

目次

- 離散信号
- 標本化とエイリアシング
- 離散フーリエ変換
- 高速フーリエ変換
- [発展] 直交展開としての複素フーリエ級数展開, 離散フーリエ変換

★7 パターン情報の成分分析(3) — 離散信号の分析

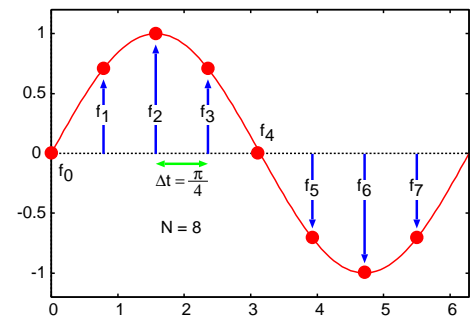
「パターン情報の成分分析」(1)ではベクトル(離散信号)で表されるパターンの直交展開の一般論と応用を学び, (2)では関数(アナログ信号)として表されるパターンの直交展開の例としてフーリエ級数展開について学んだ。今回は, アナログ信号を離散化する話からはじめて, 離散信号に対してもフーリエ級数展開と同様のこと(正弦波による成分分析)ができることを学ぶ。

★7.1 離散信号

アナログ信号を適当な標本化間隔で標本化すると, 変数が離散化された**離散信号**が得られる。N点から成る(N個の値からできた)離散信号は, N個の値をならべたN次元ベクトルとみなして, 次のように表すことができる。

$$f = (f_0, f_1, \dots, f_{N-1}) \quad (1)$$

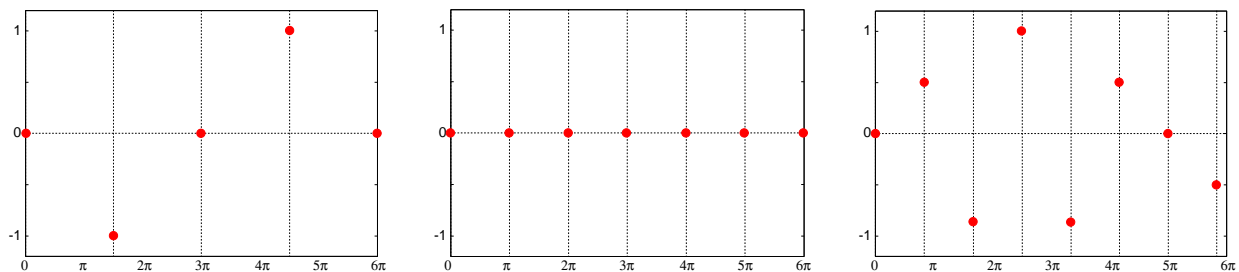
ベクトルの表記としてよく用いられるものとは違い, 添字を0からはじめていることに注意。



★7.2 標本化とエイリアシング

アナログ信号を標本化する際には, 標本化間隔の決め方が重要となる。

Q1. 下図は, ある正弦波をそれぞれ異なる標本化間隔(左から順に $\frac{3}{2}\pi, \pi, \frac{5}{6}\pi$)で標本化したものである。これらの点をつなぐ曲線を描いて, 標本化される前の正弦波がどのようなものだったか予想してみなさい。



Q1の結果から, 元の信号の周期に比べて標本化周期を大きくしすぎる(標本化周波数を低くしすぎる)と, 元の信号の性質が失われてしまうことがわかる。また, 正弦波の場合, 元の信号の周期の半分以下の標本化間隔で標本化すると, 元の正弦波を復元できそうなことがわかる。

ほげお: 信号を標本化したいけど, あまり標本化間隔を大きくとるとまずいみたい. かと
 といって間隔を小さくするとデータ量が多くなって大変. 標本化間隔はどうやって
 決めたらいい?

ふがよ: さっきの周期 2π の正弦波の場合, 標本化間隔が π より小さければこの信号の周
 期が 2π であることがわかりそう. だから, 標本化間隔は元の信号の周期の半分よ
 り小さくすればええんちゃう?

ほげお: でも正弦波じゃない一般の信号のときは?

ふがよ: むー

へなこ: フーリエ級数展開 (フーリエ変換) 考えたらどんな信号も正弦波の重ね合わせ
 なんやから, その中でいちばん周期の短い (周波数の高い) 正弦波が再現できるよ
 うにしたら大丈夫ちゃう?

ほげお&ふがよ: おおー!

ちうわけで, 標本化周期 (周波数) の決め方に関して, 次の定理が知られている.

標本化定理: 信号に含まれる有効な成分の中で最高の周波数が f である場合,
 標本化周波数を $2f$ 以上にして標本化すれば, 標本化した信号から元の信号を再
 現できる (☆1)(☆2).

この定理より, 例えば, ある信号に 500Hz より高い周波数の成分が含まれて
 いないことがわかっている場合, その信号を標本化するときの標本化周波数は
 1000Hz 以上に (標本化周期を 1/1000 秒以下に) すればよいことがわかる. 以
 前, CD の規格では標本化周波数が $44.1\text{kHz} = 44100\text{Hz}$ と定められていること
 を紹介したが, これは, 人間に聴こえる音の周波数が最高 20kHz 程度だからで
 ある. また, 人の会話音声は 4000Hz 以下の成分のみで十分に伝えられるので,
 電話の標本化周波数は 8000Hz とされることが多い (☆3).

Q1 でみたように, 標本化定理の教えを守らずに低い標本化周波数で標本化す
 ると, 本来の信号には存在しない低い周波数成分が現れてしまう. この現象を**エイ
 リアシング**という (☆4)

★ 7.3 離散フーリエ変換

周期的なアナログ信号が複素フーリエ級数展開できたように, 離散信号 (ベク
 トル) は離散フーリエ変換できる.

N 点から成る離散信号 $\mathbf{f} = (f_0, f_1, \dots, f_{N-1})$ は次式のように表せる (☆5).

$$\mathbf{f} = \sum_{k=0}^{N-1} C_k \mathbf{e}_k \quad (3)$$

ただし, $\mathbf{e}_k (k = 0, 1, \dots, N-1)$ は,

$$\mathbf{e}_k = \left(e^{i\omega_0 k \times 0}, e^{i\omega_0 k \times 1}, \dots, e^{i\omega_0 k \times (N-1)} \right) \quad (4)$$

と定義する. $\omega_0 = \frac{2\pi}{N}$ である. フーリエ係数 C_k は次式で与えられる (☆6).

$$C_k = \frac{1}{N} \sum_{t=0}^{N-1} f_t e^{-i\omega_0 kt} \quad (k = 0, 1, \dots, N-1) \quad (5)$$

ここで, \mathbf{e}_k は周期 $\frac{N}{k}$ の正弦波成分である (☆7). 例として下図に $N = 16$ の
 場合の $\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2$ の各要素の実部と虚部を描いたものを示す (見やすいように曲
 線を描いているが, 離散信号であるから本来は●の値のみから成ることに注意).

☆1) 標本化のことをサンプリングともいうので, **サンプリング定理**とも言う. C. Shannon によって 1949 年に証明された. Shannon は, 1937 年に電気回路で論理演算 (AND, OR, etc., ...) ができることを示した人でもあり, 情報理論の創始者でもある.

☆2) 周波数が f より高い成分が含まれていても, それらが雑音などの無効な成分である場合には, f より高い周波数成分をカットするローパスフィルタ (いずれ説明します) をかけてから標本化周波数 $2f$ で標本化すればよい.

☆3) なので, 高い周波数成分から成る虫の鳴き声は電話越しに聴こえなかったりする.

☆4) エイリアシング: aliasing. こんなときにも…「馬車の車輪が回転する様子をビデオカメラで撮影」, 「蛍光灯の下で扇風機の羽根が回るのを観察」

☆5) \mathbf{f} の要素毎に表すと

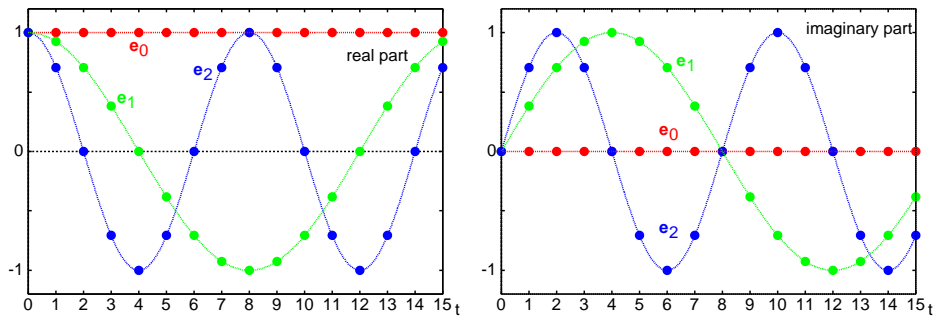
$$f_t = \sum_{k=0}^{N-1} C_k e^{i\omega_0 kt} \quad (2)$$

($t = 0, 1, \dots, N-1$).

☆6) 式 (5) は, \mathbf{f} と \mathbf{e}_k の内積 ($C_k = \mathbf{f} \cdot \mathbf{e}_k$) とみなせる. 詳しくは後ろの [発展] 参照.

☆7) $e^{i\omega_0 kt} = e^{i\frac{2\pi}{N} kt}$ の実部は $\cos \frac{2\pi}{N} kt$, 虚部は $\sin \frac{2\pi}{N} kt$.

式 (3) は、任意の離散信号 f は、定数の成分 $e_0 = (1, 1, \dots, 1)$ 、周期 N の成分 e_1 、周期 $\frac{N}{2}$ の成分 e_2, \dots 、周期 $\frac{N}{N-1}$ の成分 e_{N-1} の重ね合わせで表せるということを意味している。このように、 N 点から成る離散信号を表すには N 個の成分で済む (☆8)。



上記の中で式 (5) に注目すると、これは、 N 個の値 f_0, f_1, \dots, f_{N-1} を N 個の複素数 C_0, C_1, \dots, C_{N-1} に変換する操作を表しているとみなせる。この変換のことを **離散フーリエ変換 (DFT)** (☆9) と呼ぶ。一方、式 (2) はその逆で $\{C_k\}$ から $\{f_t\}$ を求める変換になっているので、**逆離散フーリエ変換 (IDFT)** と呼ばれる。

$$\{f_0, f_1, \dots, f_{N-1}\} \begin{matrix} \xrightarrow{\text{DFT}} \\ \xleftarrow{\text{IDFT}} \end{matrix} \{C_0, C_1, \dots, C_{N-1}\}$$

フーリエ級数展開の場合と同様に、離散信号を DFT して振幅スペクトル $\{|C_k|\}$ や位相スペクトル $\{\arg C_k\}$ を描けば、その信号がどんな成分をどれくらい含んでいるかを調べることができる。前回の資料に示した音響信号のスペクトルは、その例である。

- ☆8) アナログ信号を (誤差なしで) 表すには無限個の成分が必要だった。
- ☆9) 離散フーリエ変換: Discrete Fourier Transform. 逆離散フーリエ変換: Inverse DFT.

- Q2. $N = 4$ として式 (4) の e_k を求めなさい。
- Q3. $f = (4, 4, 0, 0)$ を離散フーリエ変換してフーリエ係数 C_k を求めなさい。
- Q4. 上で求めた $\{C_k\}$ の逆離散フーリエ変換 $\sum_{k=0}^{N-1} C_k e_k$ を求め、これが f に一致することを確かめなさい。
- Q5. $g = (0, 1, 0, 2)$ に対して Q3,4 と同じことをしなさい。

★7.4 高速フーリエ変換

式 (5) の計算を行えば DFT ができる。しかし、このように定義通りに計算を行なうアルゴリズムでは時間計算量のオーダーが N^2 となるので、実用に耐えない (N が 2 倍になったら計算に要する時間が 4 倍になる (☆10))。そこで、DFT の計算を高速に実行できるアルゴリズムが開発されている。これを **高速フーリエ変換 (FFT)** という。FFT の時間計算量のオーダーは $N \log_2 N$ である。FFT のアルゴリズムには、「 N が 2 のべき乗でなければならない」という制約がある (☆12) が、それでも非常に高速である。

例えば、 $N = 44100$ (CD の信号 1 チャンネルの 1 秒分に相当) のデータを FFT したい場合、データ数を 2 のべき乗にするため、データに余分な 0 を付け加えて $N = 65536 = 2^{16}$ とする。こうするとデータ数がずいぶん大きくなるので、元のデータを DFT した方がよさそうに思えるかもしれないが、 $44100^2 = 19$ 億に対して $65536 \times 16 = 100$ 万なので、データ数を増やしても FFT した方が圧倒的に高速に計算できる。

離散フーリエ変換のような技術が幅広く実用されるようになったのは、FFT のおかげと言ってもよい。

- ☆10) DFT に要する乗算回数はおおよそ $2N^2$ であり、 $N = 44100$ なら約 39 億回になる。
- ☆11) 高速フーリエ変換: Fast Fourier Transform.
- ☆12) 「 N が 2 のべき乗」という制約は実は少し緩められる。

★ 7.5 [発展] 直交展開としての複素フーリエ級数展開, 離散フーリエ変換

実フーリエ級数展開の場合 (☆13) と同様に, 複素フーリエ級数展開や離散フーリエ変換も正規直交基底による展開として考えることができる. ☆13) 第6回資料参照.

★ 7.5.1 複素フーリエ級数展開

周期 T の関数 $x(t)$ と $y(t)$ (☆14) の内積を

☆14) $x(t)$ も $y(t)$ も実変数複素数値関数.

$$(x(t), y(t)) = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) \overline{y(t)} dt \quad (6)$$

と定義する. $\omega_0 = \frac{2\pi}{T}$ とおく. 無限個の関数から成る集合 $\{\dots, e^{-2i\omega_0 t}, e^{-i\omega_0 t}, 1, e^{i\omega_0 t}, e^{2i\omega_0 t}, \dots\}$ はこの内積の定義のもとで正規直交基底を成し, 周期 T の関数 $f(t)$ を

$$f(t) = \sum_{k=-\infty}^{\infty} C_k e^{ikt} = \dots + C_{-1} e^{-i\omega_0 t} + C_0 + C_1 e^{i\omega_0 t} + C_2 e^{2i\omega_0 t} + \dots \quad (7)$$

と展開することができる. このとき, 係数 C_k は,

$$C_k = (f(t), e^{ikt}) = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) \overline{e^{i\omega_0 kt}} dt = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) e^{-i\omega_0 kt} dt \quad (8)$$

で求まる. すなわち, 複素フーリエ級数展開とは, 上記の内積の定義のもとで, 周期関数 $f(t)$ を正規直交基底 $\{\dots, e^{-2i\omega_0 t}, e^{-i\omega_0 t}, 1, e^{i\omega_0 t}, e^{2i\omega_0 t}, \dots\}$ を用いて展開する計算であり, フーリエ係数 C_k はその展開係数である.

この説明からわかるように, 関数の空間は無限次元と考えられる (☆15).

☆15) この辺のちゃんとしたことは, 数理情報学科4年次の数学系科目で学べるかも.

★ 7.5.2 離散フーリエ変換

複素ベクトル $\mathbf{x} = (x_0, x_1, \dots, x_{N-1})$ と $\mathbf{y} = (y_0, y_1, \dots, y_{N-1})$ の内積を

$$\mathbf{x} \cdot \mathbf{y} = \frac{1}{N} \sum_{t=0}^{N-1} x_t \overline{y_t} \quad (9)$$

と定義する (\bar{z} は z の複素共役). $\mathbf{e}_k = (e^{i\omega_0 k \times 0}, e^{i\omega_0 k \times 1}, \dots, e^{i\omega_0 k \times (N-1)})$ とおく ($\omega_0 = \frac{2\pi}{N}$) と, $\{\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_{N-1}\}$ はこの内積の定義のもとで正規直交基底 (☆16) であり, 離散フーリエ変換の式 (2) は

☆16) $i \neq j$ ならば $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ であり, $i = j$ ならば $\mathbf{e}_i \cdot \mathbf{e}_j = \|\mathbf{e}_i\|^2 = 1$ となる.

$$\mathbf{f} = \sum_{k=0}^{N-1} C_k \mathbf{e}_k = C_0 \mathbf{e}_0 + C_1 \mathbf{e}_1 + \dots + C_{N-1} \mathbf{e}_{N-1} \quad (10)$$

と表せる. この式は正規直交基底による展開であるから, 係数 C_k ($k = 0, 1, \dots, N-1$) は, $C_k = \mathbf{f} \cdot \mathbf{e}_k$ で求まる. すなわち, 離散フーリエ変換とは, 上記の内積の定義のもとで, 離散信号 \mathbf{f} を正規直交基底 $\{\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_{N-1}\}$ を用いて展開する計算であり, フーリエ係数 C_k はその展開係数である.