有限距離空間の離散凸性

平井　広志
(東京大学大学院情報理工学系研究科数理情報学専攻　現所属・京都大学数理解析研究所)
指導教官　室田一雄 教授

1.　概要

本研究では，バイオインフォマティックス等の分野で現れる有限個の要素からなる距離空間を離散的な凹関数と見なすという新しいアプローチを提案した．これによって凸解析の観点から有限距離空間の組合せ構造を抽出できるようになった．特に進化系統樹の構成法として知られるH.-J. BandeltとA W M. Dressの距離関数のスプリット分解理論[1]を，多面体的凸関数および離散凸関数のスプリット分解に一般化し，その多面体的な幾何構造をより明確にした．またDressらによる有限距離空間の組合せ論を論じる統合的枠組「T理論」[2]の中心的概念であるtight spanの双対的概念である距離複体の概念を導入した．本研究で用いた手法の多くは，室田[3]による多面体的組合せ論から離散凸解析に至るアプローチを精神的に受け継いでいる．

2.　距離関数のスプリット分解の理論

BandeltとDress[1]のスプリット分解を説明する．Vを有限集合，$d: V\times V\to \mathbf{R}$を距離関数とする．$V$上のスプリット$S=A|B$とは，非自明な$V$の2分割のことである．すなわち$V=A\cup B$, $A\cap B=\emptyset$, $A\neq\emptyset$, $B\neq\emptyset$である．V上のスプリット$A|B$に対して，スプリット距離$\delta_{A|B}$を

$$\delta_{A|B}(i,j)=\begin{cases}1 & \text{if } i\in A, j\in B \text{ or } j\in A, i\in B\\ 0 & \text{if } i,j\in A \text{ or } i,j\in B\end{cases}$$

V上のスプリット$A|B$に対して，分離指数$\alpha_{A|B}(d)$を

$$\alpha_{A|B}(d)=\frac{1}{2}\min_{\substack{i,j\in A\\ k,l\in B}}\left\{\max\left\{\begin{matrix}d(i,k)+d(j,l)\\ d(i,l)+d(j,k)\end{matrix}\right\}-d(i,j)-d(k,l)\right\}$$

と定義する．

定理1(文献[1])　$d: V\times V\to\mathbf{R}$を距離関数とすると，

$$d=\sum_{S\ \text{split},\ \alpha_S(d)>0}\alpha_S(d)\delta_S+d'$$

と分解され，d'も距離関数で任意のスプリットSに対して$\alpha_S(d')\leq 0$となる．

スプリット距離の和から対応するネットワークが構成できるので，この分解手法は進化系統樹の推定などに応用されている．

3.　多面体的凸関数のスプリット分解

$f: \mathbf{R}^n\to\mathbf{R}\cup\{\infty\}$が，有限個の線形不等式制約で記述される実効定義域を持ち，そのうえで有限個のアフィン関数の最大値として表されるとき，多面体的凸関数であるいう．$\mathbf{R}^n$の超平面Hに対してスプリット関数を

$$l_H(x)=[x \text{ と } H \text{ の距離}]/2$$

と定義する．これはHを谷としてHの法線方向にむかって大きくなるような単純な多面体的凸関数である．多面体的凸関数fとある超平面Hによるスプリット関数l_Hの「商」$c_H(f)$を

$$c_H(f)=\sup\{t\in\mathbf{R}_+ | f-tl_H \text{ が凸関数}\}$$

と定義する．$0<c_H(f)<\infty$となるとき，fはスプリット関数l_Hの成分を含んでいることを意味する．次のスプリット分解定理が成立する．

定理2　多面体的凸関数fは，

$$f=\sum_{\substack{H\ \text{hyperplane}\\ 0<c_H(f)<\infty}}c_H(f)l_H+f'$$

と分解され，f'も多面体的凸関数で任意の超平面Hに対して$c_H(f')\in\{0,\infty\}$となる．さらにこの分解は一意である．

すなわちスプリット分解とはスプリット関数の成分とスプリット関数の成分を含まない多面体的凸関数への分解である．

4.　離散凸関数のスプリット分解

$X\subseteq\mathbf{R}^n$を有限集合とする．その上の関数$f: X\to\mathbf{R}$の凸拡張を$\bar{f}: \mathbf{R}^n\to\mathbf{R}\cup\{\infty\}$と表す．また$Y\subseteq X$に

対し，f の Y への制限を $f|_Y$ で表す．f が $f=(\bar{f})|_X$ を満すとき凸拡張可能関数と呼ぶことにする．凸拡張可能関数は離散凸関数の最も基本的なクラスの一つである[3]．多面体的凸関数のスプリット分解を離散化することによって次の定理が得られた．

定理3 凸拡張可能関数 $f: X\to\mathbf{R}$ は，

$$f=\sum_{H\in\mathcal{H}_X\ \tilde{c}_H(f)>0}\tilde{c}_H(f)l_H|_X+\gamma$$

と分解され，γ も凸拡張可能で $\forall H\in\mathcal{H}_X$，$\tilde{c}_H(\gamma)\le 0$ を満す．ここで $\mathcal{H}_X$ と $\tilde{c}_H(\gamma)$ は，

$$\mathcal{H}_X=\left\{H:\text{超平面}\left|\begin{array}{l}H\cap\mathrm{conv}\,X=\mathrm{conv}\,(X\cap H)\\ H\cap\mathrm{ri\,conv}\,X\neq\emptyset\end{array}\right.\right\},$$

$$\tilde{c}_H(f)=$$

$$\min_{\substack{x\in X\cap H^{++}\\ y\in X\cap H^{--}}}\left\{\frac{f(x)-\overline{f|_{X\cap H}}(w)}{2l_H(x)}+\frac{f(y)-\overline{f|_{X\cap H}}(w)}{2l_H(y)}\right\}$$

(w は x と y を結ぶ線分と H との交点)

で定義される．

5. 離散凹関数としての距離関数

距離関数を離散凹関数とみなすことで，距離関数のスプリット分解が離散凸関数のスプリット分解の特殊ケースになっていることが示される．まず，有限集合 $X\subseteq\mathbf{R}^V$ を

$$X=\{\chi_i+\chi_j|i,j\in V\}$$

(χ_i は i 方向単位ベクトル)

と定義し，距離関数 $d: V\times V\to\mathbf{R}$ に対し，

$$d(\chi_i+\chi_j)\leftarrow d(i,j)\quad(i,j\in V)$$

と対応させることで $d: X\to\mathbf{R}$ と見なすことができる．このとき距離関数 d は X 上の凹拡張可能関数である．さらに V 上のスプリット $A|B$ に対し超平面 $H_{A|B}$ を

$$H_{A|B}=\{x\in\mathbf{R}^V|\textstyle\sum_{i\in A}x(i)=\sum_{i\in B}x(i)\}$$

と対応させると，$H_{A|B}\in\mathcal{H}_X$ である．すなわちスプリットによる V の2分割を超平面による単体の2分割に対応させるのである．この対応の下，次の定理が成立する．

定理4．V 上のスプリット $A|B$ に対し，

$$\delta_{A|B}=-\sqrt{|V|}\,l_{H_{A|B}}|_X+1,$$

$$\alpha_{A|B}(d)=\tilde{c}_{H_{A|B}}(-d)/\sqrt{|V|}$$

が成り立つ．

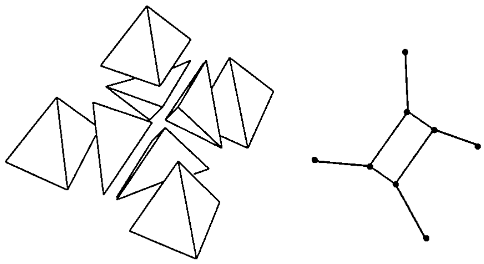

図1 距離複体(左)と tight span(右)

この定理によって距離関数のスプリット分解は離散凸関数のスプリット分解の特殊ケースになっていることかわかる．

6. 距離複体と tight span

A. W. M. Dress らによるT理論の中心的概念である tight span を説明し，本研究との関連を述べる．距離関数 d に対して，tight span $T(d)$ とは，

$$T(d)=\{p\in\mathbf{R}^V|\forall i\in V,\,p(i)=\max_{j\in V}\{d(i,j)-p(j)\}\}$$

で定義される $\mathbf{R}^V$ の部分集合である．tight span には，距離空間の組合せ構造が $\mathbf{R}^V$ の図形の構造として自然に現れる[2]．距離関数 d を $d: X\to\mathbf{R}$ とみなすことによって d の凹拡張の上側面を conv X に射影することによって，自然に conv X の多面体的分割を得る．これを距離複体と呼ぶことにする．このとき，距離複体と tight span は双対的関係にあることが分かった．図1に4点からなる距離空間の場合の例を示す．距離複体の3次元多面体が tight span においては頂点に対応し，距離複体の3次元多面体がファセットを共有するとき対応する枝が tight span に現れる．

参考文献

[1] H.-J. Bandelt and A. W. M. Dress. A canonical decomposition theory for metrics on a finite set, *Advances in Mathematics*, 92, no. 1, 47-105, 1992.

[2] A. W. M. Dress. Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: a note on combinatorial properties of metric spaces, *Advances in Mathematics*, 53, no. 3, 321-402, 1984.

[3] K. Murota. *Discrete Convex Analysis*, SIAM, Philadelphia, PA, 2003.

