

# 数学備忘録

motchy

2015 年 5 月 27 日 ~ 2022 年 5 月 25 日  
ver 0.12.0

# 目次

第 I 部	離散数学	10
第 I.1 章	表記	11
第 I.2 章	算術	12
I.2.1	諸定理	12
I.2.1.1	$a, b, c, d \in \mathbb{N}, ab = cd, \gcd(a, c) = \gcd(b, d) = 1 \Rightarrow a = d, b = c$	12
I.2.2	Extended Euclidean Algorithm	12
第 I.3 章	整数の合同	15
I.3.1	定数倍と加減乗除	15
I.3.2	$a, m \in \mathbb{N}$ が互いに素であるとき、 $0, a, 2a, \dots, (m-1)a$ を $m$ で割った余りは全て異なる	15
I.3.3	和のべき乗を展開	15
I.3.4	Fermat の小定理	16
第 I.4 章	$r$ 進法	17
I.4.1	$1/10$ が $2$ 進表記で無限小数になることの証明	17
第 I.5 章	公式	18
I.5.1	$\sum_{k=1}^n k^2 = n(n+1)(2n+1)/6$	18
I.5.2	$\sum_{k=0}^n (2k+1)^2 = (n+1)(2n+1)(2n+3)/3$	18
I.5.3	$1 \cdot 3 \cdot 5 \cdots (2n-1)2n = (2n)!/n!$	19
I.5.4	$a^n - b^n = (a-b) \sum_{k=0}^{n-1} a^{n-1-k} b^k$	19
I.5.5	Möbius の反転公式	19
第 I.6 章	数え上げ	21
I.6.1	“for $i = 1, \dots, n (\in \mathbb{N})$ ” $\iff$ “for $d \mid n$ , for $i = 1, \dots, n$ where $\gcd(i, n) = d$ ” $\iff$ “for $d \mid n$ , for $j = 1, \dots, n/d$ where $\gcd(j, n/d) = 1$ ”	21
I.6.2	包除原理	21
I.6.2.1	補題: $A \subseteq U \Rightarrow  A  = \sum_{x \in U}  A \cap \{x\} $	21
I.6.2.2	あの等式をどうやって思いついたか	23
I.6.3	配列の飛び飛びマーキング	24
第 I.7 章	置換	26
I.7.1	置換の符号の定理	26
I.7.1.1	$\sigma'(i) := \sigma(n+1-i)$ としたときの $\operatorname{sgn}(\sigma')$	26
I.7.1.2	$\sigma'(i) := n+1-\sigma(i)$ としたときの $\operatorname{sgn}(\sigma')$	26
第 I.8 章	有限幾何学	28
I.8.1	有限射影平面	28
I.8.1.1	位数 $n$ の有限射影平面には $n^2 + n + 1$ 個の点と直線が存在し、各点は $n+1$ 本の直線に含まれる。	28
第 II 部	代数学	30
第 II.1 章	群	31
II.1.1	剰余類	31
II.1.1.1	定義	31
II.1.1.2	群 $G$ の部分群 $H$ の全ての剰余類は互いに素である	31
II.1.1.3	2 つの元が同じ剰余類に属する必要十分条件	31
II.1.2	$\mathbb{Z}_p$ ( $p$ : 素数) に乗法群としての位数 $r$ の元があれば $p-1$ は $r$ の倍数	32
第 II.2 章	多項式	33
II.2.1	規約	33
II.2.2	係数が次数に関して対称な奇数次の多項式 $f(x)$ を $(x+1)$ で除したのも係数が次数に関して対称になる	33
II.2.3	$\sum_{i=0}^n a_i x^i$ が既約多項式ならば $\sum_{i=0}^n a_{n-i} x^i$ も既約多項式である	33
II.2.4	既約多項式 $\iff$ 最小多項式	34
第 II.3 章	有限体	35
II.3.1	素数位数の有限体の構成法	35
II.3.2	位数が素数の冪乗である有限体の構成法	35
II.3.2.1	既約多項式を用いる方法	35
II.3.2.2	既約多項式の根を用いる方法 (根体)	36
第 II.4 章	形式的冪級数	37
II.4.1	規約	37
II.4.2	原始多項式	37
II.4.2.1	$\sum_{i=0}^n a_i x^i$ が原始多項式ならば $\sum_{i=0}^n a_{n-i} x^i$ も原始多項式である	37
II.4.2.2	原始多項式の逆数の係数列は周期的	38
II.4.2.3	原始多項式 $f_1(x)$ の係数の順番を逆転した $f_2(x)$ の逆数の係数列は $f_1(x)^{-1}$ の係数列の逆順逆符号	39
II.4.2.3.1	例	39

II.4.2.3.2	証明	39
第 II.5 章	ブール代数	41
II.5.1	双対原理	41
II.5.2	応用	41
II.5.2.1	2 進 Gray コードへの変換	41
第 III 部	実解析	43
第 III.1 章	表記	44
第 III.2 章	離散数学との関係	45
III.2.1	Dirichlet の Diophantine 近似定理	45
III.2.2	床関数と天井関数	45
III.2.2.1	$\forall x \in \mathbb{R}, \lfloor -x \rfloor = -\lceil x \rceil, \lceil -x \rceil = -\lfloor x \rfloor$	45
III.2.2.2	$\forall x, y \in \mathbb{R}, \lfloor x \rfloor + \lfloor y \rfloor \leq \lfloor x + y \rfloor, \lceil x \rceil + \lceil y \rceil \geq \lceil x + y \rceil$	45
第 III.3 章	極限	47
III.3.1	問題例	47
III.3.1.1	$\lim_{x \rightarrow \infty} \sqrt{x^2 + 1} - x = 0$	47
III.3.1.2	$a > 0, \lim_{x \rightarrow +0} \frac{x^2}{\sqrt{x^2 + a^2} - a} = 2a$	47
第 III.4 章	数列	48
III.4.1	単調性	48
III.4.1.1	$(1 - \frac{t}{n})^n$ ( $t \in [0, n]$ ) は $n$ に関して単調増加	48
III.4.2	極限	48
III.4.2.1	$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n} + \sum_{k=2}^{\infty} \frac{c_k}{n^k}\right)^n = e$ ( $\sum_{k=2}^{\infty} c_k$ converges)	48
III.4.2.2	$\lim_{n \rightarrow \infty} (1 + \frac{1}{n} + o(1/n))^n = e$	49
III.4.3	漸化式	49
III.4.3.1	Fibonacci 数列	49
III.4.3.1.1	つがいの増殖からの漸化式の導出	49
III.4.3.2	$a_{n+1} = ra_n + p(n)$ , ( $r \neq 1, p(n): d$ 次多項式) の一般項は $a_n = Ar^n + q(n)$ , ( $A: 定数, q(n): d$ 次多項式)	49
第 III.5 章	級数	50
III.5.1	$S_l = \sum_{k=0}^{\infty} k^l r^k$ を逐次的に求める	50
III.5.2	Abel の総和公式: $\sum_{i=1}^n a_i b_i = A_n b_n - \sum_{i=1}^{n-1} A_i (b_{i+1} - b_i)$	50
III.5.3	Chebyshev の和の不等式	51
III.5.4	等比級数と 0 収束列の畳み込み: $ r  < 1, \lim_{n \rightarrow \infty} a_n = 0 \Rightarrow \lim_{t \rightarrow \infty} \sum_{\tau=0}^t r^{t-\tau} a_{\tau} = 0 = 0$	52
III.5.5	$\sum_{k=1}^n \cos kx = -\frac{1}{2} + \frac{\sin(n+1/2)x}{2 \sin(x/2)} \quad (x \neq 0)$	52
III.5.6	$\sum_{k=1}^n \sin kx = \frac{\sin \frac{nx}{2} \sin \frac{(n+1)x}{2}}{2 \sin(x/2)} \quad (x \neq 0)$	53
第 III.6 章	乗積	54
III.6.1	$a_i \geq 0$ ( $i = 1, 2, \dots$ ), $\sum_{i=1}^{\infty} (1 - a_i) = \infty \Rightarrow \prod_{i=1}^{\infty} a_i = 0$	54
第 III.7 章	位相空間	55
III.7.1	$A \subseteq B \Rightarrow \text{cl } A \subseteq \text{cl } B$	55
第 III.8 章	部分分数分解	56
III.8.1	諸公式	56
III.8.1.1	$\frac{1}{(x^2-1)^2} = \frac{1}{4} \left[ \frac{1}{x+1} - \frac{1}{x-1} + \frac{1}{(x+1)^2} + \frac{1}{(x-1)^2} \right]$	56
III.8.2	諸定理	56
III.8.2.1	定理 III.8.2.1.0	56
第 III.9 章	指示関数	58
III.9.1	諸定理	58
III.9.1.1	$\Omega$ を定義域とする写像 $f, g, T$ に対して $\forall x \in \Omega, \forall t \in T(\Omega), \mathbf{1}\{T(x) = t\} f(x) = \mathbf{1}\{T(x) = t\} g(x) \iff \forall x \in \Omega, f(x) = g(x)$	58
第 III.10 章	微分	59
III.10.1	$f(a) < 0 < f(b) \Rightarrow \exists x \in (a, b) \text{ s.t. } f(x) = 0, f'(x) \geq 0$	59
III.10.2	$\lim_{x \rightarrow \infty} f(x) = 0$ であっても $\lim_{x \rightarrow 0} f'(x) = 0$ とは限らない	59
III.10.3	soft-step 関数	59
III.10.3.1	$C^{n-1}$ 級の soft-step 関数	60
第 III.11 章	Taylor 級数	61
III.11.1	$e^{-x^2}$ は全域で Taylor 展開可能	61
第 III.12 章	凸関数	62
III.12.1	諸定理	62
III.12.1.1	$f_1, f_2$ が凸なら $\max\{f_1, f_2\}$ は凸	62
III.12.1.2	Jensen の不等式	62
III.12.1.3	重み付き相加, 相乗平均の関係と Young の不等式	63
III.12.1.4	2 階導関数と凸性の関係	64
III.12.2	Legendre 変換	65
III.12.2.1	十分滑らかな狭義凸関数に対して Legendre 変換を 2 回施すと元の関数に戻る	65
III.12.3	諸注意	65
III.12.3.1	非負の凸関数同士の積は凸関数とは限らない	65
第 III.13 章	逆関数	66
III.13.1	連続関数の逆関数は連続	66
III.13.2	可微分な関数 $f$ の逆関数は $f'(f^{-1}(x)) \neq 0$ なる点 $x$ で微分可能であり、 $\frac{d f^{-1}}{d x}(x) = 1/f'(f^{-1}(x))$ 。	67
第 III.14 章	諸定理	69

III.14.1	偶関数と奇関数への分解	69
III.14.2	$\sup_{\mathbf{x} \in X} (f(\mathbf{x}) + g(\mathbf{x})) \leq \sup_{\mathbf{x} \in X} f(\mathbf{x}) + \sup_{\mathbf{x} \in X} g(\mathbf{x}), \inf_{\mathbf{x} \in X} f(\mathbf{x}) + \inf_{\mathbf{x} \in X} g(\mathbf{x}) \leq \inf_{\mathbf{x} \in X} (f(\mathbf{x}) + g(\mathbf{x}))$	70
III.14.3	$\sup_{\mathbf{x} \in X} (-f(\mathbf{x})) = -\inf_{\mathbf{x} \in X} f(\mathbf{x}), \inf_{\mathbf{x} \in X} (-f(\mathbf{x})) = -\sup_{\mathbf{x} \in X} f(\mathbf{x})$	70
III.14.4	Hölder(ヘルダー) の不等式	70
III.14.5	Minkowski の不等式 (三角不等式の一般化)	71
第 III.15 章	2 変数関数	72
III.15.1	4 象限逆正接: $\text{Tan}^{-1}(x, y)$	72
第 III.16 章	多変数関数	73
III.16.1	Jacobi 行列	73
III.16.1.1	$J_{G \circ F} = J_G J_F$	73
III.16.1.2	$J_{f^\top g} = \mathbf{g}^\top J_f + \mathbf{f}^\top J_g$	73
III.16.2	勾配と Hessian	73
III.16.2.1	$\nabla(G \circ F) = \nabla F \nabla G$	73
III.16.2.2	$\nabla_{\mathbf{xy}} f = (\nabla_{\mathbf{yx}} f)^\top$	74
III.16.2.3	勾配ベクトルが等値面の接平面と直交することの証明	74
III.16.2.4	$\nabla A\mathbf{x} = A^\top$	74
III.16.2.5	$\nabla \mathbf{x}^\top \mathbf{x} = 2\mathbf{x}$	74
III.16.2.6	$\nabla \mathbf{x}^\top A\mathbf{x} = (A + A^\top)\mathbf{x}$	75
III.16.2.7	$\nabla \ A\mathbf{x} + \mathbf{b}\ _2^2 = 2A^\top(A\mathbf{x} + \mathbf{b})$	75
III.16.2.8	$\nabla(A\mathbf{x} + \mathbf{b})^\top B(A\mathbf{x} + \mathbf{b}) = A^\top(B + B^\top)(A\mathbf{x} + \mathbf{b})$	75
III.16.2.9	$\nabla^2 \mathbf{x}^\top A\mathbf{x} = A + A^\top$	75
III.16.2.10	公式集	75
III.16.3	逆関数定理	76
III.16.4	2 次形式	77
III.16.4.1	$\forall A \in \mathbb{R}^{n \times n}, \mathbf{x} \in \mathbb{R}^n, \exists$ 対称行列 $B \in \mathbb{R}^{n \times n}$ s.t. $\mathbf{x}^\top A\mathbf{x} = \mathbf{x}^\top B\mathbf{x}$	77
第 III.17 章	測度	78
III.17.1	$\sigma$ -加法族	78
III.17.1.1	2 つの $\sigma$ -加法族の和集合は $\sigma$ -加法族になるとは限らない	78
III.17.1.2	$\sigma$ -加法族の要素同士の直積集合の族は $\sigma$ -加法族になるとは限らない	78
III.17.1.3	$X$ を集合とする。 $A, B \subset 2^X, A \subset B \Rightarrow \sigma(A) \subset \sigma(B)$	78
III.17.2	Jordan 測度	78
III.17.2.1	Jordan 測度の有限加法性	79
III.17.2.2	$A \subset E$ ( $E$ は直方体) のとき $\underline{m}_J(A) =  E  - \overline{m}_J(A^c \cap E)$	79
III.17.2.3	系: $A \subset E$ ( $E$ は直方体) が Jordan 可測のとき $A^c \cap E$ も Jordan 可測である	79
III.17.3	Lebesgue 測度	80
III.17.3.1	$A \subset B$ ならば $\overline{m}_L(A) \leq \overline{m}_L(B)$	80
III.17.3.2	$\overline{m}_L(A \cup B) \leq \overline{m}_L(A) + \overline{m}_L(B)$	80
III.17.3.3	Lebesgue 外測度の劣加法性: $\overline{m}_L(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \overline{m}_L(A_n)$	80
III.17.3.4	有界集合 $A$ に対して $\overline{m}_L(A) \leq \overline{m}_J(A)$	80
III.17.3.5	有界集合 $A \subset \mathbb{R}^n$ に対して $\overline{m}_L(A) = 0 \iff m_J(A) = 0$	81
III.17.3.6	$A \subset \mathbb{R}^n$ とする。任意の直方体 $E$ に対して $\overline{m}_L(A \cap E) + \overline{m}_L(A^c \cap E) = \overline{m}_L(E) \iff$ 任意の $B \subset \mathbb{R}^n$ に対して $\overline{m}_L(A \cap B) + \overline{m}_L(A^c \cap B) = \overline{m}_L(B)$	81
III.17.3.7	$A_1, A_2 \subset \mathbb{R}^n$ が Jordan 可測ならば Lebesgue 可測であり、 $m_L(A) = m_J(A)$	81
III.17.3.8	$A_1, A_2 \subset \mathbb{R}^n$ が Lebesgue 可測なら $A_1 \cup A_2$ も Lebesgue 可測	81
III.17.3.9	系: $A_1, \dots, A_n \subset \mathbb{R}^n$ が Lebesgue 可測なら $A_1 \cup A_2 \cup \dots \cup A_n$ も Lebesgue 可測	82
III.17.3.10	Lebesgue 可測集合 $A_1, A_2 \subset \mathbb{R}^n$ が非交差的なら $m_L(A_1 \cup A_2) = m_L(A_1) + m_L(A_2)$	82
III.17.3.11	系: Lebesgue 測度の有限加法性	82
III.17.3.12	Lebesgue 測度の $\sigma$ -加法性	82
III.17.3.13	Lebesgue 可測の直感的理解	83
III.17.3.14	Jordan 測度と Lebesgue 測度の違い	83
III.17.4	包除原理	84
第 III.18 章	積分	86
III.18.1	変数変換: $\int_{x(t_0)}^{x(t)} f(x)dx = \int_{t_0}^t f(x)x'(t)dt$	86
III.18.2	Chebyshev の和の不等式: 可積分な単調増加関数 $f, g$ に対して $(\int_0^1 f(x)dx)(\int_0^1 g(x)dx) \leq \int_0^1 f(x)g(x)dx$	86
第 III.19 章	重積分	87
III.19.1	平均値の定理	87
III.19.2	$C^1$ 級の変換による面積 0 の集合の像の面積もまた 0	87
III.19.3	関数 $f, g$ が $\mathbb{R}$ 上で有界, 連続で $\int_{-\infty}^{\infty} f(x)dx, \int_{-\infty}^{\infty} g(x)dx$ が絶対収束するなら畳み込み $f * g$ は $\mathbb{R}$ 上で連続で絶対可積分である。	88
III.19.4	諸公式	89
III.19.4.1	$n$ 次元の単体の体積 $\int_{x_1=0}^a \int_{x_2=0}^{a-x_1} \dots \int_{x_n=0}^{a-x_1-\dots-x_{n-2}} dx_n \dots dx_2 dx_1 = a^n/n! (a > 0)$	89
第 III.20 章	線積分	90
III.20.1	定義	90
III.20.2	計算法	90
III.20.3	Green の定理の拡張	91
第 III.21 章	変数変換	92
III.21.1	三角関数	92
III.21.1.1	$x, y \in \mathbb{R}, x^2 + y^2 = 1 \iff x = \cos \theta, y = \sin \theta$	92
III.21.2	極座標表示	92
III.21.2.1	極方程式から曲線を導くときの注意: $r, \theta$ 空間の 2 つの異なる集合から $x, y$ 空間へ写像した 2 つの集合の排他的論理和が $\emptyset$ なら $x, y$ 空間において両者の違いは無い	92
第 III.22 章	Beta 関数	93
III.22.1	Beta 関数と Gamma 関数の関係	93
III.22.2	多変量 Beta 関数	93

III.22.3	多変量 Beta 関数と Gamma 関数の関係	94
第 III.23 章	冪級数関数	95
III.23.1	2 次以上の項が $o(x)$ ( $x \rightarrow 0$ ) であること	95
第 III.24 章	微分と積分の関係	96
III.24.1	$\lim_{x \rightarrow \infty} f(x)$ が存在すれば $\int_0^\infty f'(x)dx = f(\infty) - f(0)$	96
III.24.2	$\frac{d}{dx} \int_a^x f(x)dx = f(x)$	96
III.24.3	畳み込みの微分 $\frac{d}{dx} \int_{s=0}^x f(s)g(x-s)ds = f(x)g(0) + \int_{s=0}^x f(s) \frac{d}{dx} g(x-s) ds$	97
III.24.4	重積分の微分 $\frac{d}{dx} \left( \int_{x_1=0}^x \cdots \int_{x_n=0}^{x-x_1-\cdots-x_{n-1}} f_1(x_1) \cdots f_n(x_n) dx_n \cdots dx_1 \right)$ $= \int_{x_1=0}^x \cdots \int_{x_{n-1}=0}^{x-x_1-\cdots-x_{n-2}} f_1(x_1) \cdots f_{n-1}(x_{n-1}) f_n(x-x_1-\cdots-x_{n-1}) dx_{n-1} \cdots dx_1$	97
III.24.5	系 $\frac{d}{dx} \left( \int_{x_1=-\infty}^\infty \cdots \int_{x_n=-\infty}^{x-x_1-\cdots-x_{n-1}} f_1(x_1) \cdots f_n(x_n) dx_n \cdots dx_1 \right)$ $= \int_{x_1=-\infty}^\infty \cdots \int_{x_{n-1}=-\infty}^{x-x_1-\cdots-x_{n-2}} f_1(x_1) \cdots f_{n-1}(x_{n-1}) f_n(x-x_1-\cdots-x_{n-1}) dx_{n-1} \cdots dx_1$	99
第 III.25 章	図形への応用	101
III.25.1	ある超平面から距離 $d$ だけ離れた超平面の方程式	101
III.25.2	$n$ 次元単位球の体積 $V_n$ は $\pi^{n/2}(\Gamma(1+n/2))^{-1}$	101
第 III.26 章	最適化への応用	103
III.26.1	狭義凸関数の線形制約下での最小化	103
III.26.2	変分法	104
III.26.2.1	大域的最小点の十分条件	104
III.26.3	例題	105
III.26.3.1	非負かつ総和 1 の下での 2 乗和の最小化	105
III.26.3.2	正かつ総和 1 の下での逆数 and の最小化	105
第 IV 部	複素解析	106
第 IV.1 章	複素数	107
IV.1.1	複素数の定義	107
IV.1.1.1	注意	108
IV.1.1.2	コラム $i^2 = -1$ ?	108
IV.1.2	諸公式	108
IV.1.2.1	$x > 0, z \in \mathbb{C}$ のとき $ x^z  = x^{\operatorname{Re}(z)}$	108
IV.1.2.2	$\forall n \nmid k, \sum_{l=1}^n \exp(i \frac{k}{n} 2\pi l) = 0$	109
IV.1.3	代数方程式	109
IV.1.3.1	諸定理	109
IV.1.3.1.1	実係数 $n$ 次方程式が複素数解 $s = \alpha + i\beta$ を解に持つならばその複素共役 $s^* = \alpha - i\beta$ もまた解である	109
IV.1.3.2	例題	109
IV.1.3.3	$z^n + z^{n-1} + \cdots + z^2 + z + 1 = 0$	109
第 IV.2 章	複素指数凸関数	111
IV.2.1	複素指数凸関数の例	111
IV.2.1.1	$\ \sum_{i=1}^N X_i \mathbf{a}_i + \mathbf{b}\ _2^2$ ( $\mathbf{a}_i, \mathbf{b} \in \mathbb{C}^n, X_i \in \mathbb{C}^{m \times n}$ ) は $X_1, \dots, X_N$ に関して凸である。	111
IV.2.1.2	$\ A\mathbf{x} + \mathbf{b}\ _2^2$ が狭義凸 $\iff A^* A \succ O$	111
第 IV.3 章	特殊関数	113
IV.3.1	Gamma 関数	113
IV.3.1.1	$\lim_{n \rightarrow \infty} \int_0^n t^{z-1} (1+t/n)^n dt = \Gamma(z)$ ( $z \in \mathbb{C}, \operatorname{Re}(z) > 0$ )	113
第 IV.4 章	複素積分	114
IV.4.1	無限に大きい半円周上での $e^{iz}/z$ の積分	114
IV.4.2	例題	114
IV.4.2.1	$\int_0^{2\pi} \operatorname{Log}(1 - a \exp(i\theta)) d\theta = 0$ ( $C$ : 原点中心の単位円, $a \in \mathbb{R}$ )	114
第 IV.5 章	Fourier 級数	116
IV.5.1	連続関数の Fourier 級数の高周波成分は 0 に収束する	116
IV.5.2	$\int_{-\pi}^\pi \frac{\sin nx}{2 \tan \frac{x}{2}} dx = \pi$ ( $n \in \mathbb{N}$ )	117
IV.5.3	Fourier 級数の応用	117
IV.5.3.1	$\int_0^\infty \frac{\sin x}{x} dx = \pi/2$	117
第 IV.6 章	Fourier 変換	119
IV.6.1	余弦, 正弦変換	119
第 IV.7 章	Bessel 関数	121
IV.7.1	定義	121
IV.7.1.1	簡単化	121
IV.7.2	性質	121
IV.7.2.1	$J_n(-z) = J_{-n}(z)$	121
IV.7.2.2	$J_{2m}(-z) = J_{2m}(z), J_{2m+1}(-z) = -J_{2m+1}(z)$	121
IV.7.3	Maclaurin 展開	122
第 V 部	線形代数	123
第 V.1 章	表記	124
第 V.2 章	行列の演算	125

V.2.1	結合則	125
V.2.2	上(下)三角行列のべき乗	125
V.2.3	列ベクトルと行ベクトルの積の和	126
V.2.4	巡回行列の可換則	126
V.2.5	$\mathbb{R}^2$ の対称行列同士の積は対称	127
V.2.6	Kronecker 積	127
V.2.6.1	混合積: $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$	127
V.2.6.2	ユニタリ行列同士の Kronecker 積はユニタリ行列である	127
第 V.3 章	ベクトル空間	129
V.3.1	基底	129
V.3.1.1	基底変換	129
V.3.1.2	諸定理	129
V.3.1.2.1	三角関数ベクトルによる直交展開	129
V.3.2	一般のベクトル空間と配列型ベクトル空間の橋渡し	130
V.3.2.1	合成ベクトルの一次独立性と係数ベクトルの一次独立性は等価	130
V.3.3	諸定理	130
V.3.3.1	ベクトル空間 $V$ の部分空間 $V_1, V_2$ について、「 $V_1 \cup V_2$ がベクトル空間である」 $\iff$ 「 $V_1 \subseteq V_2$ または $V_1 \supseteq V_2$ 」	130
V.3.3.2	$V_i \subseteq W_i, \bigoplus_{i=1}^n W_i = \bigoplus_{i=1}^n V_i \Rightarrow V_i = W_i$	131
V.3.3.3	一次独立なベクトル $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ に $\mathbf{v} \in V$ を加えたものが一次独立 $\iff \mathbf{v}$ が $\mathbf{v}_1, \dots, \mathbf{v}_k$ の一次結合で表せない	131
V.3.3.4	$A_v := \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ と $A_w := \{\mathbf{w}_1, \dots, \mathbf{w}_l\}$ が各々一次独立で $k > l \Rightarrow \exists \mathbf{v} \in A_v$ s.t. $\mathbf{w}_1, \dots, \mathbf{w}_l, \mathbf{v}$ が一次独立	132
V.3.4	諸注意	132
V.3.4.1	$\{\mathbf{v}_1, \mathbf{v}_2\}, \{\mathbf{w}_1, \mathbf{w}_2\}, \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_1\}, \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_2\}$ が各々一次独立でも $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_1, \mathbf{w}_2\}$ が一次独立とは限らない	132
V.3.4.2	$W_i \cap W_j = \{\mathbf{0}\}$ ( $i \in \{1, \dots, n\}, i \neq j$ ) であっても $W_1 + \dots + W_n = W_1 \oplus \dots \oplus W_n$ とは限らない	132
V.3.4.3	$W = V \oplus X, W = V \oplus Y$ でも $X = Y$ とは限らない	132
V.3.4.4	$\text{span}[\mathbf{a}, \mathbf{b}, \mathbf{c}] / \text{span}[\mathbf{a}, \mathbf{b}] = \text{span}[\mathbf{c}]$ とは限らない	132
第 V.4 章	線形写像	133
V.4.1	表記の規則	133
V.4.1.1	一般のベクトルを要素とする配列と行列の積	133
V.4.1.2	一般のベクトルを要素とする行列同士の積	133
V.4.2	$\dim(V) = \dim(W), f: V \rightarrow W$ であるとき、 $f$ が単射であることと全射であることは同値。	133
V.4.3	$\dim V = \dim W$ ならば全単射線形写像 $f: V \rightarrow W$ が存在する	134
V.4.4	不変部分空間と表現行列	135
V.4.4.1	正規直交基底間での座標変換は直交変換	135
第 V.5 章	双対空間	137
V.5.1	基底変換と座標変換の関係	137
第 V.6 章	行列式	138
V.6.1	$A \in \mathbb{R}^{n \times m}, B \in \mathbb{R}^{m \times n}$ に対して $ I_m - AB  =  I_n - BA $	138
V.6.2	系: $ I_n - \mathbf{v}\mathbf{v}^\top  = 1 - \ \mathbf{v}\ _2^2$	138
V.6.3	逆対角転置行列の行列式は元の行列のそれと等しい	138
第 V.7 章	逆行列	140
V.7.1	$[\phi_1, \dots, \phi_n]^{-1} \phi_i = \mathbf{e}_i$	140
V.7.2	(系) 結合係数の抽出	140
V.7.3	上(下)三角行列が逆行列をもてば、それも上(下)三角行列である	140
V.7.4	正則行列 $A$ の列(行)が直交系を成すとき、 $A^{-1}$ の行(列)は直交系を成す	141
V.7.5	1 列(or 1 行)だけ置き換えた行列の逆行列	141
V.7.6	列(or 行)を入れ替えた行列の逆行列	142
V.7.7	Sherman-Morrison の公式の特別な場合: $(I + \mathbf{u}\mathbf{v}^*)^{-1} = (I - \mathbf{u}\mathbf{v}^*/(1 + \mathbf{v}^*\mathbf{u}))$	142
V.7.8	Sherman-Morrison の公式: $(A + \mathbf{u}\mathbf{v}^*)^{-1} = (I - A^{-1}\mathbf{u}\mathbf{v}^*/(1 + \mathbf{v}^*A^{-1}\mathbf{u}))A^{-1}$	143
V.7.9	Woodbury の公式の特別な場合: $(I_n + UV)^{-1} = I_n - U(I_k + VU)^{-1}V$	143
V.7.10	Woodbury の公式: $(A + UCV)^{-1} = [I_n - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}V]A^{-1}$	144
第 V.8 章	特性多項式	145
V.8.1	逆行列の特性多項式: $\phi_{A^{-1}}(\lambda) =  A ^{-1}(-\lambda)^n \phi_A(1/\lambda)$	145
第 V.9 章	ユニタリ行列	146
V.9.1	ユニタリ行列の複素共役はユニタリ行列	146
第 V.10 章	行列の階数	147
V.10.1	$\text{rank}(AB) = \text{rank}(BA)$ とは限らない	147
V.10.2	$T$ が正則でも $\text{rank}(ATB) = \text{rank}(AB)$ とは限らない	147
V.10.3	フルランク条件	147
V.10.4	$A \in \mathbb{C}^{m \times n}, \text{rank}(A) = \text{rank}(A^*A) = \text{rank}(AA^*)$	147
V.10.5	$A \in \mathbb{F}^{m \times n}$ の階数が $n$ ならば $A^*A$ は正則	148
V.10.6	(系) $A \in \mathbb{F}^{m \times n}$ が行フルランクならば $AA^\top$ は正則	148
V.10.7	$A \in \mathbb{F}^{m \times n}$ の階数が $r$ であるなら、ある正則行列 $T_x \in \mathbb{F}^{n \times n}, T_y \in \mathbb{F}^{m \times m}$ が存在して...	149
V.10.8	同じ型の行列の階数が等しいための必要十分条件	149
V.10.9	優対角正方行列は正則	150
V.10.10	最大ランク分解	150
第 V.11 章	固有値, 固有ベクトル	151
V.11.1	諸定理	151
V.11.1.1	相異なる固有値に対応する固有ベクトルは一次独立	151
V.11.1.2	$n \times n$ 行列 $A$ の全要素が $a \neq 0$ であるなら非零固有値は $an$ のみ	152
V.11.1.3	転置行列の固有値, 固有ベクトル	152
V.11.1.4	逆行列の固有値, 固有ベクトル	152
V.11.1.5	ユニタリ行列の全ての固有値の絶対値は 1 である	153
V.11.1.6	コンパニオン行列	153
V.11.1.7	巡回行列の固有値	154

V.11.1.8	Hadamard 行列の固有値	155
V.11.1.9	最小消去多項式が最小多項式と一致するベクトルの存在	155
V.11.2	スペクトル写像定理	156
V.11.3	対角化	157
V.11.3.1	正規行列 $\iff$ ユニタリ行列で対角化可能	157
V.11.4	一般固有空間	157
V.11.4.1	一般固有空間の階数の頭打ち	157
V.11.4.2	一般固有空間への直和分解	158
V.11.4.3	代数的重複度 $\geq$ 幾何学的重複度	159
V.11.4.4	$AB = O$ でも $\mathbb{R}^n = \text{Ker}(A) + \text{Ker}(B)$ とは限らない	159
V.11.5	Jordan 標準形	160
V.11.5.1	Jordan ブロックの逆行列	160
V.11.5.2	Jordan ブロックの指数行列	161
V.11.6	Perron の定理	162
第 V.12 章	置換行列	164
V.12.1	定義	164
V.12.2	諸定理	164
V.12.2.1	行と列の同期入れ替え	164
V.12.2.2	置換行列は直交行列	164
第 V.13 章	トレース	165
V.13.1	$\text{tr}(AB) = \text{tr}(BA)$	165
V.13.2	Hermite 行列 $A, B$ に対して $\text{tr}(AB) \in \mathbb{R}$	165
V.13.3	正定値行列 $A \in \mathbb{R}^{m \times m}$ に対して $\text{tr}(X^\top AX)$ ( $X \in \mathbb{R}^{m \times n}$ ) は $X$ の狭義凸関数である	165
第 V.14 章	定値性	167
V.14.1	正定 Hermite 対称行列の逆行列も正定	167
V.14.2	$A$ が半正定で $\mathbf{x}^\top A \mathbf{x} = 0$ ならば $\mathbf{x} \in \text{Ker}(A)$	167
V.14.3	Gram 行列は半正定	167
V.14.4	正定値行列同士の積は正定値とは限らない	167
V.14.5	対角成分が全て非負の対称行列が半正定とは限らない	168
V.14.6	非対称行列は固有値が全て正でも正定とは限らない	168
V.14.7	$A$ : Hermite 対称で正定 $\Rightarrow (\mathbf{x}^* A \mathbf{x})(\mathbf{y}^* A^{-1} \mathbf{y}) \geq  \mathbf{x}^* \mathbf{y} ^2$	168
第 V.15 章	擬逆行列	169
V.15.1	諸定理: $(A^H A)^\dagger A^H = A^\dagger$ など	169
V.15.2	行フルランクな行列の擬逆行列を用いた線形方程式の解がノルム最小であること	169
第 V.16 章	ベクトルのノルム	171
V.16.1	諸定理	171
V.16.1.1	$\ \mathbf{x}\ _{p+a} \leq \ \mathbf{x}\ _p$ ( $\mathbf{x} \in \mathbb{R}^n, p \geq 1, a \geq 0$ )	171
V.16.1.2	$1 \leq p \leq q, \mathbf{x} \in \mathbb{C}^n$ のとき $\ \mathbf{x}\ _p \leq n^{1/p-1/q} \ \mathbf{x}\ _q$	171
V.16.1.3	和のノルムとノルムの和が等しい時	172
V.16.1.4	和のノルムとノルムの和が等しい時 (2)	172
V.16.1.5	$\mathbf{x} \in \mathbb{C}^n, \ \mathbf{x}\ _2 = 1 \Rightarrow \sum_{i=1}^n  x_i  \leq \sqrt{n}$	173
V.16.2	応用	173
V.16.2.1	超平面と点の距離公式の直感的説明	173
第 V.17 章	行列のノルム	174
V.17.1	2-ノルム (2-演算子ノルム)	174
V.17.1.1	定義	174
V.17.1.2	部分行列のノルム: $A \in \mathbb{C}^{m \times n}, \ [A]_{i,j}\ _2 \leq \ A\ _2$	175
V.17.1.3	スペクトル半径の上界	175
V.17.1.4	Hermite 行列の絶対値最大固有値の絶対値は 2-演算子ノルムと一致する	175
V.17.1.5	$\max_{\ \mathbf{x}\ =1} \ A\mathbf{x}\  = \sqrt{\lambda_{\max}(A^\top A)}$ ( $A \in \mathbb{R}^{n \times n}$ )	176
V.17.1.6	逆行列のノルムは元の行列の最小特異値の逆数と等しい	177
V.17.1.7	$A \in \mathbb{C}^{m \times n}$ について $\alpha = \max  A_{ij} $ とすると $\alpha \leq \ A\ _2$	177
V.17.1.8	$A \in \mathbb{C}^{m \times n}$ の各列ベクトルのノルムが高々 $a$ ならば $\ A\ _2 \leq a\sqrt{n}$	178
V.17.1.8.1	系: $A \in \mathbb{C}^{m \times n}$ の全要素の絶対値が $\varepsilon$ 以下であれば $\ A\ _2 \leq \varepsilon\sqrt{mn}$	178
V.17.1.8.2	系: $n$ 次確率行列のノルムは高々 $\sqrt{n}$	178
第 V.18 章	凸領域	179
V.18.1	凸領域内のベクトルの凸結合は元の凸領域に属す	179
第 V.19 章	既約性	180
V.19.1	非負行列が既約であるための必要十分条件	180
第 V.20 章	漏れ確率行列	181
V.20.1	定義	181
V.20.2	漏れ確率行列のべき乗は $O$ に収束する	181
V.20.3	$o(r)$ の漏れ確率行列 $A$ のノイマン級数は収束し、各要素の和は $\frac{1}{1-r}$ 以下である。	181
第 V.21 章	最適化への応用	182
V.21.1	$\ A\mathbf{x} + \mathbf{b}\ _2^2$ の最小化条件	182
V.21.1.1	制約なしの場合	182
V.21.1.2	線形制約付きの場合	182
第 V.22 章	発想, 技巧	184
V.22.1	演算	184
V.22.1.1	積和の階数を拡張する	184

## 第 VI 部 ベクトル解析

### 第 VI.1 章 3 次元 Euclid 空間

185  
186

VI.1.1	諸定義	186
VI.1.1.1	Vector Laplacian	186
VI.1.2	諸公式	186
VI.1.2.1	ベクトル三重積の公式: $A \times (B \times C) = \langle A, C \rangle B - \langle A, B \rangle C$	186
VI.1.2.2	$\langle \mathbf{a} \times \mathbf{b}, \mathbf{c} \times \mathbf{d} \rangle = \langle \mathbf{a}, \mathbf{c} \rangle \langle \mathbf{b}, \mathbf{d} \rangle - \langle \mathbf{a}, \mathbf{d} \rangle \langle \mathbf{b}, \mathbf{c} \rangle$	187
VI.1.2.3	(系) $\langle \mathbf{a} \times \mathbf{b}, \mathbf{a} \times \mathbf{c} \rangle = \ \mathbf{a}\ ^2 \langle \mathbf{b}, \mathbf{c} \rangle - \langle \mathbf{a}, \mathbf{b} \rangle \langle \mathbf{a}, \mathbf{c} \rangle$	187
VI.1.2.4	発散の別表現: $\nabla \cdot \mathbf{A} = \frac{1}{ V } \lim_{ V  \rightarrow 0} \int_{\partial V} \mathbf{A}(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) d^2 \mathbf{r}$	187
VI.1.2.5	$\nabla_{\mathbf{r}} \cdot f(\mathbf{r}) \mathbf{A}(\mathbf{r}) = (\nabla_{\mathbf{r}} f(\mathbf{r})) \cdot \mathbf{A}(\mathbf{r}) + f(\mathbf{r}) \nabla_{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r})$	188
VI.1.2.6	$\nabla_{\mathbf{r}} \times f(\mathbf{r}) \mathbf{A}(\mathbf{r}) = (\nabla_{\mathbf{r}} f(\mathbf{r})) \times \mathbf{A}(\mathbf{r}) + f(\mathbf{r}) \nabla_{\mathbf{r}} \times \mathbf{A}(\mathbf{r})$	188
VI.1.2.7	$\nabla_{\mathbf{r}} \times (\mathbf{C} \times \mathbf{A}(\mathbf{r})) = (\nabla \cdot \mathbf{A}(\mathbf{r})) \mathbf{C} - J_{\mathbf{A}} \mathbf{C}$	189
VI.1.2.8	3 次元 Euclid 空間に於ける Vector Laplacian: $\nabla^2 \mathbf{A} = \mathbf{i}_1 \Delta A_1 + \mathbf{i}_2 \Delta A_2 + \mathbf{i}_3 \Delta A_3$	189
VI.1.2.9	$\nabla \cdot (\nabla^2 \mathbf{A}) = \Delta(\nabla \cdot \mathbf{A})$	190
VI.1.2.10	$\nabla \cdot (\nabla f, \mathbf{v}) = (\nabla^2 f) \mathbf{v}$	190
VI.1.2.11	$\nabla_{\mathbf{r}} \ \mathbf{r} + \mathbf{a}\ ^n = n \ \mathbf{r} + \mathbf{a}\ ^{n-2} (\mathbf{r} + \mathbf{a})$	191
VI.1.3	Helmholtz の定理	191
VI.1.3.1	補題: 全空間に渡る積分が存在するための十分条件	191
VI.1.3.2	補題: 与えられた発散を有し、回転が $\mathbf{0}$ であるベクトル場の構成	191
VI.1.3.3	補題: 与えられた回転を有し、発散が $\mathbf{0}$ であるベクトル場の構成	192
VI.1.3.4	Helmholtz の定理: 任意の $C^1$ 級のベクトル場は回転が $\mathbf{0}$ である $C^1$ 級の場合と発散が $\mathbf{0}$ である $C^1$ 級の場に分解できる	192
VI.1.4	諸定理	192
VI.1.4.1	球殻定理	192
VI.1.4.2	非負領域の共通部分の非負領域	193
VI.1.4.3	Stokes の定理	195
第 VI.2 章	一般の直交座標系と 3 次元 Euclid 空間の関係	198
VI.2.1	座標変換	198
VI.2.2	表記の慣習上の注意	198
VI.2.3	単位ベクトルの変換	199
VI.2.4	ベクトルの成分の変換	199
VI.2.5	勾配	199
VI.2.6	発散	200
VI.2.7	回転	202
VI.2.8	Laplacian	204
第 VI.3 章	円柱座標系	205
VI.3.1	計量係数	205
VI.3.2	単位ベクトルに関する 3 次元 Euclid 空間との関係	205
VI.3.3	微分演算に関する直交座標系との関係	206
第 VI.4 章	球座標系	207
VI.4.1	計量係数	207
VI.4.2	単位ベクトルに関する 3 次元 Euclid 空間との関係	207
第 VII 部	幾何学	209
第 VII.1 章	Euclid 幾何学	210
VII.1.1	諸公式	210
VII.1.1.1	ヘロンの公式	210
VII.1.2	諸定理	211
VII.1.2.1	$\lim_{\ \mathbf{x}\  \rightarrow \infty} \ \mathbf{x} - \mathbf{a}\  - \ \mathbf{x}\  = -\frac{\mathbf{a}}{\ \mathbf{a}\ } \cdot \mathbf{a}$	211
第 VII.2 章	球面幾何学	212
VII.2.1	余弦定理, 正弦定理について	212
VII.2.2	極三角形の極三角形が元の三角形になること	212
VII.2.3	元の三角形 $ABC$ の極三角形 $A'B'C'$ について $A' = \pi - a$ , $B' = \pi - b$ , $C' = \pi - c$ となること ( $a = BC$ , $b = CA$ , $c = AB$ )	213
第 VIII 部	確率論	214
第 VIII.1 章	表記	215
第 VIII.2 章	独立性に関する諸定理	216
VIII.2.1	$A_1, \dots, A_n$ が独立ならば、そのうちどれか 1 つ $A_l$ を $A_l^c$ に置き換えたものも独立である	216
VIII.2.2	$A_1, \dots, A_n$ が独立ならば $A_1^c, \dots, A_n^c$ も独立である	216
VIII.2.3	$A_1, \dots, A_n$ が独立ならば、そのうちどれか 2 つ $A_p, A_q$ を合併したものも独立である。	217
VIII.2.4	$A_1, \dots, A_n$ が独立ならば、 $I_1 \sqcup I_2 = \{1 : n\}$ なる $I_1, I_2$ に対して $\bigcup_{i \in I_1} A_i$ と $\bigcup_{j \in I_2} A_j$ とは独立である。	217
第 VIII.3 章	エントロピー	218
VIII.3.1	離散型確率分布の場合	218
VIII.3.1.1	一様分布がエントロピーを最大化する	218
VIII.3.2	連続型確率分布の場合	218
VIII.3.2.1	一様分布がエントロピーを最大化する	218
第 IX 部	記述統計	220
第 IX.1 章	諸定義	221
IX.1.1	平均, 分散, 共分散	221
IX.1.2	分散共分散行列	221
第 IX.2 章	諸定理	222



IX.2.1	分散共分散行列は半正定。特に分散が正なら正定。	222
第 IX.3 章	主成分分析	223
IX.3.1	分散共分散行列の対角化	223
IX.3.2	データセットの射影	223
IX.3.3	主成分	224
IX.3.3.1	主成分の分散共分散行列	224
第 IX.4 章	Fisher の線形判別分析	225
IX.4.1	クラス間分散とクラス内分散の比を最大にする $\mathbf{w}$ の導出	225

## 第 X 部 頻度論的統計 227

第 X.1 章	一般の分布	228
X.1.1	$\mathbf{x}$ の分散共分散行列を $V = \text{Var} [\mathbf{x}]$ とすると $\text{Var} [\mathbf{v}^\top \mathbf{x}] = \mathbf{v}^\top V \mathbf{v}$	228
X.1.2	ベクトル確率変数の線形写像の像の期待値と分散共分散行列	228
X.1.3	スコア関数	229
X.1.3.1	$E_{\mathbf{X}} [V(\mathbf{X}, \theta)] = 0$	229
X.1.4	Cramér-Rao の不等式	229
X.1.5	条件付き確率	230
X.1.5.1	$\Pr(A) = \sum_k \Pr(A, B_k) = \sum_k \Pr(A   B_k) \Pr(B_k)$	230
X.1.5.2	$\Pr(A, B   C) = \Pr(A   B, C) \Pr(B   C)$ (スライディング)	230
X.1.5.3	$\Pr(A   C) = \sum_k \Pr(A   B_k, C) \Pr(B_k   C)$	230
X.1.6	条件付き確率密度関数	231
X.1.7	条件付き期待値	231
X.1.7.1	$E_{X,Y} [E_Y [Y X]] = E_{X,Y} [Y]$	231
X.1.8	確率不等式	232
X.1.8.1	マルコフの不等式	232
X.1.8.2	チェビシェフの不等式	232
X.1.8.3	ヘフディングの補題	233
X.1.8.4	ヘフディングの不等式	234
X.1.9	中心極限定理	235
第 X.2 章	多項分布	236
X.2.1	周辺分布は 2 項分布	236
X.2.2	共分散	237
第 X.3 章	幾何分布	238
X.3.1	無記憶性からの導出	238
第 X.4 章	指数分布	239
X.4.1	定義	239
X.4.2	解釈	239
X.4.3	無記憶性からの導出	239
X.4.4	特性関数	240
第 X.5 章	Erlang 分布	241
X.5.1	定義	241
X.5.2	解釈	241
X.5.3	指数分布との関係: $X_1, \dots, X_k (\text{独立}) \sim \text{ExpDist}(\mu) \Rightarrow X_1 + \dots + X_k \sim \text{ErlangDist}(k, \mu)$	241
第 X.6 章	Poisson 分布	243
X.6.1	定義	243
X.6.2	解釈	243
X.6.2.1	指数分布からの導出	243
X.6.3	Erlang 分布との関係	243
X.6.4	再生性	244
第 X.7 章	beta 分布	245
X.7.1	Beta 分布に従う確率変数の生成	245
X.7.1.1	$X \sim \text{Gamma}(\alpha, 1), Y \sim \text{Gamma}(\beta, 1) \Rightarrow \frac{X}{X+Y} \sim \text{Beta}(\alpha, \beta)$	245
第 X.8 章	正規分布	246
X.8.1	裾確率の評価	246
X.8.2	スケール変換とシフト	247
X.8.3	再生性	248
X.8.4	一次結合	249
X.8.5	分散 $\sigma^2$ , 平均 $\mu$ のときの累積分布関数は $\Phi(\frac{z-\mu}{\sigma})$	249
X.8.6	零平均, 同分散の正規分布に独立に従う $n$ 個の確率変数の直交変換もまた同じ分布に従う	249
X.8.7	多変量正規分布	250
X.8.7.1	定義	250
X.8.7.2	密度関数の全空間積分が 1 になることの確認	250
X.8.7.3	期待値の導出	251
X.8.7.4	共分散行列の導出	251
X.8.7.5	特性関数は $\phi_{\mathbf{X}}(\mathbf{t}) = \exp(i\mathbf{t}^\top \boldsymbol{\mu} - \mathbf{t}^\top \Sigma \mathbf{t}/2)$	251
X.8.7.6	特性関数を用いた期待値と分散の導出	252
X.8.7.7	線形写像: $A\mathbf{X} + \mathbf{b} \sim N(A\boldsymbol{\mu} + \mathbf{b}, A\Sigma A^\top)$	253
X.8.7.8	周辺分布: $X_i \sim N(\mu_i, \sigma_i^2)$	253
第 X.9 章	Rayleigh 分布	254
X.9.1	定義	254
X.9.2	正規分布する成分を持つ 2 次元ベクトルのノルムとの関係	254
第 X.10 章	Rice 分布	255
X.10.1	定義	255
X.10.2	正規分布する成分を持つ 2 次元ベクトルと定数ベクトルとの関係	255

X.10.2.1	記号の準備	255
X.10.2.2	ノルムとの関係	255
X.10.2.3	偏角との関係	256
第 X.11 章	対数正規分布	258
X.11.1	確率密度関数の導出	258
X.11.2	最頻値	258
X.11.3	期待値	259
第 X.12 章	$\chi^2$ 分布	260
X.12.1	定義	260
X.12.2	$X \sim N(0, 1) \Rightarrow X^2 \sim \chi_1^2$	260
X.12.3	$X_i \sim N(0, 1) \Rightarrow Z_n = \sum_{i=1}^n X_i^2 \sim \chi_n^2$	260
X.12.4	$X \sim \chi_n^2 \Rightarrow E[X] = n, V[x] = 2n$	261
X.12.5	再生性	262
X.12.6	$X_1, \dots, X_n \sim N(\mu, \sigma^2) \Rightarrow \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi_{n-1}^2$	262
第 X.13 章	$t$ 分布	264
X.13.1	定義	264
X.13.2	$X.13.2.0 X, Y : \text{indep}, X \sim N(0, 1), Y \sim \chi_n^2 \Rightarrow T := \frac{X}{\sqrt{\frac{Y}{n}}} \sim t_n$	264
X.13.3	$X.13.3.0 t$ 分布の期待値	265
X.13.4	$X.13.4.0 t$ 分布の分散	265
第 X.14 章	$F$ 分布	267
X.14.1	定義	267
X.14.2	$X_1, X_2 : \text{indep}, X_1 \sim \chi_{\nu_1}^2, X_2 \sim \chi_{\nu_2}^2 \Rightarrow F := \frac{\frac{X_1}{\nu_1}}{\frac{X_2}{\nu_2}} \sim F_{\nu_1}^{\nu_2}$	267
X.14.3	自由度 $(\nu_1, \nu_2)$ の $F$ 分布 $F_{\nu_2}^{\nu_1}$ は $\nu_2 \geq 3$ の時に限り期待値が定義できてその値は $\frac{\nu_2}{\nu_2 - 2}$ である。	268
X.14.4	自由度 $(\nu_1, \nu_2)$ の $F$ 分布 $F_{\nu_2}^{\nu_1}$ の分散は $\nu_2 \leq 4$ の場合 $\infty$ 、 $\nu_2 \geq 5$ の場合 $\frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}$ である。	269
第 X.15 章	推定	270
X.15.1	標本分散	270
X.15.1.1	$E \left[ \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \right] = \frac{n-1}{n} \sigma^2$	270

第 XI 部	ベイズ統計	272
第 XI.1 章	事前分布と事後分布の考え方	273
XI.1.1	離散型	273
XI.1.2	連続型	273
第 XI.2 章	ベイズ推定量	275
第 XI.3 章	事前分布と事後分布の例	276
XI.3.1	正規分布の事後分布	276

第 XII 部	グラフ理論	277
第 XII.1 章	定義	278
第 XII.2 章	連結グラフ	279
XII.2.1	諸定理	279
XII.2.1.1	$n$ 頂点の連結グラフは少なくとも $n - 1$ 本の辺をもつ	279
XII.2.1.2	任意の 2 つの頂点間の最短距離が 2 である無向グラフは少なくとも $n - 1$ 本の辺をもつ	279
XII.2.1.3	$n$ 頂点グラフが $\frac{1}{2}(n - 1)(n - 2)$ 本より多くの辺を持つなら連結である	279
XII.2.1.4	Ore の定理	280
第 XII.3 章	木	282
XII.3.1	諸定理	282
XII.3.1.1	木は 2 部グラフである	282
XII.3.1.2	1 本以上の辺を持つ有限な木は次数 1 の枝を少なくとも 2 個持つ	282
XII.3.1.3	どこへでも行ける点が 1 個あればその点を根として木を作れる	282
XII.3.1.4	$\mathcal{I}$ : グラフ $G$ の非閉路部分グラフ全体の集合。 $I_1, I_2 \in \mathcal{I},  I_1  <  I_2  \Rightarrow \exists a \in I_2 \text{ s.t. } I_1 \cup \{a\} \in \mathcal{I}$	283
XII.3.1.5	$\mathcal{T}$ : グラフ $G$ の全域木の集合。 $T_1, T_2 \in \mathcal{T}, T_1 \neq T_2, \forall a_1 \in T_1 \setminus T_2, \exists a_2 \in T_2 \text{ s.t. } (T_1 \setminus a_1) \cup a_2 \in \mathcal{T}$	283
第 XII.4 章	有向グラフ	284
XII.4.1	諸定理	284
XII.4.1.1	$G_1 = (\nu, \varepsilon_1), G_2 = (\nu, \varepsilon_2)$ に対して $G_1^{ \nu-1 } = G_2^{ \nu-1 }$ であるとき、 $G_1$ が全域木を持つ $\iff G_2$ が''	284
XII.4.1.2	非閉路的ならば pruning のアルゴリズムはデッドロックしない	284
XII.4.1.3	スタートからゴールへの全ての道の連鎖重みの総和は 1 以下	285
XII.4.1.4	(補題) ゴールに隣接する 1 つのスタート地点からの全ての経路の連鎖重みの総和は 1 以下	287

第 XIII 部	その他	289
第 XIII.1 章	無限級数公式集	290
XIII.1.1	$\sum_{k=1}^{\infty} \frac{\sin k}{k} = \frac{\pi-1}{2}$	290
XIII.1.2	$\sum_{n=1}^{\infty} \frac{1}{1+n^2} = \frac{-1}{2} + \frac{\pi}{2} \coth \pi$	290
XIII.1.3	$\sum_{n=1}^{\infty} \frac{1}{(1+n^2)^2} = \frac{-1}{2} + \frac{\pi^2}{4 \sinh^2 \pi} + \frac{\pi}{4} \coth \pi$	291
参考文献		292

第 I 部

离散数学

## 第 I.1 章

# 表記

- $\{m, \dots, n\}$  :  $m$  以上  $n$  以下の整数の集合
- ${}_A C_k$  ( $A$ は集合) :  $A$  の  $k$  元部分集合全体から成る集合族を表す。  
例えば  $A = \{1, 2, 3\}$  に対して  ${}_A C_2 = \{\{1, 2\}, \{2, 3\}, \{3, 1\}\}$

## 第 1.2 章

# 算術

### 1.2.1 諸定理

1.2.1.1  $a, b, c, d \in \mathbb{N}$ ,  $ab = cd$ ,  $\gcd(a, c) = \gcd(b, d) = 1 \Rightarrow a = d, b = c$

*Proof.*

(最初に思い付いた証明)

$p$  を、 $a$  の素因数分解に指数  $l \geq 1$  で現れる任意の素数とする。 $a$  と  $c$  は互いに素だから  $c$  の素因数分解に於ける  $p$  の指数は 0 である。さらに  $p \nmid b$  である。なぜならば、 $b$  と  $d$  が互いに素なので、もし  $p \mid b$  とすると  $d$  の素因数分解に於ける  $p$  の指数は 0 であり、 $p \nmid cd$  となり  $ab = cd$  に矛盾するからである。よって左辺の素因数分解に於ける  $p$  の指数は  $l$  であり、前述の通り  $a$  と  $c$  は互いに素だから  $d$  の素因数分解に於ける  $p$  の指数が  $l$  でなければならない。

$a$  と  $d$  の立場を変えて同様に考えると  $d$  の素因数分解に指数  $m \geq 1$  で現れる任意の素数  $q$  は  $a$  の素因数分解に於いて同じ指数  $m$  で現れることがわかる。よって  $a = d$  である。これより  $b = c$  が得られる。

(簡潔な証明)

$ab = cd$  より  $a \mid cd$ 。  $a$  と  $c$  は互いに素だから  $a \mid d$  でなくてはならない。同様に、 $ab = cd$  より  $d \mid ab$ 。  $d$  と  $b$  は互いに素だから  $d \mid a$  でなくてはならない。以上より  $a = d$ 。これより  $b = c$  が得られる。  $\square$

### 1.2.2 Extended Euclidean Algorithm

Extended Euclidean Algorithm (EEA, 拡張されたユークリッドの互除法) とは、 $a, b \in \mathbb{Z}$  に対して  $as + bt = \gcd(a, b)$  となるような  $s, t \in \mathbb{Z}$  を求める手法である。この手法では次の規則で非負整数列  $\{r_i\}$ , および整数列  $\{q_i\}, \{s_i\}, \{t_i\}$  を定める。

$$r_0 = a, s_0 = 1, t_0 = 0, \quad r_1 = b, s_1 = 0, t_1 = 1$$

$$r_{i+1} = r_{i-1} - q_i r_i, \quad 0 \leq r_{i+1} < |r_i|$$

$$s_{i+1} = s_{i-1} - q_i s_i$$

$$t_{i+1} = t_{i-1} - q_i t_i$$

上の式は、 $i \geq 1$  については  $q_i, r_{i+1}$  はそれぞれ  $r_{i-1}$  を  $r_i$  で割った商と余りであることを言っている。 $r_2, r_3, \dots$  は真に単調減少するから、ある  $k \in \mathbb{N}$  において  $r_{k+1} = 0$  となる。この時点で数列の生成を停止する。 $a$  や  $b$  が負数の場合は  $q_i < 0$  となることもあるが、それは高々  $q_2$  までで、 $q_3$  以降は 1 以上になる。なぜならば  $r_2$  以降が非負だからである。上記の数列について次が成り立つ。

1.  $\gcd(a, b) = r_k = as_k + bt_k$
2.  $s_i, t_i$  ( $i = 0, \dots, k+1$ ) および  $s_i, s_{i+1}$  および  $t_i, t_{i+1}$  ( $i = 0, \dots, k$ ) は互いに素である
3.  $|s_{k+1}| = |b|/\gcd(a, b)$ ,  $|t_{k+1}| = |a|/\gcd(a, b)$
4.  $a, b > 0$ ,  $\gcd(a, b) < \min(a, b)$  ならば  $|s_i| < b/\gcd(a, b)$ ,  $|t_i| < a/\gcd(a, b)$  ( $i = 0, \dots, k$ )

項目 4 は、計算機で EEA を実行する際に、 $a, b$  を表現できるだけの桁数がある数値型を使えば桁溢れしないことを保証する。

*Proof.*

[https://en.wikipedia.org/wiki/Extended\\_Euclidean\\_algorithm](https://en.wikipedia.org/wiki/Extended_Euclidean_algorithm) を基に若干改変, 加筆した。

(項目 1)

普通の Euclid の互除法を既に知っているものとする。 $r_k = \gcd(a, b)$  は明らか。 $r_0 = a = as_0 + bt_0$ ,  $r_1 = b = as_1 + bt_1$  であり、 $r_2 = r_0 - q_1 r_1 = as_0 + bt_0 - q_1(as_1 + bt_1) = a(s_0 - q_1 s_1) + b(t_0 - q_1 t_1) = as_2 + bt_2$  が成り立つ。これを続けると  $r_i = as_i + bt_i$  ( $i = 0, \dots, k+1$ ) が得られる。

項目 1 の証明はこれで済んだが、 $\{s_i\}, \{t_i\}$  の漸化式がどうやって考案されたのか想像してみる。等式  $r_i = as_i + bt_i$  が  $i = 0, 1, \dots$  について成り立って欲しいのだから、 $as_{i-1} + bt_{i-1} = r_{i-1} = r_{i+1} + q_i r_i = as_{i+1} + bt_{i+1} + q_i(as_i + bt_i) = a(s_{i+1} + q_i s_i) + b(t_{i+1} + q_i t_i)$  とおいて最左辺と最右辺の  $a, b$  の係数を比較すれば漸化式を得られる。

(項目 2)

$$A_i := \begin{bmatrix} s_i & s_{i+1} \\ t_i & t_{i+1} \end{bmatrix}, \quad Q_i := \begin{bmatrix} 0 & 1 \\ 1 & -q_i \end{bmatrix}$$

とすると

$$A_i = A_0 \prod_{j=1}^i Q_j$$

となる。両辺の行列式を考えると

$$s_i t_{i+1} - s_{i+1} t_i = |A_i| = |A_0| \prod_{j=1}^i |Q_j| = (-1)^i$$

これと Bézout の等式より定理の主張が従う。

(項目 3)

$0 = r_{k+1} = as_{k+1} + bt_{k+1}$  より  $|a||s_{k+1}| = |b||t_{k+1}|$ 。両辺を  $\gcd(a, b)$  で割って次式を得る。

$$\frac{|a|}{\gcd(a, b)} |s_{k+1}| = \frac{|b|}{\gcd(a, b)} |t_{k+1}|$$

$|a|/\gcd(a, b)$  と  $|b|/\gcd(a, b)$  は互いに素であり、前述の通り  $|s_{k+1}|$  と  $|t_{k+1}|$  も互いに素だから I.2.1.1 より

$$|s_{k+1}| = \frac{|b|}{\gcd(a, b)}, \quad |t_{k+1}| = \frac{|a|}{\gcd(a, b)}$$

特に  $0 = as_{k+1} + bt_{k+1}$  を満たすのは次の組み合わせである。

$$(s_{k+1}, t_{k+1}) = \pm \left( \frac{b}{\gcd(a, b)}, -\frac{a}{\gcd(a, b)} \right)$$

(項目 4)

まず  $s_0, s_2, s_4, \dots > 0$ ,  $s_3, s_5, s_7, \dots < 0$  を示す。 $s_0$  は定義より正である。さらに  $s_2 = s_0 - q_1 s_1 = 1$ ,  $s_3 = s_1 - q_2 s_2 = -q_2$  を直接確かめられる。これより始めて帰納的に、 $l \geq 2$  に対して  $s_{2l} = s_{2l-2} - q_{2l-1} s_{2l-1} > 0$ ,  $s_{2l-1} = s_{2l-3} - q_{2l-2} s_{2l-2} < 0$  が成り立つ。

次に数列  $\{|s_i|\}$  ( $i = 1, 2, \dots$ ) が単調増加であること、特に  $|s_2| \leq |s_3|$  を除いて他は全て狭義単調増加であることを示す。まず  $0 = |s_1| < 1 = |s_2|$ ,  $|s_3| = |-q_2| \geq 1 = |s_2|$ ,  $|s_4| = |s_2 - q_3 s_3| = |1 + q_2 q_3| > |q_2| = |s_3|$  を直接確かめられる。これより始めて帰納的に、 $l \geq 2$  に対して  $|s_{2l+1}| - |s_{2l}| = -s_{2l+1} - s_{2l} = (q_{2l} - 1)s_{2l} - s_{2l-1} > 0$ ,  $|s_{2l+2}| - |s_{2l+1}| = (1 - q_{2l+1})s_{2l+1} + s_{2l} > 0$  が成り立つから  $|s_{2l+1}| > |s_{2l}|$  ( $l = 2, 3, \dots$ ) である。同様にして  $|t_0| < |t_1| \leq |t_2| < |t_3| < \dots$  が成り立つ。これと項目 3 より定理の主張が成り立つ。  $\square$

## 第 I.3 章

# 整数の合同

### I.3.1 定数倍と加減乗除

$n$  を 2 以上の整数とする。 $x_1 \equiv_n x_2$  かつ  $y_1 \equiv_n y_2$  であるとき次が成り立つ。

- 任意の整数  $a$  に対して  $ax_1 \equiv_n ax_2$
- $x_1 + y_1 \equiv_n x_2 + y_2$
- $x_1 y_1 \equiv_n x_2 y_2$
- $x_1, x_2$  が共に整数  $m$  で割り切れて  $m$  と  $n$  が互いに素ならば  $x_1/m \equiv_n x_2/m$

### I.3.2 $a, m \in \mathbb{N}$ が互いに素であるとき、 $0, a, 2a, \dots, (m-1)a$ を $m$ で割った余りは全て異なる

*Proof.*

$b_1, b_2$  ( $0 \leq b_1 < b_2 \leq m-1$ ) を任意にとる。 $a$  と  $m$  が互いに素であるから  $(m/\gcd(b_2 - b_1, m)) \nmid (a(b_2 - b_1)/\gcd(b_2 - b_1, m))$  であるので  $m \nmid a(b_2 - b_1)$ 、すなわち  $ab_1 \not\equiv_m ab_2$  である。  $\square$

### I.3.3 和のべき乗を展開

$x, y$  を整数、 $p$  を素数とすると次が成り立つ。

1.  $(x + y)^p \equiv_p x^p + y^p$
2.  $(x + y)^p \equiv_p x^p + y$

*Proof.*

まず 1 を示す。

$$(x + y)^p \equiv_p x^p + \sum_{k=1}^{p-1} {}_p C_k x^{p-k} y^k + y^p$$

であるが  $\sum_{k=1}^{p-1} {}_p C_k x^{p-k} y^k$  は  $p$  の倍数である。何故ならば

$${}_p C_k = \frac{p!}{k!(p-k)!} \quad (k = 1, \dots, p-1)$$



において  $p$  は素数であり  $k < p$  なので分子の  $p$  は約分されないから。よって次式より 1 が成り立つ。

$$x^p + \sum_{k=1}^{p-1} {}_p C_k x^{p-k} y^k + y^p \equiv_p x^p + y^p$$

次に 2 を示す。  $p = 2$  のときは、左辺と右辺の差は  $(x+y)^2 - (x^2+y) = y(y-1+2x)$  なので  $y$  の偶奇に関わらず  $p$  の倍数である。以下では  $p$  は 3 以上の素数とする。  $y$  が 0 のときは明らかに成り立つ。1 の結果を用いると、  $y$  が自然数のときは

$$(x+y)^p = [(x+y-1)+1]^p \equiv_p (x+y-1)^p + 1 \equiv_p (x+y-2)^p + 2 \equiv_p \cdots \equiv_p x^p + y$$

より成り立つ。  $y$  が負の整数のときも

$$(x+y)^p = [(x+y+1)-1]^p \equiv_p (x+y+1)^p - 1 \equiv_p (x+y+2)^p - 2 \equiv_p \cdots \equiv_p x^p + (-y)(-1) = x^p + y$$

より成り立つ。 □

### 1.3.4 Fermat の小定理

素数  $p$  と整数  $a$  が互いに素であるとき、  $a^{p-1} \equiv_p 1$  である。

*Proof.*

「**任意**の整数  $a$  に対して  $a^p \equiv_p a$ 」を示し、  $a$  と  $p$  が互いに素である場合に限定して両辺を  $a$  で割れば良いので、これを示す。  $p = 2$  のときは  $a$  の偶奇で場合分けすれば簡単に示せる。以下では  $p$  は 3 以上の素数であるとする。まず  $a = 0$  のときは明らかに成り立つ。1.3.3 を用いると、  $a$  が自然数のときは

$$a^p \equiv_p [(a-1)+1]^p \equiv_p (a-1)^p + 1 \equiv_p [(a-2)+1]^p + 1 \equiv_p (a-2)^p + 1 + 1 \equiv_p \cdots \equiv_p a$$

となり、成り立つ。  $a$  が負の整数のときは

$$a^p \equiv_p [(a+1)-1]^p \equiv_p (a+1)^p - 1 \equiv_p [(a+2)-1]^p - 1 \equiv_p (a+2)^p - 1 - 1 \equiv_p \cdots \equiv_p (-a)(-1) \equiv_p a$$

となり、やはり成り立つ。 □

## 第 1.4 章

### $r$ 進法

#### 1.4.1 $1/10$ が 2 進表記で無限小数になることの証明

*Proof.*

$\frac{1}{10} = \frac{1}{2} \times \frac{1}{5}$  なので  $1/5$  が 2 進表記で無限小数になることを示せば十分。背理法で示す。ある適当な自然数  $m$  を用いて  $\frac{1}{5} = \sum_{i=1}^m a_i 2^{-i}$  ( $a_i \in \{0, 1\}$ ) と表せると仮定する。両辺に  $2^m \times 5$  を掛けて  $2^m = 5 \sum_{i=1}^m a_i 2^{m-i} = 5 \sum_{i=0}^{m-1} a_{m-i} 2^i$  となる。左辺は 5 で割り切れないが右辺は割り切れるので矛盾している。  $\square$

## 第 1.5 章

# 公式

### 1.5.1 $\sum_{k=1}^n k^2 = n(n+1)(2n+1)/6$

*Proof.*

$S_n := \sum_{k=1}^n k^2$  とする。 $S_n$  に何らかの項を付け加え、 $T_n$  とし、 $T_{n+1} - T_n$  が  $n$  の高々 1 次式となるようにできれば、 $T_n$  を求めることができ、 $S_n$  も求まる。 $S_{n+1} - S_n = (n+1)^2$  であり、 $n^2$  の項を消去するには  $S_n$  に  $n^3$  に比例する項を付け加えればよい ( $n$  の 2 次式では不足である。なぜなら  $T_{n+1} - T_n$  の段階で  $n^2$  に比例する項が相殺して消えるから  $S_{n+1} - S_n$  由来の  $n^2$  の項を決してキャンセルできないからである)。その比例係数を  $\alpha$  とし、 $T_n := S_n + \alpha n^3$  とすると次式が成り立つ。

$$T_{n+1} - T_n = (n+1)^2 - \alpha((n+1)^3 - n^3) = (n+1)^2 + \alpha(3n^2 + 3n + 1)$$

そこで  $\alpha = -1/3$  とすると次式を得る。

$$T_{n+1} - T_n = 2/3 + n$$

よって  $n \geq 2$  のとき

$$T_n = T_1 + \sum_{k=1}^{n-1} (T_{k+1} - T_k) = 2/3 + (n-1)(n/2 + 2/3)$$

上式は  $n = 1$  のとき  $2/3$  となるが、これは  $T_1$  と等しいので、上式は  $n = 1$  のときも含めて成り立つ。以上より

$$S_n = T_n + n^3/3 = \frac{1}{6}n(n+1)(2n+1)$$

□

### 1.5.2 $\sum_{k=0}^n (2k+1)^2 = (n+1)(2n+1)(2n+3)/3$

*Proof.*

$$\begin{aligned} \sum_{k=0}^n (2k+1)^2 &= \sum_{k=0}^n (4k^2 + 4k + 1) = 4 \sum_{k=1}^n k^2 + 4 \sum_{k=1}^n k + n + 1 \\ &= \frac{2}{3}n(n+1)(2n+1) + 2n(n+1) + (n+1) \quad (\because \text{I.5.1}) \\ &= \frac{1}{3}(n+1)[2n(2n+1) + 6n + 3] = \frac{1}{3}(n+1)(2n+1)(2n+3) \end{aligned}$$

□

$$1.5.3 \quad 1 \cdot 3 \cdot 5 \cdots (2n-1)2n = (2n)!/n!$$

*Proof.*

$$1 \cdot 3 \cdot 5 \cdots (2n-1)2n = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdots (2n-2)(2n-1)2n}{\underbrace{2 \cdot 4 \cdot 6 \cdots (2n-2)2n}_{n \text{ 個}}} 2^n = \frac{(2n)!}{1 \cdot 2 \cdot 3 \cdots (n-1)n} = \frac{(2n)!}{n!}$$

□

$$1.5.4 \quad a^n - b^n = (a - b) \sum_{k=0}^{n-1} a^{n-1-k} b^k$$

*Proof.*

$$\begin{aligned} (a - b) \sum_{k=0}^{n-1} a^{n-1-k} b^k &= \sum_{k=0}^{n-1} a^{n-k} b^k - \sum_{k=0}^{n-1} a^{n-1-k} b^{k+1} \\ &= a^n + \sum_{k=1}^{n-1} a^{n-k} b^k - \sum_{k=1}^n a^{n-k} b^k = a^n - b^n \end{aligned}$$

□

## 1.5.5 Möbius の反転公式

$f, g$  を自然数から複素数への写像とする。

1. 任意の自然数  $n$  に対して  $g(n) = \sum_{d|n} f(d)$  が成り立つならば  $f(n) = \sum_{d|n} \mu(d)g(n/d) = \sum_{d|n} \mu(n/d)g(n)$  が成り立つ。
2. 任意の自然数  $n$  に対して  $g(n) = \prod_{d|n} f(d)$  が成り立つならば  $f(n) = \prod_{d|n} g(n/d)^{\mu(d)} = \prod_{d|n} g(d)^{\mu(n/d)}$  が成り立つ。

*Proof.*

(1 について)

結論の式の左から 2 つ目の等号は約数の意味から明らかである。1 つ目の等号成立を示す。

$$\begin{aligned} \sum_{d|n} \mu(d)g(n/d) &= \sum_{d|n} \mu(d) \sum_{e|n/d} f(e) = \sum_{d|n} \sum_{e|n/d} \mu(d)f(e) = \sum_{\substack{d,e \\ de|n}} \mu(d)f(e) \\ &= \sum_{e|n} \sum_{d|n/e} \mu(d)f(e) = \sum_{e|n} f(e) \sum_{d|n/e} \mu(d) = \sum_{e|n} f(e) \mathbb{1}\{e=n\} = f(n) \end{aligned}$$

(2 について)

$$\begin{aligned}
\prod_{d|n} g(n/d)^{\mu(d)} &= \prod_{d|n} \left( \prod_{e|n/d} f(e) \right)^{\mu(d)} = \prod_{d|n} \prod_{e|n/d} f(e)^{\mu(d)} = \prod_{\substack{d,e \\ de|n}} f(e)^{\mu(d)} \\
&= \prod_{e|n} \prod_{d|n/e} f(e)^{\mu(d)} = \prod_{e|n} f(e)^{\sum_{d|n/e} \mu(d)} = \prod_{e|n} f(e)^{\mathbb{1}_{\{e=n\}}} = f(n)
\end{aligned}$$

□

## 第 I.6 章

# 数え上げ

I.6.1 “for  $i = 1, \dots, n (\in \mathbb{N})$ ”  $\iff$  “for  
 $d \mid n$ , for  $i = 1, \dots, n$  where  $\gcd(i, n) = d$ ”  $\iff$  “for  
 $d \mid n$ , for  $j = 1, \dots, n/d$  where  $\gcd(j, n/d) = 1$ ”

1 から  $n$  までの自然数に対して、それと  $n$  との最大公約数は  $n$  の約数のどれかと一致するから、最初の  $\iff$  が成り立つ。2 つ目の  $\iff$  については  $i = dj$  なる変数変換を施すことで分かる。次のような使い方が考えられる。

$$\begin{aligned} \bullet \sum_{i=1}^n f(i) &= \sum_{d \mid n} \sum_{\substack{i=1, \dots, n \\ \gcd(i, n)=d}} f(i) = \sum_{d \mid n} \sum_{\substack{j=1, \dots, n/d \\ \gcd(j, n/d)=1}} f(dj) \\ \bullet \prod_{i=1}^n f(i) &= \prod_{d \mid n} \prod_{\substack{i=1, \dots, n \\ \gcd(i, n)=d}} f(i) = \prod_{d \mid n} \prod_{\substack{j=1, \dots, n/d \\ \gcd(j, n/d)=1}} f(dj) \end{aligned}$$

## I.6.2 包除原理

I.6.2.1 補題:  $A \subseteq U \Rightarrow |A| = \sum_{x \in U} |A \cap \{x\}|$

解説.

当たり前すぎて証明に困る。 $U$  の各要素について、それが  $A$  に含まれるかどうかを風潰しに調べているだけ。

任意の集合  $A_1, \dots, A_n$  に対して次の等式が成り立つ。

$$\begin{aligned} |A_1 \cup \dots \cup A_n| &= |A_1| + \dots + |A_n| \\ &\quad - |A_1 \cap A_2| - |A_1 \cap A_3| - \dots - |A_{n-1} \cap A_n| \\ &\quad + |A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| + \dots + |A_{n-2} \cap A_{n-1} \cap A_n| \\ &\quad - \dots + (-1)^{n-1} |A_1 \cap \dots \cap A_n| \end{aligned}$$

すなわち

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{k=1}^n (-1)^{k-1} \sum_{I \in \{1, \dots, n\} C_k} \left| \bigcap_{i \in I} A_i \right| \quad (1)$$

もっと短く書けば

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{I \in 2^{\{1, \dots, n\}} \setminus \{\emptyset\}} (-1)^{|I|-1} \left| \bigcap_{i \in I} A_i \right|$$

*Proof.*

式 (1) を証明する。  $U := \bigcup_{i=1}^n A_i$  とする。補題 1.6.2.1 より

$$\begin{aligned} (1) \text{ の左辺} &= \sum_{x \in U} \left| \left( \bigcup_{i=1}^n A_i \right) \cap \{x\} \right| = \sum_{x \in U} 1 \\ (1) \text{ の右辺} &= \sum_{k=1}^n (-1)^{k-1} \sum_{I \in \{1, \dots, n\} C_k} \sum_{x \in U} \left| \left( \bigcap_{i \in I} A_i \right) \cap \{x\} \right| \\ &= \sum_{k=1}^n (-1)^{k-1} \sum_{x \in U} \sum_{I \in \{1, \dots, n\} C_k} \left| \left( \bigcap_{i \in I} A_i \right) \cap \{x\} \right| \quad (\text{和の順序交換}) \\ &= \sum_{x \in U} \sum_{k=1}^n (-1)^{k-1} \sum_{I \in \{1, \dots, n\} C_k} \left| \left( \bigcap_{i \in I} A_i \right) \cap \{x\} \right| \quad (\text{和の順序交換}) \\ &= \sum_{x \in U} c(x) \quad \text{where} \quad c(x) := \sum_{k=1}^n (-1)^{k-1} \sum_{I \in \{1, \dots, n\} C_k} \left| \left( \bigcap_{i \in I} A_i \right) \cap \{x\} \right| \end{aligned}$$

任意の  $x \in U$  に対して  $c(x) = 1$  であることを示す。各  $x$  に対して、 $x$  に依存して決まる次のような自然数  $m(x)$  が存在する。

「 $x$  は  $m(x)$  個の集合  $A_{i_1}, \dots, A_{i_{m(x)}}$  に共通して属し、他の  $n - m(x)$  個の集合には属さない。」

従って、 $c(x)$  の  $\sum_{I \in \{1, \dots, n\} C_k} \left| \left( \bigcap_{i \in I} A_i \right) \cap \{x\} \right|$  は  $k > m(x) + 1$  に対して 0 であるから  $c(x)$  は次のように書き直される。

$$c(x) = \sum_{k=1}^{m(x)} (-1)^{k-1} \sum_{I \in \{1, \dots, n\} C_k} \left| \left( \bigcap_{i \in I} A_i \right) \cap \{x\} \right|$$

上式の  $\left| \left( \bigcap_{i \in I} A_i \right) \cap \{x\} \right|$  が 1 になるのは上述の  $m(x)$  個の集合から  $k$  個の集合を選んできた時だけであるか

ら、 $c(x)$  はさらに次のように書き直される。

$$\begin{aligned}
 c(x) &= \sum_{k=1}^{m(x)} (-1)^{k-1} \sum_{I \in \{i_1, \dots, i_{m(x)}\}^{C_k}} 1 = \sum_{k=1}^{m(x)} (-1)^{k-1} m(x) C_k \\
 &= - \sum_{k=1}^{m(x)} (-1)^k m(x) C_k = - \left( \sum_{k=0}^{m(x)} (-1)^k m(x) C_k - 1 \right) \\
 &= - \left( (1-1)^{m(x)} - 1 \right) = 1
 \end{aligned}$$

□

### 1.6.2.2 あの等式をどうやって思いついたか

包除原理の等式を証明することはできたが、一番最初にあの等式を思い付いた人は何を考えたのだろうか。ここでは、あの等式を導き出す自然なアルゴリズムを考えてみる。

まず我々は「メモリ」を **1つ** 用意して 0 で初期化する。次に、 $U := \bigcup_{i=1}^n A_i$  とし、 $U$  の **各要素** に「カウンタ」を取り付けて 0 で初期化する。

まず次の操作を実行する。「任意の  $A_i$  について、 $A_i$  に属する各要素のカウンタを +1 する。カウンタを操作する度にメモリも +1 する。この操作を  $i = 1, \dots, n$  すべてに対して行う。」この操作の実行後、メモリの値は  $|A_1| + \dots + |A_n|$  になっている。そしてカウンタの値は次のようになっている。

- 1つの集合に属しており、2つ以上の集合に属していない要素のカウンタは 1
- 2つの集合に属しており、3つ以上の集合に属していない要素のカウンタは 2
- ...
- $n$  個の集合に属している要素のカウンタは  $n$

当然だ。複数の集合が被さっている部分の要素はオーバーカウントされている。

そこで、「2つの集合に属しており、3つ以上の集合に属していない要素」のオーバーカウントを解消するために次の操作を行う。「任意の2つの集合  $A_i, A_j$  について、 $A_i \cap A_j$  に属する要素のカウンタを -1 する。カウンタを操作する度にメモリも -1 する。この操作を全ての  $\{i, j\} \in \{1, \dots, n\} C_2$  に対して行う。」この操作の実行後、メモリの値は  $|A_1| + \dots + |A_n| - |A_1 \cap A_2| - |A_1 \cap A_3| - \dots - |A_{n-1} \cap A_n|$  になっている。そしてカウンタの値は次のようになっている。

- 1つの集合に属しており、2つ以上の集合に属していない要素のカウンタは 1
- 2つの集合に属しており、3つ以上の集合に属していない要素のカウンタは  $2 - {}_2C_2 = 1$
- 3つの集合に属しており、4つ以上の集合に属していない要素のカウンタは  $3 - {}_3C_2 = \mathbf{0}$
- 4つの集合に属しており、4つ以上の集合に属していない要素のカウンタは  $4 - {}_4C_2 = -2$
- 5つの集合に属しており、4つ以上の集合に属していない要素のカウンタは  $5 - {}_5C_2 = -5$
- ...
- $n$  個の集合に属している要素のカウンタは  $n - {}_nC_2 < 0$

(※  ${}_nC_2$  が現れる理由は次のとおり。例えば「3つの集合に属しており、4つ以上の集合に属していない要素」は、上述の操作により、自分が属している3つの集合から2つを選んで作られた  ${}_3C_2$  通りの集合のパターン



の数だけカウンタが-1 される。)

やりすぎたようだ。3 つ以上の集合に属している要素をアンダーカウントしている。

そこで、「3 つの集合に属しており、4 つ以上の集合に属していない要素」を数え直すために次の操作を実行する。「任意の 3 つの集合  $A_i, A_j, A_k$  について、 $A_i \cap A_j \cap A_k$  に属する要素のカウンタを +1 する。カウンタを操作する度にメモリも +1 する。この操作を全ての  $\{i, j, k\} \in \{1, \dots, n\} C_3$  に対して行う。」この操作の実行後、メモリの値は  $|A_1| + \dots + |A_n| - |A_1 \cap A_2| - |A_1 \cap A_3| - \dots - |A_{n-1} \cap A_n| + |A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| + \dots + |A_{n-2} \cap A_{n-1} \cap A_n|$  になっている。そしてカウンタの値は次のようになっている。

- 1 つの集合に属しており、2 つ以上の集合に属していない要素のカウンタは 1
- 2 つの集合に属しており、3 つ以上の集合に属していない要素のカウンタは  $2 - {}_2C_2 = 1$
- 3 つの集合に属しており、4 つ以上の集合に属していない要素のカウンタは  $3 - {}_3C_2 + {}_3C_3 = 1$
- 4 つの集合に属しており、4 つ以上の集合に属していない要素のカウンタは  $4 - {}_4C_2 + {}_4C_3 = 2$
- 5 つの集合に属しており、4 つ以上の集合に属していない要素のカウンタは  $5 - {}_5C_2 + {}_5C_3 = 5$
- ...
- $n$  個の集合に属している要素のカウンタは  $n - {}_nC_2 + {}_nC_3 > 0$

4 つ以上の集合に含まれる要素はオーバーカウントされている。

もう気づいたと思うが、これまでやってきたように「足しすぎては引き、引きすぎては足し、…」を繰り返すと最終的にメモリの値は包除原理の主張である  $|A_1| + \dots + |A_n| - |A_1 \cap A_2| - |A_1 \cap A_3| - \dots - |A_{n-1} \cap A_n| + |A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| + \dots + |A_{n-2} \cap A_{n-1} \cap A_n| - \dots + (-1)^{n-1} |A_1 \cap \dots \cap A_n|$  になり、カウンタの値は次のようになる。

- 1 つの集合に属しており、2 つ以上の集合に属していない要素のカウンタは 1
- 2 つの集合に属しており、3 つ以上の集合に属していない要素のカウンタは  $2 - {}_2C_2 = 1$
- 3 つの集合に属しており、4 つ以上の集合に属していない要素のカウンタは  $3 - {}_3C_2 + {}_3C_3 = 1$
- 4 つの集合に属しており、4 つ以上の集合に属していない要素のカウンタは  $4 - {}_4C_2 + {}_4C_3 - {}_4C_4 = 1$
- 5 つの集合に属しており、4 つ以上の集合に属していない要素のカウンタは  $5 - {}_5C_2 + {}_5C_3 - {}_5C_4 + {}_5C_5 = 1$
- ...
- $n$  個の集合に属している要素のカウンタは  $n + \sum_{i=2}^n (-1)^{i-1} {}_nC_i = -\sum_{i=1}^n (-1)^i {}_nC_i = -(\sum_{i=0}^n (-1)^i {}_nC_i - 1) = -((1-1)^n - 1) = 1$

全ての要素を過不足無くカウントできていることがわかる。

### 1.6.3 配列の飛び飛びマーキング

$m < n$  を互いに素な自然数とする。自然数  $k$  を 1 から順番に増やしながら、長さ  $n$  の配列  $A$  の  $km \% n$  番目の要素をマークしていけば、 $k = n$  の時にすべての要素を丁度 1 回ずつマーク完了する。

*Proof.*

$m, n$  は互いに素であるから、 $A[n-1]$  が初めてマークされるのは  $km = \text{lcm}(m, n)$ 、すなわち  $k = n$  の時である。仮に、それまでのある時刻で  $A[0] \sim A[n-2]$  のどれかが 2 回マークされたとすると、動作の周期性から、 $k$  をどれだけ増やしていってっても、それまでにマークされた要素を何度もマークするだけで、 $A[n-1]$  は

いつまでもマークされない。これは不合理である。また、仮に  $k = n$  の時点で未マークの要素があるとする  
と、鳩の巣原理から少なくとも 1 つの要素が 2 回マークされていることになり、不合理である。  $\square$

## 第 1.7 章

# 置換

### 1.7.1 置換の符号の定理

#### 1.7.1.1 $\sigma'(i) := \sigma(n+1-i)$ としたときの $\text{sgn}(\sigma')$

$\sigma$  を  $S = \{1, \dots, n\}$  上の置換とする。  $S$  上の置換  $\sigma'$  を  $\sigma'(i) = \sigma(n+1-i)$  で定めるとき、  $\text{sgn}(\sigma')$  は次式となる。

$$\begin{aligned}\text{sgn}(\sigma') &= s \text{sgn}(\sigma) \\ s &= \begin{cases} 1 & (n \text{ is even, } n/2 \text{ is even}) \\ -1 & (n \text{ is even, } n/2 \text{ is odd}) \\ 1 & (n \text{ is odd, } (n-1)/2 \text{ is even}) \\ -1 & (n \text{ is odd, } (n-1)/2 \text{ is odd}) \end{cases} \\ &= \begin{cases} 1 & (n(n-1)/2 \text{ is even}) \\ -1 & (n(n-1)/2 \text{ is odd}) \end{cases} \\ &= (-1)^{n(n-1)/2}\end{aligned}$$

*Proof.*

Cauchy の 2 行記法に於いて  $\sigma'$  の下行は  $\sigma$  のそれを左右反転したものになっている。つまり  $\sigma'$  を作用させることは  $\sigma$  を作用させた後に逆順にする置換を施すことに等しい。逆順にする置換は、  $1, 2, \dots, n$  を左右の端同士で交換する処理を外側から内側に向かって進める方針で考えれば定理の主張が成り立つことがわかる。  $\square$

Mathematica を用いた数値例による検証は `theorem_sign_of_permutation.nb` を参照

#### 1.7.1.2 $\sigma'(i) := n+1-\sigma(i)$ としたときの $\text{sgn}(\sigma')$

$\sigma$  を  $S = \{1, \dots, n\}$  上の置換とする。  $S$  上の置換  $\sigma'$  を  $\sigma'(i) = n+1-\sigma(i)$  で定めるとき、  $\text{sgn}(\sigma')$  は次式となる。

$$\text{sgn}(\sigma') = s \text{sgn}(\sigma)$$

$s$  は 1.7.1.1 のものと等しい。

*Proof.*

$\sigma$  を互換の積で表したものを  $\sigma = b_k b_{k-1} \cdots b_1 b_0$  とする。但し  $b_0 := \epsilon$  (単位置換) とする。 $\sigma_0 := b_0$ ,  $\sigma_l := b_l \sigma_{l-1}$  ( $l = 1, 2, \dots, k$ ) とする。

$\sigma'_0$  は逆順にする置換である。 $1, 2, \dots, n$  を左右の端同士で交換する処理を外側から内側に向かって進める方針で考えれば  $\sigma'_0$  の符号は定理の主張の  $s$  に等しいことが解る。

$\sigma'_l$  と  $\sigma'_{l-1}$  を比較すると、 $b_l$  に対応する部分が交換されているので  $\text{sgn}(\sigma'_l) = -\text{sgn}(\sigma'_{l-1})$  である。これより  $\text{sgn}(\sigma') = \text{sgn}(\sigma'_k) = (-1)^k \text{sgn}(\sigma'_0) = \text{sgn}(\sigma) s$  □

Mathematica を用いた数値例による検証は `theorem_sign_of_permutation.nb` を参照

## 第 1.8 章

# 有限幾何学

### 1.8.1 有限射影平面

1.8.1.1 位数  $n$  の有限射影平面には  $n^2 + n + 1$  個の点と直線が存在し、各点は  $n + 1$  本の直線に含まれる。

*Proof.*

念の為、「位数が  $n$  である」というのは各直線が丁度  $n + 1$  個の点を含むこととして定義されていることを断っておく。求めたい点の数を  $n_p$ 、直線の数を  $n_l$  とし、各点が  $n_b$  本の直線に含まれるとして  $n_p, n_l, n_b$  を求める。

まず  $n_b$  が点に依らず一定であることを示す。任意の点  $p$  を選び、それを含む直線の数を  $n_b(p)$  とする。この  $n_b(p)$  本の直線によって全ての点がカバーされる。なぜならば、もしカバーされない点が存在すれば、有限射影平面の公理 1 「2 つの異なる任意の点が与えられたとき、それらを含むような直線が唯 1 つだけ存在する。」によりその点と  $p$  を結ぶ直線が存在して、それは先述の  $n_b(p)$  本の直線のどれとも一致しない。次に公理 2 「2 つの異なる任意の直線の交わりは唯 1 つの点を含む。」より、 $p$  は先述の  $n_b(p)$  本の直線の唯一つの共有点であり、各直線は  $p$  以外に  $n$  個の点を含むから

$$nn_b(p) + 1 = n_p \quad \therefore n_b(p) = (p - 1)/n$$

となり、 $n_b(p)$  の値は  $p$  に依存しない。この値を  $n_b$  と表すことにする。よって次式が成り立つ。

$$nn_b + 1 = n_p \tag{1}$$

次に、 $n_l$  本の直線から 2 本を選ぶ方法の数を 2 通りに表して方程式を 1 本立てる。1 つ目の表し方は明らかに  $\binom{n_l}{2}$  である。公理 2 より、2 本の直線を選ぶと、その共有点として対応する 1 点が決まる。これは一対一対応ではなく、先述のように 1 点に対してそれを含む直線が  $n_b$  本存在するので、各点に於いてそれらの直線から 2 本を選ぶ操作を全ての点に対して行えば、全ての直線から 2 本の直線を選ぶ方法を網羅できる。つまり次式が成り立つ。

$$\binom{n_l}{2} = n_p \binom{n_b}{2} \tag{2}$$

次に、 $n_p$  個の点から 2 つを選ぶ方法の数を 2 通りに表して方程式を 1 本立てる。1 つ目の表し方は明らかに  $\binom{n_p}{2}$  である。公理 1 より、2 個の点を選ぶと、それらを含む直線が 1 つ決まる。これは一対一対応ではなく、先述のように 1 直線に対してそれに含まれる点が  $n + 1$  個存在するので、各直線に於いてそれが含む点か

ら 2 つを選ぶ操作を全ての直線に対して行えば、全ての点から 2 つの点を選ぶ方法を網羅できる。つまり次式が成り立つ。

$$\binom{n_p}{2} = n_l \binom{n+1}{2} \quad (3)$$

式 (3) より

$$n_l = \frac{n_p(n_p - 1)}{n(n + 1)} \quad (4)$$

となる。これを式 (2) に代入して整理すると

$$\frac{n_p - 1}{n(n + 1)} \left( \frac{n_p(n_p - 1)}{n(n + 1)} - 1 \right)$$

となる。これに式 (1) を適用して整理すると  $(n_b - 1)(n_b - n - 1) = 0$  を得る。公理 3 「どの 3 点も同一直線上にないような 4 点集合が存在する」より  $n_b = 1$  は不適なので  $n_b = n + 1$  である。これと式 (1) より  $n_p = n^2 + n + 1$  である。これと式 (4) より  $n_l = n^2 + n + 1 = n_p$  である。□

第 II 部

代数学

## 第 II.1 章

# 群

### II.1.1 剰余類

#### II.1.1.1 定義

$G$  を群とし、 $H$  をその任意の部分群とする。 $g \in G$  を任意にとる。

「 $G$  に於ける  $H$  の左剰余類 (the left coset of  $H$  in  $G$ )」とは次の集合族のことを言う。

$$\{gH \mid g \in G\} = \{\{gh \mid h \in H\} \mid g \in G\}$$

また、「 $G$  に於ける  $H$  の右剰余類」とは次の集合族のことを言う。

$$\{Hg \mid g \in G\} = \{\{hg \mid h \in H\} \mid g \in G\}$$

特に、 $gH$  ( $g \in G$ ) を「 $G$  に於ける  $g$  に関する  $H$  の左剰余類 (the left coset of  $H$  in  $G$  with respect to  $g$ )」と呼び、 $Hg$  ( $g \in G$ ) を「 $G$  に於ける  $g$  に関する  $H$  の右剰余類」と呼ぶ。

#### II.1.1.2 群 $G$ の部分群 $H$ の全ての剰余類は互いに素である

*Proof.*

左剰余類について示す。右剰余類についても同様に示せる。

命題の対偶「2つの左剰余類  $R = gH$ ,  $R' = g'H$  ( $g \in G$ ) の共通部分が非空ならば  $R = R'$ 」を示す。 $R$  と  $R'$  の共通部分を  $gh = g'h'$  ( $h, h' \in H$ ) とすると  $g' = gh h'^{-1}$  だから  $\forall h'' \in H$ ,  $(R' \ni) g'h'' = gh h'^{-1} h'' \in gH = R$  ( $\because h h'^{-1} h'' \in H$ ) であり  $R' \subseteq R$  となる。同様に  $R' \supseteq R$  も示せるから  $R = R'$  となる。  $\square$

#### II.1.1.3 2つの元が同じ剰余類に属する必要十分条件

$G$  を群とし、 $H$  をその部分群とする。 $s_1, s_2 \in G$  が  $H$  の同じ左剰余類に属する必要十分条件は  $s_1^{-1}s_2 \in H$  である。また、同じ右剰余類に属する必要十分条件は  $s_1 s_2^{-1} \in H$  である。

*Proof.*

左剰余類について示す。右剰余類についても同様に示せる。

( $\Rightarrow$ )



仮定より  $\exists g \in G, h_1, h_2 \in H$  s.t.  $s_1 = gh_1, s_2 = gh_2$ 。よって  $s_1^{-1}s_2 = h_1^{-1}g^{-1}gh_2 = h_1^{-1}h_2 \in H$   
 $(\Leftarrow)$

仮定より  $\exists h \in H$  s.t.  $s_1^{-1}s_2 = h$ 。直前に示した定理より  $H$  の剰余類は  $G$  を分割するから、 $\exists g \in G, h_1 \in H$  s.t.  $s_1 = gh_1 (\in gH)$ 。よって  $s_2 = s_1h = gh_1h \in gH (\because h_1h \in H)$  □

## II.1.2 $\mathbb{Z}_p$ ( $p$ : 素数) に乗法群としての位数 $r$ の元があれば $p-1$ は $r$ の倍数

*Proof.*

論題の元を  $a$  とする。 $r$  は  $a^r \equiv_p 1$  となる最小の自然数であるから、 $k \in \mathbb{N}$  に対して  $a^k \equiv_p 1$  となるのは  $k$  が  $r$  の倍数であるとき、かつそのときに限る。Fermat の小定理から  $a^{p-1} \equiv_p 1$  であるので、 $p-1$  は  $r$  の倍数である。 □

## 第 II.2 章

# 多項式

### II.2.1 規約

- 以降の説明に於いて多項式中に特別な断りなく現れる  $x$  は不定元を意味する。
- 体  $F$  上の、 $x$  を不定元とする多項式全体の集合を  $F[x]$  と表記する。
- 体  $F$  上の、 $x$  を不定元とする Laurent 多項式全体の集合を  $F[x, x^{-1}]$  と表記する。

### II.2.2 係数が次数に関して対称な奇数次の多項式 $f(x)$ を $(x+1)$ で除した ものも係数が次数に関して対称になる

*Proof.*

$n \in \mathbb{N}$  を奇数とし、 $f(x) = \sum_{i=0}^n a_i x^i$ ,  $a_i = a_{n-i}$  とする。 $f(x)$  は  $(x+1)$  で割り切れる。 $g(x) = f(x)/(x+1) = \sum_{i=0}^{n-1} b_i x^i$  としたときに  $b_i = b_{n-1-i}$  となることを示す。

$$\begin{aligned} f(x) &= (x+1)g(x) = \sum_{i=0}^{n-1} (b_i x^{i+1} + b_i x^i) = b_0 + \sum_{i=1}^{n-1} b_i x^i + \sum_{i=1}^n b_{i-1} x^i \\ &= b_0 + \sum_{i=1}^{n-1} (b_{i-1} + b_i) x^i + b_{n-1} x^n \end{aligned}$$

これよりまず  $b_0 = a_0 = a_n = b_{n-1}$  となる。次に  $x, x^2, \dots, x^{n-1}$  の項の係数に関しては  $b_{i-1} + b_i = a_i = a_{n-i} = b_{n-1-i} + b_{n-i}$  すなわち  $b_i - b_{n-1-i} = b_{(n-1)-(i-1)} - b_{i-1}$  となる。既に見た  $b_0 = b_{n-1}$  から始めて帰納的に示せる。  $\square$

### II.2.3 $\sum_{i=0}^n a_i x^i$ が既約多項式ならば $\sum_{i=0}^n a_{n-i} x^i$ も既約多項式である

体  $F$  上の  $n$  次多項式  $f_1(x) := \sum_{i=0}^n a_i x^i$  が  $F$  上で既約であれば、係数の順番を逆にした多項式  $f_2(x) := \sum_{i=0}^n a_{n-i} x^i$  もまた既約である。

*Proof.*

背理法で示す。 $f_2(x)$  が既約でないと仮定すると、適当な多項式  $g_{2,1}(x), g_{2,2}(x) \in F[x]$ ,  $d_{2,1} + d_{2,2} :=$

$\deg(g_{2,1}(x)) + \deg(g_{2,2}(x)) = n$  が存在して次式が成り立つ。

$$f_2(x) = g_{2,1}(x)g_{2,2}(x)$$

上式を Laurent 多項式に拡張して両辺の  $x$  を  $x^{-1}$  で置き換えると

$$f_2(x^{-1}) = g_{2,1}(x^{-1})g_{2,2}(x^{-1})$$

両辺に  $x^n$  を掛けると

$$f_1(x) = x^n f_2(x^{-1}) = x^{d_{2,1}} g_{2,1}(x^{-1}) x^{d_{2,2}} g_{2,2}(x^{-1}) =: g_{1,1}(x) g_{1,2}(x)$$

ここに  $g_{1,1}(x), g_{1,2}(x)$  は  $g_{2,1}(x), g_{2,2}(x)$  の係数の順番を逆にしたものである。これは  $f_1(x)$  が既約であることに矛盾する。  $\square$

## 11.2.4 既約多項式 $\iff$ 最小多項式

$P$  を体  $F$  上の多項式とする。 $P$  が  $F$  上で既約であることと、 $P$  がその任意の根の最小多項式であることは同値である。

*Proof.*

( $\Rightarrow$ )

背理法で示す。 $\alpha$  を  $P$  の任意の根とする。仮に  $P$  より低次の  $F$  上の多項式  $P_2$  が存在して  $P_2(\alpha) = 0$  であるとする。 $P$  を  $P_2$  で割った商を  $Q_2$ 、余りを  $P_3$  とすると、 $0 = P(\alpha) = Q_2(\alpha)P_2(\alpha) + P_3(\alpha)$  より  $P_3(\alpha) = 0$  である。同じ要領で  $P_i$  を  $P_{i+1}$  で割った商を  $Q_{i+1}$ 、余りを  $P_{i+2}$  とすると  $0 = P(\alpha) = P_2(\alpha) = P_3(\alpha) = \dots$  となる。 $P_i$  の次数は  $i$  の増加と共に真に減少するため、次のいずれかが成り立つ。

1. ある  $i$  で  $P_{i+1} \mid P_i$
2. ある  $i$  で  $\deg(P_i) = 0$  となる

Euclid の互除法により、1. の場合は  $\gcd(P, P_2) = \gcd(P_2, P_3) = \dots = \gcd(P_i, P_{i+1}) = P_{i+1}$  となり、 $P$  が既約であるという前提に矛盾する。2. の場合も先述の通り  $P_i(\alpha) = 0$  であり、 $\deg(P_i) = 0$  に注意して  $P_i = 0$  であるので  $P_{i-1} \mid P_{i-2}$  であり、1. と同様に矛盾が生じる。

( $\Leftarrow$ )

背理法で示す。 $\alpha$  を  $P$  の任意の根とする。仮に  $P$  が既約でないとする。すると  $F$  上の  $P$  より低次の 2 つの多項式  $P_2, P_3$  を用いて  $P = P_2 P_3$  と表せる。 $\alpha$  は  $P$  の根であるから  $0 = P(\alpha) = P_2(\alpha)P_3(\alpha)$  より  $P_2(\alpha) = 0$  または  $P_3(\alpha) = 0$  であるが、これは  $P$  が  $\alpha$  の最小多項式であることに矛盾する。  $\square$

## 第 II.3 章

# 有限体

### II.3.1 素数位数の有限体の構成法

$p$  を素数とし、 $S := \{0, 1, \dots, p-1\}$  とする。 $S$  の任意の元  $a, b$  に対する演算  $+$ ' と  $*$ ' を次式で定義する。

- $a + ' b = (a + b) \% p$
- $a * ' b = (ab) \% p$

このようにすると代数構造  $(S, +', *')$  が体を成すことを説明する。

まず  $+$ ' に関する結合法則, 単位元と唯一の逆元の存在、 $*$ ' に関する結合法則と単位元の存在、また、 $*$ ' が  $+$ ' の上に分配的であることは合同式の基本的な知識があれば明らかであろう。あとは、 $*$ ' に関する唯一の逆元が存在すること、すなわち任意の  $0 \neq a \in S$  に対して唯一の  $b \in S$  が存在して  $a * ' b = 1$  となることを示せばよいのだが、I.3.2 より、それが保証される。

有限体は Galois 体とも呼ばれ、位数  $p$  の有限体を  $\text{GF}(p)$  と表す。

### II.3.2 位数が素数の冪乗である有限体の構成法

#### II.3.2.1 既約多項式を用いる方法

$p$  を素数とする。 $\text{GF}(p)$  上の既約多項式を任意に一つ選んで  $P$  とし、 $m = \deg(P)$  とする。 $S$  を、 $\text{GF}(p)$  上の次数  $m-1$  以下の全ての多項式の集合を  $S$  とする ( $|S| = p^m$ )。  $S$  の任意の元  $a, b$  に対する演算  $+$ ' と  $*$ ' を次式で定義する。

- $a + ' b = (a + b) \% P$
- $a * ' b = (ab) \% P$

このようにすると代数構造  $(S, +', *')$  が体を成すことを説明する。

まず  $+$ ' に関する結合法則, 単位元と唯一の逆元の存在、 $*$ ' に関する結合法則と単位元の存在、また、 $*$ ' が  $+$ ' の上に分配的であることは合同式, 素数位数の有限体の基本的な知識があれば明らかであろう。あとは、 $*$ ' に関する唯一の逆元が存在すること、すなわち任意の  $0 \neq a \in S$  に対して唯一の  $b \in S$  が存在して  $a * ' b = 1$  となることを示せばよい。少なくとも一つ存在することは Bézout の等式により保証される。唯一であることを背理法で示す。仮に異なる  $b_1, b_2 \in S$  が存在して  $a * ' b_1 = a * ' b_2 = 1$  であるとする。  $a * ' (b_1 - b_2) = 0$  であるが、これは  $P$  は既約であることと矛盾する。なぜならば  $a, b_1, b_2$  は  $P$  と互いに素で

あるから、 $P \nmid a *' (b_1 - b_2)$  であるからである。

### II.3.2.2 既約多項式の根を用いる方法 (根体)

$p$  を素数とする。 $\text{GF}(p)$  上の既約多項式を任意に一つ選んで  $P$  とし、 $m = \deg(P)$  とする。 $\alpha$  を  $P$  の任意の根とする。 $S := \left\{ \sum_{i=0}^{m-1} a_i \alpha^i \mid a_0, \dots, a_{m-1} \in \text{GF}(p) \right\}$  とする。 $S$  上の演算として  $\text{GF}(p)$  上の多項式に対する加法と乗法をそのまま使うと、 $S$  とこの演算の組が体を成すことを説明する。

まず加法に関する結合法則、単位元と唯一の逆元の存在、乗法に関する結合法則と単位元の存在、また、乗法が加法の上に分配的であることは  $P(\alpha) = 0$  であること、素数位数の有限体、多項式の除法の基本的な知識があれば明らかであろう。あとは、乗法に関する唯一の逆元が存在すること、すなわち任意の  $0 \neq a \in S$  に対して唯一の  $b \in S$  が存在して  $ab = 1$  となることを示せばよい。少なくとも一つ存在することは、 $a$  に現れる  $\alpha$  を不定元  $x$  で置換して多項式に戻してから Bézout の等式を考えることで理解できる。唯一であることは背理法で示せる。仮に異なる  $b_1, b_2 \in S$  が存在して  $ab_1 = ab_2 = 1$  であるとする  $a(b_1 - b_2) = 0$  であるが、一方  $a, b_1 - b_2 \neq 0$  だから  $a(b_1 - b_2) \neq 0$  であり、矛盾している。

## 第 II.4 章

# 形式的冪級数

### II.4.1 規約

- 体  $F$  上の、 $x$  を不定元とする形式的冪級数全体の集合を  $F[[x]]$  と表記する。
- 体  $F$  上の、 $x$  を不定元とする形式的 Laurent 級数全体の集合を  $F((x))$  と表記する。

### II.4.2 原始多項式

#### II.4.2.1 $\sum_{i=0}^n a_i x^i$ が原始多項式ならば $\sum_{i=0}^n a_{n-i} x^i$ も原始多項式である

$p$  を素数とする。 $n$  次多項式  $f_1(x) := \sum_{i=0}^n a_i x^i \in \text{GF}(p)[x]$  が原始多項式であれば、係数の順番を逆にした多項式  $f_2(x) := \sum_{i=0}^n a_{n-i} x^i$  もまた原始多項式である。

*Proof.*

II.2.3 より  $f_2(x)$  は  $\text{GF}(p)$  上で既約である。あとは  $f_2(x) \mid x^l - 1$  となる最小の自然数  $l$  が  $p^n - 1$  であることを示せば良い。

まず  $f_2(x) \mid x^{p^n-1} - 1$  を示す。 $f_1(x)$  が原始多項式だから適当な  $p^n - 1 - n$  次多項式  $g_1(x) \in \text{GF}(p)[x]$  が存在して次式が成り立つ。

$$x^{p^n-1} - 1 = g_1(x)f_1(x)$$

上式を Laurent 多項式に拡張して両辺の  $x$  を  $x^{-1}$  で置き換えると

$$\frac{1}{x^{p^n-1}} - 1 = g_1(x^{-1})f_1(x^{-1})$$

両辺に  $x^{p^n-1}$  を掛けると

$$1 - x^{p^n-1} = x^{p^n-1-n}g_1(x^{-1})x^n f_1(x^{-1}) =: g_2(x)f_2(x)$$

$$\therefore x^{p^n-1} - 1 \equiv 0 \pmod{f_2(x)}$$

ここに  $g_2(x)$  は  $g_1(x)$  の係数の順番を逆にしたものである。

次に  $l = 1, 2, \dots, p^n - 2$  について  $f_2(x) \nmid x^l - 1$  を背理法で示す。仮にある  $l$  に対して  $l - n$  次多項式  $h_1(x) \in \text{GF}(p)[x]$  が存在して

$$x^l - 1 = h_1(x)f_2(x)$$

が成り立つとする。上式を Laurent 多項式に拡張して両辺の  $x$  を  $x^{-1}$  で置き換えると

$$\frac{1}{x^l} - 1 = h_1(x^{-1})f_2(x^{-1})$$

両辺に  $x^l$  を掛けると

$$1 - x^l = x^{l-n}h_1(x^{-1})x^n f_2(x^{-1}) =: h_2(x)f_1(x)$$

$$\therefore x^l - 1 \equiv 0 \pmod{f_1(x)}$$

ここに  $h_2(x)$  は  $h_1(x)$  の係数の順番を逆にしたものである。これは  $f_1(x)$  が原始多項式であることに矛盾する。

以上より  $f_2(x) \mid x^l - 1$  となる最小の自然数  $l$  は  $p^m - 1$  である。  $\square$

#### II.4.2.2 原始多項式の逆数の係数列は周期的

$p$  を素数とする。 $n$  次の原始多項式  $f(x) \in \text{GF}(p)[[x]]$  の逆数  $f(x)^{-1}$  の係数列は周期的であり、その周期は  $p^n - 1$  である。

*Proof.*

まず係数列が周期的であることを示す。 $f(x)$  が原始多項式だから、適当な  $p^n - 1 - n$  次多項式  $g(x) \in \text{GF}(p)[[x]]$  が存在して次式が成り立つ。

$$x^{p^n-1} - 1 = g(x)f(x)$$

$$\therefore f(x)^{-1} = -g(x) \left(1 - x^{p^n-1}\right)^{-1} = -g(x) \sum_{i=0}^{\infty} x^{i(p^n-1)}$$

$\deg g(x) < p^n - 1$  だから右辺の係数列は周期的であり、その周期は高々  $p^n - 1$  である。「高々」と言ったのは、 $g(x)$  の係数列のパターンに依ってはより短い周期になり得るからである。この周期が真に  $p^n - 1$  であることを背理法で示す。

仮に周期が  $p^n - 2$  次以下であるとする、つまりある  $l \in \{1, 2, \dots, p^n - 2\}$  と適当な  $l - 1$  次以下の多項式  $h(x) \in \text{GF}(p)[[x]]$  が存在して

$$f(x)^{-1} = -h(x) \sum_{i=0}^{\infty} x^{il}$$

が成り立つと仮定すると次式を得る。

$$f(x)^{-1} = -h(x) \sum_{i=0}^{\infty} x^{il} = h(x)(x^l - 1)^{-1}$$

$$\therefore x^l - 1 = h(x)f(x) \equiv 0 \pmod{f(x)}$$

これは  $f(x)$  が原始多項式であることに矛盾する。  $\square$

### II.4.2.3 原始多項式 $f_1(x)$ の係数の順番を逆転した $f_2(x)$ の逆数の係数列は $f_1(x)^{-1}$ の係数列の逆順逆符号

$p$  を素数とする。 $n$  次の原始多項式  $f_1(x) \in \text{GF}(p)[[x]]$  の係数の順番を逆転させ  $f_2(x)$  とする。 $f_1(x)^{-1} =: a_0 + a_1x + \dots$  とする。このとき次式が成り立つ。

$$f_2(x)^{-1} = - \sum_{j=0}^{p^n-1-n} a_{p^n-1-n-j} x^j \sum_{i=0}^{\infty} x^{i(p^n-1)}$$

視覚的に言えば、 $f_2(x)^{-1}$  の係数列は周期的であり、 $f_1(x)^{-1}$  の係数列に現れるパターンを逆転させ、符号を逆にしたものになる。

#### II.4.2.3.1 例

$f_1(x) = 1 + x^2 + x^3 \in \text{GF}(2)[[x]]$  とすると  $f_2(x) = 1 + x + x^3$  である。

$$\begin{aligned} f_1(x)^{-1} &= 1 + 0x + 1x^2 + 1x^3 + 1x^4 + 0x^5 + 0x^6 \\ &\quad + 1x^7 + 0x^8 + 1x^9 + 1x^{10} + 1x^{11} + 0x^{12} + 0x^{13} \\ &\quad + 1x^{14} + 0x^{15} + 1x^{16} + \dots \\ &= (1 + 0x + 1x^2 + 1x^3 + 1x^4 + 0x^5 + 0x^6) \sum_{i=0}^{\infty} x^{7i} \end{aligned}$$

$f_2(x)^{-1}$  は次のようになる。

$$\begin{aligned} f_2(x)^{-1} &= 1 + 1x + 1x^2 + 0x^3 + 1x^4 + 0x^5 + 0x^6 \\ &\quad + 1x^7 + 1x^8 + 1x^9 + 0x^{10} + 1x^{11} + 0x^{12} + 0x^{13} \\ &\quad + 1x^{14} + 1x^{15} + 1x^{16} + \dots \\ &= (1 + 1x + 1x^2 + 0x^3 + 1x^4 + 0x^5 + 0x^6) \sum_{i=0}^{\infty} x^{7i} \end{aligned}$$

$f_2(x)^{-1}$  の係数列のパターンは  $f_1(x)^{-1}$  のそれを逆転させたものになっている。係数体が  $\text{GF}(2)$  なので符号反転は起こらない。

#### II.4.2.3.2 証明

*Proof.*

II.4.2.2 より  $f_1(x)^{-1}$  の係数列は周期  $p^n - 1$  であるから、適当な  $p^n - 1$  次未満の多項式  $g_1(x) \in \text{GF}(p)[[x]]$  が存在して次式が成り立つ。

$$f_1(x)^{-1} = g_1(x) \sum_{i=0}^{\infty} x^{i(p^n-1)}$$

この式を Laurent 多項式に拡張して変形してゆく。

$$\begin{aligned} f_1(x)^{-1} &= g_1(x) \sum_{i=0}^{\infty} x^{i(p^n-1)} = -g_1(x) (x^{p^n-1} - 1)^{-1} \\ \therefore x^{p^n-1} - 1 &= -g_1(x) f_1(x) = -g_1(x) x^n f_2(x^{-1}) \end{aligned}$$



ここに  $f_2(x)$  は  $f_1(x)$  の係数の順番を逆転させたものである。初めに  $\deg g_1(x) < p^n - 1$  としていたが、上式より  $\deg g_1(x) = p^n - 1 - n$  が確定する。上式の  $x$  を  $x^{-1}$  で置き換えると次式を得る。

$$x^{-(p^n-1)} - 1 = -g_1(x^{-1})x^{-n}f_2(x)$$

両辺に  $x^{p^n-1}$  を掛けると

$$1 - x^{p^n-1} = -x^{p^n-1-n}g_1(x^{-1})f_2(x) =: -g_2(x)f_2(x)$$

ここに  $g_2(x)$  は  $g_1(x)$  の係数の順番を逆転させたものである。これより次式を得る。

$$f_2(x)^{-1} = -g_2(x) \left(1 - x^{p^n-1}\right)^{-1} = -g_2(x) \sum_{i=0}^{\infty} x^{i(p^n-1)}$$

□

## 第 II.5 章

# ブール代数

### II.5.1 双対原理

*Proof.*

論理式  $P(x)$  の双対  $P(x)^d$  は  $\overline{P(\overline{x})}$  で表せる。 $P(x) = Q(x)$  ならば  $\overline{P(\overline{x})} = \overline{Q(\overline{x})}$  なので  $P(x)^d = Q(x)^d$  である。□

### II.5.2 応用

#### II.5.2.1 2進 Gray コードへの変換

$x$  の 2 進表現から Gray コードへの変換は  $x$  の 2 進表現と、それを 1 ビット右シフトして先頭に 0 をつけたものとの排他的論理和で得られる。これで上手くいく理由を考えてみる。 $x$  の  $n$  bit 2 進表現を  $b_{n-1}, b_{n-2}, \dots, b_1, b_0$  とする。Gray コードへの変換操作を乱暴に言うと、 $b_{n-1}$  から  $b_0$  に向かって見ながら、 $b_{n-1}$  はそのまま、以降は 1 と 0 が交互に現れる箇所を 1、そうでない箇所を 0 に置き換える。この操作で  $x$  と  $x+1$  の Gray コードのハミング距離が 1 になることを示す。

(1) 最下位桁が 0 である場合

$$\begin{aligned} x &= b_{n-1}, b_{n-2}, \dots, b_1, 0 && \xrightarrow{\text{変換}} && b_{n-1}, b_{n-1} \oplus b_{n-2}, \dots, b_2 \oplus b_1, \overline{b_1} \\ x+1 &= b_{n-1}, b_{n-2}, \dots, b_1, 1 && \xrightarrow{\text{変換}} && b_{n-1}, b_{n-1} \oplus b_{n-2}, \dots, b_2 \oplus b_1, \overline{\overline{b_1}} \end{aligned}$$

両者の Gray コードは最下位桁のみが異なっている。

(2) 最下位桁が 1 である場合

$x$  の下位  $k$  桁が 1 の連続であるとする。

$$\begin{aligned} x &= b_{n-1}, b_{n-2}, \dots, b_{k+1}, \underbrace{0, 1, \dots, 1}_{k \text{ 個}} && \xrightarrow{\text{変換}} && b_{n-1}, b_{n-1} \oplus b_{n-2}, \dots, b_{k+2} \oplus b_{k+1}, \overline{\overline{b_{k+1}}}, \underbrace{1, 0, \dots, 0}_{k-1 \text{ 個}} \\ x+1 &= b_{n-1}, b_{n-2}, \dots, b_{k+1}, \underbrace{1, 0, \dots, 0}_{k \text{ 個}} && \xrightarrow{\text{変換}} && b_{n-1}, b_{n-1} \oplus b_{n-2}, \dots, b_{k+2} \oplus b_{k+1}, \overline{\overline{b_{k+1}}}, \underbrace{1, 0, \dots, 0}_{k-1 \text{ 個}} \end{aligned}$$

両者の Gray コードは第  $k + 1$  桁のみが異なっている。

## 第Ⅲ部

## 実解析

## 第 III.1 章

# 表記

- $\mathbb{R}_+ := \{x \in \mathbb{R} | x \geq 0\}$  : 非負の実数全体の集合
- $\mathbb{R}_{++} := \{x \in \mathbb{R} | x > 0\}$  : 正の実数全体の集合
- $\mathbf{1}_n$  : 要素が全て 1 である  $n$  次元列ベクトル
- $\mathbf{0}_n$  : 要素が全て 0 である  $n$  次元列ベクトル
- $\mathbf{x}[i]$  : 列 (行) ベクトルの第  $i$  成分
- $A[i][j]$ ,  $A_{i,j}$  : 行列  $A$  の第  $(i, j)$  成分
- $A[i_1 : i_2][j_1 : j_2]$  : 行列  $A$  の第  $i_1$  行から第  $i_2$  行までの部分と第  $j_1$  列から第  $j_2$  列までの部分を抽出したもの
- $A[i][:]$  : 行列  $A$  の第  $i$  行
- $A[:, j]$  : 行列  $A$  の第  $j$  列
- $[A]_{i,j}$  : 行列  $A$  の第  $i$  行  $j$  列を取り除いてできる部分行列。

## 第 III.2 章

# 離散数学との関係

### III.2.1 Dirichlet の Diophantine 近似定理

$$\forall \alpha \in \mathbb{R}, \forall N \in \mathbb{N}, \exists (N \geq) q \in \mathbb{N}, p \in \mathbb{Z} \text{ s.t. } \left| \alpha - \frac{p}{q} \right| < \frac{1}{qN}$$

*Proof.*

示したい式を変形して  $|q\alpha - p| < 1/N$  と見る。 $q \in \{1, \dots, N\}$  と変化させるとき、 $q\alpha$  の中に格子点との距離が  $1/N$  未満のものが存在することを示せば良い。つまり  $q\alpha \% 1$  のうちで  $[0, 1/N)$  に含まれるものが存在することを示せば良い。区間  $I_k := [k/N, (k+1)/N)$  ( $k = 0, \dots, N-1$ ) を考える。 $q\alpha \% 1$  は  $I_0, \dots, I_{N-1}$  のどれかに入る。もし仮に  $I_0$  に入るものが無いとすると、鳩ノ巣原理よりある自然数  $1 \leq q_1 < q_2 \leq N$  が存在して  $q_1\alpha \% 1$  と  $q_2\alpha \% 1$  がともにある区間  $I_k$  に含まれる。つまりある  $m \in \mathbb{Z}, \varepsilon \in \mathbb{R}, |\varepsilon| < 1/N$  が存在して  $q_2\alpha = q_1\alpha + m + \varepsilon$  となる。このとき  $|(q_2 - q_1)\alpha - m| = |\varepsilon| < 1/N$  が成り立つ。 $q = q_2 - q_1, p = m$  とすれば定理の主張を満たす  $p/q$  が得られる。□

### III.2.2 床関数と天井関数

$$\text{III.2.2.1 } \forall x \in \mathbb{R}, \lfloor -x \rfloor = -\lceil x \rceil, \lceil -x \rceil = -\lfloor x \rfloor$$

*Proof.*

前者は次のようにして示せる。

$$n = \lfloor -x \rfloor \iff n \in \mathbb{Z}, -x - 1 < n \leq -x \iff n \in \mathbb{Z}, x \leq -n < x + 1 \iff -n = \lceil x \rceil \iff n = -\lceil x \rceil$$

前者の式で  $x$  を  $-x$  で置き換えれば後者が得られる。□

$$\text{III.2.2.2 } \forall x, y \in \mathbb{R}, \lfloor x \rfloor + \lfloor y \rfloor \leq \lfloor x + y \rfloor, \lceil x \rceil + \lceil y \rceil \geq \lceil x + y \rceil$$

*Proof.*

( $\lfloor x \rfloor + \lfloor y \rfloor \leq \lfloor x + y \rfloor$  の証明)

$x = X + m_x, y = Y + m_y, X, Y \in \mathbb{Z}, m_x, m_y \in [0, 1)$  とすると  $\lfloor x \rfloor + \lfloor y \rfloor = X + Y$  であり、一方で  $\lfloor x + y \rfloor = \lfloor X + Y + m_x + m_y \rfloor = X + Y + \lfloor m_x + m_y \rfloor \geq X + Y$  であるので、主張が従う。

( $\lceil x \rceil + \lceil y \rceil \geq \lceil x + y \rceil$  の証明)

$x = X - m_x$ ,  $y = Y - m_y$ ,  $X, Y \in \mathbb{Z}$ ,  $m_x, m_y \in [0, 1)$  とすると  $\lceil x \rceil + \lceil y \rceil = X + Y$  であり、一方で  $\lceil x + y \rceil = \lceil X + Y - m_x - m_y \rceil = X + Y + \lceil -m_x - m_y \rceil \leq X + Y$  であるので、主張が従う。  $\square$

## 第 III.3 章

# 極限

### III.3.1 問題例

III.3.1.1  $\lim_{x \rightarrow \infty} \sqrt{x^2 + 1} - x = 0$

*Proof.*

$y(x) := \sqrt{x^2 + 1} - x$  とおく。  $y(x) > 0$  は明らか。

$$y(x) + x = \sqrt{x^2 + 1} \quad \therefore y(x)^2 + 2xy(x) + x^2 = x^2 + 1 \quad \therefore y(x)^2 + 2xy(x) = 1$$

$$\therefore y(x)(y(x) + 2x) = 1 \quad \therefore 0 < y(x) = \frac{1}{2x + y(x)} < \frac{1}{2x} \rightarrow 0 \text{ as } x \rightarrow \infty$$

挟みうちの原理から  $y(x) \rightarrow 0$  である。

□

III.3.1.2  $a > 0, \lim_{x \rightarrow +0} \frac{x^2}{\sqrt{x^2 + a^2} - a} = 2a$

*Proof.*

$y = x + \sqrt{x^2 + a^2} - a$  とおき、 $x$  について解くと  $x = \frac{y(y+2a)}{2(y+a)}$  となる。また、 $\sqrt{x^2 + a^2} - a = y - x = \frac{y^2}{2(y+a)}$  となる。 $x \rightarrow +0$  のとき  $y \rightarrow +0$  であり、

$$\lim_{x \rightarrow +0} \frac{x^2}{\sqrt{x^2 + a^2} - a} = \lim_{y \rightarrow +0} \frac{(y + 2a)^2}{2(y + a)} = 4a^2/(2a) = 2a$$

□



## 第 III.4 章

# 数列

### III.4.1 単調性

III.4.1.1  $\left(1 - \frac{t}{n}\right)^n$  ( $t \in [0, n]$ ) は  $n$  に関して単調増加

*Proof.*

$a_n := \left(1 - \frac{t}{n}\right)^n$  とする。 $t = 0, n$  の場合は単調増加性は明らか。以下  $0 < t < n$  とする。

$$\begin{aligned} a_{n+1}/a_n &= \frac{(n+1-t)^{n+1}}{(n+1)^{n+1}} \times \frac{n^n}{(n-t)^n} = \left(\frac{(n+1-t)n}{(n+1)(n-t)}\right)^n \times \frac{n+1-t}{n+1} \\ &= \left(1 + \frac{t}{(n+1)(n-t)}\right)^n \times \frac{n+1-t}{n+1} = \sum_{k=0}^n \binom{n}{k} \left(\frac{t}{(n+1)(n-t)}\right)^k \times \frac{n+1-t}{n+1} \\ &\geq \left(1 + \frac{nt}{(n+1)(n-t)}\right) \times \frac{n+1-t}{n+1} = \frac{(n^2+n-t)(n+1-t)}{(n+1)^2(n-t)} \\ &= \frac{(n+1)^2(n-t) + t^2}{(n+1)^2(n-t)} > 1 \end{aligned}$$

□

### III.4.2 極限

III.4.2.1  $\lim_{n \rightarrow \infty} \left(1 + 1/n + \sum_{k=2}^{\infty} \frac{c_k}{n^k}\right)^n = e$  ( $\sum_{k=2}^{\infty} c_k$  converges)

*Proof.*

$$1 + 1/n + \sum_{k=2}^{\infty} \frac{c_k}{n^k} = 1 + \frac{1}{n} \left(1 + \sum_{k=2}^{\infty} \frac{c_k}{n^{k-1}}\right)$$

であるから  $\epsilon > 0$  を任意にとると、十分大きい  $n \in \mathbb{N}$  に対して

$$\left(1 + \frac{1-\epsilon}{n}\right)^n \leq \left(1 + \frac{1}{n} \left(1 + \sum_{k=2}^{\infty} \frac{c_k}{n^{k-1}}\right)\right)^n \leq \left(1 + \frac{1+\epsilon}{n}\right)^n$$

両辺で  $n \rightarrow \infty$  として従う。

□

### III.4.2.2 $\lim_{n \rightarrow \infty} (1 + 1/n + o(1/n))^n = e$

*Proof.*

$o(1/n)$  の部分を  $f(n)$  とおく。

$$\left(1 + \frac{1}{n} + f(n)\right)^n = (1 + 1/n)^n + \sum_{i=1}^n {}_nC_i (1 + 1/n)^{n-i} f(n)^i$$

$\lim_{n \rightarrow \infty} (1 + 1/n)^n = e$  の証明より、上式において  $(1 + 1/n)^{n-i} \leq 3$  であることがわかっているの

$$\begin{aligned} \sum_{i=1}^n {}_nC_i (1 + 1/n)^{n-i} f(n)^i &\leq 3 \sum_{i=1}^n {}_nC_i f(n)^i = 3 \sum_{i=1}^n \frac{1}{i!} \frac{n(n-1) \cdots (n-i+1)}{n^i} n^i f(n)^i \\ &\leq 3 \sum_{i=1}^n \frac{1}{i!} (nf(n))^i \leq 3nf(n) \sum_{i=1}^n \frac{1}{i!} (nf(n))^{i-1} \end{aligned}$$

$n$  を十分大きくすれば  $nf(n) \leq 1$  なので

$$\sum_{i=1}^n {}_nC_i (1 + 1/n)^{n-i} f(n)^i \leq 3nf(n) \sum_{i=1}^n \frac{1}{i!} \leq 6nf(n) \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

□

## III.4.3 漸化式

### III.4.3.1 Fibonacci 数列

#### III.4.3.1.1 つがいの増殖からの漸化式の導出

第  $n$  世代のつがいの数を  $a_n$  とし、そのうち成熟しているものを  $m_n$ 、未成熟のものを  $y_n$  とすると、 $a_{n+1} - a_n = m_n$  であり、同時に  $m_n = m_{n-1} + y_{n-1}$  なので結局  $a_{n+1} = a_n + a_{n-1}$  となる。

### III.4.3.2 $a_{n+1} = ra_n + p(n)$ , ( $r \neq 1$ , $p(n)$ : $d$ 次多項式) の一般項は $a_n = Ar^n + q(n)$ , ( $A$ : 定数, $q(n)$ : $d$ 次多項式)

*Proof.*

$p(n) = b_d n^d + \cdots + b_1 n + b_0$  とする。 $a_n = Ar^n + q(n)$  ( $q(n) = c_d n^d + \cdots + c_1 n + c_0$ ) において漸化式に代入すると次式を得る。

$$Ar^{n+1} + q(n+1) = Ar^{n+1} + rq(n) + p(n) \quad \therefore q(n+1) = rq(n) + p(n)$$

両辺の  $k$  次の項の係数を比較すると右辺では  $rc_k + b_k$ 、左辺では  $\sum_{i=k}^d \binom{i}{k} c_i$  であり、 $r \neq 1$  なので  $d$  次の係数から順に見てゆき、 $c_d, c_{d-1}, \dots, c_1, c_0$  の順に決定できる。最後に初期条件  $a_0$  (或いは  $a_1, a_2, \dots$ , 値が分かっている項なら何でも良い) の値から  $A$  を決定できる。 □

## 第 III.5 章

# 級数

### III.5.1 $S_l = \sum_{k=0}^{\infty} k^l r^k$ を逐次的に求める

$S_l := \sum_{k=0}^{\infty} k^l r^k$  とすると次式が成り立つ。

$$S_l = \frac{1}{1-r} \sum_{m=0}^{l-1} (-1)^{l-m+1} {}_l C_m (S_m - 0^m)$$

*Proof.*

$$rS_l = \sum_{k=0}^{\infty} k^l r^{k+1} = \sum_{k=1}^{\infty} (k-1)^l r^k \quad \text{より}$$

$$\begin{aligned} (1-r)S_l &= \sum_{k=1}^{\infty} [k^l - (k-1)^l] r^k = \sum_{k=1}^{\infty} \sum_{m=0}^{l-1} {}_l C_m k^m (-1)^{l-m+1} r^k = \sum_{m=0}^{l-1} {}_l C_m (-1)^{l-m+1} \sum_{k=1}^{\infty} k^m r^k \\ &= \sum_{m=0}^{l-1} {}_l C_m (-1)^{l-m+1} (S_m - 0^m) \end{aligned}$$

□

$l=2$  まで具体的に求めた結果:

$$S_0 = 1/(1-r), \quad S_1 = r/(1-r)^2, \quad S_2 = r(r+1)/(1-r)^3$$

### III.5.2 Abel の総和公式: $\sum_{i=1}^n a_i b_i = A_n b_n - \sum_{i=1}^{n-1} A_i (b_{i+1} - b_i)$

数列  $\{a_k\}, \{b_k\}$  に対して次が成り立つ。

$$\sum_{i=1}^n a_i b_i = A_n b_n - \sum_{i=1}^{n-1} A_i (b_{i+1} - b_i) \quad \left( A_i := \sum_{j=1}^i a_j \right)$$

*Proof.*

(泥臭い証明)

表 III.5.1 第2項目の寄与

$i$	$a_j$ の係数
1	0
$\vdots$	$\vdots$
$j-1$	0
$j$	$-b_{j+1} + b_j$
$j+1$	$-b_{j+2} + b_{j+1}$
$\vdots$	$\vdots$
$n-1$	$-b_n + b_{n-1}$

右辺の各  $a_j$  ( $j = 1, \dots, n$ ) について、その係数が正味  $b_i$  だけであることを示せば良い。(右辺) =  $A_n b_n - \sum_{i=1}^{n-1} \sum_{j=1}^i a_j (b_{i+1} - b_i)$  であるから、右辺に現れる  $a_j$  の係数として、まず第1項目より  $b_n$  がある。第2項目の寄与は左の表のようになり、殆ど相殺して  $b_j$  のみが残る。

(エレガントな証明)

$A_i := \sum_{j=1}^i a_j$ ,  $A_0 := 0$  と定義すると

$$\begin{aligned} \sum_{i=1}^n a_i b_i &= \sum_{i=1}^n (A_i - A_{i-1}) b_i = \sum_{i=1}^n A_i b_i - \sum_{i=2}^n A_{i-1} b_i = \sum_{i=1}^n A_i b_i - \sum_{i=1}^n A_i b_{i+1} \\ &= A_n b_n + \sum_{i=1}^n A_i (b_i - b_{i+1}) = A_n b_n - \sum_{i=1}^n A_i (b_{i+1} - b_i) \end{aligned}$$

□

### III.5.3 Chebyshev の和の不等式

1. 単調増加あるいは単調減少な数列  $\{a_n\}, \{b_n\}$  について、 $\left(\frac{1}{n} \sum_{i=1}^n a_i\right) \left(\frac{1}{n} \sum_{i=1}^n b_i\right) \leq \frac{1}{n} \sum_{i=1}^n a_i b_i$  が成り立つ。
2. 単調増加列  $\{a_n\}$  と単調減少列  $\{b_n\}$  について、 $\left(\frac{1}{n} \sum_{i=1}^n a_i\right) \left(\frac{1}{n} \sum_{i=1}^n b_i\right) \geq \frac{1}{n} \sum_{i=1}^n a_i b_i$  が成り立つ。

*Proof.*

以下の証明は [https://en.wikipedia.org/wiki/Chebyshev%27s\\_sum\\_inequality](https://en.wikipedia.org/wiki/Chebyshev%27s_sum_inequality) によるものである。

1. を示す。2. は同様に示せる。 $S := \sum_{i=1}^n \sum_{j=1}^n (a_i - a_j)(b_i - b_j)$  とする。 $\{a_n\}, \{b_n\}$  の単調性の仮定から、 $(a_i - a_j)$  と  $(b_i - b_j)$  の符号は等しく、 $S \geq 0$  である。よって次式が成り立つ。

$$\begin{aligned} 0 \leq S &= \sum_{i=1}^n \sum_{j=1}^n (a_i b_i - a_i b_j - a_j b_i + a_j b_j) = 2n \sum_{i=1}^n a_i b_i - 2 \left( \sum_{i=1}^n a_i \right) \left( \sum_{j=1}^n b_j \right) \\ &\therefore \left( \frac{1}{n} \sum_{i=1}^n a_i \right) \left( \frac{1}{n} \sum_{i=1}^n b_i \right) \leq \frac{1}{n} \sum_{i=1}^n a_i b_i \end{aligned}$$

□

### III.5.4 等比級数と0収束列の畳み込み:

$$|r| < 1, \lim_{n \rightarrow \infty} a_n = 0 \Rightarrow \lim_{t \rightarrow \infty} \sum_{\tau=0}^t r^{t-\tau} a_\tau = 0 = 0$$

$|r| < 1$  とし、数列  $\{a_n\}$  が  $\lim_{n \rightarrow \infty} a_n = 0$  を満たすとき、次が成り立つ。

$$\lim_{t \rightarrow \infty} \sum_{\tau=0}^t r^{t-\tau} a_\tau = 0$$

*Proof.*

$\lim_{n \rightarrow \infty} a_n = 0$  より  $\exists C \geq 0$  s.t.  $|a_n| \leq C$  for  $\forall n \in \{1, 2, \dots\}$ 。また、for  $\forall \varepsilon > 0, \exists T \geq 1$  s.t.  $t \geq T \Rightarrow |a_t| \leq \varepsilon$ 。そこで  $t \geq T$  とすると

$$\begin{aligned} \left| \sum_{\tau=0}^t r^{t-\tau} a_\tau \right| &\leq \left| \sum_{\tau=0}^{T-1} r^{t-\tau} a_\tau \right| + \left| \sum_{\tau=T}^t r^{t-\tau} a_\tau \right| \leq \sum_{\tau=0}^{T-1} |r|^{t-\tau} |a_\tau| + \sum_{\tau=T}^t |r|^{t-\tau} |a_\tau| \\ &\leq C |r|^{t-T+1} \frac{1 - |r|^{T-1}}{1 - |r|} + \varepsilon \frac{1 - |r|^{t-T}}{1 - |r|} \leq C \frac{|r|^{t-T+1}}{1 - |r|} + \varepsilon \frac{1}{1 - |r|} \end{aligned}$$

$|r| < 1$  であるから  $\exists T_2 \geq T$  s.t.  $t \geq T_2 \Rightarrow C |r|^{t-T+1} \leq \varepsilon$ 。そこで  $t \geq T_2$  と取り直すと、上の式より

$$\left| \sum_{\tau=0}^t r^{t-\tau} a_\tau \right| \leq \frac{2\varepsilon}{1 - |r|}$$

□

### III.5.5 $\sum_{k=1}^n \cos kx = -\frac{1}{2} + \frac{\sin(n+1/2)x}{2\sin(x/2)}$ ( $x \neq 0$ )

*Proof.*

$$\begin{aligned} (\sin x) \sum_{k=1}^n \cos kx &= \frac{1}{2} \sum_{k=1}^n (\sin(1+k)x + \sin(1-k)x) \\ &= \frac{1}{2} \left( \sum_{l=2}^{n+1} (\sin lx + \sin(-lx)) + \sin 0x + \sin(-x) + \sin nx + \sin(n+1)x \right) \\ \therefore \sum_{k=1}^n \cos kx &= -\frac{1}{2} + \frac{1}{2\sin x} (\sin nx + \sin(n+1)x) \\ &= -\frac{1}{2} + \frac{1}{4\sin \frac{x}{2} \cos \frac{x}{2}} 2\sin \frac{(2n+1)x}{2} \cos \frac{-x}{2} = -\frac{1}{2} + \frac{\sin(n+1/2)x}{2\sin \frac{x}{2}} \end{aligned}$$

□

$$\text{III.5.6} \quad \sum_{k=1}^n \sin kx = \frac{\sin \frac{nx}{2} \sin \frac{(n+1)x}{2}}{2 \sin(x/2)} \quad (x \neq 0)$$

*Proof.*

$$\begin{aligned} (\sin x) \sum_{k=1}^n \sin kx &= \frac{1}{2} \sum_{k=1}^n (\cos(1-k)x - \cos(1+k)x) \\ &= \frac{1}{2} \left( \sum_{l=2}^{n-1} (\cos lx - \cos(-lx)) + \cos 0x + \cos x - \cos nx - \cos(n+1)x \right) \\ &= \frac{1}{2} \left( 1 + \cos x - 2 \cos \frac{(2n+1)x}{2} \cos \frac{-x}{2} \right) \\ \therefore \sum_{k=1}^n \sin kx &= \frac{1}{2 \sin x} \left( 2 \cos^2 \frac{x}{2} - 2 \cos \frac{(2n+1)x}{2} \cos \frac{-x}{2} \right) \\ &= \frac{1}{2 \sin \frac{x}{2} \cos \frac{x}{2}} \left( \cos^2 \frac{x}{2} - \cos \frac{(2n+1)x}{2} \cos \frac{x}{2} \right) \\ &= \frac{1}{2 \sin \frac{x}{2}} \left( \cos \frac{x}{2} - \cos \frac{(2n+1)x}{2} \right) \\ &= \frac{1}{2 \sin \frac{x}{2}} (-2) \sin \frac{(n+1)x}{2} \sin \frac{-nx}{2} \\ &= \frac{\sin \frac{nx}{2} \sin \frac{(n+1)x}{2}}{2 \sin(x/2)} \end{aligned}$$

□

## 第 III.6 章

# 乗積

III.6.1  $a_i \geq 0$  ( $i = 1, 2, \dots$ ),  $\sum_{i=1}^{\infty} (1 - a_i) = \infty \Rightarrow \prod_{i=1}^{\infty} a_i = 0$

*Proof.*

$a_i = 0$  となる  $i \in \mathbb{N}$  が存在する場合は命題の主張は明らかに成り立つので、以下では  $a_i \in (0, 1]$  とする。  
 $b_i = -\log a_i$  とすると  $a_i = e^{-b_i}$  であり、 $\forall x \in \mathbb{R}, 1 - e^{-x} \leq x$  より  $1 - a_i = 1 - e^{-b_i} \leq b_i$  なので次式が成り立つ。

$$\infty = \sum_{i=1}^{\infty} (1 - a_i) \leq \sum_{i=1}^{\infty} b_i$$

よって

$$\prod_{i=1}^{\infty} a_i = \prod_{i=1}^{\infty} e^{-b_i} = \exp\left(-\sum_{i=1}^{\infty} b_i\right) = 0$$

□

## 第 III.7 章

# 位相空間

### III.7.1 $A \subseteq B \Rightarrow \text{cl } A \subseteq \text{cl } B$

*Proof.*

もしもある  $a \in \text{cl } A, a \notin \text{cl } B$  が存在すれば、 $a$  に収束するある点列  $\{a_k\}, a_k \in A \subseteq B \subseteq \text{cl } B$  に対してその収束先  $a$  が  $\text{cl } B$  に含まれない。これは  $\text{cl } B$  が閉集合であることに矛盾する。  $\square$



## 第 III.8 章

# 部分分数分解

### III.8.1 諸公式

$$\text{III.8.1.1} \quad \frac{1}{(x^2-1)^2} = \frac{1}{4} \left[ \frac{1}{x+1} - \frac{1}{x-1} + \frac{1}{(x+1)^2} + \frac{1}{(x-1)^2} \right]$$

$$\frac{1}{(x^2-1)^2} = \frac{1}{4} \left[ \frac{1}{x+1} - \frac{1}{x-1} + \frac{1}{(x+1)^2} + \frac{1}{(x-1)^2} \right]$$

### III.8.2 諸定理

#### III.8.2.1 定理 III.8.2.1.0

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = a_n (x - x_n)(x - x_{n-1}) \cdots (x - x_1)$$

と因数分解できるとして、

$$\frac{\alpha}{a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0} = \frac{c_n}{x - x_n} + \frac{c_{n-1}}{x - x_{n-1}} + \cdots + \frac{c_1}{x - x_1}$$

ならば

$$c_n x_n + c_{n-1} x_{n-1} + \cdots + c_1 x_1 = 0$$

*Proof.*

$$\frac{\alpha}{a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0} = \frac{c_n}{x - x_n} + \frac{c_{n-1}}{x - x_{n-1}} + \cdots + \frac{c_1}{x - x_1}$$

の両辺に  $(x - x_n)(x - x_{n-1}) \cdots (x - x_1)$  を掛けて整理すると左辺は

$$\frac{\alpha}{a_n}$$

右辺について

$x^{n-1}$  の係数は

$$\sum_{i=1}^n c_i$$

であり、これは0でなくてはならない。 $x$  の係数は

$$-\sum_{i=1}^n c_i \sum_{j=1, j \neq i}^n x_j$$

であるが、特に

$$\sum_{i=1}^n c_i \sum_{j=1, j \neq i}^n x_j = \sum_{i=1}^n c_i \sum_{j=1}^n x_j - \sum_{i=1}^n c_i x_i = - \sum_{i=1}^n c_i x_i \left( \because \sum_{i=1}^n c_i = 0 \right)$$

であるから結局  $x$  の係数は

$$\sum_{i=1}^n c_i x_i = c_n x_n + c_{n-1} x_{n-1} + \cdots + c_1 x_1$$

これは0でなくてはならない。

□

## 第 III.9 章

# 指示関数

### III.9.1 諸定理

III.9.1.1  $\Omega$  を定義域とする写像  $f, g, T$  に対して  $\forall x \in \Omega, \forall t \in T(\Omega), \mathbb{1}\{T(x) = t\} f(x) = \mathbb{1}\{T(x) = t\} g(x) \iff \forall x \in \Omega, f(x) = g(x)$

*Proof.*

( $\Rightarrow$ )

$x \in \Omega$  を任意にとる。 $t = T(x)$  を満たす  $t \in T(\Omega)$  をとって主張の左側の式に代入すると右側の式を得る。

( $\Leftarrow$ )

$x \in \Omega, t \in T(\Omega)$  を任意にとる。 $t = T(x)$  の場合は主張の右側の式より左側の式が成り立つ。 $t \neq T(x)$  の場合は主張の左側の式は  $0 = 0$  となる。  $\square$

## 第 III.10 章

# 微分

$$\text{III.10.1} \quad f(a) < 0 < f(b) \Rightarrow \exists x \in (a, b) \text{ s.t. } f(x) = 0, f'(x) \geq 0$$

区間  $[a, b]$  ( $a < b$ ) で定義された連続関数  $f$  が开区間  $(a, b)$  で微分可能であるとする。 $f(a) < 0 < f(b)$  であるとき、ある  $x \in (a, b)$  が存在して  $f(x) = 0, f'(x) \geq 0$  を満たす。

*Proof.*

$f(x) = 0$  となる  $x \in (a, b)$  の存在は中間値の定理が保証している。そのような点のうち最小のものを  $x_0$  とするとこれが定理の主張を満たすことを背理法で示す。もしも  $f'(x_0) < 0$  であれば、微分係数の定義より、十分小さい  $\delta > 0$  が存在して  $a < x_0 - \delta, f(x_0 - \delta) > 0$  なので中間値の定理より、ある  $y \in (a, x_0 - \delta)$  が存在して  $f(y) = 0$  となる。これは  $x_0$  が  $f(x) = 0$  を満たす最小の点であったことに矛盾する。  $\square$

$$\text{III.10.2} \quad \lim_{x \rightarrow \infty} f(x) = 0 \text{ であっても } \lim_{x \rightarrow 0} f'(x) = 0 \text{ とは限らない}$$

$C^1$  級関数  $f(x)$  について  $\lim_{x \rightarrow \infty} f(x) = 0$  であっても  $\lim_{x \rightarrow 0} f'(x) = 0$  とは限らない。

*Proof.*

例えば  $f(x) = \frac{\sin x^2}{x}$  は  $\lim_{x \rightarrow \infty} f(x) = 0$  であるが  $f'(x) = \frac{2x^2 \cos x^2 - x \sin^2 x^2}{x^2}$  であって  $x \rightarrow \infty$  で 0 にならない。  $\square$

## III.10.3 soft-step 関数

工学的な道具として使うための恣意的な関数を考えてみる。ここでは滑らかに 0 から 1 に変化するステップ状の関数を考えてみる。

### III.10.3.1 $C^{n-1}$ 級の soft-step 関数

$\sigma > 0, n \in \mathbb{N}$  とする。次の関数は  $\mathbb{R}$  上で  $C^{n-1}$  級である。

$$f(x) = \begin{cases} 0 & (x < -\sigma) \\ \frac{1}{2\sigma^n}(x + \sigma)^n & (-\sigma \leq x < 0) \\ \frac{(-1)^{n+1}}{2\sigma^n}(x - \sigma)^n & (0 \leq x < \sigma) \\ 1 & (x \geq \sigma) \end{cases}$$

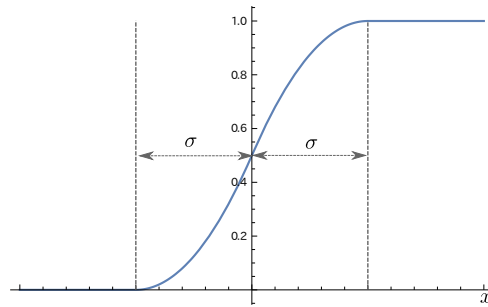


図 III.10.1 グラフ

$k(< n-1)$  階微分は次のようになる

$$f(x) = \begin{cases} 0 & (x < -\sigma) \\ \frac{1}{2\sigma^n} {}_n P_k (x + \sigma)^{n-k} & (-\sigma < x < 0) \\ \frac{(-1)^{n+1}}{2\sigma^n} {}_n P_k (x - \sigma)^{n-k} & (0 < x < \sigma) \\ 1 & (x > \sigma) \end{cases}$$

$k \in \{1, \dots, n-1\}$  に対して隣接する領域同士の境界点で左側, 右側微分係数が一致するので  $C^{n-1}$  級である。

## 第 III.11 章

# Taylor 級数

### III.11.1 $e^{-x^2}$ は全域で Taylor 展開可能

*Proof.*

$f(x) = e^{-x^2}$  ( $x \in \mathbb{R}$ ) とする。これ自体は  $f$  が複素関数として全平面で解析的であることを示して結果を実数に限定することで容易に示せるが、ここでは敢えて実解析の範囲で剰余項の評価を評価することで Taylor 展開可能であることを示す。 $f$  の  $n$  階微分は  $x$  の  $n$  次多項式と  $e^{-x^2}$  の積になり、多項式の係数の最大の絶対値が  $2^n$  であることは容易に確かめられる。よって

$$|f^{(n)}(x)| \leq e^{-x^2} \sum_{i=0}^n 2^n |x|^i$$

右辺は  $x$  を固定するごとに高々定数の  $n$  乗なので  $n!$  で除した結果は  $n \rightarrow \infty$  で 0 に収束する。 □

## 第 III.12 章

# 凸関数

### III.12.1 諸定理

#### III.12.1.1 $f_1, f_2$ が凸なら $\max\{f_1, f_2\}$ は凸

*Proof.*

$g(\mathbf{x}) := \max\{f_1(\mathbf{x}), f_2(\mathbf{x})\}$  とする。 $g$  の定義域内の任意の  $\mathbf{x}_1, \mathbf{x}_2$  と任意の  $\theta \in [0, 1]$  に対して

$$(1 - \theta)g(\mathbf{x}_1) + \theta g(\mathbf{x}_2) = (1 - \theta) \max\{f_1(\mathbf{x}_1), f_2(\mathbf{x}_1)\} + \theta \max\{f_1(\mathbf{x}_2), f_2(\mathbf{x}_2)\}$$

であるが、 $\max$  の性質から

$$\max\{f_1(\mathbf{x}_1), f_2(\mathbf{x}_1)\} \geq f_1(\mathbf{x}_1), f_2(\mathbf{x}_1) \text{ かつ } \max\{f_1(\mathbf{x}_2), f_2(\mathbf{x}_2)\} \geq f_1(\mathbf{x}_2), f_2(\mathbf{x}_2)$$

であるので

$$(1 - \theta)g(\mathbf{x}_1) + \theta g(\mathbf{x}_2) \geq (1 - \theta)f_1(\mathbf{x}_1) + \theta f_1(\mathbf{x}_2) \geq f_1[(1 - \theta)\mathbf{x}_1 + \theta\mathbf{x}_2]$$

かつ

$$(1 - \theta)g(\mathbf{x}_1) + \theta g(\mathbf{x}_2) \geq (1 - \theta)f_2(\mathbf{x}_1) + \theta f_2(\mathbf{x}_2) \geq f_2[(1 - \theta)\mathbf{x}_1 + \theta\mathbf{x}_2]$$

である。このことから

$$(1 - \theta)g(\mathbf{x}_1) + \theta g(\mathbf{x}_2) \geq \max\{f_1[(1 - \theta)\mathbf{x}_1 + \theta\mathbf{x}_2], f_2[(1 - \theta)\mathbf{x}_1 + \theta\mathbf{x}_2]\} = g[(1 - \theta)\mathbf{x}_1 + \theta\mathbf{x}_2]$$

□

#### III.12.1.2 Jensen の不等式

$f(x)$  を凸関数とし、 $w_i \geq 0 (i = 1, \dots, n)$ ,  $\sum_{i=1}^n w_i = 1$  とするとき、次式が成り立つ。

$$f\left(\sum_{i=1}^n w_i x_i\right) \leq \sum_{i=1}^n w_i f(x_i)$$

*Proof.*

$n = 1$  の時は明らか。 $n = 2$  のときは凸関数の定義より成り立つ。 $n = n_0 \in \mathbb{N}$  のとき成り立つと仮定して  $n = n_0 + 1$  のとき成り立つことを示す。 $w'_i := w_i / (1 - w_{n_0+1})$  とする (最初の  $n_0$  個の重みを規格化したも

の) と  $w'_i > 0$ ,  $w'_1 + \cdots + w'_{n_0} = 1$  だから、仮定より

$$\begin{aligned}
 f\left(\sum_{i=1}^{n_0} w'_i x_i\right) &\leq \sum_{i=1}^{n_0} w'_i f(x_i) \\
 \therefore f\left(\sum_{i=1}^{n_0} \frac{w_i}{1-w_{n_0+1}} x_i\right) &\leq \sum_{i=1}^{n_0} \frac{w_i}{1-w_{n_0+1}} f(x_i) \\
 \therefore (1-w_{n_0+1})f\left(\sum_{i=1}^{n_0} \frac{w_i}{1-w_{n_0+1}} x_i\right) &\leq \sum_{i=1}^{n_0} w_i f(x_i) \\
 \therefore w_{n_0+1}f(x_{n_0+1}) + (1-w_{n_0+1})f(X_{n_0}) &\leq \sum_{i=1}^{n_0+1} w_i f(x_i) \quad \left(X_{n_0} = \sum_{i=1}^{n_0} \frac{w_i}{1-w_{n_0+1}} x_i \text{ おいた} \right)
 \end{aligned}$$

$f$  の凸性より

$$\text{左辺} \geq f(w_{n_0+1}x_{n_0+1} + (1-w_{n_0+1})X_{n_0}) = f\left(\sum_{i=1}^{n_0+1} w_i x_i\right)$$

□

### III.12.1.3 重み付き相加, 相乗平均の関係と Young の不等式

$w_1, \dots, w_n \geq 0$ ,  $w_1 + \cdots + w_n = 1$ ,  $x_1, \dots, x_n \geq 0$  とするとき、次式が成り立つ。

$$\sum_{i=1}^n w_i x_i \geq \prod_{i=1}^n x_i^{w_i}$$

この関係は**重み付き相加, 相乗平均の関係**と呼ばれている。

*Proof.*

$x_i = 0$  である  $i$  が存在するときは明らかに成り立つ。以下では全ての  $i$  について  $x_i > 0$  とする。対数関数の凹性を利用する。Jensen の不等式より

$$\log \sum_{i=1}^n w_i x_i \geq \sum_{i=1}^n w_i \log x_i \quad \therefore \log \sum_{i=1}^n w_i x_i \geq \log \prod_{i=1}^n x_i^{w_i} \quad \therefore \sum_{i=1}^n w_i x_i \geq \prod_{i=1}^n x_i^{w_i}$$

最初の式より、等号成立条件は  $x_1 = \cdots = x_n$  である。 □

この不等式には変形版があり、**Young の不等式**と呼ばれている。一時的に上の定理の条件を少し強めて  $w_1, \dots, w_n > 0$  とし、 $y_i = x_i^{w_i}$  とすると  $x_i = y_i^{1/w_i}$  となる。これらを重み付き相加, 相乗平均の関係に代入して

$$\sum_{i=1}^n w_i y_i^{1/w_i} \geq \prod_{i=1}^n y_i$$

を得る。 $p_i = 1/w_i$  とすると次式を得る。

$$\sum_{i=1}^n \frac{y_i^{p_i}}{p_i} \geq \prod_{i=1}^n y_i$$

改めて定理の形で述べると、次のようになる。



$p_1, \dots, p_n > 1, 1/p_1 + \dots + 1/p_n = 1, x_1, \dots, x_n \geq 0$  とするとき、次式が成り立つ。

$$\sum_{i=1}^n \frac{x_i^{p_i}}{p_i} \geq \prod_{i=1}^n x_i$$

### III.12.1.4 2 階導関数と凸性の関係

$a \in \mathbb{R}$  とし、 $I$  は  $a$  の開近傍であるとする。 $f: I \rightarrow \mathbb{R}$  は  $I$  上で  $C^2$  級であるとする。次の 3 つの事項

1.  $\forall x \in I, f^{(2)}(x) > 0$
2.  $\forall (x_1, x_2 \in I, x_1 \neq x_2), f(x_2) > f(x_1) + f^{(1)}(x_1)(x_2 - x_1)$
3.  $f$  は  $I$  上で狭義に下に凸である。

に対して  $1 \Rightarrow 2, 2 \iff 3$  である。さらに  $1 \Leftarrow 2$  は偽である。

*Proof.*

$1 \Leftarrow 2$  が偽であることは反例  $f(x) = x^4$  から直ちに示される。

$1 \Rightarrow 2$  を示す。 $x \in I, x \neq x_1$  とする。 $g(x) := f(x) - (f(x_1) + f^{(1)}(x_1)(x - x_1))$  とすると  $g^{(1)}(x) = f^{(1)}(x) - f^{(1)}(x_1)$  なので  $g^{(1)}(x_1) = 0$  である。さらに  $g^{(2)}(x) = f^{(2)}(x) > 0$  である。よって  $g^{(1)}(x) < 0 (x < x_1), g^{(1)}(x) > 0 (x > x_1)$  であるので  $g(x_2) > 0$  であり、2 が成り立つ。

$2 \Rightarrow 3$  を示す。 $x_1, x_2 \in I, x_1 \neq x_2, 0 < \lambda < 1, x_3 = (1 - \lambda)x_1 + \lambda x_2$  とすると次式が成り立つ。

$$f(x_1) > f(x_3) + f^{(1)}(x_3)(x_1 - x_3), \quad f(x_2) > f(x_3) + f^{(1)}(x_3)(x_2 - x_3)$$

よって次式が成り立つ。

$$\begin{aligned} (1 - \lambda)f(x_1) + \lambda f(x_2) &> f(x_3) + (1 - \lambda)f^{(1)}(x_3)(x_1 - x_3) + \lambda f^{(1)}(x_3)(x_2 - x_3) \\ &= f(x_3) + f^{(1)}(x_3)[(1 - \lambda)(x_1 - x_3) + \lambda(x_2 - x_3)] \\ &= f(x_3) + f^{(1)}(x_3)[(1 - \lambda)x_1 + \lambda x_2 - x_3] \\ &= f(x_3) + f^{(1)}(x_3)(x_3 - x_3) = f(x_3) \end{aligned}$$

$3 \Rightarrow 2$  を示す。 $x_1 < x_2$  の場合について示すが、逆の場合も同様にして示せる。 $0 < \lambda < 1, x_3 = (1 - \lambda)x_1 + \lambda x_2$  とすると次式が成り立つ。

$$\begin{aligned} \frac{f(x_2) - f(x_1)}{x_2 - x_1} &= \frac{(1 - \lambda)f(x_1) + \lambda f(x_2) - f(x_1)}{x_3 - x_1} = \frac{(1 - \lambda)f(x_1) + \lambda f(x_2) - f(x_1)}{\lambda(x_2 - x_1)} \\ &> \frac{f(x_3) - f(x_1)}{\lambda(x_2 - x_1)} = \frac{f(x_1 + \lambda(x_2 - x_1)) - f(x_1)}{\lambda(x_2 - x_1)} \\ &\rightarrow f^{(1)}(x_1) \text{ as } \lambda \rightarrow 0 \\ \therefore \frac{f(x_2) - f(x_1)}{x_2 - x_1} &\geq f^{(1)}(x_1) \quad (\text{真に大きいとまでは言えないことに注意}) \\ f(x_2) &\geq f(x_1) + f^{(1)}(x_1)(x_2 - x_1) \end{aligned} \tag{1}$$

よって  $x_4 := (x_1 + x_2)/2$  とすると  $f(x_4) \geq f(x_1) + f^{(1)}(x_1)(x_4 - x_1)$  となる。式 (1) で  $\lambda = 1/2$  とすると

次式を得る。

$$\begin{aligned}\frac{f(x_2) - f(x_1)}{x_2 - x_1} &> \frac{f(x_4) - f(x_1)}{(x_2 - x_1)/2} \\ f(x_2) - f(x_1) &> 2f(x_4) - 2f(x_1) \geq 2f^{(1)}(x_1)(x_4 - x_1) = f^{(1)}(x_1)(x_2 - x_1) \\ \therefore f(x_2) &> f(x_1) + f^{(1)}(x_1)(x_2 - x_1)\end{aligned}$$

□

## III.12.2 Legendre 変換

### III.12.2.1 十分滑らかな狭義凸関数に対して Legendre 変換を 2 回施すと元の関数に戻る

*Proof.*

$f$  を开区間  $I$  で定義された  $C^2$  級の狭義凸関数とする。(よって 1 階導関数  $f^{(1)}$  は狭義単調増加だから逆関数が存在する。)  $f^{(1)}$  の値域を  $I_2$  とする。 $f$  の Legendre 変換  $f^*$  は次のようになる。

$$f^*(p) = pf^{(1)-1}(p) - f\left(f^{(1)-1}(p)\right) \quad (p \in I_2)$$

$f^*$  の Legendre 変換  $f^{**}$  を求めるためにまず  $f^{*(1)}$  の値域を調べる。

$$f^{*(1)}(q) \ (q \in I_2) = f^{(1)-1}(q) + q \frac{1}{f^{(2)}(q)} - f^{(1)}\left(f^{(1)-1}(q)\right) \frac{1}{f^{(2)}(q)} = f^{(1)-1}(q)$$

より  $f^{*(1)}$  の値域は  $I$  である。さらに、 $f^{(1)-1}$  は狭義単調増加なので上式より  $f^{*(1)}$  の逆関数が存在して  $f^{*(1)-1} = f^{(1)}$  である。 $f^{*(1)-1}$  の定義域は  $I$  である。以上より  $f^{**}$  は次のようになる。

$$\begin{aligned}f^{**}(q) \ (q \in I) &= qf^{*(1)-1}(q) - f^*\left(f^{*(1)-1}(q)\right) \\ &= qf^{(1)}(q) - f^*(f^{(1)}) \\ &= qf^{(1)}(q) - f^{(1)}(q)f^{(1)-1}\left(f^{(1)}(q)\right) + f\left(f^{(1)-1}\left(f^{(1)}(q)\right)\right) \\ &= f(q)\end{aligned}$$

□

## III.12.3 諸注意

### III.12.3.1 非負の凸関数同士の積は凸関数とは限らない

反例として  $x^2$  と  $(x+1)^2$  の積がある。両者は単体では下に凸であるが、両者の積は  $x = -1/2$  で上に凸である。

## 第 III.13 章

# 逆関数

### III.13.1 連続関数の逆関数は連続

*Proof.*

まず元の関数の定義域が開区間の場合について示す。 $f$  を区間  $I = (a_1, a_2)$  上の連続関数とし、 $J := f(I) =: (b_1, b_2)$  とする。 $f$  には逆関数が存在するものとする (必然的に  $f$  は狭義単調増加/減少なものに限られる)。  $f$  が狭義単調増加な場合について示すが、狭義単調減少の場合も同様にして示せる。

$a \in I$  を任意に選ぶ。 $f(a)$  に於ける  $f^{-1}$  の右連続性を否定して矛盾が生じることを示す。右連続性を否定すると、ある  $\alpha > 0$  が存在して、任意に小さい如何なる  $\gamma > 0$  に対してもある適当な  $y \in [f(a), f(a) + \gamma]$  が存在して、 $f^{-1}(y) \geq a + \alpha$  となる。

$a + \alpha$  が  $f$  の定義域に含まれることを確認しておく。 $f^{-1}(y)$  なる値が存在していることと逆関数の定義から  $f^{-1}(y)$  は  $f$  の定義域に含まれていなくてはならない。よって  $I \ni a < a + \alpha \leq f^{-1}(y) \in I$  であり、 $f$  の定義域  $I$  は区間であると冒頭で仮定しているから  $a + \alpha \in I$  でなくてはならない。

先述の通り  $f^{-1}(y) \geq a + \alpha$  であって、 $f$  は単調増加だから  $y \geq f(a + \alpha)$  である。これと  $y \leq f(a) + \gamma$  より  $f(a + \alpha) \leq f(a) + \gamma$  であり、 $\gamma$  は任意に小さくできるから  $\gamma \rightarrow 0$  とすると  $f(a + \alpha) \leq f(a)$  となり、 $f$  が狭義単調増加であるという前提に矛盾する。

同様に  $f(a)$  に於ける  $f^{-1}$  の左連続性を否定しても矛盾が生じることが示される。よって  $f^{-1}$  は  $f(a)$  で連続であり、 $a$  は  $I$  内の任意の点だから  $f(a)$  は  $J$  内の任意の点となり得るため  $f^{-1}$  は  $J$  上連続である。

次に  $f$  の定義域が閉区間  $I = [a_1, a_2]$  と表される場合について示す。半开区間の場合の証明は开区間の場合と閉区間の場合の証明を組み合わせ直ちに得られる (半开区間を隣接する閉区間と开区間に分けて考える。境界となる点での連続性は开区間の場合で証明済みである)。  $J := f(I) =: [b_1, b_2]$  とする。开区間での右連続性の証明と同様にして  $b_1$  での右連続性が示される。また开区間での左連続性の証明と同様にして  $b_2$  での左連続性が示される。以上より  $f^{-1}$  は定義域の端点を含めて連続である。

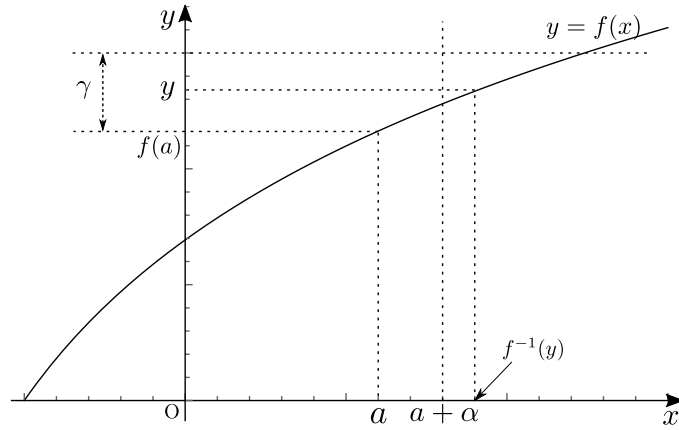


図 III.13.1  $a, \alpha, \gamma$  の様子

□

III.13.2 可微分な関数  $f$  の逆関数は  $f'(f^{-1}(x)) \neq 0$  なる点  $x$  で微分可能であり、 $\frac{d f^{-1}}{d x}(x) = 1/f'(f^{-1}(x))$ 。

*Proof.*

$f$  を開区間  $I$  上の可微分な関数とし、 $f$  には逆関数が存在するとする (必然的に  $f$  は狭義単調増加/減少なものに限られる)。まず  $I$  が開区間である場合について示す。 $f$  が狭義単調増加な場合について示すが、狭義単調減少の場合も同様にして示せる。 $b \in \text{dom}(f^{-1})$  を任意にとる。 $k > 0$  を  $b + k \in \text{dom}(f^{-1})$  となるように十分小さくとる。 $f$  の微分可能性より  $k = f'(f^{-1}(b))(f^{-1}(b + k) - f^{-1}(b)) + o(f^{-1}(b + k) - f^{-1}(b))$  であるから

$$\begin{aligned} \frac{f^{-1}(b + k) - f^{-1}(b)}{k} &= \frac{f^{-1}(b + k) - f^{-1}(b)}{f'(f^{-1}(b))(f^{-1}(b + k) - f^{-1}(b)) + o(f^{-1}(b + k) - f^{-1}(b))} \\ &= \frac{1}{f'(f^{-1}(b)) + \frac{o(f^{-1}(b + k) - f^{-1}(b))}{f^{-1}(b + k) - f^{-1}(b)}} \end{aligned}$$

$k \rightarrow 0$  とすると  $f^{-1}$  の連続性から  $f^{-1}(b + k) - f^{-1}(b) \rightarrow 0$  となるので

$$\lim_{k \rightarrow 0} \frac{f^{-1}(b + k) - f^{-1}(b)}{k} = \frac{1}{f'(f^{-1}(b))}$$

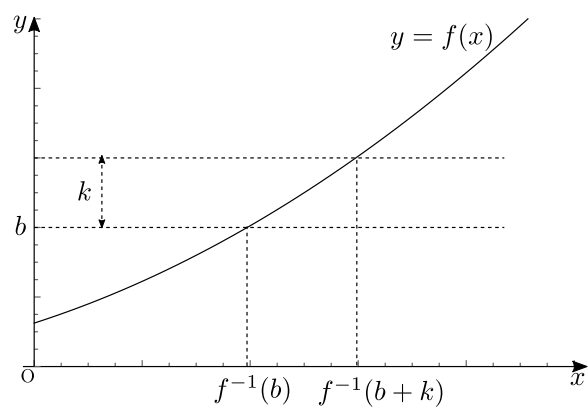


図 III.13.1 様子

□

## 第 III.14 章

# 諸定理

### III.14.1 偶関数と奇関数への分解

連続な 1 変数関数  $f(x)$  は連続な偶関数, 奇関数に一通りに分解できる

*Proof.*

(分解可能性)

$$\begin{aligned}f_e(x) &:= \frac{1}{2}(f(x) + f(-x)) \\f_o(x) &:= \frac{1}{2}(f(x) - f(-x))\end{aligned}$$

とすれば確かに  $f(x) = f_e(x) + f_o(x)$  であり  $f_e(x), f_o(x)$  はそれぞれ連続な偶, 奇関数になっている。

(一意性)

次に、分解が一通りであることを示す。 $f(x)$  の分解が次のように 2 通りあったとする。

$$\begin{cases} f(x) = f_{ea}(x) + f_{oa}(x) \\ f(x) = f_{eb}(x) + f_{ob}(x) \end{cases} \quad (1)$$

任意の  $l > 0$  に対して

$$\begin{aligned}0 &= \int_{-l}^l (f(x) - f(x))dx = \int_{-l}^l (f_{ea}(x) + f_{oa}(x) - f_{eb}(x) + f_{ob}(x))dx \\ &= 2 \int_0^l (f_{ea}(x) - f_{eb}(x))dx\end{aligned}$$

であるから両辺を  $l$  で微分すると

$$0 = f_{ea}(l) - f_{eb}(l)$$

これで  $x > 0$  において  $f_{ea}, f_{eb}$  が等しいことが示されたが、そもそも両者は偶関数であったから  $x < 0$  においても等しくならざるを得ない。残るは  $x = 0$  であるが、 $f_{oa}, f_{ob}$  は連続な奇関数なので  $f_{oa}(0) = f_{ob}(0) = 0$ 。これと式 (1) より  $f_{ea}(0) = f_{eb}(0) = f(0)$  となり、結局  $f_{ea}$  と  $f_{eb}$  は常に等しい。これと式 (1) から奇関数成分  $f_{oa}, f_{ob}$  も等しいことがわかる。□

$$\text{III.14.2} \quad \sup_{x \in X} (f(x) + g(x)) \leq \sup_{x \in X} f(x) + \sup_{x \in X} g(x), \quad \inf_{x \in X} f(x) + \inf_{x \in X} g(x) \leq \inf_{x \in X} (f(x) + g(x))$$

集合  $X$  上で定義された有界な関数  $f, g$  について次が成り立つ。

$$\sup_{x \in X} (f(x) + g(x)) \leq \sup_{x \in X} f(x) + \sup_{x \in X} g(x), \quad \inf_{x \in X} f(x) + \inf_{x \in X} g(x) \leq \inf_{x \in X} (f(x) + g(x))$$

*Proof.*

$\sup_{x \in X} (f(x) + g(x))$  を与える  $x \in X$  の一つを  $x_1$  とすると  $\sup_{x \in X} (f(x) + g(x)) = f(x_1) + g(x_1) \leq \sup_{x \in X} f(x) + \sup_{x \in X} g(x)$ 。また、 $\inf_{x \in X} (f(x) + g(x))$  を与える  $x \in X$  の一つを  $x_2$  とすると  $\inf_{x \in X} (f(x) + g(x)) = f(x_2) + g(x_2) \geq \inf_{x \in X} f(x) + \inf_{x \in X} g(x)$   $\square$

$$\text{III.14.3} \quad \sup_{x \in X} (-f(x)) = -\inf_{x \in X} f(x), \quad \inf_{x \in X} (-f(x)) = -\sup_{x \in X} f(x)$$

集合  $X$  上で定義された有界な関数  $f, g$  について次が成り立つ。

$$\sup_{x \in X} (-f(x)) = -\inf_{x \in X} f(x), \quad \inf_{x \in X} (-f(x)) = -\sup_{x \in X} f(x)$$

*Proof.*

$\forall y \in X, \inf_{x \in X} f(x) \leq f(y)$  なので  $\forall y \in X, -\inf_{x \in X} f(x) \geq -f(y)$ 。任意の  $\varepsilon > 0$  に対して  $\inf_{x \in X} f(x) + \varepsilon$  は  $f$  の下界ではないから、適当な  $x_1 \in X$  が存在して  $f(x_1) < \inf_{x \in X} f(x) + \varepsilon$ 。  $\therefore -f(x_1) > -\inf_{x \in X} f(x) - \varepsilon$ 。すなわち  $-\inf_{x \in X} f(x) - \varepsilon$  は  $-f$  の上界ではない。よって上限の定義より  $\sup_{x \in X} (-f(x)) = -\inf_{x \in X} f(x)$ 。2 つ目の主張の証明も同様。  $\square$

### III.14.4 Hölder(ヘルダー) の不等式

$p, q > 1, \frac{1}{p} + \frac{1}{q} = 1$  のとき

$$\sum_{i=1}^n |x_i y_i| \leq \|x\|_p \|y\|_q$$

*Proof.*

$x, y$  の少なくとも 1 つが零ベクトルであるときは明らかに成り立つ。以下では両者とも零ベクトルでないとする。Young の不等式より  $a, b \geq 0$  に対して  $ab \leq a^p/p + b^q/q$  が成り立つ。この式で  $a = |x_i|/\|x\|_p, b = |y_i|/\|y\|_q$  とすると

$$\frac{|x_i y_i|}{\|x\|_p \|y\|_q} \leq \frac{|x_i|^p / \|x\|_p^p}{p} + \frac{|y_i|^q / \|y\|_q^q}{q}$$

これを  $i = 1, \dots, n$  まで足し合わせると

$$\frac{1}{\|\mathbf{x}\|_p \|\mathbf{y}\|_q} \sum_{i=1}^n |x_i y_i| \leq \frac{\|\mathbf{x}\|_p^p / \|\mathbf{x}\|_p^p}{p} + \frac{\|\mathbf{y}\|_q^q / \|\mathbf{y}\|_q^q}{q} = 1/p + 1/q = 1$$

$$\therefore \sum_{i=1}^n |x_i y_i| \leq \|\mathbf{x}\|_p \|\mathbf{y}\|_q$$

□

### III.14.5 Minkowski の不等式 (三角不等式の一般化)

$1 < p < \infty$  のとき

$$\|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p$$

*Proof.*

$$\|\mathbf{x} + \mathbf{y}\|_p^p = \sum_{i=1}^n |x_i + y_i|^p = \sum_{i=1}^n |x_i + y_i| |x_i + y_i|^{p-1} \leq \sum_{i=1}^n |x_i| |x_i + y_i|^{p-1} + \sum_{i=1}^n |y_i| |x_i + y_i|^{p-1} \quad (1)$$

右辺第 1 項は Hölder の不等式より上から

$$\|\mathbf{x}\|_p \|\mathbf{v}\|_{\frac{p}{p-1}} \quad \text{where} \quad \mathbf{v}[i] := |x_i + y_i|^{p-1}$$

で抑えられる。さらに

$$\|\mathbf{x}\|_p \|\mathbf{v}\|_{\frac{p}{p-1}} = \|\mathbf{x}\|_p \left[ \sum_{i=1}^n |x_i + y_i|^{(p-1) \frac{p}{p-1}} \right]^{\frac{p-1}{p}} = \|\mathbf{x}\|_p \|\mathbf{x} + \mathbf{y}\|_p^{p-1}$$

同様に式 (1) の右辺第 2 項も上から  $\|\mathbf{y}\|_p \|\mathbf{x} + \mathbf{y}\|_p^{p-1}$  で抑えられるので

$$\|\mathbf{x} + \mathbf{y}\|_p^p \leq \|\mathbf{x}\|_p \|\mathbf{x} + \mathbf{y}\|_p^{p-1} + \|\mathbf{y}\|_p \|\mathbf{x} + \mathbf{y}\|_p^{p-1}$$

両辺を  $\|\mathbf{x} + \mathbf{y}\|_p^{p-1}$  で割ることで Minkowski の不等式を得る。

□



## 第 III.15 章

# 2 変数関数

### III.15.1 4 象限逆正接 : $\text{Tan}^{-1}(x, y)$

4 象限逆正接  $\text{Tan}^{-1}(x, y)$  を次で定義する。

$$\text{Tan}^{-1}(x, y) := \begin{cases} \text{Tan}^{-1} \frac{y}{x} + \pi & (\text{領域 1 : } x < 0, y > 0) \\ \frac{\pi}{2} & (\text{領域 2 : } x = 0, y > 0) \\ \text{Tan}^{-1} \frac{y}{x} & (\text{領域 3 : } x > 0, y \neq 0) \\ -\frac{\pi}{2} & (\text{領域 4 : } x = 0, y < 0) \\ \text{Tan}^{-1} \frac{y}{x} - \pi & (\text{領域 5 : } x < 0, y < 0) \end{cases}$$

$\text{Tan}^{-1}(x, y)$  は原点及び  $y$  軸の負の領域の領域では定義されない。

この関数は定義域で  $C^1$  級である。

*Proof.*

領域 1, 3, 5 では連続であり、偏微分可能であって偏微分係数はいずれも

$$\begin{aligned} \frac{\partial \text{Tan}^{-1}(x, y)}{\partial x} &= \frac{-y}{x^2 + y^2} \\ \frac{\partial \text{Tan}^{-1}(x, y)}{\partial y} &= \frac{x}{x^2 + y^2} \end{aligned}$$

であり連続だから  $C^1$  級である。

領域 2 における連続性を示す。領域 2 上の点  $(a, b)$  ( $a = 0, b > 0$ ) に対して

$$d = |\text{Tan}^{-1}(a, b) - \text{Tan}^{-1}(a + h, b + k)| = \left| \frac{\pi}{2} - \text{Tan}^{-1}(b + k, h) \right|$$

とおくと、 $k$  を十分小さく取れば  $b + k > 0$  なので

$$d = \begin{cases} \frac{\pi}{2} - \text{Tan}^{-1} \frac{b+k}{h} \rightarrow 0 & \text{as } h \rightarrow +0, k \rightarrow 0 \\ \frac{\pi}{2} - (\text{Tan}^{-1} \frac{b+k}{h} + \pi) & \text{as } h \rightarrow -0, k \rightarrow 0 \end{cases}$$

であり、領域 2 における連続性が示せた。領域 4 における連続性も同様に示せる。

さて、先程の偏微分係数をみると、幸運なことにこれは領域 2, 4 においても連続である。以上より結局、領域 1, 2, 3, 4, 5 全域で  $\text{Tan}^{-1}(x, y)$  は  $C^1$  級である。  $\square$

## 第 III.16 章

# 多変数関数

### III.16.1 Jacobi 行列

#### III.16.1.1 $J_{G \circ F} = J_G J_F$

$F: \mathbb{R}^l \rightarrow \mathbb{R}^m$ ,  $\mathbf{x} \mapsto \mathbf{y}(\mathbf{x})$  と  $G: \mathbb{R}^m \rightarrow \mathbb{R}^n$ ,  $\mathbf{y} \mapsto \mathbf{z}(\mathbf{y})$  の Jacobi 行列を  $J_F, J_G$  とすると、 $F, G$  の合成写像  $G \circ F$  の Jacobi 行列は  $J_{G \circ F} = J_G J_F$  である。

*Proof.*

$$J_{G \circ F}[i][j] = \left( \frac{\partial z_i}{\partial x_j} \right) (\mathbf{x}) = \sum_{k=1}^n \left( \frac{\partial z_i}{\partial y_k} \right) (\mathbf{y}) \left( \frac{\partial y_k}{\partial x_j} \right) (\mathbf{x}) = J_G[i][:] J_F[:,j]$$

□

#### III.16.1.2 $J_{\mathbf{f}^\top \mathbf{g}} = \mathbf{g}^\top J_{\mathbf{f}} + \mathbf{f}^\top J_{\mathbf{g}}$

$\mathbf{f}, \mathbf{g}: \mathbb{R}^m \rightarrow \mathbb{R}^n$  の Jacobi 行列を  $J_{\mathbf{f}}, J_{\mathbf{g}}$  とすると、 $J_{\mathbf{f}^\top \mathbf{g}} = \mathbf{g}^\top J_{\mathbf{f}} + \mathbf{f}^\top J_{\mathbf{g}}$  である。

*Proof.*

$$\begin{aligned} J_{\mathbf{f}^\top \mathbf{g}}[1, j] &= \frac{\partial}{\partial x_j} \sum_{k=1}^n f_k(\mathbf{x}) g_k(\mathbf{x}) = \sum_{k=1}^n \frac{\partial}{\partial x_j} f_k(\mathbf{x}) g_k(\mathbf{x}) = \sum_{k=1}^n \left( \frac{\partial f_k(\mathbf{x})}{\partial x_j} g_k(\mathbf{x}) + f_k(\mathbf{x}) \frac{\partial g_k(\mathbf{x})}{\partial x_j} \right) \\ &= \sum_{k=1}^n (J_{\mathbf{f}}[k, j] g_k(\mathbf{x}) + f_k(\mathbf{x}) J_{\mathbf{g}}[k, j]) = (\mathbf{g}^\top J_{\mathbf{f}})[1, j] + (\mathbf{f}^\top J_{\mathbf{g}})[1, j] \end{aligned}$$

□

### III.16.2 勾配と Hessian

#### III.16.2.1 $\nabla(G \circ F) = \nabla F \nabla G$

$F: \mathbb{R}^l \rightarrow \mathbb{R}^m$ ,  $\mathbf{x} \mapsto \mathbf{y}(\mathbf{x})$  と  $G: \mathbb{R}^m \rightarrow \mathbb{R}^n$ ,  $\mathbf{y} \mapsto \mathbf{z}(\mathbf{y})$  に対して  $F, G$  の合成写像  $G \circ F$  の勾配は  $\nabla(G \circ F) = \nabla F \nabla G$  である。

*Proof.*  $\nabla F = J_F^\top$  であることと III.16.1.1 より直ちに従う。  $\square$

### III.16.2.2 $\nabla_{\mathbf{x}\mathbf{y}} f = (\nabla_{\mathbf{y}\mathbf{x}} f)^\top$

$\mathbf{x} \in \mathbb{R}^{n_1}, \mathbf{y} \in \mathbb{R}^{n_2}$  とし、 $f : (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{n_1+n_2} \mapsto f(\mathbf{x}, \mathbf{y}) \in \mathbb{R}$  とするとき、次が成り立つ。

$$\nabla_{\mathbf{x}\mathbf{y}} f = (\nabla_{\mathbf{y}\mathbf{x}} f)^\top$$

*Proof.*

$$\nabla_{\mathbf{x}\mathbf{y}} f = \nabla_{\mathbf{x}} \nabla_{\mathbf{y}} f = \nabla_{\mathbf{x}} \left[ \frac{\partial f}{\partial y_1}, \dots, \frac{\partial f}{\partial y_{n_2}} \right]^\top = \left[ \frac{\partial^2 f}{\partial x_i \partial y_j} \right]_{n_1 \times n_2} = \left[ \frac{\partial^2 f}{\partial y_i \partial x_j} \right]_{n_2 \times n_1}^\top = (\nabla_{\mathbf{y}\mathbf{x}} f)^\top$$

$\square$

### III.16.2.3 勾配ベクトルが等値面の接平面と直交することの証明

$n$  変数関数  $f(\mathbf{x})$ ,  $(\mathbf{x} := [x_1, \dots, x_n]^\top)$  の点  $\mathbf{a} := [a_1, \dots, a_n]^\top$  における勾配  $\nabla f(\mathbf{a})$  が  $\mathbf{0}$  でないとき、等値面  $f(\mathbf{x}) = f(\mathbf{a})$  上の点  $\mathbf{a}$  における接平面と  $\nabla f(\mathbf{a})$  直交する。

*Proof.*

$\nabla f(\mathbf{a}) \neq \mathbf{0}$  だから  $f_{x_1}(\mathbf{a}), \dots, f_{x_n}(\mathbf{a})$  のうち少なくとも 1 つは 0 でない。一般性を失わずに  $f_{x_n}(\mathbf{a}) \neq 0$  であるとする (変数の添字をそのように付け替えれば良い)。このとき陰関数定理より点  $\mathbf{a}' := [a_1, \dots, a_{n-1}]^\top$  の近傍で  $f(\mathbf{x}) = f(\mathbf{a})$  の陰関数  $x_n = \varphi(\mathbf{x}')$ ,  $(\mathbf{x}' = [x_1, \dots, x_{n-1}]^\top)$  のうち  $\varphi(\mathbf{a}') = a_n$  を満たすものが唯一存在する。 $n-1$  次元超曲面  $x_n = \varphi(\mathbf{x}')$  上の点  $\mathbf{a}'$  における接平面の方程式は

$$x_n = a_n + \sum_{i=1}^{n-1} \varphi_{x_i}(\mathbf{a}')(x_i - a_i) = a_n - \sum_{i=1}^{n-1} \frac{f_{x_i}(\mathbf{a})}{f_{x_n}(\mathbf{a})}(x_i - a_i) =: \psi(\mathbf{x}')$$

であるから、点  $[\mathbf{a}', \varphi(\mathbf{a}') = a_n]^\top$  に対する接平面上の任意の点  $[\mathbf{x}', \psi(\mathbf{x}')]^\top$  の相対位置ベクトルは線形空間

$$W := \text{span} \left[ \left\{ \left[ 0, \dots, 0, \underset{i \text{ 番目}}{1}, 0, \dots, 0, -\frac{f_{x_i}(\mathbf{a}')}{f_{x_n}(\mathbf{a}')} \right]^\top \mid i = 1, \dots, n-1 \right\} \right]$$

のベクトルである。 $\nabla f(\mathbf{a}) = [f_{x_1}(\mathbf{a}'), \dots, f_{x_n}(\mathbf{a}')]^\top$  は  $W$  の任意のベクトルと直交する。  $\square$

### III.16.2.4 $\nabla A\mathbf{x} = A^\top$

*Proof.*  $\nabla(A\mathbf{x})[i][j] = \frac{\partial(A\mathbf{x})[j]}{\partial x_i} = \frac{\partial}{\partial x_i} \sum_{k=1}^n a_{jk}x_k = a_{ji}$   $\square$

### III.16.2.5 $\nabla \mathbf{x}^\top \mathbf{x} = 2\mathbf{x}$

*Proof.*  $\frac{\partial \mathbf{x}^\top \mathbf{x}}{\partial x_i} = \frac{\partial}{\partial x_i} \sum_{j=1}^n x_j^2 = 2x_i \quad \therefore \nabla \mathbf{x}^\top \mathbf{x} = 2\mathbf{x}$   $\square$

$$\text{III.16.2.6} \quad \nabla \mathbf{x}^\top A \mathbf{x} = (A + A^\top) \mathbf{x}$$

*Proof.*

$$\nabla \mathbf{x}^\top A \mathbf{x} = \nabla \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = \nabla \left( \sum_{i=1}^n a_{ii} x_i^2 + \sum_{i \neq j}^n a_{ij} x_i x_j \right)$$

よって第  $i$  要素は

$$(\nabla \mathbf{x}^\top A \mathbf{x}) [i] = 2a_{ii}x_i + \sum_{j \neq i}^n (a_{ij} + a_{ji})x_j = \sum_{j=1}^n (a_{ij} + a_{ji})x_j = ((A + A^\top) \mathbf{x}) [i]$$

□

$$\text{III.16.2.7} \quad \nabla \|A \mathbf{x} + \mathbf{b}\|_2^2 = 2A^\top (A \mathbf{x} + \mathbf{b})$$

*Proof.*

合成関数の微分法より

$$\nabla \|A \mathbf{x} + \mathbf{b}\|_2^2 = (\nabla(A \mathbf{x} + \mathbf{b})) (\nabla \mathbf{y}^\top \mathbf{y}) \big|_{\mathbf{y}=A \mathbf{x} + \mathbf{b}} = A^\top 2(A \mathbf{x} + \mathbf{b}) = 2A^\top (A \mathbf{x} + \mathbf{b})$$

□

$$\text{III.16.2.8} \quad \nabla (A \mathbf{x} + \mathbf{b})^\top B (A \mathbf{x} + \mathbf{b}) = A^\top (B + B^\top) (A \mathbf{x} + \mathbf{b})$$

*Proof.*

合成関数の微分法より

$$\nabla (A \mathbf{x} + \mathbf{b})^\top B (A \mathbf{x} + \mathbf{b}) = \nabla (A \mathbf{x} + \mathbf{b}) \left( \nabla (\mathbf{y}^\top B \mathbf{y}) \big|_{\mathbf{y}=A \mathbf{x} + \mathbf{b}} \right) = A^\top (B + B^\top) (A \mathbf{x} + \mathbf{b})$$

□

$$\text{III.16.2.9} \quad \nabla^2 \mathbf{x}^\top A \mathbf{x} = A + A^\top$$

$$\text{Proof.} \quad \nabla^2 \mathbf{x}^\top A \mathbf{x} = \nabla \nabla \mathbf{x}^\top A \mathbf{x} = \nabla (A + A^\top) \mathbf{x} = A + A^\top$$

□

### III.16.2.10 公式集

- $\nabla^2 \|A \mathbf{x} + \mathbf{b}\|_2^2 = 2A^\top A$
- $\nabla^2 (A \mathbf{x} + \mathbf{b})^\top B (A \mathbf{x} + \mathbf{b}) = A^\top (B + B^\top) A$

### III.16.3 逆関数定理

$F \in C^1(\mathbb{R}^n \rightarrow \mathbb{R}^n)$ ,  $\mathbf{x} \mapsto \mathbf{y}(\mathbf{x})$  の Jacobi 行列を  $J_F(\mathbf{x})$  とする。ある  $\mathbf{a} \in \text{dom}(F)$  に対して  $J_F(\mathbf{a})$  が正則なら  $F(\mathbf{a})$  の近傍で  $C^1$  級の  $F$  の逆写像  $F^{-1}$  が存在して  $J_{F^{-1}}(F(\mathbf{a})) = J_F(\mathbf{a})^{-1}$

*Proof.*

$n = 1$  のときは 1 変数関数の逆関数定理より従う。 $n = m$  のとき成り立つと仮定して  $n = m + 1$  のとき成り立つことを示す。

$F = [f_1, \dots, f_{m+1}]^\top$ ,  $\mathbf{u} = F(\mathbf{a})$ ,  $\hat{\mathbf{x}} := [x_1, \dots, x_m]^\top(\mathbf{a}, \mathbf{u})$  についても同様に定義) とする。

$J_F(\mathbf{a}) \in \mathbb{R}^{(m+1) \times (m+1)}$  は正則なので  $m + 1$  行目の成分の少なくとも 1 つは 0 でない。一般性を失わず第  $m + 1, m + 1$  成分  $\left(\frac{\partial f_{m+1}}{\partial x_{m+1}}\right)(\mathbf{a})$  が 0 でないとする、陰関数定理より関数  $h \in C^1(\mathbb{R}^m \rightarrow \mathbb{R})$  で  $f_{m+1}([\hat{\mathbf{x}}^\top, h(\hat{\mathbf{x}})]^\top) = u_{m+1}$ ,  $h(\hat{\mathbf{a}}) = a_{m+1}$  を満たすものが  $\hat{\mathbf{a}}$  の近傍で唯一存在する。この  $h$  は当然  $F([\hat{\mathbf{a}}^\top, h(\hat{\mathbf{a}})]^\top) = \mathbf{u}$  を満たす。

$\hat{f}_i(\hat{\mathbf{x}}) := f_i([\hat{\mathbf{x}}, h(\hat{\mathbf{x}})]^\top)$ ,  $\hat{F}(\hat{\mathbf{x}}) := [\hat{f}_1(\hat{\mathbf{x}}), \dots, \hat{f}_m(\hat{\mathbf{x}})]^\top$  とすると  $\hat{F}(\hat{\mathbf{a}}) = \hat{\mathbf{u}}$  であり、帰納法の仮定より  $J_{\hat{F}}(\hat{\mathbf{a}})$  が正則であれば  $\hat{\mathbf{u}} (= \hat{F}(\hat{\mathbf{a}}))$  の近傍で  $\hat{F}$  の逆写像  $\hat{F}^{-1} \in C^1(\mathbb{R}^m \rightarrow \mathbb{R}^m)$  が存在して  $\hat{\mathbf{u}}$  に対応する  $\hat{\mathbf{a}}$ 、さらには  $a_{m+1} = h(\hat{\mathbf{a}})$  も唯一に定まる。すなわち  $F^{-1}$  が存在する。そこで  $|J_{\hat{F}}(\hat{\mathbf{a}})| \neq 0$  を示す。列に着目すると

$$J_{\hat{F}}(\hat{\mathbf{a}})[:, j] = \left(\frac{\partial \hat{F}}{\partial x_j}\right)(\hat{\mathbf{a}}) + \left(\frac{\partial \hat{F}}{\partial x_{m+1}}\right)(h(\hat{\mathbf{a}})) \left(\frac{\partial h}{\partial x_j}\right)(\hat{\mathbf{a}}) = \mathbf{p}_j + \lambda_j \mathbf{p}_{m+1}$$

$$\text{where } \mathbf{p}_j := \left(\frac{\partial \hat{F}}{\partial x_j}\right)(\hat{\mathbf{a}}) \ (j = 1, \dots, m+1), \ \lambda_j := \left(\frac{\partial h}{\partial x_j}\right)(\hat{\mathbf{a}}) \ (j = 1, \dots, m)$$

行列式の多重線形性を用いて  $|J_{\hat{F}}(\hat{\mathbf{a}})|$  を計算するが、 $\mathbf{p}_{m+1}$  に比例する列を 2 本以上選んでできる行列はそれらの列が一次従属となり行列式は 0 だから、結局次のようになる。

$$\begin{aligned} |J_{\hat{F}}(\hat{\mathbf{a}})| &= |J_F(\mathbf{a})[1:m][1:m]| + \sum_{i=1}^m \lambda_i \left| \mathbf{p}_1, \dots, \mathbf{p}_{i-1}, \underset{i \text{ 列目}}{\mathbf{p}_{m+1}}, \mathbf{p}_{i+1}, \dots, \mathbf{p}_m \right| \\ &= |J_F(\mathbf{a})[1:m][1:m]| + \sum_{i=1}^m (-1)^{m-i} \lambda_i \left| \mathbf{p}_1, \dots, \mathbf{p}_{i-1}, \mathbf{p}_{i+1}, \dots, \mathbf{p}_m, \mathbf{p}_{m+1} \right| \end{aligned}$$

ここで

$$\lambda_i = - \frac{\left(\frac{\partial f_{m+1}}{\partial x_i}\right)(\mathbf{a})}{\left(\frac{\partial f_{m+1}}{\partial x_{m+1}}\right)(\mathbf{a})}$$

であることと  $(-1)^{m-i} = (-1)^{m+i}$  に注意して

$$\begin{aligned} |J_{\hat{F}}(\hat{\mathbf{a}})| &= \frac{1}{\left(\frac{\partial f_{m+1}}{\partial x_{m+1}}\right)(\mathbf{a})} \left[ \left(\frac{\partial f_{m+1}}{\partial x_{m+1}}\right)(\mathbf{a}) |J_F(\mathbf{a})[1:m][1:m]| \right. \\ &\quad \left. + \sum_{i=1}^m (-1)^{m+i} \left(\frac{\partial f_{m+1}}{\partial x_i}\right)(\mathbf{a}) |\mathbf{p}_1, \dots, \mathbf{p}_{i-1}, \mathbf{p}_{i+1}, \dots, \mathbf{p}_{m+1}| \right] \end{aligned}$$

上式の [ ] 内は  $|J_F(\mathbf{a})|$  の第  $m$  行に関する余因子展開と等しいので

$$|J_{\hat{F}}(\hat{\mathbf{a}})| = \frac{1}{\left(\frac{\partial f_{m+1}}{\partial x_{m+1}}\right)(\mathbf{a})} |J_F(\mathbf{a})| \neq 0$$

以上より  $F$  の逆関数が存在して  $C^1$  級であることが示された。そして  $F^{-1} \circ F$  は恒等写像だから  $I_n = J_{F^{-1} \circ F}(\mathbf{a}) = J_{F^{-1}}(\mathbf{a})J_F(\mathbf{a})$  より  $J_{F^{-1}}(\mathbf{a}) = J_F(\mathbf{a})^{-1}$  となる。  $\square$

## III.16.4 2 次形式

III.16.4.1  $\forall A \in \mathbb{R}^{n \times n}, \mathbf{x} \in \mathbb{R}^n, \exists$  対称行列  $B \in \mathbb{R}^{n \times n}$  s.t.  $\mathbf{x}^\top A \mathbf{x} = \mathbf{x}^\top B \mathbf{x}$

*Proof.*

$$\mathbf{x}^\top A \mathbf{x} = \frac{1}{2} (\mathbf{x}^\top A \mathbf{x} + \mathbf{x}^\top A \mathbf{x}) = \frac{1}{2} (\mathbf{x}^\top A \mathbf{x} + (\mathbf{x}^\top A \mathbf{x})^\top) = \frac{1}{2} (\mathbf{x}^\top A \mathbf{x} + \mathbf{x}^\top A^\top \mathbf{x}) = \mathbf{x}^\top \frac{A + A^\top}{2} \mathbf{x}$$

であるので  $B := \frac{A + A^\top}{2}$  とすれば  $B$  は対称行列であり

$$\mathbf{x}^\top A \mathbf{x} = \mathbf{x}^\top B \mathbf{x}$$

$\square$

## 第 III.17 章

# 測度

### III.17.1 $\sigma$ -加法族

#### III.17.1.1 2つの $\sigma$ -加法族の和集合は $\sigma$ -加法族になるとは限らない

例えば  $S = \{1, 2, 3\}$ ,  $X = \{\emptyset, \{1\}, \{2, 3\}, S\}$ ,  $Y = \{\emptyset, \{1, 2\}, \{3\}, S\}$  とすると  $X, Y$  は  $S$  上の  $\sigma$ -加法族であり、 $X \cup Y = \{\emptyset, \{1\}, \{3\}, \{2, 3\}, \{1, 2\}, S\}$  である。 $\{1\}, \{3\} \subset X \cup Y$  だが  $\{1\} \cup \{3\} = \{1, 3\} \notin X \cup Y$  なので  $X \cup Y$  は  $\sigma$ -加法族ではない。

#### III.17.1.2 $\sigma$ -加法族の要素同士の直積集合の族は $\sigma$ -加法族になるとは限らない

例えば  $S = \{1, 2, 3\}$ ,  $X = \{\emptyset, \{1\}, \{2, 3\}, S\}$ ,  $Y = \{\emptyset, \{1, 2\}, \{3\}, S\}$  とすると  $X, Y$  は  $S$  上の  $\sigma$ -加法族であり、その要素同士の直積集合の集合族を  $Z$  とすると  $Z = \{\emptyset, \{1\} \times \{1, 2\}, \{1, 3\}, \{1\} \times S, \{2, 3\} \times \{1, 2\}, \{2, 3\} \times \{3\}, \{2, 3\} \times S, S \times \{1, 2\}, S \times \{3\}, S \times S\}$  である。 $\{2, 3\} \times \{1, 2\} \in Z$  であるが  $(\{2, 3\} \times \{1, 2\})^c = \{(1, 1), (1, 2), (1, 3), (2, 3), (3, 3)\} \notin Z$  なので  $Z$  は  $\sigma$ -加法族ではない。

#### III.17.1.3 $X$ を集合とする。 $A, B \subset 2^X$ , $A \subset B \Rightarrow \sigma(A) \subset \sigma(B)$

*Proof.*

まず、 $\sigma(\cdot)$  の定義より  $B \subset \sigma(B)$  であり、これと  $A \subset B$  より  $A \subset \sigma(B)$  である。よって

$$\sigma(A) = A \text{ を含む全ての } \sigma\text{-加法族の共通部分} \subset \sigma(B) \text{ (} \supset B \supset A \text{)}$$

□

### III.17.2 Jordan 測度

いくつか記号の約束を決めておく。

- 閉直方体  $E$  上の有界関数  $f$  と、 $E$  の分割  $\Delta$  に対して過剰和と不足和を

$$S(f, \Delta) := \sum_{r \in \Delta} \sup_{\mathbf{x} \in r} f(\mathbf{x}) |r|, \quad s(f, \Delta) := \sum_{r \in \Delta} \inf_{\mathbf{x} \in r} f(\mathbf{x}) |r|$$

と定義する。

- 有界集合  $A$  の Jordan 外測度, 内測度, 測度をそれぞれ  $\overline{m}_J(A)$ ,  $\underline{m}_J(A)$ ,  $m_J(A)$  と書く。

### III.17.2.1 Jordan 測度の有限加法性

$A_1, A_2$  が Jordan 可測ならば  $A_1 \cap A_2$ ,  $A_1 \cup A_2$  も可測で  $m_J(A_1 \cup A_2) = m_J(A_1) + m_J(A_2) - m_J(A_1 \cap A_2)$

*Proof.*

$A_1, A_2$  を包含する直方体  $E$  を考える。 $A_1, A_2$  が Jordan 可測だから任意の正数  $\varepsilon$  に対して  $E$  のある分割  $\Delta$  が存在して次式を満たす。

$$S(\mathbb{1}_{A_1}(\cdot), \Delta) - s(\mathbb{1}_{A_1}(\cdot), \Delta), S(\mathbb{1}_{A_2}(\cdot), \Delta) - s(\mathbb{1}_{A_2}(\cdot), \Delta) < \varepsilon/2$$

$\partial(A_1 \cap A_2), \partial(A_1 \cup A_2) \subseteq A_1, A_2$  だから

$$\begin{aligned} & S(\mathbb{1}_{A_1 \cap A_2}(\cdot), \Delta) - s(\mathbb{1}_{A_1 \cap A_2}(\cdot), \Delta), \quad S(\mathbb{1}_{A_1 \cup A_2}(\cdot), \Delta) - s(\mathbb{1}_{A_1 \cup A_2}(\cdot), \Delta) \\ & \leq S(\mathbb{1}_{A_1}(\cdot), \Delta) - s(\mathbb{1}_{A_1}(\cdot), \Delta) + S(\mathbb{1}_{A_2}(\cdot), \Delta) - s(\mathbb{1}_{A_2}(\cdot), \Delta) < \varepsilon \end{aligned}$$

( $S$  と  $s$  の差は境界で発生することに注意すれば上の式に納得できる。) より  $A_1 \cap A_2$ ,  $A_1 \cup A_2$  も可測である。ここで

$$\begin{aligned} S(\mathbb{1}_{A_1}(\cdot), \Delta) + S(\mathbb{1}_{A_2}(\cdot), \Delta) - S(\mathbb{1}_{A_1 \cap A_2}(\cdot), \Delta) & \geq s(\mathbb{1}_{A_1 \cup A_2}(\cdot), \Delta) \\ & \geq s(\mathbb{1}_{A_1}(\cdot), \Delta) + s(\mathbb{1}_{A_2}(\cdot), \Delta) - s(\mathbb{1}_{A_1 \cap A_2}(\cdot), \Delta) \end{aligned}$$

が成り立っており、これとダルブーの定理より次式が従う。

$$m_J(A_1 \cup A_2) = m_J(A_1) + m_J(A_2) - m_J(A_1 \cap A_2)$$

□

### III.17.2.2 $A \subset E$ ( $E$ は直方体) のとき $\underline{m}_J(A) = |E| - \overline{m}_J(A^c \cap E)$

*Proof.*

外測度と内測度の定義より任意の正数  $\varepsilon$  に対して  $E$  のある分割  $\Delta$  が存在して  $\underline{m}_J(A) - s(\mathbb{1}_A(\cdot), \Delta), S(\mathbb{1}_{A^c \cap E}(\cdot), \Delta) - \overline{m}_J(A^c \cap E) < \varepsilon/2$  であることと  $s(\mathbb{1}_A(\cdot), \Delta) = |E| - S(\mathbb{1}_{A^c \cap E}(\cdot), \Delta)$  より従う。 □

### III.17.2.3 系: $A \subset E$ ( $E$ は直方体) が Jordan 可測のとき $A^c \cap E$ も Jordan 可測である

*Proof.*

直前の定理より

$$\underline{m}_J(A^c \cap E) = |E| - \overline{m}_J((A^c \cap E)^c \cap E) = |E| - \overline{m}_J(A) = |E| - \underline{m}_J(A) = \overline{m}_J(A^c \cap E)$$

□



### III.17.3 Lebesgue 測度

いくつか記号の約束を決めておく。

- 有界集合  $A$  の Lebesgue 外測度, 測度を  $\overline{m}_L(A), m_L(A)$  と書く。

#### III.17.3.1 $A \subset B$ ならば $\overline{m}_L(A) \leq \overline{m}_L(B)$

*Proof.*

背理法で示す。仮に  $\overline{m}_L(A) > \overline{m}_L(B)$  であるとする。任意の  $0 < \varepsilon < \overline{m}_L(A) - \overline{m}_L(B)$  に対してある直方体領域の集合  $R$  が存在して  $B \subset R$ ,  $|R| < \overline{m}_L(B) + \varepsilon$  となる。 $A \subset B \subset R$  となるので  $\overline{m}_L(A) \leq |R| < \overline{m}_L(B) + \varepsilon < \overline{m}_L(A)$  となり矛盾。□

#### III.17.3.2 $\overline{m}_L(A \cup B) \leq \overline{m}_L(A) + \overline{m}_L(B)$

*Proof.*

背理法で示す。仮に  $\overline{m}_L(A \cup B) > \overline{m}_L(A) + \overline{m}_L(B)$  とする。任意の  $0 < \varepsilon < (\overline{m}_L(A \cup B) - (\overline{m}_L(A) + \overline{m}_L(B)))/2$  に対してある直方体領域の集合  $R_1, R_2$  が存在して  $A \subset R_1, B \subset R_2$ ,  $|R_1| < \overline{m}_L(A) + \varepsilon$ ,  $|R_2| < \overline{m}_L(B) + \varepsilon$  となる。 $A \cup B \subset R_1 \cup R_2$  となるので  $\overline{m}_L(A \cup B) \leq |R_1 \cup R_2| \leq |R_1| + |R_2| < \overline{m}_L(A) + \overline{m}_L(B) + 2\varepsilon < \overline{m}_L(A \cup B)$  となり矛盾。□

#### III.17.3.3 Lebesgue 外測度の劣加法性: $\overline{m}_L(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \overline{m}_L(A_n)$

*Proof.*

背理法で示す。仮に  $\overline{m}_L(\bigcup_{n=1}^{\infty} A_n) > \sum_{n=1}^{\infty} \overline{m}_L(A_n)$  とし、その差を  $d > 0$  とおく。数列  $\{a_n\}$ ,  $a_n = \frac{d}{3} 2^{-n+1}$  ( $n = 1, 2, \dots$ ) とおくと級数  $\sum_{n=1}^{\infty} a_n = 2d/3$  となる。任意の  $n \in \mathbb{N}$  に対してある直方体領域  $R_n$  が存在して  $A_n \subset R_n$ ,  $|R_n| < \overline{m}_L(A_n) + a_n$  となる。 $\bigcup_{n=1}^{\infty} A_n \subset \bigcup_{n=1}^{\infty} R_n$  となるので  $\overline{m}_L(\bigcup_{n=1}^{\infty} A_n) \leq |\bigcup_{n=1}^{\infty} R_n| \leq \sum_{n=1}^{\infty} |R_n| = \sum_{n=1}^{\infty} \overline{m}_L(A_n) + \sum_{n=1}^{\infty} a_n < \sum_{n=1}^{\infty} \overline{m}_L(A_n) + d < \overline{m}_L(\bigcup_{n=1}^{\infty} A_n)$  となり矛盾。□

#### III.17.3.4 有界集合 $A$ に対して $\overline{m}_L(A) \leq \overline{m}_J(A)$

*Proof.*

背理法で示す。仮に  $\overline{m}_L(A) > \overline{m}_J(A)$  とし、 $0 < \varepsilon < \overline{m}_L(A) - \overline{m}_J(A)$  とする。 $A$  を包含する直方体領域を  $E$  とする。Jordan 外測度の定義より  $E$  のある分割  $\Delta$  が存在して  $S(\mathbb{1}_A(\cdot), \Delta) < \overline{m}_J(A) + \varepsilon$  となる。 $A \subset \bigcup_{r \in \Delta, r \cap A \neq \emptyset} r$  なので  $\overline{m}_L(A) \leq \sum_{r \in \Delta, r \cap A \neq \emptyset} |r| = S(\mathbb{1}_A(\cdot), \Delta) < \overline{m}_J(A) + \varepsilon < \overline{m}_L(A)$  となり矛盾。□

### III.17.3.5 有界閉集合 $A \subset \mathbb{R}^n$ に対して $\overline{m}_L(A) = 0 \iff m_J(A) = 0$

*Proof.*

$\Leftarrow$  は直前の定理より明らか。 $\Rightarrow$  を示す。 $\overline{m}_L(A) = 0$  より、任意の  $\varepsilon > 0$  に対してある直方体集合  $R = \bigcup_{n=1}^{\infty} E_n$  が存在して  $A \subset R$ ,  $|R| < \varepsilon/2$  を満たす。任意の直方体  $r$  に対して、その中心を保ったまま  $k$  倍に相似拡大した直方体を  $kr$  と書くことにする。 $\bigcup_{n=1}^{\infty} \text{Int}(2E_n)$  は  $A$  の開被覆であり、 $A$  は  $\mathbb{R}^n$  の有界閉集合なのでコンパクトであるから、ある  $N \in \mathbb{N}$  が存在して  $A \subset \bigcup_{n=1}^N \text{Int}(2E_n)$  となる。 $\bigcup_{n=1}^N \text{Int}(2E_n)$  を包含する十分大きな直方体  $B$  を考え、 $2E_n$  ( $n = 1, \dots, N$ ) を非交差的になるように適当に分割した直方体群を全て要素に持つような  $B$  の分割  $\Delta$  に対して  $S(A, \Delta) \leq 2|R| < \varepsilon$  となるから  $\overline{m}_J(A) < \varepsilon$  となる。  $\square$

### III.17.3.6 $A \subset \mathbb{R}^n$ とする。任意の直方体 $E$ に対して $\overline{m}_L(A \cap E) + \overline{m}_L(A^c \cap E) = \overline{m}_L(E) \iff$ 任意の $B \subset \mathbb{R}^n$ に対して $\overline{m}_L(A \cap B) + \overline{m}_L(A^c \cap B) = \overline{m}_L(B)$

*Proof.*

$\Leftarrow$  は明らか。 $\Rightarrow$  を示す。 $(A \cap B) \cup (A^c \cap B) = B$  なのでまず  $\overline{m}_L(A \cap B) + \overline{m}_L(A^c \cap B) \geq \overline{m}_L(B)$  が成り立つ。 $\overline{m}_L(A \cap B) + \overline{m}_L(A^c \cap B) > \overline{m}_L(B)$  と仮定して矛盾を導く。 $0 < \varepsilon < \overline{m}_L(A \cap B) + \overline{m}_L(A^c \cap B) - \overline{m}_L(B)$  とすると、ある直方体集合  $R$  が存在して  $B \subset R$ ,  $|R| < \overline{m}_L(B) + \varepsilon$  を満たす。 $B \subset R$  なので  $A \cap B \subset A \cap R$ ,  $A^c \cap B \subset A^c \cap R$  だから  $\overline{m}_L(A \cap B) + \overline{m}_L(A^c \cap B) \leq \overline{m}_L(A \cap R) + \overline{m}_L(A^c \cap R) = \overline{m}_L(R) = |R|$  となる。これと  $|R| < \overline{m}_L(B) + \varepsilon < \overline{m}_L(A \cap B) + \overline{m}_L(A^c \cap B)$  より  $|R| < |R|$  となり矛盾。  $\square$

### III.17.3.7 有界集合 $A \subset \mathbb{R}^n$ が Jordan 可測ならば Lebesgue 可測であり、 $m_L(A) = m_J(A)$

*Proof.*

$A$  を包含する任意の直方体  $R$  に対して

$$\begin{aligned} \overline{m}_L(A) &\geq |R| - \overline{m}_L(A^c \cap R) \geq |R| - \overline{m}_J(A^c \cap R) \quad (\because \text{III.17.3.4}) \\ &= |R| - m_J(A^c \cap R) = m_J(A) \quad (\because \text{III.17.2.3}) \end{aligned}$$

これと III.17.3.4 ( $\overline{m}_L(A) \leq \overline{m}_J(A) = m_J(A)$ ) より  $\overline{m}_L(A) = m_J(A)$ 。さらに

$$\overline{m}_L(A) \leq \overline{m}_J(A) = m_J(A) = |R| - m_J(A^c \cap R) \leq |R| - \overline{m}_L(A^c \cap R)$$

と  $\overline{m}_L(A) \geq |R| - \overline{m}_L(A^c \cap R)$  より  $\overline{m}_L(A) = |R| - \overline{m}_L(A^c \cap R)$  なので  $A$  は Lebesgue 可測である。以上より  $m_L(A) = \overline{m}_L(A) = m_J(A)$   $\square$

### III.17.3.8 $A_1, A_2 \subset \mathbb{R}^n$ が Lebesgue 可測なら $A_1 \cup A_2$ も Lebesgue 可測

*Proof.*

$B \in \mathbb{R}^n$  を任意の集合とする。 $\overline{m}_L((A_1 \cup A_2) \cap B) + \overline{m}_L((A_1 \cup A_2)^c \cap B) = \overline{m}_L(B)$  を示せばよい。 $\geq$  は Lebesgue 外測度の劣加法性より成り立つので  $\leq$  を示す。

$$\begin{aligned}
& \overline{m}_L((A_1 \cup A_2) \cap B) + \overline{m}_L((A_1 \cup A_2)^c \cap B) \\
&= \overline{m}_L((A_1 \cap B) \cup (A_2 \cap A_1^c \cap B)) + \overline{m}_L(A_2^c \cap A_1^c \cap B) \\
&\leq \overline{m}_L(A_1 \cap B) + \overline{m}_L(A_2 \cap A_1^c \cap B) + \overline{m}_L(A_2^c \cap A_1^c \cap B) \\
&= \overline{m}_L(A_1 \cap B) + \overline{m}_L(A_1^c \cap B) \quad (\because A_2 \text{ は可測}) \\
&= \overline{m}_L(B) \quad (\because A_1 \text{ は可測})
\end{aligned}$$

□

III.17.3.9 系:  $A_1, \dots, A_n \subset \mathbb{R}^n$  が Lebesgue 可測なら  $A_1 \cup A_2 \cup \dots \cup A_n$  も Lebesgue 可測

*Proof.* 直前の定理を繰り返し用いる。

□

III.17.3.10 Lebesgue 可測集合  $A_1, A_2 \subset \mathbb{R}^n$  が非交差的なら  $m_L(A_1 \sqcup A_2) = m_L(A_1) + m_L(A_2)$

*Proof.*

直前の定理の証明の式の 3 行目第 2 項で  $A_2 \cap A_1^c \cap B = A_2 \cap B$  となるので

$$\begin{aligned}
\overline{m}_L(B) &\leq \overline{m}_L((A_1 \sqcup A_2) \cap B) + \overline{m}_L((A_1 \sqcup A_2)^c \cap B) \\
&\leq \overline{m}_L(A_1 \cap B) + \overline{m}_L(A_2 \cap B) + \overline{m}_L((A_1 \sqcup A_2)^c \cap B) \leq \overline{m}_L(B)
\end{aligned}$$

すなわち

$$\overline{m}_L((A_1 \sqcup A_2) \cap B) = \overline{m}_L(A_1 \cap B) + \overline{m}_L(A_2 \cap B)$$

$B$  は任意なので  $B = \mathbb{R}^n$  とすると  $\overline{m}_L(A_1 \sqcup A_2) = \overline{m}_L(A_1) + \overline{m}_L(A_2)$  となるが、 $A_1, A_2, A_1 \sqcup A_2$  が可測であるので外測度=測度であるから結局

$$m_L(A_1 \sqcup A_2) = m_L(A_1) + m_L(A_2)$$

□

III.17.3.11 系: Lebesgue 測度の有限加法性

Lebesgue 可測集合  $A_1, \dots, A_n \in \mathbb{R}^n$  が非交差的なら  $m_L(\bigsqcup_{i=1}^n A_i) = \sum_{i=1}^n m_L(A_i)$

*Proof.* 直前の定理を繰り返し用いる。

□

III.17.3.12 Lebesgue 測度の  $\sigma$ -加法性

Lebesgue 可測集合  $A_1, A_2, \dots \in \mathbb{R}^n$  が非交差的なら  $m_L(\bigsqcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} m_L(A_i)$

*Proof.*

$B_n := \bigsqcup_{i=1}^n A_i$ ,  $B := \bigsqcup_{i=1}^{\infty} A_i$  とする。任意の集合  $S \in \mathbb{R}^n$  に対して

$$\overline{m}_L(S \cap B) + \overline{m}_L(S \cap B^c) = \overline{m}_L(S)$$

となることを示す。 $\geq$  は Lebesgue 外測度の劣加法性より成り立つので  $\leq$  を示す。Lebesgue 測度の有限加法性より

$$\begin{aligned} \overline{m}_L(S) &= \overline{m}_L(S \cap B_n) + \overline{m}_L(S \cap B_n^c) \\ &\geq \overline{m}_L(S \cap B_n) + \overline{m}_L(S \cap B^c) \quad (\because S \cap B^c \subset S \cap B_n^c) \\ &= \sum_{i=1}^n \overline{m}_L(S \cap A_i) + \overline{m}_L(S \cap B^c) \end{aligned}$$

上式で  $n \rightarrow \infty$  として

$$\begin{aligned} \overline{m}_L(S) &\geq \sum_{i=1}^{\infty} \overline{m}_L(S \cap A_i) + \overline{m}_L(S \cap B^c) \\ &\geq \overline{m}_L(S \cap B) + \overline{m}_L(S \cap B^c) \quad (\because \text{Lebesgue 外測度の劣加法性}) \\ &\geq \overline{m}_L(S) \end{aligned}$$

よって  $B$  は Lebesgue 可測であり  $\overline{m}_L(S \cap B) = \sum_{i=1}^{\infty} \overline{m}_L(S \cap A_i)$  である。 $S = \mathbb{R}^n$  とすると  $m_L(B) = m_L(S \cap B) = \overline{m}_L(S \cap B) = \sum_{i=1}^{\infty} \overline{m}_L(S \cap A_i) = \sum_{i=1}^{\infty} m_L(S \cap A_i) = \sum_{i=1}^{\infty} m_L(A_i)$   $\square$

### III.17.3.13 Lebesgue 可測の直感的理解

集合  $A$  に対する直方体領域を加えたり切り取ったりする (切り取るという操作は、元の直方体領域を別の複数の直方体領域の和に変換することと等しいことに注意) **可算無限回** の近似操作により作られる近似集合  $B_n$  が存在して

$$\forall x \in A, \exists N(x) \in \mathbb{N} \text{ s.t. } n \geq N(x), x \in B_n$$

$$\forall y \notin A, \exists N(y) \in \mathbb{N} \text{ s.t. } n \geq N(y), y \notin B_n$$

が成り立つとき、 $A$  は Lebesgue 可測であるのだと思う。理解が深まったら書き直すかもしれない。

### III.17.3.14 Jordan 測度と Lebesgue 測度の違い

Lebesgue 外測度をもつ劣加法性に対応する性質が Jordan 外測度にはない ( $\mathbb{Q} \cap [0, 1]$  を考えてみればわかる)。Jordan 外測度では有限の分割に対して指示関数の  $\sup$  をとってから、分割数  $\rightarrow \infty$  とすることしか許されていないのに対し、Lebesgue 外測度では個々の小直方体の体積の評価と、小直方体の増産を同時に実行することを許されている。この点が Jordan 測度と Lebesgue 測度の違いを生んでいると思う。

### III.17.4 包除原理

測度空間  $(S, \mathcal{M}, \mu)$  を考える。任意の  $A_1, \dots, A_n \in \mathcal{M}$  に対して次の等式が成り立つ。

$$\begin{aligned} \mu(A_1 \cup \dots \cup A_n) &= \mu(A_1) + \dots + \mu(A_n) \\ &\quad - \mu(A_1 \cap A_2) - \mu(A_1 \cap A_3) - \dots - \mu(A_{n-1} \cap A_n) \\ &\quad + \mu(A_1 \cap A_2 \cap A_3) + \mu(A_1 \cap A_2 \cap A_4) + \dots + \mu(A_{n-2} \cap A_{n-1} \cap A_n) \\ &\quad - \dots + (-1)^{n-1} \mu(A_1 \cap \dots \cap A_n) \end{aligned}$$

すなわち

$$\mu\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k-1} \sum_{I \in \{1, \dots, n\}^{\mathcal{C}_k}} \mu\left(\bigcap_{i \in I} A_i\right) \quad (1)$$

もっと短く書けば

$$\mu\left(\bigcup_{i=1}^n A_i\right) = \sum_{I \in 2^{\{1, \dots, n\}} \setminus \{\emptyset\}} (-1)^{|I|-1} \mu\left(\bigcap_{i \in I} A_i\right)$$

*Proof.*

式 (1) を証明する。

$$(1) \text{ の左辺} = \int_S \mathbb{1}_U(x) \mu(dx)$$

$$\begin{aligned} (1) \text{ の右辺} &= \sum_{k=1}^n (-1)^{k-1} \sum_{I \in \{1, \dots, n\}^{\mathcal{C}_k}} \int_S \mathbb{1}_{\bigcap_{i \in I} A_i}(x) \mu(dx) \\ &= \int_S \left[ \sum_{k=1}^n (-1)^{k-1} \sum_{I \in \{1, \dots, n\}^{\mathcal{C}_k}} \mathbb{1}_{\bigcap_{i \in I} A_i}(x) \right] \mu(dx) \\ &= \int_S c(x) \mu(dx) \quad \text{where } c(x) := \sum_{k=1}^n (-1)^{k-1} \sum_{I \in \{1, \dots, n\}^{\mathcal{C}_k}} \mathbb{1}_{\bigcap_{i \in I} A_i}(x) \end{aligned}$$

$U := \bigcup_{i=1}^n A_i$  とすると  $c(x) = \mathbb{1}_U(x)$  となることを示せばよい。まず  $x \notin U \Rightarrow c(x) = 0$  は明らか。各  $x \in U$  に対して、 $x$  に依存して決まる次のような自然数  $m(x)$  が存在する。

「 $x$  は  $m(x)$  個の集合  $A_{i_1}, \dots, A_{i_{m(x)}}$  に共通して属し、他の  $n - m(x)$  個の集合には属さない。」

従って、 $c(x)$  の  $\sum_{I \in \{1, \dots, n\}^{\mathcal{C}_k}} \mathbb{1}_{\bigcap_{i \in I} A_i}(x)$  は  $k > m(x) + 1$  に対して 0 であるから  $c(x)$  は次のように書き直される。

$$c(x) = \sum_{k=1}^{m(x)} (-1)^{k-1} \sum_{I \in \{1, \dots, n\}^{\mathcal{C}_k}} \mathbb{1}_{\bigcap_{i \in I} A_i}(x)$$

上式の  $\mathbb{1}_{\bigcap_{i \in I} A_i}(x)$  が 1 になるのは上述の  $m(x)$  個の集合から  $k$  個の集合を選んできた時だけであるから、

$c(x)$  はさらに次のように書き直される。

$$\begin{aligned}
c(x) &= \sum_{k=1}^{m(x)} (-1)^{k-1} \sum_{I \in \{i_1, \dots, i_{m(x)}\}^{\mathbf{C}_k}} 1 = \sum_{k=1}^{m(x)} (-1)^{k-1} \mathbf{C}_{m(x)} \mathbf{C}_k \\
&= - \sum_{k=1}^{m(x)} (-1)^k \mathbf{C}_{m(x)} \mathbf{C}_k = - \left( \sum_{k=0}^{m(x)} (-1)^k \mathbf{C}_{m(x)} \mathbf{C}_k - 1 \right) \\
&= - \left( (1-1)^{m(x)} - 1 \right) = 1
\end{aligned}$$

よって確かに  $c(x) = \mathbb{1}_U(x)$  である。

□

## 第 III.18 章

# 積分

### III.18.1 変数変換 : $\int_{x(t_0)}^{x(t)} f(x)dx = \int_{t_0}^t f(x)x'(t)dt$

$t$  の関数  $x(t)$  とその関数  $f(x)$  について

$$\int_{x(t_0)}^{x(t)} f(x)dx = \int_{t_0}^t f(x)x'(t)dt$$

*Proof.*

$t = 0$  について

$$\int_{x(t_0)}^{x(t)} f(x)dx = \int_{x(t_0)}^{x(t_0)} f(x)dx = 0$$

$$\int_{t_0}^t f(x)x'(t)dt = \int_{t_0}^{t_0} f(x)x'(t)dt = 0$$

より両者は一致する。次に

$$\frac{d}{dt} \int_{x(t_0)}^{x(t)} f(x)dx = \left( \frac{d}{dx} \int_{x(t_0)}^{x(t)} f(x)dx \right) \frac{dx}{dt} = f(x)x'(t) = \frac{d}{dt} \int_{t_0}^t f(x)x'(t)dt$$

より両者の微分係数は一致する。以上より両者は等しい。  $\square$

### III.18.2 Chebyshev の和の不等式 : 可積分な単調増加関数 $f, g$ に対して $\left( \int_0^1 f(x)dx \right) \left( \int_0^1 g(x)dx \right) \leq \int_0^1 f(x)g(x)dx$

*Proof.* 積分区間  $[0, 1]$  の  $n$  等分割のリーマン和に対して Chebyshev の和の不等式 (III.5.3) を適用すればよい。  $\square$

## 第 III.19 章

# 重積分

### III.19.1 平均値の定理

$\Omega \in \mathbb{R}^n$  を体積確定の有界閉領域、 $f : \Omega \rightarrow \mathbb{R}$  を連続関数とすると、 $\exists \mathbf{x}'$  s.t.  $\int_{\Omega} f(\mathbf{x}) d\mathbf{x} = f(\mathbf{x}') |\Omega|$

*Proof.*

$\Omega$  内での  $f$  の最小値と最大値を各々  $m(\Omega), M(\Omega)$  とすると

$$m(\Omega) |\Omega| \leq \int_{\Omega} f(\mathbf{x}) d\mathbf{x} \leq M(\Omega) |\Omega| \quad \therefore m(\Omega) \leq \frac{1}{|\Omega|} \int_{\Omega} f(\mathbf{x}) d\mathbf{x} \leq M(\Omega)$$

であるから中辺の値を  $I(\Omega)$  とすれば中間値の定理より  $\exists \mathbf{x}'$  s.t.  $I(\Omega) = f(\mathbf{x}')$  であり、定理が従う。  $\square$

### III.19.2 $C^1$ 級の変換による面積 0 の集合の像の面積もまた 0

閉超直方体  $R \subset \mathbb{R}^n$  を含む開集合  $D$  上で定義された  $C^1$  級の変換  $F : D \rightarrow \mathbb{R}^n$  を考える。集合  $C \subset \mathbb{R}^n$  の面積が 0 であるとき、 $F(C) \subset \mathbb{R}^n$  の面積もまた 0 である。

*Proof.*

$F$  は  $R$  で  $C^1$  級なので特に Lipschitz 連続である。Lipschitz 定数を  $K$  とする。 $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$  の  $m$  等分割  $\Delta$  を考える。第  $d$  軸方向の分割を  $\Delta_d : x_i^{(d)} = a_d + \frac{i}{m}(b_d - a_d)$  ( $i = 0, \dots, m, d = 1, \dots, n$ ) とし、各小超直方体を  $R_{i_1, \dots, i_n} := [x_{i_1-1}^{(1)}, x_{i_1}^{(1)}] \times \cdots \times [x_{i_n-1}^{(n)}, x_{i_n}^{(n)}]$  で表す。 $R$  の対角線の長さを  $L = \sqrt{\sum_{d=1}^n (b_d - a_d)^2}$  としておく。 $C$  は面積確定で面積が 0 であるから、任意の  $\varepsilon > 0$  に対してある  $M(\varepsilon) \in \mathbb{N}$  が存在し、 $m \geq M(\varepsilon)$  なる任意の分割に対して次式を満たす。

$$\frac{|R|}{(2KL)^n} \varepsilon > \sum_{R_{i_1, \dots, i_n} \cap C \neq \emptyset} |R_{i_1, \dots, i_n}| = \sum_{R_{i_1, \dots, i_n} \cap C \neq \emptyset} \frac{|R|}{m^n}$$

各  $R_{i_1, \dots, i_n}$  の対角線の長さは  $L/m$  であり、 $F$  が Lipschitz 連続だから各  $R_{i_1, \dots, i_n}$  について、その中の任意の 1 点の  $F$  による像 (点) の半径  $KL/m$  以内に他の全ての点の像が含まれる。つまり各  $F(R_{i_1, \dots, i_n})$  は 1 辺の長さが  $2KL/m$  の  $n$  次元超立方体  $R'_{i_1, \dots, i_n}$  で覆うことができる。よって  $F(C)$  は  $\{R'_{i_1, \dots, i_n} \mid R_{i_1, \dots, i_n} \cap C \neq \emptyset\}$  で覆うことができ、その面積は

$$\sum_{R_{i_1, \dots, i_n} \cap C \neq \emptyset} |R'_{i_1, \dots, i_n}| \leq \sum_{R_{i_1, \dots, i_n} \cap C \neq \emptyset} (2KL/m)^n \leq \frac{(2KL)^n}{|R|} \sum_{R_{i_1, \dots, i_n} \cap C \neq \emptyset} \frac{|R|}{m^n} < \varepsilon$$



□

III.19.3 関数  $f, g$  が  $\mathbb{R}$  上で有界, 連続で  $\int_{-\infty}^{\infty} f(x)dx, \int_{-\infty}^{\infty} g(x)dx$  が絶対収束するなら畳み込み  $f * g$  は  $\mathbb{R}$  上で連続で絶対可積分である。

*Proof.*

(連続であること):

$\varepsilon > 0, x \in \mathbb{R}, h \in (-1, 1)$  を任意にとる。

$$\begin{aligned} |(f * g)(x+h) - (f * g)(x)| &= \left| \int_{-\infty}^{\infty} [f(x-y+h) - f(x-y)] g(y) dy \right| \\ &\leq \int_{-\infty}^{\infty} |[f(x-y+h) - f(x-y)] g(y)| dy \end{aligned}$$

$f$  の有界性と  $g$  の絶対可積分性より、ある  $a$  が存在して

$$\int_{-\infty}^a |[f(x-y+h) - f(x-y)] g(y)| dy, \int_a^{\infty} |[f(x-y+h) - f(x-y)] g(y)| dy < \varepsilon/3$$

となる。 $f$  は  $\mathbb{R}$  上で連続だから特に  $[x-a-1, x+a+1]$  上で一様連続である。よって  $h$  を十分小さくすれば

$$|f(x-y+h) - f(x-y)| < \frac{\varepsilon}{3L} \quad \left( L := \int_{-\infty}^{\infty} |g(y)| dy \right)$$

となり、

$$\int_{-a}^a |[f(x-y+h) - f(x-y)] g(y)| dy < \varepsilon/3$$

以上より

$$|(f * g)(x+h) - (f * g)(x)| < \varepsilon$$

(絶対可積分であること):

$a > 0$  として  $I(a) := \int_{-a}^a |(f * g)(x)| dx$  とすると  $I(a)$  は  $a$  に関して単調増加であり、次式が成り立つ。

$$I(a) = \int_{-a}^a \left| \int_{-\infty}^{\infty} f(x-y) g(y) dy \right| dx \leq \int_{-a}^a \int_{-\infty}^{\infty} |f(x-y)| |g(y)| dy dx = \int_{-\infty}^{\infty} \int_{-a}^a |f(x-y)| |g(y)| dx dy$$

右端の等号は  $f$  の有界性と  $g$  の絶対可積分性より  $\int_{-\infty}^{\infty} |f(x-y)| |g(y)| dy$  が  $x$  の値に関係なく収束することによる。これより

$$\begin{aligned} I(a) &\leq \int_{-\infty}^{\infty} |g(y)| \int_{-a}^a |f(x-y)| dx dy \leq \int_{-\infty}^{\infty} |g(y)| M_1 dy \quad \left( M_1 := \int_{-\infty}^{\infty} |f(x-y)| dx \right) \\ &\leq M_1 M_2 \quad \left( M_2 := \int_{-\infty}^{\infty} |g(y)| dy \right) \end{aligned}$$

よって  $I(a)$  は単調増加かつ上に有界なので収束する。つまり  $\int_{-\infty}^{\infty} (f * g)(x) dx$  は絶対収束する。 □

### III.19.4 諸公式

III.19.4.1  $n$  次元の単体の体積  $\int_{x_1=0}^a \int_{x_2=0}^{a-x_1} \cdots \int_{x_n=0}^{a-x_1-\cdots-x_{n-2}} dx_n \cdots dx_2 dx_1 = a^n/n! \ (a > 0)$

*Proof.*

$$\begin{aligned}
 & \int_{x_1=0}^a \int_{x_2=0}^{a-x_1} \cdots \int_{x_n=0}^{a-x_1-\cdots-x_{n-2}} dx_n \cdots dx_2 dx_1 \\
 &= \int_{x_1=0}^a \int_{x_2=0}^{a-x_1} \cdots \int_{x_{n-1}=0}^{a-x_1-\cdots-x_{n-3}} (a-x_1-\cdots-x_{n-1}) dx_{n-1} \cdots dx_2 dx_1 \\
 &= \frac{1}{2} \int_{x_1=0}^a \int_{x_2=0}^{a-x_1} \cdots \int_{x_{n-2}=0}^{a-x_1-\cdots-x_{n-4}} (a-x_1-\cdots-x_{n-2})^2 dx_{n-2} \cdots dx_2 dx_1 \\
 &= \cdots \\
 &= \frac{1}{2 \times 3 \times \cdots (n-1)} \int_{x_1=0}^a (a-x_1)^{n-1} dx_1 \\
 &= \frac{a^n}{n!}
 \end{aligned}$$

□

## 第 III.20 章

# 線積分

### III.20.1 定義

$\mathbb{R}^n$  上の有界関数  $f$  と区分的に滑らかな有向曲線  $C$  に対して  $x_d$  ( $d = 1, \dots, n$ ) 軸方向の線積分を以下のようにして定義する。

$C$  上にその向きに沿って  $n+1$  個の分点 始点  $= P_0, P_1, \dots, P_n =$  終点 をとる。添字の意味で隣接する 2 点  $P_{i-1}, P_i$  は重ならないようにする。さらに添字の順番が曲線の向きと逆行しないようにする。曲線が自己交差する点では複数の点が重なることがあるが、それは構わない。これらの点による  $C$  の分割を  $\Delta$  と表し、その「幅」を  $|\Delta| := \max_{i \in \{1, \dots, n\}} \|P_{i-1}P_i\|_2$  で定義する。 $P_{i-1}$  から  $P_i$  までの部分曲線から任意に点  $\xi_i$  をとり、また、 $\overrightarrow{P_{i-1}P_i}$  の  $x_d$  軸成分を  $(\Delta x_d)_i$  とする。級数  $S(f, \Delta, \{\xi_i\}) := \sum_{i=1}^n f(\xi_i)(\Delta x_d)_i$  を考え、これを  $f$  の  $\Delta, \{\xi_i\}$  に関する **Riemann 和** と呼ぶ。分割を任意のやり方で細かくして  $|\Delta| \rightarrow 0$  としたときに  $S(f, \Delta, \{\xi_i\})$  が収束するとき、これを  $\int_C f(\mathbf{x}) dx_d$  と表し、 $x_d$  方向の線積分という。

### III.20.2 計算法

$f$  が  $C$  に関して線積分可能であるとし、パラメータ  $t$  を用いて  $\mathbf{x} = \phi(t) = [\phi_1(t), \dots, \phi_n(t)]^\top$ ,  $\phi \in C^1(\mathbb{R} \rightarrow \mathbb{R}^n)$ ,  $t: a \rightarrow b$  と表せる場合は次が成り立つ。

$$\int_C f(\mathbf{x}) dx_d = \int_a^b f(\phi(t)) \phi'_d(t) dt$$

その理由を以下で説明する。

諸々の定義は「定義」で述べたものを引き続き用いる。 $t$  を  $a$  から始めて  $b$  まで増加させると  $P_0, P_1, \dots$  と順に分点に出会う。出会った順に対応する  $t$  を  $t_0, t_1, \dots$  と決めてゆく。曲線が自己交差する点では複数の点が重なっていることがあるが、このときは添字の若い順にとる。つまり最初の通過で一番若い添字のものが対応し、次の通過で次に若い添字を対応させる。平均値の定理より、各  $i$  に対して適当な  $\mu_i \in (t_{i-1}, t_i)$  が存在して  $(\Delta x_d)_i = \phi_d(t_i) - \phi_d(t_{i-1}) = \phi'_d(\mu_i)(t_i - t_{i-1})$  を満たす。級数  $S^{(2)} := \sum_{i=1}^n f(\phi(\mu_i)) \phi'_d(\mu_i)(t_i - t_{i-1})$  を考えると、これもまた  $f$  の  $\Delta, \{\phi(\mu_i)\}$  に関する Riemann 和であるから  $f$  の線積分可能性より、分割を細かくして  $|\Delta| \rightarrow 0$  とするとき  $\int_C f(\mathbf{x}) dx_d$  に収束する。また、 $S^{(2)}$  を  $t$  についての 1 変数関数  $f(\phi(t)) \phi'_d(t)$  の Riemann 和と見れば、これは分割を細かくするときに  $\int_a^b f(\phi(t)) \phi'_d(t) dt$  に収束する。

曲線を別のパラメータ表示  $\mathbf{x} = \psi(u)$ ,  $u: \alpha \rightarrow \beta$  で表しても、線積分の値は変わらない。なぜなら、既に

見たように線積分はパラメータ表示に依存せずに定義されているからである。

### III.20.3 Green の定理の拡張

$\Omega \subset \mathbb{R}^2$  とし、その境界  $\partial\Omega$  は区分的に滑らかであるとする。 $f, u$  は  $\Omega$  上で定義された  $C^1$  級の関数であるとする。このとき、次式が成り立つ。

$$\int_{\Omega} \frac{\partial f}{\partial x_1} u d\Omega = \int_{\partial\Omega} f u dx_2 - \int_{\Omega} f \frac{\partial u}{\partial x_1} d\Omega \quad (1)$$

$$\int_{\Omega} \frac{\partial f}{\partial x_2} u d\Omega = - \int_{\partial\Omega} f u dx_1 - \int_{\Omega} f \frac{\partial u}{\partial x_2} d\Omega \quad (2)$$

*Proof.*

式 (2) を示す。式 (1) は同様に示せる。

一般性を失わずに  $\partial\Omega$  は下図のように  $x_1$  に関して単純な2つの部分に分けられる形状であるとする。

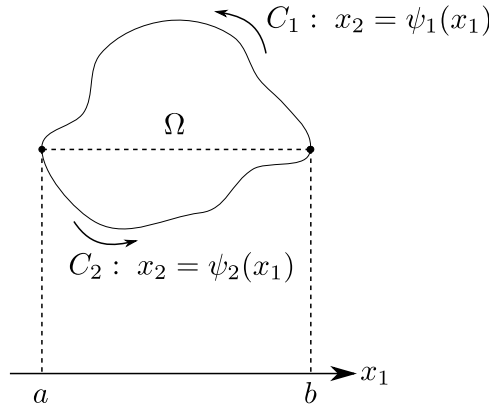


図 III.20.1 積分領域

なぜなら、そうでない場合は、条件を満たすように複数の閉曲線に分割して其々に対して以下の証明を適用して合算すればよいからである。このように積分しても、2つの閉曲線同士で共有された境界上での積分は相殺して0となるから、 $\partial\Omega$  で線積分した結果と一致する。

$$\begin{aligned} \int_{\Omega} \frac{\partial f}{\partial x_2} u d\Omega &= \int_{x_1=a}^b \int_{x_2=\psi_2(x_1)}^{\psi_1(x_1)} \frac{\partial f}{\partial x_2} u dx_2 dx_1 = \int_{x_1=a}^b \left\{ [fu]_{x_2=\psi_2(x_1)}^{\psi_1(x_1)} - \int_{x_2=\psi_2(x_1)}^{\psi_1(x_1)} f \frac{\partial u}{\partial x_2} dx_2 \right\} dx_1 \\ &= \int_{x_1=a}^b [f(x_1, \psi_1(x_1))u(x_1, \psi_1(x_1)) - f(x_1, \psi_2(x_1))u(x_1, \psi_2(x_1))] dx_1 \\ &\quad - \int_{x_1=a}^b \int_{x_2=\psi_2(x_1)}^{\psi_1(x_1)} f \frac{\partial u}{\partial x_2} dx_2 dx_1 \\ &= - \int_{C_1} f u dx_1 - \int_{C_2} f u dx_1 - \int_{\Omega} f \frac{\partial u}{\partial x_2} d\Omega \\ &= - \int_{\partial\Omega} f u dx_1 - \int_{\Omega} f \frac{\partial u}{\partial x_2} d\Omega \end{aligned}$$

□

## 第 III.21 章

# 変数変換

### III.21.1 三角関数

$$\text{III.21.1.1 } x, y \in \mathbb{R}, x^2 + y^2 = 1 \iff x = \cos \theta, y = \sin \theta$$

$$x^2 + y^2 = 1 \iff x = \cos \theta, y = \sin \theta$$

$x = \cos \theta, y = -\sin \theta$  とか  $x = -\cos \theta, y = \sin \theta$  とか  $x = -\cos \theta, y = -\sin \theta$  とか  
 $x = \sin \theta, y = \cos \theta$  とか  $x = \sin \theta, y = -\cos \theta$  とか  $x = -\sin \theta, y = -\cos \theta$  とかは全部同値である。

*Proof.*

$\theta$  に上手く定数を足せばどれも表現できる。 □

### III.21.2 極座標表示

III.21.2.1 極方程式から曲線を導くときの注意 :  $r, \theta$  空間の 2 つの異なる集合から  $x, y$  空間へ写像した 2 つの集合の排他的論理和が  $\emptyset$  なら  $x, y$  空間において両者の違いは無い

言われてみれば当たり前だと思えることだが、計算問題を解いている最中は意外と戸惑うことが多い。レムニスケート  $(x^2 + y^2)^2 = x^2 - y^2$  を例に考えてみる。極座標変換  $x = r \cos \theta, y = r \sin \theta$  を施せば  $(x^2 + y^2)^2 = x^2 - y^2 \iff r \geq 0 \wedge r^4 = r^2 \cos 2\theta \iff r \geq 0 \wedge r^2(r^2 - \cos 2\theta) = 0$  となる。多くの参考書ではこの後いきなり  $r = \sqrt{\cos 2\theta}$  としているが、その裏には次に述べる論理が隠れている。

$r \geq 0 \wedge r^2(r^2 - \cos 2\theta) = 0$  を満たす  $r, \theta$  空間の集合を考える。極座標変換だから  $\theta$  は適当な  $2\pi$  区間の部分集合を選べば良い。よって  $r \geq 0 \wedge r^2(r^2 - \cos 2\theta) = 0$  を満たす集合は  $S = \{(r, \theta) \mid r = 0, \theta \in [-\pi, \pi]\} \cup \{(r, \theta) \mid \theta \in [-\pi, \pi], r = \sqrt{\cos 2\theta}\}$  である。しかし  $\{(r, \theta) \mid r = 0, \theta \in [-\pi, \pi]\}$  の  $xy$  空間での像は原点であり、 $\{(r, \theta) \mid \theta \in [-\pi, \pi], r = \sqrt{\cos 2\theta}\}$  の像にも原点が含まれる。そこで  $S_2 := \{(r, \theta) \mid \theta \in [-\pi, \pi], r = \sqrt{\cos 2\theta}\}$  とすると、 $S, S_2$  それぞれを  $x, y$  空間に写像してできる 2 つの曲線の排他的論理和は  $\emptyset$  である。だから曲線を表現するという目的では  $r = \sqrt{\cos 2\theta}$  で十分なのである。

## 第 III.22 章

# Beta 関数

### III.22.1 Beta 関数と Gamma 関数の関係

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

*Proof.*

$$\Gamma(p)\Gamma(q) = \left( \int_0^\infty x^{p-1} e^{-x} dx \right) \left( \int_0^\infty y^{q-1} e^{-y} dy \right) = \int_{x=0}^\infty \int_{y=0}^\infty x^{p-1} y^{q-1} e^{-x-y} dy dx$$

積分経路を変更する。線分  $(0, z) \rightarrow (z, 0)$  に沿った積分を  $z : 0 \rightarrow \infty$  まで足し合わせる。上式の  $y$  は  $z - x$  に変わり

$$\int_{z=0}^\infty \left( \int_{x=0}^z x^{p-1} (z-x)^{q-1} e^{-x+x-z} dx \right) dz = \int_{z=0}^\infty e^{-z} \left[ \int_{x=0}^z x^{p-1} (z-x)^{q-1} dx \right] dz$$

(詳しく言えば、 $x = u, y = v - u$  なる変数変換を行ったのと等しい。ヤコビアンは 1 となる。)

$x = uz$  なる変数変換を行って

$$\begin{aligned} \int_{z=0}^\infty e^{-z} \left[ \int_{u=0}^1 z^{p-1} u^{p-1} z^{q-1} (1-u)^{q-1} z du \right] dz &= \int_{z=0}^\infty e^{-z} z^{p+q-1} \left[ \int_{u=0}^1 u^{p-1} (1-u)^{q-1} du \right] dz \\ &= \int_0^\infty e^{-z} z^{p+q-1} B(p, q) dz = B(p, q) \int_0^\infty e^{-z} z^{p+q-1} dz = B(p, q) \Gamma(p+q) \end{aligned}$$

□

### III.22.2 多変量 Beta 関数

$\alpha = [\alpha_1, \dots, \alpha_n]^\top > \mathbf{0}$  に対して多変量 Beta 関数を次の広義積分で定義する。

$$\begin{aligned} B(\alpha) &:= \int_{\Omega} \prod_{i=1}^{n-1} x_i^{\alpha_i-1} \left( 1 - \sum_{i=1}^{n-1} x_i \right)^{\alpha_n-1} d\Omega, \quad \Omega := \left\{ [x_1, \dots, x_{n-1}]^\top \mid x_i \geq 0, \sum_{i=1}^{n-1} x_i \leq 1 \right\} \\ &= \int_{x_1=0}^1 \int_{x_2=0}^{1-x_1} \cdots \int_{x_{n-1}=0}^{1-\sum_{i=1}^{n-1} x_i} \prod_{i=1}^{n-1} x_i^{\alpha_i-1} \left( 1 - \sum_{i=1}^{n-1} x_i \right)^{\alpha_n-1} dx_{n-1} \cdots dx_2 dx_1 \end{aligned}$$

### III.22.3 多変量 Beta 関数と Gamma 関数の関係

$$B(\boldsymbol{\alpha}) = \frac{\prod_{i=1}^n \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^n \alpha_i)}$$

*Proof.*

$$\prod_{i=1}^n \Gamma(\alpha_i) = \prod_{i=1}^n \int_0^\infty x_i^{\alpha_i-1} e^{-x_i} dx_i = \int_{x_1=0}^\infty \cdots \int_{x_n=0}^\infty \prod_{i=1}^n x_i^{\alpha_i-1} \exp\left(-\sum_{i=1}^n x_i\right) dx_n \cdots dx_1$$

ここで

$$\mathbf{y} = P\mathbf{x}, \quad P := \begin{bmatrix} \mathbf{e}_1^\top \\ \mathbf{e}_2^\top \\ \vdots \\ \mathbf{e}_{n-1}^\top \\ \mathbf{1}_n^\top \end{bmatrix}$$

なる変数変換を行うと、 $x_n = y_n - \sum_{i=1}^{n-1} x_i = y_n - \sum_{i=1}^{n-1} y_i$ ,  $|P| = 1$  なので

$$\begin{aligned} \prod_{i=1}^n \Gamma(\alpha_i) &= \int_{y_n=0}^\infty \int_{y_1=0}^{y_n} \cdots \int_{y_{n-1}=0}^{y_n - \sum_{i=1}^{n-2} y_i} \prod_{i=1}^{n-1} y_i^{\alpha_i-1} \left(y_n - \sum_{i=1}^{n-1} y_i\right)^{\alpha_n-1} e^{-y_n} |P|^{-1} dy_{n-1} \cdots dy_1 dy_n \\ &= \int_{y_n=0}^\infty e^{-y_n} \int_{y_1=0}^{y_n} \cdots \int_{y_{n-1}=0}^{y_n - \sum_{i=1}^{n-2} y_i} \prod_{i=1}^{n-1} y_i^{\alpha_i-1} \left(y_n - \sum_{i=1}^{n-1} y_i\right)^{\alpha_n-1} dy_{n-1} \cdots dy_1 dy_n \end{aligned}$$

ここで  $y_i = z_i y_n$  ( $i = 1, \dots, n-1$ ) と変数変換すると

$$\begin{aligned} \prod_{i=1}^n \Gamma(\alpha_i) &= \int_{y_n=0}^\infty y_n^{(\sum_{i=1}^n \alpha_i - 1)} e^{-y_n} dy_n \int_{z_1=0}^1 \cdots \int_{z_{n-1}=0}^{1 - \sum_{i=1}^{n-2} z_i} \prod_{i=1}^{n-1} z_i^{\alpha_i-1} \left(1 - \sum_{i=1}^{n-1} z_i\right)^{\alpha_n-1} dz_{n-1} \cdots dz_1 \\ &= \Gamma\left(\sum_{i=1}^n \alpha_i\right) B(\boldsymbol{\alpha}) \end{aligned}$$

□

## 第 III.23 章

# 冪級数関数

### III.23.1 2 次以上の項が $o(x)$ ( $x \rightarrow 0$ ) であること

冪級数関数  $f(x) = \sum_{i=0}^{\infty} a_i x^i$  の収束半径が  $r > 0$  であるとする。  $x$  の 2 次以上の項は  $o(x)$  ( $x \rightarrow 0$ ) である。 よって  $a_1 \neq 0$  ならば、 十分小さい  $x \neq 0$  に対して  $f(x)$  の 2 次以上の項は 1 次までの項に対して真に小さい。

*Proof.*

$|x| < r$  とする。 このとき  $|x| < u < r$  なる  $u$  が存在する。 収束半径の定義より  $f(u)$  は絶対収束するから  $\lim_{i \rightarrow \infty} |a_i u^i|$  なので  $\forall i, |a_i u^i| < M$  なる  $M$  が存在する。  $f(x)$  の 2 次以上の項は

$$\left| \sum_{i=2}^{\infty} a_i x^i \right| \leq \sum_{i=2}^{\infty} |a_i| |x|^i < \sum_{i=2}^{\infty} \frac{M}{u^i} |x|^i = M \frac{|x|^2/u^2}{1 - |x|/u}$$

であるから  $x \neq 0$  ならば

$$\left| \sum_{i=2}^{\infty} a_i x^i \right| / |x| < \frac{M}{u} \frac{|x|/u}{1 - |x|/u} \rightarrow 0 \text{ (as } x \rightarrow 0 \text{)}$$

□



## 第 III.24 章

# 微分と積分の関係

### III.24.1 $\lim_{x \rightarrow \infty} f(x)$ が存在すれば $\int_0^\infty f'(x)dx = f(\infty) - f(0)$

$C^1$  級関数  $f(x)$  について  $\lim_{x \rightarrow \infty} f(x)$  が存在するとき  $f'(x)$  は  $0 \sim \infty$  で可積分である。すなわち  $\int_0^\infty f'(x)dx$  が存在する。

*Proof.*

$f(\infty) := \lim_{x \rightarrow \infty} f(x)$  とすると任意の  $\epsilon > 0$  に対してある正数  $X(\epsilon)$  が存在して  $x \geq X \Rightarrow |f(x) - f(\infty)| < \epsilon$  であるから、 $x \geq X$  とすると

$$\begin{aligned} |[f(x) - f(0)] - [f(\infty) - f(0)]| &< \epsilon \\ \therefore \left| \int_0^x f'(u)du - [f(\infty) - f(0)] \right| &< \epsilon \end{aligned}$$

$\epsilon \rightarrow 0$  のとき  $X(\epsilon) \rightarrow \infty$  であるから結局

$$\int_0^\infty f'(x)dx = f(\infty) - f(0)$$

□

### III.24.2 $\frac{d}{dx} \int_a^x f(x)dx = f(x)$

連続関数  $f(x)$  について

$$\frac{d}{dx} \int_a^x f(x)dx = f(x)$$

*Proof.*

$$\frac{d}{dx} \int_a^x f(x)dx = \lim_{h \rightarrow 0} \frac{1}{h} \left( \int_a^{x+h} f(x)dx - \int_a^x f(x)dx \right) = \lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x+h} f(x)dx \quad (1)$$

$h > 0$  の場合を示す。 $h < 0$  の場合も同じ要領で示せる。積分の平均値の定理よりある  $\xi \in (x, x+h)$  が存在して

$$hf(\xi) = \int_x^{x+h} f(x)dx$$

となるから

$$(1) = \lim_{h \rightarrow 0} f(\xi) = f(x)$$

□

### III.24.3 畳み込みの微分

$$\frac{d}{dx} \int_{s=0}^x f(s)g(x-s)ds = f(x)g(0) + \int_{s=0}^x f(s) \left. \frac{dg(x)}{dx} \right|_{x=x-s} ds$$

*Proof.*

$I(x) := \int_0^x f(s)g(x-s)ds$  とおき、微分の定義に従い  $\lim_{\Delta x \rightarrow 0} \frac{I(x+\Delta x) - I(x)}{\Delta x}$  を計算する。

$$\begin{aligned} I(x+\Delta x) - I(x) &= \left( \int_0^x + \int_x^{x+\Delta x} \right) f(s)g(x+\Delta x-s)ds - \int_{s=0}^x f(s)g(x-s)ds \\ &= \int_0^x f(s)(g(x-s+\Delta x) - g(x-s))ds + \int_x^{x+\Delta x} f(s)g(x+\Delta x-s)ds \end{aligned}$$

積分の平均値の定理より  $\exists \theta \in (x, x+\Delta x)$  s.t.  $\int_x^{x+\Delta x} f(s)g(x+\Delta x-s)ds = \Delta x f(\theta)g(x+\Delta x-\theta)$  であるから

$$\begin{aligned} \lim_{\Delta x \rightarrow 0} \frac{I(x+\Delta x) - I(x)}{\Delta x} &= \lim_{\Delta x \rightarrow 0} \left[ f(\theta)g(x+\Delta x-\theta) + \int_0^x f(s) \frac{g(x-s+\Delta x) - g(x-s)}{\Delta x} ds \right] \\ &= f(x)g(0) + \int_{s=0}^x f(s) \left. \frac{dg(x)}{dx} \right|_{x=x-s} ds \end{aligned}$$

□

### III.24.4 重積分の微分

$$\begin{aligned} &\frac{d}{dx} \left( \int_{x_1=0}^x \cdots \int_{x_n=0}^{x-x_1-\cdots-x_{n-1}} f_1(x_1) \cdots f_n(x_n) dx_n \cdots dx_1 \right) \\ &= \int_{x_1=0}^x \cdots \int_{x_{n-1}=0}^{x-x_1-\cdots-x_{n-2}} f_1(x_1) \cdots f_{n-1}(x_{n-1}) f_n(x-x_1-\cdots-x_{n-1}) dx_{n-1} \cdots dx_1 \end{aligned}$$

$\omega > 0$  とする。

$$\Omega := \{(x_1, \dots, x_n) | x_1, \dots, x_n \geq 0, x_1 + \cdots + x_n \leq \omega\}$$

$$\Omega^+ := \{(x_1, \dots, x_n) | x_1, \dots, x_n \geq 0, x_1 + \cdots + x_n = \omega\} \subset \Omega$$

$f_1(x), \dots, f_n(x)$  を  $\{x | x \geq 0\}$  上の連続関数とし、 $f(\mathbf{x}) := f_1(x_1) \cdots f_n(x_n)$ ,  $F(\omega) := \int_{\Omega} f(\mathbf{x}) d\mathbf{x}$  とするとき、次が成り立つ。

$$\frac{dF(\omega)}{d\omega} = \int_{\Omega^+} f(\mathbf{x}) d\mathbf{x}$$

要するに、領域  $x_1, \dots, x_n \geq 0$  の超平面  $x_1 + \cdots + x_n = \omega$  の原点側の部分における  $f_1(x_1) \cdots f_n(x_n)$  の積分を  $\omega$  で微分したものは領域  $x_1, \dots, x_n \geq 0$  と超平面の共通部分上での積分に等しい。

この定理は 2 つの独立な非負の確率変数の和の確率密度関数を計算する時に役立つ。

*Proof.*

微分の定義に従い  $\lim_{\Delta\omega \rightarrow 0} \frac{F(\omega+\Delta\omega)-F(\omega)}{\Delta\omega}$  を計算する。

$$\begin{aligned} F(\omega + \Delta\omega) &= \int_{x_1=0}^{\omega} \cdots \int_{x_n=0}^{\omega-x_1-\cdots-x_{n-1}} f_1(x_1) \cdots f_n(x_n) dx_n \cdots dx_1 \\ &= \left( \int_{x_1=0}^{\omega} + \int_{x_1=\omega}^{\omega+\Delta\omega} \right) f_1(x_1) \left( \int_{x_2=0}^{\omega-x_1} + \int_{x_2=\omega-x_1}^{\omega-x_1+\Delta\omega} \right) f_2(x_2) \cdots \\ &\quad \left( \int_{x_n=0}^{\omega-x_1-\cdots-x_{n-1}} + \int_{x_n=\omega-x_1-\cdots-x_{n-1}}^{\omega-x_1-\cdots-x_{n-1}+\Delta\omega} \right) f_n(x_n) dx_n \cdots dx_2 dx_1 \end{aligned}$$

であり、括弧の中身を全部展開すれば  $F(\omega)$  を含む  $2^n$  個の項が出る。 $\frac{F(\omega+\Delta\omega)-F(\omega)}{\Delta\omega}$  を考えるときは  $F(\omega)$  は相殺するから放っておけば良い。 $F(\omega)$  を除いた  $2^n - 1$  項を  $\Delta\omega$  で除して  $\Delta\omega \rightarrow 0$  としたとき、ある 1 項を除いて他は 0 になり、残った項が  $\int_{\Omega^+} f(x) dx$  になることを以下で示していく。

$$g_1(\omega, x_1) := \int_{x_2=0}^{\omega-x_1+\Delta\omega} f_2(x_2) \cdots \int_{x_n=0}^{\omega-x_1-\cdots-x_{n-1}+\Delta\omega} f_n(x_n) dx_n \cdots dx_2$$

とおけば

$$F(\omega + \Delta\omega) = \underbrace{\int_{x_1=0}^{\omega} f_1(x_1) g_1(\omega, x_1) dx_1}_{=: h_{11}(\omega, \Delta\omega)} + \underbrace{\int_{x_1=\omega}^{\omega+\Delta\omega} f_1(x_1) g_1(\omega, x_1) dx_1}_{=: h_{12}(\omega, \Delta\omega)}$$

となる。 $\lim_{\Delta\omega \rightarrow 0} \frac{h_{12}(\omega, \Delta\omega)}{\Delta\omega} = 0$  を示す。積分の平均値の定理から  $\exists \theta_1 \in (\omega, \omega + \Delta\omega)$  s.t.  $h_{12}(\omega, \Delta\omega) = \Delta\omega f_1(\theta_1) g_1(\omega, \theta_1)$  であるので

$$\lim_{\Delta\omega \rightarrow 0} \frac{h_{12}(\omega, \Delta\omega)}{\Delta\omega} = \lim_{\Delta\omega \rightarrow 0} f_1(\theta_1) g_1(\omega, \theta_1) = f_1(\omega) \lim_{\Delta\omega \rightarrow 0} g_1(\omega, \theta_1) = 0$$

結局  $h_{11}(\omega, \Delta\omega)$  だけ考えれば良い。

$$g_2(\omega, x_1, x_2) := \int_{x_3=0}^{\omega-x_1-x_2+\Delta\omega} f_3(x_3) \cdots \int_{x_n=0}^{\omega-x_1-\cdots-x_{n-1}+\Delta\omega} f_n(x_n) dx_n \cdots dx_3$$

とおけば

$$\begin{aligned} h_{11}(\omega, \Delta\omega) &= \underbrace{\int_{x_1=0}^{\omega} f_1(x_1) \int_{x_2=0}^{\omega-x_1} f_2(x_2) g_2(\omega, x_1, x_2) dx_2 dx_1}_{=: h_{21}(\omega, \Delta\omega)} \\ &\quad + \underbrace{\int_{x_1=0}^{\omega} f_1(x_1) \int_{x_2=\omega-x_1}^{\omega-x_1+\Delta\omega} f_2(x_2) g_2(\omega, x_1, x_2) dx_2 dx_1}_{=: h_{22}(\omega, \Delta\omega)} \end{aligned}$$

となる。 $\lim_{\Delta\omega \rightarrow 0} \frac{h_{22}(\omega, \Delta\omega)}{\Delta\omega} = 0$  を示す。積分の平均値の定理から  $\exists \theta_2 \in (\omega - x_1, \omega - x_1 + \Delta\omega)$  s.t.  $\int_{x_2=\omega-x_1}^{\omega-x_1+\Delta\omega} f_2(x_2) g_2(\omega, x_1, x_2) dx_2 = \Delta\omega f_2(\theta_2) g_2(\omega, x_1, \theta_2)$  であるので

$$\lim_{\Delta\omega \rightarrow 0} \frac{h_{22}(\omega, \Delta\omega)}{\Delta\omega} = \lim_{\Delta\omega \rightarrow 0} \int_{x_1=0}^{\omega} f_1(x_1) f_2(\theta_2) g_2(\omega, x_1, \theta_2) dx_1$$

これと  $\lim_{\Delta\omega \rightarrow 0} g_2(\omega, x_1, \theta_2) = 0$  であり、かつ  $\int_{x_1=0}^{\omega} f_1(x_1) f_2(\theta_2) dx_1$  が有界であることから

$$\lim_{\Delta\omega \rightarrow 0} \frac{h_{22}(\omega, \Delta\omega)}{\Delta\omega} = 0$$

となるので、結局  $h_{21}(\omega, \Delta\omega)$  だけを考えれば良い。同様の議論を繰り返すと、結局  $h_{ij}(\omega, \Delta\omega)$  ( $(i, j) \in \{1, \dots, n-1\} \times \{1, 2\}$ ) に対して  $\lim_{\Delta\omega \rightarrow 0} \frac{h_{ij}(\omega, \Delta\omega)}{\Delta\omega} = 0$  であり、残る項は

$$\begin{aligned}
& h_{n-1,1}(\omega, \Delta\omega) \\
&= \int_{x_1=0}^{\omega} f_1(x_1) \cdots \int_{x_{n-1}=0}^{\omega-x_1-\cdots-x_{n-2}} f_{n-1}(x_{n-1}) \\
&\quad \left( \int_{x_n=0}^{\omega-x_1-\cdots-x_{n-1}} + \int_{x_n=\omega-x_1-\cdots-x_{n-1}}^{\omega-x_1-\cdots-x_{n-1}+\Delta\omega} \right) f_n(x_n) dx_n dx_{n-1} \cdots dx_1 \\
&= F(\omega) + \underbrace{\int_{x_1=0}^{\omega} f_1(x_1) \cdots \int_{x_{n-1}=0}^{\omega-x_1-\cdots-x_{n-2}} f_{n-1}(x_{n-1}) \int_{x_n=\omega-x_1-\cdots-x_{n-1}}^{\omega-x_1-\cdots-x_{n-1}+\Delta\omega} f_n(x_n) dx_n dx_{n-1} \cdots dx_1}_{=: h_{n2}(\omega, \Delta\omega)}
\end{aligned}$$

だけである。積分の平均値の定理から  $\exists \theta_n \in (\omega - x_1 - \cdots - x_{n-1}, \omega - x_1 - \cdots - x_{n-1} + \Delta\omega)$  s.t.  $\int_{x_n=\omega-x_1-\cdots-x_{n-1}}^{\omega-x_1-\cdots-x_{n-1}+\Delta\omega} f_n(x_n) dx_n = \Delta\omega f_n(\theta_n)$  であるので

$$\begin{aligned}
\lim_{\Delta\omega \rightarrow 0} \frac{h_{n2}(\omega, \Delta\omega)}{\Delta\omega} &= \int_{x_1=0}^{\omega} f_1(x_1) \cdots \int_{x_{n-1}=0}^{\omega-x_1-\cdots-x_{n-2}} f_{n-1}(x_{n-1}) f_n(\theta_n) dx_{n-1} \cdots dx_1 \\
&= \int_{x_1=0}^{\omega} f_1(x_1) \cdots \int_{x_{n-1}=0}^{\omega-x_1-\cdots-x_{n-2}} f_{n-1}(x_{n-1}) f_n(\omega - x_1 - \cdots - x_{n-1}) dx_{n-1} \cdots dx_1 \\
&= \int_{\overline{\Omega}^+} f(\mathbf{x}) d\mathbf{x}
\end{aligned}$$

以上より

$$\frac{dF(\omega)}{d\omega} = \lim_{\Delta\omega \rightarrow 0} \frac{F(\omega + \Delta\omega) - F(\omega)}{\Delta\omega} = \lim_{\Delta\omega \rightarrow 0} \frac{h_{n2}(\omega, \Delta\omega)}{\Delta\omega} = \int_{\overline{\Omega}^+} f(\mathbf{x}) d\mathbf{x}$$

□

### III.24.5 系

$$\begin{aligned}
& \frac{d}{d\omega} \left( \int_{x_1=-\infty}^{\infty} \cdots \int_{x_n=-\infty}^{x-x_1-\cdots-x_{n-1}} f_1(x_1) \cdots f_n(x_n) dx_n \cdots dx_1 \right) \\
&= \int_{x_1=-\infty}^{\infty} \cdots \int_{x_{n-1}=-\infty}^{x-x_1-\cdots-x_{n-2}} f_1(x_1) \cdots f_{n-1}(x_{n-1}) f_n(x - x_1 - \cdots - x_{n-1}) dx_{n-1} \cdots dx_1
\end{aligned}$$

$$\Omega := \{(x_1, \dots, x_n) | x_1 + \cdots + x_n \leq \omega\}$$

$$\overline{\Omega}^+ := \{(x_1, \dots, x_n) | x_1 + \cdots + x_n = \omega\} \subset \Omega$$

$f_1(x), \dots, f_n(x)$  を  $\mathbb{R}$  上の連続関数とし、 $f(\mathbf{x}) := f_1(x_1) \cdots f_n(x_n)$ ,  $F(\omega) := \int_{\Omega} f(\mathbf{x}) d\mathbf{x}$  とするとき、次が成り立つ。

$$\frac{dF(\omega)}{d\omega} = \int_{\overline{\Omega}^+} f(\mathbf{x}) d\mathbf{x}$$

要するに、超平面  $x_1 + \cdots + x_n = \omega$  の負領域における  $f_1(x_1) \cdots f_n(x_n)$  の積分を  $\omega$  で微分したものは超平面上での積分に等しい。

この定理は 2 つの独立な確率変数の和の確率密度関数を計算する時に役立つ。

*Proof.*

直前の定理の証明を少し改造するだけで簡単に証明できる。

□

## 第 III.25 章

# 図形への応用

### III.25.1 ある超平面から距離 $d$ だけ離れた超平面の方程式

$n \in \mathbb{N}$ ,  $\mathbf{w} \in \mathbb{R}^n \neq \mathbf{0}$ ,  $c \in \mathbb{R}$  とする。写像  $f: \mathbf{x} \in \mathbb{R}^n \mapsto \mathbf{w}^\top \mathbf{x} + c \in \mathbb{R}$  が定める  $\mathbb{R}^n$  上の超平面  $f(\mathbf{x}) = 0$  から距離  $d > 0$  だけ離れた 2 枚の超平面を求める方法を述べる。

新しい写像  $g = f / \|\mathbf{w}\|_2$  を考えると、 $\nabla g = \mathbf{w} / \|\mathbf{w}\|_2$ ,  $\|\nabla g\|_2 = 1$  であるので、超平面  $g(\mathbf{x}) = 0$  (これは超平面  $f(\mathbf{x}) = 0$  と一致) からそれと直角な方向、すなわち  $\pm \nabla g$  に距離  $d$  だけ進んだ位置では  $g$  の値が  $\pm d$  だけ変化する。逆に  $g$  の値が  $\pm d$  である位置は超平面  $g(\mathbf{x}) = 0$  から  $\pm \nabla g$  に距離  $d$  だけ離れた位置である。よって求めたい 2 平面の方程式は次式で表される。

$$g(\mathbf{x}) \pm d = 0 \iff f(\mathbf{x}) \pm d \|\mathbf{w}\|_2 = 0$$

$n = 2$ ,  $\mathbf{w} = [1, 1]^\top$ ,  $c = 0.5$ ,  $d = 0.5$  の例を [parts/realAnalysis//chapters//mathematica/平面上の直線から距離 d だけ離れた直線.nb](#) に示した。

### III.25.2 $n$ 次元単位球の体積 $V_n$ は $\pi^{n/2}(\Gamma(1 + n/2))^{-1}$

*Proof.*

$\Omega(n, r)$  ( $r \in \mathbb{R}$ )  $:= \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 \leq r\}$  とすると

$$\begin{aligned} V_n &= \int_{\Omega(n,1)} 1 d\mathbf{x} = \int_{-1}^1 \left( \int_{\Omega(n-1,1-x_n^2)} 1 d\mathbf{x} \right) dx_n = \int_{-1}^1 (\sqrt{1-x_n^2})^{n-1} V_{n-1} dx_n \\ &= V_{n-1} \int_{-1}^1 (1-x^2)^{(n-1)/2} dx = 2V_{n-1} \int_0^1 (1-x^2)^{(n-1)/2} dx \\ &= V_{n-1} \int_0^1 y^{-1/2} (1-y)^{(n-1)/2} dy \quad (y = x^2 \text{ とおいた}) \\ &= V_{n-1} \text{Beta}\left(\frac{1}{2}, \frac{n+1}{2}\right) = V_{n-1} \frac{\Gamma(\frac{1}{2}) \Gamma(\frac{n+1}{2})}{\Gamma(1 + \frac{n}{2})} \\ &= V_{n-1} \frac{\sqrt{\pi} \Gamma(\frac{n+1}{2})}{\Gamma(\frac{(n+1)+1}{2})} \\ \therefore V_n &= V_1 \prod_{i=3}^{n+1} \sqrt{\pi} \frac{\Gamma(i/2)}{\Gamma((i+1)/2)} = 2\pi^{(n-1)/2} \frac{\Gamma(3/2)}{\Gamma(1+n/2)} = \frac{\pi^{n/2}}{\Gamma(1+n/2)} \end{aligned}$$



## 第 III.26 章

# 最適化への応用

### III.26.1 狭義凸関数の線形制約下での最小化

$m, n \in \mathbb{N}$  とする。 $f: \mathbb{R}^n \rightarrow \mathbb{R}$  を狭義凸関数とする。 $\mathbf{c}_i \in \mathbb{R}^n, b_i \in \mathbb{R}$  ( $i = 1, \dots, m$ ) とし、 $g_i: \mathbf{x} \in \mathbb{R}^n \mapsto \mathbf{c}_i^\top \mathbf{x} + b_i \in \mathbb{R}$  とする。 $\boldsymbol{\lambda} \in \mathbb{R}^m$  とし、Lagrange 関数  $F$  を  $F(\mathbf{x}, \boldsymbol{\lambda}) := f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x})$  とする。 $\mathbf{x}_0, \boldsymbol{\lambda}_0$  を  $F$  の停留点、すなわち  $\nabla_{\mathbf{x}} F(\mathbf{x}_0, \boldsymbol{\lambda}_0) = \mathbf{0}$  とすると  $f$  は  $g_i(\mathbf{x}) = 0$  ( $i = 1, \dots, m$ ) なる拘束条件の下、 $\mathbf{x}_0$  で最小値をとる。

*Proof.*

$\mathbf{h} \in \mathbb{R}^n \neq \mathbf{0}$  とし、 $g_i(\mathbf{x} + \mathbf{h}) = 0$  ( $i = 1, \dots, m$ ) を満たすとする。このとき  $f(\mathbf{x}_0 + \mathbf{h}) > f(\mathbf{x}_0)$  であることを示す。

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^\top \mathbf{h} + \mathbf{h}^\top H(f, \mathbf{x}_0) \mathbf{h} + o(\|\mathbf{h}\|_2^3)$$

ここに、 $H(f, \mathbf{x}_0)$  は  $f$  の Hesse 行列の  $\mathbf{x}_0$  に於ける値である。 $\nabla_{\mathbf{x}} F(\mathbf{x}_0, \boldsymbol{\lambda}_0) = \mathbf{0}$  より  $\nabla f(\mathbf{x}_0) = -\sum_{i=1}^m \lambda_0[i] \nabla g_i(\mathbf{x}_0)$  である。 $g_i(\mathbf{x} + \mathbf{h}) = 0$  ( $i = 1, \dots, m$ ) より  $\mathbf{c}_i^\top \mathbf{h} = 0$  であるから

$$\nabla f(\mathbf{x}_0)^\top \mathbf{h} = -\sum_{i=1}^m \lambda_0[i] \nabla g_i(\mathbf{x}_0)^\top \mathbf{h} = -\sum_{i=1}^m \lambda_0[i] \mathbf{c}_i^\top \mathbf{h} = 0$$

である。よって

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \mathbf{h}^\top H(f, \mathbf{x}_0) \mathbf{h} + o(\|\mathbf{h}\|_2^3)$$

$f$  は狭義凸であるので  $H(f, \mathbf{x}_0)$  は正定であり、 $O(\|\mathbf{h}\|_2^2)$  であるから、十分小さい  $\mathbf{h}$  に対して上式の右辺は  $f(\mathbf{x}_0)$  を厳密に上回る。□



## III.26.2 変分法

### III.26.2.1 大域的最小点の十分条件

$a, b \in \mathbb{R}$ ,  $a < b$  とする。 $\Omega_1, \Omega'_1 \subseteq \mathbb{R}$  は空でない開凸領域であるとする。 $v_a, v_b \in \Omega_1$  とする。 $[a, b]$  から  $\Omega_1$  への  $C^1$  級の写像であって、 $f(a) = b_a$ ,  $f(b) = v_b$  を満たし、導関数の値域が  $\Omega'_1$  に含まれるもの全体の集合を  $\mathcal{F}_1$  とする。写像  $g : \Omega_1 \times \Omega'_1 \rightarrow \mathbb{R}$  は  $C^2$  級かつ狭義凸であるとする。汎関数  $I$  を次式で定義する。

$$I : f \in \mathcal{F}_1 \mapsto \int_a^b g(f(x), f'(x)) dx$$

このとき、積分の線形性と  $g$  の狭義凸性から  $I$  は狭義凸である。さらに次が成り立つ。

$$f_0 = \arg \min_{f \in \mathcal{F}} I(f) \iff g^{(1,0)}(f_0(x), f'_0(x)) = \frac{\partial}{\partial x} g^{(0,1)}(f_0(x), f'_0(x))$$

*Proof.*

$\Rightarrow$  については  $f_0$  が変分法の Euler 方程式を満たすことが必要条件であるため、直ちに従う。

以下、 $\Leftarrow$  を示す。 $I$  の狭義凸性から、 $I$  の極小点が存在すれば、それは最小点である。よって  $f_0$  が極小点であることを示せば十分である。今、 $\eta : [a, b] \rightarrow \Omega_1$  が  $\Omega_1$  上で  $C^1$  級かつ  $\eta(a) = \eta(b) = 0$  であるとする。さらに  $\|\eta\|_2 := \sqrt{\int_a^b [\eta(x)^2 + \eta'(x)^2] dx}$  は十分小さく、 $f_0 + \eta \in \mathcal{F}_1$  であるとする。 $\|\eta\|_2 > 0$  である任意の  $\eta$  について  $I(f_0 + \eta) > I(f_0)$  であることを示す。

$$\begin{aligned} I(f_0 + \eta) &= \int_a^b g(f_0(x) + \eta(x), f'_0(x) + \eta'(x)) dx \\ &= \int_a^b g(f_0(x), f'_0(x)) dx + \underbrace{\int_a^b (\nabla g)(f_0(x), f'_0(x))^\top [\eta(x), \eta'(x)]^\top dx}_{(1)} \\ &\quad + \underbrace{\frac{1}{2} \int_a^b [\eta(x), \eta'(x)] H(g)(\eta(x), \eta'(x)) [\eta(x), \eta'(x)]^\top dx + o(\|\eta\|_2^2)}_{(2)} \end{aligned}$$

ここに  $H(g) : \Omega_1 \times \Omega'_1 \rightarrow \mathbb{R}^2$  は  $g$  の Hesse 行列から構成され、次式で定義される。

$$H(g)(x_1, x_2) := \begin{bmatrix} g^{(2,0)}(x_1, x_2) & g^{(1,1)}(x_1, x_2) \\ g^{(1,1)}(x_1, x_2) & g^{(0,2)}(x_1, x_2) \end{bmatrix}$$

$f_0$  が Euler 方程式を満たすことから (1) = 0 である (Euler 方程式の導出と同じ要領で示せる)。  $g$  の狭義凸性より  $H(g)(\eta(x), \eta'(x))$  は正定であり、これと  $\|\eta\|_2 > 0$  より (2) > 0 である。よって、 $\|\eta\|_2^2$  が十分に小さいとき  $I(f_0 + \eta) > I(f_0)$  が成り立つ。  $\square$

### III.26.3 例題

#### III.26.3.1 非負かつ総和 1 の下での 2 乗和の最小化

$N \in \mathbb{N}$  とし、 $x_1, x_2, \dots, x_N \geq 0$ ,  $\sum_{i=1}^N x_i = 1$  であるとする。 $\sum_{i=1}^N x_i^2$  が最小となる必要十分条件は  $x_1 = x_2 = \dots = x_N = 1/N$  である。

数理最適化の分野では、この問題は線形制約付き凸最適化問題に分類され、解法は Lagrange の未定乗数法が定石であるが、ここでは敢えて初等的な証明を与えてみる。

*Proof.*

次式が成り立つ。

$$\sum_{i=1}^N x_i^2 = \sum_{i=1}^N (x_i - 1/N)^2 + 1/N \quad (1)$$

なぜならば

$$\begin{aligned} \sum_{i=1}^N (x_i - 1/N)^2 + 1/N &= \sum_{i=1}^N x_i^2 - \frac{2}{N} \sum_{i=1}^N x_i + \frac{2}{N} \\ &= \sum_{i=1}^N x_i^2 + \frac{2}{N} \left( 1 - \sum_{i=1}^N x_i \right) = \sum_{i=1}^N x_i^2 \quad \left( \because \sum_{i=1}^N x_i = 1 \right) \end{aligned}$$

式 (1) より  $\sum_{i=1}^N x_i^2$  が最小となるのは  $x_1 = x_2 = \dots = x_N = 1/N$  のとき、かつその時に限る。最小値は  $1/N$  である。  $\square$

#### III.26.3.2 正かつ総和 1 の下での逆数和の最小化

$N \in \mathbb{N}$  とし、 $x_1, x_2, \dots, x_N > 0$ ,  $\sum_{i=1}^N x_i = 1$  であるとする。 $\sum_{i=1}^N 1/x_i$  が最小となる必要十分条件は  $x_1 = x_2 = \dots = x_N = 1/N$  である。

この問題も III.26.3.1 と同じく Lagrange の未定乗数法が定石であるが、ここでは敢えて初等的な証明を与えてみる。

*Proof.*

関数  $f : x \in \{t \in \mathbb{R} \mid t > 0\} \mapsto 1/x$  は 狭義下に凸である。 $w_i \geq 0$  ( $i = 1, 2, \dots, N$ ),  $\sum_{i=1}^N w_i = 1$  とすると、Jensen の不等式より次式が成り立つ。

$$\sum_{i=1}^N w_i f(x_i) \geq f\left(\sum_{i=1}^N w_i x_i\right)$$

等号成立の必要十分条件は  $x_1 = x_2 = \dots = x_N$  である。上式で  $w_i = 1/N$  ( $i = 1, \dots, N$ ) とすると次式を得る。

$$\frac{1}{N} \sum_{i=1}^N f(x_i) \geq \frac{1}{\frac{1}{N} \sum_{i=1}^N x_i} = N \quad \left( \because \sum_{i=1}^N w_i = 1 \right) \quad \therefore \sum_{i=1}^N 1/x_i \geq N^2$$

先述の等号成立条件を考えると、 $x_i = 1/N$  ( $i = 1, 2, \dots, N$ ) が唯一の最小点である。  $\square$

## 第Ⅳ部

# 複素解析

## 第 IV.1 章

# 複素数

### IV.1.1 複素数の定義

任意の 2 つの実数  $a, b \in \mathbb{R}$  に対する 2 項組  $(a, b)$  (※ここでは  $(,)$  という記号を用いているが、どんな記号を用いるかはどうでもいい。別の記号を使っても差し支えない。) を「**複素数**」と名付ける。複素数  $(a, b)$  の第 1 要素  $a$  を「**実部**」、第 2 要素  $b$  を「**虚部**」と名付ける。複素数全体の集合を  $\mathbb{C}$  と表記することにする。

任意の 2 つの複素数  $(a_1, b_1), (a_2, b_2) \in \mathbb{C}$  に対して演算「和 (+)」, 「実数倍 (スカラー倍)」及び「積 ( $\times$ )」を次のように**定義**する。

$$\begin{aligned}\text{和} : (a_1, b_1) + (a_2, b_2) &= (a_1 + a_2, b_1 + b_2) \\ \text{実数倍} (a \in \mathbb{R}) : a(a_1, b_1) &= (aa_1, ab_1) \\ \text{積} : (a_1, b_1) \times (a_2, b_2) &= (a_1a_2 - b_1b_2, a_1b_2 + b_2a_1)\end{aligned}$$

特に、実部が 0 である複素数  $(0, b)$  を「**純虚数**」と名付ける。純虚数のうち虚部が 1 であるもの  $(0, 1)$  を特に「**虚数単位**」と名付け、 $i$  と表記することにする。

$(a, 0)$  を簡略化して  $a$  と書くことを許せば、読者が高校以来慣れ親しんでいるであろう、次のような計算が意味を持つ。

例えば

$$(a_1 + ib_1)(a_2 + ib_2) = (a_1a_2 - b_1b_2) + i(a_1b_2 + b_2a_1)$$

なぜならば、

$$\begin{aligned}(a_1 + ib_1)(a_2 + ib_2) &= [(a_1, 0) + i(b_1, 0)] \times [(a_2, 0) + i(b_2, 0)] \\ &= [(a_1, 0) + (0, 1) \times (b_1, 0)] \times [(a_2, 0) + (0, 1) \times (b_2, 0)] \\ &= [(a_1, 0) + (0, b_1)] \times [(a_2, 0) + (0, b_2)] \\ &= (a_1, b_1) \times (a_2, b_2) = (a_1a_2 - b_1b_2, a_1b_2 + b_2a_1) \\ &= (a_1a_2 - b_1b_2) + i(a_1b_2 + b_2a_1)\end{aligned}$$

また、例えば、

$$i^2 = -1$$

なぜならば

$$i^2 = (0, 1) \times (0, 1) = (0 \times 0 - 1 \times 1, 0 \times 1 + 1 \times 0) = (-1, 0) = -1$$

$i^2 = -1$  が虚数単位の定義なのではなく、そうなるように先回りして巧妙に複素数を構築しただけであることに注意して欲しい。

#### IV.1.1.1 注意

実数  $a$  と複素数  $(a, 0)$  は等しくない。イカメシく言うと実数と複素数は「型 (type)」が異なる。複素数はあくまで上で定義した 2 実数の組であって、単一の実数ではない。だから、**実数と複素数の間**に演算「和」は定義されていない。つまり、 $a + ib$  という書き方は厳密には誤りである。本来  $(a, 0) + i \times (b, 0)$  であるものを簡略化して書いているに過ぎない。

#### IV.1.1.2 コラム $i^2 = -1$ ?

「2 乗して**実数**-1 になる」数は無い!! 代わりに、「2 乗して**複素数**  $(-1, 0)$  になる」数はある。それは先程定義した虚数単位  $i = (0, 1)$  である。

いい加減な教科書では、「2 乗して-1 になる数を虚数単位として定義し、 $i$  と表記する」と言っているが、この定義は**間違っている**。複素数の集合の定義が完了していない段階で演算「2 乗」と口走っているのがおかしい。

まず数の集合があって、その上に初めて演算を定義できる。先述の「2 乗して...」の「2 乗」はこの文章中では未定義なのである!! 百歩譲って実数の集合  $\mathbb{R}$  上で定義された「2 乗」と解釈したとしても、実数の集合に属さない数に対してこの演算を適用できる保証など無い。ガソリンの給油口に紅茶を注いで「走れ」と祈るようなものだ。

高校で「2 乗して-1 になる**仮想的な数** (imaginary number) を「虚数単位」 $i$  と呼ぶ。」と教わったはいいものの、「そんな数は自然界にないよなあ。なんか気持ち悪い...」という**違和感**に取り憑かれる原因は、そもそもの定義の間違いにある。違和感があって当然である。

確かに、人類が複素数を考えた動機は「2 乗して-1 になる数はないか?」という問題であった。新しい理論がそういう動機から始まるのは自然である。しかし数学的に「ちゃんとした」理論を作るためには、既存の枠組みのままでは不可能だと分かり次第、思い切ってルールを拡張し、所望の性質が得られるように理屈を先回りして演算を機械的に定義するような break through が必要なこともある。その際に、既存の理論に矛盾しないように構築しなければならないのは言うまでもない。

### IV.1.2 諸公式

#### IV.1.2.1 $x > 0, z \in \mathbb{C}$ のとき $|x^z| = x^{\operatorname{Re}(z)}$

*Proof.*

$z = a + bi$  ( $a, b \in \mathbb{R}$ ) とする。

$$|x^z| = |x^{a+bi}| = |x^a x^{bi}| = |x^a| |x^{bi}| = |x^{\operatorname{Re}(z)}| |e^{bi \log x}| = x^{\operatorname{Re}(z)}$$

□

$$\text{IV.1.2.2} \quad \forall n \nmid k, \sum_{l=1}^n \exp\left(i\frac{k}{n}2\pi l\right) = 0$$

*Proof.*

$$\sum_{l=1}^n \exp\left(i\frac{k}{n}2\pi l\right) = \frac{\exp\left(i\frac{k}{n}2\pi\right) \left(1 - \exp\left(i\frac{k}{n}2\pi\right)^n\right)}{1 - \exp\left(i\frac{k}{n}2\pi\right)} = \frac{\exp\left(i\frac{k}{n}2\pi\right) (1 - e^{ik2\pi})}{1 - \exp\left(i\frac{k}{n}2\pi\right)} = 0$$

□

## IV.1.3 代数方程式

### IV.1.3.1 諸定理

IV.1.3.1.1 実係数  $n$  次方程式が複素数解  $s = \alpha + i\beta$  を解に持つならばその複素共役  $s^* = \alpha - i\beta$  もまた解である

実係数  $n$  次方程式が複素数解  $s = \alpha + i\beta$  を解に持つならばその複素共役  $s^* = \alpha - i\beta$  もまた解である。

*Proof.*

最高次の係数を常に 1 として一般性を失わないので方程式は

$$x^n + \sum_{k=n-1}^0 c_k x^k = 0, \quad c_k \in \mathbb{R}$$

と書ける。 $s = \alpha + i\beta$ ,  $\alpha, \beta \in \mathbb{R}$  がこれの解であるから当然

$$s^n + \sum_{k=n-1}^0 c_k s^k = 0$$

両辺の複素共役をとると

$$\begin{aligned} \left[ s^n + \sum_{k=n-1}^0 c_k s^k \right]^* &= 0^* = 0 \\ \therefore s^{*n} + \sum_{k=n-1}^0 c_k s^{*k} &= 0 \end{aligned}$$

これは  $s^*$  が解であることに他ならない。

□

### IV.1.3.2 例題

$$\text{IV.1.3.3} \quad z^n + z^{n-1} + \cdots + z^2 + z + 1 = 0$$

次の方程式を解け。

$$z^n + z^{n-1} + \cdots + z^2 + z + 1 = 0$$

解法.

$z = 1$  が解ではないから両辺に  $z - 1$  を掛けて

$$z^{n+1} - 1 = 0$$

の形にする。 $z$  の絶対値が 1 だと分かるので  $z = e^{\theta i}$  ( $0 \leq \theta \leq 2\pi$ ) とおいて解き、 $z = 1$  という不要な解を除去する。 □

## 第 IV.2 章

# 複素索引数凸関数

### IV.2.1 複素索引数凸関数の例

IV.2.1.1  $\|\sum_{i=1}^N X_i \mathbf{a}_i + \mathbf{b}\|_2^2$  ( $\mathbf{a}_i, \mathbf{b} \in \mathbb{C}^n$ ,  $X_i \in \mathbb{C}^{m \times n}$ ) は  $X_1, \dots, X_N$  に関して凸である。

*Proof.*

$T_X := (X_1, \dots, X_N)$ ,  $T_Y := (Y_1, \dots, Y_N)$ ,  $\lambda \in [0, 1]$  とし、 $f(T_X) := \|\sum_{i=1}^N X_i \mathbf{a}_i + \mathbf{b}\|_2^2$  とする。

$$\begin{aligned} f((1-\lambda)T_X + \lambda T_Y) &= \left\| \sum_{i=1}^N ((1-\lambda)X_i + \lambda Y_i) \mathbf{a}_i + \mathbf{b} \right\|_2^2 \\ &= \left\| (1-\lambda) \left( \sum_{i=1}^N X_i \mathbf{a}_i + \mathbf{b} \right) + \lambda \left( \sum_{i=1}^N Y_i \mathbf{a}_i + \mathbf{b} \right) \right\|_2^2 \\ &= (1-\lambda)^2 \left\| \sum_{i=1}^N X_i \mathbf{a}_i + \mathbf{b} \right\|_2^2 + \lambda^2 \left\| \sum_{i=1}^N Y_i \mathbf{a}_i + \mathbf{b} \right\|_2^2 \\ &\quad + 2\lambda(1-\lambda) \operatorname{Re} \left( \left( \sum_{i=1}^N X_i \mathbf{a}_i + \mathbf{b} \right)^H \left( \sum_{i=1}^N Y_i \mathbf{a}_i + \mathbf{b} \right) \right) \end{aligned}$$

よって

$$\begin{aligned} &(1-\lambda)f(T_X) + \lambda f(T_Y) - f((1-\lambda)T_X + \lambda T_Y) \\ &= \lambda(1-\lambda) \left[ \left\| \sum_{i=1}^N X_i \mathbf{a}_i + \mathbf{b} \right\|_2^2 + \left\| \sum_{i=1}^N Y_i \mathbf{a}_i + \mathbf{b} \right\|_2^2 - 2 \operatorname{Re} \left( \left( \sum_{i=1}^N X_i \mathbf{a}_i + \mathbf{b} \right)^H \left( \sum_{i=1}^N Y_i \mathbf{a}_i + \mathbf{b} \right) \right) \right] \\ &= \lambda(1-\lambda) \left\| \left( \sum_{i=1}^N X_i \mathbf{a}_i + \mathbf{b} \right) - \left( \sum_{i=1}^N Y_i \mathbf{a}_i + \mathbf{b} \right) \right\|_2^2 \geq 0 \end{aligned}$$

□

IV.2.1.2  $\|A\mathbf{x} + \mathbf{b}\|_2^2$  が狭義凸  $\iff A^*A \succ O$

$A \in \mathbb{C}^{m \times n}$ ,  $\mathbf{x} \in \mathbb{C}^n$ ,  $\mathbf{b} \in \mathbb{C}^m$  とする。 $f(\mathbf{x}) := \|A\mathbf{x} + \mathbf{b}\|_2^2$  が  $\mathbf{x}$  に関して狭義凸となる必要十分条件は  $A^*A \succ O$  である。



$\lambda \in (0, 1)$ ,  $\mathbf{x}, \mathbf{y} \in \mathbb{C}^m$ ,  $\mathbf{x} \neq \mathbf{y}$  とする。

*Proof.* まず次式が成り立つ。

$$\begin{aligned} f((1-\lambda)\mathbf{x} + \lambda\mathbf{y}) &:= \|(1-\lambda)(A\mathbf{x} + \mathbf{b}) + \lambda(A\mathbf{y} + \mathbf{b})\|_2^2 \\ &= (1-\lambda)^2 f(\mathbf{x}) + \lambda^2 f(\mathbf{y}) + 2\lambda(1-\lambda) \operatorname{Re}((A\mathbf{x} + \mathbf{b})^*(A\mathbf{y} + \mathbf{b})) \end{aligned}$$

これより

$$\begin{aligned} &(1-\lambda)f(\mathbf{x}) + \lambda f(\mathbf{y}) - f((1-\lambda)\mathbf{x} + \lambda\mathbf{y}) \\ &= \lambda(1-\lambda) [f(\mathbf{x}) + f(\mathbf{y}) - 2\operatorname{Re}((A\mathbf{x} + \mathbf{b})^*(A\mathbf{y} + \mathbf{b}))] \\ &= \lambda(1-\lambda) \| (A\mathbf{x} + \mathbf{b}) - (A\mathbf{y} + \mathbf{b}) \|_2^2 = \lambda(1-\lambda) \| A(\mathbf{x} - \mathbf{y}) \|_2^2 \\ &= \lambda(1-\lambda) (\mathbf{x} - \mathbf{y})^* A^* A (\mathbf{x} - \mathbf{y}) \end{aligned}$$

$\lambda \in (0, 1)$ ,  $\mathbf{x} \neq \mathbf{y}$  に対して上式が正となる必要十分条件は  $A^* A \succ O$  である。

□

## 第 IV.3 章

# 特殊関数

### IV.3.1 Gamma 関数

$$\text{IV.3.1.1} \quad \lim_{n \rightarrow \infty} \int_0^n t^{z-1} (1+t/n)^n dt = \Gamma(z) \quad (z \in \mathbb{C}, \operatorname{Re}(z) > 0)$$

Wikipedia で紹介されている無限乗積表示の導出過程でこの主張が使われているが、その根拠が薄かったの  
で自力で証明してみた。

*Proof.*

$\varepsilon > 0$  を任意にとる。  $G_n(z) := \int_0^n t^{z-1} (1+t/n)^n dt$  とする。

Gamma 関数  $\int_0^\infty t^{z-1} e^{-t} dt$  は収束するから、十分大きい  $T_1(\varepsilon) > 0$  を取れば

$$\left| \int_{T_1(\varepsilon)}^\infty t^{z-1} e^{-t} dt \right| \leq \int_{T_1(\varepsilon)}^\infty |t^{z-1} e^{-t}| dt < \varepsilon/3$$

III.4.1.1 より  $(1-t/n)^n$  は  $n$  について単調増加であり、かつ  $\lim_{n \rightarrow \infty} (1-t/n)^n = e^{-t}$  なので、 $n \geq T_1(\varepsilon)$   
ならば

$$\begin{aligned} \left| \int_{T_1(\varepsilon)}^n t^{z-1} (1-t/n)^n dt \right| &\leq \int_{T_1(\varepsilon)}^n |t^{z-1} (1-t/n)^n| dt \\ &= \int_{T_1(\varepsilon)}^n t^{\operatorname{Re}(z)-1} (1-t/n)^n dt \leq \int_{T_1(\varepsilon)}^\infty |t^{z-1} e^{-t}| dt < \varepsilon/3 \end{aligned}$$

また、十分大きい自然数  $N_1(\varepsilon, T_1(\varepsilon))$  をとると、 $n \geq N_1(\varepsilon, T_1(\varepsilon))$  ならば次式が成り立つ。

$$\left| \int_0^{T_1(\varepsilon)} t^{z-1} \{e^{-t} - (1-t/n)^n\} dt \right| < \varepsilon/3$$

以上より、 $n \geq \max\{T_1(\varepsilon), N_1(\varepsilon, T_1(\varepsilon))\}$  ならば次式が成り立つ。

$$\begin{aligned} |\Gamma(z) - G_n(z)| &= \left| \int_{T_1(\varepsilon)}^\infty t^{z-1} e^{-t} dt + \int_0^{T_1(\varepsilon)} t^{z-1} \{e^{-t} - (1-t/n)^n\} dt - \int_{T_1(\varepsilon)}^n t^{z-1} (1-t/n)^n dt \right| \\ &\leq \left| \int_{T_1(\varepsilon)}^\infty t^{z-1} e^{-t} dt \right| + \left| \int_0^{T_1(\varepsilon)} t^{z-1} \{e^{-t} - (1-t/n)^n\} dt \right| + \left| \int_{T_1(\varepsilon)}^n t^{z-1} (1-t/n)^n dt \right| < \varepsilon \end{aligned}$$

□

## 第 IV.4 章

# 複素積分

### IV.4.1 無限に大きい半円周上での $e^{iz}/z$ の積分

$R > 0$  とし、パラメータ表示された曲線  $C: z = Re^{i\theta}; \theta: 0 \rightarrow \pi$  を考える。このとき次式が成り立つ。

$$\left| \int_C \frac{e^{iz}}{z} dz \right| < \frac{\pi}{R} (1 - e^{-R})$$

*Proof.*

$dz = iRe^{i\theta} d\theta$  であるから

$$\begin{aligned} S &:= \int_C \frac{e^{iz}}{z} dz = \int_0^\pi \frac{1}{Re^{i\theta}} \exp iR(\cos \theta + i \sin \theta) iRe^{i\theta} d\theta = i \int_0^\pi \exp(iR \cos \theta - R \sin \theta) d\theta \\ |S| &= \left| \int_0^\pi \exp(iR \cos \theta - R \sin \theta) d\theta \right| \leq \int_0^\pi |\exp(iR \cos \theta - R \sin \theta)| d\theta \\ &= \int_0^\pi e^{-R \sin \theta} d\theta = 2 \int_0^{\pi/2} e^{-R \sin \theta} d\theta \end{aligned}$$

$2\theta/\pi < \sin \theta$  ( $\theta \in (0, \pi/2)$ ) であるから次式が成り立つ。

$$2 \int_0^{\pi/2} e^{-R \sin \theta} d\theta < 2 \int_0^{\pi/2} e^{-2R\theta/\pi} d\theta = \frac{\pi}{R} (1 - e^{-R})$$

□

### IV.4.2 例題

IV.4.2.1  $\int_0^{2\pi} \text{Log}(1 - a \exp(i\theta)) d\theta = 0$  ( $C$ : 原点中心の単位円,  $a \in \mathbb{R}$ )

解法.

$a = 0$  のときは明らかに成り立つ。 $a \neq 0$  のときは  $z = e^{i\theta}$  とすると

$$\int_0^{2\pi} \text{Log}(1 - ae^{i\theta}) d\theta = \int_C \frac{\text{Log}(1 - az)}{iz} dz \quad (C: \text{原点中心の左回り単位円})$$

被積分関数の  $C$  内部の特異点は  $z = 0$  である。原点の近傍 (具体的には  $|z| < 1/|a|$  の範囲) においては

$$\frac{\operatorname{Log}(1 - az)}{iz} = \frac{-1}{iz} \sum_{n=1}^{\infty} \frac{(az)^n}{n} = \frac{-1}{i} \sum_{n=1}^{\infty} \frac{a^n z^{n-1}}{n} = \frac{-1}{i} \sum_{n=0}^{\infty} \frac{a^{n+1}}{n+1} z^n$$

であり、特異点  $z = 0$  は除去可能であることから従う。

## 第 IV.5 章

# Fourier 級数

### IV.5.1 連続関数の Fourier 級数の高周波成分は 0 に収束する

$f$  は区間  $[-\pi, \pi]$  で連続とする。 $f$  の Fourier 係数を  $c_k$  とすると  $\lim_{k \rightarrow \infty} c_k = 0$  となる。

*Proof.*

まず

$$\lim_{k \rightarrow \infty} \int_{-\pi}^{\pi} f(x) \sin kx dx = 0$$

を示す。ここで

$$\int_{-\pi}^{\pi} f(x) \sin kx dx = \int_0^{\pi} (f(x) - f(-x)) \sin kx dx$$

であり、 $f(x) - f(-x)$  は  $[0, \pi]$  で連続なので結局のところ区間  $[0, \pi]$  の任意の連続関数  $f$  に対して

$$\lim_{k \rightarrow \infty} \int_0^{\pi} f(x) \sin kx dx = 0$$

が成り立つことを示せば良い。 $f$  は閉区間で連続だから有界であり、特に一様連続なので、任意の  $\varepsilon > 0$  に対してある  $\delta > 0$  が存在して  $\forall (x, y \in [0, \pi], |x - y| \leq \delta), |f(x) - f(y)| < \varepsilon$  が成り立つ。 $k \geq 2, \pi/k < \delta$  となるように  $k$  を十分大きくとる。

$$\begin{aligned} \int_0^{\pi} f(x) \sin kx dx &= \sum_{l=0}^{\lfloor k/2 \rfloor - 1} \left( \int_{2l\pi/k}^{(2l+1)\pi/k} f(x) \sin kx dx + \int_{(2l+1)\pi/k}^{(2l+2)\pi/k} f(x) \sin kx dx \right) \\ &\quad + \int_{2\lfloor k/2 \rfloor \pi/k}^{\pi} f(x) \sin kx dx \end{aligned}$$

最後の項は  $k \rightarrow \infty$  のとき 0 に収束する。最初の 2 項は  $y = kx$  と変数変換して

$$\begin{aligned} &\sum_{l=0}^{\lfloor k/2 \rfloor - 1} \frac{1}{k} \left( \int_{2l\pi}^{(2l+1)\pi} f(y/k) \sin y dy + \int_{(2l+1)\pi}^{(2l+2)\pi} f(y/k) \sin y dy \right) \\ &= \frac{1}{k} \sum_{l=0}^{\lfloor k/2 \rfloor - 1} \int_{2l\pi}^{(2l+1)\pi} (f(y/k) - f(y/k + \pi/k)) \sin y dy \end{aligned}$$

となる。この絶対値は

$$\begin{aligned} &\leq \frac{1}{k} \sum_{l=0}^{\lfloor k/2 \rfloor - 1} \int_{2l\pi}^{(2l+1)\pi} |f(y/k) - f(y/k + \pi/k) \sin y| dy \\ &< \frac{1}{k} \sum_{l=0}^{\lfloor k/2 \rfloor - 1} \int_{2l\pi}^{(2l+1)\pi} \varepsilon |\sin y| dy = \frac{2\varepsilon}{k} \sum_{l=0}^{\lfloor k/2 \rfloor - 1} 1 = \frac{2\varepsilon}{k} \lfloor k/2 \rfloor \leq \frac{2\varepsilon}{k} \frac{k}{2} = \varepsilon \end{aligned}$$

同様に

$$\lim_{k \rightarrow \infty} \int_{-\pi}^{\pi} f(x) \cos kx dx = 0$$

も示せる。 □

## IV.5.2 $\int_{-\pi}^{\pi} \frac{\sin nx}{2 \tan \frac{x}{2}} dx = \pi \quad (n \in \mathbb{N})$

*Proof.*

求めたい積分の値を  $I_n$  とする。分子の  $\sin nx$  について  $nx = (n-1/2)x + x/2$  として加法定理により展開すると

$$I_n = \frac{1}{2} \int_{-\pi}^{\pi} \frac{\cos^2(x/2)}{\sin(x/2)} \sin(n-1/2)x dx + \frac{1}{2} \int_{-\pi}^{\pi} \cos(x/2) \cos(n-1/2)x dx$$

第一項の分子で  $\cos^2(x/2) = 1 - \sin^2(x/2)$  とすると

$$\begin{aligned} I_n &= \frac{1}{2} \int_{-\pi}^{\pi} \frac{\sin(n-1/2)x}{\sin(x/2)} dx - \frac{1}{2} \int_{-\pi}^{\pi} \sin(x/2) \sin(n-1/2)x dx + \frac{1}{2} \int_{-\pi}^{\pi} \cos(x/2) \cos(n-1/2)x dx \\ &= \frac{1}{2} \int_{-\pi}^{\pi} \frac{\sin(n-1/2)x}{\sin(x/2)} dx + \frac{1}{2} \int_{-\pi}^{\pi} \cos nx dx \quad (\uparrow \text{の第 2,3 項を加法定理でまとめた}) \\ &= \frac{1}{2} \int_{-\pi}^{\pi} \frac{\sin(n-1/2)x}{\sin(x/2)} dx \quad (\because \text{第 2 項は 0}) \end{aligned}$$

さらに  $(n-1/2)x = (n-1)x + x/2$  として加法定理を使うと

$$I_n = \frac{1}{2} \int_{-\pi}^{\pi} \frac{\sin(n-1)x \cos(x/2)}{\sin(x/2)} dx + \frac{1}{2} \int_{-\pi}^{\pi} \cos(n-1)x dx = I_{n-1} + \begin{cases} \pi & (n=1) \\ 0 & (\text{otherwise}) \end{cases}$$

これと  $I_0 = 0$  より定理が成り立つ。 □

## IV.5.3 Fourier 級数の応用

### IV.5.3.1 $\int_0^{\infty} \frac{\sin x}{x} dx = \pi/2$

*Proof.*

複素積分を使った導出がよく知られているが、実解析の範囲で導出してみる。この積分が収束することは、 $(\sin x)/x$  が  $[0, \pi/2]$  で連続であることから  $[0, \pi/2]$  で積分可能であること、および  $\sin x = (-\cos x)'$  を用い

て部分積分することで  $[\pi/2, \infty)$  で積分が絶対収束することから示せる。よって、

$$\begin{aligned}\int_0^\infty \frac{\sin x}{x} dx &= \lim_{n \rightarrow \infty} \int_0^{n\pi} \frac{\sin x}{x} dx = \lim_{n \rightarrow \infty} \int_0^\pi \frac{\sin nt}{t} dt \quad (t = x/n \text{ とおいた}) \\ &= \lim_{n \rightarrow \infty} \frac{1}{2} \int_{-\pi}^\pi \frac{\sin nt}{t} dt = \lim_{n \rightarrow \infty} \frac{1}{2} \int_{-\pi}^\pi \left( \frac{1}{t} - \frac{1}{2 \tan(t/2)} + \frac{1}{2 \tan(t/2)} \right) \sin ntdt\end{aligned}$$

$\frac{1}{t} - \frac{1}{2 \tan(t/2)}$  は  $[-\pi, \pi]$  で連続だから [IV.5.1](#) より

$$\lim_{n \rightarrow \infty} \frac{1}{2} \int_{-\pi}^\pi \left( \frac{1}{t} - \frac{1}{2 \tan(t/2)} \right) \sin ntdt = 0$$

であること、および [IV.5.2](#) より証明が済む。

□

## 第 IV.6 章

# Fourier 変換

### IV.6.1 余弦, 正弦変換

$f(t)$  の余弦変換  $C(\omega)$ , 正弦変換  $S(\omega)$  は次のように定義される。

$$C(\omega) := \sqrt{\frac{2}{\pi}} \int_0^\infty f(x) \cos \omega t dt$$
$$S(\omega) := \sqrt{\frac{2}{\pi}} \int_0^\infty f(x) \sin \omega t dt$$

$C, S$  はそれぞれ  $\omega$  の偶関数, 奇関数になっていることに注意。そして  $t \geq 0$  において

$$f(t) \sim \sqrt{\frac{2}{\pi}} \int_0^\infty C(\omega) \cos \omega t d\omega = \sqrt{\frac{2}{\pi}} \int_0^\infty S(\omega) \sin \omega t d\omega$$

が成り立つ。注意すべきは、この逆変換は  $t < 0$  に関しては何も教えてくれないということだ。変換時に  $t < 0$  における情報を捨ててしまっているのだから当然である。以下、上の関係を示す。

*Proof.*

$f(t)$  の  $t \geq 0$  の部分を取り出して  $t < 0$  の部分に拡張することを考える。関数  $g_e(t), g_o(t)$  を次のように定義する。

$$g_e(t) := f(|t|) \quad (-\infty < t < \infty)$$

$$g_o(t) := \begin{cases} f(t) & (t \geq 0) \\ -f(-t) & (t < 0) \end{cases}$$

$g_e, g_o$  はそれぞれ  $f$  を偶関数, 奇関数として拡張したものである。両者のフーリエ変換を  $G_e(\omega), G_o(\omega)$  とすると

$$\begin{aligned} G_e(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty g_e(t) e^{-i\omega t} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty g_e(t) (\cos \omega t - i \sin \omega t) dt \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty g_e(t) \cos \omega t dt \quad (\because g_e(t) \sin \omega t \text{ は奇関数なので対称積分の結果は } 0) \\ &= \sqrt{\frac{2}{\pi}} \int_0^\infty g_e(t) \cos \omega t dt = \sqrt{\frac{2}{\pi}} \int_0^\infty f(t) \cos \omega t dt = C(\omega) \end{aligned}$$



となり余弦変換と一致する。同様にして

$$G_o(\omega) = -iS(\omega)$$

となることも確かめられる。

次に逆フーリエ変換で  $g_e, g_o$  を復元してみる。

$$\begin{aligned} g_e(t) &\sim \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} G_e(\omega) e^{i\omega t} d\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} C(\omega) e^{i\omega t} d\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} C(\omega) (\cos \omega t + i \sin \omega t) d\omega \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} C(\omega) \cos \omega t d\omega \quad (\because C(\omega) \sin \omega t \text{ は奇関数なので対称積分の結果は } 0) \\ &= \sqrt{\frac{2}{\pi}} \int_0^{\infty} C(\omega) \cos \omega t d\omega \end{aligned}$$

$$\begin{aligned} g_o(t) &\sim \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} G_o(\omega) e^{i\omega t} d\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} -iS(\omega) e^{i\omega t} d\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} -iS(\omega) (\cos \omega t + i \sin \omega t) d\omega \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} S(\omega) \sin \omega t d\omega \quad (\because S(\omega) \cos \omega t \text{ は奇関数なので対称積分の結果は } 0) \\ &= \sqrt{\frac{2}{\pi}} \int_0^{\infty} S(\omega) \sin \omega t d\omega \end{aligned}$$

$t \geq 0$  では  $f, g_e, g_o$  は等しいから

$$f(t) \sim \sqrt{\frac{2}{\pi}} \int_0^{\infty} C(\omega) \cos \omega t d\omega = \sqrt{\frac{2}{\pi}} \int_0^{\infty} S(\omega) \sin \omega t d\omega$$

□

繰り返しになるが上式は  $t \geq 0$  でしか成り立たない。 $t < 0$  のことはわからないのである。クレイジーな変換だと思うかもしれないが、ちゃんと利点もある。実験データや音声、映像データは普通、記録開始を時刻 0 として扱う。この場合は  $t < 0$  に関しては知る必要がない (どうでもいい) のだ。そういう状況下では複素数が登場する本家のフーリエ変換よりも  $\sin, \cos$  どちらか好きな方だけ使えば済むような正弦、余弦変換が役に立つことがある。

## 第 IV.7 章

# Bessel 関数

### IV.7.1 定義

$x \in \mathbb{R}$ ,  $z \in \mathbb{C}$  に対して  $e^{iz \sin x}$  の Fourier 級数展開により「 $n$  次の第一種 Bessel 関数  $J_n(z)$ 」を次のように定義する。

$$e^{iz \sin x} = \sum_{n=-\infty}^{\infty} J_n(z) e^{inx} \quad (1)$$

すなわち

$$J_n(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{iz \sin x} e^{-inx} dx \quad (2)$$

#### IV.7.1.1 簡単化

式 (2) は次のようにして簡単化できる。

$$J_n(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp i(z \sin x - nx) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(z \sin x - nx) dx + \frac{1}{2\pi} \int_{-\pi}^{\pi} i \sin(z \sin x - nx) dx$$

第 1 項は偶関数の対称積分であり、第 2 項は奇関数の対称積分なので 0 だから結局次のようになる。

$$J_n(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(z \sin x - nx) dx = \frac{1}{\pi} \int_0^{\pi} \cos(z \sin x - nx) dx \quad (1)$$

### IV.7.2 性質

#### IV.7.2.1 $J_n(-z) = J_{-n}(z)$

式 (1) より容易に示せる。

#### IV.7.2.2 $J_{2m}(-z) = J_{2m}(z)$ , $J_{2m+1}(-z) = -J_{2m+1}(z)$

$m \in \mathbb{Z}$  とするとき、次式が成り立つ。

$$J_{2m}(-z) = J_{2m}(z), \quad J_{2m+1}(-z) = -J_{2m+1}(z)$$

*Proof.*

$n = 2m$  とする。

$$\begin{aligned}
 J_{2m}(-z) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(-z \sin x - 2mx) dx \\
 &= \frac{1}{2\pi} \int_{-2\pi}^0 \cos(z \sin y - 2m(y + \pi)) dy \quad (\text{変数変換 } x = y + \pi) \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(z \sin y - 2m(y + \pi)) dy = \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos(z \sin y - 2my) dy = J_{2m}(z)
 \end{aligned}$$

同様に  $J_{2m+1}(-z) = -J_{2m+1}(z)$  も示せる。

□

### IV.7.3 Maclaurin 展開

式 (1) の左辺を指数関数に関して Maclaurin 展開すると  $x$  を変数とした Fourier 級数展開が得られることを利用して  $J_n(z)$  の Maclaurin 展開を求める。

$$e^{iz \sin x} = \exp\left(z \frac{e^{ix} - e^{-ix}}{2}\right) = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{z}{2}\right)^k (e^{ix} - e^{-ix})^k$$

$n \geq 0$  のとき、 $(e^{ix} - e^{-ix})^k$  の展開で  $e^{inx}$  が生じるのは  $k = n + 2l$  ( $l = 0, 1, 2, \dots$ ) のときで、その係数は  $(-1)^l \binom{k}{l} = (-1)^l \binom{n+2l}{l}$  である。 $n < 0$  のとき、 $(e^{ix} - e^{-ix})^k$  の展開で  $e^{inx}$  が生じるのは  $k = -n + 2l$  ( $l = 0, 1, 2, \dots$ ) のときで、その係数は  $(-1)^{k-l} \binom{k}{l} = (-1)^{-n+l} \binom{-n+2l}{l}$  である。式 (1) の両辺の Fourier 展開基底の係数同士が一致せねばならぬから

$$J_n(z) = \begin{cases} \sum_{l=0}^{\infty} \frac{(-1)^l}{(n+2l)!} \left(\frac{z}{2}\right)^{n+2l} \binom{n+2l}{l} = \sum_{l=0}^{\infty} \frac{(-1)^l}{l!(n+l)!} \left(\frac{z}{2}\right)^{n+2l} & (n \geq 0) \\ \sum_{l=0}^{\infty} \frac{(-1)^{-n+l}}{(-n+2l)!} \left(\frac{z}{2}\right)^{-n+2l} \binom{-n+2l}{l} = \sum_{l=0}^{\infty} \frac{(-1)^{-n+l}}{l!(-n+l)!} \left(\frac{z}{2}\right)^{-n+2l} & (n < 0) \end{cases}$$

第Ⅴ部

線形代数

## 第 V.1 章

# 表記

- $E_A(\lambda)$  : 行列  $A$  の固有値  $\lambda$  に対する固有空間
- $F_{A,k}(\lambda)$  : 行列  $A$  の固有値  $\lambda$  に対する  $k$  階一般固有空間
- $F_A(\lambda)$  : 行列  $A$  の固有値  $\lambda$  に対する一般固有空間

## 第 V.2 章

# 行列の演算

### V.2.1 結合則

$A, B, C$  が  $l \times m$ ,  $m \times n$ ,  $n \times o$  行列であるとき  $(AB)C = A(BC)$

*Proof.*

まず下準備。

$$A := [a_{ij}]_{l \times m}, \quad B := [b_{ij}]_{m \times n}, \quad C := [c_{ij}]_{n \times o}$$

$$M_{AB} := [m_{ABij}]_{l \times n} = AB, \quad M_{BC} := [m_{BCij}]_{m \times o} = BC$$

$$N_1 := [n_{1ij}]_{l \times o} = (AB)C = M_{AB}C, \quad N_2 := [n_{2ij}]_{l \times o} = A(BC) = AM_{BC}$$

とする。  $N_1 = N_2$  すなわち  $n_{1ij} = n_{2ij}$  を示せば良い。

まず

$$m_{ABij} = \sum_{k=1}^m a_{ik} b_{kj}, \quad m_{BCij} = \sum_{k=1}^n b_{ik} c_{kj}$$

である。これを用いて

$$n_{1ij} = \sum_{k=1}^n m_{ABik} c_{kj} = \sum_{k=1}^n \left( \sum_{l=1}^m a_{il} b_{lk} \right) c_{kj} = \sum_{k=1}^n \sum_{l=1}^m (a_{il} b_{lk} c_{kj})$$

$$n_{2ij} = \sum_{k=1}^m a_{ik} m_{BCkj} = \sum_{k=1}^m a_{ik} \left( \sum_{l=1}^n b_{kl} c_{lj} \right) = \sum_{l=1}^n \sum_{k=1}^m (a_{ik} b_{kl} c_{lj})$$

上式の  $k$  と  $l$  を交換 (純粹に紙の上で文字として書き換える) しても何ら問題はない。そうすると上式は

$$\sum_{k=1}^n \sum_{l=1}^m (a_{il} b_{lk} c_{kj}) = n_{1ij}$$

となり結局  $n_{1ij}$  と一致する。 □

### V.2.2 上 (下) 三角行列のべき乗

上三角行列  $A$  の  $n$  乗もまた上三角行列であり、特に第  $i$  対角成分は  $a_{ii}^n$  である。下三角行列についても上と同様の定理が成り立つ。

*Proof.*

上三角行列について示す。下三角行列については同様に示せる。

$n = 1$  のときは明らか。まず  $n = 2$  のときを示す。

$$A^2[i][j] = \sum_{k=1}^n a_{ik}a_{kj} = \sum_{i \leq k, k \leq j} a_{ik}a_{kj} \quad (\because a_{ij} = 0 \text{ when } i > j)$$

であるから次の 2 つの式が成り立つ。

$$A^2[i][i] = \sum_{i \leq k, k \leq i} a_{ik}a_{ki} = a_{ii}^2, \quad A^2[i][j](i > j) = 0$$

次に、 $n$  まで成り立つと仮定して  $n + 1$  のとき成り立つことを示す。

$$\begin{aligned} A^{n+1} &= AA^n = \sum_{k=1}^n a_{ik}b_{kj} \quad (b_{kj} = A^n[k][j]) \\ &= \sum_{i \leq k, k \leq j} a_{ik}b_{kj} \quad (\because \text{帰納法の仮定より } b_{ij} = 0 \text{ when } i > j) \end{aligned}$$

であるから帰納法の仮定より次の 2 つの式が成り立つ。

$$\begin{aligned} A^{n+1}[i][i] &= \sum_{i \leq k, k \leq i} a_{ik}b_{ki} = a_{ii}b_{ii} = a_{ii}a_{ii}^n = a_{ii}^{n+1} \\ A^{n+1}[i][j](i > j) &= 0 \end{aligned}$$

□

## V.2.3 列ベクトルと行ベクトルの積の和

列ベクトル  $\mathbf{a}_i \in \mathbb{F}^{m \times 1}$  と行ベクトル  $\mathbf{b}_i \in \mathbb{F}^{1 \times n}$  の積を  $C_i \in \mathbb{F}^{m \times n}$  とするとき、

$$\sum_{i \in i_1, i_2, \dots, i_l} C_i = [\mathbf{a}_{i_1}, \mathbf{a}_{i_2}, \dots, \mathbf{a}_{i_l}] \begin{bmatrix} \mathbf{b}_{i_1} \\ \mathbf{b}_{i_2} \\ \dots \\ \mathbf{b}_{i_l} \end{bmatrix}$$

*Proof.* 行列のブロック演算を考えればわかる。

□

## V.2.4 巡回行列の可換則

*Proof.*

$A = [a_{ij}], B = [b_{ij}]$  を  $n$  次の巡回行列とすると、適当な定数  $c_0, \dots, c_{n-1}, d_0, \dots, d_{n-1}$  を用いて  $a_{ij} = c_{(i-j)\%n}, b_{ij} = d_{(i-j)\%n}$  と表せる。まず

$$(AB)[i][j] = \sum_{k=0}^{n-1} a_{ik}b_{kj} = \sum_{k=0}^{n-1} c_{(i-k)\%n}d_{(k-j)\%n}$$

であり、添字に剰余を用いていることと総和の範囲に注意して、

$$\begin{aligned}
 (BA)[i][j] &= \sum_{k=0}^{n-1} b_{ik} a_{kj} = \sum_{k=0}^{n-1} d_{(i-k)\%n} c_{(k-j)\%n} = \sum_{k=0}^{n-1} d_{(i-(k+i+j))\%n} c_{((k+i+j)-j)\%n} \\
 &= \sum_{k=0}^{n-1} d_{(-k-j)\%n} c_{(k+i)\%n} = \sum_{k=0}^{n-1} d_{(-(-k)-j)\%n} c_{((-k)+i)\%n} \\
 &= \sum_{k=0}^{n-1} c_{(i-k)\%n} b_{(k-j)\%n} = (AB)[i][j]
 \end{aligned}$$

□

## V.2.5 $\mathbb{R}^2$ の対称行列同士の積は対称

*Proof.*  $A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}, \quad B = \begin{bmatrix} c & d \\ d & c \end{bmatrix}$  とすると  $AB = \begin{bmatrix} ac+bd & ad+bc \\ ad+bc & ac+bd \end{bmatrix}$

□

注意.  $\mathbb{R}^3$  以上では成り立たない。

## V.2.6 Kronecker 積

### V.2.6.1 混合積: $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$

行列  $A, B, C, D$  を、積  $AC, BD$  が定義できるものとするとき  $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$  が成り立つ。

*Proof.*

$A$  の行数を  $m_A$ , 列数を  $n_A$ ,  $C$  の列数を  $n_C$  とする。

$$\begin{aligned}
 (A \otimes B)(C \otimes D) &= \begin{bmatrix} a_{1,1}B & a_{1,2}B & \cdots & a_{1,n_A}B \\ a_{2,1}B & & \ddots & \vdots \\ \vdots & & & \\ a_{m_A,1}B & \cdots & & a_{m_A,n_A}B \end{bmatrix} \begin{bmatrix} c_{1,1}D & c_{1,2}D & \cdots & c_{1,n_C}D \\ c_{2,1}D & & \ddots & \vdots \\ \vdots & & & \\ c_{n_A,1}D & \cdots & & c_{n_A,n_C}D \end{bmatrix} \\
 &= \begin{bmatrix} (AC)[1,1]BD & (AC)[1,2]BD & \cdots & (AC)[1,n_C]BD \\ (AC)[2,1]BD & & \ddots & \vdots \\ \vdots & & & \\ (AC)[m_A,1]BD & \cdots & & (AC)[m_A,n_C]BD \end{bmatrix} = (AC) \otimes (BD)
 \end{aligned}$$

□

### V.2.6.2 ユニタリ行列同士の Kronecker 積はユニタリ行列である

*Proof.*



$A \in \mathbb{C}^{m \times m}, B \in \mathbb{C}^{n \times n}$  をユニタリ行列とする。混合積の性質から

$$(A \otimes B)^*(A \otimes B) = (A^* \otimes B^*)(A \otimes B) = (A^*A) \otimes (B^*B) = I_m \otimes I_n = I_{m+n}$$

□

## 第 V.3 章

# ベクトル空間

### V.3.1 基底

#### V.3.1.1 基底変換

$V$  の基底  $\{v_1, \dots, v_n\}$  に以下の操作を行っても依然として基底である。

- $v_i$  を非零倍する
- $v_j$  に  $v_i$  の定数倍を加える

*Proof.*

操作後のベクトルの線形和を 0 にするためには全ての係数を 0 にせねばならないことに気付く。 □

#### V.3.1.2 諸定理

##### V.3.1.2.1 三角関数ベクトルによる直交展開

$\mathbb{C}^n$  を考える。  $n$  本のベクトルの集合

$$\left\{ v_1, \dots, v_n \mid v_k[l] = \frac{1}{\sqrt{n}} \exp\left(i \frac{k}{n} 2\pi l\right) \right\}$$

は  $\mathbb{C}^n$  の正規直交基底を成す。

*Proof.*

まず各ベクトルの 2 ノルムが 1 であることが次のようにして簡単に確かめられる。

$$\|v_k\|_2 = \frac{1}{\sqrt{n}} \sqrt{\sum_{l=1}^n \left| \exp\left(i \frac{k}{n} 2\pi l\right) \right|^2} = \frac{1}{\sqrt{n}} \sqrt{n} = 1$$

次に、相異なる  $k, l$  に対して  $v_k$  と  $v_l$  が直交することは次のようにして確かめられる。

$$\langle v_k, v_l \rangle = v_k^* v_l = \frac{1}{n} \sum_{m=1}^n \exp\left(-i \frac{k}{n} 2\pi m\right) \exp\left(i \frac{l}{n} 2\pi m\right) = \frac{1}{n} \sum_{m=1}^n \exp\left(i \frac{l-k}{n} 2\pi m\right) = 0 \quad (\because \text{IV.1.2.2})$$

□

## V.3.2 一般のベクトル空間と配列型ベクトル空間の橋渡し

### V.3.2.1 合成ベクトルの一次独立性と係数ベクトルの一次独立性は等価

$n$  次元ベクトル空間  $V$  の基底  $\mathbf{b}_1, \dots, \mathbf{b}_n$  の一次結合から成る  $p \leq n$  個のベクトル  $\mathbf{v}_i = c_{1i}\mathbf{b}_1 + \dots + c_{ni}\mathbf{b}_n$  ( $i = 1, \dots, p$ ) が一次独立であることと、その係数からなる  $p$  個のベクトル  $[c_{1i}, \dots, c_{ni}]^T$  ( $i = 1, \dots, p$ ) が一次独立であることは同値。

*Proof.*

まず

$$\begin{aligned}
 d_1 \mathbf{v}_1 + \dots + d_p \mathbf{v}_p &= \mathbf{0} & (1) \\
 \iff d_1(c_{11}\mathbf{b}_1 + \dots + c_{n1}\mathbf{b}_n) + \dots + d_p(c_{1p}\mathbf{b}_1 + \dots + c_{np}\mathbf{b}_n) &= \mathbf{0} \\
 \iff (d_1c_{11} + \dots + d_pc_{1p})\mathbf{b}_1 + \dots + (d_1c_{n1} + \dots + d_pc_{np})\mathbf{b}_n &= \mathbf{0} \\
 \iff d_1c_{11} + \dots + d_pc_{1p} = 0, \dots, d_1c_{n1} + \dots + d_pc_{np} = 0 & \text{ (}\mathbf{b}_1, \dots, \mathbf{b}_n\text{は基底)} \\
 \iff \begin{bmatrix} c_{11} & \cdots & c_{1p} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{np} \end{bmatrix} \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} &= \mathbf{0}_n & (2)
 \end{aligned}$$

が成り立つから、 $\mathbf{v}_1, \dots, \mathbf{v}_p$  が一次独立なら (1) より  $d_1 = \dots = d_n = 0$  であるが、これが (2)  $\Rightarrow d_1 = \dots = d_n = 0$  と等価なので、 $[c_{1i}, \dots, c_{ni}]^T$  ( $i = 1, \dots, p$ ) が一次独立であることになる。逆に  $[c_{1i}, \dots, c_{ni}]^T$  ( $i = 1, \dots, p$ ) が一次独立なら (2) より  $d_1 = \dots = d_n = 0$  であるが、これが (1)  $\Rightarrow d_1 = \dots = d_n = 0$  と等価なので、 $\mathbf{v}_1, \dots, \mathbf{v}_p$  が一次独立であることになる。  $\square$

## V.3.3 諸定理

### V.3.3.1 ベクトル空間 $V$ の部分空間 $V_1, V_2$ について、「 $V_1 \cup V_2$ がベクトル空間である」 $\iff$ 「 $V_1 \subseteq V_2$ または $V_1 \supseteq V_2$ 」

*Proof.*

( $\Leftarrow$ ) 容易。

( $\Rightarrow$ ) 背理法で示す。  $\exists \mathbf{v}_1 \in V_1 \setminus V_2, \mathbf{v}_2 \in V_2 \setminus V_1$  と仮定する。  $V_1 \cup V_2$  がベクトル空間であるから  $\mathbf{v}_1 + \mathbf{v}_2 \in V_1 \cup V_2$  である。  $\mathbf{v}_1 + \mathbf{v}_2 \in V_1$  であるとすると

$$\mathbf{v}_2 = \underbrace{(\mathbf{v}_1 + \mathbf{v}_2)}_{\in V_1} - \mathbf{v}_1$$

となり、左辺  $\notin V_1$  であるが、右辺  $\in V_1$  となり矛盾する。同様に  $\mathbf{v}_1 + \mathbf{v}_2 \in V_2$  であるとしても矛盾が生じる。従って  $V_1 \cup V_2$  がベクトル空間であることに矛盾する。  $\square$

$$V.3.3.2 \quad V_i \subseteq W_i, \bigoplus_{i=1}^n W_i = \bigoplus_{i=1}^n V_i \Rightarrow V_i = W_i$$

$$\bigoplus_{i=1}^n W_i = \bigoplus_{i=1}^n V_i, V_i \subseteq W_i \quad (i = 1, 2, \dots, n) \Rightarrow W_i = V_i \quad (i = 1, 2, \dots, n)$$

*Proof.*

$W_i \subseteq V_i$  を示せば良い。 $x \in W_i$  を任意にとると  $x \in \bigoplus_{i=1}^n W_i = \bigoplus_{i=1}^n V_i$  でもあるから適当なベクトル  $v_i \in V_i$  が存在して

$$x = v_1 + v_2 + \cdots + v_i + \cdots + v_n$$

と一意に分解できるが、 $V_i \subseteq W_i$  より  $v_i \in W_i$  であるので

$$x = \underbrace{v_1}_{\in W_1} + \underbrace{v_2}_{\in W_2} + \cdots + \underbrace{v_i}_{\in W_i} + \cdots + \underbrace{v_n}_{\in W_n}$$

ところで  $x \in \bigoplus_{i=1}^n W_i$  であったから  $x$  は

$$x = \underbrace{0}_{\in W_1} + \underbrace{0}_{\in W_2} + \cdots + \underbrace{x}_{\in W_i} + \cdots + \underbrace{0}_{\in W_n}$$

と分解される。分解は一意でなくてはならぬから

$$v_j = \begin{cases} x & (j = i) \\ 0 & (\text{otherwise}) \end{cases}$$

よって  $x \in V_i$

□

V.3.3.3 一次独立なベクトル  $v_1, \dots, v_k \in V$  に  $v \in V$  を加えたものが一次独立  $\iff v$  が  $v_1, \dots, v_k$  の一次結合で表せない

$V$  を線形ベクトル空間とし、 $v_1, \dots, v_k \in V$  が一次独立であるとする。この時、別のベクトル  $v \in V$  を付け加えたものが一次独立であることは、 $v$  が  $v_1, \dots, v_k$  の一次結合で表せないことと同値である。

*Proof.*

$v_1, \dots, v_k, v$  が一次独立なら、 $v$  が  $v_1, \dots, v_k$  の一次結合で表せないことは一次独立性の定義から明らか。逆に  $v$  が  $v_1, \dots, v_k$  の一次結合で表せないとする。このとき  $c_1 v_1 + \cdots + c_k v_k + cv = 0$  とおくと、 $c_1 v_1 + \cdots + c_k v_k = -cv$  となるが、もし  $c \neq 0$  とすると両辺を  $-c$  で割れば  $v$  が  $v_1, \dots, v_k$  の一次結合で表せることになってしまい、仮定に矛盾するので  $c = 0$  でなければならない。よって  $c_1 v_1 + \cdots + c_k v_k = 0$  であるが、 $v_1, \dots, v_k$  は一次独立であるから  $c_1 = \cdots = c_k = 0$  である。すなわち  $c_1 = \cdots = c_k = c = 0$  なので  $v_1, \dots, v_k, v$  は一次独立である。

□

V.3.3.4  $A_v := \{v_1, \dots, v_k\}$  と  $A_w := \{w_1, \dots, w_l\}$  が各々一次独立で  $k > l \Rightarrow \exists v \in A_v$  s.t.  $w_1, \dots, w_l, v$  が一次独立

線形空間  $V$  のベクトルの集合  $A_v := \{v_1, \dots, v_k\}$  と  $A_w := \{w_1, \dots, w_l\}$  がそれぞれ一次独立で  $k > l$  であれば、ある  $v \in A_v$  が存在して  $w_1, \dots, w_l, v$  が一次独立である。

*Proof.*

背理法で示す。任意の  $v \in A_v$  に対して  $w_1, \dots, w_l, v$  が一次従属であると仮定すると、V.3.3.3 より、任意の  $v \in A_v$  が  $w_1, \dots, w_l$  の一次結合で表せる。よって  $A_v$  は  $l$  個の一次独立なベクトルの一次結合で表された  $k > l$  個以上のベクトルの集合なので一次従属であるが、これは  $A_v$  の一次独立性の仮定に矛盾する。  $\square$

## V.3.4 諸注意

V.3.4.1  $\{v_1, v_2\}, \{w_1, w_2\}, \{v_1, v_2, w_1\}, \{v_1, v_2, w_2\}$  が各々一次独立でも  $\{v_1, v_2, w_1, w_2\}$  が一次独立とは限らない

*Proof.* 例えば  $\mathbb{R}^4$  で考えて  $v_1 = e_1, v_2 = e_2, w_1 = e_3, w_2 = e_1 + e_2 + e_3$   $\square$

V.3.4.2  $W_i \cap W_j = \{0\}$  ( $i \in \{1, \dots, n\}, i \neq j$ ) であっても  $W_1 + \dots + W_n = W_1 \oplus \dots \oplus W_n$  とは限らない

*Proof.*

例えば  $W_1, W_2, W_3 \subseteq \mathbb{R}^2$ ,  $W_1 = \{e_1\}, W_2 = \{e_2\}, W_3 = \{e_1 + e_2\}$  とするとベクトル  $e_1 + e_2$  に対して分解が一意でない ( $W_1, W_2, W_3$  から  $e_1, e_2, 0$  を持ってくる方法と、 $W_1, W_2, W_3$  から  $0, 0, e_1 + e_2$  を持ってくる方法がある)。  $\square$

V.3.4.3  $W = V \oplus X, W = V \oplus Y$  でも  $X = Y$  とは限らない

*Proof.* 例えば  $W = \mathbb{R}^2$ ,  $V = \{[1, 0]^\top\}$ ,  $X = \{[0, 1]^\top\}$ ,  $Y = \{[1, 1]^\top\}$   $\square$

V.3.4.4  $\text{span}[a, b, c] / \text{span}[a, b] = \text{span}[c]$  とは限らない

*Proof.*

例えば  $a = [1, 0, 0]^\top$ ,  $b = [0, 1, 0]^\top$ ,  $c = [0, 0, 1]^\top$ 、すなわち  $\text{span}[a, b, c] = \mathbb{R}^3$  のとき、 $\text{span}[a, b, c] / \text{span}[a, b]$  は  $\mathbb{R}^3$  から第1要素が0である点だけを抜いたものであり、 $[x, *, *]^\top$ ,  $x \neq 0$  のような点が残っている。このような点は  $\text{span}[c]$  には属さない。  $\square$

## 第 V.4 章

# 線形写像

### V.4.1 表記の規則

#### V.4.1.1 一般のベクトルを要素とする配列と行列の積

$\mathbf{v}_1, \dots, \mathbf{v}_n$  をベクトルとする ( $\mathbb{R}^n$  や  $\mathbb{C}^n$  の要素とは限らないことに注意)。これと  $m \times n$  行列  $A = [a_{ij}]_{m \times n}$ ,  $n \times m$  行列  $B = [b_{ij}]_{n \times m}$  に対して次のように演算を定義する。

$$A \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{bmatrix} := \begin{bmatrix} \sum_{j=1}^n a_{1j} \mathbf{v}_j \\ \vdots \\ \sum_{j=1}^n a_{mj} \mathbf{v}_j \end{bmatrix}, \quad [\mathbf{v}_1, \dots, \mathbf{v}_n] B := \left[ \sum_{i=1}^n b_{i1} \mathbf{v}_i, \dots, \sum_{i=1}^n b_{im} \mathbf{v}_i \right]$$

これは  $\mathbf{v}_1, \dots, \mathbf{v}_n$  を形式的にスカラーと見做して普通の行列同士の積を考えたものである。

#### V.4.1.2 一般のベクトルを要素とする行列同士の積

ベクトル  $\mathbf{a}_{ij}$  ( $(i, j) \in \{1 : l\} \times \{1 : m\}$ ) と  $\mathbf{b}_{ij}$  ( $(i, j) \in \{1 : m\} \times \{1 : n\}$ ) に対して、これらを要素とする行列  $A := [\mathbf{a}_{ij}]_{l \times m}$ ,  $B := [\mathbf{b}_{ij}]_{m \times n}$  の積  $C$  を  $C := [c_{ij}]_{l \times n}$ ,  $c_{ij} := \sum_{k=1}^m \langle \mathbf{a}_{ik}, \mathbf{b}_{kj} \rangle$  で定義する。これは  $\mathbf{a}_{ij}, \mathbf{b}_{ij}$  を形式的にスカラーと見做して普通の行列同士の積を考えたものである。

### V.4.2 $\dim(V) = \dim(W)$ , $f : V \rightarrow W$ であるとき、 $f$ が単射であることと全射であることは同値。

$f$  を  $V \rightarrow W$  への線形写像とする。  $\dim(V) = \dim(W)$  であるとき、 $f$  が単射であることと全射であることは同値。

*Proof.*

$n = \dim(V) = \dim(W)$  とする。

( $\Rightarrow$ )

$W$  の任意のベクトルは  $W$  の基底の一次結合で表されるから、全射性を言うためには  $V$  の中から適当な  $n$  個のベクトルを選んで  $f$  で送ったものの集合で  $W$  の基底を構成できることを示せば十分である。 $V$  の基底を  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  として  $\{f(\mathbf{v}_1), f(\mathbf{v}_2), \dots, f(\mathbf{v}_n)\}$  が  $W$  の基底を成すことを示す。

$$c_1 f(\mathbf{v}_1) + c_2 f(\mathbf{v}_2) + \dots + c_n f(\mathbf{v}_n) = \mathbf{0}, \quad c_1, c_2, \dots, c_n \in \mathbb{F}$$

とおくと  $f$  の線形性から

$$f(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) = \mathbf{0}$$

$f$  は単射であったから  $\text{Ker}(f) = \{\mathbf{0}\}$  なので

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

$\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  は  $V$  の基底なので

$$c_1 = c_2 = \cdots = c_n = 0$$

よって  $W$  の  $n = \dim(W)$  個のベクトル  $\{f(\mathbf{v}_1), f(\mathbf{v}_2), \dots, f(\mathbf{v}_n)\}$  が一次独立なのでこれは  $W$  の基底である。

( $\Leftarrow$ )

次元定理から

$$\dim(\text{Img}(f)) + \dim(\text{Ker}(f)) = \dim(V) = \dim(W) = n$$

$f$  は全射であったから  $\text{Img}(f) = W$  であるので  $\dim(\text{Img}(f)) = n$ 。よって上式より  $\dim(\text{Ker}(f)) = 0$  なので  $f$  は単射である。

(別証)

$f$  が全射であるから、 $W$  の任意の基底  $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$  に対して適当な  $V$  のベクトル  $\mathbf{v}_1, \dots, \mathbf{v}_n$  が存在して  $f(\mathbf{v}_i) = \mathbf{w}_i$  ( $i = 1, \dots, n$ ) を満たす。ここで  $c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$  とおくと、 $f$  は線形写像なので  $f(\mathbf{0}) = \mathbf{0}$  だから  $\mathbf{0} = f(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1f(\mathbf{v}_1) + \cdots + c_nf(\mathbf{v}_n) = c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n$  となる。 $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$  は  $W$  の基底であったから  $c_1 = \cdots = c_n = 0$  である。よって  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  は一次独立であり、 $\dim V = n$  よりこれらは  $V$  の基底である。

$\mathbf{x} \in V$  に対して  $f(\mathbf{x}) = \mathbf{0}$  とおく。 $\mathbf{x}$  を先程の  $V$  の基底で展開して  $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$  とすると、

$$\mathbf{0} = f(\mathbf{x}) = f(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n$$

より  $c_1 = \cdots = c_n = 0$  となるから  $\mathbf{x} = \mathbf{0}$  である。すなわち  $\text{Ker}(f) = \{\mathbf{0}\}$  なので  $f$  は単射である。  $\square$

### V.4.3 $\dim V = \dim W$ ならば全単射線形写像 $f: V \rightarrow W$ が存在する

*Proof.*

$n = \dim V = \dim W$  とし、 $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  を  $V$  の基底、 $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$  を  $W$  の基底とする。写像  $f: V \rightarrow W$  を  $f(\mathbf{v}_i) = \mathbf{w}_i$  ( $i = 1, \dots, n$ )、 $\forall \mathbf{x}, \mathbf{y} \in V, \alpha, \beta \in \mathbb{R}, f(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y})$  で定めると  $f$  は線形写像である。さらに単射であることが次のようにして確かめられる。

$\mathbf{x}, \mathbf{y} \in V$  に対して  $f(\mathbf{x}) = f(\mathbf{y})$  であるとする。 $\mathbf{x}, \mathbf{y}$  を基底で展開して  $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{v}_i, \mathbf{y} = \sum_{i=1}^n y_i \mathbf{v}_i$  とすると  $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{w}_i = \mathbf{y} = \sum_{i=1}^n y_i \mathbf{w}_i$  すなわち  $\mathbf{x} = \sum_{i=1}^n (x_i - y_i) \mathbf{w}_i = \mathbf{0}$  であるが、 $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$  は  $W$  の基底であり一次独立だから  $x_i = y_i$  ( $i = 1, \dots, n$ ) であるので  $\mathbf{x} = \mathbf{y}$  である。直前の定理より  $f$  は全射でもあるから全単射である。  $\square$

## V.4.4 不変部分空間と表現行列

$A \in \mathbb{F}^{n \times n}$ ,  $W_1, \dots, W_m \in \mathbb{F}^n : A$  の不変部分空間,  $\mathbb{F}^n = W_1 \oplus \dots \oplus W_m$ ,  $d_i := \dim W_i$   
 $\{\mathbf{w}_{i1}, \dots, \mathbf{w}_{id_i}\} : W_i$  の基底,  $\Phi_i := [\mathbf{w}_{i1}, \dots, \mathbf{w}_{id_i}]$ ,  $\Phi := [\Phi_1, \dots, \Phi_m]$   
 以上の設定の下で、次式を満たす適当な正方行列  $A_i \in \mathbb{F}^{d_i \times d_i}$  が存在する。

$$A = \Phi \operatorname{diag}(A_1, \dots, A_m) \Phi^{-1}$$

$A_i$  は次式で得られる。

$$A_i = \left( \Phi_i^\top \Phi_i \right)^{-1} \Phi_i^\top A \Phi_i$$

特に  $W_i$  の基底を正規直交基底にしておけば、より簡単に次式で得られる。

$$A_i = \Phi_i^\top A \Phi_i$$

*Proof.*

$A = \Phi \operatorname{diag}(A_1, \dots, A_m) \Phi^{-1}$  を満たす  $A_i$  が存在することは [2] 定理 1.20 を見ればわかる。 $\Phi$  は  $\mathbb{F}$  から来る入力ベクトルを  $A$  の不変部分空間の成分に分解する変換器になっている。

$A_i$  の構成法を示す。一般の場合を示すのは別に難しくないが、式が煩雑になって無意味に難解になるので、具体的な場合を示す。例えば  $n = 4, m = 2, d_1 = d_2 = 2$  としてみる。この場合の証明が理解できれば一般の証明は楽勝である。まず

$$A = \Phi \operatorname{diag}(A_1, \dots, A_m) \Phi^{-1} = [\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4] \operatorname{diag}(A_1, A_2) [\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4]^{-1}$$

とおく。まず  $A_1$  を求めてみる。定理 V.7.1 からヒントを得て両辺に  $[\mathbf{w}_1, \mathbf{w}_2]$  を右から掛けて

$$\begin{aligned} A[\mathbf{w}_1, \mathbf{w}_2] &= [\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4] \operatorname{diag}(A_1, A_2) \begin{bmatrix} I_2 \\ O_2 \end{bmatrix} = [\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4] \begin{bmatrix} A_1 \\ O_2 \end{bmatrix} \\ &= [\mathbf{w}_1, \mathbf{w}_2] A_1 \end{aligned}$$

$[\mathbf{w}_1, \mathbf{w}_2]$  は列フルランクであるから定理 V.10.5 より  $[\mathbf{w}_1, \mathbf{w}_2]^\top [\mathbf{w}_1, \mathbf{w}_2]$  は逆行列を持つ。そこで上式に左から  $[\mathbf{w}_1, \mathbf{w}_2]^\top$  を掛けて

$$\begin{aligned} [\mathbf{w}_1, \mathbf{w}_2]^\top A[\mathbf{w}_1, \mathbf{w}_2] &= [\mathbf{w}_1, \mathbf{w}_2]^\top [\mathbf{w}_1, \mathbf{w}_2] A_1 \\ \therefore A_1 &= ([\mathbf{w}_1, \mathbf{w}_2]^\top [\mathbf{w}_1, \mathbf{w}_2])^{-1} [\mathbf{w}_1, \mathbf{w}_2]^\top A[\mathbf{w}_1, \mathbf{w}_2] \end{aligned}$$

特に、予め  $\{\mathbf{w}_1, \mathbf{w}_2\}$  を正規直交基底にしておけば  $[\mathbf{w}_1, \mathbf{w}_2]^\top [\mathbf{w}_1, \mathbf{w}_2] = I_2$  なので  $A_1$  は次式で得られる。

$$A_1 = [\mathbf{w}_1, \mathbf{w}_2]^\top A[\mathbf{w}_1, \mathbf{w}_2]$$

同様に  $A_2$  も求められる。一般の場合も同じ要領でやれば良い。 □

### V.4.4.1 正規直交基底間での座標変換は直交変換

$n$  次元線形空間  $V$  の 2 つの正規直交基底  $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ ,  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  に対して、 $\{\mathbf{u}_i\}$  座標系から  $\{\mathbf{v}_i\}$  座標系への座標変換は直交変換である。すなわち座標変換行列  $P := [\langle \mathbf{u}_i, \mathbf{u}_j \rangle]_{n \times n}$  は直交行列である。



*Proof.*

$$\mathbf{v}_i = \sum_{k=1}^n \langle \mathbf{v}_i, \mathbf{u}_k \rangle \mathbf{u}_k = \sum_{k=1}^n P[i][k] \mathbf{u}_k$$

だから

$$\begin{aligned} \delta_{ij} = \langle \mathbf{v}_i, \mathbf{v}_j \rangle &= \left\langle \sum_{k=1}^n P[i][k] \mathbf{u}_k, \sum_{l=1}^n P[j][l] \mathbf{u}_l \right\rangle = \sum_{k=1}^n \sum_{l=1}^n \langle P[i][k] \mathbf{u}_k, P[j][l] \mathbf{u}_l \rangle \\ &= \sum_{k=1}^n \sum_{l=1}^n P[i][k] P[j][l] \delta_{kl} = \sum_{k=1}^n P[i][k] P[j][k] = \sum_{k=1}^n P[i][k] P^\top[k][j] = (PP^\top)[i][j] \end{aligned}$$

□

## 第 V.5 章

# 双対空間

### V.5.1 基底変換と座標変換の関係

$V$  を  $n$  次元ベクトル空間とし ( $\mathbb{R}^n$  や  $\mathbb{C}^n$  とは限らないことに注意)、 $\{v_1, \dots, v_n\}$  と  $\{v_1', \dots, v_n'\}$  を  $V$  の基底とする。これらが  $n$  次正則行列  $A$  により次の関係式に従うとする。

$$\begin{bmatrix} v_1' \\ \vdots \\ v_n' \end{bmatrix} = A \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

$\{v_1, \dots, v_n\}$ ,  $\{v_1', \dots, v_n'\}$  の双対基底をそれぞれ  $\{v^1, \dots, v^n\}$ ,  $\{v^{1'}, \dots, v^{n'}\}$  とし、 $\underline{x} := [x_1, \dots, x_n]^\top$ ,  $\underline{x}' := [x_1', \dots, x_n']^\top$  とする。このとき  $[v^{1'}, \dots, v^{n'}]\underline{x}' = [v^1, \dots, v^n]\underline{x}$  であれば  $\underline{x}' = A\underline{x}$  である。すなわち、双対空間での座標変換と元の空間の基底間の線形変換が同じ形になる。

*Proof.*

$$[v^{1'}, \dots, v^{n'}]\underline{x}' = [v^1, \dots, v^n]\underline{x} \text{ より}$$

$$\underline{x}' = \begin{bmatrix} v_1' \\ \vdots \\ v_n' \end{bmatrix} [v^1, \dots, v^n]\underline{x} = A \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} [v^1, \dots, v^n]\underline{x} = AI_3\underline{x} = A\underline{x}$$

□

## 第 V.6 章

# 行列式

V.6.1  $A \in \mathbb{R}^{n \times m}, B \in \mathbb{R}^{m \times n}$  に対して  $|I_m - AB| = |I_n - BA|$

*Proof.*

$$M := \begin{bmatrix} I_n & A \\ B & I_m \end{bmatrix}$$

とすると

$$\det M = \det \left( M \begin{bmatrix} I_n & -A \\ O & I_m \end{bmatrix} \right) = \det \begin{bmatrix} I_n & O \\ B & I_m - BA \end{bmatrix} = |I_m - BA|$$

一方で

$$\det M = \det \left( M \begin{bmatrix} I_n & O \\ -B & I_m \end{bmatrix} \right) = \det \begin{bmatrix} I_n - AB & A \\ O & I_m \end{bmatrix} = |I_n - AB|$$

□

V.6.2 系:  $|I_n - \mathbf{v}\mathbf{v}^\top| = 1 - \|\mathbf{v}\|_2^2$

$$\mathbf{v} \in \mathbb{C}^n \text{ に対して } |I_n - \mathbf{v}\mathbf{v}^\top| = 1 - \|\mathbf{v}\|_2^2$$

V.6.3 逆対角転置行列の行列式は元の行列のそれと等しい

*Proof.*

$A$  を  $n$  次対角行列とし、その逆対角転置を  $A'$  とする。 $S_n$  を  $n$  次の置換全体の集合とする。

$$\begin{aligned} |A'| &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n A'[i, \sigma(i)] = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n A[n - \sigma(i) + 1, n - i + 1] \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^n A[j, n - \sigma^{-1}(n - j + 1) + 1] \\ &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^n A[j, \tilde{\sigma}(j)] \quad (\tilde{\sigma} : j \mapsto n + 1 - \sigma^{-1}(n - j + 1)) \end{aligned}$$

ここで逆置換の符号が元の置換の符号と等しいこと、および [I.7.1.1](#), [I.7.1.2](#) より  $\operatorname{sgn}(\tilde{\sigma}) = s^2 \operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma)$  であるから

$$|A'| = \sum_{\sigma \in S_n} \operatorname{sgn}(\tilde{\sigma}) \prod_{j=1}^n A[j, \tilde{\sigma}(j)]$$

$\sigma$  と  $\tilde{\sigma}$  は一対一に対応するので  $\sigma$  が  $S_n$  全体を動くとき  $\tilde{\sigma}$  もそうだから上式の  $\sum_{\sigma \in S_n}$  を  $\sum_{\tilde{\sigma} \in S_n}$  で置き換えて

$$|A'| = \sum_{\tilde{\sigma} \in S_n} \operatorname{sgn}(\tilde{\sigma}) \prod_{j=1}^n A[j, \tilde{\sigma}(j)] = |A|$$

□

## 第 V.7 章

# 逆行列

### V.7.1 $[\phi_1, \dots, \phi_n]^{-1} \phi_i = e_i$

$\phi_1, \dots, \phi_n \in \mathbb{F}^n$  が一次独立であるとき、 $[\phi_1, \dots, \phi_n]^{-1} \phi_i = e_i$

*Proof.*

$\Phi := [\phi_1, \dots, \phi_n]$  とする。次式より直ちに定理が証明される。

$$[e_1, \dots, e_n] = I_n = \Phi^{-1} \Phi = \Phi^{-1} [\phi_1, \dots, \phi_n] = [\Phi^{-1} \phi_1, \dots, \Phi^{-1} \phi_n]$$

□

この定理から直ちに次の系が得られる。

### V.7.2 (系) 結合係数の抽出

$\phi_1, \dots, \phi_n \in \mathbb{F}^n$  は一次独立であるとする。このとき、任意に与えられたベクトル  $x = \sum_{i=1}^n c_i \phi_i \in \mathbb{F}^n$  ( $\phi_1, \dots, \phi_n$  は  $\mathbb{F}^n$  の基底だから  $x$  は必ずこのように分解できる!) から  $c_1, \dots, c_n$  を抽出するには左から  $[\phi_1, \dots, \phi_n]^{-1}$  を掛ければよい。詳しく言うと、次式が成り立つ。

$$[c_1, \dots, c_n]^T = [\phi_1, \dots, \phi_n]^{-1} x$$

### V.7.3 上 (下) 三角行列が逆行列をもてば、それも上 (下) 三角行列である

*Proof.*

$A \in \mathbb{C}^{n \times n}$  を正則な上三角行列とすると、 $A$  の対角成分は全て非零である (さもなくば  $\det A = 0$  となり、正則でない) ので、 $[A|I_n]$  の簡約化による逆行列の計算を考えれば、 $A^{-1}$  も上三角行列であることが容易にわかる。 $A$  が下三角行列である場合は、 $A^T$  について簡約化による逆行列の計算を行えば  $(A^T)^{-1} = (A^{-1})^T$  が上三角行列になり、 $A^{-1}$  は下三角行列になる。

□

### V.7.4 正則行列 $A$ の列 (行) が直交系を成すとき、 $A^{-1}$ の行 (列) は直交系を成す

*Proof.*

$A$  の列が直交系を成すとする。 $A^{-1} (A^{-1})^{\top} = A^{-1} (A^{\top})^{-1} = (A^{\top} A)^{-1}$  であるが、仮定より  $A^{\top} A$  が正則な対角行列であるので  $(A^{\top} A)^{-1}$  は対角行列になり、定理の主張が成り立つ。 $A$  の行が直交系を成すときに  $A^{-1}$  の列が直交系を成すことも同様に示せる。  $\square$

### V.7.5 1 列 (or 1 行) だけ置き換えた行列の逆行列

$A = [\mathbf{a}_1, \dots, \mathbf{a}_n] \in \mathbb{C}^{n \times n}$  を正則な行列とし、 $c$  列目を置き換えた行列  $A' = [\mathbf{a}_1, \dots, \mathbf{a}'_c, \dots, \mathbf{a}_n]$  も正則であるとする。 $A^{-1}$  が既知であれば  $A'^{-1}$  を以下のようにして少ない手間で計算できる。

まず  $A^{-1} A'$  は  $c$  列目以外は単位行列と同じである、次のような行列になる。

$$A^{-1} A' = [A^{-1} \mathbf{a}_1, \dots, A^{-1} \mathbf{a}_{c-1}, A^{-1} \mathbf{a}'_c, A^{-1} \mathbf{a}_{c+1}, \dots, A^{-1} \mathbf{a}_n] = \begin{bmatrix} & p_1 & & \\ & \vdots & & \\ \mathbf{e}_1 \cdots \mathbf{e}_c & p_c & \mathbf{e}_{c+1} \cdots \mathbf{e}_n & \\ & \vdots & & \\ & p_n & & \end{bmatrix}$$

ここで  $p_c \neq 0$  である (もしそうでなければ  $c$  列目が他の列の一次結合で表せるので非正則になるが、これは左辺が正則行列同士の積で正則であることに矛盾する)。右辺の  $c$  列目を  $\mathbf{e}_c$  に変形するためには、両辺に次の行基本変形行列を左から掛ければよい。

$$M = \begin{bmatrix} & -p_1/p_c & & \\ & \vdots & & \\ & -p_{c-1}/p_c & & \\ \mathbf{e}_1 \cdots \mathbf{e}_c & 1/p_c & \mathbf{e}_{c+1} \cdots \mathbf{e}_n & \\ & -p_{c+1}/p_c & & \\ & \vdots & & \\ & -p_n/p_c & & \end{bmatrix}$$

そうすると  $MA^{-1} A' = I_n$  となるので  $A'^{-1} = MA^{-1}$  である。要するに  $A^{-1} \mathbf{a}'_c$  より  $p_1, \dots, p_n$  を計算して  $A^{-1}$  の第  $c$  行を  $p_c$  で割った後、第  $i \neq c$  行から第  $c$  行の  $p_i$  倍を引くだけで  $A'^{-1}$  が計算できる。

上と同様にして、1 行だけ置き換えた行列の逆行列も、元の逆行列が既知であれば少ない手間で計算できる。

$$A = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_c \\ \vdots \\ \mathbf{a}_n \end{bmatrix} \in \mathbb{C}^{n \times n} \text{ を正則行列とし、第 } c \text{ 行を入れ替えてできる行列 } A' = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}'_c \\ \vdots \\ \mathbf{a}_n \end{bmatrix} \in \mathbb{C}^{n \times n} \text{ も正則であるとする}$$

る。まず  $A'A^{-1}$  は第  $c$  列以外は単位行列と同じである、次のような行列になる。

$$A'A^{-1} = \begin{bmatrix} \mathbf{a}_1 A^{-1} \\ \vdots \\ \mathbf{a}_{c-1} A^{-1} \\ \mathbf{a}'_c A^{-1} \\ \mathbf{a}_{c+1} A^{-1} \\ \vdots \\ \mathbf{a}_n A^{-1} \end{bmatrix} = \begin{bmatrix} & & \mathbf{e}_1 & & \\ & & \vdots & & \\ & & \mathbf{e}_{c-1} & & \\ p_1 & \cdots & p_c & \cdots & p_n \\ & & \mathbf{e}_{c+1} & & \\ & & \vdots & & \\ & & \mathbf{e}_n & & \end{bmatrix}$$

ここで  $p_c \neq 0$  である (もしそうでなければ第  $c$  行が他の行の一次結合で表せるので非正則になるが、これは左辺が正則行列同士の積で正則であることに矛盾する)。右辺の第  $c$  行を  $\mathbf{e}_c$  に変形するためには、両辺に次の列基本変形行列を右から掛ければよい。

$$M = \begin{bmatrix} & & \mathbf{e}_1 & & \\ & & \vdots & & \\ & & \mathbf{e}_{c-1} & & \\ -p_1/p_c & \cdots & -p_{c-1}/p_c & 1/p_c & -p_{c+1}/p_c & \cdots & p_n \\ & & \mathbf{e}_{c+1} & & \\ & & \vdots & & \\ & & \mathbf{e}_n & & \end{bmatrix}$$

そうすると  $A'A^{-1}M = I_n$  となるので  $A'^{-1} = A^{-1}M$  である。要するに  $\mathbf{a}'_c A^{-1}$  より  $p_1, \dots, p_n$  を計算して  $A^{-1}$  の第  $c$  列を  $p_c$  で割った後、第  $i \neq c$  列から第  $c$  行の  $p_i$  倍を引くだけで  $A'^{-1}$  が計算できる。

## V.7.6 列 (or 行) を入れ替えた行列の逆行列

正方行列  $A$  の幾つかの列 (行) を入れ替えた行列  $A'$  の逆行列は、 $A^{-1}$  の行 (列) に対して同じ並べ替えを適用して得られる。

*Proof.*

まず列の置換に関して定理の主張を示す。 $A$  を  $n$  次正方行列とし、列の置換を表す列置換行列を  $\Pi = \Pi_1 \Pi_2 \cdots \Pi_c$  とする。ここに  $\Pi_1, \Pi_2, \dots, \Pi_c$  は単位列置換行列である。すると、 $A'^{-1} = (A\Pi)^{-1} = \Pi^{-1}A^{-1} = \Pi_c^{-1} \cdots \Pi_2^{-1} \Pi_1^{-1} A^{-1} = \Pi_c \cdots \Pi_2 \Pi_1 A^{-1}$  となるから定理の主張が成り立つ。

同様にして行の置換に関しても示せる。行の置換を表す行置換行列を  $\Pi = \Pi_c \cdots \Pi_2 \Pi_1$  とする ( $\Pi_1, \Pi_2, \dots, \Pi_c$  は単位行置換行列) と、 $A'^{-1} = (\Pi A)^{-1} = A^{-1} \Pi^{-1} = A^{-1} \Pi_1^{-1} \Pi_2^{-1} \cdots \Pi_c^{-1} = A^{-1} \Pi_1 \Pi_2 \cdots \Pi_c$  □

## V.7.7 Sherman-Morrison の公式の特別な場合:

$$(I + \mathbf{u}\mathbf{v}^*)^{-1} = (I - \mathbf{u}\mathbf{v}^*/(1 + \mathbf{v}^*\mathbf{u}))$$

$\mathbf{u}, \mathbf{v} \in \mathbb{C}^m$ ,  $1 + \mathbf{v}^*\mathbf{u} \neq 0$  とするとき次式が成り立つ。

$$(I + \mathbf{u}\mathbf{v}^*)^{-1} = I - \frac{1}{1 + \mathbf{v}^*\mathbf{u}} \mathbf{u}\mathbf{v}^*$$

*Proof.*

Woodbury の公式の特別な場合として片付けるのが一般的だが、ここでは Jordan 分解を用いて導出してみる。 $\mathbf{uv}^* = 0$  のときは定理の主張は明らかに成り立つ。以下では  $\mathbf{uv}^* \neq 0$  とする。仮定  $1 + \mathbf{v}^* \mathbf{u} \neq 0$  と V.6.1 より  $I + \mathbf{uv}^*$  には逆行列が存在する。 $\mathbf{uv}^*$  の階数は 1 であり ( $\text{Im}(\mathbf{uv}^*) = \text{span}[\mathbf{u}]$ )、唯一の固有値は  $\mathbf{v}^* \mathbf{u}$  である (対応する固有ベクトルは  $\mathbf{u}$ )。よって  $\mathbf{uv}^*$  を Jordan 分解すると、適当な正則行列  $J$  を用いて次のように表せる。

$$\mathbf{uv}^* = \begin{bmatrix} \mathbf{v}^* \mathbf{u} & \mathbf{0}^\top \\ \mathbf{0} & O \end{bmatrix} =: J \Lambda J^{-1}$$

よって

$$\begin{aligned} (I + \mathbf{uv}^*)^{-1} &= J(I + \Lambda)^{-1} J^{-1} = J \begin{bmatrix} 1 + \mathbf{v}^* \mathbf{u} & \mathbf{0}^\top \\ \mathbf{0} & I \end{bmatrix}^{-1} J^{-1} = J \begin{bmatrix} \frac{1}{1 + \mathbf{v}^* \mathbf{u}} & \mathbf{0}^\top \\ \mathbf{0} & I \end{bmatrix} J^{-1} \\ &= J \frac{1}{1 + \mathbf{v}^* \mathbf{u}} \begin{bmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & (1 + \mathbf{v}^* \mathbf{u})I \end{bmatrix} J^{-1} = J \frac{1}{1 + \mathbf{v}^* \mathbf{u}} ((1 + \mathbf{v}^* \mathbf{u})I - \Lambda) J^{-1} \\ &= I - \frac{1}{1 + \mathbf{v}^* \mathbf{u}} \mathbf{uv}^* \end{aligned}$$

□

### V.7.8 Sherman-Morrison の公式:

$$(A + \mathbf{uv}^*)^{-1} = (I - A^{-1} \mathbf{uv}^* / (1 + \mathbf{v}^* A^{-1} \mathbf{u})) A^{-1}$$

$A \in \mathbb{C}^{m \times m}$ ,  $\mathbf{u}, \mathbf{v} \in \mathbb{C}^m$  とする。 $A$  は可逆であるとし、 $1 + \mathbf{v}^* A^{-1} \mathbf{u} \neq 0$  とする。このとき次式が成り立つ。

$$(A + \mathbf{uv}^*)^{-1} = \left( I - \frac{1}{1 + \mathbf{v}^* A^{-1} \mathbf{u}} A^{-1} \mathbf{uv}^* \right) A^{-1}$$

*Proof.*

$A + \mathbf{uv}^* = A(I + A^{-1} \mathbf{uv}^*)$  だから  $(A + \mathbf{uv}^*)^{-1} = (I + A^{-1} \mathbf{uv}^*)^{-1} A^{-1}$  であり、2 つ目の因子に V.7.7 を適用すればよい。 □

### V.7.9 Woodbury の公式の特別な場合:

$$(I_n + UV)^{-1} = I_n - U(I_k + VU)^{-1}V$$

$I_n, I_k$  をそれぞれ  $n, k$  次の単位行列とする。 $U \in \mathbb{C}^{m \times k}$ ,  $V \in \mathbb{C}^{k \times m}$  とし、 $I_k + VU$  が可逆であるとする。このとき次式が成り立つ。

$$(I_n + UV)^{-1} = I_n - U(I_k + VU)^{-1}V$$

V.7.7 からの類推でこの等式が予想され、直接代入することで予想が正しいことが確認される。

*Proof.*



直接代入して確認する。

$$\begin{aligned}
(I_n + UV)(I_n - U(I_k + VU)^{-1}V) &= I_n - U(I_k + VU)^{-1}V + UV - UVU(I_k + VU)^{-1}V \\
&= I_n - U(I_k + VU)^{-1}V + UV - U[(I_k + VU) - I_k](I_k + VU)^{-1}V \\
&= I_n - U(I_k + VU)^{-1}V + UV - UV - U(I_k + VU)^{-1}V = I_n
\end{aligned}$$

□

#### V.7.10 Woodbury の公式:

$$(A + UCV)^{-1} = [I_n - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}V] A^{-1}$$

$I_n$  を  $n$  次の単位行列とする。  $A \in \mathbb{C}^{n \times n}$ ,  $C \in \mathbb{C}^{k \times k}$ ,  $U \in \mathbb{C}^{m \times k}$ ,  $V \in \mathbb{C}^{k \times m}$  とし、  $A$ ,  $C$ ,  $I_k + CVA^{-1}U$  が可逆であるとする (結果として  $C^{-1} + VA^{-1}U$  も可逆になる)。このとき次式が成り立つ。

$$(A + UCV)^{-1} = [I_n - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}V] A^{-1}$$

*Proof.*

$A + UCV = A(I_n + A^{-1}UCV)$  であるから  $(A + UCV)^{-1} = (I_n + A^{-1}UCV)^{-1}A^{-1}$  である。V.7.9 で  $U$  を  $A^{-1}U$  で、 $V$  を  $CV$  で置き換えて上式に適用すると次式を得る。

$$(A + UCV)^{-1} = [I_n - A^{-1}U(I_k + CVA^{-1}U)^{-1}CV] A^{-1} = [I_n - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}V] A^{-1}$$

□

## 第 V.8 章

# 特性多項式

### V.8.1 逆行列の特性多項式: $\phi_{A^{-1}}(\lambda) = |A|^{-1}(-\lambda)^n \phi_A(1/\lambda)$

$A$  を正則な  $n$  次正方行列とし、特性多項式を  $\phi_A(\lambda)$  とすると、 $\phi_{A^{-1}}(\lambda) = |A|^{-1}(-\lambda)^n \phi_A(1/\lambda)$ 。

*Proof.*

$|A(\lambda I - A^{-1})|$  を 2 通りの方法で評価する。まず  $|A(\lambda I - A^{-1})| = |A|\phi_{A^{-1}}(\lambda)$  である。次に  $|A(\lambda I - A^{-1})| = |\lambda A - I| = |(-\lambda)((1/\lambda)I - A)| = (-\lambda)^n |(1/\lambda)I - A| = (-\lambda)^n \phi_A(1/\lambda)$ 。よって主張が成り立つ。  $\square$

## 第 V.9 章

# ユニタリ行列

### V.9.1 ユニタリ行列の複素共役はユニタリ行列

*Proof.*  $U$  をユニタリ行列とすると  $(\overline{U})^* \overline{U} = \overline{(U^*)} \overline{U} = \overline{U^* U} = \overline{I} = I$

□

## 第 V.10 章

# 行列の階数

### V.10.1 $\text{rank}(AB) = \text{rank}(BA)$ とは限らない

例えば次の例では  $\text{rank}(AB) = 2 \neq \text{rank}(BA) = 0$ 。

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

### V.10.2 $T$ が正則でも $\text{rank}(ATB) = \text{rank}(AB)$ とは限らない

*Proof.*

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ のとき、} AB = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, ATB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

□

### V.10.3 フルランク条件

行列  $A \in \mathbb{C}^{m \times n} (m \leq n)$  の階数が  $m$  である ( $=A$  がフルランク)  $\Leftrightarrow (\mathbf{x}^\top A = \mathbf{0}_{1 \times n} \Rightarrow \mathbf{x} = \mathbf{0}_{m \times 1})$

*Proof.*

まず、 $\mathbf{x}^\top A = \mathbf{0}_{1 \times n}$  は  $\sum_{i=1}^m x_i \mathbf{a}_i = \mathbf{0}_{n \times 1}$  ( $\mathbf{a}_i$  は  $A$  の第  $i$  行ベクトル) と同値である。 $A$  がフルランクのとき、 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m\}$  は一次独立だから  $\mathbf{x} = \mathbf{0}_{m \times 1}$  でなくてはならない。逆に  $\mathbf{x}^\top A = \mathbf{0}_{1 \times n} \Rightarrow \mathbf{x} = \mathbf{0}_{m \times 1}$  であるとき、 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m\}$  は一次独立だから  $A$  はフルランク。 □

### V.10.4 $A \in \mathbb{C}^{m \times n}, \text{rank}(A) = \text{rank}(A^*A) = \text{rank}(AA^*)$

次元定理を用いる証明

*Proof.*

$\mathbf{x} \in \mathbb{C}^n$  とする。まず主張の左側の等号成立を示す。 $\mathbf{x} \in \ker A$  ならば  $A^*A\mathbf{x} = A^*\mathbf{0} = \mathbf{0}$  だから  $\ker A \subseteq \ker A^*A$ 。逆に  $\mathbf{x} \in \ker A^*A$  ならば  $A^*A\mathbf{x} = \mathbf{0}$  なので  $0 = \mathbf{x}^*A^*A\mathbf{x} = \|A\mathbf{x}\|_2^2$  より  $A\mathbf{x} = \mathbf{0}$ 。

つまり  $\mathbf{x} \in \ker A$  だから  $\ker A \supseteq \ker A^*A$ 。以上より  $\ker A = \ker A^*A$ 。これと次元定理より  $\text{rank}(A) = \dim(\text{Img}(A)) = n - \dim(\ker A) = n - \dim(\ker A^*A) = \dim(\text{Img}(A^*A)) = \text{rank}(A^*A)$ 。

この結果を使って定理の主張の右側の等号成立は直ちに示される:  $\text{rank}(A) = \text{rank}(A^*) = \text{rank}((A^*)^*(A^*)) = \text{rank}(AA^*)$ 。□

特異値分解を用いる証明

*Proof.*

$A$  の特異値分解を  $U\Sigma V^*$  とすると、 $\text{rank}(AA^*) = \text{rank}(U\Sigma V^*V\Sigma U^*) = \text{rank}(U\Sigma^2 U^*) = \text{rank}(\Sigma^2) = \text{rank}(\Sigma) = \text{rank}(A)$ 。同様に  $\text{rank}(A^*A) = \text{rank}(A)$  も示せる。□

## V.10.5 $A \in \mathbb{F}^{m \times n}$ の階数が $n$ ならば $A^*A$ は正則

$A \in \mathbb{F}^{m \times n}$  が列フルランクならば  $A^*A$  は正則

*Proof.* (ややこしい証明)

$A$  の第  $i$  列ベクトルを  $\mathbf{a}_i$  とすると  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$  は一次独立である。これを正規直交化したものを  $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n\}$  とする。  $\mathbf{p}_i \in \text{span}[\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]$  であるから

$$\mathbf{t}_i \in \mathbb{F}^n, \quad A\mathbf{t}_i = \mathbf{p}_i$$

なる  $\mathbf{t}_i$  は唯一に定まり、 $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_n$  は一次独立である。よって

$$\begin{aligned} T &:= [\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_n] \\ P &:= [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n] \end{aligned}$$

とすれば

$$AT = P \quad (A = PT^{-1})$$

であり、 $T$  は正則である。さらに  $P^*P = I_{n \times n}$  であるから

$$\text{rank}(A^*A) = \text{rank}\left((T^{-1})^* P^* P T^{-1}\right) = \text{rank}\left((T^{-1})^* I_{n \times n} T^{-1}\right) = \text{rank}(I_{n \times n}) = n$$

□

*Proof.* (より簡単な証明)

$\mathbf{x} \in \mathbb{F}^n$  を任意にとって  $A^*A\mathbf{x} = \mathbf{0}$  とおいてみると

$$A^*A\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{x}^* A^* A \mathbf{x} = \mathbf{0} \iff \langle A\mathbf{x}, A\mathbf{x} \rangle = 0 \iff \|A\mathbf{x}\| = 0 \iff A\mathbf{x} = \mathbf{0}$$

$A$  は列フルランクであったから  $\mathbf{x} = \mathbf{0}$  となるので  $A^*A$  は正則である。□

## V.10.6 (系) $A \in \mathbb{F}^{m \times n}$ が行フルランクならば $AA^\top$ は正則

(系)  $A \in \mathbb{F}^{m \times n}$  が行フルランクならば  $AA^\top$  は正則

*Proof.*

$B = A^\top$  とおき、 $B$  に対して直前の定理を適用する。□

### V.10.7 $A \in \mathbb{F}^{m \times n}$ の階数が $r$ であるなら、ある正則行列

$T_x \in \mathbb{F}^{n \times n}$ ,  $T_y \in \mathbb{F}^{m \times m}$  が存在して...

$A \in \mathbb{F}^{m \times n}$  の階数が  $r$  であるなら、適当な正則行列  $T_x \in \mathbb{F}^{n \times n}$ ,  $T_y \in \mathbb{F}^{m \times m}$  が存在して

$$AT_x = T_y \begin{bmatrix} I_{r \times r} & O_{r \times (n-r)} \\ O_{(m-r) \times r} & O_{(m-r) \times (n-r)} \end{bmatrix}$$

*Proof.*

$\text{Ker}(A)$  の基底を  $\{x_{r+1}, x_{r+2}, \dots, x_n\}$  とする (次元定理から  $\dim(\text{Ker}(A)) = n - \dim(\text{Im}(A)) = n - \text{rank}(A) = n - r$ )。これを含む  $\mathbb{F}^n$  の基底を  $\{x_1, x_2, \dots, x_r, x_{r+1}, \dots, x_n\}$  とすると

$$\text{span}[Ax_1, Ax_2, \dots, Ax_r] = \text{Im}(A) \subseteq \mathbb{F}^m \quad (\text{証明は次元定理を参照})$$

である。 $y_i := Ax_i$  ( $i = 1, 2, \dots, r$ ) とし、これらを含む  $\mathbb{F}^m$  の基底を  $\{y_1, y_2, \dots, y_r, y_{r+1}, \dots, y_m\}$  とすると  $T_x := [x_1, x_2, \dots, x_n]$ ,  $T_y := [y_1, y_2, \dots, y_m]$  は正則であり、

$$\begin{aligned} AT_x &= [y_1, y_2, \dots, y_r, \underbrace{0, \dots, 0}_{n-r \text{ 個}}] \\ &= [y_1, y_2, \dots, y_m][e_1, e_2, \dots, e_r, \underbrace{0, \dots, 0}_{n-r \text{ 個}}] \quad (e_i \text{ は } \mathbb{F}^m \text{ の標準基底ベクトル}) \\ &= T_y \begin{bmatrix} I_{r \times r} & O_{r \times (n-r)} \\ O_{(m-r) \times r} & O_{(m-r) \times (n-r)} \end{bmatrix} \end{aligned}$$

□

### V.10.8 同じ型の行列の階数が等しいための必要十分条件

$A, B \in \mathbb{F}^{m \times n}$  について  $\text{rank}(A) = \text{rank}(B)$  であるための必要十分条件は、適当な正則行列が存在して  $AT_1 = T_2B$  となることである。

*Proof.*

(必要)

定理 V.10.7 より、 $A, B$  それぞれに対して適当な正則行列  $T_{x,A}, T_{x,B} \in \mathbb{F}^{n \times n}$ ,  $T_{y,B}, T_{y,A} \in \mathbb{F}^{m \times m}$  が存在して

$$\begin{aligned} T_{y,A}^{-1}AT_{x,A} &= T_{y,B}^{-1}BT_{x,B} \\ \therefore AT_{x,A}T_{x,B}^{-1} &= T_{y,A}T_{y,B}^{-1}B \end{aligned}$$

上式で  $T_1 := T_{x,A}T_{x,B}^{-1}$ ,  $T_2 := T_{y,A}T_{y,B}^{-1}$  とすれば良い。

(十分)

ある行列に正則行列を掛けても階数は変化しない。

$$\text{rank}(A) = \text{rank}(AT_1) = \text{rank}(T_2B) = \text{rank}(B)$$

□

## V.10.9 優対角正方行列は正則

*Proof.*

$A = [a_{ij}] \in \mathbb{C}^{n \times n}$  を優対角行列とし、 $A\mathbf{x} = \mathbf{0}_n$ ,  $\mathbf{x} = [x_1, \dots, x_n]^\top \in \mathbb{C}^n$  とおく。 $A\mathbf{x}$  の第 1 成分の絶対値について

$$\begin{aligned} |a_{11}x_1 + \dots + a_{1n}x_n| &\geq ||a_{11}x_1| - |a_{12}x_2 + \dots + a_{1n}x_n|| \geq |a_{11}|x_1| - |a_{12}x_2 + \dots + a_{1n}x_n| \\ &\geq |a_{11}||x_1| - \dots - |a_{1n}||x_n| \end{aligned}$$

が成り立つが、ここで仮に  $|x_1| > |x_2|, \dots, |x_n|$  とすると、 $A$  の優対角性から

$$\begin{aligned} |a_{11}||x_1| - \dots - |a_{1n}||x_n| &> |a_{11}||x_1| - |a_{12}||x_1| - \dots - |a_{1n}||x_1| \\ &= (|a_{11}| - |a_{12}| - \dots - |a_{1n}|)|x_1| > 0 \end{aligned} \tag{1}$$

となり、 $A\mathbf{x} = \mathbf{0}_n$  に矛盾する。故に  $|x_1| \leq \max\{|x_1|, |x_2|, \dots, |x_n|\}$  である。同様にして  $A\mathbf{x}$  の第 2, 3, ...,  $n$  成分の絶対値を考えることで  $|x_i| \leq \max\{|x_j|\} \ (i, j \in \{1, \dots, n\})$  が示される。これは  $|x_1| = \dots = |x_n| = c$ , ( $c$  は適当な非負数) に他ならない。仮に  $c > 0$  であるとする式 (1) より  $|A\mathbf{x}$  の第 1 成分  $> 0$  となり、 $A\mathbf{x} = \mathbf{0}$  に矛盾する。故に  $c = 0$ 、つまり  $\mathbf{x} = \mathbf{0}_n$  である。よって  $A\mathbf{x} = \mathbf{0}_n \Rightarrow \mathbf{x} = \mathbf{0}_n$  が成り立つから  $A$  は正則である。  $\square$

## V.10.10 最大ランク分解

$m \times n$  行列  $A$  の階数を  $r$  とする。 $A$  の一次独立な列ベクトルを左から順に集めてできる列フルランクな  $m \times r$  行列を  $B$  とし、 $A$  の簡約化の零行ベクトルを除去した行フルランクな  $r \times n$  行列を  $F$  とすると  $A = BF$ 。

*Proof.*

簡約化に用いる行変形行列を  $R$  とすると

$$RA = \begin{bmatrix} F \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} I_r \\ \mathbf{0} \end{bmatrix} F = RBF$$

$R$  は正則だから  $A = BF$ 。

$\square$

## 第 V.11 章

# 固有値, 固有ベクトル

### V.11.1 諸定理

#### V.11.1.1 相異なる固有値に対応する固有ベクトルは一次独立

$A \in \mathbb{C}^{n \times n}$  の相異なる  $p$  個の固有値  $\lambda_i$  ( $i = 1, \dots, p$ ) に対応する固有ベクトルを  $\mathbf{v}_i$  とすると  $\mathbf{v}_1, \dots, \mathbf{v}_p$  は一次独立である。

*Proof.*

$c_1, \dots, c_p$  を定数として

$$c_1 \mathbf{v}_1 + \dots + c_p \mathbf{v}_p = \mathbf{0}_n$$

とおいて両辺に左から  $A$  を掛けると

$$\lambda_1 c_1 \mathbf{v}_1 + \dots + \lambda_p c_p \mathbf{v}_p = \mathbf{0}_n$$

となる。これを繰り返して

$$\begin{aligned} \lambda_1^2 c_1 \mathbf{v}_1 + \dots + \lambda_p^2 c_p \mathbf{v}_p &= \mathbf{0}_n \\ &\vdots \\ \lambda_1^{p-1} c_1 \mathbf{v}_1 + \dots + \lambda_p^{p-1} c_p \mathbf{v}_p &= \mathbf{0}_n \end{aligned}$$

を得る。これより

$$[c_1 \mathbf{v}_1, \dots, c_p \mathbf{v}_p] \begin{bmatrix} 1 & \lambda_1 & \dots & \lambda_1^{p-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_p & \dots & \lambda_p^{p-1} \end{bmatrix} = O_{n \times p}$$

となるが、 $\lambda_i$  ( $i = 1, \dots, p$ ) が相異なることより、上式の等号の左隣のヴァンデルモンド行列は正則だから

$$[c_1 \mathbf{v}_1, \dots, c_p \mathbf{v}_p] = O_{n \times p}$$

となる。 $\mathbf{v}_1, \dots, \mathbf{v}_p \neq \mathbf{0}_n$  より  $c_1 = \dots = c_p = 0$  である。 □



### V.11.1.2 $n \times n$ 行列 $A$ の全要素が $a \neq 0$ であるなら非零固有値は $an$ のみ

*Proof.*

まず、 $an$  が固有値であり、全ての要素が等しい  $n$  次元ベクトルが固有値  $an$  に対応する固有ベクトルであることは、実際に計算してみることで容易に確かめられる。非零固有値がこれ以外に存在しないことを示す。 $\lambda \neq 0, an$  とし、 $\lambda I - A$  の第  $j$  列ベクトルを  $\mathbf{b}_j$  とすると、

$$b_j(i) = \begin{cases} -a & (i \neq k) \\ \lambda - a & (i = k) \end{cases}$$

$\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$  が一次独立であることを示せば、 $|\lambda I - A| \neq 0$  が示せて、 $\lambda$  が固有値でないことが示せる。 $c_1, \dots, c_n$  を定数とし、 $\sum_{j=1}^n c_j \mathbf{b}_j = \mathbf{0}$  とおくと、第  $i$  成分に着目して

$$0 = c_i(\lambda - a) - a \sum_{j \neq i}^n c_j = c_i \lambda - a \sum_{j=1}^n c_j$$

これと  $\lambda \neq 0$  より  $c_1 = \dots = c_n$  である。この値を  $c$  とおくと  $c\lambda = anc$  となり、 $\lambda \neq an$  より  $c = 0$  である。□

### V.11.1.3 転置行列の固有値, 固有ベクトル

$A \in \mathbb{F}^{n \times n}$  について  $A^\top$  の固有値は  $A$  と等しく、各固有値  $\lambda_i$  に対する固有空間の次元も  $A$  の固有値  $\lambda_i$  に対するそれと等しい。

*Proof.*

$A^\top$  の特性多項式は

$$|sI - A^\top| = |(sI - A)^\top| = |sI - A|$$

の如く  $A$  の特性多項式と等しいから固有値も全て等しい。各固有値  $\lambda_i$  に対する固有空間  $E_{A^\top}(\lambda_i)$  の次元は

$$\begin{aligned} \dim E_{A^\top}(\lambda_i) &= \dim \text{Ker}(\lambda_i I - A^\top) = \dim \mathbb{F}^n - \text{rank}(\lambda_i I - A^\top) = n - \text{rank}((\lambda_i I - A)^\top) \\ &= n - \text{rank}(\lambda_i I - A) = \dim E_A(\lambda_i) \end{aligned}$$

の如く  $A$  の固有値  $\lambda_i$  に対する固有空間の次元と等しい。□

### V.11.1.4 逆行列の固有値, 固有ベクトル

正則な行列  $A \in \mathbb{F}^{n \times n}$  の逆行列の固有値は  $A$  の固有値の逆数であり、 $A$  と共通の固有ベクトルをもつ。

*Proof.*

$A^{-1}$  もまた正則だから 0 固有値は持たない。 $A^{-1}$  の固有値の一つを  $\lambda' \neq 0$  とし、それに対する固有ベクトルを  $\mathbf{y} \neq \mathbf{0}$  とおくと  $A^{-1}\mathbf{y} = \lambda'\mathbf{y}$  であり、両辺に左から  $A$  を掛けて  $\mathbf{y} = \lambda'A\mathbf{y} \quad \therefore A\mathbf{y} = \frac{1}{\lambda'}\mathbf{y}$  となる。よって  $A^{-1}$  の固有値  $\lambda'$  は  $A$  の固有値  $1/\lambda'$  の逆数であり、固有ベクトルは共通である。□

### V.11.1.5 ユニタリ行列の全ての固有値の絶対値は 1 である

*Proof.*

ユニタリ行列  $P$  とその任意の固有値  $\lambda$  およびそれに対応する固有ベクトル  $\mathbf{x} \neq \mathbf{0}$  について  $P\mathbf{x} = \lambda\mathbf{x}$  の両辺のノルムをとると  $\|\mathbf{x}\|_2 = \|P\mathbf{x}\|_2 = |\lambda|\|\mathbf{x}\|_2$  であるからわかる。  $\square$

### V.11.1.6 コンパニオン行列

$\lambda$  に関する多項式  $f_n(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0$  に対して次の行列

$$A = \begin{bmatrix} -a_{n-1} & -a_{n-2} & -a_{n-3} & \cdots & -a_0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

を  $f(\lambda)$  のコンパニオン行列という。  $\det A = (-1)^n a_0$  であり、  $\det(\lambda I - A) = f_n(\lambda)$  である。

*Proof.*

$\det A = (-1)^n a_0$  については、第 1 行をバブルソートの要領で  $n-1$  回の置換により最下行に移動させて下三角行列を作ることによって証明できる。

次に  $\det(\lambda I - A) = f(\lambda)$  を示す。  $n=1$  のとき成り立つ。  $n-1$  まで成り立つと仮定する。第 1 列に関して余因子展開すると

$$\begin{aligned} \det(\lambda I - A) &= \det \begin{bmatrix} \lambda + a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_0 \\ -1 & \lambda & 0 & \cdots & 0 \\ 0 & -1 & \lambda & \cdots & 0 \\ \vdots & & \ddots & \ddots & \\ 0 & \cdots & 0 & -1 & \lambda \end{bmatrix} \\ &= (\lambda + a_{n-1})\lambda^{n-1} + \det \begin{bmatrix} a_{n-2} & a_{n-3} & a_{n-4} & \cdots & a_0 \\ -1 & \lambda & 0 & \cdots & 0 \\ 0 & -1 & \lambda & \cdots & 0 \\ \vdots & & \ddots & \ddots & \\ 0 & \cdots & 0 & -1 & \lambda \end{bmatrix} \\ &= \lambda^n + a_{n-1}\lambda^{n-1} + f_{n-1}(\lambda)|_{a_{n-2} \rightarrow a_{n-2} - \lambda} \\ &= \lambda^n + a_{n-1}\lambda^{n-1} + \lambda^{n-1} + (a_{n-2} - \lambda)\lambda^{n-2} + a_{n-3}\lambda^{n-3} + \cdots + a_1\lambda + a_0 \\ &= \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \cdots + a_1\lambda + a_0 = f_n(\lambda) \end{aligned}$$

$\square$

### V.11.1.7 巡回行列の固有値

$N \in \mathbb{N}$ ,  $a_0, a_1, \dots, a_{N-1} \in \mathbb{C}$ ,  $\mathbf{a} := [a_0, \dots, a_{N-1}]^\top$  とする。 $\mathbf{a}$  の要素を  $m \in \{0, 1, \dots, N-1\}$  だけ巡回シフトしたベクトル  $\mathbf{a}_m$  を次式で定義する。

$$\mathbf{a}_m := \begin{cases} \mathbf{a} & (m = 0) \\ [a_{N-m}, \dots, a_{N-1}, a_0, a_1, \dots, a_{N-1-m}]^\top & (m = 1, 2, \dots, N-1) \end{cases}$$

$A \in \mathbb{C}^{N \times N}$  を次式で定義する。

$$A := [\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_{N-1}]$$

このとき  $A$  は次のように対角化可能である。

$$A = \sqrt{N} W \text{diag}(\text{DFT}(\mathbf{a})) W^*$$

ここに  $W$  は  $N \times N$  DFT 基底行列であり、第  $(n, k)$  要素は次式で定義される。

$$W_{n,k} := \frac{1}{\sqrt{N}} \exp i \frac{nk}{N} 2\pi$$

$\text{DFT}(\mathbf{a})$  は  $\mathbf{a}$  の DFT (離散 Fourier 変換) であり、次式で定義される。

$$\text{DFT}(\mathbf{a}) := W^* \mathbf{a}$$

すなわち第  $k$  要素は次式である。

$$\sum_{n=0}^{N-1} \overline{W_{k,n}} a_n$$

*Proof.*

[8] で述べられている DFT の性質を用いる。 $\mathbf{e}_m \in \mathbb{R}^N$  を第  $m$  要素が 1 でそれ以外は 0 のベクトルとする。 $\mathbf{a}_m$  は巡回畳み込みを用いて次式で表せる。

$$\mathbf{a}_m = \mathbf{a} \underset{\text{cyc}}{*} \mathbf{e}_m$$

巡回畳み込みの DFT の性質より次式が成り立つ。

$$\text{DFT}(\mathbf{a}_m) = \sqrt{N} \text{DFT}(\mathbf{a}) \odot \text{DFT}(\mathbf{e}_m) = \sqrt{N} \text{DFT}(\mathbf{a}) \odot W^*[:, m]$$

これより次式が成り立つ。

$$\begin{aligned} W^* A &= [\text{DFT}(\mathbf{a}_0), \text{DFT}(\mathbf{a}_1), \dots, \text{DFT}(\mathbf{a}_{N-1})] \\ &= \sqrt{N} [\text{DFT}(\mathbf{a}) \odot W^*[:, 0], \text{DFT}(\mathbf{a}) \odot W^*[:, 1], \dots, \text{DFT}(\mathbf{a}) \odot W^*[:, N-1]] \\ &= \sqrt{N} \text{diag}(\text{DFT}(\mathbf{a})) W^* \end{aligned}$$

両辺に左から  $W$  を掛け、 $W$  のユニタリ性  $WW^* = I$  を用いて次式を得る。

$$A = \sqrt{N} W \text{diag}(\text{DFT}(\mathbf{a})) W^*$$

□

### V.11.1.8 Hadamard 行列の固有値

$2^n$  ( $n \in \mathbb{N}$ ) 次の Hadamard 行列の固有値は  $2^{n/2}$  と  $-2^{n/2}$  が  $2^{n-1}$  個ずつである [7]。

*Proof.*

$2^n$  次の Hadamard 行列を  $H_n$  と書く。 $H_0 := [1]$  と定義する。まず、 $H_1$  の固有値と固有ベクトルの組が  $(\pm\sqrt{2}, [1 \pm \sqrt{2}, 1]^\top)$  (複合同順) であることは容易に導ける。よって  $n = 1$  のとき定理の主張が成り立つ。

$n = m$  のとき成り立つと仮定して  $n = m + 1$  のとき成り立つことを示す。 $H_m$  の固有値  $2^{m/2}$  に対応する相異なる固有ベクトルを  $\mathbf{u}_m^1, \dots, \mathbf{u}_m^{2^{m-1}}$  とし、固有値  $-2^{m/2}$  に対応する相異なる固有ベクトルを  $\mathbf{v}_m^1, \dots, \mathbf{v}_m^{2^{m-1}}$  とする。このとき次のことが成り立つ。

$i = 1, \dots, 2^{m-1}$  に対して

$$\begin{bmatrix} \mathbf{u}_m^i \\ (-1 + \sqrt{2})\mathbf{u}_m^i \end{bmatrix}, \begin{bmatrix} (1 - \sqrt{2})\mathbf{v}_m^i \\ \mathbf{v}_m^i \end{bmatrix} \text{ は } H_{m+1} \text{ の固有値 } 2^{(m+1)/2} \text{ に対応する固有ベクトルである} \quad (1)$$

$$\begin{bmatrix} (1 - \sqrt{2})\mathbf{u}_m^i \\ \mathbf{u}_m^i \end{bmatrix}, \begin{bmatrix} \mathbf{v}_m^i \\ (-1 + \sqrt{2})\mathbf{v}_m^i \end{bmatrix} \text{ は } H_{m+1} \text{ の固有値 } -2^{(m+1)/2} \text{ に対応する固有ベクトルである} \quad (2)$$

(1) の主張の半分は次のようにして確かめられる。

$$H_{m+1} \begin{bmatrix} \mathbf{u}_m^i \\ (-1 + \sqrt{2})\mathbf{u}_m^i \end{bmatrix} = \begin{bmatrix} H_m & H_m \\ H_m & -H_m \end{bmatrix} \begin{bmatrix} \mathbf{u}_m^i \\ (-1 + \sqrt{2})\mathbf{u}_m^i \end{bmatrix} = \begin{bmatrix} \sqrt{2}H_m\mathbf{u}_m^i \\ (2 - \sqrt{2})H_m\mathbf{u}_m^i \end{bmatrix} = 2^{\frac{m+1}{2}} \begin{bmatrix} \mathbf{u}_m^i \\ (-1 + \sqrt{2})\mathbf{u}_m^i \end{bmatrix}$$

他も同様に確かめられる。これら  $2^{m+1}$  個の固有ベクトルが一次独立であることを示せば、 $n$  が 1 増える毎に  $H_n$  の固有ベクトルの個数が倍になるから、定理の証明が済む。相異なる固有値に対応する固有ベクトルは一次独立 (V.11.1.1) だから、(1) で言及した固有ベクトルの一次結合と (2) で言及した固有ベクトルの一次結合は一次独立である。あとは、(1) の固有ベクトル達、および (2) の固有ベクトル達がそれぞれ一次独立であることを示せばよい。(1) の固有ベクトル達が一次独立であることを示す。(2) についても同様に示せる。(1) の列ベクトル達の上半分 (或いは下半分) は  $H_m$  の固有ベクトルであるから、相異なる固有値に対応する固有ベクトルが一次独立であることを再び用いて  $\begin{bmatrix} \mathbf{u}_m^i \\ (-1 + \sqrt{2})\mathbf{u}_m^i \end{bmatrix}$  ( $i = 1, \dots, 2^{m-1}$ ) の一次結合と

$\begin{bmatrix} (1 - \sqrt{2})\mathbf{v}_m^i \\ \mathbf{v}_m^i \end{bmatrix}$  ( $i = 1, \dots, 2^{m-1}$ ) の一次結合は一次独立である。また、 $\mathbf{u}_m^1, \dots, \mathbf{u}_m^{2^{m-1}}$  は  $H_m$  の固有ベクトルだから  $\begin{bmatrix} \mathbf{u}_m^i \\ (-1 + \sqrt{2})\mathbf{u}_m^i \end{bmatrix}$  ( $i = 1, \dots, 2^{m-1}$ ) は一次独立である。

以上より (1),(2) で言及した固有ベクトル達は一次独立である。 □

### V.11.1.9 最小消去多項式が最小多項式と一致するベクトルの存在

任意の正方行列  $A \in \mathbb{F}^{n \times n}$  につきその最小消去多項式  $\psi_{A,x}(\lambda)$  が  $A$  の最小多項式  $\psi(\lambda)$  と一致するベクトル  $\mathbf{x} \in \mathbb{F}^n$  が存在する。

*Proof.*

$$\psi(\lambda) = \prod_{i=1}^p (\lambda - \lambda_i)^{\mu_i}$$

とする。但し  $\lambda_i$  は  $A$  の相異なる  $p$  個の固有値である。このとき  $\mathbb{F}^n$  は次のように直和分解できる ([3] P77 「 $e^{Mt}$  と一般固有値問題」)。

$$\mathbb{F}^n = \bigoplus_{i=1}^p \text{Ker}((A - \lambda_i I)^{\mu_i})$$

$\psi(\lambda)$  より低次のいかなる多項式によっても消去できない  $\mathbf{x} \in \mathbb{F}^n$  を構成しよう。

$$\text{Ker}((A - \lambda_i I)^{\mu_i}) / \text{Ker}((A - \lambda_i I)^{\mu_i - 1}) \supset \{\mathbf{0}\}$$

である (さもなくば  $\psi(\lambda)$  が最小多項式であったことに矛盾する!) から、ここからひとつベクトル  $\mathbf{x}_i$  をとると自然数  $k_i$  につき

$$(A - \lambda_i I)^{k_i} \mathbf{x}_i \begin{cases} = \mathbf{0} & (\lambda = \lambda_i, k_i \geq \mu_i) \\ \neq \mathbf{0} & (\text{otherwise}) \end{cases}$$

もっと簡単に言うと、 $\mathbf{x}_i$  を消去するには  $(A - \lambda_i I)$  の  $\mu_i$  次以上の積が必須である。そこで

$$\mathbf{x} = \sum_{i=1}^p \mathbf{x}_i$$

とすれば  $\mathbf{x}$  は  $\psi(\lambda)$  で消去でき、かつ  $\mathbf{x}$  の任意の消去多項式は必ず  $\psi(\lambda)$  で割り切れなくてはならないから  $\mathbf{x}$  の最小消去多項式は  $\psi(\lambda)$  である。  $\square$

## V.11.2 スペクトル写像定理

$A \in \mathbb{F}^{n \times n}$  の固有値を重複も含めて  $\lambda_1, \dots, \lambda_n$  とし、 $f(s)$  を  $s$  の多項式とする。このとき行列多項式  $f(A)$  の固有値は重複を含めて  $f(\lambda_1), \dots, f(\lambda_n)$  となる。

*Proof.*

$A$  は正方行列であるから適当なユニタリ行列  $P \in \mathbb{C}^{n \times n}$  により次のように上三角化できる。

$$A = P U P^\top$$

但し  $U$  は第  $i$  対角成分が  $\lambda_i$  である上三角行列である ( $P$  はユニタリだから  $A^n = P U^n P^\top$  となることに注意せよ)。  $f(s)$  の次数を  $m$  とし、 $k$  次の係数を  $c_k$  とする ( $f(s) = \sum_{k=0}^m c_k s^k$ )。  $f(A)$  の固有多項式は

$$\begin{aligned} |\lambda I - f(A)| &= \left| \lambda I - \sum_{k=0}^m c_k A^k \right| = \left| \lambda I - \sum_{k=0}^m c_k P U^k P^\top \right| = \left| P \left( \lambda I - \sum_{k=0}^m c_k U^k \right) P^\top \right| \\ &= \left| \lambda I - \sum_{k=0}^m c_k U^k \right| \end{aligned}$$

定理 V.2.2 より  $U^k$  は上三角行列であり、その第  $i$  対角成分は  $\lambda_i^k$  であるから、上の  $|\cdot|$  の中身は上三角行列であり、その第  $i$  対角成分は  $\lambda - \sum_{k=0}^m c_k \lambda_i^k = \lambda - f(\lambda_i)$  である。よって上の式は

$$\left| \lambda I - \sum_{k=0}^m c_k U^k \right| = \prod_{i=1}^n (\lambda - f(\lambda_i))$$

であるから  $f(A)$  の固有値が  $f(\lambda_1), \dots, f(\lambda_n)$  であることがわかる。 □

### V.11.3 対角化

#### V.11.3.1 正規行列 $\iff$ ユニタリ行列で対角化可能

$n$  を自然数とする。  $A \in \mathbb{C}^{n \times n}$  が正規行列であるための必要十分条件は、ある対角行列  $\Lambda \in \mathbb{C}^{n \times n}$  とユニタリ行列  $Q \in \mathbb{C}^{n \times n}$  が存在して  $A = Q\Lambda Q^*$  となることである。

*Proof.*

十分性は容易に示せるので省略し、必要性のみ示す。  $A$  を正規行列とする。  $A$  の Schur 分解を  $A = SUS^*$  とする。ここに  $S, U \in \mathbb{C}^{n \times n}$  はそれぞれ適当なユニタリ行列, 上三角行列である。  $A$  が正規だから

$$\begin{aligned} O &= A^*A - AA^* = SU^*US^* - SUU^*S^* = S(U^*U - UU^*)S^* \\ \therefore U^*U - UU^* &= S^*OS = O \\ \therefore U^*U &= UU^* \end{aligned}$$

これより、まず  $U^*U$  と  $UU^*$  の第 1,1 成分が等しいから

$$\begin{aligned} (U^*U)_{1,1} &= (UU^*)_{1,1} \\ \therefore |u_{1,1}|^2 &= |u_{1,1}|^2 + |u_{1,2}|^2 + \dots + |u_{1,n}|^2 \\ u_{1,2} &= u_{1,3} = \dots = u_{1,n} = 0 \end{aligned}$$

次に、  $U^*U$  と  $UU^*$  の第 2,2 成分が等しいから

$$|u_{1,2}|^2 + |u_{2,2}|^2 = |u_{2,2}|^2 + |u_{2,3}|^2 + \dots + |u_{2,n}|^2$$

これと先程の結果  $|u_{1,2}| = 0$  より

$$u_{2,3} = u_{2,4} = \dots = u_{2,n} = 0$$

同じ要領で続けていくと  $U$  の非対角要素が全て 0 であることがわかる。 □

### V.11.4 一般固有空間

#### V.11.4.1 一般固有空間の階数の頭打ち

行列  $A$  とその固有値  $\lambda$  の  $k$  階一般固有空間  $F_{A,k}(\lambda)$  に対して、もしある自然数  $l$  が存在して  $F_{A,l}(\lambda) = F_{A,l+1}(\lambda)$  が成り立てば、任意の  $m \geq l+2$  に対して  $F_{A,m}(\lambda) = F_{A,l}(\lambda)$ 。

*Proof.*

$m = l+1$  に対して示す。それ以上の  $m$  に対しても帰納的に示せる。結論を否定すると、ある  $\mathbf{v} \in F_{A,m}(\lambda)$  が存在して  $(A - \lambda I)^{l+2}\mathbf{v} = \mathbf{0} \wedge (A - \lambda I)^{l+1}\mathbf{v} \neq \mathbf{0}$ 。すなわち  $(A - \lambda I)^{l+1}[(A - \lambda I)\mathbf{v}] = \mathbf{0} \wedge (A - \lambda I)^l[(A - \lambda I)\mathbf{v}] \neq \mathbf{0}$ 。よって  $(A - \lambda I)\mathbf{v} \in F_{A,l+1}(\lambda)$ 。これと仮定  $F_{A,l+1}(\lambda) = F_{A,l}(\lambda)$  より  $(A - \lambda I)\mathbf{v} \in F_{A,l}(\lambda)$ 。よって  $(A - \lambda I)^{l+1}\mathbf{v} = \mathbf{0}$  となり、矛盾する。 □

#### V.11.4.2 一般固有空間への直和分解

$A \in \mathbb{F}^{n \times n}$  の固有多項式が相異なる固有値  $\lambda_1, \dots, \lambda_p$  に対して  $\varphi_A(\lambda) = \prod_{i=1}^p (\lambda - \lambda_i)^{\nu_i}$  となるとき、次が成り立つ。

$$\mathbb{F}^n = \bigoplus_{i=1}^p F_A(\lambda_i), \quad \dim F_A(\lambda_i) = \nu_i$$

*Proof.*

Cayley-Hamilton の定理と [2] 補題 2.2 より最初の式  $\mathbb{F}^n = \bigoplus_{i=1}^p F_A(\lambda_i)$  が従う。次に  $\nu_i' := \dim F_A(\lambda_i)$  が  $\nu_i$  に等しいことを示す。各  $F_A(\lambda_i)$  が  $A$  の不変部分空間であること ([2] 定理 2.4) と定理 V.4.4 より、 $F_A(\lambda_i)$  の基底を  $\{\phi_{i1}, \dots, \phi_{i\nu_i'}\}$  とし、 $\Phi_i := [\phi_{i1}, \dots, \phi_{i\nu_i'}]$ ,  $\Phi := [\Phi_1, \dots, \Phi_p] \in \mathbb{F}^{n \times n}$  とすれば、次式を満たす正方行列  $A_i \in \mathbb{F}^{\nu_i' \times \nu_i'}$  ( $i = 1, \dots, p$ ) が存在する。

$$A = \Phi \operatorname{diag}(A_1, \dots, A_p) \Phi^{-1}$$

ここで下準備として  $A_i$  が唯一の固有値  $\lambda_i$  を持つことを示しておく。まず固有値  $\lambda_i$  を持つことを示す。 $\mathbf{x} \in E_A(\lambda_i) \subseteq F_A(\lambda_i)$  を任意に取り、 $F_A(\lambda_i)$  の基底で次のように展開する。

$$\mathbf{x} = \sum_{j=1}^{\nu_i'} c_j \phi_{ij}$$

これを上式に右から掛けて

$$\begin{aligned} A\mathbf{x} &= \Phi \operatorname{diag}(A_1, \dots, A_p) \Phi^{-1} \sum_{j=1}^{\nu_i'} c_j \phi_{ij} = \Phi \operatorname{diag}(A_1, \dots, A_p) \sum_{j=1}^{\nu_i'} c_j \mathbf{e}_{j+\sum_{k=1}^{i-1} \nu_k'} \quad (\because \text{定理 V.7.1}) \\ &= \Phi \operatorname{diag}(A_1, \dots, A_p) \left[ \underbrace{0, \dots, 0}_{\nu_1' + \dots + \nu_{i-1}'}, c_1, \dots, c_{\nu_i'}, \underbrace{0, \dots, 0}_{\nu_{i+1}' + \dots + \nu_p'} \right]^\top \\ &= \Phi \begin{bmatrix} \mathbf{0}_{\nu_1'} \\ \vdots \\ \mathbf{0}_{\nu_{i-1}'} \\ A_i [c_1, \dots, c_{\nu_i'}]^\top \\ \mathbf{0}_{\nu_{i+1}'} \\ \vdots \\ \mathbf{0}_{\nu_p'} \end{bmatrix} \quad (\ast \mathbf{0}_\nu \text{ は } \nu \text{ 次元の } 0 \text{ ベクトル}) \end{aligned} \quad (1)$$

ここで  $[d_1, \dots, d_{\nu_i'}]^\top = A_i [c_1, \dots, c_{\nu_i'}]^\top$  とおくと

$$A\mathbf{x} = (1) = \sum_{j=1}^{\nu_i'} d_j \phi_{ij}$$

であるが、そもそも  $A\mathbf{x} = \lambda_i \mathbf{x} = \sum_{j=1}^{\nu_i'} \lambda_i c_j \phi_{ij}$  であつたから  $d_j = \lambda_j c_j$  となり、 $A_i [c_1, \dots, c_{\nu_i'}]^\top = \lambda_i [c_1, \dots, c_{\nu_i'}]^\top$  となるから、 $A_i$  は固有値  $\lambda_i$  を持つことになる。

次に  $A_i$  の固有値が  $\lambda_i$  のみであることを背理法で示す。 $A_i$  が固有値  $\lambda' \notin \{\lambda_1, \dots, \lambda_p\}$  を持つと仮定し、 $\lambda'$  に対する固有ベクトルの一つを  $\mathbf{y} = [y_1, \dots, y_{\nu_i'}]^\top \in \mathbb{F}^{\nu_i'}$  とする ( $A_i \mathbf{y} = \lambda' \mathbf{y}$ )。ベクトル  $\mathbf{x} \in \mathbb{F}^n$  を

$$\Phi^{-1} \mathbf{x} = \left[ \underbrace{0, \dots, 0}_{\nu_1' + \dots + \nu_{i-1}'}, \mathbf{y}^\top, \underbrace{0, \dots, 0}_{\nu_{i+1}' + \dots + \nu_p'} \right]^\top$$

すなわち

$$\mathbf{x} = \Phi \left[ \underbrace{0, \dots, 0}_{\nu_1' + \dots + \nu_{i-1}'}, \mathbf{y}^\top, \underbrace{0, \dots, 0}_{\nu_{i+1}' + \dots + \nu_p'} \right]^\top = \sum_{j=1}^{\nu_i'} y_j \phi_{ij}$$

で定めると  $\phi_{i1}, \dots, \phi_{i\nu_i'}$  の和で表されているから  $\mathbf{x} \in F_A(\lambda_i)$  であるので、まず次式が成り立つ。

$$(A - \lambda_i I_n)^{\nu_i} \mathbf{x} = \mathbf{0}_n \quad (2)$$

一方で

$$\begin{aligned} (A - \lambda_i I_n) \mathbf{x} &= A \mathbf{x} - \lambda_i \mathbf{x} = \Phi \operatorname{diag}(A_1, \dots, A_p) \Phi^{-1} \mathbf{x} - \lambda_i \mathbf{x} \\ &= \Phi \operatorname{diag}(A_1, \dots, A_p) \left[ \underbrace{0, \dots, 0}_{\nu_1' + \dots + \nu_{i-1}'}, \mathbf{y}^\top, \underbrace{0, \dots, 0}_{\nu_{i+1}' + \dots + \nu_p'} \right]^\top - \lambda_i \mathbf{x} \\ &= \Phi \begin{bmatrix} \mathbf{0}_{\nu_1'} \\ \vdots \\ \mathbf{0}_{\nu_{i-1}'} \\ A_i \mathbf{y} \\ \mathbf{0}_{\nu_{i+1}'} \\ \vdots \\ \mathbf{0}_{\nu_p'} \end{bmatrix} - \lambda_i \mathbf{x} = \Phi \begin{bmatrix} \mathbf{0}_{\nu_1'} \\ \vdots \\ \mathbf{0}_{\nu_{i-1}'} \\ \lambda' \mathbf{y} \\ \mathbf{0}_{\nu_{i+1}'} \\ \vdots \\ \mathbf{0}_{\nu_p'} \end{bmatrix} - \lambda_i \mathbf{x} \\ &= \lambda' \Phi \left[ \underbrace{0, \dots, 0}_{\nu_1' + \dots + \nu_{i-1}'}, \mathbf{y}^\top, \underbrace{0, \dots, 0}_{\nu_{i+1}' + \dots + \nu_p'} \right]^\top - \lambda_i \mathbf{x} = \lambda' \mathbf{x} - \lambda_i \mathbf{x} \\ &= (\lambda' - \lambda_i) \mathbf{x} \end{aligned}$$

であるから

$$(A - \lambda_i I_n)^{\nu_i} \mathbf{x} = (\lambda' - \lambda_i)^{\nu_i} \mathbf{x} \neq \mathbf{0}_n$$

となるが、これは式 (2) に矛盾する。故に背理法の仮定が誤っており、 $A_i$  は固有値  $\lambda'$  を持たない。

以上より  $A_i$  は唯一の固有値  $\lambda_i$  を持つことがわかる。よって

$$\varphi_A(\lambda) = |\lambda I_n - A| = \prod_{i=1}^p |\lambda I_{\nu_i'} - A_i| = \prod_{i=1}^p (\lambda - \lambda_i)^{\nu_i'}$$

となり、定理の仮定の  $\varphi_A(\lambda)$  と比較して  $\nu_i' = \nu_i$  が得られる。□

この定理から直ちに次が成り立つ。

#### V.11.4.3 代数的重複度 $\geq$ 幾何学的重複度

任意の  $A \in \mathbb{F}^{n \times n}$  の固有値  $\lambda$  の代数的重複度  $\nu_i$  は幾何学的重複度  $\mu_i$  以上である。

*Proof.* 直前の定理より  $\nu_i = \dim F_A(\lambda)$  であることと、 $F_A(\lambda) \supseteq E_A(\lambda)$  であることより従う。□

#### V.11.4.4 $AB = O$ でも $\mathbb{R}^n = \operatorname{Ker}(A) + \operatorname{Ker}(B)$ とは限らない

*Proof.* 例えば  $A = B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

□



## V.11.5 Jordan 標準形

### V.11.5.1 Jordan ブロックの逆行列

$n$  次の Jordan ブロック

$$J(\lambda, n) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ & & 0 & \lambda & 1 \\ 0 & \cdots & & 0 & \lambda \end{bmatrix}$$

に対して

$$J(\lambda, n)^{-1} = \left[ \mathbb{1}\{i \leq j\} \frac{(-1)^{j-i}}{\lambda^{j-i+1}} \right]_{n \times n} = \begin{bmatrix} \frac{1}{\lambda} & \frac{-1}{\lambda^2} & \frac{1}{\lambda^3} & \cdots & \frac{(-1)^{n-1}}{\lambda^n} \\ 0 & \frac{1}{\lambda} & \frac{-1}{\lambda^2} & \cdots & \frac{(-1)^{n-2}}{\lambda^{n-1}} \\ \vdots & 0 & \frac{1}{\lambda} & \ddots & \vdots \\ 0 & & \ddots & \ddots & \frac{-1}{\lambda^2} \\ 0 & 0 & \cdots & 0 & \frac{1}{\lambda} \end{bmatrix}$$

*Proof.*

(M 演算 ([1] lesson10) を使う方法):

$$J(\lambda, n)^{-1} = M(\lambda)^{-1} = M(1/\lambda) = \left[ \mathbb{1}\{i \leq j\} \frac{1}{(i-j)!} \frac{d^{i-j}}{d\lambda^{i-j}} \frac{1}{\lambda} \right]_{n \times n} = \left[ \mathbb{1}\{i \leq j\} \frac{(-1)^{j-i}}{\lambda^{j-i+1}} \right]_{n \times n}$$

(M 演算を使わない方法):

帰納法で示す。 $n$  次の Jordan ブロックに対して定理が成り立つと仮定して  $(n+1)$  次のときも成り立つことを示す。 $J(\lambda, n+1)$  を次のように分割して

$$J(\lambda, n+1) = \begin{bmatrix} J(\lambda, n) & \mathbf{b} \\ \mathbf{0}_n^\top & \lambda \end{bmatrix}, \quad \mathbf{b} = [0 \cdots 0 \ 1]^\top \in \mathbb{C}^n$$

と表すと、この逆行列は次のように表される ([4] 問題 2.4.8)。

$$J(\lambda, n+1)^{-1} = \begin{bmatrix} J(\lambda, n)^{-1} & \mathbf{v} \\ \mathbf{0}_n^\top & 1/\lambda \end{bmatrix}, \quad \mathbf{v} = \frac{-1}{\lambda} J(\lambda, n)^{-1} \mathbf{b} = \left[ \frac{-1}{\lambda} J(\lambda, n)^{-1} [i][n] \right]_{n \times 1} = \left[ \frac{(-1)^{(n+1)-i}}{\lambda^{(n+1)-i+1}} \right]_{n \times 1}$$

よって  $n+1$  次のときも定理が成り立つ。  $\square$

### V.11.5.2 Jordan ブロックの指数行列

$n$  次の Jordan ブロック

$$J(\lambda, n) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ & & 0 & \lambda & 1 \\ 0 & \cdots & & 0 & \lambda \end{bmatrix}$$

に対して

$$e^{J(\lambda, n)t} = \begin{bmatrix} e^{\lambda t} & te^{\lambda t} & \cdots & \frac{t^{n-1}}{(n-1)!}e^{\lambda t} \\ 0 & e^{\lambda t} & \ddots & \vdots \\ \vdots & \ddots & \ddots & te^{\lambda t} \\ 0 & \cdots & 0 & e^{\lambda t} \end{bmatrix}$$

*Proof.*

$J(\lambda, n)$  を次のように対角成分とその一つ上の斜めラインに分ける。

$$J(\lambda, n) = \Gamma + U, \quad \Gamma := \lambda I_{n \times n}, \quad U := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & 0 & 1 \\ 0 & \cdots & & & 0 \end{bmatrix}$$

$\Gamma$  と  $U$  は可換であり ( $\Gamma$  が単位行列の定数倍であることから明らか)、 $\Gamma$  に  $U$  を左或いは右から掛けると  $\Gamma$  の各行が上に一段シフトする (或いは各列が右に 1 段シフトするとも言える)。溢れた行或いは列は消え去り、補充されない。溢れた分だけ反対側から空行或いは空列が注入される。よって  $[J(\lambda, n)t]^p$ ,  $p \in \mathbb{N} + \{0\}$  の対角成分より下側の成分は 0 であることがわかる。

$$[J(\lambda, n)t]^p = t^p(\Gamma + U)^p = t^p \sum_{k=0}^p \binom{p}{k} U^k \Gamma^{p-k} \quad (1)$$

$(\Gamma + U)^p$  の  $(i, j)$  成分を  $a_{i,j}(p)$  とする。先述の通り  $a_{i,j}(p) = 0$ ,  $i > j$  である。

以下  $i \leq j$  とする。式 (1) の  $U^k \Gamma^{p-k}$  において  $(i, j)$  要素が非零となるのは  $k = j - i$  のときに限る (当然、同時に  $p \geq j - i$  でなくてはならない)。各  $p$  に対する  $a_{i,j}(p)$  を表にすると次のようになる。

表 V.11.1  $p$  と  $a_{i,j}(p)$

$p$	$a_{i,j}(p)$
0	0
1	0
$\vdots$	$\vdots$
$j-i$	${}_p C_{j-i}$
$j-i+1$	$\lambda_p {}_p C_{j-i}$
$j-i+2$	$\lambda^2 {}_p C_{j-i}$
$\vdots$	$\vdots$
$p$	$\lambda^{p-(j-i)} {}_p C_{j-i}$

よって  $e^{J(\lambda,n)t}$  の  $(i,j)$  成分は

$$\begin{aligned}
 \sum_{p=0}^{\infty} \frac{1}{p!} t^p a_{i,j}(p) &= \sum_{p=j-i}^{\infty} \frac{1}{p!} t^p \lambda^{p-(j-i)} {}_p C_{j-i} \\
 &= \sum_{l=0}^{\infty} \frac{1}{(l+(j-i))!} t^{l+(j-i)} \lambda^l \frac{(l+(j-i))!}{(j-i)!l!} \quad (l = p - (j-i) \text{ とおいた}) \\
 &= \frac{t^{j-i}}{(j-i)!} \sum_{l=0}^{\infty} \frac{(t\lambda)^l}{l!} = \frac{t^{j-i}}{(j-i)!} e^{\lambda t}
 \end{aligned}$$

□

## V.11.6 Perron の定理

$n$  次正行列  $A$  について次が成り立つ

1.  $\rho(A) > 0$
2.  $\rho(A)$  は  $A$  の単純固有値の一つと等しく、 $\rho(A)$  は他のどの固有値の絶対値よりも真に大きい
3.  $\rho(A)$  と等しい固有値に対応する正の固有ベクトルが存在する

Perron-Frobenius の定理の主張との違いは 2. の「 $\rho(A)$  は他のどの固有値の絶対値よりも真に大きい」である。

*Proof.*

1. と 3. および 2. の「 $\rho(A)$  は  $A$  の単純固有値の一つと等しく」は [2] の Perron-Frobenius の定理の証明と同じ方法で証明できる。2. の「 $\rho(A)$  は他のどの固有値の絶対値よりも真に大きい」を、[2] 170 頁 定理 6.11 の Perron-Frobenius の定理の証明を利用して示す。その証明の中に

$$\lambda_i x_i = \sum_{j=1}^n a_{ij} x_j \quad (1)$$

$$|\lambda| |x_i| = \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \sum_{j=1}^n a_{ij} |x_j|$$

という式があるが、 $|\lambda| \leq \hat{\lambda}$  は既に示されているので、上式に於いて等号が成り立てば  $\lambda = \hat{\lambda}$  であることを示せばよい。

$A$  は正行列だから  $a_{ij} > 0$  である。 $\boldsymbol{x}$  は固有ベクトルであるから  $x_j$  の少なくともどれか一つは非零である。よって等号が成り立つとき  $x_j$  は全て同じ向きである。言い換えれば、ある非零な複素数  $z$  と  $b_1, \dots, b_n \geq 0$  が存在して

$$\boldsymbol{x} = z \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

である。これを式 (1) に代入して

$$\begin{aligned} \lambda z b_i &= \sum_{j=1}^n a_{ij} z b_j \\ \therefore \lambda b_i &= \sum_{j=1}^n a_{ij} b_j \end{aligned}$$

$b_1, \dots, b_n$  の少なくとも 1 つは正である。それを  $b_k$  とすると

$$\lambda = \frac{1}{b_k} \sum_{j=1}^n a_{ij} b_j > 0$$

これと  $|\lambda| = \hat{\lambda}$  より  $\lambda = \hat{\lambda}$

□

## 第 V.12 章

# 置換行列

### V.12.1 定義

定義.

$I_{n \times n}$  の 2 つの列 (第  $i, j$  列) を入れ替えてできる対称行列を基本置換行列という。これを左から掛けると第  $i, j$  行の交換になり、右から掛けると第  $i, j$  行の交換になる。基本置換行列の積を置換行列という。置換行列は各行、各列に 1 つだけ 1 をもち、他は全て 0 である。

### V.12.2 諸定理

#### V.12.2.1 行と列の同期入れ替え

$\Pi$  を  $n$  次の置換行列とする。 $A \in \mathbb{F}^{n \times n}$  に対して  $\Pi^T A \Pi$  は  $A$  の行と列に同じ入れ替えを行う。

*Proof.*

置換行列  $\Pi$  は基本置換行列  $\Pi_1, \Pi_2, \dots, \Pi_n$  の積で表される。各  $\Pi_i$  は対称行列であるから

$$\Pi^T A \Pi = \Pi_n^T \cdots \Pi_1^T A \Pi_1 \cdots \Pi_n = \Pi_n \cdots \Pi_1 A \Pi_1 \cdots \Pi_n$$

であり、明らかに行と列に同じ入れ替えを行っている。 □

#### V.12.2.2 置換行列は直交行列

置換行列は直交行列である

*Proof.*

置換行列は単位行列の列を適当に入れ替えて作ることができる (または行を入れ替えて作ることでもある)。単位行列の各列ベクトルは正規直交であるが、列を入れ替えても正規直交性は維持され、置換行列の各列ベクトルは正規直交なので直交行列である。 □

## 第 V.13 章

# トレース

### V.13.1 $\operatorname{tr}(AB) = \operatorname{tr}(BA)$

$m \times n$  行列  $A$  と  $n \times m$  行列  $B$  に対して  $\operatorname{tr}(AB) = \operatorname{tr}(BA)$

*Proof.*

$$\operatorname{tr}(BA) = \sum_{l=1}^n (BA)[l][l] = \sum_{l=1}^n \sum_{k=1}^m b_{lk} a_{kl} = \sum_{k=1}^m \sum_{l=1}^n a_{kl} b_{lk} = \sum_{k=1}^m (AB)[k][k] = \operatorname{tr}(AB)$$

□

### V.13.2 Hermite 行列 $A, B$ に対して $\operatorname{tr}(AB) \in \mathbb{R}$

$A, B \in \mathbb{C}^{n \times n}$  がエルミート行列であるとき、 $\operatorname{tr}(AB)$  は実数である。

*Proof.*

$$\overline{\operatorname{tr}(AB)} = \operatorname{tr}(\overline{AB}) = \operatorname{tr}(\overline{A^* B^*}) = \operatorname{tr}(A^\top B^\top) = \operatorname{tr}((BA)^\top) = \operatorname{tr}(BA) = \operatorname{tr}(AB)$$

□

### V.13.3 正定値行列 $A \in \mathbb{R}^{m \times m}$ に対して $\operatorname{tr}(X^\top AX)$ ( $X \in \mathbb{R}^{m \times n}$ ) は $X$ の狭義凸関数である

*Proof.*

$\lambda \in [0, 1]$ ,  $X, Y \in \mathbb{R}^{m \times n}$  とする。

$$\begin{aligned} & (1 - \lambda)f(X) + \lambda f(Y) - (f((1 - \lambda)X + \lambda Y)) \\ &= (1 - \lambda) \operatorname{tr}(X^\top AX) + \lambda \operatorname{tr}(Y^\top AY) - \operatorname{tr}(((1 - \lambda)X + \lambda Y)^\top A((1 - \lambda)X + \lambda Y)) \\ &= \lambda(1 - \lambda) \operatorname{tr}((X - Y)^\top A(X - Y)) \\ &= \lambda(1 - \lambda) \sum_{i=1}^n (X[:, i] - Y[:, i])^\top A(X[:, i] - Y[:, i]) \end{aligned}$$

$A$  は正定なので上式の  $\Sigma$  の中身は正であるから、上式は  $\lambda = 0, 1$  でない限り正である。 □

## 第 V.14 章

# 定値性

### V.14.1 正定 Hermite 対称行列の逆行列も正定

*Proof.*

$A \in \mathbb{C}^{n \times n}$  を Hermite 対称な正定行列とし、重複を含めた  $n$  個の固有値を  $\lambda_1, \dots, \lambda_n > 0$  とする。 $A$  は Hermite 対称だから適当なユニタリ行列  $P \in \mathbb{C}^{n \times n}$  が存在して  $A = P \operatorname{diag}(\lambda_1, \dots, \lambda_n) P^*$  と表せるので  $A^{-1} = P \operatorname{diag}(\lambda_1, \dots, \lambda_n)^{-1} P^* = P \operatorname{diag}(1/\lambda_1, \dots, 1/\lambda_n) P^*$  である。よって  $A^{-1}$  は対称で固有値が全て正なので正定である。  $\square$

### V.14.2 $A$ が半正定で $x^\top A x = 0$ ならば $x \in \operatorname{Ker}(A)$

*Proof.*

$A$  を  $n \times n$  実対称行列とする。 $A$  を対角化する実直交行列を  $P = [\mathbf{p}_1, \dots, \mathbf{p}_n]$  とし、 $\mathbf{p}_i$  に対応する固有値を  $\lambda_i$  とすると、 $A$  が半正定だから  $\lambda_i \geq 0$  である。 $\mathbf{x}$  を  $\mathbb{R}^n$  の正規直交基底  $\{\mathbf{p}_1, \dots, \mathbf{p}_n\}$  で展開して  $\mathbf{x} = \sum_{i=1}^n c_i \mathbf{p}_i$  とすると  $0 = \mathbf{x}^\top A \mathbf{x} = \sum_{i=1}^n \lambda_i c_i^2$  だから正の固有値に対応する  $c_i$  は全て 0 である。ゆえに  $\mathbf{x}$  は  $A$  の固有値 0 の固有空間すなわち  $\operatorname{Ker}(A)$  に属する。  $\square$

### V.14.3 Gram 行列は半正定

*Proof.*

$V$  を内積空間とし、 $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$  の Gram 行列を  $G = [\langle \mathbf{v}_i, \mathbf{v}_j \rangle]_{n \times n}$  とする。 $\mathbf{x} \in \mathbb{C}^n$  に対して

$$\mathbf{x}^\top G \mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n x_i x_j \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \sum_{i=1}^n \sum_{j=1}^n \langle x_i \mathbf{v}_i, x_j \mathbf{v}_j \rangle = \left\langle \sum_{i=1}^n x_i \mathbf{v}_i, \sum_{j=1}^n x_j \mathbf{v}_j \right\rangle = \left\| \sum_{i=1}^n x_i \mathbf{v}_i \right\|_2^2 \geq 0$$

$\square$

### V.14.4 正定値行列同士の積は正定値とは限らない

*Proof.*



(例) : 75° 左回転行列同士の積

(※ : 行列が正定値というのは、飛ばす前のベクトルと飛ばしたあとのベクトルの成す角が 90° 未満であることと同値) □

## V.14.5 対角成分が全て非負の対称行列が半正定とは限らない

*Proof.*

例えば  $A = \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix}$  とすると  $x_1 = x_2$  なる  $\mathbf{x}$  に対して  $\mathbf{x}^\top A \mathbf{x} = -2x_1^2 < 0$  □

## V.14.6 非対称行列は固有値が全て正でも正定とは限らない

*Proof.*

例えば  $A = \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}$  の固有値は全て 1 であるが  $[1, 1]A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -1$  □

## V.14.7 $A$ : Hermite 対称で正定 $\Rightarrow (\mathbf{x}^* A \mathbf{x})(\mathbf{y}^* A^{-1} \mathbf{y}) \geq |\mathbf{x}^* \mathbf{y}|^2$

$A \in \mathbb{C}^{n \times n}$  は Hermite 対称で正定であるとする。このとき、任意の  $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$  に対して次式が成り立つ。

$$(\mathbf{x}^* A \mathbf{x})(\mathbf{y}^* A^{-1} \mathbf{y}) \geq \|\mathbf{x}^* \mathbf{y}\|_2^2$$

等号が成立するのは  $\mathbf{x}$  と  $\mathbf{y}$  どちらかが  $\mathbf{0}$  であるか、そうでなければ  $\mathbf{y}$  と  $A\mathbf{x}$  が平行であるとき、かつそのときである。

*Proof.*

$A$  は Hermite 対称で正定であるから、適当なユニタリ行列で対角化可能であり、固有値が全て正であることに注意すれば、 $\sqrt{A}$ ,  $\sqrt{A^{-1}}$  を定義できる。これを用いると次式を得る。

$$(\mathbf{x}^* A \mathbf{x})(\mathbf{y}^* A^{-1} \mathbf{y}) = \|\sqrt{A}\mathbf{x}\|_2^2 \|\sqrt{A^{-1}}\mathbf{y}\|_2^2 \geq \left| \langle \sqrt{A}\mathbf{x}, \sqrt{A^{-1}}\mathbf{y} \rangle \right|^2 = |\mathbf{x}^* \mathbf{y}|^2$$

2 つ目の不等号は Cauchy-Schwartz の不等式より導かれる。 □

## 第 V.15 章

# 擬逆行列

### V.15.1 諸定理: $(A^H A)^\dagger A^H = A^\dagger$ など

以下のものは特異値分解を用いて示せる。

1.  $(A^H A)^\dagger A^H = A^\dagger$
2.  $(A^H A)^\dagger A^H = A^\dagger$

以下のものも上と同様に示せる。 $AA^H$ ,  $A^H A$  が半正定値 Hermite 行列であることと、半正定値 Hermite 行列とその擬逆行列が可換であること (半正定値 Hermite 行列は固有値が全て非負であることと、特異値分解が直交行列による対角化と一致することから解る) を用いる。

1.  $(AA^H)^\dagger A = A^{H\dagger} (= A^{\dagger H})$
2.  $(AA^H)(AA^H)^\dagger A = (AA^H)^\dagger (AA^H)A = A$
3.  $(A^H A)(A^H A)^\dagger A^H = (A^H A)^\dagger (A^H A)A^H = A^H$
4.  $A(A^H A)(A^H A)^\dagger = A(A^H A)^\dagger (A^H A) = A$

### V.15.2 行フルランクな行列の擬逆行列を用いた線形方程式の解がノルム最小であること

$A \in \mathbb{C}^{m \times n}$  は行フルランクであるとし、 $\mathbf{b} \in \mathbb{C}^m$  とする。線形方程式  $A\mathbf{x} = \mathbf{b}$  の解の一つである  $A^\dagger \mathbf{b}$  は全ての解の中で  $L^2$  ノルムが最小である。

*Proof.*

$A^\dagger \mathbf{b}$  が解であることと、 $A^\dagger = A^*(AA^*)^{-1}$  であることは既知とする (特異値分解を用いて確かめられる)。 $A\mathbf{x} = \mathbf{b}$  の任意の解を  $\mathbf{x}$  とすると、 $A(A^\dagger \mathbf{x} - \mathbf{x}) = \mathbf{0}$  だから  $A^\dagger \mathbf{x} - \mathbf{x} \in \text{Ker}(A)$  つまりある適当な  $\mathbf{v} \in \text{Ker}(A)$  が存在して  $\mathbf{x} = A^\dagger \mathbf{b} + \mathbf{v}$  である。

$$\begin{aligned}\|\mathbf{x}\|_2^2 &= \|A^\dagger \mathbf{b}\|_2^2 + \|\mathbf{v}\|_2^2 + 2\text{Re}(\langle \mathbf{v}, A^\dagger \mathbf{b} \rangle) = \|A^\dagger \mathbf{b}\|_2^2 + \|\mathbf{v}\|_2^2 + 2\text{Re}(\mathbf{v}^* A^*(AA^*)^{-1} \mathbf{b}) \\ &= \|A^\dagger \mathbf{b}\|_2^2 + \|\mathbf{v}\|_2^2 + 2\text{Re}((A\mathbf{v})^*(AA^*)^{-1} \mathbf{b}) \quad (\because \mathbf{v} \in \text{Ker}(A)) \\ &= \|A^\dagger \mathbf{b}\|_2^2 + \|\mathbf{v}\|_2^2 \geq \|A^\dagger \mathbf{b}\|_2^2\end{aligned}$$



## 第 V.16 章

# ベクトルのノルム

### V.16.1 諸定理

$$\text{V.16.1.1} \quad \|\mathbf{x}\|_{p+a} \leq \|\mathbf{x}\|_p \quad (\mathbf{x} \in \mathbb{R}^n, p \geq 1, a \geq 0)$$

*Proof.*

$a = 0$  のときは明らか。 $a > 0$  のときを考える。 $x_1, \dots, x_n$  のうち非零要素が 1 つ以下のときは明らか。非零要素が 2 つ以上のときを考える。一般性を失わずに  $|x_1|, \dots, |x_r| > 0$  ( $2 \leq r \leq n$ ) とする。 $\frac{d}{dp} \|\mathbf{x}\|_p < 0$  を示せば良い。

$$\begin{aligned} \frac{d}{dp} \|\mathbf{x}\|_p < 0 &\iff \frac{d}{dp} \log \|\mathbf{x}\|_p < 0 \iff -\frac{1}{p^2} \log \sum_{i=1}^r |x_i|^p + \frac{1}{p} \times \frac{\sum_{i=1}^r |x_i|^p \log |x_i|}{\sum_{i=1}^r |x_i|^p} < 0 \\ &\iff \frac{\sum_{i=1}^r |x_i|^p \log |x_i|}{\sum_{i=1}^r |x_i|^p} < \frac{1}{p} \log \sum_{i=1}^r |x_i|^p \\ &\iff \sum_{i=1}^r w_i \log |x_i|^p < \log M \quad \left( M := \sum_{i=1}^r |x_i|^p, w_i := |x_i|^p / M \right) \\ &\iff \sum_{i=1}^r w_i \log M w_i < \log M \iff \sum_{i=1}^r w_i (\log M + \log w_i) < \log M \\ &\iff \sum_{i=1}^r w_i \log w_i < 0 \quad \left( \because \sum_{i=1}^r w_i = 1 \right) = \text{true} \quad (\because 0 < w_i < 1) \end{aligned}$$

□

$$\text{V.16.1.2} \quad 1 \leq p \leq q, \mathbf{x} \in \mathbb{C}^n \text{ のとき } \|\mathbf{x}\|_p \leq n^{1/p-1/q} \|\mathbf{x}\|_q$$

*Proof.*

$\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$  とする。 $r = q/p, s = q/(q-p)$  とすると  $1 \leq r, s$  および  $1/r + 1/s = 1$  を満たすので Hölder の不等式より次式が成り立つ。

$$\sum_{i=1}^n |a_i b_i| \leq \left( \sum_{i=1}^n |a_i|^r \right)^{\frac{1}{r}} \left( \sum_{i=1}^n |b_i|^s \right)^{\frac{1}{s}}$$

$|a_i| = |x_i|^p$ ,  $b_i = 1$  とすると次式を得る。

$$\sum_{i=1}^n |x_i|^p \leq \left( \sum_{i=1}^n |x_i|^q \right)^{\frac{p}{q}} n^{1-p/q}$$

両辺の  $p$  乗根をとって次式を得る。

$$\|\mathbf{x}\|_p \leq n^{1/p-1/q} \|\mathbf{x}\|_q$$

□

### V.16.1.3 和のノルムとノルムの和が等しい時

$\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$  が非零であるとき  $\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\| \iff \exists k > 0 \text{ s.t. } \mathbf{y} = k\mathbf{x}$

*Proof.*

$\Leftarrow$  は明らか。  $\Rightarrow$  を示す。

$$\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$$

の両辺を 2 乗して整理すると

$$\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\| \|\mathbf{y}\| \quad (1)$$

よって

$$\langle \mathbf{x}, \mathbf{y} \rangle^2 - \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 = 0$$

ところでこれは  $s$  に関する二次方程式  $\|\mathbf{s}\mathbf{x} + \mathbf{y}\|^2 = 0$  の判別式の値が 0 であることを意味する。よって重解が存在し、それを  $-k$  とおくと

$$\mathbf{y} = k\mathbf{x} \quad (2)$$

式 (1) に代入して

$$k \|\mathbf{x}\|^2 = |k| \|\mathbf{x}\|^2$$

$\mathbf{x} \neq \mathbf{0}$  だから  $\|\mathbf{x}\| > 0$  なので  $k = |k| \therefore k \geq 0$ 。さらに  $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$  であったから式 (2) より  $k \neq 0$ 。結局  $k > 0$  □

### V.16.1.4 和のノルムとノルムの和が等しい時 (2)

$\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathbb{F}^n$  に対して  $\|\sum_{i=1}^m \mathbf{x}_i\| = \sum_{i=1}^m \|\mathbf{x}_i\|$  であるとき、任意の  $k, l \in \{1, 2, \dots, m\}$  に対して  $\|\mathbf{x}_k + \mathbf{x}_l\| = \|\mathbf{x}_k\| + \|\mathbf{x}_l\|$

*Proof.*

$k = l$  については明らか。  $k \neq l$  について示す。三角不等式より

$$\left\| \sum_{i=1}^m \mathbf{x}_i \right\| \leq \|\mathbf{x}_k + \mathbf{x}_l\| + \sum_{i \neq k, l} \|\mathbf{x}_i\| \leq \sum_{i=1}^m \|\mathbf{x}_i\|$$

仮定より最右辺が  $\|\sum_{i=1}^m \mathbf{x}_i\|$  と等しいから結局

$$\left\| \sum_{i=1}^m \mathbf{x}_i \right\| \leq \|\mathbf{x}_k + \mathbf{x}_l\| + \sum_{i \neq k, l} \|\mathbf{x}_i\| \leq \left\| \sum_{i=1}^m \mathbf{x}_i \right\|$$

$$\begin{aligned}\therefore \left\| \sum_{i=1}^m \mathbf{x}_i \right\| &= \|\mathbf{x}_k + \mathbf{x}_l\| + \sum_{i \neq k, l} \|\mathbf{x}_i\| = \sum_{i=1}^m \|\mathbf{x}_i\| \\ \therefore \|\mathbf{x}_k + \mathbf{x}_l\| &= \|\mathbf{x}_k\| + \|\mathbf{x}_l\|\end{aligned}$$

□

$$\text{V.16.1.5} \quad \mathbf{x} \in \mathbb{C}^n, \|\mathbf{x}\|_2 = 1 \Rightarrow \sum_{i=1}^n |x_i| \leq \sqrt{n}$$

*Proof.*

$x_k = r_k e^{i\theta_k}$ , ( $r_k \geq 0, \theta_k \in \mathbb{R}, k = 1, \dots, n$ ) とすると  $\|\mathbf{x}\|_2 = 1 \iff \sum_{k=1}^n r_k^2 = 1, \sum_{k=1}^n |x_k| = \sum_{k=1}^n r_k$  であるから、制約条件  $g(\mathbf{r}) := \sum_{k=1}^n r_k^2 - 1 = 0, \mathbf{r} \geq \mathbf{0}_n$  の下で関数  $f(\mathbf{r}) := \sum_{k=1}^n r_k$  の最大値が  $\sqrt{n}$  以下であること (\*) を示せばよい。

数学的帰納法で示す。 $n = 1$  のときは明らかに (\*) が成り立つ。以下  $n = m \geq 2$  とし、 $n = 1, \dots, m-1$  について (\*) が成り立つと仮定する。 $r_1, \dots, r_n$  の少なくとも 1 つが 0 のとき、すなわち制約集合の境界上では変数の個数が  $n-1$  以下の場合に帰着するから帰納法の仮定より成り立つ。

そこで  $r_1, \dots, r_n$  が全て正であるとする。Lagrange の未定乗数法を使う。Lagrange 関数を  $L(\mathbf{r}, \lambda) := f(\mathbf{r}) + \lambda g(\mathbf{r})$  とする。 $\nabla_{\mathbf{r}} g(\mathbf{r}) = 2\mathbf{r} > \mathbf{0}_n$  であるから、今関心のある領域に特異点は存在しない。 $r_n = \sqrt{1 - \sum_{k=1}^{n-1} r_k^2}$  と表されるから、 $f(\mathbf{r}) = \sum_{k=1}^{n-1} r_k + \sqrt{1 - \sum_{k=1}^{n-1} r_k^2}$  であり、これは  $r_1, \dots, r_{n-1}$  の関数として上に凸な関数である (Hessian を計算すれば対角成分が全て負の対角行列になる)。よって Lagrange の未定乗数法で極値を見つければそれは極大値であり、また最大値でもある。極値を与える点を  $\overset{\circ}{\mathbf{r}}$  とすると  $\nabla_{\mathbf{r}} L(\overset{\circ}{\mathbf{r}}, \lambda) = \mathbf{0}$  より  $\mathbf{1}_n + 2\lambda \overset{\circ}{\mathbf{r}} = \mathbf{0} \therefore \overset{\circ}{\mathbf{r}} = -\frac{1}{2\lambda} \mathbf{1}_n$ 。よって  $\overset{\circ}{\mathbf{r}}$  は全要素が等しい。これと制約条件より  $\overset{\circ}{\mathbf{r}} = \frac{1}{\sqrt{n}} \mathbf{1}_n$  であり、 $f(\overset{\circ}{\mathbf{r}}) = \sqrt{n}$  □

## V.16.2 応用

### V.16.2.1 超平面と点の距離公式の直感的説明

超平面  $\mathbf{w}^\top \mathbf{x} + b = 0$  と点  $\mathbf{y}$  の距離公式  $|\mathbf{w}^\top \mathbf{y} + b| / \|\mathbf{w}\|$  について直感的な説明を考えてみる。 $-\mathbf{w}^\top \mathbf{p} = b$  なる  $\mathbf{p}$  を考えると  $\mathbf{w}^\top \mathbf{x} + b = 0 \iff \mathbf{w}^\top (\mathbf{x} - \mathbf{p}) = 0$  であり、 $\mathbf{p}$  は超平面上の点で、 $\mathbf{w}$  は法線ベクトルである。よって  $|\mathbf{w}^\top \mathbf{y} + b| = |\mathbf{w}^\top (\mathbf{y} - \mathbf{p})|$  は  $\mathbf{y}$  から超平面に降ろした垂線の長さの  $\|\mathbf{w}\|$  倍を表しており、これを  $\|\mathbf{w}\|$  で割れば公式を得る。

## 第 V.17 章

# 行列のノルム

### V.17.1 2-ノルム (2-演算子ノルム)

#### V.17.1.1 定義

$A \in \mathbb{C}^{m \times n}$  の 2-ノルム (「2-演算子ノルム」とも呼ばれる)  $\|A\|_2$  は次の 3 つの等価な式で定義される。

$$\begin{aligned}\|A\|_2 &= \max \left\{ \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} : \mathbf{x} \neq \mathbf{0} \right\} \\ &= \max \{ \|A\mathbf{x}\|_2 : \|\mathbf{x}\|_2 = 1 \} \\ &= \min \{ c > 0 : \|A\mathbf{x}\|_2 \leq c \|\mathbf{x}\|_2, \mathbf{x} \in \mathbb{C}^n \}\end{aligned}$$

*Proof.*

上述の 3 式が等しいことを示す。第 1, 2, 3 式の与える値を  $c_1, c_2, c_3$  とし、3 つの式に対応する集合を 3 つ定義する。

$$S_1 := \left\{ \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} : \mathbf{x} \neq \mathbf{0} \right\}, \quad S_2 := \{ \|A\mathbf{x}\|_2 : \|\mathbf{x}\|_2 = 1 \}, \quad S_3 := \{ c > 0 : \|A\mathbf{x}\|_2 \leq c \|\mathbf{x}\|_2, \mathbf{x} \in \mathbb{C}^n \}$$

まず

$$S_2 = \left\{ \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} : \|\mathbf{x}\|_2 = 1 \right\} \subseteq S_1$$

より  $c_1 \geq c_2$  である。

$c_1$  を与えるベクトル  $\mathbf{x}_1$  に対して  $\mathbf{x}'_1 := \mathbf{x}_1 / \|\mathbf{x}_1\|_2$  とすれば  $\|\mathbf{x}'_1\|_2 = 1$  であり、 $\|A\mathbf{x}'_1\|_2 = \frac{\|A\mathbf{x}_1\|_2}{\|\mathbf{x}_1\|_2}$  であるからつまり  $c_1 \in S_2$  である ( $S_1$  の最大要素が  $S_2$  に属する)。よって  $c_1 \leq c_2$ 。これと  $c_1 \geq c_2$  より  $c_1 = c_2$ 。

$c \in S_3$  は任意の  $\mathbf{x} \neq \mathbf{0}$  に対して  $\frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq c$  を満たすから  $c_1 \leq c$ 。よって  $c_1 \leq c_3$ 。

$c_1$  はその定義より任意の  $\mathbf{x} \neq \mathbf{0}$  に対して  $\frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq c_1$  を満たすから任意の  $\mathbf{x}$  に対して  $\|A\mathbf{x}\|_2 \leq c_1 \|\mathbf{x}\|_2$  を満たす。よって  $c_1 \geq c_3$ 。これと  $c_1 \leq c_3$  より  $c_1 = c_3$ 。

以上より  $c_1 = c_2 = c_3$

□

以降、特に断らない限り  $\|\cdot\|$  は 2-ノルムを表す。

### V.17.1.2 部分行列のノルム: $A \in \mathbb{C}^{m \times n}$ , $\|[A]_{i,j}\|_2 \leq \|A\|_2$

*Proof.*

$\max_{\|\mathbf{x}\|=1} \|[A]_{i,j}\mathbf{x}\|$  を最大にする  $\mathbf{x}$  に対して、 $\mathbf{x}' = [\mathbf{x}[1:j-1]]^\top, 0, \mathbf{x}[j:n-1]]^\top$  とすると、 $\|\mathbf{x}'\| = 1$ ,  $\|A\mathbf{x}'\| \geq \|[A]_{i,j}\mathbf{x}\|$  □

### V.17.1.3 スペクトル半径の上界

$A \in \mathbb{C}^{n \times n}$  について  $\rho(A) \leq \|A\|$

*Proof.*

$A$  の絶対値最大固有値を  $\lambda_{\max}$ , その任意の固有ベクトルを  $\mathbf{x}$  とすると

$$\rho(A) := |\lambda_{\max}| = \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \leq \max_{\mathbf{y} \neq \mathbf{0}} \frac{\|A\mathbf{y}\|}{\|\mathbf{y}\|} =: \|A\|$$

□

### V.17.1.4 Hermite 行列の絶対値最大固有値の絶対値は 2-演算子ノルムと一致する

$A \in \mathbb{C}^{n \times n}$  が Hermite 行列であるとき  $\max |\lambda(A)| = \|A\|$

(短いが高度な知識を前提とした証明)

*Proof.*

$A$  は Hermite 行列だから固有値は全て実数で、適当なユニタリ行列  $U$  で対角化できる。 $A$  の対角化を  $A = UDU^H$  とする。ここに  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ ,  $|\lambda_1| \geq \dots \geq |\lambda_n|$  である。 $A^H A = U D^H D U^H$  は半正定 Hermite であり固有値は全て非負で、 $U D^H D U^H$  自身が  $A^H A$  の特異値分解である。よって  $\max |\lambda(A)| = |\lambda_1| = \sqrt{\|A^H A\|_2} = \sqrt{\|A\|_2^2} = \|A\|_2$  □

(初等的な証明)

*Proof.*

$A$  は Hermite 行列であるから固有値は全て実数で、 $A$  がユニタリ交行列で対角化できることは既知とする。 $A$  の異なる固有値を、絶対値が大きい順に  $\lambda_1, \lambda_2, \dots, \lambda_\sigma$  ( $\sigma \leq n$ ) とし、 $\lambda_i$  に属する固有空間  $W(\lambda_i; A)$  の次元を  $n_i$  とする。 $\sum_{i=1}^\sigma n_i = n$  であり、 $W(\lambda_i; A)$  の正規直交基底を  $\{\mathbf{x}_{1,i}, \mathbf{x}_{2,i}, \dots, \mathbf{x}_{n_i,i}\}$  とすると  $\bigcup_{i=1}^\sigma \{\mathbf{x}_{1,i}, \mathbf{x}_{2,i}, \dots, \mathbf{x}_{n_i,i}\}$  は  $\mathbb{C}^n$  の正規直交基底である。

$\|A\|$  の定義は  $\|A\| := \max\{\|A\mathbf{x}\| : \|\mathbf{x}\| = 1\}$  であった。任意の  $\mathbf{x} \in \mathbb{C}^n$ ,  $\|\mathbf{x}\| = 1$  は適当な複素数  $\bigcup_{i=1}^\sigma \{c_{1,i}, c_{2,i}, \dots, c_{n_i,i}\}$ ,  $\sum_{i=1}^\sigma \sum_{k=1}^{n_i} |c_{k,i}|^2 = 1$  を用いて  $\mathbf{x} = \sum_{i=1}^\sigma \sum_{k=1}^{n_i} c_{k,i} \mathbf{x}_{k,i}$  と表せるから、

$$A\mathbf{x} = \sum_{i=1}^\sigma \lambda_i \sum_{k=1}^{n_i} c_{k,i} \mathbf{x}_{k,i}$$

よって

$$\|A\mathbf{x}\|^2 = \langle A\mathbf{x}, A\mathbf{x} \rangle = \sum_{i=1}^\sigma \lambda_i^2 \sum_{k=1}^{n_i} |c_{k,i}|^2 = \sum_{i=1}^\sigma c_i \lambda_i^2 \quad \left( c_i := \sum_{k=1}^{n_i} |c_{k,i}|^2, \quad \sum_{i=1}^\sigma c_i = 1 \right)$$



これは  $c_1 = 1, \quad c_2 = \cdots = c_\sigma = 0$  のときに最大となる。よって

$$\|A\| = \max\{\|A\mathbf{x}\| : \|\mathbf{x}\| = 1\} = |\lambda_1| = \max|\lambda(A)|$$

□

$$\text{V.17.1.5} \quad \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| = \sqrt{\lambda_{\max}(A^\top A)} \quad (A \in \mathbb{R}^{n \times n})$$

実正方行列  $A \in \mathbb{R}^{n \times n}$  と実ベクトル  $\mathbf{x} \in \mathbb{R}^n$  について

$$\max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| = \sqrt{\lambda_{\max}(A^\top A)}$$

(簡潔だが、高度な知識を前提とする証明)

*Proof.* 左辺は  $A$  の最大特異値である。 $A$  の特異値分解を  $A = U\Sigma V^\top$  ( $U, V$  は  $n$  次実直交行列) とすると  $A^\top A = V\Sigma^2 V^\top$  であり、定理の主張の右辺が  $A$  の最大特異値であることがわかる。 □

(初等的な証明)

*Proof.*

$$\max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|^2 = \lambda_{\max}(A^\top A)$$

を示せば良い。

$$\|A\mathbf{x}\|^2 = (A\mathbf{x})^\top (A\mathbf{x}) = \mathbf{x}^\top (A^\top A\mathbf{x}) = \langle \mathbf{x}, A^\top A\mathbf{x} \rangle = \langle A^\top A\mathbf{x}, \mathbf{x} \rangle$$

であり、 $A^\top A$  は実対称行列である ( $\because (A^\top A)^\top = A^\top (A^\top)^\top = A^\top A$ ) から、固有値は全て実数であり、固有空間の基底に正規直交基を選べる。

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r$  を  $A^\top A$  の固有値とし、 $n_i := \dim(W(\lambda_i; A))$  とする。

$\mathbf{b}_{i1}, \mathbf{b}_{i2}, \dots, \mathbf{b}_{in_i}$  を  $W(\lambda_i; A)$  の正規直交基とすると、ノルムが 1 であるような任意のベクトル  $\mathbf{x} \in \mathbb{R}^n$  は適当な実数  $c_{ij}$  を用いて

$$\mathbf{x} = \sum_{i=1}^r \sum_{j=1}^{n_i} c_{ij} \mathbf{b}_{ij}$$

と表現できる。但し、 $\sum_{ij} c_{ij}^2 = 1 \cdots$  (i)。

よって

$$A^\top A\mathbf{x} = A^\top A \sum_{i=1}^r \sum_{j=1}^{n_i} c_{ij} \mathbf{b}_{ij} = \sum_{i=1}^r \sum_{j=1}^{n_i} c_{ij} A^\top A \mathbf{b}_{ij} = \sum_{i=1}^r \sum_{j=1}^{n_i} c_{ij} \lambda_i \mathbf{b}_{ij} = \sum_{i=1}^r \lambda_i \sum_{j=1}^{n_i} c_{ij} \mathbf{b}_{ij}$$

となるから、 $\mathbf{b}_{ij}$  の正規直交性に注意して

$$(A^\top A\mathbf{x}, \mathbf{x}) = \left( \sum_{i=1}^r \lambda_i \sum_{j=1}^{n_i} c_{ij} \mathbf{b}_{ij}, \sum_{i=1}^r \sum_{j=1}^{n_i} c_{ij} \mathbf{b}_{ij} \right) = \sum_{i=1}^r \lambda_i \sum_{j=1}^{n_i} c_{ij}^2 \cdots \text{(ii)}$$

条件 (i) の下での (ii) の最大値が  $\lambda_{\max}(A^\top A)$  すなわち  $\lambda_1$  と等しいことを示せばよい。ラグランジュの未定乗数法を使う。

$$\begin{cases} g(c_{11}, c_{12}, \dots, c_{1n_1}, \dots, c_{r1}, c_{r2}, \dots, c_{rn_r}) = \sum_{ij} c_{ij}^2 - 1 \\ f(c_{11}, c_{12}, \dots, c_{1n_1}, \dots, c_{r1}, c_{r2}, \dots, c_{rn_r}) = \sum_{i=1}^r \lambda_i \sum_{j=1}^{n_i} c_{ij}^2 \\ F(c_{11}, c_{12}, \dots, c_{1n_1}, \dots, c_{r1}, c_{r2}, \dots, c_{rn_r}, \lambda) = f - \lambda g \end{cases}$$

とする。

$F_{c_{ij}} = 0$  とおくと  $2\lambda_i c_{ij} - 2\lambda c_{ij} = 0$  となるから  $\lambda = \lambda_i$  or  $c_{ij} = 0$ 。 $\lambda$  が  $\lambda_1, \lambda_2, \dots, \lambda_r$  のいずれとも一致しなければ  $c_{ij}$  が全ての  $ij$  で 0 になり条件 (i) を満たさない。従って  $\lambda$  は  $\lambda_1, \lambda_2, \dots, \lambda_r$  のいずれか一つと一致しなければならない。

$\lambda = \lambda_p$  ( $1 \leq p \leq r$ ) であるとき  $c_{ij}|_{i \neq p} = 0 \dots$  (iii) となり、これと条件 (i) から

$$\sum_{j=1}^{n_p} c_{pj}^2 = 1 \dots \text{(iv)}$$

となるので

$$f = (A^\top A \mathbf{x}, \mathbf{x}) = \lambda_p \sum_{j=1}^{n_p} c_{pj}^2 = \lambda_p$$

これは  $p = 1$  のとき最大値  $\lambda_1 = \lambda_{\max}(A^\top A)$  をとる。

さて、これが果たして  $f$  の最大値なのかどうかを吟味する必要がある。条件 (i) の表す領域は (iii) に注意すると (iv) の表す  $n_1$  次元超球面であるから、最大値は極大値と同じ意味になる。(iv) より  $c_{11}, c_{12}, \dots, c_{1n_1}$  の少なくとも 1 つは 0 でない。0 でないものが  $c_{1q}$  であつたとしよう。すると  $\frac{\partial g}{\partial c_{1q}} = 2c_{1q} \neq 0$  だから陰関数定理より先述の超球面は  $c_{1q}$  の近傍で

$$c_{1q}^2 = 1 - \sum_{j=1}^{q-1} c_{1j}^2 - \sum_{j=q+1}^{n_1} c_{1j}^2$$

と表現される。 $f$  は定数関数  $\lambda_1$  となっているから、その  $c_{11}, c_{12}, \dots, c_{1(q-1)}, c_{1(q+1)}, \dots, c_{1n_1}$  による一階偏微分、二階偏微分はともに 0 であり、 $\lambda_1$  が極値たる条件を満たす。

よって  $\lambda_1 = \lambda_{\max}(A^\top A)$  は  $f$  の最大値である。 □

#### V.17.1.6 逆行列のノルムは元の行列の最小特異値の逆数と等しい

正則行列  $A \in \mathbb{C}^{n \times n}$  の特異値を  $\sigma_1 \geq \dots \geq \sigma_n > 0$  とすると  $\|A^{-1}\|_2 = 1/\sigma_n$

*Proof.*  $A$  の特異値分解より明らか。 □

#### V.17.1.7 $A \in \mathbb{C}^{m \times n}$ について $\alpha = \max |A_{ij}|$ とすると $\alpha \leq \|A\|_2$

*Proof.*

絶対値最大の列番号を  $q$  とする。第  $q$  要素のみ 1 で他は 0 の単位ベクトル  $\mathbf{e}_q$  に対して  $\alpha \leq \|A\mathbf{e}_q\|_2 \leq \|A\|_2$  □

V.17.1.8  $A \in \mathbb{C}^{m \times n}$  の各列ベクトルのノルムが高々  $a$  ならば  $\|A\|_2 \leq a\sqrt{n}$

*Proof.*

$A$  の第  $i$  列ベクトルを  $\mathbf{a}_i$  とし、 $\mathbf{x} \in \mathbb{C}^n, \|\mathbf{x}\|_2 = 1$  とすると

$$\|A\mathbf{x}\|_2 = \|x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n\|_2 \leq |x_1|\|\mathbf{a}_1\|_2 + \cdots + |x_n|\|\mathbf{a}_n\|_2 \leq (|x_1| + \cdots + |x_n|)a$$

これと V.16.1.5 より  $|x_1| + \cdots + |x_n| \leq \sqrt{n}$  なので定理が従う。 □

V.17.1.8.1 系:  $A \in \mathbb{C}^{m \times n}$  の全要素の絶対値が  $\varepsilon$  以下であれば  $\|A\|_2 \leq \varepsilon\sqrt{mn}$

*Proof.*  $A$  の各列ベクトルのノルムが高々  $\varepsilon\sqrt{m}$  なので直前の定理より従う。 □

V.17.1.8.2 系:  $n$  次確率行列のノルムは高々  $\sqrt{n}$

*Proof.*

転置してもノルムが変わらないことと、確率ベクトル (の転置) のノルムは高々 1 であることを使えば V.17.1.8 より従う。 □

## 第 V.18 章

# 凸領域

### V.18.1 凸領域内のベクトルの凸結合は元の凸領域に属す

$V$  を凸領域とし、 $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ ,  $\alpha_1, \dots, \alpha_n \geq 0$ ,  $\sum_{i=1}^n \alpha_i = 1$  であるとき、 $\sum_{i=1}^n \alpha_i \mathbf{v}_i \in V$

*Proof.*

$n = 1$  のときは明らか。 $n = m \in \mathbb{N}$  まで成り立つと仮定する。 $n = m + 1$  のとき、

$$\sum_{i=1}^n \alpha_i \mathbf{v}_i = \sum_{i=1}^{m+1} \alpha_i \mathbf{v}_i = \sum_{i=1}^m \alpha_i \mathbf{v}_i + \alpha_{m+1} \mathbf{v}_{m+1} = A_m \sum_{i=1}^m \beta_i \mathbf{v}_i + \alpha_{m+1} \mathbf{v}_{m+1} \quad \left( A_m := \sum_{i=1}^m \alpha_i, \beta_i := \frac{\alpha_i}{A_m} \right)$$

$\sum_{i=1}^m \beta_i = 1$  であるから、仮定より  $\mathbf{w}_{m+1} := \sum_{i=1}^m \beta_i \mathbf{v}_i \in V$  である。そして  $A_m \geq 0$ ,  $A_m + \alpha_{m+1} = 1$  であるから  $\sum_{i=1}^n \alpha_i \mathbf{v}_i = A_m \mathbf{w}_{m+1} + \alpha_{m+1} \mathbf{v}_{m+1} \in V$  である。  $\square$

## 第 V.19 章

# 既約性

### V.19.1 非負行列が既約であるための必要十分条件

$A \in \mathbb{R}^{n \times n}$ ,  $A \geq O$  が既約であるための必要十分条件は  $(I + A)^{n-1} > O$

*Proof.*

(必要)

[2] 170 頁 定理 6.10 を参照。

(十分)

既約でないと仮定して矛盾を導く。 $A$  が既約であることと  $I + A$  が既約であることは同値である。 $A$  が既約でないと仮定したから  $I + A$  も既約でないので、適当な置換行列  $\Pi$  が存在して  $\Pi^\top(I + A)\Pi$  が上三角行列

$$\begin{bmatrix} B_{11} & B_{12} \\ O & B_{22} \end{bmatrix}$$

になる ( $B_{11}, B_{22}$  は正方行列)。正行列の左から  $\Pi^\top$  を掛けても、右から  $\Pi$  を掛けてもその結果は正行列である。また、 $\Pi$  は直交行列であったから  $\Pi\Pi^\top = I$ 。よって前提の式

$$(I + A)^{n-1} > O$$

より

$$\begin{aligned} \Pi^\top(I + A)^{n-1}\Pi &> O \\ \Pi^\top(I + A)\Pi\Pi^\top(I + A)\Pi \cdots \Pi^\top(I + A)\Pi &> O \\ (\Pi^\top(I + A)\Pi)^{n-1} &> O \\ \begin{bmatrix} B_{11} & B_{12} \\ O & B_{22} \end{bmatrix}^{n-1} &> O \end{aligned}$$

然るに左辺の左下ブロックは  $O$  だから矛盾している。

□

## 第 V.20 章

# 漏れ確率行列

### V.20.1 定義

正方行列  $A \in \mathbb{R}^{n \times n}$  が次の条件を満たすとき、 $A$  を  $o(r)$  の漏れ確率行列と呼ぶことにする。

- 各要素が 0 以上 1 未満
- 各行ベクトルの 1 ノルムが  $r < 1$  以下である。

### V.20.2 漏れ確率行列のべき乗は $O$ に収束する

*Proof.*

$A$  を  $o(r)$  の漏れ確率行列とし、 $C = A^2$  とすると  $c_{ij} = \sum_{k=1}^n a_{ik}a_{kj}$  であるから  $C$  の第  $i$  行ベクトルの 1 ノルムは  $\sum_{j=1}^n c_{ij} = \sum_{j=1}^n \sum_{k=1}^n a_{ik}a_{kj} = \sum_{k=1}^n a_{ik} \sum_{j=1}^n a_{kj} \leq r \sum_{k=1}^n a_{ik} \leq r^2$  となるので、 $C$  が非負行列であったことに注意して、各要素  $c_{ij}$  は必ず 0 以上  $r^2$  以下である。これを繰り返すと  $A^n$  の各要素は 0 以上  $r^n$  以下であることがわかるので、 $\lim_{n \rightarrow \infty} A^n = O$  □

### V.20.3 $o(r)$ の漏れ確率行列 $A$ のノイマン級数は収束し、各要素の和は $\frac{1}{1-r}$ 以下である。

*Proof.*

直前の定理より  $A$  のノイマン級数  $I + A + A^2 + \cdots$  の各要素は 0 以上  $1 + r + r^2 + \cdots = \frac{1}{1-r}$  以下である。 □

## 第 V.21 章

# 最適化への応用

### V.21.1 $\|Ax + b\|_2^2$ の最小化条件

#### V.21.1.1 制約なしの場合

$A \in \mathbb{C}^{m \times n}$ ,  $b \in \mathbb{C}^m$ ,  $x \in \mathbb{C}^n$  とする。  $f(x) := \|Ax + b\|_2^2$  がある  $\hat{x} \in \mathbb{C}^n$  で最小となる必要十分条件は次式である。

$$A^* A \hat{x} = -A^* b$$

特に  $A$  が列フルランクであれば  $A^* A$  は正則 (V.10.5) なので  $\hat{x}$  は次式で一意に定まる。

$$\hat{x} = -(A^* A)^{-1} A^* b$$

*Proof.*

IV.2.1 より  $f$  は凸であるので、 $\hat{x}$  が  $f$  を最小化する必要十分条件は次式である。

$$f(\hat{x} + dx) - f(\hat{x}) = o(\|dx\|) \quad (dx \rightarrow 0)$$

これは  $f(\hat{x} + dx)$  の  $dx$  に関する 1 次項が 0 であることと同値である。この 1 次項を  $g(dx)$  とすると

$$g(dx) = (Adx)^*(A\hat{x} + b) + (A\hat{x} + b)^*(Adx) = 2\operatorname{Re}((\hat{x}^* A^* A + b^* A)dx)$$

であるから、 $\hat{x}$  に関する条件は  $\hat{x}^* A^* A + b^* A = 0$ 、すなわち  $A^* A \hat{x} = -A^* b$  である。 □

#### V.21.1.2 線形制約付きの場合

$f$  の定義は前小節のものとする。  $m \in \mathbb{N}$ ,  $c_i \in \mathbb{C}^n$ ,  $d_i \in \mathbb{C}$ ,  $g_i(x) := c_i^\top x + d_i$  とする。ある  $\hat{x} \in \mathbb{C}^n$ ,  $\hat{\lambda} \in \mathbb{C}^m$  に対して  $g_i(\hat{x}) = 0$  ( $i = 1, \dots, m$ ),  $(A\hat{x} + b)^* A = \sum_{i=1}^m \hat{\lambda}_i c_i^\top$  であるとき、 $f$  は制約条件  $g_i(x) = 0$  ( $i = 1, \dots, m$ ) の下で  $\hat{x}$  に於いて最小値をとる。

*Proof.*

$h \in \mathbb{C}^n$ ,  $g_i(\hat{x} + h) = 0$  ( $i = 1, \dots, m$ ) とする。  $g_i(\hat{x} + h) = 0 \iff c_i^\top h = 0$  に注意する。

$$\begin{aligned} f(\hat{x} + h) &= f(\hat{x}) + 2\operatorname{Re}((A\hat{x} + b)^* Ah) + h^* A^* Ah = f(\hat{x}) + 2\operatorname{Re}\left(\sum_{i=1}^m \hat{\lambda}_i c_i^\top h\right) + h^* A^* Ah \\ &= f(\hat{x}) + h^* A^* Ah \geq f(\hat{x}) \end{aligned}$$





## 第 V.22 章

# 発想, 技巧

### V.22.1 演算

#### V.22.1.1 積和の階数を拡張する

以下、 $\mathbb{F}$  は適当な体を、 $l, m, n$  は適当な自然数を表す。

1. スカラー同士の積:  $x, y \in \mathbb{F}$  に対して  $xy$
2. 内積 (スカラー同士の積の和):  $\mathbf{x}, \mathbf{y} \in \mathbb{F}^l$  に対して  $\mathbf{x}^\top \mathbf{y} = \sum_{i=1}^l x_i y_i$
3. トレース (内積同士の和):  $X = [\mathbf{x}_1, \dots, \mathbf{x}_n]^\top, Y = [\mathbf{y}_1, \dots, \mathbf{y}_n]^\top$  ( $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{y}_n \in \mathbb{F}^l$ ) に対して  $\text{tr}(X^\top Y) = \sum_{i=1}^l \mathbf{x}_i^\top \mathbf{y}_i$

## 第 VI 部

# ベクトル解析

## 第 VI.1 章

# 3 次元 Euclid 空間

### VI.1.1 諸定義

#### VI.1.1.1 Vector Laplacian

定義.  $A : \mathbf{r} \in \mathbb{R}^3 \mapsto A(\mathbf{r}) \in \mathbb{R}^3$  ( $A$  の各成分は  $C^2$  級) の Vector Laplacian  $\nabla^2 A$  は次式で定義される。

$$\nabla^2 A := \nabla(\nabla \cdot A) - \nabla \times (\nabla \times A)$$

### VI.1.2 諸公式

#### VI.1.2.1 ベクトル三重積の公式: $A \times (B \times C) = \langle A, C \rangle B - \langle A, B \rangle C$

*Proof.*

$A, B, C$  を次のようにおく。

$$A = i_1 A_1 + i_2 A_2 + i_3 A_3$$

$$B = i_1 B_1 + i_2 B_2 + i_3 B_3$$

$$C = i_1 C_1 + i_2 C_2 + i_3 C_3$$

すると

$$\begin{aligned}
A \times (B \times C) &= (\mathbf{i}_1 A_1 + \mathbf{i}_2 A_2 + \mathbf{i}_3 A_3) \times [\mathbf{i}_1 (B_2 C_3 - B_3 C_2) + \mathbf{i}_2 (B_3 C_1 - B_1 C_3) + \mathbf{i}_3 (B_1 C_2 - B_2 C_1)] \\
&= \mathbf{i}_3 A_1 (B_3 C_1 - B_1 C_3) - \mathbf{i}_2 A_1 (B_1 C_2 - B_2 C_1) \\
&\quad - \mathbf{i}_3 A_2 (B_2 C_3 - B_3 C_2) + \mathbf{i}_1 A_2 (B_1 C_2 - B_2 C_1) \\
&\quad + \mathbf{i}_2 A_3 (B_2 C_3 - B_3 C_2) - \mathbf{i}_1 A_3 (B_3 C_1 - B_1 C_3) \\
&= \mathbf{i}_1 [(A_2 C_2 + A_3 C_3) B_1 - (A_2 B_2 + A_3 B_3) C_1] \\
&\quad + \mathbf{i}_2 [(A_1 C_1 + A_3 C_3) B_2 - (A_1 B_1 + A_3 B_3) C_2] \\
&\quad + \mathbf{i}_3 [(A_1 C_1 + A_2 C_2) B_3 - (A_1 B_1 + A_2 B_2) C_3] \\
&= \mathbf{i}_1 [\langle A, C \rangle - A_1 C_1] B_1 - (\langle A, B \rangle - A_1 B_1) C_1 \\
&\quad + \mathbf{i}_2 [\langle A, C \rangle - A_2 C_2] B_2 - (\langle A, B \rangle - A_2 B_2) C_2 \\
&\quad + \mathbf{i}_3 [\langle A, C \rangle - A_3 C_3] B_3 - (\langle A, B \rangle - A_3 B_3) C_3 \\
&= \mathbf{i}_1 [\langle A, C \rangle B_1 - \langle A, B \rangle C_1] \\
&\quad + \mathbf{i}_2 [\langle A, C \rangle B_2 - \langle A, B \rangle C_2] \\
&\quad + \mathbf{i}_3 [\langle A, C \rangle B_3 - \langle A, B \rangle C_3] \\
&= \langle A, C \rangle B - \langle A, B \rangle C
\end{aligned}$$

□

VI.1.2.2  $\langle \mathbf{a} \times \mathbf{b}, \mathbf{c} \times \mathbf{d} \rangle = \langle \mathbf{a}, \mathbf{c} \rangle \langle \mathbf{b}, \mathbf{d} \rangle - \langle \mathbf{a}, \mathbf{d} \rangle \langle \mathbf{b}, \mathbf{c} \rangle$

*Proof.*

ベクトル三重積の公式を利用する。

$$\langle \mathbf{a} \times \mathbf{b}, \mathbf{c} \times \mathbf{d} \rangle = \langle \mathbf{a}, \mathbf{b} \times (\mathbf{c} \times \mathbf{d}) \rangle = \langle \mathbf{a}, \langle \mathbf{d}, \mathbf{b} \rangle \mathbf{c} - \langle \mathbf{b}, \mathbf{c} \rangle \mathbf{d} \rangle = \langle \mathbf{a}, \mathbf{c} \rangle \langle \mathbf{b}, \mathbf{d} \rangle - \langle \mathbf{a}, \mathbf{d} \rangle \langle \mathbf{b}, \mathbf{c} \rangle$$

□

VI.1.2.3 (系)  $\langle \mathbf{a} \times \mathbf{b}, \mathbf{a} \times \mathbf{c} \rangle = \|\mathbf{a}\|^2 \langle \mathbf{b}, \mathbf{c} \rangle - \langle \mathbf{a}, \mathbf{b} \rangle \langle \mathbf{a}, \mathbf{c} \rangle$

*Proof.* 直前の公式から直ちに導ける。

□

VI.1.2.4 発散の別表現:  $\nabla \cdot A = \frac{1}{|V|} \lim_{|V| \rightarrow 0} \int_{\partial V} A(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) d^2 \mathbf{r}$

$\mathbf{r}_0 \in \mathbb{R}^3$  とし、 $\Omega \subset \mathbb{R}^3$  上のベクトル場  $A$  が  $\mathbf{r}_0$  の近傍で  $C^1$  級であるとする。 $\mathbf{r}_0$  を含む微小な領域を  $V$  とし、 $\partial V$  を  $V$  の境界、 $|V|$  を  $V$  の体積とする。但し  $\partial V$  は  $C^1$  級であるとする。このとき次式が成り立つ。

$$(\nabla \cdot A)(\mathbf{r}_0) = \frac{1}{|V|} \lim_{|V| \rightarrow 0} \int_{\partial V} A(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) d^2 \mathbf{r}$$

ここに  $\mathbf{n}(\mathbf{r})$  は  $\mathbf{r} \in \partial V$  に於ける外向きの法線ベクトルである。

*Proof.*

発散定理より次式が成り立つ。

$$\int_{\partial V} \mathbf{A}(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) d^2 \mathbf{r} = \int_V (\nabla \cdot \mathbf{A})(\mathbf{r}) d^3 \mathbf{r}$$

$\varepsilon > 0$  を任意に小さくとる。 $\mathbf{A}$  が  $\mathbf{r}_0$  の近傍で  $C^1$  級であるという仮定から、 $V$  を充分小さくとれば、 $|(\nabla \cdot \mathbf{A})(\mathbf{r}) - (\nabla \cdot \mathbf{A})(\mathbf{r}_0)| < \varepsilon$  である。よって次式が成り立つ。

$$\left| \frac{1}{|V|} \int_V (\nabla \cdot \mathbf{A})(\mathbf{r}) d^3 \mathbf{r} - (\nabla \cdot \mathbf{A})(\mathbf{r}_0) \right| < \varepsilon$$

$|V| \rightarrow 0$  とすれば定理の主張が成り立つ。 □

$$\text{VI.1.2.5} \quad \nabla_{\mathbf{r}} \cdot f(\mathbf{r}) \mathbf{A}(\mathbf{r}) = (\nabla_{\mathbf{r}} f(\mathbf{r})) \cdot \mathbf{A}(\mathbf{r}) + f(\mathbf{r}) \nabla_{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r})$$

$\mathbf{r} \in \mathbb{R}$  とする。 $f: \mathbb{R}^3 \rightarrow \mathbb{C}$ ,  $\mathbf{A}(\mathbf{r}): \mathbb{R}^3 \rightarrow \mathbb{R}^3$  はともに  $C^1$  級とする。このとき次式が成り立つ。

$$\nabla_{\mathbf{r}} \cdot f(\mathbf{r}) \mathbf{A}(\mathbf{r}) = (\nabla_{\mathbf{r}} f(\mathbf{r})) \cdot \mathbf{A}(\mathbf{r}) + f(\mathbf{r}) \nabla_{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r})$$

*Proof.*

$$\begin{aligned} \nabla_{\mathbf{r}} \cdot f(\mathbf{r}) \mathbf{A}(\mathbf{r}) &= \sum_{i=1}^3 \frac{\partial}{\partial r_i} f(\mathbf{r}) A_i(\mathbf{r}) = \sum_{i=1}^3 \left( \frac{\partial f(\mathbf{r})}{\partial r_i} A_i(\mathbf{r}) + f(\mathbf{r}) \frac{\partial A_i(\mathbf{r})}{\partial r_i} \right) \\ &= (\nabla_{\mathbf{r}} f(\mathbf{r})) \cdot \mathbf{A}(\mathbf{r}) + f(\mathbf{r}) \nabla_{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r}) \end{aligned}$$

□

$$\text{VI.1.2.6} \quad \nabla_{\mathbf{r}} \times f(\mathbf{r}) \mathbf{A}(\mathbf{r}) = (\nabla_{\mathbf{r}} f(\mathbf{r})) \times \mathbf{A}(\mathbf{r}) + f(\mathbf{r}) \nabla_{\mathbf{r}} \times \mathbf{A}(\mathbf{r})$$

$\mathbf{r} \in \mathbb{R}$  とする。 $f: \mathbb{R}^3 \rightarrow \mathbb{C}$ ,  $\mathbf{A}(\mathbf{r}): \mathbb{R}^3 \rightarrow \mathbb{R}^3$  はともに  $C^1$  級とする。このとき次式が成り立つ。

$$\nabla_{\mathbf{r}} \times f(\mathbf{r}) \mathbf{A}(\mathbf{r}) = (\nabla_{\mathbf{r}} f(\mathbf{r})) \times \mathbf{A}(\mathbf{r}) + f(\mathbf{r}) \nabla_{\mathbf{r}} \times \mathbf{A}(\mathbf{r})$$

*Proof.*

$i + j$  を 3 で割った余りを  $[i + j]$  と表すことにする。

$$\begin{aligned} \nabla_{\mathbf{r}} \times f(\mathbf{r}) \mathbf{A}(\mathbf{r}) &= \sum_{i=1}^3 \mathbf{i}_i \left( \frac{\partial f(\mathbf{r}) A_{[i+2]}(\mathbf{r})}{\partial r_{[i+1]}} - \frac{\partial f(\mathbf{r}) A_{[i+1]}(\mathbf{r})}{\partial r_{[i+2]}} \right) \\ &= \sum_{i=1}^3 \mathbf{i}_i \left[ \left( \frac{\partial f(\mathbf{r})}{\partial r_{[i+1]}} A_{[i+2]}(\mathbf{r}) - \frac{\partial f(\mathbf{r})}{\partial r_{[i+2]}} A_{[i+3]}(\mathbf{r}) \right) + f(\mathbf{r}) \left( \frac{\partial A_{[i+2]}(\mathbf{r})}{\partial r_{[i+1]}} - \frac{\partial A_{[i+1]}(\mathbf{r})}{\partial r_{[i+2]}} \right) \right] \\ &= (\nabla_{\mathbf{r}} f(\mathbf{r})) \times \mathbf{A}(\mathbf{r}) + f(\mathbf{r}) \nabla_{\mathbf{r}} \times \mathbf{A}(\mathbf{r}) \end{aligned}$$

□

$$\text{VI.1.2.7} \quad \nabla_{\mathbf{r}} \times (\mathbf{C} \times \mathbf{A}(\mathbf{r})) = (\nabla \cdot \mathbf{A}(\mathbf{r}))\mathbf{C} - J_{\mathbf{A}}\mathbf{C}$$

$\mathbf{r}, \mathbf{C} \in \mathbb{R}$  とする。  $A: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  は  $C^1$  級とする。このとき次式が成り立つ。

$$\nabla_{\mathbf{r}} \times (\mathbf{C} \times \mathbf{A}(\mathbf{r})) = (\nabla \cdot \mathbf{A}(\mathbf{r}))\mathbf{C} - J_{\mathbf{A}}\mathbf{C}$$

ここに  $J_{\mathbf{A}}$  は  $A$  の Jacobi 行列である。

*Proof.*

$$\begin{aligned} & \nabla_{\mathbf{r}} \times (\mathbf{C} \times \mathbf{A}(\mathbf{r})) \\ &= \mathbf{i}_1 \left[ C_1 \frac{\partial}{\partial r_2} A_2 - C_2 \frac{\partial}{\partial r_2} A_1 - C_3 \frac{\partial}{\partial r_3} A_1 + C_1 \frac{\partial}{\partial r_3} A_3 \right] \\ & \quad + \mathbf{i}_2 \left[ C_2 \frac{\partial}{\partial r_3} A_3 - C_3 \frac{\partial}{\partial r_3} A_2 - C_1 \frac{\partial}{\partial r_1} A_2 + C_2 \frac{\partial}{\partial r_1} A_1 \right] \\ & \quad + \mathbf{i}_3 \left[ C_3 \frac{\partial}{\partial r_1} A_1 - C_1 \frac{\partial}{\partial r_1} A_3 - C_2 \frac{\partial}{\partial r_2} A_3 + C_3 \frac{\partial}{\partial r_2} A_2 \right] \end{aligned}$$

$\mathbf{i}_1$  の係数を変形して次式を得る。

$$C_1 \left( \frac{\partial}{\partial r_1} A_1 + \frac{\partial}{\partial r_2} A_2 + \frac{\partial}{\partial r_3} A_3 \right) - C_1 \frac{\partial}{\partial r_1} A_1 - C_2 \frac{\partial}{\partial r_2} A_1 - C_3 \frac{\partial}{\partial r_3} A_1 = C_1 \nabla \cdot \mathbf{A} - \mathbf{C} \cdot \nabla A_1$$

$\mathbf{i}_2, \mathbf{i}_3$  についても同様にして、結局次式を得る。

$$\nabla_{\mathbf{r}} \times (\mathbf{C} \times \mathbf{A}(\mathbf{r})) = (\nabla \cdot \mathbf{A})\mathbf{C} - (\mathbf{C} \cdot \nabla A_1)\mathbf{i}_1 - (\mathbf{C} \cdot \nabla A_2)\mathbf{i}_2 - (\mathbf{C} \cdot \nabla A_3)\mathbf{i}_3 = (\nabla \cdot \mathbf{A}(\mathbf{r}))\mathbf{C} - J_{\mathbf{A}}\mathbf{C}$$

□

$$\text{VI.1.2.8} \quad 3 \text{ 次元 Euclid 空間に於ける Vector Laplacian: } \nabla^2 A = \mathbf{i}_1 \Delta A_1 + \mathbf{i}_2 \Delta A_2 + \mathbf{i}_3 \Delta A_3$$

VI.1.1.1 に於いて、特に 3 次元 Euclid 空間を考える場合は次式のように簡約化される。

$$\nabla^2 A = \mathbf{i}_1 \Delta A_1 + \mathbf{i}_2 \Delta A_2 + \mathbf{i}_3 \Delta A_3$$

*Proof.*

$$\begin{aligned} \nabla(\nabla \cdot A) &= \mathbf{i}_1 \frac{\partial}{\partial r_1} (\nabla \cdot A) + \mathbf{i}_2 \frac{\partial}{\partial r_2} (\nabla \cdot A) + \mathbf{i}_3 \frac{\partial}{\partial r_3} (\nabla \cdot A) \\ &= \mathbf{i}_1 \left( \frac{\partial^2}{\partial r_1^2} A_1 + \frac{\partial^2}{\partial r_1 \partial r_2} A_2 + \frac{\partial^2}{\partial r_1 \partial r_3} A_3 \right) \\ & \quad + \mathbf{i}_2 \left( \frac{\partial^2}{\partial r_2 \partial r_1} A_1 + \frac{\partial^2}{\partial r_2^2} A_2 + \frac{\partial^2}{\partial r_2 \partial r_3} A_3 \right) \\ & \quad + \mathbf{i}_3 \left( \frac{\partial^2}{\partial r_3 \partial r_1} A_1 + \frac{\partial^2}{\partial r_3 \partial r_2} A_2 + \frac{\partial^2}{\partial r_3^2} A_3 \right) \end{aligned}$$

$$\begin{aligned}
\nabla \times (\nabla \times A) &= \nabla \times \left[ \mathbf{i}_1 \left( \frac{\partial A_3}{\partial r_2} - \frac{\partial A_2}{\partial r_3} \right) + \mathbf{i}_2 \left( \frac{\partial A_1}{\partial r_3} - \frac{\partial A_3}{\partial r_1} \right) + \mathbf{i}_3 \left( \frac{\partial A_2}{\partial r_1} - \frac{\partial A_1}{\partial r_2} \right) \right] \\
&= \mathbf{i}_1 \left( -\frac{\partial^2}{\partial r_2^2} A_1 - \frac{\partial^2}{\partial r_3^2} A_1 + \frac{\partial^2}{\partial r_2 \partial r_1} A_2 + \frac{\partial^2}{\partial r_3 \partial r_2} A_3 \right) \\
&\quad + \mathbf{i}_2 \left( -\frac{\partial^2}{\partial r_3^2} A_2 - \frac{\partial^2}{\partial r_1^2} A_2 + \frac{\partial^2}{\partial r_3 \partial r_2} A_3 + \frac{\partial^2}{\partial r_1 \partial r_2} A_1 \right) \\
&\quad + \mathbf{i}_3 \left( -\frac{\partial^2}{\partial r_1^2} A_3 - \frac{\partial^2}{\partial r_2^2} A_3 + \frac{\partial^2}{\partial r_1 \partial r_3} A_1 + \frac{\partial^2}{\partial r_2 \partial r_3} A_2 \right)
\end{aligned}$$

$A_1, A_2, A_3$  が  $C^2$  級であるから偏微分の順序が交換可能であるので、

$$\nabla(\nabla \cdot A) - \nabla \times (\nabla \times A) = \mathbf{i}_1 \Delta A_1 + \mathbf{i}_2 \Delta A_2 + \mathbf{i}_3 \Delta A_3$$

□

#### VI.1.2.9 $\nabla \cdot (\nabla^2 \mathbf{A}) = \Delta(\nabla \cdot \mathbf{A})$

$\mathbf{r} \in \mathbb{R}$  とする。  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  は  $C^3$  級とする。次式が成り立つ。

$$\nabla \cdot (\nabla^2 \mathbf{A}) = \Delta(\nabla \cdot \mathbf{A})$$

*Proof.*

$$\begin{aligned}
\nabla \cdot \nabla^2 \mathbf{A} &= \nabla \cdot (\mathbf{i}_1 \Delta A_1 + \mathbf{i}_2 \Delta A_2 + \mathbf{i}_3 \Delta A_3) = \frac{\partial}{\partial r_1} \Delta A_1 + \frac{\partial}{\partial r_2} \Delta A_2 + \frac{\partial}{\partial r_3} \Delta A_3 \\
&= \Delta \left( \frac{\partial}{\partial r_1} A_1 + \frac{\partial}{\partial r_2} A_2 + \frac{\partial}{\partial r_3} A_3 \right) = \Delta(\nabla \cdot \mathbf{A})
\end{aligned}$$

□

#### VI.1.2.10 $\nabla (\langle \nabla f, \mathbf{v} \rangle) = (\nabla^2 f) \mathbf{v}$

$\mathbf{v}$  を定数ベクトルとすると

$$\nabla (\langle \nabla f, \mathbf{v} \rangle) = (\nabla^2 f) \mathbf{v}$$

*Proof.*

$$\frac{\partial}{\partial x_k} \langle \nabla f, \mathbf{v} \rangle = \frac{\partial}{\partial x_k} \sum_{i=1}^n \frac{\partial f}{\partial x_i} v_i = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_k \partial x_i} v_i$$

であるから

$$\nabla (\langle \nabla f, \mathbf{v} \rangle) = \begin{bmatrix} \frac{\partial}{\partial x_1} \langle \nabla f, \mathbf{v} \rangle \\ \frac{\partial}{\partial x_2} \langle \nabla f, \mathbf{v} \rangle \\ \vdots \\ \frac{\partial}{\partial x_n} \langle \nabla f, \mathbf{v} \rangle \end{bmatrix} = \begin{bmatrix} \left[ \frac{\partial^2 f}{\partial x_1^2}, \frac{\partial^2 f}{\partial x_1 \partial x_2}, \dots, \frac{\partial^2 f}{\partial x_1 \partial x_n} \right] \mathbf{v} \\ \vdots \\ \left[ \frac{\partial^2 f}{\partial x_n \partial x_1}, \frac{\partial^2 f}{\partial x_n \partial x_2}, \dots, \frac{\partial^2 f}{\partial x_n^2} \right] \mathbf{v} \end{bmatrix} = (\nabla^2 f) \mathbf{v}$$

□

$$\text{VI.1.2.11} \quad \nabla_{\mathbf{r}} \|\mathbf{r} + \mathbf{a}\|^n = n \|\mathbf{r} + \mathbf{a}\|^{n-2} (\mathbf{r} + \mathbf{a})$$

$\mathbf{r} := [r_1, r_2, r_3]^\top, \mathbf{a} := [a_1, a_2, a_3]^\top \in \mathbb{R}^3$  とする。次式が成り立つ。

$$\nabla_{\mathbf{r}} \|\mathbf{r} + \mathbf{a}\|^n = n \|\mathbf{r} + \mathbf{a}\|^{n-2} (\mathbf{r} + \mathbf{a})$$

*Proof.*

$$\begin{aligned} \nabla_{\mathbf{r}} \|\mathbf{r} + \mathbf{a}\|^n &= \sum_{i=1}^3 \mathbf{i}_i \frac{\partial}{\partial r_i} \left( \|\mathbf{r} + \mathbf{a}\|^2 \right)^{n/2} = \frac{n}{2} \sum_{i=1}^3 \mathbf{i}_i \left( \|\mathbf{r} + \mathbf{a}\|^2 \right)^{n/2-1} 2(r_i + a_i) \\ &= n \|\mathbf{r} + \mathbf{a}\|^{n-2} (\mathbf{r} + \mathbf{a}) \end{aligned}$$

□

## VI.1.3 Helmholtz の定理

### VI.1.3.1 補題: 全空間に渡る積分が存在するための十分条件

$\mathbf{r} \in \mathbb{R}^3$  とする。 $f: \mathbb{R}^3 \rightarrow \mathbb{C}$  は十分遠方に於いて、その絶対値がある  $\alpha > 0$  に対して  $o(\|\mathbf{r}\|_2^{-3-\alpha})$  であるとする。このとき  $\mathbb{R}^3$  上での  $f$  の積分は存在する。すなわち次式が成り立つ。

$$R > 0, V(R) := \{\mathbf{r} \in \mathbb{R}^3 \mid \|\mathbf{r}\|_2 \leq R\}, \lim_{R \rightarrow \infty} \left| \int_{V(R)} f(\mathbf{r}) d^3\mathbf{r} \right| < \infty$$

*Proof.*

$R > 0$  は十分大きいものとし、 $I(R)$  を次式で定義する。

$$I(R) := \int_{V(R)} f(\mathbf{r}) d^3\mathbf{r}$$

球座標変換  $\mathbf{r} = r[\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta]^\top$  を用いて次式を得る。

$$I(R) = \int_0^R \int_0^\pi \int_0^{2\pi} f(\mathbf{r}) r^2 \sin \theta d\phi d\theta dr$$

$0 < R_0 < R$  なる  $R_0$  をとり、 $I(R) = I(R_0) + I_2(R)$  と分解する。ここに  $I_2(R) := I(R) - I(R_0)$  である。

$$\begin{aligned} |I_2(R)| &\leq \int_{R_0}^R \int_0^\pi \int_0^{2\pi} |f(\mathbf{r})| r^2 \sin \theta d\phi d\theta dr = 4\pi \int_{R_0}^R o(\|\mathbf{r}\|_2^{-3-\alpha}) r^2 dr \\ &= 4\pi \int_{R_0}^R o(r^{-1-\alpha}) dr < \infty \text{ as } R \rightarrow \infty \end{aligned}$$

□

### VI.1.3.2 補題: 与えられた発散を有し、回転が 0 であるベクトル場の構成

Dirac のデルタ関数を使わずに証明したいが、未達成。今後の課題。



### VI.1.3.3 補題: 与えられた回転を有し、発散が 0 であるベクトル場の構成

Dirac のデルタ関数を使わずに証明したいが、未達成。今後の課題。

### VI.1.3.4 Helmholtz の定理: 任意の $C^1$ 級のベクトル場は回転が 0 である $C^1$ 級の場合と発散が 0 である $C^1$ 級の場合に分解できる

任意に与えられた、発散および回転が十分遠方で 0 に収束するような  $C^1$  級のベクトル場  $A$  が与えられたとき、補題 VI.1.3.2 により  $A$  と同じ発散を持ち、かつ回転が 0 であるようなベクトル場  $B$  を構成できる。このとき  $A - B$  は発散が 0 であるので、補題 VI.1.3.3 により、回転がこれに一致するようなベクトル場  $C$  を構成できる。これらを用いて  $A = B + C$  と分解すればよい。

## VI.1.4 諸定理

### VI.1.4.1 球殻定理

$R > 0$ ,  $V := \{\mathbf{r} \in \mathbb{R}^3 \mid \|\mathbf{r}\|_2 \leq R\}$  とする。ベクトル場  $\mathbf{A}(\mathbf{r})$  を次式で定義する。

$$\mathbf{A}(\mathbf{r}) := \int_V \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|_2^3} d^3 \mathbf{r}'$$

このとき次が成り立つ。

$$\mathbf{A}(\mathbf{r}) = \frac{4}{3}\pi \begin{cases} \mathbf{r} & (\|\mathbf{r}\|_2 \leq R) \\ \frac{R^3}{\|\mathbf{r}\|_2^3} \mathbf{r} & (\|\mathbf{r}\|_2 > R) \end{cases}$$

*Proof.*

必要なら系全体を回転させ、一般性を失わず  $\mathbf{r} = [0, 0, r]^\top$  ( $r > 0$ ) とする。球座標変換を用いて積分を実行する。 $\mathbf{r}' := r'[\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta]^\top$  とすると

$$\begin{aligned} \mathbf{A}(\mathbf{r}) &= \int_0^R \int_0^\pi \int_0^{2\pi} \frac{\mathbf{r} - \mathbf{r}'}{\|\mathbf{r} - \mathbf{r}'\|_2^3} r'^2 (\sin \theta) d\phi d\theta dr' \\ &= 2\pi \mathbf{i}_3 \int_0^R \int_0^\pi \frac{r - r' \cos \theta}{(r^2 + r'^2 - 2rr' \cos \theta)^{3/2}} r'^2 (\sin \theta) d\theta dr' \end{aligned}$$

ここに  $\mathbf{i}_3$  は第 3 軸方向の単位ベクトルである。

$f(r) := -(r^2 + r'^2 - 2rr' \cos \theta)^{-1/2} r'^2 \sin \theta$  とすると次式が成り立つ。

$$\frac{d}{dr} f(r) = \frac{(r - r' \cos \theta) r'^2 \sin \theta}{(r^2 + r'^2 - 2rr' \cos \theta)^{3/2}}$$

上式右辺の  $(r', \theta) = (0, 0)$  近傍での振る舞いを考える。 $r' = r + \Delta r$  ( $\Delta r \neq 0$ ) とすると、 $\sin \theta = O(\theta)$ ,  $\cos \theta = 1 + O(\theta^2)$  なので分母・分子ともに  $r^3 O(\theta^3)$  となる。よって積分領域に於いて  $f$  は  $r$  の  $C^1$  級関数であるから

積分と微分の順序交換が可能であり、次式が成り立つ。

$$\begin{aligned} \mathbf{A}(\mathbf{r}) &= 2\pi \mathbf{i}_3 \frac{d}{dr} \int_0^R \int_0^\pi f(r) d\theta dr' \\ &= -2\pi \mathbf{i}_3 \frac{d}{dr} \int_0^R \int_0^\pi \frac{r'^2 \sin \theta}{\sqrt{r^2 + r'^2 - 2rr' \cos \theta}} d\theta dr' \end{aligned}$$

$g(\theta) := \frac{r'}{r} \sqrt{r^2 + r'^2 - 2rr' \cos \theta}$  とすると次式が成り立つ。

$$\frac{dg}{d\theta}(\theta) = \frac{r'^2 \sin \theta}{\sqrt{r^2 + r'^2 - 2rr' \cos \theta}}$$

よって次式が成り立つ。

$$\mathbf{A}(\mathbf{r}) = -2\pi \mathbf{i}_3 \frac{d}{dr} \int_0^R [g(\theta)]_0^\pi dr' = -2\pi \mathbf{i}_3 \frac{d}{dr} \int_0^R \frac{r'}{r} (|r + r'| - |r - r'|) dr'$$

$R < r$  のときは次式が成り立つ。

$$\mathbf{A}(\mathbf{r}) = -2\pi \mathbf{i}_3 \frac{d}{dr} \int_0^R \frac{2r'^2}{r} dr' = \frac{4\pi R^3}{3r^2} \mathbf{i}_3 = \frac{4\pi R^3}{3} \frac{\mathbf{r}}{\|\mathbf{r}\|^3}$$

$R \geq r$  のときは  $r'$  に関する積分領域を  $[0, r]$  と  $[r, R]$  に分けて考える。前者については  $R < r$  のときの結果で  $R$  を  $r$  で置き換えたものとなる。後者については  $|r + r'| - |r - r'| = 2r$  となり、次式が成り立つ。

$$\mathbf{A}(\mathbf{r}) = -2\pi \mathbf{i}_3 \frac{d}{dr} \int_0^R 2r' dr' = \mathbf{0}$$

□

これを検証した Mathematica ノートブックを `parts/vectorCalculus/mathematica/shell_theorem.nb` に置いてある。

#### VI.1.4.2 非負領域の共通部分の非負領域

$\mathbb{R}^n$  上で考える。 $\mathbf{v}_1, \dots, \mathbf{v}_m$  ( $m \leq n$ ) は一次独立であるとし、 $\mathbf{v}$  の非負領域を

$$P(\mathbf{v}) := \{\mathbf{w} \mid \langle \mathbf{v}, \mathbf{w} \rangle \geq 0\}$$

で定義する。あるベクトル  $\mathbf{u}$  が  $\bigcap_{i=1}^m P(\mathbf{v}_i)$  の非負領域にある、すなわち

$$\langle \mathbf{u}, \mathbf{w} \rangle \geq 0, \forall \mathbf{w} \in \bigcap_{i=1}^m P(\mathbf{v}_i)$$

であるとき、 $\mathbf{u}$  は適当な定数  $a_i \geq 0$  ( $i = 1, \dots, m$ ) によって

$$\mathbf{u} = \sum_{i=1}^m a_i \mathbf{v}_i$$

と表せる。

*Proof.*

まず  $\mathbf{u}$  が  $\mathbf{v}_1, \dots, \mathbf{v}_m$  の一次結合で表せることを示す。

$W := \text{span}[\mathbf{v}_1, \dots, \mathbf{v}_m]$  とすると  $\mathbb{R}^n = W \oplus W^\perp$  であるから、適当な定数  $c_1, \dots, c_m$  及び適当なベクトル  $\mathbf{w}_t \in W^\perp$  を用いて

$$\mathbf{u} = \sum_{i=1}^m c_i \mathbf{v}_i + \mathbf{w}_t \quad (1)$$

と表せる。 $\mathbf{p}$  を  $\bigcap_{i=1}^m P(\mathbf{v}_i)$  の任意のベクトルとすると

$$\langle \mathbf{v}_i, \mathbf{p} \rangle \geq 0, \forall \mathbf{v}_i \in W$$

また、 $\mathbf{w}_{t2}$  を  $W^\perp$  の任意のベクトルとすると

$$\langle \mathbf{v}_i, \mathbf{w}_{t2} \rangle = 0, \forall \mathbf{v}_i \in W$$

従って  $\mathbf{p}' := \mathbf{p} + \mathbf{w}_{t2}$  とすると

$$\langle \mathbf{v}_i, \mathbf{p}' \rangle \geq 0, \forall \mathbf{v}_i \in W$$

であるから  $\mathbf{p}'$  もまた  $\bigcap_{i=1}^m P(\mathbf{v}_i)$  のベクトルである。従って  $\langle \mathbf{u}, \mathbf{p}' \rangle \geq 0$  であるから

$$0 \leq \langle \mathbf{u}, \mathbf{p}' \rangle = \left\langle \sum_{i=1}^m c_i \mathbf{v}_i, \mathbf{p} \right\rangle + \langle \mathbf{w}_{t1}, \mathbf{w}_{t2} \rangle$$

である。然るに  $\mathbf{w}_{t2} \in W^\perp$  は任意であるから、 $\mathbf{w}_{t1} = \mathbf{0}$  でない限り  $\langle \mathbf{w}_{t1}, \mathbf{w}_{t2} \rangle$  の値をいくらでも小さくできる。よって上式が成り立つためには  $\mathbf{w}_{t1} = \mathbf{0}$  でなくてはならならず、式 (1) より結局

$$\mathbf{u} = \sum_{i=1}^m c_i \mathbf{v}_i$$

と表せることになる。

次に  $c_i$  が全て非負であることを示す。 $\mathbf{v}_1, \dots, \mathbf{v}_m$  ( $m \leq n$ ) は一次独立であったから次の行列

$$A = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_m^T \end{bmatrix}$$

の階数は  $m$  である。従って次の方程式

$$A\mathbf{x} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i$$

は解を持つ。この解を  $\mathbf{d}_i$  とすれば  $\langle \mathbf{v}_j, \mathbf{d}_i \rangle = \delta_{ij}$  であるから  $\mathbf{d}_i \in \bigcap_{i=1}^m P(\mathbf{v}_i)$ 。よって

$$0 \leq \langle \mathbf{u}, \mathbf{d}_i \rangle = c_i$$

□

### VI.1.4.3 Stokes の定理

$c$  を空間上の任意の  $C^1$  級の閉曲線、 $S$  を、 $c$  を周  $\partial S$  とする任意の  $C^2$  級の曲面、 $A$  を  $C^1$  級のベクトル場とすると

$$\oint_c A \cdot d\mathbf{c} = \iint_S \nabla \times A \cdot d\mathbf{S}$$

*Proof.*

$S$  が  $S(x, y)$  の如く  $x, y$  の関数で表せる場合について証明すれば十分である。なぜなら、 $S$  が複雑な場合は  $S(s, y), S(y, z), S(z, x)$  いずれかの型の 2 変数関数で表せる曲面に分割して、各部分で、これから証明する  $S(x, y)$  型の場合と同じ手続きを踏んで最後に合計すれば全体の結果になるからである。領域分割を行う場合は分割境界でも線積分を行うことになるが、その回数は 2 で、互いに逆向きであるから、合計する段階で相殺して 0 になるため最終的な結果には現れない。

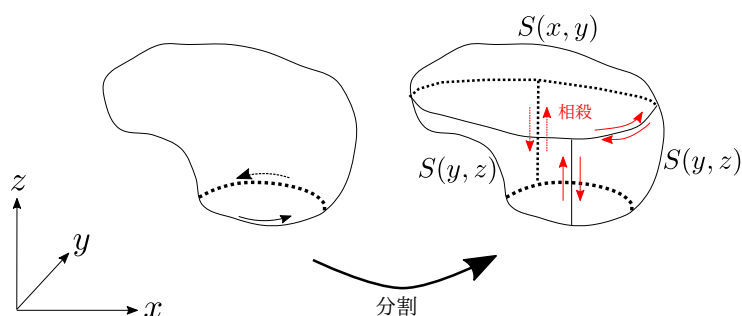


図 VI.1.1  $S$  の分割

以降  $S$  は  $S(x, y)$  型の曲面とする。 $\partial S$  を  $t$  をパラメータとする閉曲線とする。すなわち

$$\partial S : x = \phi(t), y = \psi(t), t : 0 \rightarrow 1$$

また、 $xy$  平面への  $S$  の射影を  $\Omega$  とする。

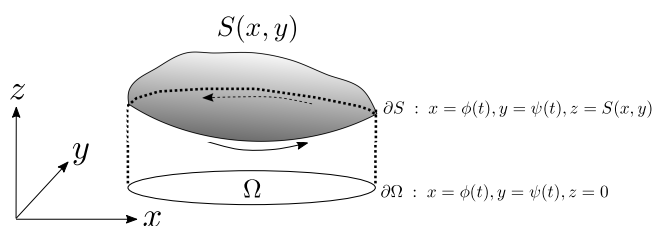


図 VI.1.2 状況

では証明しよう。線積分で表された定理の左辺を変形して面積分に置き換えて、定理の右辺と見比べる流れで行く。

関数の引数を全部書いてはスペースがいくらあっても足りないので省略して書くことになる (例えば  $A_x(x, y, z)$  を単に  $A_x$  と書いたり) のだが、こうすると独立変数と従属変数の関係を見誤る危険が生じるため

注意が必要だ。迂闊に計算すると罣にはまる。特に危ない所では敢えて引数を明示する。

$$\begin{aligned}
 \oint_c A \cdot d\mathbf{c} &= \oint_{\partial S} (\mathbf{i}_x A_x + \mathbf{i}_y A_y + \mathbf{i}_z A_z) \cdot (\mathbf{i}_x dx + \mathbf{i}_y dy + \mathbf{i}_z dz) \\
 &= \oint_{\partial S} (A_x dx + A_y dy) + \oint_{\partial S} A_z dz \\
 &= \underbrace{\oint_{\partial\Omega} (A_x(x, y, S(x, y)) dx + A_y(x, y, S(x, y)) dy)}_{\textcircled{A}} + \underbrace{\oint_{\partial S} A_z dz}_{\textcircled{B}}
 \end{aligned}$$

①で積分経路が  $\partial S$  から  $\partial\Omega$  に変更されたことに注意。Green の定理を使いたいから  $xy$  平面に落としたのだ。Green の定理より

$$\textcircled{A} = \iint_{\Omega} \left[ \underbrace{\frac{\partial}{\partial x} A_y(x, y, S(x, y))}_{\textcircled{C}} - \underbrace{\frac{\partial}{\partial y} A_x(x, y, S(x, y))}_{\textcircled{D}} \right] dx dy$$

③が罣だ。引数を明示していないと  $\frac{\partial A_y}{\partial x} \Big|_{(x, y, z)=(x, y, S(x, y))}$  と勘違いしてしまう。  $\frac{\partial}{\partial x} A_y(x, y, S(x, y))$  と  $\frac{\partial A_y}{\partial x} \Big|_{(x, y, z)=(x, y, S(x, y))}$  は別物であることに注意せねばならぬ。連鎖律に注意して

$$\begin{aligned}
 \textcircled{C} &= \frac{\partial A_y}{\partial x} \Big|_{(x, y, z)=(x, y, S(x, y))} = \frac{\partial}{\partial x} A_y(x, y, S(x, y)) \\
 &= \frac{\partial A_y}{\partial x} + \frac{\partial A_y}{\partial z} \frac{\partial S}{\partial x}
 \end{aligned}$$

同様に

$$\textcircled{D} = \frac{\partial A_x}{\partial y} + \frac{\partial A_x}{\partial z} \frac{\partial S}{\partial y}$$

よって結局

$$\begin{aligned}
 \textcircled{A} &= \iint_{\Omega} \left[ \frac{\partial A_y}{\partial x} + \frac{\partial A_y}{\partial z} \frac{\partial S}{\partial x} - \frac{\partial A_x}{\partial y} - \frac{\partial A_x}{\partial z} \frac{\partial S}{\partial y} \right] dx dy \\
 &= \iint_{\Omega} \left[ \left( \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) + \frac{\partial A_y}{\partial z} \frac{\partial S}{\partial x} - \frac{\partial A_x}{\partial z} \frac{\partial S}{\partial y} \right] dx dy
 \end{aligned}$$

次に②を見る。  $t$  による偏微分を  $\dot{\phantom{x}}$  で表すことにする。線積分の定義に従って

$$\begin{aligned}
 \textcircled{B} &= \int_0^1 A_z(\phi(t), \psi(t), S(\phi(t), \psi(t))) \left( \frac{\partial S}{\partial x} \dot{\phi}(t) + \frac{\partial S}{\partial y} \dot{\psi}(t) \right) dt \\
 &= \int_0^1 \left( A_z(\phi(t), \psi(t), S(\phi(t), \psi(t))) \frac{\partial S}{\partial x} \right) \dot{\phi}(t) dt + \int_0^1 \left( A_z(\phi(t), \psi(t), S(\phi(t), \psi(t))) \frac{\partial S}{\partial y} \right) \dot{\psi}(t) dt \\
 &= \oint_{\partial\Omega} A_z(x, y, S(x, y)) \frac{\partial S}{\partial x} dx + \oint_{\partial\Omega} A_z(x, y, S(x, y)) \frac{\partial S}{\partial y} dy \\
 &= \oint_{\partial\Omega} \left( A_z(x, y, S(x, y)) \frac{\partial S}{\partial x} dx + A_z(x, y, S(x, y)) \frac{\partial S}{\partial y} dy \right)
 \end{aligned}$$

これに Green の定理を適用して、連鎖律に注意して偏微分していく。

$$\begin{aligned}
\textcircled{B} &= \iint_{\Omega} \left[ \frac{\partial}{\partial x} \left( A_z(x, y, S(x, y)) \frac{\partial S}{\partial y} \right) - \frac{\partial}{\partial y} \left( A_z(x, y, S(x, y)) \frac{\partial S}{\partial x} \right) \right] dx dy \\
&= \iint_{\Omega} \left( \frac{\partial A_z}{\partial x} \frac{\partial S}{\partial y} + \frac{\partial A_z}{\partial z} \frac{\partial S}{\partial x} \frac{\partial S}{\partial y} + A_z \frac{\partial^2 S}{\partial y \partial x} - \frac{\partial A_z}{\partial y} \frac{\partial S}{\partial x} - \frac{\partial A_z}{\partial z} \frac{\partial S}{\partial y} \frac{\partial S}{\partial x} - A_z \frac{\partial^2 S}{\partial x \partial y} \right) dx dy \\
&= \iint_{\Omega} \left( \frac{\partial A_z}{\partial x} \frac{\partial S}{\partial y} - \frac{\partial A_z}{\partial y} \frac{\partial S}{\partial x} \right) dx dy \quad \left( \because S \text{ は } C^2 \text{ 級と仮定したから } \frac{\partial^2 S}{\partial y \partial x} = \frac{\partial^2 S}{\partial x \partial y} \right)
\end{aligned}$$

以上より

$$\oint_c A \cdot d\mathbf{c} = \iint_{\Omega} \left[ - \left( \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \frac{\partial S}{\partial x} - \left( \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \frac{\partial S}{\partial y} + \left( \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \right] dx dy$$

これが定理の右辺と等しいことを示せば良い。まず  $d\mathbf{S}$  の正体を暴こう。感覚的には、 $d\mathbf{S}$  とは、 $S$  上で  $x, y$  が微小量  $dx, dy$  だけ動いた時に舐める  $S$  の面積をノルムに持ち、向きはその微小面上の任意の代表点での上向きの法線ベクトルに一致するようなベクトルであった。所謂「面積ベクトル」である。

厳密には Riemann 和を取る段階で定義され、 $S$  上の点  $(x, y, S(x, y))$  における接平面のうち、その射影が  $xy$  平面上の矩形  $dx dy$  に収まる部分 (3 次元上の平行四辺形) の面積をノルムに持ち、向きは接平面の上向き法線ベクトルと一致するようなベクトルである。

$$d\mathbf{S} = \left( \mathbf{i}_x dx + \mathbf{i}_z \frac{\partial S}{\partial x} dx \right) \times \left( \mathbf{i}_y dy + \mathbf{i}_z \frac{\partial S}{\partial y} dy \right) = \left( -\mathbf{i}_x \frac{\partial S}{\partial x} - \mathbf{i}_y \frac{\partial S}{\partial y} + \mathbf{i}_z \right) dx dy$$

よって定理の右辺は

$$\begin{aligned}
\iint_S \nabla \times A \cdot d\mathbf{S} &= \iint_{\Omega} \left[ \mathbf{i}_x \left( \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) + \mathbf{i}_y \left( \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) + \mathbf{i}_z \left( \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \right] \cdot d\mathbf{S} \\
&= \iint_{\Omega} \left[ - \left( \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \frac{\partial S}{\partial x} - \left( \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \frac{\partial S}{\partial y} + \left( \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \right] dx dy \\
&= \oint_c A \cdot d\mathbf{c}
\end{aligned}$$

□

## 第 VI.2 章

# 一般の直交座標系と 3 次元 Euclid 空間の関係

### VI.2.1 座標変換

$T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  は可逆であり  $C^1$  級であるとする。 $T$  の Jacobi 行列の列は互いに直交するものとする。 $\mathbf{u} \in \mathbb{R}^3$ ,  $\mathbf{r} = T(\mathbf{u})$  とする。例えば球座標系では  $\mathbf{u} = (r, \theta, \varphi)$  であり、 $\mathbf{r} = \mathbf{i}_x x + \mathbf{i}_y y + \mathbf{i}_z z$ ,  $(x, y, z) = (r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta)$  である。次式で定める量を  $\mathbf{u}$  に対する  $\mathbf{r}$  の計量係数 (Lamé coefficient) と呼ぶ。

$$h_i := \left\| \frac{\partial \mathbf{r}}{\partial u_i} \right\|_2 \quad (i = 1, 2, 3)$$

$h_i$  は  $T$  の Jacobi 行列  $J$  の第  $i$  列のノルムであり、逆写像定理から  $J$  は可逆なので  $h_i \neq 0$  である。

次式で定める単位ベクトルを「 $u_i$  方向単位ベクトル」と呼ぶ。

$$\mathbf{j}_i := \frac{1}{h_i} \frac{\partial \mathbf{r}}{\partial u_i}$$

$T$  の Jacobi 行列の列が互いに直交するため、 $\mathbf{j}_1, \mathbf{j}_2, \mathbf{j}_3$  は互いに直交する。

### VI.2.2 表記の慣習上の注意

$\mathbb{R}^n$  という記号は複数の文脈で使われる。ある時は単に  $\mathbb{R}$  の直積集合を表す記号として、またある時は  $n$  次元 Euclid 空間を表す記号として使われる。例えば球座標系の  $(r, \theta, \varphi)$  が属する構造は 3 次元 Euclid 空間ではなく単に  $\mathbb{R}$  の 3 つの直積集合であり、その  $T$  による像  $(x, y, z)$  もまた  $\mathbb{R}$  の 3 つの直積集合上の量である。そして  $(x, y, z)$  と 3 次元 Euclid 空間上のベクトル  $\mathbf{i}_x x + \mathbf{i}_y y + \mathbf{i}_z z$  は一対一に対応するから、しばしばこれらを同一視する。

### VI.2.3 単位ベクトルの変換

既出の記号の定義は引き継ぐ。3次元 Euclid 空間の標準単位ベクトルは直交座標系の単位ベクトルを用いて次のように表せる。

$$\mathbf{i}_i = \sum_{j=1}^3 \frac{1}{h_j} \frac{\partial r_i}{\partial u_j} \mathbf{j}_j \quad (i = 1, 2, 3)$$

つまり行列を用いて次のように表せる。

$$[\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3] = [\mathbf{j}_1, \mathbf{j}_2, \mathbf{j}_3] J H^{-1}$$

ここに  $J$  は  $T$  の Jacobi 行列であり、 $H := \text{diag}(h_1, h_2, h_3)$  である。

*Proof.*

$\mathbf{j}_1, \mathbf{j}_2, \mathbf{j}_3$  は  $\mathbb{R}^3$  の正規直交基底であるから次式が成り立つ。

$$\mathbf{i}_i = \sum_{j=1}^3 (\mathbf{i}_i \cdot \mathbf{j}_j) \mathbf{j}_j = \sum_{j=1}^3 \left( \mathbf{i}_i \cdot \frac{1}{h_j} \frac{\partial \mathbf{r}}{\partial u_j} \right) \mathbf{j}_j = \sum_{j=1}^3 \left( \mathbf{i}_i \cdot \frac{1}{h_j} \sum_{k=1}^3 \frac{\partial r_k}{\partial u_j} \mathbf{i}_k \right) \mathbf{j}_j = \sum_{j=1}^3 \frac{1}{h_j} \frac{\partial r_i}{\partial u_j} \mathbf{j}_j$$

□

### VI.2.4 ベクトルの成分の変換

既出の記号の定義は引き継ぐ。3次元 Euclid 空間上のベクトル  $\mathbf{A}(\mathbf{r}) = \mathbf{i}_1 A_1(\mathbf{r}) + \mathbf{i}_2 A_2(\mathbf{r}) + \mathbf{i}_3 A_3(\mathbf{r})$  を  $u_1, u_2, u_3$  直交座標系で表したものを  $\mathbf{j}_1 \tilde{A}_1(\mathbf{u}) + \mathbf{j}_2 \tilde{A}_2(\mathbf{u}) + \mathbf{j}_3 \tilde{A}_3(\mathbf{u})$  とするとき、 $A_1, A_2, A_3$  と  $\tilde{A}_1, \tilde{A}_2, \tilde{A}_3$  の間には次の関係式が成り立つ。

$$[A_1, A_2, A_3]^\top = J H^{-1} [\tilde{A}_1, \tilde{A}_2, \tilde{A}_3]^\top$$

ここに  $J$  は  $T$  の Jacobi 行列であり、 $H := \text{diag}(h_1, h_2, h_3)$  である。

*Proof.*

$$\begin{aligned} [\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3][A_1, A_2, A_3] &= \mathbf{A}(\mathbf{r}) = \mathbf{j}_1 \tilde{A}_1 + \mathbf{j}_2 \tilde{A}_2 + \mathbf{j}_3 \tilde{A}_3 = [\mathbf{j}_1, \mathbf{j}_2, \mathbf{j}_3][\tilde{A}_1, \tilde{A}_2, \tilde{A}_3]^\top \\ &= [\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3] J H^{-1} [\tilde{A}_1, \tilde{A}_2, \tilde{A}_3]^\top \end{aligned}$$

□

### VI.2.5 勾配

既出の記号の定義は引き継ぐ。 $\phi: \mathbb{R}^3 \rightarrow \mathbb{C}$  は  $\mathbb{R}^3$  上で  $C^1$  級であるとする。 $\tilde{\phi}(\mathbf{u}) := \phi(T(\mathbf{u}))$  とする。 $\phi$  の勾配  $\nabla_{\mathbf{r}} \phi(\mathbf{r})$  は  $\mathbf{u}$  を用いて次式で表せる。

$$\nabla_{\mathbf{r}} \phi(\mathbf{r}) = \sum_{i=1}^3 \frac{\mathbf{j}_i}{h_i} \frac{\partial \phi}{\partial u_i} = \left[ \frac{\mathbf{j}_1}{h_1}, \frac{\mathbf{j}_2}{h_2}, \frac{\mathbf{j}_3}{h_3} \right]^\top \nabla_{\mathbf{u}} \tilde{\phi}(\mathbf{u})$$



*Proof.*

$$\begin{aligned}\mathbf{j}_i &= \frac{1}{h_i} \frac{\partial \mathbf{r}}{\partial u_i} = \frac{1}{h_i} \frac{\partial}{\partial u_i} \sum_{j=1}^3 r_j \mathbf{i}_j = \frac{1}{h_i} [\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3] \left[ \frac{\partial r_1}{\partial u_i}, \frac{\partial r_2}{\partial u_i}, \frac{\partial r_3}{\partial u_i} \right]^\top \\ \therefore [\mathbf{j}_1, \mathbf{j}_2, \mathbf{j}_3] &= [\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3] \left[ \frac{1}{h_j} \frac{\partial r_i}{\partial u_j} \right]_{(i,j) \in \{1,2,3\}^2} = JH^{-1}\end{aligned}$$

ここに  $H := \text{diag}(h_1, h_2, h_3)$  であり、 $J$  は  $T$  の Jacobi 行列である。以上より次式が成り立つ。

$$[\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3] = [\mathbf{j}_1, \mathbf{j}_2, \mathbf{j}_3] H J^{-1}$$

$\frac{\partial \phi}{\partial r_i}$  を評価すると次式を得る。

$$\begin{aligned}\frac{\partial \phi}{\partial r_i} &= \sum_{j=1}^3 \frac{\partial \tilde{\phi}}{\partial u_j} \frac{\partial u_j}{\partial r_i} \\ \therefore \nabla_{\mathbf{r}} \phi(\mathbf{r}) &= \left[ \frac{\partial \phi}{\partial r_1}, \frac{\partial \phi}{\partial r_2}, \frac{\partial \phi}{\partial r_3} \right]^\top = \left[ \frac{\partial u_j}{\partial r_i} \right]_{(i,j) \in \{1,2,3\}^2} \left[ \frac{\partial \tilde{\phi}}{\partial u_1}, \frac{\partial \tilde{\phi}}{\partial u_2}, \frac{\partial \tilde{\phi}}{\partial u_3} \right]^\top \\ &= (J^{-1})^\top \nabla_{\mathbf{u}} \tilde{\phi}(\mathbf{u})\end{aligned}$$

最後の等号は  $\mathbf{u} = T^{-1}(\mathbf{r})$  と逆関数定理による。以上より、次式が成り立つ。

$$\nabla_{\mathbf{r}} \phi(\mathbf{r}) = [\mathbf{j}_1, \mathbf{j}_2, \mathbf{j}_3] H J^{-1} (J^{-1})^\top \nabla_{\mathbf{u}} \tilde{\phi}(\mathbf{u})$$

$J^{-1} (J^{-1})^\top = J^{-1} (J^\top)^{-1} = (J^\top J)^{-1}$  であり、 $T$  に関する仮定から  $J$  の列は互いに直交し、計量係数の定義から  $J^\top J$  の第  $i$  対角成分は  $h_i^2$  である。よって  $(J^\top J)^{-1} = H^{-2}$  であり、次式が成り立つ。

$$\nabla_{\mathbf{r}} \phi(\mathbf{r}) = [\mathbf{j}_1, \mathbf{j}_2, \mathbf{j}_3] H^{-1} \nabla_{\mathbf{u}} \tilde{\phi}(\mathbf{u}) = \left[ \frac{\mathbf{j}_1}{h_1}, \frac{\mathbf{j}_2}{h_2}, \frac{\mathbf{j}_3}{h_3} \right]^\top \nabla_{\mathbf{u}} \tilde{\phi}(\mathbf{u})$$

□

## VI.2.6 発散

既出の記号の定義は引き継ぐ。さらに  $T$  は  $C^2$  級であるとし、 $\frac{\partial h_i}{\partial u_i} = 0$  ( $i = 1, 2, 3$ ) が成り立つとする (例: 円柱座標系, 球座標系)。3 次元 Euclid 空間上の  $C^1$  級のベクトル場  $\mathbf{A}(\mathbf{r}) = \mathbf{i}_1 A_1(\mathbf{r}) + \mathbf{i}_2 A_2(\mathbf{r}) + \mathbf{i}_3 A_3(\mathbf{r})$  を  $u_1, u_2, u_3$  直交座標系で表したものを  $\mathbf{j}_1 \tilde{A}_1(\mathbf{u}) + \mathbf{j}_2 \tilde{A}_2(\mathbf{u}) + \mathbf{j}_3 \tilde{A}_3(\mathbf{u})$  とするとき、次式が成り立つ。

$$\nabla_{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r}) = \frac{1}{h_1 h_2 h_3} \left[ \frac{\partial h_2 h_3 \tilde{A}_1(\mathbf{u})}{\partial u_1} + \frac{\partial h_3 h_1 \tilde{A}_2(\mathbf{u})}{\partial u_2} + \frac{\partial h_1 h_2 \tilde{A}_3(\mathbf{u})}{\partial u_3} \right]$$

*Proof.*

$$\nabla_{\mathbf{r}} \cdot \mathbf{A}(\mathbf{r}) = \sum_{i=1}^3 \frac{\partial A_i(\mathbf{r})}{\partial r_i} = \sum_{i=1}^3 \sum_{j=1}^3 \frac{\partial A_i(\mathbf{r})}{\partial u_j} \frac{\partial u_j}{\partial r_i} = \sum_{i=1}^3 \sum_{j=1}^3 \left( \frac{\partial}{\partial u_j} \sum_{k=1}^3 \frac{1}{h_k} \frac{\partial r_i}{\partial u_k} \tilde{A}_k(\mathbf{u}) \right) \frac{\partial u_j}{\partial r_i} \quad (1)$$

最後の等号成立は VI.2.4 による。

まず  $\frac{\partial u_j}{\partial r_i}$  を評価する。 $J$  を  $T$  の Jacobi 行列とすると次式が成り立つ。

$$\frac{\partial u_j}{\partial r_i} = J^{-1}[j, i] = (H^{-2}J^\top)[j, i] = \frac{1}{h_j^2}J[i, j] = \frac{1}{h_j^2}\frac{\partial r_i}{\partial u_j}$$

式 (1) の  $\tilde{A}_k(\mathbf{u})$  に関する項に着目すると、以下のように計算される。但し以下の式で  $\delta_{i,j}$  はクロネッカーのデルタを表す。

$$\begin{aligned} & \sum_{i=1}^3 \sum_{j=1}^3 \left( \frac{\partial}{\partial u_j} \frac{1}{h_k} \frac{\partial r_i}{\partial u_k} \tilde{A}_k(\mathbf{u}) \right) \frac{1}{h_j^2} \frac{\partial r_i}{\partial u_j} = \sum_{j=1}^3 \frac{1}{h_j^2} \sum_{i=1}^3 \left( \frac{\partial}{\partial u_j} \frac{1}{h_k} \frac{\partial r_i}{\partial u_k} \tilde{A}_k(\mathbf{u}) \right) \frac{\partial r_i}{\partial u_j} \\ &= \sum_{j=1}^3 \frac{1}{h_j^2} \sum_{i=1}^3 \left[ \left( \frac{\partial}{\partial u_j} \frac{1}{h_k} \right) \tilde{A}_k(\mathbf{u}) \frac{\partial r_i}{\partial u_k} \frac{\partial r_i}{\partial u_j} + \frac{1}{h_k} \frac{\partial \tilde{A}_k(\mathbf{u})}{\partial u_j} \frac{\partial r_i}{\partial u_k} \frac{\partial r_i}{\partial u_j} + \frac{\tilde{A}_k(\mathbf{u})}{h_k} \frac{\partial^2 r_i}{\partial u_j \partial u_k} \frac{\partial r_i}{\partial u_j} \right] \\ &= \sum_{j=1}^3 \frac{1}{h_j^2} \left[ \left( \frac{\partial}{\partial u_j} \frac{1}{h_k} \right) \tilde{A}_k(\mathbf{u}) \frac{\partial \mathbf{r}}{\partial u_k} \cdot \frac{\partial \mathbf{r}}{\partial u_j} + \frac{1}{h_k} \frac{\partial \tilde{A}_k(\mathbf{u})}{\partial u_j} \frac{\partial \mathbf{r}}{\partial u_k} \cdot \frac{\partial \mathbf{r}}{\partial u_j} + \frac{\tilde{A}_k(\mathbf{u})}{h_k} \frac{\partial^2 \mathbf{r}}{\partial u_j \partial u_k} \cdot \frac{\partial \mathbf{r}}{\partial u_j} \right] \\ &= \sum_{j=1}^3 \frac{1}{h_j^2} \left[ \left( \frac{\partial}{\partial u_j} \frac{1}{h_k} \right) \tilde{A}_k(\mathbf{u}) h_k^2 \delta_{k,j} + \frac{1}{h_k} \frac{\partial \tilde{A}_k(\mathbf{u})}{\partial u_j} h_k^2 \delta_{k,j} + \frac{\tilde{A}_k(\mathbf{u})}{h_k} \frac{\partial^2 \mathbf{r}}{\partial u_j \partial u_k} \cdot \frac{\partial \mathbf{r}}{\partial u_j} \right] \\ &= \frac{1}{h_k^2} \left[ \left( \frac{\partial}{\partial u_k} \frac{1}{h_k} \right) \tilde{A}_k(\mathbf{u}) h_k^2 + h_k \frac{\partial \tilde{A}_k(\mathbf{u})}{\partial u_k} \right] + \frac{\tilde{A}_k(\mathbf{u})}{h_k} \sum_{j=1}^3 \frac{1}{h_j^2} \frac{\partial^2 \mathbf{r}}{\partial u_j \partial u_k} \cdot \frac{\partial \mathbf{r}}{\partial u_j} \\ &= \frac{1}{h_k} \frac{\partial \tilde{A}_k(\mathbf{u})}{\partial u_k} + \frac{\tilde{A}_k(\mathbf{u})}{h_k} \underbrace{\sum_{j=1}^3 \frac{1}{h_j^2} \frac{\partial^2 \mathbf{r}}{\partial u_j \partial u_k} \cdot \frac{\partial \mathbf{r}}{\partial u_j}}_{(2)} \quad \left( \because \text{仮定より } \frac{\partial}{\partial u_j} \frac{1}{h_k} = 0 \right) \end{aligned} \quad (3)$$

式 (2) を評価する。 $j = k$  のとき次式より 0 である。

$$\frac{\partial^2 \mathbf{r}}{\partial u_k^2} \cdot \frac{\partial \mathbf{r}}{\partial u_k} = \frac{1}{2} \frac{\partial}{\partial u_k} \left( \frac{\partial \mathbf{r}}{\partial u_k} \cdot \frac{\partial \mathbf{r}}{\partial u_k} \right) = \frac{1}{2} \frac{\partial h_k^2}{\partial u_k} = 0$$

$j \neq k$  のとき次式を得る。

$$\begin{aligned} & \frac{1}{h_j^2} \frac{\partial^2 \mathbf{r}}{\partial u_j \partial u_k} \cdot \frac{\partial \mathbf{r}}{\partial u_j} = \frac{1}{h_j^2} \frac{\partial^2 \mathbf{r}}{\partial u_k \partial u_j} \cdot \frac{\partial \mathbf{r}}{\partial u_j} \\ &= \frac{1}{h_j^2} \times \frac{1}{2} \frac{\partial}{\partial u_k} \left( \frac{\partial \mathbf{r}}{\partial u_j} \cdot \frac{\partial \mathbf{r}}{\partial u_j} \right) = \frac{1}{h_j^2} \times \frac{1}{2} \frac{\partial h_j^2}{\partial u_k} = \frac{1}{h_j} \frac{\partial h_j}{\partial u_k} \end{aligned}$$

これらを式 (3) に適用して次式を得る。但し以下で  $[i+j]$  は  $i+j$  を 3 で割った余りを表す。

$$(3) = \frac{1}{h_k} \frac{\partial \tilde{A}_k(\mathbf{u})}{\partial u_k} + \frac{\tilde{A}_k(\mathbf{u})}{h_k} \left[ \frac{1}{h_{[k+1]}} \frac{\partial h_{[k+1]}}{\partial u_k} + \frac{1}{h_{[k+2]}} \frac{\partial h_{[k+2]}}{\partial u_k} \right] = \frac{1}{h_1 h_2 h_3} \frac{\partial h_{[k+1]} h_{[k+2]} \tilde{A}_k(\mathbf{u})}{\partial u_k}$$

□

## VI.2.7 回転

既出の記号の定義は引き継ぐ。さらに  $T$  は  $C^2$  級であるとし、次式が成り立つとする (例: 円柱座標系, 球座標系)。

$$\frac{\partial^2 \mathbf{r}}{\partial u_j \partial u_i} \cdot \frac{\partial \mathbf{r}}{\partial u_k} = 0 \quad (\{i, j, k\} = \{1, 2, 3\}) \quad (1)$$

3 次元 Euclid 空間上の  $C^1$  級のベクトル場  $\mathbf{A}(\mathbf{r}) = \mathbf{i}_1 A_1(\mathbf{r}) + \mathbf{i}_2 A_2(\mathbf{r}) + \mathbf{i}_3 A_3(\mathbf{r})$  を  $u_1, u_2, u_3$  直交座標系で表したものを  $\mathbf{j}_1 \tilde{A}_1(\mathbf{u}) + \mathbf{j}_2 \tilde{A}_2(\mathbf{u}) + \mathbf{j}_3 \tilde{A}_3(\mathbf{u})$  とするとき、次式が成り立つ。

$$\begin{aligned} & \nabla_{\mathbf{r}} \times \mathbf{A}(\mathbf{r}) \\ &= \frac{\mathbf{j}_1}{h_2 h_3} \left[ \frac{\partial h_3 \tilde{A}_3}{\partial u_2} - \frac{\partial h_2 \tilde{A}_2}{\partial u_3} \right] + \frac{\mathbf{j}_2}{h_3 h_1} \left[ \frac{\partial h_1 \tilde{A}_1}{\partial u_3} - \frac{\partial h_3 \tilde{A}_3}{\partial u_1} \right] + \frac{\mathbf{j}_3}{h_1 h_2} \left[ \frac{\partial h_2 \tilde{A}_2}{\partial u_1} - \frac{\partial h_1 \tilde{A}_1}{\partial u_2} \right] \\ &= \sum_{i=1}^3 \frac{\mathbf{j}_i}{h_{[i+1]} h_{[i+2]}} \left[ \frac{\partial h_{[i+2]} \tilde{A}_{[i+2]}}{\partial u_{[i+1]}} - \frac{\partial h_{[i+1]} \tilde{A}_{[i+1]}}{\partial u_{[i+2]}} \right] \end{aligned}$$

ここに  $[i+j]$  は  $i+j$  を 3 で割った余りを表す。

*Proof.*

$$\begin{aligned} \nabla_{\mathbf{r}} \times \mathbf{A}(\mathbf{r}) &= \nabla_{\mathbf{r}} \times \sum_{i=1}^3 A_i(\mathbf{r}) \mathbf{i}_i = \sum_{i=1}^3 \nabla_{\mathbf{r}} \times A_i(\mathbf{r}) \mathbf{i}_i = \sum_{i=1}^3 (\nabla_{\mathbf{r}} A_i(\mathbf{r})) \times \mathbf{i}_i \quad (\because \text{VI.1.2.6}) \\ &= \sum_{i=1}^3 \sum_{j=1}^3 \frac{\mathbf{j}_j}{h_j} \frac{\partial A_i(\mathbf{r})}{\partial u_j} \times \mathbf{i}_i = \sum_{i=1}^3 \sum_{j=1}^3 \frac{\mathbf{j}_j}{h_j} \left( \frac{\partial}{\partial u_j} \sum_{k=1}^3 \frac{\tilde{A}_k(\mathbf{u})}{h_k} \frac{\partial r_i}{\partial u_k} \right) \times \sum_{l=1}^3 \frac{\mathbf{j}_l}{h_l} \frac{\partial r_i}{\partial u_l} \\ & \quad (\because \text{VI.2.5, VI.2.4, VI.2.3}) \end{aligned}$$

以下では表記の簡便さのため、 $\tilde{A}_k(\mathbf{u})$  を単に  $\tilde{A}_k$  と書く。上式の第  $m \in \{1, 2, 3\}$  成分を評価する。第  $m$  成分に寄与するのは  $(j, l) = ([m+1], [m+2]), ([m+2], [m+1])$  のときであり、計算すると次式を得る。

$$\begin{aligned} & \sum_{i=1}^3 \left[ \frac{1}{h_{[m+1]}} \left( \frac{\partial}{\partial u_{[m+1]}} \sum_{k=1}^3 \frac{\tilde{A}_k(\mathbf{u})}{h_k} \frac{\partial r_i}{\partial u_k} \right) \frac{1}{h_{[m+2]}} \frac{\partial r_i}{\partial u_{[m+2]}} \right. \\ & \quad \left. - \frac{1}{h_{[m+2]}} \left( \frac{\partial}{\partial u_{[m+2]}} \sum_{k=1}^3 \frac{\tilde{A}_k(\mathbf{u})}{h_k} \frac{\partial r_i}{\partial u_k} \right) \frac{1}{h_{[m+1]}} \frac{\partial r_i}{\partial u_{[m+1]}} \right] \\ &= \frac{1}{h_{[m+1]} h_{[m+2]}} \sum_{k=1}^3 \left[ \sum_{i=1}^3 \left( \frac{\partial}{\partial u_{[m+1]}} \frac{\tilde{A}_k}{h_k} \frac{\partial r_i}{\partial u_k} \right) \frac{\partial r_i}{\partial u_{[m+2]}} - \sum_{i=1}^3 \left( \frac{\partial}{\partial u_{[m+2]}} \frac{\tilde{A}_k}{h_k} \frac{\partial r_i}{\partial u_k} \right) \frac{\partial r_i}{\partial u_{[m+1]}} \right] \quad (2) \end{aligned}$$

この式の [ ] 内第 1 項を評価すると次式を得る。

$$\begin{aligned}
& \sum_{k=1}^3 \sum_{i=1}^3 \left( \frac{\partial}{\partial u_{[m+1]}} \frac{\tilde{A}_k}{h_k} \frac{\partial r_i}{\partial u_k} \right) \frac{\partial r_i}{\partial u_{[m+2]}} \\
&= \sum_{k=1}^3 \left[ \left( \frac{\partial}{\partial u_{[m+1]}} \frac{\tilde{A}_k}{h_k} \right) \sum_{i=1}^3 \frac{\partial r_i}{\partial u_k} \frac{\partial r_i}{\partial u_{[m+2]}} + \frac{\tilde{A}_k}{h_k} \sum_{i=1}^3 \frac{\partial^2 r_i}{\partial u_{[m+1]} \partial u_k} \frac{\partial r_i}{\partial u_{[m+2]}} \right] \\
&= \sum_{k=1}^3 \left[ \left( \frac{\partial}{\partial u_{[m+1]}} \frac{\tilde{A}_k}{h_k} \right) \frac{\partial \mathbf{r}}{\partial u_k} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+2]}} + \frac{\tilde{A}_k}{h_k} \frac{\partial^2 \mathbf{r}}{\partial u_{[m+1]} \partial u_k} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+2]}} \right] \\
&= \left( \frac{\partial}{\partial u_{[m+1]}} \frac{\tilde{A}_{[m+2]}}{h_{[m+2]}} \right) h_{[m+2]}^2 + \sum_{k \in \{[m+1], [m+2]\}} \frac{\tilde{A}_k}{h_k} \frac{\partial^2 \mathbf{r}}{\partial u_{[m+1]} \partial u_k} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+2]}}
\end{aligned}$$

最後の等号成立には式 (1) を用いた。式 (2) の [ ] 内第 2 項も同様に評価すると、式 (2) の  $1/(h_{[m+1]}h_{[m+2]})$  の右側の部分は次式になる。

$$\begin{aligned}
& \left( \frac{\partial}{\partial u_{[m+1]}} \frac{\tilde{A}_{[m+2]}}{h_{[m+2]}} \right) h_{[m+2]}^2 - \left( \frac{\partial}{\partial u_{[m+2]}} \frac{\tilde{A}_{[m+1]}}{h_{[m+1]}} \right) h_{[m+1]}^2 \\
&+ \sum_{k \in \{[m+1], [m+2]\}} \frac{\tilde{A}_k}{h_k} \left( \frac{\partial^2 \mathbf{r}}{\partial u_{[m+1]} \partial u_k} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+2]}} - \frac{\partial^2 \mathbf{r}}{\partial u_{[m+2]} \partial u_k} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+1]}} \right) \quad (3)
\end{aligned}$$

この式の第 3 項の ( ) 内を評価する。 $T$  は  $C^2$  級であると仮定しているから、微分順序の入れ替えが可能であることに留意する。 $k = [m+1]$  のときは次式になる。

$$\begin{aligned}
& \frac{\partial^2 \mathbf{r}}{\partial u_{[m+1]}^2} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+2]}} - \frac{\partial^2 \mathbf{r}}{\partial u_{[m+2]} \partial u_{[m+1]}} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+1]}} \\
&= \frac{\partial}{\partial u_{[m+1]}} \left( \frac{\partial \mathbf{r}}{\partial u_{[m+1]}} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+2]}} \right) - \frac{\partial \mathbf{r}}{\partial u_{[m+1]}} \cdot \frac{\partial^2 \mathbf{r}}{\partial u_{[m+1]} \partial u_{[m+2]}} - \frac{\partial^2 \mathbf{r}}{\partial u_{[m+2]} \partial u_{[m+1]}} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+1]}} \\
&= - \frac{\partial \mathbf{r}}{\partial u_{[m+1]}} \cdot \frac{\partial^2 \mathbf{r}}{\partial u_{[m+2]} \partial u_{[m+1]}} - \frac{\partial^2 \mathbf{r}}{\partial u_{[m+2]} \partial u_{[m+1]}} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+1]}} \\
&= - \frac{\partial}{\partial u_{[m+2]}} \left( \frac{\partial \mathbf{r}}{\partial u_{[m+1]}} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+1]}} \right) = - \frac{\partial h_{[m+1]}^2}{\partial u_{[m+2]}}
\end{aligned}$$

$k = [m+2]$  のときは次式になる。

$$\begin{aligned}
& \frac{\partial^2 \mathbf{r}}{\partial u_{[m+1]} \partial u_{[m+2]}} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+2]}} - \frac{\partial^2 \mathbf{r}}{\partial u_{[m+2]}^2} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+1]}} \\
&= \frac{\partial^2 \mathbf{r}}{\partial u_{[m+1]} \partial u_{[m+2]}} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+2]}} - \frac{\partial}{\partial u_{[m+2]}} \left( \frac{\partial \mathbf{r}}{\partial u_{[m+2]}} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+1]}} \right) + \frac{\partial \mathbf{r}}{\partial u_{[m+2]}} \cdot \frac{\partial^2 \mathbf{r}}{\partial u_{[m+2]} \partial u_{[m+1]}} \\
&= \frac{\partial^2 \mathbf{r}}{\partial u_{[m+1]} \partial u_{[m+2]}} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+2]}} + \frac{\partial \mathbf{r}}{\partial u_{[m+2]}} \cdot \frac{\partial^2 \mathbf{r}}{\partial u_{[m+1]} \partial u_{[m+2]}} \\
&= \frac{\partial}{\partial u_{[m+1]}} \left( \frac{\partial \mathbf{r}}{\partial u_{[m+2]}} \cdot \frac{\partial \mathbf{r}}{\partial u_{[m+2]}} \right) = \frac{\partial h_{[m+2]}^2}{\partial u_{[m+1]}}
\end{aligned}$$

これらの式を (3) に適用すると次式を得る。

$$\begin{aligned}
(3) &= \left( \frac{\partial}{\partial u_{[m+1]}} \frac{\tilde{A}_{[m+2]}}{h_{[m+2]}} \right) h_{[m+2]}^2 + \frac{\tilde{A}_{[m+2]}}{h_{[m+2]}} \frac{\partial h_{[m+2]}^2}{\partial u_{[m+1]}} - \left( \frac{\partial}{\partial u_{[m+2]}} \frac{\tilde{A}_{[m+1]}}{h_{[m+1]}} \right) h_{[m+1]}^2 - \frac{\tilde{A}_{[m+1]}}{h_{[m+1]}} \frac{\partial h_{[m+1]}^2}{\partial u_{[m+2]}} \\
&= \frac{\partial}{\partial u_{[m+1]}} \frac{\tilde{A}_{[m+2]}}{h_{[m+2]}} h_{[m+2]}^2 - \frac{\partial}{\partial u_{[m+2]}} \frac{\tilde{A}_{[m+1]}}{h_{[m+1]}} h_{[m+1]}^2 \\
&= \frac{\partial h_{[m+2]} \tilde{A}_{[m+2]}}{\partial u_{[m+1]}} - \frac{\partial h_{[m+1]} \tilde{A}_{[m+1]}}{\partial u_{[m+2]}}
\end{aligned}$$

これを式 (2) に適用して定理の主張を得る。 □

## VI.2.8 Laplacian

既出の記号の定義は引き継ぐ。さらに  $T$  は  $C^2$  級であるとし、 $\frac{\partial h_i}{\partial u_i} = 0$  ( $i = 1, 2, 3$ ) が成り立つとする (例: 円柱座標系, 球座標系)。 $\phi: \mathbb{R}^3 \rightarrow \mathbb{C}$  は  $\mathbb{R}^3$  上で  $C^2$  級であるとする。 $\tilde{\phi}(\mathbf{u}) := \phi(T(\mathbf{u}))$  とする。 $\nabla_{\mathbf{r}}^2 \phi(\mathbf{r})$  は  $\mathbf{u}$  を用いて次式で表せる。

$$\nabla_{\mathbf{r}}^2 \phi(\mathbf{r}) = \frac{1}{h_1 h_2 h_3} \left[ \frac{\partial}{\partial u_1} \left( \frac{h_2 h_3}{h_1} \frac{\partial \tilde{\phi}(\mathbf{u})}{\partial u_1} \right) + \frac{\partial}{\partial u_2} \left( \frac{h_3 h_1}{h_2} \frac{\partial \tilde{\phi}(\mathbf{u})}{\partial u_2} \right) + \frac{\partial}{\partial u_3} \left( \frac{h_1 h_2}{h_3} \frac{\partial \tilde{\phi}(\mathbf{u})}{\partial u_3} \right) \right]$$

*Proof.* VI.2.5, VI.2.6 を用いて容易に示せる。 □

## 第 VI.3 章

# 円柱座標系

### VI.3.1 計量係数

$$h_r = 1, h_\varphi = r, h_z = 1$$

*Proof.*

位置ベクトルは

$$\mathbf{r} = \begin{bmatrix} r \cos \varphi \\ r \sin \varphi \\ z \end{bmatrix}$$

よって

$$\begin{aligned} h_r &:= \left\| \frac{\partial \mathbf{r}}{\partial r} \right\| = \left\| \begin{bmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{bmatrix} \right\| = 1 \\ h_\varphi &:= \left\| \frac{\partial \mathbf{r}}{\partial \varphi} \right\| = \left\| \begin{bmatrix} -r \cos \varphi \\ r \sin \varphi \\ 0 \end{bmatrix} \right\| = r \\ h_z &:= \left\| \frac{\partial \mathbf{r}}{\partial z} \right\| = \left\| \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\| = 1 \end{aligned}$$

□

### VI.3.2 単位ベクトルに関する 3 次元 Euclid 空間との関係

$$\begin{aligned} \mathbf{i}_r &= \mathbf{i}_x \cos \varphi + \mathbf{i}_y \sin \varphi \\ \mathbf{i}_\varphi &= -\mathbf{i}_x \sin \varphi + \mathbf{i}_y \cos \varphi \end{aligned}$$

$$\begin{aligned} \mathbf{i}_x &= \mathbf{i}_r \cos \varphi - \mathbf{i}_\varphi \sin \varphi \\ \mathbf{i}_y &= \mathbf{i}_r \sin \varphi + \mathbf{i}_\varphi \cos \varphi \end{aligned}$$

*Proof.*

位置ベクトルは

$$\mathbf{r} = \begin{bmatrix} r \cos \varphi \\ r \sin \varphi \\ z \end{bmatrix}$$

よって

$$\mathbf{i}_r := \frac{1}{h_r} \frac{\partial \mathbf{r}}{\partial r} = \begin{bmatrix} \cos \varphi \\ \sin \varphi \\ 0 \end{bmatrix} = \mathbf{i}_x \cos \varphi + \mathbf{i}_y \sin \varphi \quad (1)$$

$$\mathbf{i}_\varphi := \frac{1}{h_\varphi} \frac{\partial \mathbf{r}}{\partial \varphi} = \frac{1}{r} \begin{bmatrix} -r \sin \varphi \\ r \cos \varphi \\ 0 \end{bmatrix} = -\mathbf{i}_x \sin \varphi + \mathbf{i}_y \cos \varphi \quad (2)$$

これを逆に解いてみる。

$$(1) \times \cos \varphi - (2) \times \sin \varphi \text{ より } \mathbf{i}_x = \mathbf{i}_r \cos \varphi - \mathbf{i}_\varphi \sin \varphi$$

$$(1) \times \sin \varphi + (2) \times \cos \varphi \text{ より } \mathbf{i}_y = \mathbf{i}_r \sin \varphi + \mathbf{i}_\varphi \cos \varphi$$

を得る。 □

### VI.3.3 微分演算に関する直交座標系との関係

$$\begin{aligned} \frac{\partial}{\partial x} &= \cos \varphi \frac{\partial}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial}{\partial \varphi} \\ \frac{\partial}{\partial y} &= \sin \varphi \frac{\partial}{\partial r} + \frac{\cos \varphi}{r} \frac{\partial}{\partial \varphi} \end{aligned}$$

*Proof.*

スカラー関数  $\phi(x, y, z)$  を微分することを考える。  $d\phi := \frac{\partial \phi}{\partial r} dr + \frac{\partial \phi}{\partial \varphi} d\varphi + \frac{\partial \phi}{\partial z} dz$  であるから

$$\begin{aligned} \frac{\partial \phi}{\partial x} &= \frac{\partial \phi}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial \phi}{\partial \varphi} \frac{\partial \varphi}{\partial x} + \frac{\partial \phi}{\partial z} \frac{\partial z}{\partial x} \\ &= \frac{\partial \phi}{\partial r} \frac{\partial}{\partial x} \sqrt{x^2 + y^2} + \frac{\partial \phi}{\partial \varphi} \frac{\partial}{\partial x} \varphi \end{aligned}$$

第1項に関して

$$\frac{\partial}{\partial x} \sqrt{x^2 + y^2} = \frac{x}{\sqrt{x^2 + y^2}} = \frac{r \cos \varphi}{r} = \cos \varphi$$

第2項に関して

$$\frac{\partial}{\partial x} \varphi = \frac{\partial}{\partial x} \text{Tan}^{-1}(x, y) = \frac{-y}{x^2 + y^2} = \frac{-r \sin \varphi}{r^2} = -\frac{\sin \varphi}{r}$$

以上より

$$\frac{\partial}{\partial x} = \cos \varphi \frac{\partial}{\partial r} - \frac{\sin \varphi}{r} \frac{\partial}{\partial \varphi}$$

$y$  に関しても全く同様にして導出できる。 □

## 第 VI.4 章

# 球座標系

### VI.4.1 計量係数

$$h_r = 1, h_\theta = r, h_\varphi = r \sin \theta$$

*Proof.*

位置ベクトルは

$$\mathbf{r} = \begin{bmatrix} r \sin \theta \cos \varphi \\ r \sin \theta \sin \varphi \\ r \cos \theta \end{bmatrix}$$

よって

$$\begin{aligned} h_r &:= \left\| \frac{\partial \mathbf{r}}{\partial r} \right\| = \left\| \begin{bmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{bmatrix} \right\| = 1 \\ h_\theta &:= \left\| \frac{\partial \mathbf{r}}{\partial \theta} \right\| = \left\| \begin{bmatrix} r \cos \theta \cos \varphi \\ r \cos \theta \sin \varphi \\ -r \sin \theta \end{bmatrix} \right\| = r \\ h_\varphi &:= \left\| \frac{\partial \mathbf{r}}{\partial \varphi} \right\| = \left\| \begin{bmatrix} -r \sin \theta \sin \varphi \\ r \sin \theta \cos \varphi \\ 0 \end{bmatrix} \right\| = r \sin \theta \end{aligned}$$

□

### VI.4.2 単位ベクトルに関する 3 次元 Euclid 空間との関係

$$\begin{aligned} \mathbf{i}_r &= \mathbf{i}_x \sin \theta \cos \varphi + \mathbf{i}_y \sin \theta \sin \varphi + \mathbf{i}_z \cos \theta \\ \mathbf{i}_\theta &= \mathbf{i}_x \cos \theta \cos \varphi + \mathbf{i}_y \cos \theta \sin \varphi - \mathbf{i}_z \sin \theta \\ \mathbf{i}_\varphi &= -\mathbf{i}_x \sin \varphi + \mathbf{i}_y \cos \varphi \end{aligned}$$

$$\begin{aligned} \mathbf{i}_x &= \mathbf{i}_r \sin \theta \cos \varphi + \mathbf{i}_\theta \cos \theta \cos \varphi - \mathbf{i}_\varphi \sin \varphi \\ \mathbf{i}_y &= \mathbf{i}_r \sin \theta \sin \varphi + \mathbf{i}_\theta \cos \theta \sin \varphi + \mathbf{i}_\varphi \cos \varphi \\ \mathbf{i}_z &= \mathbf{i}_r \cos \theta - \mathbf{i}_\theta \sin \theta \end{aligned}$$



(手間が大きい証明)

*Proof.*

位置ベクトルは

$$\mathbf{r} = \begin{bmatrix} r \sin \theta \cos \varphi \\ r \sin \theta \sin \varphi \\ r \cos \theta \end{bmatrix}$$

よって

$$\mathbf{i}_r := \frac{1}{h_r} \frac{\partial \mathbf{r}}{\partial r} = \begin{bmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{bmatrix} = \mathbf{i}_x \sin \theta \cos \varphi + \mathbf{i}_y \sin \theta \sin \varphi + \mathbf{i}_z \cos \theta \quad (1)$$

$$\mathbf{i}_\theta := \frac{1}{h_\theta} \frac{\partial \mathbf{r}}{\partial \theta} = \frac{1}{r} \begin{bmatrix} r \cos \theta \cos \varphi \\ r \cos \theta \sin \varphi \\ -r \sin \theta \end{bmatrix} = \mathbf{i}_x \cos \theta \cos \varphi + \mathbf{i}_y \cos \theta \sin \varphi - \mathbf{i}_z \sin \theta \quad (2)$$

$$\mathbf{i}_\varphi := \frac{1}{h_\varphi} \frac{\partial \mathbf{r}}{\partial \varphi} = \frac{1}{r \sin \theta} \begin{bmatrix} -r \sin \theta \sin \varphi \\ r \sin \theta \cos \varphi \\ 0 \end{bmatrix} = -\mathbf{i}_x \sin \varphi + \mathbf{i}_y \cos \varphi \quad (3)$$

これを逆に解いてみる。

$$(1) \times \cos \theta - (2) \times \sin \theta \text{ より } \mathbf{i}_z = \mathbf{i}_r \cos \theta - \mathbf{i}_\theta \sin \theta$$

$$(1) \times \sin \theta + (2) \times \cos \theta \text{ より}$$

$$\mathbf{i}_r \sin \theta + \mathbf{i}_\theta \cos \theta = \mathbf{i}_x \cos \varphi + \mathbf{i}_y \sin \varphi \quad (4)$$

$$(4) \times \cos \varphi - (3) \times \sin \varphi \text{ より } \mathbf{i}_x = \mathbf{i}_r \sin \theta \cos \varphi + \mathbf{i}_\theta \cos \theta \cos \varphi - \mathbf{i}_\varphi \sin \varphi$$

$$(3) \times \cos \varphi + (4) \times \sin \varphi \text{ より } \mathbf{i}_y = \mathbf{i}_r \sin \theta \sin \varphi + \mathbf{i}_\theta \cos \theta \sin \varphi + \mathbf{i}_\varphi \cos \varphi \quad \square$$

(手間が少ない証明)

*Proof.*

$\mathbf{i}_r, \mathbf{i}_\theta, \mathbf{i}_\varphi$  の導出までは上の証明と同じ。 $\mathbf{i}_r, \mathbf{i}_\theta, \mathbf{i}_\varphi$  が直交系を成すことを利用し、 $\mathbf{i}_x = (\mathbf{i}_x \cdot \mathbf{i}_r) \mathbf{i}_r + (\mathbf{i}_x \cdot \mathbf{i}_\theta) \mathbf{i}_\theta + (\mathbf{i}_x \cdot \mathbf{i}_\varphi) \mathbf{i}_\varphi$ ,  $\mathbf{i}_y = (\mathbf{i}_y \cdot \mathbf{i}_r) \mathbf{i}_r + (\mathbf{i}_y \cdot \mathbf{i}_\theta) \mathbf{i}_\theta + (\mathbf{i}_y \cdot \mathbf{i}_\varphi) \mathbf{i}_\varphi$ ,  $\mathbf{i}_z = (\mathbf{i}_z \cdot \mathbf{i}_r) \mathbf{i}_r + (\mathbf{i}_z \cdot \mathbf{i}_\theta) \mathbf{i}_\theta + (\mathbf{i}_z \cdot \mathbf{i}_\varphi) \mathbf{i}_\varphi$  として導出できる。  $\square$

## 第 VII 部

## 幾何学

## 第 VII.1 章

# Euclid 幾何学

### VII.1.1 諸公式

#### VII.1.1.1 ヘロンの公式

辺の長さが  $a, b, c$  なる三角形の面積  $S$  は

$$S = \sqrt{s(s-a)(s-b)(s-c)}$$

但し

$$s = \frac{a+b+c}{2}$$

*Proof.* 余弦定理を用いて

$$\cos C = \frac{a^2 + b^2 - c^2}{2ab}$$

$$\therefore \sin C = \sqrt{1 - \cos^2 C} = \frac{\sqrt{4a^2b^2 - (a^2 + b^2 - c^2)^2}}{2ab}$$

$$\begin{aligned} S &= \frac{1}{2}ab \sin C \\ &= \frac{1}{2}\sqrt{4a^2b^2 - (a^2 + b^2 - c^2)^2} \end{aligned}$$

ここで

$$\begin{aligned} &4a^2b^2 - (a^2 + b^2 - c^2)^2 \\ = & - (a^4 + b^4 + c^4 - 2a^2b^2 - 2b^2c^2 - 2c^2a^2) \\ = & - [a^4 - 2(b^2 + c^2)a^2 + b^4 + c^4 - 2b^2c^2] \\ = & - [a^4 - 2(b^2 + c^2)a^2 + (b^2 + c^2)^2 - 4b^2c^2] \\ = & - \left\{ [a^2 - (b^2 + c^2)]^2 - (2bc)^2 \right\} \\ = & - [a^2 - (b^2 + c^2) + 2bc] [a^2 - (b^2 + c^2) - 2bc] \\ = & - [a^2 - (b - c)^2] [a^2 - (b + c)^2] \\ = & (a + b + c)(-a + b + c)(a - b + c)(a + b - c) \end{aligned}$$

よって

$$S = \sqrt{\frac{a+b+c}{2} \times \frac{-a+b+c}{2} \times \frac{a-b+c}{2} \times \frac{a+b-c}{2}}$$

$$= \sqrt{s(s-a)(s-b)(s-c)}$$

但し

$$s = \frac{a+b+c}{2}$$

□

## VII.1.2 諸定理

$$\text{VII.1.2.1} \quad \lim_{\|\mathbf{x}\| \rightarrow \infty} \|\mathbf{x} - \mathbf{a}\| - \|\mathbf{x}\| = -\frac{\mathbf{x}}{\|\mathbf{x}\|} \cdot \mathbf{a}$$

$\mathbf{a}, \mathbf{x} \in \mathbb{R}^3$  とするとき次式が成り立つ。

$$\lim_{\|\mathbf{x}\| \rightarrow \infty} \|\mathbf{x} - \mathbf{a}\| - \|\mathbf{x}\| = -\frac{\mathbf{x}}{\|\mathbf{x}\|} \cdot \mathbf{a}$$

*Proof.*

$\mathbf{a} = \mathbf{0}$  のときは明らかに成り立つ。以下では  $\mathbf{a} \neq \mathbf{0}$  とする。また、 $\|\mathbf{x}\|$  を十分大きくとり、 $\|\mathbf{x}\| > \|\mathbf{a}\|/2$  とする。 $f(\mathbf{x}) := \|\mathbf{x} - \mathbf{a}\| - \|\mathbf{x}\|$  とすると次式が成り立つ。

$$(f(\mathbf{x}) + \|\mathbf{x}\|)^2 = \|\mathbf{x} - \mathbf{a}\|^2 \quad \therefore f(\mathbf{x})(f(\mathbf{x}) + 2\|\mathbf{x}\|) = \|\mathbf{a}\|^2 - 2\mathbf{x} \cdot \mathbf{a}$$

$f(\mathbf{x})$  が最小となるのは  $\mathbf{a}$  が、原点と  $\mathbf{x}$  を結ぶ線分の上にあるときで、その値は  $-\|\mathbf{a}\|$  である。 $\|\mathbf{x}\| > \|\mathbf{a}\|/2$  と仮定しているから次式を得る。

$$f(\mathbf{x}) = \frac{\|\mathbf{a}\|^2 - 2\mathbf{x} \cdot \mathbf{a}}{f(\mathbf{x}) + 2\|\mathbf{x}\|} < \frac{\|\mathbf{a}\|^2 - 2\mathbf{x} \cdot \mathbf{a}}{2\|\mathbf{x}\| - \|\mathbf{a}\|} \rightarrow -\frac{\mathbf{x}}{\|\mathbf{x}\|} \cdot \mathbf{a} \quad \text{as } \|\mathbf{x}\| \rightarrow \infty$$

□

## 第 VII.2 章

# 球面幾何学

### VII.2.1 余弦定理, 正弦定理について

Wikipedia 英語版の記事に簡潔な証明がある。3 次元直角座標系の中心に単位球を置き、三角形の 1 頂点を北極に固定し、1 つの辺を本初子午線に重ねてベクトルの内積を用いて余弦定理を導出している。その結果を用いて  $(\sin^2 A = 1 - \cos^2 A)$  正弦定理が証明される。

### VII.2.2 極三角形の極三角形が元の三角形になること

*Proof.*

$xyz$  直角座標系を考え、原点を  $O$  とする。球の中心を  $O$  に据える。 $A', B', C'$  はそれぞれ  $\overrightarrow{OB} \times \overrightarrow{OC}$ ,  $\overrightarrow{OC} \times \overrightarrow{OA}$ ,  $\overrightarrow{OA} \times \overrightarrow{OB}$  と球面との交点である。よって  $A, B, C$  がそれぞれ  $\overrightarrow{OB'} \times \overrightarrow{OC'}$ ,  $\overrightarrow{OC'} \times \overrightarrow{OA'}$ ,  $\overrightarrow{OA'} \times \overrightarrow{OB'}$  の上にあることを示せば良い。すなわち、次の等式を示せば良い。

$$\begin{cases} \left[ (\overrightarrow{OB} \times \overrightarrow{OC}) \times (\overrightarrow{OC} \times \overrightarrow{OA}) \right] \times \overrightarrow{OC} = \mathbf{0} \\ \left[ (\overrightarrow{OC} \times \overrightarrow{OA}) \times (\overrightarrow{OA} \times \overrightarrow{OB}) \right] \times \overrightarrow{OA} = \mathbf{0} \\ \left[ (\overrightarrow{OA} \times \overrightarrow{OB}) \times (\overrightarrow{OB} \times \overrightarrow{OC}) \right] \times \overrightarrow{OB} = \mathbf{0} \end{cases}$$

1 つ目の等式が成り立つことを示す。他の 2 つは同様に示せる。ベクトル三重積の性質  $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$  より

$$\begin{aligned} & [(\overrightarrow{OB} \times \overrightarrow{OC}) \times (\overrightarrow{OC} \times \overrightarrow{OA})] \times \overrightarrow{OC} \\ &= \left\{ [(\overrightarrow{OB} \times \overrightarrow{OC}) \cdot \overrightarrow{OA}] \overrightarrow{OC} - [(\overrightarrow{OB} \times \overrightarrow{OC}) \cdot \overrightarrow{OC}] \overrightarrow{OA} \right\} \times \overrightarrow{OC} = - \left\{ [(\overrightarrow{OB} \times \overrightarrow{OC}) \cdot \overrightarrow{OC}] \overrightarrow{OA} \right\} \times \overrightarrow{OC} \\ &= - \left\{ [(\overrightarrow{OC} \times \overrightarrow{OC}) \cdot \overrightarrow{OB}] \overrightarrow{OA} \right\} \times \overrightarrow{OC} \quad (\text{スカラー三重積の性質を用いた}) \\ &= [(\mathbf{0} \cdot \overrightarrow{OB}) \overrightarrow{OA}] \times \overrightarrow{OC} = \mathbf{0} \end{aligned}$$

□

### VII.2.3 元の三角形 $ABC$ の極三角形 $A'B'C'$ について

$A' = \pi - a, B' = \pi - b, C' = \pi - c$  となること  
( $a = BC, b = CA, c = AB$ )

*Proof.*

$xyz$  直角座標系を考え、原点を  $O$  とする。球の中心を  $O$  に据える。前述の定理より極三角形の極三角形は元の三角形だから、 $C, B$  はそれぞれ  $\overrightarrow{OA'} \times \overrightarrow{OB'}$ ,  $\overrightarrow{OC'} \times \overrightarrow{OA'}$  上にある。よって  $\triangle OA'B'$  と  $\triangle OC'A'$  の法線ベクトル同士の成す角が  $\overrightarrow{OB}$  と  $\overrightarrow{OC}$  の成す角すなわち  $a$  となる。 $A'$  は平面  $OA'B'$  と  $OC'A'$  の成す角 (2 つあり、その和は  $\pi$ ) のうち  $\triangle ABC$  の辺  $BC(=a)$  に向かう側であるので、 $A' = \pi - a$  となる。  $\square$

## 第 VIII 部

### 確率論

## 第 VIII.1 章

# 表記

- 特に断らない限り、問題は確率空間  $(\Omega, \mathcal{F}, P)$  で議論されているものと約束する。



## 第 VIII.2 章

# 独立性に関する諸定理

VIII.2.1  $A_1, \dots, A_n$  が独立ならば、そのうちどれか 1 つ  $A_l$  を  $A_l^c$  に置き換えたものも独立である

*Proof.*

$$\begin{aligned} P\left(A_l^c \cap \left(\bigcap_{i \neq l} A_i\right)\right) &= P\left(\bigcap_{i \neq l} A_i\right) - P\left(A_l \cap \left(\bigcap_{i \neq l} A_i\right)\right) = P\left(\bigcap_{i \neq l} A_i\right) - P\left(\bigcap_{i=1}^n A_i\right) \\ &= \prod_{i \neq l}^n P(A_i) - \prod_{i=1}^n P(A_i) \quad (\because \text{独立性の仮定}) \\ &= (1 - P(A_l)) \prod_{i \neq l}^n P(A_i) = P(A_l^c) \prod_{i \neq l}^n P(A_i) \end{aligned}$$

□

VIII.2.2  $A_1, \dots, A_n$  が独立ならば  $A_1^c, \dots, A_n^c$  も独立である

*Proof.* 直前の定理を繰り返し用いる。

□

*Proof.* (別証)

包除原理を用いる。

$$\begin{aligned} P\left(\bigcap_{i=1}^n A_i^c\right) &= 1 - P\left(\left(\bigcap_{i=1}^n A_i^c\right)^c\right) = 1 - P\left(\bigcup_{i=1}^n A_i\right) = 1 - \sum_{j=1}^n (-1)^{j-1} \sum_{I \in \{1, \dots, n\}^{C_j}} P\left(\bigcap_{i \in I} A_i\right) \\ &= 1 + \sum_{j=1}^n (-1)^j \sum_{I \in \{1, \dots, n\}^{C_j}} \prod_{i \in I} P(A_i) = 1 + \sum_{I \in \text{powset}(\{1, \dots, n\}) \setminus \emptyset} (-1)^{|I|} \prod_{i \in I} P(A_i) \\ &= \prod_{i=1}^n (1 - P(A_i)) = \prod_{i=1}^n P(A_i^c) \end{aligned}$$

□

VIII.2.3  $A_1, \dots, A_n$  が独立ならば、そのうちどれか 2 つ  $A_p, A_q$  を合併した  
ものも独立である。

*Proof.*

$$\begin{aligned}
& P \left( (A_p \cup A_q) \cup \left( \bigcap_{i \neq p, q}^n A_i \right) \right) \\
&= P \left( \left( \left( \bigcap_{i \neq p, q}^n A_i \right) \cap A_p \right) \cup \left( \left( \bigcap_{i \neq p, q}^n A_i \right) \cap A_q \right) \right) \\
&= P \left( \left( \left( \bigcap_{i \neq p, q}^n A_i \right) \cap A_p \right) \right) + P \left( \left( \left( \bigcap_{i \neq p, q}^n A_i \right) \cap A_q \right) \right) \\
&\quad - P \left( \left( \left( \bigcap_{i \neq p, q}^n A_i \right) \cap A_p \right) \cap \left( \left( \bigcap_{i \neq p, q}^n A_i \right) \cap A_q \right) \right) \\
&= P \left( \bigcap_{i \neq p, q}^n A_i \right) P(A_p) + P \left( \bigcap_{i \neq p, q}^n A_i \right) P(A_q) \\
&\quad - P \left( \left( \bigcap_{i \neq p, q}^n A_i \right) \cap (A_p \cap A_q) \right) \\
&= P \left( \bigcap_{i \neq p, q}^n A_i \right) P(A_p) + P \left( \bigcap_{i \neq p, q}^n A_i \right) P(A_q) \\
&\quad - P \left( \bigcap_{i \neq p, q}^n A_i \right) P(A_p) P(A_q) \\
&= (P(A_p) + P(A_q) - P(A_p)P(A_q)) P \left( \bigcap_{i \neq p, q}^n A_i \right) \\
&= P(A_p \cup A_q) P \left( \bigcap_{i \neq p, q}^n A_i \right)
\end{aligned}$$

□

VIII.2.4  $A_1, \dots, A_n$  が独立ならば、 $I_1 \sqcup I_2 = \{1 : n\}$  なる  $I_1, I_2$  に対して  
 $\bigcup_{i \in I_1} A_i$  と  $\bigcup_{j \in I_2} A_j$  とは独立である。

*Proof.* 直前の定理を繰り返し用いる。

□

## 第 VIII.3 章

# エントロピー

### VIII.3.1 離散型確率分布の場合

#### VIII.3.1.1 一様分布がエントロピーを最大化する

$M \in \mathbb{N}$  とし、確率変数  $X$  の取り得る値が  $\{x_1, x_2, \dots, x_M\}$  であるとする。 $X$  が  $x_i$  ( $i = 1, 2, \dots, M$ ) となる確率を  $p_i$  とする。この確率分布のエントロピーが最大となるのは、 $p_1 = p_2 = \dots = p_M = 1/M$  のとき、かつそのときに限る。

*Proof.*

Lagrange の未定乗数法を用いても示せるが、もっと初等的に示せる。 $\mathbf{P} := [p_1, \dots, p_M]^\top$  とする。 $X$  の確率分布のエントロピー  $H(\mathbf{P})$  について次式が成り立つ。

$$H(\mathbf{p}) = \sum_{m=1}^M p_m \log_2(1/p_m) \leq \log_2 \left( \sum_{m=1}^M p_m/p_m \right) = \log_2 M$$

不等号の導出には対数関数が上に凸であることと Jensen の不等式を用いた。等号成立の必要十分条件は  $p_1 = p_2 = \dots = p_M = 1/M$  である。  $\square$

### VIII.3.2 連続型確率分布の場合

#### VIII.3.2.1 一様分布がエントロピーを最大化する

$a, b \in \mathbb{R}$ ,  $a < b$  とする。 $p$  を  $[a, b]$  上の確率密度関数とする。 $p$  のエントロピーが最大となるのは  $p$  が一様分布の時かつその時に限る。

*Proof.*

変分法を用いる。 $[a, b]$  上の確率分布であって密度関数が存在するもの全体の集合を  $\mathcal{F}$  とする。 $P \in \mathcal{F}$  には固定端条件  $P(a) = 0, P(b) = 1$  が課される。 $\mathcal{F}$  上の汎関数  $I$  を次式で定義する。

$$I : P \in \mathcal{F} \mapsto \int_a^b f(P(x)) dx \quad \text{where} \quad f(x) = -x \log x$$

$f$  は上に凸であることを容易に示せる。よって III.26.2.1 より  $p$  が Euler 方程式の解であることと、 $p$  がエントロピーを最大化することは同値である。Euler 方程式を変形すると  $\log P'(x) + 1 = 0$  となり、 $P'(x) = \text{const.}$

である。これと条件  $P(a) = 0, P(b) = 1$  より  $p(x) = 1/(b-a)$  となる。

□

第 IX 部

記述統計

## 第 IX.1 章

# 諸定義

### IX.1.1 平均, 分散, 共分散

$D$  個の  $n$  次元ベクトルから成るデータ集合  $\text{DS} := \{\mathbf{x}_d \in \mathbb{R}^n\}_{d=1}^D$  に対して「平均  $\bar{\mathbf{x}}$ 」, 「分散  $\text{Var}[\text{DS}]$ 」を次のように定義する。

- $\bar{\mathbf{x}} := \frac{1}{D} \sum_{d=1}^D \mathbf{x}_d$
- $\text{Var}[\text{DS}] := \frac{1}{D} \sum_{d=1}^D \|\mathbf{x}_d - \bar{\mathbf{x}}\|^2 = \frac{1}{D} \sum_{d=1}^D (\mathbf{x}_d - \bar{\mathbf{x}})^\top (\mathbf{x}_d - \bar{\mathbf{x}})$

「第  $i$  次元成分の平均  $\bar{x}_i$ 」, 「第  $i$  次元成分と第  $j$  次元成分の共分散  $\text{Cov}[x_i, x_j]$ 」を次のように定義する。

- $\bar{x}_i := \bar{\mathbf{x}}[i]$
- $\text{Cov}[x_i, x_j] := \frac{1}{D} \sum_{d=1}^D (\mathbf{x}_d[i] - \bar{x}_i)(\mathbf{x}_d[j] - \bar{x}_j)$

共分散において  $i = j$  のときは  $\text{Cov}[x_i, x_j] = \frac{1}{D} \sum_{d=1}^D (\mathbf{x}_d[i] - \bar{x}_i)^2$  となり、特に「第  $i$  次元成分の分散  $\text{Var}[x_i]$ 」と呼ぶ。

### IX.1.2 分散共分散行列

上述のデータ集合に対して、 $\text{Cov}[x_i, x_j]$  を縦と横に並べてできる行列、すなわちその第  $(i, j)$  成分が  $\text{Cov}[x_i, x_j]$  である行列を「分散共分散行列  $\Sigma_{\text{DS}}$ 」と定義する。この行列は次式で得られる。

$$\Sigma_{\text{DS}} = \frac{1}{D} \sum_{d=1}^D (\mathbf{x}_d - \bar{\mathbf{x}})(\mathbf{x}_d - \bar{\mathbf{x}})^\top$$

## 第 IX.2 章

# 諸定理

### IX.2.1 分散共分散行列は半正定。特に分散が正なら正定。

$D$  個の  $n$  次元ベクトルから成るデータ集合  $\text{DS} := \{\mathbf{x}_d \in \mathbb{R}^n\}_{d=1}^D$  の分散共分散行列  $\Sigma_{\text{DS}}$  は半正定である。特に  $\text{Var}[\text{DS}] > 0$  ならば  $\Sigma_{\text{DS}}$  は正定である。

*Proof.*

任意の  $\mathbf{v} \in \mathbb{R}^n \setminus \mathbf{0}_n$  に対して

$$\mathbf{v}^\top \Sigma_{\text{DS}} \mathbf{v} = \frac{1}{D} \sum_{d=1}^D \mathbf{v}^\top (\mathbf{x}_d - \bar{\mathbf{x}})(\mathbf{x}_d - \bar{\mathbf{x}})^\top \mathbf{v} = \frac{1}{D} \sum_{d=1}^D \|\mathbf{v}^\top (\mathbf{x}_d - \bar{\mathbf{x}})\|^2 \geq 0$$

であるから少なくとも半正定である。さらに

$$\begin{aligned} \text{等号成立} &\iff \|\mathbf{v}^\top (\mathbf{x}_d - \bar{\mathbf{x}})\|^2 = 0, \forall d \in \{1, \dots, D\}, \forall \mathbf{v} \in \mathbb{R}^n \setminus \mathbf{0}_n \iff \mathbf{x}_d = \bar{\mathbf{x}}, \forall d \in 1, \dots, D \\ &\iff \text{Var}[\text{DS}] = 0 \end{aligned}$$

□

## 第 IX.3 章

# 主成分分析

### IX.3.1 分散共分散行列の対角化

$D$  個の  $n$  次元ベクトルから成るデータ集合  $\text{DS} := \{\mathbf{x}_d \in \mathbb{R}^n\}_{d=1}^D$  を考える。IX.2.1 より、分散共分散行列  $\Sigma_{\text{DS}}$  は半正定であり、特に  $\text{Var}[\text{DS}] > 0$  であれば正定である。従って、 $\Sigma_{\text{DS}}$  は直交行列で対角化でき、対角行列には固有値が並ぶ。しかも固有値は全て実数である。

簡単のため固有値に重複はないものとする、全ての固有空間の次元は 1 である。固有値を大きい順に  $\lambda_1 > \dots > \lambda_n \geq 0$  とし、固有値  $\lambda_i$  に対応する大きさ 1 の固有ベクトルを  $\mathbf{p}_i$  とする。直交行列  $P := [\mathbf{p}_1, \dots, \mathbf{p}_n]$  と対角行列  $\Lambda := \text{diag}(\lambda_1, \dots, \lambda_n)$  を用いて  $\Sigma_{\text{DS}}$  は次のように対角化される。

$$\Sigma_{\text{DS}} = P\Lambda P^\top$$

### IX.3.2 データセットの射影

任意の  $\mathbf{x}_d \in \text{DS}$  の、単位ベクトル  $\mathbf{u} \in \mathbb{R}^n$  方向成分  $y_d := \mathbf{u}^\top \mathbf{x}_d$  から成るデータ集合  $\text{DS}_y := \{y_d\}_{d=1}^D$  の分散  $\text{Var}[\text{DS}_y]$  は  $\mathbf{u}^\top \Sigma_{\text{DS}} \mathbf{u}$  である。

*Proof.*

$\text{DS}_y$  の平均  $\bar{y}$  は

$$\bar{y} = \frac{1}{D} \sum_{d=1}^D y_d = \frac{1}{D} \sum_{d=1}^D \mathbf{u}^\top \mathbf{x}_d = \mathbf{u}^\top \frac{1}{D} \sum_{d=1}^D \mathbf{x} = \mathbf{u}^\top \bar{\mathbf{x}}$$

これを用いて

$$\begin{aligned} \text{Var}[\text{DS}_y] &= \frac{1}{D} \sum_{d=1}^D (y_d - \bar{y})^2 = \frac{1}{D} \sum_{d=1}^D (\mathbf{u}^\top \mathbf{x}_d - \mathbf{u}^\top \bar{\mathbf{x}})^2 = \frac{1}{D} \sum_{d=1}^D [\mathbf{u}^\top (\mathbf{x}_d - \bar{\mathbf{x}})]^2 \\ &= \frac{1}{D} \sum_{d=1}^D \mathbf{u}^\top (\mathbf{x}_d - \bar{\mathbf{x}}) \mathbf{u}^\top (\mathbf{x}_d - \bar{\mathbf{x}}) = \frac{1}{D} \sum_{d=1}^D \mathbf{u}^\top (\mathbf{x}_d - \bar{\mathbf{x}}) (\mathbf{x}_d - \bar{\mathbf{x}})^\top \mathbf{u} \\ &= \mathbf{u}^\top \left( \frac{1}{D} \sum_{d=1}^D (\mathbf{x}_d - \bar{\mathbf{x}}) (\mathbf{x}_d - \bar{\mathbf{x}})^\top \right) \mathbf{u} = \mathbf{u}^\top \Sigma_{\text{DS}} \mathbf{u} \end{aligned}$$

□



### IX.3.3 主成分

任意の  $\mathbf{x}_d \in \text{DS}$  について、 $\mathbf{p}_i$  方向成分  $y_d := \mathbf{p}_i^\top \mathbf{x}_d$  を「 $\mathbf{x}_d$  の第  $i$  主成分」と呼ぶ。そのような  $y_d$  の集合  $\text{DS}_y := \{y_d\}_{d=1}^D = \{\mathbf{p}_i^\top \mathbf{x}_d\}_{d=1}^D$  を「DS の第  $i$  主成分」と呼ぶ。 $\mathbf{x}_d$  の  $P$  による直交変換  $\mathbf{y}_d := P^\top \mathbf{x}_d$  の集合  $\text{DS}_y := \{\mathbf{y}_d\}_{d=1}^D = \{P^\top \mathbf{x}_d\}_{d=1}^D$  を考えると、これは DS の  $P$  による直交変換である。

#### IX.3.3.1 主成分の分散共分散行列

上述の  $\text{DS}_y$  の分散共分散行列  $\Sigma_{\text{DS}_y}$  は先述の  $\Lambda$  に等しい。

*Proof.*

$\text{DS}_y$  の平均  $\bar{\mathbf{y}}$  は

$$\bar{\mathbf{y}} = \frac{1}{D} \sum_{d=1}^D \mathbf{y}_d = \frac{1}{D} \sum_{d=1}^D P^\top \mathbf{x}_d = P^\top \frac{1}{D} \sum_{d=1}^D \mathbf{x}_d = P^\top \bar{\mathbf{x}}$$

これを用いて

$$\begin{aligned} \Sigma_{\text{DS}_y} &= \frac{1}{D} \sum_{d=1}^D (\mathbf{y}_d - \bar{\mathbf{y}})(\mathbf{y}_d - \bar{\mathbf{y}})^\top = \frac{1}{D} \sum_{d=1}^D P^\top (\mathbf{x}_d - \bar{\mathbf{x}})(\mathbf{x}_d - \bar{\mathbf{x}})^\top P = P^\top \Sigma_{\text{DS}} P \\ &= P^\top P \Lambda P^\top P = \Lambda \end{aligned}$$

□

この定理から、第  $i$  主成分の分散が  $\lambda_i$  であり、第  $i$  主成分と第  $j \neq i$  主成分の共分散は 0 であることがわかる。つまり、どんなデータ集合でも上述の直交変換で回してやることで異なる次元方向の間の共分散を 0 にできる。

第 1 主成分に対応する固有ベクトル  $\mathbf{p}_1$  は、最も散らばりが大きい方向を指している。例えば DS の分布がラグビーボールのような形をしていれば、その長軸方向が第 1 主成分の方向である。

## 第 IX.4 章

# Fisher の線形判別分析

### IX.4.1 クラス間分散とクラス内分散の比を最大にする $w$ の導出

クラス内共分散行列とクラス間共分散行列をそれぞれ  $S_W$ ,  $S_B := (\bar{x}^1 - \bar{x}^2)(\bar{x}^1 - \bar{x}^2)^\top$  とする。ここに  $\bar{x}^1, \bar{x}^2$  はクラス 1, 2 のサンプルデータの平均である。 $S_W$  の正則性と  $\bar{x}^1 \neq \bar{x}^2$  を仮定すると、 $S_W$  は正定値対称、 $S_B$  は半正定値対称で階数は 1 となる。

クラス間分散とクラス内分散の比

$$r(w) := \frac{w^\top S_B w}{w^\top S_W w}$$

は  $w$  の定数倍の差に影響されない。ある  $w^*$  が  $r$  を最大化するならば任意の  $\alpha \neq 0$  に対して  $\alpha w^*$  も  $r$  を最大化する。よって  $w^\top S_W w = 1$  という制約を加えても  $r$  の最大値は変わらないから、結局次の問題の解集合を含む最小の部分空間から  $\mathbf{0}$  を抜いたものが元の問題の解集合となる。

$$\begin{aligned} & \text{maximize} && f(w) := w^\top S_B w \\ & \text{subject to} && g(w) := w^\top S_W w - 1 = 0 \end{aligned}$$

実行可能領域は有界閉集合 ( $n$  次元の楕円) であり、目的関数は連続であるからこの最大化問題は解をもつ。 $S_W$  は正定値であるから実行可能領域に特異点は存在せず、Lagrange の未定乗数法が使える。制約条件に対する Lagrange 定数を  $\lambda$  とすると、最適解の満たすべき必要条件は次のようになる。

$$\begin{cases} (S_B - \lambda S_W)w = \mathbf{0} & (1) \\ w^\top S_W w = 1 & (2) \end{cases}$$

(2) より  $w \neq \mathbf{0}$  であるから、(1) の行列  $S_B - \lambda S_W$  が非正則になる必要がある。ここで  $\lambda = 0$  としてみると (1) より  $S_B w = \mathbf{0}$  つまり  $f(w) = 0$  となるが、 $S_B$  は階数 1 の半正定値行列なのでこのような  $w$  は明らかに最大化問題の解ではない。 $(S_B$  の固有値の一つが正で他の  $n - 1$  個が 0 なので  $f(w) > 0$  となる  $w$  が実行可能領域に必ず存在する。) よって  $\lambda \neq 0$  である。(1) の両辺に左から  $S_W^{-1}$  を掛けると上の連立方程式は次のようになる。

$$\begin{cases} (S_W^{-1} S_B - \lambda I_n)w = \mathbf{0} & (3) \\ w^\top S_W w = 1 & (4) \end{cases}$$

(3) より  $S_W^{-1} S_B w = \lambda w$  となるので  $\lambda$  は  $S_W^{-1} S_B$  の固有値であり、 $w$  は対応する固有ベクトルである。

突然だがここで

$$\begin{aligned}\mathbf{w}^* &= \frac{S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2)}{\sqrt{(S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2))^\top S_W (S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2))}} \quad (\neq \mathbf{0}) \\ \lambda^* &= (S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2))^\top S_W (S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2)) \quad (> 0)\end{aligned}$$

とおいてみると

$$\begin{aligned}S_W^{-1}S_B\mathbf{w}^* &= S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2)(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2)^\top \frac{S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2)}{\sqrt{(S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2))^\top S_W (S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2))}} \\ &= \sqrt{(S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2))^\top S_W (S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2))} S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2) \\ &= (S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2))^\top S_W (S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2)) \mathbf{w}^* \\ &= \lambda^* \mathbf{w}^*\end{aligned}$$

となり、 $\lambda^*, \mathbf{w}^*$  はそれぞれ  $S_W^{-1}S_B$  の固有値、固有ベクトルになっている。さらに  $\mathbf{w}^{*\top} S_W \mathbf{w}^* = 1$  となる。よって力任せに作った  $(\lambda^*, \mathbf{w}^*)$  は (3),(4) の解である。同様に  $(\lambda^*, -\mathbf{w}^*)$  も解である。この 2 つ以外に解が存在しないことを示す。 $S_W^{-1}S_B$  は階数 1 であるから、Jordan 分解すると次のようになる。

$$S_W^{-1}S_B = JDJ^{-1}, \quad D = \begin{bmatrix} d & 0 & \cdots & 0 \\ 0 & 0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}, \quad d \neq 0$$

ここに正則行列  $J$  は Jordan 標準形への適当な変換行列である。 $\mathbf{z} = J^{-1}\mathbf{w}$  すなわち  $\mathbf{w} = J\mathbf{z}$  なる変数変換を行うと次のようになる。

$$\begin{aligned}\begin{cases} (3) \\ (4) \end{cases} &\iff \begin{cases} (S_W^{-1}S_B - \lambda I_n)J\mathbf{z} &= \mathbf{0} \\ \mathbf{z}^\top J^\top S_W J\mathbf{z} &= 1 \end{cases} \iff \begin{cases} J^{-1}(S_W^{-1}S_B - \lambda I_n)J\mathbf{z} &= \mathbf{0} \\ \mathbf{z}^\top J^\top S_W J\mathbf{z} &= 1 \end{cases} \\ &\iff \begin{cases} (D - \lambda I_n)\mathbf{z} &= \mathbf{0} \\ \mathbf{z}^\top J^\top S_W J\mathbf{z} &= 1 \end{cases} \end{aligned} \tag{5}$$

$$\tag{6}$$

$\lambda \neq 0, d$  のとき  $D - \lambda I_n$  は正則なので連立方程式は解を持たない。 $\lambda = 0$  は先述の通り最大化問題の解ではないので除外する。 $\lambda = d$  のとき  $D - \lambda I_n$  の階数は  $n - 1$  となる。すなわち  $D$  の固有値  $d$  に対応する固有空間の次元は 1 であるから連立方程式の解は符号の違いによる 2 つしかない。よって先程作った  $\lambda^* > 0$  が実は  $d$  だったのであり、 $(\lambda^*, \mathbf{w}^*)$  と  $(\lambda^*, -\mathbf{w}^*)$  が連立方程式 (3),(4) の解の全てである。

Lagrange の未定乗数法に基づく連立方程式 (3),(4) は最大化問題の解の必要条件に過ぎないが、今回は解候補が 2 つしかなく、両方が同じ目的関数値を与えるのでこれが最適解である。

ここまでは考察のために  $\mathbf{w}^\top S_W \mathbf{w} = 1$  という制約を付けていた。しかし冒頭で述べたように元の最大化問題では  $\mathbf{w}$  の定数倍の差は関係ないので、上の議論の  $\mathbf{w}^*$  ように規格化の手間を掛ける必要はなく、 $\mathbf{w}$  のノルムに関心が無いときは  $\mathbf{w} = S_W^{-1}(\bar{\mathbf{x}}^1 - \bar{\mathbf{x}}^2)$  とするだけで構わない。

## 第 X 部

# 頻度論的統計

## 第 X.1 章

# 一般の分布

X.1.1  $x$  の分散共分散行列を  $V = \text{Var}[x]$  とすると  $\text{Var}[v^\top x] = v^\top V v$

*Proof.*

$\hat{x} = E[x]$  とすると

$$\begin{aligned}\text{Var}[v^\top x] &= E[(v^\top x - E[v^\top x])^2] = E[(v^\top x - v^\top \hat{x})^2] = E[(v^\top (x - \hat{x}))^2] \\ &= E[v^\top (x - \hat{x})(x - \hat{x})^\top v] = v^\top E[(x - \hat{x})(x - \hat{x})^\top] v = v^\top V v\end{aligned}$$

□

X.1.2 ベクトル確率変数の線形写像の像の期待値と分散共分散行列

$X_1, \dots, X_n \in \mathbb{R}$  を確率変数、 $A \in \mathbb{R}^{n \times m}$ 、 $\mathbf{X} := [X_1, \dots, X_n]^\top$ 、 $\Sigma_X := \text{Cov}[\mathbf{X}]$  とすると次が成り立つ。

$$E[A^\top \mathbf{X}] = A^\top E[\mathbf{X}], \quad \text{Cov}[A^\top \mathbf{X}] = A^\top \Sigma_X A$$

特に  $X_1, \dots, X_n$  が独立で同一の分散  $\sigma^2$  をもち、かつ  $A$  が  $n$  次直交行列であれば

$$\text{Cov}[A^\top \mathbf{X}] = \Sigma_X$$

*Proof.*

$E[A^\top \mathbf{X}] = A^\top E[\mathbf{X}]$  は期待値の線形性より明らか。  $\text{Cov}[A^\top \mathbf{X}] = A^\top \Sigma_X A$  を示す。 $\mathbf{Y} := A^\top \mathbf{X}$  とし、 $A$  の第  $i$  列ベクトルを  $\mathbf{a}_i$  とすると

$$\begin{aligned}\text{Cov}[A^\top \mathbf{X}]_{ij} &= \text{Cov}[\mathbf{Y}]_{ij} = \text{Cov}[Y_i, Y_j] = E[(Y_i - E[Y_i])(Y_j - E[Y_j])] \\ &= E[(\mathbf{a}_i^\top \mathbf{X} - E[\mathbf{a}_i^\top \mathbf{X}])(\mathbf{a}_j^\top \mathbf{X} - E[\mathbf{a}_j^\top \mathbf{X}])] \\ &= E[\mathbf{a}_i^\top (\mathbf{X} - E[\mathbf{X}]) \mathbf{a}_j^\top (\mathbf{X} - E[\mathbf{X}])] \\ &= E[\mathbf{a}_i^\top (\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])^\top \mathbf{a}_j] \\ &= \mathbf{a}_i^\top E[(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])^\top] \mathbf{a}_j = \mathbf{a}_i^\top \Sigma_X \mathbf{a}_j\end{aligned}$$

これより  $\text{Cov}[A^\top \mathbf{X}] = A^\top \Sigma_X A$ 。

特に  $X_1, \dots, X_n$  が独立で同一の分散  $\sigma^2$  をもち、かつ  $A$  が  $n$  次直交行列であれば

$$\text{Cov} [A^\top \mathbf{X}] = A^\top \Sigma_X A = A^\top \sigma^2 I_n A = \sigma^2 I_n = \Sigma_X$$

□

### X.1.3 スコア関数

$\theta$  をパラメータとする確率密度関数  $f(x; \theta)$  に従う i.i.d 確率変数  $X_1, \dots, X_n \in \Omega$  に対して「スコア関数」を次式で定義する。

$$V_n(\mathbf{X}, \theta) := \frac{\partial}{\partial \theta} \log f(\mathbf{X}; \theta) = \frac{\partial}{\partial \theta} \log \prod_{i=1}^n f(X_i; \theta) = \frac{\partial}{\partial \theta} \sum_{i=1}^n \log f(X_i; \theta) = \sum_{i=1}^n V(X_i, \theta)$$

#### X.1.3.1 $E_{\mathbf{X}} [V(\mathbf{X}, \theta)] = 0$

*Proof.*

$$\begin{aligned} E_{\mathbf{X}} [V_n(\mathbf{X}, \theta)] &= \int_{\Omega^n} V_n(\mathbf{x}, \theta) \prod_{i=1}^n f(x_i; \theta) d\mathbf{x} = \int_{\Omega^n} \frac{\partial}{\partial \theta} \prod_{i=1}^n f(x_i; \theta) d\mathbf{x} \\ &= \frac{\partial}{\partial \theta} \int_{\Omega^n} \prod_{i=1}^n f(x_i; \theta) d\mathbf{x} \quad (\text{微分と積分の順序交換可能性を仮定}) \\ &= \frac{\partial 1}{\partial \theta} = 0 \end{aligned}$$

□

### X.1.4 Cramér-Rao の不等式

$\theta$  をパラメータとする確率密度関数  $f(x; \theta)$  に従う i.i.d 確率変数  $X_1, \dots, X_n \in \Omega$  から計算した  $\theta$  の不偏推定量  $\hat{\theta}(\mathbf{X})$  について次式が成り立つ

$$\text{Var}_{\mathbf{X}} [\hat{\theta}(\mathbf{X})] \geq (n E_{\mathbf{X}} [V(\mathbf{X}, \theta)^2])^{-1} = (n \text{Var}_{\mathbf{X}} [V(\mathbf{X}, \theta)])^{-1}$$

*Proof.*

(参考: <https://to-kei.net/estimator/cramer-rao/>)

$$\begin{aligned}
\text{Var}_{\mathbf{X}} [\hat{\theta}(\mathbf{X})] \text{E}_{\mathbf{X}} [V_n(\mathbf{X}, \theta)^2] &= \int_{\Omega} (\hat{\theta}(\mathbf{x}) - \theta)^2 f(\mathbf{x}; \theta) d\mathbf{x} \int_{\Omega} V_n(\mathbf{x}, \theta)^2 f(\mathbf{x}; \theta) d\mathbf{x} \\
&= \int_{\Omega} \left( (\hat{\theta}(\mathbf{x}) - \theta) \sqrt{f(\mathbf{x}; \theta)} \right)^2 d\mathbf{x} \int_{\Omega} \left( V_n(\mathbf{x}, \theta) \sqrt{f(\mathbf{x}; \theta)} \right)^2 d\mathbf{x} \\
&\geq \int_{\Omega} (\hat{\theta}(\mathbf{x}) - \theta) V_n(\mathbf{x}, \theta) f(\mathbf{x}; \theta) d\mathbf{x} \quad (\text{Cauchy-Schwarz の不等式を用いた}) \\
&= \int_{\Omega} \hat{\theta}(\mathbf{x}) V_n(\mathbf{x}, \theta) f(\mathbf{x}, \theta) d\mathbf{x} \quad (\because \text{直前の定理より } \text{E}_{\mathbf{X}} [V_n(\mathbf{X}, \theta)] = 0) \\
&= \int_{\Omega} \hat{\theta}(\mathbf{x}) \left( \frac{\partial}{\partial \theta} \log f(\mathbf{x}; \theta) \right) f(\mathbf{x}; \theta) d\mathbf{x} = \int_{\Omega} \hat{\theta}(\mathbf{x}) \frac{\partial}{\partial \theta} f(\mathbf{x}; \theta) d\mathbf{x} \\
&= \frac{\partial}{\partial \theta} \int_{\Omega} \hat{\theta}(\mathbf{x}) f(\mathbf{x}; \theta) d\mathbf{x} \quad (\text{微分と積分の順序交換可能性を仮定した}) \\
&= \frac{\partial}{\partial \theta} \text{E}_{\mathbf{X}} [\hat{\theta}(\mathbf{X})] = \frac{\partial \theta}{\partial \theta} = 1
\end{aligned}$$

ここで

$$\begin{aligned}
\text{E}_{\mathbf{X}} [V_n(\mathbf{X}, \theta)^2] &= \text{E}_{\mathbf{X}} \left[ \left( \sum_{i=1}^n V(X_i, \theta) \right)^2 \right] = \sum_{i=1}^n \sum_{j=1}^n \text{E}_{\mathbf{X}} [V(X_i, \theta) V(X_j, \theta)] \\
&= \sum_{i=1}^n \text{E}_{\mathbf{X}} [V(X_i, \theta)^2] \quad (\because V(X_i, \theta), V(X_j, \theta), i \neq j \text{ は独立で } \text{E}_{\mathbf{X}} [V(X_i, \theta)] = 0) \\
&= \sum_{i=1}^n \text{E}_{\mathbf{X}_i} [V(X_i, \theta)^2] = n \text{E}_X [V(X, \theta)^2]
\end{aligned}$$

以上より定理の主張が成り立つ。 □

## X.1.5 条件付き確率

$$\text{X.1.5.1} \quad \Pr(A) = \sum_k \Pr(A, B_k) = \sum_k \Pr(A | B_k) \Pr(B_k)$$

$$\text{X.1.5.2} \quad \Pr(A, B | C) = \Pr(A | B, C) \Pr(B | C) \quad (\text{スライディング})$$

*Proof.*

$$\Pr(A, B | C) = \frac{\Pr(A, B, C)}{\Pr(C)} = \frac{\Pr(A | B, C) \Pr(B, C)}{\Pr(C)} = \Pr(A | B, C) \Pr(B | C)$$

□

$$\text{X.1.5.3} \quad \Pr(A | C) = \sum_k \Pr(A | B_k, C) \Pr(B_k | C)$$

*Proof.*

$$\begin{aligned}
\Pr(A | C) &= \frac{\Pr(A, C)}{\Pr(C)} = \frac{\sum_k \Pr(A, B_k, C)}{\Pr(C)} \quad (\because \text{X.1.5.1}) \\
&= \frac{\sum_k \Pr(A | B_k, C) \Pr(B_k, C)}{\Pr(C)} = \sum_k \Pr(A | B_k, C) \Pr(B_k | C)
\end{aligned}$$

□

## X.1.6 条件付き確率密度関数

連続型 2 変数  $X, Y$  に関して、条件付き確率密度関数を次のように定義する。

$$p_Y(y|X=x) := \frac{p(x,y)}{\int_{y_{\min}}^{y_{\max}} p(x,y)dy}$$

( $p(x,y)$  は  $X, Y$  の同時確率密度関数)

## X.1.7 条件付き期待値

2 変数  $X, Y$  に関して、条件付き期待値を次のように定義する。

$$\begin{aligned} (\text{離散}) : E_Y[Y|X=x] &:= \sum_y y \Pr(Y=y|X=x) = \sum_y y \frac{\Pr(Y=y, X=x)}{\Pr(X=x)} \\ (\text{連続}) : E_Y[Y|X=x] &:= \int_{y_{\min}}^{y_{\max}} yp_Y(y|X=x)dy \end{aligned}$$

### X.1.7.1 $E_{X,Y}[E_Y[Y|X]] = E_{X,Y}[Y]$

条件付き期待値について次の性質が成り立つ。

$$E_{X,Y}[E_Y[Y|X]] = E_{X,Y}[Y]$$

*Proof.*

(離散型)

$$\begin{aligned} E_{X,Y}[E_Y[Y|X]] &= \sum_x \sum_y E_Y[Y|X=x] \Pr(X=x, Y=y) \\ &= \sum_x E_Y[Y|X=x] \sum_y \Pr(X=x, Y=y) \quad (\because E_Y[Y|X=x] \text{ は } y \text{ に無関係}) \\ &= \sum_x \left( \sum_y \frac{\Pr(X=x, Y=y)}{\Pr(X=x)} \right) \Pr(X=x) \\ &= \sum_x \sum_y y \Pr(X=x, Y=y) = E_{X,Y}[Y] \end{aligned}$$

(連続型)

$X, Y$  の同時確率密度関数を  $p(x,y)$  とし、 $X$  の周辺確率密度関数を  $p_X(x)$  とする。

$$\begin{aligned} E_{X,Y}[E_Y[Y|X]] &= \int_{x_{\min}}^{x_{\max}} \int_{y_{\min}}^{y_{\max}} E_Y[Y|X=x] p(x,y) dy dx \\ &= \int_{x_{\min}}^{x_{\max}} E_Y[Y|X=x] \int_{y_{\min}}^{y_{\max}} p(x,y) dy dx \quad (\because E_Y[Y|X=x] \text{ は } y \text{ に無関係}) \\ &= \int_{x_{\min}}^{x_{\max}} \left( \int_{y_{\min}}^{y_{\max}} y \frac{p(x,y)}{p_X(x)} dy \right) p_X(x) dx = \int_{x_{\min}}^{x_{\max}} \int_{y_{\min}}^{y_{\max}} yp(x,y) dy dx = E_{X,Y}[Y] \end{aligned}$$



□

## X.1.8 確率不等式

### X.1.8.1 マルコフの不等式

$(X, \mathcal{M}, \mu)$  を測度空間、 $f : X \rightarrow [-\infty, \infty]$  を可測関数、 $a > 0$  とするとき、次が成り立つ。

$$\mu(\{x \in X \mid |f(x)| \geq a\}) \leq \frac{1}{a} \int_X |f(x)| \mu(dx)$$

*Proof.*

(引用 : <http://mathcommunication.hatenablog.com/entry/2017/06/27/001722>)

$$\begin{aligned} \int_X |f(x)| \mu(dx) &= \int_{|f(x)| \geq a} + \int_{|f(x)| < a} \\ &\geq \int_{|f(x)| \geq a} \\ &\geq a \mu(\{x \in X \mid |f(x)| \geq a\}) \end{aligned}$$

□

### X.1.8.2 チェビシェフの不等式

任意の連続型確変数  $X$  について次が成り立つ。

$$\Pr(|X - \mu| \geq a\sigma) \leq 1/a^2$$

但し  $a > 0, \mu = E[X], \sigma^2 = \text{Var}[X]$

*Proof.*

$X$  の確率密度関数を  $f(x)$  とすると

$$\begin{aligned} \Pr(|X - \mu| \geq a\sigma) &= \left( \int_{-\infty}^{\mu - a\sigma} + \int_{\mu + a\sigma}^{\infty} \right) \frac{(x - \mu)^2}{a^2 \sigma^2} f(x) dx \\ &\left( \because \frac{(x - \mu)^2}{a^2 \sigma^2} \geq 1, x \in (-\infty, \mu - a\sigma] \cup [\mu + a\sigma, \infty) \right) \\ &\leq \int_{-\infty}^{\infty} \frac{(x - \mu)^2}{a^2 \sigma^2} f(x) dx = 1/a^2 \end{aligned}$$

□

### X.1.8.3 ヘフディングの補題

$(\Omega, \mathcal{F}, P)$  を確率空間とし、 $X$  を  $E_P[X] = 0$ 、 $a \leq X \leq b$  (a.e.) を満たす確率変数とする。このとき任意の  $t > 0$  に対して次の不等式が成り立つ。

$$E_P[\exp(tX)] \leq \exp\left(\frac{1}{8}t^2(b-a)^2\right)$$

*Proof.*

(参考 : <http://mathcommunication.hatenablog.com/entry/2017/06/27/001722>)

$E_P[X] = 0$  より  $a < 0, b > 0$  であることに注意。指数関数の凸性より、

$$\exp(tX) \leq \frac{X-a}{b-a} \exp(tb) + \frac{b-X}{b-a} \exp(ta), \quad (\text{a.e.})$$

であるから

$$E_P[\exp(tX)] \leq \frac{-a}{b-a} \exp(tb) + \frac{b}{b-a} \exp(ta) \quad (1)$$

今、

$$p := \frac{-a}{b-a}, \quad u := t(b-a)$$

とおくと  $p := \frac{-a}{b-a}$ ,  $u := t(b-a)$  であり、

$$\begin{aligned} (1) &= p \exp(tb) + (1-p) \exp(ta) \\ &= p \exp((1-p)u) + (1-p) \exp(-pu) \end{aligned}$$

$$\because 1-p = \frac{b}{b-a} \text{ より } tb = t(b-a)(1-p) = (1-p)u \text{ また、 } ta = -tp(b-a) = -pu$$

さらに、 $\phi(u) := \log(\text{式}(1))$  とおくと

$$\begin{aligned} \phi(u) &= \log[p \exp((1-p)u) + (1-p) \exp(-pu)] \\ &= \log[\exp((1-p)u) \cdot (p + (1-p) \exp(-u))] \\ &= (1-p)u + \log(p + (1-p) \exp(-u)) \end{aligned}$$

$$\begin{aligned} \phi'(u) &= 1-p + \frac{-(1-p) \exp(-u)}{p + (1-p) \exp(-u)}, \\ \phi''(u) &= \frac{p(1-p) \exp(-u)}{(p + (1-p) \exp(-u))^2} \end{aligned}$$

ここで Taylor の定理より、ある  $u_0 \in (0, u)$  が存在して

$$\phi(u) = \phi(0) + \phi'(0)u + \frac{1}{2}\phi''(u_0)u^2 = \frac{1}{2}\phi''(u_0)u^2$$

であるが、相加・相乗平均の関係より

$$\phi''(u_0) = \frac{p \cdot (1-p) \exp(-u_0)}{(p + (1-p) \exp(-u_0))^2} \leq \frac{1}{4}$$

であるので

$$\phi(u) \leq \frac{1}{8}u^2$$

よって

$$E_P[\exp(tX)] \leq \exp(\phi(u)) \leq \exp\left(\frac{1}{8}u^2\right) = \exp\left(\frac{1}{8}t^2(b-a)^2\right)$$

□

#### X.1.8.4 ヘフディングの不等式

$(\Omega, \mathcal{F}, P)$  を確率空間とし、 $X_1, \dots, X_n$  を独立同分布な確率変数で  $a_i \leq X_i \leq b_i$  (a.e.) を満たすとする。このとき任意の  $\epsilon > 0$  に対して次の不等式が成り立つ。

$$P(S - E_P[S] \geq \epsilon) \leq \exp\left(-\frac{2\epsilon^2}{\sum_{i=1}^n (b_i - a_i)^2}\right),$$

$$P(S - E_P[S] \leq -\epsilon) \leq \exp\left(-\frac{2\epsilon^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$$

*Proof.*

(引用 : <http://mathcommunication.hatenablog.com/entry/2017/06/27/001722>)

任意の  $t > 0$  に対して

$$\begin{aligned} P(S - E_P[S] \geq \epsilon) &= P(\exp(t(S - E_P[S])) \geq \exp(t\epsilon)) \\ &\leq \exp(-t\epsilon) E_P[\exp(t(S - E_P[S]))] && (\text{マルコフの不等式}) \\ &= \exp(-t\epsilon) \prod_{i=1}^n E_P[\exp(t(X_i - E_P[X_i]))] && (\text{各 } X_i \text{ は独立}) \\ &\leq \exp(-t\epsilon) \prod_{i=1}^n \exp\left(\frac{t^2}{8}(b_i - a_i)^2\right) && (\text{ヘフディングの補題}) \\ &= \exp(-t\epsilon) \exp\left(\frac{t^2}{8} \sum_{i=1}^n (b_i - a_i)^2\right) \\ &= \exp\left(\frac{t^2}{8} \sum_{i=1}^n (b_i - a_i)^2 - t\epsilon\right) \end{aligned}$$

が成り立つ。

$$\frac{1}{8}t^2 \sum_{i=1}^n (b_i - a_i)^2 - t\epsilon = \frac{1}{8} \left( \sum_{i=1}^n (b_i - a_i)^2 \right) \left( t - \frac{4\epsilon}{\sum_{i=1}^n (b_i - a_i)^2} \right)^2 - \frac{2\epsilon^2}{\sum_{i=1}^n (b_i - a_i)^2}$$

であり、第1項  $\geq 0$  だから

$$P(S - E_P[S] \geq \epsilon) \leq \exp\left(-\frac{2\epsilon^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$$

同様にして

$$P(S - E_P[S] \leq -\epsilon) \leq \exp\left(-\frac{2\epsilon^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$$

□

## X.1.9 中心極限定理

標準化された標本平均  $\frac{\sqrt{n}(\hat{\mu}_n - \mu)}{\sigma}$  の分布は標準正規分布に弱収束する。すなわち

$$\lim_{n \rightarrow \infty} \Pr \left( \frac{\sqrt{n}(\hat{\mu}_n - \mu)}{\sigma} \leq x \right) = \Phi(x)$$

但し  $\Phi(x)$  は標準正規分布の累積分布関数である。

*Proof.*

(引用 : <https://ja.wikipedia.org/wiki/中心極限定理>)

$$Z_n := \frac{\sqrt{n}(\hat{\mu}_n - \mu)}{\sigma} = \frac{\sum_{j=1}^n X_j - n\mu}{\sqrt{n}\sigma} = \sum_{j=1}^n Y_j / \sqrt{n} \quad \text{where } Y_j := \frac{X_j - \mu}{\sigma}$$

とおく。  $Y_j$  の平均は 0 分散は 1 である。

$Z_n$  の特性関数  $\phi_{Z_n}(t)$  は

$$\begin{aligned} \phi_{Z_n}(t) &= E[e^{itZ_n}] = E[e^{it \sum_{j=1}^n Y_j / \sqrt{n}}] = E\left[\prod_{j=1}^n e^{itY_j / \sqrt{n}}\right] = \prod_{j=1}^n E[e^{itY_j / \sqrt{n}}] \quad (\because \text{各 } Y_j \text{ は独立}) \\ &= \prod_{j=1}^n \phi_{Y_j}(t/\sqrt{n}) = [\phi_{Y_1}(t/\sqrt{n})]^n \quad (\because \text{各 } Y_j \text{ は全部同等}) \end{aligned}$$

$\phi_{Y_1}(t)$  の Taylor 展開を考える。

$$\begin{aligned} \phi_{Y_1}(0) &= 1 \\ \phi_{Y_1}'(0) &= E[iY_1 e^{itY_1}]|_{t=0} = 0 \\ \phi_{Y_1}''(0) &= E[-Y_1^2 e^{itY_1}]|_{t=0} = -E[Y_1^2] = -1 \end{aligned}$$

これより、

$$\phi_{Y_1}(t/\sqrt{n}) = 1 - \frac{1}{2} \frac{t^2}{n} + c \frac{1}{3!} \left(\frac{t}{\sqrt{n}}\right)^3 + o\left(\left(\frac{t}{\sqrt{n}}\right)^3\right) \quad (t \rightarrow 0)$$

よって

$$\phi_{Z_n}(t) = \left[ 1 - \frac{1}{2} \frac{t^2}{n} + c \frac{1}{3!} \left(\frac{t}{\sqrt{n}}\right)^3 + o\left(\left(\frac{t}{\sqrt{n}}\right)^3\right) \right]^n \rightarrow e^{-t^2/2} \quad \text{as } n \rightarrow \infty$$

最右辺は標準正規分布の特性関数であり、特性関数と確率密度関数は 1 対 1 で対応するから  $Z_n = \frac{\sqrt{n}(\hat{\mu}_n - \mu)}{\sigma}$  は  $n \rightarrow \infty$  で標準正規分布に従う。  $\square$

## 第 X.2 章

# 多項分布

以下  $(X_1, X_2, \dots, X_n) \sim B_N(n, p_1, p_2, \dots, p_n)$  とする。

### X.2.1 周辺分布は 2 項分布

$$P(X_h = x_h) = \frac{n!}{x_h!(n-x_h)!} p_h^{x_h} (1-p_h)^{n-x_h}$$

*Proof.*

$$\begin{aligned} P(X = x_h) &= \sum_{x_1=0}^{n-x_h} \sum_{x_2=0}^{n-x_h-x_1} \cdots \sum_{x_{h-1}=0}^{n-x_h-\sum_{j=1}^{h-2} x_j} \sum_{x_{h+1}=0}^{n-\sum_{j=1}^h x_j} \cdots \sum_{x_n=0}^{n-\sum_{j=1}^{n-1} x_j} P(X_1 = x_1, \dots, X_n = x_n) \\ &= \sum_{x_1=0}^{n-x_h} \sum_{x_2=0}^{n-x_h-x_1} \cdots \sum_{x_{h-1}=0}^{n-x_h-\sum_{j=1}^{h-2} x_j} \sum_{x_{h+1}=0}^{n-\sum_{j=1}^h x_j} \cdots \sum_{x_n=0}^{n-\sum_{j=1}^{n-1} x_j} \\ &\quad \frac{n!}{(\prod_{i=1}^n x_i!) (n - \sum_{i=1}^n x_i)!} \left( \prod_{i=1}^n p_i^{x_i} \right) \left( 1 - \sum_{i=1}^n p_i \right)^{n - \sum_{i=1}^n x_i} \\ &= \frac{n!}{x_h!(n-x_h)!} p_h^{x_h} \cdots \sum_{x_{h-1}=0}^{n-x_h-\sum_{j=1}^{h-2} x_j} \frac{(n-x_h-\sum_{j=1}^{h-2} x_j)!}{x_{h-1}!(n-x_h-\sum_{j=1}^{h-2} x_j-x_{h-1})!} p_{h-1}^{x_{h-1}} \\ &\quad \sum_{x_{h+1}=0}^{n-\sum_{j=1}^h x_j} \frac{(n-\sum_{j=1}^h x_j)!}{x_{h+1}![(n-\sum_{j=1}^h x_j)-x_{h+1}]!} p_{h+1}^{x_{h+1}} \cdots \\ &\quad \sum_{x_n=0}^{n-\sum_{j=1}^{n-1} x_j} \frac{(n-\sum_{j=1}^{n-1} x_j)!}{x_n![(n-\sum_{j=1}^{n-1} x_j)-x_n]!} p_n^{x_n} \left[ \left( 1 - \sum_{j=1}^{n-1} p_j \right) - p_n \right]^{[(n-\sum_{j=1}^{n-1} x_j)-x_n]} \end{aligned}$$

式の末尾の  $\Sigma$  は

$$\left[ \left( 1 - \sum_{j=1}^{n-1} p_j \right) - p_n + p_n \right]^{n-\sum_{j=1}^{n-1} x_j} = \left[ 1 - \sum_{j=1}^{n-1} p_j \right]^{n-\sum_{j=1}^{n-1} x_j}$$

と等しい。同様のことを繰り返すと次の結論に至る。

$$P(X_h = x_h) = \frac{n!}{x_h!(n-x_h)!} p_h^{x_h} (1-p_h)^{n-x_h}$$

□

## X.2.2 共分散

$$\text{Cov}[X_i, X_j] = -np_i p_j$$

*Proof.*

一般に  $\text{Cov}[X, Y] = E[XY] - E[X]E[Y]$  であって、 $E[X_i] = np_i, E[X_j] = np_j$  は既知であるから  $E[X_i X_j]$  を計算すれば良い。定理 1 より

$$P(X_i = x_i, X_j = x_j) = \frac{n!}{x_i! x_j! (n-x_i-x_j)!} p_i^{x_i} p_j^{x_j} (1-p_i-p_j)^{n-x_i-x_j}$$

であることを利用する。 $X_i + X_j \leq n$  であるから一方が 0 或いは  $n$  であれば積が 0 になることに注意して

$$\begin{aligned} E[X_i X_j] &= \sum_{x_i=1}^{n-1} \sum_{x_j=1}^{n-x_i} x_i x_j P(X_i = x_i, X_j = x_j) \\ &= \sum_{x_i=1}^{n-1} \sum_{x_j=1}^{n-x_i} x_i x_j \frac{n!}{x_i! x_j! (n-x_i-x_j)!} p_i^{x_i} p_j^{x_j} (1-p_i-p_j)^{n-x_i-x_j} \\ &= \sum_{x_i=1}^{n-1} \frac{n!}{(x_i-1)!(n-x_i)!} p_i^{x_i} \sum_{x_j=1}^{n-x_i} \frac{(n-x_i)!}{(x_j-1)!(n-x_i-x_j)!} p_j^{x_j} (1-p_i-p_j)^{n-x_i-x_j} \\ &= \sum_{x_i=1}^{n-1} \frac{n!}{(x_i-1)!(n-x_i)!} p_i^{x_i} (n-x_i) p_j \\ &\quad \sum_{x_j'=0}^{n-x_i-1} \frac{(n-x_i-1)!}{x_j'!(n-x_i-1-x_j')!} p_j^{x_j'} (1-p_i-p_j)^{n-x_i-1-x_j'} \\ &= \sum_{x_i=1}^{n-1} \frac{n!}{(x_i-1)!(n-x_i)!} p_i^{x_i} (n-x_i) p_j (1-p_i)^{n-x_i-1} \\ &= p_i p_j n(n-1) \sum_{x_i'=0}^{n-2} \frac{(n-2)!}{x_i'!(n-2-x_i')!} p_i^{x_i'} (1-p_i)^{n-2-x_i'} \\ &= p_i p_j n(n-1) \end{aligned}$$

よって

$$\text{Cov}[X_i, X_j] = E[X_i X_j] - E[X_i]E[X_j] = p_i p_j n(n-1) - np_i np_j = -np_i p_j$$

□

## 第 X.3 章

# 幾何分布

### X.3.1 無記憶性からの導出

*Proof.*

確率変数  $X$  を、ベルヌーイ試行を繰り返して初めて成功するまでに要する試行回数とする。無記憶であるとはすなわち「試行開始から  $n$  回失敗したという事象の下、それからさらに  $k \in \mathbb{N}$  回失敗する条件付き確率」が「試行開始から連続で  $k$  回失敗する確率」に等しいということである。特に  $k = 1$  としたときの条件から次の式が成り立つ。

$$\begin{aligned}\frac{P(X > n+1)}{P(X > n)} &= P(X > 1) \\ P(X > n+1) &= P(X > 1)P(X > n) \\ P(X > n) - P(X > n+1) &= P(X > n) - P(X > 1)P(X > n) \\ P(X = n+1) &= (1 - P(X > 1))P(X > n) = P(X = 1)(1 - P(X \leq n))\end{aligned}$$

これより

$$\begin{aligned}P(X = 2) &= P(X = 1)(1 - P(X = 1)) = p(1 - p) \quad (p := P(X = 1)) \\ P(X = 3) &= P(X = 1)(1 - P(X = 1) - P(X = 2)) = p[1 - p - p(1 - p)] = p(1 - p)^2 \\ P(X = 4) &= P(X = 1)(1 - P(X = 1) - P(X = 2) - P(X = 3)) \\ &= p[1 - p - p(1 - p) - p(1 - p)^2] = p(1 - p)^3 \\ &\vdots \\ P(X = n) &= p(1 - p)^{n-1}\end{aligned}$$

□

## 第 X.4 章

# 指数分布

### X.4.1 定義

$\mu > 0$  とする。確率密度関数  $f(x)$  が次式となる分布を、尺度母数  $\mu$  の指数分布という。

$$f(x) = \begin{cases} \frac{1}{\mu} e^{-\frac{x}{\mu}} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

確率変数  $X$  がこの分布に従うことを  $X \sim \text{ExpDist}(\mu)$  と表すことにする。

### X.4.2 解釈

単位時間当たりの生起回数の期待値が  $\lambda$  であるような事象が時間  $x > 0$  の間に生起する確率の累積分布。

### X.4.3 無記憶性からの導出

*Proof.*

「観測開始から時間  $s$  の間イベントが起こらなかったという事象の下、それから時間  $\Delta t$  以内にイベントが起こる条件付き確率」が「観測開始 ( $t = 0$ ) から時間  $\Delta t$  以内にイベントが起こる確率」に等しいという条件より次式が成り立つ。

$$\begin{aligned} \frac{\int_s^{s+\Delta t} f(x) dx}{\int_s^\infty f(x) dx} &= \int_0^{\Delta t} f(x) dx \\ \therefore \int_s^{s+\Delta t} f(x) dx &= \int_0^{\Delta t} f(x) dx \int_s^\infty f(x) dx \\ \therefore f(s + \Delta t) - f(s) &= -f(s) \int_0^{\Delta t} f(x) dx \quad (\text{両辺を } s \text{ で微分した}) \\ \therefore f'(s) &= -f(0)f(s) \quad (\text{両辺を } \Delta t \text{ で割った後 } \Delta t \rightarrow +0 \text{ とした}) \end{aligned}$$

この微分方程式を解いて規格化すると  $f(x) = f(0) \exp(-f(0)x)$  ( $f(0) > 0$ ) となる。 $\lambda = f(0)$  とすれば  $f(x) = \lambda e^{-\lambda x}$  となる。□



#### X.4.4 特性関数

尺度母数  $\mu$  の指数分布の特性関数は次式である。

$$\mathbb{E} [e^{itX}] = \int_0^\infty \frac{1}{\mu} e^{-\frac{x}{\mu}} e^{itx} dx = \frac{1}{\mu} \int_0^\infty e^{(\frac{-1}{\mu} + it)x} dx = \frac{1}{\mu \frac{-1}{\mu} + it} \left[ e^{(\frac{-1}{\mu} + it)x} \right]_0^\infty = \frac{1}{1 - i\mu t}$$

## 第 X.5 章

# Erlang 分布

### X.5.1 定義

$k \in \mathbb{N}$ ,  $\mu > 0$  とする。確率密度関数  $f(x)$  が次式となる分布を、形状母数  $k$ , 尺度母数  $\mu$  の Erlang 分布という。

$$f(x) = \begin{cases} \frac{x^{k-1}}{(k-1)!\mu^k} e^{-\frac{x}{\mu}} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

確率変数  $X$  がこの分布に従うことを  $X \sim \text{ErlangDist}(k, \mu)$  と表すことにする。

### X.5.2 解釈

単位時間当たりの生起回数の期待値が  $\lambda = 1/\mu$  である事象が、時間  $x$  の間に  $k$  以上回数生起する (つまり、 $k$  回目の事象の発生までの待ち時間が  $x$  である) 確率の累積分布。

### X.5.3 指数分布との関係: $X_1, \dots, X_k$ (独立) $\sim \text{ExpDist}(\mu) \Rightarrow X_1 + \dots + X_k \sim \text{ErlangDist}(k, \mu)$

*Proof.* (直接的な証明)

数学的帰納法で示す。尺度母数  $\mu$  の指数分布の確率密度関数を  $f$  とし、 $\sum_{l=1}^k X_l$  が従う分布の確率密度関数を  $f_k$  とする。まず  $k=1$  のときに成り立つことは簡単に確かめられる。 $k=l \in \mathbb{N}$  まで成り立つと仮定して  $k=l+1$  のときに成り立つことを示す。

$$\begin{aligned} f_{l+1}(x) &= \int_0^x f_l(y) f(x-y) dy = \int_0^x \frac{y^{l-1}}{(l-1)!\mu^l} e^{-\frac{y}{\mu}} \frac{1}{\mu} e^{-\frac{(x-y)}{\mu}} dy \\ &= \frac{1}{(l-1)!\mu^{l+1}} e^{-\frac{x}{\mu}} \int_0^x y^{l-1} dy = \frac{x^l}{l!\mu^{l+1}} e^{-\frac{x}{\mu}} \end{aligned}$$

□

*Proof.* (特性関数を用いた証明)

$X_1 + \dots + X_k$  が従う分布の特性関数は次式である ( $X_1, \dots, X_k$  が独立であることに注意)。

$$\mathbb{E} \left[ e^{it(X_1 + \dots + X_k)} \right] = \prod_{l=1}^k \mathbb{E} \left[ e^{itX_l} \right] = \mathbb{E} \left[ e^{itX_1} \right]^k = \frac{1}{(1 - i\mu t)^k} \quad (\text{X.4.4 を用いた})$$

一方、ErlangDist  $(k, \mu)$  の特性関数を計算すると次式になる。

$$\begin{aligned}
 \int_0^\infty \frac{x^{k-1}}{(k-1)!\mu^k} e^{\frac{-x}{\mu}} e^{itx} dx &= \frac{1}{(k-1)!\mu^k} \int_0^\infty x^{k-1} e^{\left(\frac{-1}{\mu} + it\right)x} dx \\
 &= \frac{1}{(k-1)!\mu^k} \left[ (-1)^{k+1} (k-1)! \frac{1}{\left(\frac{-1}{\mu} + it\right)^k} e^{\left(\frac{-1}{\mu} + it\right)x} \right]_0^\infty \quad (\text{瞬間部分積分を用いた}) \\
 &= \frac{1}{(1 - i\mu t)^k}
 \end{aligned}$$

両者は一致する。

□

## 第 X.6 章

# Poisson 分布

### X.6.1 定義

$k \in \mathbb{N} \cup \{0\}$ ,  $\mu > 0$ ,  $\lambda = 1/\mu$  とする。確率質量関数  $P(X = k)$  が次式となる分布を、母数  $\lambda$  の Poisson 分布という。

$$P(X = k) = \begin{cases} \frac{\lambda^k}{k!} e^{-\lambda} & (k = 0, 1, 2, \dots) \\ 0 & (\text{otherwise}) \end{cases}$$

確率変数  $X$  がこの分布に従うことを  $X \sim \text{PoissonDist}(k)\mu$  と表すことにする。

### X.6.2 解釈

単位時間当たりの生起回数の期待値が  $\lambda = 1/\mu$  である事象が単位時間内に  $k$  回生起する確率。

#### X.6.2.1 指数分布からの導出

指数分布と Erlang 分布の関係をを用いる。X.5.3 において  $x = 1$ (= 単位時間) とすればよい。

### X.6.3 Erlang 分布との関係

単位時間当たりの生起回数の期待値が  $\lambda = 1/\mu$  である事象が単位時間内に  $k$  回以上生起する確率を Poisson 分布, Erlang 分布の両方の視点から求めてみる。Poisson 分布の視点で求めると次式となる。

$$\sum_{l=k}^{\infty} \frac{\lambda^l}{l!} e^{-\lambda}$$

一方で Erlang 分布の視点で求めると次式となり、当然だが上式と一致する。

$$\begin{aligned} \int_0^1 \frac{1}{(k-1)!\mu^k} x^{k-1} e^{-x/\mu} dx &= \frac{1}{(k-1)!\mu^k} \int_0^1 x^{k-1} e^{-x/\mu} dx \\ &= \frac{1}{(k-1)!\mu^k} \left[ -\sum_{l=0}^{k-1} k-1 P_l \mu^{l+1} x^{k-1-l} e^{-x/\mu} \right]_0^1 = \left[ -\sum_{l=0}^{k-1} \frac{1}{(k-1-l)!\mu^{k-1-l}} x^{k-1-l} e^{-x/\mu} \right]_0^1 \\ &= \left[ -\sum_{l=0}^{k-1} \frac{x^l}{l!\mu^l} e^{-x/\mu} \right]_0^1 = 1 - \sum_{l=0}^{k-1} \frac{1}{l!\mu^l} e^{-1/\mu} = \sum_{l=0}^{\infty} \frac{\lambda^l}{l!} e^{-\lambda} - \sum_{l=0}^{k-1} \frac{\lambda^l}{l!} e^{-\lambda} = \sum_{l=k}^{\infty} \frac{\lambda^l}{l!} e^{-\lambda} \end{aligned}$$

## X.6.4 再生性

$$X_1 \sim \text{PoissonDist}(\lambda_1), X_2 \sim \text{PoissonDist}(\lambda_2) \Rightarrow X_1 + X_2 \sim \text{PoissonDist}(\lambda_1 + \lambda_2)$$

*Proof.*

$$P(X_1 + X_2 = k) = \sum_{l=0}^k \frac{\lambda_1^l}{l!} e^{-\lambda_1} \frac{\lambda_2^{k-l}}{(k-l)!} e^{-\lambda_2} = \frac{1}{k!} e^{-(\lambda_1 + \lambda_2)} \sum_{l=0}^k \frac{k!}{l!(k-l)!} \lambda_1^l \lambda_2^{k-l} = \frac{1}{k!} (\lambda_1 + \lambda_2)^k e^{-(\lambda_1 + \lambda_2)}$$

□

## 第 X.7 章

# beta 分布

### X.7.1 Beta 分布に従う確率変数の生成

X.7.1.1  $X \sim \text{Gamma}(\alpha, 1), Y \sim \text{Gamma}(\beta, 1) \Rightarrow \frac{X}{X+Y} \sim \text{Beta}(\alpha, \beta)$

*Proof.*

$Z = \frac{X}{X+Y}$  とおく。

$$\begin{aligned} \text{PDF}[Z, z] &= \frac{\partial}{\partial z} \Pr(Z \leq z) = \frac{\partial}{\partial z} \Pr\left(X \leq \frac{z}{1-z}Y\right) \\ &= \frac{\partial}{\partial z} \int_{y=0}^{\infty} \int_{x=0}^{\frac{z}{1-z}y} \text{PDF}[\text{Gamma}(\alpha, 1), x] \text{PDF}[\text{Gamma}(\beta, 1), y] dx dy \\ &= \frac{\partial}{\partial z} \int_{y=0}^{\infty} \int_{x=0}^{\frac{z}{1-z}y} \frac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)} \frac{y^{\beta-1} e^{-y}}{\Gamma(\beta)} dx dy \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_{y=0}^{\infty} \frac{\partial}{\partial z} \int_{x=0}^{\frac{z}{1-z}y} x^{\alpha-1} e^{-x} y^{\beta-1} e^{-y} dx dy \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_{y=0}^{\infty} y^{\beta-1} e^{-y} \left(\frac{z}{1-z}y\right)^{\alpha-1} \exp\left(-\frac{z}{1-z}y\right) \left(\frac{\partial}{\partial z} \frac{z}{1-z}y\right) dy \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_{y=0}^{\infty} y^{\beta-1} \exp\left(\frac{-y}{1-z}y\right) \left(\frac{y}{1-z}\right)^{\alpha-1} z^{\alpha-1} \frac{1}{(1-z)^2} dy \\ &= \frac{z^{\alpha-1}(1-z)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^{\infty} u^{\alpha+\beta-1} e^{-u} du \quad \left(u = \frac{y}{1-z} \text{ とおいた}\right) \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} z^{\alpha-1}(1-z)^{\beta-1} = \frac{z^{\alpha-1}(1-z)^{\beta-1}}{B(\alpha, \beta)} = \text{PDF}[\text{Beta}(\alpha, \beta), z] \end{aligned}$$

□

## 第 X.8 章

# 正規分布

### X.8.1 裾確率の評価

標準正規分布の累積分布関数  $\Phi(x)$  に対して次が成り立つ。

$$1 - \Phi(\alpha) \leq \frac{1}{2} \exp\left(-\frac{1}{2}\alpha^2\right) \quad \text{for } \forall \alpha \geq 0$$

*Proof.* 標準正規分布の確率密度関数を  $\phi(\alpha)$  とすると、任意の  $\alpha > 0$  に対して次式が成り立つ。

$$1 - \Phi(\alpha) = \int_{\alpha}^{\infty} \phi(x) dx \leq \int_{\alpha}^{\infty} \frac{x}{\alpha} \phi(x) dx = \frac{1}{\alpha} [-\phi(x)]_{\alpha}^{\infty} = \frac{1}{\alpha} \phi(\alpha)$$

よって  $\alpha \geq 1$  に対しては

$$1 - \Phi(\alpha) \leq \phi(\alpha) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\alpha^2} \leq \frac{1}{2} e^{-\frac{1}{2}\alpha^2}$$

$\alpha < 1$  に対しては別のやり方で示す。  $1 - \Phi(\alpha) = 1/2 - \int_0^{\alpha} \phi(x) dx =: 1/2 - S(\alpha)$  であり、 $\phi(x)$  の変曲点は  $x = \pm 1$  である。

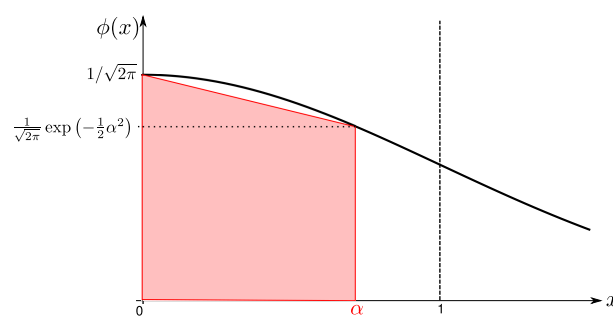


図 X.8.1  $\phi(x)$  のグラフ

$$\begin{aligned}
S(\alpha) &\geq \text{赤い台形の面積} = \frac{\alpha}{2\sqrt{2\pi}}(1 + e^{-\frac{1}{2}\alpha^2}) \\
\therefore 1 - \Phi(\alpha) &\leq \frac{1}{2} \left( 1 - \frac{\alpha}{\sqrt{2\pi}}(1 + e^{-\frac{1}{2}\alpha^2}) \right) \\
\therefore 1 - \Phi(\alpha) - \frac{1}{2}e^{-\frac{1}{2}\alpha^2} &\leq \frac{1}{2} \left( 1 - \frac{\alpha}{\sqrt{2\pi}}(1 + e^{-\frac{1}{2}\alpha^2}) - e^{-\frac{1}{2}\alpha^2} \right) =: g(\alpha)/2
\end{aligned}$$

$g(\alpha)$  は区間  $[0, 1]$  で凸であり (凸関数の和が凸関数であることと、2 階導関数の計算によりわかる)、 $g(0) = 0$  である。また、数値計算によると  $g(1) \approx 1 - \frac{4.8}{4.2} < 0$  であるから区間  $[0, 1]$  で  $g(\alpha) \leq 0$   $\square$

## X.8.2 スケール変換とシフト

$$X \sim N(\mu, \sigma^2) \Leftrightarrow Z = aX + b \sim N(a\mu + b, (a\sigma)^2) \quad (a \neq 0)$$

*Proof.*

( $\Rightarrow$ )

$X$  の確率密度関数を  $f^X(x)$  と書くことにすると

$$f^X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(x - \mu)^2}{2\sigma^2} \right]$$

(i)  $a > 0$  の場合

$$\begin{aligned}
P(-\infty \leq Z \leq z) &= P\left(\infty \leq X \leq \frac{z-b}{a}\right) = \int_{-\infty}^{\frac{z-b}{a}} f^X(x) dx \\
\therefore f^Z(z) &= \frac{d}{dz} P(-\infty \leq Z \leq z) = f^X\left(\frac{z-b}{a}\right) \cdot \frac{d}{dz} \left(\frac{z-b}{a}\right) = \frac{1}{\sqrt{2\pi}(a\sigma)} \exp \left[ -\frac{(z - a\mu - b)^2}{2(a\sigma)^2} \right] \\
\therefore f^Z(z) &= \frac{1}{\sqrt{2\pi}(a\sigma)} \exp \left[ -\frac{(z - a\mu - b)^2}{2(a\sigma)^2} \right] = \text{pdf of } N(a\mu + b, (a\sigma)^2)
\end{aligned}$$

(ii)  $a < 0$  の場合

$$\begin{aligned}
P(-\infty \leq Z \leq z) &= P\left(\frac{z-b}{a} \leq X \leq \infty\right) = \int_{\frac{z-b}{a}}^{\infty} f(x) dx \\
\therefore f^Z(z) &= \frac{d}{dz} P(-\infty \leq Z \leq z) = -f^X\left(\frac{z-b}{a}\right) \cdot \frac{d}{dz} \left(\frac{z-b}{a}\right) = \frac{-1}{\sqrt{2\pi}(a\sigma)} \exp \left[ -\frac{(z - a\mu - b)^2}{2(a\sigma)^2} \right] \\
&= \frac{1}{\sqrt{2\pi}(|a|\sigma)} \exp \left[ -\frac{(z - a\mu - b)^2}{2(a\sigma)^2} \right] = \text{pdf of } N(a\mu + b, (a\sigma)^2)
\end{aligned}$$

( $\Leftarrow$ )

$Y \sim N(a\mu + b, (b\sigma)^2)$  として、 $(Y - b)/a (= X)$  の分布を既に示された必要性に依って計算すれば明らかである。  $\square$



### X.8.3 再生性

$$X \sim N(\mu_X, \sigma_X^2) \quad , \quad Y \sim N(\mu_Y, \sigma_Y^2) \quad \Rightarrow \quad Z = X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$$

*Proof.*

$X, Y, Z$  の確率密度関数を  $f^X(x), f^Y(y), f^Z(z)$  と書くことにすると

$$\begin{aligned} f^Z(z) &= \frac{d}{dz} \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{z-x} f^X(x) f^Y(y) dy dx = \int_{-\infty}^{\infty} f^X(x) f^Y(-x+z) dx \quad (\because III.24.5) \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left[-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right] \frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left[-\frac{(-x+z-\mu_Y)^2}{2\sigma_Y^2}\right] dx \\ &= \frac{1}{2\pi\sigma_X\sigma_Y} \int_{-\infty}^{\infty} \exp\left[-\frac{(x-\mu_X)^2}{2\sigma_X^2} - \frac{(-x+z-\mu_Y)^2}{2\sigma_Y^2}\right] dx \end{aligned}$$

$\exp$  の引数を  $A$  とおくと

$$\begin{aligned} A &= -\frac{1}{2\sigma_X^2\sigma_Y^2} [\sigma_Y^2(x-\mu_X)^2 + \sigma_X^2(-x+z-\mu_Y)^2] \\ &= -\frac{1}{2\sigma_X^2\sigma_Y^2} [\sigma_X^2(x^2 + \mu_X^2 - 2\mu_X x) + \sigma_X^2(x^2 + z^2 + \mu_Y^2 - 2zx - 2z\mu_Y + 2\mu_Y x)] \\ &= -\frac{1}{2\sigma_X^2\sigma_Y^2} [(\sigma_X^2 + \sigma_Y^2)x^2 + (-2\sigma_Y^2\mu_X - 2\sigma_X^2z + 2\sigma_X^2\mu_Y)x \\ &\quad + \sigma_Y^2\mu_X^2 + \sigma_X^2z^2 + \sigma_X^2\mu_Y^2 - 2\sigma_X^2z\mu_Y] \\ &= -\frac{1}{2\sigma_X^2\sigma_Y^2} \left\{ \left[ \sqrt{\sigma_X^2 + \sigma_Y^2}x - \frac{\sigma_Y^2\mu_X + \sigma_X^2z - \sigma_X^2\mu_Y}{\sqrt{\sigma_X^2 + \sigma_Y^2}} \right]^2 \right. \\ &\quad \left. - \frac{1}{\sigma_X^2 + \sigma_Y^2} (\sigma_Y^4\mu_X^2 + \sigma_X^4z^2 + \sigma_X^4\mu_Y^2 + 2\sigma_X^2\sigma_Y^2\mu_X z - 2\sigma_X^4\mu_Y z - 2\sigma_X^2\sigma_Y^2\mu_X\mu_Y) \right. \\ &\quad \left. + \sigma_Y^2\mu_X^2 + \sigma_X^2z^2 + \sigma_X^2\mu_Y^2 - 2\sigma_X^2z\mu_Y \right\} \\ &= -\frac{1}{2\sigma_X^2\sigma_Y^2} \left\{ \left[ \sqrt{\sigma_X^2 + \sigma_Y^2}x - \frac{\sigma_Y^2\mu_X + \sigma_X^2z - \sigma_X^2\mu_Y}{\sqrt{\sigma_X^2 + \sigma_Y^2}} \right]^2 \right. \\ &\quad \left. + \frac{\sigma_X^2\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2} (\mu_X^2 + z^2 + \mu_Y^2 - 2\mu_Y z - 2\mu_X z + 2\mu_X\mu_Y) \right\} \\ &= -\frac{1}{2\sigma_X^2\sigma_Y^2} \left\{ \left[ \sqrt{\sigma_X^2 + \sigma_Y^2}x - \frac{\sigma_Y^2\mu_X + \sigma_X^2z - \sigma_X^2\mu_Y}{\sqrt{\sigma_X^2 + \sigma_Y^2}} \right]^2 + \frac{\sigma_X^2\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2} (z - \mu_X - \mu_Y)^2 \right\} \end{aligned}$$

これと (1) より

$$\begin{aligned} f^Z(z) &= \frac{1}{2\pi\sigma_X\sigma_Y} \exp\left[-\frac{(z-\mu_X-\mu_Y)^2}{2(\sigma_X^2+\sigma_Y^2)}\right] \int_{-\infty}^{\infty} \exp\left[\frac{\sigma_X^2+\sigma_Y^2}{2\sigma_X^2\sigma_Y^2} \left(x - \frac{\sigma_Y^2\mu_X + \sigma_X^2z - \sigma_X^2\mu_Y}{\sigma_X^2 + \sigma_Y^2}\right)^2\right] dx \\ &= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_X^2 + \sigma_Y^2}} \exp\left[-\frac{(z-\mu_X-\mu_Y)^2}{2(\sigma_X^2 + \sigma_Y^2)}\right] = \text{pdf of } N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2) \end{aligned}$$

□

## X.8.4 一次結合

$$X_i \sim N(\mu_i, \sigma_i^2) \quad (i = 1, 2, \dots) \Rightarrow Z = \sum_{i=1}^n c_i X_i \sim N\left(\sum_{i=1}^n c_i \mu_i, \sum_{i=1}^n (c_i \sigma_i)^2\right)$$

*Proof.*

定理 1(スケール変換&シフト), 2(和) より明らか。 □

## X.8.5 分散 $\sigma^2$ , 平均 $\mu$ のときの累積分布関数は $\Phi\left(\frac{z-\mu}{\sigma}\right)$

標準正規分布の累積分布関数を  $\Phi(z) := \int_{-\infty}^z \phi(x) dx$  とするとき、 $X \sim N(\mu, \sigma)$  の累積分布関数は  $\Phi_X(Z) = \Phi\left(\frac{z-\mu}{\sigma}\right)$  となる。

*Proof.*

$$\begin{aligned} \Phi_X(z) &= \int_{-\infty}^z \frac{1}{\sigma} \phi\left(\frac{z-\mu}{\sigma}\right) dx = \int_{-\infty}^{\frac{z-\mu}{\sigma}} d\phi(y) y \quad \left(\frac{z-\mu}{\sigma} = y \text{ とおいた}\right) \\ &= \Phi\left(\frac{z-\mu}{\sigma}\right) \end{aligned}$$

□

## X.8.6 零平均, 同分散の正規分布に独立に従う $n$ 個の確率変数の直交変換もまた同じ分布に従う

$X_1, \dots, X_n \sim N(0, \sigma^2)$ ,  $\mathbf{X} := [X_1, \dots, X_n]^\top$  とし、 $P \in \mathbb{R}^{n \times n}$  を直交行列とする。 $\mathbf{Y} := [Y_1, \dots, Y_n]^\top := P\mathbf{X}$  とすると、 $Y_1, \dots, Y_n \sim N(0, \sigma^2)$

*Proof.*

$\mathbf{X}$  と  $\mathbf{Y}$  の確率密度関数が一致することを示せば良い。両者の確率密度関数を各々  $p_{\mathbf{X}}, p_{\mathbf{Y}}$  とする。任意の点  $\mathbf{y}_0 \in \mathbb{R}^n$  における  $p_{\mathbf{Y}}$  を計算するために、 $\Omega$  を、 $\mathbf{y}_0$  を含む任意の領域とする。 $\Omega$  を直交変換  $P$  の逆写像で

写した領域を  $P^\top(\Omega)$  と表記すると

$$\begin{aligned}
 p_Y(\mathbf{y}_0) &:= \lim_{|\Omega| \rightarrow 0} \frac{\Pr(\mathbf{y} \in \Omega)}{|\Omega|} = \lim_{|\Omega| \rightarrow 0} \frac{\Pr(\mathbf{x} \in P^\top(\Omega))}{|\Omega|} = \lim_{|\Omega| \rightarrow 0} \frac{1}{|\Omega|} \int_{P^\top(\Omega)} p_X(\mathbf{x}) d\mathbf{x} \\
 &= \lim_{|\Omega| \rightarrow 0} \frac{1}{|\Omega|} \int_{P^\top(\Omega)} \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\|\mathbf{x}\|^2}{2\sigma^2}\right) d\mathbf{x} \\
 &= \lim_{|\Omega| \rightarrow 0} \frac{1}{|\Omega|} \int_{P(P^\top(\Omega))} \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\|P^\top \mathbf{y}\|^2}{2\sigma^2}\right) |\det P^\top| d\mathbf{y} \quad (1)
 \end{aligned}$$

(↑ 重積分の変数変換  $\mathbf{x} = P^\top \mathbf{y}$  を用いた)

$$\begin{aligned}
 &= \lim_{|\Omega| \rightarrow 0} \frac{1}{|\Omega|} \int_{\Omega} \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\|\mathbf{y}\|^2}{2\sigma^2}\right) d\mathbf{y} \quad (P \text{ は直交行列だから } \det P^\top = \pm 1, \|P^\top \mathbf{y}\| = \|\mathbf{y}\|) \\
 &= \lim_{|\Omega| \rightarrow 0} \frac{1}{|\Omega|} \int_{\Omega} p_X(\mathbf{y}) d\mathbf{y} \quad (2)
 \end{aligned}$$

重積分の平均値の定理より、ある  $\tilde{\mathbf{y}} \in \Omega$  が存在して

$$(1) = \lim_{|\Omega| \rightarrow 0} \frac{1}{|\Omega|} |\Omega| p_X(\tilde{\mathbf{y}}) = \lim_{|\Omega| \rightarrow 0} p_X(\tilde{\mathbf{y}}) = p_X(\mathbf{y}_0)$$

□

## X.8.7 多変量正規分布

### X.8.7.1 定義

$\Sigma$  を正定値対称行列とし、対角化が  $\text{diag}(\sigma_1^2, \dots, \sigma_n^2)$  であるとする。確率変数  $X_1, \dots, X_n$  の同時確率密度関数  $f_X(\mathbf{x})$  が

$$f_X(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right]$$

であるとき、 $X_1, \dots, X_n$  が従う分布は**多変量正規分布**と呼ばれ、 $N(\boldsymbol{\mu}, \Sigma)$  と表記される。期待値と共分散行列は  $E[\mathbf{X}] = \boldsymbol{\mu}$ ,  $\text{Cov}[\mathbf{X}] = \Sigma$  である。

### X.8.7.2 密度関数の全空間積分が 1 になることの確認

$\Sigma$  に関する仮定より、適当な直交行列  $V$  が存在して  $\Sigma = V\Lambda V^\top$ ,  $\Lambda = \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$  と変形できる。そこで  $\mathbf{x} - \boldsymbol{\mu} = V\mathbf{y}$  すなわち  $\mathbf{y} = V^\top(\mathbf{x} - \boldsymbol{\mu})$  なる直交変換を行うと、求めたい積分は次のように変形できる。

$$\begin{aligned}
 \int_{\mathbb{R}^n} f_X(\mathbf{x}) d\mathbf{x} &= \int_{\mathbb{R}^n} \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}\mathbf{y}^\top V(V\Lambda V^\top)^{-1}V\mathbf{y}\right) |\det V| d\mathbf{y} \\
 &= \int_{\mathbb{R}^n} \frac{1}{\prod_{i=1}^n \sqrt{2\pi\sigma_i}} \exp\left(-\frac{1}{2}\mathbf{y}^\top \Lambda^{-1}\mathbf{y}\right) d\mathbf{y} \quad (\because \text{は直交行列なので } |\det V| = 1) \\
 &= \int_{\mathbb{R}^n} \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left(-\frac{1}{2\sigma_i^2} y_i^2\right) d\mathbf{y} = \prod_{i=1}^n \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left(-\frac{1}{2\sigma_i^2} y_i^2\right) dy_i = 1
 \end{aligned}$$

### X.8.7.3 期待値の導出

*Proof.*

先程の直交変換をここでも用いる。

$$\begin{aligned}
 E[\mathbf{X}] &= \int_{\mathbb{R}^n} \mathbf{x} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^n} (V\mathbf{y} + \boldsymbol{\mu}) \frac{1}{\prod_{i=1}^n \sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2}\mathbf{y}^\top \Lambda^{-1} \mathbf{y}\right) d\mathbf{y} \\
 &= \boldsymbol{\mu} + V \int_{\mathbb{R}^n} \mathbf{y} \frac{1}{\prod_{i=1}^n \sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2}\mathbf{y}^\top \Lambda^{-1} \mathbf{y}\right) d\mathbf{y} \\
 &= \boldsymbol{\mu} + V \int_{\mathbb{R}^n} \mathbf{y} \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2\sigma_i^2} y_i^2\right) d\mathbf{y} \\
 &= \boldsymbol{\mu} + V \int_{y_n=-\infty}^{\infty} \cdots \int_{y_1=-\infty}^{\infty} \mathbf{y} \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2\sigma_i^2} y_i^2\right) dy_1 \cdots dy_n
 \end{aligned}$$

第二項目はベクトルの各インデックス毎に奇関数の対称積分が起り、結局全体が0になることがわかる。  $\square$

### X.8.7.4 共分散行列の導出

*Proof.*

[X.8.7.2](#) で用いた直交変換をここでも用いる。

$$\begin{aligned}
 \text{Cov}[\mathbf{X}] &= E[(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])^\top] = \int_{\mathbb{R}^n} (\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^\top f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \\
 &= \int_{\mathbb{R}^n} V\mathbf{y}\mathbf{y}^\top V^\top \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2\sigma_i^2} y_i^2\right) d\mathbf{y} \\
 &= V \left( \int_{\mathbb{R}^n} \mathbf{y}\mathbf{y}^\top \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2\sigma_i^2} y_i^2\right) d\mathbf{y} \right) V^\top
 \end{aligned}$$

積分の部分に着目すると、非対角要素では奇関数の対称積分が起り、0になる。第  $k$  対角成分は、 $y_i (i \neq k)$  に関する積分を先に実行することで

$$\int_{\mathbb{R}^n} y_k^2 \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2\sigma_i^2} y_i^2\right) d\mathbf{y} = \int_{-\infty}^{\infty} y_k^2 \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left(-\frac{1}{2\sigma_k^2} y_k^2\right) dy_k = \sigma_k^2$$

となることがわかる。よって

$$\text{Cov}[\mathbf{X}] = V \Lambda V^\top = \Sigma$$

$\square$

### X.8.7.5 特性関数は $\phi_{\mathbf{X}}(\mathbf{t}) = \exp(i\mathbf{t}^\top \boldsymbol{\mu} - \mathbf{t}^\top \Sigma \mathbf{t}/2)$

*Proof.*

X.8.7.2 で用いた直交変換をここでも用いる。

$$\begin{aligned}
\phi_{\mathbf{X}}(\mathbf{t}) &= \mathbb{E} [\exp(i\mathbf{t}^\top \mathbf{X})] = \int_{\mathbb{R}^n} \exp(i\mathbf{t}^\top \mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \\
&= \int_{\mathbb{R}^n} \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp(i\mathbf{t}^\top \mathbf{x}) \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right) d\mathbf{x} \\
&= \int_{\mathbb{R}^n} \frac{1}{(2\pi)^{n/2} \prod_{i=1}^n \sigma_i} \exp(i\mathbf{t}^\top (V\mathbf{y} + \boldsymbol{\mu})) \exp\left(-\frac{1}{2}\mathbf{y}^\top \Lambda^{-1}\mathbf{y}\right) |\det V| d\mathbf{y} \\
&= \exp(i\mathbf{t}^\top \boldsymbol{\mu}) \int_{\mathbb{R}^n} \frac{1}{(2\pi)^{n/2} \prod_{i=1}^n \sigma_i} \exp\left(-\frac{1}{2}\mathbf{y}^\top \Lambda^{-1}\mathbf{y} + i\mathbf{t}^\top V\mathbf{y}\right) d\mathbf{y} \\
&= \exp(i\mathbf{t}^\top \boldsymbol{\mu}) \int_{\mathbb{R}^n} \frac{1}{(2\pi)^{n/2} \prod_{i=1}^n \sigma_i} \exp\left(-\frac{1}{2} \sum_{i=1}^n \frac{1}{\sigma_i^2} y_i^2 + i\mathbf{t}^\top V\mathbf{y}\right) d\mathbf{y}
\end{aligned}$$

ここで  $c_i := (\mathbf{t}^\top V)[i]$  とすると  $\exp$  の引数は次のように変形できる。

$$\begin{aligned}
-\frac{1}{2} \sum_{i=1}^n \frac{1}{\sigma_i^2} y_i^2 + i\mathbf{t}^\top V\mathbf{y} &= -\frac{1}{2} \sum_{i=1}^n \frac{1}{\sigma_i^2} y_i^2 + i \sum_{i=1}^n c_i y_i = -\frac{1}{2} \sum_{i=1}^n \frac{1}{\sigma_i^2} (y_i^2 - i2\sigma_i^2 c_i y_i) \\
&= -\frac{1}{2} \sum_{i=1}^n \frac{1}{\sigma_i^2} [(y_i - i\sigma_i^2 c_i)^2 + \sigma_i^4 c_i^2] = -\frac{1}{2} \sum_{i=1}^n \sigma_i^2 c_i^2 - \frac{1}{2} \sum_{i=1}^n \frac{1}{\sigma_i^2} (y_i - i\sigma_i^2 c_i)^2
\end{aligned}$$

さらに次式が成り立つ。

$$\sum_{i=1}^n \sigma_i^2 c_i^2 = [c_1, \dots, c_n] \Lambda [c_1, \dots, c_n]^\top = \mathbf{t}^\top V \Lambda V^\top \mathbf{t} = \mathbf{t}^\top \Sigma \mathbf{t}$$

よって

$$\begin{aligned}
\phi_{\mathbf{X}}(\mathbf{t}) &= \exp\left(i\mathbf{t}^\top \boldsymbol{\mu} - \frac{1}{2} \mathbf{t}^\top \Sigma \mathbf{t}\right) \int_{\mathbb{R}^n} \frac{1}{(2\pi)^{n/2} \prod_{i=1}^n \sigma_i} \exp\left(-\frac{1}{2} \sum_{i=1}^n \frac{1}{\sigma_i^2} (y_i - i\sigma_i^2 c_i)^2\right) d\mathbf{y} \\
&= \exp\left(i\mathbf{t}^\top \boldsymbol{\mu} - \frac{1}{2} \mathbf{t}^\top \Sigma \mathbf{t}\right) \int_{\mathbb{R}^n} \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2\sigma_i^2} (y_i - i\sigma_i^2 c_i)^2\right) d\mathbf{y} \\
&= \exp\left(i\mathbf{t}^\top \boldsymbol{\mu} - \frac{1}{2} \mathbf{t}^\top \Sigma \mathbf{t}\right) \prod_{i=1}^n I_i \quad \left(I_i := \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2\sigma_i^2} (y_i - i\sigma_i^2 c_i)^2\right) dy_i\right)
\end{aligned}$$

$I_i$  は複素積分により 1 である (矩形閉路 ABCDA:  $A = (-L, 0)$ ,  $B = (L, 0)$ ,  $C = (L, i\sigma_i^2 c_i)$ ,  $D = (-L, i\sigma_i^2 c_i)$  を用いて  $L \rightarrow \infty$  とするとよい) ので

$$\phi_{\mathbf{X}}(\mathbf{t}) = \exp\left(i\mathbf{t}^\top \boldsymbol{\mu} - \frac{1}{2} \mathbf{t}^\top \Sigma \mathbf{t}\right)$$

□

### X.8.7.6 特性関数を用いた期待値と分散の導出

*Proof.*

まず、期待値については  $\nabla_{\mathbf{t}} \phi_{\mathbf{X}}(\mathbf{t}) = (i\boldsymbol{\mu} - \Sigma \mathbf{t}) \phi_{\mathbf{X}}(\mathbf{t})$  より  $\mathbb{E}[\mathbf{X}] = -i \nabla_{\mathbf{t}} \phi_{\mathbf{X}}(\mathbf{0}) = \boldsymbol{\mu}$  である。次に、共分散行列については、

$$\nabla_{\mathbf{t}}^2 \phi_{\mathbf{X}}(\mathbf{t}) = (-\Sigma + (i\boldsymbol{\mu} - \Sigma \mathbf{t})(i\boldsymbol{\mu} - \Sigma \mathbf{t})^\top) \phi_{\mathbf{X}}(\mathbf{t})$$

より  $\mathbb{E}[\mathbf{X} \mathbf{X}^\top] = -\nabla_{\mathbf{t}}^2 \phi_{\mathbf{X}}(\mathbf{0}) = \Sigma + \boldsymbol{\mu} \boldsymbol{\mu}^\top$  であるので  $\text{Cov}[\mathbf{X}] = \mathbb{E}[\mathbf{X} \mathbf{X}^\top] - \mathbb{E}[\mathbf{X}] \mathbb{E}[\mathbf{X}]^\top = \Sigma$  □

### X.8.7.7 線形写像 : $A\mathbf{X} + \mathbf{b} \sim N(A\boldsymbol{\mu} + \mathbf{b}, A\Sigma A^\top)$

*Proof.*

特性関数を用いて証明する。

$$\begin{aligned}\varphi_{A\mathbf{X}+\mathbf{b}}(\mathbf{t}) &= \mathbb{E} [\exp(i\mathbf{t}^\top (A\mathbf{X} + \mathbf{b}))] = \exp(i\mathbf{t}^\top \mathbf{b}) \mathbb{E} [\exp(i\mathbf{t}^\top A\mathbf{X})] = \exp(i\mathbf{t}^\top \mathbf{b}) \mathbb{E} [\exp(i(A^\top \mathbf{t})^\top \mathbf{X})] \\ &= \exp(i\mathbf{t}^\top \mathbf{b}) \varphi_{\mathbf{X}}(A^\top \mathbf{t}) = \exp(i\mathbf{t}^\top \mathbf{b}) \exp\left(i(A^\top \mathbf{t})^\top \boldsymbol{\mu} - \frac{1}{2}(A^\top \mathbf{t})^\top \Sigma (A^\top \mathbf{t})\right) \\ &= \exp(i\mathbf{t}^\top \mathbf{b}) \exp\left(i\mathbf{t}^\top (A\boldsymbol{\mu}) - \frac{1}{2}\mathbf{t}^\top (A\Sigma A^\top)\mathbf{t}\right) = \exp\left(i\mathbf{t}^\top (A\boldsymbol{\mu} + \mathbf{b}) - \frac{1}{2}\mathbf{t}^\top (A\Sigma A^\top)\mathbf{t}\right)\end{aligned}$$

これは期待値  $A\boldsymbol{\mu} + \mathbf{b}$ , 共分散行列  $A\Sigma A^\top$  の多変量正規分布の特性関数であるから、 $A\mathbf{X} + \mathbf{b}$  がこの多変量正規分布に従うことがわかる。  $\square$

### X.8.7.8 周辺分布 : $X_i \sim N(\mu_i, \sigma_i^2)$

*Proof.* 直前の定理で  $A = \mathbf{e}_i^\top$ ,  $\mathbf{b} = 0$  とすればよい。  $\square$

## 第 X.9 章

# Rayleigh 分布

### X.9.1 定義

$\sigma > 0$  とする。母数  $\sigma$  の Rayleigh 分布の連続確率密度関数を次式で定義する。

$$f(x; \sigma) := \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

### X.9.2 正規分布する成分を持つ 2 次元ベクトルのノルムとの関係

$\mathbf{X} \in \mathbb{R}^2$  の各成分が正規分布  $N(0, \sigma^2)$  に従う独立な確率変数であるとき、 $\|\mathbf{X}\|_2$  は母数  $\sigma$  の Rayleigh 分布に従う。

*Proof.*

$N(0, \sigma^2)$  の確率密度関数を  $p$  と表す。 $r \geq 0$  とする。 $\Omega(r)$  を半径  $r$  の円盤領域とする。

$$\begin{aligned} \Pr(\|\mathbf{X}\|_2 \leq r) &= \int_{\Omega(r)} p(x_1)p(x_2)dx_1dx_2 = \int_0^r \int_0^{2\pi} p(r' \cos \theta)p(r' \sin \theta)r'd\theta dr' \\ &= \int_0^r \int_0^{2\pi} \frac{r'}{2\pi\sigma^2} \exp\left(-\frac{r'^2}{2\sigma^2}\right) r'd\theta dr' = \int_0^r \frac{r'}{\sigma^2} \exp\left(-\frac{r'^2}{2\sigma^2}\right) dr' \end{aligned}$$

よって  $\|\mathbf{X}\|_2$  の確率密度関数は次式である。

$$\frac{d}{dr} \Pr(\|\mathbf{X}\|_2 \leq r) = \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right)$$

□

## 第 X.10 章

# Rice 分布

### X.10.1 定義

$a, \sigma > 0$  とする。Rice 分布の連続確率密度関数を次式で定義する。

$$f(x; a, \sigma) := \frac{x^2}{\sigma^2} \exp\left(-\frac{x^2 + a^2}{2\sigma^2}\right) I_0\left(\frac{ax}{\sigma^2}\right)$$

ここに  $I_0$  は 0 次の第一種変形 Bessel 関数である。

### X.10.2 正規分布する成分を持つ 2 次元ベクトルと定数ベクトルとの関係

#### X.10.2.1 記号の準備

以下で用いる記号ここを定義しておく。 $\mathbf{a} := [a_1, a_2]^\top \in \mathbb{R}^2$ ,  $a := \|\mathbf{a}\|_2$  とする。 $\tilde{\mathbf{X}} \in \mathbb{R}^2$  の各成分が正規分布  $N(0, \sigma^2)$  に従う独立な確率変数であるとする。 $\mathbf{X} := \mathbf{a} + \tilde{\mathbf{X}}$  とし、 $\mathbf{X}$  の第  $i$  成分を  $X_i$  と表す。

#### X.10.2.2 ノルムとの関係

$\|\mathbf{X}\|_2$  は母数  $\sigma$  の Rayleigh 分布に従う。

*Proof.*

座標系を適当に回転させて  $\mathbf{a} = [a, 0]^\top$  とでき、この操作は  $\|\mathbf{X}\|_2$  を変化させない。よって一般性を失わずに  $\mathbf{a} = [a, 0]^\top$  であるとする。

$X_i$  の確率密度分布を  $f_i$  とする。 $f_1$  は  $N(0, \sigma^2)$  の、 $f_2$  は  $N(a, \sigma^2)$  の確率密度関数である。 $r \geq 0$  とし、原点を中心とする半径  $r$  の円盤領域を  $\Omega(r)$  とする。

$$\begin{aligned} \Pr(\|\mathbf{X}\|_2 \leq r) &= \int_{\Omega(r)} f_1(x_1) f_2(x_2) dx_1 dx_2 = \int_0^r \int_0^{2\pi} f_1(r' \cos \theta) f_2(r' \sin \theta) r' d\theta dr' \\ &= \int_0^r \int_0^{2\pi} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2\sigma^2} ((r' \cos \theta - a)^2 + r'^2 \sin^2 \theta)\right) r' d\theta dr' \\ &= \frac{1}{\sigma^2} \int_0^r r' \exp\left(-\frac{1}{2\sigma^2} (r'^2 + a^2)\right) \underbrace{\frac{1}{2\pi} \int_0^{2\pi} \exp\left(\frac{ar' \cos \theta}{\sigma^2}\right) d\theta}_{(A)} dr' \end{aligned}$$



(A) を評価する。

$$\begin{aligned}(A) &= \frac{1}{2\pi} \int_0^{2\pi} \exp\left(\frac{ar' \sin \theta}{\sigma^2}\right) d\theta = \frac{1}{2\pi} \int_0^{2\pi} \exp\left(\frac{-ar' \sin \theta}{\sigma^2}\right) d\theta \\ &\quad (\text{被積分関数の } 2\pi \text{ 周期性を用いた}) \\ &= J_0\left(i\frac{ar'}{\sigma^2}\right) = I_0\left(\frac{ar'}{\sigma^2}\right)\end{aligned}$$

よって次式が成り立つ。

$$\Pr(\|\mathbf{X}\|_2 \leq r) = \frac{1}{\sigma^2} \int_0^r r' \exp\left(-\frac{1}{2\sigma^2} (r'^2 + a^2)^2\right) I_0\left(\frac{ar'}{\sigma^2}\right) dr'$$

$\|\mathbf{X}\|_2$  の確率密度関数は次式である。

$$\frac{\partial}{\partial r} \Pr(\|\mathbf{X}\|_2 \leq r) = \frac{1}{\sigma^2} r \exp\left(-\frac{1}{2\sigma^2} (r^2 + a^2)^2\right) I_0\left(\frac{ar}{\sigma^2}\right)$$

□

### X.10.2.3 偏角との関係

$\theta_0 := \text{Arg}(a_1 + ia_2)$ ,  $\Theta := \text{Arg}(X_1 + iX_2)$  とする。 $\Theta$  の確率密度関数は次式である。

$$\rho(\theta; \mathbf{a}, \sigma) = \frac{1}{2\pi} \exp\left(-\frac{a^2}{2\sigma^2}\right) \left[1 + \frac{a \cos(\theta - \theta_0)}{\sigma} \sqrt{\frac{\pi}{2}} \exp\left(-\frac{a^2 \cos^2(\theta - \theta_0)}{2\sigma^2}\right) \left(1 + \text{erf}\left(\frac{a \cos(\theta - \theta_0)}{\sqrt{2}\sigma}\right)\right)\right]$$

*Proof.*

まず  $\theta_0 = 0$  の場合を考え、その後で一般化する。 $X_i$  の確率密度分布を  $f_i$  とする。 $f_1$  は  $N(0, \sigma^2)$  の、 $f_2$  は  $N(a, \sigma^2)$  の確率密度関数である。系は  $\Theta$  に関して偶対称であるから、 $\Theta \in [0, \pi]$  の範囲で考えた結果を  $\Theta < 0$  の場合に偶対称に拡張できる。

$$\Pr(0 \leq \Theta \leq \theta) = \int_{\theta'=0}^{\theta} \int_{r'=0}^{\infty} \frac{r'}{2\pi\sigma^2} \exp\left(-\frac{1}{2\sigma^2} ((r' \cos \theta' - a)^2 + r'^2 \sin^2 \theta')\right) dr' d\theta'$$

よって  $\Theta$  の確率密度関数を  $\rho(\Theta; a, \sigma)$  とすると

$$\begin{aligned}\rho(\Theta; a, \sigma) &= \frac{\partial}{\partial \theta} \Pr(0 \leq \Theta \leq \theta) = \int_{r'=0}^{\infty} \frac{r'}{2\pi\sigma^2} \exp\left(-\frac{1}{2\sigma^2} ((r' \cos \theta' - a)^2 + r'^2 \sin^2 \theta')\right) dr' \\ &= \frac{1}{2\pi} \exp\left(-\frac{a^2 \sin^2 \theta}{2\sigma^2}\right) \underbrace{\int_{r'=0}^{\infty} \frac{r'}{\sigma^2} \exp\left(-\frac{(r' - a \cos \theta)^2}{2\sigma^2}\right) dr'}_{(A)}\end{aligned}$$

ここで

$$-\frac{\partial}{\partial r'} \exp\left(-\frac{(r' - a \cos \theta)^2}{2\sigma^2}\right) = \frac{r' - a \cos \theta}{\sigma^2} \exp\left(-\frac{(r' - a \cos \theta)^2}{2\sigma^2}\right)$$

を利用し、

$$\begin{aligned}
 (A) &= \left[ -\exp\left(-\frac{(r' - a \cos \theta)^2}{2\sigma^2}\right) \right]_{r'=0}^{\infty} + \frac{a \cos \theta}{\sigma^2} \underbrace{\int_{r'=0}^{\infty} \exp\left(-\frac{(r' - a \cos \theta)^2}{2\sigma^2}\right) dr'}_{(B)} \\
 (B) &= \int_{-\frac{r' - a \cos \theta}{\sqrt{2}\sigma}}^{\infty} e^{-t^2} \sqrt{2}\sigma dt \quad (\text{変数変換 } t := (r' - a \cos \theta)/(\sqrt{2}\sigma)) \\
 &= \sigma \sqrt{\frac{\pi}{2}} \frac{2}{\sqrt{\pi}} \int_{-\frac{r' - a \cos \theta}{\sqrt{2}\sigma}}^{\infty} e^{-t^2} dt = \sigma \sqrt{\frac{\pi}{2}} \left( 1 + \operatorname{erf}\left(\frac{a \cos \theta}{\sqrt{2}\sigma}\right) \right)
 \end{aligned}$$

よって

$$\rho(\Theta; a, \sigma) = \frac{1}{2\pi} \exp\left(-\frac{a^2}{2\sigma^2}\right) \left[ 1 + \frac{a \cos \theta}{\sigma} \sqrt{\frac{\pi}{2}} \exp\left(-\frac{a^2 \cos^2 \theta}{2\sigma^2}\right) \left( 1 + \operatorname{erf}\left(\frac{a \cos \theta}{\sqrt{2}\sigma}\right) \right) \right]$$

$\theta_0 = 0$  とは限らない一般の場合は分布を  $-\theta_0$  だけ並行移動したものとなるので、冒頭に掲げた  $\rho(\theta; \mathbf{a}, \sigma)$  の式が成り立つ。  $\square$

## 第 X.11 章

# 対数正規分布

### X.11.1 確率密度関数の導出

対数正規分布  $\Lambda(\mu, \sigma^2)$  の確率密度関数  $f$  は次式である。

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right)$$

確率変数  $X > 0$  の対数  $Y = \log X$  が期待値  $\mu$ 、分散  $\sigma$  の正規分布  $N(\mu, \sigma^2)$  に従うという条件から  $X$  の確率密度関数  $f_X$  を導出する。 $x_1, x_2 > 0$  とすると次式が成り立つ。

$$\Pr(x_1 \leq X < x_2) = \Pr(\log x_1 \leq Y < \log x_2) = \int_{\log x_1}^{\log x_2} f_Y(y) dy$$

ここに  $f_Y$  は  $N(\mu, \sigma^2)$  の確率密度関数である。 $f_X$  は次式となる。

$$\begin{aligned} f_X(x_2) &= \frac{\partial}{\partial x_2} \Pr(x_1 \leq X < x_2) = f_Y(\log x_2) \frac{d}{dx_2} \log x_2 = \frac{1}{x_2} f_Y(\log x_2) \\ &= \frac{1}{\sqrt{2\pi}\sigma x_2} \exp\left(-\frac{(\log x_2 - \mu)^2}{2\sigma^2}\right) \end{aligned}$$

### X.11.2 最頻値

対数正規分布  $\Lambda(\mu, \sigma^2)$  の最頻値は  $e^{\mu - \sigma^2}$  である。

*Proof.*

対数正規分布の確率密度関数  $f$  が 1 峰性であることは既知として、最頻値すなわち  $f$  の導関数が 0 となる点を求める。

$$\frac{d f_X(x)}{dx} = -\frac{1}{\sqrt{2\pi}\sigma x^2} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right) \left(1 + \frac{\log x - \mu}{\sigma^2}\right)$$

これが 0 となる  $x$  の値は  $e^{\mu - \sigma^2}$  である。

□

### X.11.3 期待値

対数正規分布  $\Lambda(\mu, \sigma^2)$  の期待値は  $e^{\mu+\sigma^2/2}$  である。

*Proof.*

$X \sim \Lambda(\mu, \sigma^2)$  とする。 $X$  の確率密度関数を  $f$  とする。

$$\begin{aligned} E[X] &= \int_{+0}^{\infty} x f(x) dx = \int_{+0}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right) dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y - \mu)^2}{2\sigma^2}\right) e^y dy \quad (\text{変数変換 } y = \log x) \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y - \mu - \sigma^2)^2}{2\sigma^2} + \mu + \sigma^2/2\right) dy \\ &= e^{\mu+\sigma^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y - \mu - \sigma^2)^2}{2\sigma^2}\right) dy = e^{\mu+\sigma^2/2} \end{aligned}$$

□

## 第 X.12 章

# $\chi^2$ 分布

### X.12.1 定義

関数  $f(x)$  を次のように定義する。

$$f(x) := \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} e^{-\frac{x}{2}} x^{\frac{n}{2}-1}$$

確率密度関数が上式で与えられるような分布を「自由度  $n$  の  $\chi^2$  分布」と呼び、 $\chi^2_n$  と書くことにする。

### X.12.2 $X \sim N(0, 1) \Rightarrow X^2 \sim \chi^2_1$

$$X \sim N(0, 1) \Rightarrow X^2 \sim \chi^2_1$$

*Proof.*

$X$  の確率密度関数を  $f^X(x)$  と書くことにすると

$$\begin{aligned} P(0 \leq Z \leq z) &= P(-\sqrt{z} \leq X \leq \sqrt{z}) = \int_{-\sqrt{z}}^{\sqrt{z}} f^X(x) dx = 2 \int_0^{\sqrt{z}} f^X(x) dx = \\ \therefore f^Z(z) &= \frac{d}{dz} P(0 \leq Z \leq z) = \left( \frac{d}{d\sqrt{z}} P(0 \leq Z \leq z) \right) \frac{d\sqrt{z}}{dz} = \frac{1}{\sqrt{z}} f^X(\sqrt{z}) = \frac{1}{\sqrt{z} \sqrt{2\pi}} e^{-\frac{z}{2}} \\ &= \frac{1}{2^{\frac{1}{2}} \Gamma(\frac{1}{2})} e^{-\frac{z}{2}} z^{\frac{1}{2}-1} = \text{pdf of } \chi^2_1 \end{aligned}$$

□

### X.12.3 $X_i \sim N(0, 1) \Rightarrow Z_n = \sum_{i=1}^n X_i^2 \sim \chi^2_n$

$$X_i \sim N(0, 1) \Rightarrow Z_n = \sum_{i=1}^n X_i^2 \sim \chi^2_n$$

*Proof.*

定理 1 より  $n = 1$  のとき成り立つ。  $n - 1$  まで成り立つと仮定して  $n$  のとき成り立つことを示せば良い。  
 $Z_n$  の確率密度関数を  $f^{Z_n}(z)$  と書くことにする。  $n - 1$  まで成り立つと仮定すると当然

$$f^{Z_{n-1}}(z) = \frac{1}{2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2}\right)} e^{-\frac{z}{2}} z^{\frac{n-1}{2}-1}$$

であるから

$$\begin{aligned} f^{Z_n}(z) &= \int_{x=0}^{\infty} \int_{y=0}^{\infty} \delta(x+y-z) f^{Z_{n-1}}(x) f^{Z_1}(y) dy dx = \int_{x=0}^{\infty} f^{Z_{n-1}}(x) \left[ \int_{y=0}^{\infty} \delta(x+y-z) f^{Z_1}(y) dy \right] dx \\ &= \int_{x=0}^z f^{Z_{n-1}}(x) f^{Z_1}(z-x) dx = \int_0^z \frac{e^{-\frac{x}{2}}}{2^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2}\right)} x^{\frac{n-1}{2}-1} \frac{e^{-\frac{z-x}{2}}}{2^{\frac{1}{2}} \Gamma\left(\frac{1}{2}\right)} (z-x)^{-\frac{1}{2}} dx \\ &= \frac{e^{-\frac{z}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{1}{2}\right)} \int_0^z x^{\frac{n-1}{2}-1} (z-x)^{\frac{n}{2}-1} dx \end{aligned}$$

$x = uz$  なる変数変換を行って

$$\begin{aligned} &\frac{e^{-\frac{z}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{1}{2}\right)} \int_0^1 (uz)^{\frac{n-1}{2}-1} z^{-\frac{1}{2}} (1-u)^{-\frac{1}{2}} z du \\ &= \frac{e^{-\frac{z}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{1}{2}\right)} z^{\frac{n}{2}-1} \int_0^1 u^{\frac{n-1}{2}-1} (1-u)^{\frac{1}{2}-1} du = \frac{e^{-\frac{z}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{1}{2}\right)} z^{\frac{n}{2}-1} B\left(\frac{n-1}{2}, \frac{1}{2}\right) \\ &= \frac{e^{-\frac{z}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{1}{2}\right)} z^{\frac{n}{2}-1} \frac{\Gamma\left(\frac{n-1}{2}\right) \Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \quad (\because \text{定理 1.3.1 : ベータ関数とガンマ関数の関係}) \\ &= \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} e^{-\frac{z}{2}} z^{\frac{n}{2}-1} = \text{pdf of } \chi^2_n \end{aligned}$$

□

$$\text{X.12.4} \quad X \sim \chi^2_n \quad \Rightarrow \quad E[X] = n, \quad V[x] = 2n$$

$$X \sim \chi^2_n \quad \Rightarrow \quad E[X] = n, \quad V[x] = 2n$$

*Proof.*

$X$  の確率密度関数を  $f(x)$  と書くことにする。

まず

$$E[X] = \int_0^{\infty} x f(x) dx = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} x^{\frac{n}{2}} e^{-\frac{x}{2}} dx$$

$\frac{x}{2} = u$  なる変数変換を行って

$$\begin{aligned} &\frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} (2u)^{\frac{n}{2}} e^{-u} 2 du = \frac{2}{\Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} u^{\frac{n}{2}} e^{-u} du = \frac{2}{\Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} u^{\left(\frac{n}{2}+1\right)-1} e^{-u} du \\ &= \frac{2\Gamma\left(\frac{n}{2}+1\right)}{\Gamma\left(\frac{n}{2}\right)} = \frac{2\frac{n}{2}\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} = n \end{aligned}$$

期待値が  $n$  になることが示された。分散を求めるために 2 乗の期待値も計算しておく。

$$E[X^2] = \int_0^{\infty} x^2 f(x) dx = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} x^{\frac{n}{2}+1} e^{-\frac{x}{2}} dx$$

$\frac{x}{2} = u$  なる変数変換を行って

$$\begin{aligned} \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \int_0^\infty (2u)^{\frac{n}{2}+1} e^{-u} 2du &= \frac{4}{\Gamma(\frac{n}{2})} \int_0^\infty u^{(\frac{n}{2}+2)-1} e^{-u} du \\ &= \frac{4}{\Gamma(\frac{n}{2})} \Gamma\left(\frac{n}{2} + 2\right) = \frac{4}{\Gamma(\frac{n}{2})} \left(\frac{n}{2} + 1\right) \left(\frac{n}{2}\right) \Gamma\left(\frac{n}{2}\right) = n^2 + 2n \end{aligned}$$

最後に分散は

$$V[X] = E[X^2] - E[X]^2 = n^2 + 2n - n^2 = 2n$$

□

## X.12.5 再生性

$$X \sim \chi^2_{n_X}, Y \sim \chi^2_{n_Y} \Rightarrow X + Y \sim \chi^2_{n_X + n_Y}$$

*Proof.*

$Z = X + Y$  とする。 $X, Y, Z$  の確率密度関数を  $f^X(x), f^Y(y), f^Z(z)$  とすると

$$\begin{aligned} f^Z(z) &= \frac{d}{dz} \int_{x=0}^z \int_{y=0}^{z-x} f^X(x) f^Y(y) dy dx = \int_0^z f^X(x) f^Y(z-x) dx \quad (\because III.24.4) \\ &= \int_0^z \frac{1}{2^{\frac{n_X}{2}} \Gamma(\frac{n_X}{2})} x^{\frac{n_X}{2}-1} e^{-\frac{x}{2}} \frac{1}{2^{\frac{n_Y}{2}} \Gamma(\frac{n_Y}{2})} (z-x)^{\frac{n_Y}{2}-1} e^{-\frac{z-x}{2}} dx \\ &= \frac{1}{2^{\frac{n_X+n_Y}{2}} \Gamma(\frac{n_X}{2}) \Gamma(\frac{n_Y}{2})} e^{-\frac{z}{2}} \int_0^z x^{\frac{n_X}{2}-1} (z-x)^{\frac{n_Y}{2}-1} dx \end{aligned}$$

積分値の係数が煩わしいのでこれを  $A$  と書くことにし、 $u = \frac{x}{z}$  なる変数変換を行って

$$\begin{aligned} f^Z(z) &= Az^{\frac{n_X+n_Y}{2}-1} \int_0^1 u^{\frac{n_X}{2}-1} (1-u)^{\frac{n_Y}{2}-1} du \\ &= Az^{\frac{n_X+n_Y}{2}-1} B\left(\frac{n_X}{2}, \frac{n_Y}{2}\right) \\ &= Az^{\frac{n_X+n_Y}{2}-1} \frac{\Gamma(\frac{n_X}{2}) \Gamma(\frac{n_Y}{2})}{\Gamma(\frac{n_X+n_Y}{2})} \\ &= \frac{1}{2^{\frac{n_X+n_Y}{2}} \Gamma(\frac{n_X}{2}) \Gamma(\frac{n_Y}{2})} e^{-\frac{z}{2}} z^{\frac{n_X+n_Y}{2}-1} \frac{\Gamma(\frac{n_X}{2}) \Gamma(\frac{n_Y}{2})}{\Gamma(\frac{n_X+n_Y}{2})} \\ &= \frac{1}{2^{\frac{n_X+n_Y}{2}} \Gamma(\frac{n_X+n_Y}{2})} z^{\frac{n_X+n_Y}{2}-1} e^{-\frac{z}{2}} \\ &= \text{pdf of } \chi^2_{n_X+n_Y} \end{aligned}$$

□

## X.12.6 $X_1, \dots, X_n \sim N(\mu, \sigma^2) \Rightarrow \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi^2_{n-1}$

$$X_1, \dots, X_n \sim N(\mu, \sigma^2) \text{ とし、} \bar{X} \text{ を標本平均とすると } \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi^2_{n-1}$$

*Proof.*

$Z_i := \frac{X_i - \mu}{\sigma}$  とすると  $Z_i \sim N(0, 1)$  であり

$$\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n (Z_i - \bar{Z})^2 \quad \text{where } \bar{Z} := \frac{1}{n} \sum_{i=1}^n Z_i \quad (1)$$

$\mathbf{Z} := [Z_1, \dots, Z_n]^\top$  とする。最下行が全て  $1/\sqrt{n}$  である  $n$  次直交行列を  $P$  として  $\mathbf{Y} := [Y_1, \dots, Y_n]^\top := P\mathbf{Z}$  とすると、[X.8.6](#) より、 $Y_1, \dots, Y_n \sim N(0, 1)$  であり、 $Y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i = \sqrt{n}\bar{Z} \therefore \bar{Z} = Y_n/\sqrt{n}$ 。

$$(1) = \sum_{i=1}^n (Z_i^2 + \bar{Z}^2 - 2\bar{Z}Z_i) = \|\mathbf{Z}\|^2 - n\bar{Z}^2 = \|\mathbf{Y}\|^2 - Y_n^2 = \sum_{i=1}^{n-1} Y_i^2$$

右から 2 つ目の等号で、直交変換がノルムを保存する性質を用いた。このように、標準正規分布に従う独立な  $n-1$  個の確率変数の 2 乗和で表されたので、[X.12.3](#) よりこれは自由度  $n-1$  のカイ 2 乗分布に従う。  $\square$

**解説.**

なぜ独立変数の数を減らせたのか考えてみよう。元々の  $\mathbf{Z}$  では  $Z_1, \dots, Z_n$  が一斉に同じ数だけ増えても  $\sum_{i=1}^n (Z_i - \bar{Z})^2$  の値は変わらない。これは、確率密度関数を求めるための重積分中の  $\mathbf{Z}$  の拘束領域が、ベクトル  $\mathbf{1}^n$  方向に伸び放題であるということだ。 $\mathbf{Z}$  はこの方向に無制限に動ける。上述の直交変換で  $Y_n$  がこの方向に割り当てられたため、確率密度関数を求めるための重積分の中のこの方向に関する  $(-\infty, \infty)$  の積分結果が  $Y_n$  の確率密度関数の全域積分すなわち 1 になり、我々が相手にすべき次元が 1 個減ったのである。



## 第 X.13 章

# $t$ 分布

### X.13.1 定義

関数  $f(t)$  を次のように定義する。

$$f(t) := \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$$

確率密度関数が上式で与えられるような分布を「自由度  $n$  の  $t$  分布」と呼び、 $t_n$  と書くことにする。

$$\text{X.13.2} \quad \text{X.13.2.0} \quad X, Y : \text{indep}, \quad X \sim N(0, 1), \quad Y \sim \chi^2_n \quad \Rightarrow \quad T := \frac{X}{\sqrt{\frac{Y}{2}}} \sim t_n$$

$$X, Y : \text{indep}, \quad X \sim N(0, 1), \quad Y \sim \chi^2_n \quad \Rightarrow \quad T := \frac{X}{\sqrt{\frac{Y}{2}}} \sim t_n$$

*Proof.*

$X, Y, T$  の確率密度関数を  $f^X(x), f^Y(y), f^T(t)$  と書くことにする。

$$\begin{aligned} P(-\infty \leq T \leq t) &= P\left(-\infty \leq X \leq t\sqrt{\frac{y}{n}}\right) = \int_0^\infty \int_{-\infty}^{t\sqrt{\frac{y}{n}}} f^Y(y) f^X(x) dx dy \\ &= \int_0^\infty f^Y(y) \left( \int_{-\infty}^{t\sqrt{\frac{y}{n}}} f^X(x) dx \right) dy \end{aligned}$$

$$\begin{aligned}
\therefore f^T(t) &= \frac{\partial}{\partial t} P(-\infty \leq T \leq t) = \int_0^\infty f^Y(y) f^X\left(t\sqrt{\frac{y}{n}}\right) \sqrt{\frac{y}{n}} dy \quad (\text{微分と積分の順序を交換した}) \\
&= \frac{1}{\sqrt{2n\pi} 2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \int_0^\infty y^{\frac{n}{2}-1} e^{-\frac{y}{2}} \exp\left(-\frac{t^2 y}{2n}\right) y^{\frac{1}{2}} dy \\
&= \frac{1}{\sqrt{2n\pi} 2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \int_0^\infty y^{\frac{n-1}{2}} \exp\left[-\frac{y}{2} \left(1 + \frac{t^2}{n}\right)\right] dy \\
&\quad \left(\frac{y}{2} \left(1 + \frac{t^2}{n}\right) = u \text{ なる変数変換を行って}\right) \\
&= \frac{1}{\sqrt{2n\pi} 2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \left(\frac{2}{1 + \frac{t^2}{n}}\right)^{\frac{n-1}{2}} \int_0^\infty u^{\frac{n-1}{2}} e^{-u} \frac{2}{1 + \frac{t^2}{n}} du \\
&= \frac{1}{\sqrt{2n\pi} 2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \left(\frac{2}{1 + \frac{t^2}{n}}\right)^{\frac{n-1}{2}} \int_0^\infty u^{\frac{n+1}{2}-1} e^{-u} du = \frac{1}{\sqrt{2n\pi} 2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \left(\frac{2}{1 + \frac{t^2}{n}}\right)^{\frac{n+1}{2}} \Gamma\left(\frac{n+1}{2}\right) \\
&= \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi} \Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} = \text{pdf of } t_n
\end{aligned}$$

□

### X.13.3 X.13.3.0t 分布の期待値

$t$  分布は自由度  $n \geq 2$  の場合に限り期待値が存在し  $E[T] = 0$  である。

*Proof.*

$tf(t)$  が  $t \rightarrow \infty$  で  $\frac{1}{t}$  オーダーより小さくなる条件が  $n \geq 2$  であることは式を見れば容易にわかる。この時に限り  $\int_0^\infty tf(t)dt$  および  $\int_{-\infty}^0 tf(t)dt$  が存在し  $E[T] = \int_{-\infty}^\infty tf(t)dt$  が存在する。被積分関数が奇関数であることから積分値は明らかに 0 である。

□

### X.13.4 X.13.4.0t 分布の分散

$t$  分布の分散は自由度  $n \leq 2$  の場合  $\infty$ 、自由度  $n \geq 3$  の場合  $\frac{n}{n-2}$

*Proof.*

$n \leq 2$  の場合、 $t \rightarrow \pm\infty$  で  $t^2 f(t)$  が  $\frac{1}{t}$  オーダーを下回らないから  $\int_{-\infty}^\infty t^2 f(t)dt = \infty$   
 $n \geq 3$  の場合、積分は収束する。このとき分散が  $\frac{n}{n-2}$  であることを示す。  
確率密度関数は偶関数だから

$$\int_{-\infty}^\infty t^2 f(t)dt = 2 \int_0^\infty t^2 f(t)dt = \frac{2\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} \int_0^\infty t^2 \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt \quad (1)$$

先に積分の値を計算する。 $t^2 = u$  なる変数変換を行って

$$\int_0^\infty t^2 \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} dt = \int_0^\infty u \left(1 + \frac{u}{n}\right)^{-\frac{n+1}{2}} \frac{1}{2\sqrt{u}} du = \frac{1}{2} \int_0^\infty u^{\frac{1}{2}} \left(1 + \frac{u}{n}\right)^{-\frac{n+1}{2}} du$$

次に  $v = \frac{\frac{u}{n}}{1 + \frac{u}{n}} = \frac{u}{n+u}$  なる変数変換を行う。  $u = \frac{nv}{1-v}$  となって  $du = \frac{n}{(1-v)^2} dv$  となる。積分範囲は  $v = 0 \rightarrow 1$  になる。

(※ (1) で積分の左側に括りだした定数群にガンマ関数が居る。最終的にこれらが消えてほしいわけだから積分結果がガンマ関数あるいはベータ関数にならなければならない。そこでベータ関数の出現を期待して積分範囲を  $0 \rightarrow \infty$  から  $0 \rightarrow 1$  に変更するような変数変換を行ったのである。)

上式は

$$\begin{aligned} \frac{1}{2} \int_0^1 \left( \frac{nv}{1-v} \right)^{\frac{1}{2}} \left( 1 + \frac{v}{1-v} \right)^{-\frac{n+1}{2}} \frac{n}{(1-v)^2} dv &= \frac{n\sqrt{n}}{2} \int_0^1 \sqrt{v}(1-v)^{\frac{n}{2}-2} dv \\ &= \frac{n\sqrt{n}}{2} \int_0^1 v^{\frac{3}{2}-1} (1-v)^{\frac{n-2}{2}-1} dv = \frac{n\sqrt{n}}{2} B\left(\frac{3}{2}, \frac{n-2}{2}\right) = \frac{n\sqrt{n}}{2} \frac{\Gamma(\frac{3}{2})\Gamma(\frac{n-2}{2})}{\Gamma(\frac{n+1}{2})} \end{aligned}$$

ここで

$$\Gamma\left(\frac{3}{2}\right) = \Gamma\left(\frac{1}{2} + 1\right) = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}$$

および

$$\Gamma\left(\frac{n-2}{2}\right) = \Gamma\left(\frac{n}{2} - 1\right)$$

であり

$$\Gamma\left(\frac{n}{2}\right) = \Gamma\left(\frac{n}{2} - 1 + 1\right) = \left(\frac{n}{2} - 1\right) \Gamma\left(\frac{n}{2} - 1\right)$$

であるから

$$\Gamma\left(\frac{n}{2} - 1\right) = \frac{2\Gamma(\frac{n}{2})}{n-2}$$

以上の結果を (1) に適用して

$$\int_{-\infty}^{\infty} t^2 f(t) dt = \frac{n}{n-2}$$

最後に

$$V[T] = E[T^2] - E[T]^2 = \int_{-\infty}^{\infty} t^2 f(t) dt - 0 = \frac{n}{n-2}$$

□

## 第 X.14 章

# F 分布

### X.14.1 定義

関数  $f(f)$  を次のように定義する。

$$f(f) = \frac{\Gamma\left(\frac{\nu_1 + \nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} f^{\frac{\nu_1 - 2}{2}} \left(1 + \frac{\nu_1}{\nu_2} f\right)^{-\frac{\nu_1 + \nu_2}{2}}$$

where  $\nu_1, \nu_2 \in \mathbb{N}, \quad f > 0$

確率密度関数が上式で与えられるような分布を「自由度  $(\nu_1, \nu_2)$  の  $F$  分布」と呼び、 $F_{\nu_2}^{\nu_1}$  と書くことにする。

$$\text{X.14.2} \quad X_1, X_2 : \text{indep}, \quad X_1 \sim \chi^2_{\nu_1}, \quad X_2 \sim \chi^2_{\nu_2} \quad \Rightarrow \quad F := \frac{\frac{X_1}{\nu_1}}{\frac{X_2}{\nu_2}} \sim F_{\nu_2}^{\nu_1}$$

$$X_1, X_2 : \text{indep}, \quad X_1 \sim \chi^2_{\nu_1}, \quad X_2 \sim \chi^2_{\nu_2} \quad \Rightarrow \quad F := \frac{\frac{X_1}{\nu_1}}{\frac{X_2}{\nu_2}} \sim F_{\nu_2}^{\nu_1}$$

*Proof.*

$X_1, X_2, F$  の確率密度関数を  $f^{X_1}(x_1), f^{X_2}(x_2), f(f)$  と書くことにする。

$$\begin{aligned} P(0 \leq F \leq f) &= P\left(0 \leq X_1 \leq \frac{\nu_1}{\nu_2} x_2 f\right) = \int_{x_2=0}^{\infty} \int_{x_1=0}^{\frac{\nu_1}{\nu_2} x_2 f} f^{X_1}(x_1) f^{X_2}(x_2) dx_1 dx_2 \\ &= \int_{x_2=0}^{\infty} f^{X_2}(x_2) \left( \int_{x_1=0}^{\frac{\nu_1}{\nu_2} x_2 f} f^{X_1}(x_1) dx_1 \right) dx_2 \end{aligned}$$

$$\begin{aligned} \therefore f(f) &= \frac{\partial}{\partial f} P(0 \leq F \leq f) = \int_{x_2=0}^{\infty} f^{X_2}(x_2) f^{X_1} \left( \frac{\nu_1}{\nu_2} x_2 f \right) \frac{\nu_1}{\nu_2} x_2 dx_2 \quad (\text{微分と積分の順序を交換した}) \\ &= \frac{1}{2^{\frac{\nu_2}{2}} \Gamma\left(\frac{\nu_2}{2}\right)} \frac{\nu_1}{\nu_2} \int_0^{\infty} x^{\frac{\nu_2}{2}-1} e^{-\frac{x}{2}} \left(\frac{\nu_1}{\nu_2} f\right)^{\frac{\nu_1}{2}-1} x^{\frac{\nu_1}{2}-1} \exp\left[-\frac{x}{2} \left(1 + \frac{\nu_1}{\nu_2} f\right)\right] dx \\ &= \frac{1}{2^{\frac{\nu_1 + \nu_2}{2}} \Gamma\left(\frac{\nu_2}{2}\right) \Gamma\left(\frac{\nu_1}{2}\right)} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} f^{\frac{\nu_1 - 2}{2}} \int_0^{\infty} x^{\frac{\nu_1 + \nu_2}{2} - 1} \exp\left[-\frac{x}{2} \left(1 + \frac{\nu_1}{\nu_2} f\right)\right] dx \end{aligned} \quad (1)$$

先に積分の値を計算する。 $\frac{x}{2} \left(1 + \frac{\nu_1}{\nu_2} f\right) = u$  なる変数変換を行って

$$\begin{aligned} \int_0^\infty x^{\frac{\nu_1+\nu_2}{2}-1} \exp\left[-\frac{x}{2} \left(1 + \frac{\nu_1}{\nu_2} f\right)\right] dx &= \int_0^\infty \left(\frac{2}{1 + \frac{\nu_1}{\nu_2} f} u\right)^{\frac{\nu_1+\nu_2}{2}-1} e^{-u} \frac{2}{1 + \frac{\nu_1}{\nu_2} f} du \\ &= \left(1 + \frac{\nu_1}{\nu_2} f\right)^{-\frac{\nu_1+\nu_2}{2}} 2^{\frac{\nu_1+\nu_2}{2}} \int_0^\infty u^{\frac{\nu_1+\nu_2}{2}-1} e^{-u} du = \left(1 + \frac{\nu_1}{\nu_2} f\right)^{-\frac{\nu_1+\nu_2}{2}} 2^{\frac{\nu_1+\nu_2}{2}} \Gamma\left(\frac{\nu_1+\nu_2}{2}\right) \end{aligned}$$

これを (2) に適用して

$$f(f) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} f^{\frac{\nu_1-2}{2}} \left(1 + \frac{\nu_1}{\nu_2} f\right)^{-\frac{\nu_1+\nu_2}{2}} = \text{pdf of } F_{\nu_2}^{\nu_1}$$

□

### X.14.3 自由度 $(\nu_1, \nu_2)$ の $F$ 分布 $F_{\nu_2}^{\nu_1}$ は $\nu_2 \geq 3$ の時に限り期待値が定義できてその値は $\frac{\nu_2}{\nu_2-2}$ である。

自由度  $(\nu_1, \nu_2)$  の  $F$  分布  $F_{\nu_2}^{\nu_1}$  は  $\nu_2 \geq 3$  の時に限り期待値が定義できてその値は  $\frac{\nu_2}{\nu_2-2}$  である。

*Proof.*

$\nu_2 \leq 2$  の場合、 $x \rightarrow \infty$  で  $xf(x)$  が  $\frac{1}{x}$  オーダーを下回らないので  $\int_0^\infty xf(x)dx$  は発散し期待値は定義できない。 $\nu_2 \geq 3$  の場合、 $x \rightarrow \infty$  で  $xf(x)$  が  $\frac{1}{x}$  オーダーを下回るので積分は収束して期待値が定義できる。この値が  $\frac{\nu_2}{\nu_2-2}$  であることを示す。

$$E[F] = \int_0^\infty xf(x)dx = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} \int_0^\infty x^{\frac{\nu_2}{2}} \left(1 + \frac{\nu_1}{\nu_2} x\right)^{-\frac{\nu_1+\nu_2}{2}} dx \quad (1)$$

積分の値を先に計算する。 $u = \frac{\frac{\nu_1}{\nu_2} x}{1 + \frac{\nu_1}{\nu_2} x} = \frac{\nu_1 x}{\nu_2 + \nu_1 x}$  なる変数変換を行うと  $x = \frac{\nu_2}{\nu_1} \frac{u}{1-u}$ ,  $dx = \frac{\nu_2}{\nu_1} \frac{1}{(1-u)^2} du$  となり積分範囲は  $u = 0 \rightarrow 1$  になる。

(※ (1) で積分の左側に括りだした定数群にガンマ関数が居る。最終的にこれらが消えてほしいわけだから積分結果がガンマ関数あるいはベータ関数にならなければならない。そこでベータ関数の出現を期待して積分範囲を  $0 \rightarrow \infty$  から  $0 \rightarrow 1$  に変更するような変数変換を行ったのである。)

(1) の積分は

$$\begin{aligned} & \int_0^1 \left(\frac{\nu_2}{\nu_1}\right)^{\frac{\nu_1}{2}} \left(\frac{u}{1-u}\right)^{\frac{\nu_1}{2}} \left(1 + \frac{u}{1-u}\right)^{-\frac{\nu_1+\nu_2}{2}} \frac{\nu_2}{\nu_1} \frac{1}{(1-u)^2} du \\ &= \left(\frac{\nu_1}{\nu_2}\right)^{-\frac{\nu_1}{2}-1} \int_0^1 u^{\frac{\nu_1}{2}} (1-u)^{\frac{\nu_2}{2}-2} du = \left(\frac{\nu_1}{\nu_2}\right)^{-\frac{\nu_1}{2}-1} \int_0^1 u^{\frac{\nu_1}{2}+1-1} (1-u)^{\frac{\nu_2}{2}-1-1} du \\ &= \left(\frac{\nu_1}{\nu_2}\right)^{-\frac{\nu_1}{2}-1} B\left(\frac{\nu_1}{2} + 1, \frac{\nu_2}{2} - 1\right) = \left(\frac{\nu_1}{\nu_2}\right)^{-\frac{\nu_1}{2}-1} \frac{\Gamma\left(\frac{\nu_1}{2} + 1\right)\Gamma\left(\frac{\nu_2}{2} - 1\right)}{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)} \\ &= \left(\frac{\nu_1}{\nu_2}\right)^{-\frac{\nu_1}{2}-1} \frac{\frac{\nu_1}{2}\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}{\left(\frac{\nu_2}{2} - 1\right)\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)} \end{aligned}$$

これを (1) に適用して

$$E[F] = \frac{\nu_2}{\nu_2 - 2}$$

□

**X.14.4 自由度  $(\nu_1, \nu_2)$  の  $F$  分布  $F_{\nu_2}^{\nu_1}$  の分散は  $\nu_2 \leq 4$  の場合  $\infty$ 、 $\nu_2 \geq 5$  の場合  $\frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}$  である。**

自由度  $(\nu_1, \nu_2)$  の  $F$  分布  $F_{\nu_2}^{\nu_1}$  の分散は  $\nu_2 \leq 4$  の場合  $\infty$ 、 $\nu_2 \geq 5$  の場合  $\frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}$  である。

*Proof.*

$\nu_2 \leq 4$  の場合、 $x \rightarrow \infty$  で  $x^2 f(x)$  が  $\frac{1}{x}$  オーダーを下回らないので  $\int_0^\infty x^2 f(x) dx = \infty$   
 $\nu_2 \geq 5$  の場合、 $x \rightarrow \infty$  で  $x^2 f(x)$  が  $\frac{1}{x}$  オーダーを下回るので積分は収束する。このとき分散が  $\frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}$  であることを示す。  
 $V[F] = E[F^2] - E[F]^2$  であり第 2 項は定理 2 より既知である。第 1 項を計算する。

$$E[F^2] = \int_0^\infty x^2 f(x) dx = \frac{\Gamma(\frac{\nu_1 + \nu_2}{2})}{\Gamma(\frac{\nu_1}{2})\Gamma(\frac{\nu_2}{2})} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} \int_0^\infty x^{\frac{\nu_2}{2}+1} \left(1 + \frac{\nu_1}{\nu_2}x\right)^{-\frac{\nu_1 + \nu_2}{2}} dx \quad (1)$$

積分の値を先に計算する。 $u = \frac{\frac{\nu_1}{\nu_2}x}{1 + \frac{\nu_1}{\nu_2}x} = \frac{\nu_1 x}{\nu_2 + \nu_1 x}$  なる変数変換を行うと  $x = \frac{\nu_2}{\nu_1} \frac{u}{1-u}$ ,  $dx = \frac{\nu_2}{\nu_1} \frac{1}{(1-u)^2} du$  となり積分範囲は  $u = 0 \rightarrow 1$  になる。

(1) の積分は

$$\begin{aligned} & \int_0^1 \left(\frac{\nu_2}{\nu_1}\right)^{\frac{\nu_1}{2}+1} \left(\frac{u}{1-u}\right)^{\frac{\nu_1}{2}+1} \left(1 + \frac{u}{1-u}\right)^{-\frac{\nu_1 + \nu_2}{2}} \frac{\nu_2}{\nu_1} \frac{1}{(1-u)^2} du \\ &= \left(\frac{\nu_1}{\nu_2}\right)^{-\frac{\nu_1}{2}-2} \int_0^1 u^{\frac{\nu_1}{2}+1} (1-u)^{\frac{\nu_2}{2}-3} du = \left(\frac{\nu_1}{\nu_2}\right)^{-\frac{\nu_1}{2}-2} B\left(\frac{\nu_1}{2} + 2, \frac{\nu_2}{2} - 2\right) \\ &= \left(\frac{\nu_1}{\nu_2}\right)^{-\frac{\nu_1}{2}-2} \frac{\Gamma(\frac{\nu_1}{2} + 2)\Gamma(\frac{\nu_2}{2} - 2)}{\Gamma(\frac{\nu_1 + \nu_2}{2})} \\ &= \left(\frac{\nu_1}{\nu_2}\right)^{-\frac{\nu_1}{2}-2} \frac{(\frac{\nu_1}{2} + 1) \frac{\nu_1}{2} \Gamma(\frac{\nu_1}{2})}{\Gamma(\frac{\nu_1 + \nu_2}{2})} \frac{4\Gamma(\frac{\nu_2}{2})}{(\nu_2 - 4)(\nu_2 - 2)} \end{aligned}$$

これを (1) に適用して

$$E[F^2] = \frac{\nu_2^2(\nu_1 + 2)}{\nu_1(\nu_2 - 4)(\nu_2 - 2)}$$

定理 2 で示された  $E[F]$  を用いて

$$V[F] = E[F^2] - E[F]^2 = \frac{\nu_2^2(\nu_1 + 2)}{\nu_1(\nu_2 - 4)(\nu_2 - 2)} - \frac{\nu_2^2}{(\nu_2 - 2)^2} = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}$$

□

## 第 X.15 章

# 推定

### X.15.1 標本分散

$$\text{X.15.1.1} \quad E \left[ \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \right] = \frac{n-1}{n} \sigma^2$$

$X$  が平均  $\mu$ , 分散  $\sigma^2$  の母集団に属しているとき、そこからランダムに抽出した標本  $(X_1, X_2, \dots, X_n)$  の標本分散  $S^2 := \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$  の期待値は  $E(S^2) = \frac{n-1}{n} \sigma^2$  である。

*Proof.*

$$E[S^2] = E \left[ \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \right] = \frac{1}{n} \sum_{i=1}^n E[(X_i - \bar{X})^2] = \frac{1}{n} \sum_{i=1}^n \left( E[X_i^2] + E[\bar{X}^2] - 2E[\bar{X}X_i] \right) \quad (1)$$

然るに

$$\text{Var}[X_i] = E[X_i^2] - E^2[X_i] \text{ より}$$

$$E[X_i^2] = \text{Var}[X_i] + E^2[X_i] = \sigma^2 + \mu^2 \quad (2)$$

$$\begin{aligned} E[\bar{X}^2] &= \frac{1}{n^2} E \left[ \left( \sum_{i=1}^n X_i \right)^2 \right] = \frac{1}{n^2} E \left[ \sum_{i=1}^n X_i^2 + \sum_{i \neq j} X_i X_j \right] = \frac{1}{n^2} \sum_{i=1}^n E[X_i^2] + \frac{1}{n^2} \sum_{i \neq j} E[X_i X_j] \\ &= \frac{1}{n^2} \sum_{i=1}^n (\sigma^2 + \mu^2) + \frac{1}{n^2} \sum_{i \neq j} E[X_i] E[X_j] = \frac{1}{n^2} \sum_{i=1}^n (\sigma^2 + \mu^2) + \frac{1}{n^2} \sum_{i \neq j} \mu^2 \\ &= \frac{1}{n} (\sigma^2 + \mu^2) + \frac{1}{n^2} (n^2 - n) \mu^2 = \frac{1}{n} \sigma^2 + \mu^2 \end{aligned} \quad (3)$$

$$\begin{aligned} E[\bar{X}X_i] &= \frac{1}{n} E[(X_1 + X_2 + \dots + X_n)X_i] = \frac{1}{n} E \left[ X_i^2 + \sum_{j=1, \neq i}^n X_j X_i \right] \\ &= \frac{1}{n} E[X_i^2] + \frac{1}{n} \sum_{j=1, \neq i}^n E[X_j X_i] = \frac{1}{n} (\sigma^2 + \mu^2) + \frac{1}{n} \sum_{j=1, \neq i}^n E[X_j] E[X_i] \\ &= \frac{1}{n} (\sigma^2 + \mu^2) + \frac{1}{n} \sum_{j=1, \neq i}^n \mu^2 = \frac{1}{n} (\sigma^2 + \mu^2) + \frac{1}{n} (n-1) \mu^2 = \frac{1}{n} \sigma^2 + \mu^2 \end{aligned} \quad (4)$$

(2),(3),(4) を (1) に適用して

$$E[\sigma^2] = \frac{1}{n} \sum_{i=1}^n \left( \sigma^2 - \frac{1}{n} \sigma^2 \right) = \frac{n-1}{n} \sigma^2$$

□



## 第 XI 部

# ベイズ統計

## 第 XI.1 章

# 事前分布と事後分布の考え方

### XI.1.1 離散型

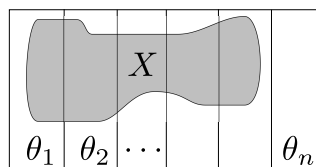


図 XI.1.1 事象  $\theta_i$  と  $X$

巷で使われている式は次のようにして導かれる。

$$\Pr(\theta_i|X) = \frac{\Pr(\theta_i \cap X)}{\Pr(X)} = \frac{\Pr(X|\theta_i) \Pr(\theta_i)}{\Pr(X)} = \frac{\Pr(X|\theta_i) \Pr(\theta_i)}{\sum_{i=1}^n \Pr(X \cap \theta_i)} = \frac{\Pr(X|\theta_i) \Pr(\theta_i)}{\sum_{i=1}^n \Pr(X|\theta_i) \Pr(\theta_i)}$$

### XI.1.2 連続型

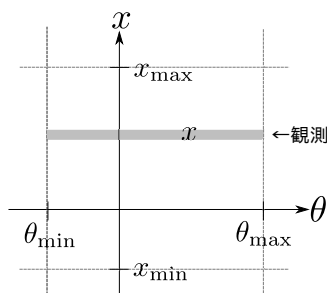


図 XI.1.1 事象  $\theta$  と  $x$  の同時確率密度

巷で使われている式を導出する。連続型の場合は少し厄介だ。まずいくつか定義を行う。

- $p(x, \theta)$ :  $x, \theta$  の同時確率密度関数。  $\int_{x_{\min}}^{x_{\max}} \int_{\theta_{\min}}^{\theta_{\max}} p(x, \theta) d\theta dx = 1$
- $\pi(\theta) := p_{\Theta}(\theta) := \int_{x_{\min}}^{x_{\max}} p(x, \theta) dx$  :  $\Theta$  の周辺確率密度。これが所謂「事前分布」。
- $f(x) := p_X(x) := \int_{\theta_{\min}}^{\theta_{\max}} p(x, \theta) d\theta$  :  $X$  の周辺確率密度。

- $f(x|\theta) := p_X(x|\Theta = \theta) = \frac{p(x,\theta)}{\pi(\theta)}$  :  $X$  の条件付き確率密度。これが所謂「**尤度**」。  $\int_{x_{\min}}^{x_{\max}} f(x|\theta)dx = 1$
- $\pi(\theta|x) := p_{\Theta}(\theta|X = x) = \frac{p(x,\theta)}{f(x)}$  :  $\Theta$  の条件付き確率密度。これが所謂「**事後分布**」。  
 $\int_{\theta_{\min}}^{\theta_{\max}} \pi(\theta|x)d\theta = 1$

以上の定義の下で、巷で使われている式が導かれる。

$$\pi(\theta|x) = \frac{p(x,\theta)}{f(x)} = \frac{f(x|\theta)\pi(\theta)}{\int_{\theta_{\min}}^{\theta_{\max}} p(x,\theta)d\theta} = \frac{f(x|\theta)\pi(\theta)}{\int_{\theta_{\min}}^{\theta_{\max}} f(x|\theta)\pi(\theta)d\theta}$$

上式から分かるように、尤度  $f(x|\theta)$  には定数倍の任意性があるので、単純に  $f(x|\theta) := p(x,\theta)$  としても上式は成り立つ。

実際の計算では、 $\pi(\theta)$  と  $f(x|\theta)$  がわかれば、その積を  $\theta$  に関して規格化すれば  $\pi(\theta|x)$  が求まる。

ベイズ推定では  $\pi(\theta)$  を我々が勝手に決める。そして  $\Theta = \theta$  というパラメータの下で母集団からのサンプルが従う分布  $f(x|\theta)$  を計算して先述の定義式の4つ目から  $p(x,\theta)$  を逆算し、最後に事後分布を得ている。

## 第 XI.2 章

# ベイズ推定量

結論から言うと、**ベイズ推定量**は**期待損失**を最小化する推定量として定義される。以下で、期待損失の定義とベイズ推定量の導出を行う。まずいくつか定義を行う。

- $\Omega_X$ : 標本空間。ここからの標本  $X$  は、我々が推定したいパラメータ  $\Theta$  の値  $\theta$  により決まる分布に従う。
- $\Omega_\Theta$ : パラメータ  $\Theta$  が属する確率空間。
- $p(x, \theta)$ :  $X$  と  $\Theta$  の同時確率密度
- $\delta: \Omega_X \rightarrow \mathbb{R}$ : 推定関数。標本  $X$  を入力すると  $\theta$  の推定値を返す。
- $L(\theta, x): \mathbb{R}^2 \rightarrow \mathbb{R}_+$ : 損失関数。 $\theta$  の推定値  $x$  に対して「損失」を返す。通常、 $\theta$  と  $x$  の差が大きいほど損失も大きくなるような損失関数を想定する。

以上の準備の下で、事前分布  $\pi(\theta)$  と推定関数  $\delta$  に対する期待損失  $E_{\Theta, X}[L(\Theta, \delta(X))]$  は次のように定義される。

$$E_{\Theta, X}[L(\Theta, \delta(X))] := \int_{\Omega_X} \int_{\Omega_\Theta} L(\theta, \delta(x)) p(x, \theta) d\theta dx$$

損失関数が  $\theta$  の凸関数であれば、上式で定義された期待損失を最小化する推定関数は事後分布の期待値と一致することを示す。期待損失を最小化するためには、各  $x$  に対して  $\int_{\Omega_\Theta} L(\theta, \delta(x)) p(x, \theta) d\theta$  を最小化する  $\delta(x)$  を構成すれば良い。Jensen の不等式を使うために、 $p(x, \theta)$  を  $\theta$  に関して規格化した  $\frac{p(x, \theta)}{f(x)}$  を使うと、 $\int_{\Omega_\Theta} L(\theta, \delta(x)) p(x, \theta) d\theta$  の最小化は  $\int_{\Omega_\Theta} L(\theta, \delta(x)) \frac{p(x, \theta)}{f(x)} d\theta$  の最小化と等価である。Jensen の不等式より

$$\int_{\Omega_\Theta} L(\theta, \delta(x)) \frac{p(x, \theta)}{f(x)} d\theta \geq L\left(\int_{\Omega_\Theta} \theta \frac{p(x, \theta)}{f(x)} d\theta, \delta(x)\right)$$

常識的な損失関数であれば  $\delta(x) = \int_{\Omega_\Theta} \theta \frac{p(x, \theta)}{f(x)} d\theta$  のときに損失が最小になり、[XI.1.2](#) でやったように、

$$\int_{\Omega_\Theta} \theta \frac{p(x, \theta)}{f(x)} d\theta = \int_{\Omega_\Theta} \theta \pi(\theta|x) d\theta = E_{\Theta \sim \pi(\Theta|x)}[\Theta]$$

となる。これは、事後分布の平均値を使えば期待損失を最小化できることを主張している。そして、期待損失を最小化する推定量、すなわち事後分布の平均値を**ベイズ推定量**と呼ぶ。

## 第 XI.3 章

# 事前分布と事後分布の例

### XI.3.1 正規分布の事後分布

平均  $\mu$ , 分散  $\sigma^2$  の標準正規分布に従う母集団から  $n$  個の標本を抽出し、標本平均を  $\bar{x}$  とする。母平均  $\mu$  の事前分布として平均  $\eta$ , 分散  $\tau^2$  の正規分布を用いると  $\mu$  の事後分布  $\pi(\theta|\bar{x})$  は

$$\text{平均: } \frac{n\tau^2\bar{x} + \sigma^2\eta}{n\tau^2 + \sigma^2}, \quad \text{分散: } \frac{\tau^2\sigma^2}{n\tau^2 + \sigma^2}$$

の正規分布になる。

*Proof.*

今、 $\pi(\theta) = \frac{1}{\sqrt{2\pi\tau^2}} \exp\left(-\frac{(\theta-\eta)^2}{2\tau^2}\right)$ ,  $f(\bar{x}|\theta) = \frac{1}{\sqrt{2\pi\frac{\sigma^2}{n}}} \exp\left(-\frac{(\bar{x}-\theta)^2}{2\frac{\sigma^2}{n}}\right)$  である。両者の積を  $\theta$  に関して規格化すれば求めたい事後分布が得られる。よって両者の積に現れる定数因子はどうでもいいから、指数部分のみに着目すると

$$\begin{aligned} & \exp\left(-\frac{(\bar{x}-\theta)^2}{2\frac{\sigma^2}{n}} - \frac{(\theta-\eta)^2}{2\tau^2}\right) \\ &= \exp\left\{\frac{-1}{2\left(\frac{\sigma^2\tau^2}{n}\right)} \left[\left(\tau^2 + \frac{\sigma^2}{n}\right)\theta^2 - 2\left(\tau^2\bar{x} + \frac{\sigma^2}{n}\eta\right)\theta + \left(\tau^2\bar{x}^2 + \frac{\sigma^2}{n}\eta^2\right)\right]\right\} \end{aligned}$$

上式の  $\exp()$  中の  $\left(\tau^2\bar{x}^2 + \frac{\sigma^2}{n}\eta^2\right)$  は定数因子になるから無視すると

$$\begin{aligned} \text{exp の引数} &= \frac{-1}{2\left(\frac{\sigma^2\tau^2}{n}\right)} \left[\left(\tau^2 + \frac{\sigma^2}{n}\right)\theta^2 - 2\left(\tau^2\bar{x} + \frac{\sigma^2}{n}\eta\right)\theta\right] \\ &= -\frac{1}{2} \times \frac{\tau^2 + \frac{\sigma^2}{n}}{\frac{\sigma^2\tau^2}{n}} \left[\left(\theta - \frac{\tau^2\bar{x} + \frac{\sigma^2}{n}\eta}{\tau^2 + \frac{\sigma^2}{n}}\right)^2 - \underbrace{\left(\frac{\tau^2\bar{x} + \frac{\sigma^2}{n}\eta}{\tau^2 + \frac{\sigma^2}{n}}\right)^2}_{\text{定数倍にしか関係ないので無視}}\right] \\ &= -\frac{1}{2} \left(\frac{\tau^2\sigma^2}{n\tau^2 + \sigma^2}\right)^{-1} \left(\theta - \frac{n\tau^2\bar{x} + \sigma^2\eta}{n\tau^2 + \sigma^2}\right)^2 \end{aligned}$$

□

## 第 XII 部

# グラフ理論

## 第 XII.1 章

# 定義

- $n$  個の頂点の集合  $\nu = \{i_1, i_2, \dots, i_n\}$  と辺集合  $\varepsilon \in \nu \times \nu$  からなるグラフ  $G$  を  $G(\nu, \varepsilon)$  と表記する。 $\nu$  の要素数を  $|\nu|$  と表記する ( $|\nu| = n$ )。
- 歩道, 道, 小道 : [\[6\]](#) の定義に従う。

## 第 XII.2 章

# 連結グラフ

### XII.2.1 諸定理

#### XII.2.1.1 $n$ 頂点の連結グラフは少なくとも $n - 1$ 本の辺をもつ

*Proof.*

$n$  に関する帰納法で示す。 $n = 2$  の時は明らか。 $n - 1$  まで成り立つと仮定して  $n$  のとき成り立つことを示す。

グラフから辺を任意に 1 本ずつ取り去ってゆき、 $c$  本目で初めて非連結になったとする。このときグラフは 2 つの連結成分  $A, B$  に分かれる (3 つ以上に分かれることはない。3 つ以上の連結成分を 1 本の辺で橋渡しするなど不可能である)。両者の点数をそれぞれ  $a, n - a$  ( $1 \leq a \leq n - 1$ ) とすると帰納法の仮定より  $A$  は  $a - 1$  本以上の、 $B$  は  $n - a - 1$  本以上の辺をそれぞれもつ。よって元々のグラフは  $(a - 1) + (n - a - 1) + c = n - 2 + c \geq n - 1$  本以上の辺をもっていたことになる。□

#### XII.2.1.2 任意の 2 つの頂点間の最短距離が 2 である無向グラフは少なくとも $n - 1$ 本の辺をもつ

*Proof.*  $n = 2, 3$  の場合は明らかである。以下  $n \geq 4$  とする。無向グラフ  $G$  が  $n$  個の点を持ち、任意の頂点間の最短距離が 2 であり、 $G$  から 1 つでも辺を取り去ると、ある 2 点間の最短距離が 3 以上になってしまうとする。 $G$  から任意の点を 1 つ選んで  $v_0$  とする。 $v_0$  に隣接しているから成る集合を  $S_0$  とする。 $S_0$  に属していない点のうち、「 $S_0$  中にある点が存在して、それと隣接している」全ての点について、その点から  $S_0$  に繋がる辺 1 本の接続先を  $v_0$  に変更する。それらの点を  $S_0$  に加えた集合を  $S_1$  とする。以降同様に  $S_2, S_3, \dots, \dots$  としてゆくと、 $v_0$  以外の全ての点が  $v_0$  に隣接する。そして  $v_0$  以外のいくつかの点同士の間にも辺が存在し得る。以上の操作は  $G$  の辺の数を保つ。 $v_0$  とそれ以外の点を結ぶ辺のみを残して他の辺を除去したグラフは、任意の 2 点間の最短距離が 2 である。以上より、定理の主張が成り立つ。□

#### XII.2.1.3 $n$ 頂点グラフが $\frac{1}{2}(n - 1)(n - 2)$ 本より多くの辺を持つなら連結である

*Proof.*

$n$  に関する帰納法で示す。 $n = 1, 2$  のときは明らかに成り立つ。 $n = m - 1$  ( $m \geq 3$ ) のとき成り立つものと



して  $n = m$  のとき成り立つことを示す。

$n = m$  のときのグラフを  $G$  とする。辺の数  $k$  について  $k > \frac{1}{2}(m-1)(m-2)$  という前提のもとで仮に  $G$  が非連結であるとする。 $k > 0$  なので次数が 1 以上の点が存在するので任意に 1 つ選んで  $v$  とし、 $G$  から  $v$  とそれに接続する辺を取り去ったグラフを  $G'$  とするとこれは非連結である ( $\cdots$ ①)。(※  $\deg(v) \geq 1$  であるところがミソ。 $\deg(v) = 0$  だと、もし不運なことに  $v$  が唯一の孤立点だったら  $G'$  が連結になってしまう。)  $G'$  は非連結であるから  $\deg(v) \leq m-2$  である (もし  $\deg(v) = m-1$  なら  $v$  は他の全ての頂点と隣接していることになり、 $G$  が連結になってしまう)。 $G'$  の辺の数は  $k' = k - \deg(v)$  であり、

$$k' = k - \deg(v) > \frac{1}{2}(m-1)(m-2) - \deg(v) > \frac{1}{2}(m-1)(m-2) - (m-2)$$

ここで

$$\left[ \frac{1}{2}(m-1)(m-2) - (m-2) \right] - \frac{1}{2}[(m-1)-1][(m-1)-2] = m-2 - (m-2) = 0$$

であるから

$$k' > \frac{1}{2}[(m-1)-1][(m-1)-2]$$

これと帰納法の仮定から  $G'$  は連結でなくてはならないが、このことは①と矛盾する。よって  $G$  が非連結であるという仮定は否定される。□

#### XII.2.1.4 Ore の定理

Ore の定理

$n(\geq 3)$  点以上からなる単純グラフ  $G$  の隣接していない任意の 2 点  $v, w$  について  $\deg(v) + \deg(w) \geq n$  であれば  $G$  はハミルトングラフである。

*Proof.*

まず、 $G$  が連結であることを示そう。仮に  $G$  が 2 つの連結成分から成るとする。一方 ( $A$  とする) の点数を  $a$  とすれば他方 ( $B$  とする) の点数は  $n - a$  である。 $A, B$  から任意の点を 1 つずつ選んで  $v, w$  とすると  $\deg(v) \leq a-1$ ,  $\deg(w) \leq n-a-1$  であるから  $\deg(v) + \deg(w) \leq n-2$  となって定理の仮定に反する。このように  $G$  が 2 つの連結成分から成る場合でさえ許されないのだから、3 つ以上など以ての外である。よって  $G$  は連結である。

次に  $G$  がハミルトングラフであることを示す。 $G$  の単純で最長な道を  $p: (u_0, u_1, \dots, u_{k-1}, u_k)$  とする。この中の点は全て異なるので  $k+1 \leq n$  である。この道は最長だから  $u_0$  および  $u_k$  の隣接接点は全てこの道の中にある。集合  $A: \{0 \leq i \leq k-1 | u_{i+1} \text{ は } u_0 \text{ の隣接接点} \}$ ,  $B: \{0 \leq i \leq k-1 | u_i \text{ は } u_k \text{ の隣接接点} \}$  を考えると、当然  $n(A \cup B) \leq k \leq n-1$  である。然るに定理の仮定から  $n(A) + n(B) \geq n$  であるので  $A$  と  $B$  は共通部分を持つ、すなわち、ある  $i$  ( $0 \leq i \leq k-1$ ) が存在して  $u_i$  は  $u_k$  の隣接接点で、 $u_{i+1}$  は  $u_0$  の隣接接点である。

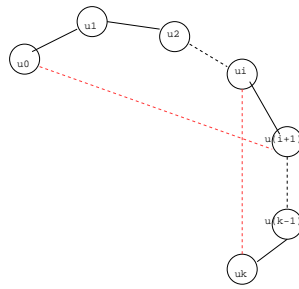


図 XII.2.1 様子

ここで閉路  $c : (u_0, u_{i+1}) + p[u_{i+1}, u_k] + (u_k, u_i) + p[u_i, u_0]$  を考える。但し  $p[u, v]$  は道  $p$  の  $u$  から  $v$  への部分を表す。実はこの閉路  $c$  がハミルトン閉路になっている。

$c$  がハミルトン閉路であることを示そう。仮にこれがハミルトン閉路でないとすると、 $G$  上のある点  $v$  が存在してこれは  $c$  に含まれない。先述の如く  $G$  は連結なので  $v$  はどうにかして  $c$  と繋がっている、すなわち、これまたある点  $w$  が存在して (※たまたま  $v$  と一致することもある) これは  $c$  に属さないが  $c$  上のある点  $u_t$  に隣接している。このとき、 $(w, u_t)$  から始めて  $c$  を上手く回って  $u_t$  の手前まで行く単純な道を考えると (※簡単に考えられる。例えば上の図で  $w$  が  $u_2$  に隣接している場合を考えてみればいい。) これは  $p$  より長くなってしまうので、 $p$  が最長であったという前提に矛盾する。よって  $G$  上の点で  $c$  に含まれないものは無かったということになる。以上より  $c$  はハミルトン閉路である。  $\square$

## 第 XII.3 章

# 木

### XII.3.1 諸定理

#### XII.3.1.1 木は 2 部グラフである

*Proof.*

木の任意の 1 点  $v_{0,0}$  を選んでこれを赤く塗る。次に  $v_{0,0}$  に隣接する全ての接点  $v_{1,0}, v_{1,1}, \dots, v_{1,k}$  を白く塗る。そして  $v_{1,0}, v_{1,1}, \dots, v_{1,k}$  に隣接する全ての無色の接点  $v_{2,0}, v_{2,1}, \dots, v_{2,l}$  を赤く塗る。そして  $v_{2,0}, v_{2,1}, \dots, v_{2,l}$  に隣接する全ての無色の接点を白く塗る。そして...。これを繰り返す。

ある点  $v_{i,j}$  に色を塗るとき、 $v_{0,0}$  から  $v_{i,j}$  への道は唯一だから塗り方は唯一に決まる (道の上で赤→白→赤→白→... と繰り返す)。

全ての点に色がつくまでこれを繰り返せば木は赤と白の 2 つの部分に分けられる 2 部グラフであることがわかる。 □

#### XII.3.1.2 1 本以上の辺を持つ有限な木は次数 1 の枝を少なくとも 2 個持つ

*Proof.*

**XII.3.1.1** と同じように色を塗ることを考える。点数  $n$  が有限なので有限時間内に塗り終わる。塗り終わる (色塗りの旅 (複数に分岐することもある) が行き止まりになって終わる) のは必ず次数 1 の点である。 $v_{0,0}$  の次数が 1 であれば塗り終わりの点 (次数 1) が必ず 1 つ以上ある。 $v_{0,0}$  の次数が 2 以上であれば色塗りの旅は 2 本以上に分岐するが、全て必ず終わるので塗り終わりの点 (次数 1) が  $v_{0,0}$  の次数以上必ずある。いずれにせよ次数 1 の点が少なくとも 2 個ある。 □

#### XII.3.1.3 どこへでも行ける点が 1 個あればその点を根として木を作れる

グラフ  $G(\nu, \varepsilon)$  にある点  $i$  があり、そこから  $G$  の任意の点への有向道があるとき、 $G$  は  $i$  を根とする全域木を持つ。

*Proof.*

$G$  の全ての点と辺を白く塗る。件の点を  $i$  とし、それ以外の点を 1 つ任意に選んで  $j_0$  とする。仮定より  $i$  から  $j_0$  への有向道が存在するから、その点と辺を全て黒く塗ると、その部分は木たる条件を満たしている。 $G$

に白い点が残っていなければ証明は完了。残っていればそこから1つ任意を選んでそれを  $j_1$  と呼ぶ。仮定より  $i$  から  $j_1$  への有向道  $l_1$  が存在する。 $l_1$  の末尾の連続した最長の白い部分 (別の言い方をすれば、「 $j_1$  から  $i$  に向かって  $l_1$  を (逆向きに) 歩いたときに初めて黒い点に踏み込む一歩手前の白い点」から  $j_1$  までの道) を黒く塗ると、 $G$  の黒い点と辺だけを通して  $i$  から  $j_1$  まで行けるようになっているし、黒い部分は依然として木たる条件を満たす。

以上の操作を白い点なくなるまで繰り返した暁には、黒い点と辺から成る木が出来上がっている。  $\square$

**XII.3.1.4  $\mathcal{I}$  : グラフ  $G$  の非閉路部分グラフ全体の集合。**  $I_1, I_2 \in \mathcal{I}, |I_1| < |I_2| \Rightarrow \exists a \in I_2$  s.t.  $I_1 \cup \{a\} \in \mathcal{I}$

*Proof.*

背理法で示す。 $I_2$  のどの辺を  $I_1$  に加えても  $I_1$  に閉路ができると仮定する。そのような辺数最小の  $I_1$  は、下図のように  $I_2$  に含まれない1点から  $I_2$  の各点に1本ずつ辺を伸ばしたものである。

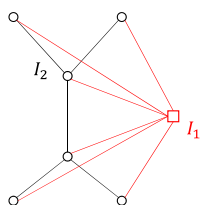


図 XII.3.1 最小の  $I_1$

このとき  $|I_1| = |I_2| + 1$  となる。よって仮定が正しくない。  $\square$

**XII.3.1.5  $\mathcal{T}$  : グラフ  $G$  の全域木の集合。**  $T_1, T_2 \in \mathcal{T}, T_1 \neq T_2, \forall a_1 \in T_1 \setminus T_2, \exists a_2 \in T_2$  s.t.  $(T_1 \setminus a_1) \cup a_2 \in \mathcal{T}$

*Proof.*

全域木の性質より、 $T_2$  に  $a_1$  を加えると  $T_2$  に閉路  $C$  がただ1つできる。その閉路上に、 $T_1$  に属さない辺が少なくとも1つ存在する (そうでなければ  $T_1$  がその閉路を含んでしまう)。それを  $a_2$  とする。 $T_1$  は木だから、 $a_1$  の両端点から  $C$  上の各点に至る経路がそれぞれただ1つずつ存在する。いくつかは  $C$  と共通しているかもしれない。

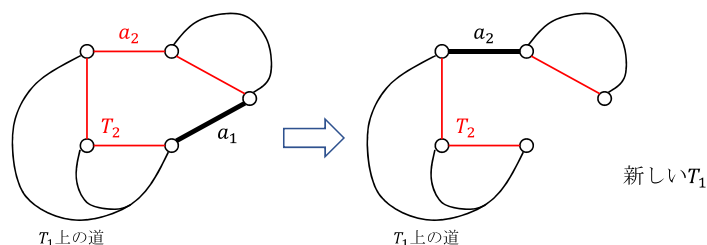


図 XII.3.1 辺を取り替えて新しい  $T_1$  を作る

$T_1$  から  $a_1$  を除去して  $a_2$  を加えると新しい木ができる。  $\square$

## 第 XII.4 章

# 有向グラフ

### XII.4.1 諸定理

XII.4.1.1  $G_1 = (\nu, \varepsilon_1), G_2 = (\nu, \varepsilon_2)$  に対して  $G_1^{|\nu-1|} = G_2^{|\nu-1|}$  であるとき、 $G_1$  が全域木を持つ  $\iff G_2$  が "

*Proof.*

$$\begin{aligned} G_1 \text{ が全域木を持つ} &\iff \exists i \in \nu \text{ s.t. } \forall j \in \nu, \exists (i, j) \in \varepsilon_1^{|\nu-1|} \quad (\because [5] \text{ 定理 2.17}) \\ &\iff \exists i \in \nu \text{ s.t. } \forall j \in \nu, \exists (i, j) \in \varepsilon_2^{|\nu-1|} \quad (\because \text{仮定より } \varepsilon_2^{|\nu-1|} = \varepsilon_1^{|\nu-1|}) \\ &\iff G_2 \text{ が全域木を持つ} \end{aligned}$$

□

#### XII.4.1.2 非閉路的ならば pruning のアルゴリズムはデッドロックしない

前提として、対象としているグラフ  $D$  内の任意の点はスタート点から到達可能であるとする。 $D$  が非閉路的なら pruning のアルゴリズムはデッドロック (ある段階で未決定点ばかりになって先に進まなくなること) しないことを示す。

*Proof.*

対偶「デッドロックするなら閉路が存在する」を示す。デッドロックが起こって  $m$  個の未知点  $u_1, u_2, \dots, u_m$  が残ったとする。 $u_1$  に進入する未知点が  $\{u_2, \dots, u_m\}$  の中に少なくとも 1 つある (そうでなければ  $u_1$  はもはや未知ではない)。必要ならば  $u_2$  以降の番号を振り直して、 $u_1$  に進入する未知点の 1 つを  $u_2$  とする。するとやはり  $u_2$  に進入する未知点が 1 つ以上ある。もしそれが  $u_1$  であれば閉路  $(u_1, u_2, u_1)$  が発生して証明は終了する。そうでないとき必要ならば  $u_3$  以降の番号を振り直して、 $u_2$  に進入する未知点の 1 つを  $u_3$  とする。するとやはり  $u_3$  に進入する未知点が 1 つ以上ある。もしそれが  $\{u_1, u_2\}$  の中にあれば閉路が発生して証明は終了する。そうでないとき必要ならば  $u_4$  以降の番号を振り直して、 $u_3$  に進入する未知点の 1 つを  $u_4$  とする…。これを  $m-1$  回繰り返すと、途中で閉路が発生して証明が終了するか、或いは  $u_{m-1}$  に進入する未知点が  $u_m$  であるところまでくる。やはり  $u_m$  に進入する未知点が 1 つ以上あるのだが、それは  $u_1, u_2, \dots, u_{m-1}$  の中のどれかである。もう他に選択肢がない。ここまでの議論で道  $(u_{m-1}, u_{m-2}, \dots, u_2, u_1)$  が存在するから、 $u_m$  に進入する未知点がどれであろうと必ず閉路が発生する。 □

### XII.4.1.3 スタートからゴールへの全ての道の連鎖重みの総和は 1 以下

有向グラフ  $G(V, E)$  を考える。ある点集合  $S = \{s_1, \dots, s_{N_s}\} \in V$  と、ある一つの頂点  $g \in V$  を選ぶ。 $S$  と  $g$  は次の条件を満たすものとする。

- $S$  のある 1 点から  $S$  の 1 点 (自分自身も含む) へ至る道は存在しない
- $g$  を含む閉路が存在したとしても、その閉路上に  $S$  の点は存在しない

このとき、 $S$  の点から  $g$  へ至る経路全ての連鎖重みを足し上げると、結果は 1 以下である。

「連鎖重み」というのは、辺を通過するたびにその辺の重みを掛け算していくという意味である。 $G$  に閉路がなければ、例えば  $s_1$  から  $g$  へ至るパスは有限通りしかない。しかし  $g$  を含む閉路が存在すると、一度  $g$  へ到着してから閉路を何度か回って  $g$  に落ち着く場合があるので、パターンは無限通りある。しかし閉路を通る毎に重みが掛け算されてどんどん小さくなっていくので、総和が収束しそうである。上の定理の主張は、閉路があっても総和が収束し、しかも 1 以下になるということである。

*Proof.*

まず重要な要素を網羅する具体例を挙げ、それを解析して、証明すべき核心要素を炙り出そう。どうやら閉路が重要そうだ。

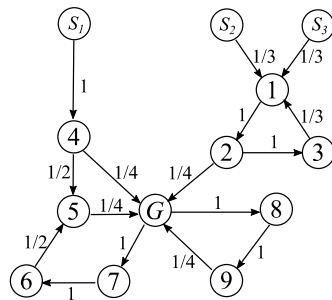


図 XII.4.1  $G(V, E)$

上のグラフは閉路を含んでいる。 $g \rightarrow 8 \rightarrow 9 \rightarrow g$  はまあなんとかなりそうだが。しかし  $g \rightarrow 7 \rightarrow 6 \rightarrow 5 \rightarrow g$  では 5 において 4 経由で  $s_1$  からの入り込みがあってヤバい感じがする。

今求めたい総和というのは結局次のものと同じである。まず  $s_1, \dots, s_3$  に立ったプレイヤーが初め LIFE1 を持っており、そこから出発して  $g$  に向かうとして、辺を通る度にプレイヤーの LIFE に重みを掛け算していき、 $g$  に落ち着いた時点でのプレイヤーの LIFE を評価する。プレイヤーの可能な道の通り方全てについて評価点を足し算した値が今求めたい総和である。

このグラフを次のように 5 段階で変形していく。各段階で何をやっているかは下で説明する。

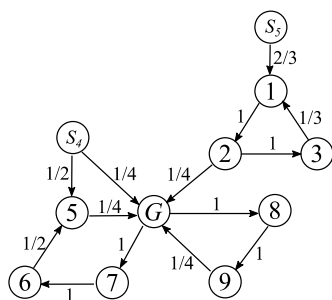


図 XII.4.2 step1

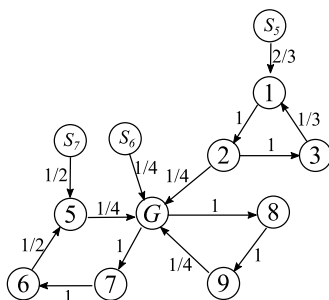


図 XII.4.3 step2

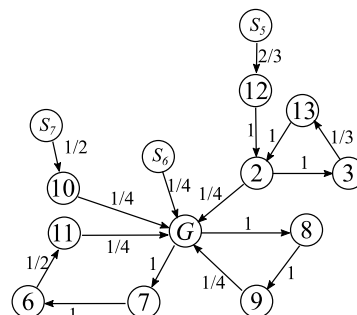


図 XII.4.4 step3

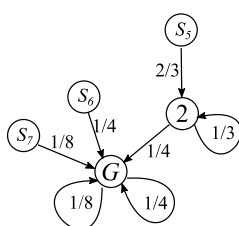


図 XII.4.5 step4

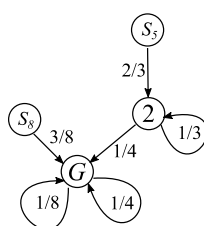


図 XII.4.6 step5

## Step1

$s_1$  から  $g$  へ至るパスを考えると、4 へ着いた時点での LIFE に、4 から  $g$  までのパスの重みを掛け算することになる。だから最初から  $s_1$  を 4 に移動させておけばいい。ここでは新しいノード  $s_4$  で 4 を置き換えている。

$s_2$  から  $g$  へのパスと  $s_3$  から  $g$  へのパスのパターンをもし全部列挙したなら、1 に着いてから先は両者で共通である。今扱っている問題では総和を考えているから、グラフ上で  $s_2$  と  $s_3$  を統合しても計算結果は変わらない。ここでは新しいノード  $s_5$  を作っている。

## Step2

$s_4$  から  $g$  へのパスを考えると、重み  $1/2$  の辺を通る場合と  $1/4$  の辺を通る場合を別々に考えることができる。しかも  $s_4$  を一度離れるともう戻ってこない。だからグラフ上で  $s_4$  からの行き先の数だけ  $s_4$  のコピーを作っても計算結果は変わらない。ここでは新しく  $s_6, s_7$  を作っている。

## Step3

ここが一番ややこしい。 $s_7$  から  $g$  へのパスを考えると、 $s_5$  を経由して  $g$  に着いてそこでお終いというパターンと、その後で閉路に入って何度か回って  $g$  に落ちて着くパターンがある。全てのパターンは別々に計算されるから、グラフ上で 5 をコピーして 2 つにしても計算結果は変わらない。ここでは新しく 10, 11 ノードを作っている。

同じことを 1 ノードに対しても行っている。ここでは新しく 12, 13 ノードを作っている。

## Step4

$s_7$  からは必ず 10 を通るから、グラフ上で  $s_7 \rightarrow 10$  を予め計算しておいても問題ない。 $s_5$  についても同じことをしている。

図 4 には 3 つのループがあるが、ループ上に割り込みが無いので、グラフ上で 1 本の辺に単純化しても計算結果は変わらない。

### Step5

$s_6$  と  $s_7$  を統合して  $s_8$  を作った。Step1 でやったことと同じ。

ここまでくると、次のことが証明できれば定理が証明できたことになるのがわかる。

有向グラフ  $G(V, E)$  を考える。ある頂点  $s$  と  $g$  を選ぶ。これらは次の条件を満たすものとする。

- $s$  は  $g$  に隣接する。
- $g$  を含む閉路が存在したとしても、その閉路上に  $s$  は存在しない

このとき、 $S$  の点から  $g$  へ至る経路全ての連鎖重みを足し上げると、結果は 1 以下である。

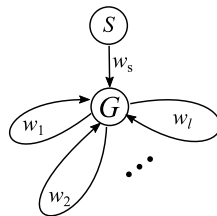


図 XII.4.7  $G(V, E)$

この補題の証明は次の小節で行う。ここではこの補題を認めることにする。この補題を図 6 の 2 ノードに対して適用すると、 $s_5$  ノードから 2 ノードへの全てのパスについて連鎖的重みを足し合わせると 1 になることがわかる。よって 2 ノードを新しく、例えば  $s_9$  で置き換えてから、さらに  $s_8$  と統合させて再び補題を適用すると定理が証明できたことになる。

上の図の場合は連鎖的重みの全ての経路についての総和がちょうど 1 になったが、これは  $S$  の点から  $g$  への全ての経路上のノードにおいて、「 $S$  からも  $g$  からも入れないような経路」からの割り込みがなかったからである。もし、例えば 8 に  $x$  ノードが隣接し、 $x$  の入次数が 0 であると、辺  $x \rightarrow 8$  の重みのせいで他の辺の重みが下がる (重み和は 1 でないといけなかった!) から結果は 1 を下回る。□

#### XII.4.1.4 (補題) ゴールに隣接する 1 つのスタート地点からの全ての経路の連鎖重みの総和は 1 以下

有向グラフ  $G(V, E)$  を考える。ある頂点  $s$  と  $g$  を選ぶ。これらは次の条件を満たすものとする。

- $s$  は  $g$  に隣接する。
- $g$  を含む閉路が存在したとしても、その閉路上に  $s$  は存在しない

このとき、 $S$  の点から  $g$  へ至る経路全ての連鎖重みを足し上げると、結果は 1 以下である。

状況を図に表すと次のようになる。



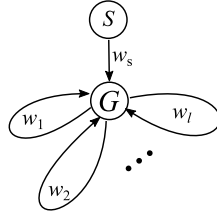


図 XII.4.1  $G(V, E)$

ゴール  $g$  に隣接するスタート  $s$  があり、 $g$  を含む閉路が  $l$  本存在する。もちろん  $w_s + \sum_{i=1}^l w_i \leq 1$  である。

*Proof.*

まず  $s$  から  $g$  へ来てそれでお終いというパターンの重みは  $w_s$  である。次に、一旦  $g$  へ来てから閉路を何度か回るパターンを全列挙するには、 $l$  本の閉路から重複を許して  $k$  本選んで回るパターンを  $k = 1, \dots, \infty$  ままで考えればよい。各パターンの重みは、最初に  $s \rightarrow g$  を通っていることに注意すると

$$\begin{aligned} w_s \sum_{k=1}^{\infty} (w_1 + \dots + w_l)^k &= w_s \left( \frac{1}{1 - (w_1 + \dots + w_l)} - 1 \right) = w_s \frac{w_1 + \dots + w_l}{1 - (w_1 + \dots + w_l)} \\ &\leq w_s \frac{w_1 + \dots + w_l}{w_s} \quad (\because w_s + w_1 + \dots + w_l \leq 1) \\ &= w_1 + \dots + w_l \end{aligned}$$

よって、最初に考えた  $s \rightarrow g$  で終わりのパターンと合わせると総和は  $w_s + w_1 + \dots + w_l \leq 1$  となる。  $\square$

## 第 XIII 部

### その他

## 第 XIII.1 章

# 無限級数公式集

$$\text{XIII.1.1} \quad \sum_{k=1}^{\infty} \frac{\sin k}{k} = \frac{\pi-1}{2}$$

*Proof.*

下図のような鋸波  $y = f(x)$  をフーリエ級数展開して求める。

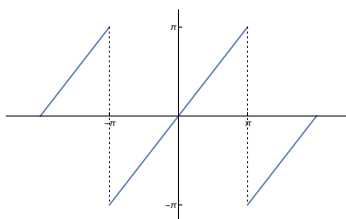


図 XIII.1.1 鋸波

$f(x)$  は奇関数なので  $\cos$  成分はゼロ、つまり  $\forall n, a_n = 0$ 。  $\sin(nx)$  の係数は

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x \sin(nx) dx = \frac{2}{\pi} \int_{-\pi}^0 x \sin(nx) dx = (-1)^{n-1} \frac{2}{n}$$

よって

$$x \sim 2 \left( \frac{\sin x}{1} - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \dots \right)$$

$x$  を  $x - \pi$  で置き換えると

$$x - \pi \sim -2 \left( \frac{\sin x}{1} + \frac{\sin 2x}{2} + \frac{\sin 3x}{3} + \dots \right)$$

特に  $x=1$  とすれば公式を得る。

□

$$\text{XIII.1.2} \quad \sum_{n=1}^{\infty} \frac{1}{1+n^2} = \frac{-1}{2} + \frac{\pi}{2} \coth \pi$$

大阪大学 大学院 工学研究科 平成 21 年度 院試 数学 問題 4 の解答を参照。

$$\text{XIII.1.3} \quad \sum_{n=1}^{\infty} \frac{1}{(1+n^2)^2} = \frac{-1}{2} + \frac{\pi^2}{4 \sinh^2 \pi} + \frac{\pi}{4} \coth \pi$$

大阪大学 大学院 工学研究科 平成 20 年度 院試 数学 問題 4 の解答を参照。

## 参考文献

- [1] 池辺八洲彦、池辺淑子、浅井信吉、宮崎佳典 (2009) 『現代線形代数入門—分解定理を主軸に整理整頓』
- [2] 谷野哲三 (2013) 『システム線形代数』 朝倉書店.
- [3] 笠原皓司 (1995) 『新微分方程式対話』 日本評論社
- [4] 三宅敏恒 著 (1991) 『入門線形代数』 培風館
- [5] 東 俊一, 永原 正章 編著, 石井秀明, 林 直樹, 桜間 一徳, 畑中 健志 共著 (2015) 『マルチエージェントシステムの制御』 コロナ社
- [6] Seymour Lipschitz 著, 成嶋 弘 監訳 (2010) 『マグローヒル大学演習 離散数学 コンピュータサイエンスの基礎数学』 オーム社
- [7] R. Yarlagadda, J. Hershey 共著 “A Note on the Eigenvectors of Hadamard Matrices of Order  $2^n$ ”  
<https://core.ac.uk/download/pdf/81967428.pdf>
- [8] motchy (2022) 『信号処理備忘録 v0.4.0』 <https://github.com/motchy869/Signal-Processing-Memorandum/releases/tag/v0.4.0>