# Anotações sobre Correlação e Análise de Sinais

**CompMus 2019 – Prof. Marcelo Queiroz**

No exemplo CompMus108 vimos como a medida de correlação entre dois sinais com $N$ amostras $s_1$ e $s_2$, definida como

$$\mathrm{corr}(s_1, s_2) = \sum_{i=0}^{N-1} s_1[i]s_2[i],$$

pode servir para identificar componentes senoidais, considerando um dos sinais ($s_1$) como o sinal a ser analisado e o outro ($s_2$) como um sinal de teste senoidal. Essa identificação é possível quando o sinal de análise tem a forma $s_1[i] = \sin(2\pi f \frac{i}{N})$ e o sinal de teste tem a forma $s_2[i] = \sin(2\pi g \frac{i}{N})$, porque nesse caso

$$\begin{aligned}\mathrm{corr}(s_1, s_2) &= \sum_{i=0}^{N-1} s_1[i]s_2[i] = \sum_{i=0}^{N-1} \sin(2\pi f \frac{i}{N}) \sin(2\pi g \frac{i}{N}) \\ &= \sum_{i=0}^{N-1} \frac{1}{2} \cos(2\pi(f-g)\frac{i}{N}) - \frac{1}{2}\cos(2\pi(f+g)\frac{i}{N});\end{aligned}$$

assim, quando $f \neq g$ os dois cossenos terão somas próximas de zero (dentro de cada período do cosseno, cada valor positivo cancela um valor correspondente negativo), ao passo que quando $f = g$ o primeiro cosseno (de frequência $f - g = 0$) é constante $= 1$, enquanto o outro terá uma soma próxima de zero. Disso se conclui que

$$\mathrm{corr}(s_1, s_2) = \begin{cases} \frac{N}{2} & \text{se } f = g \\ \approx 0 & \text{se } f \neq g \end{cases}$$

Isso permite então descobrir automaticamente o conteúdo de um sinal senoidal fazendo uma varredura por todas as frequências possíveis, construindo sinais de teste e calculando correlações. Esse esquema permite até analisar sinais que são misturas de várias componentes senoidais, pois cada componente senoidal só "ativa" a correlação com um sinal de teste que possua a mesma frequência daquela componente. Finalmente, se uma componente senoidal tem uma amplitude genérica $A \in [0, 1]$, sua correlação com um sinal de teste de mesma frequência será $\mathrm{corr}(s_1, s_2) = A\frac{N}{2}$, de onde também é possível detectar as amplitudes das componentes através da expressão $A = \frac{\mathrm{corr}(s_1,s_2)}{N/2}$, de forma independente para cada frequência de teste.

Esse esquema de detecção é razoavelmente fácil de explicar, entender e aplicar, mas ele esconde alguns detalhes e limitações que não são óbvios à primeira vista. Um destes detalhes é o fato de que seria impossível fazer uma varredura de todas as frequências sem conhecer um conjunto *finito* de frequências candidatas, como $g \in \{1, 2, \ldots, 1000\}$ na implementação que fizemos. A solução para essa questão é tecnicamente elaborada e não vamos aprofundá-la nesse momento. O outro detalhe tem a ver com *onde* os sinais senoidais começam suas trajetórias, e isso tem a ver com os experimentos da segunda aula prática.

Na aula prática #2, vimos que o esquema de detecção proposto só funciona porque os dois sinais possuem exatamente aquela expressão do seno, isso é, são componentes senoidais com a mesma *fase inicial*: o mesmo esquema não funciona para uma função senoidal genérica, deslocada para a direita ou para a esquerda no tempo, ou seja, com uma expressão da forma

$$s_1[i] = \sin(2\pi f \frac{i}{N} + \varphi)$$

onde $\varphi \neq 0$ é um ângulo ou fase inicial[1] qualquer. Vimos o exemplo onde $s_1[i] = \cos(2\pi f \frac{i}{N})$ (que é o mesmo que $s_1[i] = \sin(2\pi f \frac{i}{N} + \frac{\pi}{2})$), para o qual todas as correlações construídas com funções da forma $s_2[i] = \sin(2\pi g \frac{i}{N})$ resultavam sempre em zero. Porém, se todas as funções de teste também tivessem a forma de cosseno, então o esquema voltava a funcionar; isso é consequência de uma propriedade parecida com a anterior:

$$\begin{aligned}
\operatorname{corr}(s_1, s_2) &= \sum_{i=0}^{N-1} s_1[i] s_2[i] = \sum_{i=0}^{N-1} \cos(2\pi f \frac{i}{N}) \cos(2\pi g \frac{i}{N}) \\
&= \sum_{i=0}^{N-1} \frac{1}{2} \cos(2\pi (f-g) \frac{i}{N}) + \frac{1}{2} \cos(2\pi (f+g) \frac{i}{N});
\end{aligned}$$

de onde se conclui outra vez que

$$\operatorname{corr}(s_1, s_2) = \begin{cases} \frac{N}{2} & \text{se } f = g \\ \approx 0 & \text{se } f \neq g \end{cases}$$

Usando a notação simplificada $\sin f$ e $\cos f$ para os sinais senoidais e cossenoidais de frequência $f$, e denotando por $g$ uma frequência arbitrária $g \neq f$, podemos combinar todas as propriedades acima na seguinte tabela:

| $s_1$ | $s_2$ | corr |
|---|---|---|
| $\sin f$ | $\sin f$ | $\frac{N}{2}$ |
| $\sin f$ | $\sin g$ | 0 |
| $\sin f$ | $\cos f$ | 0 |
| $\sin f$ | $\cos g$ | 0 |
| $\cos f$ | $\cos g$ | 0 |
| $\cos f$ | $\cos f$ | $\frac{N}{2}$ |

A generalização do método se deu através da constatação de que uma função senoidal genérica sempre pode ser escrita como uma combinação de um seno e um cosseno "puros" (ou seja, com fase inicial zero), por causa da expressão

$$s_1(t) = A \sin(2\pi f t + \varphi) = \alpha \sin(2\pi f t) + \beta \cos(2\pi f t)$$

onde $\alpha = A\cos(\varphi)$ e $\beta = A\sin(\varphi)$. Assim, calcular as correlações de $s_1(t)$ com todas as funções $\sin f$ e $\cos f$ possíveis produz todos os valores de $\alpha$ e $\beta$ (em função de $f$):

$$\alpha(f) = \frac{\operatorname{corr}(s_1, \sin f)}{N/2} \qquad \text{e} \qquad \beta(f) = \frac{\operatorname{corr}(s_1, \cos f)}{N/2}$$

[1] chamado assim porque $s_1[0] = \sin(\varphi)$, ou seja, em $i = 0$ o valor do seno é determinado por $\varphi$.

e consequentemente podem ser obtidos todos os valores de $A = \frac{\sqrt{\alpha^2+\beta^2}}{N/2}$ e $\varphi = \tan^{-1}\left(\frac{\beta}{\alpha}\right)$, também em função da frequência.

## Mas afinal, o que significam essas operações entre os sinais?

O nome *correlação* vem da Estatística e representa uma medida de dependência ou relação linear entre duas fontes de informação (variáveis aleatórias), e é tanto maior quanto mais forte essa dependência entre os valores, sendo máxima quando o conhecimento de um dos valores permite calcular o outro por uma regra de 3 simples, como ocorre quando usamos duas senoides de mesma frequência e fase inicial para gerar duas listas de valores (que serão iguais nesse caso). Nossos sinais senoidais não fornecem um exemplo muito bom de geradores de valores aleatórios, mas essa ideia de duas fontes de informação que estão em "perfeita sintonia" ou "completamente independentes" caracteriza bastante bem os exemplos observados: duas senoides idênticas possuem correlação alta, porém ao se mudar a frequência de uma delas, ou a fase de seno para cosseno, elas perdem a "sintonia" e passam a se correlacionar como fontes independentes de informação.

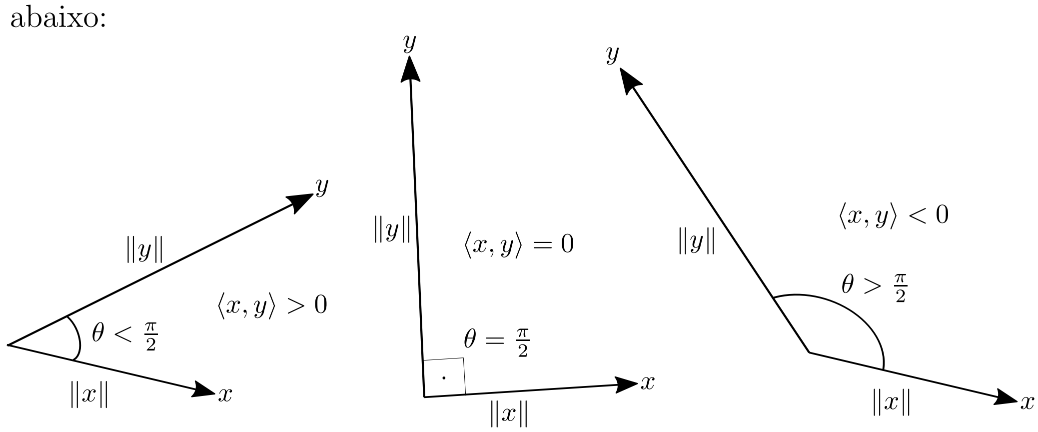
A operação de correlação tem uma outra interpretação, geométrica, que também nos ajuda a compreender os exemplos estudados. A expressão da correlação equivale à operação de *produto interno* entre os vetores, que no caso de vetores de 2 ou 3 dimensões está relacionada à noção de ângulo: dois vetores no plano $x = (x_1, x_2)$ e $y = (y_1, y_2)$ possuem produto interno

$$\langle x, y\rangle = x_1 y_1 + x_2 y_2$$

que satisfaz a propriedade

$$\langle x, y\rangle = \|x\| \cdot \|y\| \cdot \cos(\theta),$$

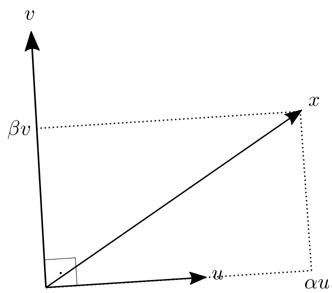
onde $\|x\| = \sqrt{x_1^2 + x_2^2}$ é a *norma* ou tamanho do vetor e $\theta$ é o ângulo formado entre $x$ e $y$, como na figura abaixo:




Em particular, se $x$ e $y$ são *ortogonais* (possuem ângulo $\theta = \frac{\pi}{2}$) então o produto interno entre eles é zero; se por outro lado $x = y$, então o ângulo entre eles é $\theta = 0$ e o produto interno será $\|x\|^2$.

Nossos sinais são vetores de dimensão bastante alta, mas a interpretação geométrica acima também é perfeitamente válida naquele contexto: em particular, verificamos empiricamente (e também formalmente através das contas acima) que senos de frequências diferentes geram vetores *ortogonais*, bem como cossenos de frequências diferentes, e além disso um seno é sempre ortogonal a um cosseno, independentemente de terem frequências iguais ou diferentes. Isso significa que coleções de senos e cossenos formam conjuntos de vetores ortogonais, e coleções de vetores ortogonais servem como *bases* para representar elementos (sinais) arbitrários pertencentes ao mesmo *espaço vetorial* (coleção de sinais de mesma duração).

Voltando a um contexto bidimensional, mais simples, a principal operação para representar um vetor $x = (x_1, x_2)$ dado em uma base ortogonal formada por vetores $u$ e $v$ é a chamada *projeção ortogonal*: essencialmente, presume-se que $x = \alpha u + \beta v$ para algum par de pesos $\alpha$ e $\beta$, e constrói-se os produtos internos dessa identidade com os vetores $u$ e $v$:




$$\langle x, u\rangle = \langle \alpha u + \beta v, u\rangle = \alpha\langle u, u\rangle + \beta\langle v, u\rangle = \alpha\|u\|^2$$

$$\langle x, v\rangle = \langle \alpha u + \beta v, v\rangle = \alpha\langle u, v\rangle + \beta\langle v, v\rangle = \beta\|v\|^2$$

de onde se conclui que os pesos $\alpha$ e $\beta$ são dados pelos produtos internos $\langle x, u\rangle$ e $\langle x, v\rangle$, normalizados respectivamente pelo vetor correspondente da base:

$$\alpha = \frac{\langle x, u\rangle}{\|u\|^2} \qquad \text{e} \qquad \beta = \frac{\langle x, v\rangle}{\|v\|^2}.$$

Veja que essa equação é idêntica às equações no alto da página 3 que definiram $\alpha(f)$ e $\beta(f)$ no método de análise que usamos, considerando $x = s_1$, $u = \sin f$ e $v = \cos f$: basta observar que $\langle x, u\rangle = \langle s_1, \sin f\rangle = \mathrm{corr}(s_1, \sin f)$ e $\|u\|^2 = \|\sin f\|^2 = \frac{N}{2}$ (analogamente para $\langle x, v\rangle$ e $\|v\|^2$).

É possível então afirmar que o método de análise que desenvolvemos empiricamente equivale a representar cada componente senoidal com frequência $f$ do sinal $s_1$ na base formada por um seno e um cosseno puros com a mesma frequência $f$, usando projeção ortogonal, e de forma independente para cada frequência $f$, de maneira geometricamente análoga à situação no plano ilustrada na figura acima, porém em um espaço de dimensão $N$ arbitrária. O nome dessa técnica é *Análise de Fourier*, que é um tema recorrente em Computação Musical, e teremos várias oportunidades de explorar suas interpretações e aplicações ao longo do semestre.

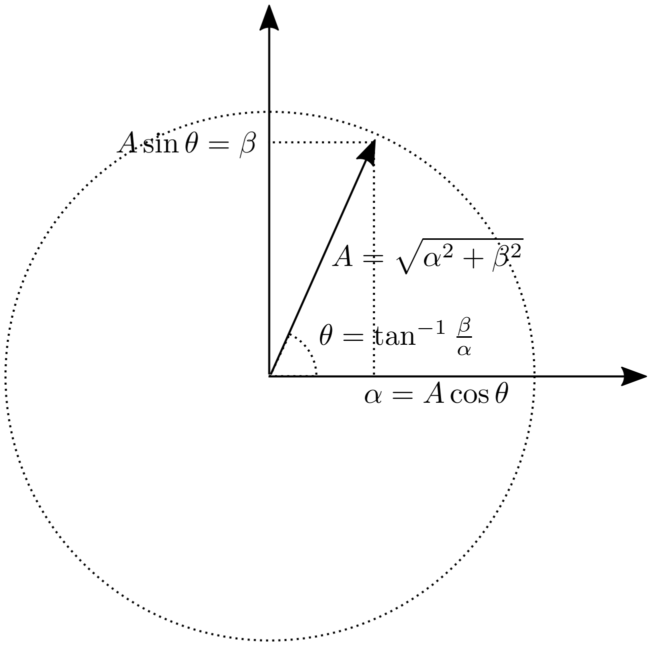
Um último comentário em relação ao método da aula prática #2 diz respeito às representações das componentes senoidais através das fórmulas equivalentes

$$A \sin(2\pi f t + \varphi)$$

e

$$\alpha \sin(2\pi f t) + \beta \cos(2\pi f t);$$

a primeira usa dois parâmetros $A$ e $\varphi$ denominados *amplitude* e *fase inicial* (interpretados como o raio do círculo trigonométrico e o ângulo inicial de giro) enquanto a segunda usa dois parâmetros $\alpha$ e $\beta$ associados respectivamente aos pesos das componentes senoidal pura e cossenoidal pura[2]. As relações entre esses dois pares de parâmetros são as mesmas que definem a conversão entre os sistemas de coordenadas *Cartesiano* e *Polar*, como pode ser visto na figura abaixo.




Essa observação está no cerne da explicação de por que a transformada de Fourier é geralmente definida a partir de uma representação usando uma base de funções exponenciais complexas, cujos pesos (complexos) são diretamente interpretáveis na representação polar como amplitudes e fases, e indiretamente interpretáveis na representação cartesiana como "peso de seno" e "peso do cosseno".

[2]Isso é só um detalhe técnico: na análise de Fourier os valores de $\alpha$ e $\beta$ correspondem aos pesos das componentes cossenoidal pura e senoidal pura, respectivamente; ou seja, os nossos pesos estão trocados. Obteríamos exatamente os mesmos pesos se refizéssemos a derivação do método tomando um sinal $s_1$ genérico em forma de cosseno.