球面上の配置と符号理論

宗政　昭弘

情報基礎科学専攻
情報基礎数理学講座

2018 年 12 月 6 日
情報科学談話会

研究室紹介

- 2003 年に着任, 15 年間で 8 人の博士号
- 専門は代数学と組合せ論の境界領域「代数的組合せ論」
- 代数的手法を用いて、球面などの中にある、良い部分集合を見つける、分類する、非存在を示す

地球の平均気温はどうやって測るの？

たくさんの地点で測って平均
偏らないように計測点を配置
偏らないとはどういうこと？
近すぎてはいけない？
真の平均と言える根拠は？

Stone–Weierstrass の定理

温度は地球上を連続的に変化すると考えられる（連続関数）

定理 (Stone–Weierstrass)

球面上の連続関数は、多項式関数を用いて一様に近似できる。

「多項式関数」とは $f(x, y, z) = x^3yz + 2xyz^6 + \cdots$ のようなもの

「一様に近似」とは全ての点で誤差が抑えられる（例えば ± 0.1 度以内）

多項式関数の数値積分

球面のような無限個の点全体で「平均をとる」のが「積分」

積分の計算は簡単な関数なら公式を使えるが、そうでない場合は実際のところ計算不可能。近似的には

$$\frac{1}{4\pi}\int_{S^2} f(x)d\sigma \approx \frac{1}{N}(f(x_1)+\cdots+f(x_N))$$

と期待できるが、根拠は？

$\approx$ を $=$ にすることは一般にはできないが、考えている関数 f の種類を制限すればその範囲で $=$ は可能。
これが球デザイン (spherical t-design) の概念。

球デザイン (spherical t-design)

S^2 = 3 次元空間内の単位球面 $\supset X$ 有限部分集合

t は正の整数とする。

X が spherical t-design であるとは、任意の t 次以下の多項式関数 $f(x,y,z)$ に対して

$$\frac{1}{4\pi}\int_{S^2} f(x,y,z)d\sigma = \frac{1}{|X|}\sum_{p\in X} f(p)$$

左辺は球面上での理論的な平均、右辺は有限個の点 X での平均

球デザインの研究課題

存在問題　要求を満たす点集合が存在するか
Seymour–Zaslavsky (1984)
Bondarenko–Radchenko–Viazovska (2013)

構成問題　要求を満たす点集合をどうやって構成するか
Kuperberg (2005)

非存在問題　少なくとも何点必要か
Delsarte–Goethals–Seidel (1977)

近似の精度と距離

直感的には、地球上の温度の平均を求めるための観測点は数が多いと同時に局所的に集まっていてはいけないことは想像できるが、

- 良い近似 (spherical *t*-design)
- 互いに離れている

という 2 つの性質は、関係がありそうだが、全く別の条件。

実例を想像すると両方の性質を持つものを考えてしまいがち。

立方体

球に内接させると
８頂点は球面上
$\frac{1}{\sqrt{3}}(\pm 1, \pm 1, \pm 1)$

立方体

地球の平均気温の近似のために、球に内接する立方体の８頂点だけでは少なすぎる気がするが、それでも spherical 3-design (３次式までは誤差なし)。

立方体とは、座標が

$$(\pm 1, \pm 1, \pm 1) \text{ の８点}$$

または、

$$(0,0,0),(0,0,1),(0,1,0),(0,1,1),$$
$$(1,0,0),(1,0,1),(1,1,0),(1,1,1)$$

の８点。この $0,1$ の列の長さを増やしたものについても、球面と同じように、「近似」と「距離」を考えることができる。

語と符号理論

「語」とは $a, b, c, \ldots, z$ からなる文字列のこと。

簡単のため、0, 1 だけからなる文字列を考える。長さ n の文字列は 2^n 個ある。n が大きいと非常に多い $\rightarrow$ 近似で済ませたい。

良い近似とは？
球面との類似を考えると、多項式関数（この場合は n 変数 $x_1, \ldots, x_n$）で次数 t までは誤差なし。つまり

$$\text{全体での平均値} = X \text{ での平均値}$$

となるように 2^n 個よりはるかに少ない語の部分集合 X が欲しい。

語全体の近似

例えば

$$(0,0,0),(0,0,1),(0,1,0),(0,1,1),$$
$$(1,0,0),(1,0,1),(1,1,0),(1,1,1)$$

の８点の代わりに

$$\begin{gathered}(\color{red}{0},\color{red}{0},\color{red}{0}),\\(\color{red}{0},\color{red}{1},\color{red}{1}),\\(\color{red}{1},\color{red}{0},\color{red}{1}),\\(\color{red}{1},\color{red}{1},\color{red}{0}).\end{gathered}$$

全ての座標を同時に見ず、２つしか見なければ、全体を忠実に表している。

直交配列

n を正の整数、X を $0,1$ からなる長さ n の語からなる部分集合とする。X が強さ t の直交配列 であるとは t 個の座標だけ見ると、長さ t の全ての語が同じ回数だけ現れているときをいう。

$$\begin{matrix} (0,0,0), \\ (0,1,1), \\ (1,0,1), \\ (1,1,0). \end{matrix} \qquad (t=2)$$

実はこれは t 次以下の多項式関数について

$$全体での平均値 = X での平均値$$

ということと同値。

双対

n を正の整数、X を $0, 1$ からなる長さ n の語からなる部分集合とする。語をベクトルと考えて、mod 2 演算で内積を考えた時の X の直交補空間を $X^{\perp}$ で表すと次は同値。

- X は強さ t の直交配列。
- $X^{\perp}$ の最小ハミング距離は t より大きい。

ハミング距離とは、異なる成分の数：

$(0,0,0)$	$(0,0,0)$	$(1,0,1)$	パンダ
$(1,0,0)$	$(0,1,1)$	$(0,1,0)$	かんだ
距離 1	距離 2	距離 3	距離 1

聞き間違い（通信時の誤り）を避けるためにはハミング距離を大きくすべき

球も語も同じようなもの

数学者にとっては、球も語も同じようなもの。どちらも

- 近似、多項式関数、構成問題など
- 距離

語についてこのような問題を研究するのが（代数的）符号理論。

球においての距離に関する問題の例。球の詰め込み問題：角度が６０度以上になるように球面上に何個の点を配置できるか？

ニュートンが 17 世紀に考えたとされているが、12 個（正 20 面体の頂点数）と決定されたのは 1953 年。

球の詰め込み問題

球や語の一般化

高次元の球面。例えば、４次元空間内の単位球面

$$\{(x, y, z, w) \in \mathbb{R}^4 \mid x^2 + y^2 + z^2 + w^2 = 1\}$$

４次元でも「角度が６０度以上」は定義できて、

- 角度が６０度以上になるように球面上に何個の点を配置できるか？

という問題が意味を持つ。2013 年に Musin によって、24 個であることが示された。しかし、それが本質的に

$$\{\frac{1}{\sqrt{2}}(\pm 1, \pm 1, 0, 0), \frac{1}{\sqrt{2}}(\pm 1, 0, \pm 1, 0), \dots\}$$

に限るかどうかは未解決。

4 次元空間内の球面とユニタリ群

$$S^3 = \{(x, y, z, w) \in \mathbb{R}^4 \mid x^2 + y^2 + z^2 + w^2 = 1\}$$

は特殊ユニタリ群

$$SU(2) = \left\{ \begin{bmatrix} x + iy & z + iw \\ -z + iw & x - iy \end{bmatrix} \mid x^2 + y^2 + z^2 + w^2 = 1 \right\}$$

と同じ。ユニタリ群を有限集合で近似することは量子情報理論でも必要（らしい）。

M. S. Baladram (2018): S^3 における spherical t-design の具体的構成法を発見した → ユニタリ群を通じて量子情報理論に応用できるかも。