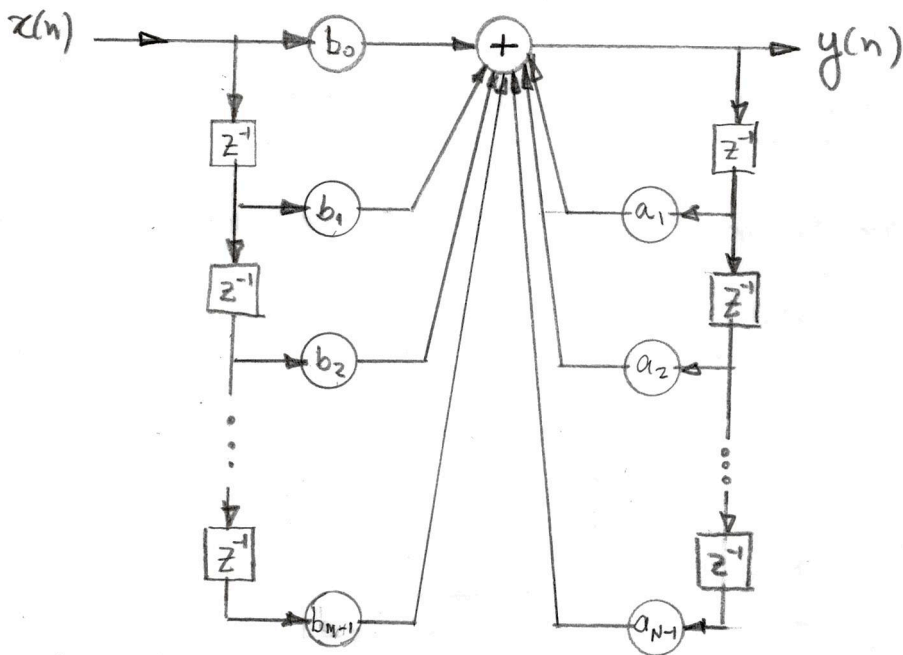


Filtragem Adaptativa IIR

Filtros IIR (Infinite Impulse Response) são estruturas de filtragem que apresentam tanto zeros quanto polos em sua função de transferência.

A principal vantagem de um filtro IIR em relação a um FIR é que ele pode proporcionar um desempenho superior com o mesmo número de coeficientes, devido à realimentação de saída.



Saída do filtro:

$$y(n) = b_0 x(n) + b_1 x(n-1) + \dots + b_{M-1} x(n-M+1) + a_1 y(n-1) + a_2 y(n-2) + \dots + a_{N-1} y(n-N+1)$$

$$y(n) = \sum_{j=0}^{M-1} b_j x(n-j) + \sum_{j=1}^{N-1} a_j y(n-j) = \underline{b}^T \underline{x}(n) + \underline{a}^T \underline{y}(n-1)$$

$$\underline{x}(n) = [x(n) \dots x(n-M+1)]^T$$

$$\underline{y}(n-1) = [y(n-1) \dots y(n-N+1)]^T$$

Função de transferência:

$$W(z) = \frac{b_0 + b_1 z^{-1} + \dots + b_{M-1} z^{-M+1}}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_{N-1} z^{-N+1}} = \frac{B(z)}{1 - A(z)}$$

em que:

$$A(z) = a_1 z^{-1} + a_2 z^{-2} + \dots + a_{N-1} z^{-N+1} = \sum_{j=1}^{N-1} a_j z^{-j}$$

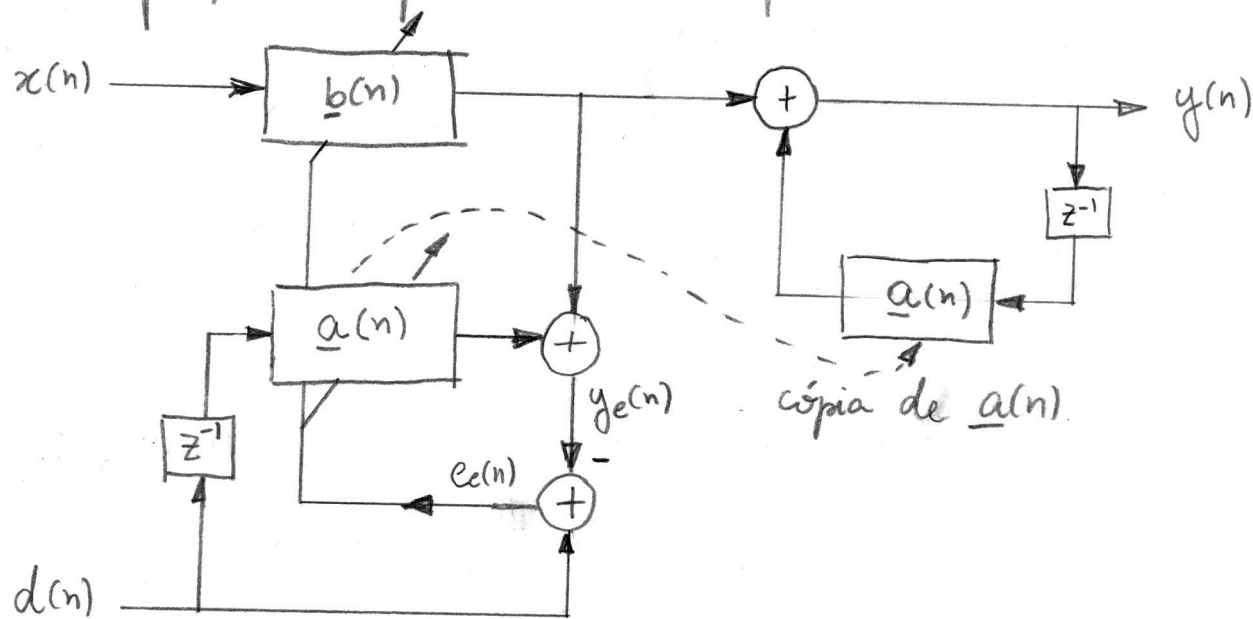
$$B(z) = b_0 + b_1 z^{-1} + \dots + b_{M-1} z^{-M+1} = \sum_{j=0}^{M-1} b_j z^{-j}$$

Desvantagens dos filtros IIR:

- 1) Instabilidade do filtro (pólos fora da CRU)
- 2) Convergência lenta dos algoritmos adaptativos
- 3) Existência de mínimos locais na função objetivo.
- 4) Ótimo global polarizado na abordagem equacion error.

Formulação Equation Error

Na formulação equation error é adotado o seguinte esquema para adaptação dos parâmetros do filtro:



Note que o sinal de saída do filtro é substituído pelo sinal desejado na adaptação dos parâmetros. Após a atualização, os novos valores de $\underline{a}(n)$ e $\underline{b}(n)$ são usados para gerar a saída $y(n)$. O erro usado na adaptação é dado por:

$$e_e(n) = d(n) - y_e(n)$$

em que

$$y_e(n) = \underline{b}^T(n) \underline{x}(n) + \underline{a}^T(n) \underline{d}(n-1)$$

com $\underline{d}(n-1) = [d(n-1) \ d(n-2) \ \dots \ d(n-N+1)]^T$

$y_e(n)$ pode ser visto como a saída de um filtro com duas entradas, $x(n)$ e $d(n)$, e uma única saída. Nesse caso não há realimentação da saída e a adaptação do filtro pode ser feita utilizando-se o formalismo desenvolvido para os filtros FIR.

Definindo-se os vetores:

$$\underline{\phi}_e(n) = [d(n-1) \dots d(n-N+1) \ x(n) \ x(n-1) \dots \tilde{x}(n-M+1)]^T$$

$$\underline{\theta}(n) = [a_1(n) \dots a_{N-1}(n) \ b_0(n) \ b_1(n) \dots b_{M-1}(n)]^T$$

Temos que

$$y_e(n) = \underline{\theta}^T(n) \underline{\phi}_e(n)$$

Qualquer versão dos algoritmos LMS e RLS pode ser aplicada na adaptação do vetor de parâmetros $\underline{\theta}(n)$, com a condição que todas as raízes do polinômio $1-A(z)$ estejam dentro da CRV. A abordagem equação erro pode convergir para um resultado que apresente uma polarização em relação à solução ótima. Por exemplo, em um problema de identificação de sistema, o algoritmo pode convergir para $E\{\underline{\theta}(n)\} = \underline{\theta}_0 + \text{bias}$ para $n \rightarrow \infty$, sendo $\underline{\theta}_0$ a solução ótima.

Pode ser mostrado que a polarização é nula quando:

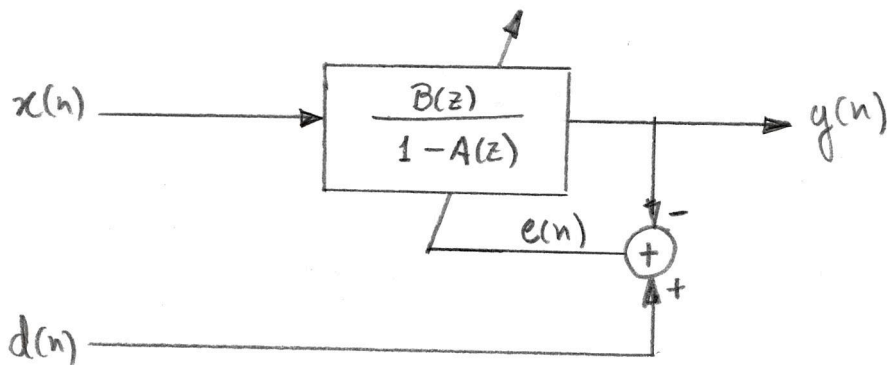
- 1- o ruído de medida aditivo presente na saída do sistema é nulo.
- 2- quando $\underline{a}(n) = 0$, ou seja, quando o filtro é FIR.

A polarização está relacionada ao uso das amostras passadas do sinal desejado ao invés das amostras de saída do filtro no cálculo do erro. Pois essa abordagem despreza a contribuição do ruído filtrado no processo.

Formulação Output Error

A abordagem output error corresponde à formulação usual do cálculo do erro, ou seja, à diferença entre o sinal de referência e a saída do filtro adaptativo:

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



Definindo-se o vetor:

$$\phi(n) = [y(n-1) \ y(n-2) \ \dots \ y(n-N+1) \ x(n) \ x(n-1) \ \dots \ x(n-M+1)]^T$$

A saída do filtro é dada por

$$y(n) = \underline{\theta}^T(n) \phi(n)$$

Devido à realimentação das amostras passadas da saída, $y(n)$ é uma função não-linear do vetor de parâmetros $\underline{\theta}(n)$. Pode-se perceber isso expandindo-se a expressão de $y(n)$:

$$y(n) = a_1(n)y(n-1) + b_0(n)x(n) = a_1(n) \left[a_1(n-1)y(n-2) + b_0(n-1)x(n-1) \right] + b_0(n)x(n)$$

$$= \underbrace{a_1(n)a_1(n-1)}_{\text{produto dos coeficientes de } \theta(n) \text{ em instantes distintos}} y(n-2) + \underbrace{a_1(n)b_0(n-1)}_{\text{produto dos coeficientes de } \theta(n) \text{ em instantes distintos}} x(n-1) + b_0(n)x(n)$$

produtos dos coeficientes de $\theta(n)$ em instantes distintos

Como $y(n)$ é uma função não linear dos parâmetros $\theta(n)$, o erro $e(n)$ também o é. Portanto, a superfície de MSE não é quadrática e pode apresentar mínimos locais. Entretanto, diferentemente do que ocorre na formulação equation error, o mínimo global da abundância output error não é polarizado.

Para o problema de identificação de sistemas, pode ser notado que nunca existem mínimos locais se:

- 1) A ordem do filtro adaptativo é igual ou superior à do sistema desconhecido.
- 2) A entrada $x(n)$ é branca.
- 3) A ordem do numerador do filtro adaptativo é maior que a ordem do denominador da função de transferência do sistema desconhecido.

Cálculo do gradiente

Como vimos anteriormente, é usual utilizar em filtros de adaptação uma aproximação instantânea do erro quadrático médio, ou seja, usar o erro quadrático instantâneo como função objetivo para adaptação dos parâmetros do filtro.

Neste caso, o gradiente de $e^2(n)$ é:

$$\nabla_{\Theta} e^2(n) = \frac{\partial e^2(n)}{\partial \Theta(n)} = 2e(n) \nabla_{\Theta} e(n) = -2e(n) \nabla_{\Theta} y(n)$$

em que

$$\nabla_{\Theta} y(n) = \left[\frac{\partial y(n)}{\partial a_1(n)}, \dots, \frac{\partial y(n)}{\partial a_{N-1}(n)}, \frac{\partial y(n)}{\partial b_0(n)}, \dots, \frac{\partial y(n)}{\partial b_{M-1}(n)} \right]^T$$

A derivada parcial da saída do filtro em relação ao coeficiente $a_j(n)$ é

$$\frac{\partial y(n)}{\partial a_j(n)} = \frac{\partial \sum_{i=0}^{M-1} b_i(n) x(n-i)}{\partial a_j(n)} + \frac{\partial \sum_{i=1}^{N-1} a_i(n) y(n-i)}{\partial a_j(n)}$$

$\rightarrow 0$ pois $b_i(n)$ e $x(n-i)$ não são funções de $a_j(n)$

Observe que

$$\frac{\partial a_i(n) y(n-i)}{\partial a_j(n)} = \frac{\partial a_i(n)}{\partial a_j(n)} y(n-i) + a_i(n) \frac{\partial y(n-i)}{\partial a_j(n)} = \begin{cases} a_i(n) \frac{\partial y(n-i)}{\partial a_j(n)}, & i \neq j \\ y(n-j) + a_j(n) \frac{\partial y(n-j)}{\partial a_j(n)}, & i = j \end{cases}$$

As derivadas parciais de $y(n-i)$ em relação aos parâmetros não se anulam pois o filtro tem realimentação. Assim, os coeficientes $a_i(n)$ atuais são funções dos valores passados da saída do filtro.

Temos então:

$$\frac{\partial y(n)}{\partial a_j(n)} = y(n-j) + \sum_{i=1}^{N-1} a_i(n) \frac{\partial y(n-i)}{\partial a_j(n)}$$

De modo semelhante, podemos calcular a derivada parcial de saída do filtro em relação aos coeficientes b_j 's:

$$\frac{\partial y(n)}{\partial b_j(n)} = \frac{\partial \sum_{i=0}^{N-1} b_i(n) x(n-i)}{\partial b_j(n)} + \frac{\partial \sum_{i=1}^{N-1} a_i(n) y(n-i)}{\partial b_j(n)}$$

$$= x(n-j) + \sum_{i=1}^{N-1} a_i(n) \frac{\partial y(n-i)}{\partial b_j(n)}$$

em que as derivadas parciais $\frac{\partial y(n-i)}{\partial b_j(n)}$ não se anulam devido à realimentação presente no filtro.

Considere que os valores dos parâmetros do filtro variam lentamente de modo que a seguinte aproximação é válida:

$$\underline{\theta}(n) \approx \underline{\theta}(n-1) \approx \dots \approx \underline{\theta}(n-N+1)$$

Desse modo, podemos dizer que $a_j(n) \approx a_j(n-i)$ e que $b_j(n) \approx b_j(n-i)$ para $1 \leq i \leq N-1$. Usando essas aproximações nas derivadas parciais, temos:

$$\frac{\partial y(n)}{\partial a_j(n)} \approx y(n-j) + \sum_{i=1}^{N-1} a_i(n) \frac{\partial y(n-i)}{\partial a_j(n-i)}$$

$$\frac{\partial y(n)}{\partial b_j(n)} \approx x(n-j) + \sum_{i=1}^{N-1} a_i(n) \frac{\partial y(n-i)}{\partial b_j(n-i)}$$

Observe que ao usar a aproximação as derivadas parciais se tornam recursivas em relação às suas versões atrasadas.

Definindo $D_{a_j}(z)$ como a transformada z da derivada parcial $\frac{\partial y(n)}{\partial a_j(n)}$ e $D_{b_j}(z)$ como a transformada z da derivada parcial $\frac{\partial y(n)}{\partial b_j(n)}$, temos

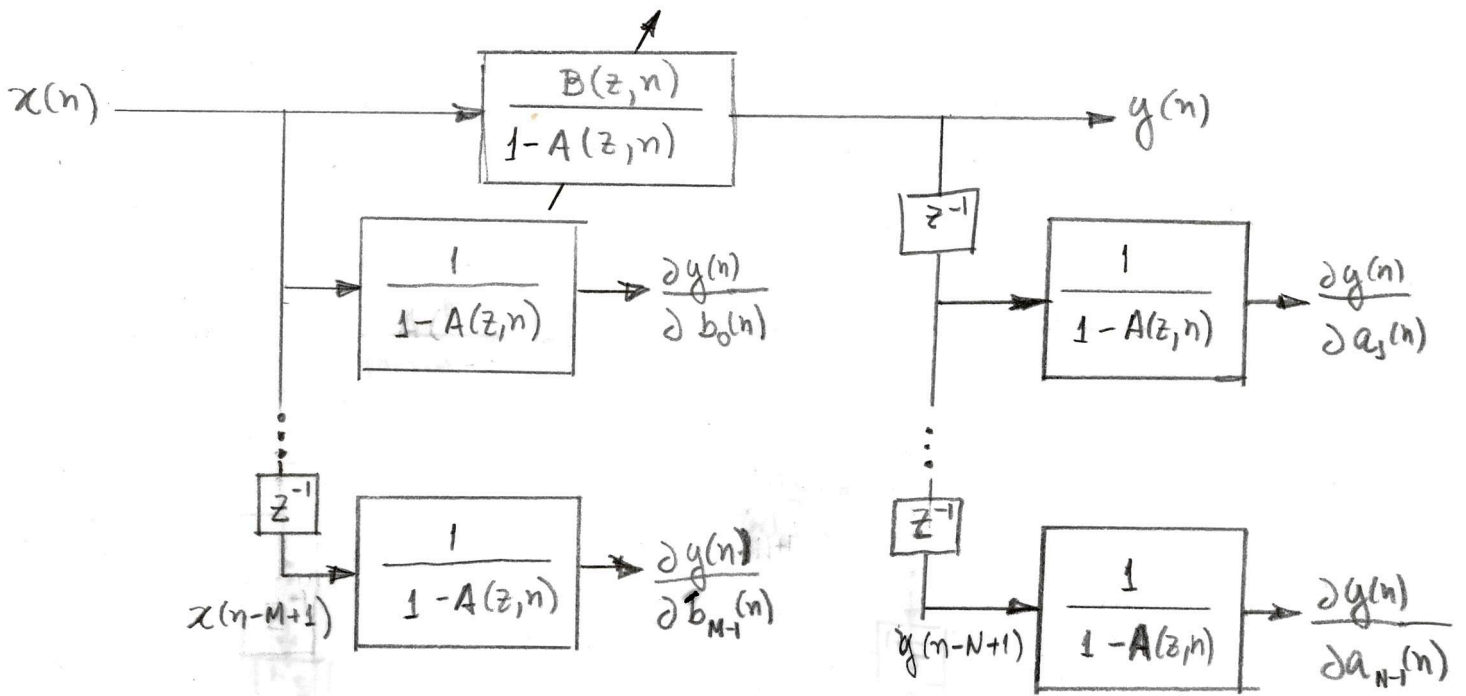
$$D_{a_j}(z) = Y(z)z^{-j} + \sum_{i=1}^{N-1} a_i(n) D_{a_j}(z) z^{-i} \Rightarrow \frac{D_{a_j}(z)}{Y(z)} = z^{-j} \left[\frac{1}{1 - \sum_{i=1}^{N-1} a_i(n) z^{-i}} \right]$$

$$\Rightarrow \frac{D_{a_j}(z)}{Y(z)} = z^{-j} \cdot \frac{1}{1 - A(z, n)}$$

$$\frac{D_{b_j}(z)}{X(z)} = z^{-j} \cdot \frac{1}{1 - A(z, n)}$$

em que $A(z, n)$ é o polinômio $A(z)$ no instante n .

Observe que as derivadas parciais podem ser obtidas a partir da filtragem de versões atrasadas da entrada e da saída por um filtro com os mesmos polos que o filtro adaptativo, conforme ilustrado no diagrama a seguir.

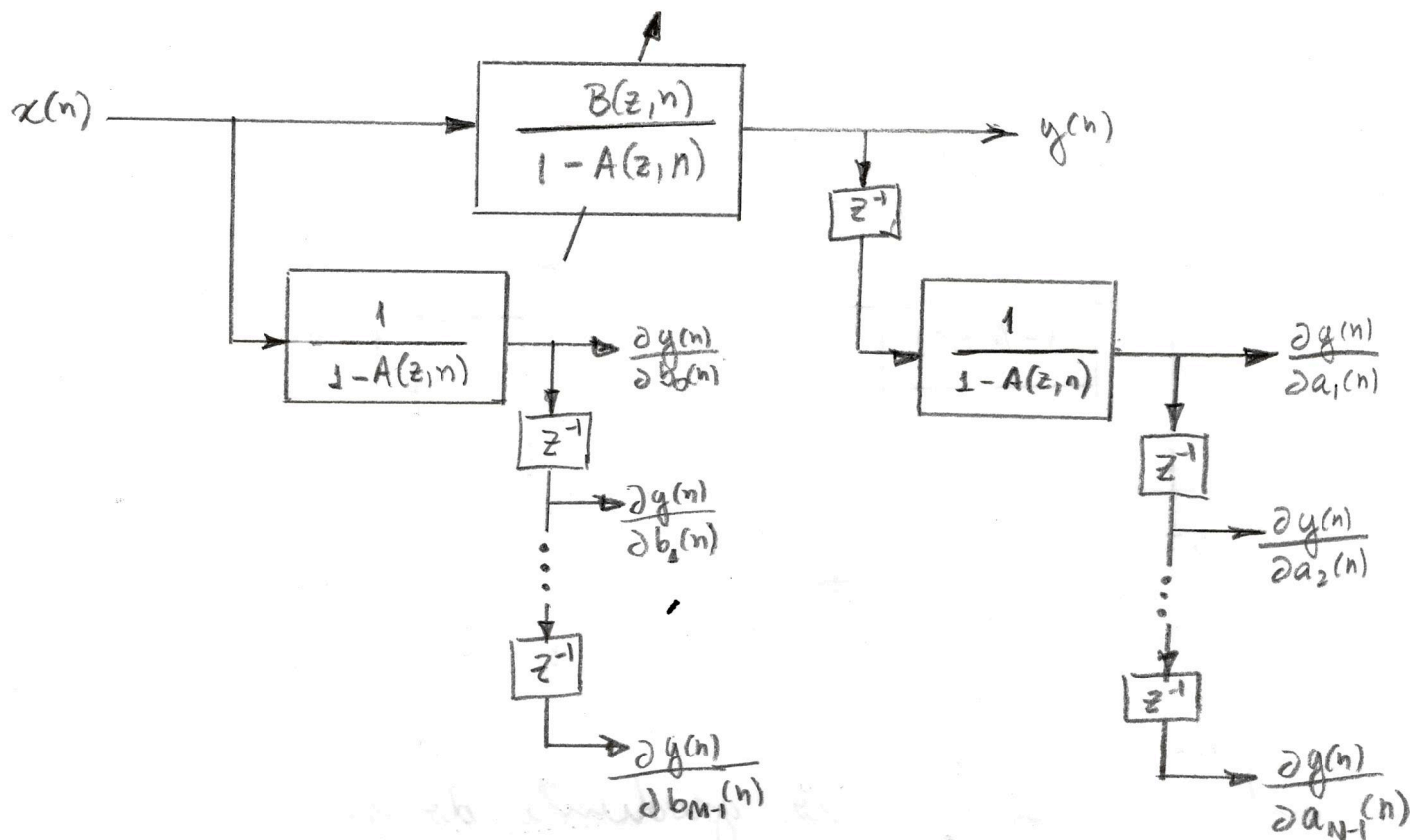


Portanto, o cálculo do gradiente do erro quadrático instantâneo requer o uso de $M+N-1$ filtros idênticos, o que representa uma complexidade computacional considerável em relação à filtragem do sinal de entrada propriamente dita.

Se considerarmos as seguintes aproximações:

$$\frac{\partial y(n)}{\partial a_j(n)} \approx \frac{\partial y(n-j+1)}{\partial a_1(n-j+1)} \quad \text{e} \quad \frac{\partial y(n)}{\partial b_j(n)} \approx \frac{\partial y(n-j)}{\partial b_0(n-j)}$$

ou seja, que as derivadas parciais da saída em relação aos coeficientes $a_j(n)$'s são versões atrasadas de $\partial y(n) / \partial a_1(n)$ e que as derivadas parciais da saída em relação aos coeficientes $b_j(n)$'s são versões atrasadas de $\partial y(n) / \partial b_0(n)$, podemos obter uma versão simplificada do gradiente que utiliza apenas dois filtros no seu cálculo, conforme ilustra o diagrama a seguir.



Essa simplificação no cálculo do gradiente introduz uma degradação muito pequena no desempenho dos algoritmos de adaptação do filtro IIR e por isso ele costuma ser usado na prática.

Algoritmos Adaptativos

Uma vez calculado o gradiente do erro quadrático instantâneo, podemos usá-lo para apresentar os algoritmos para ajuste adaptativo dos parâmetros dos filtros IIR.

LMS

A equação de adaptação dos parâmetros baseada no gradiente estocástico é dada por

$$\underline{\theta}(n+1) = \underline{\theta}(n) - \frac{\mu}{2} [\nabla_{\theta} e^2(n)]$$

em que $\nabla_{\theta} \tilde{e}(n) = -2e(n) \underline{\varphi}(n)$, com

$$\underline{\varphi}(n) = \nabla_{\theta} y(n) = \left[\frac{\partial y(n)}{\partial a_1(n)}, \dots, \frac{\partial y(n)}{\partial a_{N-1}(n)}, \frac{\partial y(n)}{\partial b_0(n)}, \dots, \frac{\partial y(n)}{\partial b_{M-1}(n)} \right]^T$$

Assim

$$\underline{\theta}(n+1) = \underline{\theta}(n) + \mu \underline{\varphi}(n) e(n)$$

O vetor $\underline{\varphi}(n)$, composto pelos derivados parciais da saída do filtro em relação aos seus parâmetros, pode ser obtido por um dos métodos apresentados anteriormente.

A cada atualização é necessário verificar a posição dos polos do filtro, isto é, as raízes do polinômio $1 - A(z, n)$, de modo a evitar instabilidade.

RLS

A versão do algoritmo RLS para adoção de filtros FIR usa a seguinte equação de atualização dos parâmetros:

$$\underline{\theta}(n+1) = \underline{\theta}(n) + R^{-1}(n) \underline{\varphi}(n) e(n)$$

com $R^{-1}(n)$ sendo calculado recursivamente a partir de equação

$$R^{-1}(n) = \frac{1}{\lambda} \left[R^{-1}(n-1) - \frac{R^{-1}(n-1) \underline{\varphi}(n) \underline{\varphi}^T(n) R^{-1}(n-1)}{\lambda + \underline{\varphi}^T(n) R^{-1}(n-1) \underline{\varphi}(n)} \right]$$

em que $0 \ll \lambda < 1$ é o fator de esquecimento.

Assim como no algoritmo LMS, a cada iteração é necessário verificar a estabilidade do filtro.

Observações:

* Existe um compromisso entre as abordagens equation error e output error. Por um lado, o equation error é uma função linear dos parâmetros de forma que a função objetivo é quadrática e apresenta um único mínimo. Os algoritmos adaptativos são os mesmos usados na adaptação de filtros FIR, apresentam convergência relativamente rápida, porém convergem para uma solução polarizada pois sempre há algum ruído aditivo. Por outro lado, o output error é uma função não-linear dos coeficientes e sua função objetivo é multimodal e possui pontos de sela e mínimos locais. Os algoritmos adaptativos geralmente convergem mais lentamente e podem convergir para um mínimo local dependendo de suas inicializações.

Referências

Shynk, John J.; "Adaptive IIR Filtering." ASSP Magazine,

IEEE 6.2 (1989): 4-21.

Diniz, Paulo S., "Adaptive Filtering: Algorithms and Practical Implementation", 4ed., 2013.