

統計学とグレブナー基底

日比孝之

大阪大学大学院情報科学研究科情報基礎数学専攻

本稿^{*1}は、科学技術振興機構の戦略的創造研究推進事業の研究領域「数学と諸分野の協働によるブレークスルーの探索」に属するCREST^{*2}研究課題「現代の産業社会とグレブナー基底の調和」の研究の枠組の紹介^{*3}である。

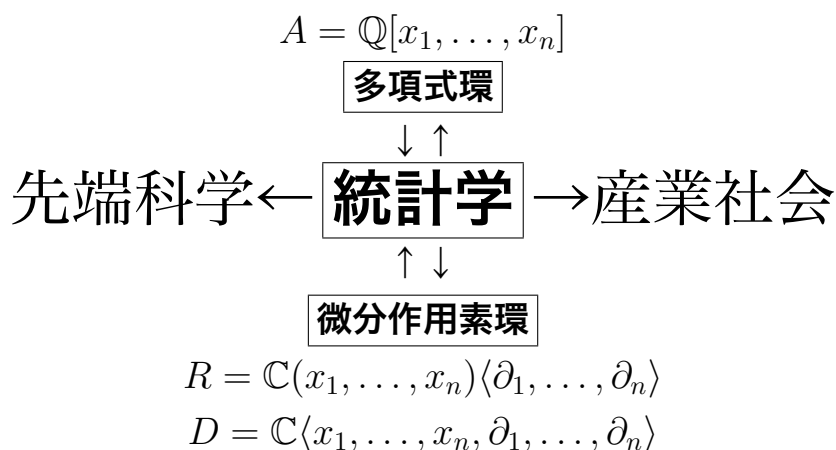
JST CREST

現代の産業社会とグレブナー基底の調和

2008年10月～2014年3月

<http://www.math.jst.go.jp/ja/scientists/teamhibi/index.html>

理論系グループ	応用系グループ	計算系グループ
日比孝之 (大阪大学)	竹村彰通 (東京大学)	高山信毅 (神戸大学)
大杉英史 (立教大学)	青木 敏 (鹿児島大学)	野呂正行 (神戸大学)
		濱田龍義 (福岡大学)



多項式環と微分作用素環のグレブナー基底の理論と計算を飛躍的に深化させ、純粋数学の範疇において、統計学の根幹に劇的な変革を導く。その変革は、大規模な計算を可能にするアルゴリズムの進化と相俟って、統計学を不可欠とする産業社会と先端科学の諸分野に広範な影響を及ぼす。

^{*1} 代数学シンポジウムの講演スライドが <http://www2.rikkyo.ac.jp/~ohsugi/Kyoto200812.pdf> にある。

^{*2} Core Research for Evolutional Science and Technology

^{*3} [2] も参照されたい。

§1. グレブナー基底とは何か？

有理数体 \mathbb{Q} 上の n 変数多項式環 $A = \mathbb{Q}[x_1, \dots, x_n]$ がグレブナー基底^{*4}の舞台である。変数の積 $x_1^{a_1} \cdots x_n^{a_n}$ を単項式と呼ぶ。特に、1 も単項式である。単項式の集合 \mathcal{M} があつたとき、 $u \in \mathcal{M}$ が \mathcal{M} の**極小元**であるとは、条件「 $v \in \mathcal{M}$ が u を割り切るならば $v = u$ である」が満たされる時に言う。

(1.1) **定理** (Dickson の補題) 単項式の集合の極小元は高々有限個である。

多項式環 A の単項式全体の集合における全順序 $<$ が、条件

- $1 < u$ ($\forall u$)
- $u < v$ ならば $u \cdot w < v \cdot w$ ($\forall w$)

を満たすとき、 $<$ を A 上の**単項式順序**と言う。たとえば、辞書式順序 $<_{\text{lex}}$ 、逆辞書式順序 $<_{\text{rev}}$ など。

多項式 $0 \neq f \in A$ に現れる単項式のなかで $<$ について最大の単項式を f の $<$ に関する**イニシャル単項式**と呼び $\text{in}_{<}(f)$ と表す。たとえば、 $f = x_1x_4 - x_2x_3$ とすると、 $\text{in}_{<_{\text{lex}}}(f) = x_1x_4$ 、 $\text{in}_{<_{\text{rev}}}(f) = x_2x_3$ である。

以下、多項式環 $A = \mathbb{Q}[x_1, \dots, x_n]$ の単項式順序 $<$ を一つ固定する。イデアル I ($\neq 0$) があつたとき、単項式 (が生成する) イデアル

$$\text{in}_{<}(I) = (\{\text{in}_{<}(f) : 0 \neq f \in I\})$$

を I の $<$ に関する**イニシャルイデアル**と呼ぶ。有限集合 $\mathcal{G} = \{g_1, \dots, g_s\} \subset I$ が ($<$ に関する I の)**グレブナー基底**であるとは

$$\text{in}_{<}(I) = (\text{in}_{<}(g_1), \dots, \text{in}_{<}(g_s))$$

が成立するときに言う^{*5}。

(1.2) **例** 多項式 $f = x_1x_4 - x_2x_3$ と $g = x_4x_7 - x_5x_6$ が生成するイデアル $I = (f, g)$ を考える。このとき、 $\{f, g\}$ は I の $<_{\text{lex}}$ に関するグレブナー基底ではない。実際、 $h = x_7f - x_1g = x_1x_5x_6 - x_2x_3x_7 \in I$ だから $x_1x_5x_6 \in \text{in}_{<_{\text{lex}}}(I)$ である。ところが、 $x_1x_5x_6 \notin (x_1x_4, x_4x_7)$ であるから $(\text{in}_{<_{\text{lex}}}(f), \text{in}_{<_{\text{lex}}}(g)) \neq \text{in}_{<_{\text{lex}}}(I)$ である。

(1.3) **定理** (Hilbert 基底定理) グレブナー基底はそのイデアルの生成系である。

Buchberger 判定法とは、イデアルの有限な生成系があつたとき、その生成系がグレブナー基底であるか否かを判定する。

Buchberger アルゴリズムとは、イデアルの有限な生成系が既知なとき、その生成系からグレブナー基底を探索する計算手続きである。

(1.4) **定理** (Macaulay の定理) イニシャルイデアル $\text{in}_{<}(I)$ に属さない単項式 (**標準単項式**と呼ぶ) の全体は、剰余環 A/I の \mathbb{Q} 上の線型空間としての基底である。

割り算アルゴリズムとは、グレブナー基底が既知なとき、多項式を (剰余環において) 標準単項式の線型結合として表す計算手続きである。

^{*4} グレブナー基底の文献と歴史的背景については、[1, 第1章] を参照されたい。

^{*5} グレブナー基底の存在は Dickson の補題から従う。

(1.5) 例 多項式環 $\mathbb{Q}[x, y, z]$ と辞書式順序 $<_{\text{lex}}$ ^{*6} を考える。いま、 $g_1 = x^2 - z, g_2 = xy - 1$ とし、 $I = (g_1, g_2)$ とすると、 $\{g_1, g_2\}$ は I の $<_{\text{lex}}$ に関するグレブナー基底ではないが、 $g_3 = -yz + x$ とすると、 $\{g_1, g_2, g_3\}$ がグレブナー基底となる。たとえば、多項式 $f = x^3 - x^2y - x^2 - 1$ は、割り算アルゴリズムを使うと、 $f = (x - y - 1)g_1 + g_3 + (xz - x - z - 1)$ と表示できる^{*7}。すると、剰余環 $\mathbb{Q}[x, y, z]/I$ において、 $f = xz - x - z - 1$ である。

§2. トーリックイデアルとマルコフ基底

非負整数を成分とする行列 $\mathbf{A} = (a_{ij})_{\substack{i=1, \dots, d \\ j=1, \dots, n}}$ が配置行列であるとは、 \mathbb{Q}^d の原点を通過しない超平面で、 \mathbf{A} の列ベクトルのすべてを含むものが存在するときに言う。多項式環 $\mathbb{Q}[x_1, \dots, x_n, y_1, \dots, y_d]$ において、多項式

$$x_j - y_1^{a_{1j}} y_2^{a_{2j}} \cdots y_d^{a_{dj}}, \quad j = 1, \dots, n$$

が生成するイデアル $J_{\mathbf{A}}$ を考える。このとき、多項式環 $A = \mathbb{Q}[x_1, \dots, x_n]$ のイデアル $J_{\mathbf{A}} \cap A$ を配置行列 \mathbf{A} の**トーリックイデアル**と呼び、 $I_{\mathbf{A}}$ と表す。

(2.1) 例

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \end{bmatrix}$$

とすると、

$$J_{\mathbf{A}} = (x_1 - y_1, x_2 - y_1 y_2, x_3 - y_1 y_2^2 y_3, x_4 - y_1 y_2 y_3^2, x_5 - y_1 y_3)$$

である。すると、配置行列 \mathbf{A} のトーリックイデアルは

$$I_{\mathbf{A}} = (x_2 x_5^2 - x_1^2 x_4, x_2 x_4 - x_3 x_5, x_2^2 x_5 - x_1^2 x_3)$$

である。

配置行列 \mathbf{A} の核 $\text{Ker} \mathbf{A}$ に属する整数ベクトル $b = [b_1, \dots, b_n]^{\top}$ に、 $A = \mathbb{Q}[x_1, \dots, x_n]$ の**二項式**

$$f_b = \prod_{b_j > 0} x_j^{b_j} - \prod_{b_j < 0} x_j^{-b_j}$$

を対応させる。たとえば、 $b = [-2, 1, 0, -1, 2]^{\top}$ ならば、 $f_b = x_2 x_5^2 - x_1^2 x_4$ である。すると、

(2.2) 補題 $I_{\mathbf{A}} = (\{f_b : b \in \text{Ker} \mathbf{A} \cap \mathbb{Z}^n\})$

(2.3) 系 トーリックイデアルの**被約グレブナー基底**^{*8}は二項式から成る。

トーリックイデアルの概念を提唱したのは、整数計画問題をグレブナー基底を使って解くアルゴリズムを披露した Conti と Traverso である^{*9}。本節の残りでは、トーリックイデアルのグレブナー基底がどのように統計学の世界に君臨するかを紹介する。

^{*6} 但し、 $x > y > z$ である。

^{*7} $f = (x - 1)g_1 - xg_2 + (xz - x - z - 1)$ なる表示も可能である。

^{*8} 被約グレブナー基底とは、単項式順序を固定したとき、無駄がなく、一意的に決まるグレブナー基底である。

^{*9} もっとも、歴史を遡ると、トーリックイデアルの着想は、1970年、Jürgen Herzog が、既に、導入している。

(2.4) 例 (Hardy-Weinberg モデル) 何らかの特定の病気に罹患している患者 100 名の ABO 式血液型の遺伝子型を調べたところ、下記の表の結果を得た。その病気と血液型は何らかの関連があるか？

遺伝子型	AA	AB	AO	BB	BO	OO
人数 (観測値)	23	10	15	6	17	29

いま、その患者 100 名が持つ遺伝子 (A 遺伝子、B 遺伝子、O 遺伝子) の個数を調べる。単なる個数の計算に過ぎないけれども、敢えて、下記のような行列と列ベクトルを使って記載する。そうすると、A 遺伝子 71 個、B 遺伝子 39 個、O 遺伝子 90 個となる。

$$\begin{bmatrix} 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 23 \\ 10 \\ 15 \\ 6 \\ 17 \\ 29 \end{bmatrix} = \begin{bmatrix} 71 \\ 39 \\ 90 \end{bmatrix}$$

統計学の世界では、 $[71, 39, 90]^T$ を **十分統計量** と呼ぶ。一般に、無作為に選ばれた 100 人の集団があったとき、その 100 人が持つ 200 個の遺伝子が、A 遺伝子 71 個、B 遺伝子 39 個、O 遺伝子 90 個とすると、遺伝子型は **Hardy-Weinberg の法則**^{*10} に従うと考えられる。すると、遺伝子型の期待値は

AA	AB	AO	BB	BO	OO
12.6025	13.845	31.95	3.8025	17.55	20.25

となる。観測値と期待値を眺めると、観測値には偏りがあるように思われるが、実際に偏りがあるのか、それとも偶然の結果と考えても不自然がないのかを判定するのが **統計的仮説検定** である。詳しい解説は、たとえば、[2、第4章] に譲るが、統計的仮説検定の常套手段の選択は、

- 人数が“少ない”から **漸近分布論** は (?)
- **Fisher の正確確率検定** は人数が“多い”から (?)
- **マルコフ連鎖モンテカルロ法** が一つの有効な方法 (!)

となる。Fisher の正確確率検定を実施するには、連立線型方程式

$$\begin{bmatrix} 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_{AA} \\ x_{AB} \\ x_{AO} \\ x_{BB} \\ x_{BO} \\ x_{OO} \end{bmatrix} = \begin{bmatrix} 71 \\ 39 \\ 90 \end{bmatrix} \quad (1)$$

の非負整数解 $[x_{AA}, x_{AB}, x_{AO}, x_{BB}, x_{BO}, x_{OO}]^T$ を全列挙し、いわゆる p 値 (有意確率)^{*11} を計算する。しかし、観測データが大きくなると全列挙をすることは絶望的であるから、非負整数解の一部をサンプリングし、 p 値の近似値を計算する。そのサンプリングを遂行するのに必要な概念が **マルコフ基底** である。マルコフ基底とは、連立線型方程式 (1) の係数行列の核に属する有限個の整数ベクトルの集合 $\{b^{(1)}, b^{(2)}, \dots, b^{(s)}\}$ であって、条件「連立線型方程式 (1) の任意の非負整数解 \mathbf{x} と \mathbf{y} があつた

^{*10} 生物種の個体群における対立遺伝子の遺伝子頻度は世代が移り変わっても変化しないことを前提とし、個体群内の遺伝子型の構成について説明する法則である。

^{*11} Hardy-Weinberg の法則に従うという仮説 (帰無仮説) の下で、観測データの事象が起こる確率が p 値である。

とき、

$$\mathbf{y} = \mathbf{x} + \varepsilon_1 b^{(1)} + \varepsilon_2 b^{(2)} + \cdots + \varepsilon_s b^{(s)}$$

となる $\varepsilon_i \in \{0, 1, -1\}$ が存在する」を満たすものである。下記の表の右側に記載されている 6 個の非負整数ベクトルがマルコフ基底である。他方、下記の表の左側は、連立線型方程式 (1) の係数行列のトーリックイデアルのグレブナー基底である。両者を眺めれば、トーリックイデアルのグレブナー基底とマルコフ基底の対応は一目瞭然であろう。

$x_1 x_4 - x_2^2$	$[1, -2, 0, 1, 0, 0]^\top$
$x_1 x_6 - x_3^2$	$[1, 0, -2, 0, 0, 1]^\top$
$x_4 x_6 - x_5^2$	$[0, 0, 0, 1, -2, 1]^\top$
$x_2 x_3 - x_1 x_5$	$[-1, 1, 1, 0, -1, 0]^\top$
$x_2 x_5 - x_3 x_4$	$[0, 1, -1, -1, 1, 0]^\top$
$x_3 x_5 - x_2 x_6$	$[0, -1, 1, 0, 1, -1]^\top$

一般に、現実の社会で遭遇する統計の問題では、膨大なデータの処理を必要とし、グレブナー基底が計算できれば解析が可能になるとしても、計算量の問題が克服できず、現状では、その実用化が困難である。ところが、可換代数の世界では、従来から、多種多様なグレブナー基底が生息している。すると、逆転の発想ではあるが、**既知のグレブナー基底をマルコフ基底とする統計的モデル**（よしんば人工的であれ）**設定すれば、その検定を実施することが可能となる。**

例 (2.4) の係数行列は **Veronese 配置** と呼ばれる配置行列である。分割表の検定^{*12}には、**Segre 配置** と呼ばれる配置行列も使われる。両者を一般化した概念に、**Segre-Veronese 配置** がある。そのグレブナー基底も既知であるから、Segre-Veronese 配置を係数行列とする統計モデルを創ることができれば、マルコフ連鎖モンテカルロ法が実施できる。実際、大学入試センター試験の科目選択の統計的検証、遺伝子モデルの検証など、Segre-Veronese 配置が有効となる統計モデルは自然に存在する。逆に、統計学から代数学へのフィードバックであるが、更に複雑な統計モデルを検証する必要から、Segre-Veronese 配置の一般化である**入れ子配置**の概念を提唱するに至った。

§3. ホロノミック勾配降下法

有理関数体 $\mathbb{C}(x_1, \dots, x_n)$ 上の n 変数微分作用素環 $R = \mathbb{C}(x_1, \dots, x_n)\langle \partial_1, \dots, \partial_n \rangle$ を考える。微分作用素環は非可換であり、関係式

$$\partial_i \partial_j = \partial_j \partial_i, \quad \partial_i a(x) = a(x) \partial_i + \frac{\partial a(x)}{\partial x_i}$$

で定義される。但し、 $a(x) \in \mathbb{C}(x_1, \dots, x_n)$ である。微分作用素環 R の単項式 $\partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}$ の全体は、 R の $\mathbb{C}(x_1, \dots, x_n)$ 上の線型空間としての基底である。

微分作用素環 R の元は x_1, \dots, x_n の滑らかな関数 $f(x)$ に自然に作用する。すなわち、

$$a(x) \partial^\alpha \bullet f(x) = a(x) \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$$

^{*12} [1, 例 (4.1.12)]

定義 微分作用素環 R の左イデアル I が **0次元左イデアル** であるとは、剰余環 R/I の $\mathbb{C}(x_1, \dots, x_n)$ 上の線型空間としての次元が有限であるときに言う。

多項式環 $A = \mathbb{Q}[x_1, \dots, x_n]$ のグレブナー基底の理論は**ほとんどそのまま**（もちろん、細部の補正は必要であるが）微分作用素環 $R = \mathbb{C}(x_1, \dots, x_n)\langle \partial_1, \dots, \partial_n \rangle$ においても成立する*13。

以下、 $I \subset R$ を 0次元左イデアルとし、 $r = \dim_{\mathbb{C}(x_1, \dots, x_n)} R/I$ と置く*14。すると、

- (i) 有限個の単項式（**標準単項式**と呼ぶ） $t_1 (= 1), t_2, \dots, t_r$ で、 R/I の $\mathbb{C}(x_1, \dots, x_n)$ 上の基底となるものと、更に、
- (ii) $\mathbb{C}(x_1, \dots, x_n)$ の元を成分とする $r \times r$ 行列 P_1, P_2, \dots, P_n （**Pfaffian 行列**と呼ぶ）で、剰余環 R/I において、関係式

$$\begin{bmatrix} \partial_i t_1 \\ \partial_i t_2 \\ \vdots \\ \partial_i t_r \end{bmatrix} = P_i \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_r \end{bmatrix} \quad (2)$$

を満たすものを**計算することが可能**である*15。

(3.1) 例 変数の個数を $n = 2$ とし、 $\mathbb{C}(x, y)\langle \partial_x, \partial_y \rangle$ の 0次元左イデアル $I = (\partial_x^2 + y^2, \partial_y^2 + x^2)$ を考える。標準単項式は、単項式順序に依存するから、Pfaffian 行列も単項式順序に依存する。標準単項式は $1, \partial_y$ となるか、あるいは、 $1, \partial_x$ となるかの何れかである。前者の Pfaffian 行列は

$$\begin{bmatrix} \partial_x \cdot 1 \\ \partial_x \cdot \partial_y \end{bmatrix} = \begin{bmatrix} 0 & y/x \\ -xy & 1/x \end{bmatrix} \begin{bmatrix} 1 \\ \partial_y \end{bmatrix} \quad \begin{bmatrix} \partial_y \cdot 1 \\ \partial_y \cdot \partial_y \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -x^2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ \partial_y \end{bmatrix}$$

であり、後者の Pfaffian 行列は

$$\begin{bmatrix} \partial_x \cdot 1 \\ \partial_x \cdot \partial_x \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -y^2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ \partial_x \end{bmatrix} \quad \begin{bmatrix} \partial_y \cdot 1 \\ \partial_y \cdot \partial_x \end{bmatrix} = \begin{bmatrix} 0 & x/y \\ -xy & 1/y \end{bmatrix} \begin{bmatrix} 1 \\ \partial_x \end{bmatrix}$$

である。

定義 \mathbb{C}^n の開集合で正則な函数 $f = f(x_1, \dots, x_n)$ が**ホロノミック函数***16であるとは、 $I \bullet f = 0$ *17を満たす 0次元左イデアル $I \subset R$ が存在するときに言う。

ホロノミック函数 f と $I \bullet f = 0$ となる 0次元左イデアル $I \subset R$ があったとき、函数 f の局所的な最小値の近似値を計算する効果的なアルゴリズム（**ホロノミック勾配降下法**）を創ることができる。

*13 詳細は [1, 第6章] に譲る。

*14 非負整数 r を I のホロノミック階数と呼ぶ。

*15 (i) は Macaulay の定理、(ii) は割り算アルゴリズムである。

*16 ホロノミック函数の解説は [1, 第6章] を参照されたい。

*17 すなわち、任意の $\ell \in I$ が $\ell \bullet f = 0$ を満たす。

(3.2) **アルゴリズム** (ホロノミック勾配降下法 [1, pp. 355 – 356])

(Step 1) 標準単項式 $t_1 (= 1), t_2, \dots, t_r$ を探す。

(Step 2) Pfaffian 行列 P_1, P_2, \dots, P_n から Pfaffian 系を導く*¹⁸。

$$\frac{\partial \mathbf{F}}{\partial x_i} = P_i \mathbf{F} \quad (i = 1, 2, \dots, n) \quad (3)$$

但し、

$$\frac{\partial \mathbf{F}}{\partial x_i} = \begin{bmatrix} \partial f / \partial x_i \\ \partial_i t_2 \bullet f \\ \vdots \\ \partial_i t_r \bullet f \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} f \\ t_2 \bullet f \\ \vdots \\ t_r \bullet f \end{bmatrix}$$

(Step 3) Take a point \mathbf{c} as a starting point and evaluate numerically \mathbf{F} at \mathbf{c} *¹⁹.

Let $\mathbf{e} = \mathbf{c}$.

(Step 4) $\mathbf{F}(\mathbf{e})$ は既知だから (3) から勾配がわかる*²⁰。

$$(\nabla(f))(\mathbf{e}) = \begin{bmatrix} (\partial f / \partial x_1)(\mathbf{e}) \\ (\partial f / \partial x_2)(\mathbf{e}) \\ \vdots \\ (\partial f / \partial x_n)(\mathbf{e}) \end{bmatrix}$$

(Step 5) If $(\nabla(f))(\mathbf{e}) = 0$, then stop the algorithm.

Let $(\nabla(f))(\mathbf{e}) \neq 0$. If $\varepsilon > 0$ is small, then it follows that*²¹

$$\mathbf{F}(\mathbf{e} - \varepsilon(\nabla(f))(\mathbf{e})) \approx \mathbf{F}(\mathbf{e}) - \sum_{i=1}^n \varepsilon (\partial f / \partial x_i)(\mathbf{e}) (\partial \mathbf{F} / \partial x_i)(\mathbf{e}) \quad (4)$$

$$= \mathbf{F}(\mathbf{e}) - \sum_{i=1}^n \varepsilon (\partial f / \partial x_i)(\mathbf{e}) (P_i \mathbf{F})(\mathbf{e}). \quad (5)$$

Thus we can evaluate numerically \mathbf{F} at $\mathbf{e} - \varepsilon(\nabla(f))(\mathbf{e})$.

Update \mathbf{e} to $\mathbf{e} - \varepsilon(\nabla(f))(\mathbf{e})$.

GO TO (Step 4) *²².

*¹⁸ 関係式 (2) は剰余環 R/I において成立する。ところが、 $I \bullet f = 0$ であるから (2) の両辺を f に作用させると Pfaffian 系 (3) が導ける。なお、 $t_1 = 1$ であることに注意する。

*¹⁹ 初期値の計算である。

*²⁰ (3) の左辺の第 1 成分に着目する。

*²¹ テイラー展開の一次近似である。

*²² (Step 3) に戻るのではない。換言すると、 \mathbf{F} の $\mathbf{e} - \varepsilon(\nabla(f))(\mathbf{e})$ に於ける値を数値計算するのではない。初期値 $\mathbf{F}(\mathbf{e})$ が既知だから、テイラー展開の近似式から $\mathbf{F}(\mathbf{e} - \varepsilon(\nabla(f))(\mathbf{e}))$ が計算可能であり、従って、 $\mathbf{e} - \varepsilon(\nabla(f))(\mathbf{e})$ における勾配もすぐわかる。すると、(Step 4) と (Step 5) を繰り返せば局所的な最小値に到達する。

本節の残りでは、**方向統計学**における円周 S^1 上の**フォン・ミーゼス分布**のパラメータ最尤推定の具体例を使って、ホロノミック勾配降下法の統計学への応用を紹介する。フォン・ミーゼス分布とは、**確率密度関数** $p(t; \theta_1, \theta_2)$ が

$$p(t; \theta_1, \theta_2) \propto f(t; \theta_1, \theta_2) = \exp(\theta_1 \cos t + \theta_2 \sin t), \quad 0 \leq t \leq 2\pi$$

と表される分布である。但し、 θ_1 と θ_2 はパラメータである。その**正規化定数**は

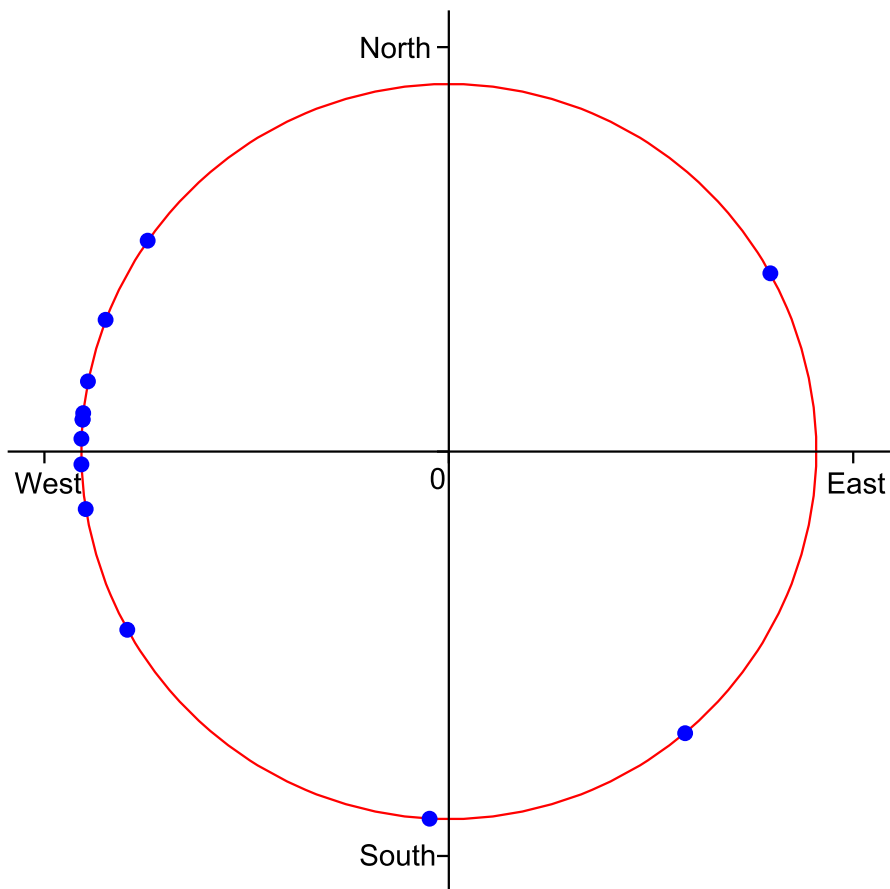
$$F(\theta_1, \theta_2) = \int_0^{2\pi} f(t; \theta_1, \theta_2) dt$$

である。観測データ T の確率は

$$p(T; \theta_1, \theta_2) = \frac{f(T; \theta_1, \theta_2)}{F(\theta_1, \theta_2)}$$

である。

下記のデータは札幌上空約 1 万メートル地点における午前 9 時の風向きの 13 日間のデータである*²³。一般に、固定された地点における固定された時刻の風向きの確率は、フォン・ミーゼス分布であることが知られている。



*²³ 2011 年 1 月 1 日から 14 日 (但し、11 日は除外) のデータ (気象庁の気象統計情報) である。

最尤法^{*24}では、観測データ T_1, \dots, T_N があつたとき尤度函数^{*25}

$$L(\theta_1, \theta_2) = \prod_{i=1}^N p(T_i; \theta_1, \theta_2) = \prod_{i=1}^N \frac{\exp(\theta_1 \cos T_i + \theta_2 \sin T_i)}{F(\theta_1, \theta_2)} \quad (6)$$

を最大化する θ_1 と θ_2 をパラメータの推定値とする。

以下、札幌上空の風向きのデータ ($N = 13$) の尤度函数 (6) を最大化するパラメータ θ_1 と θ_2 を、ホロノミック勾配降下法を使って計算する。

- 変数変換 $\cos t = \frac{1-u^2}{1+u^2}$, $\sin t = \frac{2u}{1+u^2}$ を施すと $F(\theta_1, \theta_2) = \int_{-\infty}^{\infty} g(u, \theta_1, \theta_2) du$, となる。但し、

$$g(u, \theta_1, \theta_2) = \frac{2}{1+u^2} \exp\left(\frac{(1-u^2)\theta_1 + 2u\theta_2}{1+u^2}\right)$$

である。

- 多項式係数の微分作用素環^{*26} $D_2 = \mathbb{C}\langle \theta_1, \theta_2, \partial_1, \partial_2 \rangle$ と $D_3 = \mathbb{C}\langle u, \theta_1, \theta_2, \partial_u, \partial_1, \partial_2 \rangle$ を考え、 g を零化するホロノミックイデアル^{*27} $J \subset D_3$ を探す。すると、積分イデアル $I = (J + \partial_u D_3) \cap D_2$ は F を零化するホロノミックイデアル (Bernstein) である。このとき、D加群の積分アルゴリズム^{*28}を使うと、 $I \subset D_2$ が計算可能である。

$$I = ((\theta_1^2 + \theta_2^2)\partial_2^2 - \theta_1\partial_1 + \theta_2\partial_2 - \theta_2^2, \theta_1\partial_1\partial_2 + \theta_2\partial_2^2 + \partial_2 - \theta_2, \theta_2\partial_1 - \theta_1\partial_2, \partial_1^2 + \partial_2^2 - 1)$$

- イデアル I を $\mathbb{C}(\theta_1, \theta_2)\langle \partial_1, \partial_2 \rangle$ のイデアルと考えると 0 次元左イデアルである。もちろん、 $I \bullet F = 0$ である。逆辞書式順序に関する I のグレブナー基底を計算する。

$$\mathcal{G} = \{(\theta_1^2 + \theta_2^2)\partial_2^2 - \theta_1\partial_1 + \theta_2\partial_2 - \theta_2^2, \theta_2\partial_1 - \theta_1\partial_2\}$$

- 標準単項式 $1, \partial_2$ から $F(\theta_1, \theta_2)$ の Pfaffian 系を計算する。

$$\begin{aligned} \begin{bmatrix} \frac{\partial F}{\partial \theta_1} \\ \frac{\partial^2 F}{\partial \theta_1 \partial \theta_2} \end{bmatrix} &= \begin{bmatrix} 0 & \frac{\theta_1}{\theta_2} \\ \frac{\theta_1\theta_2}{\theta_1^2 + \theta_2^2} & \frac{-2\theta_1}{\theta_1^2 + \theta_2^2} \end{bmatrix} \begin{bmatrix} F \\ \frac{\partial F}{\partial \theta_2} \end{bmatrix} \\ \begin{bmatrix} \frac{\partial F}{\partial \theta_2} \\ \frac{\partial^2 F}{\partial \theta_2^2} \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ \frac{\theta_2}{\theta_1^2 + \theta_2^2} & \frac{\theta_1^2 - \theta_2^2}{\theta_2(\theta_1^2 + \theta_2^2)} \end{bmatrix} \begin{bmatrix} F \\ \frac{\partial F}{\partial \theta_2} \end{bmatrix} \end{aligned}$$

- 尤度函数 $L(\theta_1, \theta_2)$ の最大化を $L(\theta_1, \theta_2)^{-1/N}$ の最小化に変更する。

$$\begin{aligned} L(\theta_1, \theta_2)^{-1/N} &= F(\theta_1, \theta_2) \left(\prod_{i=1}^N p(T_i; \theta_1, \theta_2) \right)^{-1/N} \\ &= F(\theta_1, \theta_2) \exp\left(-\frac{\theta_1}{N} \sum_{i=1}^N \cos T_i - \frac{\theta_2}{N} \sum_{i=1}^N \sin T_i\right) \end{aligned}$$

*24 maximum likelihood estimation

*25 likelihood function

*26 ワイル代数とも呼ばれる。

*27 [1, p. 370]

*28 [1, p. 375]

● $F(\theta_1, \theta_2)$ の Pfaffian 系から $L^* = L(\theta_1, \theta_2)^{-1/N}$ の Pfaffian 系を導く。

$$\begin{bmatrix} \frac{\partial L^*}{\partial \theta_1} \\ \frac{\partial^2 L^*}{\partial \theta_1 \partial \theta_2} \end{bmatrix} = P \begin{bmatrix} L^* \\ \frac{\partial L^*}{\partial \theta_2} \end{bmatrix} \quad (7)$$

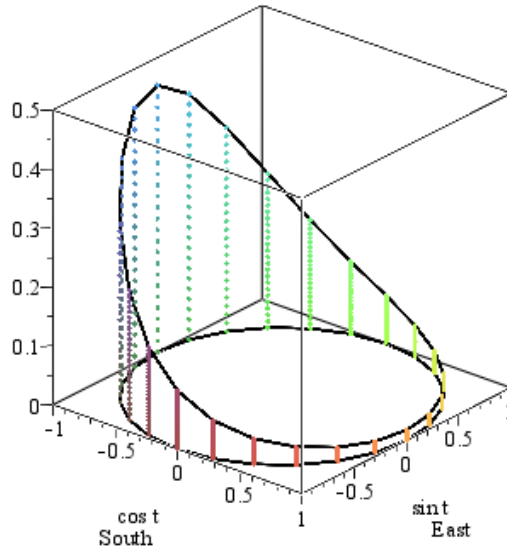
$$\begin{bmatrix} \frac{\partial L^*}{\partial \theta_2} \\ \frac{\partial^2 L^*}{\partial \theta_2^2} \end{bmatrix} = Q \begin{bmatrix} L^* \\ \frac{\partial L^*}{\partial \theta_2} \end{bmatrix} \quad (8)$$

但し、

$$P = \begin{bmatrix} -\frac{1}{N} \sum_{i=1}^N \cos(T_i) & \frac{\theta_1}{\theta_2} \\ \frac{\theta_1 \theta_2}{\theta_1^2 + \theta_2^2} & -\frac{2\theta_1}{\theta_1^2 + \theta_2^2} - \frac{1}{N} \sum_{i=1}^N \cos(T_i) \end{bmatrix}$$

$$Q = \begin{bmatrix} -\frac{1}{N} \sum_{i=1}^N \sin(T_i) & 1 \\ \frac{\theta_2^2}{\theta_1^2 + \theta_2^2} & \frac{\theta_1^2 - \theta_2^2}{\theta_2(\theta_1^2 + \theta_2^2)} - \frac{1}{N} \sum_{i=1}^N \sin(T_i) \end{bmatrix}$$

である。函数 $L^* = L(\theta_1, \theta_2)^{-1/N}$ の Pfaffian 系から、ホロノミック勾配降下法を使えば、 L^* を最小化するパラメータ θ_1 と θ_2 の値を近似することができる。



$$p(t; \theta_1, \theta_2) \quad \text{with} \quad \theta_1 = -1.62, \theta_2 = -0.10$$

ホロノミック勾配降下法は、Zeilberger を創始者とする holonomic function approach の潮流に乗る研究であり、指数型分布族の基準化定数、最尤推定量の計算のための、従来の手法の限界を遥かに越える、全く新しい汎用的な方法である^{*29}。更に、**Risa/Asir** などにおける、微分作用素環のグレブナー基底の高速計算の実装の下、球面 S^n 上の**フィッシャー・ビンガム分布**など**空間統計学**で重要なパラメータ次元の高い分布でも、ホロノミック勾配降下法が適用可能である。

^{*29} D加群のグレブナー基底と積分アルゴリズムが統計学に応用できることが判明したことは、特筆に値する。

参考文献

プロジェクト研究の成果となる研究論文、招待講演のリストなどはホームページで公開されている。以下、参考文献として、プロジェクト研究に関連して出版された単行本などを紹介する。

単行本 [1] は、研究成果の社会への還元を目指し、プロジェクトのメンバーが総力を結集して執筆している。帯には「グレブナー基底の奥義を伝授」と記載されている。本著は xii + 577 ページと分厚く、定価（本体 5,600 円+税）も高い。それにもかかわらず、初版第 1 刷の発行（2011 年 9 月 20 日）からわずか 5 ヶ月ほどで、第 1 刷は完売している。誤植などを可能な限り修正し、初版第 2 刷が 2012 年 3 月 20 日に発行されている。第 1 章から第 6 章の内容と執筆者は、グレブナー基底の基礎理論（日比孝之）、KNOPPIX/Math の紹介（濱田龍義）、計算の効率化（野呂正行）、計算代数統計の展開（青木敏、竹村彰通）、凸多面体の三角形分割（大杉英史）、微分作用素環のグレブナー基底（高山信毅）となっている。しかし、本著の特徴は、何と言っても、140 余ページに及ぶ豊富な例題と演習問題の集積である第 7 章（中山洋将、西山絢太）の存在である。グレブナー基底に関連する 10 種類を越えるソフトウェアを駆使し、多種多様な問題の計算機による計算方法を完璧に紹介している。第 7 章を地道に学べば、グレブナー基底に関連する計算ソフトウェアを使うノウハウは、ひとまず、万全と言える。

数学セミナーの特集 [2, pp. 7–52] は、魅惑的なグレブナーワールドに読者を招待する企画であり、8 編の解説記事から構成されている。計算ソフトの開発秘話、「素手で計算する」とは何か、計算代数統計の世界、D 加群の積分アルゴリズムと推定理論などが素人にもわかりやすく紹介されている。

文献 [3] は、2010 年 6 月 28 日～7 月 2 日、ホテル阪急エキスポパークで開催された国際会議 “Harmony of Gröbner Bases and the Modern Industrial Society” の報告集であり、20 編の論文（査読有）が掲載されている。

[1] JST CREST 日比チーム（編）『グレブナー道場』（共立出版）2011 年

[2] 特集「グレブナー基底の新天地」『数学セミナー 2012 年 2 月号』（日本評論社）

[3] T. Hibi (Ed.), “Harmony of Gröbner Bases and the Modern Industrial Society,” World Scientific, Singapore, 2012.