

情報幾何の双対空間と超幾何多項式  
高山信毅 (神戸大学) (栗木, 竹村との共同研究より)  
arxiv: 1510.02269

hgm OpenXM search.

整数成分の  $d \times n$  行列  $A = (a_{ij})$ .  $A$  の第  $i$  列ベクトル  $a_i$  を  $\mathbb{Z}^d$  の元とみたとき, この列ベクトル達は  $\mathbb{Z}^d$  を生成していると仮定する. ある行  $i$  について  $a_{ij} > 0$  と仮定.

$\beta \in \mathbf{N}_0 A = \mathbf{N}_0 a_1 + \cdots + \mathbf{N}_0 a_n$  に対して多項式

$$Z_A(\beta; p) = \sum_{Au = \beta, u \in \mathbf{N}_0^n} \frac{p^u}{u!} \quad (1)$$

を  $A$ -超幾何多項式とよぶ.  $u! = \prod_{i=1}^n u_i!$ ,  $p^u = \prod_{i=1}^n p_i^{u_i}$ .  
 $Z = Z_A$  は多項分布を  $Au = \beta$  へ制限した条件付き分布の正規化定数 (分配関数).

$$Z_A(\beta; p) = \sum_{Au=\beta, u \in \mathbf{N}_0^d} \frac{p^u}{u!}$$

例 22 ( $2 \times 2$  分割表):  $A = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$ ,  $\beta = (37, 36, 12)^T$ .  $u$

を次の形式で書く:  $\begin{pmatrix} u_1 & u_2 \\ u_3 & u_4 \end{pmatrix}$ .  $Au = \beta$  を満たす  $u$  達は

$$\hat{u} = u = \begin{pmatrix} 11 & 0 \\ 25 & 12 \end{pmatrix}, \dots, u = \begin{pmatrix} 4 & 7 \\ 32 & 5 \end{pmatrix}, \dots, u = \begin{pmatrix} 0 & 11 \\ 36 & 1 \end{pmatrix}.$$

$$Z_A(\beta; p) = \frac{p_1^{11} p_3^{25} p_4^{12}}{11! 25! 12!} {}_2F_1(-11, -12, 26; y), y = \frac{p_2 p_3}{p_1 p_4} \text{ (odds 比)},$$

$$\text{Ker } A = \mathbb{Z}\bar{a}, \bar{a} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, u = \hat{u} + v\bar{a} \text{ の確率は } \frac{y^v / (\hat{u} + v\bar{a})!}{{}_2F_1(y) / \hat{u}!}.$$

$p(\xi) = (\exp \xi_1, \dots, \exp \xi_n)$ ,  $\psi(\xi) = \log Z(\beta; p(\xi))$  とおく.

$$E[U_i] = \frac{p_i \partial_i \bullet Z}{Z} \Big|_{p=p(\xi)} = \frac{\partial \psi(\xi)}{\partial \xi_i},$$

ここで  $\partial_i = \frac{\partial}{\partial p_i}$ .

$$\frac{\partial \psi(\xi)}{\partial \xi_i} = \sum_{Au=\beta} u_i \frac{\exp(u \cdot \xi) / u!}{Z}$$

(青の部分は  $u$  が出現する確率) なので  $E[U_i]$  は  $u_i$  の期待値.

**MLE (Maximal likelihood estimation) 問題**とは:  $\xi$  の関数としての  $E[U]$  の逆像計算 ( $Au = \beta$  を満たす  $u$  のみを考えるので条件付き MLE とも呼ばれている).

- ① 応用上なぜこの逆像計算が興味ある?
- ② MLE はいつ解ける?  $E[U]$  の像は何?
- ③ 順像および逆像を効率的に計算せよ.
- ④ この写像の性質は?

情報幾何 (参考: 甘利, 情報幾何学の新展開, 2014, サイエンス社) の記号法を使い  $\eta_i = E[U_i]$ ,  $\eta = (\eta_i)$  とおく. 情報幾何の考え方では  $\xi$ -空間と  $\eta$ -空間は双対空間 (双対座標系と見なすのが本式) でありこれらを結ぶ写像が *moment map*  $E[U]$  である.

情報幾何の一般論の紹介:  $\psi(\xi)$  下に (strictly に) 凸な関数.

$$\xi \mapsto \nabla\psi(\xi) =: \xi^* \text{ または } \eta \text{ と書く}$$

$$\psi^*(\xi^*) = \max_{\xi} (\xi \cdot \xi^* - \psi(\xi))$$

とすると  $\psi^*(\xi^*) = \xi$ .

$\xi$ -空間の直線を  $\xi^*$  空間 ( $\eta$  空間) の直線に移したものを, dual flat line.  $D[\xi : \xi'] = \psi(\xi) - \psi(\xi') - \nabla\psi(\xi')(\xi - \xi')$  (divergence) と定義すると,  $D[\xi : \xi'] = D^*[\xi'^*, \xi^*]$ .

例:  $\psi(\xi) = \log Z(\beta; p(\xi))$  とおくと指数型分布族の一般論よりこれは下に凸.

$U = (U_1, \dots, U_n)$  を観測したとして, Fisher の最尤推定 (Fisher's maximal likelihood estimation, MLE) は,  $U$  が起きる (条件付き) 確率  $\frac{p^U}{U!} / Z(\beta, p)$  を最大化する  $p$  を  $p$  の推定値とする.

### Proposition (統計の基礎)

$p$  が Fisher's MLE ならば  $p$  は  $E_i(p) = p_i \frac{\partial Z}{\partial p_i} / Z = U_i$  を満たす.

証. 最大化したい関数  $\frac{p^U}{U!} / Z(\beta, p)$  の対数をとると  
 $\sum U_i \log p_i - \log U! - \log Z$  これを  $F$  とおこう. 最大値をとる  $p$  で,  
 $\frac{\partial F}{\partial p_i} = \frac{U_i}{p_i} - \frac{\partial Z / \partial p_i}{Z} = 0$ . よって,  $U_i = p_i \frac{\partial Z / \partial p_i}{Z}$ .

## 例: 正規分布にみる双対性

M.Michalek, B.Sturmfels, C.Uhler, P.Zwiernik, Exponential Varieties, arxiv:1412.6185 の例:  $\xi$  を positive definite  $m \times m$  symmetric matrix とする.  $x$  を  $m$ -次元のたてベクトルとするととき,

$$Z(\xi) = \int_{\mathbf{R}^m} \exp\left(-\frac{1}{2}x^T \xi x\right) dx = \frac{(2\pi)^{m/2}}{\det(\xi)^{1/2}}.$$

$\log Z = \frac{m}{2} \log(2\pi) - \frac{1}{2} \log \det(\xi)$  となり,

$$\nabla \log Z = -\frac{1}{2}\xi^{-1}$$

Positive definite symmetric matrix 全体  $\text{PD}_m$  は convex cone. 像は  $-\text{PM}_m$  であり convex cone.

$$\frac{\partial}{\partial \xi_{ij}} \log Z = -\frac{1}{2} \int_{\mathbf{R}^n} x_i x_j \exp\left(-\frac{1}{2}x^T \xi x\right) dx = -\frac{1}{2} E[X_i X_j]$$

$\xi$  の最尤推定 (maximal likelihood estimation) は

$$\log \prod_{x \in \text{データ集合}} \exp\left(-\frac{1}{2}x^T \xi x\right) / Z$$

の最大化.

# 応用上なぜこの逆像計算 (条件付き MLE) が興味ある？

	アセトアミノフェン	ジクロフェナクナトリウム
死亡	4	7
生存	32	5

この (4, 7, 32, 5) できまる  $\beta$  を fix して  $\nabla\psi$  の逆像を求めたい。

逆像は 一意ではない が,  $\frac{p_2(\xi)p_3(\xi)}{p_1(\xi)p_4(\xi)}$  は一意にきまり, 10.4167. 条件付き MLE による odds 比とよぶ。

(聞きかじりでは) 同じ  $\xi$  をもつ分割表が多数あるとき, 条件付き MLE による odds 比が同じ程度な分割表は合併して考察してよい。これが異なる場合は合併してはいけない。

- ① 青木 敏, 大津起夫, 竹村彰通, 沼田泰英: 大学入試センター試験科目選択データの統計解析, 応用統計学会誌 39. 71-100 (2010).
- ② 小川光紀, Algebraic Statistical Methods for Conditional Inference of Discrete Statistical Models, 2015/3月, 博士論文 (東京大学).

$R$  ( $\mathbb{Z}^n$  の unimodular 行列),  $S$  ( $\mathbb{Z}^d$  の unimodular 行列) をうまく選べば

$$SAR = \begin{pmatrix} \alpha_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \alpha_d & \end{pmatrix}, \quad \alpha_i \neq 0, \alpha_i | \alpha_{i+1}.$$

$\text{Ker}(A : \mathbb{Z}^n \rightarrow \mathbb{Z}^d)$  の  $\mathbb{Z}$ -加群としての基底は  $\{Re_{d+1}, \dots, Re_n\}$ ,  $(Re_{d+i})^T$  を  $\bar{a}_i$  と書いて,

$$\bar{A} = \begin{pmatrix} \bar{a}_1 \\ \vdots \\ \bar{a}_{n-d} \end{pmatrix}_{(n-d) \times n},$$

$\lambda = \bar{A}\xi$  とおく.  $\exp(\lambda_i) = p^{\bar{a}_i} = \prod_{j=1}^n p_j^{\bar{a}_{ij}}$  を一般化 odds 比とよぶ.

例 22:  $\bar{A} = (-1, 1, 1, -1)$ ,  $\frac{p_2 p_3}{p_1 p_4}$  (元祖 odds 比=一般化 odds 比)



# MLE はいつ解ける? $E[U]$ の像は何?

性質 (不変性).  $\xi - \xi' \in \text{Im } A^T$  の時.  $E[U](\xi) = E[U](\xi')$ .

## Theorem

*Newton* 多面体  $\text{New}(Z)$  (ただし  $Z$  は  $\xi$  の式でなく  $p$  の多項式とみなす) の次元が  $n - d$  と仮定する. このとき, (一般化 odds 比の  $\log$  の空間で考えた) *moment map* は次の同型を与える.

$$E[U] : \mathbb{R}^n / \text{Im } A^T \longrightarrow \text{relint}(\text{New}(Z)) \subset \mathbb{R}^n$$

なお “relint” は相対的内部を意味する.  
証明は,

$$f(\xi) = \eta \cdot \xi - \log Z(\beta, p(\xi))$$

の最大化.

例 22:  $25 < E[U_{21}] < 36$ .

(1) この定理は MLE の可解条件の答え. 境界の対応は複雑.

# moment map $E$

例.  $2 \times 3$  分割表.

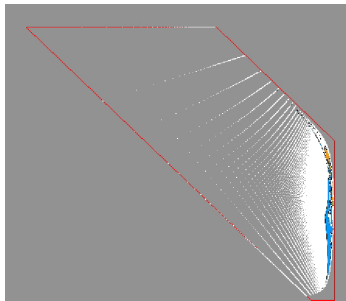
	アセトアミノフェン	ジクロフェナクナトリウム	メフェナム酸
死亡	4	7	2
生存	32	5	6

この場合は, generalized odds の空間は 2 次元.  $Z$  の Newton polytope は五角形で 2 次元.

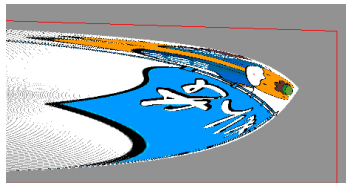
写像のイメージを得るには, 絵がどのように写像されるのを見  
るのも有益. 次の絵が  $p$  空間の商空間  $\mathbf{R}^2$  にあるとする.



# moment map $E[U](p)$ による像



歪みが激しいことがわかる.



拡大図.

# $\nabla \log Z_A$ の順像および逆像を効率的に計算せよ. 歴史.

- ① M.Saito, B.Sturmfels, N.Takayama, Hypergeometric polynomials and Integer Programming, *Compositio Mathematica*, 115 (1999), 185–204. 隣接関係式. 特性多項式の多面体による特徴づけ.
- ② Holonomic Gradient Method (HGM) の登場 (2011–). 正規化定数  $Z$  およびその微分の近似計算、または正確計算. HGM=漸化式 (差分方程式) または微分方程式による  $Z$  やその微分の数値評価.
- ③ M.Ogawa, A.Takemura, N.Takayama, An Application of  $A$ -hypergeometric Equations to Conditional Maximal Likelihood Estimation of  $2 \times m$  Contingency Tables, in preparation. 小川 D 論 (2015/3月) より.
- ④ Y.Goto, Contiguity Relations of Lauricella's  $F_D$  Revisited, arxiv:1412.3256.  $F_D$  と  $2 \times m$  分割表が対応.
- ⑤ K.Ohara, N.Takayama, Pfaffian Systems of  $A$ -Hypergeometric Systems II — Holonomic Gradient Method, arxiv:1505.02947.  $A$ -超幾何多項式の数値評価アルゴリズム. Macaulay 型行列.

$E[U]$  の順像計算も計算量の壁があり近似が主に研究されてきたのが before HGM.

## この写像 $E[U](\xi)$ の性質は?

次の ( $k$  を増やして  $\beta$  部をどんどん大きくしていく) 確率分布の列を考える.

$$P_k(u, \xi) = \frac{\exp(u \cdot \xi)}{u! Z_k(\xi)}, \quad u \in S_k, \quad k = 1, 2, \dots,$$

ここで

$$S_k = \{u \in \mathbb{N}^n \mid Au = k\beta\}, \quad Z_k(\xi) = \sum_{u \in S_k} \frac{\exp(u \cdot \xi)}{u!},$$

**問題:**  $k \rightarrow \infty$  の時これはどのような分布になるか?

- ① J.Cornfield, A Statistical Problem Arising from Retrospective Studies, Proceedings of 3rd Berkeley Symposium on Mathematical Statistics and Probability 4 (1956), 135–148.
- ② 広津, 離散データ解析, 教育出版, 1982.
- ③ R.L.Plackett, Analysis of Categorical Data, 2nd ed, Griffin, 1981. (pp. 41 ( $2 \times 2$  table), pp. 65–66 ( $r \times s$  table))

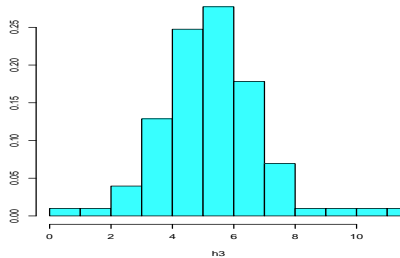
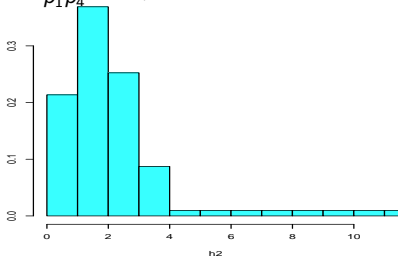
問題:  $k \rightarrow \infty$  の時  $\frac{p^u/u!}{Z(k\beta;p)}$  はどのような分布になるか?

$\lambda = \bar{A}\xi$ ,  $\exp(\xi) = p$  とおく.  $m = m(\lambda)$  を次の連立代数方程式系の (一意的な正の実数) 解とする (解法は伝統的に IPS (iterative proportional scaling) と呼ばれている).

$$\begin{cases} \beta = Am, \\ \lambda = \bar{A} \log m. \end{cases} \quad (2)$$

答:  $km$  を平均とする正規分布に近づく.

例 22,  $\frac{p_2 p_3}{p_1 p_4} = 1/3$  と 3 の時の分布.



## Theorem

$\beta \in \mathbb{N}_0 A \cap \text{int}(\mathbb{R}_{\geq 0} A)$  と仮定する.  $M = \text{diag}(m_i)$ .

$$\widehat{P}_k(u, \xi) = \frac{\det(\bar{A}M^{-1}\bar{A}^T)^{1/2}}{(2\pi k)^{(n-d)/2}} \exp\left(-\sum_{i=1}^n \frac{(u_i - km_i)^2}{2km_i}\right)$$

と置くとき,

$$\sup_{\forall i |u_i - km_i| < \varphi(k)} \left| \frac{P_k(u, \xi)}{\widehat{P}_k(u, \xi)} - 1 \right| \rightarrow 0 \quad (k \rightarrow \infty),$$

ここで  $\varphi(k)$  は次を満たす正の値をもつ関数:  $\varphi(k) = o(k^{2/3})$ ,  $k/\varphi(k)^2 = o(1)$ .

# 条件付き MLE の効率的計算への応用

- ①  $p = u/|u|$  ( $|u| = u_1 + \dots + u_n$ ) とおく. ( $m$  を観測データで決めた odds 比  $\lambda$  からきめる. 近似定理 2 より  $E[U](p) \simeq m$  なのでこれは答えに近いと仮定できる.)
- ②  $\xi$  空間の dual flat line にそって  $p$  を動かして逆像に近づける.

最近の進展 ( $\dot{E}(U)(\xi)$  の有理数による正確数値計算):

- ① Y.Goto, K.Matsumoto, Pfaffian equations and contiguity relations of the hypergeometric function of type  $(k+1, k+n+2)$  and their applications, arxiv:1602.01637. 松本-後藤 による  $r_1 \times r_2$  contingency tables の正規化定数の HGM による高速計算法 (Aomoto-Gel'fand の超幾何系  $E(k, n)$  の twisted cohomology group の組み合わせ論的考察を活用).
- ② modular method による高速化実装は, 後藤-橘-高山.



例.

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

$2 \times 2 \times 2$  分割表.

$p_1$	$p_2$	$p_4$	$p_5$
0	$p_3$	$p_6$	$p_7$

周辺和は“平面”でとる.  $p_4 + p_5 + p_6 + p_7$ ,  $p_1 + p_4 + p_6$ ,  $p_2 + p_3 + p_5 + p_7$ ,  $p_1 + p_2 + p_4 + p_5$ .  
次の観測データを与えるもっともらしい cell の確率  $p = \exp(\xi)$  は? (MLE, 逆像).

19	132	11	52
0	9	6	97

$p_1, p_2, p_3, p_4$  を固定. ( $p_5, p_6, p_7$ ) を動かす.  $\eta = (E[U_5], E[U_6], E[U_7])$

$p$	$\eta$
$P_0 = (19, 132, 9, 11, 52, 6, 97)/326$ 第一近似	(51.9194, 5.99193, 97.0891)
$P_0 + (0, 0, 0, 0, h_1, h_2, h_3)$ ここで $h = (0.000256154, -0.000152585, -0.00310983)$	(52.0006, 6.00006, 96.9993)