

第57回代数学シンポジウム

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
JST* CREST†

2012年8月20日

* Japan Science and Technology Agency

† Core Research for Evolutional Science and Technology



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

$$A = \mathbb{Q}[x_1, \dots, x_n]$$

多項式環



先端科学 ← **統計学** → 産業社会



微分作用素環

$$R = \mathbb{C}(x_1, \dots, x_n) \langle \partial_1, \dots, \partial_n \rangle$$

- § 1. グレブナー基底とは何か？
- § 2. トーリックイデアルとマルコフ基底
- § 3. ホロノミック勾配降下法

§ 1. グレブナー基底とは何か？

$$A = \mathbb{Q}[x_1, \dots, x_n] \quad \text{多項式環}$$
$$x_1^{a_1} \cdots x_n^{a_n} \quad \text{単項式} \quad (\forall a_i \in \mathbb{Z}_{\geq 0})$$

単項式の（有限あるいは無限）集合 \mathcal{M} があったとき、
 $u \in \mathcal{M}$ 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」
が満たされるときに言う。

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。

$A = \mathbb{Q}[x_1, \dots, x_n]$ 多項式環

$x_1^{a_1} \cdots x_n^{a_n}$ 単項式

Dicksonの補題

多項式環 A の単項式全体の集合における全順序 $<$ が、条件

- $1 < u$ ($\forall u \neq 1$)
- $u < v$ 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 $f = x_1x_4 - x_2x_3$ ($x_1 > x_2 > x_3 > x_4$) とすると
 $\text{in}_{<_{\text{lex}}}(f) = x_1x_4$ $\text{in}_{<_{\text{rev}}}(f) = x_2x_3$

$$A = \mathbb{Q}[x_1, \dots, x_n]$$

$$x_1^{a_1} \cdots x_n^{a_n}$$

Dickson の補題

$$\text{in}_{<}(f)$$

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。

多項式環 A のイデアル I ($\neq 0$) があつたとき、イデアル $\text{in}_{<}(I) =_{\text{def}} (\{ \text{in}_{<}(f) : 0 \neq f \in I \})$ を I の $<$ に関する **イニシャルイデアル** と呼ぶ。

定義

有限集合 $G = \{g_1, \dots, g_s\} \subset I$ が ($<$ に関する I の) **グレブナー基底** であるとは $\text{in}_{<}(I) = (\text{in}_{<}(g_1), \dots, \text{in}_{<}(g_s))$ が成立するときを言う。

すると、グレブナー基底の存在は Dickson の補題から従う。

$A = \mathbb{Q}[x_1, \dots, x_n]$ $x_1^{a_1} \cdots x_n^{a_n}$ Dicksonの補題 $\text{in}_{<}(f)$ $\text{in}_{<}(I)$ グレブナー基底

例 変数の個数 $n = 7$ 辞書式順序 $<_{\text{lex}}$ $I = (f, g)$

$$\begin{aligned} f &= x_1x_4 - x_2x_3 & \text{in}_{<_{\text{lex}}}(f) &= x_1x_4 \\ g &= x_4x_7 - x_5x_6 & \text{in}_{<_{\text{lex}}}(g) &= x_4x_7 \end{aligned}$$

いま

$$h = x_7f - x_1g = x_1x_5x_6 - x_2x_3x_7$$

とすると

$$h \in I \quad \text{in}_{<_{\text{lex}}}(h) = x_1x_5x_6 \in \text{in}_{<_{\text{lex}}}(I)$$

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。従って、

$\{f, g\}$ はグレブナー基底ではない。

$A = \mathbb{Q}[x_1, \dots, x_n]$ $x_1^{a_1} \cdots x_n^{a_n}$ Dicksonの補題 $\text{in}_<(f)$ $\text{in}_<(I)$ グレブナー基底

定理 (Hilbert基底定理) 多項式環のイデアルの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

イデアルの有限な生成系があったとき、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

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

$A = \mathbb{Q}[x_1, \dots, x_n]$ $x_1^{a_1} \cdots x_n^{a_n}$ Dickson の補題 $\text{in}_{<}(f)$ $\text{in}_{<}(I)$ グレブナー基底
Hilbert 基底定理 Buchberger 判定法 Buchberger アルゴリズム

定理 (Macaulay の定理) $\text{in}_{<}(I)$ 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

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

§ 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。

定義

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。

$\mathbf{A} = (a_{ij})_{\substack{i=1,\dots,d \\ j=1,\dots,n}}$ 配置行列

$I_{\mathbf{A}} = J_{\mathbf{A}} \cap A$ トーリックイデアル

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

$$\begin{aligned} x_1 - y_1 & & x_2 - y_1 y_2 & & x_3 - y_1 y_2^2 y_3 \\ x_4 - y_1 y_2 y_3^2 & & & & x_5 - y_1 y_3 \end{aligned}$$

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

$A = (a_{ij})_{\substack{i=1,\dots,d \\ j=1,\dots,n}}$ 配置行列

$I_A = J_A \cap A$ トーリックイデアル

配置行列 A の核 $\text{Ker} A$ 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 $f_b = x_2 x_5^2 - x_1^2 x_4$

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

特に、 I_A の被約グレブナー基底は二項式から成る。

Hardy-Weinbergモデル

(ABO 式血液型)

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

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

Veronese 配置

十分統計量

遺伝子型と期待値

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

- 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(!)

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

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

$$\begin{array}{ll}
x_1x_4 - x_2^2 & [1, -2, 0, 1, 0, 0]^\top \\
x_1x_6 - x_3^2 & [1, 0, -2, 0, 0, 1]^\top \\
x_4x_6 - x_5^2 & [0, 0, 0, 1, -2, 1]^\top \\
x_2x_3 - x_1x_5 & [-1, 1, 1, 0, -1, 0]^\top \\
x_2x_5 - x_3x_4 & [0, 1, -1, -1, 1, 0]^\top \\
x_3x_5 - x_2x_6 & [0, -1, 1, 0, 1, -1]^\top
\end{array}$$

グレブナー基底

マルコフ基底

2 × 2 分割表

(代数と統計の成績)

代数 \ 統計	5	4	3	2	1	計
5	2	1	1	0	0	4
4	8	3	3	0	0	14
3	0	2	1	1	1	5
2	0	0	0	1	1	2
1	0	0	0	0	1	1
計	10	6	5	2	3	26

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を（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。

$$\begin{bmatrix} 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 1 & 1 & 1 & 0 & 0 & 0 & 2 & 2 & 2 & 1 & 1 & 1 & 0 & 0 & 0 & 2 & 2 & 2 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 2 & 1 & 0 & 2 & 1 & 0 & 2 & 1 & 0 & 2 & 1 & 0 & 2 & 1 & 0 & 2 & 1 & 0 & 2 & 1 & 0 & 2 & 1 & 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \end{bmatrix}$$

Segre – Veronese 配置

- 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

統計学からのフィードバック

更に複雑な統計モデルを検証するため、Segre – Veronese 配置の一般化である**入れ子配置**の概念を提唱

§ 3. ホロノミック勾配降下法

$\mathbb{C}(x_1, \dots, x_n)$ 有理函数体

$R = \mathbb{C}(x_1, \dots, x_n) \langle \partial_1, \dots, \partial_n \rangle$ **微分作用素環**

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

但し、 $a(x) \in \mathbb{C}(x_1, \dots, x_n)$

例

$$\begin{aligned} \partial_1^2 x_1^2 x_2^2 &= x_2^2 \partial_1^2 x_1^2 = x_2^2 \partial_1 (\partial_1 x_1^2) = x_2^2 \partial_1 (x_1^2 \partial_1 + 2x_1) \\ &= \dots \\ &= x_1^2 x_2^2 \partial_1^2 + 4x_1 x_2^2 \partial_1 + 2x_2^2 \end{aligned}$$

$$\partial^\alpha = \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n} \quad \text{単項式}$$

補題

単項式の全体は、 R の $\mathbb{C}(x_1, \dots, x_n)$ 上の線型空間としての基底である。

$R = \mathbb{C}(x_1, \dots, x_n) \langle \partial_1, \dots, \partial_n \rangle$ 微分作用素環

$\partial^\alpha = \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n}$ 単項式

微分作用素環 R の元は x_1, \dots, x_n の滑らかな
関数 $f(x)$ に作用する。

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

例

$$(x_1^3 x_2^2 \partial_1^2 \partial_2^3) \bullet \cos(x_1^2 + x_2^2) = x_1^3 x_2^2 \frac{\partial^5 \cos(x_1^2 + x_2^2)}{\partial x_1^2 \partial x_2^3}$$

$$R = \mathbb{C}(x_1, \dots, x_n) \langle \partial_1, \dots, \partial_n \rangle$$

$$\partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}$$

$$a(x) \partial^\alpha \bullet f(x)$$

定義 左イデアル $I \subset R$ 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。

$$\dim_{\mathbb{C}(x_1, \dots, x_n)} R/I < \infty$$

例 ($n = 2$) $x = x_1$ $y = x_2$ $\partial_x = \partial_1$ $\partial_y = \partial_2$

$$R = \mathbb{C}(x, y) \langle \partial_x, \partial_y \rangle \quad I = (\partial_x^2 + y^2, \partial_y^2 + x^2)$$

すると、 $\dim_{\mathbb{C}(x, y)} R/I = 2$

$$\begin{aligned} \text{実際、} R/I &= \mathbb{C}(x, y) \cdot 1 + \mathbb{C}(x, y) \cdot \partial_y \\ &= \mathbb{C}(x, y) \cdot 1 + \mathbb{C}(x, y) \cdot \partial_x \end{aligned}$$

$$R = \mathbb{C}(x_1, \dots, x_n) \langle \partial_1, \dots, \partial_n \rangle \quad \partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} \quad a(x) \partial^\alpha \bullet f(x) \quad 0 \text{次元左イデアル}$$

多項式環のグレブナー基底の理論は**ほとんどそのまま**微分作用素環においても成立する。

$$I \subset R \quad 0 \text{次元左イデアル} \quad r = \dim_{\mathbb{C}(x_1, \dots, x_n)} R/I$$

- 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、
- $\mathbb{C}(x_1, \dots, x_n)$ 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

$$\begin{bmatrix} \partial_i t_1 \\ \partial_i t_2 \\ \vdots \\ \partial_i t_r \end{bmatrix} = P_i \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_r \end{bmatrix}$$

を満たすものを**計算することが可能**である。

$R = \mathbb{C}(x_1, \dots, x_n) \langle \partial_1, \dots, \partial_n \rangle \quad \partial^\alpha = \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n} \quad a(x) \partial^\alpha \bullet f(x) \quad 0$ 次元左イデアル
 $r = \dim_{\mathbb{C}(x_1, \dots, x_n)} R/I \quad t_1 (= 1), t_2, \dots, t_r$ 標準単項式 P_1, P_2, \dots, P_n Pfaffian 行列

例 $(n = r = 2) \quad R = \mathbb{C}(x, y) \langle \partial_x, \partial_y \rangle \quad I = (\partial_x^2 + y^2, \partial_y^2 + x^2)$

標準単項式 $1, \partial_y$

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

標準単項式 $1, \partial_x$

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

$R = \mathbb{C}(x_1, \dots, x_n) \langle \partial_1, \dots, \partial_n \rangle \quad \partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} \quad a(x) \partial^\alpha \bullet f(x) \quad 0$ 次元左イデアル
 $r = \dim_{\mathbb{C}(x_1, \dots, x_n)} R/I \quad t_1 (= 1), t_2, \dots, t_r$ 標準単項式 P_1, P_2, \dots, P_n Pfaffian 行列

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 であるとは、0次元左イデアル
 $I \subset R$ で $I \bullet f = 0$ (i.e., $l \bullet f = 0, \forall l \in I$)
 となるものが存在するときに言う。

例 $(n = 2) \quad R = \mathbb{C}(x, y) \langle \partial_x, \partial_y \rangle \quad f = -\cos(x^2 + y^2)$
 $l_1 \bullet f = 0 \quad l_2 \bullet f = 0$
 但し、 $l_1 = y\partial_x - x\partial_y \quad l_2 = \partial_x\partial_y + 4xy$
 $I = (l_1, l_2) \quad \dim_{\mathbb{C}(x, y)} R/I = 2$

練習問題 ホロノミック関数はどれか？

(a) $\sin x + \sin y \cos z$

(b) $1/\sin x$

(c) $\tan x$

(d) $\exp(x + y)$

(e) $f(x, y) = \int_0^{2\pi} \exp(x \cos t + y \sin t) dt$

解答

(a) ○ (b) × (c) × (d) ○ (e) ○

問題

ホロノミック関数 f と $I \bullet f = 0$ となる
0次元左イデアル $I \subset R$ があつたとき、
関数 f の局所的な最小値の近似値を計算せよ。

アルゴリズム (ホロノミック勾配降下法)

Let ε be a small positive number.

(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。

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

但し、

$$\frac{\partial \mathbf{F}}{\partial x_i} = \begin{bmatrix} \partial_i t_1 \bullet f \\ \partial_i t_2 \bullet f \\ \vdots \\ \partial_i t_r \bullet f \end{bmatrix} \quad \mathbf{F} = \begin{bmatrix} t_1 \bullet f \\ t_2 \bullet f \\ \vdots \\ t_r \bullet f \end{bmatrix}$$

(Step 3) Take a point c as a starting point and evaluate numerically \mathbf{F} at c Let $e = c$

アルゴリズム (ホロノミック勾配降下法)

Let ε be a small positive number.

(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。

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

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 c as a starting point and evaluate numerically \mathbf{F} at c Let $e = c$

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

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

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

Let $(\nabla(f))(\mathbf{e}) \neq 0$. Since $\varepsilon > 0$ is small,

$$\begin{aligned} \mathbf{F}(\mathbf{e} - \varepsilon(\nabla(f))(\mathbf{e})) &\approx \mathbf{F}(\mathbf{e}) - \sum_{i=1}^n \varepsilon (\partial f / \partial x_i)(\mathbf{e}) (\partial \mathbf{F} / \partial x_i)(\mathbf{e}) \\ &= \mathbf{F}(\mathbf{e}) - \sum_{i=1}^n \varepsilon (\partial f / \partial x_i)(\mathbf{e}) (P_i \mathbf{F})(\mathbf{e}) \end{aligned}$$

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)

例 ($n = 2$) $R = \mathbb{C}(x, y) \langle \partial_x, \partial_y \rangle$ $I = (y\partial_x - x\partial_y, \partial_x\partial_y + 4xy)$
 $r = \dim_{\mathbb{C}(x,y)} R/I = 2$ $I \bullet f = 0$

(Step 1) 標準單項式 $1, \partial_y$

(Step 2) Pfaffian 系

$$\frac{\partial \mathbf{F}}{\partial x} = \begin{bmatrix} 0 & x/y \\ -4xy & 0 \end{bmatrix} \mathbf{F}, \quad \frac{\partial \mathbf{F}}{\partial y} = \begin{bmatrix} 0 & 1 \\ -4y^2 & 1/y \end{bmatrix} \mathbf{F}, \quad \mathbf{F} = \begin{bmatrix} f \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}$$

(Step 3) starting point $\mathbf{c} = (1.4, 1.4)^\top$

$$\mathbf{F} = (0.7120, -1.9660)^\top \quad \text{Let } \mathbf{e} = \mathbf{c}$$

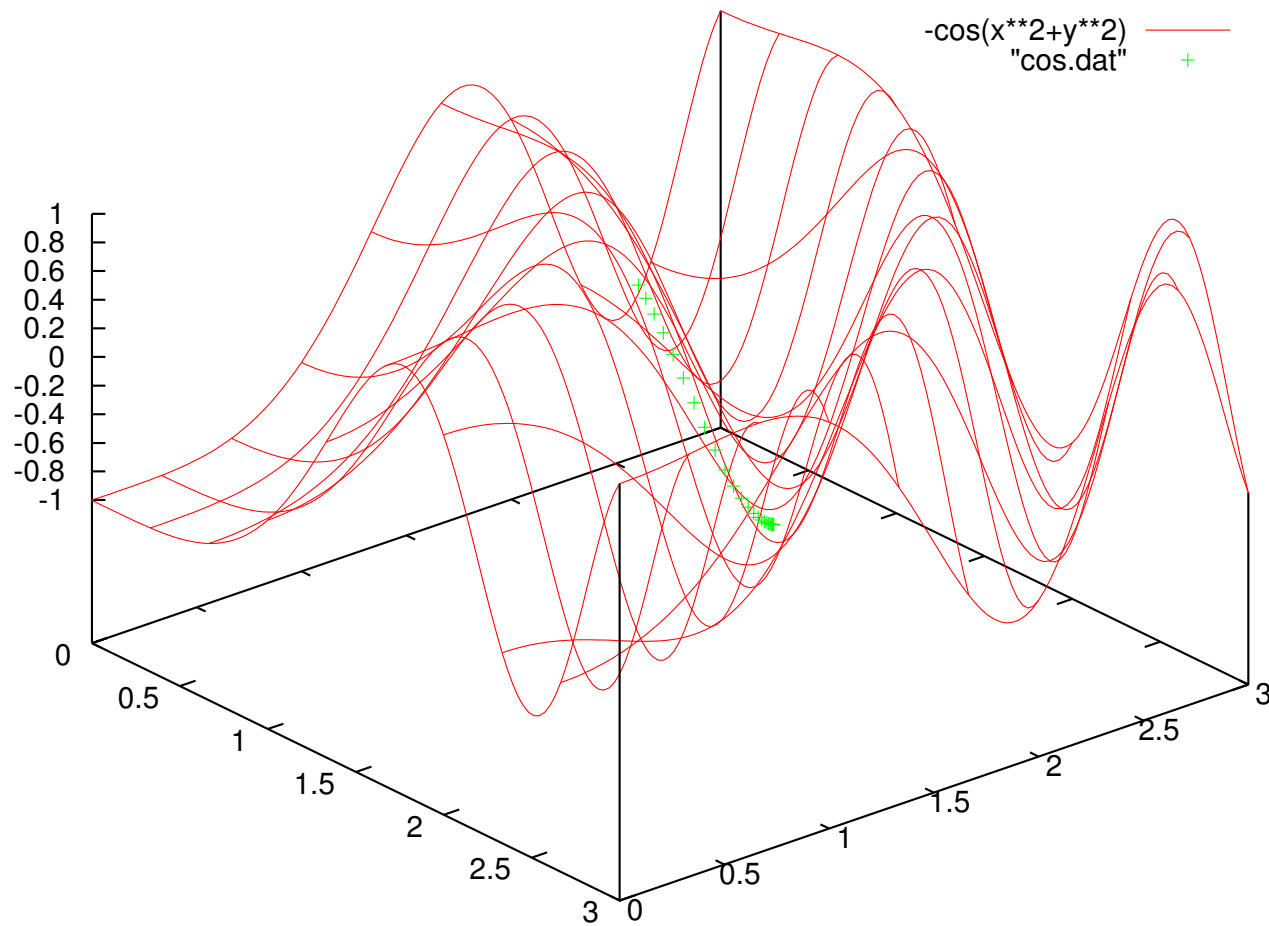
(Step 4)

$$\nabla(f) = \begin{bmatrix} (x/y) \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} (x/y) F_2 \\ F_2 \end{bmatrix} \quad (\nabla(f))(\mathbf{e}) = \begin{bmatrix} -1.9660 \\ -1.9660 \end{bmatrix}$$

(Step 5) $(\nabla(f))(\mathbf{e}) \neq 0$ $\varepsilon = 0.01$

$$\mathbf{e} - \varepsilon(\nabla(f))(\mathbf{e}) = (1.4197, 1.4197)^\top$$

$$\mathbf{F}(\mathbf{e} - \varepsilon(\nabla(f))(\mathbf{e})) = (0.6300, -2.2051)^\top$$



$$(x, y) = (1.7685, 1.7685)^{\top} \quad \mathbf{F} = (-0.9996, -0.09865)^{\top}$$

local minimum -0.9996

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。

確率密度関数

$$p(t; \theta_1, \theta_2) \propto f(t; \theta_1, \theta_2) = \exp(\theta_1 \cos t + \theta_2 \sin t)$$

正規化定数

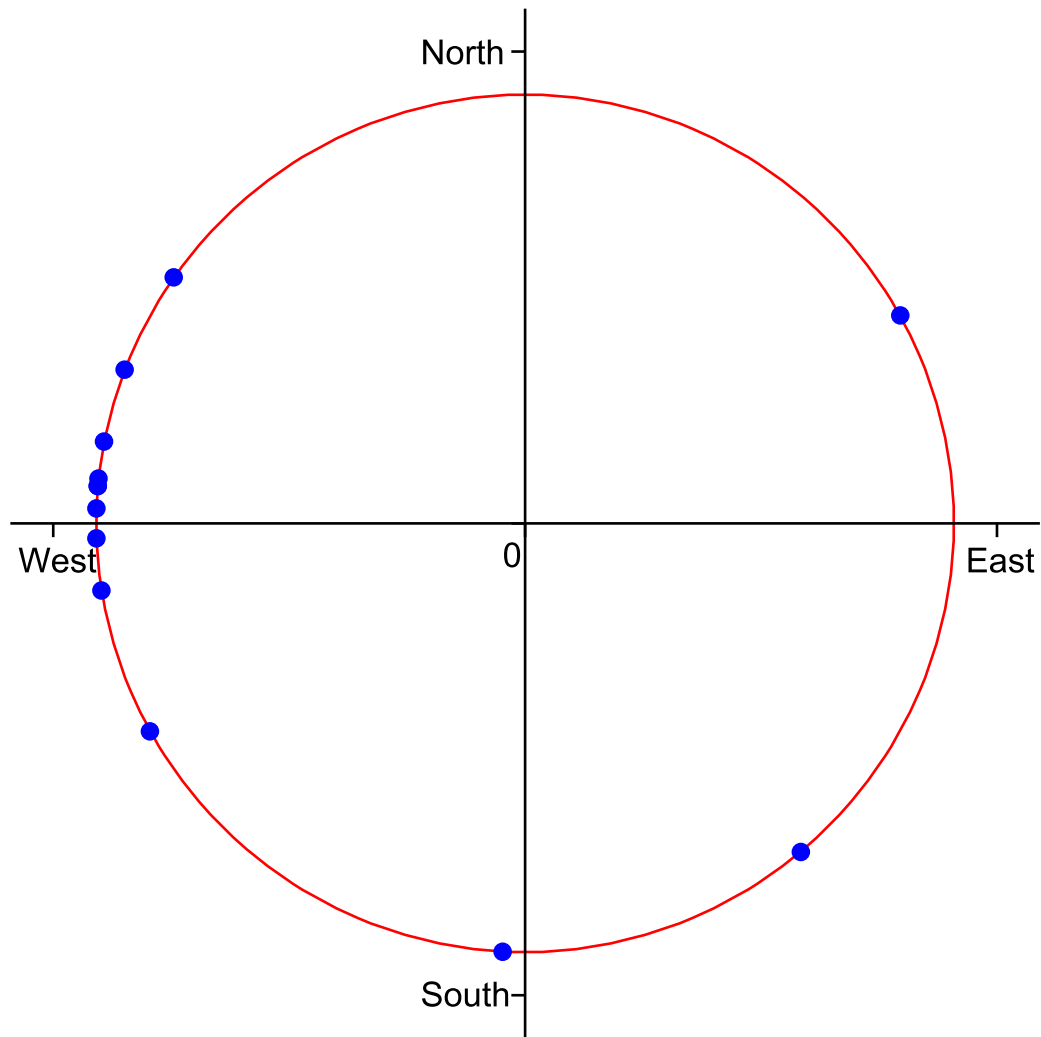
$$F(\theta_1, \theta_2) = \int_0^{2\pi} f(t; \theta_1, \theta_2) dt$$

観測データ T の確率

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

定義

非正規化確率密度関数がホロノミック関数である確率分布を**ホロノミック分布**と呼ぶ。



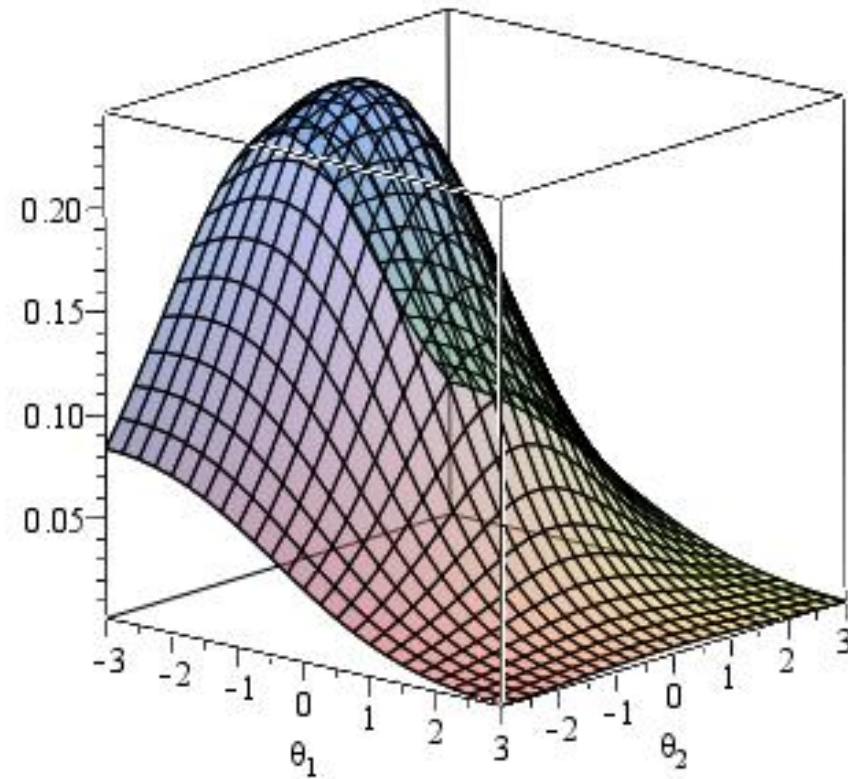
札幌上空約1万メートル地点における午前9時の風向きの13日間のデータ
(2011年1月1日から14日(11日は除外); 気象庁の気象統計情報から)

最尤法 (maximum likelihood estimation) とは何か？

観察データ T_1, \dots, T_N ($N = 13$) があつたとき
尤度関数 (likelihood function)

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

を最大化する θ_1 と θ_2 をパラメータの推定値とする。



likelihood function $L(\theta_1, \theta_2)$

ホロノミック勾配降下法をどのように使うか？

- 変数変換

$$\cos t = \frac{1-u^2}{1+u^2} \quad \sin t = \frac{2u}{1+u^2}$$

を施すと

$$F(\theta_1, \theta_2) = \int_{-\infty}^{\infty} g(u, \theta_1, \theta_2) du,$$

となる。但し、

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

- 多項式係数の微分作用素環（ワイル代数）を考える。

$$D_2 = \mathbb{C}\langle \theta_1, \theta_2, \partial_1, \partial_2 \rangle$$

$$D_3 = \mathbb{C}\langle u, \theta_1, \theta_2, \partial_u, \partial_1, \partial_2 \rangle$$

函数 g を零化するホロノミックイデアル $J \subset D_3$ 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)。

このとき、D加群の**積分アルゴリズム** (大阿久俊則) を

使うと、 $I \subset D_2$ が計算可能。

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

- 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。

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

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

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

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

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

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

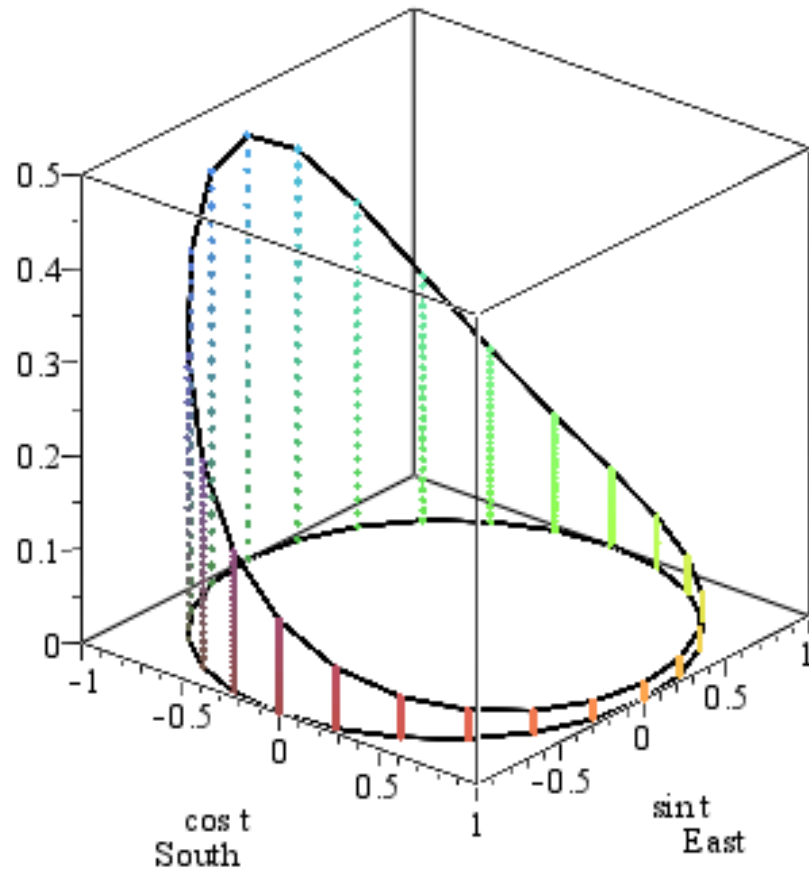
$$\begin{bmatrix} \frac{\partial L^*}{\partial \theta_1} \\ \frac{\partial^2 L^*}{\partial \theta_1 \partial \theta_2} \end{bmatrix} = P \begin{bmatrix} L^* \\ \frac{\partial L^*}{\partial \theta_2} \end{bmatrix}$$

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

但し、 $L^* = L(\theta_1, \theta_2)^{-1/N}$

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

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



$p(t; \theta_1, \theta_2)$ with $\theta_1 = -1.62, \theta_2 = -0.10$

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。

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、適用可能である。