目次

★4 パターン情報の成分分析 (1) — 直交展開（承前）

★4.2 ベクトルで表されるパターンの直交基底による展開

N 次元ベクトル $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$ をいくつかの成分 $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots$ を用いて $\boldsymbol{x} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \ldots$ のように表すことを考える．ただし，$\boldsymbol{x}$ も $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots$ も N 次元実ベクトルとする (☆1)．したがって係数 $c_1, c_2, \ldots$ も実数である．このとき，線形代数で学んだように，次のことが言える．

☆1) 実ベクトル: 要素が全て実数のみから成るベクトル．

N 次元ベクトル $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_N$ が**線形独立**であれば，任意の N 次元ベクトル $\boldsymbol{x}$ をこれらの**線形結合**として

$$\boldsymbol{x} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_N\boldsymbol{u}_N \tag{1}$$

と一意に表せる．このようなベクトルの集合 $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_N\}$ を**基底**という．

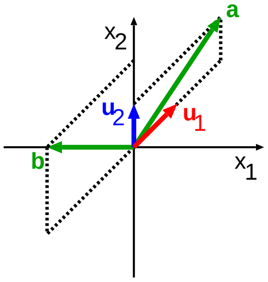
線形独立: linearly independent， 線形結合: linear combination， 基底: basis （複数形は bases)，展開: expansion（名），expand（動）．

言い換えればこれは，N 個の値から成る任意のパターンを $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N$ という成分の和で表せるということである．式 (1) のように表すことを,「ベクトル $\boldsymbol{x}$ を基底 $\{\boldsymbol{u}_k\}$ で**展開**する」という．

下図に，2 次元ベクトルの例を示す．(1),(2) いずれの場合も $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ は基底であり，$\boldsymbol{a} = (2, 3)$ や $\boldsymbol{b} = (-2, 0)$ のようなベクトルを式 (1) の形に表せていることがわかる．

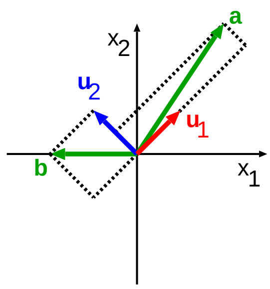
(1) $\boldsymbol{u}_1 = (1, 1), \boldsymbol{u}_2 = (0, 1)$ の場合:

$$\begin{aligned} \boldsymbol{a} &= 2\boldsymbol{u}_1 + \boldsymbol{u}_2 \\ &= 2(1, 1) + (0, 1) \\ \boldsymbol{b} &= -2\boldsymbol{u}_1 + 2\boldsymbol{u}_2 \end{aligned}$$



(2) $\boldsymbol{u}_1 = (1, 1), \boldsymbol{u}_2 = (-1, 1)$ の場合（$\boldsymbol{u}_1$ と $\boldsymbol{u}_2$ は直交してる）：

$$\begin{aligned} \boldsymbol{a} &= \frac{5}{2}\boldsymbol{u}_1 + \frac{1}{2}\boldsymbol{u}_2 \\ &= \frac{5}{2}(1, 1) + \frac{1}{2}(-1, 1) \\ \boldsymbol{b} &= -\boldsymbol{u}_1 + \boldsymbol{u}_2 \end{aligned}$$



Q1. $\boldsymbol{u}_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$,$\boldsymbol{u}_2 = \left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$,$\boldsymbol{u}_3 = \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$ とおくと，これらは互いに直交し，ノルム（長さ）が 1 であることを示せ．

Q2. $\boldsymbol{x} = (2, 4, -1)$ に対して，上記の $\{\boldsymbol{u}_k\}$ を用いて 3 つの実数 $c_k (k = 1, 2, 3)$ を

$$c_k = \boldsymbol{x} \cdot \boldsymbol{u}_k \tag{2}$$

と定義する．このとき c_k の値を求めなさい．さらに，$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + c_3\boldsymbol{u}_3$ を求めなさい．

基底は，互いに直交するベクトルで作ると都合がよい (☆ 2)．つまり，

$$i \neq j \text{ ならば } \boldsymbol{u}_i \cdot \boldsymbol{u}_j = 0 \text{ であり，} i = j \text{ ならば } \boldsymbol{u}_i \cdot \boldsymbol{u}_j \neq 0$$

となるような $\{\boldsymbol{u}_k\}$ を基底とするということである．なぜなら，この場合

$$\begin{aligned} \boldsymbol{x} \cdot \boldsymbol{u}_k &= (c_1\boldsymbol{u}_1 + \cdots + c_N\boldsymbol{u}_N) \cdot \boldsymbol{u}_k & (3) \\ &= c_1\boldsymbol{u}_1 \cdot \boldsymbol{u}_k + \cdots + c_k\boldsymbol{u}_k \cdot \boldsymbol{u}_k + \cdots + c_N\boldsymbol{u}_N \cdot \boldsymbol{u}_k & (4) \\ &= c_k\boldsymbol{u}_k \cdot \boldsymbol{u}_k = c_k||\boldsymbol{u}_k||^2 & (5) \end{aligned}$$

となるので，係数を求める計算が

$$c_k = \frac{\boldsymbol{x} \cdot \boldsymbol{u}_k}{||\boldsymbol{u}_k||^2} \tag{6}$$

と簡単になるからである．さらに，基底を構成するベクトルのノルム（長さ）を全て 1 に，すなわち $||\boldsymbol{u}_k|| = 1$ としておけば，係数の計算はもっと簡単に，

$$c_k = \boldsymbol{x} \cdot \boldsymbol{u}_k \tag{7}$$

とするだけになる．

☆ 2) 成分分析という観点からは，直感的にも成分同士が直交している方がよいことがわかる．例えば,「塩」と「こしょう」を成分とする方が,「塩」と「塩こしょう（塩とこしょうの混ざった調味料）」を成分とするよりよさそうだろう．

互いに直交するベクトルから成る基底を**直交基底**といい，ノルムが 1 で互いに直交するベクトルから成る基底を**正規直交基底**という．$\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N\}$ が正規直交基底だった場合，上述のように展開係数 c_k は次式で求められる．

直交する: orthogonal（形）.
直交基底: orthogonal basis.
正規直交基底: orthonormal basis.

$$c_k = \boldsymbol{x} \cdot \boldsymbol{u}_k \qquad (k = 1, 2, \ldots, N) \tag{8}$$

Q3. Q1 の $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ が正規直交基底であることを利用して，$\boldsymbol{x} = (2\sqrt{3}, 0, -2\sqrt{3})$ を次式のように展開したときの展開係数 c_1, c_2, c_3 を求めなさい．

$$\boldsymbol{x} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + c_3\boldsymbol{u}_3 \tag{9}$$

★ 4.3　離散コサイン変換

パターン情報の分析やデータ圧縮によく用いられる直交展開の例をあげる．

離散コサイン変換 (☆ 3):N 次元ベクトル $\boldsymbol{e}_k$ $(k = 0, 1, \ldots, N-1)$ を次のように定めると，$\{\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_{N-1}\}$ は正規直交基底となる．

$$\boldsymbol{e}_0 = \sqrt{\frac{1}{N}}\,(1, 1, \ldots, 1) \tag{10}$$

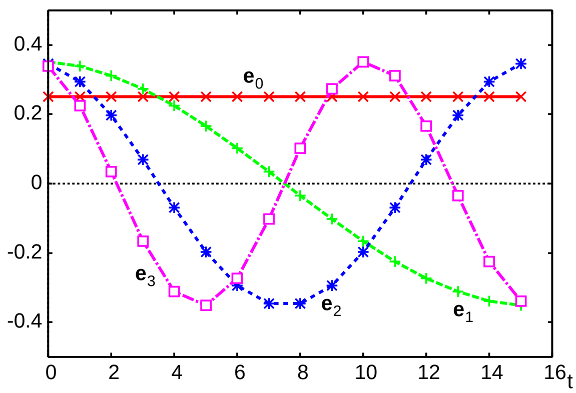
$$\boldsymbol{e}_k = \sqrt{\frac{2}{N}}\left(\cos\frac{\pi k}{2N}, \cos\frac{3\pi k}{2N}, \ldots, \cos\frac{(2N-1)\pi k}{2N}\right) \quad (k = 1, \ldots, N-1) \tag{11}$$

したがって，$\{\boldsymbol{e}_k\}$ を用いると，任意の N 次元実ベクトル $\boldsymbol{f} = (f_0, f_1, \ldots, f_{N-1})$ を次式のように一意に展開できる．

$$\boldsymbol{f} = c_0\boldsymbol{e}_0 + c_1\boldsymbol{e}_1 + \cdots + c_{N-1}\boldsymbol{e}_{N-1} \tag{12}$$

ただし，展開係数 c_k は $c_k = \boldsymbol{f} \cdot \boldsymbol{e}_k$ で与えられる．$\boldsymbol{f}$ から $c_0, c_1, \ldots, c_{N-1}$ を求める演算を離散コサイン変換という．

☆ 3)　離散コサイン変換: Discrete Cosine Transform. DCT と略す．DCT は条件の設定の仕方によっていくつか異なる定義があり，定数倍した形で表すこともあるので，ここに示した式とは異なる式で定義されることも多い．



離散コサイン変換は，DFT（離散フーリエ変換，後の回で解説予定）をベースとして，基底も展開係数も実数の範囲で表せるように修正したものとなっている（DFT に対する FFT と同様に高速演算アルゴリズムが知られている）．右図に，$N = 16$ のときの基底 $\{\boldsymbol{e}_k\}$ のうち番号 k の小さいもの（$k = 0, 1, 2, 3$）を示す．DFT と同様に，0 番目が直流成分（t によらず一定の成分）であり，番号の大きな基底ほど周期の短い（周波数の高い）波の成分に相当していることがわかる．この性質を利用して，DCT はデータ圧縮等に利用される．

音楽などのデータ圧縮手法として有名ないわゆる MP3(MPEG Audio Layer-3) もその一つである．CD に記録された音楽データを MP3 で圧縮すると，典型的にはデータ量を 1/10 程度にすることができる．

$N = 16$ 次元ベクトルに DCT を適用した例を示す．下図左は,

$$\boldsymbol{f} = (0, 0, 0, 0, 10, 10, 10, 10, 8, 6, 4, 2, 0, 0, 0, 0)$$

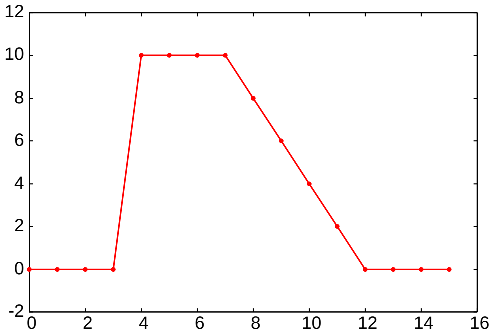
というベクトルの要素を，要素番号を横軸としてグラフに描いたものである．$\boldsymbol{f}$ を DCT して得られる係数 c_k は，$c_k = \boldsymbol{f} \cdot \boldsymbol{e}_k$ $(k = 0, 1, \ldots, 15)$ より，次のようになる（以下の表示は小数点以下 2 桁で打ち切っている）．

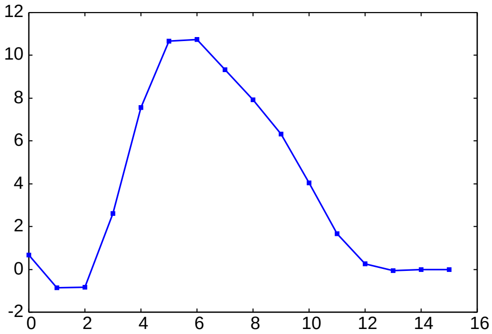
$$\begin{aligned}(c_0, c_1, \ldots, c_{15}) =&(15, 3.27, -14.52, -5.90, 2.23, 2.27, 3.02, 1.73,\\ &0, -1.09, -2.48, -1.55, 0.16, 1.32, 1.95, 0.97)\end{aligned}$$

これら 16 個の係数値を全て用いて式 (12) の右辺を計算すると，元のベクトル $\boldsymbol{f}$ を完全に復元できる．一方，c_8 から c_{15} までの係数を無視して，つまり c_0 から c_7 までの 8 個の係数のみを用いて,

$$\widetilde{\boldsymbol{f}} = c_0\boldsymbol{e}_0 + c_1\boldsymbol{e}_1 + \cdots + c_7\boldsymbol{e}_7$$

という計算をしてみると，得られるベクトル $\widetilde{\boldsymbol{f}}$ は下図右のようになる．DCT の基底は，番号 k が小さいものほど周波数の低い波の成分となっているため，上記のように小さい番号の基底のみでは，高い周波数の成分を表すことができない．そのため，$\widetilde{\boldsymbol{f}}$ は $\boldsymbol{f}$ よりもなだらかな変化をするものとなっている．





★ 4.4　直交展開とデータ圧縮

直交展開は，パターンの成分分析だけでなく，データ圧縮にも利用することができる．

N 個の値から成るパターン $\boldsymbol{x}$ が，基底 $\{\boldsymbol{u}_k\}$ を用いて次のように展開できるとする．

$$\boldsymbol{x} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_N\boldsymbol{u}_N \tag{13}$$

ここで，各基底ベクトルに重要なものから順に番号がついている（k が小さいものほど重要な成分である）としよう (☆ 4)．すると，番号の大きな基底を無視して

$$\widetilde{\boldsymbol{x}} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_H\boldsymbol{u}_H \qquad (H < N) \tag{14}$$

というベクトルを考えると，$\widetilde{\boldsymbol{x}}$ は $\boldsymbol{x}$ の近似になる（両者の差は小さい）と期待できる．このことを利用してデータ圧縮を実現できる (☆ 5)．

具体的に，$N = 1000$ としてみよう．この場合，$\boldsymbol{x}$ をそのまま扱うならその要素である 1000 個の値を記憶する必要がある．しかし，$H = 10$ 個の基底で近似できるなら，c_1 から c_{10} までの 10 個の値を記憶しておくだけで済む．必要なときに式 (14) の計算を行なって近似値 $\widetilde{\boldsymbol{x}}$ を得られるからである．

次回の講義では，実際のデータ圧縮に用いられる直交基底の例を紹介する．

☆ 4) 厳密には「重要な」成分とはどういうものか定義しなければならないが，ここでは直感的な説明にとどめておく．

☆ 5) 近似であることからわかるように，一般に不可逆圧縮である．