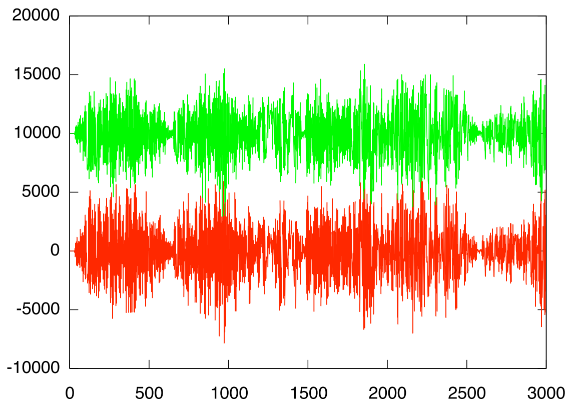
1.ランダムネス(4/27)

1.音声=時間変化する信号

ステレオの音楽の波形を図で描くと次のようになる。



2つのグラフは左右の音を表している。(見易いように片方を上にずらしてあるが、音声信号の時間平均は0)これを、横軸と縦軸に左右の音波の振幅を描くと、次のような図が描ける。

グラフからもわかるように、左右の音はよく似たパターンに見える。そこで、実際どれぐらい似ているか、共分散共分散(相関関数)、相関係数を計算してみる。共分散は、次の式であらわされる。

$$C = \frac{1}{N}\sum_{i=1}^{N}(A_i - \mu_A)(B_i - \mu_B)$$

音声の場合は、平均μは0なので、もっと単純に

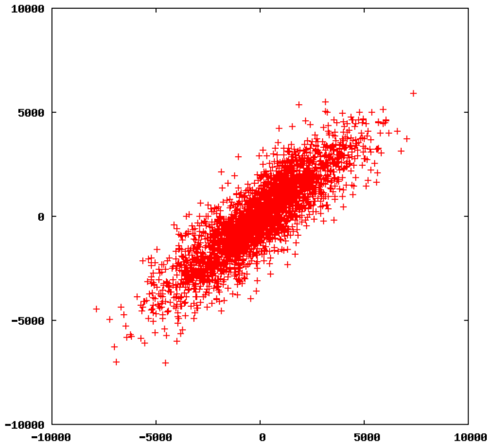
$$C = \frac{1}{N}\sum_{i=1}^{N}A_iB_i$$

と書ける。連続信号なら、平均を積分におきかえて、

$$C = \frac{1}{\tau}\int_0^{\tau}A(t)B(t)\mathrm{d}t$$

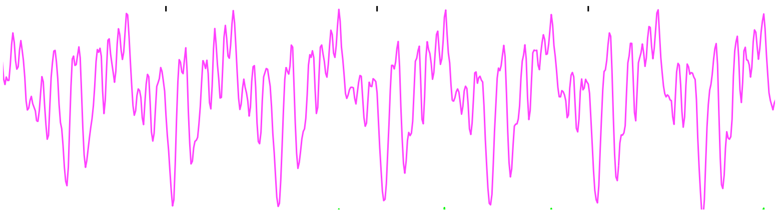
と書けるだろう。(τは音声信号の長さ)

相関係数は共分散をそれぞれの信号の標準偏差で割って規格化したもので、全く同一の信号なら1、逆位相(逆相関)なら–1、線形相関がなければ0になる量である。左右の音声の相関係数は0.86であり、非常に良くにていることがわかる。

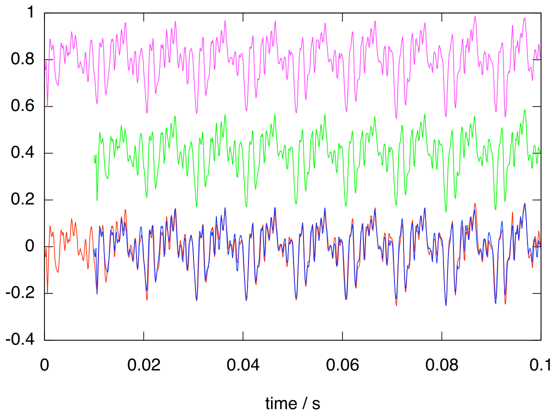


このように、共分散をとると、2つの信号がどれぐらい似ているかがわかる。

次に、音声波形をもっと詳しく観察する。次の波形は、Bachの無伴奏チェロソナタの冒頭の音の一部を拡大したものである。



良く似た形が繰り返しているのがわかる。実際、波形をずらして重ねてみると、0.01秒ずらすとほぼ完全に重なることがわかった。

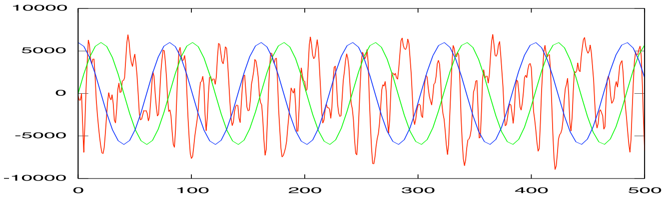


つまり、この音の振動数は100 Hzである。

さて、フーリエは、このような波形を目にして、すべての波が、実はサイン波の重ねあわせで表現できるのではないかと考え、フーリエ変換というアイデアにたどりついたと思われる。

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(t) \exp(-ikt) \mathrm{d}t$$

チェロの、複雑な波形を見て、これがまるっこいサイン波の重ねあわせで表現できるとは想像しがたい。(下図は、100 Hzのサイン波、コサイン波を重ねてみたもの)



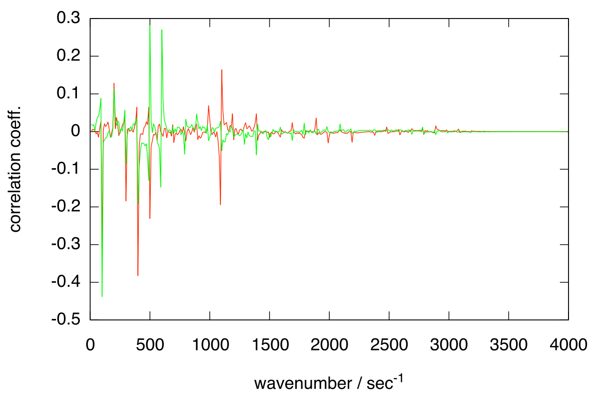
そこで、実際に、チェロの音と、100 Hzの正弦波がどれぐらい似ているかを、共分散で評価してみる。音声信号をf(t)とし、積分範囲や前の定数はとりあえずてきとうに決めるとして、

$$C_s(\omega) = \int_0^\tau f(t)\sin\omega t\mathrm{d}t$$

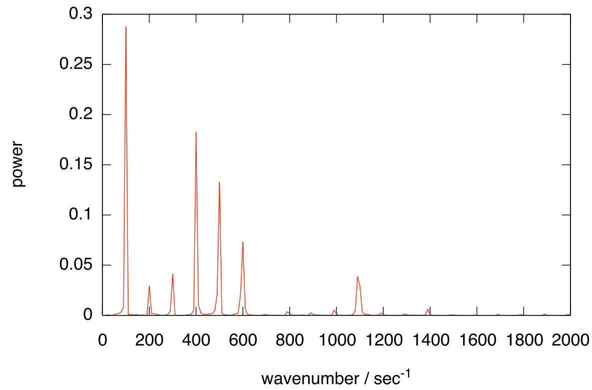
$$C_c(\omega) = \int_0^\tau f(t)\cos\omega t\mathrm{d}t$$

こんな風に書ける。100Hzのサイン波との相関係数は–0.3、コサイン波との相関係数は–0.4と求まった。たしかに一定の相関があることがわかる。同じように、正弦波の振動数ωをいろいろ変えてみて、どんな波とだと共分散が大きくなるかを見ると、次のようなグラフが描ける。

赤はサイン波との共分散、緑はコサイン波との共分散である。これを見ると、確かに100Hzに大きなピークがあり、共分散で評価する限り、チェロの波形は100 Hzの正弦波と似ていると言える。また、ほかにも倍音にあたる400 Hzや、共鳴する500 Hz、600 Hzに大きなピークがあること、一方で110 Hzや170 Hzといった、倍数以外の振動はほとんど含まれていないことがわかる。サインとコサインを分けると面倒なので、両者の二乗和をとってプロットすると次のグラフのようになる。



このグラフの縦軸は波の振幅の二乗に比例する量であり、波のエネルギーに比例することから、パワースペクトルと呼ばれている。分光装置で観測されるスペクトルに相当する。



さて、フーリエ変換とは何だろうか。関数f(t)のフーリエ変換の式は次のように書ける。

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)\exp(-i\omega t)\mathrm{d}t$$

虚数の指数関数の部分は以下のように書換えることができる。

$$\exp(-i\omega t) = \cos\omega t - i\sin\omega t$$

何のことはない。フーリエ変換というのは、もとの関数とサイン波の共分散だった!

2.フーリエ変換の性質

逆フーリエ変換は次のように定義される。

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\omega)\exp(i\omega t)\mathrm{d}\omega$$

フーリエ変換の式とほとんどそっくりに見えるが、指数関数の中の符号が違う。そして、こちらは時間ではなく振動数で積分している。このように、もとの関数とフーリエ変換された関数の間で、自由に変換逆変換をすることができる。なお、積分の前の定数や、指数関数の中身にかかる係数は、本によって定義が違う(統計力学と数学と情報科学で異なる)ので、フーリエ変換関連の公式を調べる場合は、フーリエ変換と逆変換がどのように定義されるかを先にチェックすること。

いくつかの代表的な関数のフーリエ変換を表に示す。(松本が自分で計算したものは、間違っている可能性があるので注意!)

$f(t)$	$\hat{f}(\omega)$	コメント
$f(t-a)$	$\exp(-i\omega a)\hat{f}(\omega)$	
$af(t)$	$a\hat{f}(\omega)$	
$f(t)+g(t)$	$\hat{f}(\omega)+\hat{g}(\omega)$	

$\delta(t)$	1	
1	$2\pi\delta(\omega)$	
$\cos at$	$\delta(\omega - a)\,2\pi$	
$\sin at$	$-i\delta(\omega - a)\,2\pi$	
$\exp(-a\|x\|)$	$\dfrac{2a}{a^2 + \omega^2}$	
$\exp(-ax^2)$	$\dfrac{\sqrt{\pi}\exp\left(-\frac{k^2}{4a}\right)}{2\sqrt{a}}$	不確定性原理*1
$\dfrac{\mathrm{d}^n}{\mathrm{d}t^n} f(t)$	$(i\omega)^n \hat{f}(\omega)$	導関数のフーリエ変換*2
$\displaystyle\int_{-\infty}^{\infty} f(s)g(t+s)\mathrm{d}s$	$\hat{f}(-\omega)\hat{g}(\omega)$	たたみこみ定理*3
$\displaystyle\int_{-\infty}^{\infty} f(s)f(t+s)\mathrm{d}s$	$\|\hat{f}(\omega)\|^2$	Wiener-Khinchinの定理*4

*1 ガウス分布関数と見比べると、左側は分散1/2aのガウス分布、右側は分散2aのガウス分布であることがわかる。つまり、aをどのように選んでも、両者の分散の積は1になる。左側を座標、右側を運動量とみれば、座標のゆらぎを小さくするほど、運動量の不確定性が増えることを意味している。

*2 ある関数を微分することは、そのフーリエ変換に係数を掛けることと等価である。この関係を使うと、積分や微分が極めて容易に計算できる。また、nを小数や負数にすることで、小数回微分や積分も簡単に定義できる。

*3 左側のような形の積分をたたみ込み積分(convolution)と呼ぶ。たたみ込みは、相関関数や信号処理で頻繁に目にする積分である。たとえば、エコーの重ねあわせ、フォーカス外れによるボケ、手振れによる画像のずれを計算したり、逆にそれを取り除く計算はすべてたたみこみ積分である。(“deconvolution”というキーワードで検索してみよ) たたみ込み積分は、フーリエ変換してしまえば、ただの積になるので、積分が不要になる。

*4 自己相関関数をフーリエ変換すると、パワースペクトルが得られる。

*1〜*4は自分で導出できるように練習しておくこと。