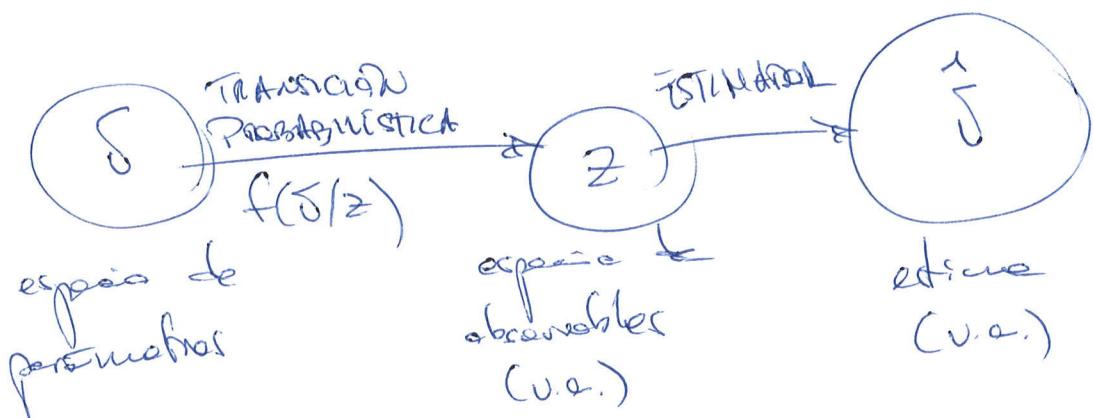
El proceso de estimación tiene 3 partes:

- Modelo de la realidad: de este modelo surge un espacio de parámetros, \mathcal{S}
Por ejemplo, la media de una señal o la respuesta a un impulso de un sistema.
- Modelo observación: va lleva a un espacio de observación

2.

Ejemplo: me conozco la señal aleatoria
en 6 medias del campo que produce

Proceso proporcional de estimación: se usa el parámetro estimado como $\hat{\gamma}$



La regla que permite, a partir de un resultado, obtener un estimador del parámetro deseado es la que se conoce como estimador. Es una regla determinista.

La relación entre el espacio de parámetros y el de observables se llama función probabilística $f(S|Z)$

Los observables medidos son variables aleatorias, por lo que las estimaciones también.

$$\text{Si fuese ideal: } E[\hat{\gamma}] = S$$

$\hat{\gamma}$ es una estimación de S - saber de los datos observados en Z .

Vamos a ver los estadísticos más importantes que caracterizan un estimador:

- Sesgo: mide la desviación del valor medio del pronóstico estimado con respecto al valor que queremos medir

$$B(\hat{\delta}) = \bar{\delta} - E\{\hat{\delta}\} \quad (\bar{\delta} \text{ : bias})$$

Hech: $B(\hat{\delta}) = 0$

→ estimador sesgado $\Rightarrow B(\hat{\delta}) \neq 0$

→ estimador insesgado $\Rightarrow B(\hat{\delta}) = 0$

- Varianza: es la varianza de la variable aleatoria estimada.

$$\text{Var}(\hat{\delta}) = E[(\hat{\delta} - E[\hat{\delta}])^2] = E\{\hat{\delta}^2\} - E\{\hat{\delta}\}^2$$

Es un parámetro importante, ya que cuantos el estimador sea insesgado, si la varianza es grande, entonces la dispersión será alta
 \Rightarrow estimador mal

- Error cuadrático medio (MSE): da una medida de la dispersión del estimador con respecto a la media del valor real

$$\text{MSE}(\hat{\delta}) = E\left\{ (\hat{\delta} - \delta)^2 \right\} = \text{Var}(\hat{\delta}) + [B(\hat{\delta})]^2$$

- Eficiencia: relacionada con la varianza. Dadas 2 reglas para estimar un espacio observable, se dice que la primera ($\hat{\delta}_1$) es más eficiente que la segunda ($\hat{\delta}_2$) si

$$\text{Var}(\hat{\delta}_1) < \text{Var}(\hat{\delta}_2)$$

- Consistencia: relacionada con el ECM. Una regla de estimación es consistente cuando el ECM se hace nulo al tender a infinito el número de medidas realizadas

$$\text{MSE}(\hat{\delta}) = 0 \quad \text{si} \quad N \rightarrow \infty$$

o bien:

$$\lim_{N \rightarrow \infty} \hat{\delta} = \delta$$

Ejercicio: Medida

$$z = \delta + u$$

+ ruido propio del proceso
de medida

\downarrow
Valor real

Modelo estadístico: dar características del ruido

$$\begin{cases} E[u] = 0 \\ E[u_i u_j] = 0 \\ E[u^2] = \sigma_u^2 \end{cases}$$

queremos, a partir de una serie
de medidas,

$$z_i = \delta + u_i \quad i = 1 : N$$

obtener una estimación usando
las medidas:

$$\hat{\delta} = \frac{1}{N} \sum_{i=1}^N z_i$$

Calcular la media, la varianza y el ECE del
estimador

3- ESTIMACIÓN APLICADA A LAS SEÑALES

Nos fijaremos en las señales estocásticas, y que para ellas podemos obtener una caracterización estadística (\bar{x} y S_x) con su realización.

Sobre la realización:

$x(n)$ puede ser

- señal aleatoria
- realización de una señal aleatoria
- segmento de duración finita de una realización

Supongamos que tenemos que modelar una señal aleatoria estacionaria $x(n)$. El estimador parte de una medida que es una realización:

$$\{x(0), x(1), \dots, x(N-1)\}$$

medida $\rightarrow x(n)$, $E\{x(n)\} = \bar{x}_x$

medida $\rightarrow \{x(0), x(1), \dots, x(N-1)\}$

- ESTIMADOR DE LA MÉDIA:

$$\hat{x}_x = \frac{1}{N} \sum_{n=0}^{N-1} x(n)$$

$x(0), x(1), \dots, x(N-1)$ v.s.

$\Rightarrow \hat{x}_x$ es v.s.

$$E[\hat{u}_x] = \frac{1}{N} \sum_{n=0}^{N-1} E[x(n)] = u_x$$

$$\Rightarrow B(\hat{u}_x) = 0 \quad \Rightarrow \text{estimador insesgado}$$

-Varianza del estimador de la media:

Como es insesgado, su varianza coincide con el valor cuadrático medio:

$$Var(\hat{u}_x) = \frac{1}{N} \sum_{n=-N+1}^{N-1} \left(1 - \frac{|n|}{N}\right) C_x(n)$$

Las correlaciones del estimador dependen de las características de la señal aleatoria

¿Es éste un l.s.e. estimado?

- al hacer $N \rightarrow \infty$, debe tender la varianza a cero.

Para ello, hay que confiar suficiente para que sea éste un estimador consistente \rightarrow l.m.e.:

$$\lim_{N \rightarrow \infty} \hat{u}_x = u_x \iff \lim_{N \rightarrow \infty} Var(\hat{u}_x) = 0$$

\uparrow
estimador consistente

Entonces, la condición suficiente para la ergodicidad:

$$\left| \sum_{m=-\infty}^{\infty} |C_x(m)| < \infty \right| \quad (\text{estimador de la media})$$

en los procesos naturales lo cumplen, aunque no hay que pensar que cualquier señal es ergódica.

Ejemplo: generador de tensión que entrega

$+IV$ o $-IV$ con probabilidad $\frac{1}{2}$.

Tensión se da media mil

El valor cero no se calcula de ninguna rectificación, y que de ellos se obtiene $+IV = -IV$, y además se cumple la condición suficiente \Rightarrow no es ergódica en la media

Ejemplo: veamos como nació el estimador de la media con ruido blanco:

$$E[x(u)] = a$$

$$C_x(uu) = Y_x(uu) = E\{x(u) \times (a-u)\} = \begin{cases} 0 & u \neq 0 \\ \sigma_x^2 & u = 0 \end{cases}$$

$$\Rightarrow C_x(u) = \sigma_x^2 S(u)$$

$$\text{Var}(\hat{m}_x) = \frac{1}{N} \sigma_x^2$$

ESTIMADOR DE LA AUTOCORRELACIÓN:

Partimos del mismo modelo y la misma medida.

Calculando el operador esperanza para el
proyecto temporal:

$$r_x(u) = \frac{1}{N} \sum_{n=u}^{N-1} x(n)x^*(n-u), \quad 0 \leq u \leq N-1$$

$$= \frac{1}{N} \sum_{n=0}^{N-1-|u|} x(n)x^*(n-u), \quad -N+1 \leq u \leq 0$$

Estamos promediando y normalizando respecto a N , que
va ser el número de elementos que se miden.

Por ese motivo otra estimación:

$$r'_x(u) = \frac{1}{N-|u|} \sum_{n=u}^{N-1} x(n)x^*(n-u), \quad 0 \leq u \leq N-1$$

Sólo cambia el factor de normalizar:

$$r'_x(u) = \frac{N}{N-|u|} r_x(u) \quad | \quad (\text{Se usan subsec})$$

- Estadísticos:

$$E[r_x(u)] = \frac{1}{N-|u|} \sum_{n=-\infty}^{N-1} E[x(n)x^*(n-u)]$$

$\overbrace{\quad \quad \quad}$
 $\hat{Y}_x(u)$
Sumado $N-|u|$ veces

$$= Y_x(u) \rightsquigarrow \text{estimador insesgado}$$

$$E[r_x(u)] = \frac{N-|u|}{N} Y_x(u) = \left(1 - \frac{|u|}{N}\right) Y_x(u)$$

estimador sesgado, salvo para $u=0$:

$$B(r_x(u)) = \frac{|u|}{N} Y_x(u)$$

- Varianza: en general depende de estadísticos de 4º orden de la señal original, aunque para señales cuasi estacionarias se prueba demostrar que es:

$$\left. \begin{aligned} \text{Var}(r_x(u)) &= \frac{1}{N} \sum_{u=-(N-1-|u|)}^{N-1-|u|} \left(1 - \frac{u+|u|}{N}\right) \left(|Y_x(u)|^2 + Y_x^*(u+|u|) Y_x(u+|u|) \right) \\ \text{Var}(r_x'(u)) &= \frac{N^2}{(N-|u|)^2} \text{Var}(r_x(u)) \end{aligned} \right\}$$

¿Cuando es exacta? $\rightsquigarrow N \rightarrow \infty \Rightarrow \text{Var}(\cdot) = 0$

La condición suficiente para la estabilidad respecto a la autocorrelación es:

$$\left| \sum_{m=-\infty}^{\infty} |r_x(m)|^2 < \infty \right|$$

El primer estimador tiene la característica que la diferencia de dos estimadores sesgados, y es que al crecer indefinidamente el número de valores observados, el sesgo tiende a cero, por lo que se le llama asintóticamente insesgado

Ejercicio: demostrar que para una señal de ruido blanco y gaussiano, el sesgo del estimador $r_x(m)$ es nulo y la varianza es:

$$Var(r_x(m)) = \begin{cases} \frac{2}{N} \sigma_x^4 & m=0 \\ \frac{1}{N} \left(1 - \frac{|m|}{N}\right) \sigma_x^4 & m \neq 0 \end{cases}$$

4.- FILTROS FIR EN CELOSÍA (Opp. 529 - 528)

Se trata de otra estructura para filtros FIR

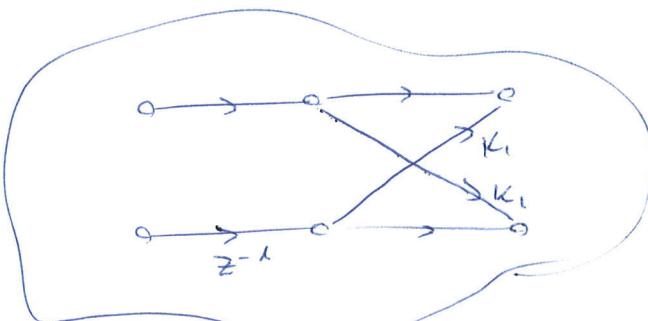
✗ Computacionalmente insuficiente

✓ Roberto enciende a la identificación de los coeficientes

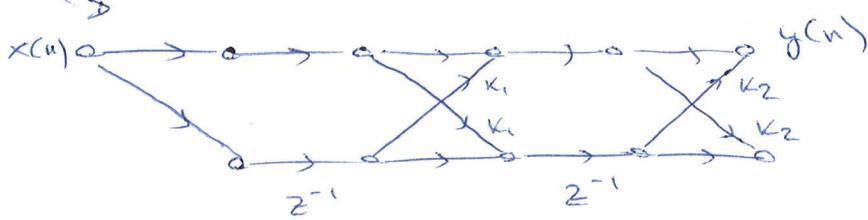
✓ Estímulos en medidas y predicción lineal

✓ Definir una familia de filtros que queden expresada en una forma más compacta

Célula básica:



se conectan
células básicas
en cascada



sin retardo nulo \Rightarrow filtro FIR : $H(z) = \sum_{k=0}^2 a_k z^{-k}$

a_k = "parámetros k "

"coeficientes de reflexión"

"PARCOR" (correlación parcial)

$$h(0) = a_0 = 1$$

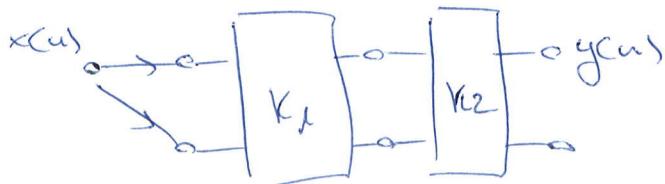
$$h(1) = k_1 (1 + k_2)$$

$$h(2) = k_2$$

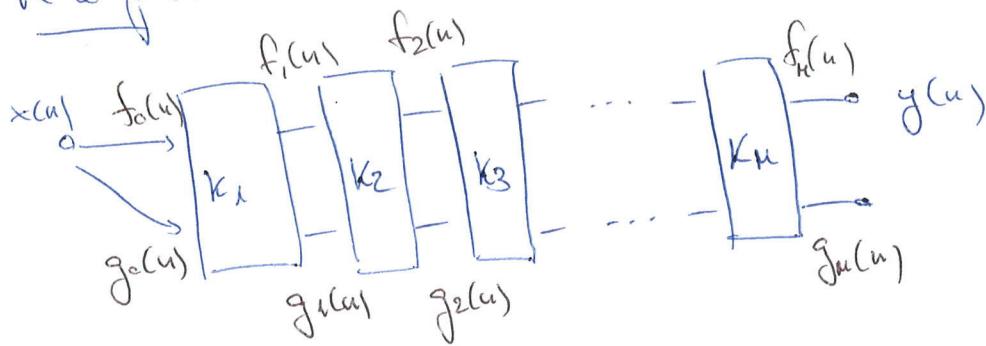
2 retardos \Rightarrow respuesta impulso
limitada $\rightarrow h(2)$

En todos los filtros en celofán, $h(0) = 1$, y la última muestra de $h(n)$ coincide con el parámetro k de la última etapa.

Forma más compacta de representarla :



M stages:



Nos puede interesar caracterizar cada sistema por separado, por ejemplo para prefijar líneas.

$A_i(z)$ = función del sistema cuya entrada es $x(u)$ y salida $f_i(u)$

$$A_i(z) = \frac{F_i(z)}{f_o(z)}$$

$$B_i(z) = \frac{G_i(z)}{G_o(z)}$$

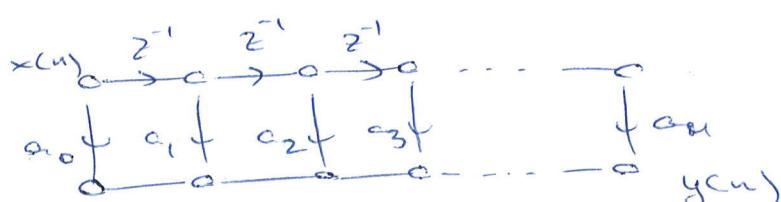
$$F_o(z) = G_o(z) = X(z)$$

Funció de transferencia del sistema completo:

$$H(z) = \frac{Y(z)}{X(z)} = A_M(z) \hookrightarrow \text{filtro FIR de orden } M$$

$$= 1 + \sum_{i=1}^M a_i z^{-i} \quad \begin{cases} a_M(0) = 1 \\ a_M(M) = k_M \end{cases}$$

a_i = coeficientes de la realización directa:



La estructura es celosa
necesita menos parámetros
para caracterizarla

La familia de filtros A_i necesita $\frac{M+1}{2}$ parámetros
para la forma trapezoide, mientras que la forma en
celosas sólo necesitan M parámetros.

$$B_{M-1}(z) = \sum_{k=0}^{M-1} b_{M-k}(k) z^{-k}$$

Filtre FIR de orden M
Parcial en 6 pasos
inferir

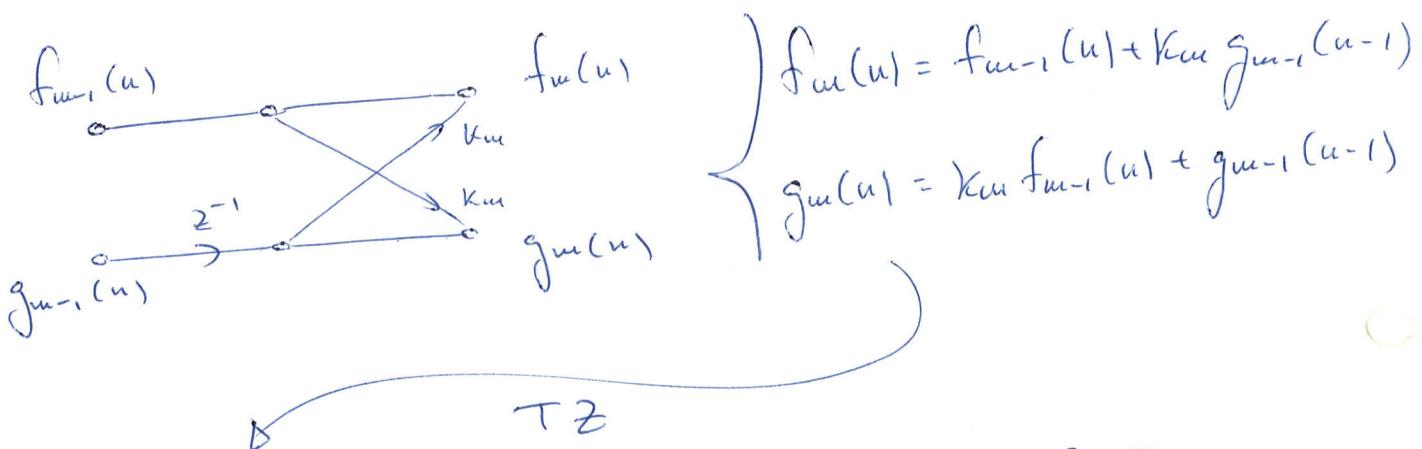
Hoy ver relación entre ambos:

$$\boxed{b_M(k) = a_M(M-k) \quad k=0 : m}$$

Se cambia el orden $\rightarrow b_M(z) = k_m$
 $b_M(u) = 1$

Por tanto: $\boxed{B_M(z) = A_M(z^{-1}) z^{-M}}$

(consecuencia de las propiedades de la transformada Z)



$$\left. \begin{array}{l} F_m(z) = F_{m-1}(z) + k_m z^{-1} G_{m-1}(z) \\ G_m(z) = k_m F_{m-1}(z) + z^{-1} G_{m-1}(z) \end{array} \right\} \quad \begin{array}{l} [1] \\ [2] \end{array}$$

Dividir por $F_0(z) = G_0(z)$:

$$\{1\} \Rightarrow \frac{F_m(z)}{F_0(z)} = \frac{F_{m-1}(z)}{F_0(z)} + k_m z^{-1} \frac{G_{m-1}(z)}{G_0(z)}$$

$$\Rightarrow \boxed{A_m(z) = A_{m-1}(z) + k_m z^{-1} B_{m-1}(z)}$$

$$\{2\} \Rightarrow \frac{G_{m-1}(z)}{G_0(z)} = K_m \frac{F_{m-1}(z)}{F_0(z)} + z^{-1} \frac{G_{m-1}(z)}{G_0(z)}$$

$$\Rightarrow \boxed{B_m(z) = K_m A_{m-1}(z) + z^{-1} B_{m-1}(z)}$$

Con estas ecuaciones se puede hacer una recursión y obtener la función del sistema a partir de los de orden inferior.

$$m=0 \quad \begin{cases} A_0(z) = 1 \\ B_0(z) = 1 \end{cases} \quad m=1 \quad \begin{cases} A_1(z) = 1 + k_1 z^{-1} \\ B_1(z) = k_1 + z^{-1} \end{cases}$$

Aquí se obtienen los coeficientes de la realización directa:

$$b_0(0) = k_1 \quad a_0(0) = 1$$

$$b_0(1) = 1 \quad a_0(1) = k_1$$

Usando la relación entre $A_m(z)$ y $B_m(z)$:

$$\boxed{A_m(z) = A_{m-1}(z) + K_m z^{-m} A_{m-1}(z^{-1})}$$

Podemos obtener así una expresión para los coeficientes de la forma lineal:

$$a_{m-k}(k) = a_{m-1}(k) + K_m a_{m-1}(m-k), \quad 1 \leq k \leq m-1$$

$$a_{m-1}(0) = 1 \quad a_{m-1}(m) = K_m$$

Ejemplo:

$$K_1 = 1/4 \quad K_2 = 1/2 \quad K_3 = 1/3$$

$$A_1(z) = 1 + 1/4 z^{-1}$$

$$A_2(z) = 1 + 1/4 z^{-1} + 1/2 z^{-2} \left(1 + 1/4 z \right) = 1 + 3/8 z^{-1} + 1/2 z^{-2}$$

$$A_3(z) = 1 + 3/8 z^{-1} + 1/2 z^{-2} + 1/3 z^{-3} \left(1 + 3/8 z^{-1} + 1/2 z^{-2} \right) =$$

$$= 1 + 13/24 z^{-1} + 5/8 z^{-2} + 1/3 z^{-3}$$

Ahora buscamos el proceso inverso: a partir de los coeficientes de la respuesta impulsiva, a_i , obtener los de la estructura en celosías, K_i .

$$A_m(z) = A_{m-1}(z) + K_m z^{-1} B_{m-1}(z)$$

$$B_m(z) = B_{m-1}(z) \cdot z^{-1} + K_m A_{m-1}(z)$$

$$\Rightarrow z^1 B_{m-1}(z) = B_m(z) - K_m A_{m-1}(z)$$

$$\Rightarrow A_m(z) = A_{m-1}(z) + K_m (B_m(z) - K_m A_{m-1}(z))$$

$$\Rightarrow A_{m-1}(z) = \frac{A_m(z) - K_m B_m(z)}{1 - K_m^2}$$

Hemos invertido la operación, con lo que podemos obtener las

fuerzas de orden inferior a partir de las de orden superior.

$$A_{m-1}(z) = \frac{A_m(z) - K_m B_m(z)}{1 - K_m^2} = \frac{A_m(z) - K_m z^{-m} A_m(z^{-1})}{1 - K_m^2}$$

Cada vez que se hace la recursión hacia abajo,

se obtiene un K_m

$$\Rightarrow K_m = A_{m,m} \quad \begin{array}{l} \text{(el coeficiente que multiplica} \\ = z^{-m}) \end{array}$$

$$A_m(z) = 1 + \sum_{k=1}^m a_m(k) z^{-k}$$

el primer coeficiente debe ser
1 para que $K_m = a_m(m)$

Pasos:

$$1.- K_m = a_m(m)$$

$$2.- A_{m-1}(z) = 1 + \sum_{k=1}^{m-1} a_{m-1}(k) z^{-k} \rightarrow K_{m-1} = a_{m-1}(m-1)$$

...

No siempre se puede obtener la estructura en celosía

Por ejemplo, si algún x_i tiene módulo 1

Se puede afirmar que, dada una filtro $A_m(z)$, si su inverso, $1/A_m(z)$, es estable, entonces se puede transformarse a estructura en celosía.

- Propiedad: si partiendo de un polinomio $A_m(z)$ se obtienen $|K_m| \leq 1$, $\forall m=1:M$, entonces las raíces de $A_m(z)$ estarán todos dentro del círculo unitario.

Esto es importante para sistemas cuya descripción sea $A_m(z)$, ya que entonces el sistema es estable, al tener todas sus poles dentro del círculo unitario.

Esto se conoce como **Test de Schur-Cohn** y sirve para verificar la estabilidad de un sistema.

La revisión se puede escribir también en función de los coeficientes de la respuesta impulso:

$$K_m = a_m(m)$$

$$a_{m-1}(k) = \frac{a_m(k) - K_m a_m(m-k)}{1 - K_m^2} \quad k=1:m-1$$

RESUMEN:

- De la estructura en celosía muchas veces no interesa sólo la estructura, sino también los subfiltros.
- No se puede cambiar el orden de las etapas.
- Si tenemos un filtro en celosía, siempre se puede obtener el filtro en forma directa.

5.- PREDICCIÓN LINEAL

Es una forma de predecir secuencias numéricas.

Se intenta averiguar el valor de una señal en un instante de tiempo en función de su valor en instantes anteriores:

$$\hat{x}(u) = f(x(u-1), x(u-2), \dots)$$

Además hay que tener en cuenta que sea una combinación lineal:

$$\hat{x}(u) = \sum_{k=1}^p a(k)x(u-k)$$

Por último, será combinación de los p valores anteriores:

$$\hat{x}(u) = - \sum_{k=1}^p a_p(k)x(u-k)$$

Predictor lineal
de orden p

El objetivo de todo predecir es equivocarse lo menos posible

\Rightarrow función de coste: error cuadrático medio (ECM)

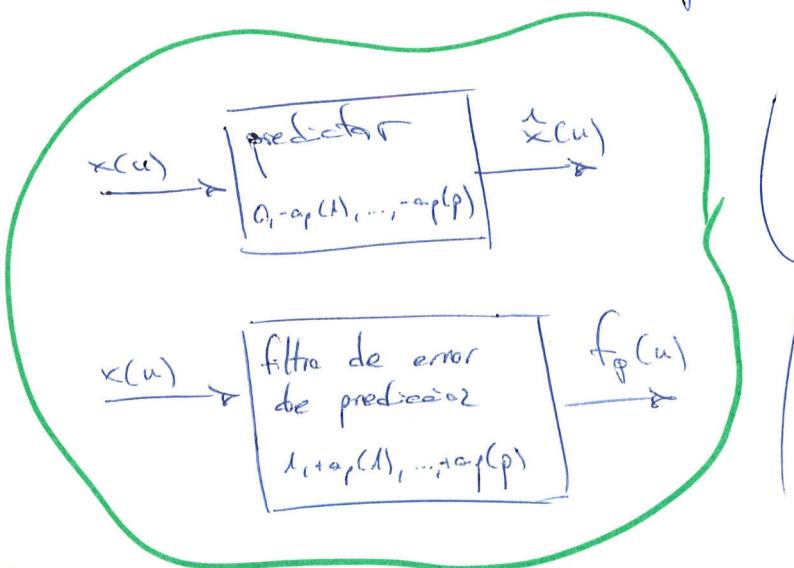
$$e_p(u) = x(u) - \hat{x}(u)$$

señal de error de
predicción

Por convención, en el predictor expresaremos los coeficientes α_p de forma que $\hat{x}(u) = - \sum \dots$, como se definió antes.

$$\begin{aligned} f_p(u) &= x(u) + \sum_{k=1}^p \alpha_p(k) x(u-k) \\ &= \sum_{k=0}^p \alpha_p(k) x(u-k) \quad \alpha_p(0) = b \end{aligned}$$

esto es un filtro tipo FIR



en otras cosas se buscan coeficientes $\alpha_p(i)$ que optimicen el criterio MSE

- Error cuadrático medio:

$$MSE = E\{f_p^2(u)\}$$

potencia de la señal de error

Si la señal es estacionaria, dependerá del instante u , pero a partir de ahora supondremos que $x(u)$ es aleatoria estacionaria

Entonces el objetivo es obtener la fórmula de precios ($a_p(k)$) para lograr un error cuadrático medio mínimo
 \Rightarrow criterio MSE

¿Qué coeficientes cumplen el criterio?

$$f_p(u) = \sum_{k=0}^P a_p(k)x(u-k)$$

$$\text{MSE} = E\left\{f_p^2(u)\right\} = E\left\{\sum_{k=0}^P \sum_{j=0}^P a_p(k)a_p(j)x(u-k)x(u-j)\right\}$$

$$= \sum_{k=0}^P \sum_{j=0}^P a_p(k)a_p(j) \underbrace{E\left\{x(u-k)x(u-j)\right\}}$$

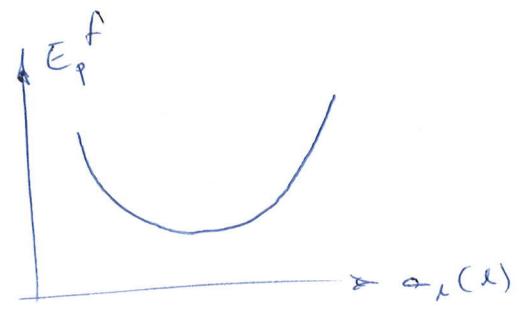
s.c.e. \Rightarrow autocorrelación $\gamma_x(j-k)$

$$\text{MSE} = \sum_{k=0}^P \sum_{j=0}^P a_p(k)a_p(j) \gamma_x(j-k)$$

Ahora hay que obtener los coeficientes que minimizan este fórmula, que es una fórmula cuadrática, y por tanto presenta un único mínimo (absoluto): E_f

Tiene forma de parábola si hay un coeficiente, de parabolíde si hay 2.

Así, se le da la menor de las derivadas parciales con respecto a todos los coeficientes:



$$\frac{\partial \bar{E}_p^f}{\partial \alpha_p(i)} = \underbrace{\sum_{k=0}^p \alpha_p(k) Y_x(i-k)}_{i=j} + \underbrace{\sum_{j=0}^p \alpha_p(j) Y_x(j-i)}_{i=k}, \quad i=1:p$$

$$= 0$$

(\bar{E}_p^f = error cuadrático medio mínimo)

De que la autocorrelación de un señal real es simétrica:

$$\frac{\partial \bar{E}_p^f}{\partial \alpha_p(i)} = 2 \sum_{k=0}^p \alpha_p(k) Y_x(i-k) = 0 \quad i=1:p$$

$$\sum_{k=0}^p \alpha_p(k) Y_x(i-k) = 0 \quad | \quad i=1:p$$

Ahora, como $\alpha_p(0) = 1$:

$$\sum_{k=1}^p \alpha_p(k) Y_x(i-k) = -Y_x(i) \quad | \quad i=1:p$$

J

Tenemos un sistema de p ecuaciones lineales, que puede escribirse así:

$$\begin{bmatrix} Y_x(0) & Y_x(-1) & \dots & Y_x(-p+1) \\ Y_x(1) & Y_x(0) & \dots & Y_x(-p+2) \\ \vdots & \ddots & \vdots & \vdots \\ Y_x(p-1) & \dots & \dots & Y_x(0) \end{bmatrix} \begin{bmatrix} a_p(1) \\ a_p(2) \\ \vdots \\ a_p(p) \end{bmatrix} = \begin{bmatrix} Y_x(1) \\ Y_x(2) \\ \vdots \\ Y_x(p) \end{bmatrix}$$

ECUACIONES NORMALES o DE YULE-WALKER

$$\underline{T_p} = \text{matriz de autocorrelación} = \{Y_x(i-k)\}$$

a_p = vector de coeficientes de predicción

γ = vector de autocorrelación

$$\underline{T_p} \cdot a_p = -\gamma \Rightarrow a_p = -\underline{T_p}^{-1} \cdot \gamma$$

predicción lineal \Rightarrow invertir una matriz \rightarrow problema complejo

Ventaja: la matriz $\underline{T_p}$ admite un algoritmo más sencillo para invertirla gracias a sus características

Valores α (valores del error de predicción):

$$e_p^f = \sum_{k=0}^P \sum_{j=0}^P \alpha_p(k) \alpha_p(j) Y_x(j-k)$$

$$= \sum_{k=0}^P \alpha_p(k) \underbrace{\sum_{j=0}^P \alpha_p(j)}_{\text{excedentes netales de } k=1: p_1} Y_x(j-k)$$

se anula cuando el predictor es óptimo \rightarrow sólo queda para $k=0$

$$MSE = \sum_{j=0}^P \alpha_p(j) Y_x(j)$$

error cuadrático medio
mínimo

- Características de la matriz de autocorrelación:

1.- Es simétrica: $T_p^t = T_p$

esto es porque las señales son reales

2.- Es de Toepitz: se puede obtener la matriz completa a partir de una sola fila.

3.- Es semi-definida positiva:

$$H \subseteq, c^t \cdot T_p \cdot c \geq 0 \Leftrightarrow \lambda_i \geq 0$$

Por tanto, el filtro de predicción óptimo tiene un implementador en donde se verifica $|K_{ik}| \leq b$

Lo único que garantiza que \underline{T}_k es semidefinita positiva son las p. primas reales de $\chi_k(u)$

Estas propiedades de \underline{T}_k están ligadas a
2 propiedades de la autocorrelación:

a) Simetría: $\chi_k(u) = \chi_k(-u)$

b) TF $\{ \chi_k(u) \} \geq 0$

5.1.- ALGORITMO DE LEVINSOHN-DURBIN

(pp. 873-877)

Dado una matriz de orden p , su inversa requiere un trabajo de cálculo de orden p^3 .

Con este algoritmo se consigue orden p^2 , gracias a las propiedades de la matriz de autocorrelación.

Se basa en un método iterativo, que a partir de la solución de orden $n-1$ obtiene la de orden n .

Se obtiene una solución similar a la de los coeficientes de la estructura en cohomo.

$$\underline{T}_p \cdot \underline{\alpha}_p = -\underline{Y}_p \Rightarrow \underline{T}_p \cdot \underline{\alpha}_p^h = -\underline{Y}_p^h$$

$$\begin{bmatrix} Y(0) & Y(1) & \dots & Y(p-2) & Y(p-1) \\ Y(1) & Y(2) & \dots & Y(p-3) & Y(p-2) \\ \vdots & \vdots & & \vdots & \vdots \\ Y(p-1) & Y(p-2) & \dots & Y(1) & Y(0) \end{bmatrix} \begin{bmatrix} \alpha_p(1) \\ \alpha_p(2) \\ \vdots \\ \alpha_p(p) \end{bmatrix} = - \begin{bmatrix} Y(1) \\ Y(2) \\ \vdots \\ Y(p) \end{bmatrix}$$

↓ descomponemos la matriz

$$\begin{bmatrix} \underline{T}_{p-1} & \underline{Y}_{p-1}^h \\ Y_{p-1}^h & Y(0) \end{bmatrix} \quad \text{donde} \quad \underline{Y}_{p-1}^h = \begin{bmatrix} Y(p-1) \\ \vdots \\ Y(1) \end{bmatrix}$$

invertir el orden de los elementos

Recordando: $E_p^f = \sum_{k=0}^p Y(k) \alpha_p(k)$

Intentaremos encontrar la solución de orden m en
según la solución de orden $m-1$:

$$\underline{a}_{m-1} = \begin{bmatrix} a_{m-1} \\ 0 \end{bmatrix} + \begin{bmatrix} d_{m-1} \\ k_m \end{bmatrix} = \begin{bmatrix} a_{m-1}(1) \\ a_{m-1}(2) \\ \vdots \\ a_{m-1}(m) \end{bmatrix}$$

Este k_m tendrá relación con los filtros en cascada

$$\begin{bmatrix} \underline{T_{m-1}} & \underline{Y_{m-1}} \\ \underline{Y_{m-1}}^h & Y(0) \end{bmatrix} \cdot \begin{pmatrix} \underline{\alpha_{m-1}} \\ 0 \end{pmatrix} + \begin{pmatrix} \underline{\beta_{m-1}} \\ \underline{k_m} \end{pmatrix} = - \begin{bmatrix} \underline{Y_{m-1}} \\ Y(m) \end{bmatrix} \quad [3]$$

Tenemos 2 sistemas de ecuaciones, uno de orden $m-1$ y otra ecuación (la m). Si operamos en la expresión anterior:

$$\underline{T_{m-1}} \cdot \underline{\alpha_{m-1}} + \underline{T_{m-1}} \cdot \underline{\beta_{m-1}} + \underline{k_m} \underline{Y_{m-1}}^h = -\underline{Y_{m-1}}$$

Dado que $\underline{T_{m-1}} \cdot \underline{\alpha_{m-1}} = -\underline{Y_{m-1}}$, se tiene la solución:

$$\underline{\beta_{m-1}} = -\underline{T_{m-1}}^{-1} \cdot \underline{Y_{m-1}}^h \cdot \underline{k_m} = \underline{k_m} \underline{\alpha_{m-1}}^h$$

La parte incremental de $\underline{\alpha_m}$ depende de $\underline{\alpha_{m-1}}$

De la ecuación [3] se obtiene la segunda ecuación:

$$\underline{Y_{m-1}}^h \cdot \underline{\alpha_{m-1}} + \underline{Y_{m-1}}^h \cdot \underline{\beta_{m-1}} + \underline{k_m} Y(0) = -Y(m)$$

Usando la anterior expresión para $\underline{\beta_{m-1}}$ y despejando $\underline{k_m}$ se obtiene:

$$\underline{k_m} = -\frac{Y(m) + \underline{Y_{m-1}}^h \cdot \underline{\alpha_{m-1}}}{Y(0) + \underline{Y_{m-1}}^h \cdot \underline{\alpha_{m-1}}} = -\frac{Y(m) + \underline{Y_{m-1}} \cdot \underline{\alpha_{m-1}}^h}{Y(0) + \underline{Y_{m-1}} \cdot \underline{\alpha_{m-1}}^h}$$

$$E_{m-1}^f = Y(0) + \sum_{k=1}^{m-1} Y(k) a_{m-1}(k) = Y(0) + \underbrace{Y_{m-1}}_{h^t} \cdot \underbrace{a_{m-1}}_{\alpha_{m-1}}$$

$$k_m = - \frac{Y(m) + \underbrace{Y_{m-1}}_{h^t} \cdot \underbrace{a_{m-1}}_{\alpha_{m-1}}}{E_{m-1}^f}$$

La recursive givs:

$$a_m(m) = k_m$$

$$a_m(k) = a_{m-1}(k) + k_m a_{m-1}(m-k)$$

$$k=1:m-1$$

Pare numbers obtained via formula recursive per \bar{E}_m^f :

$$\bar{E}_m^f = \sum_{k=0}^m Y(k) a_{m-1}(k) = Y(0) + Y(m) k_m + \sum_{k=1}^{m-1} Y(k) a_{m-1}(k) =$$

$$= Y(0) + k_m Y(m) + \sum_{k=1}^{m-1} Y(k) (a_{m-1}(k) + k_m a_{m-1}(m-k)) =$$

$$= \bar{E}_{m-1}^f + k_m \left(Y(m) + \sum_{k=1}^{m-1} Y(k) a_{m-1}(k) \right)$$

$$- k_m \bar{E}_{m-1}^f$$

$$\boxed{\bar{E}_m^f = \bar{E}_{m-1}^f (1 - k_m^2)}$$

-Resumen: Algoritmo de Levinson-Durbin:

1.- $E_0^f = \gamma(0)$ (prediccion nula)

2.- $a_{uu}(0) = 1$

$$a_{uu}(k) = a_{uu}(k-1) + k u a_{uu}(u-k), \quad k=1, u-1$$

$$k u = - \frac{\sum_{k=1}^{u-1} a_{uu}(u-k) \gamma(k)}{E_{uu-1}^f} = c_{uu}(u)$$

3.- $E_u^f = E_{uu}^f (1 - k_u^2)$

¿Cómo ha de ser k_u ?

Se debe verificar

$$E_u^f \leq E_{uu-1}^f, \text{ por lo que}$$

$$|k_u| \leq 1$$

Ojo: signo del problema

Con eso se asegura que $\underline{\underline{\Gamma}}_u$ es una matriz de autoresolvible simétrica.

, w>

En el caso de $|K_{kk}| = 1$, tendremos una señal totalmente predecible. Ocurre con señales (sus realizaciones) periódicas, que solo tienen componentes espectrales discretas:

$$x(u) = \sum_i \alpha(i) e^{j(\omega_0 u + \phi_i)}$$

Si el algoritmo se ejecutase sobre p procesadores en paralelo, requeriría un orden $\mathcal{O}(p \log p)$ de unidades de tiempo.

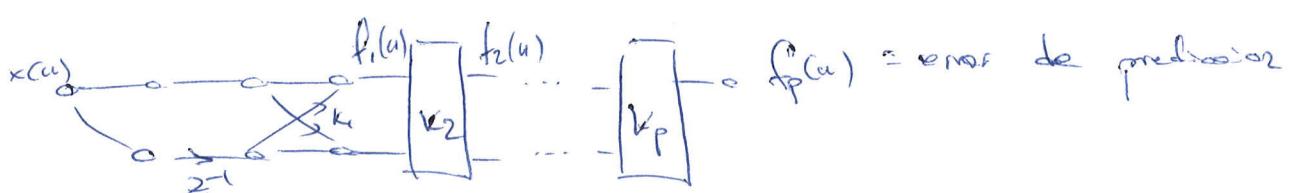
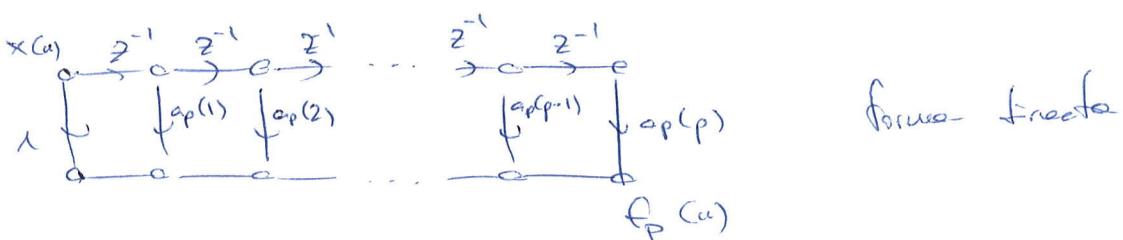
- Relación entre la predicción lineal y los filtros en cebolla:

algoritmo de Leimkufer-Durbin:

$$\left| \begin{array}{l} a_{kk}(0) = 1 \\ a_{kk}(k) = a_{k-1,k}(k) + K_{kk} a_{k-1,k}(k-k) \\ a_{kk}(m) = K_{kk} = \dots \end{array} \right.$$

K_{kk} son los coeficientes de la realización en cebolla del predictor óptimo, por lo que al hacer el algoritmo de Leimkufer-Durbin se tiene directamente el filtro en cebolla.

Predictor de orden p:



Una ventaja de la estructura en celada es que en la salida de cada etapa esté el error de predicción del predictor óptimo para los ordenes inferiores.

En la forma directa ^{no} aparece por ningún lado ese error.

Otra ventaja es que la estructura varía el error de predicción hacia atrás ("backwards"), en los ordenes inferiores.

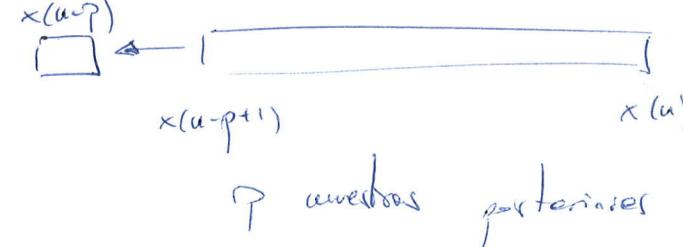
- PREDICCIÓN LINEAL HACIA ATRÁS:

Es igual, pero con el tiempo en sentido inverso

Presencia en un predictor de orden p :

$$\begin{matrix} x(u) & x(u-1) & x(u-2) & \dots & x(u-p+1) & x(u-p) \end{matrix}$$

predicar este dato usando los p anteriores

Predicción local atras: 

$$\hat{x}(u-p) = - \sum_{k=0}^{p-1} b_p(k) x(u-k)$$

$$x(u-p) - \hat{x}(u-p) = g_p(u) = \text{error de predicción}$$

Se pone $g_p(u)$ para que el filtro sea causal

$$g_p(u) = \sum_{k=0}^p b_p(k) x(u-k) \quad b_p(p) = 1$$

Criterio: minimizar $E[g_p^2(u)] = \text{MSE}$

$$\begin{bmatrix} y(0) & \dots & y(p-1) \end{bmatrix} \begin{bmatrix} b_p(p-1) \\ \vdots \\ b_p(0) \end{bmatrix} = - \begin{bmatrix} y(1) \\ \vdots \\ y(p) \end{bmatrix}$$

$$\Rightarrow \boxed{b_p(k) = \alpha_p(p-k) \quad k=0:p}$$

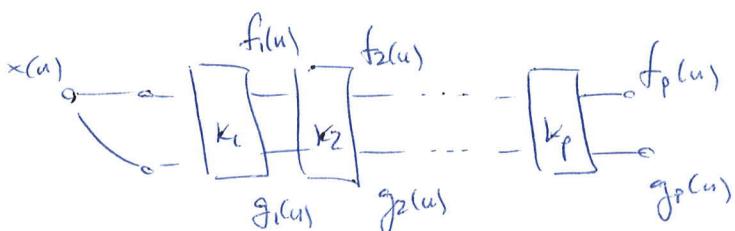
Es decir, la predicción local atras es la misma solución que local adelante, en el orden de los coeficientes cambiado.

Al ser matemáticamente idénticos, el error de predicción será el mismo:

$$E_p^f = E_p^b \quad \begin{cases} b = \text{backward} \\ f = \text{forward} \end{cases}$$

$$\boxed{E_p^f = E_p^b = \sum_{k=0}^P b_p(p-k) Y(k) = \sum_{k=0}^P a_p(k) Y(k)}$$

El error $g_p(u)$ aparece también en la estructura en celosía, en los salidas inferiores:



5.2.- PROPIEDAD DE LA SEÑAL DE ERROR DE PREDICCIÓN

\Rightarrow ORTOGONALIDAD DE LA SEÑAL DE ERROR DE PREDICCIÓN CON LOS DATOS

$$x(u) \quad x(u-1) \quad \dots \quad x(u-p)$$

$\hat{x}(u)$: se cumple para el predictor óptimo

$f_p(u) = x(u) - \hat{x}(u)$ es ortogonal a los datos usados para obtener la predicción (los p anteriores):

$$E[f_p(u) \times (u-k)] = 0 \quad k = l:p$$

El predictor óptimo saca todo la información posible de las muestras pesadas para hacer la predicción, por lo que no queda información en el sonido de error de las demás muestras.

Demostrar:

$$\begin{aligned} E\left[\sum_{i=0}^P \alpha_p(i) \times (u_i - i) \times (u - k)\right] &= \sum_{i=0}^P \alpha_p(i) E[(u_i - i) \times (u - k)] = \\ &= \sum_{i=0}^P \alpha_p(i) \gamma(k-i) = 0 \quad \text{= ecuaciones de Cole-Watson,} \\ &\quad \Leftrightarrow \text{el predictor es óptimo, vale cero.} \end{aligned}$$

$$k = l:p$$

- OTRA FORMA DE DEFINIR K_{ui} :

$$K_{ui} = - \frac{E[f_{ui-1}(u) g_{ui-1}(u-1)]}{\sqrt{E_{ui-1}^f} \sqrt{E_{ui-1}^g}}$$

Así se justifica el nombre de PARCOR (Correlación Parcial)

$$f_{m-1}(u) = x(u) + \sum_{k=1}^{m-1} a_{m-1}(k) \times (u-k)$$

$$g_{m-1}(u-1) = x(u-m) + \sum_{k=1}^{m-1} a_{m-1}(k) \times (u-m+k)$$

$$f_{m-1}(u) \rightarrow \begin{matrix} x(u) & x(u-1) \\ \circ & \circ \\ \dots & \dots \\ x(u-m+1) & x(u-m) \\ \circ & \circ \end{matrix}$$

$$g_{m-1}(u-1) \rightarrow \begin{matrix} \dots & \dots & \dots \end{matrix}$$

$$\text{óptima} \Rightarrow E_{m-1} = E_{m-1}^b$$

$$E \left[f_{m-1}(u) \left(\sum_{k=1}^{m-1} a_{m-1}(k) \times (u-m+k) + x(u-m) \right) \right] =$$

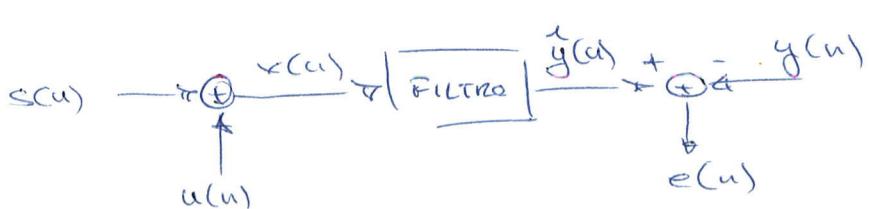
$$= E \left[f_{m-1}(u) \times (u-m) \right] = E \left[\sum_{k=0}^{m-1} a_{m-1}(k) \times (u-k) \times (u-m) \right] =$$

$$= \sum_{k=0}^{m-1} a_{m-1}(k) \times (u-k) \quad \Rightarrow$$

esta es la que aparece en el numerador en el algoritmo de Leimer-Durbin

S.3.- FILTRADO LINEAL ÓPTIMO (FILTRO DE WIENER) (pp 889-898)

Tenemos parte del conocimiento de una señal y queremos obtener una estimación de la señal mediante filtro:



Filtro de Wiener: minimiza la varianza de $e(u)$

$$J = E[e^2(u)] \Rightarrow \text{minimizar } J$$

Veamos el caso en que la estima sea un filtro de medias, P :

$$\hat{y}(u) = \sum_{k=0}^{P-1} a_p(k) x(u-k)$$

filtro FIR de p coeficientes

$$J = E[(y(u) - \hat{y}(u))^2] = E\left[\left(y(u) - \sum_{k=0}^{P-1} a_p(k) x(u-k)\right)^2\right] =$$

$$= E[y^2(u)] + \sum_{k=0}^{P-1} \sum_{j=0}^{P-1} a_p(k) a_p(j) E[x(u-k)x(u-j)] -$$

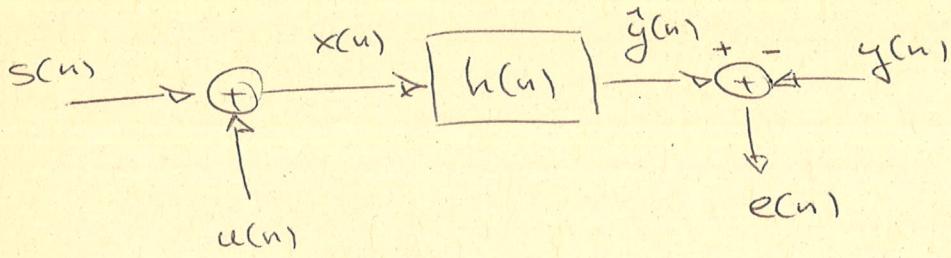
$$- 2 \sum_{k=0}^{P-1} a_p(k) E[y(u)x(u-k)]$$

$$J = P_y + \sum_{k=0}^{P-1} \sum_{j=0}^{P-1} a_p(k) a_p(j) \gamma(j-k) - 2 \sum_{k=0}^{P-1} a_p(k) \gamma_{xy}(-k)$$

Ahora se deriva respecto a los coeficientes:

$$\frac{\partial J}{\partial a_i} = \underbrace{\sum_{k=0}^{P-1} a_p(k) \gamma(j-k)}_{\text{simetría}} + \underbrace{\sum_{i=0}^{P-1} a_p(i) \gamma(i-j)}_{\text{simetría}} - 2 \gamma_{xy}(-i) = 0$$

FILTRADO PARA DE WIENER



$$J = E[e^2(u)]$$

$$\boxed{\hat{y} = \sum_{k=0}^{P-1} h(k) \times (u-k)}$$

$$J = E[(y(u) - \hat{y}(u))^2] = E\left[\left(y(u) - \sum_{k=0}^{P-1} h(k) \times (u-k)\right)^2\right] =$$

$$= E[y^2(u)] + E\left[\sum_{k=0}^{P-1} \sum_{j=0}^{P-1} h(k) h(j) \times (u-k) \times (u-j)\right] - \\ - 2 \sum_{k=0}^{P-1} h(k) \underbrace{E[x(u-k) y(u)]}_{Y_{xy}(-k)}$$

$$\boxed{J = P_y + \sum_{k=0}^{P-1} \sum_{j=0}^{P-1} h(k) h(j) Y_x(j-k) - 2 \sum_{k=0}^{P-1} h(k) Y_{xy}(-k)}$$

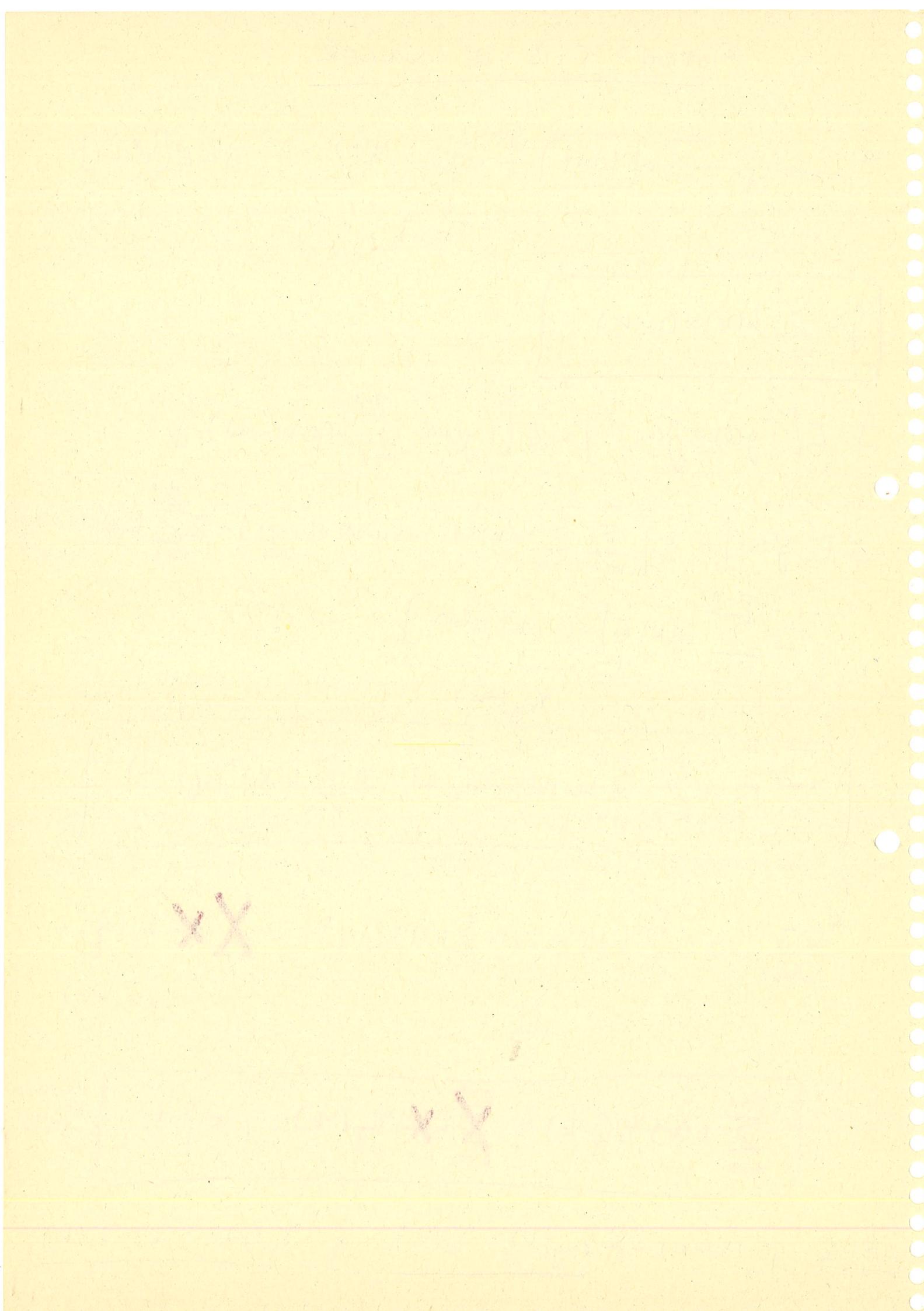
$$\frac{\partial J}{\partial h(j)} = 0 = \sum_{k=0}^{P-1} h(k) Y_x(j-k) + \sum_{i=0}^{P-1} h(i) Y_x(i-j) - 2 \cancel{\sum_{k=0}^{P-1} h(k) Y_{xy}(-k)}$$

↓

$j = 0 : P-1$

$$\boxed{\sum_{k=0}^{P-1} h(k) Y_x(j-k) = \cancel{\sum_{k=0}^{P-1} h(k) Y_{xy}(-k)} \quad j = 0 : P-1}$$

Filtreado: $y(u) = s(u) \Rightarrow Y_{xy}(k) = Y_s(k)$ y $Y_x(k) = Y_s(k) + Y_u(k)$



$$\Rightarrow \sum_{k=0}^{p-1} a_p(k) Y_{xg}(j-k) = Y_{xg}(-j) \quad j=0:p-1$$

$$\begin{bmatrix} Y(0) & \dots & Y(p-1) \\ \vdots & \ddots & \vdots \\ Y(p-1) & \dots & Y(0) \end{bmatrix} \begin{bmatrix} a_p(0) \\ \vdots \\ a_p(p-1) \end{bmatrix} = \begin{bmatrix} Y_{xg}(0) \\ \vdots \\ Y_{xg}(-p+1) \end{bmatrix}$$

○ igual que en predicción lineal pero es una correlación cruzada

La solución de este sistema es el filtro de Wiener:

$$a_p = T_p^{-1} \cdot Y_{xg}$$

La matriz T_p es siempre
Toeplitz (matriz de filas
y de columnas)

6.- MODELO DE SEÑALES

Objetivo: ser capaz de caracterizar estadísticamente una señal de forma que sea más a su vez capaces de generar procesos que tengan las mismas estadísticas definidas previamente.

La forma de hacer este modelo es el filtro de riido blanco.

En este apartado nos referiremos a la autocorrelación (estadística de 2º orden) o a su transformada de Fourier, la DFT.

$$\text{Notación: } Y_x(k) \xleftarrow{\text{TF}} T_x(f)$$

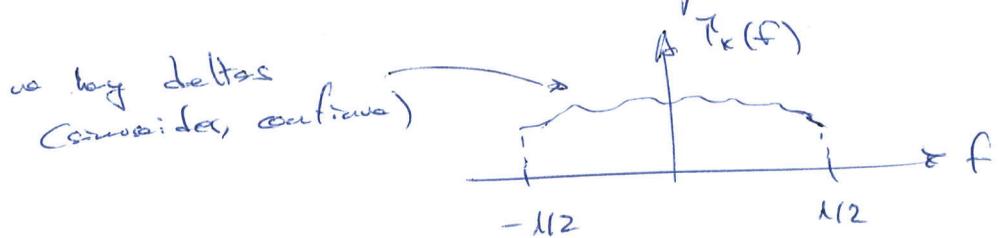
$$T_x(f) = \sum_{k=-\infty}^{\infty} Y_x(k) e^{-j2\pi fk}$$

Basado en estas condiciones, determinadas señales admiten este resultado:

- Condiciones de Paley-Wiener:

Partidas de una señal aleatoria estacionaria, $x(u)$, y su DFT, que cumple:

a) No tiene componentes discretas

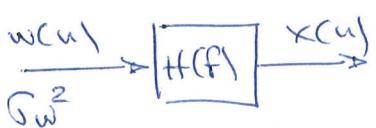


$$b) \int_{-1/2}^{1/2} |T_x(f)| df < \infty$$

Otra condición que se exige es integrable, para que vemos

- proporcionalidad: $T_x(f) \neq 0 \forall f$

ruido blanco



cuando se cumplen las condiciones, se puede generar x(u) partiendo de ruido blanco y haciendo uso de filtros

Matemáticamente, esto quiere decir que podemos expresar la DEP de $x(u)$ como:

$$T_x(z) = G_w^2 H(z) H(l(z))$$

Oj@: transformada Z

Note:



$$x(u), h(u), y(u) \in \mathbb{R}$$

$$\begin{cases} Y_y(u) = Y_x(u) * h(u) * h(-u) \\ \Rightarrow T_y(f) = T_x(f) \cdot |H(f)|^2 \\ T_y(z) = T_x(z) H(z) H(l(z)) \end{cases}$$

Bajo estas condiciones, $H(z)$ y $H(l(z))$ son tales que $H(f)$ corresponde a un filtro causal y estable, y ademas el filtro inverso tambien es causal y estable.

$$Y_x(k) = Y_x(-k) \xrightarrow{TZ} T_x(z) = T_x(l(z))$$

la DEP de una señal real verifica este igualdad

Ahora nos centramos en señales cuya DEP o TZ se puede expresar como cocientes de polinomios:

$$T_x(z) = \frac{P(z)}{Q(z)}$$

Al calcular el filtro deseado, con polinomios finitos, se tiene un filtro IIR realizable.

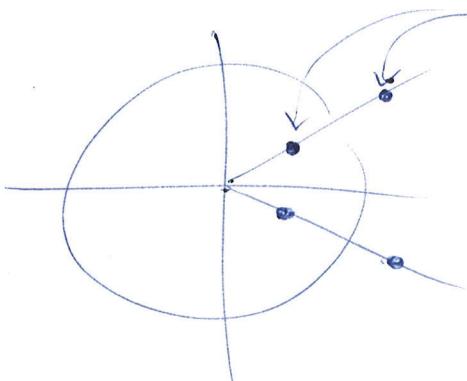
Esto se consigue con un orden del polinomio suficientemente alto.

$$T_x(z) = \frac{P(z)}{Q(z)} = \frac{P(z^{-1})}{Q(z^{-1})}$$

$P(z) = P(z^{-1})$ sólo si los coeficientes del polinomio son simétricos.

Si tenemos una raíz $z=c$ ($P(c)=0$), entonces c^{-1} debe ser una raíz.

Sólo las reales \Rightarrow coeficientes reales \Rightarrow raíces que aparecen en pares complejos conjugados



por lo tanto: por cada cero en el interior del círculo unitario hay otro fuera con la misma fase

Se puede expresar entonces $T_x(z)$ de la siguiente forma:

$$T_x(z) = \frac{B(z)}{A(z)} \frac{B(z^{-1})}{A(z^{-1})}$$

→ se distribuyen los zeros de los polinomios

Hacemos que $A(z)$ y $B(z)$ tengan sus zeros dentro del círculo unitario.

Se construye entonces el sistema de la siguiente forma:

$$\frac{w(u)}{\zeta_w^2} \rightarrow \boxed{H(z) = \frac{B(z)}{A(z)}} \xrightarrow{x(u)} \text{estable y de fase unitaria}$$

La descomposición de $A(z)$ y $B(z)$ es única.

Se fuerza que el filtro tenga $h(0) = 1$

Razón: $\cancel{h(0)} = \lim_{z \rightarrow \infty} H(z) = 1$

Así se consigue que la descomposición sea única

También identificamos así ζ_w^2 , $A(z)$ y $B(z)$ inmediatamente.

Ejemplo: $T_x(z) = -\frac{4z^2 - \lambda z + 4z^{-2}}{4z^2 + \lambda z + 4z^{-2}}$

$\left. \begin{array}{l} \text{poles: } \pm 1/2i, \pm 2i \\ \text{zeros: } \pm \lambda/2, \pm 2 \end{array} \right\}$

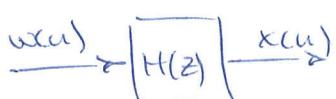
$$T_x(z) = -\frac{4(z - \lambda/2z^{-1})(z + \lambda/2z^{-1})(z - 2z^{-1})(z + 2z^{-1})}{4(z - i/2z^{-1})(z + i/2z^{-1})(z - i/2z)(z + i/2z)}$$

$$= -\frac{(z - \lambda/2z^{-1})(z + \lambda/2z^{-1})}{(z - i/2z^{-1})(z + i/2z^{-1})} \frac{(z - \lambda/2z)(z + \lambda/2z)}{(z - i/2z)(z + i/2z)} = \tilde{\gamma}_w^2 H(z) H(z^{-1})$$

$$\tilde{\gamma}_w^2 = \lambda \quad H(z) = \frac{1 - \lambda/4z^{-2}}{1 + \lambda/4z^{-2}}$$

————— o —————

Queremos modelar:



$$T_x(z) = \tilde{\gamma}_w^2 H(z) H(z^{-1})$$

$$Y_x(u) = \bar{T} z^{-1} \quad \left. \right\} T_x(z)$$

$w(u)$ ruido blanco gaussiano con $S_w(u) = \tilde{\gamma}_w^2 S(u)$

señal observada: $x(0), \dots, x(N-1)$

• estimar los parámetros del modelo:

$$H(z) = \frac{B(z)}{A(z)} = \frac{\lambda + \sum_{k=1}^q b_k z^{-k}}{\lambda + \sum_{k=1}^p a_k z^{-k}}$$

$$x(u) = \sum_{k=0}^q b_k w(u-k) - \sum_{k=1}^p a_k x(u-k)$$

→ filtro IIR

Este modelo se denomina ARMA, y es el más general que veremos, ya que tiene poles y ceros.

- Modelo MA ("Moving Average"):

Es un predecir con ventana deslizante

$$H(z) = 1 + \sum_{k=1}^q b_k z^{-k}$$

$$x(u) = \sum_{k=0}^q b_k w(u-k)$$

(filtro FIR)

- Modelo AR (Autoregresivo):

$$H(z) = \frac{1}{1 + \sum_{k=1}^p a_k z^{-k}}$$

$$x(u) = w(u) - \sum_{k=1}^p a_k x(u-k)$$

(filtro IIR)

- TEOREMA DE WOLD: todos los modelos tienden a ser iguales asintóticamente

Matemáticamente, el más simple es el AR.

Queremos ver relación entre la autocorrelación y los parámetros del modelo.

Considerando el caso general (AR(p)) y sin las restricciones de estacionariedad:

$$\begin{aligned} \gamma_x(u) &= E[x(u)x(u-u)] = \\ &= \sum_{k=0}^q b_k E[w(u-k)x(u-u)] - \sum_{k=1}^p a_k E[x(u-k)x(u-u)] = \\ &= \sum_{k=0}^q b_k Y_{wx}(u-k) - \sum_{k=1}^p a_k \gamma_x(u-k) \end{aligned}$$

Siendo $x(u) = \sum_{i=0}^{\infty} h(i) w(u-i)$, $Y_{wx}(u-k)$ queda:

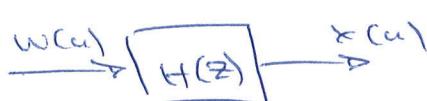
$$\begin{aligned} Y_{wx}(u-k) &= E[w(u)x(u-u)] = \\ &= \sum_{i=0}^{\infty} h(i) E[w(u)w(u-u-i)] = \underbrace{\gamma_w^2 h(-u)}_{\gamma_w(u+i)} \end{aligned}$$

Fundamente:

$$Y_x(u) = T_w^2 \sum_{k=u}^{q} b_k h(k-u) - \sum_{k=1}^P a_k Y_x(u-k) ; 0 \leq u \leq q$$

$$u > q \Rightarrow 0$$

Se resuelve es compleja, debido a la relación de $h(u)$ con los coeficientes.



$$H(z) = \frac{\sum_{k=0}^{q} b_k z^{-k}}{\sum_{k=1}^P a_k z^{-k}}$$

$$a_0 = b_0 = 1$$

$$Y(u) = - \sum_{k=1}^P a_k Y(u-k) \quad u > q$$

$$Y(u) = - \sum_{k=1}^P a_k Y(u-k) + T_w^2 \sum_{k=u}^{q} b_k h(k-u) \quad 0 \leq u \leq q$$

(ADM)

relaciones no lineales
($h(u)$ = respuesta := pulsiva)

- Modelo AR: $g=0 \Rightarrow$ desaparecen las reacciones no lineales

$$\left\{ \begin{array}{l} Y(u) = - \sum_{k=1}^P a_k Y(u-k) \quad u > 0 \\ Y(0) = - \sum_{k=1}^P a_k Y(-k) + \sigma_w^2 \underbrace{\sum_{k=0}^0 b_k h(k)}_{b_0 h(0)} \\ = - \sum_{k=1}^P a_k Y(-k) + \sigma_w^2 \end{array} \right.$$

Las primeras líneas ($u > 0$), para $u = t$, se corresponde con las ecuaciones lineales del predictor lineal óptimo.

Además:

$$\sigma_w^2 = \sum_{k=1}^P a_k Y(k) = \sum_{k=0}^P a_k Y(k)$$

Expresar que nos daña el error de predicción para el predictor óptimo.

Por tanto, hay una relación entre el modelo AR y el predictor lineal

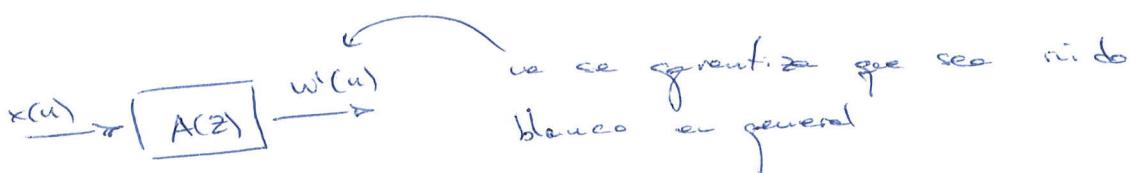
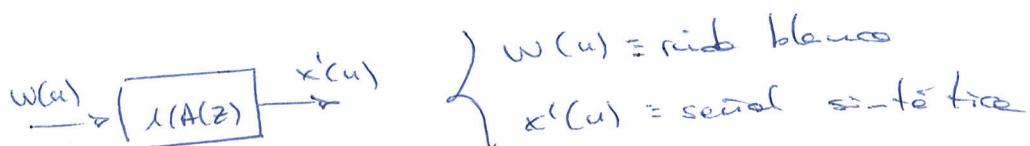
En el problema de modelado, partimos de una señal $x(u)$, de la cual podemos tener $y(0) : y(p)$, y se dan datos o estimaciones de las mismas a partir de un segmento de la señal. Tenemos que construir un modelo que permite reconstruir la señal a partir de su autocorrelación

\Rightarrow construir un predictor óptimo

$$x(u) \rightarrow y(0) \dots y(p) \rightarrow \text{predictor óptimo: } \{a_k, k=1:p\}$$

\downarrow

$$x(0) \dots x(N-1) \qquad \qquad \qquad E_p \rightarrow \sigma_w^2$$



$$\text{P}_w(w'(u)) = \text{P}_w(w(u))$$

Algo que tenemos el ~~que~~ ~~que~~: tenemos una señal aleatoria que se obtiene mediante un modelo, $x(u)$. Nos preguntamos cómo podemos hacer la predicción.

$$\frac{w(u)}{\sigma_w^2} \rightarrow \boxed{\lambda/A(z)} \rightarrow x(u)$$

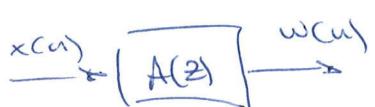
$$A(z) = \sum_{k=0}^p a_k z^{-k}$$

modelo AR de orden p

Predictor óptimo de orden p' para $x(u)$:

- En primer lugar, consideremos el caso $\underline{p'} = p$:

⇒ ecuaciones normales



el filtro predictor óptimo es
el inverso del de modelo

Todo es así porque $a_0 = 1$ ($h(0) = 1$)

$$E_p^f = G_w^2$$

- $\underline{p'} < p$:

Coefficientes del predictor óptimo: $a_{p'}(k) \neq 0 \forall k$

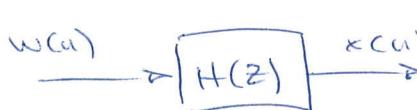
Son coeficientes de la forma directa, por lo que
no tienen rebotes. Lo tienen los de la celosía

- $\underline{p'} > p$:

El modelo AR complica las ecuaciones hasta el
infinito, pero los sumandos van hasta p , por tanto:

$$a_{p'}(k) = 0 \quad k = p+1 : p'$$

De forma intuitiva: si el predictor de orden p nos permite
obtener $w(u)$, no podemos predecir nada más, por lo
que no tiene de sentido aumentar el orden del filtro
de predicción.



$$T_x(f) = \sigma_w^2 |H(f)|^2$$

$$H(z) = \frac{\sum_{k=0}^{\infty} b_k z^{-k}}{\sum_{k=0}^{\infty} a_k z^{-k}}$$

Para una señal con cierta DFP, $T_x(f)$, ¿cuál se prefiere predecir (mínimo error) o predicción lineal?

$H(z)$ genérica $\rightarrow h(u)$ genérica

$$H(z) = \sum_{k=0}^{\infty} h(k) z^{-k} = \frac{1}{\sum_{k=0}^{\infty} a_k z^{-k}} = \frac{1}{A(z)}$$

modelos AR y ARMA asintóticamente equivalentes

\Rightarrow el filtro de predicción óptima es el inverso, $A(z)$



σ_w^2 = potencia del error nómico

La potencia con que se excite el modelo será la mínima que se puede obtener con un predictor lineal óptimo de orden de predicción arbitrariamente grande

$$\int_{-1/2}^{1/2} (\ln |H(f)|)^2 df = 0$$

$$\int_{-1/2}^{1/2} \ln G_w^2 df = \int_{-1/2}^{1/2} \ln T_x(f) df$$

$$\ln G_w^2 = \int_{-1/2}^{1/2} \ln T_x(f) df \quad \text{Se nos permite obtener la petencia de error mínima a partir de la DEP de la señal}$$

Si queremos producir una señal blanca:

$$T_x(f) = c = \int_{-1/2}^{1/2} T_x(f) df$$

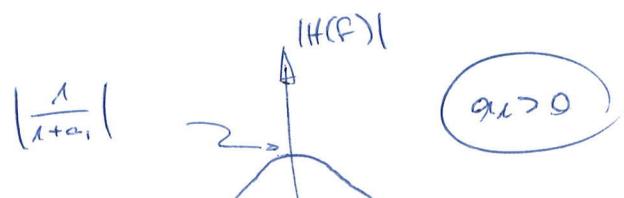
$$\ln G_w^2 = \ln c$$

\Rightarrow se se puede producir una

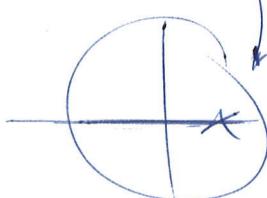
6.1.- Ejemplos de modelos

a) AR(1)

$$H(z) = \frac{1}{1 + \alpha_1 z^{-1}} \quad \alpha_1 \in \mathbb{R} \quad \Rightarrow z = \alpha_1 \text{ es polo único (estable)}$$

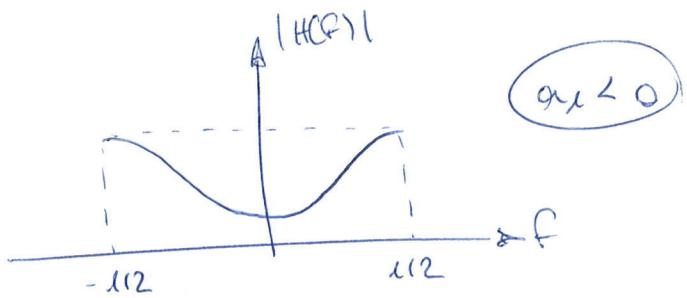


$$|\alpha_1| > 0$$



$$\left| \frac{1}{1 - \alpha_1} \right| \quad \text{at } f = 0$$

pico en $f=0$



$$T_k(f) = \zeta \omega^2 |H(f)|^2$$

picos en $f = \pm 1/2$

Ecuaciones normales

$$\left. \begin{array}{l} Y(0) + Y(-1) \alpha_1 = \zeta \omega^2 \\ Y(1) + \alpha_1 Y(0) = 0 \end{array} \right\}$$

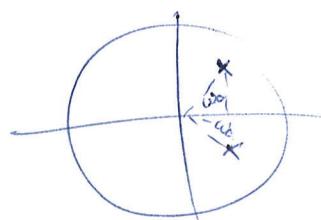
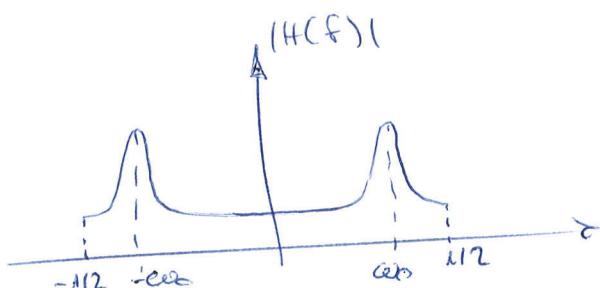
$$\alpha_1 = -\frac{Y(1)}{Y(0)}$$

$$\zeta \omega^2 = Y(0) - \frac{Y^2(1)}{Y(0)}$$

b) AR(2)

$$H(z) = \frac{1}{1 + \alpha_1 z^{-1} + \alpha_2 z^{-2}}$$

→ 2 poles conjugados ($\alpha_k \in \mathbb{R}$)



Con AR(2) podemos modelar procesos que tengan un pico.

Ecuaciones normales:

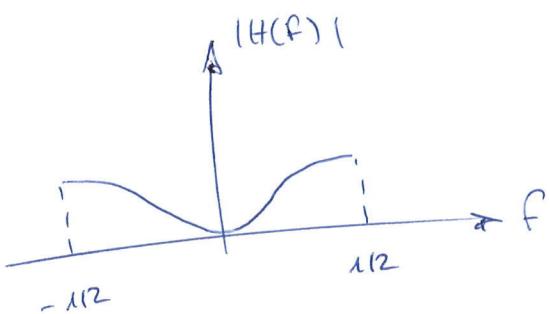
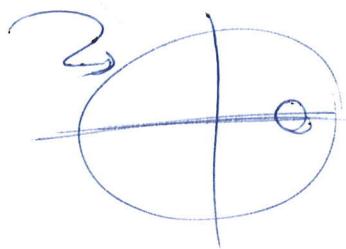
$$\left. \begin{array}{l} Y(0) + Y(1) \alpha_1 + Y(2) \alpha_2 = \zeta \omega^2 \\ Y(1) + Y(0) \alpha_1 + Y(1) \alpha_2 = 0 \\ Y(2) + Y(1) \alpha_1 + Y(0) \alpha_2 = 0 \end{array} \right\}$$

$$\alpha_1 = \frac{Y(1)(Y(2) - Y(0))}{Y^2(0) - Y^2(1)}$$

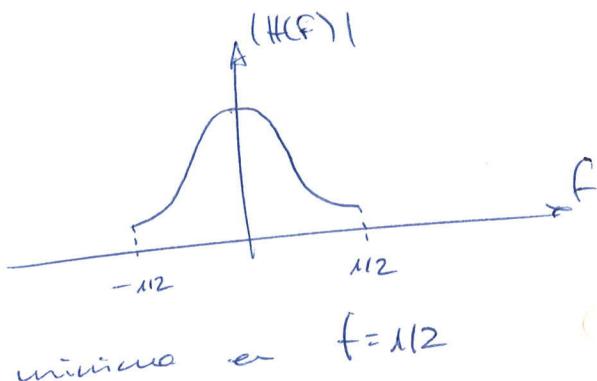
$$\alpha_2 = \frac{Y^2(1) - Y(0)Y(2)}{Y^2(0) - Y^2(1)}$$

$$c) \underline{HAT(\lambda)} \Rightarrow h(k) = b_k$$

$$H(z) = \lambda + b_1 z^{-1} \Rightarrow \lambda < 0$$



mínimo en $f=0$
 \Rightarrow cero próximo al circuito



Ecuaciones normales:

$$\left\{ \begin{array}{l} Y_k(u) = \Gamma_w^2 \sum_{k=u}^q b_k h(k-u) \quad 0 \leq u \leq q \\ Y_k(u) = 0 \quad u > q \end{array} \right.$$

$$Y(0) = \Gamma_w^2 (b_0^2 + b_1^2) = \Gamma_w^2 (\lambda + b_1^2)$$

$$Y(1) = \Gamma_w^2 b_1 b_0 = \Gamma_w^2 b_1$$

$$\Rightarrow b_1^2 - \frac{Y(0)}{Y(1)} b_1 + \lambda = 0$$

En este modelo, las relaciones entre los autores de los sistemas
y los coeficientes no son lineales, sino de 2º orden en el
caso de $HAT(\lambda)$

$$b_1 = \frac{\frac{Y(0)}{Y(1)} \pm \sqrt{\frac{Y^2(0)}{Y^2(1)} - 4}}{2}$$

Para $b_1 \in \mathbb{R}_+$ hay una restricción en la ordenadas:

$$\left| \frac{\gamma(0)}{\gamma(1)} \right| \geq 2$$

Los modelos Arost y RGA no se resuelven directamente de las ecuaciones.

A medida que se sube el orden, se cierran degeneraciones en la respuesta.

7.- Análisis Espectral Clásico

Bibliografía: Proakis - Manolakis 916-929

Se basa en el periodograma: básicamente consiste en calcular la TF y decor suavizar al cuadrado.

Propuesto por Schwert en 1898

Hay 3 métodos básicos:

1.- Periodograma

Bartlett, 1948

2.- Preceded de periodogramas

Welch, 1967

3.- Señales del periodograma: Blackman-Tukey, 1958

Análisis espectral $\xrightarrow{\text{mejor}}$ estimación espectral

$x(u)$ $\xrightarrow{\text{model}}$ s.a.e., ergfica perfecta + b autocorrelación

$x(0) \dots x(N-1)$

queremos estimar la D.E.P de ese señal, $\tilde{T}_x(f)$, a partir de una realización de duración N muestras

7.1.- Periodograma

$$\tilde{T}_x(f) \longleftrightarrow Y_x(u)$$

$$r_x(u) = \frac{1}{N} \sum_{n=-N+1}^{N-1} x(n)x(n-u) \quad (\text{estimador corregido})$$

$$r_x^1(u) = \frac{1}{N-|u|} \sum_{n=0}^{N-1} x(n) \quad (\text{estimador inscrito})$$

$$\tilde{T}_x(f) = \sum_u Y_x(u) e^{-j2\pi fu}$$

$$r_x(u) \Rightarrow P_x^*(f) = \sum_u r_x(u) e^{-j2\pi fu}$$

$\overbrace{x(0) \dots x(N-1)}$
 $\downarrow T_f$ (señal de duración finita)

$$\tilde{X}(f) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi fu} \Rightarrow |\tilde{X}(f)|^2 = \text{Densidad Espectral de Energía}$$

8.- ANÁLISIS ESPECTRAL PARAMÉTRICO

- REPRESENTACIÓN DE WOLD: señales aleatorias estacionarias en sentido amplio.

Estas señales se pueden representar (considerar otra señal que se viene media y autocorrelación) como la salida de un filtro excitado por ruido blanco. Además, ese filtro y su inverso serán considerados.

$$\begin{array}{c} w(u) \xrightarrow{\text{H}(f)} x(u) \\ \left\{ \begin{array}{l} Y_x(u) = X_x(u) \\ u_k = u_{kY} \end{array} \right. \end{array} \Rightarrow T_x(f) = T_{x^1}(f)$$

$$T_{x^1}(f) = |H(f)|^2 T_w(f) = \sigma_w^2 |H(f)|^2$$

La representación de Wold es la base del análisis paramétrico

- BASADO EN MODELOS AR:

Se establece el espectro de una señal como si fuera un modelo AR:

$$T_x(f) = \frac{\sigma_w^2}{\left| 1 + \sum_{k=1}^P a_k e^{-j2\pi f k} \right|^2} \stackrel{(1)}{=} \sum_{k=-\infty}^{\infty} Y_x(k) e^{-j2\pi f k} \stackrel{(2)}{=}$$

Nos dan los datos: $[Y_X(0) \dots Y_X(p)]$

Opciones:

1.- A partir de ellos, resolver las ecuaciones de Yule-Walker.

Por ejemplo, con Leamer-Durbin

$$\rightarrow a_k \quad k=1:p \quad y \text{ se usa} \quad ①$$

$$2.- V(u) = - \sum_{k=1}^p a_k Y_X(u-k) \quad u=p+1:\infty$$

Con esto, se usa ②

Tambien se pueden usar estos datos para hacer un análisis espectral clásico:

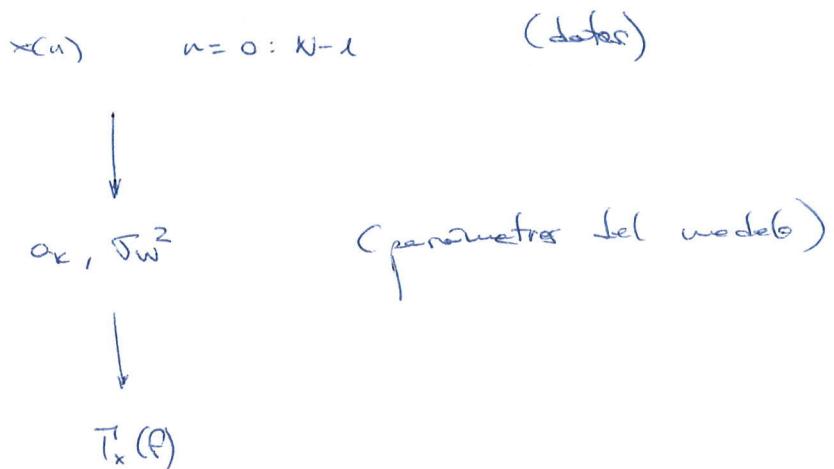
$$T_X(f) = \sum_{k=-p}^p Y_X(k) e^{-j2\pi f k}$$

Se ve la diferencia entre ambas:

- La clásica usa los datos de la autocorrelación que se dan
- La geométrica usa los datos de la autocorrelación

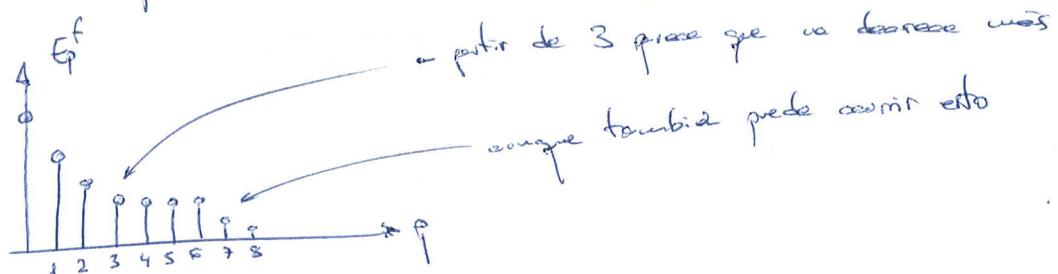
La limitación de la clásica viene por el anafaseo que hace implicitamente a los datos, lo que provoca baja resolución espectral.

Por otro lado, partir de los autovalores no es el problema real, sino partir de una secuencia de datos:



Hay que hacer una hipótesis sobre el orden, ya que no es una relación sencilla. Para ello hay algunas fórmulas, aunque ninguna funciona adecuadamente.

También es posible basarse en la evolución del error de predicción.



No se puede incrementar mucho el orden cuando se tienen pocos datos.

Ahora veremos 4 métodos para pasar de los datos a los parámetros del modelo:

- 1.- Método de Yule-Walker
- 2.- Método de los coeficientes de reflexión
- 3.- Método de la covarianza
- 4.- Método de la covarianza modificada

8.1.- MÉTODO DE YULE-WALKER (Pgs 933-934)

Consiste en resolver las ecuaciones de Yule-Walker, para lo que hay que estimar la autocorrelación en primer lugar

$$1) r_x(u) = \frac{1}{N} \sum_{k=u}^{N-1} x(k) x(k-u)$$

$$r'_x(u) = \frac{N}{N-|u|} r_x(u)$$

- 2) Resolver las ecuaciones de Yule-Walker, por ejemplo
con el algoritmo de Levinson-Durbin

Si se usa la estimación, la matriz de autocorrelación
que se obtiene es semidefinida positiva. Es recomendable
usar esta estimación para diagonalizarla. Así, el modelo AR es
estable, aunque no es crítico en análisis espectral.

8.2.- MÉTODO DE LOS COEFICIENTES DE REFLEXIÓN

(Pgs. 934-937)

Se basa en la fórmula de los coeficientes PARCOR

$$k_{uu} = - \frac{E[f_{uu}(u) g_{uu}(u-1)]}{\sqrt{E[f_{uu}^2(u)] E[g_{uu}^2(u-1)]}} = - \frac{c_{u-1}}{\sqrt{f_{uu} \cdot b}}$$

Lo : Debe obtener los coeficientes de forma iterativa:

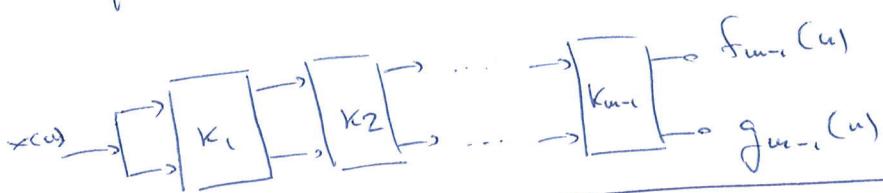
$$V_1, V_2, \dots$$

$$f_{uu}(u) = x(u) - \hat{x}(u) \quad g_{uu}(u) = x(u-u) - \hat{x}^*(u-u)$$

$$\hat{x}(u) = - \sum_{k=1}^m a_{uu}(k) x(u-k)$$

$$\hat{x}^*(u-u) = - \sum_{k=1}^m a_{uu}^*(k) x(u+k-u)$$

Se hace metiendo la secuencia de datos al filtro de errores de predicción óptima con estructura en celulas:



$$K_{n-1} = \sqrt{\frac{^bf}{^bE_{n-1}}}$$

Análisis verificar la
recursividad de la ecuación -
Dibujar

$$\hat{C}_{n-1} = \sum_{u=n-1}^{N-1} f_{nu}(u) g_{n-1}(u-1)$$

$$\hat{E}_{n-1} = \sum_{u=n-1}^{N-1} f_{nu}^2(u)$$

$$\hat{b}_E = \sum_{u=n-1}^{N-1} g_{n-1}^2(u-1)$$

Sustituimos los generadores esperan que promuevan temporales si
normalizar, y que se cancelen

Los índices empiezan en n-1 para eliminar el tránsito del
filtro y que las células de reflejo se inicializan a cero

Hay 2 variantes de este método:

Método geométrico: usa la expresión que hemos visto

$$K_{n-1} = \sqrt{\frac{\hat{C}_{n-1}}{\hat{f}_{n-1} \hat{b}_{E_{n-1}}}}$$

es una meth geométrica

$$\hat{E}_n = \sqrt{\hat{f}_{n-1} \hat{b}_{E_{n-1}}}$$

Método de Burg: usa la media aritmética

$$\hat{E}_n = \frac{1}{2} (\hat{f}_{n-1} + \hat{b}_{E_{n-1}})$$

$$K_{n-1} = \frac{\hat{C}_{n-1}}{\frac{\hat{f}_{n-1} + \hat{b}_{E_{n-1}}}{2}}$$

$$T_x(f) = \frac{\hat{E}_n}{\left| 1 + \sum_{k=1}^n \alpha_k k^{-1} e^{j2\pi f k} \right|^2}$$

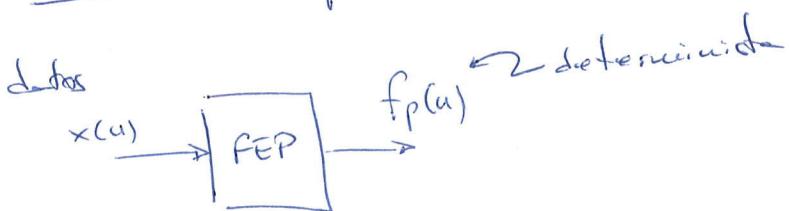
8.3.- MÉTODO DE LA COVARIANZA

Pgs. 937-938

Utiliza la técnica de mínimos cuadrados (LS = Least Square).

Es un método directo, que consiste en, usando un filtro de error de predicción, obtener la señal de error de predicción, pero no introduciendo una señal aleatoria, sino los datos (deterministas).

- Cómo resolver el predictor lineal con mínimos cuadrados:



Hay que minimizar la energía de $f_p(u)$

$$f_p(u) = \sum_{k=0}^p \alpha_p(k) x(u-k)$$

$\Rightarrow \alpha_p(k)$ tales que $E_p = \sum_{u=u_I}^{u_F} f_p^2(u)$ sea mínima

no se impone la restricción
de Levens-Durbin

intervalo $[u_I, u_F]$ genérico

Varios casos según cuales sean u_I, u_F :

- a) $u_I = 0$ { se calcula la energía de $f_p(u)$ en todos los instantes en que $f_p(u) \neq 0$
- $u_F = N+p-2$

\Rightarrow Método de la autocorrelación: es igual que el método de

Yule-Walker con estímulos $(x(u))$

b) $u_I = P$ } se elimina el tránsito de correg del filtro
 $u_F = N-1$ } y tambien el de finalizar

\Rightarrow Método de la covarianza

Desde el punto de vista de las prestaciones, es mejor el de la covarianza, ya que elimina los tránsitos.

- Método de la covarianza:

Consiste en resolver el problema de predicción lineal (error hacia adelante) por métodos análogos al método de la covarianza

covarianza

8.4.-MÉTODO DE LA COVARIANZA MODIFICADA

Se puede resolver con el filtro de error de predicción hacia delante. La covarianza modificada combina ambas formas.

$$\epsilon_p^f = \epsilon_p^f + \epsilon_p^b \quad \text{objetivo: minimizar } \epsilon_p^f \\ (\text{forward}) \quad (\text{backward}) \quad \left(\epsilon_p^f + \epsilon_p^b \leq u_I = P, u_F = N-1 \right)$$

Tiene algunas mejoras respecto a la covarianza.

Todos los métodos visto tienen implementación en Matlab.

$$T_x(f) = \frac{\epsilon_p}{\left| 1 + \sum_{k=1}^P a_p(k) e^{-j2\pi f k} \right|^2}$$

RESOLUCIÓN DE PROBLEMAS POR MÉTODOS CUADRADOS:

$$\left. \begin{array}{l} \text{MÉSE} \rightarrow \min E\{f_p^2(u)\} \quad (\text{estadístico}) \\ \text{LS} \rightarrow \min \sum_{u=u_1}^{u_F} f_p^2(u) \quad (\text{determinista}) \end{array} \right\}$$

MÉSE:

$$T_p \alpha = -Y_p$$

LS:

$$R_p \cdot \alpha = -r_p$$

$$R_p = \begin{bmatrix} r_p(1,1) & \dots & r_p(1,p) \\ \vdots & \ddots & \vdots \\ r_p(p,1) & \dots & r_p(p,p) \end{bmatrix}$$

$$r_p = \begin{bmatrix} r_p(0,1) \\ \vdots \\ r_p(0,p) \end{bmatrix}$$

error de predicción
baja debilidad

$$r_p(i,j) = \sum_{u=u_1}^{u_F} x(u-i) \times (u-j)$$

R_p:

$$- r_p(i,j) = r_p(j,i)$$

- no es, en general, una matriz de Toeplitz

- no es, en general, una matriz de Levinson-Durbin
 \Rightarrow no se puede usar Levinson-Durbin

R_p para el caso en que el error de predicción sea baja:

$$r_p(i,j) = \sum_{u=u_1}^{u_F} x(u-p+i) \times (u-j) \quad (\text{covarianza})$$

Covarianza modificada \Rightarrow $r_p(i,j) = \sum_{u=u_1}^{u_F} (x(u-i) \times (u-j) + x(u-p+i) \times (u-p+j))$

(*) $u_1 = p, u_F = N-1$

Para el método de los cuadrados, habrá que usar la suma de los dos errores al resolver el sistema:

$$\underline{R_p} = \underline{R_p^f} + \underline{R_p^b}$$

$$r_p = r_p^f + r_p^b$$

$$\sum_{k=1}^P \alpha_p(k) r_p(i, k) = -r_p^{(0), i} \quad i = f, p$$

MSE

$$\hat{\epsilon}_p^f = \gamma(q) + \underline{Y_p^t} \cdot \alpha_p$$

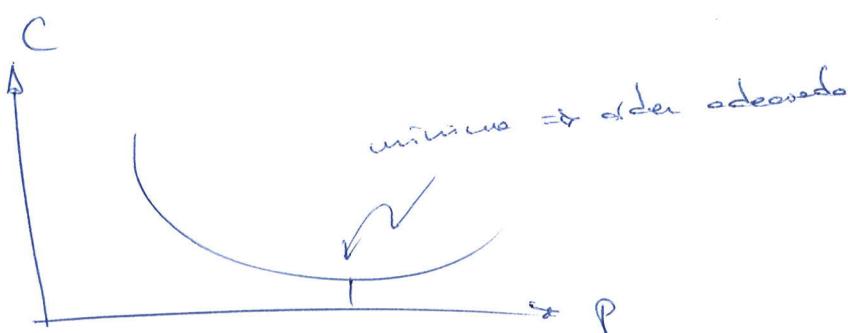
LS

$$\begin{aligned} \min \hat{\epsilon}_p^f &= \hat{\epsilon}_p^f \\ \hat{\epsilon}_p^f &= r_p(0, 0) + \underline{Y_p^t} \cdot \alpha_p \\ &= r_p(0, 0) + \sum_{k=1}^P \alpha_p(k) r_p(0, k) \end{aligned}$$

8.5.- ELECCIÓN DEL ORDEN DEL MODELO

Se trata de proponer funciones de coste que sean función de los datos de los que se dispone:

$$C = f(N, \hat{\epsilon}_p^f, q) \quad \left\{ \begin{array}{l} N \text{ fijo} \\ \hat{\epsilon}_p^f \text{ fija una vez elegido} \\ \text{el método} \end{array} \right.$$



Distintos autores han propuesto varias fórmulas:

1.- Final prediction error: $FPE(\rho) = \hat{E}_p^f \frac{N + \rho + 1}{N - \rho - 1}$

2.- Akaike information criterion: $AIC(\rho) = (\ln \hat{E}_p^f + 2 \frac{p}{N})$

3.- Minimum description length: $MDL(\rho) = N \ln \hat{E}_p^f + p \ln N$

4.- Criterio autoregresivo transformador: $CAT(\rho) = \frac{1}{N^2} \sum_{k=1}^p \frac{N-k}{\hat{E}_p^f} - \frac{1}{\hat{E}_p^f}$

Ejercicios: Practicando

Cap 11: 1-7, 9, 19

Cap 12: 13, 14, 17, 19, 21, 23-28, 31, 34, 36, 38

8.6.- APLICACION DEL ALGORITMO DE LEUNSON-PURBIN

Nos quedan las las autocorrelaciones a los datos

$$\gamma(0) = 3 \quad \gamma(1) = 0.9 \quad \gamma(2) = -0.8 \quad \gamma(3) = -0.7$$

obtener el predictor óptimo de orden 3.

Solución:

$$1.- E_0^f = \gamma(0) = 3$$

$$2.- m=1: \quad a_1(0) = 1$$

$$a_1(1) = k_1 = - \frac{a_0(0) \gamma(1)}{E_0^f} = - \frac{\gamma(1)}{\gamma(0)} = -0,3$$

$$E_1^f = E_0^f (1 - k_1^2) = 2,73$$

$$m=2: \quad a_2(0) = 1$$

$$a_2(2) = k_2 = - \frac{a_1(1) \gamma(1) + a_0(0) \gamma(2)}{E_1^f} = 0,3 a_1(1)$$

$$E_2^f = E_1^f (1 - k_2^2) = 2,31$$

$$a_2(1) = -0,4176$$

$$m=3: \quad k_3 = -0,6002$$

$$E_3^f = 1,4783$$

$$|k_i| < 1$$

$$\left. \begin{array}{l} \gamma(0) = 1 \\ \gamma(1) = 0,95 \\ \gamma(2) = 0,9 \\ \gamma(3) = 0,6 \end{array} \right\} \Rightarrow \begin{array}{l} k_1 = -0,95, \quad E_1^f = 0,0975 \\ k_2 = 0,0256, \quad E_2^f = 0,9974 \\ k_3 = 2,6, \quad E_3^f > E_2^f ! \end{array}$$

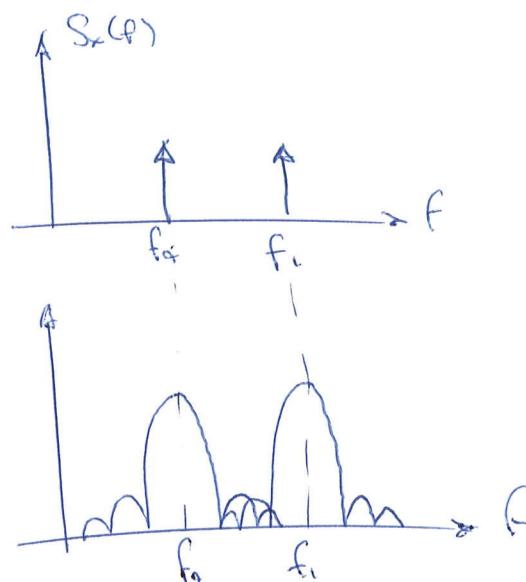
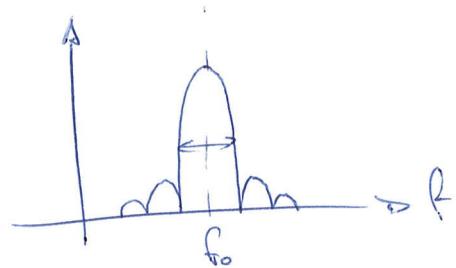
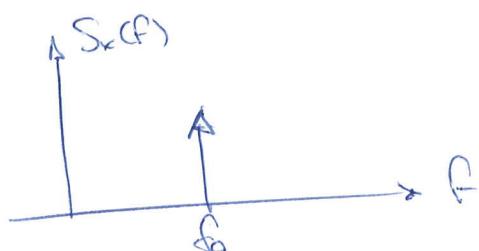
⇒ los valores de la matriz de autocorrelación no pueden ser una secuencia de autocorrelación (al crecer con el 4°)

⇒ los datos no son ordenados.

9.- CARACTERIZACIÓN DE LOS ESTIMADORES ESTUDIADOS

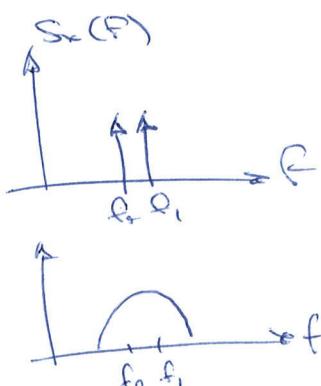
Caracterizar los estimadores no es sencillo, ya que depende de la señal.

- ANÁLISIS ESPECTRAL CLÁSICO: $x(u) \quad u = 0 : N-1$



No se verán los rayos espectrales, sino los efectos de los vértices desplazados

Si tenemos focos demasiado juntos: $(*)1.3*$



→ no se puede distinguir los focos

$(*)1.2*$ → inconsistencia del período gama
Si la varía la velocidad del rizado, se levará la varianza

(*1.4*) \rightarrow Bottlett

(*1.5*) \rightarrow efecto de los bloques laterales. Puedes suficiente
resolver para separar las 2 fases, pero los bloques
secundarios enmascaran el efecto de veces amplificado

Solución: método de Welch \rightarrow útil para señales
con grandes cambios espectrales

- ANÁLISIS ESPECTRAL PARÁMÉTRICO:

(*1.6*)

(*1.8*) \rightarrow representación da que, sintéticamente, todos
los modelos se equivalentes

(*1.10*) \rightarrow AR(9), 3 fases

) SNR A \Rightarrow buena resultado

) SNR B \Rightarrow se enmascaran las fases

(*1.11, 1.12*) \rightarrow comparativa de las técnicas de estimación
Yule-Walker incluye los tránsitos, por eso tiene mejor
resultado.

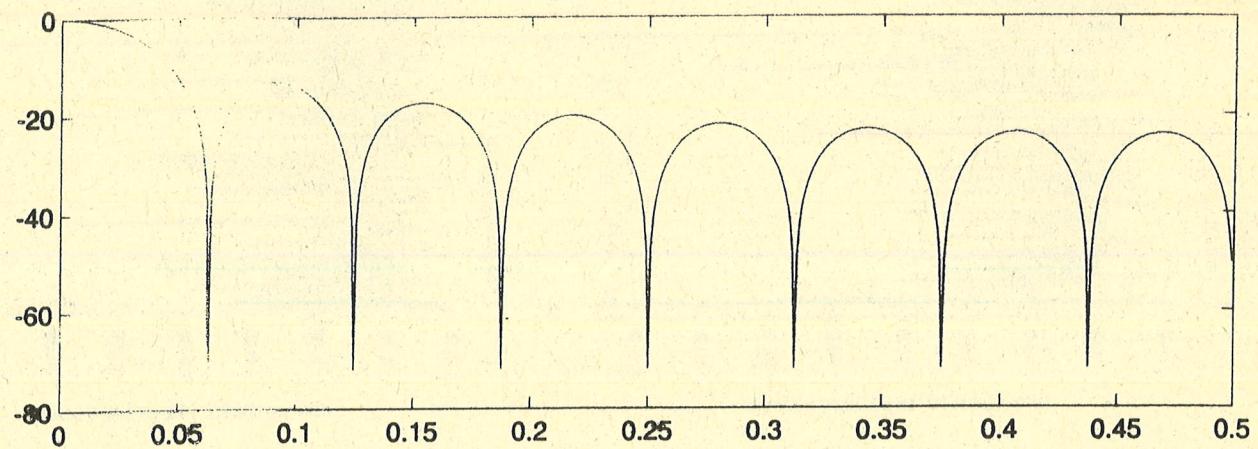
Picos + valles por el aumento del orden para mejor resultado

\Rightarrow 1065 y cercas

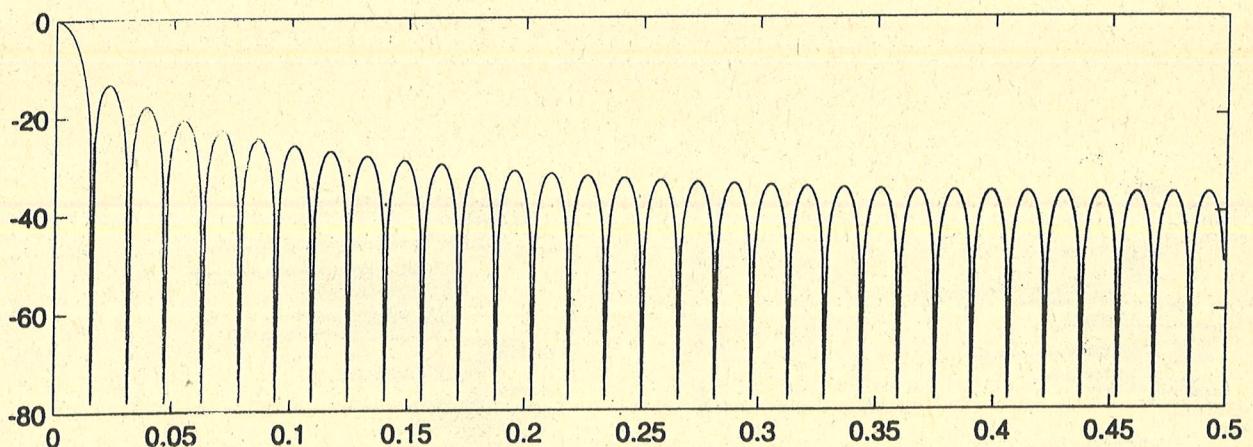
- Características del A.E.P. basado en modelo AR:

- Apariencia de picos espíneos al aumentar el orden
→ relación entre picos y resolución
- Para el método de Burg, aparece un sesgo en la estimación de la frecuencia de tocas con ruido añadido,
y depende de la fase inicial del toque
(* 1.13*, fig 8.4)
- Desdoblamiento de picos espectrales (Burg): (* 1.13*, fig 8.3)

T. F. VENTANA TRIANGULAR
 $(2N-1)$



$N = 16$



$N = 64$

INCONSISTENCIA DEL PERIODOGRAFO

O-2

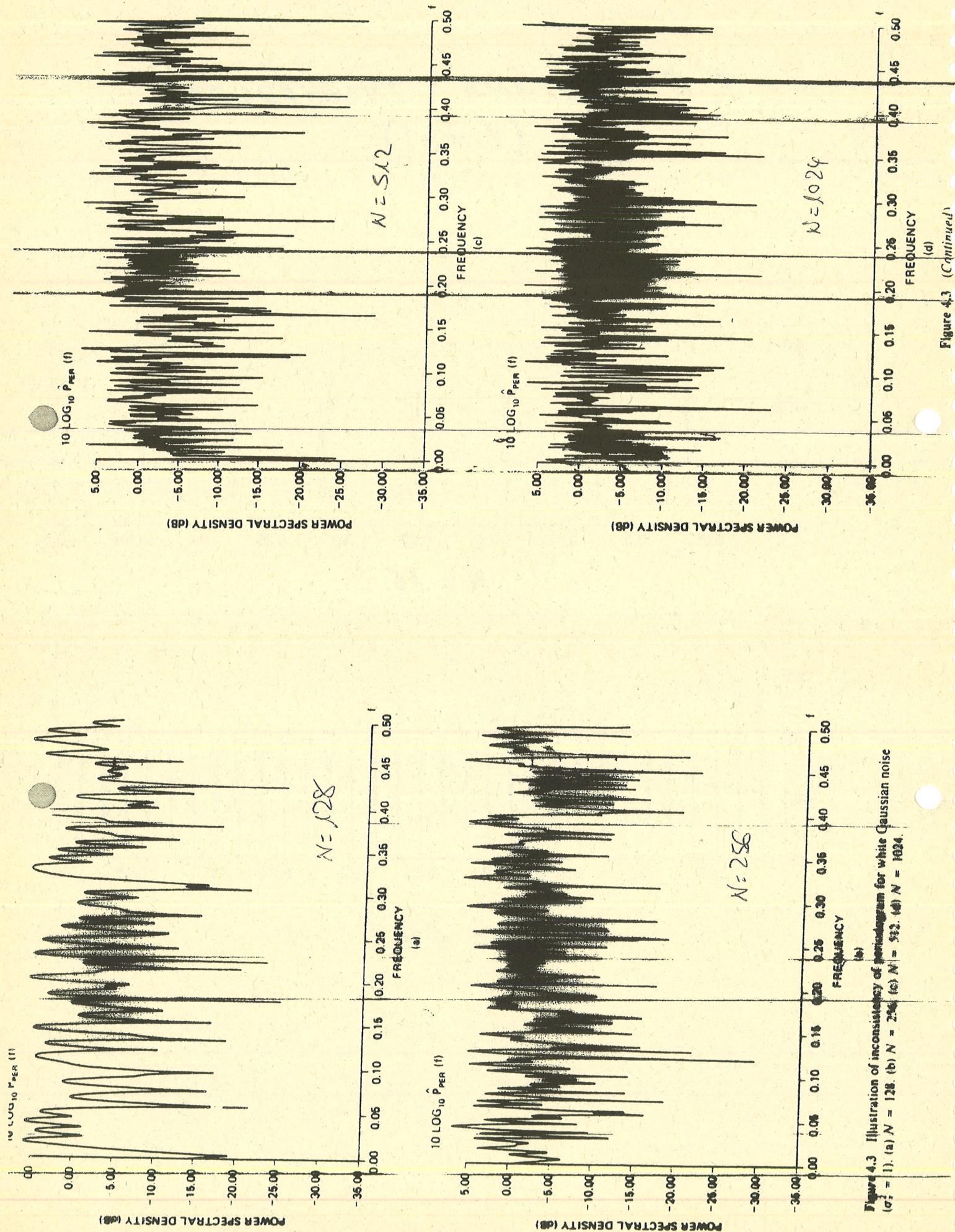
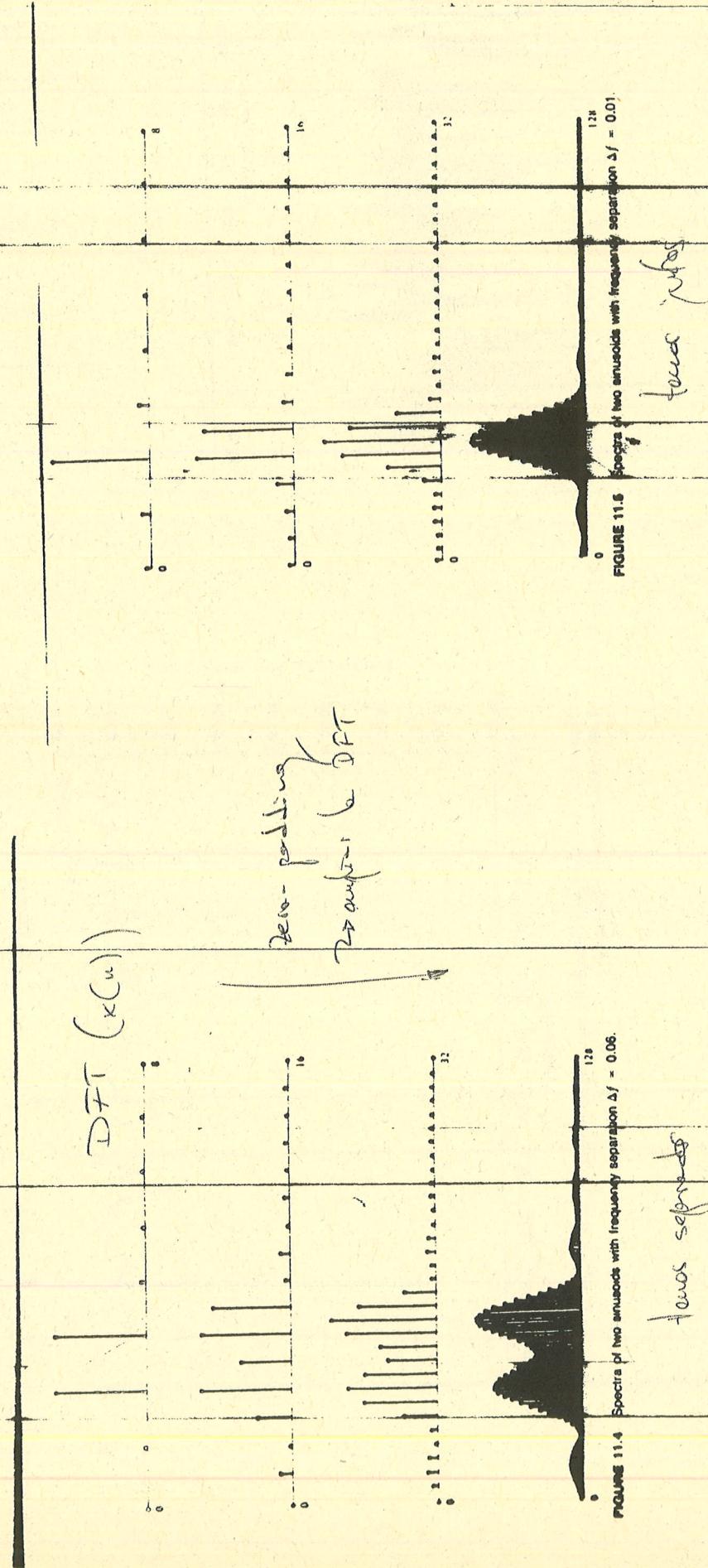


Figure 4.3 Illustration of inconsistency of periodogram for white Gaussian noise
($\sigma^2 = 1$). (a) $N = 128$. (b) $N = 256$. (c) $N = 512$. (d) $N = 1024$.

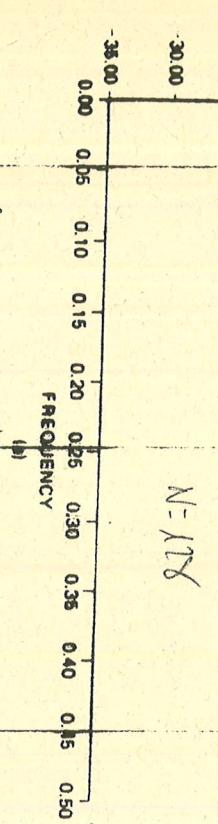
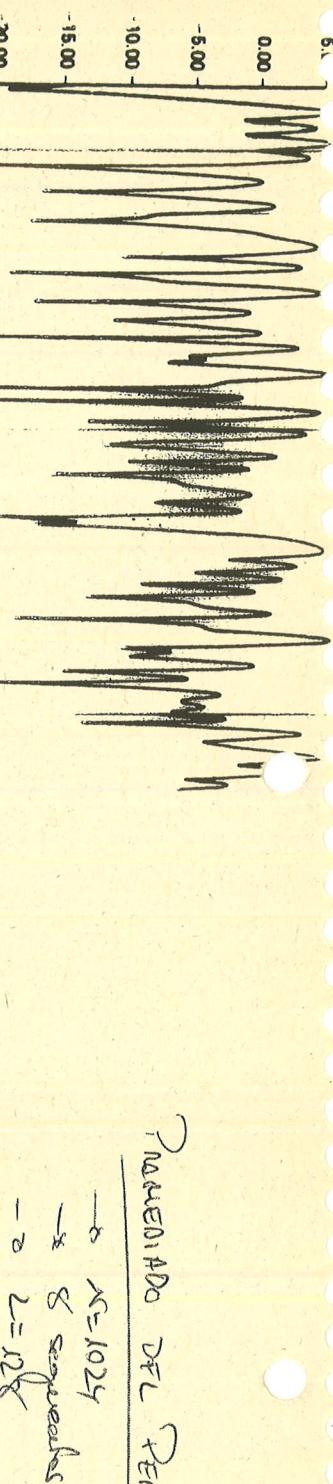
Figure 4.3 (Continued)

FIGURE 11.4 Spectra of two sinusoids with frequency separation $\Delta f = 0.06$.FIGURE 11.5 Spectra of two sinusoids with frequency separation $\Delta f = 0.01$.

(Rockis-Henselius pp. 912, 918)

four "nbs"

NAMEADA DFL PERIODGRAM (BARTLETT)



POWER SPECTRAL DENSITY (dB)

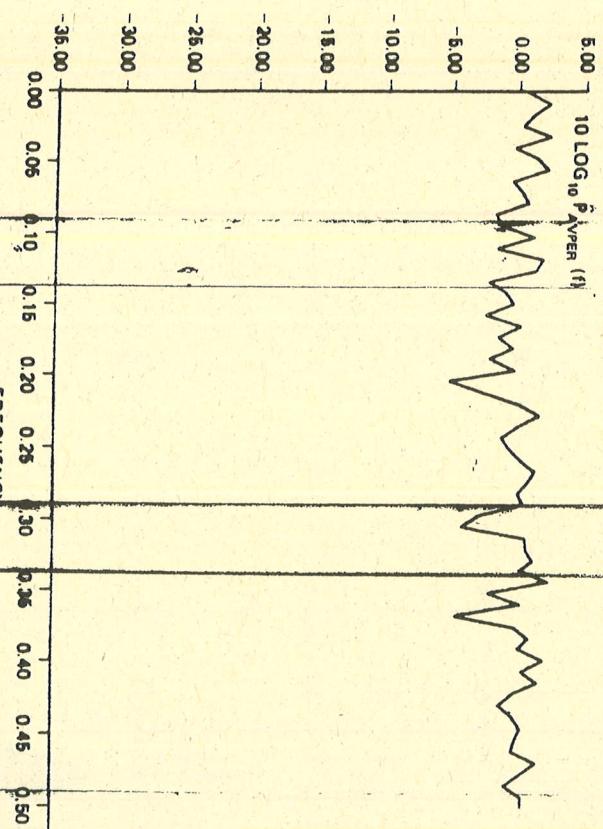


Figure 4.5 Averaged Periodogram for white noise with $\sigma^2 = 1$, $L = 128$, $K = 8$ illustrating reduction in variance.

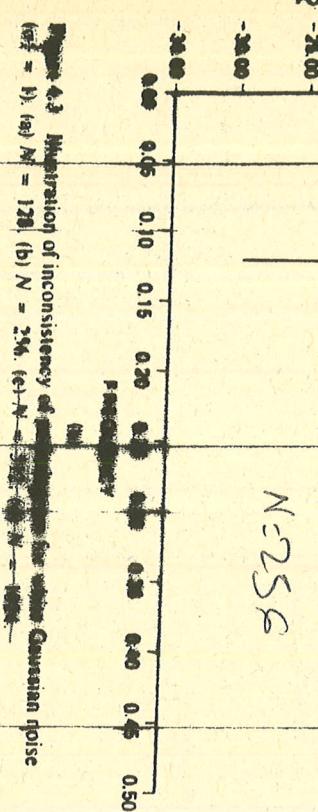


Figure 4.3 Illustration of inconsistency of periodograms for white Gaussian noise
(a) $N = 16$, (b) $N = 128$, (c) $N = 256$.

2 TONOS CON DIFERENTE AMPLITUD

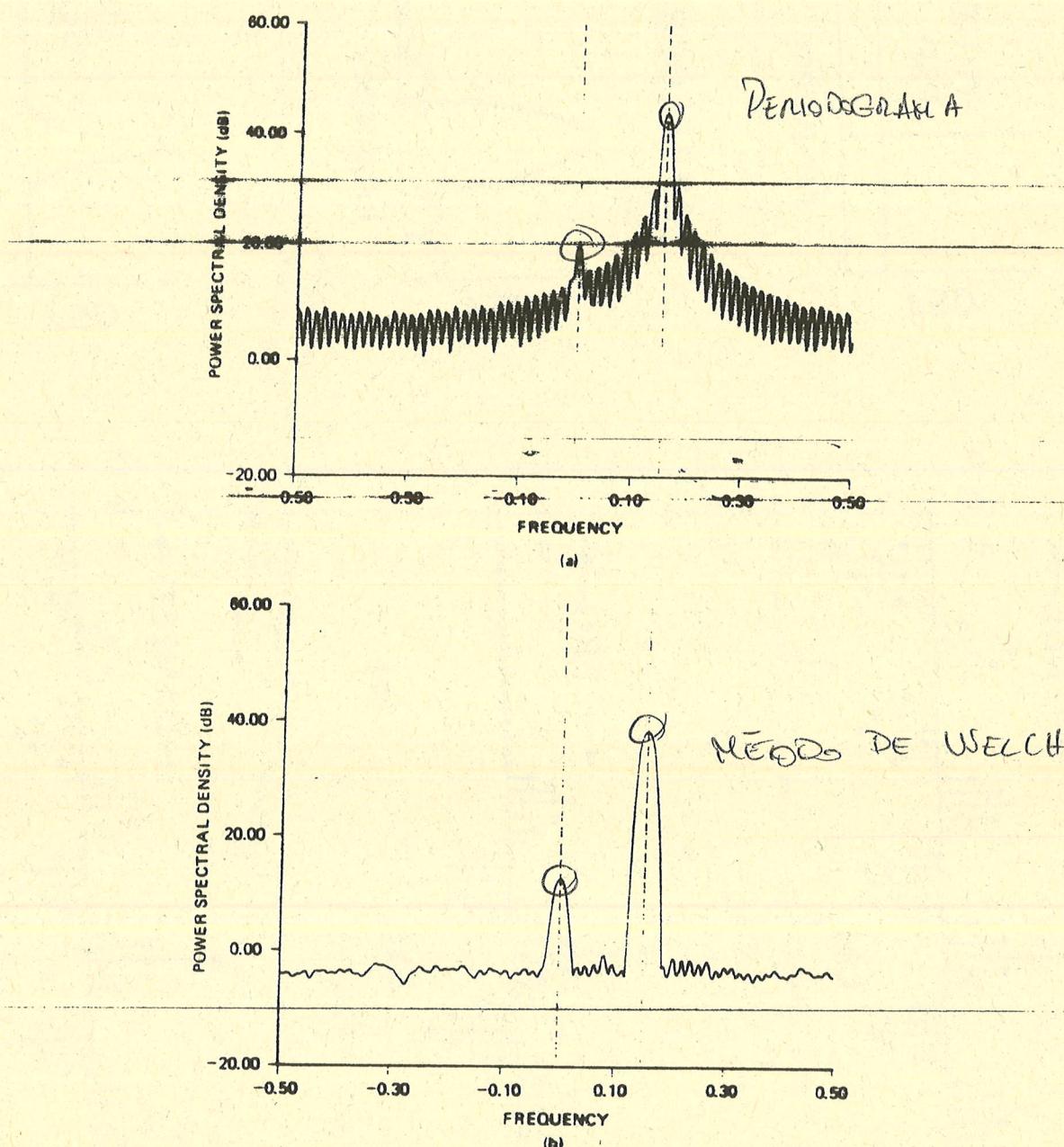
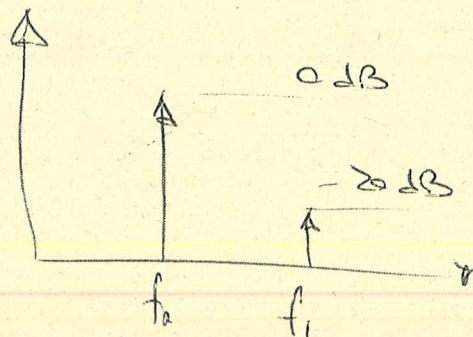


Figure 4.4 Use of data windowing to reduce sidelobes in periodogram of narrowband processes. (a) No data window. (b) Hamming data window.



Suficiente Δf , pero media diferencia de amplitud \Rightarrow efecto de lobas extendidas

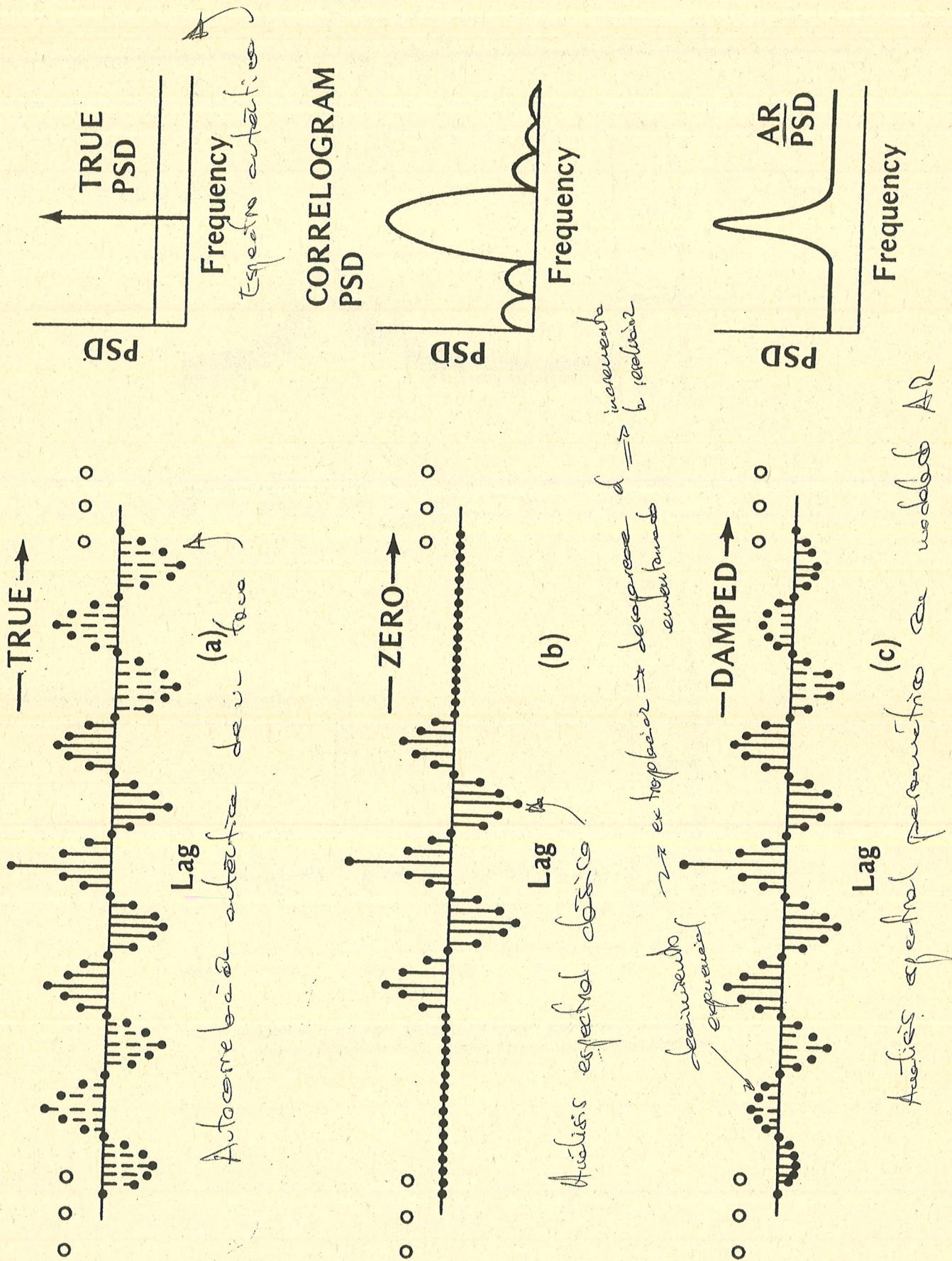


Figure 7.4 Implied extrapolation of the autocorrelation sequence (ACS) (a)

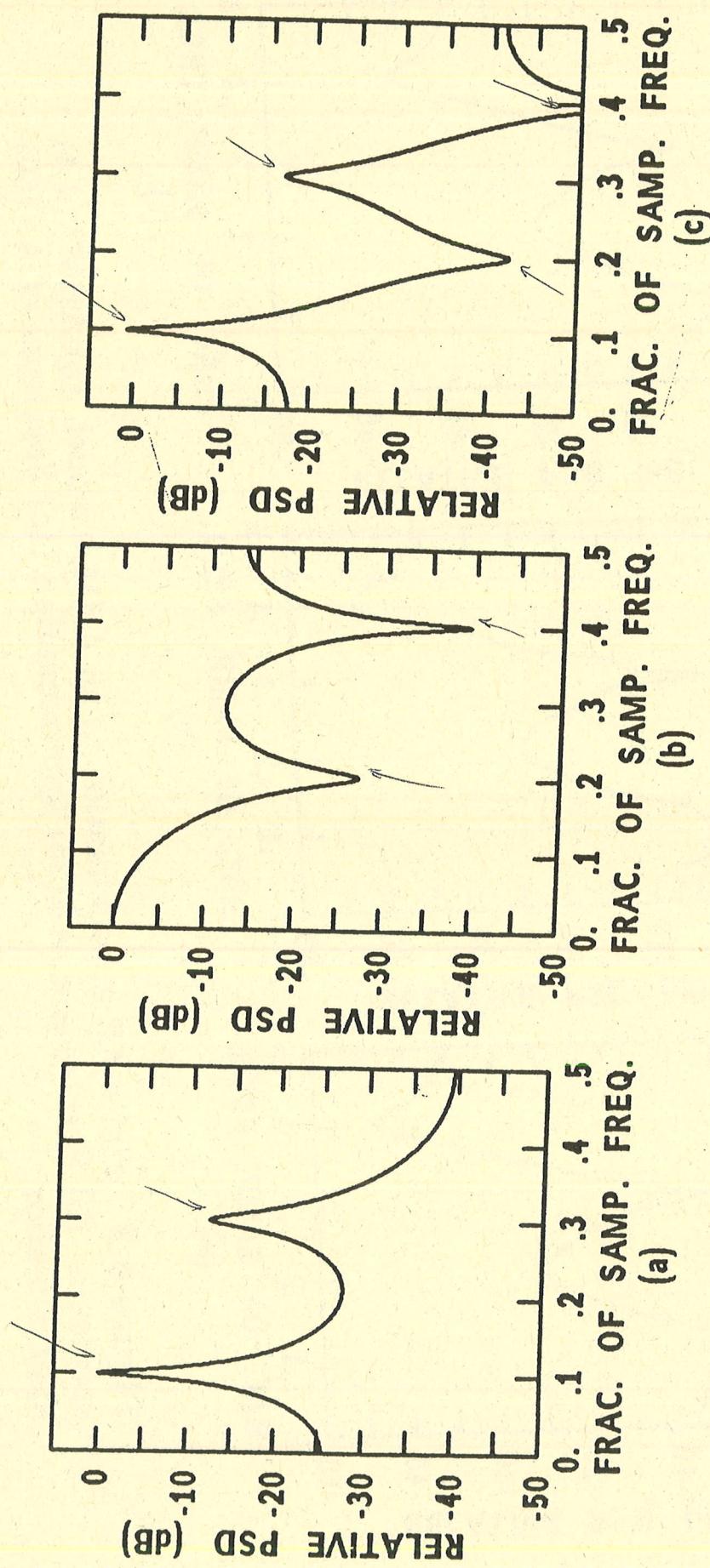


Figure 6.2 Typical parametric model spectra. (a) AR(4) spectrum. (b) MA(4) spectrum. (c) ARMA(4,4) spectrum.

$A_R(Y) \rightarrow 2 \cos(\omega_c)$
 $\text{MA}(4) \rightarrow 2 \text{ zeros} (\omega_c)$
 $\text{ARMA}(4,4) \rightarrow \text{doublet poles}$ (poles + zeros)

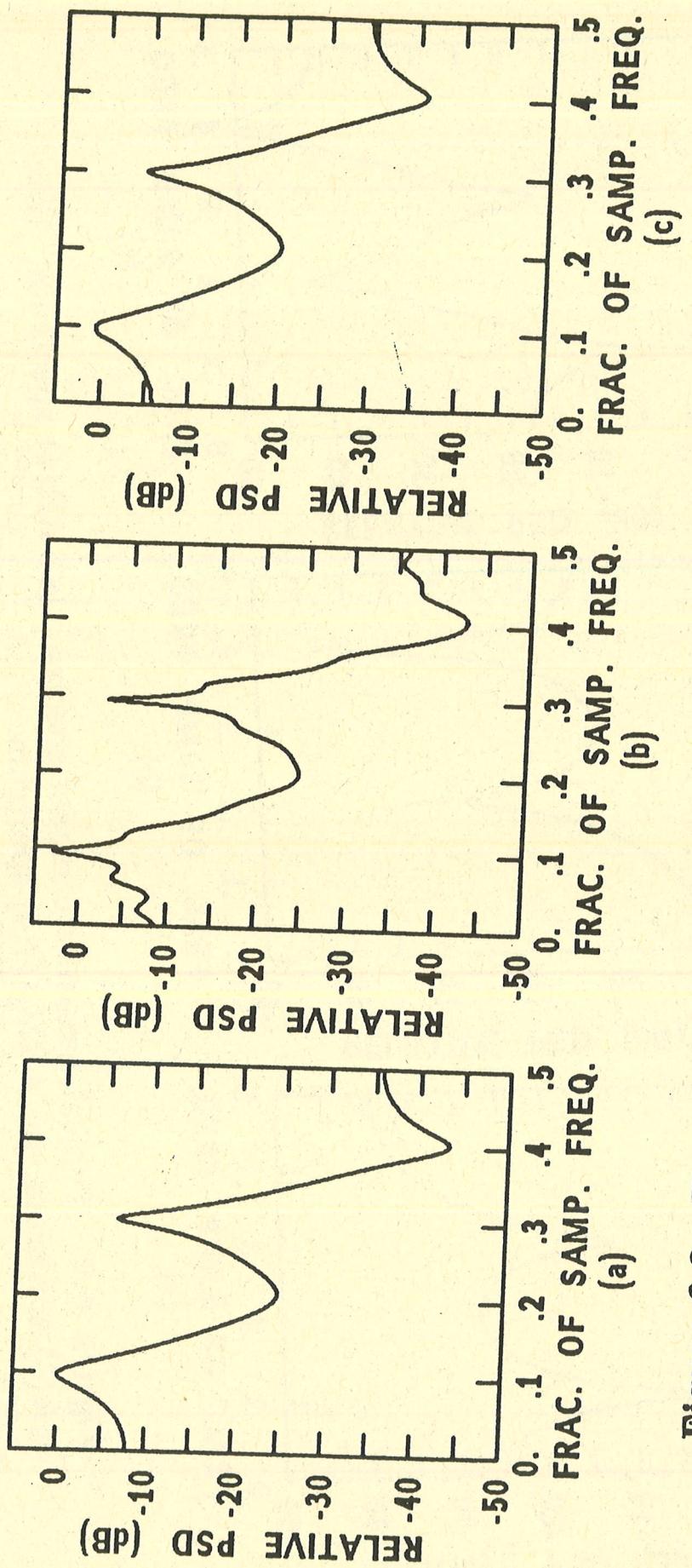


Figure 6.3 Finite AR approximation to an ARMA model. (a) ARMA(4,4) spectrum. (b) Spectrum of AR(25) approximation. (c) Spectrum of AR(50) approximation.

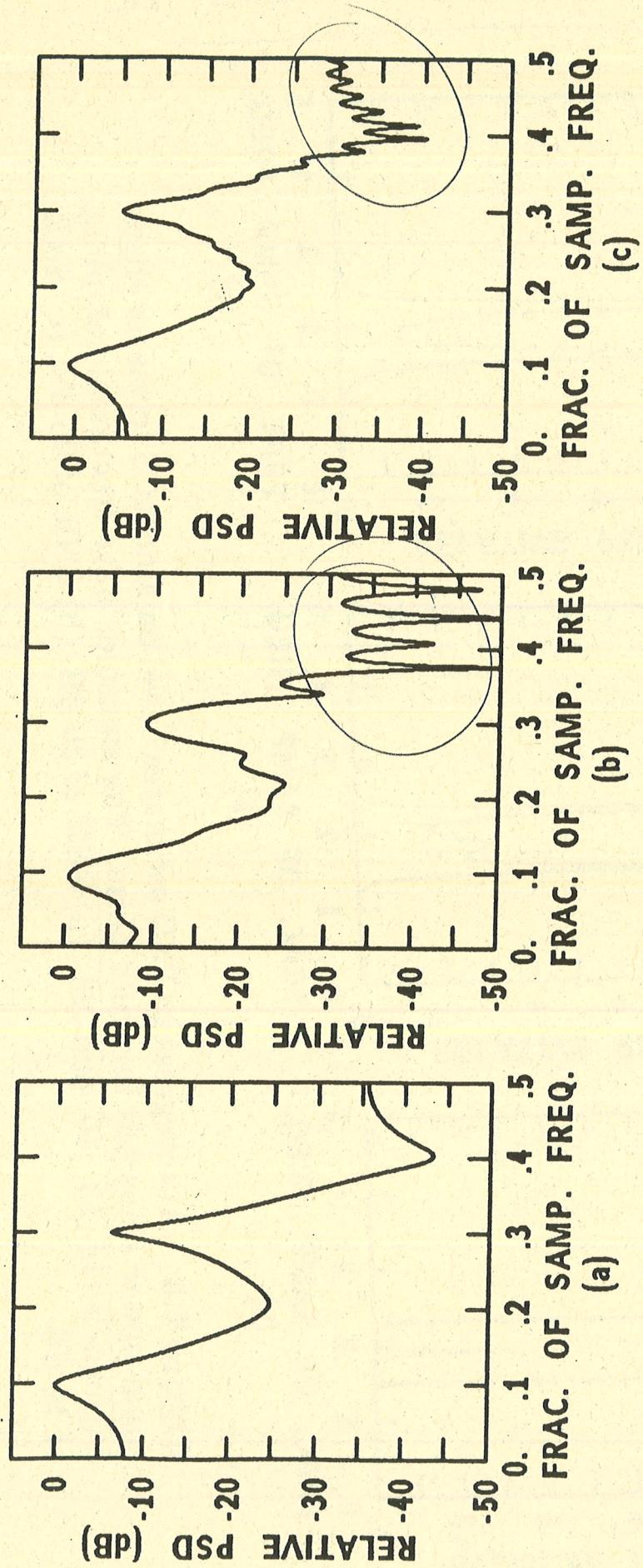


Figure 6.4 Finite MA approximation to an ARMA model. (a) $\text{ARMA}(4,4)$ spectrum. (b) Spectrum of $\underline{\text{MA}(25)}$ approximation. (c) Spectrum of $\underline{\text{MA}(50)}$ approximation.

MA → envejecimiento de la autocorrelación ⇒ ríos

DEFICIENCIAS DEL MODELO ADO

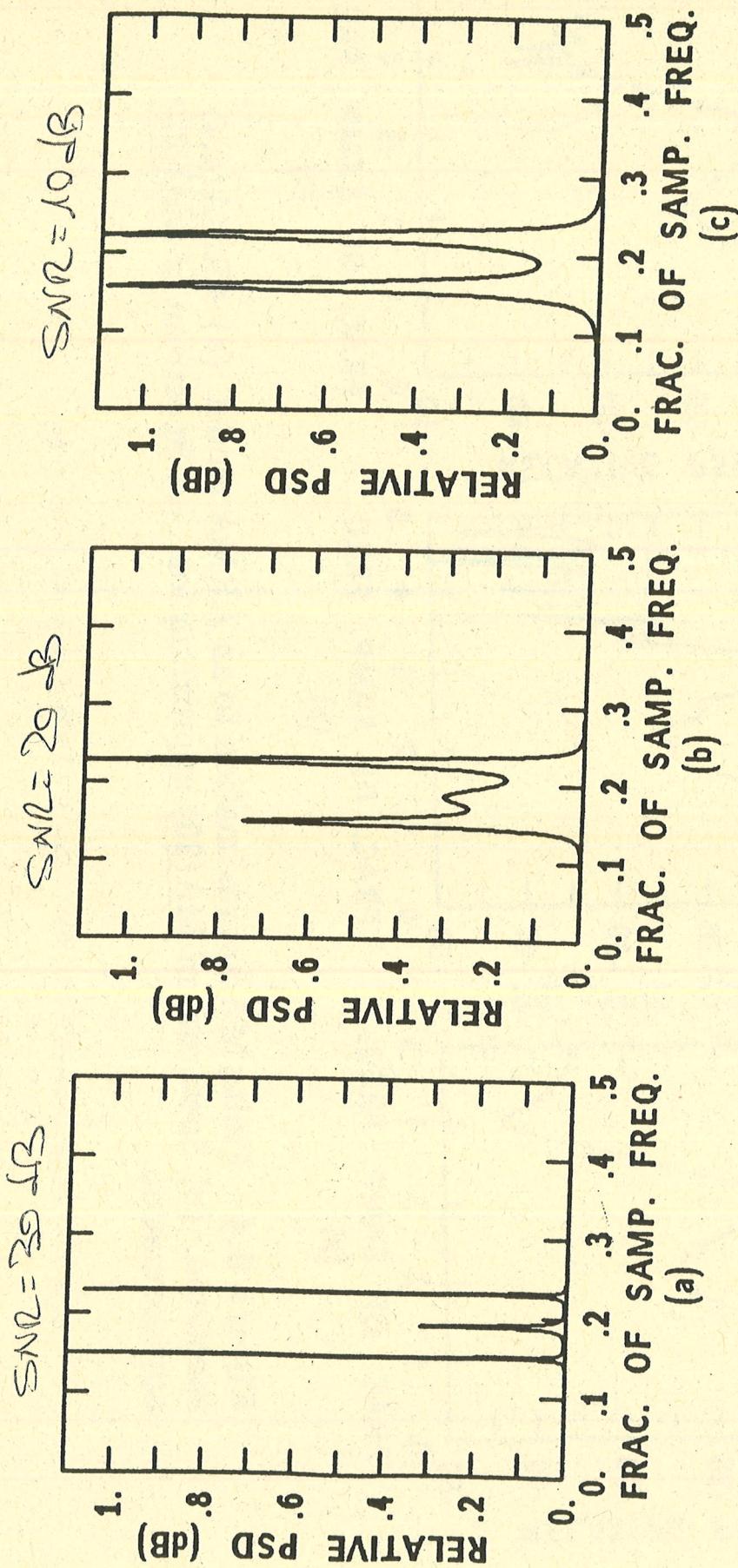


Figure 7.5 Autoregressive PSD estimates based on nine ACS terms for three equal-power sinusoids in white noise. The sinusoids have fractional sampling frequencies of .15, .19, and .23. (a) SNR of 30 dB. (b) SNR of 20 dB. (c) SNR of 10 dB.

$\mathcal{AR}(9)$

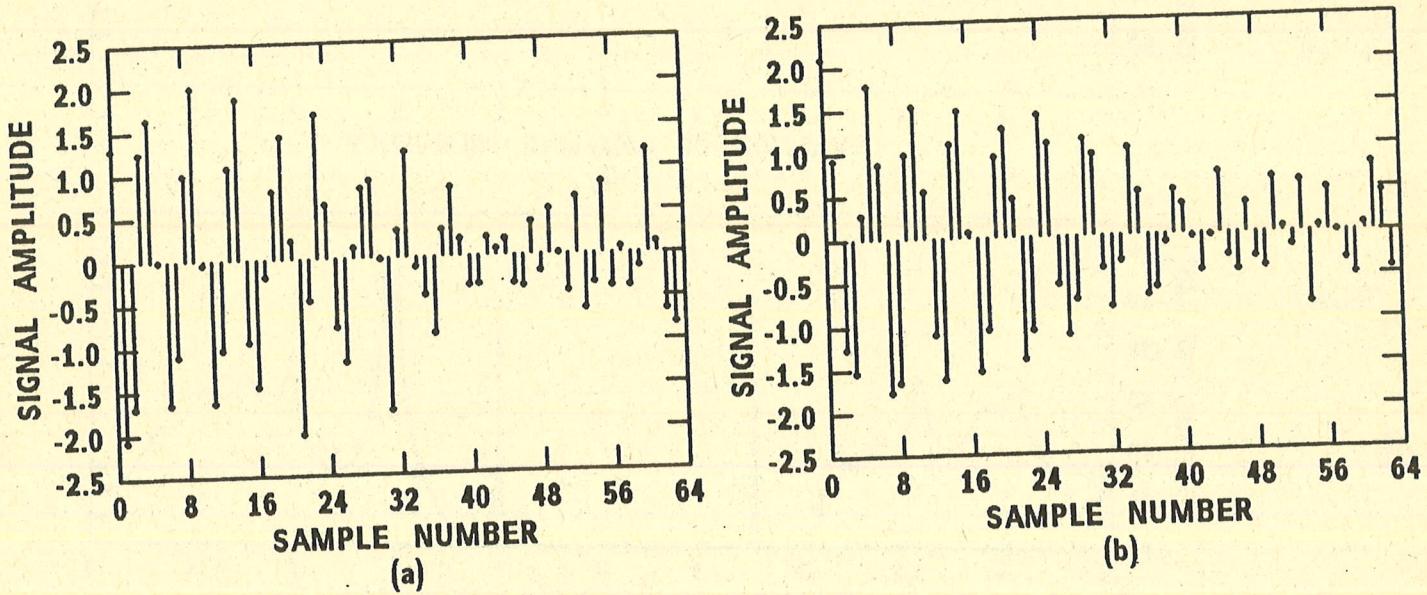


Figure 1.7 Complex test data sequence. (a) Real part. (b) Imaginary part.

Senal de muestra

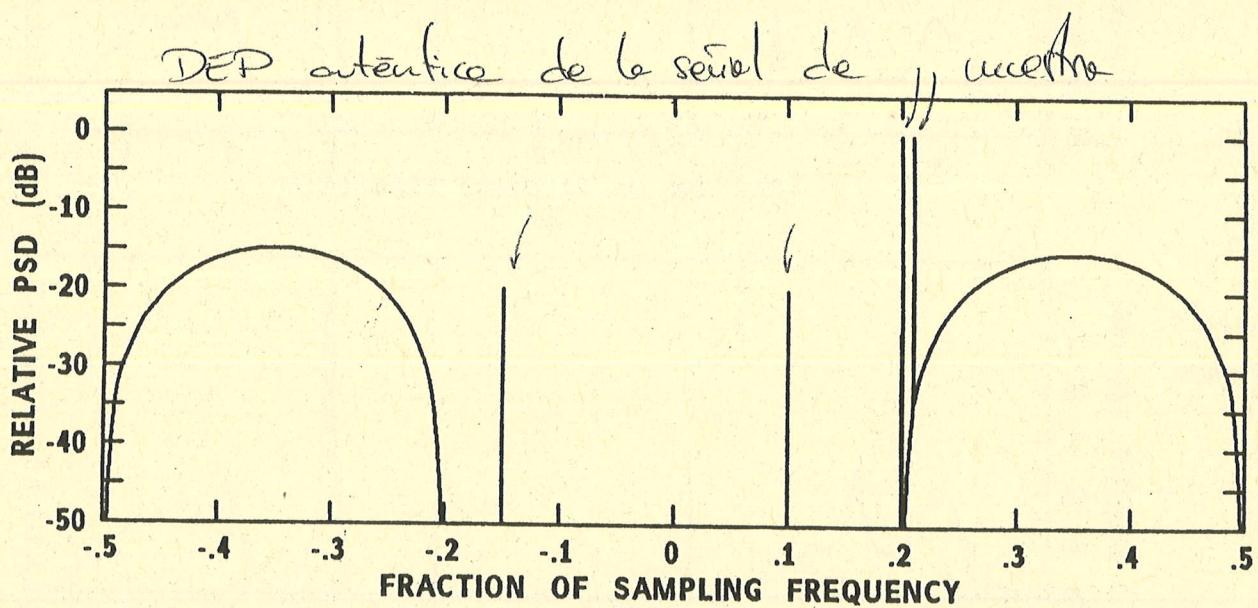


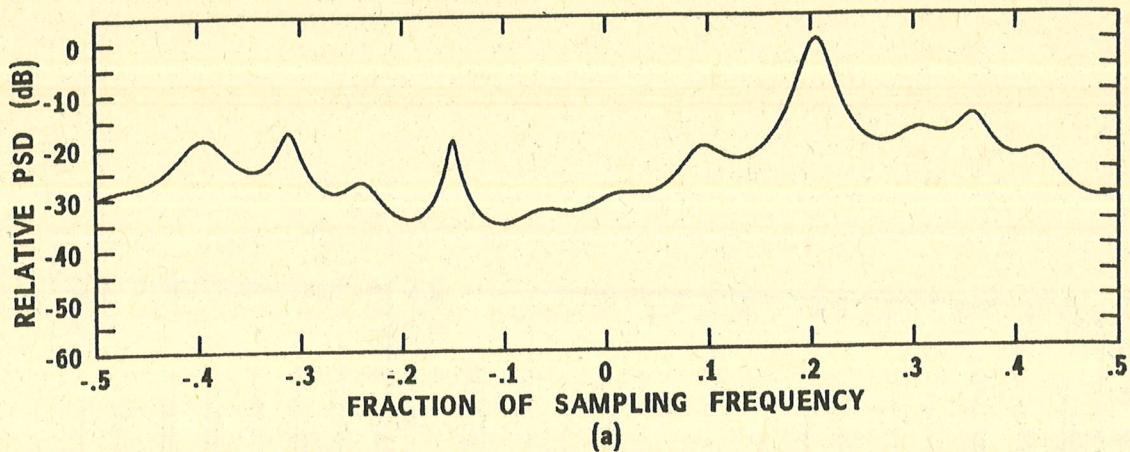
Figure 1.8 Analytically determined spectrum of test data process.

) = falso

AR-15

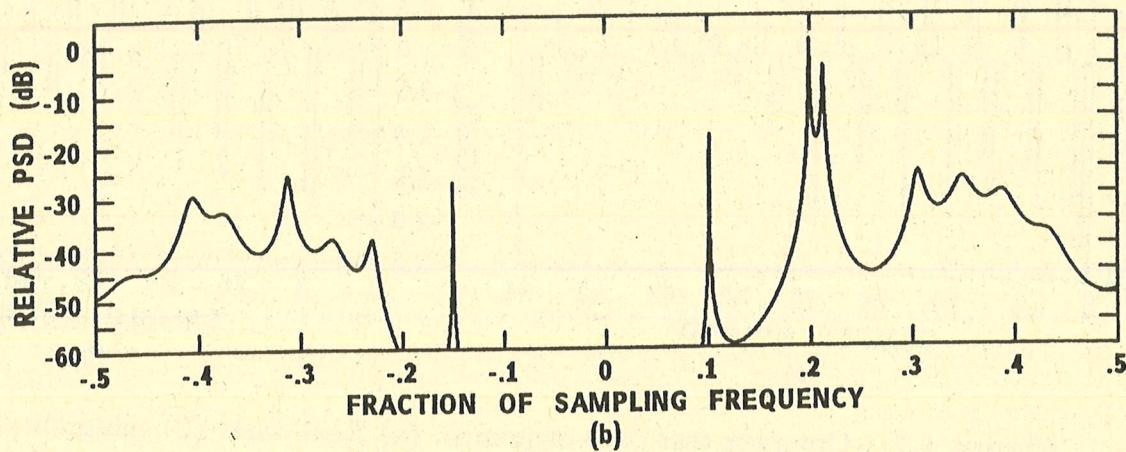
1-2

Yule-Walker



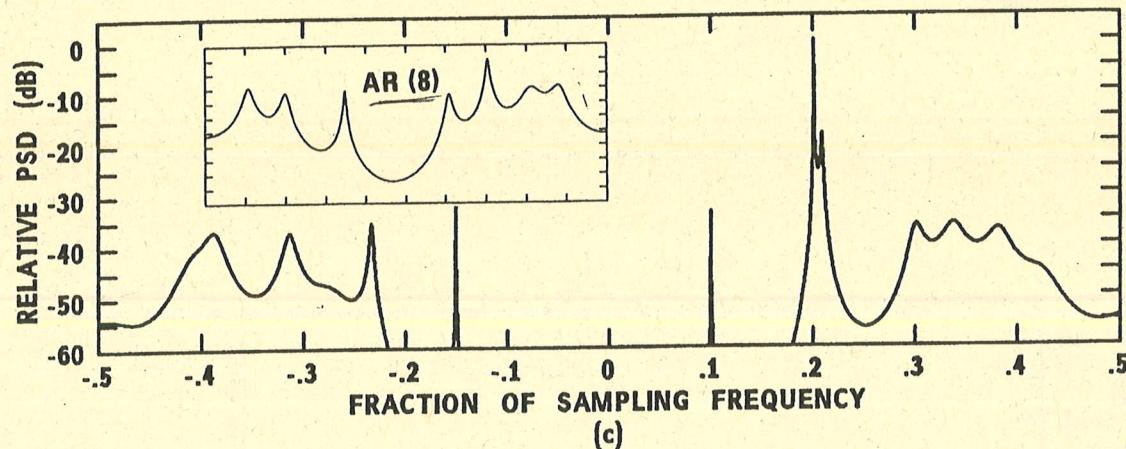
(a)

Burg

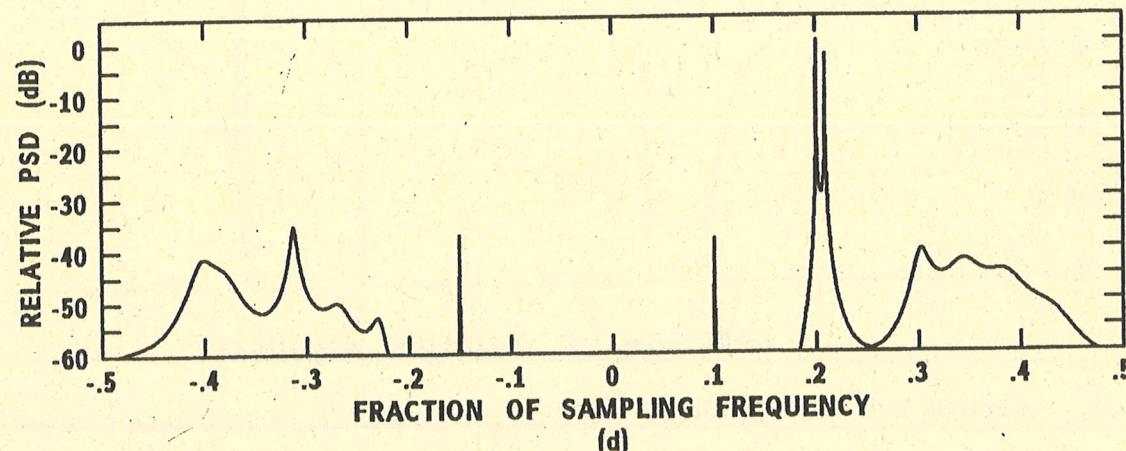


(b)

Covarianza



(c)

Covarianza
Modificada

(d)

AR(8) = 0 resolved

Figure 8.2 Examples of four AR(15) spectral estimates made from the 64-point test data sequence. (a) Yule-Walker method. (b) Burg method. (c) Covariance method. (d) Modified covariance method.

DESDOBRAMIENTO DE PICOS ESPECTRALES (BURG)

AR(25)

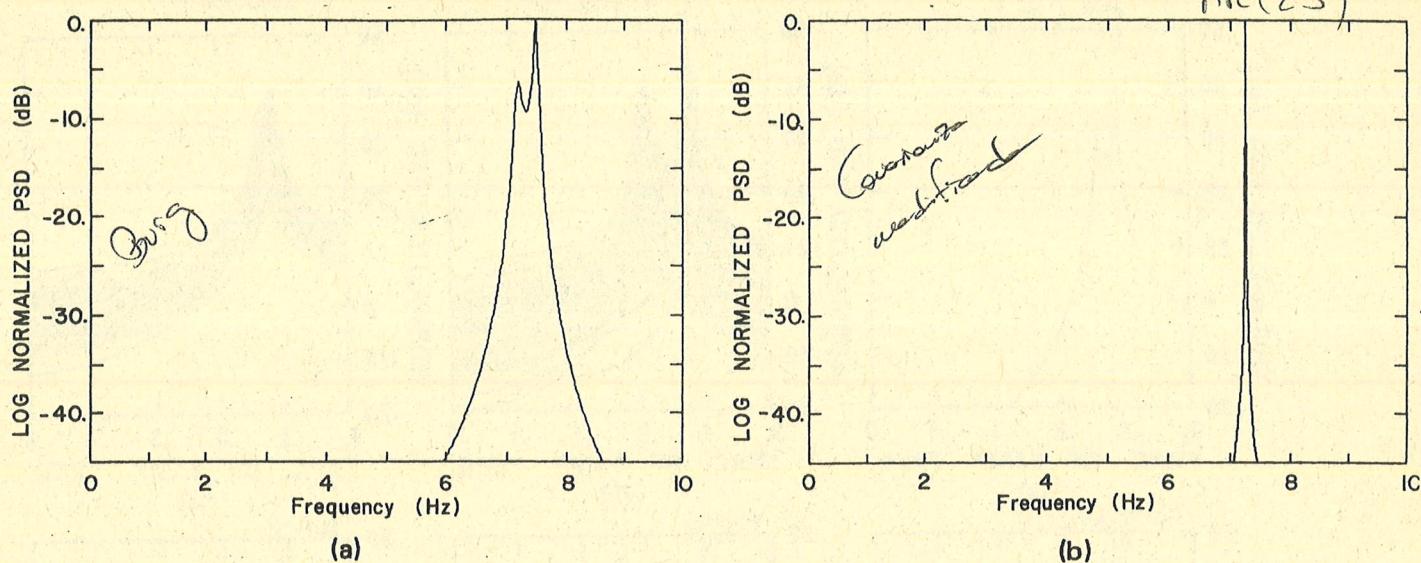


Figure 8.3 Response of two autoregressive spectral estimators to 101 samples of a 7.25-Hz sinusoid in white noise (SNR = 50). The sample rate was 100 Hz and the initial phase was 45° . (a) Burg spectrum with line splitting. (b) Modified covariance spectrum of same data with no line splitting.

Parece que la frecuencia de un solo espectral mejor el método de la covarianza modificada

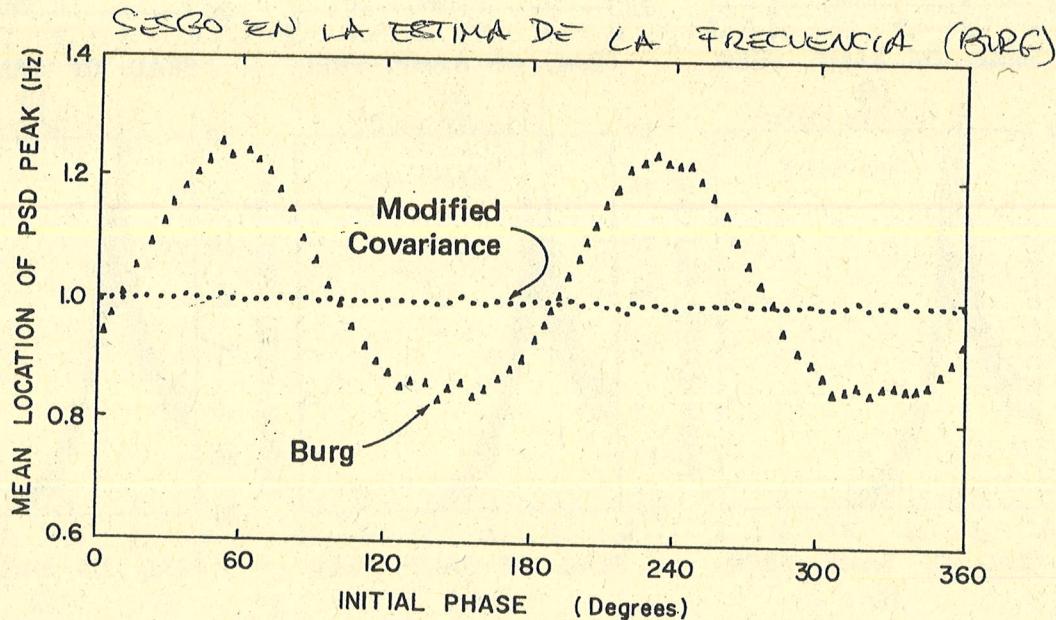


Figure 8.4 Frequency bias of two AR spectral methods. Plot represents the ensemble mean of spectral peak location as a function of initial phase for a cosine signal in white noise (SNR = 10) for Burg and modified covariance methods, using order 9.

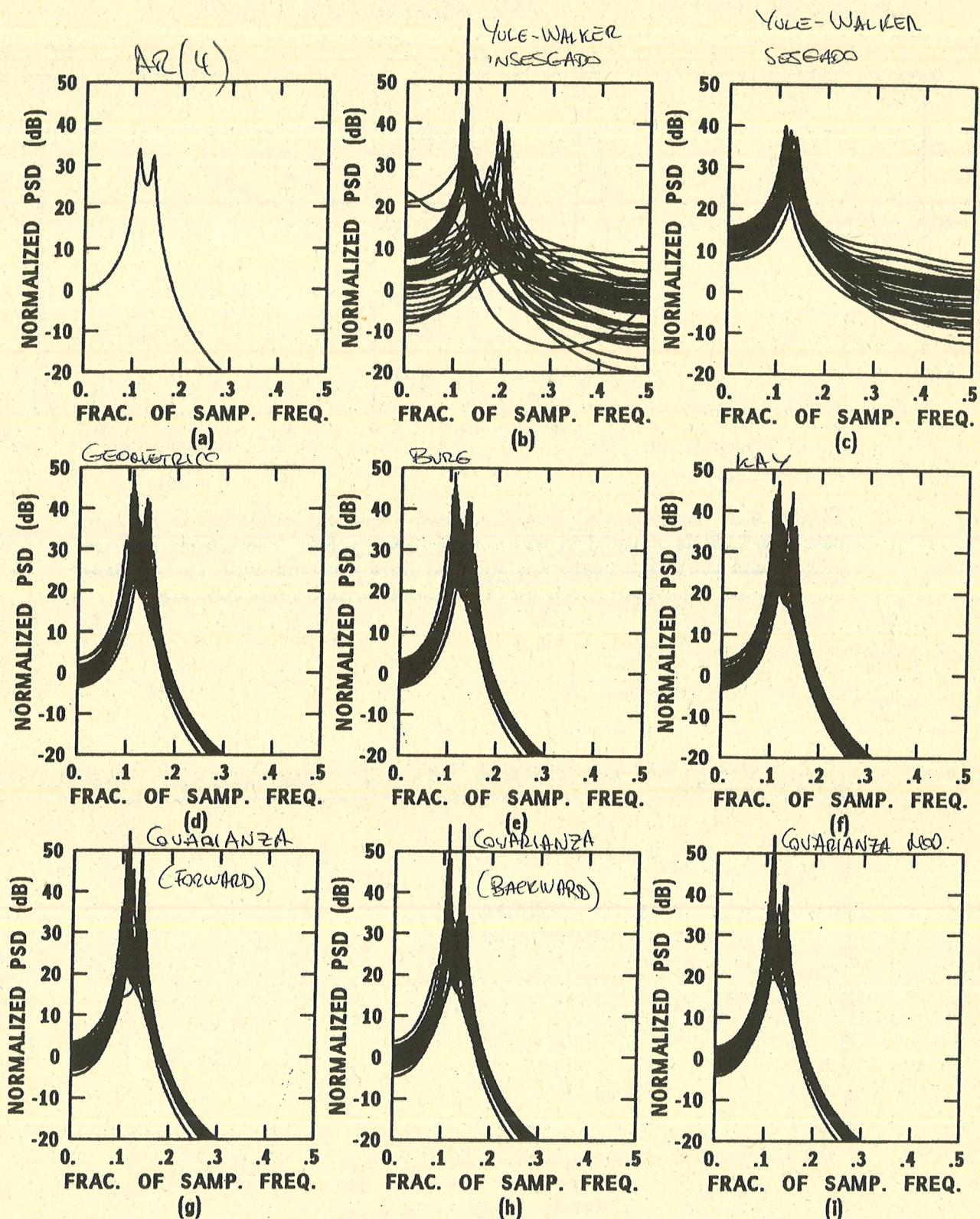


Figure 8.5 Response of five AR spectral estimators to a fourth-order AR process. Overlapped estimates of 50 realizations of 40-sample records are shown for each method. (a) True spectrum of AR(4) process. (b) Yule-Walker, unbiased autocorrelation estimates. (c) Yule-Walker, biased autocorrelation estimates. (d) Geometric method. (e) Burg method. (f) Kay method. (g) Covariance method, forward prediction coefficients. (h) Covariance method, backward prediction coefficients. (i) Modified covariance method.

Figure 6.5 Illustration of spectral factorization and the stability issue.

