Espectro de potência de um audio & FFT

Eduardo F. Costa, Luiz H. Romero

1 Introdução

A imagem abaixo é o espectro de potência de um sinal de áudio. No eixo horizontal temos as frequências e no eixo vertical, a "magnitude". Podemos perceber que as magnitudes estão concentradas em certas frequências. Qual será o sinal de áudio que tem esse espectro?

Nesse trabalho, você vai aprender mais detalhes sobre como obter esse gráfico, sua relação com mínimos quadrados / regressão linear visto em aula, e como gravar sons e obter seu espectro com o notebook usando Octave ou software similar.

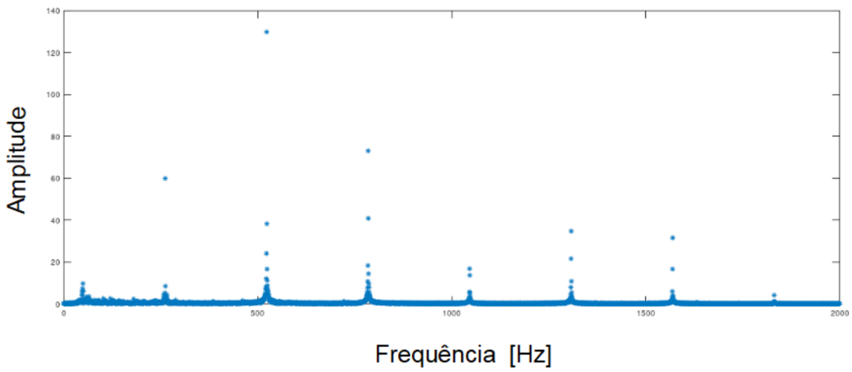

Figura 1: Magnitudes versus frequências (em Hz) de um sinal de áudio.

2 Mínimos quadrados trigonométricos e sua relação com FFT

Nessa primeira parte do trabalho, mais teórica, vamos ver a relação entre a teoria que vimos e a chamada "fast Fourier transform" (FFT). Seja uma base de funções na forma que vimos em aula: $\phi_0 = 1$, $\phi_1 = sin(\frac{2\pi}{T}x)$, $\phi_2 = cos(\frac{2\pi}{T}x)$, $\phi_3 = sin(\frac{2\pi}{T}2x), \ldots$. Sabemos que, para que essa base seja ortogonal no intervalo $[-(T/2) \quad (T/2)]$ no caso discreto, basta tomarmos uma malha igualmente espaçada nesse intervalo e dispensar a que coincide com $(T/2)$. Sendo ortogonal, os coeficientes do polinômio aproximador

$$P_n(x) = a_0\phi_0(x) + a_1\phi_1(x) + \cdots + a_n\phi_n(x)$$

são dados por

$$a_j = \frac{\langle f, \phi_j \rangle}{\langle \phi_j, \phi_j \rangle}$$

onde representamos por f a função que estamos aproximando.

Exercício 1. Seja $T = 10$, f a função degrau unitário, 1000 amostras no intervalo $[-5\ \ 5]$.

a) gere a malha x = -5:(10/1000):5, dispense o último ponto, calcule os valores de f nesses pontos e armazene num vetor y.

b) Calcule $\langle \phi_j, \phi_j \rangle$ para $j = 1, 2, 3, ..., 10$ no computador, e baseado nos resultados descubra a relação com o número de amostras.

c) Calcule $a_0, a_1, \ldots, a_{10}$. No Octave/Matlab, peça f = fft(y), descarte a última metade do vetor f, e compare os coeficientes com os a obtidos. Qual a relação entre eles?

Como vemos pelo Exercício 1, há uma relação muito próxima entre FFT e mínimos quadrados trigonométricos no caso discreto. Os coeficientes a associados aos senos são os mesmos da parte real da FFT, e os a associados aos cossenos são iguais às partes imaginárias da FFT. Assim, são essencialmente a mesma coisa, com formas de escrever um pouco diferentes.

A FFT, no entanto, é mais rápida, pois tem alguns "truques" que aceleram seu cálculo. Por isso, deste ponto em diante, sugerimos usar o comando pronto fft.

3 Quais frequencias estão associadas aos valores obtidos na fft?

Como já notamos no Exercício 1, ao pedir f = fft(y), obtemos um vetor de números complexos, cujos valores estão associados a um polinômio trigonométrico. Por exemplo, digamos que $f(2) = 3 + 4i$. No polinômio aproximador, isso está associado com $3\sin(\frac{2\pi}{T}x) + 4\cos(\frac{2\pi}{T}x)$. Então, concluímos que a frequência associada a esse termo é $f = 1/T$.

Mais um exemplo: $f(3) = 5+6i$ está ligada com $5\sin(\frac{2\pi}{T}2x)+6\cos(\frac{2\pi}{T}2x)$, que por sua vez é um sinal com frequência $f = 2/T$. E assim por diante, cada elemento de fft(y) está associado a uma frequência.

Vale lembrar que $\phi_0 = 1 = cos(0x)$, pode ser interpretado como um sinal de frequência zero. O primeiro termo da fft é sempre um número real, associado e esse ϕ_0 com frequência zero.

Exercício 2. Se tomarmos mil amostras de um sinal entre $x = -10$ e $x = 10$, quais serão as frequências associadas a cada elemento da fft(y)? Note que a resposta não depende do número de amostras, mas somente do "tamanho do intervalo" T.

4 Como obter o espectro de um sinal?

Para fazer o gráfico do espectro, temos que associar uma magnitude a cada frequência. Vamos considerar o mesmo exemplo acima, $3sin(\frac{2\pi}{T}x)+4cos(\frac{2\pi}{T}x)$ e, digamos, $T = 10$. A frequência será $f = 1/10[Hz]$ (assumindo que T é em segundos). Mas temos dois coeficientes, 3 e 4. Para convertermos num só coeficiente, vamos usar a fórmula:

$$b\sin(x) + c\cos(x) = m\sin(x + \phi)$$

em que $m = \sqrt{b^2 + c^2}$. No exemplo, agora, temos a magnitude $m = \sqrt{3^2 + 4^2} = 5$ para a frequência $0.1Hz$.

Juntando tudo: para as magnitudes, pedimos f=fft(y), pegamos apenas a primeira metade de f, tomamos o valor absoluto m=abs(f) (o que já calcula os quadrados e raizes quadradas necessários); para as frequências, definimos um vetor x com os valores de frequência, por exemplo x = 0:(1/T):100, e tracamos o gráfico com plot(x,m,'*'). Precisa tomar o cuidado de que os vetores x e m tem que ter mesma dimensão para fazer o gráfico - pode ser preciso alterar o x.

Exercício 3. Trace o espectro do sinal degrau unitário, tomando mil amostras iniciando em $x = -10$ até $x = 10$. O gráfico deve conter m versus frequências, como o da Fig. 1.

5 Brincando com sua própria voz e outros sons

Exercício 4. Tente "cantar uma só nota" e grave, por exemplo usando o comando record no Octave. Esse comando permite alterar a frequência amostral e a duração. Sugerimos 5 segundos e 16.000 Hz (freq. amostral). Obtenha a figura do espectro da sua voz. Ele provavelmente vai conter picos, análogos aos da Fig. 1, só que, normalmente, os picos estão mais espalhados, mas ainda estão em múltiplos de uma frequência.

Exercício 5. Agora simplesmente grave sua fala. Obtenha a figura do espectro. Agora, as magnitudes devem estar mais espalhadas, sem picos tão evidentes. Compare com o resultado do exercício anterior.

Exercício 6 - opcional. Obtenha o espectro do som de um instrumento musical, tocando uma só nota. Ele se parece com o do Ex. 4?

Exercício 7. Filtre sua fala, retirando todos componentes de frequência acima de 500 Hz. Reproduza sua voz. Para isso, obtenha o sinal através do polinômio interpolador $P_n = a_0 + \cdots + a_N \sin(2\pi 1000x) + a_{N+1}\cos(2\pi 1000x)$, truncado na frequência que corresponde a mil Hz.

Exercício 8. Filtre sua fala, mantendo componentes entre 500Hz e 1000Hz. Reproduza sua voz.

Respondendo a pergunta do início: é um sintetizador, tocando uma só nota de uma sanfona. Vemos a frequência principal da nota em torno de 250 Hz, e outras frequências múltiplos de 250Hz, fazendo o som ser harmonioso. Um bom instrumento tem muitos múltiplos, o que dá uma "profundidade" e define o timbre do instrumento. Um instrumento ruim tem poucos desses harmônicos, deixando o som "raso" e sem timbre/personalidade.

6 Instrução para entrega

Entregue um pdf contendo títuo, data, nome dos integrantes (até 3), uma breve introdução ao assunto, solução dos exercícios e uma conclusão. No caso dos exercícios 7 e 8, explique o que aconteceu - ex., se sua voz ficou mais grave ou aguda, se ainda é possível entender o que falou, se o resultado é compatível com o esperado.

A entrega será no e-disciplinas, bastando que um membro do grupo faça a entrega.