Proceedings of the Twenty-Seventh RAMP Symposium
Shizuoka University, Shizuoka, October 15–16, 2015

ホロノミック勾配法

Holonomic Gradient Method

高山 信毅[1*]

Nobuki Takayama

概要　ホロノミック勾配法 (holonomic gradient method, HGM) は各種の確率分布の正規化定数およびその高階微分を数値評価するための新しい手法である．この手法では holonomic 系とよばれる線形偏微分方程式系を用いる．理論, アルゴリズム, と共に Fisher-Bingham 分布, Bingham 分布, Wishart 分布, Fisher 分布, 象限確率, 分割表, 多面体確率等への適用例および実装などを紹介する．この新しい数値計算技法はこれらの分布に対する最適化問題への適用も期待できる．

キーワード　ホロノミック勾配法, holonomic gradient method, HGM, holonomic 系, グレブナー基底, Runge-Kutta 法, Fisher-Bingham 分布, Bingham 分布, Wishart 分布, Fisher 分布, 象限確率, 分割表, 多面体確率

1.　準備

ホロノミック勾配法 (holonomic gradient method, HGM) については, さまざまな入門的な文献がある [20], [2], [3]. ここでは入門的な話題はこれらに譲り, 現状でどのような分布に対してどの程度 HGM の適用が可能かを念頭に公式集的に HGM の概説を行なう．なお後述の式 (2) の両辺を微分すると Hessian が計算できる．HGM は各種の最適化アルゴリズムとの相性がよく, 目的関数が陽に表せず積分などを含む場合に有効である．

ある点 $x = a$ の近傍でなめらかな多変数関数 $f(x) = f(x_1, \ldots, x_n)$ が n 個の多項式係数の常微分方程式

$$\sum_{k=0}^{r_i} a_{ik}(x)\partial_i^k \bullet f = 0, \quad (i = 1, \ldots, n) \tag{1}$$

を満すとする． ここで a_{ik} は $x = (x_1, \ldots, x_n)$ の多項式である． また $\partial_i = \frac{\partial}{\partial x_i}$ であり, $\partial_i^k \bullet f = \frac{\partial^k f}{\partial x_i^k}$ と $\bullet$ 記号で ∂_i^k の f への作用を表す．上記の常微分方程式は x_i についての常微分方程式で, その他の x_j はパラメータとみなせる．このような性質を満す多変数関数の解析接続を, holonomic 解析関数と呼ぶ．たとえば, g, h を n 変数の多項式とするとき $\exp(g(x)/h(x))$ は holonomic 解析関数である．

応用のためには holonomic 超関数を導入しておく必要があるが, まず応用で重要となる

1　神戸大学 大学院理学研究科数学専攻〒657-8501 神戸市灘区六甲台 1-1
Department of Mathematics, Kobe University Rokkodai 1-1, Kobe, 657-8501, Japan

∗　E-mail address: takayama@math.kobe-u.ac.jp

holonomic 超関数のクラスを二つ紹介しておく．$f(x)$ を holonomic 解析関数とする．$h(x)$ を n 変数の実数係数の多項式とし, h の実零点集合 $V(h) = \{z \in \mathbf{R}^n \,|, h(z) = 0\}$ がなめらかであり (特異点がない), かつ f の特異点集合と $V(h)$ の交わりは空であるとする．このとき $\delta_{V(h)}(x)f(x)$ は holonomic 超関数である．ここで $\delta_{V(h)}(x)$ は $V(h)$ に台をもつデルタ関数である．

二つ目の例は, (ヘビサイド関数) $\times$ (holonomic 解析関数), である．$h_j(x)$, $(j = 1, \ldots, m)$ を実数係数の多項式とする．$h_j > 0$ 達の定義する領域, つまり $\mathcal{H} = \{z \in \mathbf{R}^n \,|\, h_j(z) > 0, j = 1, \ldots, m\}$ で値 1, この外で値 0 をとる関数を $\{h_j\}$ の定義するヘビサイド関数とよび $H_{\mathcal{H}}(x)$ と書く．$f(x)$ を holonomic 解析関数とする．f の特異点集合と $\mathcal{H}$ の境界との交わりは空であるとする．このとき, $H_{\mathcal{H}}(x)f(x)$ は holonomic 超関数である．holonmic 解析関数, holonomic 超関数を総称して, 誤解のない場合は, 以下 holonomic 関数とよぼう．

holonomic 系の概念を導入しよう．この概念の技術的細部は煩雑なので詳しくは参考書 [2, §6.8] を参照．定義は煩雑だがさまざまな良い性質をもつ素晴らしい概念であることは強調しておく．D を n 変数の多項式係数の微分作用素環とする．$\{F_k\}$ を D の Bernstein filtration, つまり F_k を次数 (x_i 達の次数と ∂_i 達の次数をすべて足したもの) が k 以下の多項式係数偏微分作用素の全体の集合とする．I を D^m の 左部分加群とする．$F_k^m/(I \cap F_k^m)$ の線形空間としての次元が k が十分大で $O(k^n)$ となるとき, D^m/I は holonomic であるという．また I の生成元達を本概説では holonomic 系とよぶ．holonomic 系の解となっている超関数が holonomic 超関数の一般的な定義である．holonomic 関数とそれに対するアルゴリズムは Zeilberger により (1990), 組み合わせ論の公式の機械証明のために導入された [2, 6 章].

f を holonomic 関数とするとき $\mathrm{Ann}\,(f) = \{\ell \in D \,|, \ell \bullet f = 0\}$ を f の零化イデアルとよぶ．$D/\mathrm{Ann}\,(f)$ は holonomic である．また, このイデアルは上記の常微分方程式 (1) に現れる形の作用素を含むが, 零化イデアルがこれらの作用素で生成されるとは限らないことが holonomic の概念の妙である．generic な点の近傍で holonomic 系のなめらかな解全体はベクトル空間となる．このベクトル空間の次元を holonomic 系の rank と呼ぶ.

常微分方程式は連立一階の常微分方程式に書換えられる．このことから推察できるように holonomic 関数 f に対して, f とその適当な階数までの微分, あわせて $r = \mathrm{rank}$ 個をベクトル F (c 基底ベクトルと呼ぶ) とすると, 有理式成分の $r \times r$ 行列 $P_i(x)$ が存在して,

$$\partial_i \bullet F = P_i(x)F, i = 1, \ldots, n \tag{2}$$

が成り立つ．この方程式系を Pfaffian 系とよぶ [2, §6.2].

本論文は筆者がある程度関わった HGM 関連の仕事に限定した概説である．本概説で触れてない分布等に関しては [1] を参照してほしい．紙面の都合で文献はかならずしもオリジナルを引用していない．必要に応じて文献表からの孫引きの労をお願いしたい.

2. Holonomic 分布

2.1. 分布と正規化定数

$f(\theta, t)$ を $n+m$ 個の変数 $\theta = (\theta_1, \ldots, \theta_n)$, $t = (t_1, \ldots, t_m)$ についての holonomic 関数とする．この関数は $t \in \mathbf{R}^m$ 全体で積分可能であるとする．このとき $Z(\theta) = \int_{\mathbf{R}^m} f(\theta, t) dt$ を正規化定数とする t 空間の分布 $\frac{1}{Z} f(\theta, t) dt$, $dt = dt_1 \cdots dt_n$ を holonomic 分布と呼ぶ．ここで θ は分布のパラメータである．

2.2. Holonomic system

holonomic 系の一般論より積分についてのある条件のもと, $Z(\theta)$ もまた holonomic 系を満す [2, §6.10]. $\mathrm{Ann}\, f(\theta, t)$ の生成元がわかっている場合, Z の満す holonomic 系を求めるアルゴリズムの一般形は大阿久の積分アルゴリズム (1997) である [2, §6.10]．より制限された状況でより高速に Z の満す微分方程式系を求めるアルゴリズムは creative telescoping 型アルゴリズムとして, さまざまな研究が行なわれている．

準備で述べた (ヘビサイド関数) $\times$ (holonomic 解析関数) なる形の holonomic 関数に対する holonomic 系を求める一般アルゴリズムは [14] で与えられている．

2.3. Pfaffian system

有理式係数の微分作用素環で Buchberger アルゴリズムを適用すれば Pfaffian 系を求めることができる．また正規化定数の Hessian なども Pfaffian 系の副産物として得られる [13, Appendix], [2, §6.2, §6.5].

2.4. アルゴリズムと実装

HGM の一般論をはじめて展開したのは [13] である．一般論の HGM では次の 3 step で正規化定数およびその微分達を数値計算する．

1. $f(\theta, t)$ の t についての積分 $Z(\theta)$ の満す holonomic 系を上記アルゴリズム達で求める．さらにそれを Buchberger アルゴリズムで Pfaffian 系に変換する．
2. 問題に応じた工夫をおこないうまい点 θ_0 で Z および必要なランクまでの Z の偏微分を数値計算する．数値積分すればいつでも近似値を計算できるがこれは最後の手段.
3. Pfaffian 系は常微分方程式系なので Runge-Kutta 法などの標準的な数値解析手法を用いて θ_0 での値を初期値として, 求めたい点 θ まで延長する．

Step 1 を遂行する一般的なパッケージとしては, `Risa/Asir` の `nk_restriction.rr` や `Macaulay 2` のパッケージ `Dmodules` がある．例については [2, 7 章] を参照．Creative Telescoping 型では Mathematica の HolonomicFunction パッケージや Maple の MGfun がある．Step 3 については我々は GSL や統計システム R の deSolve パッケージを利用している．残念ながら一般的なアルゴリズムでは解ける問題のサイズが限定されている．Step 1 についても, 問題に応じた工夫をして, 興味あるサイズの問題を解けるようにしたい．このような工夫が成功しているのが, 以下のページからはじまる分布達である．

3. Fisher-Bingham 分布, Bingham 分布

3.1. 分布と正規化定数

Fisher-Bingham 分布は d 次元の球面 $S^d(r) = \{(t_1, \ldots, t_{d+1}) \mid \sum_{i=1}^{d+1} t_i^2 = r^2, r > 0\}$ に多次元の正規分布を制限したものである. 分布密度関数は $\frac{1}{Z(x,y,r)} \exp\left(\sum_{1 \le i \le j \le d+1} x_{ij} t_i t_j + \sum_{i=1}^{d+1} y_i t_i\right) |dt|$ である. ここで Z は正規化定数で,

$$Z(x,y,r) = \int_{S^d(r)} \exp\left(\sum_{1 \le i \le j \le d+1} x_{ij} t_i t_j + \sum_{i=1}^{d+1} y_i t_i\right) |dt| \tag{3}$$

また $|dt|$ は半径 r の球面上の不変測度で $\int_{S^d(r)} |dt| = r^d \frac{2\pi^{(d+1)/2}}{\Gamma((d+1)/2)}$ を満たすものとする. Bingham 分布は Fisher-Bingham 分布をさらに $r = 1$, $y_i = 0$, $x_{ij} = 0$, $i \neq j$ へ制限したものである. [19] では x_{ii} を θ_i と記述している.

3.2. Holonomic system

$\partial_{ij} = \partial/\partial x_{ij}$, $\partial_k = \partial/\partial y_k$, $\partial_r = \partial/\partial r$, とおくと Z は次の holonomic 系を満たす [13].

$$\partial_{ij} - \partial_i \partial_j, \quad \sum_{i=1}^{d+1} \partial_i^2 - r^2, \quad r\partial_r - 2\sum_{i \le j} x_{ij} \partial_i \partial_j - \sum_i y_i \partial_i - d,$$
$$x_{ij}\partial_i^2 + 2(x_{jj} - x_{ii})\partial_i \partial_j - x_{ij}\partial_j^2 + \sum_{s \neq i,j} (x_{sj}\partial_i\partial_s - x_{is}\partial_j\partial_s) + y_j\partial_i - y_i\partial_j.$$

Bingham 分布の正規化定数は以下の holonomic 系を満たす.

$$\sum_{i=1}^{d} \partial_{ii} - 1, \quad 2(x_{ii} - x_{jj})\partial_{ii}\partial_{jj} - (\partial_{ii} - \partial_{jj}), \ (1 \le i < j \le d)$$

3.3. Pfaffian system

$x_{ij} = 0$, $i \neq j$ に制限した場合, r 方向の Pfaffian は $\tilde{F} = (\partial_1, \ldots, \partial_{d+1}, \partial_1^2, \ldots, \partial_{d+1}^2)^T \bullet Z$ とおいたとき, $\partial_r \bullet \tilde{F} = P^{(r)}\tilde{F}$ となる [8]. $(2d+2) \times (2d+2)$ 行列 $P^{(r)} = (p_{ij}^{(r)})$ の ij 成分は

$$\begin{aligned} rp_{ij}^{(r)} &= (2x_{ii}r^2 + 1)\delta_{ij} + \sum_{k=1}^{d+1} y_i \delta_{j(k+d+1)} \quad (1 \le i \le d+1), \\ rp_{(i+d+1)j}^{(r)} &= y_i r^2 \delta_{ij} + (2x_i r^2 + 2)\delta_{j(i+d+1)} + \sum_{k \neq i} \delta_{j(k+d+1)} \quad (1 \le i \le d+1) \end{aligned}$$

Bingham 分布の正規化定数については [19, Theorem 1] を参照.

3.4. アルゴリズムと実装

Bingham 分布の正規化定数は R パッケージ hgm の関数 `hgm.ncBingham` が計算する. Fisher-Bingham 分布の正規化定数については [10] で高速な数値計算法が与えられている. その試験実装は [11] で公開されている.

3.5. 例と応用

引用の論文を参照.

4. Wishart 分布の最大固有値

4.1. 分布と正規化定数

自由度 n, $m \times m$ 共分散行列 Σ できまる Wishart 分布に従う $m \times m$ 行列の最大固有値 ℓ_1 が x より小さい確率は C を定数として [5]

$$P[\ell_1 < x] = C \exp\left(-\frac{x}{2}\mathrm{Tr}\,\Sigma^{-1}\right) x^{\frac{1}{2}nm} {}_1F_1\left(\frac{m+1}{2}; \frac{n+m+1}{2}; \frac{x}{2}\Sigma^{-1}\right), \tag{4}$$

と書けることが知られている. ${}_1F_1$ は行列引数の超幾何関数で次の積分で定義される.

$${}_1F_1(a;c;Y) = \frac{\Gamma_m(c)}{\Gamma_m(a)\Gamma_m(c-a)} \int_{0<X<I_m} \exp(\mathrm{Tr}\,XY)|X|^{a-(m+1)/2}|I_m - X|^{c-a-(m+1)/2} dX, \tag{5}$$

なお $0 < X < I_m$ は行列 X および $I_m - X$ が正定値対称行列であることを意味する. $dX = \prod_{i \le j} dx_{ij}$ は X の 上三角成分に渡る積分を意味する. また

$$\Gamma_m(a) = \pi^{\frac{1}{4}m(m-1)} \prod_{i=1}^{m} \Gamma\left(a - \frac{i-1}{2}\right).$$

4.2. Holonomic system

${}_1F_1(a;c;Y)$ の Y が対角成分が $y_1, \ldots, y_m$ である対角行列であるとする. Muirhead (1970) は ${}_1F_1$ が満たす次の微分方程式を導いた.

$$\left[y_i\,\partial_i^2 + \left\{ c - \frac{m-1}{2} - y_i + \frac{1}{2}\sum_{j=1, j\neq i}^{m} \frac{y_i}{y_i - y_j} \right\} \partial_i - \frac{1}{2}\sum_{j=1, j\neq i}^{m} \frac{y_j}{y_i - y_j}\partial_j - a \right] F = 0, \quad (i = 1, \ldots, m)$$

なおこの方程式系は Pfaffian を求めるためには十分であるが, $\prod_{1 \le i < j \le m}(y_i - y_j)$ をかけて多項式係数にした方程式系の holonomic 性は $m \ge 4$ で壊れていると予想されている.

4.3. Pfaffian system

[5] では上記の微分方程式系を sparsity をもつ Pfaffian へ変換して数値解析に活用している.

4.4. アルゴリズムと実装

x が小さい時の初期値の近似値を zonal 多項式を用いた ${}_1F_1$ の展開で計算し, それを Pfaffian 方程式の数値解析でより大きな x へ延長していくことにより上記の確率およびそれの微分達を計算している. 現状のアルゴリズムの限界: $y_i = y_j$ の近傍, y_i の値が極端に小さいものと大きいものが混在している場合は Pfaffian から導かれる常微分方程式の数値解析が不安定.

4.5. 例と応用

統計システム R のパッケージ hgm に実装されている. 下記で `beta` は Σ の逆行列の固有値達のベクトルである. `q` は上の x である.

```
library(hgm)
hgm.pwishart(m=3,n=5,beta=c(1,2,3),q=10)
b<-hgm.pwishart(m=4,n=10,beta=c(1,2,3,4),q0=1,q=10,approxdeg=20,mode=c(1,1,(16+1)*100));
## $q$ を増やしていった時の確率の graph を書くために option mode
## を使い途中の確率も戻す. 16 は 2^m で Pfaffian 系の rank である.
cc<-matrix(b,ncol=16+1,byrow=1);
plot(cc)
```

5. 象限確率

5.1. 分布と正規化定数

正規分布を第一象限 $t_i \geq 0$ へ制限した分布を考える. 上記分布の正規化定数, つまり

$$Z(x,y) = \int_{\mathbf{R}^d} H(t) \exp\left(\sum_{1 \leq i \leq j \leq d} x_{ij} t_i t_j + \sum_{i=1}^{d} y_i t_i \right) dt_1 \cdots dt_d \tag{6}$$

を象限確率とよぶ. ここで $H(t)$ は 第一象限で値 1, その外で値 0 を持つヘビサイド関数である.

5.2. Holonomic system

[9, Theorem 2, Theorem 5] で $H(t)$ が一般の holonomic 超関数の場合の holonomic 系が以下のように与えれている. $P_1, \ldots, P_s$ が H に対する holonomic 系とすると,

$$\begin{array}{ll} \varphi(P_k) & (1 \leq k \leq s), \quad \varphi \text{ の定義は [9] 参照} \\ \partial_{x_{ij}} - 2\partial_{y_i}\partial_{y_j} & (1 \leq i < j \leq d), \\ \partial_{x_{ii}} - \partial_{y_i}^2 & (1 \leq i \leq d) \end{array}$$

が上記積分 Z が一般の holonomic 超関数 $H(t)$ の場合に満たす holonomic 系となる.

5.3. Pfaffian system

$J \subset [d] = \{1, \ldots, d\}$ に対して

$$\begin{aligned} h_J(x,y,t) &= \sum_{i \in J} \sum_{j \in J} x_{ij} t_i t_j + \sum_{k \in J} y_k t_k, \\ g_J(x,y) &= \int_0^\infty \cdots \int_0^\infty \exp\left(h_J(x,y,t)\right) dt_J, \end{aligned}$$

とおく. ここで $dt_J = \prod_{j \in J} dt_j$ であり J が空集合の時は $g_\emptyset = 1$ とおく. Pfaffian 系は次の式である.

$$\partial_{y_i} g_J = \begin{cases} \mu_i^J g_J + \sum_{j \in J} \sigma_{ij}^J g_{J \setminus \{j\}} & i \in J \\ 0 & i \notin J \end{cases} \tag{7}$$

$$\partial_{x_{ij}} g_J = \begin{cases} 2\partial_{y_i}\partial_{y_j} g_J & \{i,j\} \subset J,\ i < j \\ \partial_{y_i}^2 g_J & \{i\} \subset J,\ i = j \\ 0 & else. \end{cases} \tag{8}$$

5.4. アルゴリズムと実装

統計システム R のパッケージ hgm の関数 `hgm.ncorthant` が正規化定数を評価する. 現状のアルゴリズムの限界: y が原点から離れると HGM による計算精度を保つのがだんだんと困難になる.

5.5. 例と応用

```
x <- matrix(c(15,26,23,19,  26,47,46,35,
       23,46,65,38,  19,35,38,33), nrow =4)
y <- c(1,2,3,4); hgm.ncorthant(x,y)
```

6. Fisher 分布

6.1. 分布と正規化定数

$SO(p) = \{X \in \mathbf{R}^{p\times p} \mid X^T X = I_p, \det(X) = 1\}$ の上の Fisher 分布は $\exp(\mathrm{Tr}\ \Theta^\top X)\delta_{SO(n)}(X)/Z(\Theta)$ と定義する．ここで，$\delta_{SO(p)}(X)$ は $SO(p)$ に台をもつデルタ関数，$\Theta = (\theta_{ij})$，Z は正規化定数で，

$$Z(\Theta) = \int_{X=(X_{ij})\in\mathbf{R}^{p\times p}} \exp(\mathrm{Tr}\ \Theta^\top X)\delta_{SO(p)}(X)dX$$

6.2. Holonomic system

$$A_{ij}^{(1)} = \sum_{k=1}^{p} \partial_{ik}\partial_{jk} - \delta_{ij}, \quad \tilde{A}_{ij}^{(1)} = \sum_{k=1}^{p} \partial_{ki}\partial_{kj} - \delta_{ij} \quad (i \le j), \quad A^{(2)} = \det(\partial_{ij}) - 1,$$

$$A_{ij}^{(3)} = \sum_{k=1}^{p} (-\theta_{jk}\partial_{ik} + \theta_{ik}\partial_{jk}), \quad \tilde{A}_{ij}^{(3)} = \sum_{k=1}^{p} (-\theta_{kj}\partial_{ki} + \theta_{ki}\partial_{kj}) \quad (i < j),$$

ここで $\partial_{ij} = \partial/\partial\theta_{ij}$．この系が holonomic であることは [12] で証明されている．

この系を Θ が対角行列の場合に制限した方程式系は重要である．$p = 3$ の時は

$$\ell_{ij} = (\theta_{ii}^2 - \theta_{jj}^2)\partial_{ii}\partial_{jj} - (\theta_{jj}\partial_{ii} - \theta_{ii}\partial_{jj}) - (\theta_{ii}^2 - \theta_{jj}^2)\partial_{kk}, \quad i \ne j$$

が制限方程式系である [18]．なお k は $\{i, j, k\} = \{1, 2, 3\}$ となるように選ぶ．

6.3. Pfaffian system

$p > 3$ の時の Pfaffian は知られていない．$p = 3$ で Θ が対角行列の場合の Pfaffian は [18, Theorem 2] に与えられている．特に，θ_{ii} を定数 a, b, c で $\theta_{11} = at$, $\theta_{22} = bt$, $\theta_{33} = ct$ とパラメータづけした場合の，t についての常微分方程式は

$$\frac{dC}{dt} = \left(A - \frac{2}{t}\mathrm{diag}(0, 1, 1, 1)\right)C, \tag{9}$$

となる．ここで $A = [[0, a, b, c], [a, 0, c, b], [b, c, 0, a], [c, b, a, 0]]$ であり，[*1] $C = ((1, \partial_{11}, \partial_{22}, \partial_{33})^T \bullet Z)(at, bt, ct)$．

6.4. アルゴリズムと実装

正規化定数のべき級数展開が [18] で与えられている．θ_{ij} が小さい時，このべき級数で正規化定数およびその微分を計算して，常微分方程式 (9) で値を延長する．R パッケージ hgm の関数 `hgm.ncso3` がこの計算法を実装している．

6.5. 例と応用

```
 hgm.ncso3(a=1,b=2,c=3,t0=0,q=3);
[1] 11.8731
```

値は (9) で $(a, b, c) = (1, 2, 3)$, $t = \mathtt{q} = 3$ の時の正規化定数の値である．

---

*1 なおこの概説では行列 (a_{ij}) をコンパクトに記述するため $[[a_{11}, a_{12}, \cdots], [a_{21}, a_{22}, \ldots], \ldots]$ と計算代数ソフトウエア風の書き方をする場合がある．

7. 離散 A-超幾何分布

7.1. 分布と正規化定数

整数成分の $d \times n$ 行列 $A = (a_{ij})$ を考える．A の第 i 列ベクトル a_i を $\mathbf{Z}^d$ の元とみたとき，この列ベクトル達は $\mathbf{Z}^d$ を生成していると仮定する．ある行 i について $a_{ij} > 0,\ \forall j$ と仮定しよう．$\mathbf{N}_0$ を非負整数の集合とし，$\beta \in \mathbf{N}_0 A = \mathbf{N}_0 a_1 + \cdots + \mathbf{N}_0 a_n$ に対して x の多項式

$$Z_A(\beta; x) = \sum_{Au=\beta, u \in \mathbf{N}_0^d} \frac{x^u}{u!} \tag{10}$$

を A-超幾何多項式とよぶ [17]．x をパラメータ，$\{U \in \mathbf{N}_0^n \,|, AU = \beta\}$ を標本空間 (全事象) としたとき，事象 u がおきる確率が $\frac{x^u}{u!}/Z_A(\beta; x)$ となる分布を離散 A-超幾何分布とよぶ．

7.2. Holonomic system

$c_1, \ldots, c_d$ を不定元としよう．S_i を $S_i \bullet f(c_i) = f(c_i - 1)$ と作用する差分作用素．微分差分作用素環 $D = \mathbf{C}\langle c_1, \ldots, c_d, S_1, \ldots, S_d\rangle\langle x_1, \ldots, x_n, \partial_1, \ldots, \partial_n\rangle$ において

$$\sum_{j=1}^{n} a_{ij} x_j \partial_j - c_i =: E_i - c_i, \quad \partial_i - S^{a_i} \qquad (i = 1, \ldots, d) \tag{11}$$

は $Z_A(c; x)$ を解にもつ．

7.3. Pfaffian system

A が p-単体と q-単体の直積の場合は $p \times q$ 分割表で周辺和を固定したものの分布が対応する．$p \times q$ 分割表の場合は, [15], [4] および後藤の準備中の論文に Pfaffian の具体形が与えれている．超幾何多項式の数値計算は項の数が増えると急速に困難になる．一般の A の場合は [16] によるアルゴリズムが漸近的に最速．

7.4. アルゴリズムと実装

Macaulay type algorithm による微分差分 Pfaffian の効率的構成法が研究 [6], [16] および実装 (`Risa/Asir`, `ot_hgm_ahg.rr`) されている．$2 \times q$ 分割表の時の実装は `tk_fd.rr`, $p \times q$ 分割表の場合は後藤が公開準備中．

7.5. 例と応用

MLE 等統計への応用については [15]．u_i の期待値は, $x_i \frac{\partial Z}{\partial x_i}$ となるので, HGM で下記のように期待値が計算できる．

```
load("ot_hgm_ahg.rr");
A= [[1,1,1,1,1,0,0,0,0,0],
    [1,0,0,0,0,1,0,0,0,1],
    [0,1,0,0,0,1,1,0,0,0],
    [0,0,1,0,0,0,1,1,0,0],
    [0,0,0,1,0,0,0,1,1,0],
    [0,0,0,0,1,0,0,0,1,1]];
U= [23,21,20,34,30,42,11,10,20,29];
B=vtol(matrix_list_to_matrix(A)*ltov(U)); SumU = oh_base.sum(U);
X = vtol((1/SumU)*ltov(U)); R = [x1,x2];//u1,u2 についての期待値を計算.
EV=hgm_ahg_expected_values_contiguity(A,B,X,R|msize=3);
```

8. 多面体確率

8.1. 分布と正規化定数

正規分布を多面体に制限した分布. [7] がこの分布についての HGM の基本文献. 多面体が

$$P = \{x \in \mathbf{R}^d : \sum_{i=1}^{d} a_{ij} x_i + b_j \geq 0,\, 1 \leq j \leq n\}, \tag{12}$$

とパラメーター a_{ij}, b_j で定義されるとき正規化定数は以下の関数.

$$\frac{1}{(2\pi)^{d/2}} \int_{x \in P} \exp(-\frac{1}{2} \sum_{i=1}^{d} x_i^2) dx_1 \cdots dx_d. \tag{13}$$

8.2. Holonomic system

Pfaffian 形が知られているので省略.

8.3. Pfaffian system

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.

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

ここで $(\alpha_F^{ij}(a))_{i,j \in F}$ は a から作った行列 $\alpha_F(a)$ の逆行列の ij 成分. $\alpha_F(a)$ の ij 成分は $\sum_{k=1}^{d} a_{ki} a_{kj}$. これが Pfaffian の係数の分母に来ることに注意.

8.4. アルゴリズムと実装

P が単体の場合は [7] の著者小山による試験実装が [11] で公開されている. 初期値の決め方等の論文は準備中だがこの実装にはそれがすでに含まれている.

8.5. 例と応用

$a = [[1, 0], [0, 1], [-1, -1]], b = [0, 0, 2]$ の場合 $\mathcal{F} = \{\emptyset, 1, 2, 3, 12, 13, 23\}$ となる. ここでたとえば 12 は index 集合 $\{1, 2\}$ を表すものとする. この三角形 $x > 0, y > 0, -x - y + 2 > 0$ へ正規分布を制限した積分を上記の実装で計算すると下記のようになる. (なお多面体が空となるとプログラムが正しく動作しないので注意.)

```
./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
```

参考文献

[1] http://www.math.kobe-u.ac.jp/OpenXM/Math/hgm/ref-hgm.html HGM 関連文献リスト.

[2] JST CREST 日比チーム編, グレブナー道場, 共立出版, 2011.

[3] 日比, 竹村, 原, 東口, 清, グレブナー道場著者一同, グレブナー教室 — 計算代数統計学への招待, 共立出版, 2015.

[4] Y.Goto, Contiguity Relations of Lauricella's F_D Revisited, arxiv:1412.3256.

[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, K. Nishiyama, N. Takayama, Pfaffian Systems of A-Hypergeometric Systems I — Bases of Twisted Cohomology Groups, arxiv:1212.6103

[7] T.Koyama, Holonomic Modules Associated with Multivariate Normal Probabilities of Polyhedra, arxiv:1311.6905, to appear in Funkcialaj Ekvacioj

[8] T. Koyama, H. Nakayama, K. Nishiyama, N. Takayama, Holonomic Gradient Descent for the Fisher-Bingham Distribution on the d-dimensional Sphere, Computational Statistics (2013)

[9] T.Koyama, A.Takemura, Calculation of Orthant Probabilities by the Holonomic Gradient Method, arxiv:1211.6822

[10] T.Koyama, A.Takemura, Holonomic Gradient Method for Distribution Function of a Weighted Sum of Noncentral Chi-square Random Variables, arxiv: 1503.00378

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

[12] T.Koyama, The Annihilating Ideal of the Fisher Integral, arxiv: 1503.05261

[13] 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.

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

[15] M.Ogawa, Algebraic Statistical Methods for Conditional Inference of Discrete Statistical Models, PhD. Thesis, University of Tokyo, 2015.

[16] K. Ohara, N. Takayama, Pfaffian Systems of A-Hypergeometric Systems II — Holonomic Gradient Method, arxiv:1505.02947.

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

[18] T.Sei, H.Shibata, A.Takemura, K.Ohara, N.Takayama, Properties and applications of Fisher distribution on the rotation group, Journal of Multivariate Analysis, 116 (2013), 440–455.

[19] T.Sei, A.Kume, Calculating the Normalising Constant of the Bingham Distribution on the Sphere using the Holonomic Gradient Method, Statistics and Computing 2013.

[20] 高山, D 加群の積分アルゴリズムと推定理論. 数学セミナー 2012 年 2 月号, 41–46.

謝辞: 原稿を読み有益なコメントをよせて頂いた, 小山民雄氏, 清智也氏に感謝したい.