

図形情報のスペクトル解析と エントロピー解析

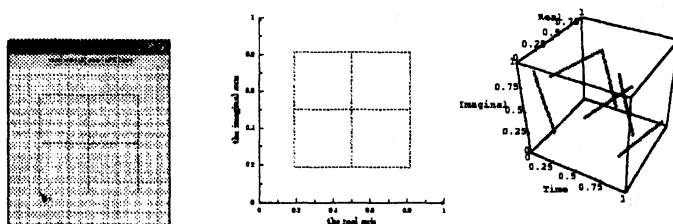
上坂吉則 (山梨英和大学), 塚田 真 (東邦大学)

1 序論

ここで扱う図形は、2次元平面に描かれた線図形のことで、

- (1) 図形は幾つかの一筆書きできる部分図形が重なって出来ている
- (2) これらの部分図形はある順序で描かれる

という性質を持つものを考える。従って、図形は t_{start} を書き始めの時刻、 t_{end} を書き終わりの時刻として、区間 $[t_{\text{start}}, t_{\text{end}}]$ で定義され複素平面 $\mathbb{C}$ に値をとる関数と見なすことができる。



上のような図形の最も代表的なものは文字 (特に漢字やひらがな、カタカナ) である。我々は楷書で書かれた文字と行書或いは草書で書かれた同じ文字を、同じ文字であると認識することができる。楷書の文字と行書或いは草書の文字を比べると、前者は1画1画の微妙な位置、長さ、傾き、跳ねや留めも正確に書かれるが、後者はそうではない。しかし行書或いは草書で文字を書く場合、書き手は楷書の文字の筆順に従って書いているものと思われる。行書或いは草書の文字は一般的に緩やかで連続した曲線で構成されている。行書或いは草書で文字は楷書の文字に比べて高周波成分 (不連続点や角度が急激に変化する点) が少ないとも言える。このことは、楷書で書かれた文字を Fourier 展開して周波数分析を行い、高周波成分を取り除けば行書或いは草書で書いたのと似たような文字が得られるのではないかと予想ができる。

2 三角級数による展開

図形 $z : [0, 1] \rightarrow \mathbb{C}$ は可測関数で2乗可積分、即ち

$$\int_0^1 |z(t)|^2 dt < \infty$$

であるとする。このような図形の全体は Hilbert 空間 $L^2[0, 1]$ をなす。この Hilbert 空間の正規直交基底として最もなじみ深いものは

$$e_k(t) = e^{2\pi i k t}, \quad (0 \leq t \leq 1; k \in \mathbb{Z})$$

で定義される関数族 $\{e_k\}_{k \in \mathbb{Z}}$ である。 z の離散 Fourier 変換は

$$\hat{z}(k) = \int_0^1 e^{-2\pi i k t} z(t) dt, \quad (k \in \mathbb{Z})$$

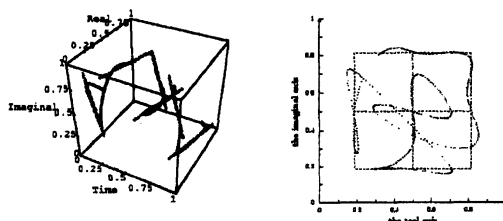
で定義され、

$$z(t) = \sum_{k \in \mathbb{Z}} \hat{z}(k) e^{2\pi i k t}$$

が 2 乗平均収束の意味で成立する (Fourier 展開)。

$$z_m(t) = \sum_{k=-m}^m \hat{z}(k) e^{2\pi i k t}$$

が、ローパスフィルタを通した関数である。これで得られる図形は始点と終点が繋がった閉曲線となる。



z そのものではなく z の微分 z' をローパスフィルタを通して得られる z'_m に対して、

$$z(0) + \int_0^t z'_m(t) dt$$

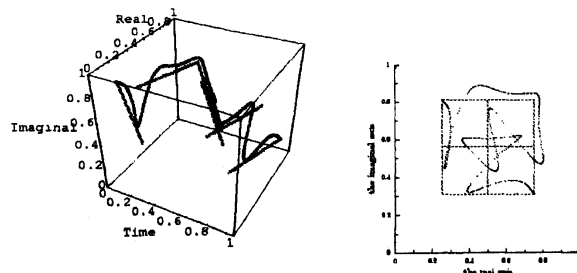
を考える。 z'_m は入力図形と始点および終点は等しくなる。実際、出力される図形の始点と終点の変位は、

$$z(0) + \int_0^1 z'_m(t) dt - z(0) = \int_0^1 z'_m(t) dt$$

であるが、

$$\begin{aligned} \int_0^1 z'_m(t) dt &= \int_0^1 \sum_{k=-m}^m \hat{z}'(k) e^{2\pi i k t} dt \\ &= \sum_{k=-m}^m \hat{z}'(k) \int_0^1 e^{2\pi i k t} dt \\ &= \hat{z}'(0) \\ &= \int_0^1 z'(t) dt \\ &= z(1) - z(0) \end{aligned}$$

となり、入力図形の始点と終点の変位に等しくなる (正規直交基底に定値関数 1 を含んでいることが本質的に効いている)。



z の微分 z' の離散 Fourier 変換は

$$\begin{aligned}\hat{z}'(k) &= \int_0^1 e^{-2\pi i k t} z'(t) dt \\ &= [e^{-2\pi i k t} z(t)]_0^1 + 2\pi i k \int_0^1 e^{-2\pi i k t} z(t) dt \\ &= z(1) - z(0) + 2\pi i k \hat{z}(k)\end{aligned}$$

であるので、 z' をローパスフィルタを通したものを積分すると、

$$\int_0^t z'_m(t) dt = \int_0^t \sum_{k=-m}^m \hat{z}'(k) e^{2\pi i k t} dt = \sum_{k=-m}^m \hat{z}'(k) \int_0^t e^{2\pi i k t} dt$$

であり、ここで

$$\int_0^t e^{2\pi i k t} dt = \begin{cases} t, & (k=0), \\ \frac{1}{2\pi i k} e^{2\pi i k t}, & (k \neq 0) \end{cases}$$

であるから

$$\int_0^t z'_m(t) dt = (z(1) - z(0))t + (z(1) - z(0)) \sum_{|k|=1}^m \frac{e^{2\pi i k t}}{2\pi i k} + \sum_{k=-m}^m \hat{z}(k) e^{2\pi i k t} - \int_0^1 z(t) dt$$

を得る。従って、得られた図形は

$$z(0) + (z(1) - z(0))t + (z(1) - z(0)) \sum_{|k|=1}^m \frac{e^{2\pi i k t}}{2\pi i k} + z_m(t) - \int_0^1 z(t) dt \quad \cdots (*)$$

で表される関数である。この関数の意味を考えてみる。今、

$$w(t) = z(t) - (z(1) - z(0))t - z(0)$$

で定義される関数を考える。この関数は、 $w(0) = w(1) = 0$ を満たす。 w の離散 Fourier 変換は

$$\hat{w}(k) = \hat{z}(k) - (z(1) - z(0)) \int_0^1 e^{-2\pi i k t} t dt - z(0) \int_0^1 e^{-2\pi i k t} dt, \quad (k \in \mathbb{Z})$$

であるが、

$$\hat{w}(0) = \hat{z}(0) - \frac{z(1) - z(0)}{2} - z(0) = \hat{z}(0) - \frac{z(1) + z(0)}{2}$$

であり、 $k \neq 0$ のときは

$$\int_0^1 e^{-2\pi i k t} t dt = \left[-\frac{e^{-2\pi i k t}}{2\pi i k} t \right]_0^1 + \frac{1}{2\pi i k} \int_0^1 e^{-2\pi i k t} dt = -\frac{1}{2\pi i k}$$

より

$$\hat{w}(k) = \hat{z}(k) + \frac{z(1) - z(0)}{2\pi i k}$$

となる。よって、 w をローパスフィルタに通したものは

$$w_m(t) = z_m(t) + (z(1) - z(0)) \sum_{|k|=1}^m \frac{e^{2\pi i k t}}{2\pi i k} - \frac{z(1) + z(0)}{2}$$

であり、これに最初の変換の逆を施せば

$$(z(1) - z(0))t + z(0) + z_m(t) + (z(1) - z(0)) \sum_{|k|=1}^m \frac{e^{2\pi i k t}}{2\pi i k} - \frac{z(1) + z(0)}{2} \quad \dots (**)$$

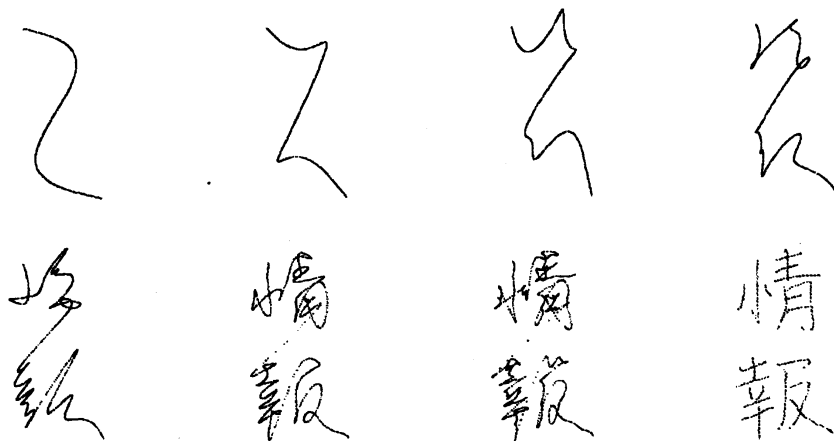
が得られる。これは、(*) の式の定数

$$\int_0^1 z(t) dt$$

(これは図形の重心) が、始点と終点の中点

$$\frac{z(1) + z(0)}{2}$$

に置き換わっただけのものである。



3 一般の関数系による展開

前節では関数 z を或いはそれを変換したものを関数族 $\{e^{-2\pi i k t}\}_{k=-m}^m$ が張る部分空間への直交射影 (最小 2 乗近似) を計算していた訳であるが、関数族は何も正規直交族である必要はない。例えば、多項式系

$$e_m(t) = t^m, \quad m = 0, 1, 2, \dots$$

や、Haar 関数系

$$e_{2^m+k}(t) = \begin{cases} 1, & \frac{k}{2^m} \leq t < \frac{k+1}{2^m} \\ 0, & \text{otherwise} \end{cases} \quad (m = 0, 1, 2, \dots; k = 0, 1, \dots, 2^m - 1)$$

などが考えられる。一般に関数系 $\{e_0, e_1, \dots, e_{m-1}\} \subseteq L^2[0, 1]$ が与えられたとき

$$z \mapsto (\langle e_0 | z \rangle, \langle e_1 | z \rangle, \dots, \langle e_{m-1} | z \rangle)$$

で定義される線形写像 $A : L^2[0, 1] \rightarrow \mathbb{C}^m$ の一般化逆写像 $A^\dagger : \mathbb{C}^m \rightarrow L^2[0, 1]$ を考えると、 $A^\dagger A : L^2[0, 1] \rightarrow L^2[0, 1]$ は関数族 $\{e_0, e_1, \dots, e_{m-1}\}$ が生成する部分空間上への直交射影となる。

実際の計算では次のように行う。一定間隔の時刻

$$t_{\text{start}} = t_0 < t_1 < \dots < t_{n-1} = t_{\text{end}}$$

でサンプリングされた $\mathbb{C}^n$ のベクトルと見なす。それを $\mathbf{v}$ で表すことにする。即ち、

$$\mathbf{v} = (v_0, v_1, \dots, v_{n-1}) = (z(t_0), z(t_1), \dots, z(t_{n-1}))$$

である。同様に、関数族 $\{e_0, e_1, \dots, e_{m-1}\}$ も

$$\mathbf{e}_i = (e_{i,0}, e_{i,1}, \dots, e_{i,n-1}) = (e_i(t_0), e_i(t_1), \dots, e_i(t_{n-1})), \quad (i = 0, 1, \dots, m-1)$$

とする。行列 $\mathbf{A}$ を

$$\mathbf{A} = \begin{bmatrix} e_{0,0} & e_{0,1} & \cdots & e_{0,n-1} \\ e_{1,0} & e_{1,1} & \cdots & e_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ e_{m-1,0} & e_{m-1,1} & \cdots & e_{m-1,n-1} \end{bmatrix}$$

において、 $\mathbf{A}$ の一般化逆行列を $\mathbf{A}^\dagger$ とする。このとき、

$$\mathbf{v} \mathbf{A}^* \mathbf{A}^{\dagger*} = \mathbf{v} \mathbf{A}^\dagger \mathbf{A} \quad (***)$$

は ($*$ は共役転置)、 $\mathbf{v}$ の $\{e_0, e_1, \dots, e_{m-1}\}$ が生成する部分空間の上への直交射影となっている。

$$\mathbf{A}^{\dagger*} = \begin{bmatrix} f_{0,0} & f_{0,1} & \cdots & f_{0,n-1} \\ f_{1,0} & f_{1,1} & \cdots & f_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ f_{m-1,0} & f_{m-1,1} & \cdots & f_{m-1,n-1} \end{bmatrix}$$

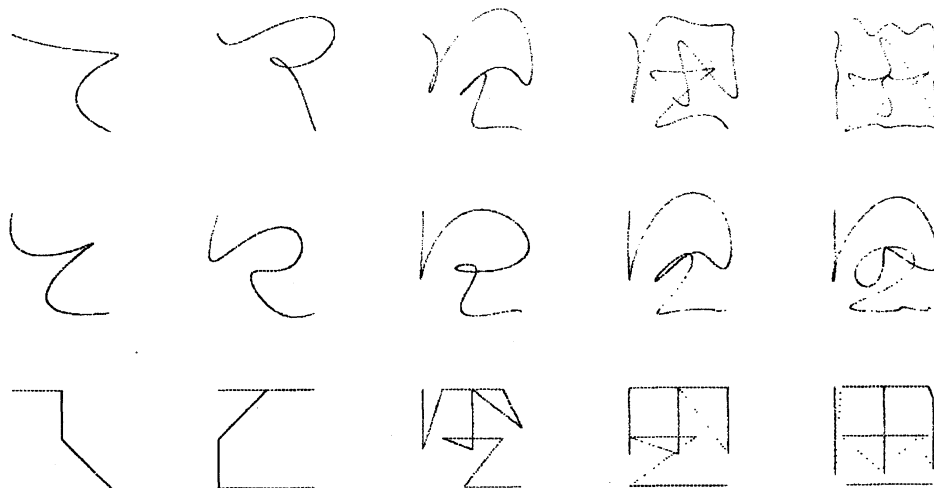
として

$$\mathbf{f}_i = (f_{i,0}, f_{i,1}, \dots, f_{i,n-1}), \quad (i = 0, 1, \dots, m-1)$$

としたとき、(***) は

$$\sum_{k=0}^{m-1} \langle \mathbf{e}_k | \mathbf{v} \rangle \mathbf{f}_k = \sum_{k=0}^{m-1} \langle \mathbf{f}_k | \mathbf{v} \rangle \mathbf{e}_k$$

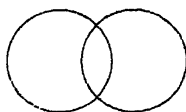
と書き下せる。即ち、 $\{e_0, e_1, \dots, e_{m-1}\}$ が生成する部分空間上へ $\mathbf{v}$ の直交射影をとったとき、それは $\{e_0, e_1, \dots, e_{m-1}\}$ の線形結合で表現できるが、そのとき $\mathbf{e}_k$ の係数が $\langle \mathbf{f}_k | \mathbf{v} \rangle$ と標準的に表すことができる。なおこれは、 $\{e_0, e_1, \dots, e_{m-1}\}$ が線形独立でなくても構わない。従って、 $\{\langle \mathbf{f}_0 | \mathbf{v} \rangle, \langle \mathbf{f}_1 | \mathbf{v} \rangle, \dots, \langle \mathbf{f}_{m-1} | \mathbf{v} \rangle\}$ は $\mathbf{v}$ の $\{e_0, e_1, \dots, e_{m-1}\}$ に関する一般化されたスペクトルと考えることができる。 $\{e_0, e_1, \dots, e_{m-1}\}$ が特に正規直交系であれば $\mathbf{f}_k = \mathbf{e}_k$ ($k = 0, \dots, m-1$) であり、通常のスペクトルに他ならない。



上の図は、「田」という文字の微分 (差分) に対して上から三角関数系、多項式系、Haar 関数系によりローパスフィルタをかけて積分して得られた文字である。左から右にいくに従って関数系の次元は高くなっている。

4 図形の複雑さとエントロピー

次のような図形を考える。



この図形は左のような円が2つ重なった図形と考えるか、それとも右のような欠けた2つの円が接してできた図形と考えるか



これは人間は図形をなるべく単純な部分図形の集まりと捉えようとするからでるとからであると考えられる。図形 (一筆書きの線図形) の複雑さを表す一つの量としてエントロピーを導入する。前節までに述べたように、図形は関数 $z: [0, 1] \rightarrow \mathbb{C}$ とみなして、

$$v = \sum_{k=-\infty}^{\infty} \langle e_k | v \rangle e_k$$

のように Fourier 展開できる。このとき、Parseval の等式

$$\|v\|^2 = \sum_{k=-\infty}^{\infty} |\langle e_k | v \rangle|^2$$

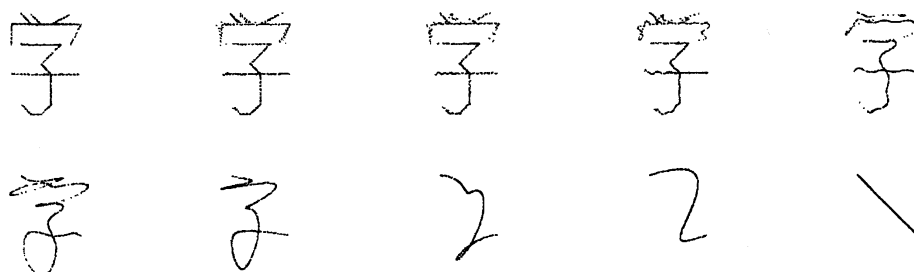
が成立するので、

$$p_k = \frac{|\langle e_k | z \rangle|^2}{\|v\|^2}, \quad k \in \mathbb{Z}$$

とおくと、 $p = (\dots, p_{-1}, p_0, p_1, p_2, \dots)$ は確率となる。そこでこの確率のエントロピー

$$S(p) = - \sum_{k=-\infty}^{\infty} p_k \log p_k$$

を与えられた図形の複雑さを表す量とする。



ローパスフィルタの帯域周波数を $m = 256, 128, 64, 32, 16, 8, 4, 2, 1, 0$ と変化させると、エントロピーは 8.43851206659, 7.43699605519, 6.3851247849, 5.37973143625, 4.33385826257, 3.57958881023, 2.58333074983, 2.07500184344, 1.52890228363, 0.0 と変化する。楷書から行書、草書になるにつれてエントロピーは小さくなる。直線のエントロピーは 0 である。



この節の最初で取り上げた図形を、左は 2 つの円が重なっていると一筆書きで書いた図形でエントロピーは 1.96791731393、右は 2 つの欠けた円が接していると一筆書きで書いた図形でエントロピーは 2.78499248473。

参考文献

- [1] 上坂, 田嶋, "図形分節を説明するためのモデル," 電子通信学会論文誌, 59-D, 1, 1-8, 1976
- [2] 上坂吉則, "かたちのスペクトル分析," 数理科学, 246, 38-45, 1983.
- [3] 上坂吉則, "開曲線にも適用できる新しいフーリエ記述子," 電子通信学会論文誌, J67-A, 3, 166-173, 1984.
- [4] Yoshinori Uesaka, "Spectral Analysis of Form based on Fourier Descriptors, Science on Form," Proceedings of the First International Symposium for Science on Form, KTK Scientific Publishers, pp.405-412, 1986