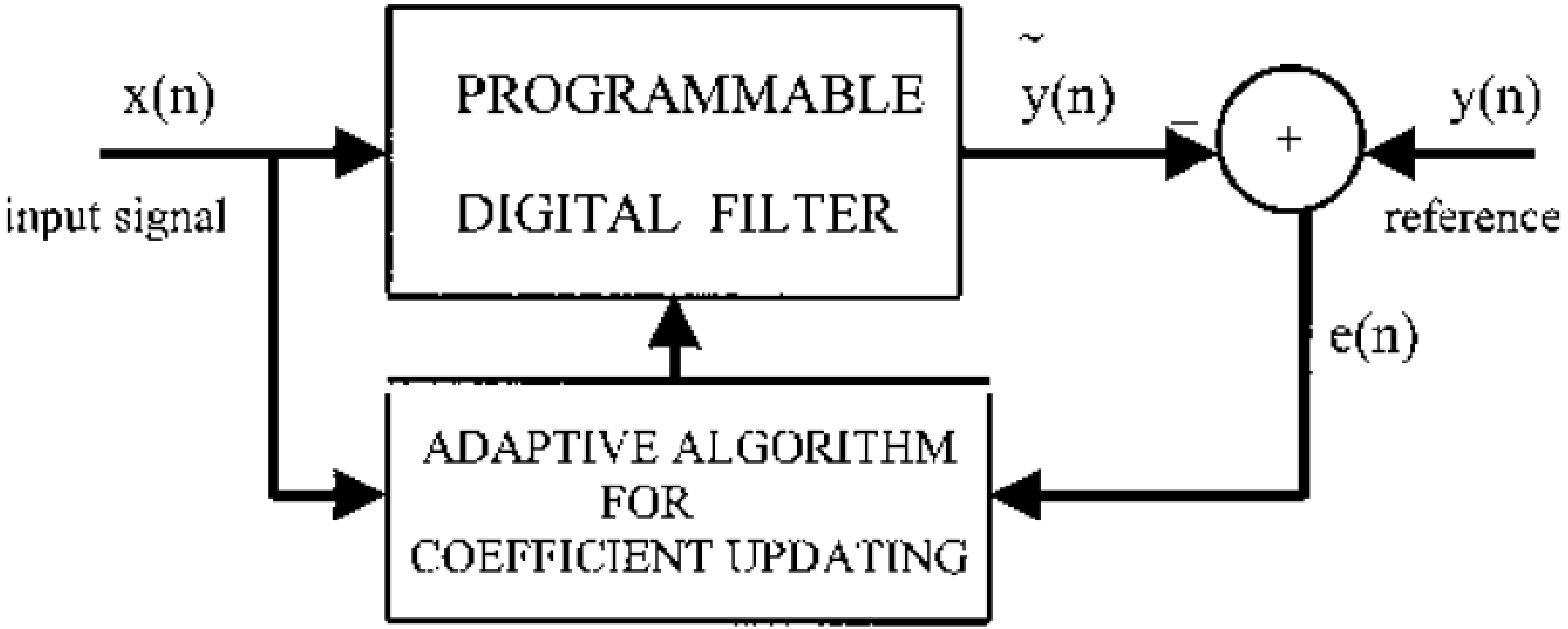
IC 512 Procesamiento Digital de Señales

Unidad 4: Filtrado Adaptativo y Aplicaciones

Filtros Adaptativos y sus Aplicaciones

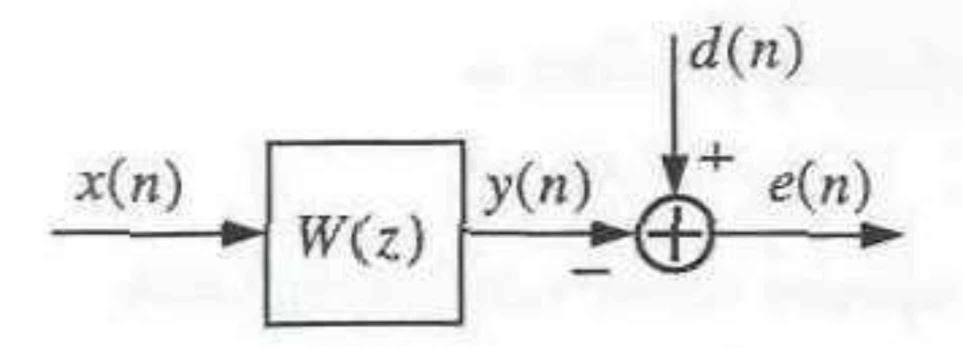
- Esta unidad introduce a los conceptos del filtrado adaptativo (el algoritmo LMS de forma específica) e ilustra como aplicarlos a la resolución de problemas.
- Un filtro adaptativo es un filtro digital con la característica de autoajuste de coeficientes.
- El ajuste de coeficientes se hace mediante un algoritmo de adaptación o algoritmo adaptativo

Esquema General de un Filtro Adaptativo



Teoría del filtro de Wiener y el algoritmo de los mínimos cuadrados medios

- Muchos filtros adaptativos pueden verse como aproximaciones al filtro de Wiener discreto
- El filtro de Wiener ajusta los coeficientes para producir una salida del filtro $y(n)$, que debería ser lo más parecida a una señal deseada, $d(n)$. Así, al substraer la salida y la señal deseada, el resultado es un error comparable al nivel de ruido $\eta(n)$



Filtro de Wiener

- Considerando el caso de un solo coeficiente,

$$y(n) = wx(n)$$

la señal de error está dada por: $e(n) = d(n) - wx(n)$

Tomando como estrategia la minimización del error cuadrático medio:

$$e^2(n) = [d(n) - wx(n)]^2 = d^2(n) - 2d(n)wx(n) + w^2x^2(n)$$

$$E\{e^2(n)\} = E\{d^2(n)\} - 2wE\{d(n)x(n)\} + w^2E\{x^2(n)\}$$

(E{.} es el valor esperado)

Filtro de Wiener

- Haciendo una interpretación estadística

$$J = E\{e^2(n)\}: \text{MSE (error cuadrático medio)}$$

$$\sigma^2 = E\{d^2(n)\}: \text{ La potencia de la señal deseada}$$

$$P = E\{d(n)x(n)\}: \text{ Correlación cruzada entre } d(n) \text{ y } x(n)$$

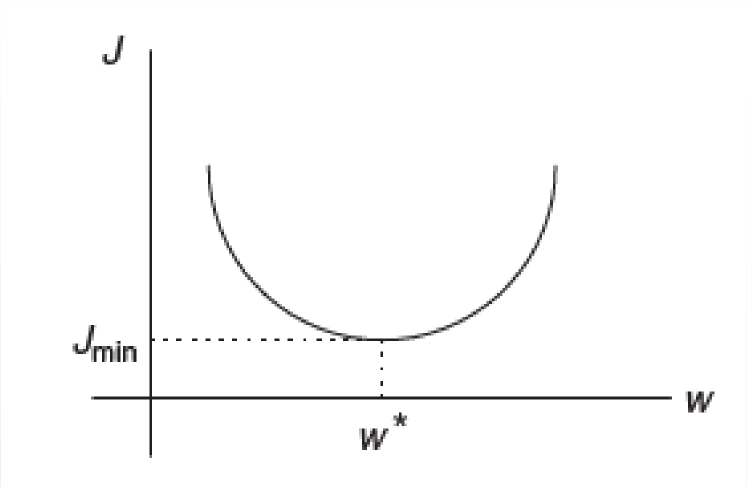
$$R = E\{x^2(n)\}: \text{ Autocorrelación de la señal de entrada}$$

- Se llega a la ecuación:

$$J = \sigma^2 - 2wP + w^2 R$$

Filtro de Wiener

- Función costo del error cuadrático medio:

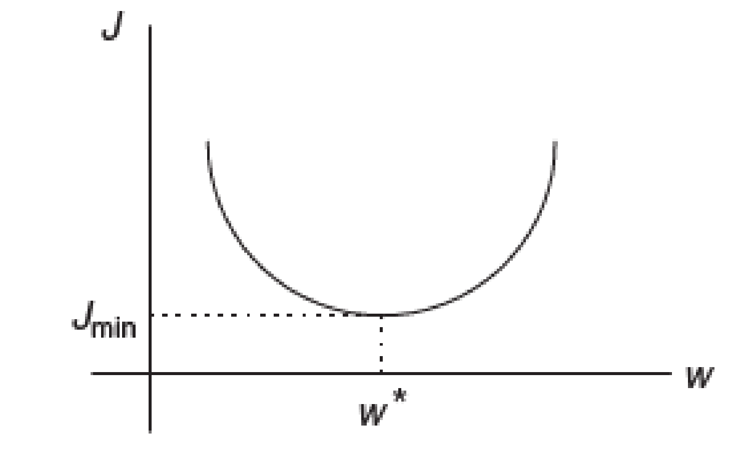


- Función cuadrática que presenta un mínimo global

Filtro de Wiener

- Ejemplo: Dado

$$J = 40 - 20w + 10w^2$$



encontrar la solución óptima

Solución:

Tomando la derivada del MSE e igualando a cero: $\frac{dJ}{dw} = -20 + 10.2w = 0 \rightarrow w_{op} = 1$

Finalmente, se puede obtener el error mínimo:

$$J_{\min} = J\Big|_{w=w_{op}} = 40 - 20.1 + 10.1 = 30$$

Filtro de Wiener

- Considerando el caso de varios coeficientes,

$$y(n) = \sum_{k=0}^{N-1} w_k x(n-k) = \mathbf{w}^{\mathrm{T}}\mathbf{x}(n), \text{ con:} \begin{cases} \mathbf{w} = [w_0 \quad w_1 \quad \ldots \quad w_{N-1}]^{\mathrm{T}} \\ \mathbf{x}(n) = [x(n)\ x(n-1) \ldots\ x(n-N+1]^{\mathrm{T}} \end{cases}$$

la señal de error está dada por: $e(n) = d(n) - \mathbf{w}^{\mathrm{T}}\mathbf{x}(n)$

Tomando como estrategia la minimización del error cuadrático medio:

$$e^2(n) = [d(n) - \mathbf{w}^{\mathrm{T}}\mathbf{x}(n)]^2 = d^2(n) - 2d(n)\mathbf{w}^{\mathrm{T}}\mathbf{x}(n) + \mathbf{w}^{\mathrm{T}}\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)\mathbf{w}$$

$$E\{e^2(n)\} = E\{d^2(n)\} - 2\mathbf{w}^{\mathrm{T}}E\{d(n)\mathbf{x}(n)\} + \mathbf{w}^{\mathrm{T}}E\{\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)\}\mathbf{w}$$

(E{.} es el valor esperado)

Filtro de Wiener

- Haciendo una interpretación estadística

$$J = E\{e^2(n)\} : \text{MSE (error cuadrático medio)}$$

$$\sigma^2 = E\{d^2(n)\} : \text{ potencia de la señal deseada}$$

$$\mathbf{p} = E\{d(n)\mathbf{x}(n)\} : \text{ vector de correlación cruzada entre } d(n) \text{ y } x(n)$$

$$\mathbf{R} = E\{\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)\} : \text{ matriz de autocorrelación de la señal de entrada}$$

- Se llega a la ecuación:

$$J = \sigma^2 - 2\mathbf{w}^{\mathrm{T}}\mathbf{p} + \mathbf{w}^{\mathrm{T}}\mathbf{R}\mathbf{w}$$

Filtro de Wiener

- Aplicando el gradiente e igualando a cero:

$$J = \sigma^2 - 2\mathbf{w}^T\mathbf{p} + \mathbf{w}^T\mathbf{R}\mathbf{w} \longrightarrow \nabla J = 2\mathbf{R}\mathbf{w} - 2\mathbf{p} = 0 \rightarrow \mathbf{w}_{op} = \mathbf{R}^{-1}\mathbf{p}$$

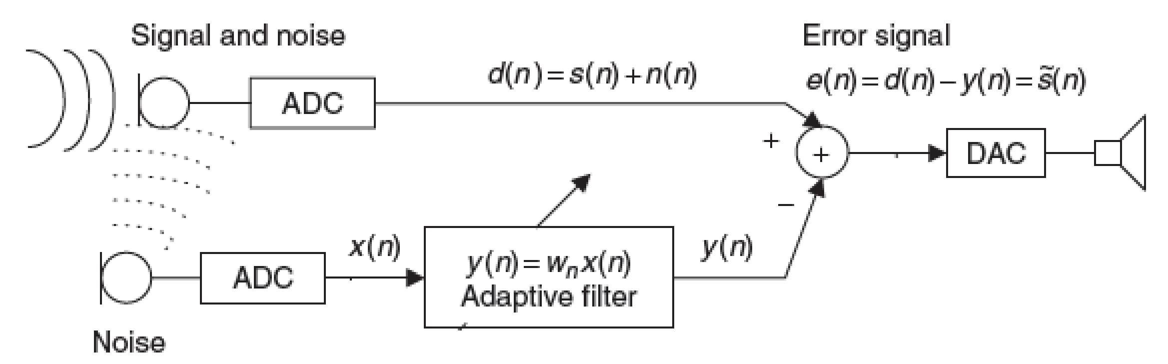
La solución óptima se basa en estadísticas

Los coeficientes óptimos pueden ser diferentes para cada bloque de datos (la autocorrelación y correlación cruzada pueden ir cambiando)

La estimación estadística con bloques suficientemente grandes puede causar un atraso que imposibilite la implementación en tiempo real.

Si se usa un número grande de coeficientes, se requiere del cálculo de la inversa de una matriz de autocorrelación que puede estar mal condicionada

Ejemplo: Cancelación de Ruido



Se supone que $x(n)$ es una señal relacionada solamente al ruido, $\eta(n)$.

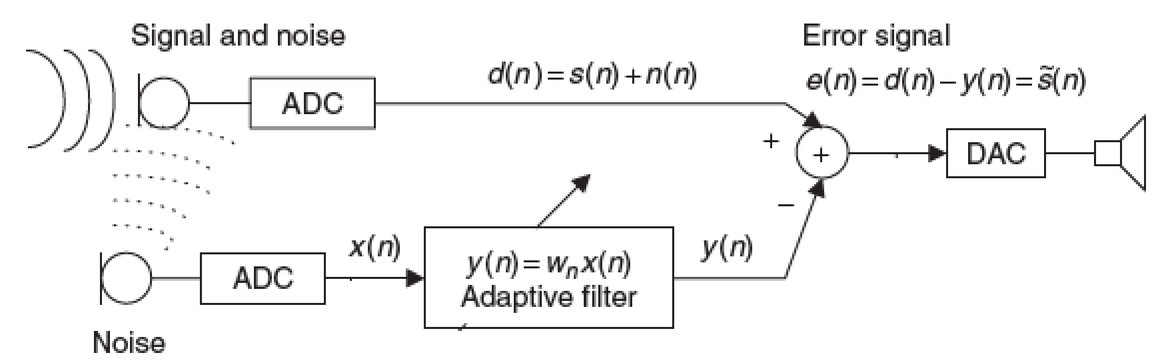
Este sistema funciona si resulta válida la relación: $\eta(n) = \mathbf{w}_{\mathrm{op}}^{\mathrm{T}}\mathbf{x}(n)$, si es así:

$$e(n) = d(n) - y(n) = s(n) + \eta(n) - y(n) = s(n) + \mathbf{w}_{\mathrm{op}}^{\mathrm{T}}\mathbf{x}(n) - \mathbf{w}^{\mathrm{T}}\mathbf{x}(n)$$

Si $\mathbf{w} = \mathbf{w}_{\mathrm{op}}$, el error corresponde a la señal libre de ruido.

Cómo estimar los coeficientes?

Ejemplo: Cancelación de Ruido



Cómo estimar los coeficientes?... La propuesta de Winner es

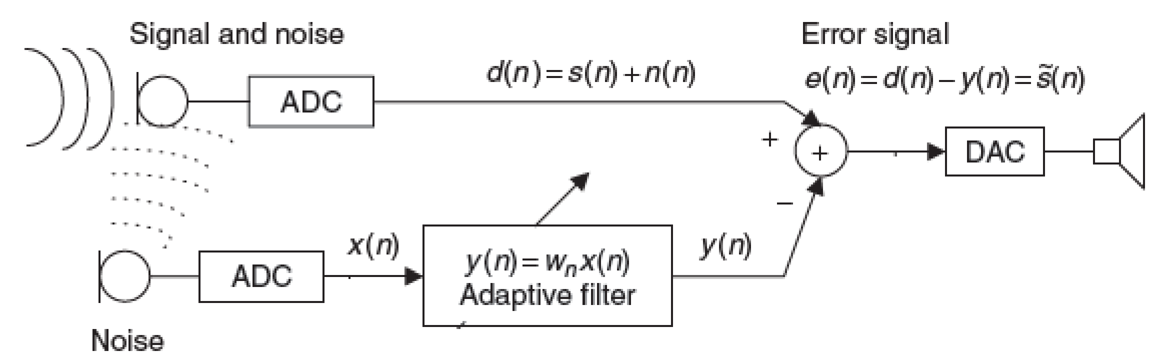
$$\hat{\mathbf{w}}_{\mathrm{op}} = \mathbf{R}^{-1}\mathbf{p} \text{ (se enfatiza que es una estimación)}$$

con:

$$\mathbf{p} = E\{d(n)\mathbf{x}(n)\}: \text{ vector de correlación cruzada entre } d(n) \text{ y } x(n)$$

$$\mathbf{R} = E\{\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)\}: \text{ Matriz de autocorrelación de la señal de entrada}$$

Ejemplo: Cancelación de Ruido



Para este caso,

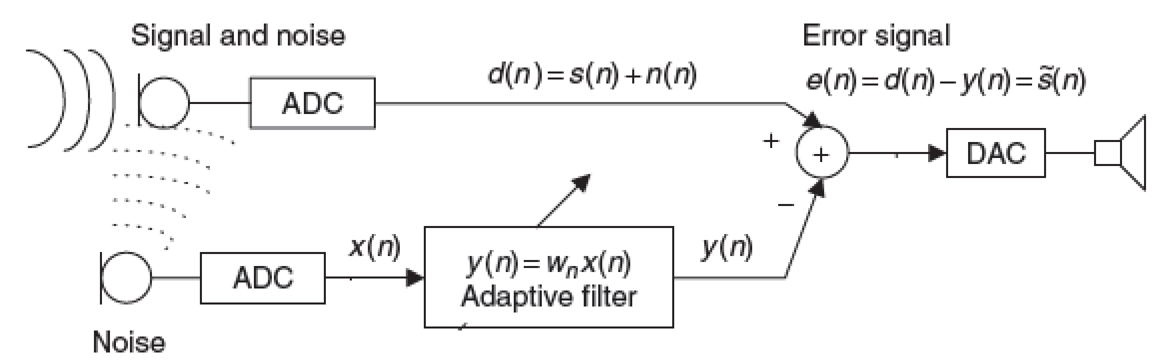
$$\mathbf{p} = E\{d(n)\mathbf{x}(n)\} = E\{[s(n)+\eta(n)]\mathbf{x}(n)\} = E\{s(n)\mathbf{x}(n)\} + E\{\eta(n)\mathbf{x}(n)\}$$

Considerando que $x(n)$ y $s(n)$ no están correlacionadas $E\{s(n)\mathbf{x}(n)\} = E\{s(n)\}E\{\mathbf{x}(n)\}$

Suponiendo valor medio nulo, resulta:

$$\mathbf{p} = E\{d(n)\mathbf{x}(n)\} = E\{\eta(n)\mathbf{x}(n)\} = E\{\mathbf{w}_{\text{op}}^{\text{T}}\mathbf{x}(n)\mathbf{x}(n)\} = E\{\mathbf{x}(n)\mathbf{x}^{\text{T}}(n)\mathbf{w}_{\text{op}}\} = \mathbf{R}\mathbf{w}_{\text{op}}$$

Ejemplo: Cancelación de Ruido



Substituyendo **p** en la estimación de Winner:

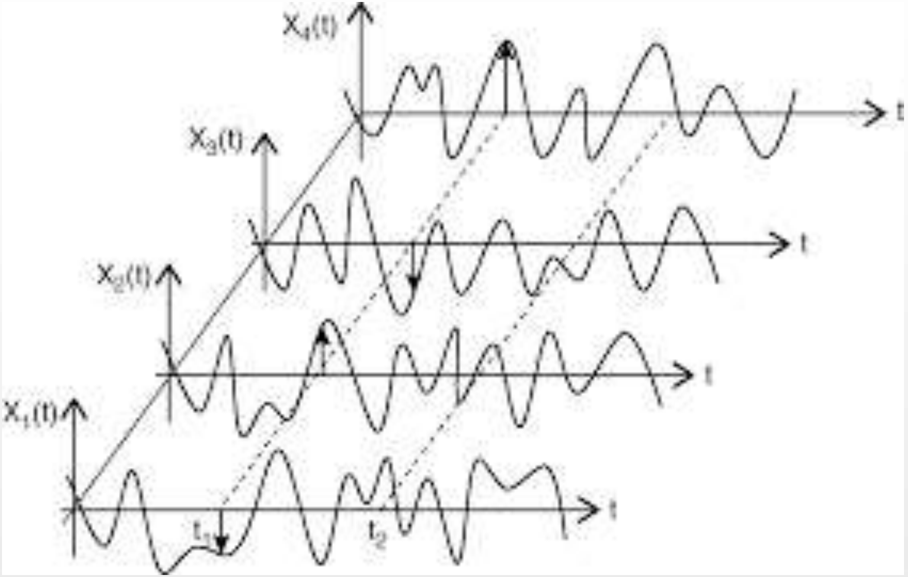
$$\hat{\mathbf{w}}_{op} = \mathbf{R}^{-1}\mathbf{p} = \mathbf{R}^{-1}\mathbf{R}\mathbf{w}_{op} = \mathbf{w}_{op}$$

Coincide plenamente bajo las consideraciones realizadas, sin embargo, el problema radica en la estimación de las estadísticas, por ejemplo:

$$E(d(n)x(n)) = \frac{d(0)x(0) + d(1)x(1) + \cdots + d(N-1)x(N-1)}{N}$$

Se supone ergodicidad

Valor esperado en procesos estocásticos



$$E\{x(t)\} \neq \lim_{T\to\infty} \frac{1}{T}\int_0^T x(t)dt$$

El valor medio es en el sentido temporal
El valor esperado es en el sentido de las repeticiones

Steepest Descent Algorithm (rumbo al LMS)

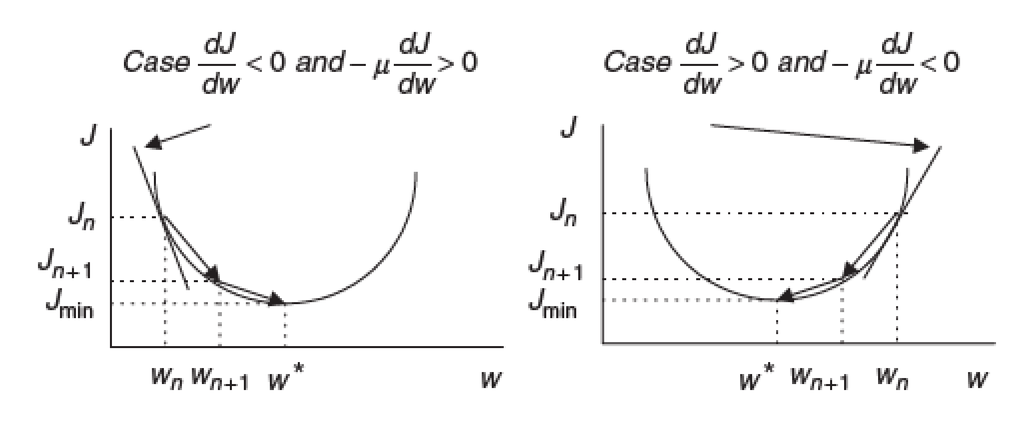
- Los métodos optimización basados en el gradiente fueron propuestos inicialmente por Cauchy en 1.847
- En la década de 1.940 fueron estudiados con más profundidad en sus propiedades.
- En 1.960 inspiró a Widrow y Hoff para inventar el algoritmo LMS
- El algoritmo LMS traccionó el desarrollo de las redes neuronales elementales, como ADALINE y MADALINE y la técnica de backpropagation.

Steepest Descent Algorithm

- Se resume en la siguiente ecuación:

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu \nabla_{\mathbf{w}} J$$

- Ejemplo unidimensional



Algoritmo LMS

- Least mean squares
- El algoritmo steepest descent requiere del cálculo de estadísticas para determinar el gradiente de *J*
- El LMS está basado en un cálculo con muestras actuales (es una estimación)

$$J = e^2(n) = \left(d(n) - wx(n)\right)^2$$

$$\frac{dJ}{dw} = 2\left(d(n) - wx(n)\right)\frac{d\left(d(n) - wx(n)\right)}{dw} = -2e(n)x(n).$$

Algoritmo LMS

1. Initialize $w(0), w(1), \ldots, w(N-1)$ to arbitrary values.
2. Read $d(n)$, $x(n)$, and perform digital filtering:

$$y(n) = w(0)x(n) + w(1)x(n-1) + \cdots + w(N-1)x(n-N+1).$$

FIR filter!

3. Compute the output error:

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

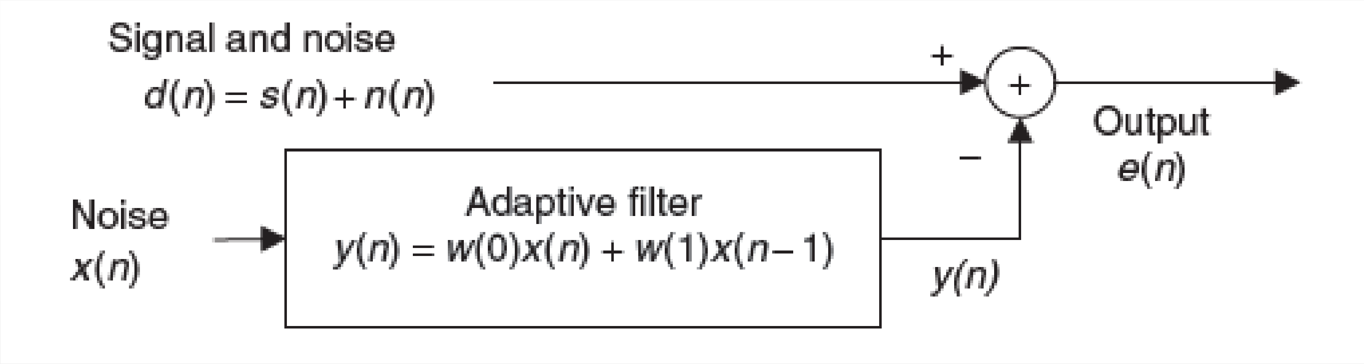
4. Update each filter coefficient using the LMS algorithm:

$$\text{for } i = 0, \ldots, N-1$$
$$w(i) = w(i) + 2\mu e(n)x(n-i).$$

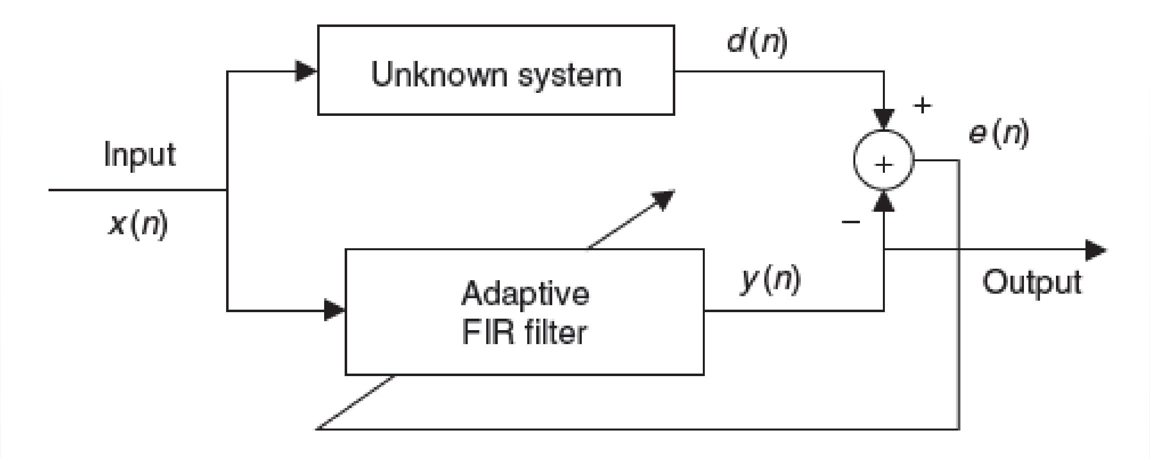
Aplicaciones:

- Cancelación de ruido
- Modelado de sistemas
- Arreglo de sensores (filtrado espacial) o *beamforming*

Cancelación de ruido

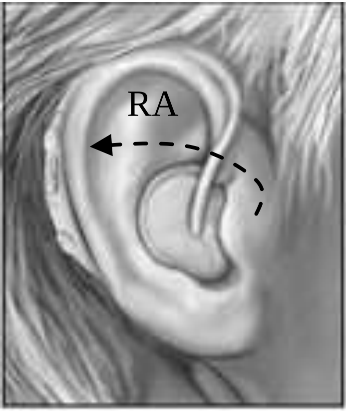


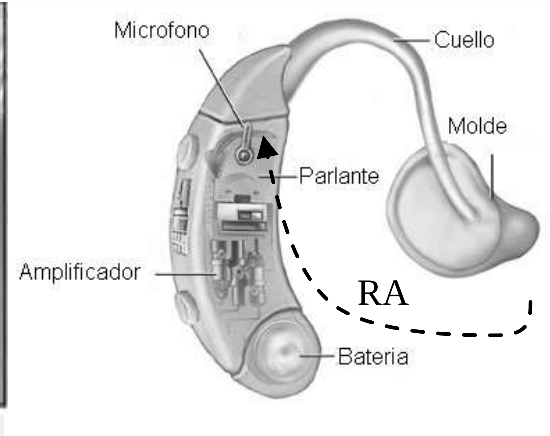
Modelado de Sistemas

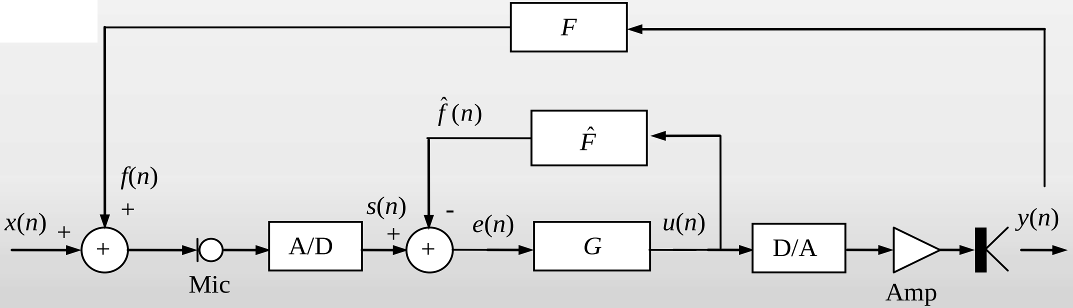


Modelado de sistemas

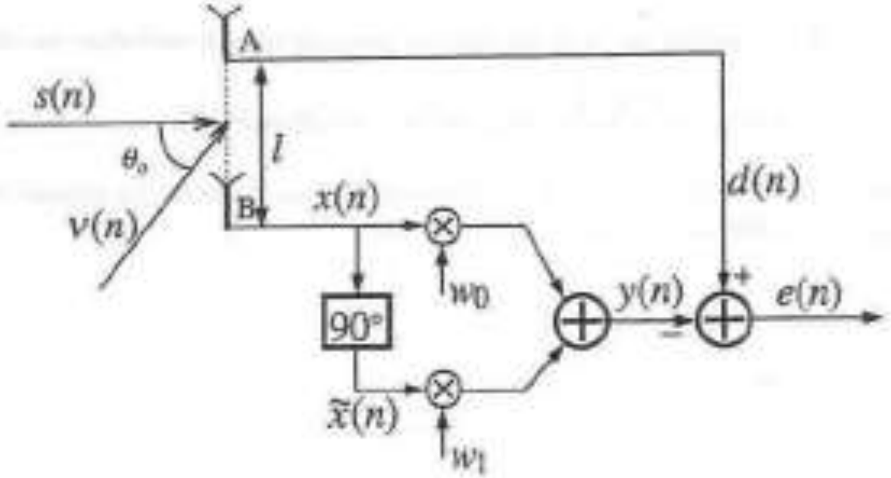
- Realimentación acústica (RA) en audífonos

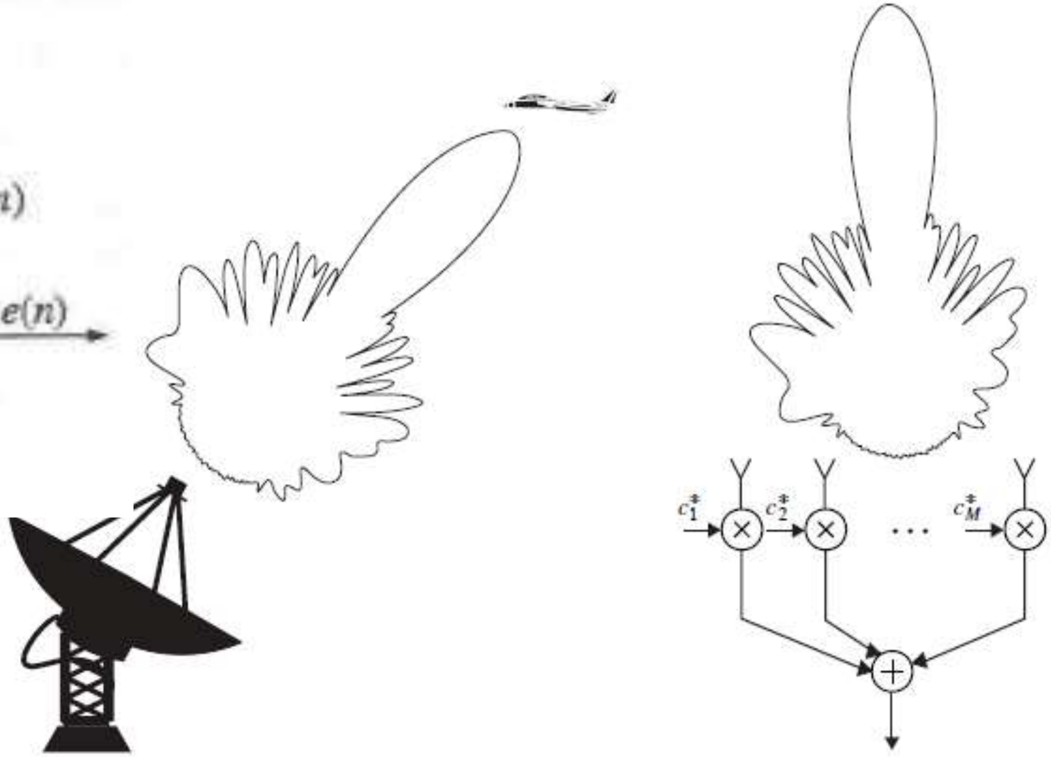






Arreglo de sensores: Beamforming





(a) Parabolic dish antenna
(continuous aperture)

(b) Sensor array antenna
(discrete spatial aperture)

Steepest Descent Algorithm

- Ejercicio 3:

$$J = 11{,}5 - 20w + 8w^2 \longrightarrow \frac{dJ}{dw} = -20 + 16w$$

- Se resume en la siguiente ecuación:

$$w(n+1) = w(n) - \mu \nabla_{\mathbf{w}} J = w(n) - 0{,}03 * (-20 + 16w(n))$$