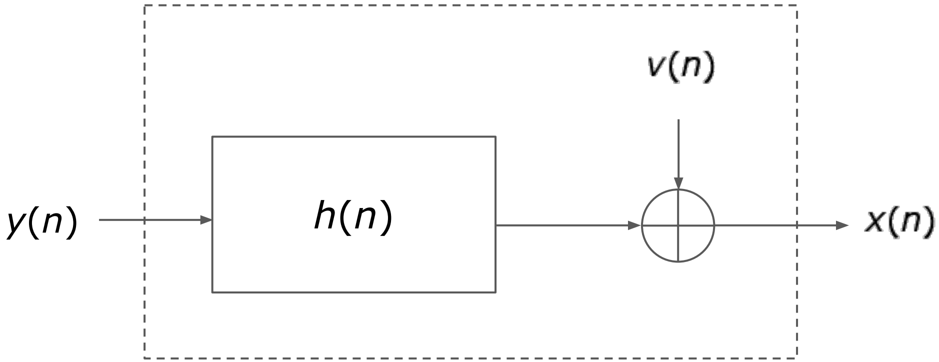
Identificación de un sistema

Identificación de un sistema



$$x(n) = \sum_{k=0}^{N-1} h_k y(n-k) + v(n)$$

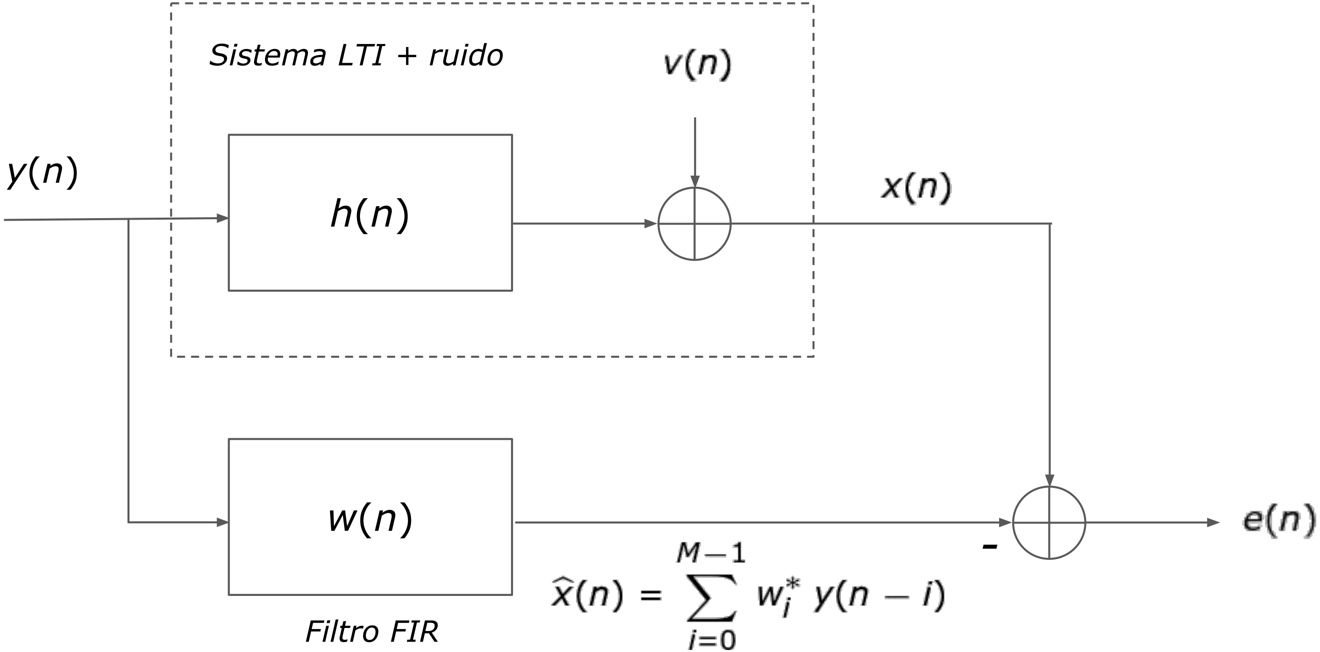
Desconocemos h(n)

Coeficientes de un sistema LTI FIR de largo N

$h(n)=\{h_0, h_1, h_1, \ldots, h_{N-1}\}$

$v(n)$ *Ruido blanco aditivo de media nula*

Identificación de un sistema



$$\hat{x}(n) = \sum_{i=0}^{M-1} w_i^* \, y(n-i)$$

Identificación de un sistema

Forma vectorial

$$x(n) = \sum_{k=0}^{N-1} h_k y(n-k) + v(n) \quad \Longrightarrow \quad x(n) = \mathbf{h}^H \mathbf{y}_N(n) + v(n)$$

$$\hat{x}(n) = \sum_{i=0}^{M-1} w_i^* \, y(n-i) \quad \Longrightarrow \quad \hat{x}(n) = \mathbf{w}^H \mathbf{y}_M(n)$$

$$\mathbf{h} = [h_0, h_1, \cdots, h_{N-1}]^T$$

$$\mathbf{w} = [w_0 \; w_1 \; \cdots \; w_{M-1}]^T$$

$$\mathbf{y}_N(n) = [y(n), y(n-1), \cdots, y(n-N+1)]^T$$

$$\mathbf{y}_M(n) = [y(n) \; y(n-1) \; \cdots \; y(n-M+1)]^T$$

Identificación de un sistema

¿Qué ocurre si el largo del sistema LTI es igual al largo del filtro?

$$\mathbf{R}_{yx} = \mathbf{R}_M \mathbf{w}_o$$ Condición óptima

$$E[\mathbf{y}_M(n)x(n)^*] = E[\mathbf{y}_M(n)\mathbf{y}_M(n)^H]\mathbf{w}_o$$

$$E[\mathbf{y}_M(n)\left(\mathbf{h}^H\mathbf{y}_N(n) + v(n)\right)^*] = E[\mathbf{y}_M(n)\mathbf{y}_M(n)^H]\mathbf{w}_o$$

$$E[\mathbf{y}_M(n)\mathbf{y}_N(n)^H]\mathbf{h} + E[\mathbf{y}_M(n)v(n)^*] = E[\mathbf{y}_M(n)\mathbf{y}_M(n)^H]\mathbf{w}_o$$

$$E[\mathbf{y}_M(n)\mathbf{y}_N(n)^H]\mathbf{h} = E[\mathbf{y}_M(n)\mathbf{y}_M(n)^H]\mathbf{w}_o$$

M=N

$$\cancel{E[\mathbf{y}_M(n)\mathbf{y}_M(n)^H]}\mathbf{h} = \cancel{E[\mathbf{y}_M(n)\mathbf{y}_M(n)^H]}\mathbf{w}_o$$

$$\boxed{\mathbf{h} = \mathbf{w}_o}$$

Identificación de un sistema

*¿Cuánto vale la potencia de error en la condición óptima (**w**=**w**$_o$)?*

$$J_{min} = \sigma_x^2 - \mathbf{R}_{yx}^H \mathbf{w}_o - \mathbf{w}_o^H \mathbf{R}_{yx} + \mathbf{w}_o^H \mathbf{R}_M \mathbf{w}_o$$

$$\mathbf{R}_{yx} = \mathbf{R}_M \mathbf{w}_o$$

$$\boxed{J_{min} = \sigma_x^2 - \mathbf{w}_o^H \mathbf{R}_M \mathbf{w}_o}$$

Busquemos σ_x^2

Identificación de un sistema

Varianza de señal x(n)

$$\sigma_x^2 = E[|x(n)|^2] = E[x(n)x(n)^*] = E[(\mathbf{h}^H\mathbf{y}_N(n) + v(n))(\mathbf{y}_N(n)^H\mathbf{h} + v^*(n))] =$$

$$\sigma_x^2 = E[|v(n)|^2] + \mathbf{h}^H \underbrace{E[(\mathbf{y}_N(n)\mathbf{y}_N(n)^H]}_{\mathbf{R}_N}\mathbf{h}$$

$$\Rightarrow \boxed{\sigma_x^2 = \sigma_v^2 + \mathbf{h}^H\mathbf{R}_N\,\mathbf{h}}$$

Identificación de un sistema

Potencia del error en la condición óptima

$$J_{min} = \sigma_X^2 - \mathbf{w}_o^H \mathbf{R}_M \mathbf{w}_o$$ Potencia de error mínima

$$\sigma_X^2 = \sigma_V^2 + \mathbf{h}^H \mathbf{R}_N \mathbf{h}$$ Varianza de señal

$$\Rightarrow \boxed{J_{min} = \sigma_V^2 + \mathbf{h}^H \mathbf{R}_N \mathbf{h} - \mathbf{w}_o^H \mathbf{R}_M \mathbf{w}_o}$$

Identificación de un sistema

¿Qué ocurre si el largo del sistema LTI es igual al largo del filtro?

$$J_{min} = \sigma_v^2 + \mathbf{h}^H \mathbf{R}_N \mathbf{h} - \mathbf{w}_o^H \mathbf{R}_M \mathbf{w}_o$$

M=N

$$J_{min} = \sigma_v^2 + \mathbf{h}^H \mathbf{R}_M \mathbf{h} - \mathbf{w}_o^H \mathbf{R}_M \mathbf{w}_o$$

$$\Downarrow$$

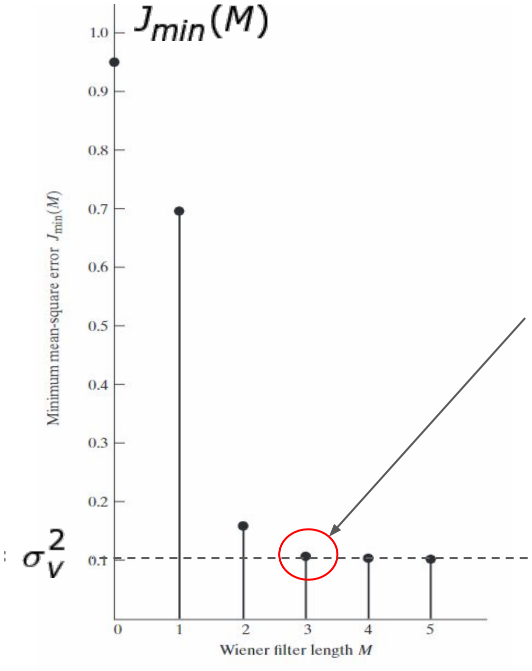
$$\mathbf{h} = \mathbf{w}_o$$

$$J_{min} = \sigma_v^2 + \mathbf{w}_o^H \mathbf{R}_M \mathbf{w}_o - \mathbf{w}_o^H \mathbf{R}_M \mathbf{w}_o$$

$$\boxed{J_{min} = \sigma_v^2}$$

Identificación de un sistema

Jmin para sistema LTI con N=3



$$J_{min} = \sigma_v^2 + \mathbf{h}^H \mathbf{R}_N \mathbf{h} - \mathbf{w}_o^H \mathbf{R}_M \mathbf{w}_o$$

Potencia de error mínima

Para M=N Jmin alcanza el mínimo valor posible, que es la varianza del ruido del modelo

Si M>N, se alcanza el mismo mínimo pero los coeficientes del filtro poseen más componentes innecesarias (de hecho se rellena con ceros: w_o=[h 0 .. 0])

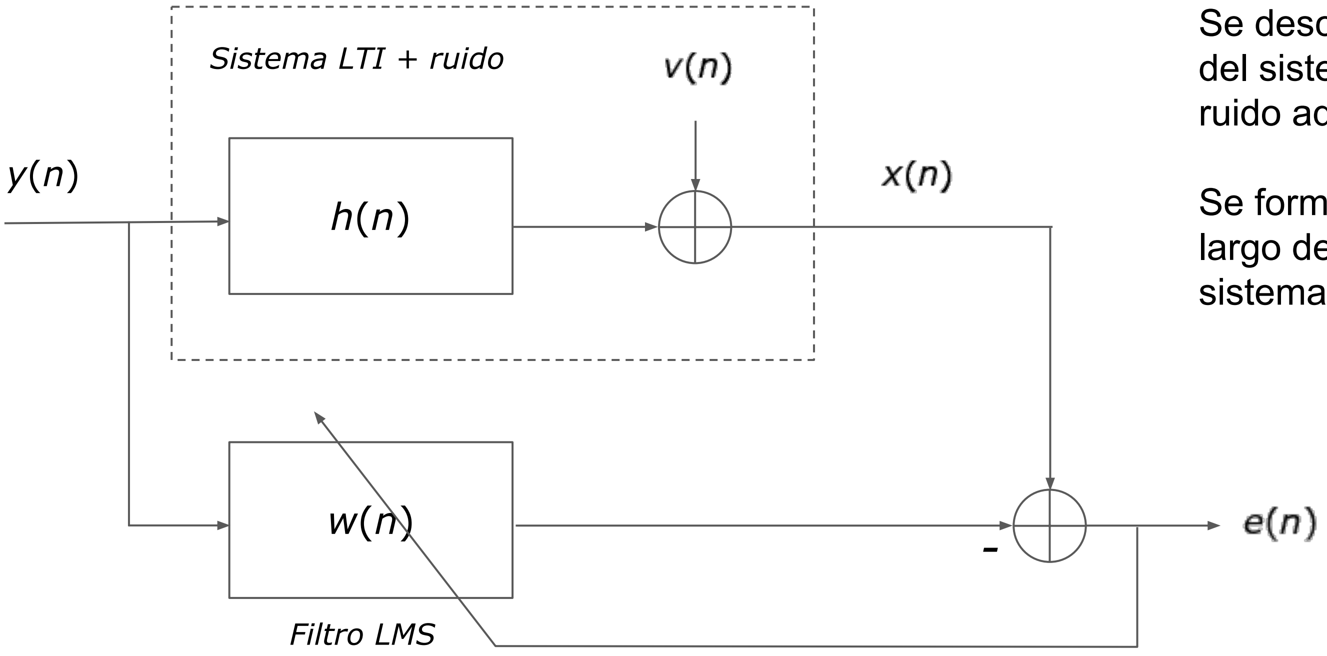
Ejercicio 3
Identificación de un sistema

Ejercicio 3

Se requiere identificar un sistema de comunicaciones a través del cual se transmiten datos $y(n)$ que atraviesan un canal con respuesta impulsiva $h(n) = \{1 \quad 0{,}4 \quad 0{,}3 \quad 0{,}1 \quad -0{,}2 \quad 0{,}05\}$ más ruido blanco aditivo gaussiano de media cero y varianza $\sigma_v^2 = 0{,}02$. La salida del sistema puede modelarse de acuerdo a la siguiente ecuación:

$$x(n) = \sum_{i=0}^{N-1} h_i \, y(n-i) + v(n)$$

Ejercicio 3



Se desconocen los coeficientes del sistema LTI y la varianza de ruido aditivo v(n)

Se formula el problema para un largo del filtro igual al largo del sistema LTI

Ejercicio 3

(a) Suponiendo que los datos se modelan como $y(n) = 2\,b(n) - 1$, con $b(n) \sim \text{Bernoulli}(0{,}5)$. Genere la salida del canal $x(n)$ a partir de la información disponible. Considere un largo $L = 2000$ muestras para la señales.

Ejercicio 3

(b) Utilizando un filtro LMS con $\mu = 0{,}1$, estime los coeficientes del canal. ¿Qué valor de M utilizaría para el largo del filtro $\mathbf{w} = [w_0, w_1, ..., w_{M-1}]^T$? Grafique en un mismo plot los M coeficientes estimados $\widehat{\mathbf{w}}_n$ del filtro en función de las iteraciones del algoritmo. Grafique también las líneas de los coeficientes reales del sistema a identificar $h(n)$ superpuestos $\widehat{\mathbf{w}}_n$.

Ejercicio 3

(c) Haga un gráfico de la curva de aprendizaje $\widehat{J}(n)$ en función de las iteraciones (considere al menos 100 realizaciones).

Ejercicio 3

(d) Repita los dos puntos anteriores para $\mu = 0{,}005$. ¿Qué diferencias observa?

Filtros Adaptativos (parte 2)

Procesamiento de señales

.UBA fiuba FACULTAD DE INGENIERÍA

Ing. Débora P. Copa ● Dra. Cecilia G. Galarza
IIBM LPSC

Ejercicio 5
NLMS

Ejercicio 5

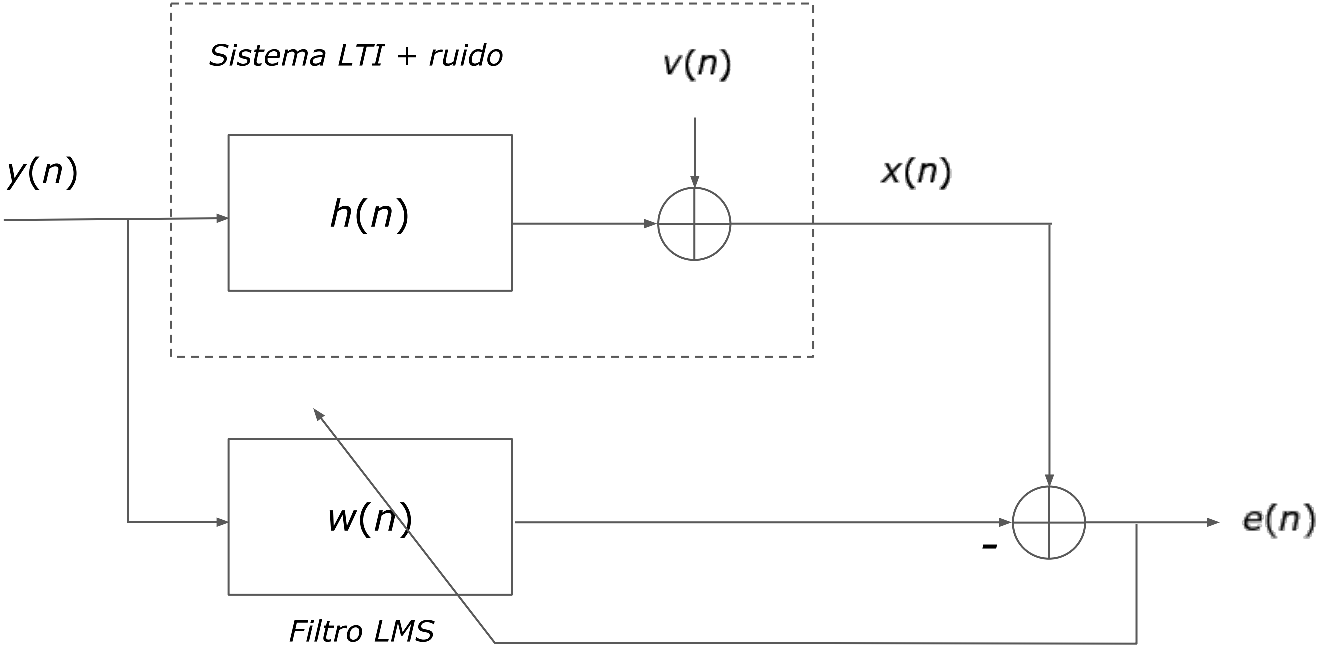
NLMS: Una variante para al algotimo LMS, es NLMS (Normalized LMS). En este caso los coeficientes del filtro pueden obtenerse mediante la siguiente ecuación recursiva:

$$\widehat{\mathbf{w}}_{n+1} = \widehat{\mathbf{w}}_n + \frac{\tilde{\mu}}{||\mathbf{y}(n)||^2}\mathbf{y}(n)e^*(n)$$

Problema: Para analizar las diferencias entre LMS y NLMS, como ejemplo vamos a utilizar un filtro adaptativo para identificar un canal de comunicaciones.

Ejercicio 5

$$h(n) = [1\ 0.4\ 0.3\ 0.1\ -0.2\ 0.05]^t$$



Ejercicio 5

(a) Considere un canal de comunicaciones con la misma respuesta impulsiva $h(n)$ que en el Ejercicio 3, donde el ruido aditivo es blanco de varianza $\sigma_v^2 = 0{,}02$ y la entrada una señal ruidosa $y(n) \sim N(0; 36)$, de largo $L = 2000$. Implemente el algoritmo NLMS y compute simultáneamente el algoritmo LMS ($\mu = 0{,}005$) y NLMS ($\tilde{\mu} = 0{,}4$). Grafique la curva de aprendizaje $\widehat{J}(n)$ para ambos algoritmos (superpuestas) para $m = 500$ realizaciones. ¿Qué diferencias observa?

Ejercicio 5

(b) Ahora suponga que la señal de entrada es tal que para ciertos instantes de tiempo $||\mathbf{y}(n)||^2$ posee una amplitud muy cercana a cero. Para probar este caso, genere una secuencia de entrada de la forma $y(n) = 1 + \cos(0{,}004\pi n)$. Simule el algoritmo NLMS ($\tilde{\mu} = 0{,}05$) para esta entrada y grafique la curva de aprendizaje $\widehat{J}(n)$. ¿Qué ocurre en los instantes donde la entrada $y(n)$ es cercana a cero?.

Ejercicio 5

(c) Una forma de mitigar el problema del punto anterior es agregando un parámetro $\delta > 0$ que evita la amplificación del gradiente cuando $||\mathbf{y}(n)||^2$ es muy pequeño, para lo cual se puede modificar la ecuación recursiva:

$$\widehat{\mathbf{w}}_{n+1} = \widehat{\mathbf{w}}_n + \frac{\tilde{\mu}}{\delta + ||\mathbf{y}(n)||^2}\mathbf{y}(n)e^*(n)$$

Repita la simulación del punto anterior, pero con esta modificación. Considere $\delta = 0{,}001$.

NLMS

- LMS tiene problemas de **amplificación de ruido de gradiente**.
- NLMS **normaliza** el vector de entradas

$$\hat{\mathbf{w}}_{n+1} = \hat{\mathbf{w}}_n + \frac{\alpha}{||\mathbf{y}(n)||^2}\mathbf{y}(n)e^*(n) \qquad 0 < \alpha < 2.$$

- Ese factor de normalización en NLMS puede producir un paso muy grande para entradas muy pequeñas
- Se soluciona agregando un parámetro $\boldsymbol{\delta}$ que mantiene acotado al término de actualización

$$\hat{\mathbf{w}}_{n+1} = \hat{\mathbf{w}}_n + \frac{\alpha}{\delta + ||\mathbf{y}(n)||^2}\mathbf{y}(n)e^*(n)$$

Método de Mínimos Cuadrados Recursivo (RLS)

Algoritmos adaptativos

Se basa en la **estadística de las entradas y salidas** (se toma la media E[.])

$J = E[\,|e(n)|^2]$ **Wiener** $\xrightarrow[\mathbf{R}_y \ \mathbf{R}_{yx}]{\text{Iterativo}}$ **Steepest descent** $\xrightarrow[\widehat{\mathbf{R}}_y \ \widehat{\mathbf{R}}_{yx}]{\text{Estimaciones instantáneas}}$ **LMS** $\xrightarrow[\frac{\alpha}{||\mathbf{y}(n)||^2}]{\text{normalización gradiente}}$ **NLMS**

Se basa en un **conjunto determinístico de entradas y salidas** (se toma el promedio)

$\mathcal{E} = \sum_{i=1}^{N} |e(i)|^2$ **LS** $\xrightarrow{\text{Recursivo}}$ **RLS** $\xrightarrow{\text{factor de olvido}}$ $\mathscr{E}_n = \sum_{i=1}^{n} \beta(n,i)|e(i)|^2$ **RLS con FO**

RLS

Observaciones

- Steepest descent es una implementación recursiva del filtro de Wiener a partir del gradiente (conociendo R y p).
- LMS es una implementación recursiva que busca el mínimo de la función costo a partir de estimaciones simples del gradiente (R y p rudimentarias)
- LMS puede no converger suficientemente rápido en ciertos casos
- LMS puede llegar a tener un error en exceso significativo.

Alternativa

- Basarse en promedios en vez de valores esperados (LS).
- Minimizar como función costo la energía del error (LS).
- Modificar la función costo con un parámetro de olvido
- Obtener recursivamente los coeficientes

RLS

RLS

Función costo

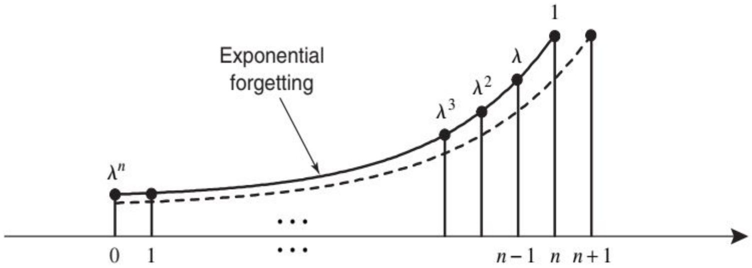
$$\mathcal{E}(n) = \sum_{i=1}^{n} \beta(n,i)\,|e(i)|^2$$

$$\beta(n,i) = \lambda^{n-i}, \qquad i = 1, 2, \ldots, n,$$

$$e(n) = d(n) - \hat{\mathbf{w}}^H(n)\mathbf{u}(n)$$

Factor de ponderación o factor de olvido

El factor de ponderación β(n, i) garantiza que los datos del pasado lejano se "olviden" para poder *seguir las variaciones estadísticas* de los datos observados.



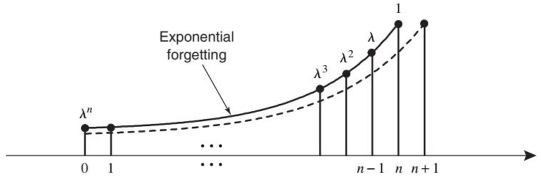
Curva de aprendizaje - En RLS se usa el error a priori

$$J'(n) = \mathbb{E}[\,|\xi(n)|^2] \qquad \xi(n) = d(n) - \widehat{\mathbf{w}}^H(n-1)\,\mathbf{u}(n)$$

pag 449 Haykin
pag 549 Manolakis

RLS

Factor de olvido



λ=1

En este caso, todos los valores del error, desde el inicio hasta el presente, tienen la **misma influencia en la función de costo** (memoria infinita).
El filtro pierde su capacidad de seguimiento (adaptación), lo que no es importante si el filtro se utiliza en un ambiente estacionario.

λ<1

En este caso, los valores más recientes de las observaciones tienen mayor influencia en la formación de la estimación LS de los coeficientes del filtro.
La memoria del filtro, o número efectivo de muestras utilizadas en cada estimación, es aproximadamente $1/(1 - \lambda)$ para $0{,}95 < \lambda < 1$.

RLS: Resumen del algoritmo

Condiciones iniciales

- $\widehat{\mathbf{w}}_0 = \mathbf{0} \qquad ; \qquad \mathbf{P}_0 = \delta^{-1} I_{M\times M}$

Resto de las iteraciones: n = 1,2,...,L-1

- $\mathbf{k}(n) = \frac{\lambda^{-1}\mathbf{P}_{n-1}\mathbf{y}(n)}{1+\lambda^{-1}\mathbf{y}(n)^H\,\mathbf{P}_{n-1}\,\mathbf{y}(n)}$ Ganancia
- $\xi(n) = x(n) - \widehat{\mathbf{w}}_{n-1}^H\,\mathbf{y}(n)$ Error a priori
- $\widehat{\mathbf{w}}_n = \widehat{\mathbf{w}}_{n-1} + \mathbf{k}(n)\xi^*(n)$ actualización
- $\mathbf{P}_n = \lambda^{-1}\mathbf{P}_{n-1} - \lambda^{-1}\mathbf{k}(n)\mathbf{y}(n)^H\,\mathbf{P}_{n-1}$ Matriz inversa de correlación

Ejercicio 6
RLS

Ejercicio 6

Considerando el algoritmo adaptativo RLS (ver Apéndice), suponga que se quiere resolver el problema de identificación del Ejercicio 3.

Ejercicio 6

(a) Implemente el algoritmo RLS para la determinación de los coeficientes $\mathbf{w}_n$. Considere $\lambda = 0{,}99$, $\delta = 0{,}001$. Grafique los coeficientes del filtro estimados $\widehat{\mathbf{w}}_n$ en función de las iteraciones.

Ejercicio 6

(b) Repita la simulación para al menos 500 realizaciones, calcule y grafique la curva de aprendizaje $\widehat{J}(n) = \frac{1}{m}\sum|\xi(n)|^2$.

Ejercicio 6

(d) Para la misma señal $x(n)$ del punto anterior, encuentre los coeficientes del filtro mediante LMS y RLS y compare las curvas de aprendizaje. Considere $\lambda = 0{,}995$ para RLS y $\mu = 0{,}1$ para LMS.

(e) Repita el punto anterior pero para $\lambda = 0{,}95$. Analice las diferencias.

(f) Reinicie la matriz P_n en el instante justo en que se genera la perturbación del sistema. Observe la respuesta y saque conclusiones.