

(212)

01149  
0936

1/8

SECCION DE CONTROL  
DIVISION DE ESTUDIOS DE POSGRADO  
FACULTAD DE INGENIERIA  
UNAM

" FILTRADO LINEAL OPTIMO "

TRABAJO DESARROLLADO POR :

OSCAR GIRON CABRERA

noviembre , 1980.

**TESIS CON  
FALLA DE ORIGEN**



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.

# TESIS CON FALLA DE ORIGEN

**TESIS SIN PAGINACION**

## I. INTRODUCCION

Nuestro contacto con los sistemas físicos es primordialmente a partir de las mediciones. Si el objetivo vital es construir un modelo del sistema de nuestra atención, ya sea para propósitos de diseño, análisis o síntesis, entonces se hace necesario de la utilización de técnicas de identificación.

Entonces el Ingeniero de Sistemas hace uso de estas técnicas y decide su aplicación en su problema en cuestión, generalmente esto implica el conocimiento de las leyes físicas que gobiernan el sistema, que puede ser biológico, económico, sociológico, ingenieril, aunque a pesar de las diversas disciplinas el problema será básicamente el de modelar o analizar series de tiempo y a pesar de la terminología cambiante en las disciplinas, la metodología de solución será la misma (Buman K. Mehra, 1971).

En este trabajo se referirá el problema de filtrado óptimo encauzado a sistemas lineales, aunque esto ya es una limitante ya que en realidad el mundo físico se presenta con una propiedad no lineal de comportamiento; pero como se sabe un problema no lineal puede reducirse a uno lineal mediante técnicas de linealización, y aunque resulte esto una simplificación del problema, pero no plega, ya que ocurre que a veces a pesar de trabajar con sistemas lineales el problema resulta complejo; en fin esto conduce a un conocimiento más claro del sistema para los propósitos de estudio.

Como se verá el problema de filtrado comprende la tarea de identificar la señal del ruido, en el contexto de los estudios de Wiener (1949), y Kalman (1960, 1961); en otros terrenos tal como predicción lineal enfocado a procesamiento de voz humana y otras aplicaciones, se desea obtener un filtro inverso tal que si le hacemos pasar ruido blanco, obtenemos una versión cercana de la señal original.

Entonces no perderemos rigurosidad en el tratamiento, o más aún confundirnos con las ideas con respecto a los problemas de filtrado y Predicción lineal, ya que como nuestro P. Oberlin (1971), estos problemas pueden ser formulados como un problema de mínimos cuadrados.

Para el desarrollo de este trabajo se utilizarán secciones destinadas para el tratamiento de los temas de interés en conjunto éste. Las secciones serán:

|      |                           |
|------|---------------------------|
| I.   | INTRODUCCION              |
| II.  | EL PROBLEMA DE ESTIMACION |
| III. | EL PROBLEMA DE FILTRADO   |
| IV.  | PREDICCION LINEAL         |
| V.   | FILTRADO LINEAL           |
| VI.  | SUMARIO                   |

Las secciones se expandirán ante las necesidades que surjan para el estudio de las materias en cuestión. En la sección II estudiamos el problema de estimación, como la identificación requiere de estimadores que minimicen en algún sentido una función, se ven los estimadores más usuales en la literatura, muy importante es ubicar la ecuación de Wiener-Hopf y el lema de Proyección Ortogonal.

En la sección III se estudia un trabajo relacionado a cancelación de ruido de Lidrov y colaboradores para establecer ideas y motivaciones con respecto al filtrado de señales.

En la sección IV se obtiene un modelo autorregresivo (AR) por el método de mínimos cuadrados, para señales discretas, particularmente señales de voz.

En la sección V se estudian los filtros óptimos de Wiener y Kalman, se incluyen problemas de aplicación, finalizamos con un sumario.



Un esquema de estimación de parámetros y de estados con un conocimiento a priori (función de densidad de probabilidad), se ve en la figura 2.a

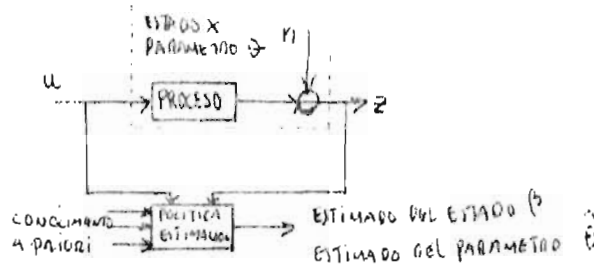


Figura 2.a

A partir de aquí se considera que el ruido acompaña al proceso en una forma aditiva. Considerando solamente estimación de parámetros podemos calcular una regla o un estimador

$$\hat{\theta} = \hat{\theta} \{ u(1), \dots, u(N), z(1), \dots, z(N) \}$$

$$\hat{x} = \hat{x} \{ u, z \}$$

Es decir el estimado depende de las observaciones, cuando tenemos una información a priori del sistema. Estudiemos ahora los estimadores de acuerdo a la cantidad de conocimiento a priori que se requiere para los estimadores, un orden será

- ESTIMACION BAYESIANA
- ESTIMACION DE MAXIMA SIMILITUD
- ESTIMACION DE VARIANZA MINIMA LINEAL
- ESTIMADOR DE MINIMOS CUADRADOS

### II.1. ESTIMACION BAYESIANA.

En este método se considera que existe un conocimiento a priori del proceso, con respecto a las variables desconocidas. Considerando únicamente estimación de parámetros, el estimador de Bayes se basa en la regla

$$P_{x|z}(\theta|\alpha) = \frac{P_{z|x}(\alpha|\theta) P_x(\theta)}{P_z(\alpha)}$$

Donde el lado izquierdo se conoce como la función de densidad de probabilidad a posteriori, es decir una función del estimado  $x = \hat{\theta}$  en base al resultado de la medición  $z = \alpha$ . Si asignamos un costo de aceptar o estimar una variable como  $\theta$ ; cuando  $\theta$  es lo cierto en el problema de estimación, lo hacemos como  $C(\tilde{\theta})$ .

$$B = \int_{-\infty}^{\infty} C(\tilde{\theta}) P_{z|\theta}(\alpha, \theta) d\alpha d\theta$$



El error de estimación es  $\tilde{\theta} = \theta - \hat{\theta}$ ,  $z$  es la observación,  $P(\alpha, \theta)$  es conocimiento a priori, vemos que  $\mathcal{B}$  es también

$$\mathcal{B} \triangleq E\{c(\tilde{\theta})\} = E\{c(\theta - \hat{\theta})\}$$

Como deseamos obtener  $\hat{\theta}$ , un estimador óptimo, dependerá de la elección del costo que asignemos a nuestro problema. Los más conocidos son

$$c(\tilde{\theta}) = \|\tilde{\theta}\|_2^2; \quad c(\tilde{\theta}) = \|\tilde{\theta}\|_1; \quad c(\tilde{\theta}) = \begin{cases} 0 & \|\tilde{\theta}\|_1 < \epsilon/2 \\ \frac{1}{\epsilon} \|\tilde{\theta}\|_1 & \|\tilde{\theta}\|_1 \geq \epsilon/2 \end{cases}$$

Con  $S$  Matriz definida positiva y simétrica. El estimador que minimice esta expresión se denomina el estimador de costo mínimo. Consideremos ahora dos estimadores que son un caso especial del estimador de bayes.

### II.1-A. ESTIMADOR DE MINIMA VARIANZA.

Este método minimiza el  $\mathcal{B}$

$$\mathcal{B}_{mv} = \iint c(\tilde{\theta}) P_{\theta, z}(\theta, \alpha) d\theta d\alpha$$

utilizando la regla de bayes  $P_{\theta, z}(\theta, \alpha) = P_{\theta|z}(\theta|\alpha) P_z(\alpha)$

$$\mathcal{B}_{mv} = \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} c(\tilde{\theta}) P_{\theta|z}(\theta|\alpha) d\theta \right\} P_z(\alpha) d\alpha$$

como  $P_z(\alpha) \geq 0$ , vemos que minimizando  $\mathcal{B}_{mv}$  es haciendo la integral encerrada por corchetes muy pequeño, es decir:

$$\min_{\tilde{\theta}} \int_{-\infty}^{\infty} c(\tilde{\theta}) P_{\theta|z}(\theta|\alpha) d\theta$$

que es también

$$\min_{\tilde{\theta}} E\{c(\tilde{\theta}) | z = \alpha\}$$

y

$$\mathcal{B}_{mv}(\hat{\theta}|\alpha) = \int c(\tilde{\theta}) P_{\theta|z}(\theta|\alpha) d\theta$$

una condición necesaria para minimizar la expresión anterior es

$$\frac{\partial}{\partial \tilde{\theta}} \mathcal{B}_{mv}(\hat{\theta}|\alpha) \Big|_{\tilde{\theta} = \hat{\theta}_{mv}} = 0$$

donde queda que

$$\hat{\theta}_{mv}(\alpha) = \int_{-\infty}^{\infty} \theta P_{\theta|z}(\theta|\alpha) d\theta$$

es decir, el valor esperado condicional dadas las observaciones.

### II.1.B. ESTIMADOR MAXIMO A POSTERIORI (MAP)

Este estimador,  $\hat{\theta}_{map}$  se define como

$$\frac{\partial P_{\theta|z}(\theta|\alpha)}{\partial \theta} \Big|_{\theta = \hat{\theta}_{map}(\alpha)} = 0$$

y se determina maximizando la probabilidad a posteriori  $P_{\theta|z}(\theta|\alpha)$  con respecto a  $\theta$ , que es:

$$\left. \frac{\partial P_{\theta|z}(\theta|\alpha)}{\partial \theta} \right|_{\theta = \hat{\theta}_{MAP}(\alpha)} = \left. \frac{\partial P_{z|\theta}(\alpha|\theta) P_{\theta}(\theta)}{\partial \theta} \right|_{\theta = \hat{\theta}_{MAP}} = 0$$

## II.2. ESTIMADOR DE MÁXIMA SIMILITUD

En este estimador se asume no tener un conocimiento a priori de  $\theta$ , digamos que es un valor constante y no variable aleatoria, se expresa como

$$\frac{\partial}{\partial \theta} P_{z|\theta}(\alpha|\theta) \Big|_{\theta = \hat{\theta}_{ML}} = 0$$

ó

$$\frac{\partial}{\partial \theta} \ln P_{z|\theta}(\alpha|\theta) \Big|_{\theta = \hat{\theta}_{ML}} = 0$$

Hasta aquí hemos tratado con estimadores que requieren de un estudio en términos estadísticos o probabilísticos para las observaciones y parámetros. El método que sigue, el de mínima varianza el estimador es definido como una combinación pesada de las observaciones.

## II.3. ESTIMADOR DE VARIANZA MÍNIMA LINEAL.

En este método es muy importante destacar el resultado de la ecuación de Wiener-Hopf y el lema de proyección ortogonal. En este método se trabaja solamente con los primer y segundo momentos estadísticos, es decir la esperanza matemática y la varianza, y el estimado óptimo es una combinación lineal pasada de las observaciones. Asumiendo el modelo de observación

$$z(t) = H(t)x(t) + v(t) \quad 0 \leq t \leq T_f \quad (2.3-1)$$

Donde la observación  $z(t)$  y el ruido de la observación son  $M$  vectores.  $H(t)$  es una matriz y se conoce;  $x(t)$  es el parámetro aleatorio a estimar y  $v$  es ruido blanco con media cero, aunque no necesariamente gaussiano con

$$\text{cov}\{v(t), v(\tau)\} = \Psi_v(t) \delta_{\theta}(t-\tau) \quad (2.3-2)$$

Los dos primeros momentos estadísticos de  $X$  son

$$E\{x(t)\} = \mu_x(t), \quad \text{Var}\{x(t)\} = V_x(t) \quad (2.3-3)$$

obteniendo el estimado de  $x(t)$  una clase de estimadores lineales no sesgados para el cual

$$\hat{x}(t) = \xi(t) + \int_0^t \Xi(t,\lambda) z(\lambda) d\lambda \quad (2.3-4)$$

El que minimiza la varianza del error

$$E \{ \|\tilde{x}_L(t)\|^2 \} = \text{Tr}(\text{var} \{ \tilde{x}_L(t) \}) = \text{Tr}(\text{var} \{ x(t) - \hat{x}_L(t) \}) \quad (2.3-5)$$

El estimador resultante se denomina estimador de mínima varianza lineal. Tomando el valor esperado en ec. (2.3.4), tenemos

$$E \{ \hat{x}_L(t) \} = \xi(t) + \int_0^t \Xi(t, \lambda) H(\lambda) M_x(\lambda) d\lambda = M_x(t) \quad (2.3.6)$$

y  $\xi(t)$  queda

$$\xi(t) = M_x(t) - \int_0^t \Xi(t, \lambda) H(\lambda) M_x(\lambda) d\lambda \quad (2.3-7)$$

sustituyendo en (2.3-4)

$$\tilde{x}_L(t) = M_x(t) + \int_0^t \Xi(t, \lambda) [z(\lambda) - H(\lambda) M_x(\lambda)] d\lambda \quad (2.3-8)$$

minimizando la expresión de varianza del error (2.3-5), por cálculo de variaciones, obtenemos

$$\text{cov} \{ z(\lambda), x(t) \} = \text{cov} \{ z(\lambda), \hat{x}_{Luv}(t) \} \quad (2.3-9)$$

$$\text{cov} \{ z(\lambda), \hat{x}_{Luv}(t) \} = 0 \quad (2.3-10)$$

donde  $\hat{x}_{Luv}$  es el estimador óptimo con  $\hat{\Xi}$  valor óptimo de  $\Xi$  igual a  $\Xi$  es decir  $E(x) = \hat{E}(t, \lambda)$ . La ecuación (2.3.9) usualmente se escribe en una forma integral

$$\text{cov} \{ x(t), z(\lambda) \} = \int_0^t \Xi(t, \tau) \text{cov} \{ z(\tau), z(\lambda) \} d\tau \quad (2.3-11)$$

(2.3-9) y (2.3-11) se denominan como ecuaciones de Wiener-Hopf, donde para obtener  $\Xi$  se debe resolver la integral el cual no es fácil. Para obtener el lema de proyección ortogonal se premultiplica ambos lados de (2.3-8) por  $\Xi(t, \lambda)$  e integrando sobre  $\lambda$ , de 0 a  $T$ , obteniendo

$$\text{cov} \{ \hat{x}_{Luv}(t), x(t) \} = \text{cov} \{ \hat{x}_{Luv}, \hat{x}_{Luv}(t) \} \quad (2.3-12)$$

$$\text{cov} \{ \hat{x}_{Luv}(t), \hat{x}_{Luv}(t) \} = 0. \quad (2.3-13)$$

lo que significa que el estimador de mínima varianza lineal óptimo es ortogonal al error en la estimación. Ahora determinemos una forma de  $\Xi$  y nos ser

virá en lo sucesivo, particularmente en la solución del filtro de Kalman en el caso discreto.

Aprovechando el lema de proyección ortogonal, deseamos obtener  $\Xi$  tal que el estimador sea

$$\hat{X}_{LHV} = \Xi z$$

entonces tenemos que

$$\begin{aligned} \text{COV} \{ \hat{X}_{LHV}, X - \hat{X}_{LHV} \} &= 0 \\ E \{ X \cdot \hat{X}_{LHV}^T \} &= E \{ \hat{X}_{LHV} \hat{X}_{LHV}^T \} \\ E \{ X, z^T \Xi^T \} &= E \{ \Xi z z^T \Xi^T \} \\ \Xi^T E \{ X z^T \} &= \Xi^T E \{ z z^T \} \Xi^T \\ \Xi &= E \{ X z^T \} (E \{ z z^T \})^{-1} \\ \hat{X}_{LHV} &= \Xi z = E \{ X z^T \} (E \{ z z^T \})^{-1} z \end{aligned}$$

#### II.4-ES IMPLICACION POR MÍNIMOS CUADRADOS

En este método no se necesita ninguna descripción probabilística o estadística. Aquí el problema es seleccionar un estimador  $\hat{X}_k$  de  $X$  tal que se minimice

$$J(\hat{X}_k) = \frac{1}{2} (z_k - H_k \hat{X}_k)^T R_k^{-1} (z_k - H_k \hat{X}_k) \quad (2.4-1)$$

El estimador óptimo que minimiza (2.4-1), se denomina estimador de mínimos cuadrados  $\hat{X}_{k|k}$ . Entonces el problema de optimización es determinístico, y una condición necesaria para que (2.4-1) sea mínimo es

$$\left. \frac{\partial J(\hat{X}_k)}{\partial \hat{X}_k} \right|_{\hat{X}_k = \hat{X}_{k|k}} = 0 \quad (2.4-2)$$

Assumiendo el modelo de observaciones

$$z_{k+1} = H_{k+1} X + V_{k+1} \quad (2.4-3)$$

el estimado es según (2.4-2)

$$\hat{X}_{k+1|k+1} = (H_{k+1}^T R_{k+1}^{-1} H_{k+1})^{-1} H_{k+1}^T R_{k+1}^{-1} z_{k+1}$$

$$\hat{X}_{k+1|k+1} = \hat{X}_{k|k} + P_{k+1} H_{k+1}^T R_{k+1}^{-1} (z_{k+1} - H_{k+1} \hat{X}_{k|k})$$

$$P_{k+1} = (H_{k+1}^T R_{k+1}^{-1} H_{k+1})^{-1}$$

A continuación mostramos un ejemplo tomado de Sage y Melsa (1971), que ilustra la solución para los estimadores estudiados anteriormente

EjemPlo. Consideremos el problema de estimar un parámetro constante  $X$ , de la observación

$$Z = Hx + v$$

con  $X$  un vector  $N$ -dimensional,  $Z$  y  $V$  son  $M$ -dimensional ( $M \geq N$ ),  $H$  es matriz  $M \times N$ .  $X$  y  $V$  son independientes y gaussianos, sus valores esperados y varianzas son

$$\begin{aligned} E\{X\} &= \mu_x & \text{Var}\{X\} &= V_x \\ E\{V\} &= 0 & \text{Var}\{V\} &= V_v \end{aligned}$$

Para la estimación bayessiana tenemos:

la densidad a posteriori es

$$P_{X|Z}(\beta|d) = \frac{[\det(HV_xH^T + V_v)]^{1/2}}{(2\pi)^{M/2} (\det V_x)^{1/2} (\det V_v)^{1/2}} \exp\left[-\frac{1}{2}(\beta - \bar{\beta})^T \Xi^{-1} (\beta - \bar{\beta})\right]$$

donde

$$\begin{aligned} \bar{\beta} &= (V_x^{-1} + H^T V_v^{-1} H)^{-1} (H^T V_v^{-1} \alpha + V_x^{-1} \mu_x) = \Xi (H^T V_v^{-1} \alpha + V_x^{-1} \mu_x) \\ \Xi^{-1} &= V_x^{-1} + H^T V_v^{-1} H \end{aligned}$$

el estimador de varianza mínima es

$$\hat{X}_{MV} = \int_{\omega} \beta P_{X|Z}(\beta|d) d\beta = \bar{\beta}, \quad V_{\hat{X}} = V_x - V_x H^T (V_v + H V_x H^T)^{-1} H V_x$$

el estimador máximo a posteriori es

$$\hat{X}_{MAP} = \bar{\beta}, \quad V_{\hat{X}} = V_x - V_x H^T (V_v + H V_x H^T)^{-1} H V_x$$

el estimador de máxima relación de similitud es

$$\frac{\partial \log P_{X|Z}(\alpha|d)}{\partial \alpha} \Big|_{\alpha = \hat{X}_{ML}} = 0, \quad \hat{X}_{ML} = (H^T V_v^{-1} H)^{-1} H^T V_v^{-1} \alpha, \quad V_{\hat{X}} = (H^T V_v^{-1} H)^{-1}$$

el estimador de mínima varianza lineal es

$$\hat{X}_{LHV} = (V_x^{-1} + H^T V_v^{-1} H)^{-1} (H^T V_v^{-1} z + V_x^{-1} \mu_x), \quad V_{\hat{X}} = (H^T V_v^{-1} H)^{-1}$$

El estimador de mínimos cuadrados es

$$\hat{X}_{LS} = (H^T R^{-1} H)^{-1}$$

Vemos que

$$\hat{X}_{MV} = \hat{X}_{MAP}$$

Lo que muestra que el estimado óptimo de Bayes es invariante, en este caso e igual al estimador de varianza mínima.

Para  $\hat{X}_{MV} = \hat{X}_{ML}$  si,  $V_X^{-1} = M_X = 0$

el estimador de varianza mínima degenera al estimador de máxima verosimilitud, además el estimador de varianza mínima es no sesgado, como lo es el estimador de máxima verosimilitud y

Para  $\hat{X}_{LM} = (H^T V_v^{-1} H)^{-1} H^T V_v^{-1} z$  con  $M_X = V_X^{-1} = 0$

y  $\hat{X}_{LS} = X_{LM}$ ,  $R = V_v$

una comparación de estos estimadores, muestra que con R=V

$\hat{X}_{LS} = X_{LM}$   
y son no sesgados.

Hasta aquí se ha tratado brevemente el problema de la estimación, herramientas como la ecuación de Wiener-Hopf y el lema de proyección ortogonal son de utilidad en la sección de filtrado.

### III. EL PROBLEMA DE FILTRADO

Para fijar ideas y como motivación del problema de filtrado describamos el método propuesto por Widrow y colaboradores (1975), para la eliminación del ruido. La filosofía de este método es la de separar el ruido aditivo por cancelación.

Una señal  $S$  se transmite en un canal, que se recibe en un sensor juntamente con un ruido aditivo  $N_0$ , a la suma  $S+N_0$  se le denomina entrada primaria; en otro sensor se recibe un ruido  $N_1$ , no correlacionado con  $S$  pero si en alguna forma con  $N_0$ . Este sensor provee una guía de referencia al cancelador, el ruido  $N_1$  es filtrado para obtener una salida del filtro  $Y$  y esta salida  $Y$  es restada de la entrada primaria para producir una salida del sistema

$$Z = S + N_0 - Y \quad (3.1)$$

veamos que  $Z = S$  cuando  $N_0 = Y$ , lo que significa un conocimiento de las características de los canales de transmisión. Pero esto no es posible, por lo que trataremos de minimizar la potencia de salida del sistema, es decir

$$\min E\{Z^2\} = E\{S^2\} + E\{(N_0 - Y)^2\} \quad (3.2)$$

donde vemos que minimizando la energía de la salida, la energía de la señal no cambia, mientras que para la expresión  $E\{(N_0 - Y)^2\}$  si ocurre la minimización. notamos que  $Z$  es un estimador de la señal. Un resultado importante en el artículo es el siguiente "desde que minimizando  $E\{Z^2\}$  minimizamos también  $E\{(N_0 - Y)^2\}$ , minimizando la potencia total de salida, minimizamos la potencia de salida del ruido  $E\{(N_0 - Y)^2\}$  y desde que la señal en la salida permanece constante minimizando la potencia total de salida maximizamos la relación señal a ruido". Un esquema del cancelador de ruido lo vemos en la figura 3.1

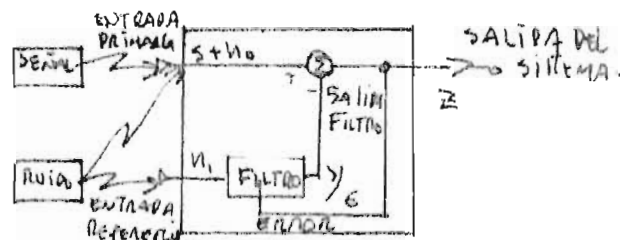


FIGURA 3.1. CANCELADOR DE RUIDO

Más aún vemos que si

$$E\{z^2\} = E\{s^2\}, \text{ por consiguiente } Y=N_0, E=S$$

lo que significa en

$$E\{(N_0 - Y)^2\} = 0$$

en este caso resulta que minimizando la potencia de salida, la señal de salida estará libre de ruido. Una serie de aplicaciones de la técnica de cancelación de ruido se han incluido en la cancelación de interferencia de 60-Hz en electrocardiografía, cancelación del electrocardiograma del donador en trasplante de corazón, cancelación del electrocardiograma materno en electrografía fetal, cancelación de ruido en señales de voz, entre otros. Describamos el método para la cancelación de interferencia de 60-Hz en electrocardiografía.

En el registro de electrocardiogramas (ECGs), es inevitable la aparición de interferencia de 60-Hz en la salida. Existen otros métodos para la eliminación de interferencia, este método puede aplicarse. La entrada primaria se toma del amplificador del electrocardiograma; la entrada de referencia de 60-Hz se toma de un contacto de pared.

El filtro contiene dos variables, uno aplicado directamente a la entrada de referencia y la otra una versión retrasada  $90^\circ$ , las dos variables se suman para formar la salida del filtro, que se resta de la entrada primaria. Mostremos un resultado típico de un grupo de experimentos ejecutados con una computadora en tiempo real en la figura 3.2. la relación de muestreo fue de 1000 Hz. 3.2-a muestra la entrada primaria, un electrocardiograma con cantidad excesiva de interferencia, 3.2-b, muestra la entrada de referencia; y 3.2-c es la salida del cancelador de ruido.



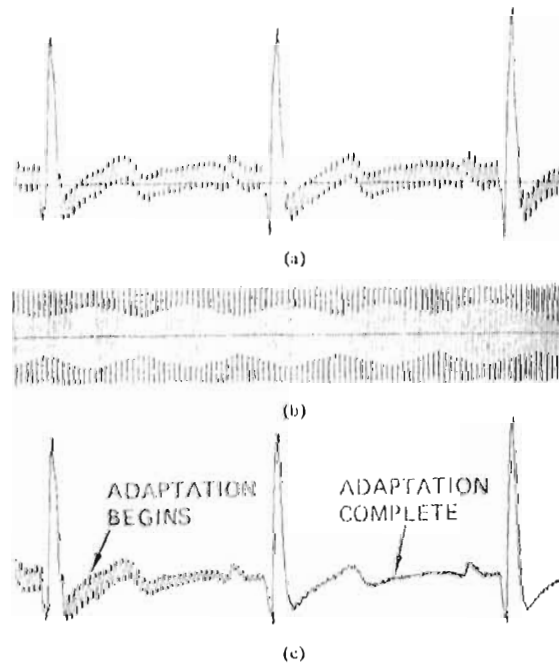


Fig. 3.1 Result of electrocardiographic noise cancelling experiment.  
 (a) Primary input. (b) Reference input. (c) Noise canceller output.

#### IV. PREDICCIÓN LINEAL

En esta sección concerniente a predicción lineal, estudiaremos la aplicación básicamente en el análisis de voz. Entonces se hace necesario un comentario breve sobre la producción de voz, detalles sobre fisiología y características acústicas se encuentran en Fant(1960) y Flanagan(1972). Labiner(1978), muestra que la voz puede ser modelado como la salida de un sistema lineal, variante con el tiempo, excitado por pulsos quasi-periódicos para la producción de los sonidos sonoros, tales como las vocales y los sonidos nasales; y excitado por ruido aleatorio para los sonidos no sonoros como los sonidos fricativos tales como los f, th, s, sh. Podemos relacionar de inmediato un problema en predicción lineal, tal como identificar si un sonido es sonoro o fricativo, para el sonoro implica la determinación del paso o distancia entre pulsos. Un modelo de producción de voz se muestra en la figura 4.1.

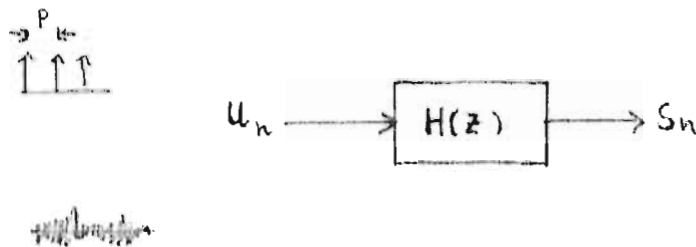


FIGURA 4.1. MODELO DE PRODUCCIÓN DE VOZ DISCRETO.

Al aplicar análisis de series de tiempo, cada señal continua  $s(t)$  es muestreada para obtener una señal discreta  $s(nT)$ , conocido como serie de tiempo, donde  $n$  es un entero y  $T$  es el intervalo de muestreo, la frecuencia de muestreo es  $f_s = 1/T$ . Abreviemos  $s(nT)$  por  $s_n$ . Consideremos la señal  $s_n$ , como la salida de un sistema con una entrada desconocida  $u_n$ .

$$s_n = -\sum_{k=1}^p a_k s_{n-k} + \sum_{l=0}^q b_l u_{n-l}, \quad b_0 = 1 \quad (4.1)$$

$a_k, 1 \leq k \leq p, \quad b_l, 1 \leq l \leq q,$

La ecuación 4.1, muestra que la salida  $s_n$  es una función lineal de las salidas y las entradas presentes y pasadas, es decir,  $s_n$  es predecible de las com-

La ecuación 4.1, muestra que la salida  $s_n$  es una función lineal de las salidas pasadas, y las entradas presentes y pasadas, es decir que  $s_n$  es predecible de las combinaciones lineales de pasadas salidas y entradas, de aquí la derivación del término Predicción Lineal (Makhoul, 1975). Sin embargo la idea básica de Predicción lineal lleva a un conjunto de técnicas de análisis que pueden ser usados para la estimación de parámetros de un modelo de voz, este conjunto de técnicas de análisis es referido como Linear predictive Coding o LPC (Sabiner 1978). Tomando transformada-Z ambos lados de 4.1, tenemos la función de transferencia

$$H(z) = \frac{S(z)}{u(z)} = \frac{1 + \sum_{k=1}^q b_k z^{-k}}{1 + \sum_{k=1}^p a_k z^{-k}} \quad (4.2)$$

donde

$$S(z) = \sum_{n=-\infty}^{\infty} s_n z^{-n} \quad (4.3)$$

ecuación 4.2 es un modelo general de polos y ceros, se derivan dos casos especiales:

- 1) modelo de ceros
- 2) modelo de polos

El modelo de ceros es conocido en la literatura estadística como modelo Moving Average (MA), el modelo de polos como autorregresivo (AR); el modelo general es un modelo autorregresivo moving average (ARMA). El modelo más aplicado a voz es el AR, pero el modelo ARMA el problema resulta no lineal, aunque se han realizado trabajos en el dominio del tiempo (Makhoul, 1975). Un enfoque semejante aplicado a problemas distintos de voz es referido por Mehra (1971). Varias formulaciones, aunque equivalentes han sido aplicados a voz, tales como

1. El método de covarianza
2. El método de autocorrelación
3. El método de rejilla

Procedamos primero a una formulación intuitiva de mínimos cuadrados debido a Makhoul (1975), para incluir brevemente las formulaciones citadas.

Considerando un modelo de polos (AR), asumimos que  $s_n$  es predecible por una combinación lineal de valores pasados y alguna entrada  $u_n$

$$s_n = -\sum_{k=1}^p a_k s_{n-k} + G u_n \quad (4.4)$$

donde  $G$  es un factor de ganancia, el problema es determinar los coeficientes predictores  $a_k$  y la ganancia  $G$ , la función de transferencia  $H(z)$  es

$$H(z) = \frac{G}{1 + \sum_{k=1}^p a_k z^{-k}} \quad (4.5)$$

asumiendo un predictor dependiente de las señales pasadas tenemos

$$\tilde{s}_n = -\sum_{k=1}^p a_k s_{n-k} \quad (4.6)$$

el error entre el valor actual  $s_n$  y el valor predecido  $\tilde{s}_n$  es

$$e_n = s_n - \tilde{s}_n = s_n + \sum_{k=1}^p a_k s_{n-k} \quad (4.7)$$

$e_n$  se conoce también como residual. los coeficientes predictores  $a_k$  son obtenidos en el método de mínimos cuadrados, por la minimización del error total al cuadrado, con respecto a los parámetros. Considerando que  $s$  es una señal determinística, la derivación para señal estocástica es análogo, derivemos para la primera.

queremos minimizar  $E$  donde

$$E = \sum_n e_n^2 = \sum_n \left( s_n + \sum_{k=1}^p a_k s_{n-k} \right)^2 \quad (4.8)$$

una condición necesaria para el mínimo es

$$\frac{\partial E}{\partial a_i} = 0 \quad 1 \leq i \leq p \quad (4.9)$$

de (4.8) y (4.9) obtenemos

$$\sum_{k=1}^p a_k \sum_n s_{n-k} s_{n-i} = -\sum_n s_n s_{n-i}, \quad 1 \leq i \leq p \quad (4.10)$$

este conjunto de ecuaciones se conocen como ecuaciones normales, para cuales

cualquier definición de la señal  $s_n$ , (4.10) forman un conjunto de  $p$ -ecuaciones con  $p$ -incógnitas, que se pueden resolver para los coeficientes predictores  $a_k$  que minimizan  $E$  en (4.8)

El error total mínimo  $E_p$ , es obtenido por expandiendo (4.8) y sustituyendo (4.10)

$$E_p = \sum_n s_n^2 + \sum_{k=1}^p a_k \sum_n s_n s_{n-k} \quad (4.11)$$

El rango de suma sobre  $n$  en (4.8) nos lleva a dos métodos distintos para la estimación de parámetros. El primer método es el de autocorrelación.

Método de autocorrelación. Se asume que el error en 4.8 se minimiza sobre el intervalo infinito  $-\infty < n < \infty$ . Las ecuaciones (4.10) y (4.11) se reducen a

$$\sum_{k=1}^p a_k R(i-k) = -R(i), \quad 1 \leq i \leq p \quad (4.12)$$

$$E_p = R(0) + \sum_{k=1}^p a_k R(k) \quad (4.13)$$

donde

$$R(i) = \sum_{n=-\infty}^{\infty} s_n s_{n+i}, \quad \text{función Autocorrelación.} \quad (4.14)$$

En la práctica, la señal  $s_n$  se conoce sobre un intervalo finito, por lo que debemos multiplicar la señal por una ventana  $w_n$ , para trancar la señal, es decir que la señal  $s'_n$  es cero fuera del intervalo  $0 \leq n \leq N-1$ :

$$s'_n = \begin{cases} s_n w_n & 0 \leq n \leq N-1 \\ 0 & \text{otra forma} \end{cases} \quad (4.16)$$

La función de autocorrelación es

$$R(i) = \sum_{n=0}^{N-1-i} s'_n s'_{n+i}, \quad i \geq 0 \quad (4.17)$$

Método de covarianza. Aquí el error  $E$  en (4.8) se minimiza sobre un intervalo finito,  $0 \leq n \leq N-1$ . Las ecuaciones (4.10) y (4.11) se reducen a

$$\sum_{k=1}^p a_k \psi_{ki} = -\psi_{0i}, \quad 1 \leq i \leq p \quad (4.18)$$

$$E_p = \psi_{00} + \sum_{k=1}^p a_k \psi_{0k} \quad (4.19)$$

donde

$$\psi_{ik} = \sum_{n=0}^{N-1} s_{n-i} s_{n-k} \quad (4.20)$$

Es la covarianza de la señal  $s_n$ .

### IGUALAMIENTO DE CORRELACION

Se desea un igualamiento entre la correlación de secuencia de la señal  $\{s_n\}$  y la respuesta a impulso del filtro  $H(z)$  en muchos puntos como sea posible. Asumiendo un modelo causal  $H(z)$  de la forma

$$H(z) = \frac{G}{A(z)} \quad (4.21)$$

donde como antes

$$A(z) = \sum_{k=0}^p a_k z^{-k}, \quad a_0 = 1 \quad (4.22)$$

multiplicación de ambos lados de (4.21) y un transformación  $-z$  inversa nos da

$$\sum_{k=0}^p a_k h_{n-k} = G \delta_{n0} \quad (4.23)$$

donde la secuencia  $\{h_n\}$  tiene la transformada- $z$ ,  $H(z)$ .

Como el filtro se asumió causal,  $h_n = 0$  para  $n < 0$ , y es estable, la secuencia de autocorrelación para el impulso unitario es

$$\rho(k-l) = \sum_{n=-\infty}^{\infty} h_{n-k} h_{n-l} = \sum_{n=0}^{\infty} h_n h_{n+|k-l|} \quad (4.24)$$

si (4.23) se multiplica por  $h_{n-l}$  y sumado para toda  $n$ , y aplicando (4.24) tenemos

$$\sum_{k=0}^p a_k \rho(k-l) = G h_{-l} \quad (4.25)$$

Como el filtro se asumió causal el lado derecho de (4.25) es cero para  $l > 0$  y de (4.23),  $h_0 = G$ , haciendo el lado derecho de (4.25) igual a  $G^2$  para  $l=0$ .

$$\sum_{k=0}^p a_k \rho(k-l) = 0 \quad \text{PARA } l > 0 \quad (4.26)$$

$$\sum_{k=0}^p a_k \rho(k) = G^2. \quad (4.27)$$

Para determinar los  $p+1$  parámetros del modelo  $H(z)$ ,  $\{G, a_1, a_2, \dots, a_p\}$  los

primeros  $P+1$  muestras de autocorrelación de la respuesta impulso unitario del filtro son elegidos para igualar los primeros muestras de autocorrelación de la secuencia de la señal  $\{s_n\}$  es decir

$$Q(l) = R(l) \quad \text{Para } l=0,1,\dots,P \quad (4.28)$$

conde  $Q(k,l)$  es dado por (4.24) y

$$R(k-l) = \sum_{n=-\infty}^{\infty} s(n-k)s(n-l) = \sum_{n=0}^{N-|k-l|} s_n s_{n+|k-l|} \quad (4.29)$$

combinando (4.23) con (4.26) para  $l=0,1,\dots,P$ , tenemos exactamente las ecuaciones que deben ser resueltas en el método de autocorrelación. combinando (4.23) con (4.27) result. en una ecuación que iguala la energía del espectro de la señal a la energía del la respuesta impulso del modelo  $h(z)$  es decir

$$G^2 = \sum_{k=0}^P a_k R(k) \quad (4.30)$$

Antes de pasar a otra formulación para LPC, vemos dos formas recursivas de determinar los coeficientes predictores en los métodos de autocorrelación y covarianza.

#### Solución de las ecuaciones de LPC.

Para que los esquemas de LPC, como son el método de correlación y covarianza funcionen de una manera más eficiente, es necesario la implementación de métodos eficientes para la determinación de los parámetros predictores. Mostremos dos métodos aplicables como son el de Durbin para el método de autocorrelación y descomposición de Cholesky para el método de covarianza.

Procedimiento recursivo de Durbin. Para el método de autocorrelación la ecuación a resolver para determinar los parámetros predictores es

$$(4.31)$$

que expandido en forma matricial es

$$\begin{bmatrix} R_0 & R_1 & R_2 & \dots & R_{p-1} \\ R_1 & R_0 & R_1 & \dots & R_{p-2} \\ R_2 & R_1 & R_0 & \dots & R_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R_{p-1} & R_{p-2} & R_{p-3} & \dots & R_0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ \vdots \\ R_p \end{bmatrix} \quad (4.32)$$

que es de naturaleza Toeplitz, Levinson y Robinson derivan un método de solución aunque el más eficiente es el procedimiento recursivo de Durbin. El algoritmo es

$$\begin{aligned} E_0 &= R(0) & 4.33. \quad (a) \\ k_i &= - \left[ R(i) + \sum_{j=1}^{i-1} a_j^{(i-1)} R(i-j) \right] / E_{i-1} & (b) \\ a_i^{(i)} &= k_i & (c) \\ a_j^{(i)} &= a_j^{(i-1)} + k_i a_{i-j}^{(i-1)} \quad 1 \leq j \leq i-1 & (d) \\ E_i &= (1 - k_i^2) E_{i-1} & (e) \end{aligned}$$

Las ecuaciones 4.33 b, c se resuelven recursivamente para  $i=1, 2, \dots, P$  y la solución final es dado por

$$a_j = a_j^{(P)}, \quad 1 \leq j \leq P \quad (4.34)$$

Notar que en el proceso de resolver para los coeficientes predictores para un coeficiente de orden  $P$ , las soluciones para los coeficientes de los ordenes menos que  $P$  han sido obtenidos. Los  $E_i$  en la ecuación 4.33 e, es el error de predicción para un predictor de orden  $i$ . La solución de 4.32 no es afectada si normalizamos por dividiendo las  $R(i)$  por  $R(0)$ ; estos coeficientes resultantes son denominados coeficientes de autocorrelación normalizadas  $V(i)$ :



$$r(i) = \frac{R(i)}{R(0)} \quad (4.35)$$

El error  $E_i$  es ahora un error normalizado

$$V_i = E_i / R(0) = 1 + \sum_{k=1}^p a_k r(k) \quad (4.36)$$

$$0 \leq V_i \leq 1, \quad i \geq 0 \quad (4.37)$$

y el error normalizado  $V_p$  es

$$V_p = \prod_{i=1}^p (1 - \kappa_i^2) \quad (4.38)$$

donde las cantidades  $\kappa_i$  están en el rango

$$-1 \leq \kappa_i \leq 1 \quad (4.39)$$

Esta condición sobre los parámetros  $\kappa_i$  es importante, desde que es condición necesaria y suficiente para que todas las raíces del polinomio  $A(z)$ , estén dentro del círculo unitario, garantizando la estabilidad del filtro  $H(z)$

Procedimiento de la descomposición de Cholesky. Para el método de covarianza el conjunto de ecuaciones a resolver es de la forma

$$\sum_{k=1}^p a_k \psi_{ki} = -\psi_{0i} \quad 1 \leq i \leq p \quad (4.40)$$

En forma matricial

$$\phi a = -\psi \quad (4.41)$$

con  $\phi$  una matriz simétrica positiva definida,  $a$  y  $\psi$  son vectores columna

El sistema de ecuaciones de la ecuación (4.41) pueden ser resueltos en una manera eficiente desde que  $\Phi$  es una matriz simétrica positiva definida, un método de solución es el de la descomposición de Cholesky, para este método la matriz  $\Phi$  se expresa en la forma

$$\Phi = VDV^T \quad (4.42)$$

con  $V$ , una matriz triangular inferior (con la diagonal principal formada con  $\underline{u}$  nos),  $D$  es una matriz diagonal. De (4.41) y (4.42) tenemos

$$VDV^T a = \psi \quad (4.43)$$

también en la forma  $VY = \psi \quad (4.44)$

$$DV^T a = Y \quad (4.45)$$

$$V^T a = D^{-1} Y \quad (4.46)$$

Para  $V$ , ecuación (4.44) se resuelve utilizando una recursión para  $Y$

$$Y_i = \psi_i + \sum_{j=1}^{i-1} V_{ij} Y_j \quad p \geq i \geq 2 \quad (4.47)$$

condición inicial  $Y_1 = \psi_1 \quad (4.48)$

para  $Y$ , ecuación (4.47) se resuelve utilizando una recursión para  $a$

$$a_i = Y_i / d_i + \sum_{j=i+1}^p V_{ji} a_j \quad 1 \leq i \leq p-1 \quad (4.49)$$

con condición inicial  $a_p = Y_p / d_p \quad (4.50)$

El error de predicción es

$$E_p = \psi_0 + \sum_{k=1}^p Y_k^2 / d_k \quad (4.51)$$

En las páginas anteriores hemos estudiado los métodos de autocorrelación y covarianza, en general en estos métodos requerimos de una matriz de correlaciones, y una solución de ecuaciones lineales. El método a continuación, el de Rejilla se relaciona en alguna forma con estos métodos. La figura 4.2 muestra la implementación del método de rejilla.

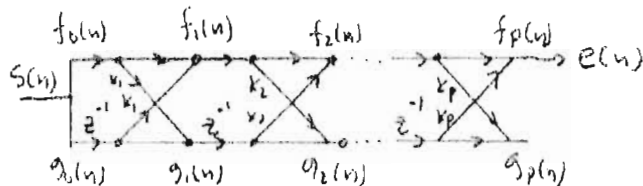


FIGURA 4.2 -- MÉTODO REJILLA

Las siguientes relaciones se mantienen

$$f_0(n) = g_0(n) = s(n) \quad (4.52)$$

$$f_m(n) = f_{m-1}(n) + k_m g_{m-1}(n-1) \quad (4.53)$$

$$g_m(n) = k_m f_{m-1}(n) + g_{m-1}(n-1) \quad (4.54)$$

\$s\_n\$ es la señal de entrada, \$f\_m(n)\$ es el residual "para adelante", y \$g\_m(n)\$ es el residual "para atrás", ambos en la etapa \$m\$, por transformación-\$z\$ tenemos

$$F_0(z) = G_0(z) = S(z) \quad (4.55)$$

$$F_m(z) = F_{m-1}(z) + k_m z^{-1} G_{m-1}(z) \quad (4.56)$$

$$G_m(z) = k_m F_{m-1}(z) + z^{-1} G_{m-1}(z) \quad (4.57)$$

Las funciones de transferencia para adelante y atrás es

$$A_m(z) = \frac{F_m(z)}{X(z)} = \frac{F_m(z)}{F_0(z)} \quad (4.58)$$

$$B_m(z) = \frac{G_m(z)}{X(z)} = \frac{G_m(z)}{G_0(z)} \quad (4.59)$$

Entonces las relaciones de recurrencia se mantienen para

$$\lambda_0(z) = B_0(z) = 1 \quad (4.60)$$

$$A_m(z) = A_{m-1}(z) + K_m z^{-1} B_{m-1}(z) \quad (4.61)$$

$$B_m(z) = K_m A_{m-1}(z) + z^{-1} B_{m-1}(z) \quad (4.62)$$

$$B_m(z) = z^{-m} A_m(z^{-1}) \quad (4.63)$$

En este método no aparecen explícitamente los parámetros  $a_k$ , los parámetros  $K_k$  se relacionan directamente con los errores de predicción para adelante y para atrás, y además pueden obtenerse sin calcular los parámetros  $a_k$ . Esta estructura es una consecuencia directa del procedimiento recursivo de Durbin y los  $K_k$  pueden obtenerse como en el método de autocorrelación. Una expresión de  $K_k$  para esta forma de rejilla es

$$K_k = - \frac{\sum_{n=0}^{M-1} f_{m-1}(n) g_{m-1}(n-1)}{\left\{ \sum_{n=0}^{M-1} (f_{m-1}(n))^2 \sum_{n=0}^{M-1} (f_{m-1}(n+1))^2 \right\}^{1/2}} \quad (4.64)$$

Que representa un coeficiente de correlación entre los errores de predicción para adelante y atrás, debido a esto los  $K_k$  son denominados también coeficientes de correlación parcial. La ecuación para  $K_k$  en el procedimiento de Durbin se puede reemplazar por (4.64), y los coeficientes predictores  $a_k$ , pueden obtenerse como antes. Cabe mencionar que el conjunto de coeficientes de correlación parcial es equivalente a un conjunto de coeficientes predictores que minimizan el error de predicción para adelante al cuadrado. Esta forma de rejilla ha motivado el estudio de otras formulaciones como son la formulación de producto interno, filtro inverso donde la solución al problema, es decir la determinación de los coeficientes predictores es por forzando el filtro  $A(z)$  a ser ortogonal a las potencias de  $Z^{-1}$  de  $Z^{-1}$  hasta  $Z^{-M}$ , y esta expresión formará las expresiones recursivas para el problema, para mayores detalles consultar Bartel y Gray (1976). Una formulación de máxima similitud se encuentra también en Bartel y Gray.

Una comparación de los tres métodos lo podemos hacer en el contexto de estabilidad del sistema resultante

$$H(z) = \frac{U}{\Lambda(z)}$$

En el método de autocorrelación, cuando los polos están dentro del círculo unitario,  $H(z)$  es garantizado ser estable, aun ue esta garantía puede no mantenerse en la práctica si la función de autocorrelación es calculada sin suficiente exactitud. En tal caso los redondeos encontrados para calcular la función de autocorrelación pueden causar una matriz mal condicionada, por lo que habrá necesidad de aplicar una ventana a la señal, de donde los polinomios resultantes permanecerán estables. Una condición necesaria y suficiente para probar estabilidad del polinomio es

$$-1 \leq K \leq 1 \quad (4.65)$$

Para el método de covarianza la estabilidad del polinomio predictor no puede ser garantizado; sin embargo en la práctica si el número de muestras de la señal es suficientemente grande, entonces el polinomio predictor resultante permanecerá estable. Esto se debe a que para un gran número de muestras, el método de covarianza y el método de autocorrelación llevan a los mismos resultados.

Para el método de rejilla el polinomio predictor es garantizado a ser estable, desde que los coeficientes son determinados de los coeficientes de correlación parcial, ya que los  $K$  satisfacen la condición de estabilidad  $-1 \leq K \leq 1$  aunque computacionalmente este método es menos eficiente para resolver las ecuaciones de LPC, (Rabiner, 1978).

### Análisis espectral.

Un importante motivo de hacer un análisis espectral sobre predicción lineal en voz, es el de determinar como el espectro del modelo o filtro, ajusta en frecuencia a la señal original. El análisis espectral se basará en el método de autocorrelación.

Considerando un conjunto de coeficientes predictoros obtenidos por el método de autocorrelación, el error de predicción  $E$  se expresa en el dominio del tiempo como

$$E_n = \sum_{m=0}^{N+P-1} e_n^2(m) \quad (4.66)$$

en el dominio de la frecuencia es

$$E_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} |S_n(e^{j\omega})|^2 |A(e^{j\omega})|^2 d\omega \quad (4.67)$$

donde  $S_n(e^{j\omega})$  es la transformada de Fourier del segmento de voz,  $S_n(m)$  y  $A(e^{j\omega})$  es

$$A(e^{j\omega}) = \sum_{k=1}^P a_k e^{-j\omega k} \quad (4.68)$$

Donde también

$$H(e^{j\omega}) = \frac{G}{A(e^{j\omega})} \quad (4.69)$$

Entonces el error puede escribirse como

$$E_n = \frac{G^2}{2\pi} \int_{-\pi}^{\pi} \frac{|S_n(e^{j\omega})|^2}{|H(e^{j\omega})|^2} d\omega \quad (4.70)$$

El filtro  $A(z)$  se denomina también filtro "blanqueador", esto lo podemos ver si comparamos las dos expresiones

$$|H(e^{j\omega})|^2 = \frac{G^2}{|A(e^{j\omega})|^2} ; |S(e^{j\omega})|^2 = \frac{|E(e^{j\omega})|^2}{|A(e^{j\omega})|^2} \quad (4.71) \quad (4.72)$$

Entonces el espectro del error  $|E(e^{j\omega})|^2$  es modelado por un espectro plano  $G^2$ , la señal  $e_n$  es blanco.

Desde que el integrando por  $E_n$  es positivo, para minimizar  $E_n$  es equivalen-

a minimizar la integral de la relación del espectro de potencia del segmento de voz a la magnitud cuadrada de la respuesta en frecuencia del sistema en el modelo. Con el método de autocorrelación la correlación de la señal de voz  $R(k)$ , y la correlación de la respuesta impulso  $\beta(k)$ , son iguales para los primeros  $P+L$  valores, así que por cuando  $P \rightarrow \infty$  las respectivas correlaciones son iguales para todos los valores  $k$  y por consiguiente:

$$\lim_{p \rightarrow \infty} |H(e^{j\omega})|^2 = |S_n(e^{j\omega})|^2 \quad (4.73)$$

Esto implica que si  $P$  es grande, podemos aproximar el espectro de la señal con pequeño error, con el modelo  $H(z)$ . Hay que notar que si la expresión para  $P \rightarrow \infty$ ,  $|H(e^{j\omega})|^2 = |S_n(e^{j\omega})|^2$ , no ocurre que  $H(e^{j\omega}) = S_n(e^{j\omega})$  es decir la respuesta en frecuencia del modelo no es igual a la transformada de Fourier de la señal. Esto es porque  $S_n(e^{j\omega})$  no necesita ser de fase mínima, mientras que  $H(e^{j\omega})$  sí es de fase mínima. El modo en que el espectro del modelo ajuste al espectro de la señal de voz, depende del error con que se contribuya en la expresión (4.70), por lo que el espectro del modelo iguala más estrechamente en las regiones de mayor energía (es decir cerca de los picos espectrales), que cerca de las regiones de baja energía (es decir los valles espectrales). Esto es entendido ya que en regiones cuando  $|S_n(e^{j\omega})| > |H(e^{j\omega})|$  contribuyen más el error total que en las regiones donde  $|S_n(e^{j\omega})| < |H(e^{j\omega})|$ ; y así el criterio de error espectral de LPC favorece un buen ajuste cerca de los picos espectrales mientras que en los valles espectrales es en ninguna parte tan bueno. Veamos un ejemplo debido a Makhoul, se compara  $20 \log_{10} |H(e^{j\omega})|$  y  $20 \log_{10} |S_n(e^{j\omega})|$ , el espectro de la señal fue obtenido por calculando con transformada de Fourier (FFT) de 20 ms de una señal de voz (muestreando a 20 Kz), pesado por una ventana de Hamming. El espectro del modelo fue calculado a (4.71) por dividiendo  $G^2$  por la magnitud cuadrada de la transformada de Fourier de la secuencia  $1, a_1, \dots, a_p$ . El sonido de voz fue la vocal /e/. el espectro de LPC fue con un predictor<sup>p</sup> de 28 polos ( $p=28$ ), obtenido por el método de autocorrelación. VER FIGURA 4.3

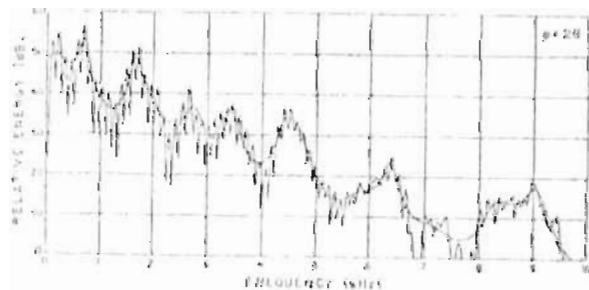


Fig. 4.3. 28-pole spectral fit to an FFT-computed signal

## SINTESIS DE VOZ DE LOS PARAMETROS PREDICTIVOS LINEALES.

Hemos obtenido hasta ahora los parámetros que caracterizan un modelo  $H(z)$ , además estudiamos su comportamiento espectral en relación a la señal original de voz. Este mismo modelo de LPC puede ser utilizado para la síntesis de voz, el sintetizador de la figura es el más simple

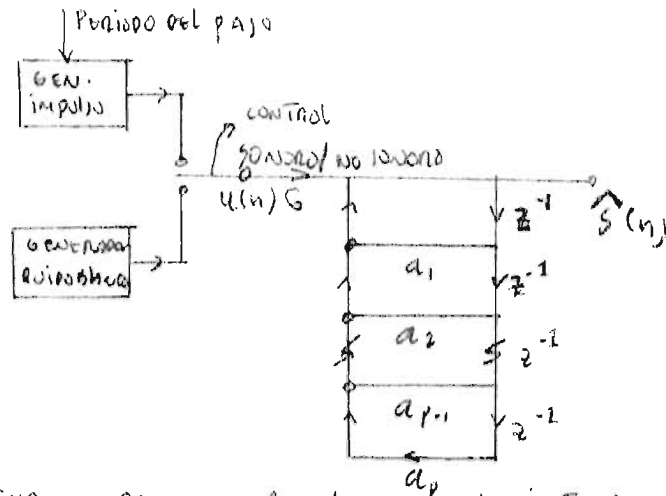


FIGURA. DIAGRAMA DE BLOQUES DEL SINTETIZADOR

Aquí los parámetros de control varían con el tiempo que se necesitan para el sintetizador son el período del paso, un switch de decisión sonoro/no sonoro, una ganancia y los  $p$  coeficientes predictores. El generador de impulsos actúa como la excitación para sonidos sonoros produciendo un pulso de amplitud unitaria en el comienzo de cada período del paso. El generador de ruido actúa como la excitación para sonidos no sonoros produciendo ruido aleatorio blanco con desviación estándar unitaria y media cero. La selección entre las dos fuentes o generadores es hecho por un control sonoro/no sonoro. El control de ganancia  $G$  determina la amplitud de la excitación. Las muestras de voz sintéticas son determinadas por

$$\hat{s}(n) = \sum_{k=1}^p a_k \hat{s}(n-k) + G u_n \quad (4.74)$$

Esta esquema utiliza un total de  $p$  multiplicadores y  $p$  sumadores para generar cada muestra de salida.



## V. FILTRADO LINEAL

En esta sección se estudiarán los filtros de Wiener y Kalman, la señal contaminada con ruido se hace pasar por estos filtros que tienden a suprimir el ruido, y así obtener un estimado de la señal. El esquema a utilizar es el de la figura 5.1

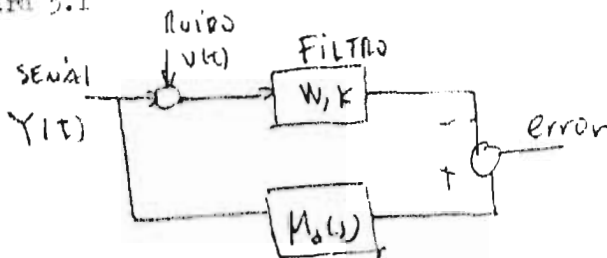


FIGURA 5.1

Para el caso de filtrado  $M_d=1$ . Desarrollemos en un principio el filtro de Wiener.

### A. FILTRO DE WIENER.

W. Wiener (1949) considera el problema de separar la señal del ruido, utilizando un caso escalar de observaciones, tanto la señal y el ruido son estacionarios (Papoulis, 1965). Para la determinación de la respuesta impulso del filtro se requiere la solución de la ecuación de Wiener-Hopf, (Kailath, 1977), aquí se utilizará respuesta en frecuencia, otros métodos se mencionan en Eykoff (1974), el filtro obtenido es lineal, invariante con el tiempo, físicamente realizable. Como en la figura 5.1, la señal o mensaje  $y(t)$  es contaminado con un ruido  $v(t)$ , donde  $y(t)$  y  $v(t)$  son procesos estacionarios con media cero, sus densidades espectrales son  $R_y(s)$  y  $R_v(s)$  y son no correlacionados. La señal  $Z(t)$ , u observación se pasa a través de un filtro lineal, invariante con el tiempo, el filtro tiene función de transferencia  $W(s)$ , denotando la salida del filtro como  $\hat{y}(t)$ , se desea obtener el filtro  $W(s)$  tal que la salida  $\hat{y}(t)$  es el mejor estimado en el sentido de varianza mínima de la señal  $y(t)$ , esto se escribe como

$$MSE = E\{\tilde{y}(t)\tilde{y}(t)\} = \text{Tr}\{E\{\tilde{y}(t)\tilde{y}^T(t)\}\} = \text{Tr}[Cov\{\tilde{y}(t)\}] \quad A.1$$

Donde  $\tilde{y}(t) = y(t) - \hat{y}(t)$  por el uso del teorema de Parseval MSE es

$$MSE = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} R_{\hat{y}}(\omega) d\omega \quad A.2$$

Lo que permite un enfoque en el dominio de la frecuencia, y por lo tanto limitado para el caso estacionario. La densidad espectral para el error es

$$R_{\hat{y}}(\omega) = [I(\omega) - W(\omega)] R_Y(\omega) [I(-\omega) - W(-\omega)]^T + W(\omega) R_V(\omega) W^T(-\omega) \quad A.3$$

sustituyendo en A.2, obtenemos

$$MSE = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \left\{ [I(\omega) - W(\omega)] R_Y(\omega) [I(-\omega) - W(-\omega)]^T + W(\omega) R_V(\omega) W^T(-\omega) \right\} d\omega \quad A.4$$

Como el problema es seleccionar una matriz  $W(s)$  tal que MSE se minimize, podemos utilizar cálculo de variaciones y haciendo

$$W(s) = W_0(s) + \epsilon \eta(s) \quad A.5$$

donde  $W_0(s)$  es la función de transferencia óptima,  $\eta(s)$  es una matriz de transferencia arbitraria, y  $\epsilon$  es un escalar. El filtro óptimo se obtiene por resolviendo

$$\left. \frac{\partial MSE}{\partial \epsilon} \right|_{\epsilon=0} = 0 \quad A.6$$

para cualquier admisible  $\eta(s)$ . Realizando A.6, obtenemos

$$\int_{-j\infty}^{j\infty} \left\{ W_0(\omega) [R_Y(\omega) + R_V(\omega)] - I(\omega) R_Y(\omega) \right\} \eta^T(-\omega) d\omega = 0 \quad A.7$$

Donde seleccionando  $W_0(s)$  tal que satisface A.7

$$W_0(s) = I(s) R_Y(s) [R_Y(s) + R_V(s)]^{-1} \quad A.8$$

Esta solución es un filtro de Wiener no realizable, es decir que  $W_0(s)$  posee polos en el lado derecho del plano-s. Esta complicación se solucionará en adelante. Notamos que podemos satisfacer con  $W(s)$  de A.6 a A.7, y es también posible satisfacer A.7 sin satisfacer A.6, donde que  $\eta(-s)$  y de  $W(s)\eta(s)$ , no son completamente arbitrarios. Para  $W(s) = \int_{-\infty}^{\infty} W_0(s)\eta(s) ds$  SEA UNA SOLUCIÓN ADMISIBLE

$W(s), W_0(s), \eta(s)$  deben ser físicamente realizables, es decir, deben tener todos sus polos en el plano izquierdo-s. Entonces es posible seleccionar  $W(s)$  que satisfaga A.7, y además sea físicamente realizable. Factorizando  $R_2(s)$  como

$$R_2(s) = R_Y(s) + R_V(s) = \Delta(s)\Delta^T(-s) \quad A.9$$

Donde la matriz  $\Delta(s)$  es tal que la función racional  $\det(\Delta(s))$  tiene todos sus polos y ceros en el plano izquierdo-s. Sustituyendo en A.7

$$\text{Tr} \left\{ \int_{-\infty}^{\infty} [W_0(s)\Delta(s) - I(s)R_Y(s)\Delta^T(-s)]\Delta^T(-s)\eta^T(-s) ds \right\} = 0 \quad A.10$$

haciendo una expansión de fracciones parciales a

$$I(s)R_Y(s)\Delta^T(-s) = A(s) + B(-s) \quad A.11$$

Donde  $A(s)$  contiene todos los términos con polos en el plano izquierdo,  $B(-s)$  contiene todos los términos con polos en el lado derecho.  $A(s)$  es la parte realizable físicamente (PR) es decir

$$A(s) = [I(s)R_Y(s)\Delta^T(-s)]_{PR} \quad A.12$$

sustituyendo A.11 en A.10

$$\text{Tr} \left\{ \int_{-\infty}^{\infty} [W_0(s)\Delta(s) - A(s)]\Delta^T(-s)\eta^T(-s) ds + \int_{-\infty}^{\infty} B(-s)\Delta^T(-s)\eta^T(-s) ds \right\} = 0 \quad A.13$$

Donde la segunda integral es cero porque todos los polos de  $B(-s)$  se encuentran en el lado derecho, por lo que la A.13 es

$$\text{Tr} \left\{ \int_{-\infty}^{\infty} [W_0(s)\Delta(s) - A(s)]\Delta^T(-s)\eta^T(-s) ds \right\} = 0 \quad A.14$$

El filtro óptimo realizable es

$$W_0(s) = A(s) \Delta^{-1}(s) \quad A.15$$

$$\text{ó} \quad W_0(s) = [I(s) R_Y(s) \Delta^{-T}(-s)]_{PR} \Delta^{-1}(s) \quad A.16$$

Que es el filtro de Wiener para estimación estacionaria multivariable. Cuando  $y(t)$  y  $v(t)$  no están no correlacionados, el filtro óptimo es

$$W_0(s) = [I(s) R_{ZY}(s) \Delta^{-T}(-s)]_{PR} \Delta^{-1}(s) \quad A.17$$

Donde  $\Delta(s) \Delta^{-T}(-s) = R_Z(s) = R_Y(s) + R_V(s) + R_{YV}(s) + R_{VY}(s) \quad A.18$

$$R_{ZY}(s) = R_Y(s) + R_{VY}(s) \quad A.19$$

para el caso escalar

$$W_0(s) = \frac{[I(s) R_{ZY}(s) / \Delta(-s)]_{PR}}{\Delta(s)} \quad A.20$$

Y CUANDO  $Y(\cdot)$  y  $V(\cdot)$  son no correlacionados

$$W_0(s) = [I(s) R_Y(s) / \Delta(-s)]_{PR} / \Delta(s) \quad A.21$$

Tambié podemos escribir la función de transferencia  $W(s)$ , cuando  $y(\cdot)$  y  $v(\cdot)$  están no correlacionados como:

$$W_0(s) = \frac{1}{\Phi_{ZZ}^+(s)} \left[ \frac{\Phi_{ZY}(s)}{\Phi_{ZZ}^-(s)} \right]_+ \quad A.22$$

Donde

$$\Phi_{ZZ}(s) = \Phi_{ZZ}^+(s) \Phi_{ZZ}^-(s)$$

$\Phi_{ZZ}^+(s)$  REPRESENTA los polos y ceros EN EL SEMI-PLANO IZQUIERDO

$\Phi_{ZZ}^-(s)$  REPRESENTA los polos y ceros EN EL SEMI-PLANO DERECHO

$$[G(s)] = [G(s)]_+ + [G(s)]_-$$

$[ \quad ]_+$  DONDE SOLAMENTE SE CONSIDERAN los polos EN el SEMIPLANO IZQUIERDO - S.

## B. FILTRO DE KALMAN.

Kalman, (1960, 1961) resuelve el problema de Wiener, y el objetivo es obtener la especificación de un sistema lineal dinámico, el cual realiza la detección, separación y predicción de una señal aleatoria. Considera los casos discreto y continuo. Aquí para el caso discreto se desarrollará basado en proyecciones ortogonales, el caso continuo se obtendrá por el método de las innovaciones debido a Kalman (1963).

### CASO DISCRETO.

El modelo del mensaje se describe por la ecuación de diferencia

$$X(k+1) = \Phi(k+1, k)X(k) + \Gamma(k)W(k) \quad \text{B.1}$$

Donde la entrada  $w$  es un proceso ruido blanco, con covarianza

$$\text{COV}\{W(k), W(j)\} = V_W(k) \delta_k(k-j) \quad \text{B.2}$$

el modelo de observación es

$$Z(k) = H(k)X(k) + V_k \quad \text{B.3}$$

el ruido en la observación  $v$  es un ruido blanco con covarianza

$$\text{COV}\{V(k), V(j)\} = V_V(k) \delta_k(k-j) \quad \text{B.4}$$

Los procesos ruido blanco  $w$  y  $v$  son no correlacionados, es decir

$$\text{COV}\{W(k), V(j)\} = 0 \quad \text{PARA TODA } j, k \quad \text{B.5}$$

el valor inicial de  $X$  es una variable aleatoria con media  $\mu_X(0)$  y varianza  $V_X(0)$

$$E\{X(0)\} = \mu_X(0) \quad \text{Var}\{X(0)\} = V_X(0) \quad \text{B.6}$$

la condición inicial  $X(0)$  está no correlacionado con  $w(k)$

$$\text{COV}\{X(0), W(k)\} = 0, \quad \text{PARA TODA } k \geq 0 \quad \text{B.7}$$

Las observaciones forman un conjunto  $Z(k) = \{z(1), z(2), \dots, z(k)\}$ , basado sobre esto deseamos obtener un estimado de  $X(j)$ ,  $\hat{X}(j|k)$ , el error de estimación es:

$$\hat{X}(j|k) = X(j) - \tilde{X}(j|k) \quad B.7$$

El estimado será no sesgado condicionalmente e incondicionalmente, es decir,  $E\{\hat{X}(j|k)|Z(k)\} = E\{X(j)|Z(k)\}$ ,  $E\{\tilde{X}(j|k)\} = E\{X(j)\}$ , y una función lineal de las observaciones. El estimador minimizará  $Var\{\tilde{X}(j|k)|Z(k)\}$

El estimado de varianza mínima lineal de  $X$  dado un conjunto de observaciones  $Z$ , es dado por la proyección ortogonal de  $X$  sobre  $Z$ , es decir,  $\hat{X} = \hat{E}\{X|Z\}$ , podrían asumirse distribuciones gaussianas, lo que haría  $\hat{E}\{X|Z\} = E\{X|Z\}$ ; en el caso no gaussiano el estimador obtenido, es solamente el mejor estimador lineal y no necesariamente el mejor estimador. Si una secuencia ortogonal  $\{\alpha_1, \dots, \alpha_m\}$  forma una base para  $Z$ , entonces  $\hat{X}$  puede ser representado por

$$\hat{X} = \hat{E}\{X|Z\} = \sum_{i=1}^m E\{X\alpha_i^T\} E\{\alpha_i\alpha_i^T\}^{-1} \alpha_i \quad B.8$$

de este mismo resultado, si  $\beta$  es ortogonal a  $Z$ , es decir  $E\{\beta_i^T \alpha_j\} = 0$  para  $i=1, \dots, m$  donde  $\{\alpha_1, \dots, \alpha_m\}$  es ortogonal para  $Z$ , entonces

$$\hat{E}\{X|Z, \beta\} = \hat{E}\{X|Z\} + \hat{E}\{X|\beta\} \quad B.9$$

Aunque estamos interesados en el estimado filtrado de  $X$ , es decir,  $\hat{X}(j|j) = \hat{E}\{X(j)|Z(j)\}$  espesaremos por predicción de una etapa. Procediendo por inducción por asumiendo que  $\hat{X}(j|j-1)$  se conoce, y entonces calcular  $\hat{X}(j+1|j)$  en términos de  $\hat{X}(j|j-1)$  y la nueva información  $z(j)$ . Sin embargo  $z(j)$  es en general no ortogonal a  $Z(j-1)$ , y antes de usar B.8, debemos determinar la porción de  $z(j)$  el cual es ortogonal a  $Z(j-1)$

1. innovación:

$$\tilde{z}(j|j-1) = z(j|j-1) = z(j) - H(j)\hat{X}(j|j-1) \quad B.10$$
 ES ORTOGONAL A  $Z(j-1)$  Y ES UN ERROR EN LAS OBSERVACIONES, Y ES LA NUEVA INFORMACION EN  $Z(j)$ .

LA SOLUCION SERA PRIMAMENTE PARA UNA ETAPA DE ESTIMACION, ENTONCES LA SEÑAL  $\hat{X}(j|j+1)$  ES ESTIMADO CON EL CONJUNTO  $Z(j-1)$  Y ADENAS DE LA NUEVA INFORMACION ESTO SE ESCRIBE POR B.10

$$\hat{x}(j+1|j) \triangleq \hat{E}\{x(j+1) | Z(j-1), z(j)\} = \hat{E}\{x(j+1) | Z(j-1), \tilde{z}(j|j-1)\}$$

ESTA ECUACION QUEDA POR LINEALIDAD (B.9)

$$\hat{x}(j+1|j) = \hat{E}\{x(j+1) | Z(j-1)\} + \hat{E}\{x(j+1) | \tilde{z}(j|j-1)\}$$

COMO  $\hat{E}\{x(j+1) | Z(j-1)\} = \hat{x}(j+1|j-1)$ , LA EXPRESION ANTERIOR ES

$$\hat{x}(j+1|j) = \hat{x}(j+1|j-1) + \hat{E}\{x(j+1) | \tilde{z}(j|j-1)\} \quad (B.11)$$

LO CUAL ES CLARO, YA QUE LA SEÑAL  $\hat{x}(\cdot)$  ES OBTENIDO POR PREDICION-DO EL VALOR DE  $x(j+1)$ , BASADO EN OBSERVACIONES PREVIAS Y SUMANDO A LA PREDICION UNA NUEVA INFORMACION.

DE B.11, CONSIDERANDO POR SEPARADO LOS TERMINOS DEL LADO DERECHO. POR INTRODUCIENDO LA ECUACION DEL MENSAJE EN  $\hat{E}\{x(j+1) | Z(j-1)\}$

$$\begin{aligned} \hat{x}(j+1|j-1) &= \hat{E}\{\Phi(j+1, j)x(j) + \Gamma(j)w(j) | Z(j-1)\} \\ &= \Phi(j+1, j) \hat{E}\{x(j) | Z(j-1)\} + \hat{E}\{\Gamma(j)w(j) | Z(j-1)\} \\ &= \Phi(j+1, j) \hat{x}(j|j-1) + \hat{E}\{\Gamma(j)w(j) | Z(j-1)\} \end{aligned}$$

COMO LA MEDIA DE  $w(\cdot)$  ES CERO

$$\hat{x}(j+1|j-1) = \Phi(j+1, j) \hat{x}(j|j-1) + 0 \quad B.12$$

UN RESULTADO QUE UTILIZAREMOS MAS TARDE SERA

$$\hat{x}(k+1|k) = \Phi(k+1, k) \hat{E}\{x(k) | Z(k)\}.$$

EL SEGUNDO SUMANDO DE B.11 QUEDA, POR EL USO DE B.8

$$\begin{aligned} \hat{E}\{x(j+1) | \tilde{z}(j|j-1)\} &= \hat{E}\{x(j+1) \tilde{z}^T(j|j-1)\} [\hat{E}\{\tilde{z}(j|j-1) \\ &\quad \times \tilde{z}^T(j|j-1)\}]^{-1} \tilde{z}(j|j-1) \quad (B.13) \end{aligned}$$

DONDE  $E\{x(j+1)\tilde{z}^T(j|j-1)} = E\{\Phi(j+1,j)x(j) + \Gamma(j)w(j)\} \tilde{z}^T(j|j-1)$   
 $\tilde{z}(j|j-1) = H(j)x(j) + v(j) - H(j)\hat{x}(j|j-1) = H(j)\tilde{x}(j|j-1) + v(j)$   
 con  $\tilde{x}(j|j-1) = x(j) - \hat{x}(j|j-1)$

$$E\{x(j+1)\tilde{z}^T(j|j-1)\} = E\{\Phi(j+1,j)x(j) + \Gamma(j)w(j)\} [H(j)\tilde{x}(j|j-1) + v(j)]^T$$

BAJADO DE LAS SUPOSICIONES SOBRE LOS RUIDOS  $w, v$

$$E\{x(j+1)\tilde{z}^T(j|j-1)\} = \Phi(j+1,j) E\{x(j)\tilde{x}^T(j|j-1)\} H^T(j) \quad B.13$$

EL TÉRMINO  $E\{x(j)\tilde{x}^T(j|j-1)\}$  ES POR  $x(j) = \tilde{x}(j|j-1) + \hat{x}(j|j-1)$

$$E\{x(j)\tilde{x}^T(j|j-1)\} = E\{\hat{x}(j|j-1) + \tilde{x}(j|j-1)\} \tilde{x}^T(j|j-1)$$

$$= E\{\hat{x}(j|j-1)\tilde{x}^T(j|j-1)\} + E\{\tilde{x}(j|j-1)\tilde{x}^T(j|j-1)\}$$

POR PROYECCIÓN ORTOGONAL

$$= 0 + E\{\tilde{x}(j|j-1)\tilde{x}^T(j|j-1)\} = \text{Var}\{\tilde{x}\}$$

$$= V_{\tilde{x}}(j|j-1)$$

ENTONCES B.14 ES

$$E\{x(j+1)\tilde{z}^T(j|j-1)\} = \Phi(j+1,j) V_{\tilde{x}}(j|j-1) H^T(j) \quad B.15$$

$$\text{PARA } E\{\tilde{z}(j|j-1)\tilde{z}^T(j|j-1)\} = E\{[H(j)\tilde{x}(j|j-1) + v(j)] [ \tilde{x}^T(j|j-1) H^T(j) + v^T(j) ]\}$$

$$= H(j) V_{\tilde{x}}(j|j-1) H^T(j) + V_v(j) \quad B.16$$

B.13 ES ENTONCES, SUSTITUYENDO B.15, B.16 Y LA INNOVACIÓN EN

$$E\{x(j+1)\tilde{z}(j|j-1)\} = \Phi(j+1,j) V_{\tilde{x}}(j|j-1) H^T(j) [H(j) V_{\tilde{x}}(j|j-1) H^T(j) + V_v(j)]^{-1} [z(j) - H(j)\hat{x}(j|j-1)]$$



$$B.11 \text{ es } \hat{x}(j+1|j) = \Phi(j+1,j) \hat{x}(j|j-1) + \Phi(j+1,j) V_{\hat{x}}(j|j-1) H^T(j) [H(j) V_{\hat{x}}(j|j-1) H^T(j) + V_w(j)]^{-1} [z(j) - H(j) \hat{x}(j|j-1)]$$

$$\text{HACIENDO } K(j+1,j) = \Phi(j+1,j) V_{\hat{x}}(j|j-1) H^T(j) [H(j) V_{\hat{x}}(j|j-1) H^T(j) + V_w(j)]^{-1} \quad (B.12)$$

$$\hat{x}(j+1|j) = \Phi(j+1,j) \hat{x}(j|j-1) + K(j+1,j) [z(j) - H(j) \hat{x}(j|j-1)] \quad (B.13)$$

$K(\cdot, \cdot)$  ES LA GANANCIA DE KALMAN. PARA UNA ETAPA DE PREDICCIÓN.

LA EXPRESIÓN ANTERIOR ES UN ALGORITMO SECUENCIAL PARA CALCULAR  $\hat{x}(j+1|j)$ , BASADO SOBRE  $\hat{x}(j|j-1)$  Y UNA NUEVA OBSERVACIÓN  $z(j)$ . DETERMINEMOS AHORA LA VARIANZA DEL ERROR. PARA OBTENER LA GANANCIA  $K(\cdot, \cdot)$ .

$$\begin{aligned} \text{COMO } \tilde{x}(j+1|j) &\triangleq x(j+1) - \hat{x}(j+1|j) = \\ &= \Phi(j+1,j) x(j) + \Gamma(j) w(j) - \Phi(j+1,j) \hat{x}(j|j-1) \\ &\quad + K(j+1,j) [z(j) - H(j) \hat{x}(j|j-1)] \\ &= [\Phi(j+1,j) \hat{x}(j|j-1) - K(j+1,j) H(j)] \tilde{x}(j|j-1) \\ &\quad + \Gamma(j) w(j) - K(j+1,j) v(j) \end{aligned}$$

$$\begin{aligned} \text{DENDE } V_{\tilde{x}}(j+1|j) &= \Phi(j+1,j) V_{\tilde{x}}(j|j-1) \Phi^T(j+1,j) + \Gamma(j) V_w(j) \Gamma^T(j) \\ &\quad - \Phi(j+1,j) V_{\tilde{x}}(j|j-1) H^T(j) K [H(j) V_{\tilde{x}}(j|j-1) H^T(j) \\ &\quad + V_w(j)]^{-1} [H(j) V_{\tilde{x}}(j|j-1) \Phi^T(j+1,j)] \quad (B.14) \end{aligned}$$

QUE LA ECUACIÓN DE VARIANZA, PARA EL CASO DISCRETO.

(B.14), (B.13) Y (B.12) DEFINEN EL PROBLEMA DE PREDICCIÓN DE UNA ETAPA, DEL FILTRO DE KALMAN.

LAS CONDICIONES INICIALES SON

$$\hat{x}(1|0) = \Phi(1,0) \hat{x}(0|0) = \Phi(1,0) \mu_{x(0)}$$

$$V_{\hat{x}}(1|0) = \text{Var}\{x(1) - \hat{x}(1|0)\} = \text{Var}\{x(1) - \Phi(1,0)\mu_{x(0)}\}$$

$$= \Phi(1,0) V_{\hat{x}}(0|0) \Phi^T(1,0) + \Gamma(0) V_w(0) \Gamma^T(0)$$

$$\hat{x}(0|0) = \mu_{x(0)}, \quad V_{\hat{x}}(0|0) = V_{x(0)}$$

PARA EL ALGORITMO SECUENCIAL DE UNA ETAPA - UNA FIGURA QUE REPRESENTA EL MODELO DE MENSAJE Y OBSERVACION, Y EL PROBLEMA DE PREDICION EN LA FIGURA B.1

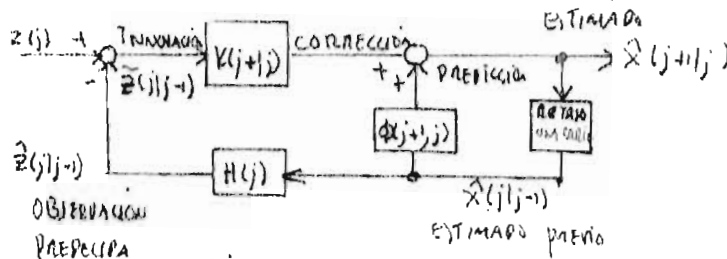
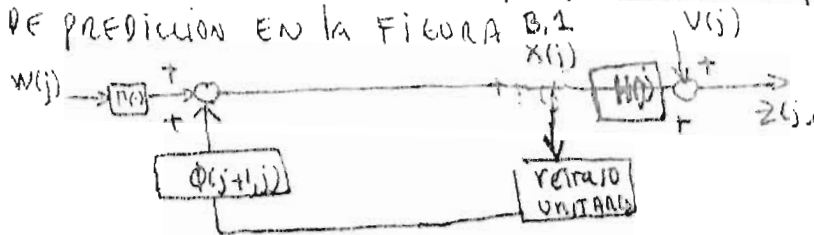


FIGURA B.1

FRECUENTEMENTE PREFERIREMOS DETERMINAR  $\hat{x}(k|k) = \hat{x}(k) = E\{x(k) | \mathcal{F}(k)\}$  A UNA ETAPA DE PREDICION  $\hat{x}(k+1|k)$ . VEMOS QUE SI EL ESTIMADO FILTRADO DE  $x(k)$ ,  $\hat{x}(k|k)$  SE CONOCE, PODEMOS OBTENER  $\hat{x}(k+1|k)$  COMO

$$\hat{x}(k+1|k) = \Phi(k+1, k) \hat{x}(k|k) \quad \text{B.20}$$

UTILIZANDO  $\hat{x}(j|j) = \hat{x}(j)$

LA EXPRESION B.20 ES  $\hat{x}(k+1|k) = \Phi(k+1, k) \hat{x}(k)$

SUSTITUYENDO EN B.18 CON  $k=j$  Y  $j-1$

$$\Phi(j+1, j) \hat{x}(j) = \Phi(j+1, j) \Phi(j, j-1) \hat{x}(j-1) + K(j+1, j) \times [z(j) - H(j) \Phi(j, j-1) \hat{x}(j-1)]$$

MULTIPLICANDO POR  $\Phi(j+1, j)^{-1} = \Phi(j, j+1)$

$$\hat{x}(j) = \Phi(j, j-1) \hat{x}(j-1) + \Phi(j, j+1) K(j+1, j) [z(j) - H(j) \Phi(j, j-1) \hat{x}(j-1)]$$

DEFINIENDO  $K(j) \triangleq \Phi(j, j+1) K(j+1, j)$

$$= V_{\hat{x}(j|j-1)} H^T(j) [H(j) V_{\hat{x}(j|j-1)} H^T(j) + V_v(j)]^{-1} \quad \text{B.21}$$

Y EL ESTIMADO ES

$$\hat{x}(j) = \Phi(j, j-1) \hat{x}(j-1) + K(j) [z(j) - H(j) \Phi(j, j-1) \hat{x}(j-1)] \quad \text{[B.22]}$$

USANDO  $\hat{x}(j|j-1) = \Phi(j+1, j) \hat{x}(j)$

B.21 ES

$$\hat{x}(j) = \hat{x}(j|j-1) + K(j) [z(j) - H(j) \hat{x}(j|j-1)] \quad \text{B.23}$$

$$\hat{x}(j) = \hat{x}(j|j-1) + K(j) \tilde{z}(j|j-1) \quad \text{[B.24]}$$

DONDE LAS ECUACIONES B.21, B.22, B.23 Y B.4, JUNTO CON B.19 DAN LA SOLUCION AL PROBLEMA DE FILTRADO Y SUS CONDICIONES INICIALES SON LAS MENCIONADAS ANTERIORMENTE. LA FORMULA DE VARIANZA  $V_{\hat{x}(j|j-1)}$  SE REFIERE COMO VARIANZA A PRIORI, ES DECIR ANTES DE LA OBSERVACION. CALCULEMOS AHORA LA VARIANZA PARA EL ERROR  $\hat{x}(j)$ , PERO LLAMADA VARIANZA A POSTERIORI.

EL ERROR  $\tilde{x} \triangleq x(j) - \hat{x}(j)$ , SUSTITUYENDO B.24

$$\begin{aligned} &= x(j) - \hat{x}(j|j-1) - K(j) \tilde{z}(j|j-1) \\ &= \tilde{x}(j|j-1) - K(j) \tilde{z}(j|j-1) \end{aligned}$$

ENTONCES 
$$\begin{aligned} V_{\tilde{x}}(j) &= E\{\left[\tilde{x}(j|j-1) - K(j) \tilde{z}(j|j-1)\right] \left[\tilde{x}(j|j-1) - K(j) \tilde{z}(j|j-1)\right]^T\} \\ &= V_{\tilde{x}}(j|j-1) + K(j) E\{\tilde{z}(j|j-1) \tilde{z}^T(j|j-1)\} K^T(j) - \\ &\quad K(j) E\{\tilde{z}(j|j-1) \tilde{x}^T(j|j-1)\} - E\{\tilde{x}(j|j-1) \tilde{z}^T(j|j-1)\} K^T(j) \end{aligned}$$

DONDE SUSTITUYENDO B.21,  $E\{\tilde{x}(j|j-1) \tilde{z}^T(j|j-1)\} = E\{x(j) \tilde{z}^T(j|j-1)\} \gamma$   
 $E\{\tilde{z}(j|j-1) \tilde{z}^T(j|j-1)\}$

$$V_{\tilde{x}}(j) = V_{\tilde{x}}(j|j-1) - V_{\tilde{x}}(j|j-1) H^T(j) \left[ H(j) V_{\tilde{x}}(j|j-1) H^T(j) + V_v(j) \right]^{-1} H(j) V_{\tilde{x}}(j|j-1)$$

USANDO K(j)

$$V_{\tilde{x}}(j) = [I - K(j) H(j)] V_{\tilde{x}}(j|j-1). \quad \text{B.25.}$$

UNA EXPRESION PARA  $V_{\tilde{x}}(j+1|j)$  PUEDE SER DESARROLLADO.  
 B.19 SE PUEDE ESCRIBIR COMO

$$V_{\tilde{x}}(j+1|j) = \Phi(j+1, j) V_{\tilde{x}}(j) \Phi^T(j+1, j) + \Gamma(j) V_w(j) \Gamma^T(j)$$

[B.26]

PODEMOS RESUMIR LAS ECUACIONES DEL FILTRO DE KALMAN, COMO SIGUE

|                       |   |
|-----------------------|---|
| MODELO DEL MENSAJE    | $x(j+1) = \Phi(j+1, j)x(j) + \Gamma(j)w(j)$   |
| MODELO DE OBSERVACION | $z(j) = H(j)x(j) + v(j)$  |
| CONDICIONES A PRIORI. | $E\{w(j)\} = E\{v(j)\} = E\{x(0)}$ $\text{COV}\{w(j), w(k)\} = V_w(j) \delta_{jk}$ $\text{COV}\{v(j), v(k)\} = V_v(j) \delta_{jk}$ $\text{COV}\{w(j), v(k)\} = \text{COV}\{x(0), w(k)\} = \text{COV}\{x(0), v(k)\} = 0$ $\text{VAR}\{x(0)\} = V_x(0)$ |
| ALGORITMO DEL FILTRO  | $\hat{x}(j) = \Phi(j, j-1)\hat{x}(j-1) + K(j)[z(j) - H(j)\Phi(j, j-1)\hat{x}(j-1)]$   |
| ECUACION DE GANANCIAS | $K(j) = V_x(j j-1)H^T(j)[H(j)V_x(j j-1)H^T(j) + V_v(j)]^{-1}$   |
| VARIANZA A PRIORI     | $V_x(j+1 j) = \Phi(j+1, j)V_x(j)\Phi^T(j+1, j) + \Gamma(j)V_w(j)\Gamma^T(j)$  |
| VARIANZA A POSTERIORI | $V_x(j) = [I - K(j)H(j)]V_x(j j-1)$   |
| CONDICIONES INICIALES | $\hat{x}(0 0) = \hat{x}(0) = \mu_x(0)$ $V_x(0 0) = V_x(0) = V_x(0)$   |

DE ESTE ALGORITMO VEMOS QUE COMO LA VARIANZA Y LAS ECUACIONES DE GANANCIAS, SON INDEPENDIENTES DE LA OBSERVACION POR LO QUE PUEDEN CALCULARSE ANTES PARA EL PROCEDIMIENTO

DE ESTIMACIÓN. LAS MATRICES  $V_X(j)$ ,  $V_X(j|j-1)$  SON SIMÉTRICAS, DE SU MAL CONDICIONAMIENTO POR ERRORES DE REDONDEO EN LA IMPLEMENTACIÓN DE LAS ECUACIONES, PUEEN AFECTAR EL PROCESO DE ESTIMACIÓN. EN LA FIGURA B.2 VEMOS UN DIAGRAMA DE BLOQUES DEL FILTRO DE KALMAN. EL DIAGRAMA ES ILUSTRATIVO Y SE PUEDE SEGUIR CON ÉL, EL PROCESO DE ESTIMACIÓN.

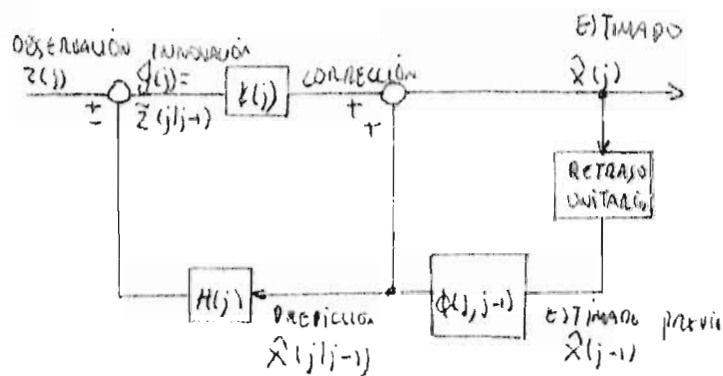


FIGURA B.2. DIAGRAMA DE BLOQUES DEL FILTRO DE KALMAN.

## CASO CONTINUO.

AQUI SE ESTUDIARA EL CASO CONTINUO, PARA EL FILTRO DE KALMAN, APLICANDO EL METODO DE LAS INNOVACIONES (KAILATH, 1968).

LA IDEA DEL METODO DE LAS INNOVACIONES ES A REMPLAZAR LAS OBSERVACIONES POR UN PROCESO QUE CONTIENE LA MISMA INFORMACION COMO EL MISMO PROCESO DE OBSERVACION, Y ASI EL PROBLEMA ES RESUELTO MAS FACILMENTE QUE EL PROBLEMA ORIGINAL.

EL MODELO O PROCESO DE OBSERVACION ES DE LA FORMA

$$Z(t) = Y(t) + V(t), \quad t \in [a, b] \quad -1$$

$V(\cdot)$  = ruido blanco con covarianza

$$E\{V(t)V^T(s)\} = \Psi_V(t)\delta(t-s), \quad \Psi_V(t) > 0$$

y media cero

$Y(\cdot)$  = señal del proceso, con media cero y covarianza finita

$$E\{Y(t)Y^T(s)\} < \infty, \quad t \in [a, b]$$

$[a, b]$  = un intervalo finito sobre la línea real

ASUMIREMOS QUE EL RUIDO "FUTURO"  $V(\cdot)$  ES NO CORRELACIONADO CON LA SEÑAL PASADA  $Y(\cdot)$ . ES DECIR

$$E\{V(t)Y^T(s)\} = 0 \quad a \leq s < t < b \quad -2$$

DESEAMOS OBTENER UN ESTIMADOR DE MÍNIMOS CUADRADOS DEL PROCESO  
X(t) DE LA FORMA

$$\hat{X}(t|b) = \text{UNA FUNCIÓN LINEAL DE LOS DATOS } \{z(s), a \leq s \leq b\}$$

$$\text{QUE MINIMIZA EL ERROR } \begin{bmatrix} Y(t) - \hat{Y}(t|b) \\ Y(t) - \hat{Y}(t|b) \end{bmatrix}^T \quad .3$$

EL ERROR DE ESTIMACIÓN ES

$$\hat{Y}(t|b) = \hat{Y}(t|b) - \hat{Y}(t|b) = Y(t) - \hat{Y}(t|b) \quad .4$$

PARA EL CASO DE FILTRADO NOS IMPORTA EL CASO CUANDO  $b = t$ .

EN ESTE MÉTODO SON DE IMPORTANCIA DOS TEOREMAS:

1. TEOREMA DE PROYECCIÓN: EL MEJOR ESTIMADO  $\hat{Y}(t|b)$  ES ÚNICO  
Y SATISFACE

$$\hat{Y}(t) \triangleq Y(t) - \hat{Y}(t|b) \perp Z(s), \quad a \leq s < b \quad \text{ES ÚNICO} \quad .5$$

DONDE  $\perp$  SIGNIFICA QUE  $E\{UV^T\} = 0$  .6

ES DECIR EL ERROR INSTANTÁNEO ES NO CORRELACIONADO CON LAS OBSERVACIONES.

TEOREMA 2 - TEOREMA DE LAS INNOVACIONES:

EL PROCESO  $V(\cdot)$  DEFINIDO COMO

$$V(t) = z(t) - \hat{Y}(t|t) = Y(t) + V(t) - \hat{Y}(t|t)$$

$$= \tilde{Y}(t) + V(t) \quad a \leq t < b \quad .7$$

Y ES DENOMINADO EL PROCESO DE INNOVACIÓN DE  $z(\cdot)$ , ES UN RUIDO BLANCO CON LA MISMA COVARIANZA QUE  $V(\cdot)$  ES DECIR

$$E\{V(t)V(s)^T\} = E\{V(t)V(s)^T\}, \quad a \leq t, s < b \quad .8$$



$z(\cdot)$  y  $v(\cdot)$  SON "EQUIVALENTES", ES DECIR TIENEN LA MISMA INFORMACIÓN ESTADÍSTICA.

ENTONCES TENEMOS QUE LA OBSERVACION ES

$$z(s) = \gamma(s) + v(s), \quad a \leq s \leq T \quad .9$$

DESEAMOS OBTENER EL ESTIMADO  $\hat{x}(t|t)$  POR MÍNIMOS CUADRADOS, EN LA VARIABLE  $x(t)$ .

DE .7 LA INNOVACION ES  $v(t) = z(t) - \hat{\gamma}(t|t)$ ,  $E\{v(t)v^T(s)\} = \Psi_v(t,s)$  COMO  $v(\cdot)$  Y  $z(\cdot)$  SON EQUIVALENTES

$$\hat{x}(t|t) = \int_a^T g(t,s) v(s) ds \quad .10$$

EL FILTRO  $g(t, \cdot)$  ES TAL QUE, ~~ES~~

$$E\{\tilde{x}(t|t)v(s)\} = 0 = E\{x(t) - \hat{x}(t|t)\}v^T(s) = 0 \quad .11$$

DE .10 Y .11 TENEMOS

$$E\{x(t)v^T(s)\} = E\{\hat{x}(t|t)v^T(s)\} \quad .12$$

$$\begin{aligned} E\{x(t)v^T(s)\} &= E\left\{\int_a^T g(t,\sigma)v(\sigma) d\sigma v^T(s)\right\} = \int_a^T g(t,\sigma) E\{v(\sigma)v^T(s)\} d\sigma \\ &= \int_a^T g(t,\sigma) R(\sigma) \delta(\sigma-s) d\sigma = g(t,s) R(s) \quad .13 \\ &\quad a \leq s \leq T. \end{aligned}$$

$$\text{como } g(t,s) = E\{x(t)v^T(s)\} R^{-1}(s)$$

POREMOS ESCRIBIR EL ESTIMADO A PARTIR DE ESTOS RESULTADOS

$$\hat{x}(t|t) = \int_a^T E\{x(t)v^T(s)\} R^{-1}(s) v(s) ds \quad .14$$

Y OBTENEMOS EL ESTIMADO POR MÍNIMOS CUADRADOS DE UNA SEÑAL  $X(t)$

SI  $X(t)$  SATISFACE LA ECUACION DIFERENCIAL

$$\dot{X}(t) = F(t) X(t) + G W(t), \quad t \geq a, \quad X(t_0) = X_a(-15)$$

$W(t)$  ES UN PROCESO DE RUIDO BLANCO CON COVARIANZA

$$E\{W(t)W^T(s)\} = V_w(t) \delta(t-s)$$

QUE ESTA NO CORRELACIONADO CON  $V(\cdot)$ , ES DECIR

$$E\{W(t)V(s)\} = 0$$

ADemás

$$E\{X_a(t)\} = 0 \quad E\{X_a X_a^T\} = P_a$$

$$E\{W(s)X_a^T\} = 0 \quad a \leq s \leq b$$

PODEMOS DERIVAR EL ESTIMADO, PARA SUSTITUIR LA ECUACION DE ESTADO.

UTILIZANDO LA REGLA DE LEIBNITZ, PARA DERIVAR FUNCIONES DE LA FORMA

$$\frac{\partial}{\partial t} \int_{\alpha(t)}^{\beta(t)} f(t, \tau) d\tau = \int_{\alpha(t)}^{\beta(t)} \frac{\partial f(t, \tau)}{\partial t} d\tau - f(t, \alpha) \frac{d\alpha(t)}{dt} + f(t, \beta) \frac{d\beta(t)}{dt}$$

DONDE  $f(t, \tau)$ ,  $\alpha(t)$ ,  $\beta(t)$  SON CONTINUOS Y DIFERENCIABLES CON RESPECTO A  $t$ .

DERIVAMOS (14)

$$\dot{\hat{X}}(t) = \int_a^t \frac{d}{dt} E\{X(t)W^T(s)\} R^{-1}(t) V(s) ds - 0 + E\{X(t)W^T(s)\} R^{-1}(t) V(t) \cdot 1 = \quad \cdot 16$$

SUSTITUYENDO 15 EN 16

$$\begin{aligned}\hat{\dot{X}}(t) &= E\{X(t)V'(t)\}R^{-1}(t)V(t) + \int_0^t E\{[F(\tau)X(\tau) + W(\tau)][V'(\tau)]\}R^{-1}(\tau)V(\tau)d\tau \\ &= E\{X(t)V'(t)\}R^{-1}(t)V(t) + [F(t) \int_0^t E\{X(\tau)V'(\tau)\}R^{-1}(\tau)V(\tau)d\tau \\ &\quad + \int_0^t E\{G(\tau)W(\tau)V'(\tau)\}R^{-1}(\tau)V(\tau)d\tau],\end{aligned}$$

como  $V(\tau)$  es independiente de  $X$ ,  $E\{W(\tau)V'(\tau)\} = 0$

$$\hat{\dot{X}}(t) = E\{X(t)V'(t)\}R^{-1}(t)V(t) + F(t)\hat{X}(t) \quad \cdot 17$$

HACIENDO  $K \triangleq E\{X(t)V'(t)\}R^{-1}(t)$

$$\hat{\dot{X}}(t) = F(t)\hat{X}(t) + K(t)V(t) \quad \cdot 18$$

$$V(t) = Z(t) - \hat{Y}(t) \quad \cdot 19$$

PARA CONCILIAR CON KALMAN,  $Y(t) = H(t)X(t)$ , ENTONCES

$$\hat{Y}(t) = H(t)\hat{X}(t) \quad \cdot 20$$

$$V(t) = H(t)X(t) + v(t) - H(t)\hat{X}(t) = H(t)[\tilde{X}(t)] + v(t)$$

ENTONCES  $K(t) = E$

$$\begin{aligned}K(t) &= E\{X(t)V'(t)\}R^{-1}(t) \\ &= E\{X(t)[\tilde{X}^T(t)H^T(t) + V^T(t)]\}R^{-1}(t) \\ &= E\{[\tilde{X}(t) + \hat{X}(t)][\tilde{X}^T(t)H^T(t) + V^T(t)]\}R^{-1}(t) + 0 \\ &= E\{\tilde{X}(t)\tilde{X}^T(t)\}H^T(t)R^{-1}(t) + 0 + 0 \\ &= P(t)H^T(t)R^{-1}(t)\end{aligned}$$

DONDE  $P(t) = E\{\tilde{X}(t)\tilde{X}^T(t)\}$  · 21

PARA DETERMINAR LA GANANCIA  $K(\cdot)$ , DEBEMOS RESOLVER TAMBIÉN PARA  $P(\cdot, \cdot)$ .

OBTENEMOS PRIMERAMENTE UNA EXPRESIÓN PARA EL ERROR.

$$\begin{aligned} \text{COMO } \tilde{x}(t|t) &= x(t) - \hat{x}(t|t), \quad \dot{\tilde{x}}(t|t) = \dot{x}(t) - \dot{\hat{x}}(t|t) \\ \dot{\tilde{x}}(t|t) &= F(t)x(t) + Gw(t) - [F(t)\hat{x}(t|t) + K(t)v(t)] \\ &= F(t)\tilde{x}(t|t) - K(t)v(t) + Gw(t) \end{aligned} \quad .22$$

$$\begin{aligned} \text{COMO } v(t) &= z(t) - \hat{y}(t|t) = Y(t)x(t) + v(t) - \hat{y}(t|t) \\ &= \tilde{y}(t|t) + v(t) \\ &= H(t)\tilde{x}(t|t) + v(t) \end{aligned} \quad .23$$

$$\begin{aligned} \dot{\tilde{x}}(t|t) &= F(t)\tilde{x}(t|t) - K(t)[H(t)\tilde{x}(t|t) + v(t)] + G(t)w(t) \\ &= [F(t) - K(t)H(t)]\tilde{x}(t|t) - K(t)v(t) + G(t)w(t) \end{aligned} \quad .24$$

$$E\{\tilde{x}(a|a)\} = X_a$$

OBTENIENDO UNA EXPRESIÓN GENERAL PARA UNA VARIANZA DE  $\tilde{x}(\cdot)$  Y LUEGO UTILIZARLO PARA  $\hat{x}(t|t)$  TENEMOS

$$\begin{aligned} \dot{x}(t) &= F(t)x(t) + G(t)w(t) \quad x(a) = X_a \quad t > a \\ E\{w(t)\} &= 0, \quad E\{w(t)w^T(s)\} = 0 \quad t \geq a. \end{aligned}$$

UNA SOLUCIÓN PARA  $x(t)$  ES

$$x(t) = \Phi(t, a)x(a) + \int_a^t \Phi(t, s)G(s)w(s)ds$$

$\Phi(\cdot, \cdot)$  TIENE LA PROPIEDAD

$$\dot{\Phi}(t, s) = F(t)\Phi(t, s)$$

$$\Phi(a, a) = I, \quad a \leq t$$

$$\begin{aligned} E\{x(t)x^T(t)\} &= P(t, t) = E\left\{ \left[ \Phi(t, a)x(a) + \int_a^t \Phi(t, s)G(s)w(s)ds \right] \right. \\ &\quad \left. \cdot \left[ x^T(a)\Phi^T(t, a) + \int_a^t w^T(s)G^T\Phi^T(t, s)ds \right] \right\} \end{aligned}$$

Y  $\dot{P}(t)$  es

$$\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + G\Psi_w G^T.$$

UTILIZANDO ESTA ECUACION PARA DETERMINAR LA VARIANZA DEL ERROR

PARA  $\hat{X}(t) = [F(t) - K(t)H(t)]\hat{X}(t) - K(t)V(t) + GwG^T$

HACIENDO  $\bar{F}(t) = F(t) - K(t)H(t)$

$$\bar{W}(t) = -K(t)V(t) + GwG^T$$

TENEMOS LA EXPRESION PARA  $\dot{P}(t)$  DEL ERROR.

$$\dot{P}(t) = \bar{F}(t)P(t) + P(t)\bar{F}^T(t) + G\Psi_v G^T - P(t)H^T(t)\Psi_v^{-1}(t)H(t)P(t).$$

EXISTEN OTROS ENFOQUES PARA LA SOLUCION DEL FILTRO DE WIENER Y KALMAN. PARA EL FILTRO DE WIENER EN EL CASO DISCRETO SE PUEDE MENCIONAR EL DESARROLLO EN KUD [1963]. PARA EL FILTRO DE KALMAN, EN EL CASO DISCRETO, MENCIONEMOS LA SOLUCION POR MAXIMO A POSTERIORI [SAGE, CHELSEA WHITE, 1977], UN METODO PURAMENTE ALGEBRAICO ES PROPUESTO POR I. A GURA [1969]. PARA

EL CASO CONTINUO ADEMAS DE LAS INNOVACIONES, HAY UNA SOLUCION POR CALCULO DE VARIACIONES (ATHANI TSE, 1967).

ALGUNAS CARACTERISTICAS QUE PODEMOS ENCONTRAR EN LOS FILTROS LO DAMOS EN FORMA DE TABLA.

## CARACTERÍSTICAS DE LOS FILTROS DE WIENER Y KALMAN.

|                          | <u>FILTRO DE WIENER</u>  | <u>FILTRO DE KALMAN.</u>  |
|--------------------------|--|---|
| CONOCIMIENTO A PRIORI:   | La señal y el ruido se dan en términos de funciones de correlación o espectro de potencia. | La señal y el ruido se dan en términos de ecuación de estado.                                     |
| ESTACIONALIDAD:          | Aplicable a procesos estacionarios por su solución de respuesta en frecuencia.             | Se puede extender a procesos no-estacionarios o arrancar las observaciones en cualquier instante. |
| CÁLCULO COMPUTACIONAL:   | NO FACILMENTE.   | POR SU FORMA RECURSIVA SE ADAPTA FACILMENTE A UN MÉTODO COMPUTACIONAL DE SOLUCIÓN.                |
| SÍNTESIS:                | COMO SE DETERMINA $H(s)$ , LA SÍNTESIS RESULTA LABORIOSA.                                  | EL ALGORITMO DEL FILTRO ES UNA DESCRIPCIÓN DIRECTA DE LA REALIZACIÓN DEL FILTRO.                  |
| PROCESOS MULTIVARIABLES: | se puede adaptar pero requiere mucho cálculo computacional.                                | Se puede adaptar fácilmente.  |

En las páginas anteriores hemos desarrollado los algoritmos del filtro de Wiener y Kalman, para resolver el problema de la estimación. El filtro de Wiener es resuelto para el caso estacionario, para una equivalencia con el filtro de Kalman podemos obtener una versión estacionaria del filtro de Kalman. Antes de pasar a este problema veamos unos resultados obtenidos por Kalman (1961), concerniente a una solución degenerada de la ecuación de Riccati o de varianza y la estabilidad del filtro obtenido.

Assumiendo el modelo del mensaje para el caso continuo

$$\dot{X}(t) = F X(t) + G W(t)$$

$$Z(t) = H X(t) + V(t) \quad (0.1)$$

con

$$\begin{aligned} E\{W(t)\} &= E\{V(t)\} = E\{W(t), V(t)\} = 0 \\ E\{W(t), W(t)^T\} &= \delta(t-\tau), \quad E\{V(t), V(t)^T\} = \delta(t-\tau) \end{aligned}$$

con  $F, G, H$  matrices constantes

además

La ecuación de varianza para el caso continuo fue

completamente controlable  $R [G, FG, F^2G, \dots, F^{n-1}G] = n$   
 completamente observable  $R [H^T, F^T H^T, \dots, (F^T)^{n-1} H^T] = n$

$$\dot{V}_X(t) = F(t) V_X(t) + V_X(t) F^T(t) - V_X(t) H^T(t) \Psi_V^{-1} H(t) V_X(t) + G(t) \Psi_W(t) G^T(t) \quad (0.2)$$

para cualquier tiempo inicial finito  $t_0$  y una matriz positiva definida  $V_X(t_0) = V_X^0$ , (0.2) tiene una solución única

$$V_X(t) = V_X(t) | V_X^0, t_0 \quad (0.3)$$

para toda  $t$ , tal que  $|t-t_0|$  es suficientemente pequeño. La existencia y unicidad de la solución para  $|t-t_0|$  suficientemente pequeña sigue directamente del efecto de que la ecuación (0.2) satisface una condición de Lipschitz (Kalman y Bertram, 1960).

(0.2) es no lineal, la existencia de una solución para  $|t-t_0|$  suficientemente pequeña no garantiza que una solución exista para toda  $t$ , (Kalman, 1963), muestra que la solución de (0.2) existe para toda  $t \geq t_0 > -\infty$  y es determinada únicamente por la selección de  $V_X^0$ . por lo tanto, para todos los problemas en el cual la observación es finita, estamos asegurando de la existencia y unicidad de la solución de la ecuación de varianza, y por consiguiente la ganancia  $K$ , por la especificación de  $V_X^0$ . hasta aquí no se han puesto restricciones

sobre el modelo del mensaje y de la observación. Para tratar el problema de la observación infinita  $t_0 = -\infty$  es necesario restringir ligeramente ya sea el mensaje o las observaciones. Por ejemplo si  $H(t) = D$  entonces  $V_{\hat{X}}(t) = V_X(t)$ , la varianza de  $\hat{X}$  es el mismo de  $X$  para toda  $t$ , y si el modelo del mensaje es inestable entonces  $V_X(t)$  y de aquí  $V_{\hat{X}}(t)$  será no acotada cuando  $t_0 \rightarrow -\infty$ . La solución en el intervalo infinito existirá en el sentido que

$$\lim_{t_0 \rightarrow \infty} V_{\hat{X}}(t | 0, t_0) = V_{\hat{X}}^* \quad (C.4)$$

existe para toda  $t$  y es una solución de C.2, si uno u otro el modelo del mensaje es asintóticamente uniformemente estable o el sistema es completamente observable.

Por el requerimiento de que el modelo del mensaje sea asintóticamente uniformemente estable, aseguramos que  $V_X(t)$  permanezca acotada y entonces el problema de que  $V_{\hat{X}}(t)$  crezca no se presenta. Si el sistema es completamente observable entonces podemos estimar  $X$  con error finito, y  $V_{\hat{X}}(t)$  permanece finito. La solución  $V_{\hat{X}}^*(t)$  es constante y es un estado de equilibrio de C.2, esto por igualando a cero el lado derecho de C.2, y resolviendo la ecuación cuadrática resultante; por consiguiente las restricciones sobre estabilidad u observabilidad garantizan la existencia de una solución del problema del filtro de Wiener. La existencia de una solución de estado permanente para intervalos de observación infinita como en C.4 es una forma de estabilidad que el algoritmo del filtro puede poseer. Si arrancamos en  $t_0 = -\infty$  con  $V_{\hat{X}}^0 = 0$  la ecuación de varianza tiene una solución para toda  $t$ ; lo mismo sucede si arrancamos con otra varianza inicial, la solución es dada por  $V_{\hat{X}}^*(t)$ . Para establecer que una solución estado permanente  $V_{\hat{X}}^*(t)$  dado la ecuación C.4 es una trayectoria estable, en el sentido de que una perturbación de la trayectoria decae a cero cuando  $t \rightarrow \infty$ , se hace necesario hacer alguna restricción más severa sobre el modelo del mensaje y la observación.

Si el mensaje y la observación son completamente observable y completamente controlable, si  $\Psi_v^{-1}(t)$ ,  $\Psi_v(t)$ ,  $G(t)$ ,  $\Psi_w(t)$ ,  $G^T(t)$  son positivos semidefinidos respectivamente, además acotadas para toda  $t$ , entonces el filtro óptimo es uniformemente asintóticamente estable. En suma para cada matriz definida positiva  $V_{\hat{X}}^0$  la solución  $V_{\hat{X}}(t) = V_{\hat{X}}(t | V_{\hat{X}}^0, t_0)$  converge a la solución estado permanente  $V_{\hat{X}}^*(t)$  dado por C.4, cuando  $t \rightarrow \infty$ .

Ahora una versión estacionaria del Filtro de Kalman es haciendo cero la ecuación de varianza A.2 es decir

$$\begin{aligned} V_{\hat{X}}(0) &= V_{\hat{X}}(0) = \text{CONSTANTE} \\ 0 &= F V_{\hat{X}}(0) + V_{\hat{X}}(0) F^T - V_{\hat{X}}(0) H^T \Psi_v^{-1} H V_{\hat{X}}(0) + G \Psi_w G^T \quad (C.5) \end{aligned}$$



la ganancia es  $K(0) = V \hat{x}(0) H^T \Psi_v^{-1}$  C.6

el filtro  $\dot{\hat{x}}(t) = F \hat{x}(t) + K(0) [z(t) - H \hat{x}(t)]$  C.7  
 $\hat{x}(-\infty) = 0$

La ecuación de varianza como en A.5 se denomina Matriz de Riccati degenerada, una expresión algebraica cuadrática donde la correcta solución es positiva definida.

Una manera de ver la relación del filtro de Wiener y Kalman en el contexto de procesos estacionarios es como lo hacen Sage y Melsa (1977), se especifica el modelo de observaciones y mensaje en la forma de Kalman y se encuentra la equivalencia espectral necesaria para el enfoque de Wiener. El modelo del mensaje y la observación es como en C.1 y C.2. Asumiendo que C.1 es asintóticamente estable y controlable,

La densidad espectral de  $y(t)$  en forma de Wiener es

$$R_y(s) = H R(s) G \Psi_w G^T \mathcal{L}(-s) H^T \quad C.8$$

$$\mathcal{L}(s) = (sI - F)^{-1}$$

$\Psi_w$  es definida positiva:

$$\Psi_{vv} = NN \quad C.9$$

con  $N$ , matriz positiva definida simétrica. Con esto el filtro de Kalman es formulado ahora como un problema de Wiener multivariable, y por lo tanto podemos aplicar lo mismo estudiado anteriormente. La matriz espectral  $R_z(s)$  es

$$R_z(s) = R_y(s) + R_v(s) = N^2 + H R(s) G \Psi_w G^T \mathcal{L}(-s) H^T \quad C.10$$

factorizando en la forma

$$R_z(s) = \Delta(s) \Delta^T(-s), \quad \Delta(s) = N + H R(s) K N \quad C.11$$

$$K = R H^T \Psi_v^{-1}$$

y  $R$  es una solución simétrica positiva definida de la ecuación de Riccati degenerada

$$FR + RF^T - K \Psi_v K^T + G \Psi_w G^T = 0 \quad C.12$$

para esta ecuación se garantiza una solución positiva definida pues C.1 es estable. Sustituyendo en la ecuación del filtro óptimo de Wiener

$$W_0(s) = H (sI - F + KH)^{-1} K \quad C.13$$

Vemos que los filtros son equivalentes para el problema de filtrado estacionario, pues de

$$c.11 \text{ y } c.6$$

$$K(0) = K$$

$$c.12 \text{ y } c.8$$

$$V_{\hat{x}}(0) = R$$

Rhodes(1971), llega a un resultado interesante, aprovechando la dualidad existente entre la estimación y control óptimo. Examina la propiedad del filtro de Kalman y su asociada ecuación de varianza en estado permanente por reduciendo a la solución conocida del problema del regulador óptimo y muestra que el problema de Kalman viene a ser idéntico al de Wiener, y el filtro de Kalman es el filtro óptimo realizable de Wiener.

Hasta aquí hemos encontrado una correspondencia entre el filtro de Kalman y el filtro de Wiener en el sentido estacionario, desde que las matrices de covarianza de los dos filtros son equivalentes, entonces la diferencia en el comportamiento de los dos filtros reside en las componentes transitorias de las varianzas del de los filtros de Kalman y Wiener, en base a esto estudiamos los aspectos más relevantes del trabajo de R.A. Singer y Frost relacionado al desempeño "relativo" de los dos filtros. Consideremos el caso continuo, el sistema del mensaje y la observación serán

$$\begin{aligned} \dot{x}(t) &= Fx(t) + Gu(t) \\ y(t) &= Hx(t) + v(t) \end{aligned} \quad \delta$$

$$\begin{aligned} E\{u(t)\} &= E\{v(t)\} = E\{u(t)v(t)\} = 0 \\ E\{u(t)u^T(s)\} &= Q\delta(t-s), \quad E\{v(t)v^T(s)\} = R\delta(t-s) \end{aligned}$$

Donde  $R$  es positiva definida, el estado inicial es independiente del proceso  $u(t)$  y  $v(t)$ , para  $t \geq 0$  y con media cero. Los valores característicos de  $F$  tienen partes reales negativas, así que  $\delta$  tiende a la estacionariedad cuando  $t \rightarrow \infty$

Las ecuaciones de Kalman para este sistema es:

$\dot{\hat{x}}_k(t) = [F - P_k H^T R^{-1} H] \hat{x}_k(t) + P_k(t) H^T R^{-1} y(t), \quad \hat{x}_k(0) = 0$   
 donde  $\hat{x}_k$  es el estimado de Kalman de  $x(t)$ ,  $P_k$  es la matriz de varianza del error, y que satisface la ecuación de Riccati

$$\begin{aligned} \dot{P}_k(t) &= F P_k(t) + P_k(t) F^T + G Q G^T - P_k(t) H^T R^{-1} H P_k(t) \\ P_k(0) &= P_0 \end{aligned}$$

La matriz  $P_0$  satisface la ecuación de Riccati en estado permanente

$$0 = F P_{\infty} + P_{\infty} F^T + G Q G^T - P_{\infty} A^T R^{-1} A P_{\infty}$$

El filtro de Wiener de coeficientes constantes, el cual es el límite en estado permanente del filtro de Kalman es

$$\dot{\hat{X}}_w(t) = [F - P_{\infty} H^T R^{-1} H] \hat{X}_w(t) + P_{\infty} H^T R^{-1} Y(t), \quad \hat{X}(0) = 0$$

donde  $\hat{X}_w$  es el estimado de Wiener de  $X(t)$ . La matriz de varianza del error de Wiener es

$$\dot{P}_w(t) = F'' P_w(t) + P_w F''^T + G Q G^T + P_{\infty} H^T R^{-1} H P_{\infty}$$

con  $F'' = F - P_{\infty} H^T R^{-1} H$ .

Desde que las matrices de varianza en estado permanente de ambos filtros son idénticas, la diferencia entre los dos filtros se encuentra en las componentes transitorias de  $P_k$  y  $P_w$ . Estos componentes transitorios son definidos por

$$\tilde{P}_k(t) = P_k(t) - P_{\infty}, \quad \tilde{P}_w(t) = P_w(t) - P_{\infty}$$

Definiendo  $M_0 = P_0 - P_{\infty}$  se obtiene ecuaciones diferenciales sobre los transitorios

$$\dot{\tilde{P}}_k(t) = F'' \tilde{P}_k(t) + \tilde{P}_k(t) F''^T - \tilde{P}_k(t) H^T R^{-1} H \tilde{P}_k(t), \quad \tilde{P}_k(0) = M_0$$

$$\dot{\tilde{P}}_w(t) = F'' \tilde{P}_w(t) + \tilde{P}_w(t) F''^T, \quad \tilde{P}_w(0) = M_0$$

se asume que con  $M_0 > 0$ , implica que filtrando se mejora el estimado de cada una de las variables.

Una cantidad obtenida por Singer y Frost es

$$P_w(t) - P_k(t) = e^{F''^T t} M_0 A(0, t) [M_0^{-1} + A(0, t)]^{-1} e^{F'' t}, \quad A(0, t) = \int_0^t e^{F''^T (t-\tau)} H^T R^{-1} H e^{F'' \tau} d\tau$$

que significa una diferencia entre las matrices de varianza de los filtros de Wiener y Kalman y se <sup>usa</sup> para medir el mejoramiento en el desempeño logrado por usando Kalman antes que Wiener. Entonces se requieren acotaciones inferiores y superiores sobre las matrices de varianza de ambos filtros, y que nos darán una medida de la relativa exactitud de estos filtros. La primera cota Superior formulada es el siguiente

$$\|P_w(t) - P_k(t)\| \leq \frac{\|P_0 - P_{\infty}\|^2 \|H^T R^{-1} H\| e^{2\alpha_{\max} t}}{2|\alpha_{\max}|} (1 - e^{-2\alpha_{\min} t})$$

donde  $\alpha_{\max}$  es la máxima parte real de los valores característicos de  $F''$ , o la recíproca negativa de la más grande constante de tiempo del filtro de Wiener.

Esta desigualdad indica que cuatro cantidades determinan el desempeño relativo del filtro de Kalman y Wiener y son: Diferencia entre la varianza inicial y de estado permanente  $\|P_0 - P_\infty\|$ , la covarianza del ruido, la más lenta constante de tiempo del filtro de Wiener ( $= 1/\alpha_{\min}$ ) y la longitud del intervalo de filtrado. La dependencia del tiempo en la cota ilustra que para periodos de filtrado ya sea más grande o más pequeño que la constante de tiempo (más lenta),  $\tau$  del filtro de Wiener, poca mejora se obtiene por el uso de filtro de Kalman. El producto de la dependencia del tiempo alcanza un máximo cuando  $\tau = 1/2 \alpha_{\max}^{-1} \tau_{\max}$  y el máximo valor de la cota superior es  $\frac{1}{4} \|P_0 - P_\infty\| \|H^T Q^{-1} H\| \tau$ . Cuando el intervalo de filtrado es tal que la cota no es despreciable, este producto debe examinarse antes de seleccionar un filtro. La cota es más grande cuando la exactitud de la estimación en estado permanente es considerablemente mejor que el estimado en tiempo inicial, es decir  $\|P_0 - P_\infty\|$  es grande cuando la constante de tiempo del filtro de Wiener  $\tau$  es grande (así que la convergencia del error en el filtro de Wiener es lenta), y cuando la intensidad del ruido de observación es pequeño. Cuando estas condiciones se satisfacen el filtro de Kalman provee una significativa mejora en la exactitud de la estimación del intervalo de filtrado. En otro caso la cota superior no puede ser grande comparado a  $\|P_k(t)\|$  y así el filtro de Kalman proveerá una mejora ligera sobre el filtro de Wiener.

Otra cota derivada de la anterior es

$$\|P_w(t) - P_k(t)\| \leq \|P_0 - P_\infty\| e^{\alpha_{\max} t} \cdot \min \left\{ 1, \frac{\|P_0 - P_\infty\| \cdot \|H^T Q^{-1} H\| (1 - e^{-2\alpha_{\min} t})}{2|\alpha_{\max}|} \right\}$$

que establece un límite numérico sobre la mejora en el desempeño que puede obtenerse por usando Kalman antes que Wiener, vemos que con  $\tau=0$ ,  $\tau=\infty$  esta cota se reduce a cero, pues los dos filtros tienen idéntica varianza.

Una cota inferior es

$$c \| \tilde{P}_w(t) \| \leq \| \tilde{P}_k(t) \| \leq \| \tilde{P}_w(t) \|$$

donde  $c = \frac{1}{1 + \|u_0\| \|H^T Q^{-1} H\| [e^{\alpha_{\max} \alpha_{\min} t} - 1]} \frac{2\alpha_{\min} t}{2|\alpha_{\max}|}$

$\alpha_{\min} = \epsilon$  la parte real más negativa de los valores característicos del filtro de Wiener.

esta cota es para  $pe$  uña  $t$ , el período durante el cual el transitorio es aún significativo. cuando  $C=1$ , ambos filtros tienen esencialmente el mismo desempeño. Esta cota es influenciada por las mismas cantidades que determinan a el desempeño relativo de ambos filtros, que las dos anteriores. Otra cota inferior es

$$\|P_0\| \geq \xi_2 [\sqrt{1 + \xi_1 \xi_2} - 1] = \sqrt{\xi_1 \xi_2} \cdot \sqrt{\frac{\sqrt{1 + \xi_1 \xi_2} - 1}{1 + \xi_1 \xi_2 + 1}}$$

Esta relaciones indican que a más grande intensidad del ruido, más grande es el error de filtrado en estado permanente. Si interpretamos  $\xi_1 \xi_2$  como el producto de los ruidos de la observación y el estado,  $\xi_1 \xi_2$  como la relación de los ruidos, el efecto del producto  $\xi_1 \xi_2$  es más significativo que la relación  $\xi_1 \xi_2$ . Si el producto es constante, pero la relación de ruido se incrementa el error de filtrado en estado permanente se incrementa. Si la relación de ruido es constante, pero el producto del ruido se incrementa, vemos que el error en estado permanente se incrementa. Estos resultados indican que el ruido de excitación del estado tiene más influencia que el ruido de la observación.

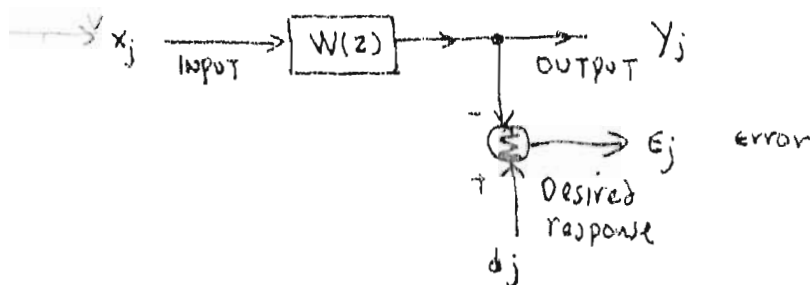
Un ejemplo de aplicación se resuelve en el artículo original, también para mayores detalles consultar R.A. Singer y Frost (1969).

## APLICACIONES.

LA OBTENCIÓN DE LOS ALGORITMOS DEL FILTRO DE WIENER Y KALMAN SE VAN DESARROLLANDO EN LAS PÁGINAS ANTERIORES. NOS RESTA: ENUNCIAR APLICACIONES. ESTO HAREMOS EN SIGUIENTE

UNA APLICACIÓN PARALELA AL FILTRO DE WIENER ES EL TRABAJO DE R. JAFFE AND E. RECHTIN (1955), EN LO RELATIVO A DISEÑO DE CIRCUITOS DE PHASE-LOCK. ESTOS CIRCUITOS SE UTILIZAN PARA LA DETECCIÓN Y RASTREO DE SEÑALES DE BANDA ANGOSTA EN LA PRESENCIA DE RUIDO DE BANDA ANCHA. CONSIDERACIONES DE DESCRIPCIÓN DEL CIRCUITO, ASPECTOS DE DISEÑO Y OTRAS CARACTERÍSTICAS ADECUADAS DE RESULTADO, ES RECOMENDABLE RECORRER AL ARTÍCULO ORIGINAL. PODEMOS MENCIONAR SOLAMENTE QUE EN LO RELACIONADO A LA OBTENCIÓN DEL FILTRO OPTIMO, SE HACE UN DESARROLLO SIMILAR A WIENER (DE VIMOS EN LA SECCIÓN DE FILTRO DE WIENER, ES DEBIDO POR FACTORIZACIÓN ESPECTRAL O RESPUESTA EN FRECUENCIA. EL FILTRO OBTENIDO AQUÍ ES FÍSICAMENTE REALIZABLE.

UNA APLICACIÓN INTERESANTE DE LA TEORÍA DEL FILTRO DE WIENER ES EN PROBLEMAS DE CANCELACIÓN DE RUIDO, WIDROW, ET AL (1975). ANTES DE ENTRAR A UN EJEMPLO VEAMOS LA OBTENCIÓN DE UN DESARROLLO DEL ENFOQUE DE WIENER PARA SU APLICACIÓN A LA CANCELACIÓN DEL RUIDO. EL FILTRO DE WIENER PARA CASO DISCRETO, CON UN ESQUEMA SEMEJANTE AL CONTINUO ESTUDIADO ANTERIORMENTE COMO EN LA FIGURA



LA FUNCIÓN DE AUTOCORRELACIÓN DISCRETA DE LA ENTRADA ES

$$\phi_{xx}(k) \triangleq E\{x_j x_{j+k}\}$$

La función de autocorrelación cruzada es

$$\phi_{xd}(k) \triangleq E\{x_j d_{j+k}\}$$

La ecuación de Wiener-Hopf discreto es

$$\sum_{l=-\infty}^{\infty} W^*(l) \phi_{xx}(l-k) = \phi_{xd}(k)$$

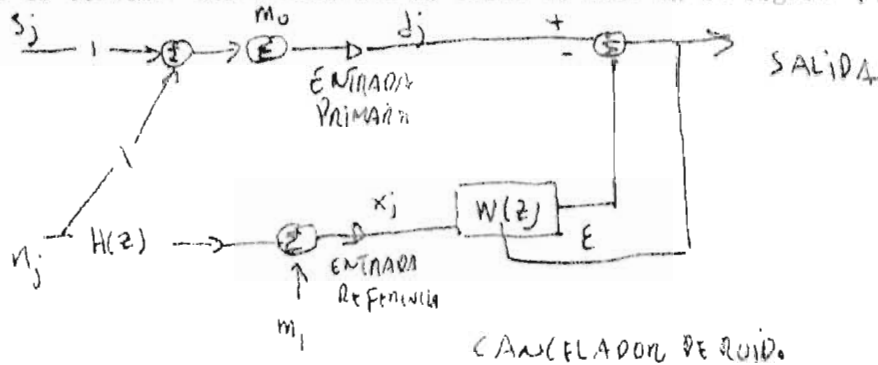
$$W(k) * \phi_{xx}(k) = \phi_{xd}(k)$$

Entonces el filtro de Wiener no restringido a ser físicamente realizable es

$$W^*(z) \triangleq \sum_{-\infty}^{\infty} W^*(k) z^{-k} = \frac{\phi_{xd}(z)}{\phi_{xx}(z)} = \frac{\sum \phi_{xx} z^{-k}}{\sum \phi_{xx}(z) z^{-k}}$$

donde  $\phi(\cdot)$  son espectros de potencia

si el esquema del cancelador de ruido es como en la figura F.2



La entrada primaria al cancelador es  $S_j + m_0 + n_j$  la entrada de referencia al cancelador es  $m_1 + n_j * H(j)$  la salida del cancelador de ruido. Para poder utilizar Wiener se asume que el proceso ha convergido y la solución del error mínimo se ha obtenido. Entonces  $W^*(z)$  se puede utilizar, para determinar la función de transferencia no restringida óptima. Sustituyendo  $\phi_{xd}(z), \phi_{xx}(z)$  en  $W^*$  tenemos

$$W^*(z) = \frac{\phi_{nn}(z) H(z^*)}{\phi_{m_1 m_1}(z) + \phi_{nn}(z) |H(z)|^2}$$

donde  $W^*$  es independiente de la espectro de la señal  $\phi_{ss}(z)$ , y  $\phi_{m_0 m_0}(z)$  si el ruido  $n_j = 0$ ,  $\phi_{m_1 m_1}(z)$  es cero y

$$W^*(z) = \frac{1}{H(z)}$$

lo que significa que el ruido  $n_j$  se cancela en la salida del cancelador de ruido.

El filtro obtenido es no restringido es decir que es no causal y no realizable, en el caso continuo obtuvimos un filtro físicamente realizable, para este caso podemos aproximar filtros no causales en una forma de retrasos, que aseguren una respuesta en tiempo real estable. Si en la entrada del cancelador en la primaria, insertamos un retraso  $\Delta$ , el retraso causa un retraso igual en la respuesta impulso del filtro no restringido. El valor de  $\Delta$  se elige para causar que el pico de la respuesta a impulso se centre en la línea de retraso. Widrow et al, comentan que esta elección produce la menor potencia de salida del ruido. La figura 3.2 muestra este esquema con retraso.

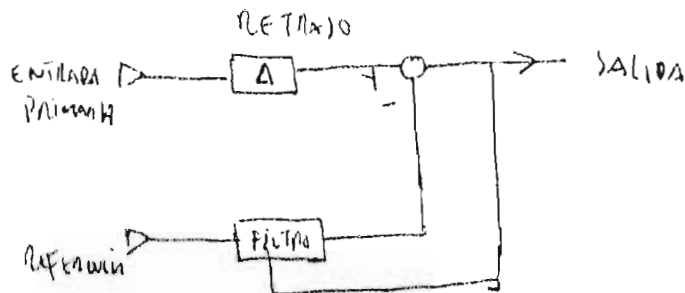


FIGURA 3.2

Vemos un problema estudiado por Widrow.

Considerando una solución de la forma

$$W(z) = \frac{1}{H(z)}$$

La función de transferencia es

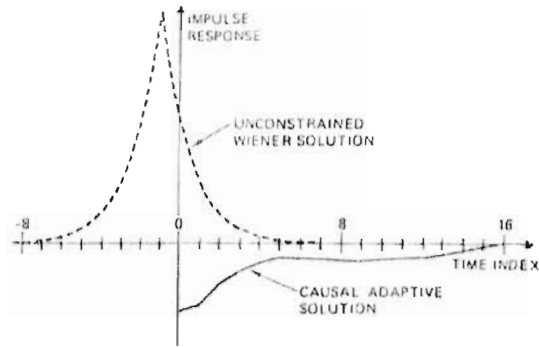
$$H(z) = 2z^{-1}(1 - \frac{1}{2}z^{-1})(1 - \frac{1}{2}z)$$

que es de fase nóminal, filtro transversal para bajas con dos ceros y ningún polo. El filtro no restringido es entonces

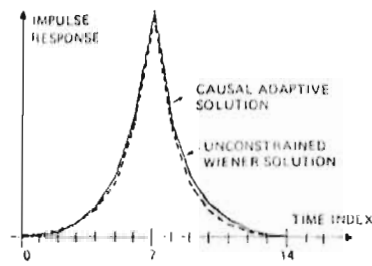
$$W(z) = \frac{1}{H(z)}$$

tiene un polo dentro del círculo unitario y otro polo fuera del círculo unitario. La entrada primaria consiste de un onda triangular, con un ruido coloreado. La entrada de referencia es un ruido coloreado correlacionado con el ruido primario. La respuesta impulso del filtro óptimo de Wiener es como en la figura 3.2. La solución causal sin retraso en la entrada se ve en la figura 3.2a en la parte inferior, podemos anticipar que el cancelador de ruido dará una versión pobre de la señal triangular, como en la figura 3.2c. Utilizando 8 unidades de retraso (la mitad de la longitud del filtro), la solución de la respuesta impulso y la solución del filtro de Wiener son más parecidas, por lo que la señal triangular obtenida con un retraso es más parecida a la señal triangular original. FIGURA 3.2 b, d.

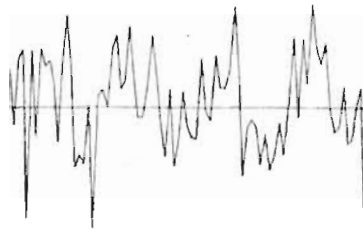




(a)



(b)



(c)



(d)

Fig. F.3 Results of noise cancelling experiment with delay in primary input path. (a) Optimal solution and adaptive solution found without time delay. (b) Optimal solution and adaptive solution found with delay of eight time units. (c) Noise canceller output without delay. (d) Noise canceller output with delay.

FIGURA F.3

## APLICACIONES DEL FILTRO DE KALMAN.

Una extensa descripción de aplicaciones de la estimación de parámetros y estados se pueden encontrar en Eykoft(1974), que de alguna forma tienen una relación con la estimación de estados por el método de Kalman, estas aplicaciones comprenden diversas disciplinas de las ciencias físicas, de la Ingeniería, etc.

Además las aplicaciones del filtrado por Kalman han sido hechos en el campo de navegación espacial, por ejemplo un trabajo de Thomas R. Blackburn(1971), En Schwartz(1975) podemos encontrar una aplicación del filtro de Kalman al tráfico aéreo y otras aplicaciones más, en el algoritmo recursivo de Kalman se requiere de un conocimiento a priori de la estadística de la señal y el ruido, además de conocer las matrices del sistema  $F, G, H$ ; lo que significa que el problema de modelado o identificación se han resuelto antes de resolver el problema de filtrado de Kalman(Mehra, 1971).

Mehra(1971), en vista de esto propone un enfoque de solución, en donde todos los parámetros son desconocidos, por lo que hay que identificarlos antes de aplicar estimación de estados por Kalman. En este trabajo se encontrarán similitudes entre el enfoque de series de tiempo (modelos ARMA), y el método de estimación de Kalman. Adoptemos este problema como una aplicación a la estimación, que puede abrir cauces para la solución de un problema más real.

### ESTABLECIMIENTO DEL PROBLEMA

Sea un sistema dinámico lineal estocástico discreto

$$\begin{aligned}x_{i+1} &= \Phi x_i + \Gamma u_i \\ z_i &= H x_i + v_i\end{aligned}\quad E.1$$

$i = 0, 1, \dots$

$x_i$      $n \times 1$

$\Phi$      $n \times n$

$\Gamma$      $n \times p$

$u_i$      $p \times 1$

$z_i$      $r \times 1$

$H$      $r \times n$

$v_i$      $r \times 1$

$$E\{u_i\} = 0$$

$$E\{v_i\} = 0$$

$$E\{u_i v_j^T\} = 0 \quad \text{PARA TODA } i, j$$

$I$     MATRIZ IDENTIDAD  $p \times p$

$R$     MATRIZ POSITIVA DEFINIDA SIMÉTRICA  
( $r \times r$ )

se desea determinar el orden  $n$  del sistema, e identificar las matrices usando datos de salida  $\{z_i, i=0, \dots, N\}$  bajo las siguientes

1) El sistema es completamente controlable y completamente observable

$$\begin{aligned} \text{rango} [r, \phi r, \dots, \phi^{n-1} r] &= n \\ \text{rango} [H^T, \phi^T H^T, \dots, (\phi^{n-1})^T H^T] &= n \end{aligned} \quad E.2$$

2) La matriz  $\phi$  es no singular y estable.

3) El sistema es de fase mínima

4) El esquema de identificación se arranca una vez que el sistema ha alcanzado de estado permanente.

Considerando el caso escalar para  $Z_i$

a) En la literatura estadística  $Z_i$  es un proceso ARMA.

b) 2 y 3 son análogos en series de tiempo; 1, es diferente ya que Walker asume que la transformada  $z$  del sistema no tiene cancelaciones de polos y ceros, una condición la cual es equivalente a 3 y 4.

c) las ecuaciones 1) y 2) pueden representar una serie de tiempo autorregresivo, por haciendo  $R=0$ , y tomando a  $\Gamma$  como una forma especial.

#### OBTENCIÓN DE LOS PARÁMETROS

Escribamos únicamente los pasos para la determinación de los parámetros.

PARA  $r=p=1$

Definiendo la correlación de  $Z_i$  como

$$C_k = E\{Z_i Z_{i-k}\} \quad E.3$$

Un estimado de , usando propiedad ergódica de  $Z_i$  es

$$\hat{C}_k = \frac{1}{N} \sum_{i=k}^N Z_i Z_{i-k} \quad E.4$$

y tenemos

$$\begin{bmatrix} \hat{C}_1 \\ \vdots \\ \hat{C}_n \end{bmatrix} = \begin{bmatrix} \hat{C}_1 \\ \vdots \\ \hat{C}_n \end{bmatrix} \quad \begin{bmatrix} \hat{C}_n \\ \vdots \\ \hat{C}_{2n-1} \end{bmatrix}^{-1} \begin{bmatrix} \hat{C}_{n+1} \\ \vdots \\ \hat{C}_{2n} \end{bmatrix} \quad E.5$$

DONDE  $\hat{A} = \frac{1}{\hat{\phi}_1} \sum_{j=0}^n \hat{\phi}_{j+1} \hat{c}_j$  E.6

Usando una representación canónica para el sistema E.1, por la transformación

$$T = \begin{bmatrix} H \\ H\Phi \\ \vdots \\ H\Phi^{n-1} \end{bmatrix},$$

$$y_i = T x_i$$

Y obtenemos

$$y_{i+1} = \Phi^* y_i + \Gamma^* u_i$$

$$z_i = H^* y_i + v_i$$

E.7

DONDE

$$\Phi^* = T\Phi T^{-1} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_1 & -a_2 & -a_3 & \dots & -a_n \end{bmatrix}$$

$$H^* = HT^{-1} = [1, 0, \dots, 0]$$

$$\Gamma^* = T\Gamma$$

ESTA REPRESENTACION CANONICA, TIENE EL MENOR NUMERO DE ELEMENTOS DESCONOCIDOS

$$P^* \text{ SE DETERMINA EN LAS ECUACIONES } P^* H^{*T} = G = (\Phi^*)^{-1} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

$$P^* = \Phi^* P^* \Phi^{*T} + P^* P^{*T} \quad \} \text{ (E.8)}$$

MEHRA (1970.1) RESUELVE ESTA ECUACION POR UN METODO ITERATIVO.

APLICACION DEL FILTRO DE KALMAN.

EL ESTIMADOR ES

$$\hat{y}_{i+1|i} = \Phi^* [\hat{y}_{i|i-1}] + K^* v_i \quad \text{E.9}$$

$$z_i = H^* \hat{y}_{i|i-1} + v_i$$

Donde  $v_i$  es la INNOVACION.

PARA UN FILTRO OPTIMO  $v_i$  ES BLANCO

$$E\{v_i\} = 0, \quad E\{v_i v_j\} = 0 \quad i \neq j \quad \text{E.10}$$

De E.9 necesitamos calcular  $K^*$  un método algorítmico para su determinación es:

1) Arrancar con cualquier ganancia  $K_0$  y filtrar el conjunto  $\{z_1, \dots, z_n\}$  usando E.10, y la ganancia  $K_0$ .

2) Autocorrelacionar la secuencia de innovación  $v_i$  y probar su "blancura", ya que si el filtro es óptimo, el proceso de innovación  $v_i$  es blanco. Esto puede ser probado estadísticamente utilizando el método de Jenkins y Watts; si  $c_k = E\{v_i v_{i-k}\}$  para un proceso ergódico

$$\hat{c}_k = \left(\frac{1}{N}\right) \sum_{i=k}^N v_i v_{i-k}^T, \quad N = \text{PUNTOS DE MUESTRA.}$$

un estimador de los coeficientes de autocorrelación  $\{c_k\}$  es

$$[\hat{P}_k]_{ij} = \frac{[\hat{c}_k]_{ij}}{\left\{[\hat{c}_0]_{ii} [\hat{c}_0]_{jj}\right\}^{1/2}} \quad \text{E.11}$$

$\{\hat{v}_i\}_i$  como  $\{\hat{c}_i\}_i$  son asintóticamente normales : 1) prueba para blancura de  $\{v_i\}$  es chequear el conjunto de valores  $\{\hat{p}_i\}_i, i > 0$ , y chequear el número de veces que ellos yacen fuera de la banda  $\pm 1.96/\sqrt{N}$  si el número de veces es menor que 5% del total, la secuencia  $v_i$  es blanca.

3) si la secuencia no es blanca, cambiar la ganancia del filtro a  $K_1$ , donde

$$K_1 = K_0 + \begin{bmatrix} H^* \Phi^* \\ H^* \Phi^* (I - K_0 H^*) \Phi^* \\ \vdots \\ H^* [\Phi^* (I - K_0 H^*) \Phi^*]^{n-1} \end{bmatrix}^{-1} \begin{bmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_n \end{bmatrix}$$

donde  $\hat{y}_1, \hat{y}_n$  representan las correlaciones normalizadas de  $v_i$ .

4) Repetir el procedimiento precedente hasta que la secuencia  $\{v_i\}$  es blanca.

La secuencia  $K_0, K_1, \dots$  converge a  $K$

En Predicción lineal el orden  $p$  del filtro permite un ajuste espectral del modelo a la señal. Si el orden de la matriz del sistema es importante para un control de ajuste. La prueba para la determinación del orden es como en la determinación de la blancura de  $v_i$ . Se chequea la optimalidad del filtro desde que el filtro de Kalman es óptimo cuando el orden del sistema es correcto, también sirve como una prueba para probar el orden del sistema.

Se procede como sigue

- 1) sea  $n=1$ , usar E.5, E.6 y E.8, para determinar  $\Phi, R$  y  $P^*$
- 2) filtrar los  $Z_i$  usando E.9 y E.12.  $K^*$  se determina como el límite de la secuencia  $k_1, k_2, \dots$ , donde  $k_1$  puede ser derivado de la ecuación de varianza de Kalman, usando  $\Phi, R, P^*$
- 3) chequear si la innovación es blanca.
- 4) Si la secuencia  $v_i$  no es blanca, repetir los pasos 1, 2-3, con  $n=2, 3, \dots$

Un ejemplo numérico es resuelto por Mehr. El sistema es de cuarto orden los valores actuales de  $\Phi, P, H, R$  son

$$\Phi = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -0.656 & 0.709 & -0.118 & 1.0 \end{bmatrix} \quad \Gamma = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad q = 1.0$$

$$H = [1, 0, 0, 0]$$

$$R = 0.25$$

Se tomaron 1000 puntos de muestr. de la señal  $z$ :

estos datos se procesaron como sigue

- i) estimados  $\hat{c}_k, k=0, \dots, 3$ , se obtuvieron de E.4
- ii) estimados  $\hat{\phi}_1, \dots, \hat{\phi}_4$  se generaron de E.5, usando estos estimados  $R$  se obtuvo de E.8. Los valores se muestran en la primera iteración de la tabla E
- iii) Usando estos valores de iteración i) y las ecuaciones de varianza de Kalman se calcula  $K_1$ .
- iv) Muestra la prueba de blancura para la secuencia de innovación, que viola la condición de 5% por lo tanto el filtro no es óptimo.
- v) se calcula una nueva  $K_2$ , usando E.12; los resultados se encuentran en la 4 tabla con iteración 2, aquí la innovación es blanca, se muestran los valores  $q, r$  obtenidos de  $K_2$ .
- vi) el procedimiento precedente se repite una vez más para checar convergencia, que se logra en tres iteraciones.
- vii) para comparación se ven en la tabla un proceso de estimación usando estimación de máxima similitud, además los valores actuales de  $\Phi, q, R$  y la ganancia  $K^*$  obtenida usando estos valores y las ecuaciones de varianza de Kalman.





## SUMARIO.

Con el material presentado aquí en este trabajo, se ha tratado de situar más que nada en los aspectos de Estimación de Sistemas, Filtrado y Predicción Lineal son conceptos que nos llevan a esta comprensión. Aunque Filtrado y Predicción autónomamente hayan alcanzado un desarrollo bastante grande en la resolución de problemas, podemos encontrar una interrelación estrecha entre ellos y que favorecen más los aspectos más vitales de Estimación de Sistemas.

El material de este trabajo cubre estos aspectos sin perder de vista los esquemas de estimación determinar por ejemplo los parámetros predictores de una función de transferencia  $H(z)$  en LPC; estimar una señal del ruido en el esquema de cancelación de ruido; extraer una función de transferencia en la formulación de Wiener; estimar el estado con el Filtro de Kalman.

REFERENCES.

- 1.-M. Athans and Tse, "A Direct Derivation of the Optimal Linear Filter Using The Maximum Principle", IEEE Trans. on Automatic Control, AC-12, No. 6, pp. 690-698, Dec. 1967.
- 2.-Thomas R. Blackburn and D. R. Vaughan, "Applications of linear optimal Control and Filtering Theory to the Saturn V Launch Vehicle", IEEE Trans. on Automatic Control, AC-16, No. 6, Dec. 1971.
- 3.-P. Eykoff, "System Identification", New York: Wiley, 1974.
- 4.-Fant, "Acoustic Theory of speech Production", Netherlands, 1960.
- 5.-Planagan, "Speech Analysis Synthesis and Perception, 2nd Edition (Springer-Verlag, Berlin, Heidelberg, New York, 1972.
- 6.-I. A. Gura, "An Algebraic Solution of The State Estimation Problem", AIAA J. 7, 1242-1247.
- 7.-R. Jaffe and E. Rechin, "Design and Performance of Phase-Lock Circuits Capable of near-Optimum Performance Over a Wide Range of Input Signal and Noise Level", IRE Trans. Information Theory IT-1, pp. 66-76, 1955.
- 8.-T. Kailath, "An Innovations to Least Squares Estimation, part I: Linear Filtering in Additive White Noise", IEEE Trans. on Automatic Control, AC-13, No. 6, Dec. 1968, pp. 646-654.
- 9.- "Linear Least Squares Estimation", Edited by T. Kailath, D Dowden Hutchinson Ross, Inc., Stroudsburg, Pennsylvania 1977.
- 10.-R. E Kalman, "A new Approach to Linear Filtering and Prediction Problem", Journal Basic Engineering, Vol. 82, pp. 34-45, Mar. 1960.
- 11.-R.E. Kalman and R.S Bucy, "New Results in Linear Filtering and Prediction Theory", Trans. ASME, Ser. D. J. Basic Engr., Vol. 83, pp. 95-107, Dec. 1961.

- 12.-R. E. Kalman, "New Methods in Wiener Filtering Theory," Proc. First Symp. Eng. Apl. Random Functions Theory Probability, 1963.
- 13.-B. C. Kuo, "Analysis and Synthesis of Sampled-Data Control Systems", Prentice-Hall, Inc., Englewood Cliffs, N. J., 1963.
- 14.-J. Makhoul, "Linear Prediction ; A tutorial Review", Proceedings of the IEEE, Vol 63, No. 4, Apr 1975, pp. 561-580.
- 15.------, "Spectral linear prediction: Properties and Application. Trans. ASSP, Jun 1975, pp. 283-296.
- 16.-J. D. Markel and A. Gray, "Linear Prediction of Speech", Springer Verlag, Berlin Heidelberg, New York, 1976.
- 17.-R. K. Mehra, "On Line Identification of Linear Dynamic Systems with A Applications to Kalman Filtering", IEEE Trans on Automatic Control, AC-16, No. 1, Feb. 1971.
- 18.------<sup>(1)</sup>, "An Algorithm to solve Matrix Equations  $PH^T=C$  and  $P=\Phi P\Phi^T +G$ ", IEEE Trans. on Automatic Control AC-16, 1971.
- 19.-Papoulis, "Probability, Random Variables and Stochastic Processes", Mc Graw Hill, 1965.
- 20.-L.R. Rabiner and R.W. Schaefer (Englewood Cliffs, NJ: Prentice-Hall 1978.
- 21.-I. B. Rhodes, "A tutorial Introduction to Estimation and Filtering"; IEEE Trans on Automatic Control, AC-16, No. 6, Dec. 1971, pp.688-706.
- 22.-A P Sage, J. L. Melsa, "Estimation Theory, with Applications to Communications and Control", Mc Graw-Hill Book Company, 1971.
- 23.-A. P. Sage, C.C. White III, "Optimum System Control", Prentice Hall, 1977.
- 24.-R. A. Singer and Frost, "On the Relative Performance of The Kalman and Wiener Filters", IEEE, Trans. on Automatic Control AC-August 1969, pp# 390-394.

25.-M. Schwartz, L. Shaw, "Signal Processing", Mc Graw-Hill, 1975.

26.-P. Swerling, "Modern State Estimation Method From The View Point of The Method of Least Squares", IEEE Trans. on Automatic Control, AC16, 1971.

27.-B. Widrow y colaboradores, "Adaptive Noise Cancelling: Principles and Applications", Proceedings of the IEEE, Vol 63, No. 12, Dec. 1975, pp. 1692-1716.

28.-N. Wiener, "Extrapolation, Interpolation and Smoothing of Stationary Time Series", New York: Tecnology Press and Wiley, 1949.