

FILTRAGEM ÓTIMA

No âmbito de filtragem adaptativa, três passos fundamentais são necessários para o projeto de um filtro:

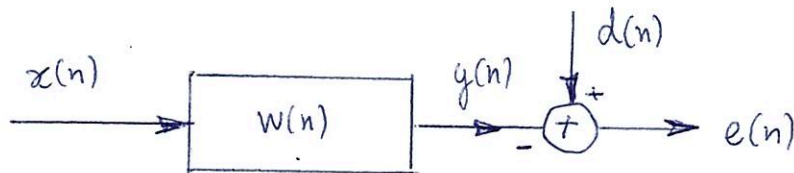
- (i) escolha de uma estrutura adequada de filtragem: linear (FIR, IIR), não-linear (Volterra, MLPs, RBFs), feedforward recurrente.
- (ii) definição de um critério a ser otimizado: uma expressão matemática que exprime o objetivo que se deseja atingir com o filtro. Por exemplo, o mínimo erro quadrático médio entre a saída do filtro e um sinal de referência.
- (iii) emprego de um algoritmo para determinar os parâmetros ótimos do filtro segundo o critério de filtragem.

O alicerce teórico de filtragem adaptativa baseia-se no uso de filtros discretos cuja estrutura é a de um combina-
dor linear (FIR) e que são projetados de acordo com o critério de mínimo erro quadrático médio (MMSE) (ou quadrados mínimos, quando a estimação explore um conjunto de amostras).

Primeiramente, vamos abordar o problema de filtragem ótima. Ou seja, vamos determinar e analisar a solução fechada para o problema de filtragem supervisionada.

Posteriormente, iremos derivar técnicas adaptativas que tentam encontrar os parâmetros ótimos do filtro até mesmo quando as propriedades estatísticas dos sinais mudam com o tempo (não-estacionário).

* Filtragem Supervisionada

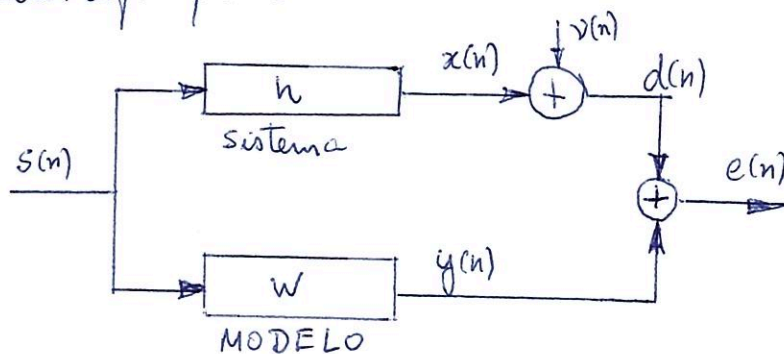


• o filtro realiza uma combinação linear de um conjunto de amostras de entrada ($x(n)$), produzindo uma saída ($y(n)$) de acordo com os valores de seus parâmetros.

Supervisionado: o processo de busca pelos parâmetros ótimos é guiado por uma sequência que representa a resposta desejada do filtro para um conjunto de padrões de entrada, o que permite uma comparação entre a saída do filtro ($y(n)$) e a saída desejada ($d(n)$).

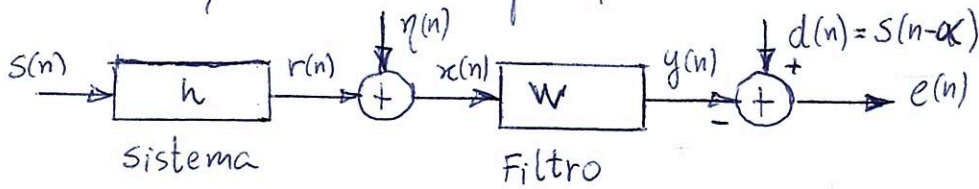
Aplicações:

a) Identificação de sistemas



queremos modelar (aproximar) um sistema cujos parâmetros não são conhecidos através de um filtro linear.

b) Desconvolução - identificação inversa.



queremos cancelar (invertir) o efeito de um sistema desconhecido

c) Predição linear

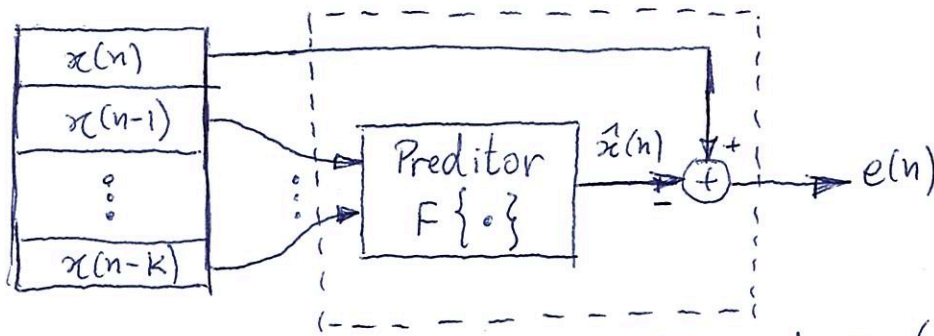
A partir da amostra atual de um conjunto de valores passados, queremos estimar valores futuros de uma série temporal.

Seja $\underline{x}(n-1) = [x(n-1), \dots, x(n-k)]^T$. Genericamente, um preditor realiza um mapeamento do tipo $\hat{x}(n) = F\{\underline{x}(n-1)\}$

Caso linear: $\hat{x}(n) = \underline{a}^T \underline{x}(n-1) = \sum_{k=0}^{K-1} a_k x(n-1-k)$

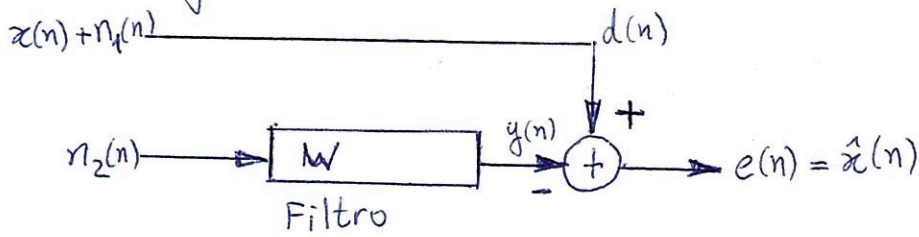
a_0, \dots, a_{k-1} são os coeficientes do preditor

Erro: $e(n) = x(n) - \hat{x}(n)$



Filtro de erro de predição (FEP)

d) Signal Enhancement

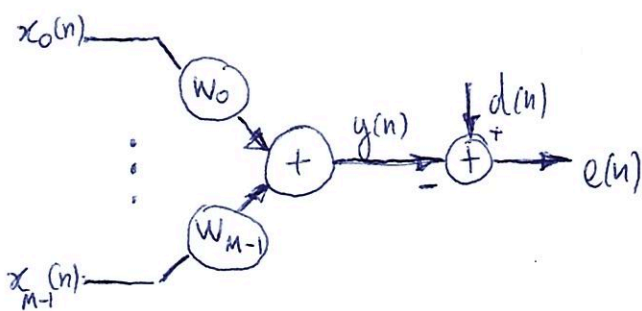


Um sinal $x(n)$ é corrompido por ruído ($n_1(n)$) e um sinal $n_2(n)$, correlacionado com $n_1(n)$, está disponível (pode ser medido). Usando $n_2(n)$ como entrada do filtro e o sinal corrompido ($x(n) + n_1(n)$) como referência, o erro de saída deve convergir para uma versão enfatizada (com ruído removido) do sinal $x(n)$.

* Filtragem de Wiener

O problema pode ser descrito por duas configurações equivalentes:

Combinador linear



Filtragem temporal

x_0, x_1, \dots, x_{M-1} - amostras temporais de um mesmo processo



$$W(z) = w_0 + w_1 z^{-1} + \dots + w_{M-1} z^{-(M-1)}$$

Em ambos os casos, podemos definir:

ENTRADA: $\underline{x}(n) = [x_0(n) \ x_1(n) \ \dots \ x_{M-1}(n)]^T$ ou $\underline{x}(n) = [x(n) \ x(n-1) \ \dots \ x(n-M+1)]^T$

Parâmetros do filtro: $\underline{w} = [w_0 \ w_1 \ \dots \ w_{M-1}]^T$

Assim, $y(n) = \underline{w}^T \underline{x}(n) = \underline{x}^T(n) \underline{w}$ e o erro é dado por $e(n) = d(n) - y(n)$.

CRITÉRIO: minimizar o erro quadrático médio entre o sinal desejado e sua estimativa.

$$\min_{\underline{w}} J(\underline{w}) = \min_{\underline{w}} E\{e^2(n)\}$$

HIPÓTESE: $x(n)$ e $d(n)$ são processos aleatórios ^{reais} discretos no tempo e estacionários no sentido amplo (WSS) e de média nula.

SOLUÇÃO: $\nabla_{\underline{w}} J(\underline{w}) = \underline{0}$ (gradiente nulo) ou $\frac{\partial J(\underline{w})}{\partial w_i} = 0$ para $i=0, 1, \dots, M-1$.

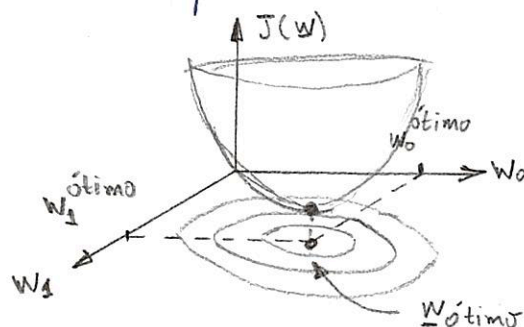
$$\begin{aligned} J(\underline{w}) &= E\{e^2(n)\} = E\{[d(n) - y(n)]^2\} = E\{[d(n) - \underline{w}^T \underline{x}(n)]^2\} \\ &= E\{d^2(n) - 2d(n)\underline{w}^T \underline{x}(n) + \underline{w}^T \underline{x}(n) \cdot \underline{x}^T(n) \underline{w}\} \\ &= E\{d^2(n)\} - 2E\{d(n)\underline{w}^T \underline{x}(n)\} + E\{\underline{w}^T \underline{x}(n) \underline{x}^T(n) \underline{w}\} \\ &= \sigma_d^2 - 2\underline{w}^T E\{d(n) \underline{x}(n)\} + \underline{w}^T E\{\underline{x}(n) \underline{x}^T(n)\} \underline{w} \end{aligned}$$

Definimos: $\underline{p}_{xd} = E\{d(n) \underline{x}(n)\}$ = vetor de correlação cruzada entre o sinal de entrada do filtro e o sinal desejado

$R_x = E\{\underline{x}(n) \underline{x}^T(n)\}$ = matriz de autocorrelação de entrada.

∴, $J(\underline{w}) = \sigma_d^2 - 2\underline{w}^T \underline{p}_{xd} + \underline{w}^T R_x \underline{w} \rightarrow$ curva quadrática em função de \underline{w}

Exemplo: filtro com 2 coeficientes $\underline{w} = [w_0 \ w_1]^T$ ($M=2$)



Cálculo do filtro ótimo:

$$\frac{\partial J(\underline{w})}{\partial \underline{w}^T} = -2 \underline{p}_{xd} + 2 R_x \underline{w} = 0$$

$$R_x \cdot \underline{w} = \underline{p}_{xd} \Rightarrow \text{equações de Wiener-Hopf.}$$

Finalmente,

$$\boxed{\underline{w}_0 = R_x^{-1} \underline{p}_{xd}} \quad \text{solução de Wiener.}$$

Observações:

- o filtro ótimo depende das estatísticas de segunda ordem dos sinais envolvidos.
- é preciso conhecer as médias estatísticas em R_x e \underline{p}_{xd}

Etapas para obtenção do filtro ótimo:

- 1) aquisição de dados (offline)
- 2) estimação da autoconexão do processo de entrada e da conexão cruzada entre $x(n)$ e $d(n)$.
- 3) otimização do filtro usando a equação $\underline{w}_0 = \hat{R}_x^{-1} \hat{\underline{p}}_{xd}$

Quando minimizamos o MSE, isto não necessariamente quer dizer que o sinal de erro alcançado será nulo, ou seja, que $y(n) = d(n)$. O valor mínimo do MSE, ou $J(\underline{w}_0)$, é dado por:

$$J(\underline{w}_0) = \sigma_d^2 - \underline{w}_0^T \underline{p}_{xd} - \underline{p}_{xd}^T \underline{w}_0 + \underbrace{\underline{w}_0^T \underline{R}_x \underline{w}_0}_{= \underline{p}_{xd}} = \underline{\sigma_d^2 - \underline{p}_{xd}^T \underline{R}_x^{-1} \underline{p}_{xd}}$$

O erro quadrático médio pode não ser nulo pois: (1) existe ruído nas amostras, (2) a estrutura de filtragem não tem flexibilidade suficiente para o problema, (3) há informações insuficientes sobre o sinal desejado nas amostras de entrada do filtro, etc.

PROPRIEDADE: $J(\underline{w}) = E\{e^2(n)\}$

$$\nabla_{\underline{w}} J(\underline{w}) = E\left\{2e(n) \frac{\partial e(n)}{\partial \underline{w}}\right\} = 2E\left\{e(n) \frac{\partial \{d(n) - \underline{w}^T \underline{x}(n)\}}{\partial \underline{w}}\right\} = -2E\{e(n) \underline{x}(n)\}$$

No ponto ótimo, o gradiente se anula $\Rightarrow -2E\{e(n) \underline{x}(n)\} = 0$

• Isso significa que o sinal de erro gerado pelo filtro de Wiener é ortogonal às entradas. Em certo sentido, o erro $e(n)$ obtido pode ser visto como a parcela do sinal desejado que não pode ser extraído de $\underline{x}(n)$ uma vez que este é decorrelacionado do conjunto de entradas.

• Se $E\{e(n) \underline{x}(n)\} = 0$, então $E\{e(n) y(n)\} = 0$ também. Ou seja, o sinal de erro produzido pelo filtro ótimo é ortogonal à saída do filtro de Wiener.

Tanto na solução de Wiener quanto no caso de alguns algoritmos adaptativos que veremos, a matriz de autocorrelação da entrada R_x exerce papel importante. Por isso, é pertinente conhecer algumas de suas propriedades.

* Propriedades da matriz de autocorrelação

Definindo $\underline{x}(n) = [x(n) \ x(n-1) \ \dots \ x(n-M+1)]^T$, $R_x = E\{\underline{x}(n)\underline{x}^T(n)\}$:

$$R_x = \begin{bmatrix} E\{x^2(n)\} & E\{x(n)x(n-1)\} & \dots & E\{x(n)x(n-M+1)\} \\ E\{x(n)x(n-1)\} & E\{x^2(n-1)\} & \dots & E\{x(n-1)x(n-M+1)\} \\ \vdots & \vdots & \ddots & \vdots \\ E\{x(n-M+1)x(n)\} & E\{x(n-M+1)x(n-1)\} & \dots & E\{x^2(n-M+1)\} \end{bmatrix}$$

Seja $r_x(i) = E\{x(n)x(n-i)\}$, temos:

$$R_x = \begin{bmatrix} r_x(0) & r_x(1) & \dots & r_x(M-1) \\ r_x(1) & r_x(0) & & \vdots \\ \vdots & & \ddots & r_x(1) \\ r_x(M-1) & \dots & r_x(1) & r_x(0) \end{bmatrix}$$

Obs: $r_x(i) = r_x(-i)$ no caso real.

i) R_x é simétrica ($R_x = R_x^T$). No caso complexo, R_x é hermitiana ($R_x = R_x^H$)

ii) Quando o sinal de entrada é composto por versões atrasadas de $x(n)$ tomadas de um processo WSS, R_x é uma matriz Toeplitz.

Toeplitz: é possível inverter com algoritmos de complexidade computacional proporcional ao quadrado da ordem (dimensão) da matriz.

iii) R_x é semidefinida positiva:
 $\underline{w}^T R_x \underline{w} \geq 0$ para $\forall \underline{w}$

Autovalores e Autovetores:

$$R_x \underline{q}_i = \lambda_i \underline{q}_i \quad \begin{cases} \lambda_i \text{ é o autovalor de } R_x \text{ associado a } \underline{q}_i \\ \underline{q}_i \text{ é um autovetor de } R_x \end{cases}$$

$$\underline{q}_i^T R_x \underline{q}_i = \underline{q}_i^T \lambda_i \underline{q}_i = \lambda_i \underline{q}_i^T \underline{q}_i \Rightarrow \lambda_i = \frac{\underline{q}_i^T R_x \underline{q}_i}{\underline{q}_i^T \underline{q}_i} \geq 0 \quad \hookrightarrow \text{propriedade (iii)}$$

iv) Os autovalores de R_x são reais não-negativos.

Por questões práticas, considere-se que os autovetores têm norma unitária: $\underline{q}_i^T \underline{q}_i = 1$, $\forall i$

v) Se R_x possui M autovetores linearmente independentes, então $Q^{-1} R_x Q = \Lambda$, onde

$$Q = [\underline{q}_0 \quad \underline{q}_1 \quad \dots \quad \underline{q}_{M-1}] \quad \text{e} \quad \Lambda = \begin{bmatrix} \lambda_0 & 0 & \dots & 0 \\ 0 & \lambda_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_{M-1} \end{bmatrix}$$

A prova para pela percepção de que $R_x Q = R_x [q_0 \dots q_{M-1}] =$
 $= [\lambda_0 q_0 \dots \lambda_{M-1} q_{M-1}] = Q \Delta$

vii) Os autovetores de R_x são ortogonais entre si:

$$q_i^T q_j = 0, \quad i \neq j$$

viii) A matriz de autovetores (normalizados) Q é unitária,
 i.e., $Q Q^T = I$ ou $Q^T Q = I \Rightarrow Q^{-1} = Q^T$

Decomposição Espectral

segundo as propriedades (vi) e (viii):

$$R_x = Q \Delta Q^T = \sum_{i=0}^{M-1} \lambda_i q_i q_i^T \quad (\text{note que } \Delta = Q^T R_x Q)$$

• traço da matriz R_x :

$$\text{tr}(\Delta) = \sum_{i=0}^{M-1} \lambda_i$$

$$\text{Mas, } \text{tr}(\Delta) = \text{tr}(Q^T R_x Q)$$

Usando a propriedade $\text{tr}(AB) = \text{tr}(BA)$, temos

$$\text{tr}(Q^T R_x Q) = \text{tr}(R_x Q Q^T) = \text{tr}(R_x) = M \pi_x(0) = \text{tr}(\Delta) = \sum_{i=0}^{M-1} \lambda_i$$

Portanto,

$$\boxed{\pi_x(0) = \frac{1}{M} \sum_{i=0}^{M-1} \lambda_i}$$

• Condicionamento ou espalhamento de R_x : é possível estabelecer uma importante relação entre a razão dos autovalores máximo e mínimo de R_x e o perfil da densidade de espectral de potência $S_x(e^{j\omega})$.

Vimos que $\lambda_i = \frac{\mathbf{q}_i^T R_x \mathbf{q}_i}{\mathbf{q}_i^T \mathbf{q}_i}$

O numerador pode ser escrito como $\mathbf{q}_i^T R_x \mathbf{q}_i = \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} q_{ik} r(l-k) q_{il}$

onde q_{ik} é o k -ésimo elemento do vetor linha \mathbf{q}_i^T , $r(l-k)$ é o elemento $[R_x]_{k,l}$ da matriz de autocorrelação e q_{il} é o l -ésimo elemento do vetor coluna \mathbf{q}_i .

Podemos escrever a seguinte relação:

$$r(l-k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_x(e^{j\omega}) e^{j\omega(l-k)} d\omega \rightarrow \text{Transf. inversa de PSD.}$$

Então,

$$\begin{aligned} \mathbf{q}_i^T R_x \mathbf{q}_i &= \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} q_{ik} r(l-k) q_{il} = \\ &= \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} q_{ik} \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} S_x(e^{j\omega}) e^{j\omega(l-k)} d\omega \right] q_{il} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} S_x(e^{j\omega}) \sum_{k=0}^{M-1} q_{ik} e^{-j\omega k} \sum_{l=0}^{M-1} q_{il} e^{j\omega l} d\omega \end{aligned}$$

Seja $Q_i(e^{j\omega}) = \sum_{k=0}^{M-1} q_{ik} e^{-j\omega k}$. Então,

$$q_i^T R_x q_i = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_x(e^{j\omega}) Q_i(e^{j\omega}) Q_i^*(e^{+j\omega}) d\omega$$

$$q_i^T R_x q_i = \frac{1}{2\pi} \int_{-\pi}^{\pi} |Q_i(e^{j\omega})|^2 S_x(e^{j\omega}) d\omega$$

Usando o mesmo raciocínio, é possível mostrar que

$$q_i^T q_i = \frac{1}{2\pi} \int_{-\pi}^{\pi} |Q_i(e^{j\omega})|^2 d\omega$$

$$\text{Assim, } \lambda_i = \frac{\int_{-\pi}^{\pi} |Q_i(e^{j\omega})|^2 S_x(e^{j\omega}) d\omega}{\int_{-\pi}^{\pi} |Q_i(e^{j\omega})|^2 d\omega}$$

Sejam S_{\min} e S_{\max} os valores absolutos mínimo e máximo da PSD de $x(n)$, $S_x(e^{j\omega})$. Então,

$$\int_{-\pi}^{\pi} |Q_i(e^{j\omega})|^2 S_x(e^{j\omega}) d\omega \geq S_{\min} \int_{-\pi}^{\pi} |Q_i(e^{j\omega})|^2 d\omega \Rightarrow \lambda_i \geq S_{\min}$$

e

$$\int_{-\pi}^{\pi} |Q_i(e^{j\omega})|^2 S_x(e^{j\omega}) d\omega \leq S_{\max} \int_{-\pi}^{\pi} |Q_i(e^{j\omega})|^2 d\omega \Rightarrow \lambda_i \leq S_{\max}$$

$$\Rightarrow \boxed{S_{\min} \leq \lambda_i \leq S_{\max}}$$

Consequentemente,

$$\chi(R_x) = \frac{\lambda_{\max}}{\lambda_{\min}} \leq \frac{S_{\max}}{S_{\min}}$$

em que $\chi(R_x)$ é denominado o número de condicionamento da matriz de autocondição.

Casos extremos:

1) $\chi(R_x) = 1$ — todos os autovalores são iguais $\Rightarrow S_{\min} = S_{\max}$

O espectro de $x(n)$ é pleno, ou seja, $x(n)$ é um ruído branco, o cálculo de R_x^{-1} é imediato (matriz bem condicionada).

2) $\chi(R_x) \rightarrow \infty$ — $\lambda_{\min} \rightarrow 0$
 $S_{\min} \rightarrow 0$

O espectro de $x(n)$ possui nulos espectrais. Nesse caso não há inversa de R_x .

* Superfície de Erro Quadrático Médio

$$J(\underline{w}) = \sigma_d^2 - 2\underline{w}^T \underline{p}_{xd} + \underline{w}^T R_x \underline{w}$$

É conveniente realizar algumas trocas de variáveis para observar as propriedades da superfície MSE em relação ao condicionamento da matriz de autocorrelação do sinal de entrada do filtro.

$$\underline{\Delta w} = \underline{w} - \underline{w}_0 \quad \begin{array}{l} \rightarrow \text{filtro ótimo} \\ \rightarrow \text{translação do ponto ótimo de } J(\underline{w}) \\ \text{para a origem} \end{array}$$

$$J_{\min} = J(\underline{w}_0) = \sigma_d^2 - \underline{w}_0^T \underline{p}_{xd}$$

Para somar e subtrair $\underline{w}_0^T \underline{p}_{xd}$ da função MSE:

$$\begin{aligned} J(\underline{w}) &= \underbrace{\sigma_d^2 - \underline{w}_0^T \underline{p}_{xd}}_{J_{\min}} + \underline{w}_0^T \underline{p}_{xd} - 2\underline{w}^T \underline{p}_{xd} + \underline{w}^T R_x \underline{w} \\ &= J_{\min} + \underbrace{\underline{w}_0^T \underline{p}_{xd} - \underline{w}^T \underline{p}_{xd} - \underline{w}_0^T \underline{p}_{xd}}_{(-\underline{\Delta w}^T \underline{p}_{xd})} + \underline{w}^T R_x \underline{w} \\ &= J_{\min} + (-\underline{\Delta w}^T \underline{p}_{xd}) - \underline{w}_0^T \underline{p}_{xd} + \underline{w}^T R_x \underline{w} \end{aligned}$$

Como $R_x \underline{w}_0 = \underline{p}_{xd}$:

$$\begin{aligned} J(\underline{w}) &= J_{\min} - \underline{\Delta w}^T \underline{p}_{xd} - \underbrace{\underline{w}_0^T R_x \underline{w}_0}_{\underline{w}_0^T \underline{p}_{xd}} + \underline{w}^T R_x \underline{w} \\ &= J_{\min} - \underline{\Delta w}^T \underline{p}_{xd} + \underline{w}^T R_x \underline{\Delta w} \end{aligned}$$

Como $\underline{\Delta w}^T \underline{p}_{xd}$ é um escalar, $\underline{\Delta w}^T \underline{p}_{xd} = (\underline{\Delta w}^T \underline{p}_{xd})^T = \underline{p}_{xd}^T \underline{\Delta w} = \underline{w}_0^T R_x \underline{\Delta w}$

Logo, $J(\underline{\Delta w}) = J_{\text{MIN}} - \underline{w}_0^T R_x \underline{\Delta w} + \underline{w}^T R_x \underline{\Delta w}$

$$= \underline{J_{\text{MIN}} + \underline{\Delta w}^T R_x \underline{\Delta w}}$$

2ª troca de variáveis:

Podemos escrever $R_x = Q \Lambda Q^T$

$$J(\underline{\Delta w}) = J_{\text{MIN}} + \underline{\Delta w}^T \underbrace{Q \Lambda Q^T}_{\underline{\nu} = Q^T \underline{\Delta w}} \underline{\Delta w}$$

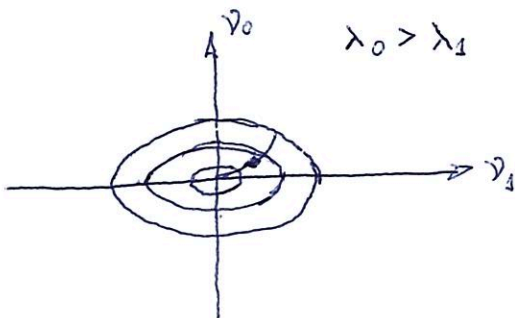
→ parâmetros rotacionados.

Então, $J(\underline{\nu}) = J_{\text{MIN}} + \underline{\nu}^T \Lambda \underline{\nu}$

↓
matriz diagonal → há apenas elementos quadráticos do tipo ν_0^2, ν_1^2, \dots

Portanto, $J(\underline{\nu}) = J_{\text{MIN}} + \lambda_0 \nu_0^2 + \lambda_1 \nu_1^2 + \dots + \lambda_{n-1} \nu_{n-1}^2$

As curvas de nível de $J(\underline{\nu})$ são hiperelipses:



⇒ eixos de elipse alinhados com ν_0 e ν_1 .

$\lambda_0 \gg \lambda_1$ → curvas de nível mais agudas.

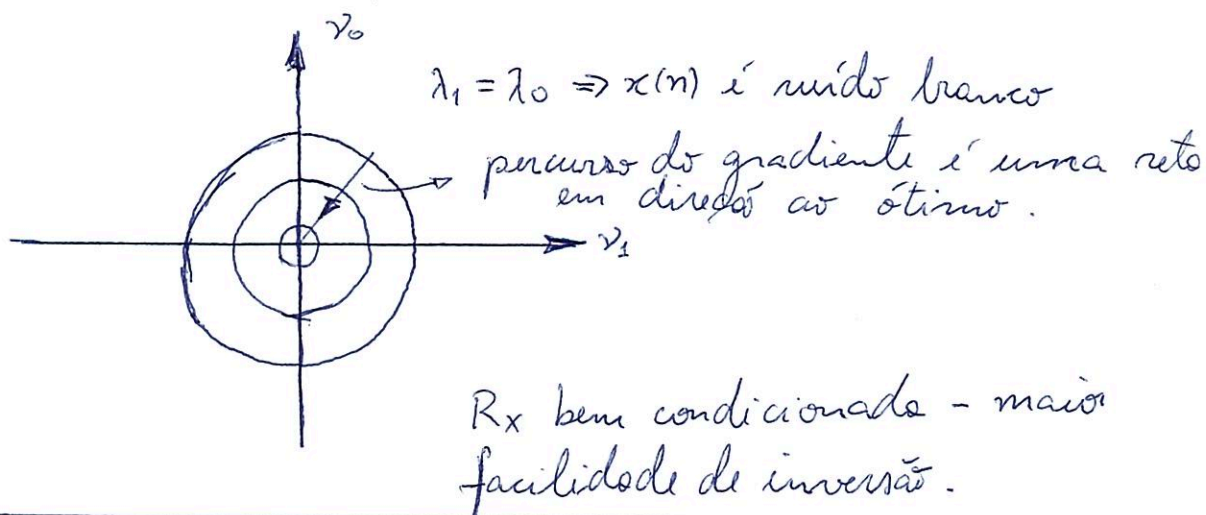
→ $x(n)$ possui espectro espalhado $S_{\text{max}} \gg S_{\text{min}}$

→ Matriz R_x mal condicionada

→ maior dificuldade de inversão de R_x

→ percurso do gradiente constantemente corrigido.

Quando temos autovalores iguais; as curvas de nível da superfície MSE se transformam em círculos concêntricos.



* Filtragem com restrições lineares.

Até este ponto, vimos que a existência de um sinal de referência, junto com um critério apropriado, leva a uma solução linear e a uma função custo convexa.

O uso explícito de um sinal de referência nem sempre é possível ou necessário em alguns problemas práticos. Tais problemas podem ser separados em duas categorias distintas:

a) o sinal de referência é desejável, mas não é acessível. Nesse caso, devemos recorrer a uma certa porção de conhecimento a priori a respeito de sua natureza, bem como de suas propriedades estatísticas. Esta é a essência do processamento não-supervisionado (cego).

b) O sinal de referência não é necessário p/ a tarefa, podendo ser substituído propositalmente por um conjunto de restrições sobre os parâmetros do filtro que viabilize a otimização e/ou o processo de adaptação.

Problema típico que se enquadra na segunda categoria: filtro de mínima variância com restrições lineares ou Linearly Constrained Minimum Variance (LCMV) filter.

$$y(n) = \underline{w}^H \underline{x}(n)$$

Nesse caso, minimizar $E\{|y(n)|^2\}$ nos leva à solução trivial $\underline{w} = \underline{0}$.

Um conjunto não-trivial de coeficientes pode ser obtido mediante o seguinte procedimento:

$$\begin{cases} \arg \min_{\underline{w}} E\{|y(n)|^2\} = E\{\underline{w}^H \underline{x}(n) \underline{x}^H(n) \underline{w}\} = \underline{w}^H R_x \underline{w} \\ \text{s.a. } C^H \underline{w} = \underline{g} \end{cases}$$

↳ Minimização de norma quadrática sujeita a um conjunto de restrições lineares.

Solução usando multiplicadores de Lagrange:

$$\arg \min_{\underline{w}} \frac{1}{2} \underline{w}^H R_x \underline{w} + \underline{\lambda}^H (C^H \underline{w} - \underline{g})$$

$$\text{gradiente: } R_x \underline{w} + C \underline{\lambda} = 0$$

Pré-multiplicando por $C^H R_x^{-1}$:

$$\underbrace{C^H R_x^{-1} R_x \underline{w}} + C^H R_x^{-1} C \underline{\lambda} = 0 \Rightarrow C^H R_x^{-1} C \underline{\lambda} + \underline{g} = 0$$

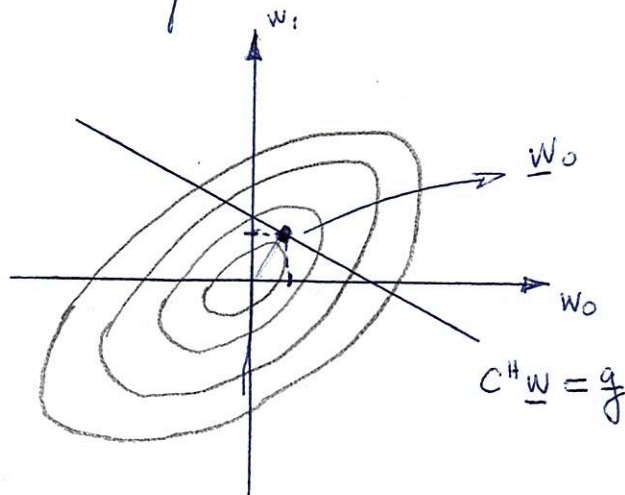
$$C^H \underline{w} = \underline{g}$$

$$\text{Então; } \underline{\lambda} = -(C^H R_x^{-1} C)^{-1} \underline{g}$$

Portanto, a solução ótima é

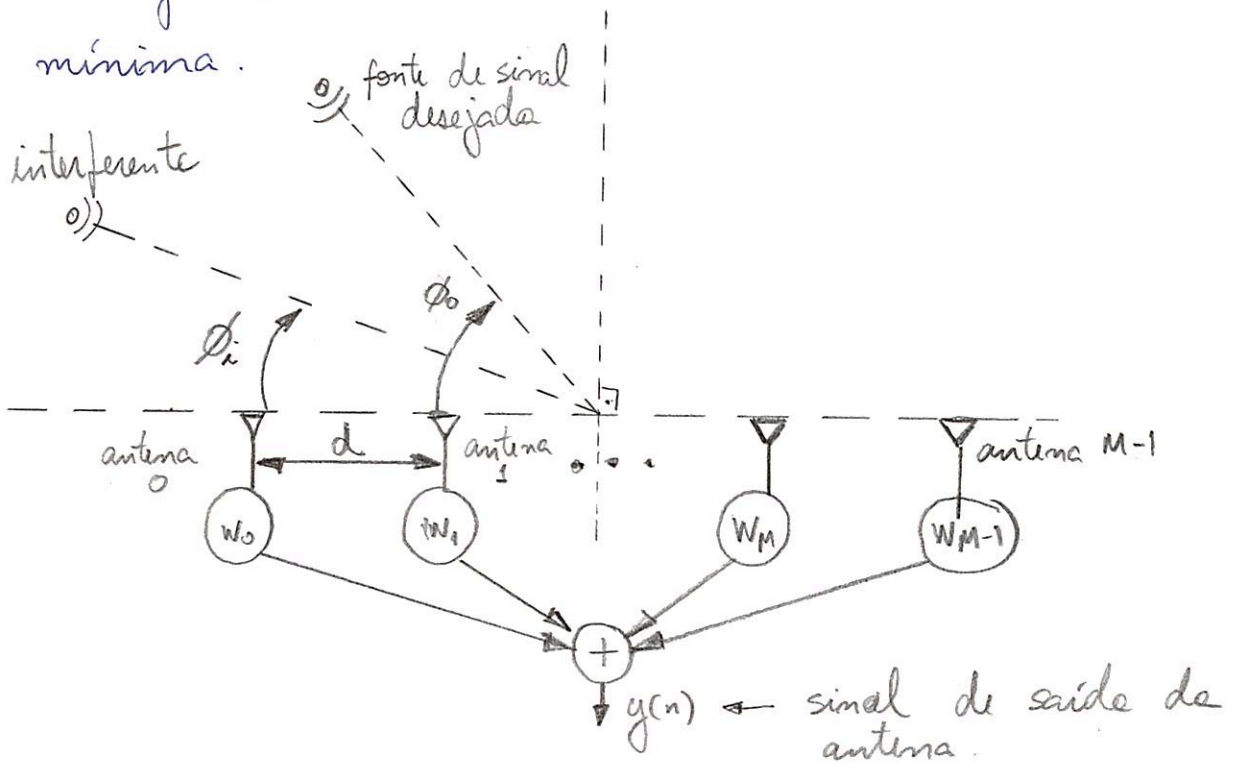
$$\underline{w}_0 = R_x^{-1} C (C^H R_x^{-1} C)^{-1} \underline{g}$$

Visualização do efeito de restrições:



$C^H \underline{w} = \underline{g}$ → as restrições definem um hiperplano que intercepta o parabolóide heliptico que caracteriza a função custo.

Exemplo: Arranjo de antenas (Formatador de Feixes) de variância mínima.



- arranjo linear uniforme omnidirecional (modelo de sinal em banda base)
- distância da n -ésima antena em relação à antena 0: nd

Objetivos do filtro: (1) preservar um determinado sinal alvo e (2) minimizar a interferência de outras fontes.

• o sinal de interesse se propaga numa direção conhecida, ϕ_0 ; as demais fontes de sinal possuem outros ângulos de chegada ϕ_j .

• A defasagem entre os sinais que atingem os elementos de antena é dada pelo ângulo elétrico:

$$\theta_j = \frac{2\pi d}{\lambda} \cdot \phi_j$$

em que:

- d é a distância entre elementos adjacentes da antena.
- λ é o comprimento de onda do sinal incidente
- ϕ_i é a direção de chegada, pertencente ao intervalo $(-\pi/2, \pi/2]$

O vetor $\underline{x}(n)$ recebido no arranjo produzido por uma fonte $s_i(n)$, posicionada a um ângulo de incidência ϕ_i (direção de chegada) é dado por

$$\underline{x}(n) = \underline{a}(\theta_i) s_i(n)$$

em que:

$$\theta_i = (2\pi d/\lambda)\phi_i$$

• o vetor $\underline{a}(\theta_i) = [1 \quad e^{-j\theta_i} \quad e^{-2j\theta_i} \quad \dots \quad e^{-j(M-1)\theta_i}]^T$ é chamado de steering vector.

Considerando que há N sinais (fontes), o vetor de sinal recebido no arranjo é dado por:

$$\underline{x}(n) = \underline{a}(\theta_1) s_1(n) + \dots + \underline{a}(\theta_N) s_N(n)$$

$$= \begin{bmatrix} \underline{a}(\theta_1) & \underline{a}(\theta_2) & \dots & \underline{a}(\theta_N) \end{bmatrix} \begin{bmatrix} s_1(n) \\ s_2(n) \\ \vdots \\ s_N(n) \end{bmatrix} = A \underline{s}(n)$$

em que $A = \begin{bmatrix} 1 & & & 1 \\ e^{-j\theta_1} & & & e^{-j\theta_N} \\ & \dots & & \vdots \\ e^{-j(M-1)\theta_1} & & & e^{-j(M-1)\theta_N} \end{bmatrix}$

Suponha um sinal de interesse com ângulo elétrico θ_0 . Os demais sinais incidentes no arranjo de antenas são considerados interferentes. Nesse caso, o vetor recebido devido ao sinal $s_0(n)$ é dado por $\underline{a}(\theta_0)D_0(n)$.

Nosso objetivo é eliminar (atenuar) todos os sinais interferentes, preservando o sinal desejado $s_0(n)$.

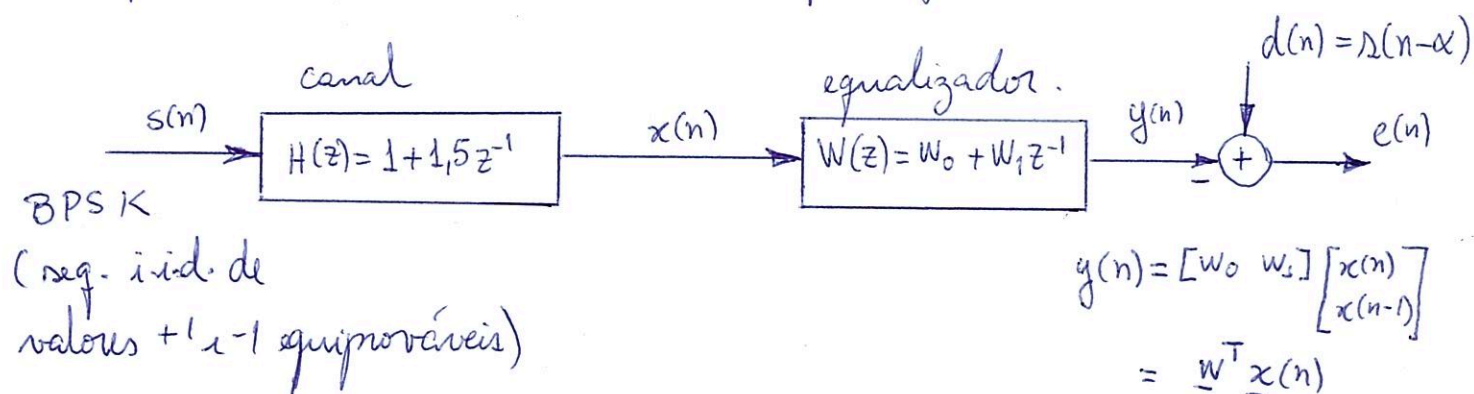
Podemos formular, então, um problema de otimização da forma:

$$\min_{\underline{w}} E\{|y(n)|^2\} = E\{\underline{w}^H \underline{x}(n) \underline{x}^H(n) \underline{w}\} = \underline{w}^H \underline{R}_x \underline{w}$$

$$\text{s.a. } \underline{w}^H \underline{a}(\theta_0) = 1$$

$$\underline{\text{Solução}} : \underline{w}_{\text{otimo}} = \frac{\underline{R}_x^{-1} \underline{a}(\theta_0)}{\underline{a}^H(\theta_0) \underline{R}_x^{-1} \underline{a}(\theta_0)}$$

Exemplo de aplicação da teoria de filtragem ótima: EQUALIZAÇÃO



signal recebido: $x(n) = \Delta(n) + 1,5\Delta(n-1)$

a) Vamos determinar a matriz de autocorrelação $R_X = \begin{bmatrix} r_X(0) & r_X(1) \\ r_X(1) & r_X(0) \end{bmatrix}$

$$\begin{aligned} r_X(1) &= E\{x(n)x(n-1)\} = E\{[\Delta(n) + 1,5\Delta(n-1)][\Delta(n-1) + 1,5\Delta(n-2)]\} \\ &= E\{\Delta(n)\Delta(n-1)\} + E\{\Delta(n) \cdot 1,5 \cdot \Delta(n-2)\} + 1,5 E\{\Delta^2(n-1)\} + E\{1,5^2 \cdot \Delta(n-1)\Delta(n-2)\} \end{aligned}$$

Como $\Delta(n)$ e $\Delta(n-1)$ são independentes e de média nula, $E\{\Delta(n)\Delta(n-1)\} = E\{\Delta(n)\}E\{\Delta(n-1)\} = 0$

O mesmo vale para $E\{\Delta(n-1)\Delta(n-2)\}$ e $E\{\Delta(n)\Delta(n-2)\}$.

Como $\Delta(n)$ e $\Delta(n-1)$ possuem a mesma distribuição, $E\{\Delta^2(n)\} = E\{\Delta^2(n-1)\}$

Assim, $r_X(1) = 1,5 E\{\Delta^2(n-1)\} = 1,5 \cdot \left[\frac{1}{2}(1)^2 + \frac{1}{2}(-1)^2 \right] = 1,5 //$

$$\begin{aligned} r_X(0) &= E\{x(n)x(n)\} = E\{[\Delta(n) + 1,5\Delta(n-1)][\Delta(n) + 1,5\Delta(n-1)]\} \\ &= E\{\Delta^2(n)\} + 3E\{\Delta(n)\Delta(n-1)\} + E\{1,5^2 \cdot \Delta^2(n-1)\} \\ &= 1 + (1,5)^2 = 3,25 // \end{aligned}$$

Logo, $R_X = \begin{bmatrix} 3,25 & 1,5 \\ 1,5 & 3,25 \end{bmatrix}$

b) Vamos determinar o vetor de correlação cruzado f_{xd} para os atrasos $\alpha = 0, 1$ e 2 .

$$\begin{aligned} \bullet E\{x(n)\Delta(n-\alpha)\} &= E\{[\Delta(n) + 1,5\Delta(n-1)]\Delta(n-\alpha)\} \\ &= \underbrace{E\{\Delta(n)\Delta(n-\alpha)\}}_{\neq 0 \text{ se } \alpha=0} + 1,5 \underbrace{E\{\Delta(n-1)\Delta(n-\alpha)\}}_{\neq 0 \text{ se } \alpha=1} \end{aligned}$$

Logo, $\alpha=0 \Rightarrow E\{x(n)\Delta(n)\} = E\{s^2(n)\} = 1$

$\alpha=1 \Rightarrow E\{x(n)\Delta(n-1)\} = 1,5 E\{s^2(n-1)\} = 1,5$

$\alpha=2 \Rightarrow E\{x(n)\Delta(n-2)\} = 0$

$$\begin{aligned} \bullet E\{x(n-1)\Delta(n-\alpha)\} &= E\{[\Delta(n-1) + 1,5\Delta(n-2)]\Delta(n-\alpha)\} \\ &= \underbrace{E\{\Delta(n-1)\Delta(n-\alpha)\}}_{\neq 0 \text{ se } \alpha=1} + 1,5 \underbrace{E\{\Delta(n-2)\Delta(n-\alpha)\}}_{\neq 0 \text{ se } \alpha=2} \end{aligned}$$

Logo, $\alpha=0 \Rightarrow E\{x(n-1)\Delta(n)\} = 0$

$\alpha=1 \Rightarrow E\{x(n-1)\Delta(n-1)\} = E\{s^2(n-1)\} = 1$

$\alpha=2 \Rightarrow E\{x(n-1)\Delta(n-2)\} = 1,5 E\{s^2(n-2)\} = 1,5$

Ou seja, $f_{xd} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ($\alpha=0$) ; $f_{xd} = \begin{bmatrix} 1,5 \\ 1 \end{bmatrix}$ ($\alpha=1$) ; $f_{xd} = \begin{bmatrix} 0 \\ 1,5 \end{bmatrix}$ ($\alpha=2$)

c) Vamos determinar a solução de Wiener, assim como o erro quadrático médio residual, para os três atrasos

$$\bullet \boxed{\alpha=0} \quad \underline{w}_0 = R_x^{-1} \underline{p}_{xd} = \begin{bmatrix} 3,25 & 1,5 \\ 1,5 & 3,25 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0,3910 \\ -0,1805 \end{bmatrix}$$

$$J(\underline{w}_0) = E\{\hat{s}^2(n)\} - \underline{p}_{xd} R_x^{-1} \underline{p}_{xd} = 1 - [1 \ 0] R_x^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \underline{0,6090}$$

$$\bullet \boxed{\alpha=1} \quad \underline{w}_0 = R_x^{-1} \underline{p}_{xd} = \begin{bmatrix} 3,25 & 1,5 \\ 1,5 & 3,25 \end{bmatrix}^{-1} \begin{bmatrix} 1,5 \\ 1 \end{bmatrix} = \begin{bmatrix} 0,4060 \\ 0,1203 \end{bmatrix}$$

$$J(\underline{w}_0) = E\{\hat{s}^2(n-1)\} - \underline{p}_{xd} R_x^{-1} \underline{p}_{xd} = 1 - [1,5 \ 1] R_x^{-1} \begin{bmatrix} 1,5 \\ 1 \end{bmatrix} = \underline{0,2707}$$

$$\bullet \boxed{\alpha=2} \quad \underline{w}_0 = R_x^{-1} \underline{p}_{xd} = \begin{bmatrix} 3,25 & 1,5 \\ 1,5 & 3,25 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1,5 \end{bmatrix} = \begin{bmatrix} -0,2707 \\ 0,5865 \end{bmatrix}$$

$$J(\underline{w}_0) = E\{\hat{s}^2(n-2)\} - \underline{p}_{xd} R_x^{-1} \underline{p}_{xd} = 1 - [0 \ 1,5] R_x^{-1} \begin{bmatrix} 0 \\ 1,5 \end{bmatrix} = \underline{0,1203}$$

Portanto, nesse caso, a melhor estratégia, em termos de MSE, é utilizar o filtro ótimo que recupera o sinal transmitido com atraso 2 (atraso de equalização $\alpha=2$).