

微分作用素環のグレブナー基底と統計学への応用

小原 功任 (金沢大学理工研究域)

1 計算代数システムと微分作用素環

斉次線形偏微分方程式系を扱うのに、微分作用素からなる非可換環や層を考え、それらを係数とする加群やイデアルと微分方程式系を同一視する方法がある。計算機の発達にともない、そのような数学的対象を計算機で厳密に計算したいというのは、自然な考えであると思われる。そのような操作が行えるシステムはいくつか存在するが、ここでは、著者も開発に参加している Risa/Asir を取り上げる。Risa/Asir は、野呂正行氏 (現・立教大学) を中心に富士通で開発された国産の計算代数システムであり、現在はフリーソフトウェアとして、有志によって開発・改良が進められている。Risa/Asir にはさまざまな機能が実装されているが、特に多項式環上のイデアルのグレブナー基底計算が高速であることはよく知られている。また、グレブナー基底計算によって実現される、多項式イデアルの準素イデアル分解や加群のシチジーなどの有用な計算もサポートされている。

さて、微分作用素環などの非可換環にはさまざまなものがあるが、Risa/Asir においては、それぞれの環の計算はそれぞれ実現しているプログラムが異なり、統一的な枠組みは存在しない。そこでそれぞれの環に対する対応状況をみていこう。

もっとも基本的な非可換環はワイル代数 D_n である。Risa/Asir では D_n における基本演算とグレブナー基底計算については、C 言語で書かれた組み込み関数が用意されており、高速に実行可能である。 D_n 加群およびイデアルの制限と積分、零化イデアル、大域 $\mathfrak{m}$ 関数、特異点集合など、応用上重要な計算が実行可能である。ワイル代数を拡張したものとしては、Poincaré-Birkhoff-Witt 代数がある。この環ではグレブナー基底計算が実行可能である。ユーザー言語による拙作のプログラムがあり、零化イデアル、大域 $\mathfrak{m}$ 関数を計算する Briançon-Maisonobe の算法が実装されている。

ワイル代数を拡張したものとしては、有理関数係数の微分作用素環 $R_n = \mathbb{C}(x)\langle \partial_1, \dots, \partial_n \rangle$ と、その拡張である、微分差分作用素環がある。ユーザー言語で書かれたプログラム (yang.rr, 拙作) が利用できる。 R_n で可能な計算としては、微分方程式系のホロノミックランクの計算や、Pfaffian 方程式の導出がある。

これら以外のものとしては、収束巾級数環係数の微分作用素環 $\mathbb{C}\{x\}\langle \partial_1, \dots, \partial_n \rangle$ が重要である。この環における、いわゆる Mora の割算が、東海大学の中山洋将氏によって実装されており、グレブナー基底および局所 $\mathfrak{m}$ 関数も計算可能である。

また、応用上は parametric な計算が必要になることがある。これは、parameter を generic だと仮定して実行する計算ではなく、parameter にいかなる値が入っていても正しい計算のことである。グレブナー基底においては、包括的グレブナー基底 (comprehensive Gröbner bases) もしくは、包括的グレブナー基底系 (comprehensive Gröbner system) の概念が対応する。多項式環やワイル代数などで CGS を計算するプログラムとしては、徳島大学の鍋島克輔氏によるものがある。

2 統計学への応用: 最尤推定とホロノミック勾配法

以下, Fisher 分布の場合を例にとり, 微分作用素環のグレブナー基底理論の統計学への応用について簡単に述べる. 詳細については, 文献 [1, 2, 4] を参照していただきたい.

パラメータ (母数) Θ を 3 次実正方行列とし, 3 次特殊直交群 $SO(3)$ 上の関数 $f(\Theta, X) = \exp(\text{tr}(\Theta^T X))$ を考える ($X \in SO(3)$). 全空間 $SO(3)$ における積分 (正規化定数) は

$$c(\Theta) = \int_{SO(3)} \exp(\text{tr}(\Theta^T Y)) d\mu(Y)$$

であり, Fisher 積分とも呼ばれる. このとき, 比 $f(\Theta, X)/c(\Theta)$ は $SO(3)$ 上の確率分布とみなすことができる. この分布を Fisher 分布という.

母数の分からない確率分布 $f(\Theta, X)/c(\Theta)$ と, 標本 $X_1, \dots, X_n$ が与えられたとき, 尤度 $\prod_{i=1}^n \frac{f(\Theta, X_i)}{c(\Theta)}$ が最大となるような母数 Θ をみつきたい. この問題を最尤推定といい, 統計学における基本的問題のひとつである.

Fisher 分布の場合, 標本は $X_1, \dots, X_n \in SO(3)$ である. また尤度関数の n 乗根は

$$\sqrt[n]{\prod_{i=1}^n \frac{f(\Theta, X_i)}{c(\Theta)}} = \frac{f(\Theta, \bar{X})}{c(\Theta)}, \quad \bar{X} = \frac{X_1 + \dots + X_n}{n} \in M_3(\mathbb{R})$$

となる. したがって尤度が最大となる点は $c(\Theta)f(\Theta, -\bar{X})$ の最小点を与える. この問題を解くためには, Fisher 積分 $c(\Theta)$ の関数としての振る舞いを知る必要がある. しかしながら, $f(\Theta, X)$ に比べて $c(\Theta)$ はずっと難しい関数である.

ところで数値解析において関数の最小点を数値的に求める手法として, 最急勾配降下法と呼ばれる方法がよく知られている. この方法を適用するには, 任意の直線上で, 与えられた関数の各点での値と勾配を評価する必要がある. Fisher 分布の場合にはこの点が難しい. $c(\Theta)$ の性質をうまく用いて, この部分を解決するのがホロノミック勾配法である.

$f(\Theta, X)$ は, Θ および X の各成分について, 線形常微分方程式を満たす. このような関数はホロノミック関数と呼ばれる. ホロノミック関数のサイクル上での積分は, またホロノミック関数であることが知られている. したがって, Fisher 積分もまたホロノミック関数である. 原理的には, $f(\Theta, X)$ の満たす線形常微分方程式系 (つまりワイル代数における零化イデアル) から, $c(\Theta)$ の満たす線形常微分方程式系 (のグレブナー基底) を求めることができる. この方法は大阿久俊則氏によって開発されたもので, 積分アルゴリズムと呼ばれ, Risa/Asir にも実装がある. ホロノミック関数に対しては, 適当な個数 (ホロノミックランクという) の導関数を並べたベクトル値関数 $F(t)$ を考えると ($t = (t_1, \dots, t_m)$ は独立変数), 一階偏微分方程式系 $\frac{\partial}{\partial t_i} F = P_i(t)F$, ($i = 1, \dots, m$) に変形できる (Pfaffian 方程式). 各方向をすべて含んでいるので, 直線に制限すると, 関数値と勾配を与える常微分方程式となる. よって, Runge-Kutta 法などの数値解法により, 数値的に関数値と勾配が求まる.

原理的には以上の方法で, 最尤推定ができるはずである. しかしながら実際には, 主として計算機の能力の問題でそのままではこのプランは実行することはできない (グレブナー基底や Pfaffian 方程式が計算できない). 次は, Fisher 積分の場合にわれわれがとった方法について述べよう.

3 Fisher 積分の満たす微分方程式

実は, Fisher 積分は, 行列変数の超幾何級数 ${}_0F_1$ によって, $c(\Theta) = {}_0F_1(\frac{3}{2}; \frac{1}{4}\Theta^T \Theta)$ と表される. 行列変数の超幾何級数はゾーナル多項式の級数であり, 単純な巾級数ではない. 行列変数の

超幾何級数 ${}_0F_1$ は一般に次の方程式を満たすことが知られている (Muirhead, 1970).

定理 1. 対称行列 Z の固有値を $z_1, \dots, z_m$ で表すと, $u = {}_0F_1(\gamma; Z)$ は微分方程式系

$$L_i u = 0 \quad (i = 1, \dots, m)$$

を満たす. ここで,

$$L_i = z_i \frac{\partial^2}{\partial z_i^2} + \left\{ \gamma - \frac{m-1}{2} + \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^m \frac{z_i}{z_i - z_j} \right\} \frac{\partial}{\partial z_i} - \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^m \frac{z_j}{z_i - z_j} \frac{\partial}{\partial z_j} - 1.$$

この方程式系を用いることの問題点は, (1) 固有値に関する方程式である (2) 超幾何級数のパラメータが半整数 $3/2$ のときは, 級数の定義から退化するはずである, の 2 点である. したがって, この方程式系は $c(\Theta)$ の零化イデアルの生成系を与えていない可能性がある. なお, この方程式系のホロノミックランクは 8 である.

一方, 積分に対する考察から, 次の方程式系が零化イデアルを与える.

定理 2. $c(\Theta)$ は微分方程式系

$$A_{ij}^{(k)} c(\Theta) = \tilde{A}_{ij}^{(k)} c(\Theta) = A^{(2)} c(\Theta) = 0.$$

を満たす. ここで,

$$\begin{aligned} A_{ij}^{(1)} &= \sum_{k=1}^3 \partial_{ik} \partial_{jk} - \delta_{ij}, \quad \tilde{A}_{ij}^{(1)} = \sum_{k=1}^3 \partial_{ki} \partial_{kj} - \delta_{ij} \quad (i \leq j), \\ A^{(2)} &= \det(\partial_{ij}) - 1, \\ A_{ij}^{(3)} &= \sum_{k=1}^3 (-\theta_{jk} \partial_{ik} + \theta_{ik} \partial_{jk}), \quad \tilde{A}_{ij}^{(3)} = \sum_{k=1}^3 (-\theta_{kj} \partial_{ki} + \theta_{ki} \partial_{kj}) \quad (i < j), \end{aligned}$$

ただし, $\Theta = (\theta_{ij})_{ij}$, $\partial_{ij} = \frac{\partial}{\partial \theta_{ij}}$, δ_{ij} : クロネッカーのデルタ

この方程式系は, ワイル代数 D_9 の左イデアルとみなすことができるが, 現在の計算機ではグレブナー基底を求めることはできそうにない (18 変数多項式環と同程度かそれ以上の計算量と推測できる). したがって, そのままではホロノミック勾配法は適用できない.

さて, Fisher 分布の問題では, 標本平均行列 X の特異値分解について次の補題が成り立つ.

補題 1. $P, Q \in \text{SO}(3)$ とする. $X = P \hat{X} Q^T$ が特異値分解であるとき, $f(\Theta, X)/c(\Theta)$ の最大値を与える Θ は, ある対角行列 $\hat{\Theta}$ を用いて $\Theta = P \hat{\Theta} Q^T$ と表される.

したがって, 対角行列の全体 $\{\Theta \mid \theta_{ij} = 0 \ (i \neq j)\}$ の上への, 制限イデアルがわかれば問題を小さくできる. 制限イデアルの計算アルゴリズムはよく知られていて, Risa/Asir にも実装があるが, しかし, その計算にはやはりグレブナー基底が必要である. われわれはグレブナー基底を使わずに, 次の定理を得た.

定理 3. 集合 $\{\Theta \mid \theta_{ij} = 0 \ (i \neq j)\}$ の上への, D -加群としての制限イデアルは, $I = \langle \ell_{ij} \mid 1 \leq i \leq j \leq 3 \rangle$ を含む. ここで

$$\begin{aligned} \ell_{ii} &= \prod_{k \neq i} (\theta_{ii}^2 - \theta_{kk}^2) \left(\partial_{ii}^2 + \sum_{j \neq i} \frac{1}{\theta_{ii}^2 - \theta_{jj}^2} (\theta_{ii} \partial_{ii} - \theta_{jj} \partial_{jj}) - 1 \right), \quad (i = 1, 2, 3) \\ \ell_{ij} &= (\theta_{ii}^2 - \theta_{jj}^2) \partial_{ii} \partial_{jj} + \theta_{ii} \partial_{jj} - \theta_{jj} \partial_{ii} - (\theta_{ii}^2 - \theta_{jj}^2) \sum_{k \neq i, j} \partial_{kk}. \quad (1 \leq i < j \leq 3) \end{aligned}$$

である. さらに, イデアル I のホロノミックランクは 4 である.

この定理は制限イデアルの部分イデアルを与える (グレブナー基底がないので, ここまでしか示せない). また Muirhead の微分方程式を含むことも容易に分かる. 以下, $X = \text{diag}(x_{11}, x_{22}, x_{33})$, $\Theta = \text{diag}(\theta_{11}, \theta_{22}, \theta_{33})$ であると仮定し, I を $R_3 = \mathbb{C}(\theta_{ii})\langle \partial_{11}, \partial_{22}, \partial_{33} \rangle$ の左イデアルとみなす. このグレブナー基底は計算可能である. よって $c(\Theta)$ の Pfaffian 方程式が具体的に求まる.

定理 4. ベクトル値関数 $c(\Theta) = \left(c, \frac{\partial c}{\partial \theta_{11}}, \frac{\partial c}{\partial \theta_{22}}, \frac{\partial c}{\partial \theta_{33}} \right)^\top$ は一階偏微分方程式系

$$\frac{\partial c}{\partial \theta_{ii}} = P_i(\Theta) c, \quad (i = 1, 2, 3)$$

を満す. ここで, $u_{ij} = \theta_{jj}/(\theta_{ii}^2 - \theta_{jj}^2)$ とおくと

$$P_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & u_{21} + u_{31} & u_{12} & u_{13} \\ 0 & u_{12} & u_{21} & 1 \\ 0 & u_{13} & 1 & u_{31} \end{pmatrix}, \quad P_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & u_{12} & u_{21} & 1 \\ 1 & u_{21} & u_{12} + u_{32} & u_{23} \\ 0 & 1 & u_{23} & u_{32} \end{pmatrix},$$

$$P_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & u_{13} & 1 & u_{31} \\ 0 & 1 & u_{23} & u_{32} \\ 1 & u_{31} & u_{32} & u_{13} + u_{23} \end{pmatrix}$$

4 数値計算

われわれが求めたいものは $c(\Theta)f(\Theta, -X)$ の最小点であった. ベクトル値関数を $g(X; \Theta) = \exp(-\sum_{i=1}^3 x_{ii}\theta_{ii}) c(\Theta)$ ととる. このとき, $g(X; \Theta)$ は Pfaffian 方程式

$$\frac{\partial g}{\partial \theta_{ii}} = (P_i(\Theta) - x_{ii}I_4) g, \quad (i = 1, 2, 3)$$

を満す. この方程式の特異点は 24 個の超平面である.

Θ が原点近傍にある場合には, $c(\Theta) = {}_0F_1(\frac{3}{2}; \frac{1}{4}\Theta^\top \Theta)$ を用いて, ${}_0F_1$ の級数展開によって, $c(\Theta)$ を計算できる. また, Pfaffian 方程式の正則点では, 直線 $(\theta_{11}, \theta_{22}, \theta_{33}) = (at, bt, ct)$ では, 常微分方程式

$$\frac{dc}{dt} = \left(A - \frac{2}{t} \text{diag}(0, 1, 1, 1) \right) c$$

が成り立つ. よって結局, 原点近傍から出発して Runge-Kutta 法によって正則点における近似値を得ることができる. なお, Pfaffian 方程式の特異点においては, さらにイデアルの制限をとることで同様に関数値を計算することができる.

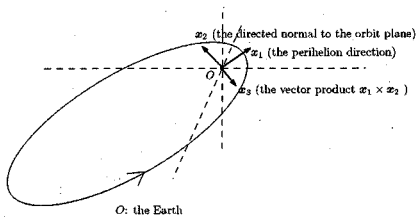
例 1 (適用例: comets). NASA 地球近傍天体プログラムで公開されている 85 個の彗星の軌道データから標本 $X = (x_1, x_2, x_3) \in \text{SO}(3)$ を作成した.

標本平均行列 X の特異値分解

$$X = \begin{pmatrix} 0.257 & 0.044 & 0.189 \\ 0.158 & -0.052 & -0.146 \\ 0.079 & 0.765 & 0.004 \end{pmatrix} = P \hat{X} Q^\top, \quad \hat{X} = \text{diag}(0.774, 0.318, 0.210)$$

に対し, $f(X; \Theta)$ の最大値を与える Θ は, ホロノミック勾配法により

$$\Theta = \begin{pmatrix} 2.953 & 0.200 & 0.871 \\ -0.423 & -0.317 & 2.390 \\ 0.378 & 5.568 & 0.251 \end{pmatrix} = P \hat{\Theta} Q^\top, \quad \hat{\Theta} = \text{diag}(5.617, 3.078, -2.387)$$



参考文献

- [1] T. Hibi (edited), Gröbner Bases: Statistics and Software Systems, Springer, 2014.
- [2] 高山信毅, D 加群の積分アルゴリズムと推定理論, 特集 グレブナー基底の新天地, 数学セミナー 2012 年 2 月号, 41–46.
- [3] H. Nakayama et al., Holonomic Gradient Descent and its Application to the Fisher-Bingham Integral, Advances in Applied Mathematics, **47**(2011), 639–658.
- [4] T. Sei et. al. Properties and applications of Fisher distribution on the rotation group, Journal of Multivariate Analysis **116** (2013), 440–455.