多変数超幾何系と統計

高山 信毅　(神戸大学)*

1. はじめに

推定および検定は統計における基本問題であり数理統計の入門教科書ではエラー関数や不完全ガンマ関数等, 一変数の特殊関数が推定や検定のために活躍している. また統計ソフトウエア R [22] にはこれらを含む数多くの特殊関数の数値評価関数が実装されている. これらの一般化としての多変数超幾何積分は統計においてもさまざまな理論的研究が活発に行われてきたが, 計算の立場からは乱数によるシミュレーション手法が主になっていると筆者は理解している. シミュレーションでなく積分の解析で推定や検定ができないであろうか? というのは自然な問である. 非正規化確率密度関数が具体的な関数として与えられている場合, 分布のパラメータ推定や統計的検定には (多変数) 超幾何積分の数値計算が必要となる. 近年の多変数超幾何積分の研究, この積分のみたす函数方程式系である超幾何系の研究, などによりこの数値計算が可能となってきており, ソフトウエアとしての実装もいくつか提供されている [7], [21, gtt_ekn.rr]. 統計で重要な対象をよく調べてみると, 多変数超幾何系の重要な対象であったりすることを筆者は何度か経験した. おもわぬ関係に出会うこともこの題材のおもしろさである.

2. 推定と HGM

S を $\mathbf{R}^d$ の適当なクラスの集合 (たとえば semi-algebraic set) とし, $T_i(t)$ を $t \in S$ についての滑らかな関数とする. パラメータ (母数) $\theta \in \mathbf{R}^n$ と t についての以下の関数 u を S 上の指数型非正規化確率分布, $Z(\theta)$ を u の正規化定数と呼ぶ.

$$u(\theta, t) = \exp\left(\sum_{i=1}^{n} \theta_i T_i(t)\right), \quad Z(\theta) = \int_S u(\theta, t) dt \tag{1}$$

$u(\theta,t)/Z(\theta)$ は θ でパラメータづけされた S 上の確率分布を定義する. これは指数型分布族に属する (たとえば [3]).

例: $A = (a_1, \ldots, a_n)$, $a_i \in \mathbf{N}_0^d$ (列ベクトル). $T_i(t) = t^{a_i} := \prod_{j=1}^{d} (t_j)^{(a_i)_j}$ とおく [1]. $Z(\theta)$ は A-超幾何積分の仲間である (S は rapid decay twisted cycle とは限らない).

例 (一変数の正規分布): $d = 1$, $A = (1, 2)$, $S = \mathbf{R}^2$. $u(\theta, t) = \exp(\theta_1 t + \theta_2 t^2)$, $\theta_2 < 0$. $Z(\theta) = \frac{\sqrt{\pi}}{\sqrt{-\theta_2}\exp(\theta_1^2/(4\theta_2))}$.

　与えられたデータ $t(j) \in S$, $j = 1, \ldots, N$ に対して, 確率の積

$$\prod_{j=1}^{N} u(\theta, t(j))/Z(\theta)$$

本研究は JST CREST JP19209317 の助成を受けたものである。
2010 Mathematics Subject Classification: MSC-33C70, MSC-62Hxx
キーワード：Holonomic gradient method, hypergeometric integral, random matrix, maximal likelihood estimation, statistical test
* 〒657-8501 兵庫県神戸市灘区六甲台町 1-1 神戸大学 大学院理学研究科
e-mail: takayama@math.kobe-u.ac.jp
web: http://www.math.kobe-u.ac.jp/HOME/taka

[1] t_j で t の j-成分を表す. 同様の記法を混乱の恐れがない限り用いる. a_i は行列 A の i 列.

を最大化するパラメータ θ を見つけることを, Fisher の最尤推定という (Fisher's maximal likelihood estimation, MLE). これには以下をみたす θ を求めればよい.

$$
\begin{aligned}
& \operatorname{maxarg}_\theta \prod_{j=1}^{N} u(\theta, t(j))/Z(\theta) \\
= \ & \operatorname{maxarg}_\theta \sum_j \log u(\theta, t(j)) - N \log Z(\theta) \\
= \ & \operatorname{maxarg}_\theta \sum_i \theta_i \frac{1}{N} \sum_j T_i(t(j)) - \log Z(\theta)
\end{aligned} \tag{2}
$$

$-\log Z(\theta)$ は上に凸な関数であることが知られている. したがって θ は次の MLE 方程式の解となる.

$$
\frac{\partial}{\partial \theta_i} \log Z = \frac{1}{N} \sum_j T_i(t(j)) \tag{3}
$$

ここで右辺はデータを十分統計量に代入したものの平均である.
例 (一変数正規分布): $A = (1, 2)$, $\theta_2 = -\frac{1}{2}$ とすると $\frac{1}{N}\sum_j t(j) = \frac{\partial}{\partial \theta_1} \log Z = -\frac{\theta_1}{2\theta_2} = \theta_1$.
この例では $\log Z$ が有理式だが, 一般には超越的な方程式となる.
この MLE 方程式を解くことは, 統計の基本問題である. 解の存在と一意性の定理を紹介しよう.

定理 1 (たとえば Brown の本 [3, Th 3.6], Michalek, Sturmfels, Uhler, Zwiernik [15, Th 2.2] 参照)
$C = \{\theta \in \mathbf{R}^n \mid \log Z(\theta) < \infty\}$, $K = \operatorname{conv} T(S)$ とおく [2]. ある条件のもと, $\nabla \log Z$ は C と K の内部の全単射.

離散版の Z とその MLE 問題の例も一つあげよう. この問題には青本-Gel'fand 超幾何関数や A-超幾何関数 (GKZ 超幾何関数, たとえば [24], 式 (5), (6) 参照) が出現する.

A を $d \times n$ 行列で成分は 0 以上の整数とする. また $\operatorname{rank} A = d$ で $(1, 1, \ldots, 1)$ なる形の行を含むとする. $\beta \in \mathbf{N}_0^d$ を一つ固定する. $p_i = \exp(\theta_i)$ をパラメータとする. $\frac{p^c}{c!}$ は有限集合 $S = \{c \in \mathbf{N}_0^n \mid Ac = \beta\}$ の上の非正規化確率分布となる.

$$
Z(\beta; p) = \sum_{Ac=\beta, c \in \mathbf{N}_0^n} \frac{p^c}{c!} = \sum_{Ac=\beta, c \in \mathbf{N}_0^n} \frac{\exp(\sum_{i=1}^n \theta_i c_i)}{c!} \tag{4}
$$

はその正規化定数.

定理 2 ([27]) Newton 多面体 $\mathrm{New}(Z)$ (Z は p の多項式とみなす) の次元が $n - d$ なら $\nabla \log Z$ は $\mathbf{R}^n / \operatorname{Im} A^T$ と $\operatorname{relint}(\mathrm{New}(Z))$ の全単射.

なおこの多項式 Z は A-超幾何多項式であり, 行列 A とパラメータ β できまる A-超幾何系の解となっている.

存在と一意性を明らかにすることは理論的な基本問題だが, 統計への応用には, 与えられたデータから MLE を計算して求めることが必要となる. (2) にあるようにここで考えている MLE は最大値最小値問題を解くことにより求めることができる. 最大値問題の解法は最適化の分野でさまざまな研究がなされているが, もっとも基本的なアイディアは gradient descent である. たとえば θ が一変数なら, 微分が正の点 $\theta = \theta_0$ で

[2] $T(t) = (T_1(t), \ldots, T_d(t))$

はθ を増やせば目的関数は増えるし, 微分が負の点 $\theta = \theta_0$ ではθ を減らせば目的関数は増える. このようにして極大となる点を探索するのが gradient descent (この場合はむしろ ascent) アルゴリズムのアイディアである. gradient descent には Z, $\frac{\partial Z}{\partial \theta_i}$ の数値, できたら 2 階微分の数値が必要となる. このような数値を計算するには下記のような方法が考えられる.
1. Z をモンテカルロ法などで数値積分する. ただしこの方法では高精度の計算は期待できない. 積分変数の次元が低い場合はさまざまな数値積分アルゴリズムを適用できる.
2. Z の級数展開を求める. 一変数ではこの方法およびそのバリエーションで高精度かつ高速な計算が可能. しかし多変数では計算量が膨大. 級数展開を求める問題, また大域的計算をおこなうには接続公式が有用.
3. Z の近似を用いる. 鞍点法など.

約 10 年ほど前に我々はこの数値計算の問題に対して holonomic gradient method (HGM) および holonomic gradient descent (HGD) という新しい方法を提案した [16], [5] [3]. HGM は非正規化確率分布関数や S に台を持つ Heaviside 関数が holonomic 系を生成する場合に対する方法である.

HGM の 3 段とび

1. 正規化定数の満たす holonomic 系を計算する. 後述の積分アルゴリズムを用いるか理論的に導出 (素手での導出) [4]. この方程式系を Pfaffian (Pfaff 方程式) に変形.
2. Pfaffian を数値的に解くための初期値ベクトルを計算. これは正規化定数とその微分をある点で数値評価することにほかならない. うまい点を見つけることが必要でここは素手.
3. 常微分方程式の数値解法アルゴリズムを用いて上記の初期値問題を解く.

ここでは一変数正規分布について Z の具体形をあえて計算せず HGM の流れを説明しよう: $Z(\theta) = \int_{-\infty}^{\infty} \exp(\theta_1 t + \theta_2 t^2) dt$.

1. Z は次の微分作用素で零化される (微分方程式を満たす).
$$\theta_1 \partial_1 + 2\theta_2 \partial_2 + 1, \partial_1^2 - \partial_2$$
$F = (Z, \partial_2 \bullet Z)^T$ とおくと次の Pfaffian を得る.
$$\frac{\partial F}{\partial \theta_1} = PF, \quad \frac{\partial F}{\partial \theta_2} = QF$$
,
$$P = \begin{pmatrix} \frac{-1}{\theta_1} & \frac{-2\theta_2}{\theta_1} \\ \frac{1}{\theta_1 \theta_2} & \frac{2\theta_2 - 1/2\theta_1^2}{\theta_1 \theta_2} \end{pmatrix}, \; Q = \begin{pmatrix} 0 & 1 \\ \frac{\frac{-1}{2}}{\theta_2^2} & \frac{-5/2\theta_2 + 1/4\theta_1^2}{\theta_2^2} \end{pmatrix}$$
2. $Z(0, -1/2) = \sqrt{\pi}$.

[3] HGM についての文献は [20] にまとめてある. 入門書としては, 日比編, グレブナー道場 (2012) およびその英語版 [6]. 概説記事としては [25] などもわかりやすい.
[4] 微分作用素環での Gröbner 基底を用いたアルゴリズムがある. §4 参照.

3. $F(\theta_1+h_1,\theta_2+h_2) \sim F(\theta)+h_1P(\theta)F(\theta)+h_2Q(\theta)F(\theta)$,　h_i は小さい数. たとえば Euler の差分法 は h_i を増やしながらこの近似を漸化式として $F(\theta)$ の値を次々と求めていく.

この場合のパラメータ θ は連続量であるが, (4) のパラメータ β は整数ベクトルである. このような場合には上記のような微分方程式でなくパラメータについての差分方程式を正規化定数 Z は満たす. 差分方程式を用いて HGM のアナロジーを遂行することを差分 HGM と呼んでいる. 後藤, 松本は Aomoto-Gel'fand 系のみたす差分方程式系を twisted cohomology 群の交点数の計算で具体的に求めた [4]. これを用いると統計の 2 元分割表の Z を正確計算できるようになりかつ MLE をもとめることができる [26], [21, gtt_ekn.rr]. 間野 [14] は direct sampler を構成するのに差分 HGM を用いている. 離散 A-分布以外に HGM の適用が研究された分布は Fisher-Bingham 分布, Bingham 分布, Fisher 分布, Wishart 分布などがある. [20] を参照してほしい.

3. 検定と HGM

統計的検定では事象がどの程度まれであるか? 確率を計算することが必要となる. 例をひとつあげよう. 次のような仮定を考える. K 大学の理学部の 4 回生 100 名の中で "教室の温度を快適と感じる人の割合は 25 度を平均として標準偏差が 3 である正規分布に従っている". さて, 20 度に暖房設定した教室でこの 100 人の学生が講義を受けている. 講義を終了したあと, 目があった一人の学生に教室の温度が快適かどうか聞いたら "とても快適!" と答えた. さて上の仮定が正しいとすると 20 度以下を快適と答える人は全体の 5 パーセント程度の 5 人のはずである. なぜなら

$$\int_{-\infty}^{20} N(25,3;t)dt \sim 0.05, \quad N(m,s;t)=e^{-\frac{(t-m)^2}{2s^2}}/Z, Z=\sqrt{2\pi s^2}$$

一人ランダムに選んだ人が快適と答えた出来事は 5 パーセント以下の確率でおきる "まれ" な出来事ということになる. したがって, 上記のインタビューの結果からは仮定を否定するのがより妥当であろう, ということになる. なお, 3 名の人に同じ質問をして, 一人以上が快適と答える確率は仮定が正しいとして, $1-0.95^3=0.14\cdots$ と以外に大きい数になるので注意が必要である. 私としては, このような論法の良し悪しの議論には距離を置きたいが, ある確率分布のもとで, "まれな出来事" の範囲を指定してそれが起きる確率を計算する, という問題は数学の問題となる. 冒頭に述べた指数型分布族の場合を考える. 積分領域 P を指定した次の積分を考える.

$$I(P;\theta)=\int_P \exp\left(\sum_{i=1}^n \theta_i T_i(t)\right) dt.$$

S を全事象の空間とし, $P \subset S$ をその部分とする時, 上記の確率の計算は $I(P;\theta)/I(S;\theta)$ を計算せよという問題にほかならない.

ここで Gel'fand, Kapranov, Zelevinsky が 1980 年代後半に研究した A 超幾何系 (GKZ 系, [24] などを参照) が登場することとなる. $T_i(t)=t^{a_i}$ とおいて次の積分を考える.

$$I(P;\theta)=\int_P \exp\left(\sum_{i=1}^n \theta_i T_i(t)+(-\beta-1)\log t\right) dt.$$

次の微分方程式系 $H_A(\beta, g)$ を 不完全 $\mathcal{A}$-超幾何系とよぶ:

$$\left(\sum_{j=1}^{n} a_{ij}\theta_j\partial_j - \beta_i\right) \bullet f \;=\; g_i, \qquad (i = 1, \ldots, d) \tag{5}$$

$$\left(\prod_{i=1}^{n} \partial_i^{u_i} - \prod_{j=1}^{n} \partial_j^{v_j}\right) \bullet f \;=\; 0 \tag{6}$$

with $u, v \in \mathbf{N}_0^n$ running over all u, v such that $Au = Av$).

ここで $\mathbf{N}_0 = \{0, 1, 2, \ldots\}$, および $\beta = (\beta_1, \ldots, \beta_d) \in \mathbf{C}^d$ はパラメターであり, $g = (g_1, \ldots, g_d)$ の各 g_i は holonomic function (定義 1).

定理 3 ([17]) . P が polytope であるとき, 関数 $I(P;\theta)$ は不完全 A-超幾何系 (5), (6) を満たす.

特に $g_i = 0$ の時は不完全 A-超幾何系は A-超幾何系であり, 過去30年の研究によりさまざまな性質が知られているが, 不完全超幾何系の研究はまだあまり進んでいない. なお P が polyhedron の場合も適当な収束条件を課せば同じような定理が成り立つ. 古典的な例として χ^2 検定に用いる χ^2 分布は不完全 A-超幾何関数である. この分布では積分

$$\int_x^\infty t^{k/2-1} \exp(-t/2) dt$$

の値の計算が検定に必要となる.

さて上記の定理は一般的すぎて実際のHGM計算に使うには不足点が多い. 統計的に重要でかつ実際の計算もできる場合としては小山, 竹村 [11] や小山 [9] による結果と実装の提供がある [7], [10]. a_{ij}, b_j で定義される次のような polyhedron を考えよう.

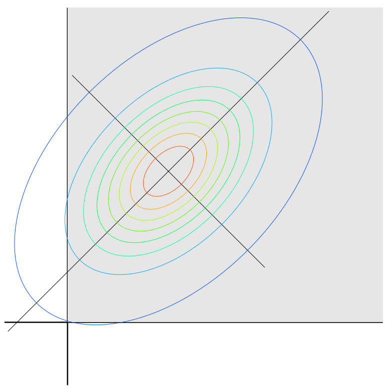
$$P = \{x \in \mathbf{R}^d : \sum_{i=1}^{d} a_{ij}x_i + b_j \geq 0,\ 1 \leq j \leq n\}, \tag{7}$$

小山が考察したのは次の積分である.

$$\mathrm{Prob}(x \in P) = \frac{1}{(2\pi)^{d/2}} \int_{x\in P} \exp\left(-\frac{1}{2}\sum_{i=1}^{d} x_i^2\right) dx_1 \cdots dx_d. \tag{8}$$

Orthant probability by R [7].

```
library(hgm);
hgm.ncorthant(
 matrix(c(2,1,1,2),nrow=2),
 y=c(1,1.5));
[1] 0.6930033
```



第一象限が積分領域の場合, [11].

F_j を P の facet で $\sum_{i=1}^d a_{ij}x_i + b_j = 0$ に対応するものとする. Facet の交わりに対応する index を集めたものを $\mathcal{F}$ と置く. つまり, $\mathcal{F} = \{J \subset \{1,\ldots,n\} \,|\, F_J \neq \emptyset\}$, $F_J = \cap_{j\in J} F_j$ ただし $\mathcal{F}$ は空集合も含むものとし, これは P に対応する. Pfaffian の基底は $\{g^J,\, J \in \mathcal{F}\}$ と $\mathcal{F}$ で index づけできる. Pfaffian は以下のようになる.

定理 4 (小山 [9])

$$
\begin{aligned}
\partial_{a_{ij}} g^J &= \sum_{k=1}^n a_{ik} \partial_{b_k} \partial_{b_j} g^J \quad (1 \le i \le d,\, 1 \le j \le n,\, J \in \mathcal{F}), \\
\partial_{b_j} g^J &= g^{J\cup\{j\}} \quad (j \in J^c,\, J \in \mathcal{F}), \\
\partial_{b_j} g^J &= -\sum_{k\in J} \alpha_J^{jk}(a) \left(b_k g^J + \sum_{\ell \in J^c} \alpha_{k\ell}(a) g^{J\cup\{\ell\}} \right) \quad (j \in J,\, J \in \mathcal{F}).
\end{aligned}
$$

ここで $\alpha_F(a)$ は ij 成分が $\sum_{k=1}^d a_{ki}a_{kj}$ である行列. $\alpha_F(a)$ の逆行列の ij 成分を $(\alpha_F^{ij}(a))_{i,j\in F}$ とおく.

例 [10]: $a = [[1,0],[0,1],[-1,-1]], b = [0,0,2]$, つまり P が $x > 0, y > 0, -x-y+2 > 0$ (三角形) の場合 $\mathcal{F} = \{\emptyset, 1, 2, 3, 12, 13, 23\}$ で $\sharp\mathcal{F} = 7$.

```
./a.out 1 2 1 0 0 0 1 0 -1 -1 2
result:
   0.1775361552    0.4772498666    0.4772498666
   1.0000000000    0.3100122972    0.1353352805
   0.1353352805
Probability = 0.177536
```

4. Holonomic distribution と計算代数

さて HGM のポイントの一つは $Z(\theta)$ の満たす微分方程式系を導出する一般的なアルゴリズムが保証されていることである. holonomic 関数, holonomic distribution の概念がこのような代数解析的アルゴリズムの基礎となっている. このクラスの関数を定義しておこう. まず 1 変数の場合. $f(x)$ を $\mathbf{R}$ の区間 U で定義された滑らかな関数としよう. この関数が多項式係数の線形常微分方程式 L の解であるときこの関数 f およびこの関数の解析接続を 1 変数 holonomic 関数とよぶ. L を f の annihilator と呼ぶ.
例: $Z(\theta) = \int_0^{+\infty} \exp(\theta x - x^3)dx$ は θ の関数として holonomic 関数である. 微分方程式 $(3\partial_\theta^2 - \theta) \bullet Z(\theta) = 1$ を満たす.
次の定理は $\mathbf{C}(x)$ の上の線形代数で簡単に示せる.

定理 5 Holonomic 関数の和および積はまた holonomic 関数である. Holonomic 関数の微分はまた holonomic 関数である.

Holonmic 関数の商は holonomic 関数とは限らないことを注意しておく. たとえば $1/\sin(x)$ は無限個の極を持つが, 多項式係数の線形常微分方程式の解は有限個の特異点しか持たないので, $\sin(x)$ は holonomic 関数であるが $1/\sin(x)$ は holonomic 関数でない.

次に n 変数の holonomic 関数を定義しよう. $f(x) = f(x_1,\ldots,x_n)$ を $\mathbf{R}^n$ の開集合 U で定義された滑らかな関数としよう.

定義 1 関数 $f(x)$ が各変数についてパラメータ付きの線形常微分方程式 n 個の共通解であるときこの関数 f およびこの関数の解析接続を n 変数 holonomic 関数とよぶ.

“各変数についてパラメータ付きの線形常微分方程式” を式で説明すると,

$$L_i = a^i_{m_i}(x)\partial_i^{m_i} + a^i_{m_i-1}(x)\partial_i^{m_i-1} + \cdots + a^i_0(x), \quad \partial_i = \frac{\partial}{\partial x_i} \tag{9}$$

$a^i_j(x) \in \mathbf{C}(x)$ なる形の方程式であり, $L_i \bullet f = 0$, $i = 1, \ldots, n$ が成り立っている.

定理 6 (Zeilberger project, [29] など) $f(x_1, \ldots, x_n)$ が x についての holonomic 関数であるとする. この時パラメータ付き定積分 $\int_\Omega f(x)dx_n$ は $(x_1, \ldots, x_{n-1})$ についての holonomic 関数となる (ただし Ω について条件が必要).

これは holonomic D-加群が積分関手でまた holonomic D-加群にうつされるという一般論からわかることであるが, このような概念がアルゴリズム的アプローチに有効であることを 1980 年代終わりに最初に提唱したのは D.Zeilberger であり, その後の組み合わせ論的等式の機械証明の基礎となった.

n 変数の holonomic distribution を定義しようとすると, 微分作用素環の代数的取扱が必須となる. またたとえば検定の解説で出てきた多面体 P を台とするヘビサイド関数を annihilate する微分作用素全体を取り扱うことができるようになる. D_n を多項式係数の微分作用素のなす環とする. D_n は有理式を係数とする微分作用素環 $R_n = \mathbf{C}(x)\langle\partial_1, \ldots, \partial_n\rangle$ の部分環である. $L = \sum_{(\alpha,\beta)\in E} a_{\alpha,\beta}x^\alpha\partial^\beta \in D_n$ に対して次のように (u, v) 次数および (u, v) initial term を定義する.

$$\begin{aligned} \mathrm{ord}_{(u,v)}(L) &= \max_{(\alpha,\beta)\in E}(u\alpha + v\beta) & (10) \\ \mathrm{in}_{(u,v)}(L) &= \sum_{\mathrm{ord}(x^\alpha\partial^\beta)=\mathrm{ord}(L),(\alpha,\beta)\in E} a_{\alpha,\beta}x^\alpha\xi^\beta \in \mathbf{C}[x, \xi] & (11) \end{aligned}$$

ここで $u = (1, \ldots, 1)$ and $v = (1, \ldots, 1)$.
例: $L = (x_1 - x_2)\partial_1\partial_2 + \partial_1 + \partial_2$. とするとき $\mathrm{ord}_{(u,v)}(L) = 3$, $\mathrm{in}_{(u,v)}(L) = (x_1 - x_2)\xi_1\xi_2$
D_n の左イデアル I に対して $\mathrm{in}_{(u,v)}(I) = \langle \mathrm{in}_{(u,v)}(L) \,|\, L \in I\rangle$ を I の (u, v)-initial ideal と呼ぶ I は $\mathrm{in}_{(u,v)}(I)$ の (Krull) 次元が n であるとき holonomic イデアル とよぶ. $\dim_{\mathbf{C}(x)} R_n/R_nI$ を I の holonomic rank と定義する. holonomic イデアルに対する holonomic rank は必ず有限となり, (generic な点における) 解空間の次元に等しい.

$\mathbf{R}^n$ 上の Schwartz distribution f が holonomic ideal で零化される (holonomic ideal を連立線形偏微分方程式として見た時の解の) とき, f を holonomic distribution とよぶ. D-加群の理論より次の事実がなりたつ ([6] などを参照).

定理 7 $f(x_1, \ldots, x_n)$ が holonomic distribution の時, ある条件の元,
$g(x') = \int_{\mathbf{R}^{n-m}} f(x)dx_{m+1}\cdots dx_n$ も m-変数 x' の holonomic distribution となる.

$g(x')$ が満たす微分方程式系を $f(x)$ の満たす微分方程式系から求めるアルゴリズムは積分アルゴリズムと呼ばれている. 具体的には D_n の左イデアル I と右イデアル $\partial_1 D_n + \cdots + \partial_m D_n$ に対して

$$(I + \partial_1 D_n + \cdots + \partial_m D_n)) \cap \mathbf{C}\langle x_{m+1}, \ldots, x_n, \partial_{m+1}, \ldots, \partial_n\rangle$$

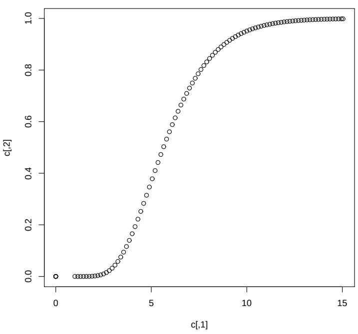
なる消去計算をする必要がある. これについてはさまざまなアルゴリズムおよび実装が研究されてきている. 全部は列挙できていないが, (1) Creative telescoping (D.Zeilberger, 1980’s – 2010’s, [29]), (2) D-加群の積分 (大阿久 [18] など), (3) F.Chyzak’s heuristics

(2000's), C.Kouchan's heuristics (2010's) は Mathematica のパッケージ HolonomicFunctions.m [12] に実装されている. (4) Risa/Asir, Macaulay 2, Singular, などの計算代数システムの D-module パッケージ, (5) semi-algebraic set の Heaviside 関数の満たす holonomic 系 (大阿久 [19] など), など多くの研究がある.

5. ランダム行列と HGM

x が Polyhedron P の上にある確率 $P(x \in P)$ を計算する問題を紹介したが, ランダム行列 W に対してその最大固有値が x 以下となる確率 $P(\lambda_1(W) < x)$ を求める問題も積分の計算に帰着する場合がある. $X_1, \ldots, X_n$ を m 次元 (縦ベクトル) の正規分布 (平均 0, 共分散行列 Σ は固定) に従う独立な n 個のランダムベクトルとする. $m \times n$ 行列 $X = (X_1, \ldots, X_n)$ に対して $P(\lambda_1(XX^T) < x)$ が行列超幾何関数 ${}_1F_1$ で書けるということは James や Constantine による 1960 年代の古典的な結果であり, また微分方程式系は Muirhead により 1970 年頃には与えられている [23]. しかしながらその級数による数値計算は m の増加とともに急激に困難になる. HGM の考え方を用いると $m = 10$ 程度でも Pfaffian の数値解析で十分高い精度で計算できることを示しその R での実装を提供したのが [5] である.

```
library("hgm");
c<-hgm.pwishart(m=4,n=10,beta=c(1,2,3,4),q0=1,q=15,approxdeg=20,
                err=c(3e-13,1e-10),autoplot=1);
> plot(c)
```



$\Sigma^{-1} = \mathrm{diag}(1, 2, 3, 4)$.
左は R の hgm パッケージ [7] による

$$P(\lambda_1(XX^T) < q)$$

のグラフ.

関連する問題として最近の結果を一つ紹介しておこう. A を $m \times n$ 実 random 行列. $g \in S^{m-1}$, $h \in S^{n-1}$ とする. g でパラメータ付けした $SO(m)$ の元を $(g, G(g))$, h でパラメータ付けした $SO(n)$ の元を $(h, H(h))$ として, $\sigma = g^T A h$, $B = G^T A H$ と A を変数変換する. この時 $M_x = \{hg^T \mid g^T A h \geq x\}$ の Euler 標数の期待値は次のような積分表示を持つ.

定理 8 ([28]) 行列 A は分布 $p(A)$ に従うとする. ある条件のもと, $\mathrm{E}[\chi(M_x)]$ は

$$\frac{1}{2}\int_x^\infty \sigma^{n-m} d\sigma \int_{\mathbf{R}^{(m-1)(n-1)}} dB \int_{S^{m-1}} G^\top dg \int_{S^{n-1}} H^\top dh \ \det(\sigma^2 I_{m-1} - BB^\top) p(A). \quad (12)$$

に等しい. ここで $G^\top dg = \wedge_{i=1}^{m-1} G_i^\top dg$, $H^\top dh = \wedge_{i=1}^{n-1} H_i^\top dh$, であり, G_i と H_i はそれぞれ G と H の i 番目の列ベクトルである. また $dg = (dg_1, \ldots, dg_m)^\top$, $dh = (dh_1, \ldots, dh_n)^\top$.

Euler 標数法の一般論 ([1], [13] 参照) によるとこれは AA^T の最大固有値の分布を $x \to +\infty$ で近似している. われわれはさまざまな条件のもとで, この積分の満たす x についての常微分方程式を理論的方法や計算代数の方法で求めた.

$$p(A)dA = \frac{1}{(2\pi)^{mn/2} \det(\Sigma)^{n/2}} \exp\Big\{-\frac{1}{2}\mathrm{Tr}\,(A-M)^\top \Sigma^{-1}(A-M)\Big\} dA. \quad (13)$$

なる場合を考えよう. A の最大特異値はnon-central Wishart matrix $W_m(n, \Sigma, \Sigma^{-1}MM^\top)$ の最大固有値の平方根となる.

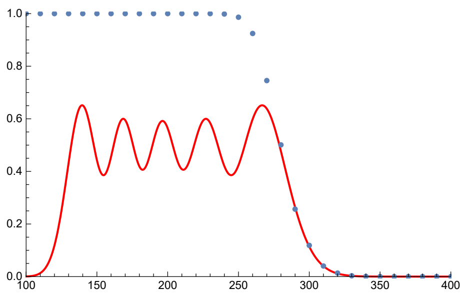
理論的な結果: $M=0$ となる central かつ共分散行列がスカラー行列 I_m/s となる場合は, 青本や金子により研究された Selberg 型積分 [2], [8] の合流型になり次の定理が成り立つ.

定理 9 ([28])

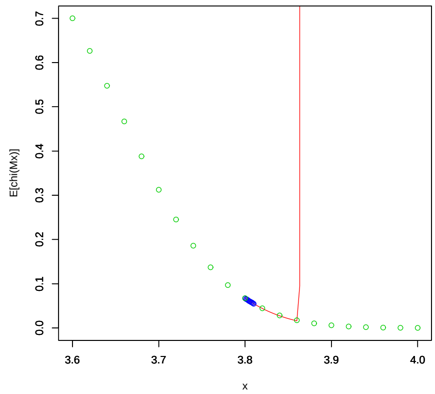
$$\mathrm{E}[\chi(M_x(s))] = \prod_{i=1}^{5} c_i \int_x^{+\infty} \exp\left(-\frac{s}{2}\sigma^2\right) \sigma^{n-m} {}_1F_1(-(m-1), 1+n-m; s\sigma^2) d\sigma,$$

ここで c_1, c_2, c_3, c_4, c_5 は Γ 関数を用いて書ける定数(具体形は論文参照).

計算による結果: non-central な場合はまだまだ未知な部分が多い. 上記で解説した, 最新の積分アルゴリズムを用いると $n=m=2$ の時に x についての 11 階の線形常微分方程式を具体的に導出できて, その数値解析で Euler 標数の期待値を計算できる. しかしながら, この数値解析はかなり困難であり, また n, m の大きくなったときの微分方程式の具体形は不明で, 積分アルゴリズムが組み合わせ論の公式証明で活躍したように, このような問題も今後の計算代数の研究の進展で攻略可能となることを期待している.



$m=10,\ n=200,\ M=0,\ \Sigma=I_m$.



$m=n=2,\ M=\begin{pmatrix}1 & 3\\ 2 & 0\end{pmatrix}$, $\Sigma=\mathrm{diag}(10^{-3}, 10^{-2})$.

参考文献

[1] R. J. Adler, J. E. Taylor, Random fields and geometry, Springer, 2007.

[2] K. Aomoto, Jacobi Polynomials associated with Selberg Integrals, SIAM Journal on Mathematical Analysis 18 (1987), 545–549.

[3] L.D.Brown, Fundamentals of Statistical Exponential Families with Applicatioins in Statistical Decision Theory, Lecture Notes-Monograph Series 9, Institute of Mathematical Statistics, 1986.

[4] Y. Goto, K. Matsumoto, Pfaffian equations and contiguity relations of the hypergeometric function of type$(k+1, k+n+2)$ and their applications, Funkcialaj Ekvacioj 61 (2018), 315–347.

[5] H. Hashiguchi, Y. Numata, N. Takayama, A. Takemura, Holonomic gradient method for the distribution function of the largest root of a Wishart matrix, Journal of Multivariate Analysis, 117 (2013), 296–312.

[6] T. Hibi et al., Gröbner Bases: Statistics and Software Systems, Springer, (2013).

[7] `https://cran.r-project.org/web/packages/hgm/index.html`

[8] J. Kaneko, Selberg Integrals and Hypergeometric Functions associated with Jack Polynomials, SIAM Journal on Mathematical Analysis 24 (1993), 1086–1110.

[9] T.Koyama, Holonomic Modules Associated with Multivariate Normal Probabilities of Polyhedra, Funkcialaj Ekvacioj 59 (2015), 217–242.

[10] `http://www.github.com/tkoyama-may10`

[11] T.Koyama, A.Takemura, Calculation of Orthant Probabilities by the Holonomic Gradient Method, Japan Journal of Industrial and Applied Mathematics 32 (2012), 187–204.

[12] C. Koutschan, HolonomicFunctions user's guide, Technical Report 10-01, RISC Report Series, Johannes Kepler University Linz, Austria, (2010). `http://www.risc.jku.at/publications/download/risc_3934/hf.pdf`

[13] S. Kuriki, A. Takemura, Volume of tubes and the distribution of the maximum of a Gaussian random field, Selected Papers on Probability and Statistics, American Mathematical Society Translations Series 2, 227 (2009), 25–48.

[14] S.Mano, Partitions, Hypergeometric Systems, and Dirichlet Processes in Statistics, JSS Research Series in Statistics, 2018.

[15] M.Michalek, B.Sturmfels, C.Uhler, P.Zwiernik, Exponential Varieties, Proceedings of the London Mathematical Society 112 (2016), 27–56.

[16] H.Nakayama, K.Nishiyama, M.Noro, K.Ohara, T.Sei, N.Takayama, A.Takemura , Holonomic Gradient Descent and its Application to Fisher-Bingham Integral, Advances in Applied Mathematics 47 (2011), 639–658.

[17] K.Nishiyama, N.Takayama, Incomplete A-Hypergeometric Systems, Ed. by T.Hibi, Harmony of Groebner Bases and the Modern Industrial Society, World Scientific, 193–212.

[18] T.Oaku, algorithms for b-functions, restrictions, and algebraic local cohomology groups, Advances in Applied Mathematics 19(1997), 61–105.

[19] T.Oaku, Algorithms for Integrals of Holonomic Functions over Domains defined by Polynomial Inequalities, Journal of Symbolic Computation 50 (2013), 1–27.

[20] References for HGM `http://www.math.kobe-u.ac.jp/OpenXM/Math/hgm/ref-hgm.html`. HGM のサーベイ `http://www.math.kobe-u.ac.jp/HOME/taka/2015/hgm-dic.pdf`.

[21] Risa/Asir, a computer algebra system. `http://www.math.kobe-u.ac.jp/Asir`

[22] `http://www.r-project.org`

[23] R. J. Muirhead, Aspects of Multivariate Statistical Theory, Wiley, 2005.

[24] M. Saito, B. Sturmfels, N. Takayama, Gröbner Deformations of Hypergeometric Differential Equations, Springer, 2000.

[25] A.L.Sattelberger, B.Sturmfels, D-Modules and Holonomic Functions, arxiv:1910.01395

[26] Y.Tachibana, Y.Goto, T.Koyama, N.Takayama, Holonomic Gradient Method for Two Way Contingency Tables, arxiv:1803.0417, to appear in Algebraic Statistics.

[27] N.Takayama, S.Kuriki, A.Takemura, A-Hpergeometric Distributions and Newton Polytopes, Advances in Applied Mathematics 99 (2018) 109–133.

[28] N.Takayama, L.Jiu, S.Kuriki, Y.Zhang, Computations of the Expected Euler Characteristic for the Largest Eigenvalue of a Real Wishart Matrix, Journal of Multivariate Analysis 179 (2020), `https://doi.org/10.1016/j.jmva.2020.104642`.

[29] M.Petkovsek, H.Wilf, D.Zeilberger, $A = B$, AK Peters/CRC Press, 1996.