IEE352 - Procesamiento Digital de Señales

Clase 12: Filtros Adaptivos 1

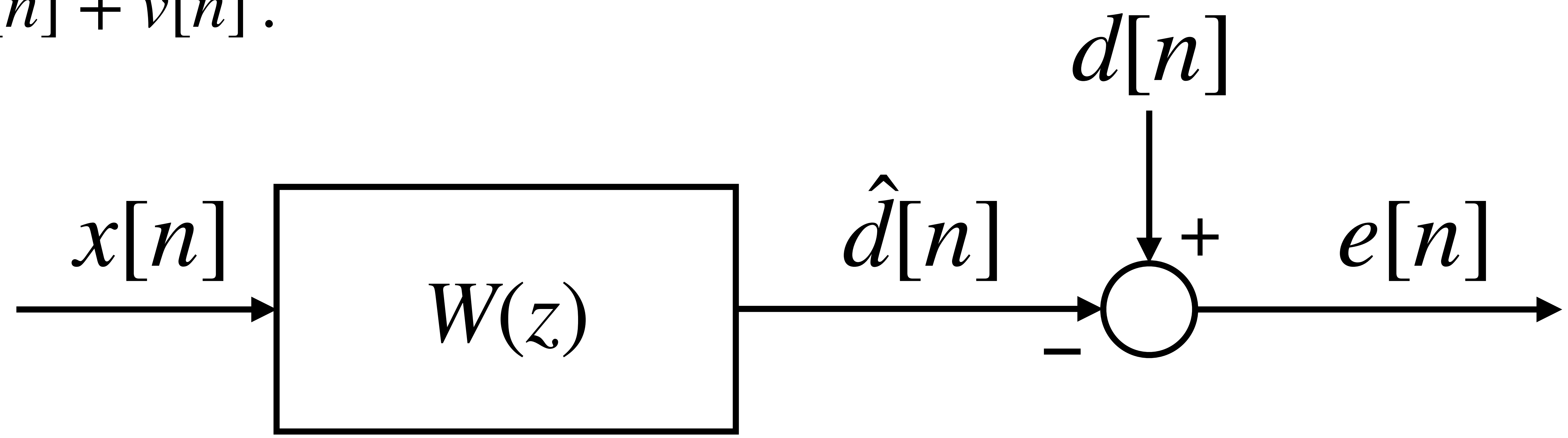
Dr. Marco A. Milla
Sección Electricidad y Electrónica (SEE)
Pontificia Universidad Católica del Perú (PUCP)

email: milla.ma@pucp.edu.pe

Filtros Óptimos

Filtro Wiener FIR

Diseñar un filtro $W(z)$ tipo FIR causal que permita recuperar una señal $d[n]$ a partir de una señal observada $x[n] = d[n] + v[n]$.



Considerando que $x[n]$ y $d[n]$ son procesos estacionarios en el sentido amplio (WSS), vamos a resolver el problema minimizando el error cuadrático medio:

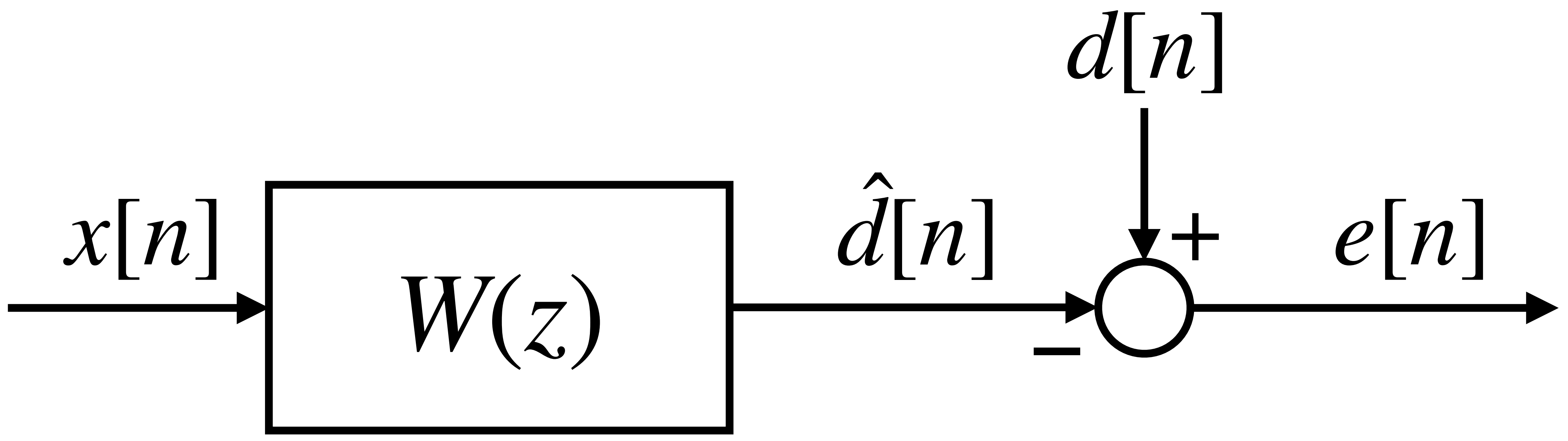
$$\xi = E\{\,|e[n]|^2\,\}\ ,$$

donde la señal de error es la diferencia entre el valor deseado $d[n]$ y el estimado $\hat{d}[n]$,

$$e[n] = d[n] - \hat{d}[n]\ .$$

Filtros Óptimos

Filtro Wiener FIR



La solución del problema se obtiene a partir de resolver la siguiente ecuación

$$\mathbf{R}_x \bar{\mathbf{w}} = \bar{\mathbf{r}}_{dx} \implies \bar{\mathbf{w}}_{\text{opt}} = \mathbf{R}_x^{-1} \bar{\mathbf{r}}_{dx}$$

donde $\mathbf{R}_x$ es la matriz de autocorrelaciones de $x[n]$, y $\bar{\mathbf{r}}_{dx}$ es el vector de correlaciones cruzadas entre las señales $d[n]$ y $x[n]$.

Algunas dificultades:

- Para resolver el problema se requiere conocer las correlaciones $r_x[k]$ y $r_{dx}[k]$, o, alternativamente, haber realizado suficientes ensayos (muestras) para calcularlas.
- Si la matrix $\mathbf{R}_x$ es muy grande, la inversión de esta matriz es computacionalmente costosa, inclusive $\mathbf{R}_x$ puede ser no invertible.

Filtros Adaptivos

Conceptos previos

Autovalores (eigenvalues): Los valores propios (autovalores) de una matriz cuadrada pueden usarse para determinar si esta es invertible, o si el cálculo de la inversa será sensible a los errores numéricos. Dada una matriz $\mathbf{A}$, sus autovalores λ deben cumplir la siguiente ecuación

$$\mathbf{A}\bar{x} = \lambda\bar{x} \ ,$$

cuyas soluciones se encuentran a partir de resolver la siguiente expresión

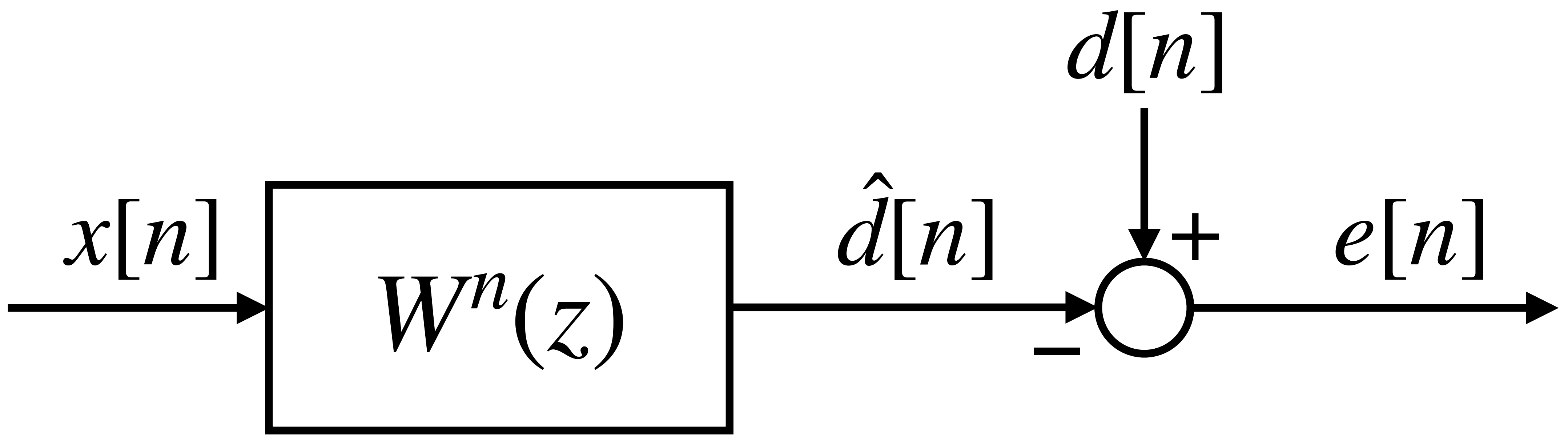
$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0 \ .$$

Norma L^p : La norma-p de un vector $\bar{x} = [x_1, x_2, \ldots, x_N]$ es la raíz p-ésima de la suma de los valores absolutos de los elementos del vector elevados a la potencia p, es decir,

$$\|\bar{x}\|_p = \left(\sum_{n=1}^{N} |x_n|^p \right)^{1/p} \ .$$

Filtros Adaptivos

Introducción



Si $x[n]$ y $d[n]$ son dos procesos aleatorios no estacionarios, los coeficientes del filtro que minimizan el error cuadrático medio $E\{\,|e[n]|^2\,\}$ van a depender de n, es decir

$$\hat{d}[n] = \sum_{k=0}^{M-1} w_n[k]x[n-k]$$

donde $w_n[k]$ es el valor en el tiempo n del k-ésimo coeficiente del filtro.

Alternativa 1: A partir de la derivación clásica de Wiener los coeficientes pueden ser estimados de la siguiente forma:

$$\mathbf{R}_x^n\,\bar{\mathbf{w}}^n = \bar{\mathbf{r}}_{dx}^n \implies \bar{\mathbf{w}}^n = (\mathbf{R}_x^n)^{-1}\,\bar{\mathbf{r}}_{dx}^n\,.$$

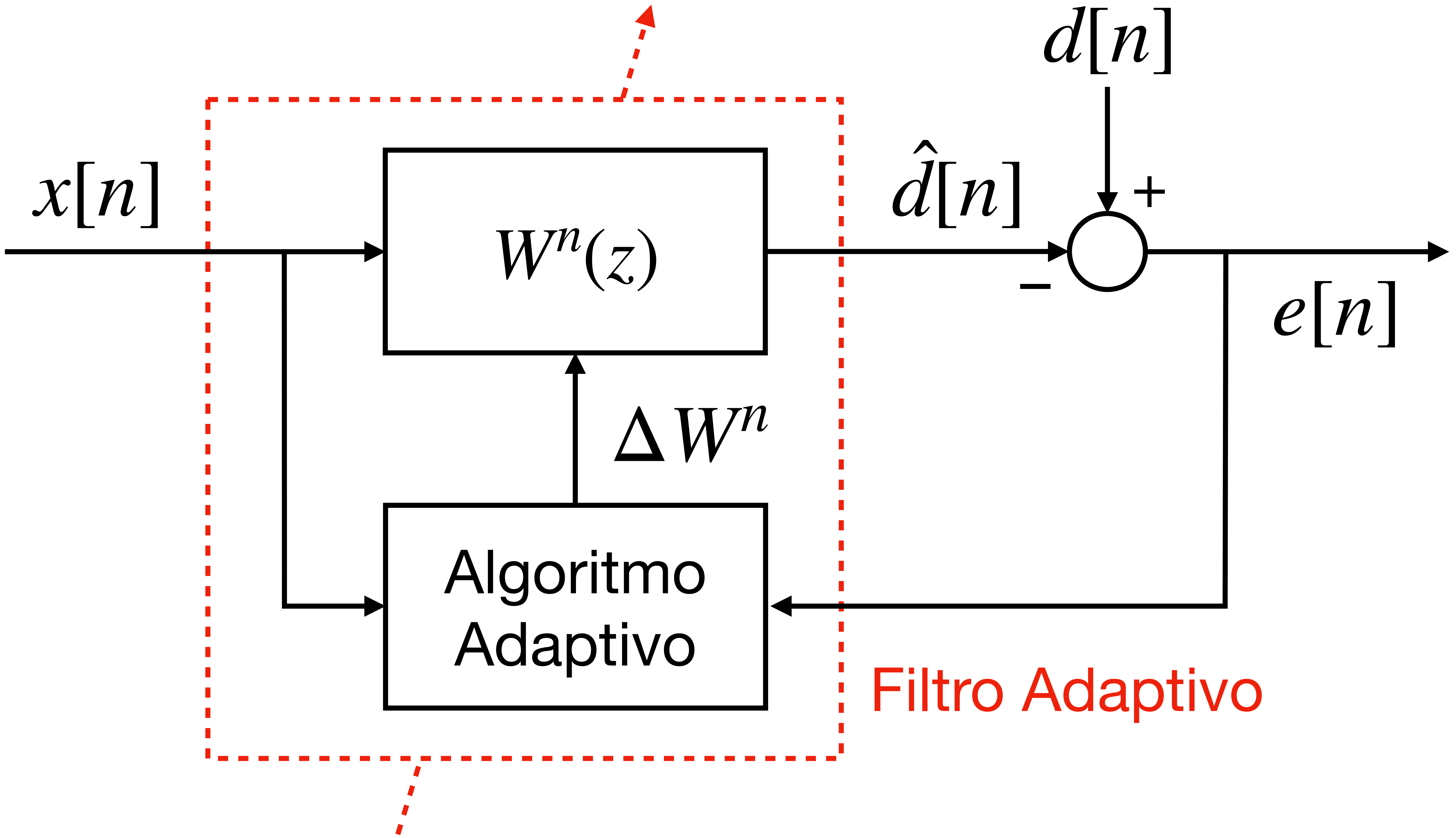
Alternativa 2: Relajando la condición que $\bar{\mathbf{w}}^n$ minimiza el error cuadrático medio, se puede implementar una ecuación de actualización de coeficientes, tal que,

$$\bar{\mathbf{w}}^{n+1} = \bar{\mathbf{w}}^n + \Delta\bar{\mathbf{w}}^n\,.$$

donde $\Delta\bar{\mathbf{w}}^n$ es una corrección aplicada a $\bar{\mathbf{w}}^n$ para formar los nuevos coeficientes $\bar{\mathbf{w}}^{n+1}$.

Filtros Adaptivos

Esquema general de Filtro Adaptivo



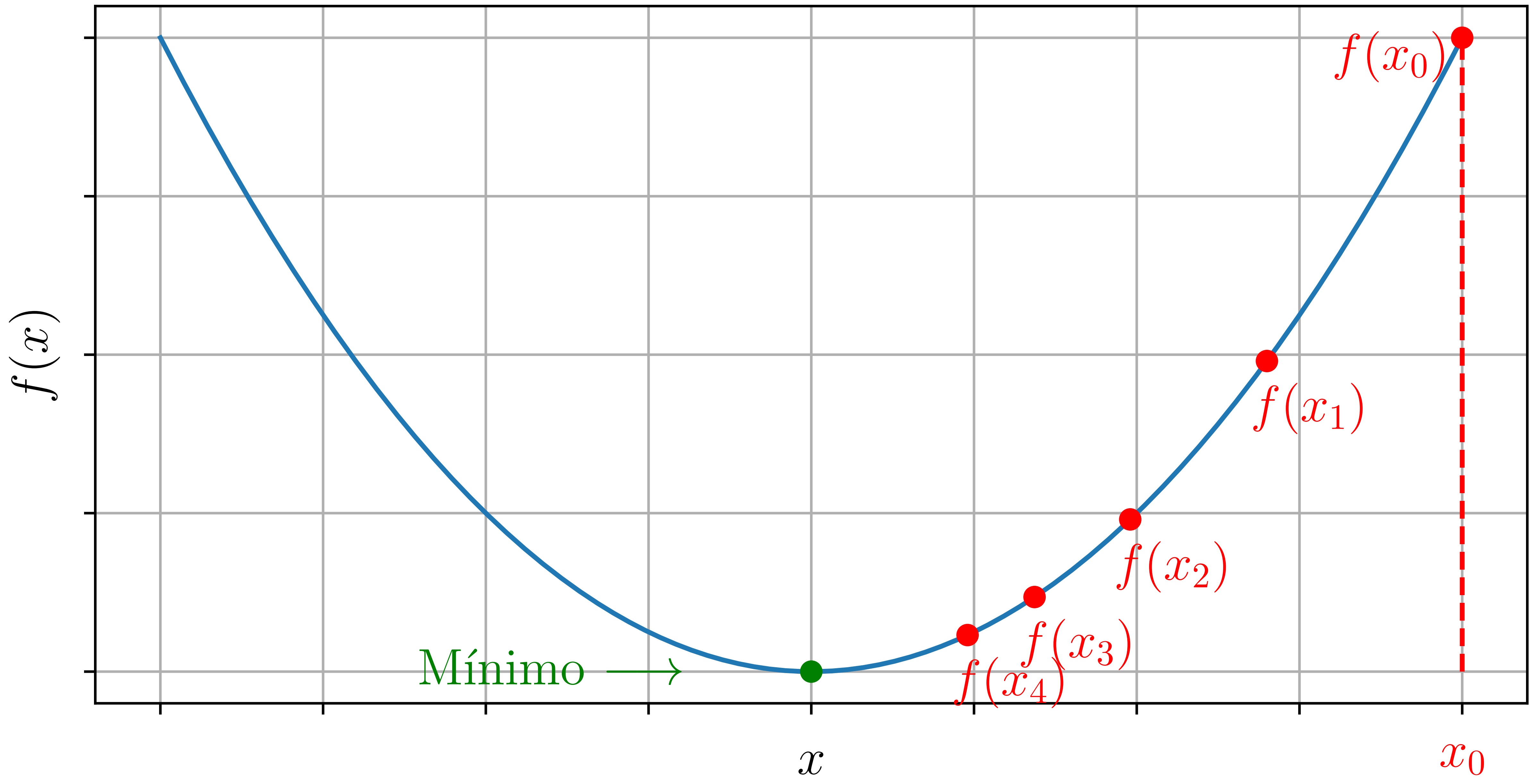
Gradiente Descendente

Descripción

Gradiente descendente es un algoritmo de optimización iterativa que sirve para encontrar un mínimo local de una función diferenciable $F(x)$. La secuencia iterativa generada es

$$x_{k+1} = x_k - \mu \nabla F(x_k)$$

donde $\nabla F(x_k)$ es la gradiente de la función a minimizar y $\mu > 0$ es el tamaño del paso (step size).



Gradiente Descendente

Ejemplo - Una variable

Dada la función $f(x) = x^2 + 1$, encontrar el punto x que minimiza $f(x)$.

Algoritmo: $x_{k+1} = x_k - \mu \nabla F(x_k)$

Gradiente: $\nabla f(x) = \dfrac{\partial f}{\partial x} = 2x$ (dirección)

Inicializando: $x_0 = 8$ y $\mu = 0.25$

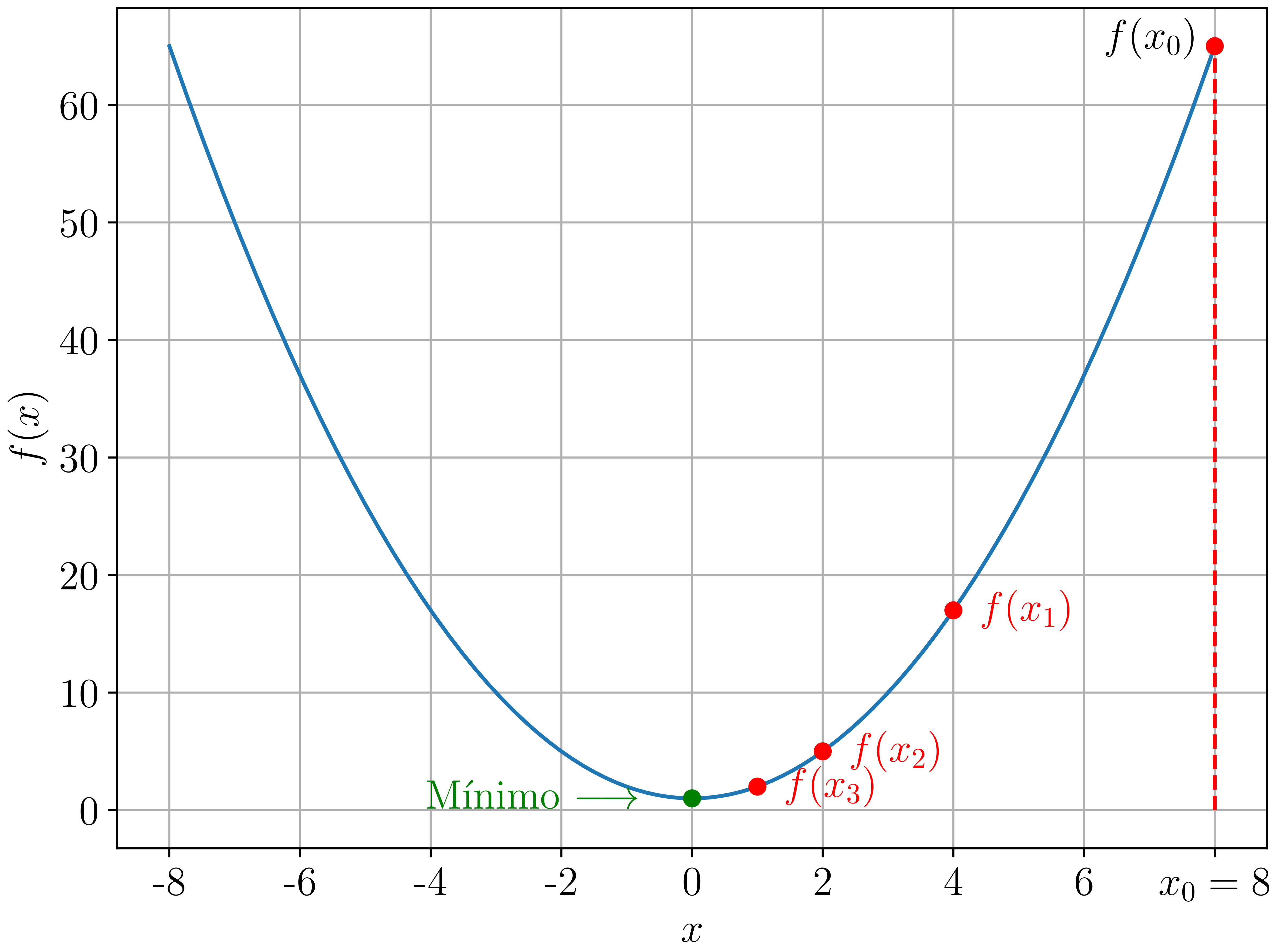
$$x_1 = x_0 - \mu \nabla f(x_0) = 8 - 0.25 \times 16 = 4$$

$$x_2 = x_1 - \mu \nabla f(x_1) = 4 - 0.25 \times 8 = 2$$

$$x_3 = x_2 - \mu \nabla f(x_2) = 2 - 0.25 \times 4 = 1$$

$$x_4 = x_3 - \mu \nabla f(x_3) = 1 - 0.25 \times 2 = 0.5$$

$$\vdots$$



Gradiente Descendente

Ejemplo - Dos variables

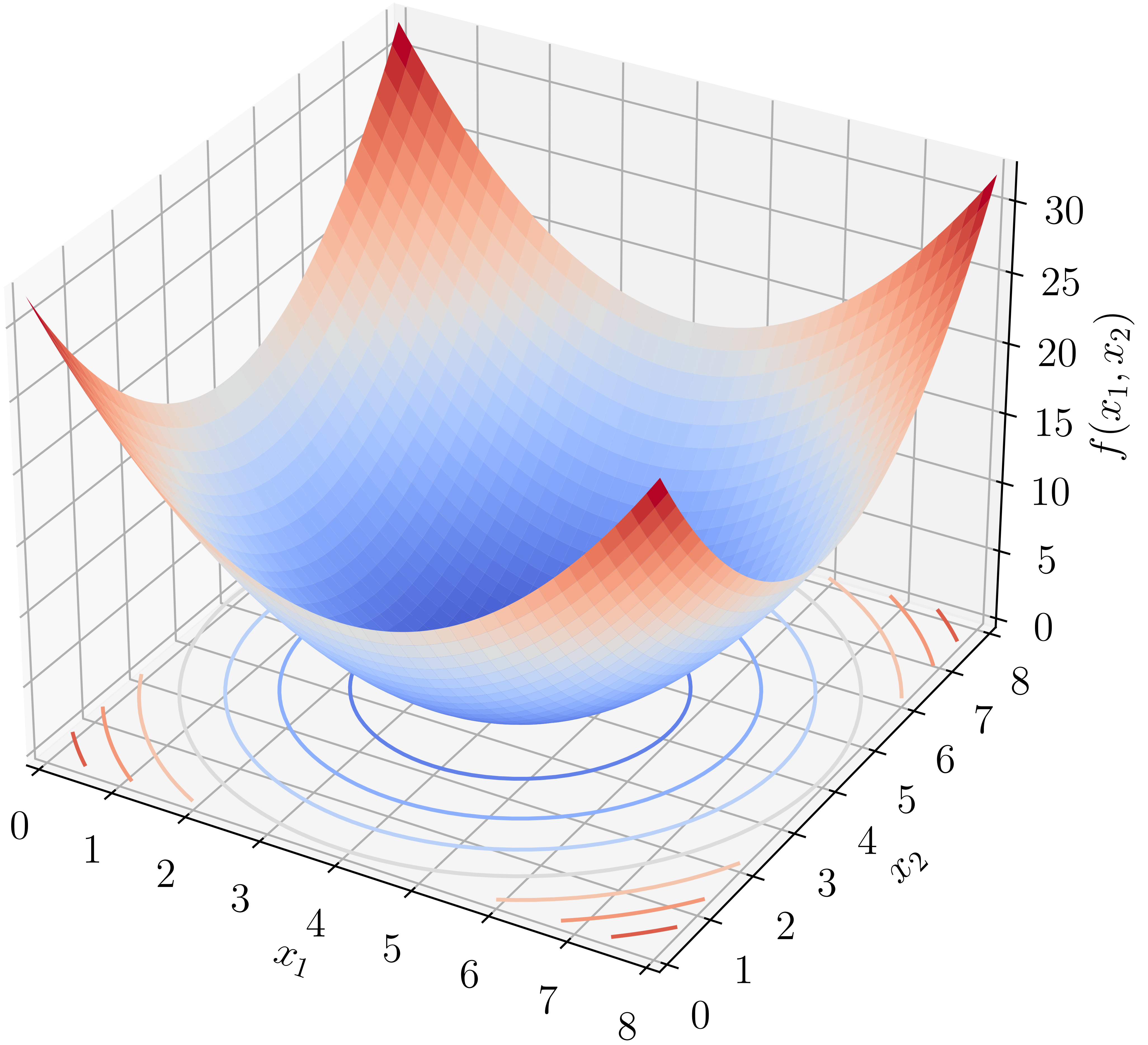
Dada la función

$$f(x_1, x_2) = (x_1 - 4)^2 + (x_2 - 4)^2 \,,$$

encontrar el punto (x_1, x_2) que minimiza la función.

Algoritmo:

$$\underbrace{\begin{bmatrix} x_1^{k+1} \\ x_2^{k+1} \end{bmatrix}}_{\bar{\mathbf{x}}^{k+1}} = \underbrace{\begin{bmatrix} x_1^{k} \\ x_2^{k} \end{bmatrix}}_{\bar{\mathbf{x}}^{k}} - \mu \underbrace{\begin{bmatrix} \dfrac{\partial f(x_1, x_2)}{\partial x_1} \\ \dfrac{\partial f(x_1, x_2)}{\partial x_2} \end{bmatrix}}_{\nabla f}$$



Gradiente Descendente

Ejercicio computacional

Desarrollar un programa basado en el concepto de gradiente descendente para minimizar la función $f(x) = \frac{1}{2}x_1^2 + \alpha x_2^2$ para distintos valores del paso μ,

$$\underbrace{\begin{bmatrix} x_1^{k+1} \\ x_2^{k+1} \end{bmatrix}}_{\bar{\mathbf{x}}^{k+1}} = \underbrace{\begin{bmatrix} x_1^{k} \\ x_2^{k} \end{bmatrix}}_{\bar{\mathbf{x}}^{k}} - \mu \underbrace{\begin{bmatrix} x_1^{k} \\ 2\alpha \cdot x_2^{k} \end{bmatrix}}_{\nabla f} .$$

Pregunta: ¿Cómo el paso μ y el factor α alteraran la convergencia?

Algoritmo de mínimos cuadrados (LMS)

Planteamiento

Dada la función error cuadrático medio $\xi[n] = E\{\,|e[n]|^2\,\}$, se puede usar el método de gradiente descendente para calcular los coeficientes de los filtros de manera iterativa

$$\bar{\mathbf{w}}^{n+1} = \bar{\mathbf{w}}^{n} - \mu\,\nabla\,\xi[n]\,.$$

Considerando que $W(z)$ es real y que $e[n] = d[n] - \hat{d}[n] = d[n] - \sum_{k=0}^{M-1} w[k]x[n-k]$, entonces la gradiente de $\xi[n]$ está dada por

$$\nabla\,\xi[n] = \nabla\,E\{\,|e[n]|^2\,\} = 2E\{e[n]\,\nabla\,e[n]\} = -\,2E\{e[n]\bar{\mathbf{x}}^{n}\}$$

$$\nabla\,e[n] = \begin{bmatrix} \dfrac{\partial e[n]}{\partial w^n[0]} \\ \dfrac{\partial e[n]}{\partial w^n[1]} \\ \vdots \\ \dfrac{\partial e[n]}{\partial w^n[M-1]} \end{bmatrix} = \begin{bmatrix} -x[n-0] \\ -x[n-1] \\ \vdots \\ -x[n-M+1] \end{bmatrix} = -\,\bar{\mathbf{x}}^{n}\,.$$

Algoritmo de mínimos cuadrados (LMS)

Planteamiento

La secuencia iterativa sería

$$\bar{\mathbf{w}}^{n+1} = \bar{\mathbf{w}}^{n} + 2\mu E\{e[n]\,\bar{\mathbf{x}}^{n}\}\ .$$

En la práctica $E\{e[n]\bar{\mathbf{x}}^{n}\}$ es desconocido, por ello, se reemplaza con una aproximación como la media muestral,

$$\hat{E}\{e[n]\,\bar{\mathbf{x}}^{n}\} = \frac{1}{L}\sum_{l=0}^{L-1} e[n-l]\,\bar{\mathbf{x}}^{n-l}\,.$$

Incorporando la aproximación en la secuencia iterativa tenemos

$$\bar{\mathbf{w}}^{n+1} = \bar{\mathbf{w}}^{n} + \frac{2\mu}{L}\sum_{l=0}^{L-1} e[n-l]\,\bar{\mathbf{x}}^{n-l}\,.$$

Algoritmo de mínimos cuadrados (LMS)

Planteamiento

Caso especial: Considerando solo un punto para el promedio muestral ($L = 1$) tenemos

$$\hat{E}\{e[n]\,\bar{\mathbf{x}}^n\} = e[n]\,\bar{\mathbf{x}}^n \,.$$

Finalmente el algoritmo LMS está dado por

$$\bar{\mathbf{w}}^{n+1} = \bar{\mathbf{w}}^n + 2\mu e[n]\,\bar{\mathbf{x}}^n$$

$$\underbrace{\begin{bmatrix} w^{n+1}[0] \\ w^{n+1}[1] \\ \vdots \\ w^{n+1}[M-1] \end{bmatrix}}_{\bar{\mathbf{w}}^{\mathbf{n+1}}} = \underbrace{\begin{bmatrix} w^{n}[0] \\ w^{n}[1] \\ \vdots \\ w^{n}[M-1] \end{bmatrix}}_{\bar{\mathbf{w}}^{\mathbf{n}}} + 2\mu e[n] \underbrace{\begin{bmatrix} x[n] \\ x[n-1] \\ \vdots \\ x[n-M+1] \end{bmatrix}}_{\bar{\mathbf{x}}^{\mathbf{n}}}$$

Alternativamente, cada coeficiente del filtro cumple la siguiente relación iterativa

$$w^{n+1}[k] = w^n[k] + 2\mu e[n]x[n-k], \qquad k = 0,1,\ldots,M-1 \,.$$

Algoritmo de mínimos cuadrados (LMS)

Algoritmo 1: Algoritmo LMS para un filtro adaptivo FIR de longitud M.

Parámetros: M = Longitud del filtro

μ = Tamaño de paso (step size)

Inicialización: $\bar{\mathbf{w}}^0 = [0,0,\ldots,0]^T$

Computar: For $n = 0,1,2,\ldots$

(a) $\hat{d}[n] = (\bar{\mathbf{w}}^n)^T\,\bar{\mathbf{x}}^n$

(b) $e[n] = d[n] - \hat{d}[n]$

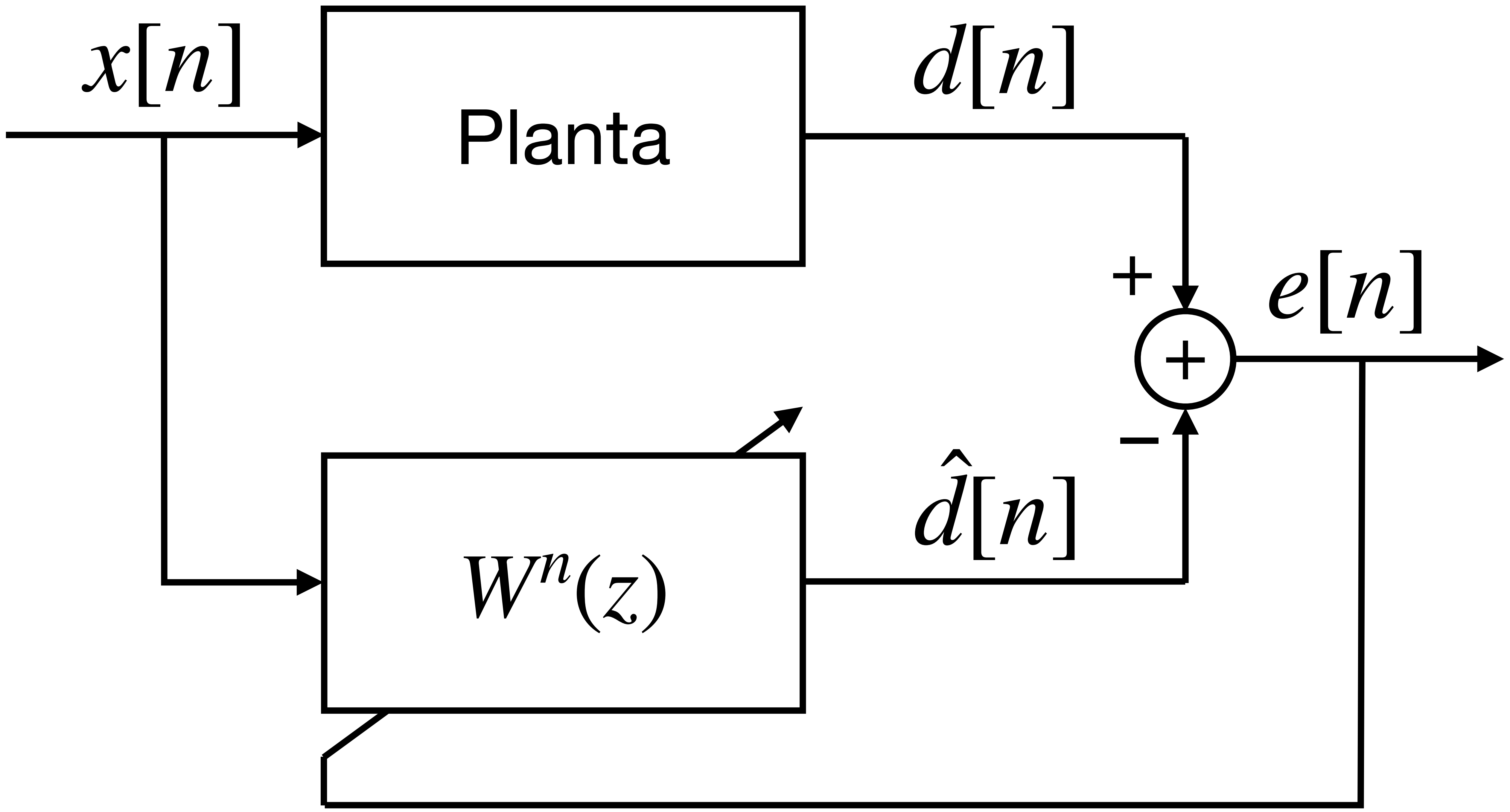
(c) $\bar{\mathbf{w}}^{n+1} = \bar{\mathbf{w}}^n + 2\mu e[n]\bar{\mathbf{x}}^n$

Nota: Señal estimada $\hat{d}[n] = \sum_{k=0}^{M-1} w^n[k]x[n-k] = (\bar{\mathbf{w}}^n)^T\,\bar{\mathbf{x}}^n$.

Filtros Adaptivos - Casos

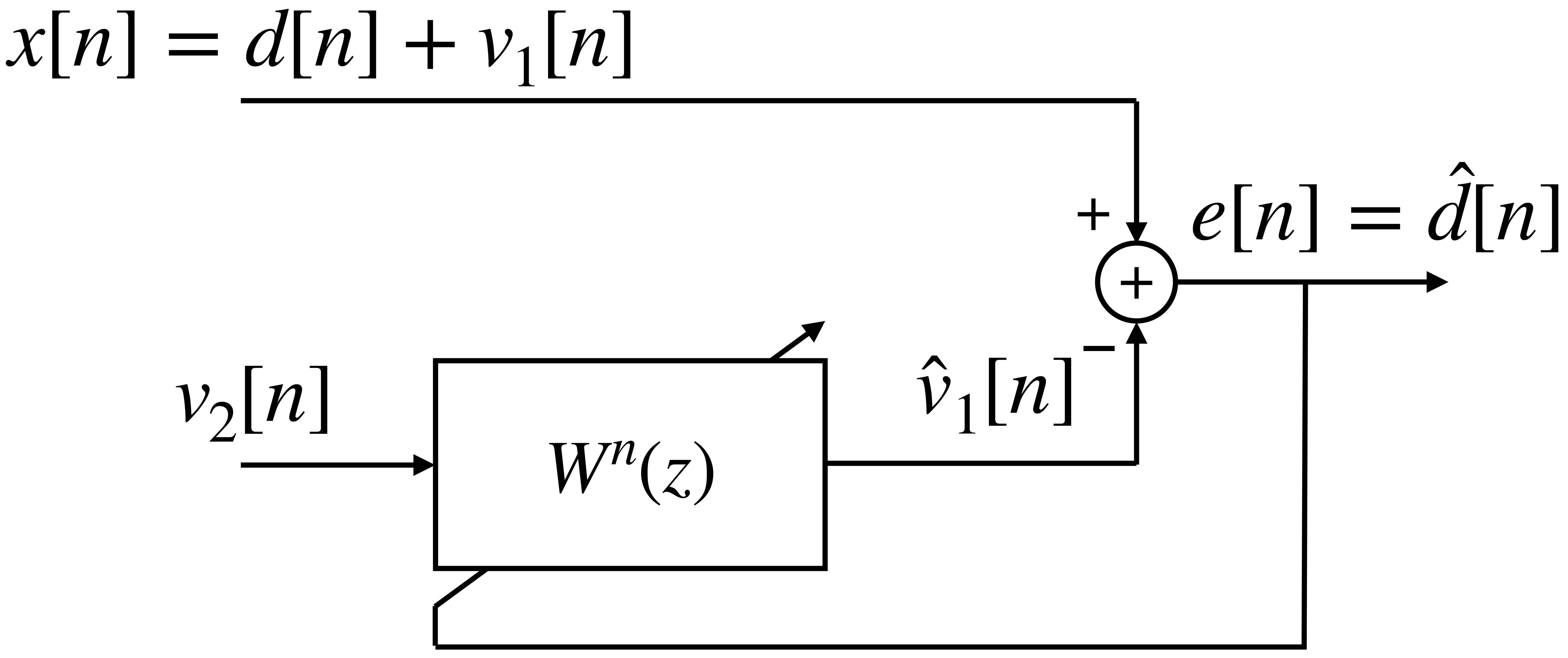
Identificador de sistema

Recordar: El algoritmo LMS estima filtros de la forma $W^n(z) = \sum_{k=0}^{M-1} w^n[k]z^{-k}$.



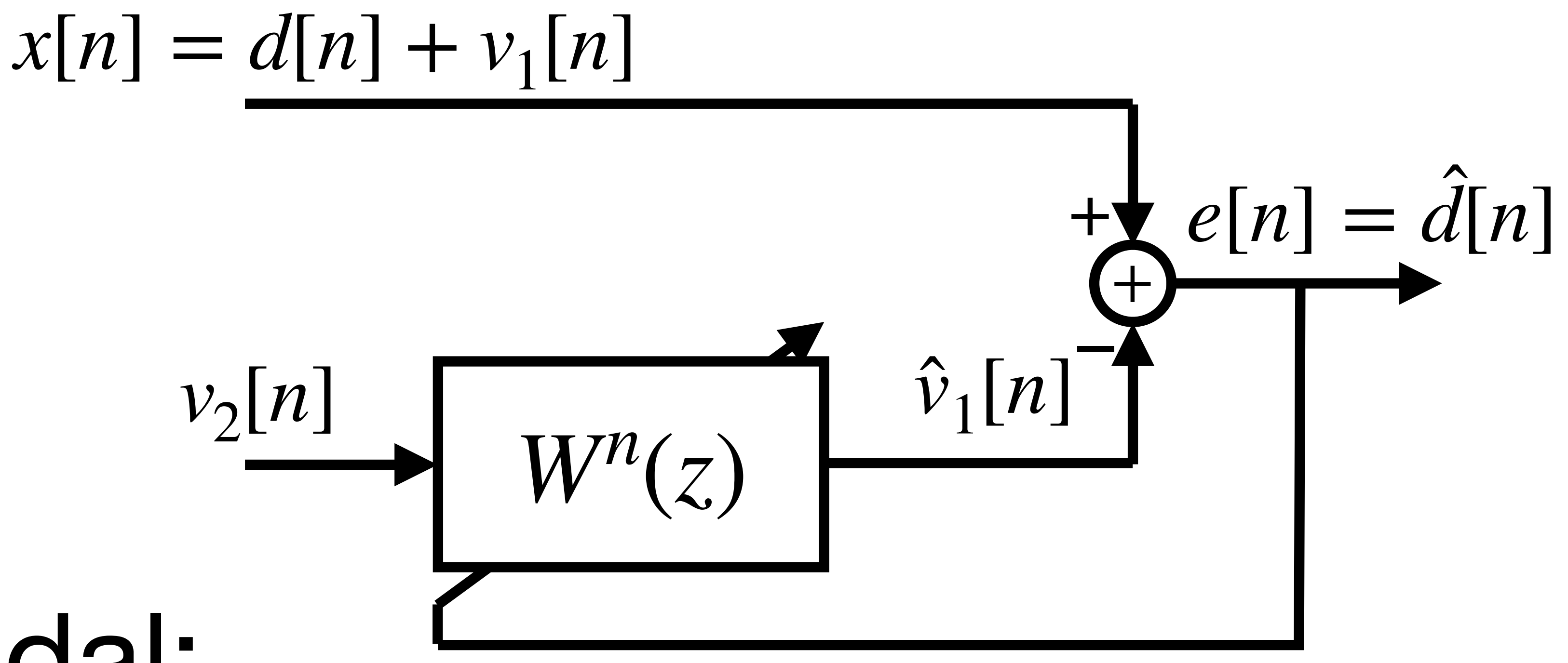
Filtros Adaptivos - Casos

Cancelador de ruido



Filtros Adaptivos - Casos

Cancelador de ruido



Suponiendo que la señal deseada $d[n]$ es una sinusoidal:

$$d[n] = \sin(n\omega_0 + \phi)$$

que la secuencia de ruido $v_1[n]$ y $v_2[n]$ son procesos AR(1) cuyas ecuaciones de diferencias son

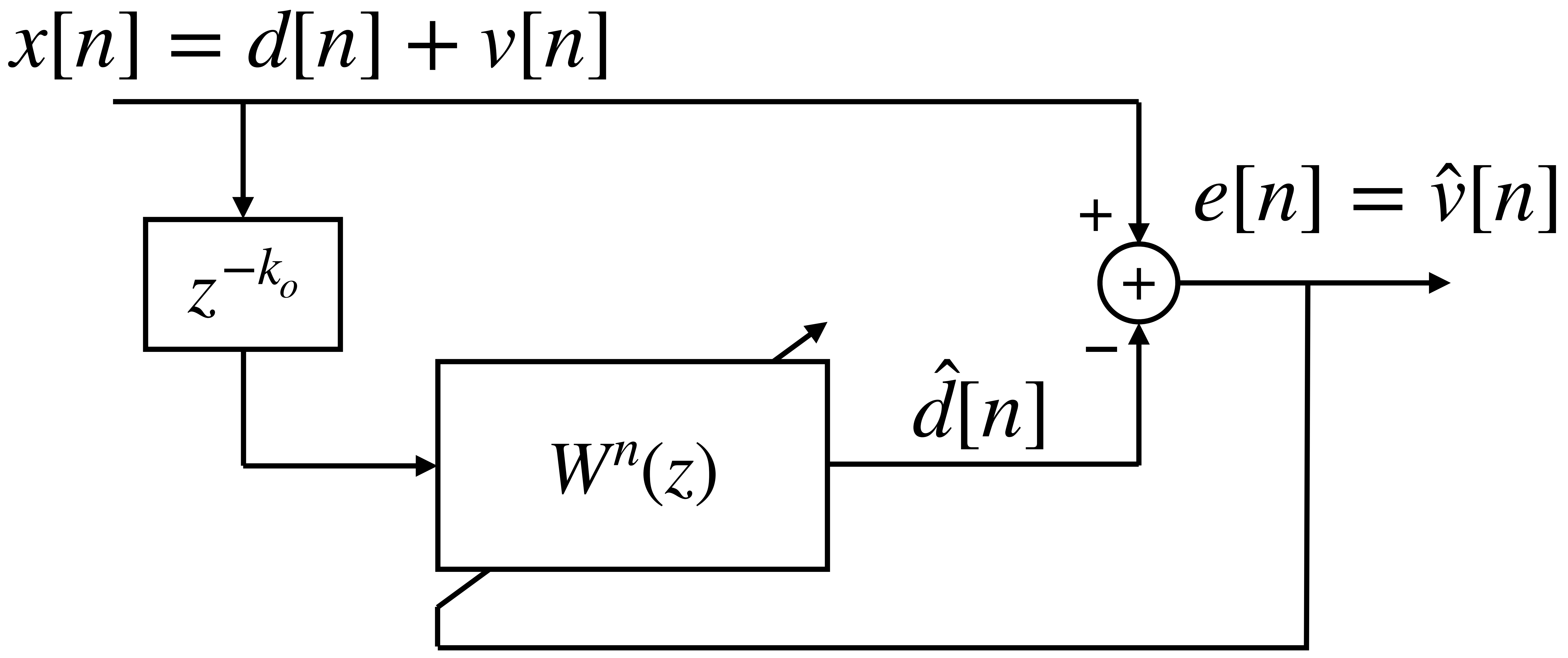
$$v_1[n] = 0.8\, v_1[n-1] + g[n] \implies H_1(z) = \frac{1}{1 - 0.8\, z^{-1}}$$

$$v_2[n] = -\,0.6\, v_2[n-1] + g[n] \implies H_2(z) = \frac{1}{1 + 0.6\, z^{-1}}$$

donde $g[n]$ es ruido blanco gaussiano con varianza igual a uno.

Filtros Adaptivos - Casos

Esquema de un predictor lineal



Algoritmo LMS

Datos adicionales

Aspectos de la convergencia:

- Dependiendo de μ, la convergencia puede ser más rápida o más lenta.
- Cuando $n \to \infty$, MSE temporal será constante.
- Cuando $n \to \infty$, $\bar{\mathbf{w}}^n \to \bar{\mathbf{w}}_{\text{opt}}$ (en la media).
- Si $x[n]$ y $d[n]$ son procesos WSS $\Longrightarrow \bar{\mathbf{w}}_{\text{opt}} = \mathbf{R}^{-1}\bar{\mathbf{r}}_{dx}$ (filtro Wiener).

Criterios de parada

- Número fijo de iteraciones
- Señal de error menor a un umbral: $|e[n]| < \tau$
- Error absoluto de filtros: $||\bar{\mathbf{w}}^{n+1} - \bar{\mathbf{w}}^n||_1 < \tau$

Algoritmo LMS

Convergencia

- Para procesos WSS, el algoritmo LMS converge en la media si

$$0 < \mu < \frac{1}{\lambda_{\max}},$$

  donde $\lambda_{\max}$ es el mayor autovalor de la matriz $\mathbf{R}_x$.

- Para procesos no WSS, el algoritmo LMS converge en la media cuadrática si

$$0 < \mu < \frac{1}{\sum_k \lambda_k} = \frac{1}{\mathrm{tr}\{\mathbf{R}_x\}}.$$

Nota: El autovalor $\lambda_{\max}$ está limitado por la traza de $\mathbf{R}_x$: $\lambda_{\max} \leq \sum_n \lambda_n = \mathrm{tr}\{\mathbf{R}_x\}$. Notar que la traza de $\mathbf{R}_x$ es la potencia del vector de entrada $\bar{\mathbf{x}}^n$.

LMS Normalizado (N-LMS)

- En la implementación de un filtro adaptivo LMS es muy importante la selección del paso μ :

$$0 < \mu < \frac{1}{\text{Potencia de entrada}} .$$

En procesos WSS: La potencia del vector de entrada es $M\, r_x[0]$.

En procesos no WSS: La potencia del vector de entrada es $\mathrm{tr}\{\mathbf{R}_x\}$.

- Si $x[n]$ es WSS entonces $\mathrm{tr}\{\mathbf{R}_x\} = M\, r_x[0] = M\, E\{\,|x[n]|^2\,\}$, luego, el límite más conservativo es

$$0 < \mu < \frac{1}{M\, E\{\,|x[n]|^2\,\}} .$$

- El valor esperado $E\{\,|x[n]|^2\,\}$ se puede estimar de la siguiente forma

$$\hat{E}\{\,|x[n]|^2\,\} = \frac{1}{M}\sum_{k=0}^{M-1} |x[n-k]|^2 .$$

LMS Normalizado (N-LMS)

Expresado matricialmente

$$0 < \mu < \frac{1}{(\bar{\mathbf{x}}^n)^T \bar{\mathbf{x}}^n} \,.$$

Esta cota se incorpora en el algoritmo LMS al usar un paso de tamaño variable

$$\mu = \frac{\beta}{(\bar{\mathbf{x}}^n)^T \bar{\mathbf{x}}^n} = \frac{\beta}{\|\bar{\mathbf{x}}^n\|^2} \,,$$

donde β es el tamaño de paso normalizado con cotas $0 < \beta < 1$.

La nueva ecuación de actualización de los coeficientes del filtro sería

$$\bar{\mathbf{w}}^{n+1} = \bar{\mathbf{w}}^n + 2\beta \frac{\bar{\mathbf{x}}^n}{\|\bar{\mathbf{x}}^n\|^2} e[n] \,.$$

La normalización $\|\bar{\mathbf{x}}^n\|^2$ altera la magnitud, no la dirección (gradiente).

LMS Normalizado (N-LMS)

En el algoritmo LMS, la gradiente es proporcional al vector $\bar{\mathbf{x}}[n]$.

- **Problema anterior:** Si $\bar{\mathbf{x}}^n$ es grande $\Longrightarrow$ amplificación de ruido de gradiente.

  Solución: La normalización mitiga este problema

- **Nuevo problema:** Si $||\bar{\mathbf{x}}^n||^2$ es muy pequeño $\Longrightarrow$ la gradiente explotaría.

  Solución:

  $\bar{\mathbf{w}}^{n+1} = \bar{\mathbf{w}}^n + 2\beta \dfrac{\bar{\mathbf{x}}^n}{||\bar{\mathbf{x}}^n||^2 + \epsilon} e[n]$, donde ϵ es un número pequeño positivo.

Nota: Comparando con algoritmo LMS, N-LMS requiere el computo adicional de $||\bar{\mathbf{x}}^n||^2$.

Solución recursiva:

$$||\bar{\mathbf{x}}^{n+1}||^2 = ||\bar{\mathbf{x}}^n||^2 + |x[n+1]|^2 - |x[n-M+1]|^2 \ .$$

¡Muchas gracias!