



Universidade Federal do Ceará
Centro de Tecnologia
Departamento de Engenharia de Teleinformática
Engenharia de Teleinformática

Filtragem Adaptativa
Listas de Exercícios Propostos

Student	Kenneth Brenner dos Anjos Benício – 519189
Professor	Charles Casimiro Cavalcante e Guilherme Barreto
Course	Filtragem Adaptativa - TIP7188

Fortaleza, 2021

Estatísticas de Segunda Ordem

1. Determine a média e a função de autocorrelação para o processo aleatório em que $v(n)$ é uma sequência de variáveis aleatórias independentes com média μ e variância σ^2 . $x(n)$ é estacionário? Justifique.

$$x(n) = v(n) + 3v(n-1), \quad (1)$$

Solução:

Para o cálculo da média é possível escrever

$$E[x(n)] = E[v(n) + 3v(n-1)], \quad (2)$$

mas uma vez que todas as variáveis aleatórias possuem mesma média o resultado será

$$E[x(n)] = \mu + 3\mu = 4\mu. \quad (3)$$

De forma semelhante a variância pode ser obtida por

$$\mathbb{E}\{[x(n) - \mu_X]^2\}, \quad (4)$$

$$\mathbb{E}\{[x(n) - 4\mu]^2\} = \mathbb{E}\{[v(n) - 3v(n-1) - 4\mu]^2\}, \quad (5)$$

$$\mathbb{E}\{[x(n) - 4\mu]^2\} = \mathbb{E}\{[v(n) - \mu + 3v(n-1) - 3\mu]^2\}, \quad (6)$$

e ao desenvolver o valor quadrado da expressão chega-se em

$$\mathbb{E}\{[x(n) - 4\mu]^2\} = \mathbb{E}\{[v(n) - \mu]^2 + 9[v(n-1) - \mu]^2 + 6[v(n) - \mu][v(n-1) - \mu]\}, \quad (7)$$

$$\mathbb{E}\{[x(n) - 4\mu]^2\} = \sigma^2 + 9\sigma^2 + \mathbb{E}\{6[v(n) - \mu][v(n-1) - \mu]\}. \quad (8)$$

Sabendo que os sinais são descorrelacionados uma vez que temos duas variáveis aleatórias independentes temos

$$6\mathbb{E}\{[v(n) - \mu][v(n-1) - \mu]\} = \mathbb{E}\{v(n)v(n-1) - \mu v(n) - \mu v(n-1) + \mu^2\}, \quad (9)$$

$$6\mathbb{E}\{[v(n) - \mu][v(n-1) - \mu]\} = \mathbb{E}\{v(n)v(n-1)\} - \mu^2 - \mu^2 + \mu^2, \quad (10)$$

$$6\mathbb{E}\{[v(n) - \mu][v(n-1) - \mu]\} = \mathbb{E}\{v(n)\}\mathbb{E}\{v(n-1)\} - \mu^2, \quad (11)$$

$$6\mathbb{E}\{[v(n) - \mu][v(n-1) - \mu]\} = \mu^2 - \mu^2 = 0, \quad (12)$$

Portanto, a média pode ser dada por

$$\mathbb{E}\{[x(n) - 4\mu]^2\} = 10\sigma^2. \quad (13)$$

Além disso, a função de correlação pode ser adquirida por meio da seguinte operação

$$\mathbb{E}\{x(n)x^*(n)\} = \mathbb{E}\{[v(n) + 3v(n-1)][v(n) + 3v(n-1)]^*\}, \quad (14)$$

$$\mathbb{E}\{x(n)x^*(n)\} = \mathbb{E}\{[v(n)v^*(n)] + 3[v(n)v^*(n-1)] + 3[v(n-1)v^*(n)] + [9v(n-1)v^*(n-1)]\}, \quad (15)$$

onde é possível reescrever a equação acima ao considerarmos novamente a decorrelação entre as variáveis aleatórias

$$\mathbb{E}\{x(n)x^*(n)\} = \mathbb{E}\{\mu^2 + 3\mu^2 + 3\mu^2 + 9\mu^2\}, \quad (16)$$

$$\mathbb{E}\{x(n)x^*(n)\} = 16\mu^2. \quad (17)$$

Uma vez que as estatísticas de primeira e de segunda ordem são independentes quanto ao deslocamento temporal é possível afirmar que o processo descrito é Estacionário no Sentido Amplo(WSS). Entretanto, não é possível afirmar que o processo é Estacionário no Sentido Estrito(SSS) uma vez que não se tem conhecimento da função que descreve a densidade de probabilidade desse sistema.

2. Sejam os processos aleatórios $x(n)$ e $y(n)$ definidos por

$$x(n) = v_1(n) + 3v_2(n-1), \quad (18)$$

$$y(n) = v_2(n+1) + 3v_1(n-1), \quad (19)$$

em que $v_1(n)$ e $v_2(n)$ são processos de ruído branco independentes cada um com variância igual a 0,5.

(a) Quais são as funções de autocorrelação de x e de y ? Os processos são WSS?

Solução:

Antes de tudo é de interesse definir o processo estocástico de ruído branco. Um processo desse tipo é formado por um número qualquer de amostras independentes entre si e com média nula. Desse modo, uma vez que a média de processos de ruído branco são nulas é possível obter facilmente a média dos novos processos formados pela combinação dos ruídos

$$\mathbb{E}\{x(n)\} = \mathbb{E}\{v_1(n) + 3v_2(n-1)\} = \mu_1 + 3\mu_2 = 0, \quad (20)$$

$$\mathbb{E}\{y(n)\} = \mathbb{E}\{v_2(n+1) + 3v_1(n-1)\} = \mu_2 + 3\mu_1 = 0. \quad (21)$$

Para a variância teremos uma resposta similar

$$\mathbb{E}\{[x(n) - \mu]^2\} = \mathbb{E}\{[x(n) - 0]^2\} = \mathbb{E}\{[x(n)]^2\},$$

$$\mathbb{E}\{[x(n) - \mu]^2\} = \mathbb{E}\{[v_1(n) + 3v_2(n-1)]^2\} = \mathbb{E}\{[v_1^2(n)] + 6[v_1(n)v_2(n-1)] + 9[v_2^2(n-1)]\},$$

$$\mathbb{E}\{[y(n) - \mu]^2\} = \mathbb{E}\{[y(n) - 0]^2\} = \mathbb{E}\{[y(n)]^2\},$$

$$\mathbb{E}\{[y(n) - \mu]^2\} = \mathbb{E}\{[v_2(n+1) + 3v_1(n-1)]^2\} = \mathbb{E}\{[v_2^2(n+1)] + 6[v_2(n+1)v_1(n-1)] + 9[v_1^2(n-1)]\}.$$

Rescrevendo a expressão para obter a formula da variância e utilizando o fato dos processos de ruído serem decorrelacionados uma nova expressão é obtida

$$\mathbb{E}\{[x(n) - \mu]^2\} = \mathbb{E}\{[v_1^2(n) - 0] + 0 + 9[v_2^2(n-1) - 0]\}, \quad (22)$$

$$\mathbb{E}\{[x(n) - \mu]^2\} = \sigma_1^2 + 9\sigma_2^2 = 10 * 0.5 = 5, \quad (23)$$

$$\mathbb{E}\{[y(n) - \mu]^2\} = \mathbb{E}\{[v_2^2(n+1) - 0] + 0 + 9[v_1^2(n-1) - 0]\}, \quad (24)$$

$$\mathbb{E}\{[y(n) - \mu]^2\} = \sigma_2^2 + 9\sigma_1^2 = 10 * 0.5 = 5. \quad (25)$$

A função de autocorrelação para X é calculada como se segue

$$r_x(\tau) = \mathbb{E}\{x(n)x(n+\tau)\} = \mathbb{E}\{[v_1(n) + 3v_2(n-1)][v_1(n+\tau) + 3v_2(n-1+\tau)]\}, \quad (26)$$

$$r_x(\tau) = \mathbb{E}\{[v_1(n)v_1(n+\tau)] + 3[v_1(n)v_2(n-1+\tau)] + 3[v_2(n-1)v_1(n+\tau)] \quad (27)$$

$$+ 9[v_2(n-1)v_2(n-1+\tau)]\}. \quad (28)$$

Sabendo que amostras de um mesmo processo são descorrelacionadas e que ambos os processos também são descorrelacionados, além de termos conhecimento de que ambos possuem média nula, então

$$\mathbb{E}\{[v_1(n)v_1(n+\tau)]\} = \mathbb{E}\{[v_1(n)]\}\mathbb{E}\{[v_1(n+\tau)]\} = 0, \quad (29)$$

$$3\mathbb{E}\{[v_1(n)v_2(n-1+\tau)]\} = 3\mathbb{E}\{[v_1(n)]\}\mathbb{E}\{[v_2(n-1+\tau)]\} = 0, \quad (30)$$

$$3\mathbb{E}\{[v_2(n-1)v_1(n+\tau)]\} = 3\mathbb{E}\{[v_2(n-1)]\}\mathbb{E}\{[v_1(n+\tau)]\} = 0, \quad (31)$$

$$9\mathbb{E}\{[v_2(n-1)v_2(n-1+\tau)]\} = 9\mathbb{E}\{[v_2(n-1)]\}\mathbb{E}\{[v_2(n-1+\tau)]\} = 0, \quad (32)$$

$$r_x(\tau) = 0. \quad (33)$$

Repetindo o procedimento para Y e omitindo informações redundantes chega-se ao seguinte resultado

$$r_y(\tau) = 0. \quad (34)$$

Sendo as estatísticas de primeira e de segunda ordem independentes quanto ao tempo para os dois processos então é possível afirmar que eles são WSS.

- (b) Qual é a função de correlação cruzada $r_{xy}(n_1, n_0)$? Estes processos são conjuntamente estacionários (no sentido amplo)? Justifique.

Solução:

$$r_{x,y}(n_1, n_0) = \mathbb{E}\{[x(n_1)y^*(n_0)]\} = \mathbb{E}\{[v_1(n_1) + 3v_2(n_1-1)][v_2(n_0+1) + 3v_1(n_0-1)]^*\}, \quad (35)$$

$$r_{x,y}(n_1, n_0) = \mathbb{E}\{[v_1(n_1)v_2^*(n_0+1)] + 3[v_1(n_1)v_1^*(n_0-1)] + 3[v_2(n_1-1)v_2^*(n_0+1)] \quad (36)$$

$$+ 9[v_2(n_1-1)v_1^*(n_0-1)]\}. \quad (37)$$

Novamente considerando que os processos de ruído branco são descorrelacionadas entre si e que as amostras individuais são também independentes então podemos simplificar a expressão acima da seguinte forma

$$r_{x,y}(n_1, n_0) = 0 \quad (38)$$

Sendo assim, os processos podem ser considerados conjuntamente estacionários uma vez que a correlação cruzada é independente do instante temporal e que os processos que compoem o processo conjunto podem ser considerados estacionários em sentido amplo isoladamente. Entretanto, vale notar que foi possível realizar tal afirmação apenas por termos conhecimento da relação de independência entre os processos de ruído branco, pois de outro modo não seria possível realizar afirmações quanto a correlação cruzada do processo conjunto.

3. Quais as condições que os elementos de uma matriz

$$\mathbf{R} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad (39)$$

devem satisfazer tal que \mathbf{R} seja uma matriz de autocorrelação válida de

(a) Um vetor aleatório bidimensional?

Solução:

Em suma, precisamos estar atentos às seguintes propriedades para garantir que temos em mãos uma matriz de autocorrelação válida

$$- \mathbf{R}_{\mathbf{x}} = \mathbf{R}_{\mathbf{x}}^H$$

$$- \mathbf{a}^H \mathbf{R}_{\mathbf{x}} \mathbf{a} \geq 0$$

$$- \mathbf{A} \mathbf{x} = \lambda \mathbf{x}, \forall \lambda \geq 0 \text{ and } \mathbf{x} \in \mathbb{R}$$

Desse modo, considerando um vetor aleatório bidimensional descrito por $\mathbf{X} = (x_1, x_2)$ podemos então

$$\mathbf{R} = \begin{bmatrix} \mathbb{E}\{[x_1^2]\} & \mathbb{E}\{[x_1 x_2^*]\} \\ \mathbb{E}\{[x_2 x_1^*]\} & \mathbb{E}\{[x_2^2]\} \end{bmatrix}, \quad (40)$$

$$\mathbf{R}^H = \begin{bmatrix} \mathbb{E}\{[x_1^2]\} & \mathbb{E}\{[x_2 x_1^*]\} \\ \mathbb{E}\{[x_1 x_2^*]\} & \mathbb{E}\{[x_2^2]\} \end{bmatrix}. \quad (41)$$

Inicialmente, podemos garantir a simetria quanto ao hermitiano se fizermos

$$\mathbb{E}\{[x_1 x_2^*]\} \triangleq \mathbb{E}\{[x_2 x_1^*]\}, \quad \mathbb{E}\{[x_2 x_1^*]\} \triangleq \mathbb{E}\{[x_1 x_2^*]\}. \quad (42)$$

Entretanto, sabemos que o operador esperança é linear tornando assim as expressões acima equivalentes. Em continuidade do problema, podemos considerar que a restrição as quais os autovalores estão submetidos pode ser facilmente atingida ao garantirmos que o determinante da matriz de correlação seja maior que a nulidade

$$\mathbb{E}\{[x_1^2]\} \mathbb{E}\{[x_2^2]\} - \mathbb{E}\{[x_1 x_2]\} \mathbb{E}\{[x_2 x_1]\} > 0, \quad (43)$$

$$\mathbb{E}\{[x_1^2]\} \mathbb{E}\{[x_2^2]\} > \mathbb{E}\{[x_1 x_2]\} \mathbb{E}\{[x_2 x_1]\}. \quad (44)$$

(b) Um processo estocástico estacionário escalar?

Solução:

Considerando um processo estocástico estacionário escalar do tipo $\mathbf{X}_t = x(t)$ e uma versão atrasada desse processo definida por $\mathbf{X}_{t+\tau} = x(t + \tau)$ temos que a matriz de correlação pode ser escrita da seguinte forma

$$\mathbf{R} = \begin{bmatrix} \mathbb{E}\{[x^2(t)_1]\} & \mathbb{E}\{[x(t)x^*(t + \tau)]\} \\ \mathbb{E}\{[x(t + \tau)x^*(t)]\} & \mathbb{E}\{[x^2(t + \tau)]\} \end{bmatrix}, \quad (45)$$

$$\mathbf{R}^H = \begin{bmatrix} \mathbb{E}\{[x^2(t)_1]\} & \mathbb{E}\{[x(t + \tau)x^*(t)]\} \\ \mathbb{E}\{[x(t)x^*(t + \tau)]\} & \mathbb{E}\{[x^2(t + \tau)]\} \end{bmatrix}. \quad (46)$$

Em sequência, podemos garantir a simetria quanto ao hermitiano se fizermos

$$\mathbb{E}\{[x(t)x^*(t + \tau)]\} \triangleq \mathbb{E}\{[x(t + \tau)x^*(t)]\}, \quad (47)$$

$$\mathbb{E}\{[x(t + \tau)x^*(t)]\} \triangleq \mathbb{E}\{[x(t)x^*(t + \tau)]\}, \quad (48)$$

mas, mais uma vez, considerando que o operador esperança é linear então as duas expressões são equivalentes. Já considerando a restrição imposta aos autovalores da matriz temos novamente

$$\mathbb{E}\{[x^2(t)]\} \mathbb{E}\{[x^2(t + \tau)]\} > \mathbb{E}\{[x(t)x^*(t + \tau)]\} \mathbb{E}\{[x(t + \tau)x^*(t)]\}. \quad (49)$$

4. Assuma que a inversa $\mathbf{R}_{\mathbf{x}}^{-1}$ da matriz de autocorrelação de um vetor coluna N -dimensional exista. Mostre que

$$\mathbb{E}\{\mathbf{x}^H \mathbf{R}_{\mathbf{x}}^{-1} \mathbf{x}\} = N \quad (50)$$

Solução:

Inicialmente podemos escrever a expressão regular para a matriz de autocorrelação e supondo que de fato existe uma inversa bem definida para a matriz de autocorrelação

$$\mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \mathbf{R}_x, \quad (51)$$

$$\mathbb{E}\{\mathbf{x}\mathbf{x}^H\} \mathbf{R}_x^{-1} = \mathbf{R}_x \mathbf{R}_x^{-1}, \quad (52)$$

$$\mathbb{E}\{\mathbf{x}\mathbf{x}^H \mathbf{R}_x^{-1}\} = \mathbf{I}_N, \quad (53)$$

$$(54)$$

onde a última passagem foi possível pois consideramos que, embora de natureza aleatória, a matriz de autocorrelação possui medições fixas após o término da coleta das amostras. Desse modo, podemos aplicar o operador traço de matriz e por meio das propriedades relevantes é possível alterar de forma conveniente a ordem do produto dos termos, somente ao aplicarmos uma permutação cíclica, obtendo portanto o seguinte resultado

$$\text{Tr}\{\mathbb{E}\{\mathbf{x}\mathbf{x}^H \mathbf{R}_x^{-1}\}\} = \text{Tr}\{\mathbf{I}_N\}, \quad (55)$$

$$\text{Tr}\{\mathbb{E}\{\mathbf{x}^H \mathbf{R}_x^{-1} \mathbf{x}\}\} = \text{Tr}\{\mathbf{I}_N\}, \quad (56)$$

$$\text{Tr}\{\mathbb{E}\{\mathbf{x}^H \mathbf{R}_x^{-1} \mathbf{x}\}\} = N, \quad (57)$$

5. Mostre que as matrizes de correlação e covariância satisfazem as relações abaixo:

(a) $\mathbf{R}_{\mathbf{x}} = \mathbf{C}_{\mathbf{x}} + \mu_{\mathbf{x}} \mu_{\mathbf{x}}^H$

Solução:

$$C_X = \mathbb{E}\{(x - \mu)(x - \mu)^H\} = \mathbb{E}\{xx^H - x\mu^H - \mu x^H + \mu\mu^H\}, \quad (58)$$

$$C_X = \mathbb{E}\{xx^H\} - \mathbb{E}\{x\mu^H\} - \mathbb{E}\{\mu x^H\} + \mathbb{E}\{\mu\mu^H\}. \quad (59)$$

Considerando a matriz de correlação pode ser escrita por $R_X = \mathbb{E}\{xx^H\}$ e que o valor médio de um escalar é o próprio escalar temos

$$C_X = R_X - \mathbb{E}\{x\mu^H\} - \mathbb{E}\{\mu x^H\} + \mu\mu^H, \quad (60)$$

$$C_X = R_X - \mu^H \mathbb{E}\{x\} - \mu \mathbb{E}\{x^H\} + \mu\mu^H, \quad (61)$$

$$C_X = R_X - \mu\mu^H - \mu\mu^H + \mu\mu^H, \quad (62)$$

$$C_X = R_X - \mu\mu^H, \quad (63)$$

$$R_X = C_X + \mu\mu^H. \quad (64)$$

- (b) $\mathbf{C}_{\mathbf{x}+\mathbf{y}} = \mathbf{C}_{\mathbf{x}} + \mathbf{C}_{\mathbf{y}}$, para \mathbf{x} e \mathbf{y} descorrelacionados

Solução:

$$\mathbf{C}_{\mathbf{x}+\mathbf{y}} = \mathbf{C}_{\mathbf{x}} + \mathbf{C}_{\mathbf{xy}} + \mathbf{C}_{\mathbf{yx}} + \mathbf{C}_{\mathbf{y}}. \quad (65)$$

Onde os termos de correlação cruzada $\mathbf{C}_{\mathbf{xy}}$ e $\mathbf{C}_{\mathbf{yx}}$ podem ser obtidos da seguinte maneira

$$\mathbf{C}_{\mathbf{xy}} = \mathbb{E}\{[x - \mu_x][y - \mu_y]\}, \quad (66)$$

$$\mathbf{C}_{\mathbf{xy}} = \mathbb{E}\{[xy]\} - \mathbb{E}\{[x\mu_y]\} - \mathbb{E}\{[\mu_x y]\} + \mathbb{E}\{[\mu_x \mu_y]\}, \quad (67)$$

$$\mathbf{C}_{\mathbf{xy}} = \mathbb{E}\{[xy]\} - \mu_y \mu_x - \mu_x \mu_y + \mu_x \mu_y, \quad (68)$$

$$\mathbf{C}_{\mathbf{xy}} = \mathbb{E}\{[xy]\} + \mu_x \mu_y, \quad (69)$$

$$\mathbf{C}_{\mathbf{xy}} = \mu_x \mu_y, \quad (70)$$

$$(71)$$

$$\mathbf{C}_{\mathbf{yx}} = \mathbb{E}\{[y - \mu_y][x - \mu_x]\}, \quad (72)$$

$$\mathbf{C}_{\mathbf{yx}} = \mathbb{E}\{[yx]\} - \mathbb{E}\{[y\mu_x]\} - \mathbb{E}\{[\mu_y x]\} + \mathbb{E}\{[\mu_y \mu_x]\}, \quad (73)$$

$$\mathbf{C}_{\mathbf{yx}} = \mathbb{E}\{[yx]\} - \mu_y \mu_x - \mu_x \mu_y + \mu_y \mu_x, \quad (74)$$

$$\mathbf{C}_{\mathbf{yx}} = \mathbb{E}\{[yx]\} - \mu_x \mu_y, \quad (75)$$

$$\mathbf{C}_{\mathbf{yx}} = -\mu_x \mu_y. \quad (76)$$

Portanto, é possível afirmar que a soma dos termos de correlação cruzada é nula e que a correlação da soma de dois vetores descorrelacionados é de fato dada por $\mathbf{C}_{\mathbf{x}+\mathbf{y}} = \mathbf{C}_{\mathbf{x}} + \mathbf{C}_{\mathbf{y}}$

$$\mathbf{C}_{\mathbf{xy}} + \mathbf{C}_{\mathbf{yx}} = \mu_x \mu_y - \mu_x \mu_y, \quad (77)$$

$$\mathbf{C}_{\mathbf{xy}} + \mathbf{C}_{\mathbf{yx}} = 0. \quad (78)$$

6. Processos aleatórios $v_1(n)$ e $v_2(n)$ são independentes e têm a mesma função de correlação

$$r_v(n_1, n_0) = 0.5\delta(n_1 - n_0) \quad (79)$$

(a) Qual é a função de correlação do processo aleatório

$$x(n) = v_1(n) + 2v_1(n+1) + 3v_2(n-1)? \quad (80)$$

Este é um processo WSS? Justifique.

Solução:

Para simplificar o desenvolvimento iremos considerar que os processos possuem media nula e são descorrelacionados

$$r_x = \mathbb{E}\{x(n)x^*(n)\},$$

$$r_x = \mathbb{E}\{[v_1(n) + 2v_1(n+1) + 3v_2(n-1)][v_1(n) + 2v_1(n+1) + 3v_2(n-1)]^*(n)\},$$

$$r_x = \mathbb{E}\{[v_1(n)v_1^*(n) + 2v_1(n)v_1^*(n+1) + 3v_1(n)v_2^*(n-1) + 2v_1(n+1)v_1^*(n) + 4v_1(n+1)v_1^*(n+1) + 6v_1(n+1)v_2^*(n-1) + 3v_2(n-1)v_1^*(n) + 6v_2(n-1)v_1^*(n+1) + 9v_2(n-1)v_2^*(n-1)]\},$$

$$r_x = r_v(n, n) + 2r_v(n, n+1) + 0 + 2r_v(n+1, n) + 4r_v(n+1, n+1) + 0 + 0 + 0 + 9r_v(n-1, n-1),$$

$$r_x = r_v(n, n) + 2r_v(n, n+1) + 2r_v(n+1, n) + 4r_v(n+1, n+1) + 9r_v(n-1, n-1).$$

Após uma breve análise na expressão $r_v(n_1, n_0)$ é possível verificar que determinados termos anulam-se ao considerar um único momento temporal devido a presença da função degrau unitário, permitindo a seguinte simplificação:

$$r_x = 2r_v(n, n+1) + 2r_v(n+1, n), \quad (81)$$

$$r_x = \delta(n - n - 1) + \delta(n + 1 - n). \quad (82)$$

Onde a generalização pode ser descrita por:

$$r_x(n_1, n_2) = \delta(n_1 - n_2) + \delta(n_2 - n_1). \quad (83)$$

Uma vez que a correlação é dependente apenas de um deslocamento temporal, então podemos classificar esse processo como WSS.

- (b) Encontre a a matrix de correlação de um vetor aleatório consistindo de oito amostras consecutivas de $x(n)$.

Solução:

$$\mathbf{R}_x = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \quad (84)$$

Para chegar a esse resultado foi utilizado a expressão para a correlação obtida no item anterior. Uma vez que a função delta de dirac terá um valor não nulo apenas quando o argumento for zero, isso irá acontecer apenas quando os momentos n_1 e n_2 forem iguais e isso irá acontecer apenas com os elementos da diagonal principal.

Filtragem Linear Ótima

1. Considere um problema de filtragem de Wiener conforme caracterizado a seguir. A matriz de correlação \mathbf{R}_x de um vetor de entrada $\mathbf{x}(n)$ é dada por

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}.$$

O vetor de correlação cruzada \mathbf{p}_{xd} entre o vetor de entrada \mathbf{x} e a resposta desejada $d(n)$ é

$$\mathbf{p}_{xd} = \begin{bmatrix} 0.5 \\ 0.25 \end{bmatrix}$$

- (a) Encontre o vetor de coeficientes do filtro de Wiener.

Solução:

Isso pode ser realizado de forma simples pelo uso da equação do filtro ótimo de wiener

$$\mathbf{w}_{\text{opt}} = \mathbf{R}_X^{-1} \mathbf{p}_{xd}, \quad (85)$$

$$\mathbf{w}_{\text{opt}} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0.5 \\ 0.25 \end{bmatrix}, \quad (86)$$

$$\mathbf{w}_{\text{opt}} = \frac{4}{3} \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.25 \end{bmatrix}, \quad (87)$$

$$\mathbf{w}_{\text{opt}} = \begin{bmatrix} 0.5 \\ 0.0 \end{bmatrix}. \quad (88)$$

- (b) Qual é o mínimo erro médio quadrático fornecido por este filtro?

Solução:

$$\mathbb{E}\{e^2(n)\} = \sigma_d^2 - 2\mathbf{w}^T \mathbf{p}_{xd} + \mathbf{w}^T \mathbf{R}_X \mathbf{w} \quad (89)$$

Ao aplicar o vetor descoberto no item anterior obtem-se o erro mínimo

$$e_{\min} = \sigma_d^2 - 2\mathbf{w}_{\text{opt}}^T \mathbf{p}_{xd} + \mathbf{w}_{\text{opt}}^T \mathbf{R}_X \mathbf{w}_{\text{opt}}, \quad (90)$$

$$e_{\min} = \sigma_d^2 - 2 \begin{bmatrix} 0.5 & 0.0 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.25 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.0 \end{bmatrix} \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.0 \end{bmatrix}, \quad (91)$$

$$e_{\min} = \sigma_d^2 - 2 * 0.25 + 0.25, \quad (92)$$

$$e_{\min} = \sigma_d^2 - 0.25. \quad (93)$$

Dessa forma o erro é dependente da variância do sinal desejado.

- (c) Formule uma representação do filtro de Wiener em termos dos autovalores da matriz \mathbf{R}_x e de seus autovetores associados.

Solução:

Utilizando a decomposição matricial em autovalores (EVD) é possível reescrever a matrix de correlação como abaixo

$$\mathbf{R}_X = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}. \quad (94)$$

A matriz $\mathbf{\Lambda}$ contém os autovalores λ_i e a matriz \mathbf{Q} os respectivos autovetores. Em posse dessa relação é possível reescrever a equação do filtro ótimo de Wiener como

$$\mathbf{w}_{\text{opt}} = \mathbf{R}_X^{-1} \mathbf{p}_{xd}, \quad (95)$$

$$\mathbf{w}_{\text{opt}} = (\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1})^{-1} \mathbf{p}_{xd}, \quad (96)$$

$$\mathbf{w}_{\text{opt}} = \mathbf{Q}^{-1} \mathbf{\Lambda}^{-1} \mathbf{Q} \mathbf{p}_{xd}. \quad (97)$$

É possível verificar de imediato que essa propriedade é bem útil uma vez que a inversa de uma matriz diagonal é certamente menos custosa que a inversa da matriz de autocorrelação completa.

2. Mostre que a equação do erro mínimo pode se escrita da seguinte maneira

$$\mathbf{A} \begin{bmatrix} 1 \\ -\mathbf{w} \end{bmatrix} = \begin{bmatrix} J_{\min} \\ \mathbf{0} \end{bmatrix},$$

em que J_{\min} é o mínimo erro médio quadrático, \mathbf{w} é o filtro de Wiener, e \mathbf{A} é a matriz de correlação do vetor aumentado

$$\begin{bmatrix} d(n) \\ \mathbf{x}(n) \end{bmatrix},$$

em que $d(n)$ é o sinal desejado e $\mathbf{x}(n)$ é o sinal de entrada do filtro de Wiener.

Solução:

Inicialmente deve-se calcular a matriz de correlação do vetor aumentado

$$\mathbf{A} = \mathbb{E} \left\{ \begin{bmatrix} d(n) \\ x(n) \end{bmatrix} \begin{bmatrix} d(n)^T & x(n)^T \end{bmatrix} \right\}, \quad (98)$$

$$\mathbf{A} = \begin{bmatrix} \mathbb{E}\{d(n)d(n)^T\} & \mathbb{E}\{d(n)x(n)^T\} \\ \mathbb{E}\{x(n)d(n)^T\} & \mathbb{E}\{x(n)x(n)^T\} \end{bmatrix}. \quad (99)$$

Dando as devidas nomeações aos termos a expressão acima reduz-se a

$$\mathbf{A} = \begin{bmatrix} \sigma_d^2 & \mathbf{p}_{xd}^T \\ \mathbf{p}_{xd} & \mathbf{R}_X \end{bmatrix}. \quad (100)$$

Multiplicando-se a expressão acima pelo vetor $[1 - \mathbf{w}]^T$

$$\mathbf{A} \begin{bmatrix} 1 \\ -\mathbf{w} \end{bmatrix} = \begin{bmatrix} \sigma_d^2 & \mathbf{p}_{xd}^T \\ \mathbf{p}_{xd} & \mathbf{R}_X \end{bmatrix} \begin{bmatrix} 1 \\ -\mathbf{w} \end{bmatrix}, \quad (101)$$

$$\mathbf{A} \begin{bmatrix} 1 \\ -\mathbf{w} \end{bmatrix} = \begin{bmatrix} \sigma_d^2 - \mathbf{p}_{xd}^T \mathbf{w} \\ \mathbf{p}_{xd} - \mathbf{R}_X \mathbf{w} \end{bmatrix}. \quad (102)$$

Por fim, resta apenas aplica a equação ótima do filtro de wiener $\mathbf{w}_{\text{opt}} = \mathbf{R}_X^{-1} \mathbf{p}_{xd}$

$$\mathbf{A} \begin{bmatrix} 1 \\ -\mathbf{w} \end{bmatrix} = \begin{bmatrix} \sigma_d^2 - \mathbf{p}_{xd}^T \mathbf{R}_X^{-1} \mathbf{p}_{xd} \\ \mathbf{p}_{xd} - \mathbf{R}_X \mathbf{R}_X^{-1} \mathbf{p}_{xd} \end{bmatrix}, \quad (103)$$

$$\mathbf{A} \begin{bmatrix} 1 \\ -\mathbf{w} \end{bmatrix} = \begin{bmatrix} \sigma_d^2 - \mathbf{p}_{xd}^T \mathbf{R}_X^{-1} \mathbf{p}_{xd} \\ \mathbf{p}_{xd} - \mathbf{I}_X \mathbf{p}_{xd} \end{bmatrix}, \quad (104)$$

$$\mathbf{A} \begin{bmatrix} 1 \\ -\mathbf{w} \end{bmatrix} = \begin{bmatrix} \sigma_d^2 - \mathbf{p}_{xd}^T \mathbf{R}_X^{-1} \mathbf{p}_{xd} \\ 0 \end{bmatrix}, \quad (105)$$

$$\mathbf{A} \begin{bmatrix} 1 \\ -\mathbf{w} \end{bmatrix} = \begin{bmatrix} J_{\min} \\ 0 \end{bmatrix}. \quad (106)$$

3. Em várias aplicações práticas há uma necessidade de cancelar ruído que foi adicionado a um sinal. Por exemplo, se estamos usando o telefone celular dentro de um ruído e o ruído do carro ou rádio é adicionado à mensagem que estamos tentando transmitir. A Figura abaixo ilustra as situações de contaminação de ruído. Calcule o filtro de Wiener (filtro ótimo) de tal configuração em relação às estatísticas dos sinais envolvidos que você dispõe (conhece).

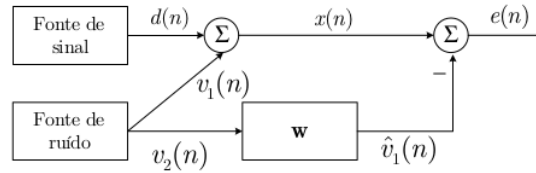


Figure 1: Esquema de cancelamento de ruído.

Figura 1: Estrutura de Equalização de Canal

Solução:

Inicialmente é necessário calcular a equação de erro do sistema aqui proposto

$$e(n) = x(n) - \hat{v}_1 = x(n) - \mathbf{w}^T v_2(n) \quad (107)$$

Em seguida faz-necessário calcular a função mean square error(MSE) que é facilmente fornecida pela manipulação algébrica abaixo

$$e^2(n) = [x(n) - \mathbf{w}^T v_2(n)][x(n) - \mathbf{w}^T v_2(n)]^T, \quad (108)$$

$$e^2(n) = x^2(n) - 2x(n)\mathbf{w}^T v_2(n) + \mathbf{w}^T v_2(n)v_2^T \mathbf{w}. \quad (109)$$

Sendo considerado que o filtro apresenta coeficientes constantes é possível aplicar o operador Valor Esperado de forma a obter a seguinte relação

$$\mathbb{E}\{e^2(n)\} = \mathbb{E}\{x^2(n)\} - 2\mathbf{w}^T \mathbb{E}\{x(n)v_2(n)\} + \mathbf{w}^T \mathbb{E}\{v_2(n)v_2(n)^T\} \mathbf{w}, \quad (110)$$

$$\mathbb{E}\{e^2(n)\} = \sigma_x^2 - 2\mathbf{w}^T \mathbf{p}_{xv_2} + \mathbf{w}^T \mathbf{R}_{v_2} \mathbf{w}. \quad (111)$$

Por fim, basta encontrar o \mathbf{w} que minimiza o MSE acima. Para chegar a esse fim, calcula-se o gradiente quanto ao \mathbf{w} igualando-se o resultado da operação a zero

$$\nabla_{\mathbf{w}} \mathbb{E}\{e^2(n)\} = -2\mathbf{p}_{xv_2} + 2\mathbf{R}_{v_2} \mathbf{w} = 0, \quad (112)$$

$$-\mathbf{p}_{xv_2} + \mathbf{R}_{v_2} \mathbf{w} = 0, \quad (113)$$

$$\mathbf{R}_{v_2} \mathbf{w} = \mathbf{p}_{xv_2}. \quad (114)$$

Utilizando a identidade matricial abaixo é possível resolver a equação acima para obter o seguinte resultado

$$\mathbf{R}_{v_2}^{-1} \mathbf{R}_{v_2} \mathbf{w} = \mathbf{R}_{v_2}^{-1} \mathbf{p}_{xv_2}, \quad (115)$$

$$\mathbf{I} \mathbf{w} = \mathbf{R}_{v_2}^{-1} \mathbf{p}_{xv_2}, \quad (116)$$

$$\mathbf{w} = \mathbf{R}_{v_2}^{-1} \mathbf{p}_{xv_2}. \quad (117)$$

Onde é possível reescrever o termo final como

$$\mathbf{w} = \mathbf{R}_{v_2}^{-1} (\mathbf{p}_d + \mathbf{p}_{v_1} + \mathbf{p}_{v_2}) \quad (118)$$

4. Seja um processo estocástico dado pela expressão abaixo onde $S(n)$ é um processo estocástico WSS dado e a é uma constante.

$$x(n) = s(n+a) + s(n-4a),$$

Deseja-se filtrar o processo de tal forma obter-se um processo $D(s) = s(n-a)$, o qual também sabe-se que é um processo WSS. Suponha que o sinal $d(n)$ possua média nula e variância unitária.

- (a) Calcule o filtro, com dois coeficientes, que fornece a solução ótima em relação ao erro médio quadrático.

Solução:

O filtro linear ótimo que minimiza o erro médio quadrático é descrito pela solução das equações de Wiener. Portanto, inicialmente definir a matriz de autocorrelação para o processo descrito por $x(n)$

$$\mathbf{R}_x = \begin{bmatrix} \mathbb{E}\{x(n)x^*(n)\} & \mathbb{E}\{x(n-1)x^*(n)\} \\ \mathbb{E}\{x(n)x^*(n-1)\} & \mathbb{E}\{x(n-1)x^*(n-1)\} \end{bmatrix}, \quad (119)$$

se considerarmos que o processo $S(n)$ é WSS com variância σ_s^2 podemos calcular as correlações como se segue

$$\begin{aligned} \mathbb{E}\{x(n)x^*(n)\} &= \mathbb{E}\{s(n-a)s^*(n-a) + s(n-a)s^*(n-4a) + s(n-4a)s^*(n-a) \\ &\quad + s(n-4a)s^*(n-4a)\} = \sigma_s^2 + 0 + 0 + \sigma_s^2 = 2\sigma_s^2, \\ \mathbb{E}\{x(n-1)x^*(n)\} &= \mathbb{E}\{s(n-1-a)s^*(n-a) + s(n-1-a)s^*(n-4a) + s(n-1-4a)s^*(n-a) \\ &\quad + s(n-1-4a)s^*(n-4a)\} = 0 + 0 + 0 + 0 = 0, \\ \mathbb{E}\{x(n)x^*(n-1)\} &= \mathbb{E}\{s(n-a)s^*(n-1-a) + s(n-a)s^*(n-1-4a) + s(n-4a)s^*(n-1-a) \\ &\quad + s(n-4a)s^*(n-1-4a)\} = 0 + 0 + 0 + 0 = 0, \\ \mathbb{E}\{x(n-1)x^*(n-1)\} &= \mathbb{E}\{s(n-1-a)s^*(n-1-a) + s(n-1-a)s^*(n-1-4a) \\ &\quad + s(n-1-4a)s^*(n-1-a) + s(n-1-4a)s^*(n-1-4a)\} = \sigma_s^2 + 0 + 0 + \sigma_s^2 = 2\sigma_s^2, \end{aligned}$$

obtendo assim

$$\mathbf{R}_x = \begin{bmatrix} 2\sigma_s^2 & 0 \\ 0 & 2\sigma_s^2 \end{bmatrix}, \quad (120)$$

Entretanto, considerando que o processo $D(n)$ têm média nula então temos na verdade um vetor de correlação cruzada nulo. Desse modo, o filtro linear ótimo para esse processo seria o próprio vetor nulo. Sendo assim

$$\mathbf{w}_{\text{opt}} = \mathbf{R}_x^{-1} \mathbf{p}_{xd} = \begin{bmatrix} 2\sigma_s^2 & 0 \\ 0 & 2\sigma_s^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad (121)$$

- (b) Calcule o preditor direto ótimo de passo unitário, com dois coeficientes, que fornece a solução ótima em relação ao erro médio quadrático.

Solução:

O filtro preditor direto é aquele que utiliza o conhecimento de amostras passadas para prever amostras futuras de algum processo. Desse modo, podemos definir o filtro direto de passo unitário como

$$\hat{x}(n) = \sum_{i=k}^{M+k-1} w_{f,i} x(n-i) = \sum_{i=1}^2 w_{f,i} x(n-i), \quad (122)$$

$$\hat{x}(n) = \mathbf{w}_f^T \mathbf{x}(n-1), \quad (123)$$

onde o erro quadrático médio pode ser dado por

$$\mathbb{E}\{e^2(n)\} = \mathbb{E}\{(x(n) - \hat{x}(n))^2\} = \mathbf{r}_x(0) - 2\mathbf{w}_f^T \mathbf{r}_{x,f} + \mathbf{w}_f^T \mathbf{R}_x \mathbf{w}_f, \quad (124)$$

e, portanto, de forma similar a solução de Wiener para o filtro linear ótimo temos

$$\mathbf{w}_{f,\text{opt}} = \mathbf{R}_x^{-1} \mathbf{r}_{x,f}. \quad (125)$$

Onde mais uma vez teremos a mesma matriz de autocorrelação definida no item anterior, mas agora teremos um vetor de correlação cruzada definido por

$$\mathbf{r}_{x,f} = \begin{bmatrix} r_x(1) \\ r_x(2) \end{bmatrix} = \begin{bmatrix} \mathbb{E}\{x(n)x(n-1)\} \\ \mathbb{E}\{x(n)x(n-2)\} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \quad (126)$$

Desse modo, temos resultado idêntico ao item anterior com a solução ótima sendo o próprio vetor nulo.

- (c) Compare as soluções dos dois.

Solução:

Os resultados são contraintuitivos e exibem limitações de ordem técnica que podemos encontrar durante a solução de certos problemas. Talvez uma solução para esse problema fosse de alguma forma obter as correlações reais entre amostras dos processos descritos por $X(n)$ e $D(n)$. Entretanto, não consigo visualizar com tanta clareza como tal procedimento poderia ser realizado.

5. Suponha que foram encontrados os seguintes coeficientes de autocorrelação: $r_x(0) = 1$ e $r_x(1) = 0$. Tais coeficientes foram obtidos de amostras corrompidas com ruído. Além disso, a variância do sinal desejado é $\sigma_d^2 = 24.40$ e o vetor de correlação cruzada é $\mathbf{p}_{xd} = [2 \ 4.5]^T$. Encontre:

- (a) O valor dos coeficientes do filtro de Wiener.

Solução:

A partir dos coeficientes fornecidos é possível escrever a matrix de correlação necessário para o filtro ótimo de wiener como uma matriz identidade de ordem 2

$$\mathbf{R}_X = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (127)$$

Ao utilizar a solução fechada do problema chega-se ao seguinte vetor resultado

$$\mathbf{w}_{\text{opt}} = \mathbf{R}_x^{-1} \mathbf{p}_{xd} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4.5 \end{bmatrix} = \begin{bmatrix} 2 \\ 4.5 \end{bmatrix}. \quad (128)$$

- (b) A superfície definida por $J(\mathbf{w})$. Faça um gráfico da mesma.

Solução:

Para obter a expressão que define a superfície basta desenvolver a expressão para o erro médio

$$\mathbf{J}(w) = \mathbb{E}\{e^2(n)\} = \sigma_d^2 - 2\mathbf{w}^T \mathbf{p}_{xd} + w^T \mathbf{R}_X \mathbf{w}. \quad (129)$$

Substituindo os valores encontrados anteriormente na expressão da superfície

$$\mathbf{J}(w_0, w_1) = 24.40 - 2 \begin{bmatrix} w_0 & w_1 \end{bmatrix} \begin{bmatrix} 2 \\ 4.5 \end{bmatrix} + \begin{bmatrix} w_0 & w_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}, \quad (130)$$

$$\mathbf{J}(w_0, w_1) = 24.40 - 4w_0 - 9w_1 + w_0^2 + w_1^2. \quad (131)$$

Utilizando um software gráfico é possível obter a Figura 2 onde é traçada a superfície de erro MSE expressa na Equação (129).

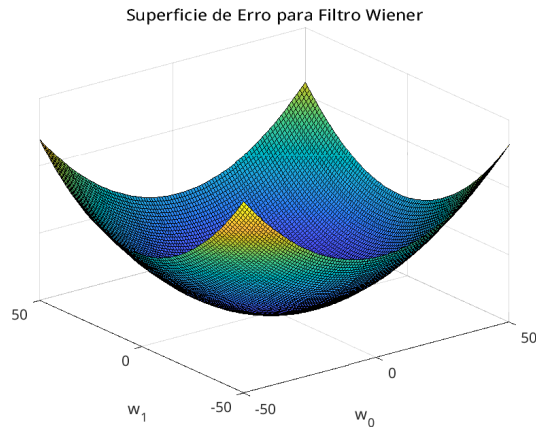


Figura 2: Superfície de erro MSE

Algoritmos Recursivos

1. Deseja-se minimizar a função objetivo $\mathbb{E}\{e^4(n)\}$ utilizando-se um algoritmo do gradiente estocástico do tipo LMS. O algoritmo resultando é chamado de algoritmo least mean fourth (LMF). Derive tal algoritmo. Derive também o filtro ótimo para tal critério e compare as soluções.

Solução:

Podemos inicialmente definir a função erro para esse filtro como

$$e(n) = d(n) - y(n), \quad (132)$$

$$e(n) = d(n) - \mathbf{w}^T(n)\mathbf{x}(n), \quad (133)$$

e para obtermos a nova expressão de recorrência precisamos primeiro obter o equacionamento para o vetor gradiente instantâneo de $\mathbb{E}\{e^4(n)\}$. Isso pode ser prontamente obtido por

$$\nabla_{\mathbf{w}}\mathbb{E}\{e^4(n)\} = \frac{\partial\mathbb{E}\{e^4(n)\}}{\partial\mathbf{w}} = \mathbb{E}\left\{\frac{\partial e^4(n)}{\partial\mathbf{w}}\right\} = \mathbb{E}\{4e^3(n)\frac{\partial e(n)}{\partial\mathbf{w}}\}, \quad (134)$$

$$\nabla_{\mathbf{w}}\mathbb{E}\{e^4(n)\} = \mathbb{E}\{4e^3(n)\frac{\partial(d(n) - \mathbf{w}^T(n)\mathbf{x}(n))}{\partial\mathbf{w}}\} = \mathbb{E}\{4e^3(n)(0 - \mathbf{x}(n))\}, \quad (135)$$

$$\nabla_{\mathbf{w}}\mathbb{E}\{e^4(n)\} = -4\mathbb{E}\{e^3(n)\mathbf{x}(n)\}. \quad (136)$$

Desse modo, para que a minimização de $\mathbb{E}\{e^4(n)\}$ seja atingida precisamos garantir apenas que o vetor $\mathbf{x}(n)$ tenha entradas ortogonais ao vetor erro $e(n)$. Assim, é necessário que a seguinte equação seja verdade

$$\mathbb{E}\{e^3(n)\mathbf{x}(n)\} = 0, \quad (137)$$

$$\mathbb{E}\{(d(n) - \mathbf{w}^T(n)\mathbf{x}(n))^3\mathbf{x}(n)\} = 0, \quad (138)$$

onde é possível demonstrar que existe convergência em média para essa expressão se definirmos o passo de aprendizado no seguinte intervalo¹

$$1 < \mu < \frac{1}{6\sigma_z^2\lambda_{\max}}, \quad (139)$$

onde σ_z^2 é a variância do ruído presente e λ_{\max} é o maior autovalor da matriz de autocorrelação \mathbf{R}_x . Por fim, a partir dessas observações podemos escrever a expressão de recorrência para o LMF utilizando a expressão padrão para o algoritmo do gradiente descendente²

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu\mathbf{g}_w(n), \quad (140)$$

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 4\mu e^3(n)\mathbf{x}(n), \quad (141)$$

onde sabemos que o erro é dado por $e(n) = d(n) - \mathbf{w}^T(n)\mathbf{x}(n)$.

¹A demonstração da propriedade foi extensivamente explicada no artigo *The Least Mean Square Fourth (LMF) Algorithm and Its Family* de 1984 por Eugene Walash e Bernard Widrow.

²Essa expressão é brevemente introduzida no livro texto da disciplina.

2. Considere o uso de um a sequência de ruído branco com média nula e variância σ^2 como entrada do algoritmo LMS. Avalie

(a) a condição para convergência do algoritmo em média.

Solução:

A condição de convergência está diretamente associada com o erro nos coeficientes do filtro adaptativo para cada iteração. Desse modo, podemos iniciar o estudo desse tópico com a seguinte expressão que relaciona o erro dos coeficientes do filtro de um iteração k para a solução ótima

$$\Delta \mathbf{w}(n) = \mathbf{w}(n) - \mathbf{w}_{\text{opt}}, \quad (142)$$

e assim podemos reescrever a função de recorrência do LMS do seguinte modo

$$\Delta \mathbf{w}(n+1) = \Delta \mathbf{w}(n) + 2\mu e(n) \mathbf{x}(n), \quad (143)$$

$$\Delta \mathbf{w}(n+1) = \Delta \mathbf{w}(n) + 2\mu \mathbf{x}(n) [\mathbf{x}^T(n) \mathbf{w}_{\text{opt}} + z(n) - \mathbf{x}^T(n) \mathbf{w}(n)], \quad (144)$$

$$\Delta \mathbf{w}(n+1) = \Delta \mathbf{w}(n) + 2\mu \mathbf{x}(n) [e_{\text{opt}}(n) - \mathbf{x}^T(n) \Delta \mathbf{w}(n)], \quad (145)$$

$$\Delta \mathbf{w}(n+1) = [\mathbf{I} - 2\mu \mathbf{x}(n) \mathbf{x}^T(n)] \Delta \mathbf{w}(n) + 2\mu e_{\text{opt}}(n) \mathbf{x}(n) \quad (146)$$

onde $z(n) \in \mathcal{N}(0, \sigma^2)$ e $e_{\text{opt}}(n) = z(n)$. Indicando que, para o caso ideal, teríamos que nos preocupar apenas com o erro introduzido pelas componentes ruidosas do sistema. Ademais, também utilizamos o fato de que $e(n) = e^T(n) = \mathbf{w}_{\text{opt}}^T \mathbf{x}(n) + z(n) - \mathbf{w}^T(n) \mathbf{x}(n) = \mathbf{x}^T(n) \mathbf{w}_{\text{opt}} + z(n) - \mathbf{x}^T(n) \mathbf{w}(n)$. Em sequência, podemos escrever o valor esperado para o erro como

$$\mathbb{E}\{\Delta \mathbf{w}(n+1)\} = \mathbb{E}\{[\mathbf{I} - 2\mu \mathbf{x}(n) \mathbf{x}^T(n)] \Delta \mathbf{w}(n) + 2\mu e_{\text{opt}}(n) \mathbf{x}(n)\}, \quad (147)$$

$$\mathbb{E}\{\Delta \mathbf{w}(n+1)\} = \mathbb{E}\{[\mathbf{I} - 2\mu \mathbf{x}(n) \mathbf{x}^T(n)] \Delta \mathbf{w}(n)\} + 2\mu \mathbb{E}\{e_{\text{opt}}(n) \mathbf{x}(n)\} \quad (148)$$

onde utilizamos a propriedade de linearidade do operador valor esperado. Ao assumirmos que $\mathbf{x}(n)$ é simultaneamente ortogonal aos elementos $e_{\text{opt}}(n)$ e $\Delta \mathbf{w}(n)$ podemos simplificar a expressão acima por

$$\mathbb{E}\{\Delta \mathbf{w}(n+1)\} = [\mathbf{I} - 2\mu \mathbb{E}\{\mathbf{x}(n) \mathbf{x}^T(n)\}] \mathbb{E}\{\Delta \mathbf{w}(n)\}, \quad (149)$$

$$\mathbb{E}\{\Delta \mathbf{w}(n+1)\} = (\mathbf{I} - 2\mu \mathbf{R}_x) \mathbb{E}\{\Delta \mathbf{w}(n)\}, \quad (150)$$

onde podemos utilizar a noção de que $\mathbb{E}\{\mathbf{I}\} = \mathbf{I}$ uma vez que temos um operador linear. Em continuidade, podemos supor que existe uma matriz \mathbf{Q} unitária que diagonaliza \mathbf{R}_x de tal modo que temos

$$\mathbb{E}\{\mathbf{Q}^T \Delta \mathbf{w}(n+1)\} = (\mathbf{I} - 2\mu \mathbf{Q}^T \mathbf{R}_x \mathbf{Q}) \mathbb{E}\{\Delta \mathbf{w}(n)\}, \quad (151)$$

$$\mathbb{E}\{\mathbf{Q}^T \Delta \mathbf{w}(n+1)\} = (\mathbf{I} - 2\mu \mathbf{Q}^T \mathbf{R}_x \mathbf{Q}) \mathbf{Q} \mathbf{Q}^T \mathbb{E}\{\Delta \mathbf{w}(n)\}, \quad (152)$$

$$\mathbb{E}\{\mathbf{Q}^T \Delta \mathbf{w}(n+1)\} = (\mathbf{I} - 2\mu \mathbf{Q}^T \mathbf{R}_x \mathbf{Q}) \mathbb{E}\{\mathbf{Q}^T \Delta \mathbf{w}(n)\}, \quad (153)$$

$$\mathbb{E}\{\Delta \mathbf{w}'(n+1)\} = (\mathbf{I} - 2\mu \mathbf{\Lambda}) \mathbb{E}\{\Delta \mathbf{w}'(n)\}. \quad (154)$$

Podemos, por fim, expandir os termos à esquerda e chegar na seguinte expressão para a análise do comportamento de convergência dos coeficientes de filtro

$$\mathbb{E}\{\Delta \mathbf{w}'(n+1)\} = (\mathbf{I} - 2\mu\mathbf{\Lambda})^{n+1} \mathbb{E}\{\Delta \mathbf{w}'(0)\}, \quad (155)$$

$$\mathbb{E}\{\Delta \mathbf{w}'(n+1)\} = \begin{bmatrix} (1-2\mu\lambda_1)^{n+1} & 0 & \cdots & 0 \\ 0 & (1-2\mu\lambda_2)^{n+1} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (1-2\mu\lambda_N)^{n+1} \end{bmatrix} \mathbb{E}\{\Delta \mathbf{w}'(0)\}, \quad (156)$$

onde temos que $\lambda_n \forall n \in \{1, \dots, N\}$ são os autovalores da matriz de autocorrelação. Portanto, vemos que para garantir a estabilidade da convergência é necessário apenas que o passo de aprendizado do algoritmo seja definido pela seguinte inequação

$$0 < \mu < \frac{1}{\lambda_{\max}}, \quad (157)$$

pois desse modo conseguimos garantir que os valores da diagonal irão tender a zero a medida que o número de iterações do algoritmo tende ao infinito. Ademais, é interessante ressaltar que a escolha do valor de μ deve também levar em consideração o espalhamento de energia da matriz de correlação. Dessa forma, se não há grande diferença entre os valores numéricos dos autovalores, então seria aconselhável escolher um passo de aprendizado muito menor do que aquele definido pelo limite superior da expressão obtida acima.

(b) o erro em excesso em média quadrática.

Solução:

O erro em excesso é normalmente ocasionado pelos termos ruidosos presentes no gradiente, impedindo que os coeficientes converjam de forma exata para a solução ótima. Podemos iniciar essa análise escrevendo a equação para o erro de estimação para um determinado instante n como se segue

$$e(n) = d(n) - \mathbf{w}_{\text{opt}}^T \mathbf{x}(n) - \Delta \mathbf{w}^T(n) \mathbf{x}(n), \quad (158)$$

$$e(n) = e_{\text{opt}}(n) - \Delta \mathbf{w}^T(n) \mathbf{x}(n), \quad (159)$$

onde o erro quadrático é expresso por

$$e^2(n) = e_{\text{opt}}^2(n) - 2e_{\text{opt}}(n)\Delta \mathbf{w}^T(n) \mathbf{x}(n) + \Delta \mathbf{w}^T(n) \mathbf{x}(n) \mathbf{x}^T(n) \Delta \mathbf{w}(n), \quad (160)$$

se definirmos o erro quadrático médio como $\xi(n) = \mathbb{E}\{e^2(n)\}$ e o erro ótimo, leia-se mínimo, como $\xi_{\min} = \mathbb{E}\{e_{\text{opt}}^2(n)\}$ então podemos escrever

$$\xi(n) = \xi_{\min} - 2\mathbb{E}\{e_{\text{opt}}(n)\Delta \mathbf{w}^T(n) \mathbf{x}(n)\} + \mathbb{E}\{\Delta \mathbf{w}^T(n) \mathbf{x}(n) \mathbf{x}^T(n) \Delta \mathbf{w}(n)\}, \quad (161)$$

considerando mais uma vez que \mathbf{x} é ortogonal ao $\Delta \mathbf{w}^T(n)$ e ao $e_{\text{opt}}(n)$, simultaneamente, podemos simplificar a expressão ainda mais

$$\xi(n) = \xi_{\min} - 2\mathbb{E}\{\Delta \mathbf{w}^T(n)\} \mathbb{E}\{e_{\text{opt}}(n) \mathbf{x}(n)\} + \mathbb{E}\{\Delta \mathbf{w}^T(n) \mathbf{x}(n) \mathbf{x}^T(n) \Delta \mathbf{w}(n)\}, \quad (162)$$

onde podemos utilizar a propriedade $\mathbb{E}\{\Delta \mathbf{w}^T(n) \mathbf{x}(n) \mathbf{x}^T(n) \Delta \mathbf{w}(n)\} = \text{tr}(\mathbb{E}\{\Delta \mathbf{w}^T(n) \mathbf{x}(n) \mathbf{x}^T(n) \Delta \mathbf{w}(n)\})$ uma vez que sabemos que o traço de um escalar é o próprio escalar. Sendo assim, utilizando a propriedade cíclica do operador traço escrevemos

$$\xi(n) = \xi_{\min} - 2\mathbb{E}\{\Delta\mathbf{w}^T(n)\}\mathbb{E}\{e_{\text{opt}}(n)\mathbf{x}(n)\} + \text{tr}(\mathbb{E}\{\Delta\mathbf{w}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\Delta\mathbf{w}(n)\}), \quad (163)$$

$$\xi(n) = \xi_{\min} - 2\mathbb{E}\{\Delta\mathbf{w}^T(n)\}\mathbb{E}\{e_{\text{opt}}(n)\mathbf{x}(n)\} + \mathbb{E}\{\text{tr}[\Delta\mathbf{w}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\Delta\mathbf{w}(n)]\}, \quad (164)$$

$$\xi(n) = \xi_{\min} - 2\mathbb{E}\{\Delta\mathbf{w}^T(n)\}\mathbb{E}\{e_{\text{opt}}(n)\mathbf{x}(n)\} + \mathbb{E}\{\text{tr}[\mathbf{x}(n)\mathbf{x}^T(n)\Delta\mathbf{w}(n)\Delta\mathbf{w}^T(n)]\}, \quad (165)$$

$$\xi(n) = \xi_{\min} + \mathbb{E}\{\text{tr}[\mathbf{x}(n)\mathbf{x}^T(n)\Delta\mathbf{w}(n)\Delta\mathbf{w}^T(n)]\}, \quad (166)$$

$$\xi(n) = \xi_{\min} + \text{tr}(\mathbb{E}\{\mathbf{x}(n)\mathbf{x}^T(n)\Delta\mathbf{w}(n)\Delta\mathbf{w}^T(n)\}), \quad (167)$$

$$\xi(n) = \xi_{\min} + \text{tr}(\mathbb{E}\{\mathbf{x}(n)\mathbf{x}^T(n)\}\mathbb{E}\{\Delta\mathbf{w}(n)\Delta\mathbf{w}^T(n)\}), \quad (168)$$

$$\xi(n) = \xi_{\min} + \text{tr}(\mathbf{R}_x\mathbb{E}\{\Delta\mathbf{w}(n)\Delta\mathbf{w}^T(n)\}), \quad (169)$$

$$\xi(n) = \xi_{\min} + \text{tr}(\mathbb{E}\{\mathbf{R}_x\Delta\mathbf{w}(n)\Delta\mathbf{w}^T(n)\}), \quad (170)$$

onde o processo justifica-se pela intercambialidade entre os operadores traço e valor esperador e pela ortogonalidade mencionada anteriormente. Em sequência podemos definir o erro em excesso por

$$\xi(n) - \xi_{\min} = \text{tr}(\mathbb{E}\{\mathbf{R}_x\Delta\mathbf{w}(n)\Delta\mathbf{w}^T(n)\}), \quad (171)$$

$$\Delta\xi(n) = \text{tr}(\mathbb{E}\{\mathbf{R}_x\Delta\mathbf{w}(n)\Delta\mathbf{w}^T(n)\}), \quad (172)$$

e adicionalmente, pela definição de uma transformação de similaridade onde garantimos que exista $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$ capaz de diagonalizar a matriz de autocorrelação, podemos reescrever a Equação (172) da seguinte forma

$$\Delta\xi(n) = \text{tr}(\mathbb{E}\{\mathbf{Q}\mathbf{Q}^T\mathbf{R}_x\mathbf{Q}\mathbf{Q}^T\Delta\mathbf{w}(n)\Delta\mathbf{w}^T(n)\mathbf{Q}\mathbf{Q}^T\}), \quad (173)$$

$$\Delta\xi(n) = \text{tr}(\mathbb{E}\{\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T\Delta\mathbf{w}(n)\Delta\mathbf{w}^T(n)\mathbf{Q}\mathbf{Q}^T\}), \quad (174)$$

$$\Delta\xi(n) = \text{tr}(\mathbb{E}\{\mathbf{Q}\mathbf{\Lambda}_{\text{cov}}[\Delta\mathbf{w}'(n)]\mathbf{Q}^T\}), \quad (175)$$

$$\Delta\xi(n) = \text{tr}(\mathbb{E}\{\mathbf{\Lambda}_{\text{cov}}[\Delta\mathbf{w}'(n)]\mathbf{Q}^T\mathbf{Q}\}), \quad (176)$$

$$\Delta\xi(n) = \text{tr}(\mathbb{E}\{\mathbf{\Lambda}_{\text{cov}}[\Delta\mathbf{w}'(n)]\}), \quad (177)$$

onde é sabido que $\text{cov}[\Delta\mathbf{w}(n)] = \mathbb{E}\{(\mathbf{w}(n) - \mathbf{w}_{\text{opt}})(\mathbf{w}(n) - \mathbf{w}_{\text{opt}})^T\}$ e $\Delta\mathbf{w}'(n) = \mathbf{Q}^T\Delta\mathbf{w}(n)\Delta\mathbf{w}^T(n)\mathbf{Q}$. Enfim, a partir das orientações presentes no livro texto da disciplina a respeito da matriz de covariância do vetor de erros é possível ainda escrever a Equação (177) como

$$\Delta\xi(n) = \sum_{i=1}^N \lambda_i v'_i(n) = \lambda^T \mathbf{v}'(n), \quad (178)$$

onde λ é um vetor contendo todos os autovalores da matriz de autocorrelação e $\mathbf{v}(n)$ é um vetor que contém os elementos da diagonal de $\text{cov}[\Delta\mathbf{w}'(n)]$. Em suma, podemos expressar o i th elemento do vetor diagonal $\mathbf{v}'(n+1)$, seguindo orientações disponíveis no livro texto da disciplina, com o seguinte equacionamento

$$v'_i(n+1) = (1 - 4\mu\lambda_i + 8\mu^2\lambda_i^2)v'_i(n) + 4\mu^2\lambda_i \sum_{j=0}^N \lambda_j v'_j(n) + 4\mu^2\sigma_z^2\lambda_i. \quad (179)$$

Entretanto, se considerarmos que $v'_i(n+1) \approx v'_i(n)$ quando $n \rightarrow \infty$ podemos simplificar a expressão, além de que também podemos realizar uma operação de soma total dos parâmetros para obter o erro em excesso total

$$\begin{aligned}
\sum_{i=1}^N v'_i(n) &= \sum_{i=1}^N (1 - 4\mu\lambda_i + 8\mu^2\lambda_i^2)v'_i(n) + 4\mu^2 \sum_{i=1}^N \lambda_i \sum_{j=0}^N \lambda_j v'_j(n) + 4\mu^2\sigma_z^2 \sum_{i=1}^N \lambda_i, \\
\sum_{i=1}^N v'_i(n) &= \sum_{i=1}^N v'_i(n) - 4\mu \sum_{i=1}^N \lambda_i v'_i(n) + 8\mu^2 \sum_{i=1}^N \lambda_i^2 v'_i(n) + 4\mu^2 \sum_{i=1}^N \lambda_i \sum_{j=0}^N \lambda_j v'_j(n) + 4\mu^2\sigma_z^2 \sum_{i=1}^N \lambda_i, \\
\sum_{i=1}^N v'_i(n) - \sum_{i=1}^N v'_i(n) + 4\mu \sum_{i=1}^N \lambda_i v'_i(n) - 4\mu^2 \sum_{i=1}^N \lambda_i \sum_{j=0}^N \lambda_j v'_j(n) &= 4\mu^2\sigma_z^2 \sum_{i=1}^N \lambda_i + 8\mu^2 \sum_{i=1}^N \lambda_i^2 v'_i(n), \\
4\mu \sum_{j=1}^N \lambda_j v'_j(n) (1 - \mu \sum_{i=1}^N \lambda_i) &= 4\mu(\mu\sigma_z^2 \sum_{i=1}^N \lambda_i + 2\mu \sum_{i=1}^N \lambda_i^2 v'_i(n)), \\
\sum_{j=1}^N \lambda_j v'_j(n) &= \frac{\mu\sigma_z^2 \sum_{i=1}^N \lambda_i + 2\mu \sum_{i=1}^N \lambda_i^2 v'_i(n)}{1 - \mu \sum_{i=1}^N \lambda_i}, \\
\sum_{j=1}^N \lambda_j v'_j(n) &\approx \frac{\mu\sigma_z^2 \sum_{i=1}^N \lambda_i}{1 - \mu \sum_{i=1}^N \lambda_i}, \\
\sum_{j=1}^N \lambda_j v'_j(n) &= \frac{\mu\sigma_z^2 \text{tr}(\mathbf{R}_x)}{1 - \mu \text{tr}(\mathbf{R}_x)}, \tag{180}
\end{aligned}$$

onde foi considerado que o termo $2\mu \sum_{i=1}^N \lambda_i^2 v'_i(n)$ apresenta contribuição insignificante para o valor absoluto do numerador. Entretanto, é mencionado no livro texto que tal aproximação é de prova complexa, mas que normalmente se verifica verdadeira para pequenos valores do passo de aprendizado μ . Portanto, o erro em excesso pode ser prontamente descrito pela expressão que se segue

$$\xi_{\text{excesso}} = \lim_{n \rightarrow \infty} \Delta\xi(n) \approx \frac{\mu\sigma_z^2 \text{tr}(\mathbf{R}_x)}{1 - \mu \text{tr}(\mathbf{R}_x)}. \tag{181}$$

Vale ainda expor que podemos considerar $1 - \mu \text{tr}(\mathbf{R}_x) \approx 1$ para valores muito pequenos de μ , obtendo assim uma versão aproximada para o erro em excesso dada por

$$\xi_{\text{excesso}} = \lim_{n \rightarrow \infty} \Delta\xi(n) \approx \mu\sigma_z^2 \text{tr}(\mathbf{R}_x). \tag{182}$$

3. Avalie a questão anterior para o caso do algoritmo LMS-Normalizado. Compare os dois casos.

Solução:

4. Considere um sinal branco gaussiano de variância unitária transmitido por um canal de comunicação de função de transferência $H(z) = 1 + 1.6z^{-1}$. Para compensar este canal utiliza-se um equalizador dado por $W(z) = w_0 + w_1z^{-1}$.

- (a) Forneça o equalizador ótimo segundo o critério de Wiener. Esboce a posição dos zeros do canal e do equalizador no plano Z.

Solução:

Considerando um sinal gaussiano branco $x(n)$ a saída do canal pode ser prontamente obtida por

$$y(n) = x(n) + 1.6x(n-1), \quad (183)$$

e a matriz de autocorrelação será então dada por

$$\mathbf{R}_y = \begin{bmatrix} \mathbb{E}\{y(n)y^H(n)\} & \mathbb{E}\{y(n)y^H(n-1)\} \\ \mathbb{E}\{y(n-1)y^H(n)\} & \mathbb{E}\{y(n-1)y^H(n-1)\} \end{bmatrix}, \quad (184)$$

onde podemos calcular os valores teóricos para as correlações da seguinte forma se assumirmos que existe independência entre amostras distintas e que o sinal é média nula

$$\begin{aligned} \mathbb{E}\{y(n)y^H(n)\} &= \mathbb{E}\{\mathbf{x}^2(n) + 1.6\mathbf{x}(n)\mathbf{x}^H(n-1) + 1.6\mathbf{x}(n-1)\mathbf{x}^H(n) \\ &\quad + 2.56\mathbf{x}^2(n-1)\} = 3.56, \\ \mathbb{E}\{y(n)y^H(n-1)\} &= \mathbb{E}\{\mathbf{x}(n)\mathbf{x}^H(n-1) + 1.6\mathbf{x}(n)\mathbf{x}^H(n-2) + 1.6\mathbf{x}(n-1)\mathbf{x}^H(n-1) \\ &\quad + 2.56\mathbf{x}(n-1)\mathbf{x}^H(n-2)\} = 1.60, \\ \mathbb{E}\{y(n-1)y^H(n)\} &= \mathbb{E}\{\mathbf{x}(n-1)\mathbf{x}^H(n) + 1.6\mathbf{x}(n-1)\mathbf{x}^H(n-1) + 1.6\mathbf{x}(n-2)\mathbf{x}^H(n) \\ &\quad + 2.56\mathbf{x}(n-2)\mathbf{x}^H(n-1)\} = 1.60, \\ \mathbb{E}\{y(n-1)y^H(n-1)\} &= \mathbb{E}\{\mathbf{x}^2(n-1) + 1.6\mathbf{x}(n-1)\mathbf{x}^H(n-2) + 1.6\mathbf{x}(n-2)\mathbf{x}^H(n-1) \\ &\quad + 2.56\mathbf{x}(n-2)^2\} = 3.56, \end{aligned}$$

podendo assim descrever a matriz de autocorrelação teórica e sua inversa como

$$\mathbf{R}_y = \begin{bmatrix} 3.56 & 1.60 \\ 1.60 & 3.56 \end{bmatrix}, \quad (185)$$

$$\mathbf{R}_y^{-1} = \frac{1}{3.56^2 - 1.6^2} \begin{bmatrix} 3.56 & -1.60 \\ -1.60 & 3.56 \end{bmatrix} = \begin{bmatrix} 0.35 & -0.16 \\ -0.16 & 0.35 \end{bmatrix}. \quad (186)$$

Já o vetor de correlação cruzada teórico pode ser descrito por

$$\mathbf{p}_{yd} = \begin{bmatrix} \mathbb{E}\{y(n)d(n)\} \\ \mathbb{E}\{y(n-1)d(n)\} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad (187)$$

pois queremos que o sinal de saída tenha a maior correlação possível com o sinal desejado de um mesmo instante mas continue sendo independente de um sinal de um instante temporal diferente. Desse modo, podemos obter o equalizador ótimo segundo o critério de Wiener como

$$\mathbf{w}_{\text{opt}} = \mathbf{R}_y^{-1} \mathbf{p}_{yd} = \begin{bmatrix} 0.35 \\ -0.16 \end{bmatrix}. \quad (188)$$

Por fim, abaixo segue o traçado para os zeros das funções de transferência tanto do canal quanto do filtro ótimo em azul e em vermelho, respectivamente.

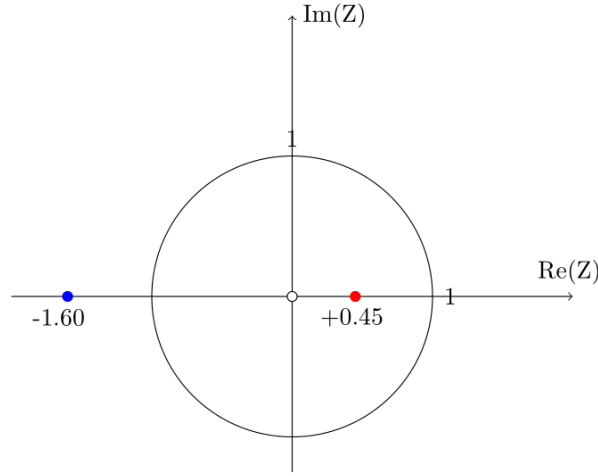


Figura 3: Zeros do canal e do equalizador no plano-z.

- (b) Obtenha o filtro de erro de predição direta de passo unitário, correspondente ao sinal à saída do canal. Calcule os zeros deste filtro e compare com os do equalizador.

Solução:

De forma semelhante ao abordado anteriormente temos que o filtro de predição direta de passo unitário é dado por

$$\hat{x}(n) = \sum_{i=k}^{M+k-1} w_{f,i} x(n-i) = \sum_{i=1}^2 w_{f,i} x(n-i) = \mathbf{w}_f^T \mathbf{x}(n-1), \quad (189)$$

onde o erro quadrático médio é dado por

$$\mathbb{E}\{e^2(n)\} = \mathbb{E}\{(x(n) - \hat{x}(n))^2\} = \mathbf{r}_x(0) - 2\mathbf{w}_f^T \mathbf{r}_{x,f} + \mathbf{w}_f^T \mathbf{R}_x \mathbf{w}_f, \quad (190)$$

onde a solução ótima será dada por

$$\mathbf{w}_{f,\text{opt}} = \mathbf{R}_x^{-1} \mathbf{r}_{x,f}. \quad (191)$$

Onde teremos que a mesma matriz de autocorrelação e o vetor de correlação cruzada serão definidos por

$$\mathbf{R}_y = \begin{bmatrix} 3.56 & 1.60 \\ 1.60 & 3.56 \end{bmatrix}, \quad (192)$$

$$\mathbf{r}_{y,f} = \begin{bmatrix} r_y(1) \\ r_y(2) \end{bmatrix} = \begin{bmatrix} \mathbb{E}\{y(n)y(n-1)\} \\ \mathbb{E}\{y(n)y(n-2)\} \end{bmatrix} = \begin{bmatrix} 1.60 \\ 0 \end{bmatrix}. \quad (193)$$

Por essa razão temos

$$\mathbf{w}_{f,\text{opt}} = \begin{bmatrix} 0.35 & -0.16 \\ -0.16 & 0.35 \end{bmatrix} \begin{bmatrix} 1.60 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.56 \\ -0.26 \end{bmatrix}. \quad (194)$$

Em seguida, podemos obter os zeros do filtro como

$$W(z) = 0.56 - 0.26z^{-1}, \quad (195)$$

$$0 = 0.56 - 0.26z^{-1}, \quad (196)$$

$$z = 0.45, \quad (197)$$

que é o mesmo zero do equalizador definido anteriormente.

- (c) Obtenha as trajetórias sobre as curvas de nível, tendo condições iniciais nulas para os coeficientes do equalizador, para os seguintes algoritmos: (a) Algoritmo de Newton, (b) Gradiente Determinístico, (c) Least Means Square e (d) Least Means Square Normalizado

Solução:

Antes de tudo é necessário definir a superfície de erro que servirá como referência para traçar as curvas de nível. Desse modo, podemos prontamente definir a superfície MSE como

$$\mathbf{J}(w) = \mathbb{E}\{e^2(n)\}, \quad (198)$$

$$\mathbf{J}(w) = \sigma_d^2 - 2\mathbf{w}^T \mathbf{p}_{xd} + w^T \mathbf{R}_X \mathbf{w}. \quad (199)$$

Desse modo, nas Figuras 5, 6, 7 e 8 temos as curvas de convergência do algoritmo sobre a superfície MSE descrita na Equação (199).

- (d) Obtenha também a evolução do erro quadrático médio para cada um dos algoritmos anteriores.

Solução:

A evolução do erro quadrático médio está disponível nas Figuras 9, 10, 11 e 12 para cada um dos algoritmos requeridos.

- (e) Qual o número de condicionamento para o problema em questão?

Solução:

o número de condicionamento pode ser prontamente obtido pela expressão

$$\mathbb{C}(\mathbf{R}_x) = \frac{\lambda_{\max}}{\lambda_{\min}}, \quad (200)$$

onde λ_{\max} e λ_{\min} são os autovalores máximo e mínimo da matriz de autocorrelação, respectivamente. Por meio de um software matemático foi possível obter os seguinte autovalores para a matriz de autocorrelação teórica

$$\mathbb{C}(\mathbf{R}_x) = \frac{5.16}{1.96} = 2.63, \quad (201)$$

onde talvez seja importante ressaltar que também poderíamos ter obtido os autovalores resolvendo a equação do polinômio característico da matriz de autocorrelação que é dada por

$$\lambda^2 - 7.12\lambda + 10.11 = 0. \quad (202)$$

- (f) Qual deveria ser o canal para que o número de condicionamento fosse menor/maior que 5? Comente os resultados.

Solução:

Inicialmente podemos escrever a matriz de correlação contabilizando a contribuição dos coeficientes do canal para os elementos individuais

$$\mathbf{R}_y = \begin{bmatrix} a_0 + a_1^2 & a_1 \\ a_1 & a_0 + a_1^2 \end{bmatrix}, \quad (203)$$

onde a função de transferência do canal seria dada por $H(z) = a_0 + a_1 z^{-1}$. A partir dessa matriz de autocorrelação genérica podemos então definir o seguinte polinômio característico

$$(\lambda - a_0 + a_1^2)^2 - a_1^2 = 0, \quad (204)$$

$$\lambda^2 - \underbrace{2(a_0 + a_1^2)}_b \lambda + \underbrace{(a_0 + a_1^2)^2 - a_1^2}_c = 0, \quad (205)$$

$$\lambda^2 + b\lambda + c = 0, \quad (206)$$

onde sabemos que a solução é facilmente obtida pela fórmula de Bháskara. A partir disso podemos definir o número de condicionamento como

$$\mathbb{C}(\mathbf{R}_x) = \frac{\lambda_{\max}}{\lambda_{\min}}, \quad (207)$$

$$\mathbb{C}(\mathbf{R}_x) = \frac{-b + \sqrt{b^2 - 4c}}{-b - \sqrt{b^2 - 4c}}, \quad (208)$$

$$\mathbb{C}(\mathbf{R}_x) = \frac{2(a_0 + a_1^2) + \sqrt{4(a_0 + a_1^2)^2 - 4(a_0 + a_1^2)^2 + 4a_1^2}}{2(a_0 + a_1^2) - \sqrt{4(a_0 + a_1^2)^2 - 4(a_0 + a_1^2)^2 + 4a_1^2}}, \quad (209)$$

$$\mathbb{C}(\mathbf{R}_x) = \frac{2(a_0 + a_1^2) + 2a_1}{2(a_0 + a_1^2) - 2a_1}, \quad (210)$$

$$\mathbb{C}(\mathbf{R}_x) = \frac{a_0 + a_1^2 + a_1}{a_0 + a_1^2 - a_1}, \quad (211)$$

assim temos agora uma fórmula para o número de condicionamento da matriz de autocorrelação com base nos coeficientes de canal. A partir disso basta que as seguintes inequações sejam atendidas para que obtenhamos um número de condicionamento maior ou menor do que o requerido

$$a_0 + a_1^2 + a_1 \geq 5(a_0 + a_1^2 - a_1), \quad (212)$$

$$a_0 + a_1^2 + a_1 \leq 5(a_0 + a_1^2 - a_1). \quad (213)$$

5. Utilize o algoritmo LMS para identificar um sistema com a função de transferência dada abaixo.

$$H(z) = \frac{1 - z^{-12}}{1 - z^{-1}} \quad (214)$$

(a) Calcule o limite superior para μ (ou seja μ_{\max}) para garantir a estabilidade do algoritmo.

Solução:

Para garantirmos a estabilidade do algoritmo precisamos apenas obter o valor numérico do maior autovalor definido pela matriz de autocorrelação do problema. Desse modo, podemos rearranjar a função de transferência acima para o seguinte formato

$$H(z) = \frac{1}{1 - z^{-1}} - \frac{z^{-12}}{1 - z^{-1}}, \quad (215)$$

$$H(z) = \frac{z^{-1}}{z^{-1}} \frac{z}{z - 1} - \frac{z^{-1}}{z^{-1}} \frac{z^{-11}}{z - 1}, \quad (216)$$

$$H(z) = \frac{z}{z - 1} - \frac{z^{-11}}{z - 1}, \quad (217)$$

e assim podemos utilizar uma tabela de transformadas z para chegar na seguinte equação para a saída do canal

$$y(n) = x(n) - x(n-12), \quad (218)$$

Assim, temos uma vez mais um sinal gaussiano branco $x(n)$ e uma saída $y(n)$ cuja matriz de autocorrelação é dada por

$$\mathbf{R}_y = \begin{bmatrix} \mathbb{E}\{y(n)y^*(n)\} & \mathbb{E}\{y(n)y^*(n-1)\} \\ \mathbb{E}\{y(n-1)y^*(n)\} & \mathbb{E}\{y(n-1)y^*(n-1)\} \end{bmatrix}, \quad (219)$$

onde temos

$$\begin{aligned} \mathbb{E}\{y(n)y^*(n)\} &= \mathbb{E}\{x(n)x^*(n) - x(n)x^*(n-12) - x(n-12)x^*(n) \\ &\quad + x(n-12)x^*(n-12)\} = 1 - 0 - 0 + 1 = 2, \\ \mathbb{E}\{y(n)y^*(n-1)\} &= \mathbb{E}\{x(n)x^*(n-1) - x(n)x^*(n-13) - x(n-12)x^*(n-1) \\ &\quad + x(n-12)x^*(n-13)\} = 0 - 0 - 0 + 0 = 0, \\ \mathbb{E}\{y(n-1)y^*(n)\} &= \mathbb{E}\{x(n-1)x^*(n) - x(n-1)x^*(n-12) - x(n-13)x^*(n) \\ &\quad + x(n-13)x^*(n-12)\} = 0 - 0 - 0 + 0 = 0, \\ \mathbb{E}\{y(n-1)y^*(n-1)\} &= \mathbb{E}\{x(n-1)x^*(n-1) - x(n-1)x^*(n-13) - x(n-13)x^*(n-1) \\ &\quad + x(n-13)x^*(n-13)\} = 1 - 0 - 0 + 1 = 2, \end{aligned}$$

resultando na seguinte matriz de autocorrelação

$$\mathbf{R}_y = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad (220)$$

onde sabemos que os autovalores de \mathbf{R}_y são $\lambda_1 = 2$ e $\lambda_2 = 2$. Portanto, o intervalo de convergência servirá

$$0 < \mu < \frac{1}{\lambda_{\max}} = 0.5, \quad (221)$$

e assim $\mu_{\max} = \frac{1}{\lambda_{\max}} = 0.5$.

- (b) Execute o algoritmo para $\frac{\mu_{\max}}{2}$, $\frac{\mu_{\max}}{10}$ e $\frac{\mu_{\max}}{50}$. Comente sobre o comportamento da convergência de cada caso.

Solução:

Nas Figuras 13, 14 e 15 podemos verificar o comportamento da convergência para esse problema. Podemos notar que não existiu convergência para os dois primeiros caso enquanto o terceiro podemos verificar a estabilização do filtro. Assim como explicado anteriormente na derivação dos limites de convergência do LMS, quando temos uma matriz de autocorrelação com uma distribuição de autovalores desconcentrada demais, significando que o autovalor máximo não possui a quase totalidade da energia, então o limiar de convergência torna-se extremamente impreciso. Desse modo, deve-se considerar um valor consideravelmente menor do que o limite superior máximo. Assim, podemos explicar por qual motivo houve a convergência apenas para o caso no qual o passo de aprendizado se distanciou consideravelmente do limite superior estabelecido.

- (c) Meça o desajuste (misadjustment) em cada exemplo e comparar com os resultados obtidos pela solução teórica (Eq. (3.50) do livro texto)

Solução:

O desajuste pode ser aproximado por

$$M = \frac{\xi_{\text{excesso}}}{\xi_{\text{min}}} \approx \frac{\mu \text{tr}(\mathbf{R}_x)}{1 - \mu \text{tr}(\mathbf{R}_x)}. \quad (222)$$

A partir disso a seguinte tabela é disponibilizada

	Empírico	Téorico
$\frac{\mu_{\max}}{2}$	-65.65	∞
$\frac{\mu_{\max}}{10}$	0.249	0.25
$\frac{\mu_{\max}}{50}$	0.044	0.04

Os resultados empíricos foram obtidos por software matemático e estão disponíveis juntamente com esse relatório.

- (d) Mostre o gráfico da resposta em frequência do filtro FIR em qualquer uma das iterações após a convergência ser obtida e compare com o sistema desconhecido.

Solução:

A resposta em frequência do filtro está disponível na Figura 16.

6. Seja o canal de comunicações dado por

$$H(z) = 0.5 + 1.2z^{-1} + 1.5z^{-2} + z^{-3}, \quad (223)$$

e deseja-se projetar um equalizador para o mesmo. A estrutura do equalizador é mostrada na Figura abaixo. Os símbolos $s(n)$ são transmitidos através de um canal e corrompidos por ruído aditivo gaussiano branco complexo $v(n)$. O sinal recebido $x(n)$ é processado pelo equalizador FIR para gerar estimativas $\tilde{s}(n - \delta)$, as quais são passadas por um dispositivo decisor gerando símbolos $\hat{s}(n - \delta)$. O equalizador possui dois modos de operação: um modo de treinamento durante o qual uma versão atrasada e replicada da sequência de entrada é usada como o sinal de referência (desejado) e um modo dirigido por decisão no qual a saída do dispositivo de decisão substitui a sequência de referência. O sinal de entrada $s(n)$ é escolhido de uma constelação QAM (por exemplo, 4-QAM, 16-QAM, 64-QAM ou 256-QAM).

- (a) Faça um programa que treine o filtro adaptativo com 500 símbolos de uma constelação 4-QAM, seguindo de uma operação dirigida por decisão de 5000 símbolos de uma constelação 16-QAM. Escolha a variância do ruído σ_v^2 de maneira que ela promova uma relação sinal ruído de 30 db na entrada do equalizador. Note que os símbolos escolhidos não têm variância unitária. Por esta razão, a a variância do ruído necessita ser ajustada adequadamente para cada uma das diferentes modulações (constelações) QAM para fornecer o nível de SNR desejado. Escolha $\delta = 15$ e o comprimento do equalizador $M = 15$. Mostre os gráficos da evolução temporal de $s(n)$, $x(n)$ e $\tilde{s}(n - \delta)$. Use o LMS-normalizado com um fator de passo de $\mu = 0.4$.

Solução:

Os resultados estão nas Figuras 17 e 18.

- (b) Para os mesmos parâmetros do item (a), plote e compare os gráficos de evolução que seriam resultante se o equalizador fosse treinado com 150, 300 e 500 iterações. Use o LMS com um $\mu = 0.001$.

Solução:

Os resultados estão na Figura 19.

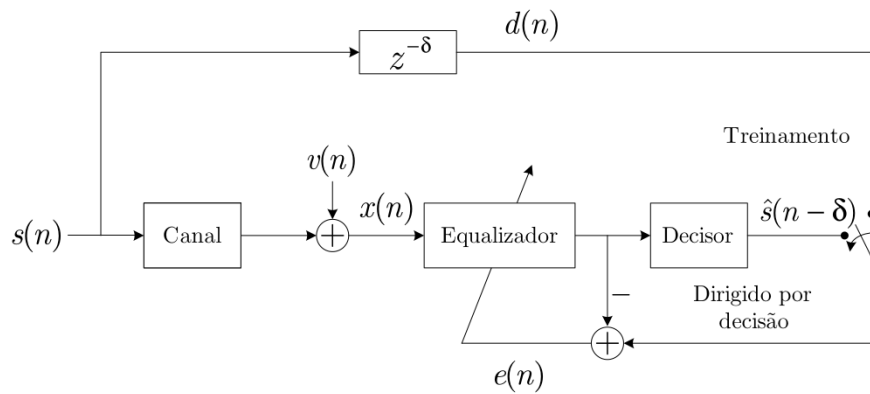


Figura 4: Equalizador Linear

- (c) Assuma agora que os dados transmitidos foram gerados de uma constelação 256-QAM ao invés de 16-QAM. Plote os gráficos da evolução do sinal na saída do equalizador quando treinado usando o LMS-normalizado e 500 símbolos de treinamento.

Solução:

Os resultados estão nas Figuras 20 e 21.

- (d) Gerar as curvas de taxa de erro de símbolo (SER, do inglês Symbol Error Rate) versus SNR na entrada do equalizador para símbolos de constelações 4, 16, 64 e 256-QAM. Faça SNR variar de 5dB a 30dB.

Solução:

Os resultados estão na Figura 22.

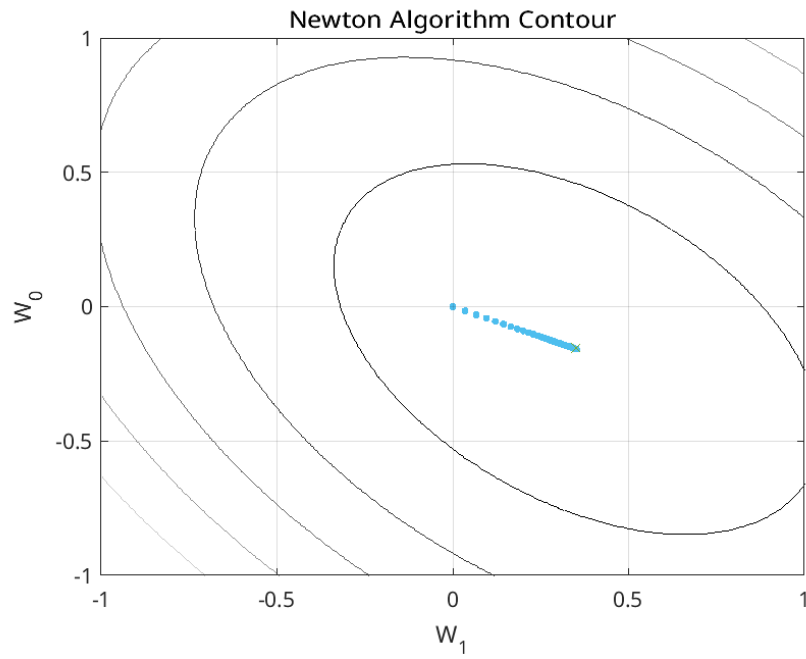


Figura 5: Caminho de convergência na superfície MSE para o Método de Newton. Amostras = 5000, $M = 2$, $\mu = 10^{-3}$

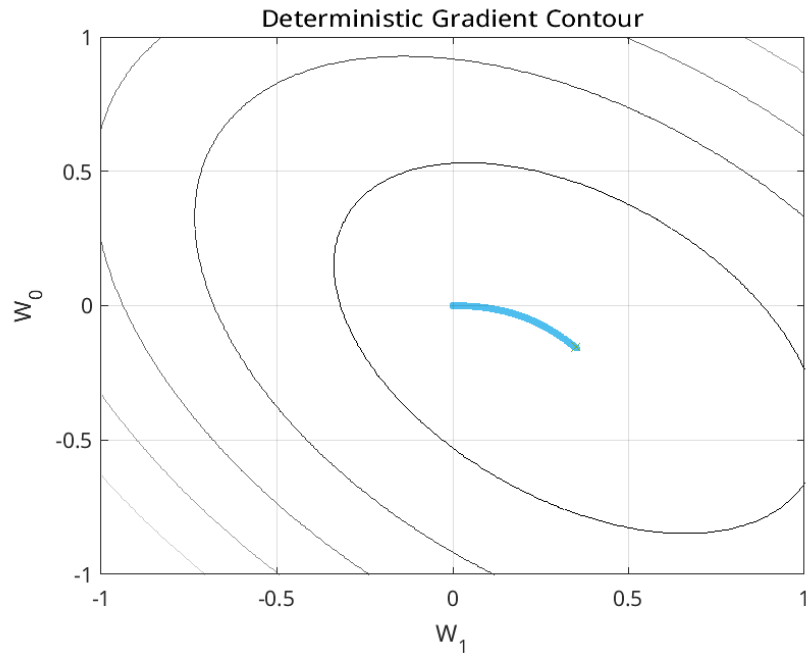


Figura 6: Caminho de convergência na superfície MSE para o Gradiente Determinístico. Amostras = 5000, $M = 2$, $\mu = 10^{-3}$

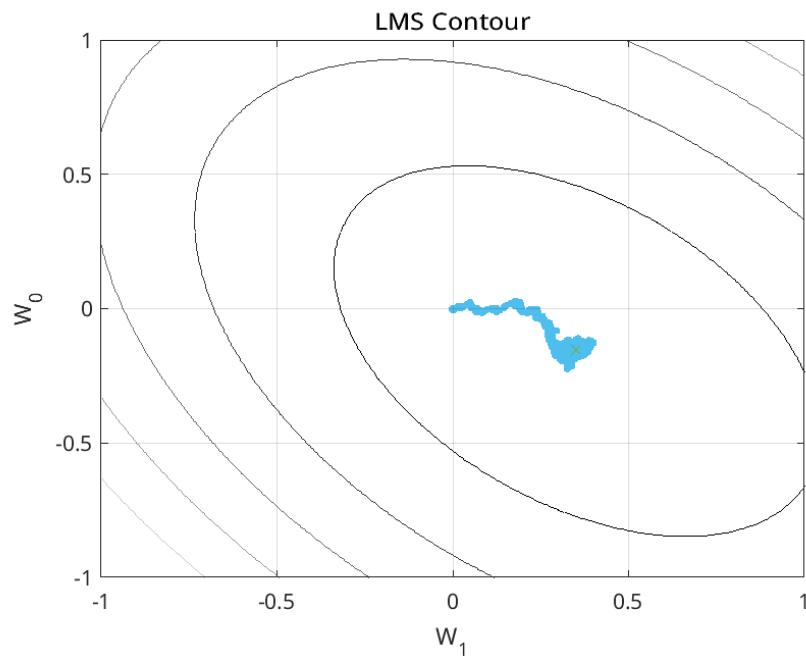


Figura 7: Caminho de convergência na superfície MSE para o LMS. Amostras = 5000, $M = 2$, $\mu = 10^{-3}$

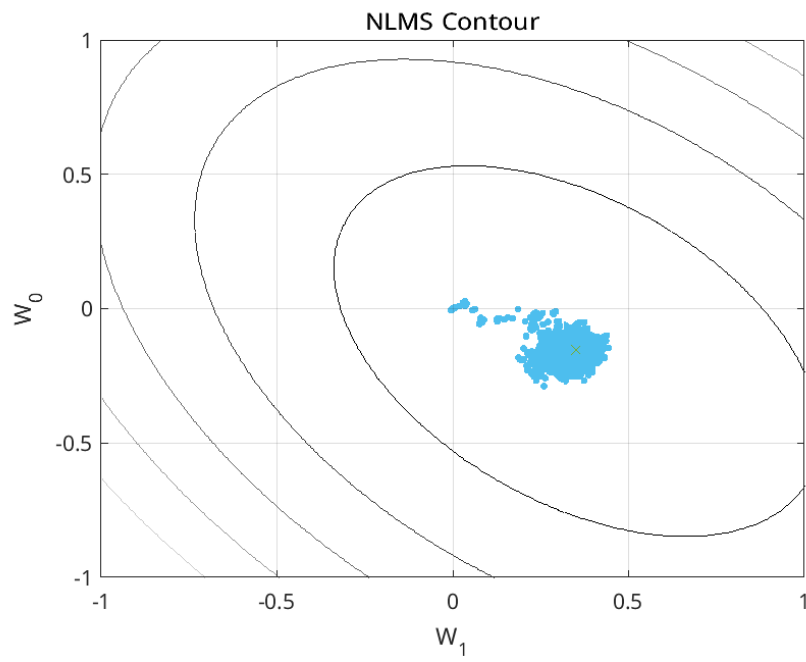


Figura 8: Caminho de convergência na superfície MSE para o NLMS. Amostras = 5000, $M = 2$, $\mu = 10^{-3}$

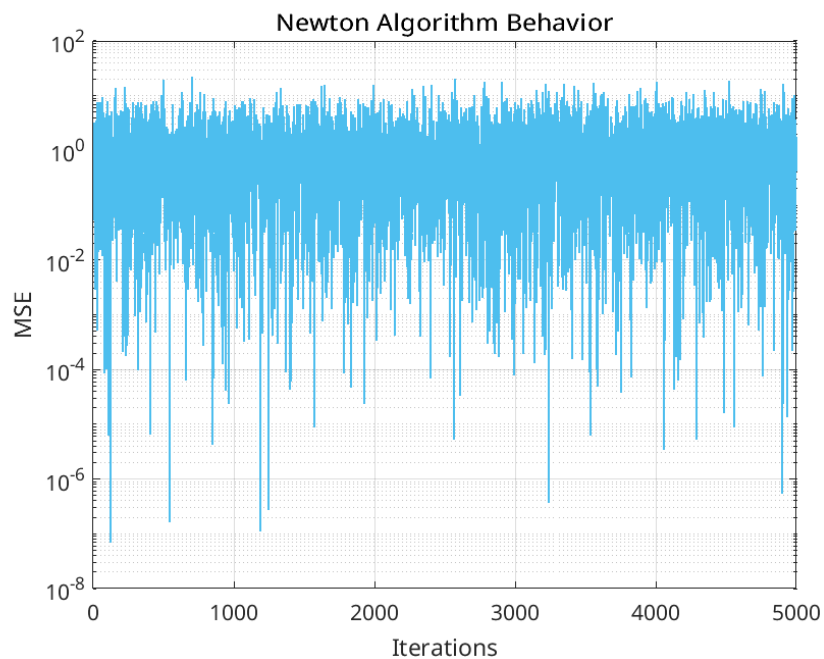


Figura 9: Caminho de convergência na superfície MSE para o Método de Newton. Amostras = 5000, $M = 2$, $\mu = 10^{-3}$

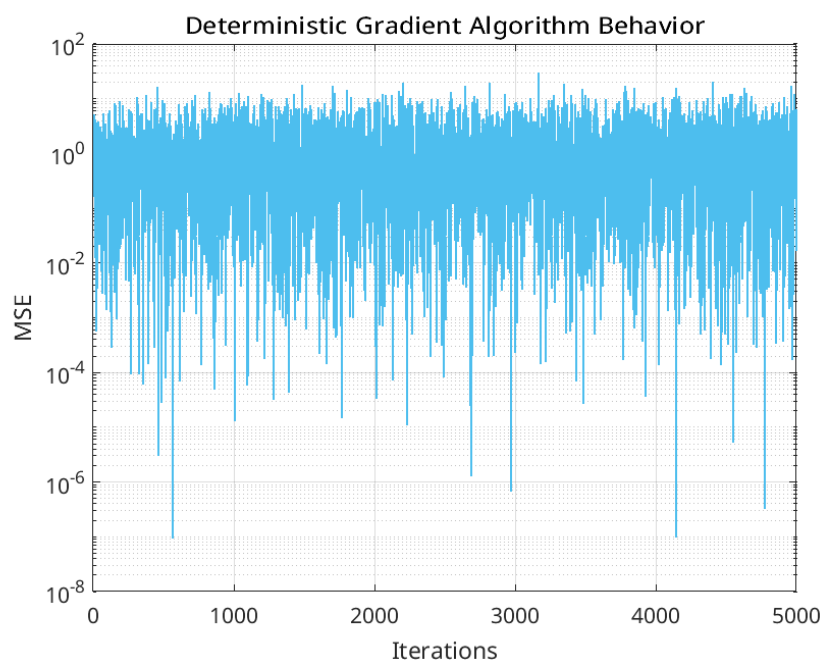


Figura 10: Caminho de convergência na superfície MSE para o Gradiente Determinístico. Amostras = 5000, $M = 2$, $\mu = 10^{-3}$

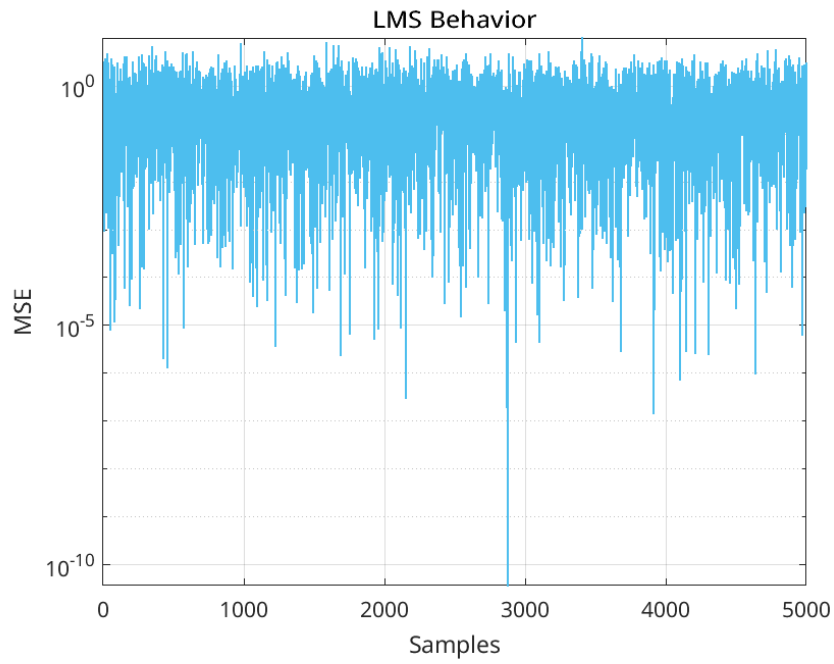


Figura 11: Caminho de convergência na superfície MSE para o LMS. Amostras = 5000, $M = 2$, $\mu = 10^{-3}$

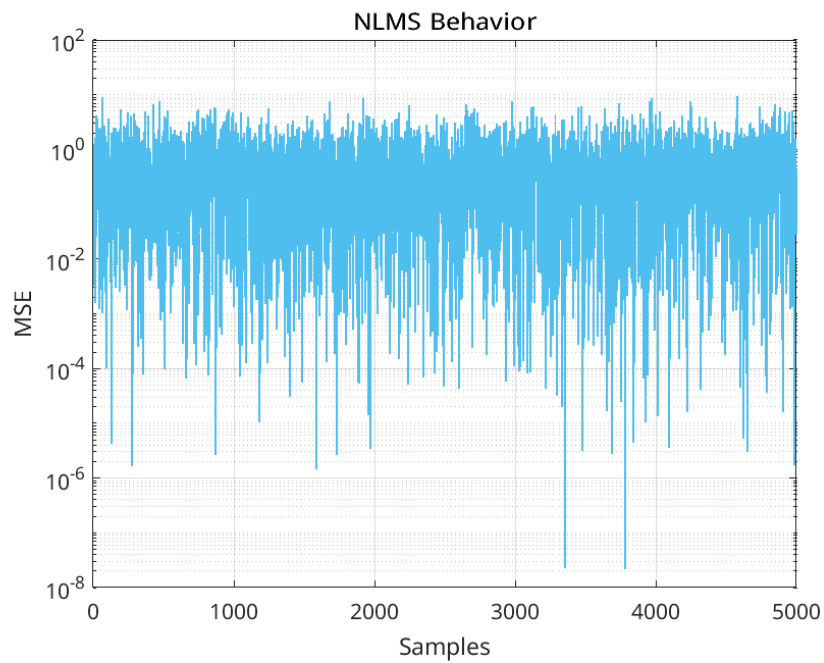


Figura 12: Caminho de convergência na superfície MSE para o NLMS. Amostras = 5000, $M = 2$, $\mu = 10^{-3}$

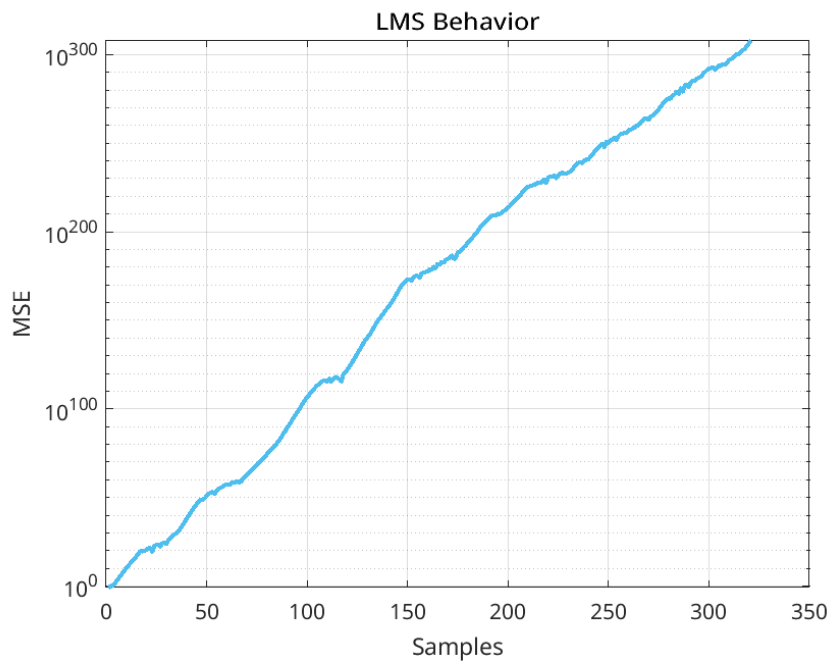


Figura 13: Amostras = 1500, $M = 15$, $\mu = \frac{\mu_{\max}}{2}$

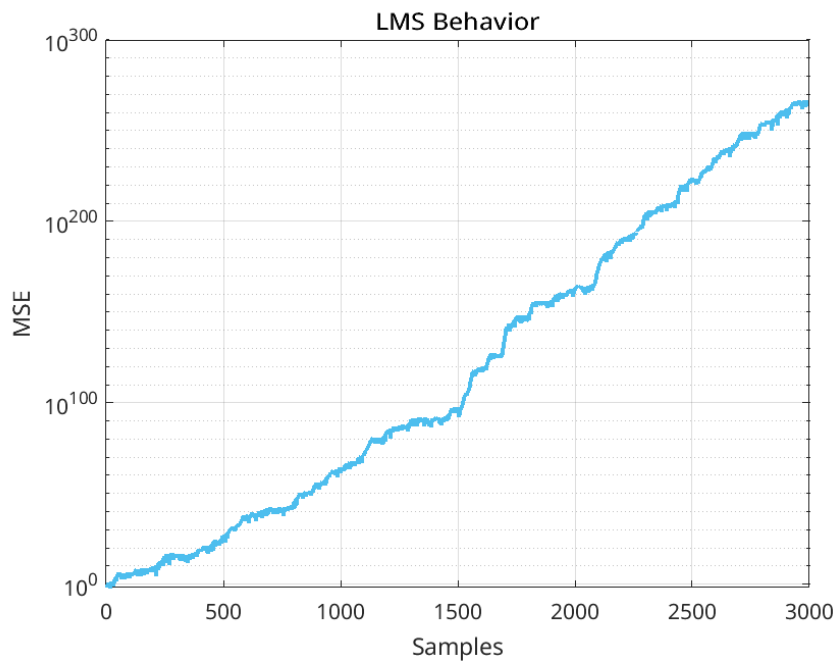


Figura 14: Amostras = 1500, $M = 15$, $\mu = \frac{\mu_{\max}}{10}$

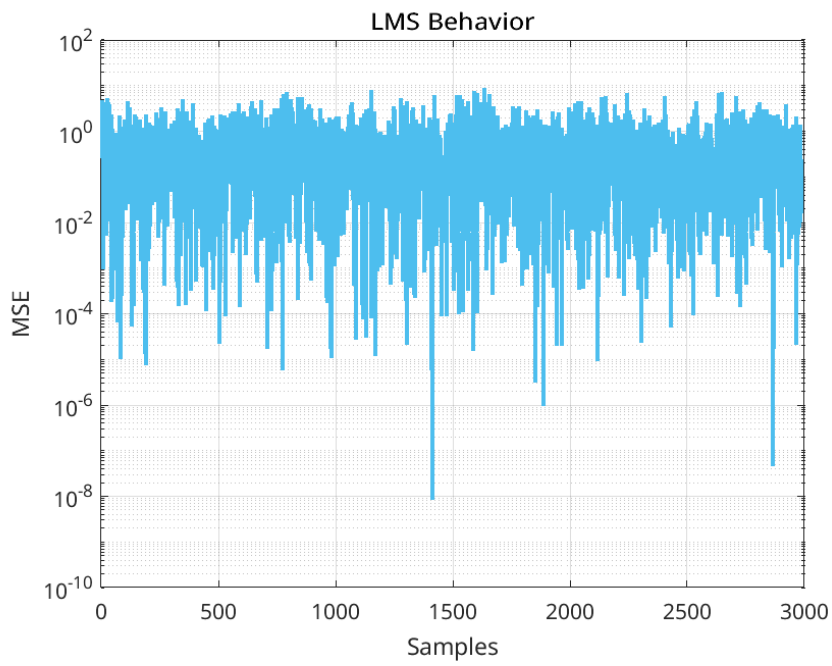


Figura 15: Amostras = 1500, $M = 15$, $\mu = \frac{\mu_{\max}}{50}$

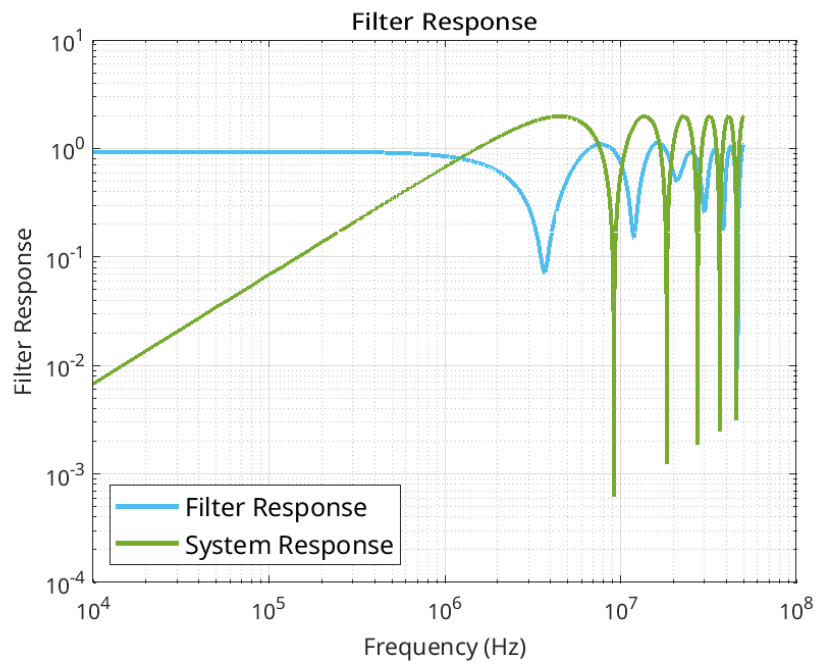


Figura 16: Amostras = 1500, $M = 15$, $\mu = \frac{\mu_{\max}}{2}$

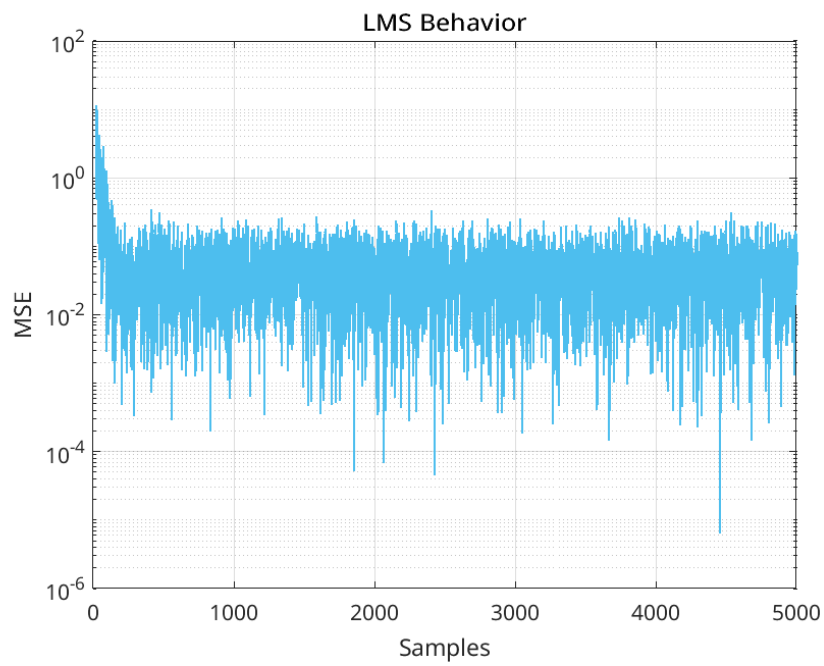


Figura 17: Amostras = 1500, $M = 15$, $\mu = \frac{\mu_{\max}}{2}$

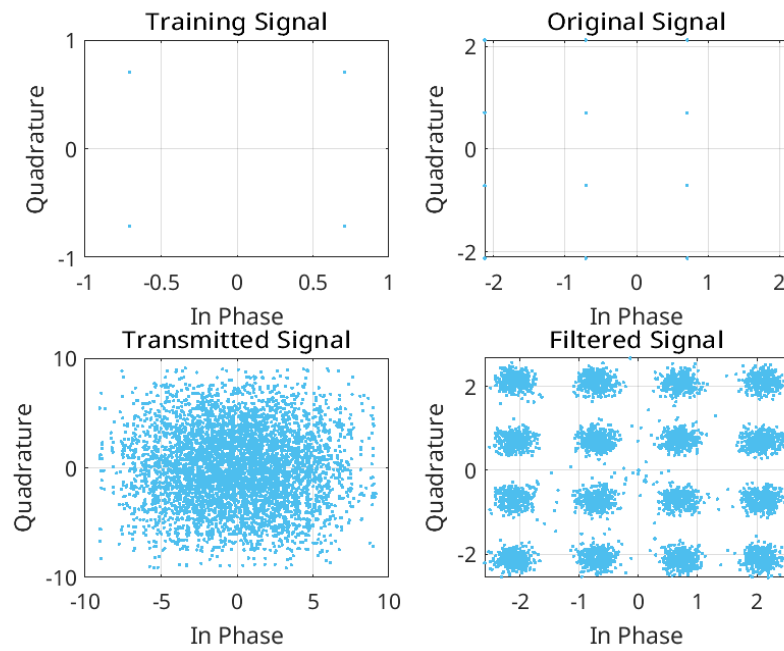


Figura 18: Amostras = 5000, $M = 15$, $\mu = 0.4$

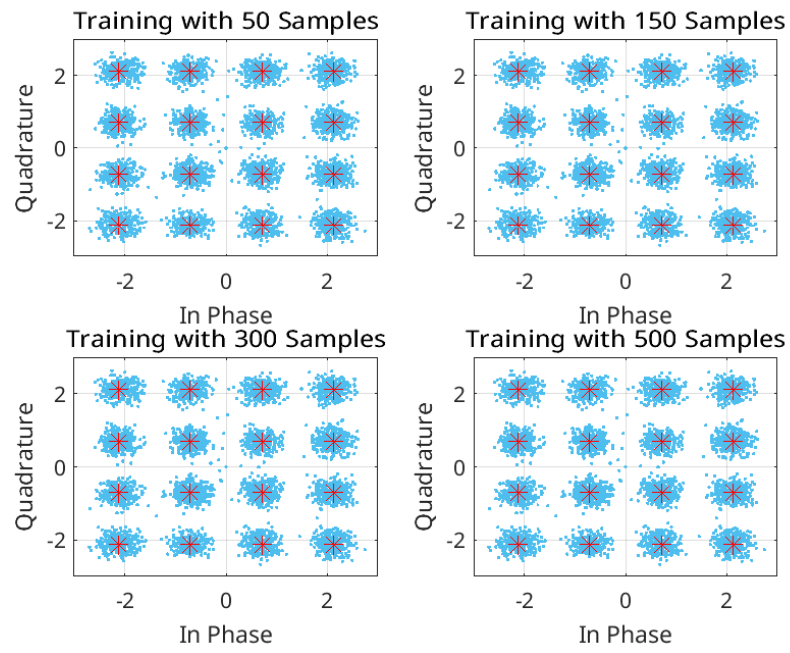


Figura 19: Amostras = 5000, $M = 15$, $\mu = 0.001$

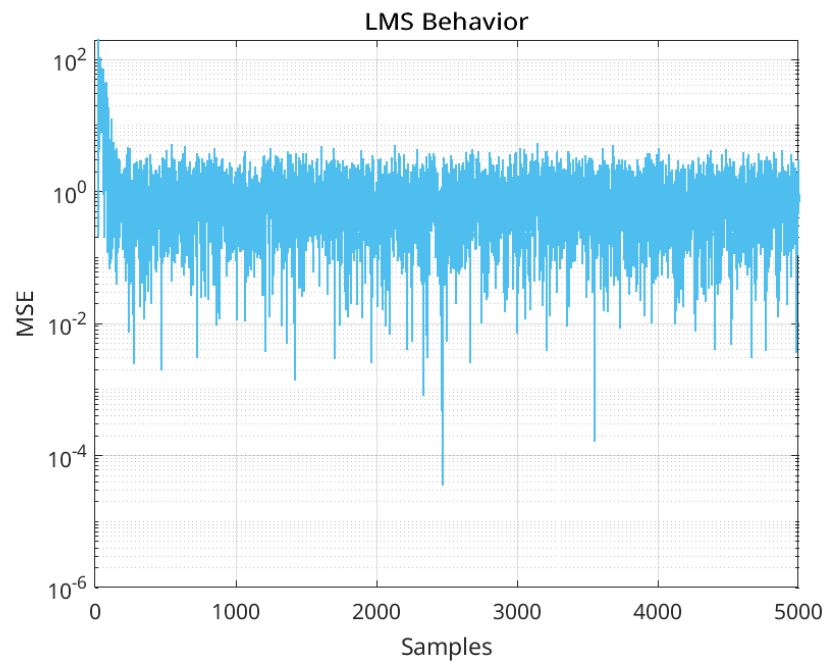


Figura 20: Amostras = 5000, $M = 15$, $\mu = 0.4$

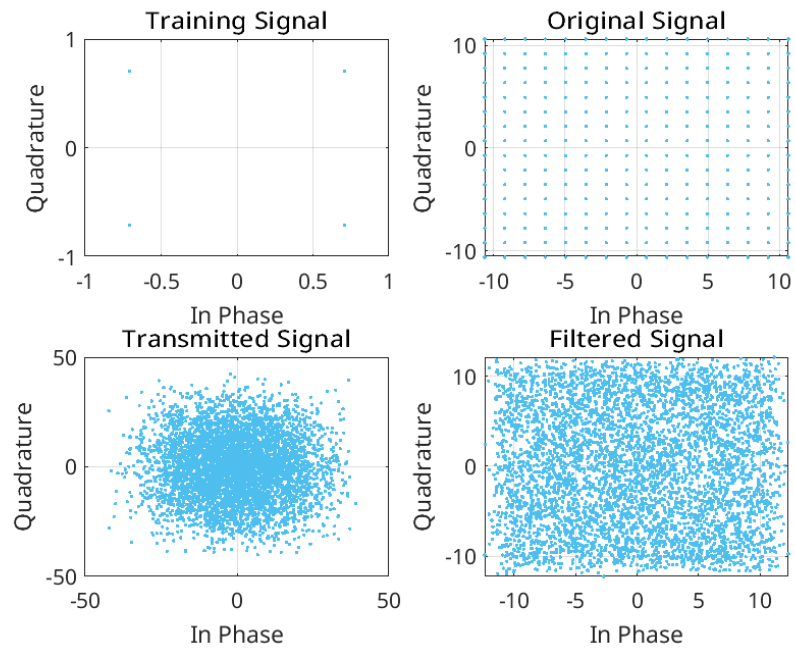


Figura 21: Amostras = 5000, $M = 15$, $\mu = 0.4$

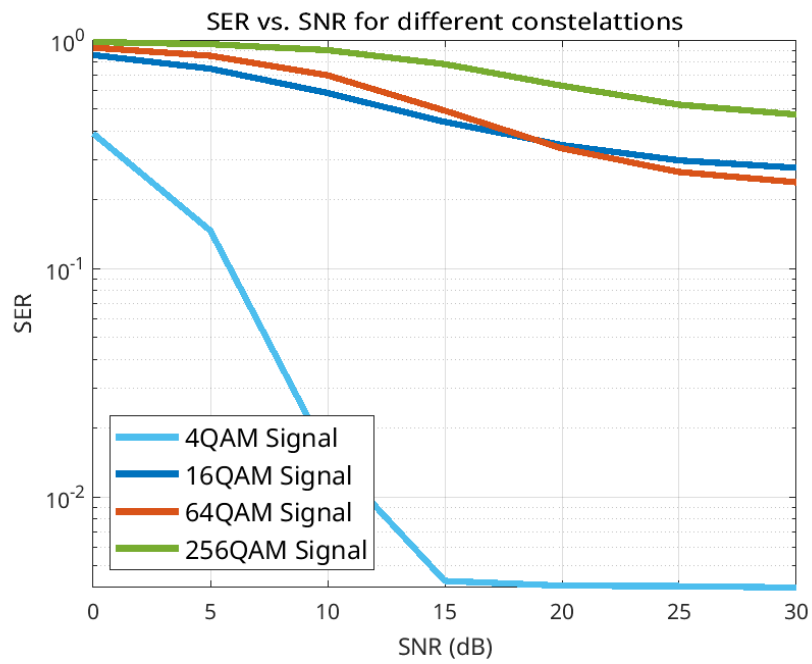


Figura 22: Amostras = 5000, $M = 15$, $\mu = 0.4$