

Procesado Estadístico de Señal

Filtro de Wiener



Departament de Teoria
del Senyal i Comunicacions



Índice

1. Señales y procesos estocásticos
2. Estimación espectral
3. Filtro de Wiener
 - 3.1. Mínimos cuadrados y ecuaciones normales
 - 3.2. Filtro de Wiener en frecuencia
 - 3.3. Filtro óptimo a partir de las muestras
 - 3.4. Predicción lineal
 - 3.5. Algoritmo de Levinson y estructura en celosía
 - 3.6. Aplicaciones
4. Filtrado adaptativo
5. Métodos transformados en codificación



Esquema general del filtro de Wiener

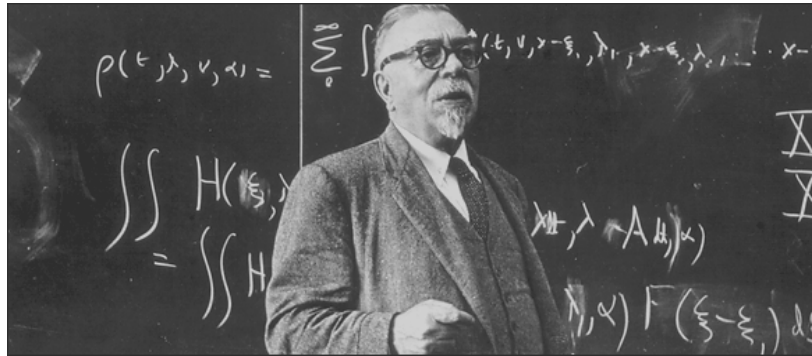
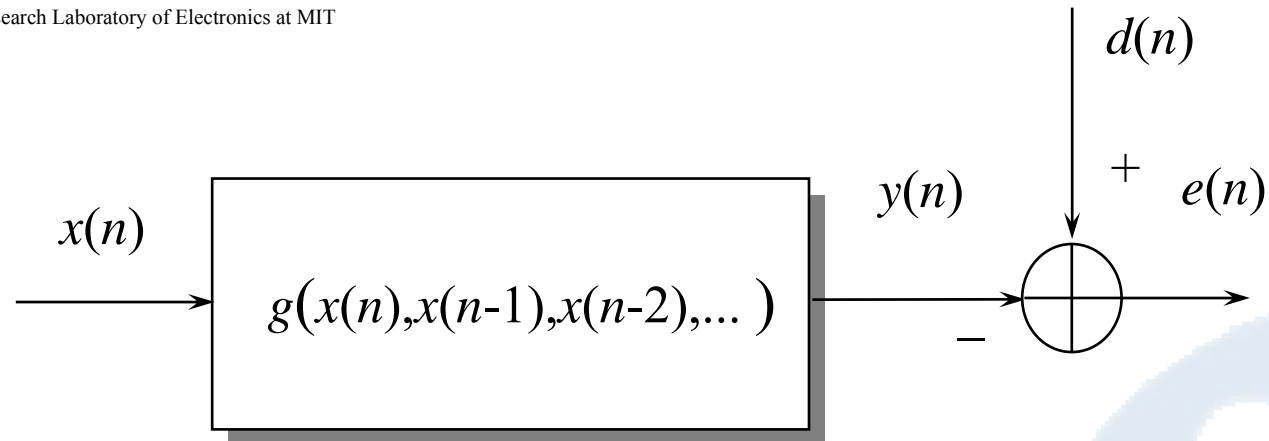


Image courtesy of the Research Laboratory of Electronics at MIT



$x(n)$

$y(n)$

$d(n)$

$e(n) = d(n) - y(n)$

entrada al filtro

salida del filtro

respuesta deseada

error de estimación

Estimación cuadrático media: nota histórica

Gauss (1795) y Legendre (1805)

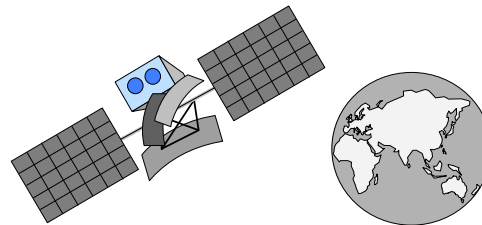
Estudio del movimiento de los cuerpos celestes: método de mínimos cuadrados



Kolmogorov (1939) y Wiener (1949)

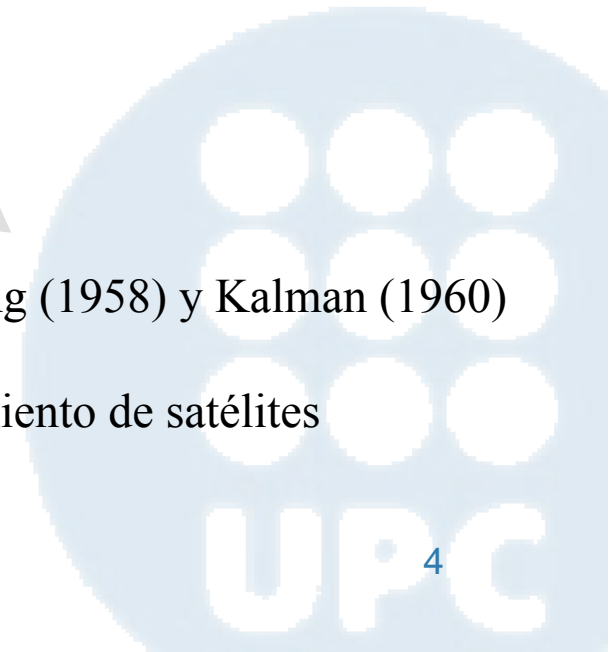
http://en.wikipedia.org/wiki/Norbert_Wiener

Mejora de las defensas antiaéreas a partir de observaciones de posición y velocidad



Swerling (1958) y Kalman (1960)

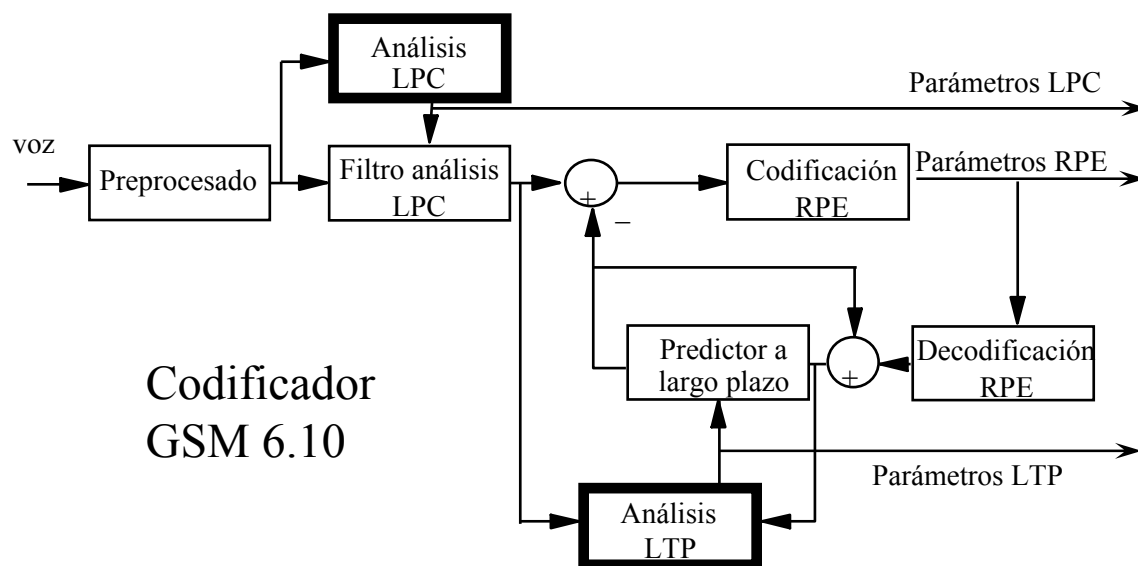
Seguimiento de satélites



Ejemplo 1. Predicción lineal en la codificación de fuente

Codificadores de voz, audio e imagen basados en la cuantificación diferencial.

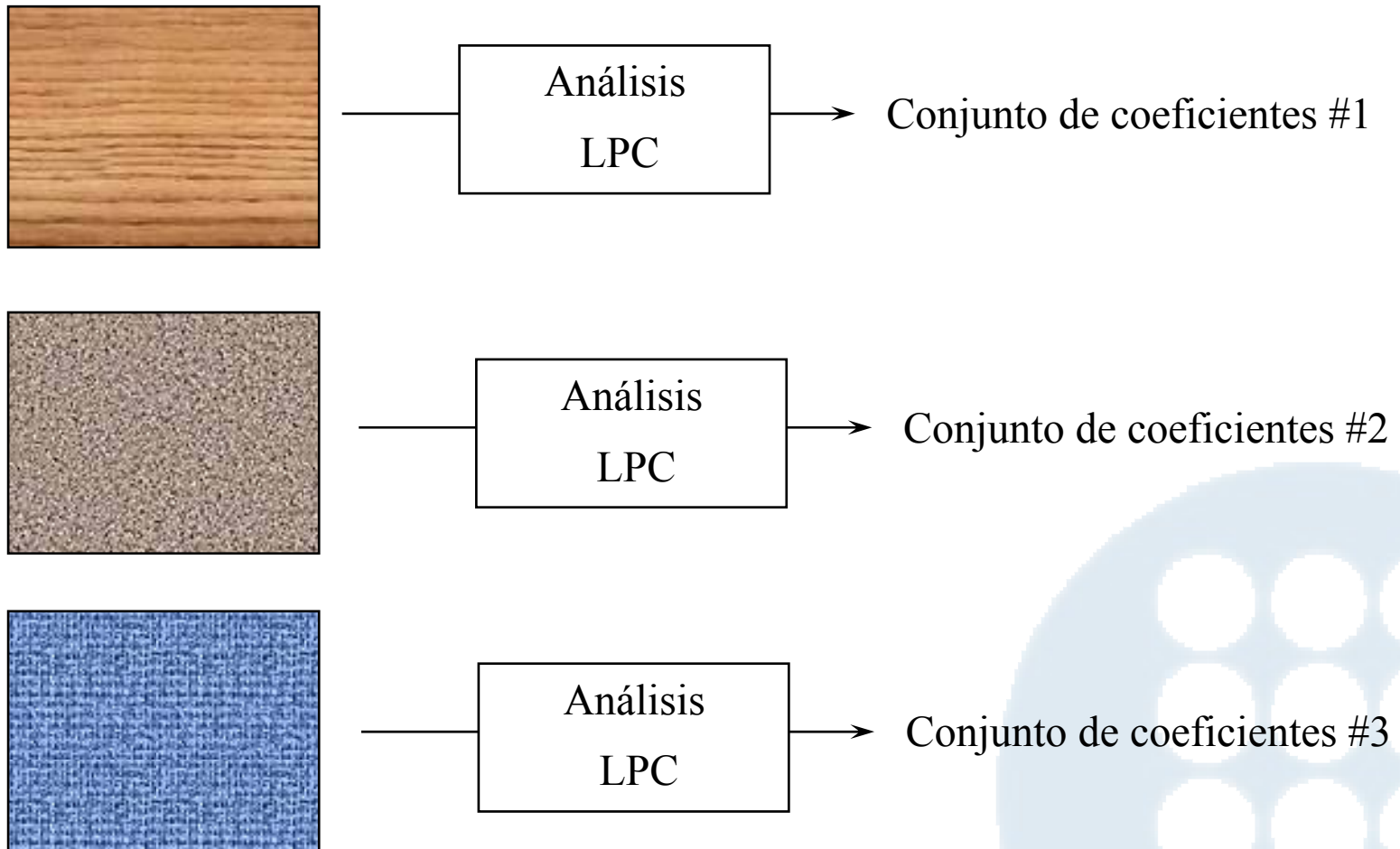
G.721, G.728 voz telefónica
G.722 audio y voz alta calidad



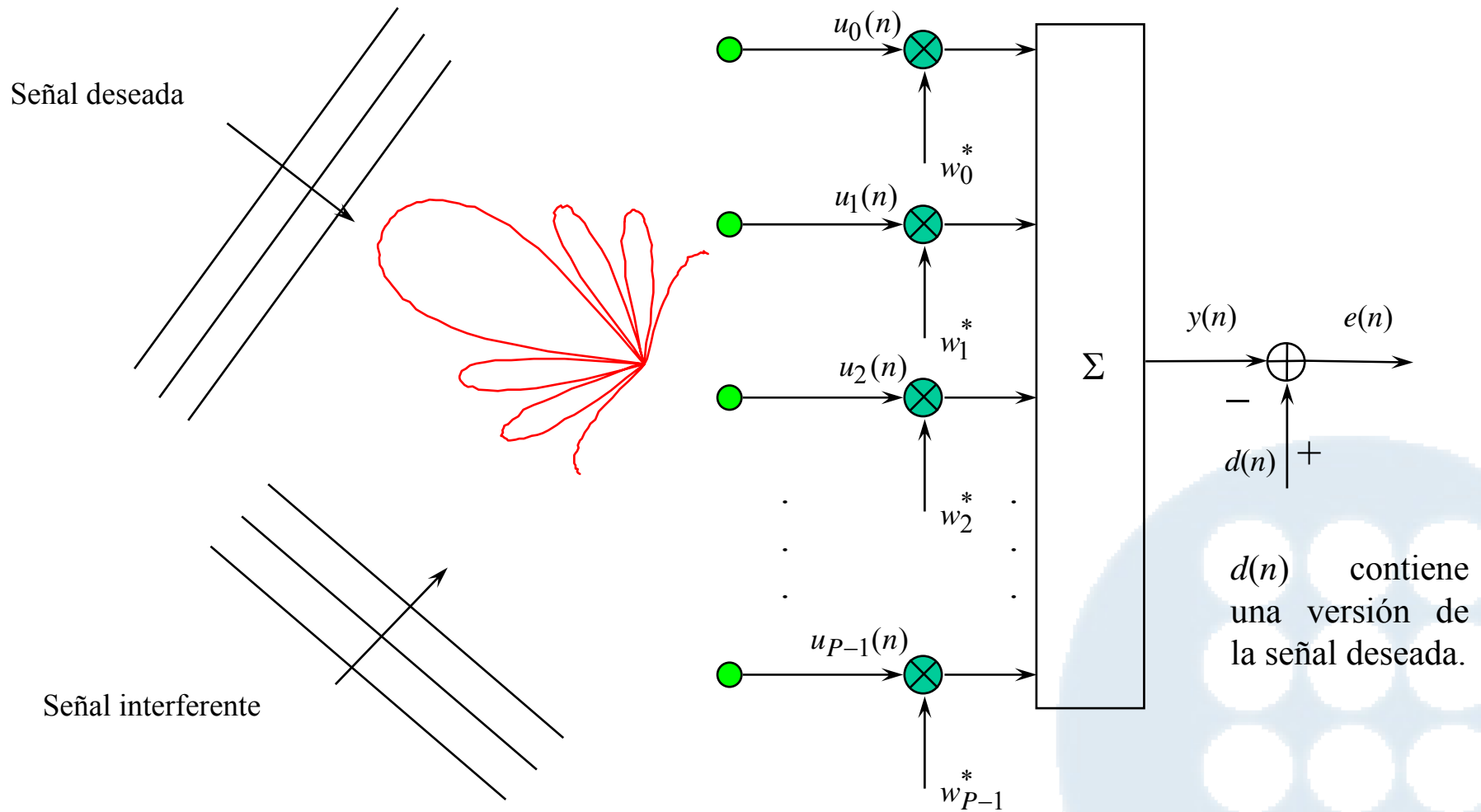
Codificadores basados en la transmisión de los coeficientes de predicción lineal.

GSM 6.10
LPC-10
CELP (GSM half-rate)

Ejemplo 2. Predicción en la codificación de texturas

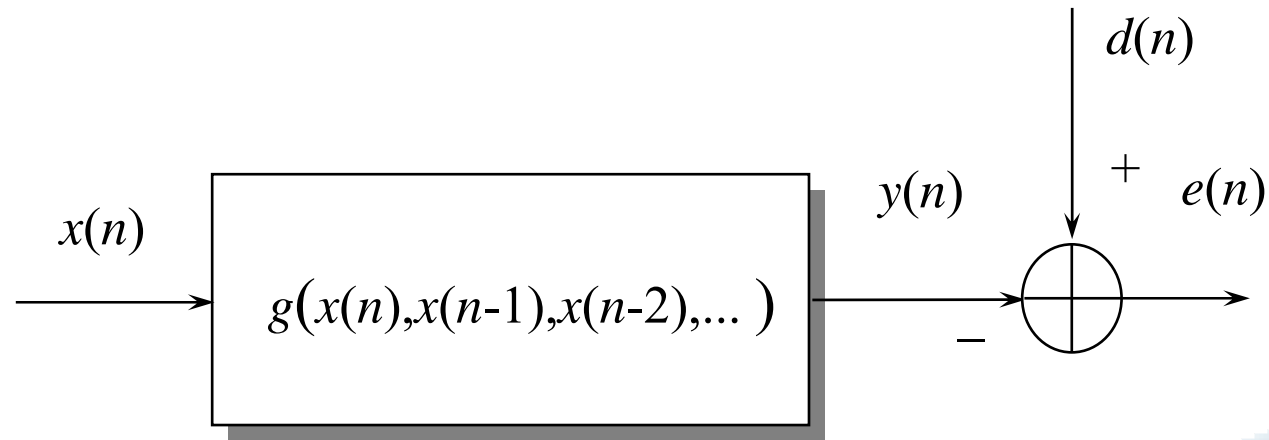


Ejemplo 3. Cancelación espacial de interferencia



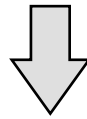
Esquema general del filtro de Wiener

Todos las configuraciones anteriores pueden generalizarse:



¿Con qué criterio escogemos $g(\cdot)$?

$$\min_g E \left\{ \left[d(n) - g(x(n), x(n-1), x(n-2), \dots) \right]^k \right\} \quad k = 2$$



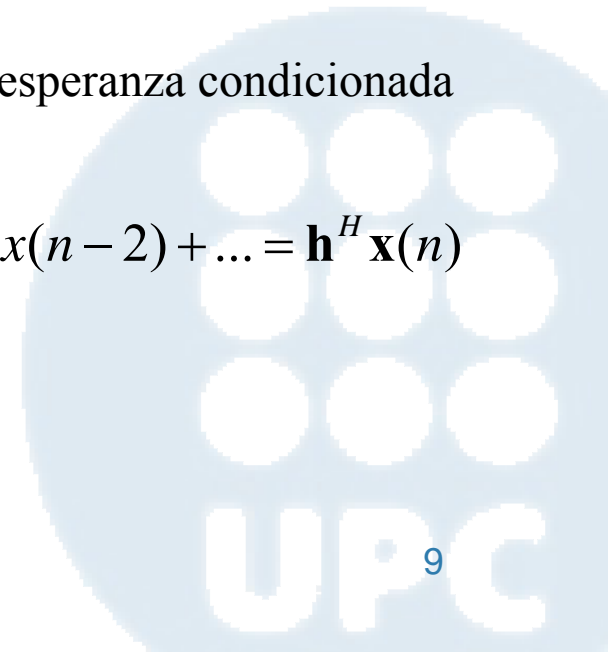
$$g(x(n), x(n-1), x(n-2), \dots) = E_d \{ d(n) / x(n), x(n-1), x(n-2), \dots \}$$

Si el error $e(n)$ es gaussiano, la minimización con $k=2$ es lo óptimo,
¡pero normalmente no conocemos la naturaleza de $e(n)$!

Si $d(n)$ y $x(n)$ son procesos conjuntamente gaussianos, la esperanza condicionada resulta ser:

$$g(x(n), x(n-1), x(n-2), \dots) = h_0^* x(n) + h_1^* x(n-1) + h_2^* x(n-2) + \dots = \mathbf{h}^H \mathbf{x}(n)$$

En estas condiciones el filtro óptimo es un **filtro lineal**.



Minimización del error cuadrático

Minimicemos la potencia del error:

$$\begin{aligned}\xi &= E\left\{|\varepsilon(n)|^2\right\} = E\left\{\left(d(n) - \mathbf{h}^H \mathbf{x}(n)\right)\left(d^*(n) - \mathbf{x}(n)^H \mathbf{h}\right)\right\} = \\ &= E\left\{|d(n)|^2\right\} + \mathbf{h}^H E\left\{\mathbf{x}(n)\mathbf{x}(n)^H\right\} \mathbf{h} - \mathbf{h}^H E\left\{\mathbf{x}(n)d^*(n)\right\} - E\left\{d(n)\mathbf{x}(n)^H\right\} \mathbf{h} = \\ &= P_d + \mathbf{h}^H \mathbf{R} \mathbf{h} - \mathbf{h}^H \mathbf{r}_{dx} - \mathbf{r}_{dx}^H \mathbf{h}\end{aligned}$$

$$\nabla_{\mathbf{h}^H} \xi = \mathbf{R} \mathbf{h} - \mathbf{r}_{dx} = \mathbf{0} \Rightarrow \boxed{\mathbf{h}_{opt} = \mathbf{R}^{-1} \mathbf{r}_{dx}} \quad \text{Ecuación de Wiener-Hopf}$$

$$\xi_{min} = P_d - \mathbf{r}_{dx}^H \mathbf{R}^{-1} \mathbf{r}_{dx} = P_d - \mathbf{r}_{dx}^H \mathbf{h}_{opt} = P_d - \mathbf{h}_{opt}^H \mathbf{R} \mathbf{h}_{opt}$$

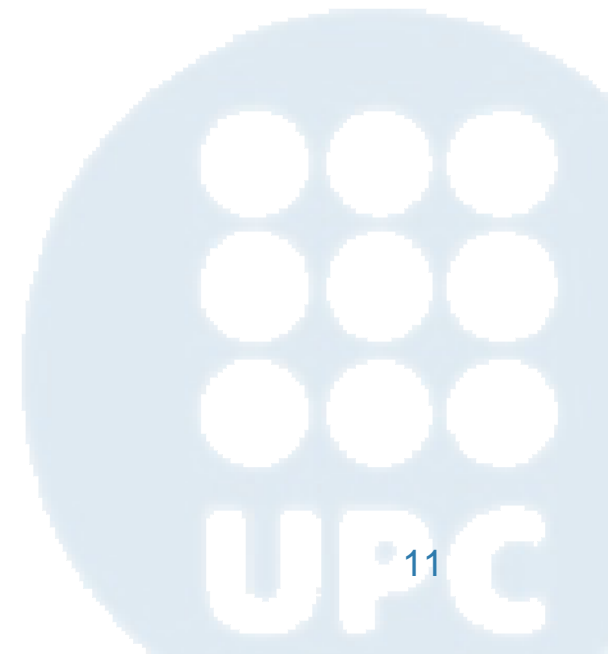
- La potencia del error es menor o igual que la potencia de $d(n)$, excepto si $x(n)$ y $d(n)$ están incorrelados, en cuyo caso son iguales.
- El MSE depende del error cometido en la implementación de los coeficientes y del escenario de entrada o datos (correlación)

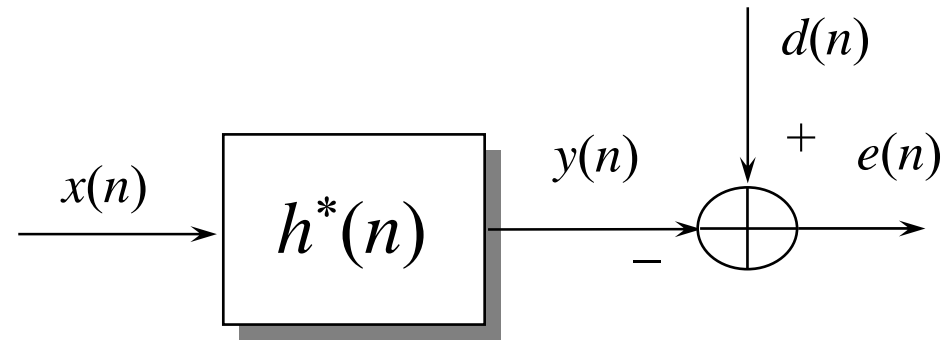
$$\xi = \xi_{min} + \left(\mathbf{h} - \mathbf{h}_{opt}\right)^H \mathbf{R} \left(\mathbf{h} - \mathbf{h}_{opt}\right)$$

- Si por errores en la implementación de los coeficientes (ej: cuantificación) los coeficientes son variables aleatorias cuya media es el valor óptimo, el MSE será variable aleatoria con valor esperado:

$$\begin{aligned} E\{\xi\} &= \xi_{\min} + E\left\{\left(\mathbf{h} - \mathbf{h}_{opt}\right)^H \mathbf{R} \left(\mathbf{h} - \mathbf{h}_{opt}\right)\right\} = \\ &= \xi_{\min} + \text{traza}\left[\mathbf{R} E\left\{\left(\mathbf{h} - \mathbf{h}_{opt}\right)\left(\mathbf{h} - \mathbf{h}_{opt}\right)^H\right\}\right] = \xi_{\min} + \text{traza}[\mathbf{R}\Sigma] \end{aligned}$$

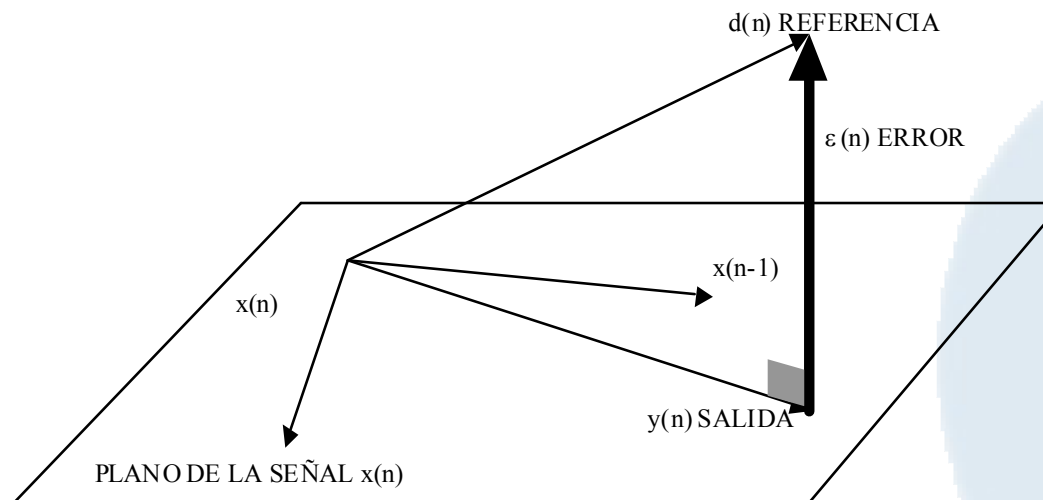
Σ es la covarianza de los coeficientes





El error es ortogonal a $x(n)$:

$$\min_{\mathbf{h}^*} E \left\{ |e(n)|^2 \right\} \Rightarrow \nabla_{\mathbf{h}^*} E \left\{ |e(n)|^2 \right\} = E \left\{ e^*(n) \nabla_{\mathbf{h}^*} e(n) \right\} = E \left\{ e^*(n) \mathbf{x}(n) \right\} = \mathbf{0}$$



Formulación matricial

En general, el filtro puede ser IIR. Si imponemos la restricción de que sea **FIR**, **siempre será estable**.

En este caso, las matrices y vectores serán de longitud finita y podemos escribir la ecuación de Wiener-Hopf como:

$$\begin{bmatrix} r_{dx}(0) \\ r_{dx}(-1) \\ \vdots \\ r_{dx}(-Q+1) \end{bmatrix} = \underbrace{\begin{bmatrix} r_{xx}(0) & r_{xx}(1) & \cdots & r_{xx}(Q-1) \\ r_{xx}(-1) & r_{xx}(0) & \cdots & r_{xx}(Q-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}(-Q+1) & r_{xx}(-Q+2) & \cdots & r_{xx}(0) \end{bmatrix}}_{\text{Matriz de correlación}} \begin{bmatrix} h(0) \\ h(1) \\ \vdots \\ h(Q-1) \end{bmatrix}$$

$$\mathbf{r}_{dx} = \mathbf{R}_{xx} \mathbf{h}$$

Matriz de correlación

Filtro óptimo en el dominio de la frecuencia

$$S_{dx}(f) = H(-f)S_{xx}(f)$$

$$\begin{aligned}\xi_{\min} &= P_d - \mathbf{r}_{dx}^H \mathbf{h}_{opt} = \int_{-1/2}^{1/2} [S_d(f) - S_{xd}(-f)H(f)] df = \\ &= \int_{-1/2}^{1/2} \left[S_d(f) - \frac{S_{xd}(f)S_{dx}(f)}{S_{xx}(f)} \right] df = \\ &= \int_{-1/2}^{1/2} S_d(f) \left[1 - \frac{S_{xd}(f)S_{dx}(f)}{S_x(f)S_d(f)} \right] df = \int_{-1/2}^{1/2} S_d(f) [1 - |\gamma(f)|^2] df\end{aligned}$$

↓
coherencia espectral entre
referencia y datos

- Permite anticipar la efectividad del filtro de Wiener

Coherencia espectral (I)

- Si existe una relación lineal entre los datos y la referencia la coherencia espectral es igual a la unidad

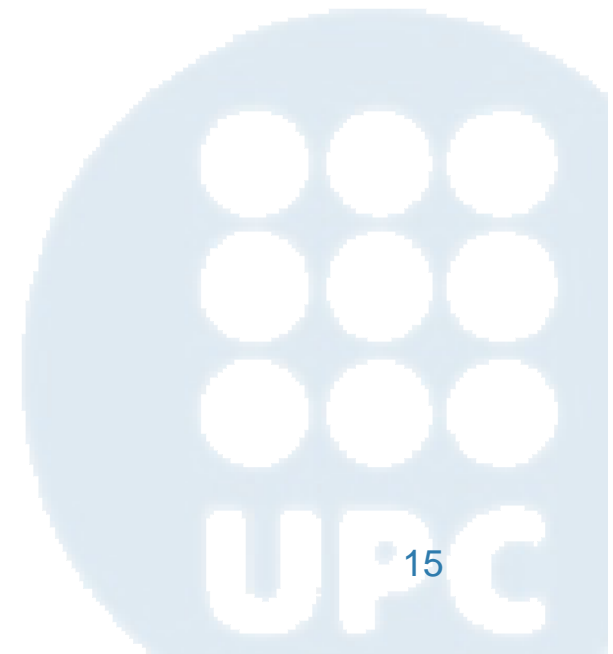
$$\text{si } x(n) = d(n) * g^*(n) \Leftrightarrow X(f) = D(f)G(f)$$

$$\text{entonces } S_{xd}(f) = G(-f)S_d(f)$$

$$S_{dx}(f) = G^*(-f)S_d(f)$$

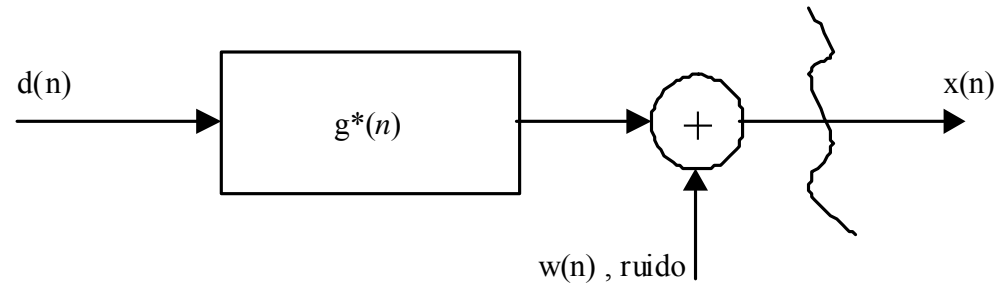
$$S_x(f) = |G(-f)|^2 S_d(f)$$

$$\text{y en consecuencia } |\gamma|^2 = 1 \quad \forall f$$



Coherencia espectral (II)

- En presencia de ruido aditivo



$$S_x(f) = |G(-f)|^2 S_d(f) + S_w(f)$$

$$S_{xd}(f) = G(-f) S_d(f)$$

$$S_{dx}(f) = G^*(-f) S_d(f)$$

$$|\gamma(f)|^2 = \frac{|G(-f)|^2 S_d^2(f)}{|G(-f)|^2 S_d^2(f) + S_w(f) S_d(f)} = \frac{1}{1 - \frac{S_w(f)}{|G(-f)|^2 S_d(f)}} = \frac{SNR(f)}{1 + SNR(f)}$$

- $SNR(f)$ es la relación entre la parte de señal coherente con la referencia y la no coherente (ruido)

$$\begin{aligned}
 H(-f) &= \frac{S_{dx}(f)}{S_x(f)} = \frac{G^*(-f)S_d(f)}{|G(f)|^2 S_d(f) + S_w(f)} = \\
 &= \frac{1}{G(-f)} \frac{S_d(f)}{S_d(f) + \frac{S_w(f)}{|G(-f)|^2}} = \frac{1}{G(-f)} \frac{1}{1 + \frac{1}{SNR(f)}}
 \end{aligned}$$

- Para valores elevados de $SNR(f)$ el filtro de Wiener ecualiza:

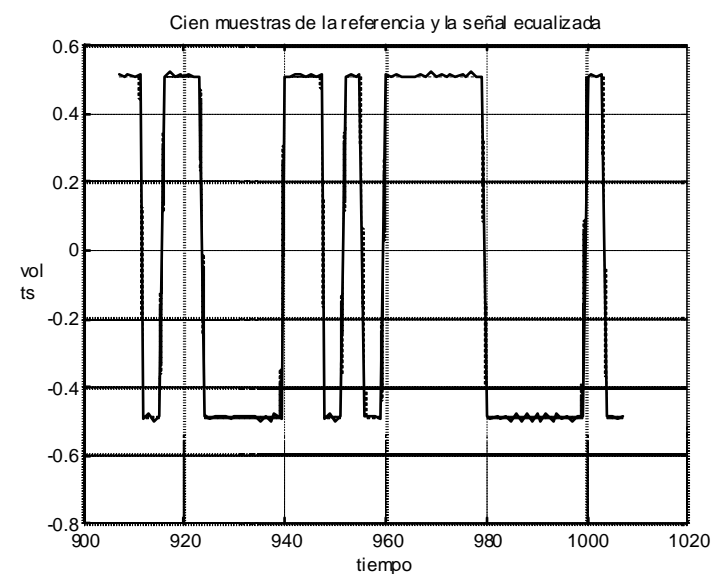
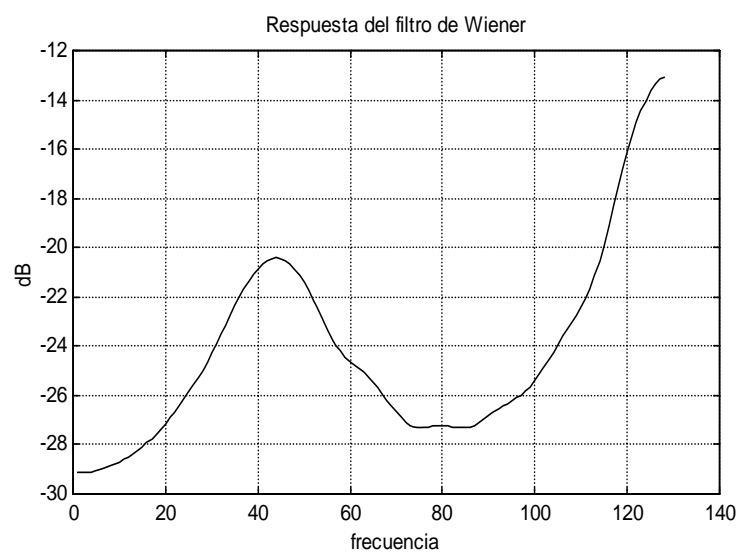
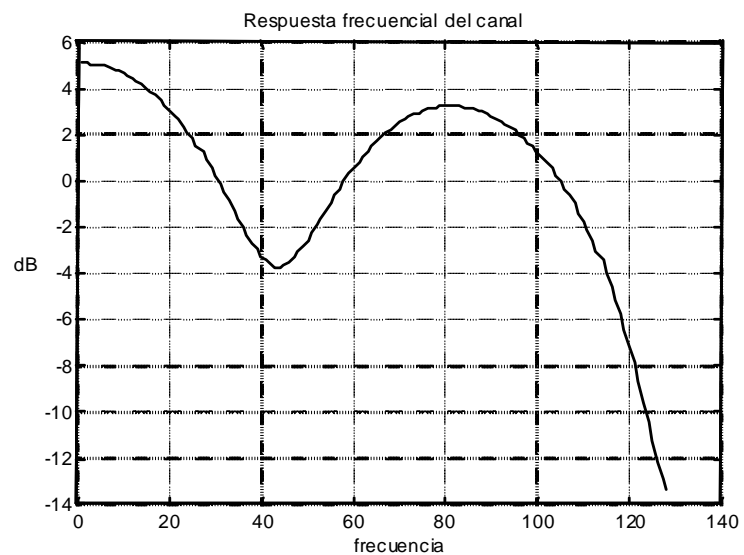
$$H(f) \approx \frac{1}{G(f)}$$

- Para valores reducidos de $SNR(f)$ el filtro de Wiener es el filtro adaptado a la respuesta del canal

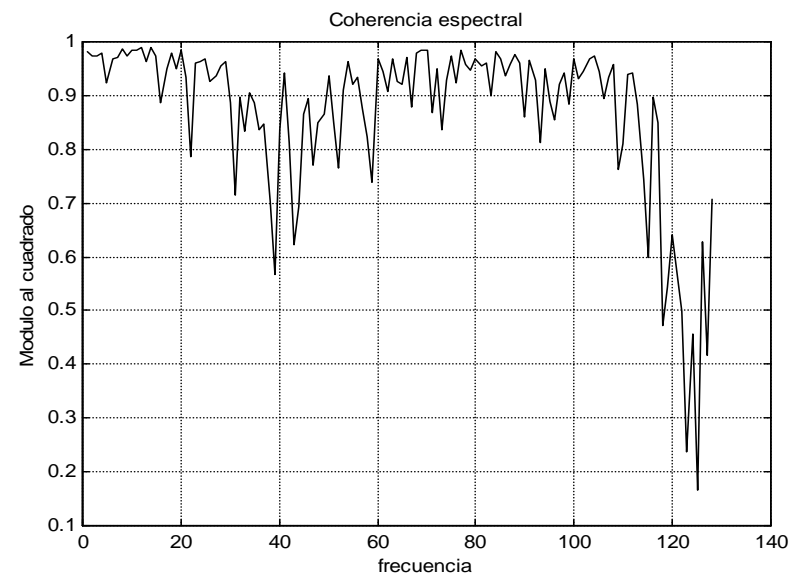
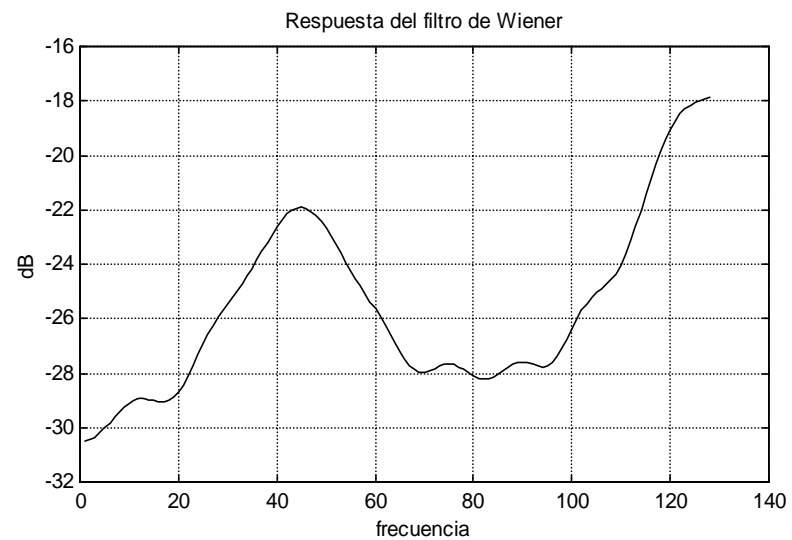
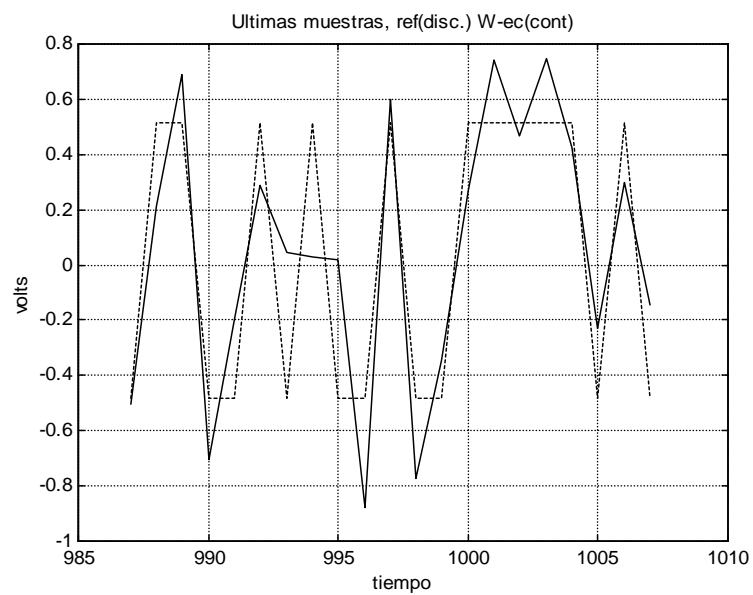
$$H(-f) \approx \frac{SNR(f)}{G(-f)} = G^*(-f) \frac{S_d(f)}{S_w(f)}$$

Ejemplo 4. Ecualización de canal

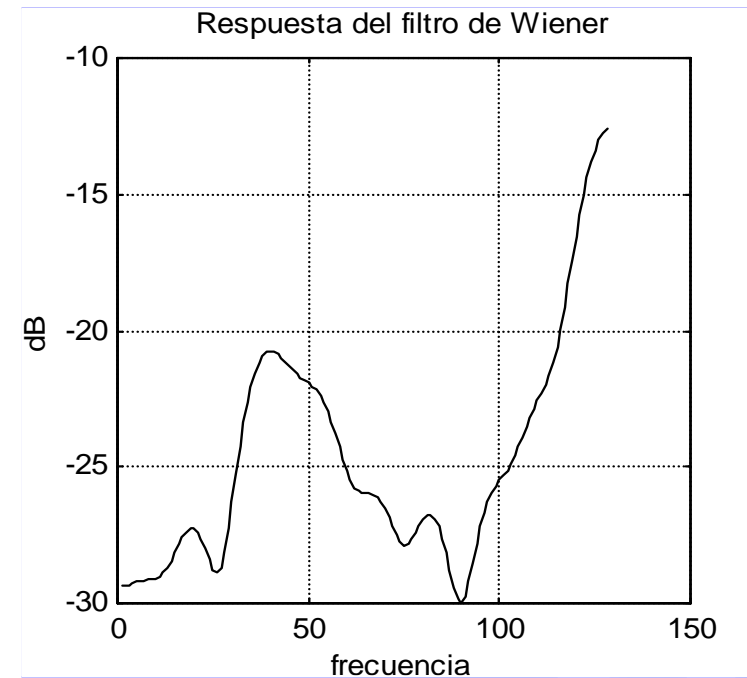
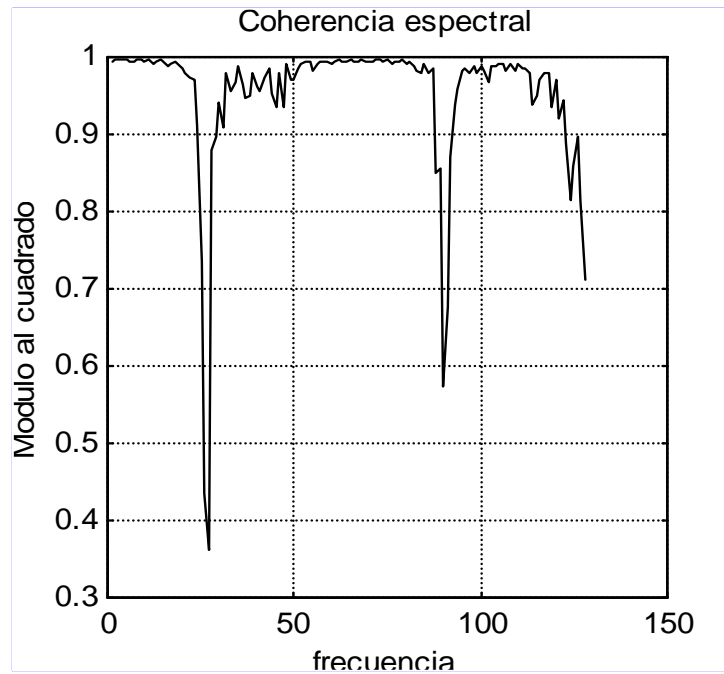
SNR > 60 dB



SNR > 11 dB



En presencia de dos señales interferentes de banda estrecha



Filtro óptimo a partir de las muestras

- Se desea diseñar un filtro de Q coeficientes
- Disponemos de N muestras de señal en $x(n)$ y en $d(n)$
- La salida del filtro $y(n)$ tendrá $N+Q-1$ de las que $Q-1$ serán el denominado pre-transitorio y otras $Q-1$ serán el post-transitorio.

$$\mathbf{y}^T = \underbrace{[y(0) \cdots y(Q-2)]}_{\text{pre-transitorio}} \underbrace{[y(Q-1) \cdots y(N-1)]}_{\text{permanente}} \underbrace{[y(N) \cdots y(N+Q-2)]}_{\text{post-transitorio}}$$

$$\mathbf{X} = \begin{bmatrix} x(0) & x(1) & \cdots & x(Q-2) & x(Q-1) & \cdots & x(N-1) & 0 & \cdots & 0 \\ 0 & x(0) & \cdots & x(Q-3) & x(Q-2) & \cdots & x(N-2) & x(N-1) & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & x(0) & x(1) & \cdots & x(N-Q-1) & x(N-Q) & \cdots & 0 \\ 0 & 0 & \cdots & 0 & x(0) & \cdots & x(N-Q) & x(N-Q-1) & \cdots & x(N-1) \end{bmatrix}$$

$$\mathbf{y}^T = \mathbf{h}^H \mathbf{X}$$

- Definición del MSE

$$MSE \equiv \sum_n |\varepsilon(n)|^2 = \|\mathbf{d}^T - \mathbf{y}^T\|^2 = \|\mathbf{d}^T - \mathbf{h}^H \mathbf{X}\|^2$$

- Gradiente respecto al vector de coeficientes del filtro

$$\mathbf{X}(\mathbf{d}^* - \mathbf{X}^H \mathbf{h}) = \mathbf{0}$$

$$\mathbf{h} = (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{d}^*$$

- Definición de los nuevos estimadores correlación

$$\hat{\mathbf{R}} = \mathbf{X}\mathbf{X}^H = \sum_{n=0}^{N+Q-2} \mathbf{x}(n)\mathbf{x}(n)^H \quad \hat{\mathbf{r}}_{dx} = \mathbf{X}\mathbf{d}^*$$

los vectores $\mathbf{x}(n)$ son las columnas de \mathbf{X}

- ¿Cómo se forma el vector referencia disponiendo sólo de N muestras?

1. $Q-1$ ceros y a continuación las N muestras disponibles de la referencia $d(n)$.
 $n=0, \dots, N+Q-2$
Incluye pre-transitorio de mínima energía
2. N muestras disponibles de la referencia $d(n)$ y $Q-1$ ceros al final.
 $n=0, \dots, N+Q-2$
Incluye post-transitorio de mínima energía

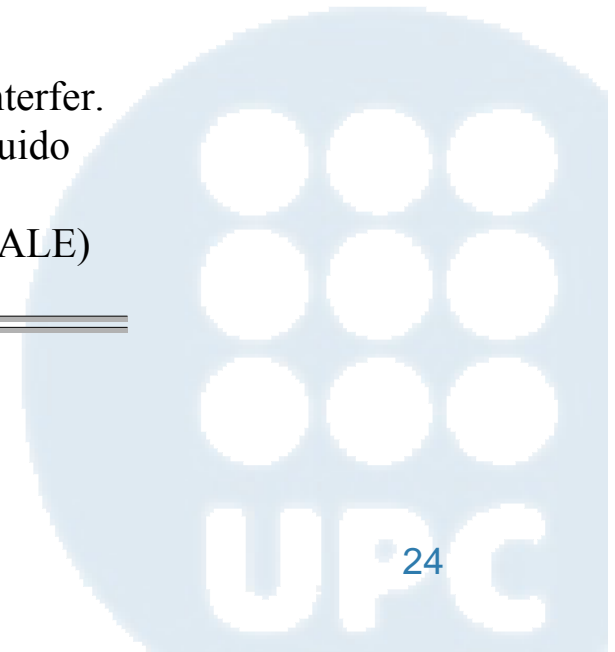
Método de covarianza

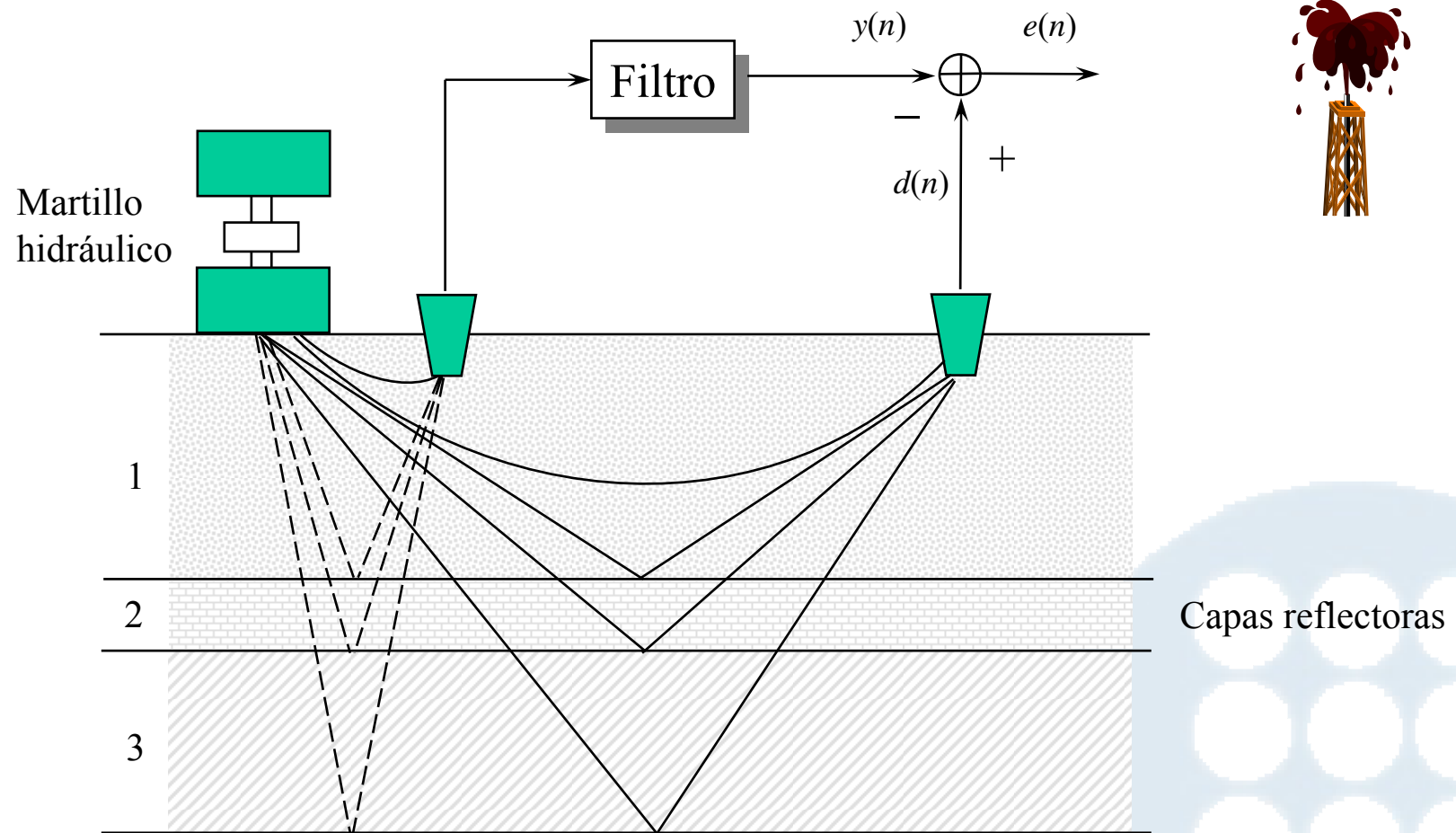
- Se considera sólo $N-Q+1$ muestras del permanente a la salida: $n=Q-1, \dots, N-1$
- No fija transitorios. No asume nada respecto a la señal de referencia donde se desconoce. Mejor calidad
 - Matriz de correlación no Toeplitz: elevada carga computacional

Método de correlación

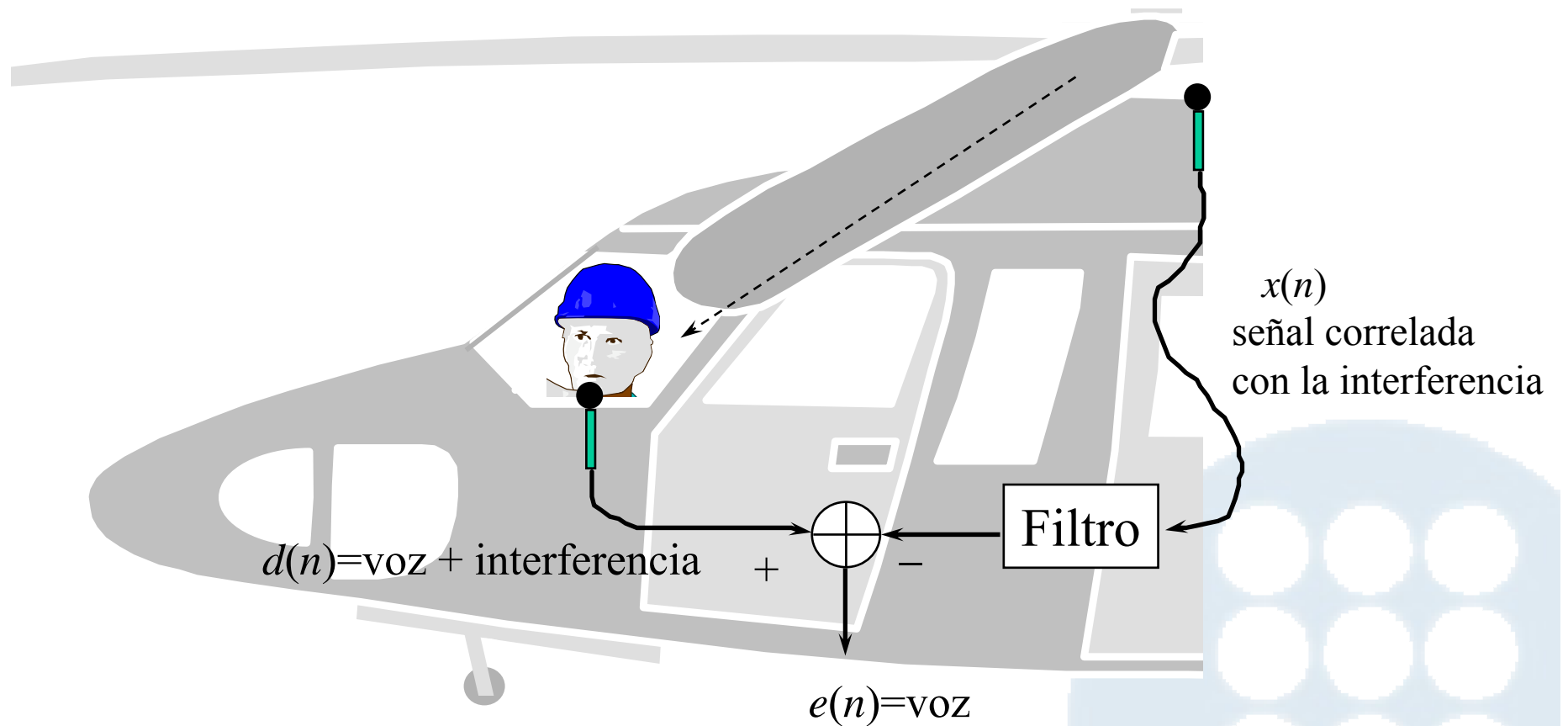
- $Q-1$ ceros, a continuación $N-Q+1$ muestras de la referencia $d(n)$ (empezando en $d(0)$) y $Q-1$ ceros al final: $n=0, \dots, N+Q-2$
- Mínima energía en los dos transitorios
 - Matriz \mathbf{XX}^H Toeplitz: ventaja computacional. Es el método más empleado en proceso segmentado por tramas.

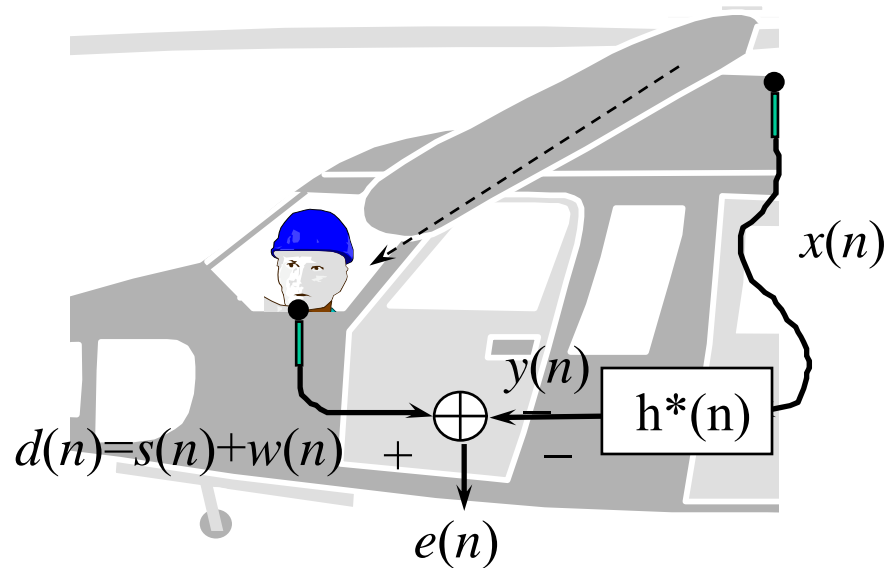
Tipo de filtrado	Aplicaciones
1. Identificación	Ident. de sistemas
2. Modelado inverso	Ecualización
3. Predicción	Codif. por predicción lineal DPCM-DM Estimación espectral Detección de señales
4. Cancelación	Cancelación de ruido/interfer. Cancelación activa de ruido Cancelación de eco Separación de señales (ALE) Conformación de haz





Ejemplo 4.





Busquemos los coeficientes del filtro $H(z)$:

$$H(z) = \sum_{i=0}^{Q-1} h_i^* z^{-i}$$

que nos permitan estimar $w(n)$ a partir de $x(n)$. Queremos pues minimizar:

$$J_1 = E \left\{ |w(n) - y(n)|^2 \right\}$$

Optimizando llegamos a las ecuaciones normales:

$$E \left\{ \left(w^*(n) - y^*(n) \right) \mathbf{x}(n) \right\} = E \left\{ \left(w^*(n) - \mathbf{x}^H(n) \mathbf{h} \right) \mathbf{x}(n) \right\} = \mathbf{0}$$



Llegamos a una solución que no podemos utilizar por el desconocimiento de \mathbf{r}_{wx} .

Veamos si podemos llegar a una solución minimizando la potencia de $e(n)$:

$$J_2 = E \left\{ |d(n) - y(n)|^2 \right\}$$

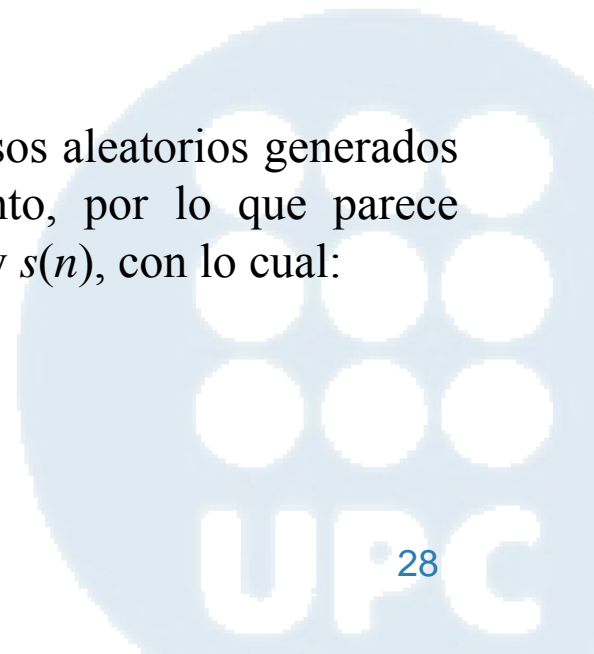
Las ecuaciones normales que obtenemos son:

$$E \left\{ \left(d^*(n) - y^*(n) \right) \mathbf{x}(n) \right\} = \mathbf{0}$$

$$E \left\{ \left(s(n) + w(n) - y(n) \right)^* \mathbf{x}(n) \right\} = \mathbf{0}$$

La voz del piloto $s(n)$ y el ruido del motor $w(n)$ son procesos aleatorios generados por mecanismos físicos de origen completamente distinto, por lo que parece razonable formular la hipótesis de incorrelación entre $x(n)$ y $s(n)$, con lo cual:

$$E \left\{ \left(w(n) - y(n) \right)^* \mathbf{x}(n) \right\} = \mathbf{0}$$



Obtenemos las mismas ecuaciones que con el criterio J_1

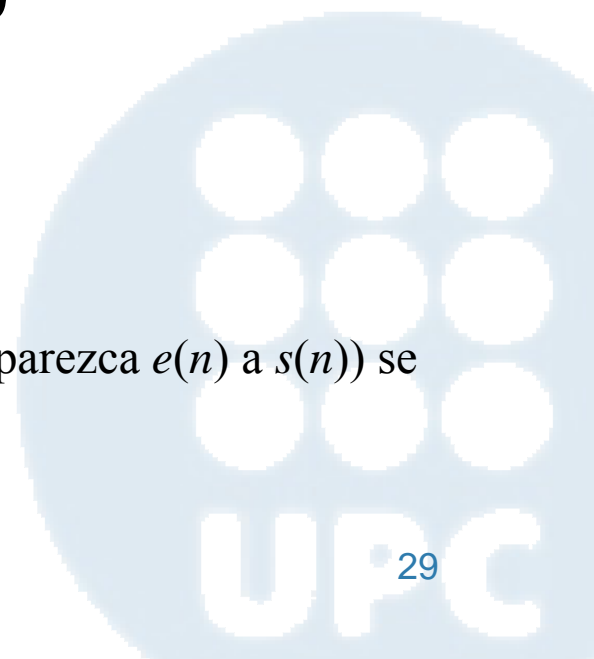


Podemos utilizar las ecuaciones que derivamos a partir de J_2 siempre y cuando la referencia y la señal que queremos recuperar sean incorreladas

$$E \left\{ \left(d(n) - y(n) \right)^* \mathbf{x}(n) \right\} = \mathbf{0}$$

$$\mathbf{r}_{dx} - \mathbf{R}_{xx} \mathbf{h} = \mathbf{0}$$

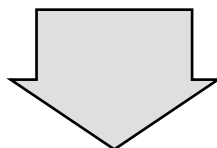
En la medida que el ruido quede cancelado (cuanto más se parezca $e(n)$ a $s(n)$) se verificará la hipótesis de incorrelación.



Si por algún motivo, $w(n)$ y $x(n)$ son procesos incorrelados, la ecuación:

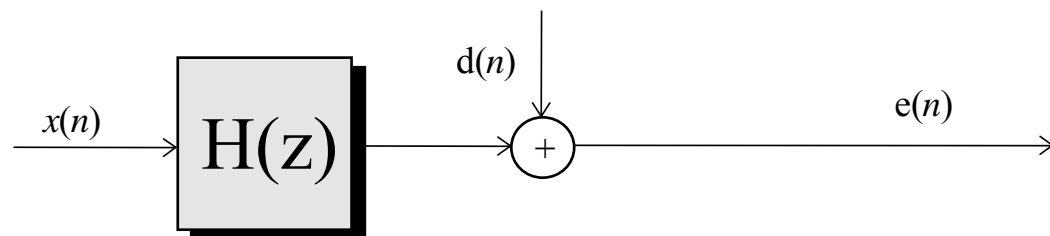
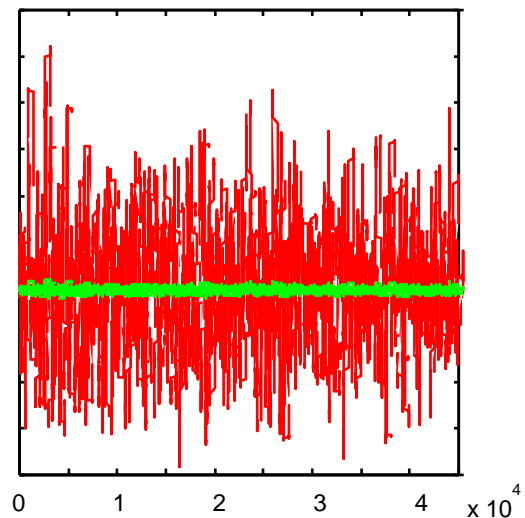
$$\mathbf{r}_{(s+w)x} - \mathbf{R}_{xx} \mathbf{h} = \mathbf{0}$$

nos dice que el término de la izquierda es nulo, y que por tanto, los coeficientes del filtro serán cero (siempre que la matriz de correlación no sea singular).

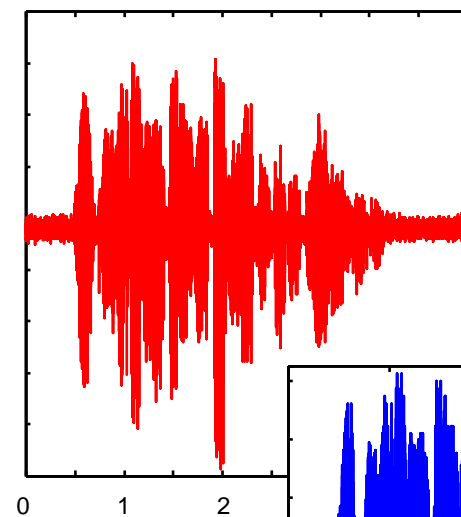


El filtro se comporta inteligentemente, de forma que no se añade más ruido a $e(n)$

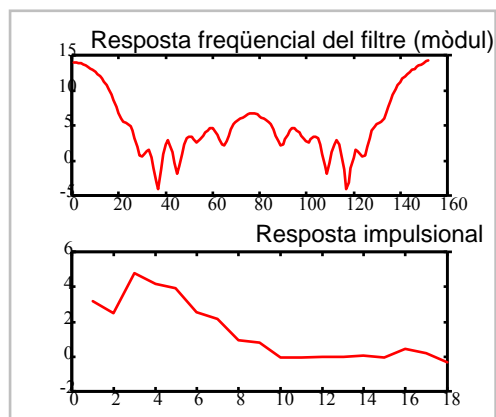
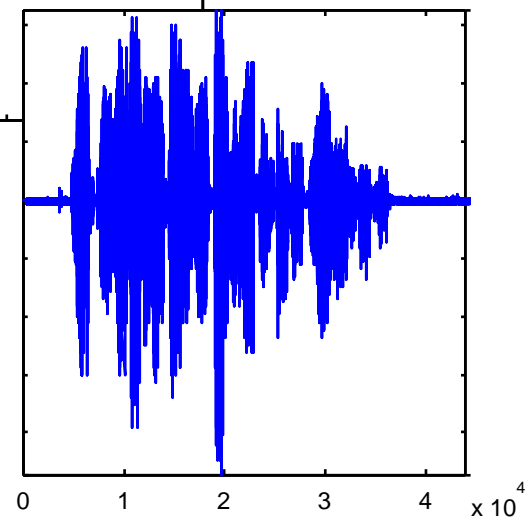
Senyal + interferencia



Senyal restaurat

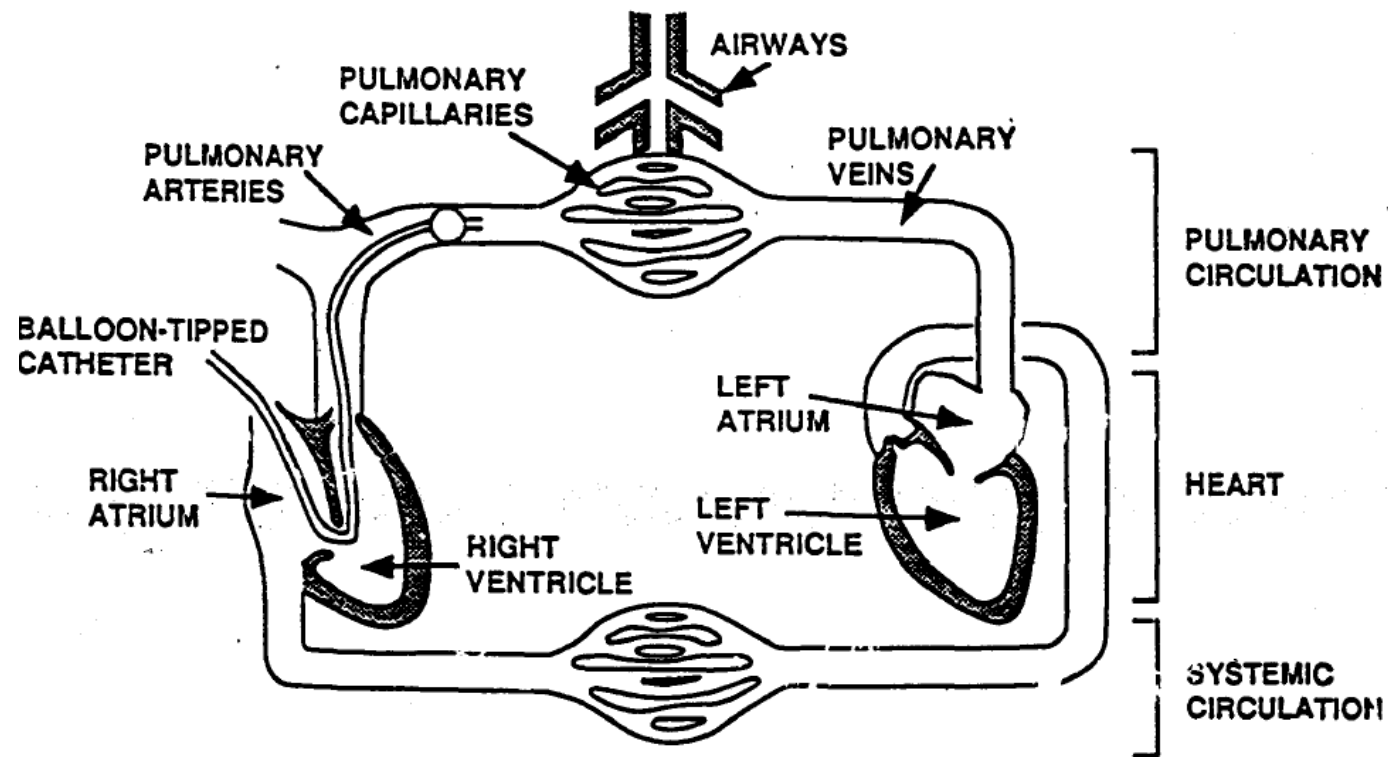


Senyal original

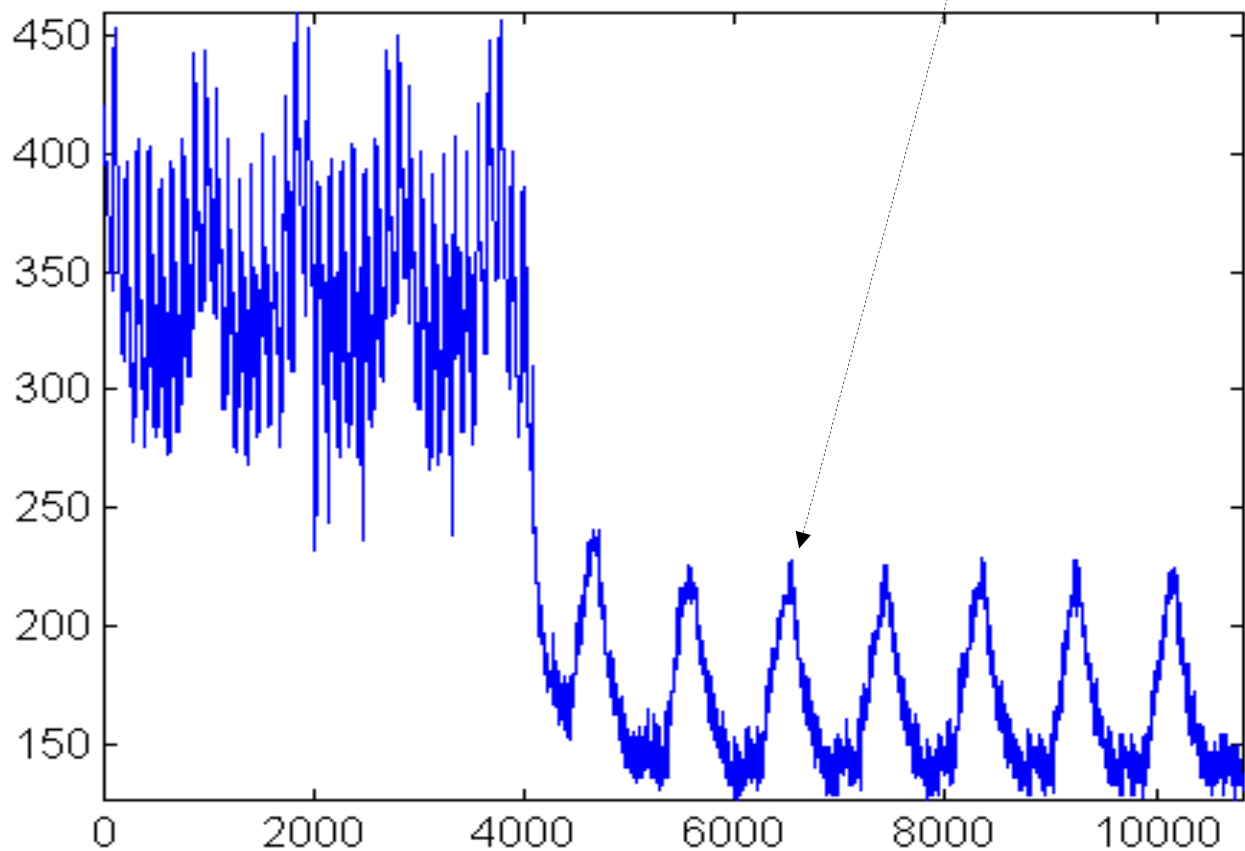


Filtro de Wiener

Ejemplo 5. Cancelación de interferencia en señales biológicas

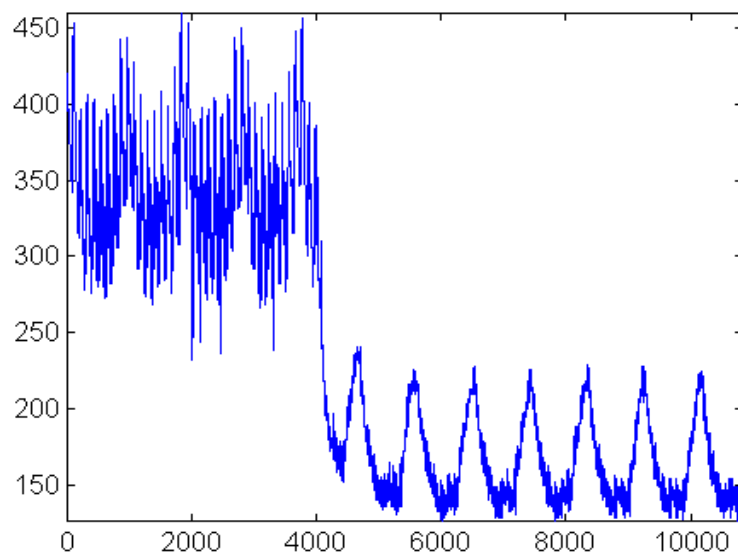


Modelo de señal: Señal exponencial decreciente + interferencia respiratoria + latido cardiaco

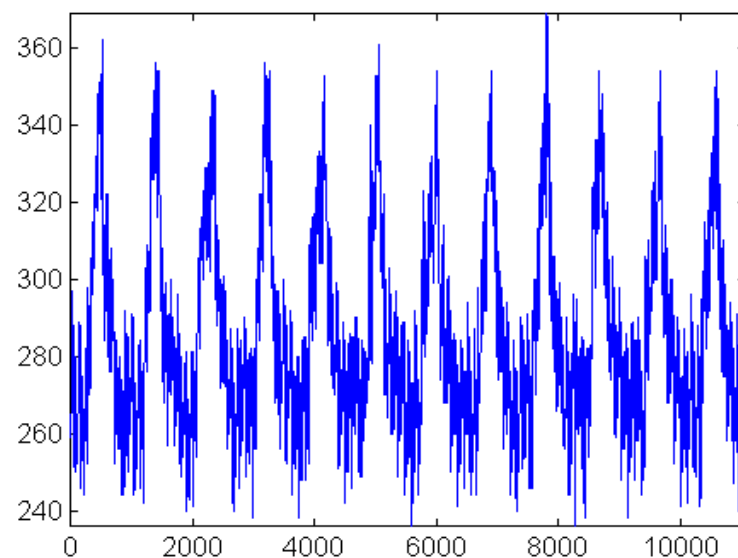


Señal de presión arterial ($f_m = 200$ Hz, cuantificada con 8 bits)

$d(n)$

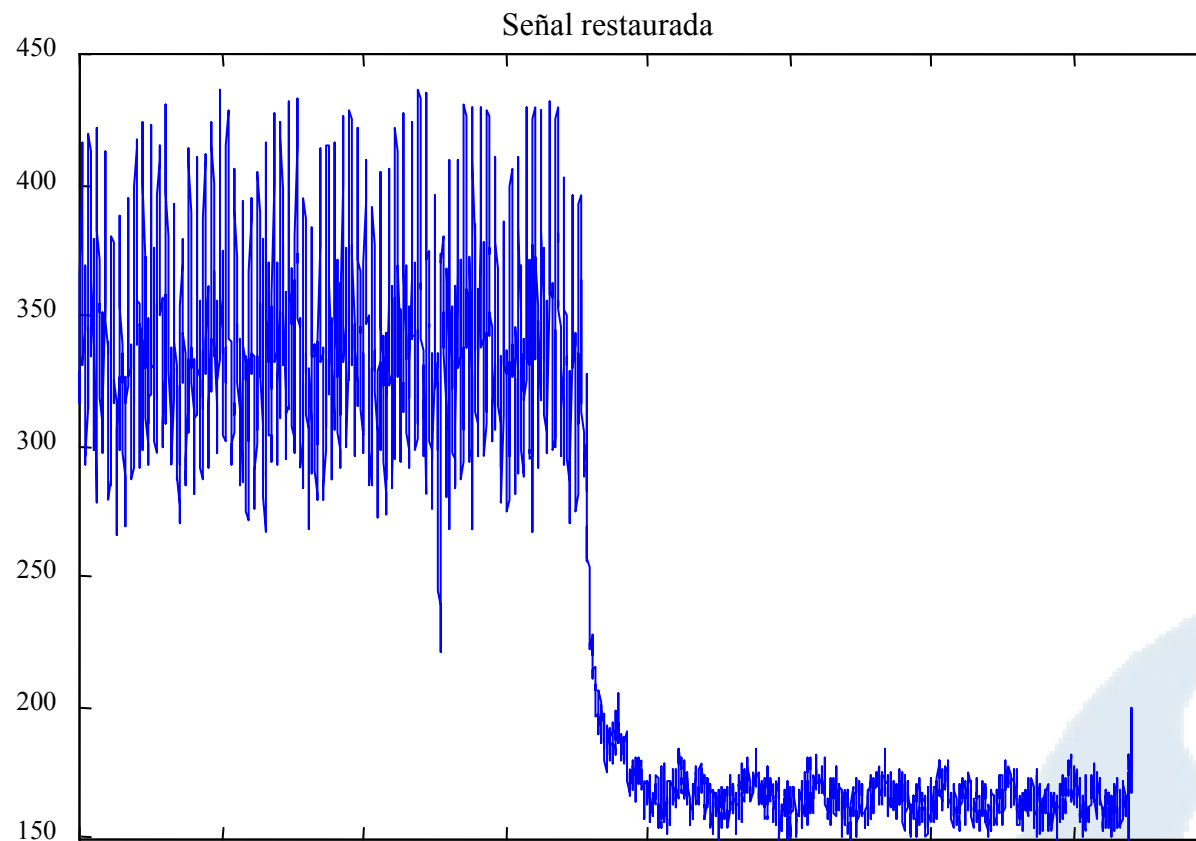


$x(n)$



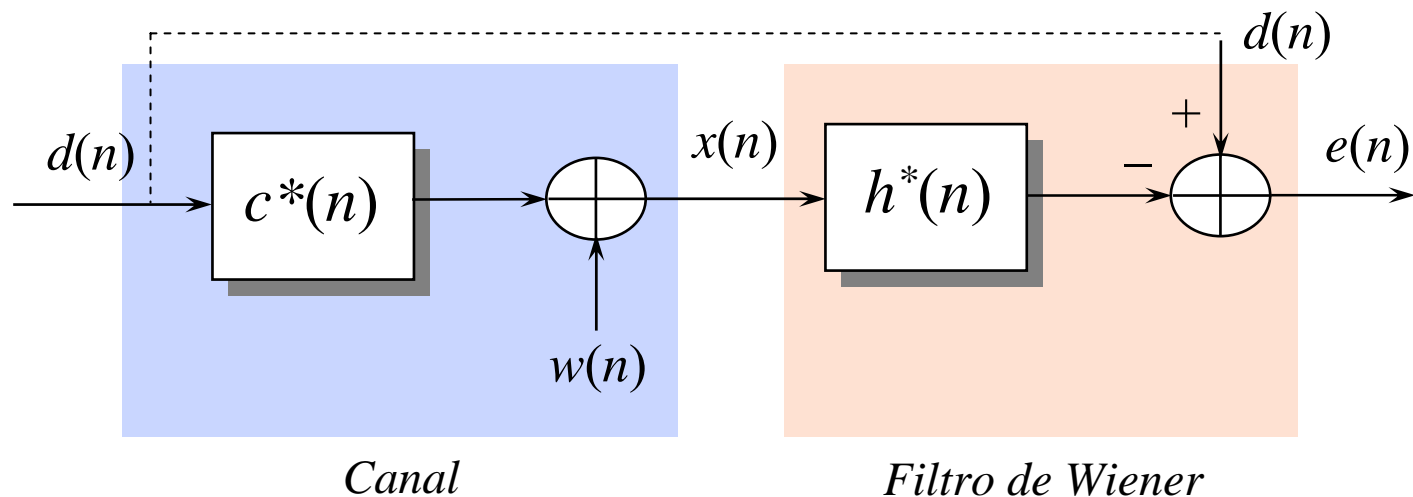
...medida en la arteria pulmonar

...medida en la aurícula derecha, y correlada con la interferencia respiratoria. Usando una versión filtrada de ésta, será posible eliminar la interferencia de la primera.



Sobre la señal limpia de interferencias es posible estimar con más precisión los parámetros del modelo (entre los que está la presión sanguínea en los capilares pulmonares).

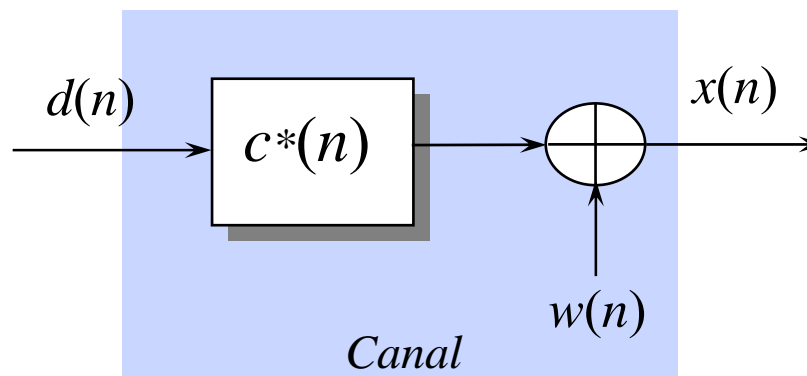
Ejemplo 6. Ecualizador de error cuadrático medio mínimo



Conocido el canal de propagación $c(n)$ y la correlación del ruido $r_w(k)$, diseñemos un filtro receptor $h^*(n)$ que minimice el error cuadrático medio entre la salida del filtro y los símbolos transmitidos $d(n)$:

$$\hat{\mathbf{h}} = \arg \min_{\mathbf{h}^*} E \left\{ \left| \mathbf{h}^H \mathbf{x}(n) - d(n) \right|^2 \right\}$$

Modelo de señal para $x(n)$



$$\begin{array}{c}
 \text{Columna } L\text{-ésima} \\
 \downarrow \\
 \begin{bmatrix} x(n) \\ x(n-1) \\ x(n-2) \\ \vdots \\ x(n-L+1) \end{bmatrix} = \begin{bmatrix} c^*(0) & c^*(1) & \dots & c^*(L-2) & c^*(L-1) & 0 & 0 & 0 & \dots & 0 \\ 0 & c^*(0) & c^*(1) & \dots & c^*(L-2) & c^*(L-1) & 0 & 0 & \dots & 0 \\ 0 & 0 & c^*(0) & c^*(1) & \dots & c^*(L-2) & c^*(L-1) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & c^*(0) & c^*(1) & \dots & c^*(L-2) & c^*(L-1) & \end{bmatrix} \begin{bmatrix} d(n) \\ d(n-1) \\ \vdots \\ d(n-L+2) \\ d(n-L+1) \\ d(n-L) \\ d(n-L-1) \\ \vdots \\ d(n-N-L+3) \\ d(n-N-L+2) \end{bmatrix} + \begin{bmatrix} w(n) \\ w(n-1) \\ w(n-2) \\ \vdots \\ w(n-L+1) \end{bmatrix}
 \end{array}$$

$$\mathbf{x}(n) = \mathbf{C}^H \mathbf{d}(n) + \mathbf{w}(n)$$

$$\mathbf{C}^H \in \mathbb{C}^{L \times (N+L-1)}$$

La solución para el filtro de Wiener es conocida: $\mathbf{h} = \mathbf{R}_x^{-1} \mathbf{r}_{dx}$

El filtro puede expresarse en función de \mathbf{C} y de \mathbf{R}_w :

- Si los símbolos transmitidos y el ruido están incorrelados

$$\begin{aligned} \mathbf{R}_x &= E \left\{ \mathbf{x}(n) \mathbf{x}^H(n) \right\} = \mathbf{C}^H E \left\{ \mathbf{d}(n) \mathbf{d}^H(n) \right\} \mathbf{C} + E \left\{ \mathbf{w}(n) \mathbf{w}^H(n) \right\} = \\ &= \mathbf{C}^H \mathbf{R}_d \mathbf{C} + \mathbf{R}_w \end{aligned}$$

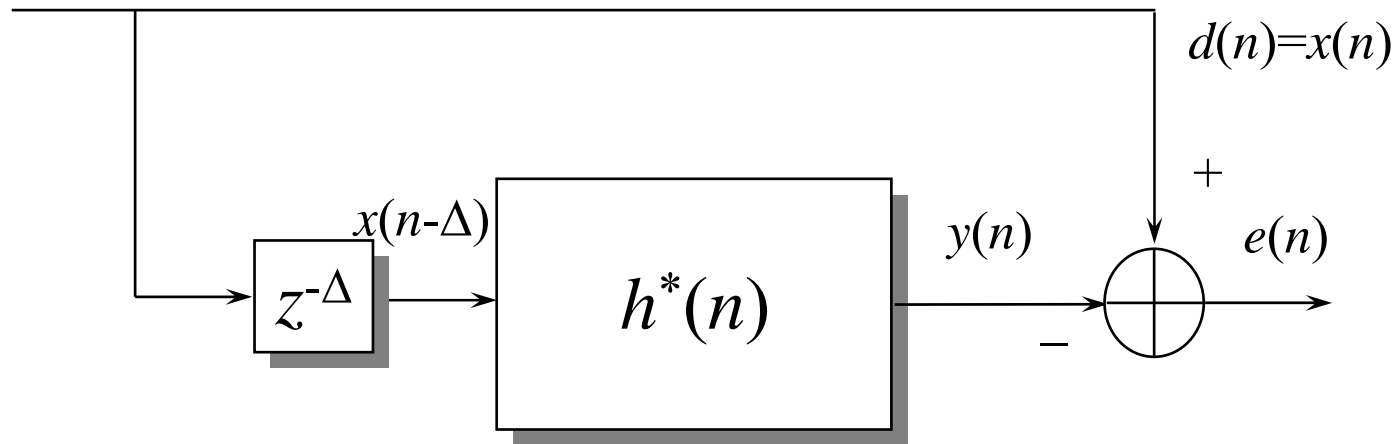
- Si los símbolos transmitidos están incorrelados

$$\mathbf{r}_{dx} = E \left\{ d^*(n-L+1) \mathbf{x}(n) \right\} = \sigma_d^2 \begin{bmatrix} c^*(L-1) \\ c^*(L-2) \\ \vdots \\ c^*(0) \end{bmatrix} = \sigma_d^2 \mathbf{c}^*$$

$$\mathbf{R}_d = \sigma_d^2 \mathbf{I}$$

$$\mathbf{h} = \left(\mathbf{C}^H \mathbf{C} + \frac{1}{\sigma_d^2} \mathbf{R}_w \right)^{-1} \mathbf{c}^*$$

- Casos límite: cuando $\text{SNR} \rightarrow \infty$ obtenemos el forzador de ceros
cuando $\text{SNR} \rightarrow 0$ obtenemos filtro adaptado



Particularizando las ecuaciones de Wiener-Hopf para $\Delta=1$

$$\left. \begin{aligned} \mathbf{R}_{xx} \mathbf{h} &= \mathbf{r}_x \\ \xi_{\min} &= r_d(0) - \mathbf{r}_x^H \mathbf{h} \end{aligned} \right\} \longrightarrow \begin{bmatrix} r_x(0) & r_x(1) & \cdots & r_x(Q) \\ r_x(-1) & r_x(0) & \cdots & r_x(Q-1) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(-Q) & r_x(-Q+1) & \cdots & r_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ -h(0) \\ \vdots \\ -h(Q-1) \end{bmatrix} = \begin{bmatrix} \xi_{\min} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\mathbf{r}_x = [r_x(-1) \ r_x(-2) \ \cdots \ r_x(-Q)]^T$$

- La señal más impredecible es el ruido blanco:
 - coeficientes del predictor nulos
 - potencia del error de predicción igual a la potencia de la señal a predecir
- La potencia del error de predicción siempre es menor o igual que la potencia de la señal

$$x(n) = \hat{x}(n) + e(n)$$

$$E\{|x(n)|^2\} = E\{|\hat{x}(n)|^2 + |e(n)|^2\} + 2 \operatorname{Re}\{E\{e^*(n)\hat{x}(n)\}\}$$

$$\xi_{\min} = E\{|x(n)|^2\} - E\{|\hat{x}(n)|^2\}$$

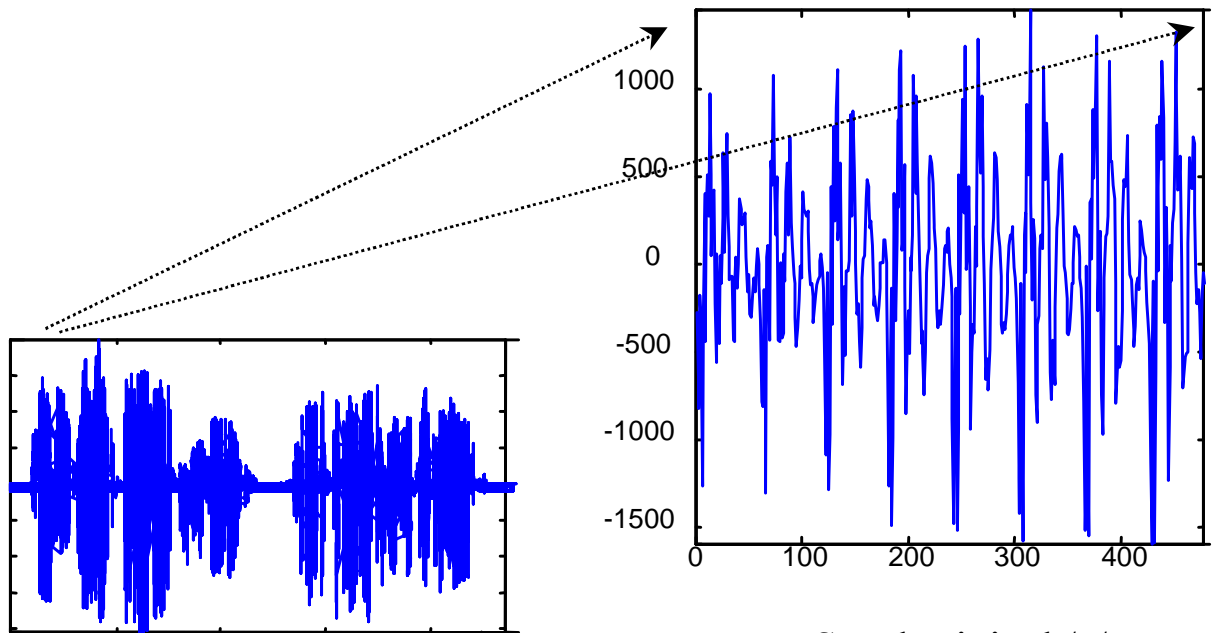
Cero, por el principio de ortogonalidad

- El filtro que genera el error de predicción es de fase mínima

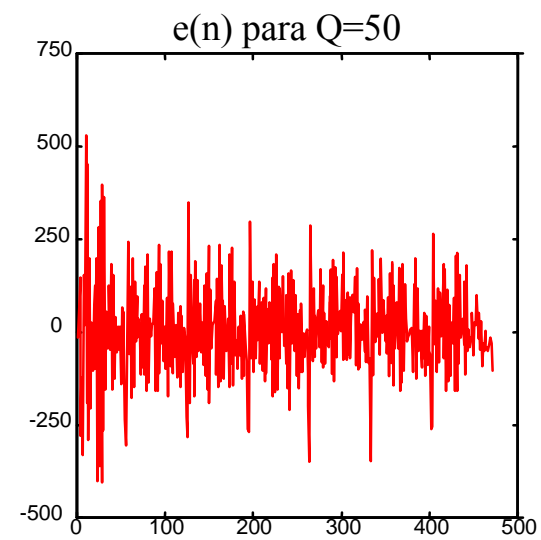
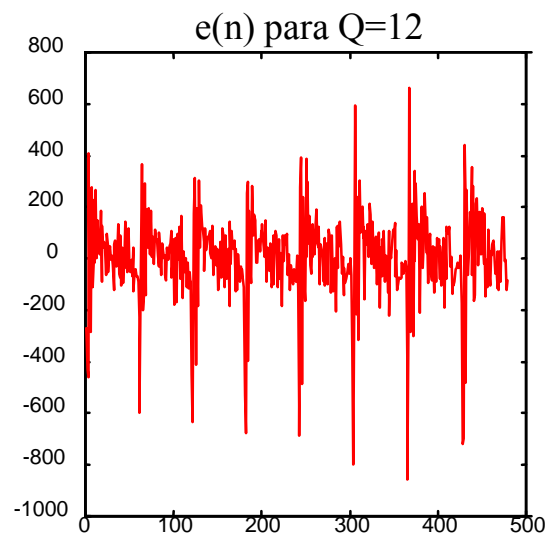
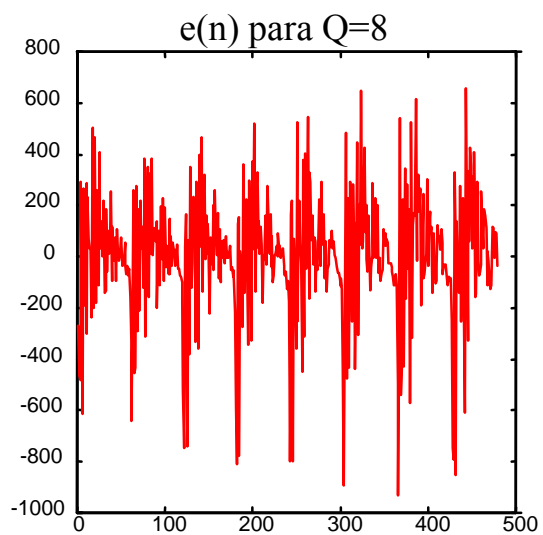
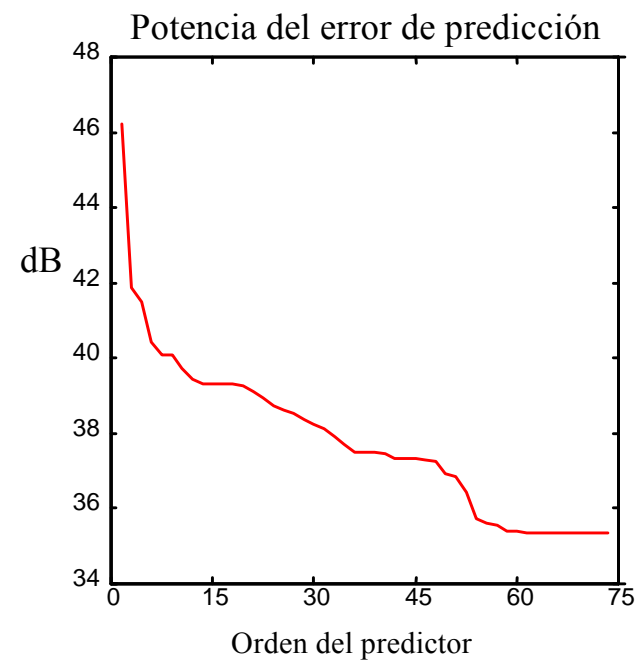
$$P(z) = 1 - h(0)^* z^{-1} - \dots - h(Q-1)^* z^{-Q+1}$$

Ejemplo 7. Predicción lineal de un proceso AR(p)

- El diseño de un predictor lineal es equivalente al ajuste de un modelo AR del proceso, si $Q-1 \geq p$
 - Ecuaciones de Yule-Walker (AR): $\mathbf{R}_x^T \mathbf{a} = -\mathbf{r}_x^*$
 - Predictor lineal: $\mathbf{R}_x \mathbf{h} = \mathbf{r}_x$
 - Equivalencia:
$$h(i) = \begin{cases} -a(i+1)^* & \text{si } 0 \leq i \leq p \\ 0 & \text{si } p \leq i \leq Q-1 \end{cases}$$
 - Filtro que genera el error de predicción $P(z) = 1 - h(0)^* z^{-1} - \dots - h(Q-1)^* z^{-Q+1}$
 - Estimador espectral AR
$$S(f) = \frac{\sigma^2}{|A(f)|^2} = \frac{\sigma^2}{|P(f)|^2}$$
- El error de predicción será blanco cuando el proceso a predecir sea AR y su orden cumpla $Q-1 \geq p$



Señal original $/e/$



Algoritmo de Levinson

- Explota la estructura Toeplitz de la matriz de correlación en el cálculo de la inversa: método de correlación
- Predictor lineal de Q coeficientes hacia delante (forward) :

$$\begin{bmatrix} r(0) & r(1) & \cdots & r(Q) \\ r(-1) & r(0) & \cdots & r(Q-1) \\ \vdots & \vdots & \ddots & \vdots \\ r(-Q) & r(-Q+1) & \cdots & r(0) \end{bmatrix} \begin{bmatrix} 1 \\ -h^Q(1) \\ \vdots \\ -h^Q(Q) \end{bmatrix} = \begin{bmatrix} \sigma_Q^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

- Al pasar a un predictor de orden Q+1 sólo aparece un nuevo valor de la correlación r(Q+1) y su conjugado r(-Q-1)

$$\begin{bmatrix} \boxed{\begin{matrix} r(0) & r(1) & \cdots & r(Q) \\ r(-1) & r(0) & \cdots & r(Q-1) \\ \vdots & \vdots & \ddots & \vdots \\ r(-Q) & r(-Q+1) & \cdots & r(0) \end{matrix}} & r(Q+1) \\ r(-Q-1) & \boxed{\begin{matrix} r(-Q) & \cdots & r(-1) \\ r(-Q) & \cdots & r(-1) \end{matrix}} \end{bmatrix} \begin{bmatrix} 1 \\ -h^{Q+1}(1) \\ \vdots \\ -h^{Q+1}(Q) \\ -h^{Q+1}(Q+1) \end{bmatrix} = \begin{bmatrix} \sigma_{Q+1}^2 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

- La estructura posibilita un procedimiento eficiente de cálculo de los coeficientes de un predictor a partir de los coeficientes del predictor de orden inferior
- Conjugando y dando la vuelta al vector de coeficientes se define la predicción hacia atrás (backward) con mismos coeficientes conjugados que la predicción hacia delante y mismo error de predicción

$$\begin{bmatrix} r(0) & r(1) & \cdots & r(Q) & r(Q+1) \\ r(1) & r(0) & \cdots & r(Q-1) & r(Q) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ r(-Q) & r(-Q+1) & \cdots & r(0) & r(1) \\ r(-Q-1) & r(-Q) & \cdots & r(-1) & r(0) \end{bmatrix} \begin{bmatrix} -h^{Q+1}(Q+1)^* \\ -h^{Q+1}(Q)^* \\ \vdots \\ -h^{Q+1}(1)^* \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \sigma_{Q+1}^2 \end{bmatrix}$$

- Si el proceso es estacionario es irrelevante el sentido de la predicción

- Se descompone el nuevo vector de coeficientes en una combinación del predictor forward y backward

$$\mathbf{a}_f^{Q+1} = \begin{bmatrix} 1 \\ -h^{Q+1}(1) \\ \vdots \\ -h^{Q+1}(Q) \\ -h^{Q+1}(Q+1) \end{bmatrix} = \begin{bmatrix} 1 \\ -h^Q(1) \\ \vdots \\ -h^Q(Q) \\ 0 \end{bmatrix} + K_{Q+1} \begin{bmatrix} 0 \\ -h^Q(Q)^* \\ \vdots \\ -h^Q(1)^* \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{a}_f^Q \\ 0 \end{bmatrix} + K_{Q+1} \begin{bmatrix} 0 \\ \mathbf{a}_b^Q \end{bmatrix}$$

$$\mathbf{a}_b^{Q+1} = \begin{bmatrix} -h^{Q+1}(Q+1)^* \\ -h^{Q+1}(Q)^* \\ \vdots \\ -h^{Q+1}(1)^* \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -h^Q(Q)^* \\ \vdots \\ -h^Q(1)^* \\ 1 \end{bmatrix} + K_{Q+1}^* \begin{bmatrix} 1 \\ -h^Q(1) \\ \vdots \\ -h^Q(Q) \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{a}_b^Q \end{bmatrix} + K_{Q+1}^* \begin{bmatrix} \mathbf{a}_f^Q \\ 0 \end{bmatrix}$$

$$\mathbf{R}_{Q+1} \mathbf{a}_f^{Q+1} = \begin{bmatrix} \sigma_{Q+1}^2 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \sigma_Q^2 \\ 0 \\ \vdots \\ 0 \\ \Delta_Q \end{bmatrix} + K_{Q+1} \begin{bmatrix} \Delta_Q^* \\ 0 \\ \vdots \\ 0 \\ \sigma_Q^2 \end{bmatrix}$$

$$\Delta_Q = r(-Q-1) - \sum_{q=1}^Q h^Q(q) r(-Q-1+q) = \begin{bmatrix} r(-Q-1) & r(-Q) & \cdots & r(0) \end{bmatrix} \begin{bmatrix} \mathbf{a}_f^Q \\ 0 \end{bmatrix}$$

Parcor (correlación parcial) $K_{Q+1} = -\frac{\Delta_Q}{\sigma_Q^2}$

Error de predicción $\sigma_{Q+1}^2 = \left(1 - |K_{Q+1}|^2\right) \sigma_Q^2$

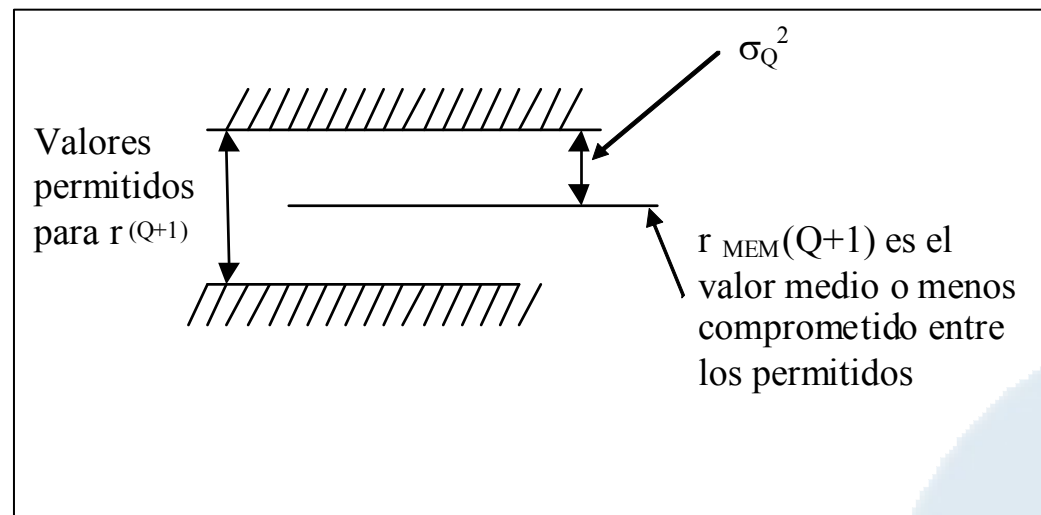
Coeficientes $h^{Q+1}(q) = h^Q(q) + K_{Q+1} h^Q(Q-q+1)^*$
 $h^{Q+1}(Q+1) = -K_{Q+1}$

Inicialización $\begin{bmatrix} r(0) & r(1) \\ r(-1) & r(0) \end{bmatrix} \begin{bmatrix} 1 \\ -h(1) \end{bmatrix} = \begin{bmatrix} \sigma_1^2 \\ 0 \end{bmatrix}$ $h^1(1) = \frac{r(-1)}{r(0)}$ $\sigma_1^2 = r(0) - \frac{|r(1)|^2}{r(0)}$

- El valor del parcor indica si la disminución del error al aumentar el orden del predictor es significativa

$$\sigma_{Q+1}^2 = \left(1 - |K_{Q+1}|^2\right) \sigma_Q^2 \geq 0 \Rightarrow |K_{Q+1}|^2 \leq 1$$

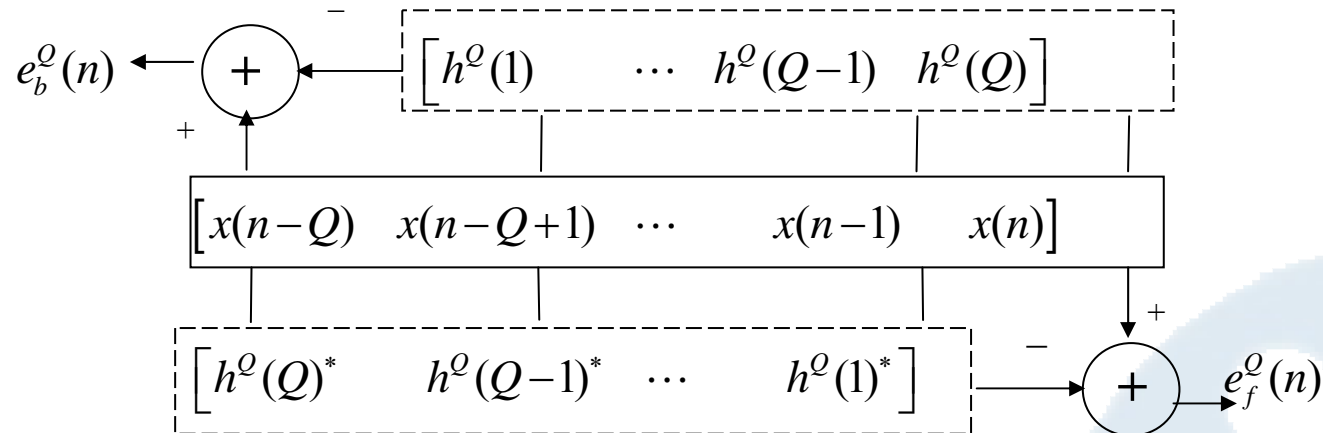
$$r(Q+1) = \Delta_Q^* + \sum_{q=1}^Q h^Q(q)^* r(Q+1-q) = -K_{Q+1}^* \sigma_Q^2 + r_{MEM}(Q+1)$$



- Los coeficientes parcors definen la correlación de los datos y permiten su extrapolación.
- Secuencia de infinitos valores de correlación de duración infinita pueden resumirse en pocos valores de parcors. Si el proceso es $AR(Q)$, sólo Q parcors no nulos definen la autocorrelación de duración infinita

Estructura en celosía (lattice)

- Se definen los errores forward y backward a partir de un segmento de datos de $Q+1$ muestras
 - Error forward en el instante n es el error al predecir $x(n)$
 - Error backward en el instante n es el error al predecir hacia atrás $x(n-Q)$



$$e_f^Q(n) = \mathbf{a}_f^{QH} \begin{bmatrix} x(n) \\ x(n-1) \\ \vdots \\ x(n-Q) \end{bmatrix}$$

$$e_b^Q(n) = \mathbf{a}_b^{QH} \begin{bmatrix} x(n) \\ x(n-1) \\ \vdots \\ x(n-Q) \end{bmatrix}$$

- Error forward para el orden $Q+1$:

$$e_f^{Q+1}(n) = \mathbf{a}_f^{Q+1H} \begin{bmatrix} x(n) \\ \vdots \\ x(n-Q) \\ x(n-Q-1) \end{bmatrix}$$

- El nuevo error usa los mismos datos que el error forward de orden Q , en el instante n , y el error backward en el instante $n-1$
- Recursión de Levinson:

$$e_f^{Q+1}(n) = \begin{bmatrix} \mathbf{a}_f^{QH} & 0 \end{bmatrix} \begin{bmatrix} x(n) \\ \vdots \\ x(n-Q) \\ x(n-Q-1) \end{bmatrix} + K_{Q+1}^* \begin{bmatrix} 0 & \mathbf{a}_b^{QH} \end{bmatrix} \begin{bmatrix} x(n) \\ \vdots \\ x(n-Q) \\ x(n-Q-1) \end{bmatrix} = e_f^Q(n) + K_{Q+1}^f e_b^Q(n-1)$$

- Se define el parcor forward que relaciona los errores forward y backward cuando se incrementa el orden del predictor

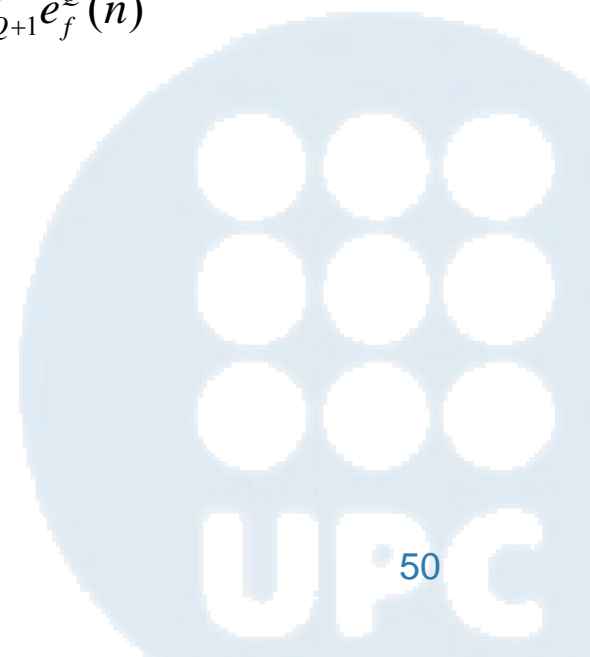
- El parcor forward óptimo puede derivarse minimizando la potencia del error forward de orden $Q+1$ dados los errores anteriores (Wiener de 1 coeficiente)

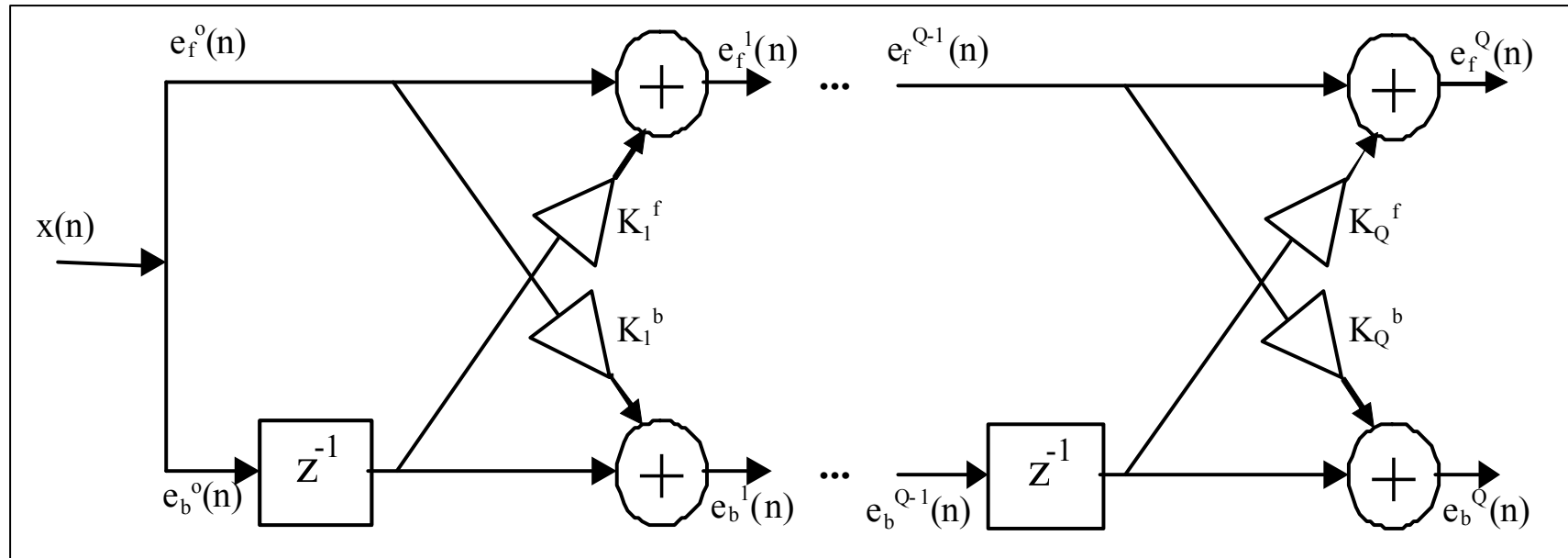
$$K_{Q+1}^f = -\frac{E\{e_f^Q(n)e_b^Q(n-1)^*\}}{E\{|e_b^Q(n-1)|^2\}}$$

- Análogamente para el error backward

$$e_b^{Q+1}(n) = e_b^Q(n-1) + K_{Q+1}e_f^Q(n) = e_b^Q(n-1) + K_{Q+1}^b e_f^Q(n)$$

$$K_{Q+1}^b = -\frac{E\{e_f^Q(n)e_b^Q(n-1)^*\}}{E\{|e_f^Q(n)|^2\}}$$





- Al incrementar el orden del predictor no se rehace el filtro, sólo se añade una sección
- Diseño de un predictor de orden Q igual al diseño de filtros de orden 1, añadiendo secciones se monitoriza el error de predicción
- Si los paráctores son en módulo menores que la unidad se garantiza que el FIR es de fase mínima

- Si el proceso es estacionario, la potencia del error forward y backward serán iguales y los coeficientes parcors también
- Si el número de datos es finito las estimaciones de los parcors forward y backward no coinciden, se elegirá:

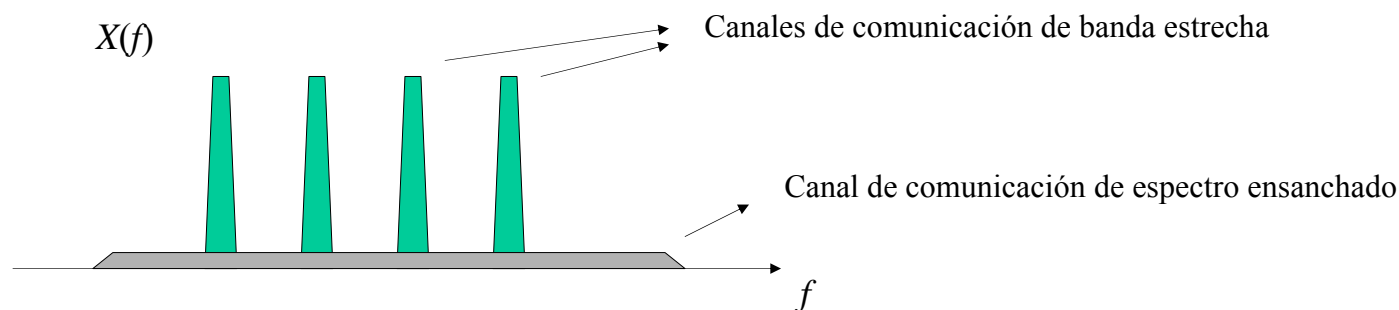
$$K_{Q+1} = \frac{2}{\frac{1}{K_{Q+1}^f} + \frac{1}{K_{Q+1}^b}} = - \frac{2E\{e_f^Q(n)e_b^Q(n-1)^*\}}{E\{|e_f^Q(n)|^2\}E\{|e_b^Q(n-1)|^2\}}$$

- Esta expresión siempre proporciona parcors en módulo inferiores a la unidad independientemente del número de datos

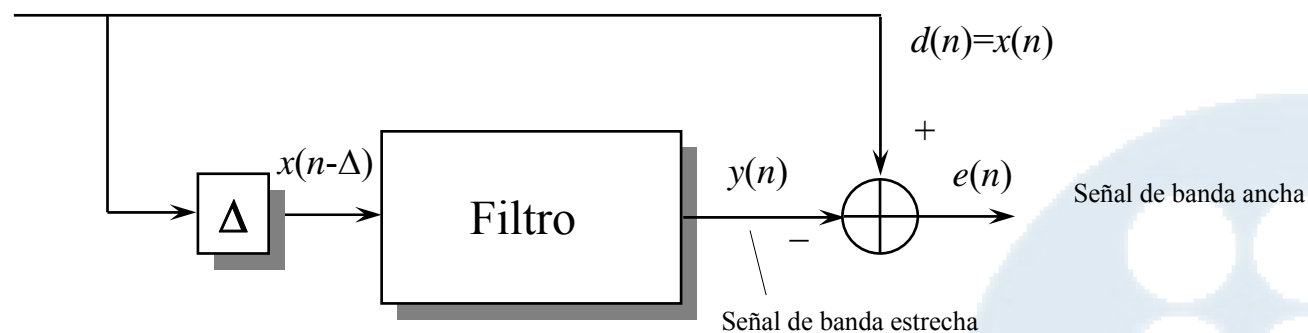


Ejemplo 8. Adaptive Line Enhancer

Sistemas de comunicación trabajando en *overlay*:

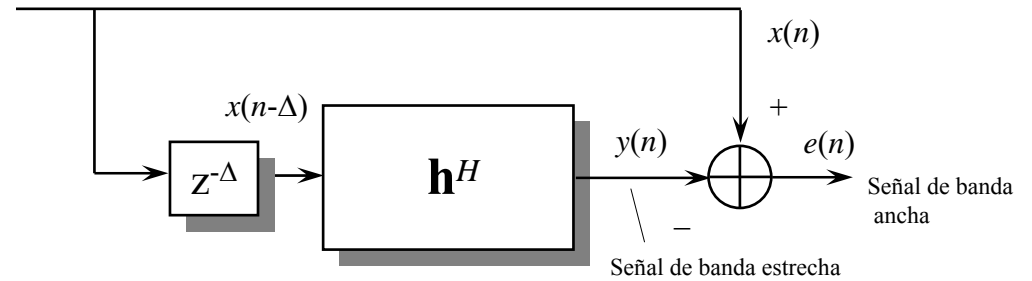
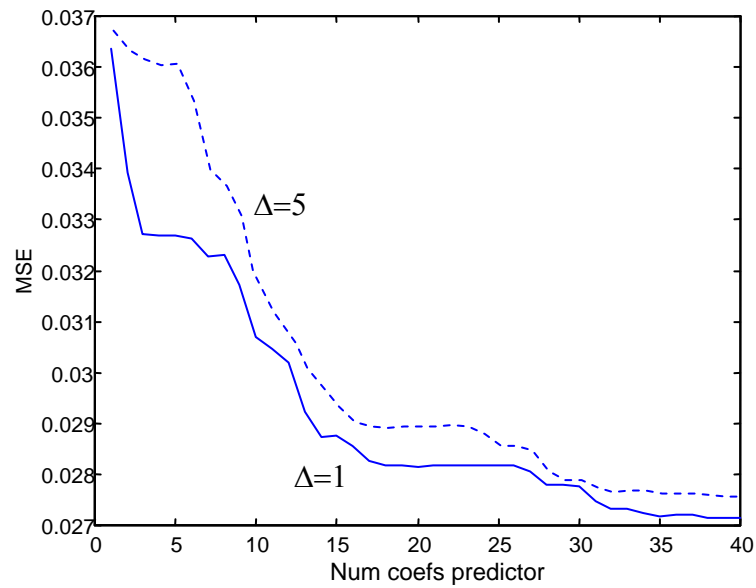
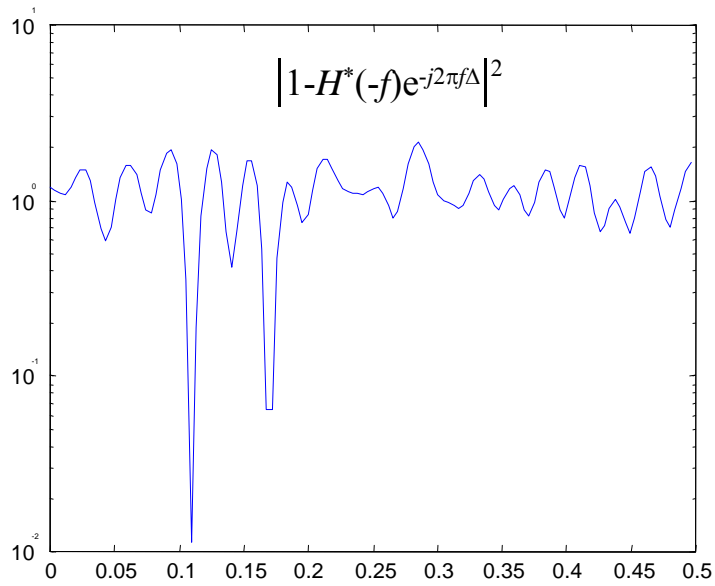


Si las señales de banda estrecha están incorreladas con la señal de banda ancha pueden separarse



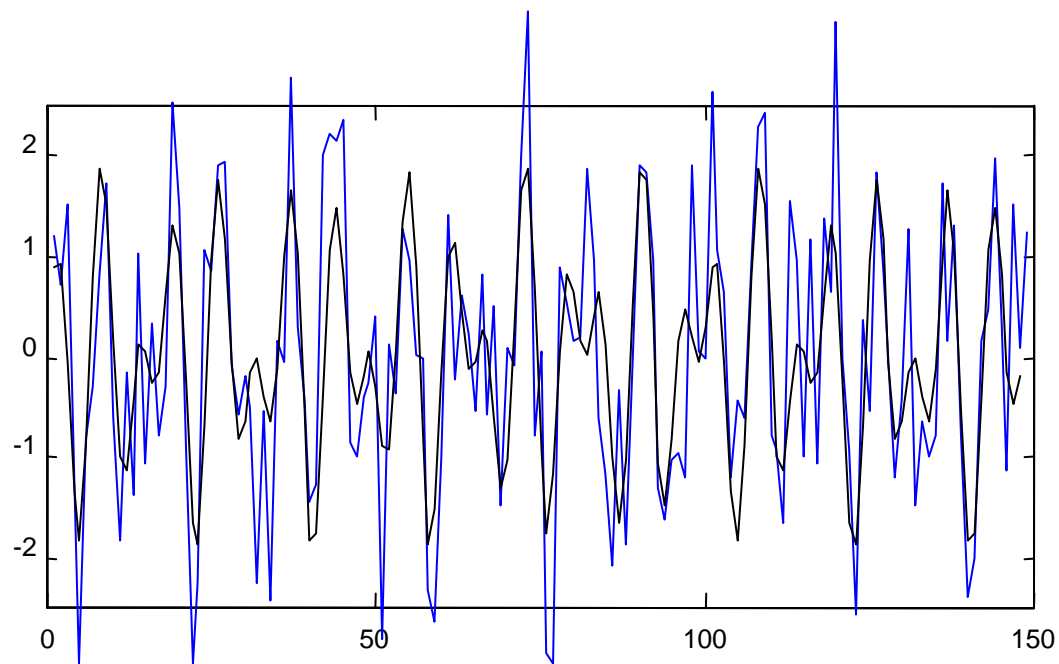
Si la señal de banda ancha tiene una potencia muy inferior a las señales de banda estrecha:

1. La calidad de los canales de banda estrecha no se ve afectada por la presencia de la señal de banda ancha y por tanto,
2. Se puede reutilizar el espectro radioeléctrico.



$x(n)$: Dos sinusoides a frecuencias
0.11 y 0.17 más ruido blanco
 SNR : 0 dB

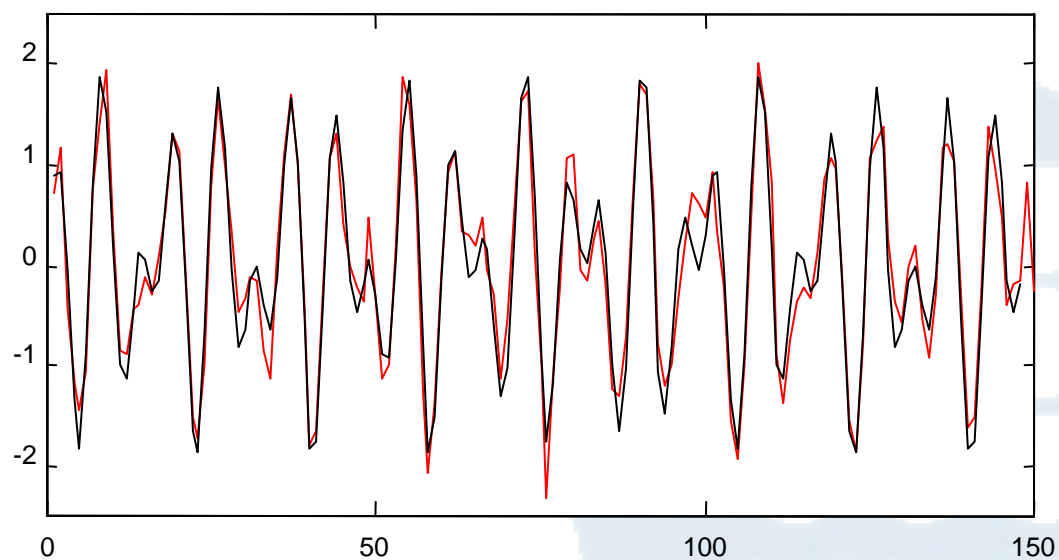
Azul: Señal + ruido $x(n)$
 Negro: Señal sin ruido



$$Q = 40$$

$$\Delta = 1$$

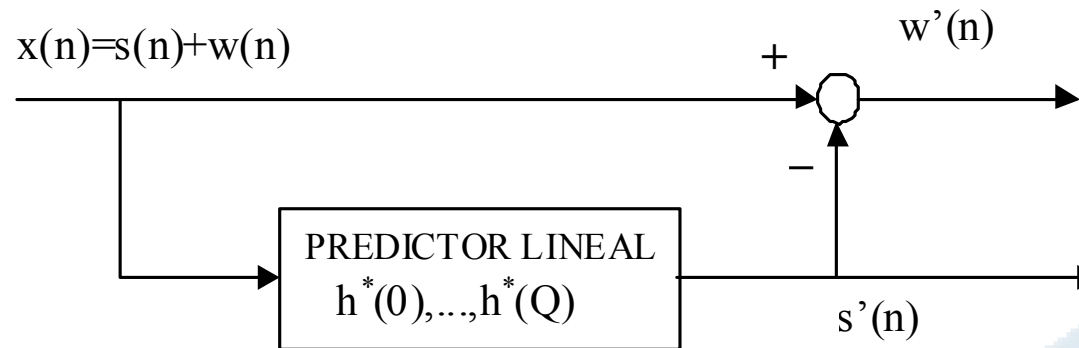
Rojo: Estimación
 de la señal, $y(n)$
 Negro: Señal sin ruido



Ejemplo 9. Recuperación de portadora en ruido blanco

- Una señal sinusoidal es perfectamente predecible
- El ruido blanco es impredecible

$$s(n) = A \exp(j\omega_o n + \theta) = A_o \exp(j\omega_o n)$$



$$\hat{x}(n) = \begin{bmatrix} h(1)^* & \dots & h(Q)^* \end{bmatrix} \begin{bmatrix} x(n-1) \\ \vdots \\ x(n-Q) \end{bmatrix}$$

$$\hat{x}(n) = \mathbf{h}^H \mathbf{x}(n-1) = \mathbf{h}^H (A_o \exp(j\omega_o n) \mathbf{s} + \mathbf{w}(n-1))$$

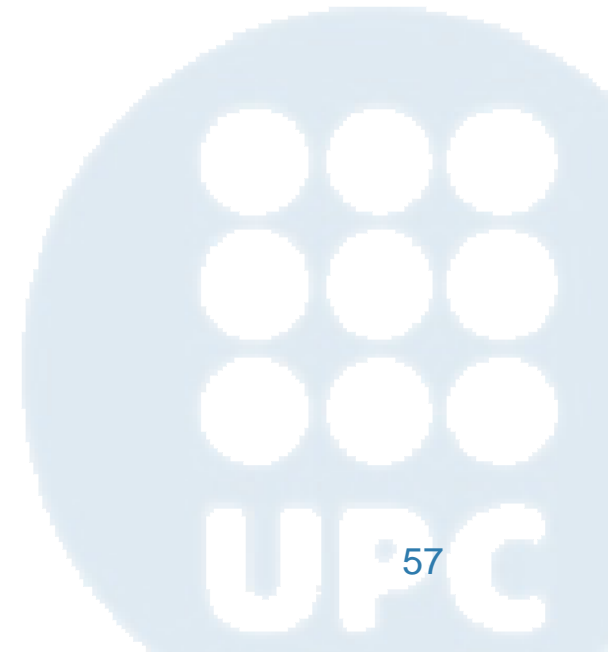
$$\mathbf{s} = \begin{bmatrix} \exp(-j\omega_o) \\ \vdots \\ \exp(-j\omega_o Q) \end{bmatrix}$$

- Planteamiento alternativo del predictor como minimización de la potencia del error de predicción

$$\xi = E \left\{ \left[\begin{array}{cccc} h(0)^* & h(1)^* & \cdots & h(Q)^* \end{array} \right] \begin{bmatrix} x(n) \\ x(n-1) \\ \vdots \\ x(n-Q) \end{bmatrix} \right\}^2 = \mathbf{h}^H \mathbf{R} \mathbf{h} \Big|_{\min}$$


- Vector de coeficientes de dimensión $Q+1$.
- Para evitar la solución trivial se impone la restricción de primer coeficiente igual a la unidad

$$\mathbf{h}^H \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{h}^H \boldsymbol{\delta} = 1$$



$$L = \mathbf{h}^H \mathbf{R} \mathbf{h} - \lambda (\mathbf{h}^H \boldsymbol{\delta} - 1)$$

$$\frac{\partial L}{\partial \mathbf{h}^H} = \mathbf{R} \mathbf{h} - \lambda \boldsymbol{\delta} = \mathbf{0} \Rightarrow \mathbf{h} = \lambda \mathbf{R}^{-1} \boldsymbol{\delta}$$

$$\lambda \boldsymbol{\delta}^T \mathbf{R}^{-1} \boldsymbol{\delta} = 1 \Rightarrow \lambda = \frac{1}{\boldsymbol{\delta}^T \mathbf{R}^{-1} \boldsymbol{\delta}}$$


$$\mathbf{h} = \frac{\mathbf{R}^{-1} \boldsymbol{\delta}}{\boldsymbol{\delta}^T \mathbf{R}^{-1} \boldsymbol{\delta}}$$

$$\xi_{\min} = \frac{1}{\boldsymbol{\delta}^T \mathbf{R}^{-1} \boldsymbol{\delta}}$$

$$\mathbf{R} = |A|^2 \mathbf{S} \mathbf{S}^H + \sigma^2 \mathbf{I} \quad \mathbf{R}^{-1} = \sigma^{-2} \left[\mathbf{I} - \frac{|A|^2}{\sigma^2 + (Q+1)|A|^2} \mathbf{S} \mathbf{S}^H \right]$$

$$\xi_{\min} = \sigma^2 \frac{\sigma^2 + (Q+1)|A|^2}{\sigma^2 + Q|A|^2} = \sigma^2 \frac{1 + (Q+1)SNR_{in}}{1 + QSNR_{in}}$$

$$SNR_{out} = Q SNR_{in}$$

Se obtiene la ganancia de un filtro adaptado a la sinusoide de manera incoherente

Ejemplo 10. Cancelador de ecos en comunicaciones dúplex

- Problema de aislamiento entre locutores en el paso de la conexión de dos a cuatro hilos.

