

Análise de Similaridade

Alexandre Gomes Caldeira - 96701
Departamento de Engenharia Elétrica,
Universidade Federal de Viçosa, Viçosa - MG
E-mail: alexandre.caldeira@ufv.br

Abstract—Neste relatório são apresentadas as funções de correlação cruzada e da coerência como métodos de análise de similaridade entre sinais no domínio da frequência. Exemplos numéricos computacionais são utilizados para demonstrar as medidas de similaridade através de funções conhecidas e plantas didáticas.

I. INTRODUÇÃO

Aprofundando na análise de sinais e sistemas no domínio da frequência, diante dos trabalhos anteriores em relação a análise, processamento e filtragem de áudios e imagens no domínio da frequência, um tópico importante neste contexto é a análise de similaridade de sinais, principalmente através de correlação cruzada da coerência [1].

A correlação cruzada é uma estatística objetiva que representa a similaridade entre dois sinais mediante um dado número de amostras. Similarmente, ao passo que a coerência tem o mesmo objetivo, se trata de uma medida em função da frequência. Ambas tem ampla aplicação na análise estatística exploratória de dados em sistemas altamente ruidosos como sonares, radares, sistemas biomédicos (ECG, EEG) e engenharia de características em aprendizado de máquinas.

II. OBJETIVOS

Este trabalho tem como objetivo geral a introdução à análise de similaridade de sinais [2]. Em termos específicos, será apresentada uma demonstração numérica e aplicação da correlação cruzada e da coerência no contexto de um sistema simulado de radar. Finalmente, serão comparadas as vantagens e desvantagens do uso de cada tipo de estatística de similaridade.

III. MATERIAIS E MÉTODOS

De forma similar aos trabalho anteriores na disciplina, será empregado o software MATLAB para construção das imagens e respostas dos sistemas estudados neste trabalho.

A. Correlação e Coerência

A correlação cruzada tem definição similar à convolução, dada por:

$$\hat{R}_{x_1, x_2}(k) = x_1(t), x_2(t) = \frac{1}{N-k} \sum_{n=-1}^{N-k} \{x_1[k] \cdot x_2[n+k]\} \quad (1)$$

ao passo que a coerência entre dois sinais é dada por:

$$\hat{\gamma}_{x_1, x_2}^2(f) = \frac{\left| \sum_{i=-1}^M X_i^*(f) \cdot Y_u(f) \right|^2}{\left| \sum_{i=-1}^M X_i^*(f) \right|^2 \cdot \left| \sum_{i=-1}^M Y_i^*(f) \right|^2} \quad (2)$$

Apesar de aparentar idêntica à definição de convolução, a correlação cruzada tem propriedades distintas devido ao sinal invertido no índice k de x_2 , como por exemplo, assumindo $f(t)$ e $g(t)$ como funções de correlação cruzada, tem-se que sua convolução é dada por:

$$[f(t) * g(t)](t) = [\overline{f(-t)} * g(t)](t)$$

porém, assim como no Teorema da Convolução, a correlação cruzada satisfaz a propriedade de que:

$$\mathcal{F}\{f * g\} = \overline{\mathcal{F}\{f\}} \cdot \mathcal{F}\{g\}$$

Assim, a correlação cruzada mostra-se como ferramenta de importante compreensão para análise de similaridade no contexto de Sinais e Sistemas. Neste trabalho serão exemplificadas a correlação cruzada e coerência de sinais entre senóides, cossenóides, ruídos gaussianos e senóides com ruídos.

B. Sonar

A partir da planta de sonar descrita no relatório, será simulada a entrada e saída do sistema, além da estimativa resultante de distância entre o sonar e o objeto proposto, através do código:

```
fs = 255; t = 0:fs; w0 = 3.5; bin = 40;  
alfa = 0.65; alfa2 = 1;  
r = randn(1,256); x = zeros(1,256);  
x(bin) = 20; y = zeros(1,length(x));  
y(33:256) = x(1:224)*alfa + r(33:256)*alfa2;  
y(1:31) = r(1:31)*alfa2;
```

```
figure; subplot(2,2,1); plot(t,x)  
title('Sinal x'); xlabel('Amostras [n]')  
ylabel('x[n]'); axis tight; grid on
```

```
subplot(2,2,3:4); plot(t,y);  
title('Sinal atrasado em 32 amostras')  
xlabel('Amostras [n]'), ylabel('y[n]')  
axis tight; grid on  
subplot(2,2,2); plot(t,r)  
title('Ruido'), xlabel('Amostras [n]')  
ylabel('r[n]'); axis tight; grid on
```

```
N = 256; [Y,lag]=xcorr(y,x); R=Y(1:N); Rrx=fliplr(R);
```

```
figure; plot(lag,Y); axis tight;
grid on; xlabel('Atraso')
title('Correlao entre sinal enviado e recebido')
```

C. Resposta ao impulso

De maneira similar à planta do sonar é proposto um sistema, sem perda de generalidade, ao qual será aplicado um sinal de impulso unitário na entrada. Analisando a saída como resposta a esse impulso, somada de interferências representadas por atrasos, empregaremos a correlação para aproximar qual é a verdadeira resposta do sistema ao impulso, através da média de múltiplas amostragens da resposta.

IV. RESULTADOS E DISCUSSÃO

Diante das definições descritas anteriormente, esta seção apresenta alguns exemplos numéricos da correlação cruzada em funções conhecidas, além de simular a planta de um sonar a nível de aplicação didática da correlação e de estimativa da resposta de um sistema ao impulso unitário.

A. Correlação e Coerência

Aplicando a Eq. 1 referente à correlação, calculou-se em código a correlação cruzada entre funções afim na Figura 1 e similarmente calculou-se a correlação cruzada entre seno e cosseno de mesma frequência apresentados na Figura 2. Nota-se que a frequência da correlação é a mesma dos sinais, e na Figura 3, é apresentado numericamente que a correlação entre x_1, x_2 é o oposto aritmético da correlação entre x_1, x_2 no caso das senóides, mas o contrário ocorre no caso da função afim.

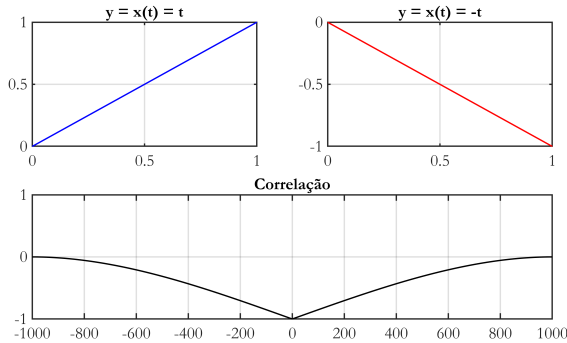


Fig. 1. Correlação entre funções $y = -t$ e $y = t$.

Assim, a correlação se apresenta sensível à fase e periodicidade de sinais. Ainda, calculando a correlação entre dois ruídos gaussianos gerados aleatoriamente, nota-se que esta se apresenta aproximadamente nula na Figura 6 conforme esperado, pois não há similaridade entre os sinais no tempo (apesar de terem sido amostrados de uma mesma distribuição estatística).

De forma similar, na Figura 5, são propostos dois sinais compostos por somas de senóides em diferentes frequências e um ruído gaussiano, como pode ser visto em seu espectro

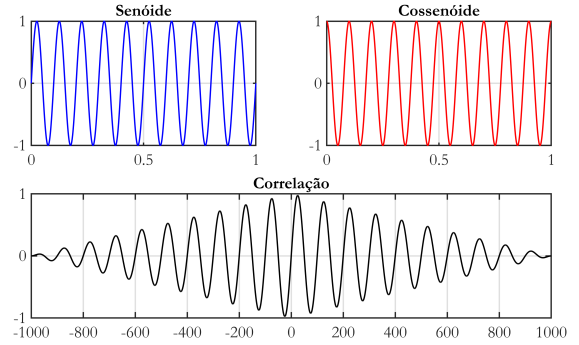


Fig. 2. Correlação entre senóide e cossenóide.

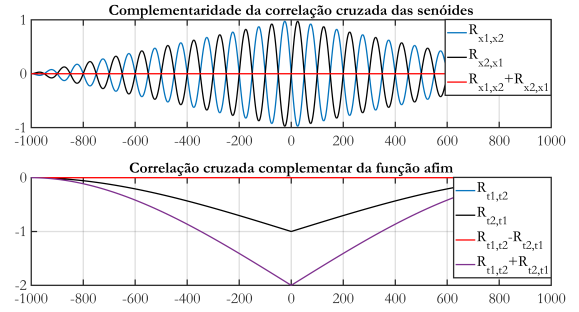


Fig. 3. Comparação entre a complementaridade da correlação de sinais periódicos e não periódicos.

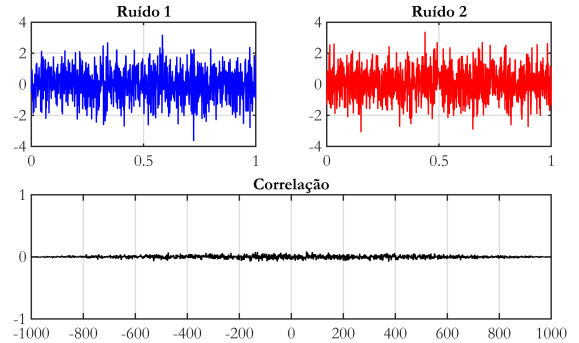


Fig. 4. Correlação entre dois ruídos gaussianos gerados aleatoriamente.

na frequência e amplitude no tempo. Nota-se que a senóide à esquerda possui componentes em 3 Hz e 7 Hz, ao passo que a senóide à direita possui componentes a 7 Hz e 11 Hz. Calculando a coerência entre esses dois sinais, nota-se portanto na Figura 2 um pico na coerência a 7 Hz, onde ambos sinais possuem harmônicas de elevada amplitude, ou seja, onde são similares.

Dessa maneira, ambas a correlação e a coerência se apresentam como ferramentas úteis para o estudo da similaridade entre sinais em termos de frequência, porém possuem características distintas em termos de sensibilidade à variações na fase.

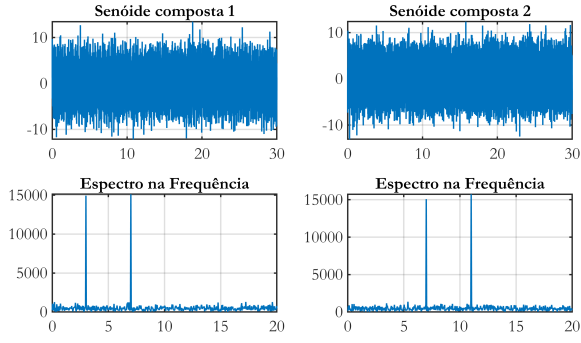


Fig. 5. Espectro de dois sinais compostos por mais de uma senóide em frequências diferentes.

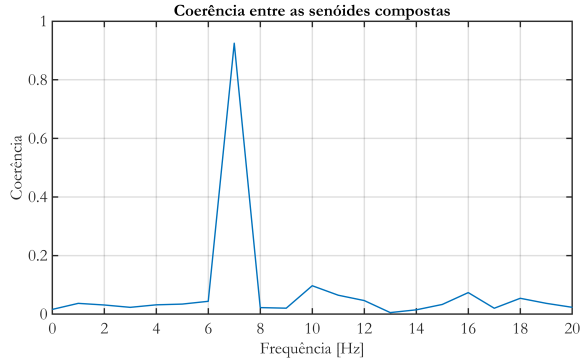


Fig. 6. Coerência calculada a partir dos sinais da Figura ??.

B. Sonar

Simulando a planta do sonar conforme definido, obteve-se a Figura 7, em que um pulso do sonar é somado à um ruído gaussiano e atrasado um número de amostras, correspondendo ao sinal recebido após reflexão no objeto medido. Aferindo então a correlação entre o sinal x de entrada e o sinal atrasado de saída do sistema, obtemos um pico de correlação na amostra de tempo discreto 32, correspondendo *exatamente* ao atraso aplicado manualmente ao sistema. Portanto, a correlação demonstra-se amplamente aplicável para verificação de respostas em sistemas ou ambientes ruidosos.

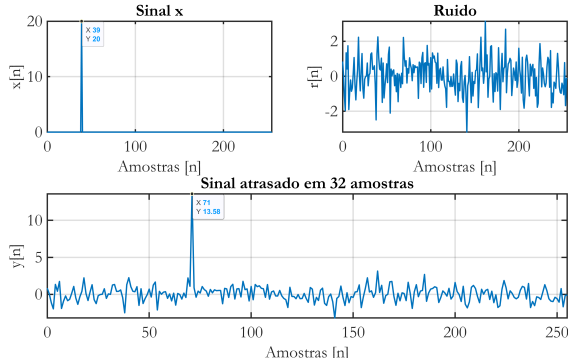


Fig. 7. Simulação da planta de sonar proposta.

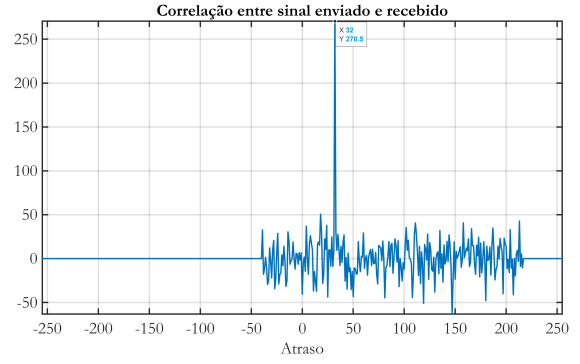


Fig. 8. Estimativa de distância pelo método da correlação cruzada.

C. Resposta ao impulso

Finalmente, aplicando ao sistema delineado no roteiro [2], calculou-se a resposta analiticamente e contruiu-se a resposta média de 150 amostras de resposta ao impulso estimadas pela correlação (de forma idêntica ao que foi feito na seção anterior). Na Figura 9, vê-se que a resposta ao impulso estimada ao fim do processo é visualmente idêntica e numericamente atingiu erro percentual médio de 0.5, com máximo erro absoluto observado de 2% na estimativa resultante. Assim, confirma-se novamente e sem perda de generalidade a aplicabilidade da correlação para estimativa de resposta ao impulso de plantas que podem ser modeladas como Sistemas Lineares Invariantes no Tempo.

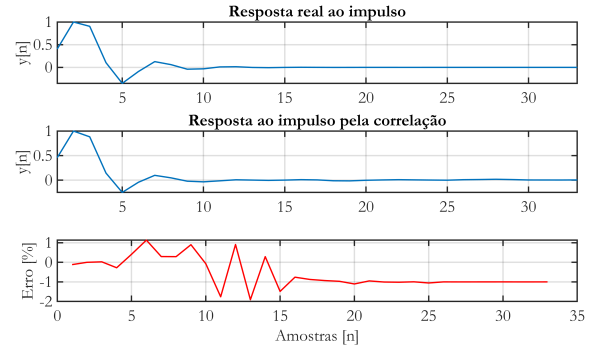


Fig. 9. Estimativa de resposta ao impulso de uma planta, através da média de 150 amostras diferentes da resposta.

V. CONSIDERAÇÕES FINAIS

O presente trabalho introduz os conceitos de análise estatística de similaridade na frequência entre sinais, utilizando exemplos numéricos e didáticos para demonstração da sua aplicabilidade. Mediante definição da correlação e da coerência, suas propriedades são introduzidas e discutidas em termos de aplicações, notando a eficácia destas medidas para identificação de similaridades entre sinais no domínio da frequência. Vale ressaltar que este trabalho não estuda o efeito da relação sinal-ruído na estimativa destas medidas, o que cabe a trabalhos futuros.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] A. V. Oppenheim, *Sinais e sistemas*. Prentice-Hall, 2010.
- [2] L. B. Felix, *Roteiro da Aula Prática 8: Análise de Similaridade*. Departamento de Engenharia Elétrica Centro de Ciências Exatas e Tecnológicas, Universidade Federal de Viçosa Universidade, 2022.