統計数理 (1994)
第 42 巻 第 2 号 197-214

分布の起源

―― ノンパラメトリックな統計的不確定性関係と統計基礎方程式 ――

統計数理研究所　松　縄　　　規

(1993 年 12 月　受付)

要　旨

ノンパラメトリックな多変量統計基礎モデルを構築する．本稿でノンパラメトリックとは我々がモデルの中に特に関心のあるパラメータを考えていないことを意味している．$\boldsymbol{A}$ を実行列の確率空間上で定義されるランダム行列とし，その観測値に基づいて構築される確率モデルの分布 P を本稿では統計的基礎モデルと呼ぶ．P は σ-有限測度 μ に関し絶対連続な多変量密度 p から成るものとする．$\boldsymbol{\varDelta}$ で観測対象行列をモデルで観測した時の誤差行列を表す．モデルの持つデータ記述能力として $\boldsymbol{A}$ に関する性能比強度，ここでは，$\mathscr{P}_A=\nabla_A \ln p$ を考える．但し $\nabla_A=\partial/\partial\boldsymbol{A}$．また誤差と観測機構（＝モデル）について，夫々の平均変動（＝不確定性）を $\boldsymbol{\Sigma}$ および $\boldsymbol{I}$ とし，それらの間の相互作用の平均変動 $\boldsymbol{J}$ とそれらの逆行列の存在を仮定する：この時，統計的不確定性関係 (a) $\boldsymbol{\Sigma}-\boldsymbol{J}^t\boldsymbol{I}^{-1}\boldsymbol{J}\gg\boldsymbol{0}$ (非負値定符号行列)，およびこの等号条件として統計基礎方程式 (b) $|\nabla\rangle p=\boldsymbol{K}|\boldsymbol{\varDelta}\rangle p$ (μ-a.e.) を得る．但し $\nabla=\nabla_A$ を，$\boldsymbol{K}$ は観測精度行列を表す．不等式 (a) は『観測対象とモデルからなる統計的複合システムで両者を同時に決定論的因果関係により厳密に記述する事の不可能性』を示している．しかし特別な場合として，統計基礎方程式(b)が統計的意味での因果関係の成立による客観的基礎モデルの構築を可能にし，この系にマクロな秩序を形成できることを意味している．$\boldsymbol{\varXi}$, $\boldsymbol{\varDelta}$ 等を適切に選択して代表的な多変量分布を含むいくつかの最小不確定性分布族を誘導する．

1.　なぜ分布の起源なのか？

物事の起源とその発展について考究することは現代の諸科学にかなり共通する興味あるテーマである．一般にも知られるようになった例として，Hawking の最近の宇宙論，Eldredge and Gould の生物の系統発生と進化に関する断続平衡説，Prigogine の非平衡熱力学に見るゆらぎからの秩序パターンの形成などを挙げ得るのではと思う．独立した分野の一流の科学者達によって殆ど時を同じくして提唱された興味ある仕事の中に，上記の意味で非常に共通する思想の存在を感じる．こう言った共通なものの見方は統計学に於ても考えられるであろうか？　この事について筆者は Bayes の定理を含め肯定的な考えを持っている．本稿では統計学に於ける物事の起源として，その本質的部分を担っている統計的不確定性関係をノンパラメトリックな場合について論じる．またそれに密接に係わる統計基礎微分方程式を多変量の場合に導入してノンパラメトリックな統計的基礎モデルの構築（＝分布の起源）を考える．物事の発展あるいは進化に関し，統計学に於ては基礎モデルから出発する “分布の発展” が対応する．これも非常に興味のある課題であるが本稿では扱わない．

現代統計学思想の代表的なものとして推測統計学とBayes統計学がまず挙げられる．これらはいわゆる記述統計学を乗り越えた数理統計学として位置付けられている．しかし両者とも問題の初期設定に，関連諸科学の問題意識の高い研究者達には受け入れがたい，大きな疑問点がいくつかあるように思える．特に前者に於ける母集団**分布**の想定，後者の事前分布の想定である．これらの問題は統計学も真なるものの存在をどのように認識し検証するのか，主観-客観の二元論を克服すべきなのかと言った問題を避けて通れないことを意味している．上記の本質的問題に正面から取り組み何らかの積極的提案がなされない限り，両統計学は諸科学の広範囲な分野から真の信頼を受けることが難しい様に思える．これらの統計学を単なる思索の学問とせず，またデータを設定の明確さを欠いた仮説の追認の為に利用されていると曲解されないように問題を根本的に検討する必要を感じる．なお，上記の二つの統計学と思想を異にしている記述統計学あるいはデータ解析学では，それらの守備範囲は別として，データが賢明に使われ上記の問題は表面的には巧みに回避されていることが多い点は参考にすべきことである．いずれにせよこれらの統計学の対立する部分が強調されがちであるがもっと相補的な役割を見出し，理論，応用の両面でより信頼できる統計学の共同した創造が望まれる．そのための第一歩として本稿では統計的意味での客観的な原理に基づく分布の起源と言うことを考える．

母集団分布の想定と言う点に関しFisher (1922) の“On the mathematical foundations of theoretical statistics”に言及しないわけに行かない．彼はこの論文で統計的方法の目的はデータの縮約に在り，その際の統計学の主要問題として (1) 分布の特定化の問題 (Problems of Specification), (2) 推定の問題 (Problems of Estimation) 及び (3) 標本分布の問題 (Problems of Distribution) を挙げている．このことは彼の著書にも転記されており統計学に対する彼の基本的立場を表明していると言えよう．後の二問題についてはそれらの意味も明確であり，彼を含め現在迄に膨大な研究がなされてきた．一方，(1) は「標本がそこから抽出されたと見做されるべき仮説母集団 (hypothetical population) の分布の数学的型を特定すること」を意味する．Fisherは (1) の統計学における重大さを指摘しながらも，これは全く，実際的な統計家の考える問題だとして，理論家がこの問題をどう解決すべきかについての積極的な提案は生涯を通じしなかった．しかし彼は上記の問題 (1) の説明に際しK. Pearsonの分布システムを高く評価している．また同論文のまとめの項で他の多くの事柄の中で「特定化の問題は統計科学 (Statistical Science) の発展に伴って急激に変化し得る考慮すべき事柄に支配されている事が分かる」と極く手短に述べている．筆者はここに，この問題 (1) にいつか全く新しい解釈と展開もあり得ると言う彼の予感を感じる．Fisherの1922年論文のもう一つの気掛りな点は，問題 (3) に関連してGibbsやPlanck等の名前に言及した上でそれらの研究方法を統計的見地からして比較的単純なものとして言外に低い評価で片付けている事である．GibbsやPlanckの仕事こそ実は問題 (1) に関係していると認識すべきであったが，Fisherにはそれらの深い内容，特にエントロピーの意味と役割の理解が不十分であった様に思える．

ところで，Fisher (1936) は，Harvard大学の開学300年祭のArts and Scienceの会議で開会記念講演“Uncertain Inference”を行なった．その一部分で母集団分布の型の決定についてその重要性に再度触れ，「いつの日かデータから母集団の関数型を与える方法を議論する，非常に広い適用範囲を持つ帰納的論証法が展開される可能性があるかもしれない．しかし現時点ではそのような理論は確立されていないと言うことを明確にしておくことが重要である」といった事を述べている．注目すべき事にこの講演においてFisherは1922年論文を参考文献に挙げておらず，“specification”及び“hypothetical population”という1922年論文で多用した言葉も殆ど使っていない．その代わり，上記の様に“データから母集団の関数型を与える”と言う表現が加わって来ている．ここに彼の母集団の記述に対する考え方が微妙に変化している様に筆

者には思える．即ち彼の脳裏に母集団の決定とはデータにもとづいて統計基礎モデルを構築することではないかという疑問が浮かび始めたのかもしれない．しかし明確な原理無しにこれを認めると彼の作り上げた推測統計学の枠組みが崩れることは避けられなかったであろう．この箇所を例外として結局彼は自分の枠組みを守り続けようとした様に見受ける．この数理統計学の創始者の姿勢は統計学の理論を大局的には諸科学から一歩離れた所に位置させる事になったように思える．よく知られているように1926年と1927年に，本質的に分布の起源の問題と密接に関連する，科学史上でも特筆すべき大発見があった．量子力学に於ける Schrödinger (1926) の基礎方程式と Heisenberg (1927)の不確定性原理の発見である．一方，統計学では Fisher は前記1922年論文と1925年の論文 “Theory of statistical estimation” の二つによってほぼ推測統計学の枠組みを作り上げてしまった．上記の物理学の成果が周辺分野に知られ，理解されるにはある程度の年数が必要であったであろう．1935年頃には Fisher にも周囲の世界で何かが起こっていることを感じていたかもしれない．しかし彼が築いた推測統計学の枠組みが障壁となって彼自身にもそれを乗り越えられない高さになっていたかのようである．上述の1936年の講演に周辺の学問の持つ潜在能力と彼の推測統計の枠組みの初期設定の不完全さがいつか表面化するかもしれないという予感が一瞬覗いた様に見える．それにしても，推測統計学が70年以上にも渡って基本的に Fisher の枠組みの中で研究されてきたことは，かなりの本質的部分で統計学と関連諸科学との組織的交流を難しくして来たのではないだろうか．

以上の考察から，分布の特定化 (Specification of distribution) と言う用語は主に Fisher によって導入されはしたが，このことに関する適切な提案はなかったと結論してよいであろう．筆者はこの用語を『データから或いは思考実験を通じて統計基礎モデル分布族を構築すること』と言う意味で使用するのが現代統計学に於てより適切と考える．もともと統計学の本流の中にこの解釈に近いものがあった．Gauss (1809), Keynes (1911, 1962), Poincaré (1912) は極く自然に独立で同一の分布に従う観測値に基づいて未知のパラメトリックな分布の型を見いだす問題を研究している．彼らの問題解決の接近法は本質的にいずれも主要なパラメータに関連した未知の**分布型推定**のための最尤推定法に，観測値の平均に関する条件を加味して改良した修正最尤法と言ったものである．この方法によって彼らは一変量分布に対する基礎微分方程式の原形を与えた．Fisher の最尤推定法が分布型既知としてその中に含まれる未知パラメータの推定であることと著しい対比をなしている．現代諸科学との関連を考える時，Gauss 等の方法は統計学に取ってより基本的であろう．

さて，上記の修正最尤法には，今のところ，使用する観測値が独立で同一分布という仮定を要することや，その意味づけなどの点でいくつかの限界がある．本稿では別の接近法を提案する．統計的不確定性関係に基づく統計基礎方程式による方法である．これらは先述の Heisenberg および Schrödinger の結果に密接に関係する．しかし本稿での結果は数学的なセットアップも含め，後で見るように，量子力学の成果の単なるマクロの世界での表現では決してない．ここで統計的不確定性とは，後述するように，我々の観測対象と我々が作る観測装置（＝多変量統計モデル）に関係するランダム行列値関数の，モデル分布による平均変動によって与えられる．そして “統計的不確定性関係” とは観測対象と観測装置が持つ不確定性の間に成立する不等式関係で表され，『観測対象とモデルからなる統計的複合システムで両者を同時に決定論的因果関係により厳密に記述する事の不可能性』を意味する．この不等式で特に等号が成立する場合に “統計基礎方程式” を得る．この方程式が，あるいはそれを解くことにより得る分布族が，我々に最小不確定な統計基礎モデルの記述を可能にする．これは “観測の精度と言う物差し” を見つけ得る場合の統計的な意味での因果関係，あるいは統計的な意味での主観と客観の合一と見ることも出来よう．

統計的不確定性関係は**形式的には** Edgeworth (1908) 論文で一変量ノンパラメトリック分布の分散の評価を扱った，Love の不等式に遡ることが出来る．論文から問題の提起が Edgeworth によるようにも読み取れるので，以後 Love-Edgeworth 不等式と呼ぶことにする．この不等式が Heisenberg の不確定性関係のほぼ 20 年も前に発見されていたのは驚くべきことであり，その後の統計研究者達が殆どこれに注目せずその解釈をしなかったことは統計学が科学の世界で重大な機会を逸したのではないかと非常に残念に思える．パラメトリックな分布に関する対応する不等式は Aitken and Silverstone (1942) により実質的に与えられた．その数年後，この論文の引用なしに，より整った正則条件の下で C.R. Rao (1945) と Cramér (1946) が独立にこの不等式及びその一般化を行なった．今日 Cramér-Rao の不等式の名で呼ばれているものである．しかしながら，これらの著者達およびその後の統計家達による不等式への主な関心事は，未知ではあるが素朴に存在する（即ち未知ではあるが確定している）と仮定された母集団分布あるいは真の分布の分散の下限の存在とその大きさの評価ということであり，上記の意味での，観測対象とモデルの間の不確定性関係という意味に全く気付かなかった様に見える．言い換えれば通常のいわゆる Cramér-Rao の不等式は形式的にはともかく，その解釈に於て，上記の統計的不確定性関係とは大きく異なったものである．今後この点の吟味を，形式的のみでなく内容的にも十分なされれば，推測統計学の理論を量子力学へ生かそうと言う試みは意義のあるものになると思われる．

統計基礎方程式が統計的不確定性関係の等号条件として得られることは上に述べた．これに関連して，これまでの統計学に於て，パラメトリックな場合に Aitken and Silverstone が上記論文で未知パラメータに関する推定基礎方程式として提案し，Koopmann 型の指数分布族が誘導出来ること，そこで十分統計量が許容されることを見出している．だがここでも統計家の問題意識が推定論の特殊な分野や分布の characterization あるいは分布の identification の問題と言った独特な狭い枠内に留まってしまった．しかし筆者には統計基礎方程式こそが統計理論を進展させる源であり，物理学における Schrödinger 方程式に対応，諸科学との相互理解の鍵を握っているように思える．

本稿の以下の構成は次の通りである．次章では必要となる記号および仮定を準備する．第 3 章ではノンパラメトリックな場合の統計的不確定性関係，統計的基礎方程式を提案する．これに基づいて第 4 章で実対称ランダム行列の分布の特定化について考察し，第 5 章で代表的な多変量分布を統計基礎モデルとして組織的に誘導出来ること，即ちモデルの特定化が可能なことを示す．このことは Bayes 統計解析に於ける事前分布を統計的客観性を持って与え得ることも意味する．

2. 記号，仮定

本章では以下の多変量分布の特定化の問題の考察に関して必要となる記号と基礎モデルが満たすべき主要な仮定を挙げる．

$\boldsymbol{G}=(\boldsymbol{g}_1, \boldsymbol{g}_2,\ldots,\boldsymbol{g}_n)$ を列ベクトル $\boldsymbol{g}_i=(g_{1i}, g_{2i},\ldots,g_{mi})^t$, $(i=1, 2,\ldots,n)$ を持つ大きさ $m\times n$ の行列とする．各 i について，$\boldsymbol{g}_i$ から m_i 個の要素を取り出して部分列ベクトル $\boldsymbol{g}_i^*=(g_{1i}^*, g_{2i}^*,\ldots, g_{m_ii}^*)^t$ を構成する．この部分列ベクトルを積み上げる事により行列 $\boldsymbol{G}$ の列ベクトル化を Dirac のケット記号に倣って定義する：

$$|\boldsymbol{G}\rangle=(g_{11}^*, g_{21}^*,\ldots,g_{m_11}^*, g_{12}^*, g_{22}^*,\ldots,g_{m_22}^*,\ldots,g_{1n}^*, g_{2n}^*,\ldots,g_{m_nn}^*)^t,$$

ここで $\sum_{i=1}^{n} m_i =: k$ は生成された列ベクトルの次元を表す．またこれを転置した行ベクトルを Dirac のブラ記号に倣って次のように表す：

$$\langle \boldsymbol{G}| \ [\equiv(|\boldsymbol{G}\rangle)^t].$$

多変量統計解析における行列のベクトル化として次の三種が重要である．

(1) 全要素積み上げ型 (full stacking)：$m\times n$ 行列 $\boldsymbol{X}$ の $k=mn$ 個の全要素を上の方式で積み上げる．これを次の様に表す．

$$|\mathrm{fls}\ \boldsymbol{X}\rangle=(x_{11}, x_{21},\dots,x_{m1}, x_{12}, x_{22},\dots,x_{m2},\dots,x_{1n}, x_{2n},\dots,x_{mn})^t$$
$$[\equiv(\langle \mathrm{fls}\ \boldsymbol{X}|)^t].$$

(2) 対称三角要素積み上げ型 (triangular stacking)：$m\times m$ 対称行列 $\boldsymbol{Y}$ の対角要素を含む上または下の三角部分の $k=m(m+1)/2$ 個の要素を上の方式で積み上げ次の様に表す．

$$|\mathrm{trs}\ \boldsymbol{Y}\rangle=(y_{11}, y_{21},\dots,y_{m1}, y_{22},\dots,y_{m2},\dots,y_{m-1,m-1}, y_{m-1,m}, y_{m,m})^t$$
$$[\equiv(\langle \mathrm{trs}\ \boldsymbol{Y}|)^t].$$

(3) 歪対称要素積み上げ型 (super-diagonal or sub-diagonal triangular stacking)：$m\times m$ 歪対称行列 $\boldsymbol{Z}$ の対角要素を除く上または下の三角部分の $k=m(m-1)/2$ 個の要素を上の方式で積み上げ次の様に表す．

$$|\mathrm{sds}\ \boldsymbol{Z}\rangle=(z_{21}, z_{31},\dots,z_{m1}, z_{32},\dots,z_{m2},\dots,z_{m-1,m-2}, z_{m,m-2}, z_{m,m-1})^t$$
$$[\equiv(\langle \mathrm{sds}\ \boldsymbol{Z}|)^t].$$

さて，$|\boldsymbol{A}\rangle$ をランダム行列 $\boldsymbol{A}$ をベクトル化した k-次元ランダムベクトルとし，測度空間 $(R^k, \mathbb{B}^k, \mu)$ 上で定義されているものとする．ここに R^k は k-次元実空間，$\mathbb{B}^k$ は R^k の部分集合のボレル σ-集合体，μ は可測空間 $(R^k, \mathbb{B}^k)$ 上の σ-有限測度を表す．以下で混乱のおそれがないならば $|\boldsymbol{A}\rangle$ の代わりに $\boldsymbol{A}$ と記す．

$\mathbb{P}=\{P^A\}$ を上記の可測空間で定義される確率分布からなるモデル分布族とする．確率分布 P^A は μ に関して絶対連続，即ち $P^A(d\boldsymbol{A})=p(\boldsymbol{A})\mu(d\boldsymbol{A})$ と仮定する．ここに $p(\boldsymbol{A})$ は μ に関する Radon-Nikodym 導関数を表す．ここで以下で必要になる行列偏微分作用素と平均作用素を定義しておこう．行列偏微分作用素は次の偏微分列ベクトルで定義する：

$$\begin{aligned}\nabla_A &:= |\nabla_{|A\rangle}\rangle \\ &= \left(\frac{\partial}{\partial a_{11}^*} \cdots \frac{\partial}{\partial a_{m_11}^*} \ \frac{\partial}{\partial a_{12}^*} \cdots \frac{\partial}{\partial a_{m_22}^*} \cdots \frac{\partial}{\partial a_{1n}^*} \cdots \frac{\partial}{\partial a_{m_nn}^*}\right)^t.\end{aligned}$$

この作用素は行ベクトルに前から作用し行列を作る．ランダム行列の平均は平均の行列として定義する：$|\boldsymbol{A}\rangle$ に対し

$$\begin{aligned}\underset{p}{E}[|\boldsymbol{A}\rangle] &:= \left(\underset{p}{E}[a_{11}^*]\cdots\underset{p}{E}[a_{m_11}^*] \ \underset{p}{E}[a_{12}^*]\cdots\underset{p}{E}[a_{m_22}^*]\cdots\underset{p}{E}[a_{1n}^*]\cdots\underset{p}{E}[a_{m_nn}^*]\right)^t \\ &= \left(\int_R a_{11}^* p(\boldsymbol{A})da_{11}^* \cdots \int_R a_{m_11}^* p(\boldsymbol{A})da_{m_11}^* \int_R a_{12}^* p(\boldsymbol{A})da_{12}^* \cdots \int_R a_{m_nn}^* p(\boldsymbol{A})da_{m_nn}^*\right)^t.\end{aligned}$$

二つのランダム行列 $\boldsymbol{A}, \boldsymbol{B}$ のベクトル化から生成される行列 $\boldsymbol{G}:=|\boldsymbol{A}\rangle\langle\boldsymbol{B}|=(g_{ij})$ に対して

$$\underset{p}{E}[\boldsymbol{G}]:=\begin{pmatrix} \underset{p}{E}[g_{11}] & \underset{p}{E}[g_{12}] & \cdots & \underset{p}{E}[g_{1m}] \\ \underset{p}{E}[g_{21}] & \underset{p}{E}[g_{22}] & \cdots & \underset{p}{E}[g_{2m}] \\ \vdots & \vdots & \ddots & \vdots \\ \underset{p}{E}[g_{n1}] & \underset{p}{E}[g_{n2}] & \cdots & \underset{p}{E}[g_{nm}] \end{pmatrix}$$

で定義する．以上の記号上の約束の下，モデル（＝観測機構）の密度 $p(\boldsymbol{A})$ に対して次の仮定を置く：

(A.1) 全ての要素が確率1で有限値を取る行列 $\nabla_A p(\boldsymbol{A})$ が存在する．

次に，$\boldsymbol{A}$ の行列値ランダム関数 $\boldsymbol{\Phi}(\boldsymbol{A}\,;\boldsymbol{\Lambda})$ で我々の観測対象関数を表す．$\boldsymbol{\Phi}$ には $\boldsymbol{A}$ と関数的に独立な潜在パラメーター $\boldsymbol{\Lambda}$ が存在してよいものとする．この関数に対する近似として同じ大きさを持つ行列値ランダム関数 $\boldsymbol{\Psi}(\boldsymbol{A}\,;\boldsymbol{\Lambda})$ を考える．

(A.2) $\boldsymbol{\Phi}(\boldsymbol{A}\,;\boldsymbol{\Lambda})$ および $\boldsymbol{\Psi}(\boldsymbol{A}\,;\boldsymbol{\Lambda})$ は $\boldsymbol{A}$ を実数行列とするとき $\boldsymbol{A}$ に関し積分可能．

さて，観測対象関数 $\boldsymbol{\Phi}$ を近似関数 $\boldsymbol{\Psi}$ で近似する時の測定誤差を次の量で考える：

$$\boldsymbol{\Delta}\equiv\boldsymbol{\Delta}(\boldsymbol{A}\,;\boldsymbol{\Lambda}):=\boldsymbol{\Phi}(\boldsymbol{A}\,;\boldsymbol{\Lambda})-\boldsymbol{\Psi}(\boldsymbol{A}\,;\boldsymbol{\Lambda}).$$

$\boldsymbol{\Delta}$ を測定する観測装置の測定能力，即ち統計モデルの持つデータの記述性能を表現する一つの尺度として

$$\ln p(\boldsymbol{A})\left(=\ln\frac{P^A(d\boldsymbol{A})}{\mu(d\boldsymbol{A})}\right)$$

が考えられる．これを μ に関するモデル分布の**情報比強度** (information specific intensity) と呼ぶ．これをモデル分布で平均すれば一般化された Kullback-Leibler 情報量を得るからであり，ノンパラメトリックな対数尤度と言った曖昧な表現を使わないことにする．更に $\boldsymbol{A}$ の微小変化に対する情報比強度の変化，即ち，モデルのデータ記述性能の $\boldsymbol{A}$ に関する変化をランダム行列

$$\mathscr{P}:=\frac{\partial}{\partial\boldsymbol{A}}\ln p(\boldsymbol{A}\,;\boldsymbol{\Lambda})\left[=:\nabla_A\ln p\equiv\left(\frac{\partial\ln p}{\partial a_{ij}}\right),\ \text{where }(a_{ij})=\boldsymbol{A}\right]$$

で捉えることにする．これをデータ $\boldsymbol{A}$ の記述に関するノンパラメトリックモデルの**性能比強度** (performance specific intensity) と呼ぼう．ここでもノンパラメトリックなスコア関数と言った表現を用いない．本稿を通じて次のことも仮定する：

(A.3) μ に関する Radon-Nikodym 導関数 $p(\boldsymbol{A}\,;\cdot)$ による期待値として次の正則な行列が存在する．

$\boldsymbol{\Sigma}:=\underset{p}{E}[|\boldsymbol{\Delta}\rangle\langle\boldsymbol{\Delta}|]$：(観測対象の平均変動＝観測対象の不確定性),

$\boldsymbol{I}:=\underset{p}{E}[|\mathscr{P}\rangle\langle\mathscr{P}|]$：(性能比強度の平均変動＝性能比強度の不確定性＝モデルによる測定に於ける単位測定尺度),

$\boldsymbol{J}^t:=\underset{p}{E}[|\boldsymbol{\Delta}\rangle\langle\mathscr{P}|]$：(観測対象からモデルへの作用による平均変動),

$\boldsymbol{J}:=\underset{p}{E}[|\mathscr{P}\rangle\langle\boldsymbol{\Delta}|]$：(モデルから観測対象への作用による平均変動),

$\boldsymbol{K}^{-1}:=\boldsymbol{J}^t\boldsymbol{I}^{-1}$：(観測対象に合わせたモデル性能の調整測定尺度),

$\boldsymbol{L}:=\boldsymbol{J}\boldsymbol{\Sigma}^{-1}$:（モデルの性能に合わせた観測対象の調整尺度）.

以上の設定の下に次章でノンパラメトリックな統計的不確定性関係と統計基礎方程式の一般的な事柄を考察する.

3. ノンパラメトリックな統計的不確定性関係と統計基礎方程式

前章の記号，仮定の下で次の定理が成立する：

定理 3.1. $\boldsymbol{y}\in R^k$ を任意の k-次元実数ベクトル，$\boldsymbol{\Lambda}\in\mathscr{S}_{\Lambda}$ を，もし存在するなら，潜在パラメータとする．この時次の不等式（＝統計的不確定性関係）が成立する：

$$\boldsymbol{y}^t\boldsymbol{\Sigma}\boldsymbol{y}\geq\boldsymbol{y}^t\boldsymbol{J}^t\boldsymbol{I}^{-1}\boldsymbol{J}\boldsymbol{y}\qquad[\text{resp. }\boldsymbol{y}^t\boldsymbol{J}^{-1}\boldsymbol{I}(\boldsymbol{J}^{-1})^t\boldsymbol{y}\geq\boldsymbol{y}^t\boldsymbol{\Sigma}^{-1}\boldsymbol{y}]\tag{3.1}$$

すなわち

$$\boldsymbol{\Sigma}-\boldsymbol{J}^t\boldsymbol{I}^{-1}\boldsymbol{J}\gg\boldsymbol{0}\qquad[\text{resp. }\boldsymbol{J}^{-1}\boldsymbol{I}(\boldsymbol{J}^{-1})^t-\boldsymbol{\Sigma}^{-1}\gg\boldsymbol{0}]\qquad(\text{非負値定符号行列}).\tag{3.2}$$

上記不等式で等号が成り立つ為の必要十分条件は，正則な測定精度行列 $\boldsymbol{K}=\boldsymbol{I}(\boldsymbol{J}^t)^{-1}\neq\boldsymbol{0}$ が存在して

$$\left|\frac{\partial\ln p(\boldsymbol{A}\,;\,\boldsymbol{\Lambda})}{\partial\boldsymbol{A}}\right\rangle=\boldsymbol{K}(\boldsymbol{\Lambda})\cdot|\boldsymbol{\Delta}(\boldsymbol{A}\,;\,\boldsymbol{\Lambda})\rangle\qquad(\mu\text{-a.e.})\tag{3.3}$$

で表現される統計基礎方程式が成立する時そしてその時のみに限られる.

証明. 観測対象のランダム関数と統計モデル（＝観測装置）による測定の間に生じる記述誤差として次の k-次元実ランダム列ベクトルを考える：

$$\boldsymbol{d}\equiv|\boldsymbol{\Delta}\rangle-\boldsymbol{J}^t\boldsymbol{I}^{-1}|\mathscr{S}\rangle\qquad[\text{resp. }\boldsymbol{\delta}\equiv|\mathscr{S}\rangle-\boldsymbol{J}\boldsymbol{\Sigma}^{-1}|\boldsymbol{\Delta}\rangle].$$

以下では $\boldsymbol{d}$ について議論すれば十分である．そこで自明な性質 $\boldsymbol{d}\boldsymbol{d}^t\gg\boldsymbol{0}$ (μ-a.e.) を用いれば，任意の非ゼロベクトル $\boldsymbol{y}\in R^k$ に対して $\boldsymbol{y}^t\underset{p}{E}[\boldsymbol{d}\boldsymbol{d}^t]\boldsymbol{y}\geq 0$ を得る．ここに等号が成立するのは $\boldsymbol{d}=\boldsymbol{0}_{k\times 1}$ (μ-a.e.) の時その時のみに限る．（ここでの Cauchy-Schwarz の不等式に依らない，より数学的に単純明快な原理に基づく証明のアイディアは Heisenberg(1930)によって本質的に使用された.）

我々は，従って，$\underset{p}{E}[\boldsymbol{d}\boldsymbol{d}^t]\gg\boldsymbol{0}$（非負値定符号）であることを，前章で導入した不確定性および測定精度の尺度に関する諸量を用いた，意味のある表現で示せばよい：

$$\begin{aligned}\underset{p}{E}[\boldsymbol{d}\boldsymbol{d}^t]&=\underset{p}{E}[\{|\boldsymbol{\Delta}\rangle-\boldsymbol{J}^t\boldsymbol{I}^{-1}|\mathscr{S}\rangle\}\{\langle\boldsymbol{\Delta}|-\langle\mathscr{S}|(\boldsymbol{I}^{-1})^t\boldsymbol{J}\}]\\&=\underset{p}{E}[|\boldsymbol{\Delta}\rangle\langle\boldsymbol{\Delta}|]-\underset{p}{E}[|\boldsymbol{\Delta}\rangle\langle\mathscr{S}|](\boldsymbol{I}^{-1})^t\boldsymbol{J}\\&\quad-\boldsymbol{J}^t\boldsymbol{I}^{-1}\underset{p}{E}[|\mathscr{S}\rangle\langle\boldsymbol{\Delta}|]+\boldsymbol{J}^t\boldsymbol{I}^{-1}\underset{p}{E}[|\mathscr{S}\rangle\langle\mathscr{S}|](\boldsymbol{I}^{-1})^t\boldsymbol{J}\\&=\boldsymbol{\Sigma}-\boldsymbol{J}^t(\boldsymbol{I}^{-1})^t\boldsymbol{J}-\boldsymbol{J}^t\boldsymbol{I}^{-1}\boldsymbol{J}+\boldsymbol{J}^t\boldsymbol{I}^{-1}\boldsymbol{I}(\boldsymbol{I}^{-1})^t\boldsymbol{J}=\boldsymbol{\Sigma}-\boldsymbol{J}^t\boldsymbol{I}^{-1}\boldsymbol{J}\gg\boldsymbol{0}.\end{aligned}\tag{3.4}$$

上述の様に $\underset{p}{E}[\boldsymbol{d}\boldsymbol{d}^t]=\boldsymbol{0}_{k\times k}$ となるのは $\boldsymbol{d}=|\boldsymbol{\Delta}\rangle-\boldsymbol{J}^t\boldsymbol{I}^{-1}|\mathscr{S}\rangle\equiv\boldsymbol{0}_{k\times 1}$ (μ-a.e.) となる時その時に限る．即ち，$\boldsymbol{J}\neq\boldsymbol{0}_{k\times k}$ ならば等号成立の必要十分条件として次の方程式が従う：

$$|\mathscr{S}\rangle=\boldsymbol{I}(\boldsymbol{J}^t)^{-1}|\boldsymbol{\Delta}\rangle\quad[\text{resp. }|\mathscr{S}\rangle=\boldsymbol{J}\boldsymbol{\Sigma}^{-1}|\boldsymbol{\Delta}\rangle]\quad(\mu\text{-a.e.}).\tag{3.5}$$

換言すると，正則な行列 $\boldsymbol{K}=\boldsymbol{I}(\boldsymbol{J}^t)^{-1}\neq\boldsymbol{0}_{k\times k}$ [or $\boldsymbol{L}=\boldsymbol{J}(\boldsymbol{\Sigma})^{-1}\neq\boldsymbol{0}$] が存在して次の偏微分方程式が成立する：

$$\left|\frac{\partial \ln p(\boldsymbol{A}\,;\,\boldsymbol{\Lambda})}{\partial \boldsymbol{A}}\right\rangle=\boldsymbol{K}(\boldsymbol{\Lambda})\cdot|\boldsymbol{\varDelta}(\boldsymbol{A}\,;\,\boldsymbol{\Lambda})\rangle \quad (\mu\text{-a.e.}). \tag{3.6}$$ ▮

注 3.1. 一変量の場合の不等式が前述の Love-Edgeworth (Edgeworth (1908)) の不等式である．彼らは不等式の等号条件には全く触れなかった．その後この不等式は Kagan et al. (1973) によっても独立に発見され，等号条件の考察と，分布の分散が既知の時ノンパラメトリックな Fisher 情報量最小化原理をいくつかの分布の特徴付けに適用した．しかし上記いずれの研究者たちも彼らの不等式の背後には統計的不確定性関係や統計基礎方程式と言う統計学にとって根本的に重要な事柄が関係していることに全く気がつかなかった様に見える．

注 3.2. (3.5) 式は以下の (3.8) のようにも表現できる：

$$|\mathscr{P}\rangle=\frac{\partial p(\boldsymbol{A}\,;\boldsymbol{\Lambda})}{\partial|\boldsymbol{A}\rangle}\cdot\frac{1}{p(\boldsymbol{A}\,;\boldsymbol{\Lambda})}=:|\nabla_A p\rangle p^{-1}$$

と記せば，これと (3.5) から，

$$|\boldsymbol{\varDelta}\rangle p=\boldsymbol{K}^{-1}|\nabla_A\rangle p\ \ [\Leftrightarrow|\nabla_A\rangle \ln p=\boldsymbol{K}|\boldsymbol{\varDelta}\rangle]\,;\ [\boldsymbol{K}^{-1}=(\boldsymbol{J}^t)\boldsymbol{I}^{-1}]. \tag{3.7}$$

ここで，$|\phi|^2:=p$ と置くと

$$|\boldsymbol{\varDelta}\rangle|\phi|=2\boldsymbol{K}^{-1}|\nabla_A\rangle|\phi| \tag{3.8}$$

を得る．そこで次の対応を考える：

$$|\boldsymbol{\varDelta}\rangle\leftrightarrow\mathscr{H}(\text{Hamiltonian operator}),\qquad \sqrt{p}=|\phi|\leftrightarrow\psi\ (\text{Wave function}),$$

$$\boldsymbol{K}^{-1}=\boldsymbol{J}^t\boldsymbol{I}^{-1}\leftrightarrow i\hbar/2(\hbar:\text{ Dirac constant}),\qquad |\boldsymbol{A}\rangle\leftrightarrow t,\qquad |\nabla_A\rangle\left(=\frac{\partial}{\partial|\boldsymbol{A}\rangle}\right)\leftrightarrow\frac{\partial}{\partial t}.$$

この結果 (3.8) に対応して次の方程式を得る：

$$\mathscr{H}\psi=i\hbar\frac{\partial}{\partial t}\psi \qquad (\text{Schrödinger 方程式}). \tag{3.9}$$

この方程式は物理学の重要な基礎方程式である事がよく知られている．このことに対応して，定理の方程式 (3.5) あるいは (3.6) をノンパラメトリックな統計基礎モデル構築の為の統計基礎方程式と呼んでもよいであろう．

注 3.3. 量子力学に於て Planck 定数 $h(=2\pi\hbar)$ はミクロの世界での測定可能限界目盛と考え得るが，その存在が古典力学に於ける因果関係を断ち切っていることはよく知られている．これに対し本稿で扱っている世界では，h に対応して $\boldsymbol{I}^{-1}$ がある意味でノンパラメトリックモデル自身が持つ測定能力の限界目盛と見なせる．この目盛を使用できる時データ $\boldsymbol{A}$ の記述に関するモデルの性能比強度 $\mathscr{P}$ は，本稿の正則条件の下，最小不確定性と言う意味で最良の関係式 (3.5) を満たす．また，h との類推で $\boldsymbol{K}^{-1}$ の存在が統計的因果関係の鍵を握っている．ただし h と異なり，$\boldsymbol{K}^{-1}$ は絶対定数行列ではなく基礎のモデル分布に固有の $\boldsymbol{I}^{-1}$ と，観測対象からモデルへの影響 $\boldsymbol{J}^t$ に依存する．

注 3.4. 上記注 3.2 の基礎方程式は様々な重要な微分方程式を含んでいる．例えば，一変量の

密度関数 $f(x)$ に対して

$$\frac{df(x)}{dx}=\varDelta\cdot f(x),\ \text{ここに}\ \varDelta=\varDelta(x\,;a,b_0,b_1,b_2)=(x+a)/(b_0+b_1x+b_2x^2).$$

これはよく知られたPearsonシステムの基本方程式を表している (cf. Pearson (1895, 1916), Elderton and Johnson (1969)). Pearsonの特殊な導入に比べ上記の解釈は，このシステムをより合理的に説明している．離散型分布の場合も本稿の基礎方程式はCarver-Ordシステムと関係する (cf. Ord (1967)).

系 3.1. 定理3.1と同じ条件の下で，$k\times k$ 実対称行列 $\boldsymbol{Y}$ に対して $\boldsymbol{y}\boldsymbol{y}^t=\boldsymbol{Y}=\boldsymbol{Y}^t$ となる k-次元実ベクトル $\boldsymbol{y}$ が存在するとき，次の行列二次形式での不確定性関係が成立する．

$$\boldsymbol{Y}^t(\boldsymbol{\Sigma}-\boldsymbol{J}^t\boldsymbol{I}^{-1}\boldsymbol{J})\boldsymbol{Y}\gg 0.$$

証明． 定理より任意の $\boldsymbol{y}\in R^k$ に対し $\boldsymbol{y}^t\boldsymbol{\Sigma}\boldsymbol{y}\geq\boldsymbol{y}^t\boldsymbol{J}^t\boldsymbol{I}^{-1}\boldsymbol{J}\boldsymbol{y}$ であるからこの不等式の両辺に前後からスカラー $\boldsymbol{y}^t\boldsymbol{y}$ を掛けて $\boldsymbol{y}\boldsymbol{y}^t=\boldsymbol{Y}=\boldsymbol{Y}^t$ に注意すれば

$$\boldsymbol{y}^t\boldsymbol{Y}^t\boldsymbol{\Sigma}\ \boldsymbol{Y}\boldsymbol{y}\geq\boldsymbol{y}^t\boldsymbol{Y}^t\boldsymbol{J}^t\boldsymbol{I}^{-1}\boldsymbol{J}\ \boldsymbol{Y}\boldsymbol{y}.$$ ∎

さて，定理3.1のノンパラメトリック統計基礎方程式あるいは (3.7) を積分することにより次の**最小不確定性分布** (minimum uncertainty distribution) の一般型を得る：

定理 3.2. 前定理と同じ記号と条件の下で，ランダム行列 $\boldsymbol{A}$ のノンパラメトリック統計基礎モデル分布の密度関数は次式で与えられる：

$$p(|\boldsymbol{A}\rangle\,;\boldsymbol{\Lambda})=c(\boldsymbol{\Lambda})\cdot\exp\left\{\int\langle d\boldsymbol{A}|\boldsymbol{K}|\boldsymbol{\varDelta}\rangle\right\}=c(\boldsymbol{\Lambda})\cdot\exp\left\{\int\langle\boldsymbol{\varDelta}|\boldsymbol{K}^t|d\boldsymbol{A}\rangle\right\} \tag{3.10}$$

ここで $\langle d\boldsymbol{A}|$ [resp.$|d\boldsymbol{A}\rangle$] は k-次元微分横ベクトル [resp. 縦ベクトル] で

$$\langle d\boldsymbol{A}|=(|d\boldsymbol{A}\rangle)^t=(da^*_{11},\dots,da^*_{m_11},da^*_{12},\dots,da^*_{m_22},\dots,da^*_{1n},\dots,da^*_{m_nn}),$$

但し $\sum_{i=1}^n m_i=k$. $\boldsymbol{K}(\boldsymbol{\Lambda})=\boldsymbol{I}(\boldsymbol{J}^t)^{-1}$ は内積 $\langle d\boldsymbol{A}|\boldsymbol{K}|\boldsymbol{\varDelta}\rangle$ の係数行列 $(k\times k)$ で $\boldsymbol{\Lambda}$ に依存してよい．また，$c(\boldsymbol{\Lambda})>0$ は次の規格化条件を満たすスカラー値関数を表す：

$$\int_{R^k}p(|\boldsymbol{A}\rangle\,;\boldsymbol{\Lambda})\,\mu\{(d|\boldsymbol{A}\rangle)\}=1.$$

なお $\mu\{\cdot\}$ は $|\boldsymbol{A}\rangle$ の定義される可測空間 $(R^k,\mathbb{B}^k)$ 上の σ-有限測度を，$(d|\boldsymbol{A}\rangle)$ は k 重積分の基本体積要素

$$(d|\boldsymbol{A}\rangle)=da^*_{11}\cdots da^*_{m_11}da^*_{12}\cdots da^*_{m_22}\cdots da^*_{1n}\cdots da^*_{m_nn},\ \left[\sum_{i=1}^n m_i=k\right]$$

を表す．

4. 実対称ランダム行列の分布の特定化

本章では μ をLebesgue測度として，それに関して絶対連続な統計基礎モデルの密度関数 $f(|\boldsymbol{A}\rangle\,;\boldsymbol{\Lambda})$ の構築を考える．以下では通常の多変量解析に於て重要な役割を演じている実対称ランダム行列 $\boldsymbol{A}(m\times m)$ の分布の特定化に限定する．まず以下の議論で二種類の積分が必要と

なるため記号 $[dA]$ 及び (dA) を導入する：

$$[d\boldsymbol{A}]=\begin{pmatrix} da_{11} & da_{12} & \cdots & da_{1m} \\ da_{12} & da_{22} & \cdots & da_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ da_{1m} & da_{2m} & \cdots & da_{mm} \end{pmatrix},$$

$$(d\boldsymbol{A})=\bigwedge_{1\le i\le j\le m} da_{ij}:\ d\boldsymbol{A}\ \text{の}\ k=m(m+1)/2\ \text{個の要素の積.}$$

さて，我々が必要なのは (3.10) に見る内積部分に相当する行列表示である．以下では実対称行列の対角可能性を考慮して (3.3) に対応する行列表示として次の型にまとめられる場合から出発する：

$$\partial \ln f(\boldsymbol{A}, \boldsymbol{\Lambda})/\partial \boldsymbol{A}_{\Delta}=\boldsymbol{\Xi}(\boldsymbol{\Lambda})\boldsymbol{\Delta}^*(\boldsymbol{A}_{\Delta}, \boldsymbol{\Lambda}) \qquad (\mu\text{-a.e.}). \tag{4.1}$$

ここで $\boldsymbol{A}=\boldsymbol{T}\boldsymbol{A}_{\Delta}\boldsymbol{T}^t$, ($\boldsymbol{T}$: $m\times m$: 直交行列) を，$\boldsymbol{\Delta}^*, \boldsymbol{\Xi}$ はそれぞれ (3.3) の $\boldsymbol{\Delta}, \boldsymbol{K}$ に対応する実対称行列で積は実対角とする．この設定で上記 (4.1) を積分して次の結果を得る：

定理 4.1. 基礎方程式 (4.1) を満たす実対称ランダム行列 $\boldsymbol{A}$ のノンパラメトリックな最小不確定性分布の密度関数型は次のように与えられる：

$$f(\boldsymbol{A}\,;\boldsymbol{\Lambda})=C(\boldsymbol{\Lambda})\cdot \mathrm{etr}\left\{\boldsymbol{\Xi}(\boldsymbol{\Lambda})\int \boldsymbol{\Delta}^*(\boldsymbol{A}_{\Delta}\,;\boldsymbol{\Lambda})[d\boldsymbol{A}_{\Delta}]\right\},$$

但し $C(\boldsymbol{\Lambda})$ は $\boldsymbol{A}$ に依存しない規格化スカラー値関数で次の条件を満たすように定まることが要求される(これが実現されない時，モデルは構築出来ない)：

$$\int_{\mathfrak{S}} f(\boldsymbol{A}\,;\ \boldsymbol{\Lambda})\,(d\boldsymbol{A})=1,$$

ここに積分は $m\times m$ の実対称行列の空間 $\mathfrak{S}$ 上でなされる．特に $\boldsymbol{A}$ が正定値の時

$$\int_{\boldsymbol{A}>0} f(\boldsymbol{A}\,;\ \boldsymbol{\Lambda})(d\boldsymbol{A})=1$$

と記す．

この定理により，$\boldsymbol{\Delta}^*$ とそれに対する $\boldsymbol{\Xi}$ を適切に設定した時 $C(\boldsymbol{\Lambda})$ が決まるなら基礎モデルが求まる (cf. 例 5.1)．実際に $\boldsymbol{\Delta}^*$ をどう与えるかは基礎モデル構築という立場から重要である．手掛かりとして $\boldsymbol{\Delta}^*(\boldsymbol{A}\,;\ \boldsymbol{\Lambda})=\boldsymbol{\Delta}_0^*(\boldsymbol{A}\,;\ \boldsymbol{\Lambda})+\boldsymbol{\Delta}_1^*(\boldsymbol{A}\,;\ \boldsymbol{\Lambda})$ と分解し主要部分を $\boldsymbol{\Delta}_0^*(\boldsymbol{A}\,;\ \boldsymbol{\Lambda})$ とする時それを各種の平均と関連させて考えると，基礎モデル構築の一助になり得る．即ち，$\{\boldsymbol{A}_i\}(i=1,2,\ldots,N)$ を $m\times m$ の実対称ランダム行列の列とし，$|\boldsymbol{A}_i\rangle$ が独立に同一な基礎モデル密度関数 $f(|\boldsymbol{A}\rangle;\ \boldsymbol{\Lambda})$ を構築しているものとする．今 $\boldsymbol{\Lambda}$ が平均行列 $\boldsymbol{M}(m\times m)$ とその他のパラメータ行列から構成されているとする．この設定の下で測定誤差の主要部分，$\boldsymbol{\Delta}_0^*(\boldsymbol{A}\,;\ \boldsymbol{\Lambda})$ として興味のあるいくつかの例が考え得る：

$$\Delta_0^*(\boldsymbol{A}\,;\,\boldsymbol{\Lambda})=\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{A}_{i\Delta}-\boldsymbol{M}_\Delta=:\boldsymbol{A}_\Delta-\boldsymbol{M}_\Delta \quad (\boldsymbol{A}_\Delta:\text{ arithmetic mean}),$$

$$\Delta_0^*(\boldsymbol{A}\,;\,\boldsymbol{\Lambda})=\ln\Big(\prod_{i=1}^{N}\boldsymbol{A}_{i\Delta}\Big)^{1/N}-\ln\boldsymbol{M}_\Delta=:\ln\boldsymbol{A}_\Delta-\ln\boldsymbol{M}_\Delta$$
$$(\boldsymbol{A}_\Delta:\text{ geometric mean}\,;\,\boldsymbol{A}_{i\Delta},\boldsymbol{A}_\Delta,\boldsymbol{M}_\Delta>0),$$

$$\Delta_0^*(\boldsymbol{A}\,;\,\boldsymbol{\Lambda})=N\Big(\sum_{i=1}^{N}\boldsymbol{A}_{i\Delta}^{-1}\Big)^{-1}-\boldsymbol{M}_\Delta^{-1}=:\ \boldsymbol{A}_\Delta^{-1}-\boldsymbol{M}_\Delta^{-1} \quad (\boldsymbol{A}_\Delta:\text{ harmonic mean}\,;\,\boldsymbol{A}_\Delta,\boldsymbol{M}_\Delta>0),$$

$$\Delta_0^*(\boldsymbol{A}\,;\,\boldsymbol{\Lambda})=\boldsymbol{M}_\Delta\Big\{\Big(\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{A}_{i\Delta}\Big)\Big(N\Big(\sum_{i=1}^{N}\boldsymbol{A}_{i\Delta}^{-1}\Big)^{-1}\Big)\Big\}^{-1}-\boldsymbol{M}_\Delta^{-1}$$
$$=:\boldsymbol{M}_\Delta\boldsymbol{A}_\Delta^{-2}-\boldsymbol{M}_\Delta^{-1} \quad (\boldsymbol{A}_\Delta:\text{ trinity mean}\,;\,\boldsymbol{A}_{i\Delta},\boldsymbol{A}_\Delta,\boldsymbol{M}_\Delta>0).$$

次章では上記またはその変形したものを基に，定理 4.1 を用いてノンパラメトリックな最小不確定性分布の幾つかの例を与える．その際必要となる定理に現われる不定積分の計算に関し次の対称行列の対角化表示を利用する：$\boldsymbol{A}$ が実対称だから

$$\boldsymbol{A}^r=\boldsymbol{T}\boldsymbol{A}_\Delta^r\boldsymbol{T}^t\ \ (r=+1,\pm2,\ldots)\quad \text{ここに } \boldsymbol{A}_\Delta=\mathrm{diag}(\alpha_1,\ldots,\alpha_m),\ \boldsymbol{T}:\text{ 直交行列}$$

と表現できる．

注 4.1. 次の対角化も次章で用いる：『$\boldsymbol{C}(m\times m)>0$, $\boldsymbol{X}(m\times m)$ が実対称の時 $\boldsymbol{C}=\boldsymbol{Q}\boldsymbol{Q}^t$, $\boldsymbol{X}=\boldsymbol{Q}\boldsymbol{X}_\Delta\boldsymbol{Q}^t$ となる正則行列 $\boldsymbol{Q}(m\times m)$ が存在する．但し $\boldsymbol{X}_\Delta=\boldsymbol{X}_\Delta(\boldsymbol{C})=\mathrm{diag}(\lambda_1,\ldots,\lambda_m)$ を意味し，$\lambda_1,\ldots,\lambda_m$ は $|\boldsymbol{C}^{-1}\boldsymbol{X}-\lambda\boldsymbol{I}_{m\times m}|=0$ の根を表す．』(cf. Bellman (1970), p. 58)．これにより実対称ランダム行列が正定符号の係数行列を持つ場合の不定積分の計算が可能になる．また，『二つの可換な実対称行列 $\boldsymbol{E},\boldsymbol{F}$ を同時に対角化する直交行列 $\boldsymbol{T}$ が存在する』事 (cf. Bellman (1970), p. 56) も次章で利用する．その際に可換性は，$\mathrm{tr}\boldsymbol{EF}=\mathrm{tr}\boldsymbol{FE}$ である事と定理 4.1 の規格化条件がチェックされるので，成立しているものとして計算できる．

なお，次章でも添字 Δ を上で定義した各種の対角化に共通して使用する．

5. 実対称ランダム行列の最小不確定性分布の例

前章の結果に基づいて種々の最小不確定性分布を求める．Δ^* および Ξ を適切に選ぶことにより以下の重要な分布を産み出すことができる．本章では通常の見慣れた型で表現するため $|\boldsymbol{A}\rangle$ の代わりに $\boldsymbol{A}$ を用いる．なお，$\boldsymbol{A}$ や $\boldsymbol{\Lambda}$ 等が一変量変数あるいは定数とする時，以下の諸分布は対応する一次元分布を表す．

例 5.1. 多変量対称正準分布 (Multivariate symmetric canonical distribution)

$$f_{\mathrm{SC}}(\boldsymbol{A}\,;\,\boldsymbol{\Lambda})=\mathrm{etr}\Big\{-\sum_{\nu=0}^{N}\boldsymbol{\Lambda}_\nu\boldsymbol{A}^\nu\Big\},\ [\boldsymbol{A},\boldsymbol{\Lambda}_\nu>0,\ \boldsymbol{A}^0=\boldsymbol{I}_{m\times m}\,;\ \boldsymbol{\Lambda}=(\boldsymbol{\Lambda}_0,\boldsymbol{\Lambda}_1,\ldots,\boldsymbol{\Lambda}_N)].$$

ここで $\boldsymbol{\Lambda}_\nu(\nu=1,\ldots,N)$ は

$$\boldsymbol{A}:=\frac{1}{N}\sum_{\nu=1}^{N}\boldsymbol{A}_\nu,\ (\boldsymbol{A}_\nu\in\mathfrak{S})$$

の冪行列 $\boldsymbol{A}^\nu\ (\nu=1,\ldots,N)$ の $m\times m$ の係数行列を表し，また $\boldsymbol{\Lambda}_0$ は次式で与えられる：

$$\boldsymbol{\Lambda}_0:=\frac{1}{m}\boldsymbol{I}_{m\times m}\ln\Big\{\int_{\mathfrak{S}}\mathrm{etr}\Big(-\sum_{\nu=1}^{N}\boldsymbol{\Lambda}_\nu\boldsymbol{A}^\nu\Big)(d\boldsymbol{A})\Big\}.$$

誘導.

$$\boldsymbol{\Delta}^*(\boldsymbol{A}_\Delta;\boldsymbol{\Lambda})=-\sum_{\nu=0}^{N}(\nu+1)\boldsymbol{A}_\Delta^\nu\ [\boldsymbol{A}_\Delta^0=\boldsymbol{I}_{m\times m}]\,;\ \boldsymbol{\Xi}(\boldsymbol{\Lambda})=\boldsymbol{I}_{m\times m}$$

と置く．ここで，$\boldsymbol{A}_\Delta^\nu=\mathrm{diag}(\lambda_1^\nu,\ldots,\lambda_m^\nu)$，但し $(\lambda_1^\nu,\ldots,\lambda_m^\nu)$ は $|\boldsymbol{\Lambda}_{\nu+1}^{-1}\boldsymbol{A}^\nu-\lambda\boldsymbol{I}_{m\times m}|=0$ の根とする．明らかに $\boldsymbol{\Xi}(\boldsymbol{\Lambda})\boldsymbol{\Delta}^*(\boldsymbol{A}_\Delta;\boldsymbol{\Lambda})$ は対角行列である．よって定理 4.1 から

$$f_{\mathrm{SC}}(\boldsymbol{A};\boldsymbol{\Lambda})=C_{\mathrm{SC}}(\boldsymbol{\Lambda})\cdot\mathrm{etr}\left\{-\int\sum_{\nu=0}^{N-1}(\nu+1)\boldsymbol{A}_\Delta^\nu[d\boldsymbol{A}_\Delta]\right\}=C_{\mathrm{SC}}(\boldsymbol{\Lambda})\cdot\mathrm{etr}\left(-\sum_{\nu=1}^{N}\boldsymbol{A}_\Delta^\nu\right).$$

注 4.1 から，$\boldsymbol{\Lambda}_\nu^{-1}=\boldsymbol{Q}_\nu\boldsymbol{Q}_\nu^t$, $\boldsymbol{A}^\nu=\boldsymbol{Q}_\nu\boldsymbol{A}_\Delta^\nu\boldsymbol{Q}_\nu^t$ $(\nu=1,\ldots,N)$ となる正則行列 $\boldsymbol{Q}_\nu(m\times m)$ が存在して $\mathrm{tr}\boldsymbol{A}_\Delta^\nu=\mathrm{tr}\{\boldsymbol{Q}_\nu^{-1}\boldsymbol{A}^\nu(\boldsymbol{Q}_\nu^t)^{-1}\}=\mathrm{tr}[(\boldsymbol{Q}_\nu^t)^{-1}\{\boldsymbol{Q}_\nu^{-1}\boldsymbol{A}^\nu(\boldsymbol{Q}_\nu^t)^{-1}\}\boldsymbol{Q}_\nu^t]=\mathrm{tr}\{\boldsymbol{\Lambda}_\nu\boldsymbol{A}^\nu\}$ であるから

$$f_{\mathrm{SC}}(\boldsymbol{A};\boldsymbol{\Lambda})=C_{\mathrm{SC}}(\boldsymbol{\Lambda})\cdot\mathrm{etr}\left(-\sum_{\nu=1}^{N}\boldsymbol{\Lambda}_\nu\boldsymbol{A}^\nu\right).$$

よって

$$C_{\mathrm{SC}}(\boldsymbol{\Lambda})=1\Big/\int_{\mathfrak{S}}\mathrm{etr}\left(-\sum_{\nu=1}^{N}\boldsymbol{\Lambda}_\nu\boldsymbol{A}^\nu\right)(d\boldsymbol{A})=\mathrm{etr}(-\boldsymbol{\Lambda}_0).$$

多変量対称正準分布の簡単なものを例示する：

(i) $N=1$

$$f(\boldsymbol{A};\boldsymbol{\Lambda})=\mathrm{etr}\{-\boldsymbol{\Lambda}_1\boldsymbol{A}\}\Big/\int_{\boldsymbol{A}>0}\mathrm{etr}\{-\boldsymbol{\Lambda}_1\boldsymbol{A}\}(d\boldsymbol{A})=\frac{1}{\Gamma_m\left(\frac{m+1}{2}\right)}|\boldsymbol{\Lambda}_1|^{(m+1)/2}\,\mathrm{etr}\{-\boldsymbol{\Lambda}_1\boldsymbol{A}\},$$
$$(\boldsymbol{A},\boldsymbol{\Lambda}_1>0).$$

これは多変量指数分布の密度関数である．

(ii) $N=2$

$$\begin{aligned}f(\boldsymbol{A};\boldsymbol{\Lambda})&=\mathrm{etr}\{-(\boldsymbol{\Lambda}_1\boldsymbol{A}+\boldsymbol{A}^2)\}\Big/\int_{\mathfrak{S}}\mathrm{etr}\{-(\boldsymbol{\Lambda}_1\boldsymbol{A}+\boldsymbol{A}^2)\}(d\boldsymbol{A})\\&=2^{m(m-1)/4}\pi^{-m(m+1)/4}\,\mathrm{etr}\left\{-\left(\boldsymbol{A}+\frac{1}{2}\boldsymbol{\Lambda}_1\right)^2\right\},\qquad(\boldsymbol{A},\boldsymbol{\Lambda}_1\in\mathfrak{S}).\end{aligned}$$

何故なら，

$$\begin{aligned}\int_{\mathfrak{S}}\mathrm{etr}\{-(\boldsymbol{\Lambda}_1\boldsymbol{A}+\boldsymbol{A}^2)\}(d\boldsymbol{A})&=\int_{\mathfrak{S}}\mathrm{etr}\left\{-\left(\boldsymbol{A}^2+\frac{1}{2}\boldsymbol{\Lambda}_1\boldsymbol{A}+\frac{1}{2}\boldsymbol{A}\boldsymbol{\Lambda}_1\right)\right\}(d\boldsymbol{A})\\&=\int_{\mathfrak{S}}\mathrm{etr}\left\{-\left(\boldsymbol{A}+\frac{1}{2}\boldsymbol{\Lambda}_1\right)^2+\frac{1}{4}\boldsymbol{\Lambda}_1^2\right\}(d\boldsymbol{A}),\end{aligned}$$

変数変換 $\sqrt{2}\left(\boldsymbol{A}+\frac{1}{2}\boldsymbol{\Lambda}_1\right)=\boldsymbol{Y}, J(\boldsymbol{A}\to\boldsymbol{Y})=2^{-m(m+1)/4}$ を行なって

$$=2^{-m(m+1)/4}\,\mathrm{etr}\left(\frac{1}{4}\boldsymbol{\Lambda}_1^2\right)\cdot I^*.$$

ここで

$$I^* := \int_{\mathfrak{S}} \mathrm{etr}\left(-\frac{1}{2}\boldsymbol{Y}^2\right)(d\boldsymbol{Y}) = \int \exp\left(-\langle d\boldsymbol{Y} | \frac{1}{2}\boldsymbol{I}_{k\times k} | \boldsymbol{Y}^2\rangle\right) \quad [k=m(m+1)/2]$$
$$= \prod_{i=1}^{m}\left[\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}\sum_{i=1}^{m} y_{ii}^2\right)dy_{ii}\right]\cdot\prod_{i<j}^{m}\left[\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}2\sum_{i<j}^{m} y_{ij}^2\right)dy_{ij}\right]$$
$$= \prod_{i=1}^{m}\left[\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}\sum_{i=1}^{m} y_{ii}^2\right)dy_{ii}\right]\cdot\prod_{i<j}^{m}\left[\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}\sum_{i<j}^{m}\left(\frac{y_{ij}}{\sqrt{2}}\right)^2\right)dy_{ij}\right]$$
$$= (2\pi)^{m/2}\cdot(2\pi)^{m(m-1)/4}\left(\frac{1}{2}\right)^{m(m-1)/4} = 2^{m/2}\pi^{m(m+1)/4}$$

と計算できる（早川 毅教授との私信 (1994)）．これより所要の結果を得る．

例 5.2. 多変量ガンマ分布 (Multivariate gamma distribution)

$$f_\Gamma(\boldsymbol{A}\,;\boldsymbol{\Lambda}) = \frac{\varkappa^{bm}}{\Gamma_m(b)}|\boldsymbol{\Theta}^{-1}|^b\,|\boldsymbol{A}|^{b-(m+1)/2}\cdot\mathrm{etr}(-\varkappa\boldsymbol{\Theta}^{-1}\boldsymbol{A}),$$
$$[\boldsymbol{A}>0\,;\boldsymbol{\Lambda} := (\boldsymbol{\Theta}, \varkappa, b),\ \boldsymbol{\Theta}=\boldsymbol{\Theta}_{m\times m}>0,\ \varkappa,\ b>0,\ b>(m-1)/2],$$

ここに

$$\Gamma_m(b) = \int_{A>0}|\boldsymbol{C}|^b\,|\boldsymbol{A}|^{b-(m+1)/2}\,\mathrm{etr}\,(-\boldsymbol{C}\boldsymbol{A})(d\boldsymbol{A}) = \pi^{m(m-1)/4}\prod_{k=1}^{m}\Gamma\left(b-\frac{k-1}{2}\right), \qquad [\boldsymbol{C}>0].$$

誘導． 次の様に設定する：

$$\boldsymbol{\Delta}^*(\boldsymbol{A}_\Delta\,;\boldsymbol{\Lambda}) = \boldsymbol{A}_\Delta^{-1} - \boldsymbol{\Theta}_\Delta^{-1} + \left\{1-\frac{\varkappa}{b-(m+1)/2}\right\}\boldsymbol{\Theta}_\Delta^{-1}, \qquad \boldsymbol{\Xi} = \{b-(m+1)/2\}\boldsymbol{I}_{m\times m}.$$

ここで $\boldsymbol{A}_\Delta = \boldsymbol{T}^t\boldsymbol{A}\boldsymbol{T}$, $\boldsymbol{\Theta}_\Delta^{-1} = \boldsymbol{T}^t\boldsymbol{\Theta}^{-1}\boldsymbol{T}$, $\boldsymbol{T}^t\boldsymbol{T} = \boldsymbol{I}_{m\times m}$. よって $\boldsymbol{\Xi}(\boldsymbol{\Lambda})\boldsymbol{\Delta}^*(\boldsymbol{A}_\Delta\,;\boldsymbol{\Lambda})$ は対角．定理 4.1 より

$$\ln f_\Gamma(\boldsymbol{A}\,;\boldsymbol{\Lambda}) \propto \mathrm{tr}\left[\int\{b-(m+1)/2\}\boldsymbol{A}_\Delta^{-1}[d\boldsymbol{A}_\Delta] - \int\varkappa\boldsymbol{\Theta}_\Delta^{-1}[d\boldsymbol{A}_\Delta]\right]$$
$$= \mathrm{tr}\left[\frac{1}{m}\left(b-\frac{m+1}{2}\right)\ln|\boldsymbol{A}_\Delta| - \varkappa\boldsymbol{\Theta}_\Delta^{-1}\boldsymbol{A}_\Delta\right] = \mathrm{tr}\left[\frac{1}{m}\left(b-\frac{m+1}{2}\right)\ln|\boldsymbol{A}| - \varkappa\boldsymbol{T}^t\boldsymbol{\Theta}^{-1}\boldsymbol{T}\boldsymbol{T}^t\boldsymbol{A}\boldsymbol{T}\right].$$

即ち

$$f_\Gamma(\boldsymbol{A}\,;\boldsymbol{\Lambda}) \propto |\boldsymbol{A}|^{b-(m+1)/2}\cdot\mathrm{etr}[-\varkappa\boldsymbol{\Theta}^{-1}\boldsymbol{A}].$$

規格化定数を次の多変量ガンマ関数の定義式 (cf. Muirhead (1982)) から定める：

$$\int_{A>0}|\boldsymbol{A}|^{b-(m+1)/2}\,\mathrm{etr}\,(-\boldsymbol{L}\boldsymbol{A})(d\boldsymbol{A}) = \Gamma_m(b)|\boldsymbol{L}|^{-b}, \qquad (\boldsymbol{L}>0).$$

$\boldsymbol{L} = \varkappa\boldsymbol{\Theta}^{-1}$ と置いて

$$\frac{\varkappa^{bm}}{\Gamma_m(b)}\int_{A>0}|\boldsymbol{\Theta}^{-1}|^b|\boldsymbol{A}|^{b-(m+1)/2}\,\mathrm{etr}\,(-\varkappa\boldsymbol{\Theta}^{-1}\boldsymbol{A})(d\boldsymbol{A}) = 1,$$

故に

$$C_\Gamma(\boldsymbol{\Theta}, \varkappa, b) = \varkappa^{bm}|\boldsymbol{\Theta}^{-1}|^b/\Gamma_m(b).$$

注 5.1. $\varkappa=1/2$, $b=n/2$, $\boldsymbol{\Theta}=1/2\cdot\boldsymbol{\Sigma}$, $\boldsymbol{A}=\boldsymbol{W}$ と置けば Wishart 分布の pdf.を得る：

$$f_{\mathrm{W}}(\boldsymbol{W};\boldsymbol{\Sigma},n,p)=\frac{1}{\Gamma_m\left(\frac{n}{2}\right)2^{mn/2}|\boldsymbol{\Sigma}|^{n/2}}\operatorname{etr}\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\right)\cdot|\boldsymbol{W}|^{(n-m-1)/2},$$

$$(\boldsymbol{W},\boldsymbol{\Sigma}>0,\ n>0,\ n>m-1).$$

例 5.3. 多変量逆ガウス分布 (Multivariate inverse Gaussian distribution)

$$\begin{aligned}f_{\mathrm{IG}}(\boldsymbol{A};\boldsymbol{\Lambda})&=C_{\mathrm{IG}}(\boldsymbol{\Theta},b,\delta)\cdot|\boldsymbol{A}|^{\delta-(m+1)/2}\operatorname{etr}\left(-\frac{b}{2}\boldsymbol{\Theta}^{-2}\boldsymbol{A}\right)\cdot\operatorname{etr}\left(-\frac{b}{2}\boldsymbol{A}^{-1}\right)\\&=C_{\mathrm{IG}}(\boldsymbol{\Theta},b,\delta)\cdot\operatorname{etr}\left\{\frac{1}{m}\left(\delta-\frac{m+1}{2}\right)\boldsymbol{I}_{m\times m}\ln|\boldsymbol{A}|-\frac{b}{2}\boldsymbol{\Theta}^{-2}\boldsymbol{A}-\frac{b}{2}\boldsymbol{A}^{-1}\right\}\\&\qquad[\boldsymbol{A}>0;\boldsymbol{\Lambda}:=(\boldsymbol{\Theta},b,\delta),\ \boldsymbol{\Theta}>0,\ b>0,\ -\infty<\delta<\infty],\end{aligned}$$

ここに

$$C_{\mathrm{IG}}(\boldsymbol{\Theta},b,\delta)=\left[\int_{A>0}|\boldsymbol{A}|^{\delta-(m+1)/2}\operatorname{etr}\left(-\frac{b}{2}\boldsymbol{A}^{-1}\right)\operatorname{etr}\left(-\frac{b}{2}\boldsymbol{\Theta}^{-2}\boldsymbol{A}\right)(d\boldsymbol{A})\right]^{-1}.$$

誘導. 次の様に設定する：

$$\boldsymbol{\Delta}^*(\boldsymbol{A}_\Delta;\boldsymbol{\Lambda})=\boldsymbol{\Theta}_\Delta\boldsymbol{A}_\Delta^{-2}-\boldsymbol{\Theta}_\Delta^{-1}+\frac{2}{b}\boldsymbol{\Theta}_\Delta\cdot\frac{1}{m}\left(\delta-\frac{m+1}{2}\right)\boldsymbol{A}_\Delta^{-1},\qquad \boldsymbol{\Xi}=\frac{b}{2}\boldsymbol{\Theta}_\Delta^{-1}.$$

ここで $\boldsymbol{A}_\Delta=\boldsymbol{T}^t\boldsymbol{A}\boldsymbol{T}$, $\boldsymbol{\Theta}_\Delta^{-2}=\boldsymbol{T}^t\boldsymbol{\Theta}^{-2}\boldsymbol{T}$, $\boldsymbol{T}^t\boldsymbol{T}=\boldsymbol{I}_{m\times m}$. よって $\boldsymbol{\Xi}(\boldsymbol{\Lambda})\boldsymbol{\Delta}^*(\boldsymbol{A}_\Delta;\boldsymbol{\Lambda})$ は対角. 定理 4.1 より

$$\ln f_{\mathrm{IG}}(\boldsymbol{A};\boldsymbol{\Lambda})\propto\operatorname{tr}\left[\int\left\{\frac{b}{2}\boldsymbol{A}_\Delta^{-2}-\frac{b}{2}\boldsymbol{\Theta}_\Delta^{-2}+\frac{1}{m}\left(\delta-\frac{m+1}{2}\right)\boldsymbol{A}_\Delta^{-1}\right\}[d\boldsymbol{A}_\Delta]\right].$$

よって

$$\begin{aligned}f_{\mathrm{IG}}(\boldsymbol{A};\boldsymbol{\Lambda})&\propto\operatorname{etr}\left\{-\frac{b}{2}\boldsymbol{A}_\Delta^{-1}-\frac{b}{2}\boldsymbol{\Theta}_\Delta^{-2}\boldsymbol{A}_\Delta+\frac{1}{m}\left(\delta-\frac{m+1}{2}\right)\boldsymbol{I}_{m\times m}\ln|\boldsymbol{A}_\Delta|\right\}\\&=\operatorname{etr}\left(-\frac{b}{2}\boldsymbol{A}^{-1}\right)\operatorname{etr}\left(-\frac{b}{2}\boldsymbol{\Theta}^{-2}\boldsymbol{A}\right)|\boldsymbol{A}|^{\delta-(m+1)/2}.\end{aligned}$$

規格化定数 $C_{\mathrm{IG}}(\boldsymbol{\Theta},b,\delta)$ を求める. このために Herz (1955) の多変量変形ベッセル関数の結果を利用する：

$$B_\delta(\boldsymbol{\Theta},\alpha,\beta)=\frac{1}{2}\left(\frac{\beta}{\alpha}\right)^\delta|\boldsymbol{\Theta}|^\delta\int_{A>0}\operatorname{etr}(-\alpha\boldsymbol{A}^{-1})\operatorname{etr}\left(-\frac{\beta^2}{\alpha}(\boldsymbol{\Theta}^t\boldsymbol{\Theta})^{-1}\boldsymbol{A}\right)\cdot|\boldsymbol{A}|^{\delta-(m+1)/2}(d\boldsymbol{A})$$

$$(-\infty<\delta<\infty,\ \alpha>0,\ \boldsymbol{\Theta}>0)$$

$\alpha=\beta=b/2$ と置いて

$$B_\delta(\boldsymbol{\Theta},b/2,b/2)=\frac{1}{2}|\boldsymbol{\Theta}|^\delta\int_{A>0}\operatorname{etr}\left(-\frac{b}{2}\boldsymbol{A}^{-1}\right)\operatorname{etr}\left(-\frac{b}{2}\boldsymbol{\Theta}^{-2}\boldsymbol{A}\right)\cdot|\boldsymbol{A}|^{\delta-(m+1)/2}(d\boldsymbol{A})$$

よって規格化定数は

$$\begin{aligned}C_{\mathrm{IG}}(\boldsymbol{\Theta},b,\delta)&=\frac{1}{2}|\boldsymbol{\Theta}|^\delta\cdot[B_\delta(\boldsymbol{\Theta},b/2,b/2)]^{-1}\\&=\left[\int_{A>0}\operatorname{etr}\left(-\frac{b}{2}\boldsymbol{A}^{-1}\right)\operatorname{etr}\left(-\frac{b}{2}\boldsymbol{\Theta}^{-2}\boldsymbol{A}\right)\cdot|\boldsymbol{A}|^{\delta-(m+1)/2}(d\boldsymbol{A})\right]^{-1}.\end{aligned}$$

例 5.4. 多変量ランダムウオーク分布（多変量逆数逆ガウス分布）(Multivariate random-walk distribution; Multivariate reciprocal inverse Gaussian distribution)

$$f_{\mathrm{RW}}(\boldsymbol{A};\boldsymbol{\Lambda})=C_{\mathrm{RW}}(\boldsymbol{\Theta},b,\delta)\,|\boldsymbol{A}|^{-\delta-(m+1)/2}\mathrm{etr}\left\{-\frac{b}{2}(\boldsymbol{A}+\boldsymbol{\Theta}^2\boldsymbol{A}^{-1})\right\}$$
$$(\boldsymbol{A},\boldsymbol{\Theta}>0;\boldsymbol{\Lambda}:=(\boldsymbol{\Theta},b,\delta),\ b>0,\ -\infty<\delta<\infty),$$

ここに

$$C_{\mathrm{RW}}(\boldsymbol{\Theta},b,\delta)=\left[\int_{A>0}|\boldsymbol{A}|^{-\delta-(m+1)/2}\,\mathrm{etr}\left(-\frac{b}{2}(\boldsymbol{A}+\boldsymbol{\Theta}^2\boldsymbol{A}^{-1})\right)(d\boldsymbol{A})\right]^{-1}.$$

誘導. 次の様に設定する．

$$\boldsymbol{\Delta}^*(\boldsymbol{A}_\Delta;\boldsymbol{\Lambda})=\boldsymbol{\Theta}_\Delta\boldsymbol{A}_\Delta^{-2}-\boldsymbol{\Theta}_\Delta^{-1}-\frac{2}{b}\boldsymbol{\Theta}_\Delta^{-1}\frac{1}{m}\left(\delta+\frac{m+1}{2}\right)\boldsymbol{A}_\Delta^{-1},\qquad \boldsymbol{\Xi}=\frac{1}{2}b\boldsymbol{\Theta}_\Delta.$$

但し $\boldsymbol{A}_\Delta^{-1}=\boldsymbol{T}^t\boldsymbol{A}^{-1}\boldsymbol{T}$, $\boldsymbol{\Theta}_\Delta^2=\boldsymbol{T}^t\boldsymbol{\Theta}^2\boldsymbol{T}$, $\boldsymbol{T}^t\boldsymbol{T}=\boldsymbol{I}_{m\times m}$. ここでの $\boldsymbol{T}$ は可換な $\boldsymbol{A}^{-1}$, $\boldsymbol{\Theta}^2$ を同時に対角化する直交行列を表す (cf. 注 4.1)．従って，定理 4.1 より

$$f_{\mathrm{RW}}(\boldsymbol{A};\boldsymbol{\Lambda})\propto\mathrm{etr}\left[\int\left\{\frac{1}{2}b\boldsymbol{\Theta}_\Delta^2\boldsymbol{A}_\Delta^{-2}-\frac{1}{2}b-\frac{1}{m}\left(\delta+\frac{m+1}{2}\right)\boldsymbol{A}_\Delta^{-1}\right\}[d\boldsymbol{A}_\Delta]\right]$$
$$=\mathrm{etr}\left[-\frac{1}{2}b(\boldsymbol{\Theta}_\Delta^2\boldsymbol{A}_\Delta^{-1}+\boldsymbol{A}_\Delta)-\frac{1}{m}\left(\delta+\frac{m+1}{2}\right)\boldsymbol{I}_{m\times m}\ln|\boldsymbol{A}_\Delta|\right].$$

よって

$$f_{\mathrm{RW}}(\boldsymbol{A};\boldsymbol{\Lambda})\propto|\boldsymbol{A}|^{-\delta-(m+1)/2}\mathrm{etr}\left\{-\frac{b}{2}(\boldsymbol{A}+\boldsymbol{\Theta}^2\boldsymbol{A}^{-1})\right\}.$$

規格化定数を多変量変形ベッセル関数を利用して定める；

$$B(\boldsymbol{L},b,\delta)=\int_{U>0}\mathrm{etr}\left(-\frac{b}{2}\boldsymbol{U}^{-1}\right)\mathrm{etr}\left(-\frac{b}{2}\boldsymbol{L}^{-2}\boldsymbol{U}\right)\cdot|\boldsymbol{U}|^{\delta-(m+1)/2}(d\boldsymbol{U}),\qquad(\boldsymbol{L}>0).$$

$\boldsymbol{L}^{-1}=\boldsymbol{\Theta}$, $\boldsymbol{U}^{-1}=\boldsymbol{A}$ と置けばヤコビアンは $\boldsymbol{J}(\boldsymbol{U}\to\boldsymbol{A})=|\boldsymbol{A}|^{-(m+1)}$ であるから

$$B(\boldsymbol{\Theta}^{-1},b,\delta)=\int_{A>0}|\boldsymbol{A}|^{-\delta-(m+1)/2}\,\mathrm{etr}\left[-\frac{b}{2}(\boldsymbol{A}+\boldsymbol{\Theta}^2\boldsymbol{A}^{-1})\right](d\boldsymbol{A}).$$

従って

$$C_{\mathrm{RW}}(\boldsymbol{\Lambda})=\{B(\boldsymbol{\Theta}^{-1},b,\delta)\}^{-1}=\left[\int_{A>0}|\boldsymbol{A}|^{-\delta-(m+1)/2}\,\mathrm{etr}\left\{-\frac{b}{2}(\boldsymbol{A}+\boldsymbol{\Theta}^2\boldsymbol{A}^{-1})\right\}(d\boldsymbol{A})\right]^{-1}.$$

例 5.5. 多変量逆数ガンマ分布 (Multivariate reciprocal gamma distribution)

$$f_{\mathrm{R\Gamma}}(\boldsymbol{A};\boldsymbol{\Lambda})=\frac{\tau^{m\alpha}}{\Gamma_m(\alpha)}\cdot|\boldsymbol{\Theta}|^{\alpha}|\boldsymbol{A}^{-1}|^{\alpha+(m+1)/2}\cdot\mathrm{etr}(-\tau\boldsymbol{\Theta}\boldsymbol{A}^{-1})$$
$$=\frac{\tau^{m\alpha}}{\Gamma_m(\alpha)}\cdot|\boldsymbol{\Theta}|^{\alpha}\cdot\mathrm{etr}\left\{-\tau\boldsymbol{\Theta}\boldsymbol{A}^{-1}-\left(\alpha+\frac{m+1}{2}\right)\ln|\boldsymbol{A}|\cdot\frac{1}{m}\boldsymbol{I}_{m\times m}\right\},$$
$$(\boldsymbol{A},\boldsymbol{\Theta}>0;\boldsymbol{\Lambda}:=(\boldsymbol{\Theta},\tau,\alpha);\tau,\alpha>0).$$

誘導. 次の様に設定する：

$$\boldsymbol{\Delta}^*(\boldsymbol{A}_\Delta;\boldsymbol{\Lambda})=\boldsymbol{\Theta}_\Delta\boldsymbol{A}_\Delta^{-2}-\{\alpha+(m+1)/2\}/\tau\cdot\boldsymbol{A}_\Delta^{-1},\qquad \boldsymbol{\Xi}=\tau\boldsymbol{I}_{m\times m},$$

ここで $\boldsymbol{A}_\Delta^{-1}=\boldsymbol{T}^t\boldsymbol{A}^{-1}\boldsymbol{T}$, $\boldsymbol{\Theta}_\Delta=\boldsymbol{T}^t\boldsymbol{\Theta}\boldsymbol{T}$, $\boldsymbol{T}^t\boldsymbol{T}=\boldsymbol{I}_{m\times m}$ と同時対角化した．従って $\boldsymbol{\Xi}(\boldsymbol{\Lambda})\boldsymbol{\Delta}^*(\boldsymbol{A}_\Delta;\boldsymbol{\Lambda})$ は対角となる．定理 4.1 から

$$f_{\mathrm{RT}}(\boldsymbol{A};\boldsymbol{\Lambda})\propto \mathrm{etr}\left[\int \tau[\boldsymbol{\Theta}_\Delta\boldsymbol{A}_\Delta^{-2}-\{\alpha+(m+1)/2\}/\tau\cdot\boldsymbol{A}_\Delta^{-1}][d\boldsymbol{A}_\Delta]\right]$$
$$=\mathrm{etr}[-\tau\boldsymbol{\Theta}_\Delta\boldsymbol{A}_\Delta^{-1}-\{\alpha+(m+1)/2\}\ln\boldsymbol{A}_\Delta].$$

従って

$$f_{\mathrm{RT}}(\boldsymbol{A};\boldsymbol{\Lambda})\propto \exp[-\mathrm{tr}\,\tau(\boldsymbol{T}^t)^{-1}\boldsymbol{\Theta}_\Delta(\boldsymbol{T})^{-1}(\boldsymbol{T}^t)^{-1}\boldsymbol{A}_\Delta^{-1}(\boldsymbol{T})^{-1}-\mathrm{tr}\{\alpha+(m+1)/2\}\ln\boldsymbol{A}]$$
$$=|\boldsymbol{A}|^{-\alpha-(m+1)/2}\cdot\mathrm{etr}(-\tau\boldsymbol{\Theta}\boldsymbol{A}^{-1}).$$

規格化定数 $C_{\mathrm{RT}}(\boldsymbol{\Lambda})$ を次の多変量ガンマ関数の公式を利用して求める：$\boldsymbol{U},\boldsymbol{K}$ を $m\times m$ の対称正則行列とすると

$$\int_{U>0}|\boldsymbol{U}|^{\alpha-(m+1)/2}\,\mathrm{etr}\,(-\boldsymbol{K}\boldsymbol{U})(d\boldsymbol{U})=\Gamma_m(\alpha)|\boldsymbol{K}|^{-\alpha}\qquad[\mathrm{Re}(\alpha)>(m-1)/2,\ \boldsymbol{K}>0].$$

$\boldsymbol{U}=\boldsymbol{A}^{-1}$ と置けば $J(\boldsymbol{U}\to\boldsymbol{A})=|\boldsymbol{A}|^{-(m+1)}$ だから

$$\int_{A>0}|\boldsymbol{A}|^{-\alpha-(m+1)/2}\,\mathrm{etr}\,(-\boldsymbol{K}\boldsymbol{A}^{-1})(d\boldsymbol{A})=\Gamma_m(\alpha)|\boldsymbol{K}|^{-\alpha}.$$

$\boldsymbol{K}=\tau\boldsymbol{\Theta}$ と置けば

$$C_{\mathrm{RT}}(\boldsymbol{\Lambda})=|\boldsymbol{\Theta}|^{\alpha}\tau^{m\alpha}/\Gamma_m(\alpha).$$

注 5.2. $\alpha=\tau=m/2$ の時，本例は多変量片側安定分布になる．

6. あとがき

本稿ではノンパラメトリックな場合について統計的基礎方程式を導入することにより，いくつかの重要な多変量確率モデルを作り出せることを示した．パラメトリックな場合についても並行する議論が可能である (cf. Matsunawa (1992))．パラメータという調整可能な部分を有する基礎モデルを考える分だけ，きめ細かな議論が可能となることが多い．いずれの場合も，母集団分布の存在を仮定する Fisher の特定化の問題を見直し，データあるいは適切な思考実験を通じて統計基礎モデルを構築するという意味での分布の特定化の問題に取り組むための有力な道具立となることが分かった．このようにして得られた統計基礎モデル系の状態の変化にともなってどう発展するかは，関連する統計的特性関数の Legendre 変換を利用し，正準パラメータ等の統計的モデルの変化を正準情報量規準を導入して考察することができる．それについては稿を改めて議論する．

謝 辞

本稿を細部に渡り検討し貴重なコメントをして頂いた査読者および担当編集委員に深く感謝します．また例 5.1 (ii) の積分に関しご教示頂いた早川 毅氏に感謝します．

参 考 文 献

Aitken, A.C. and Silverstone, H. (1942). On the estimation of statistical parameters, *Proc. Roy. Soc.*

Edinburgh Sect. A, **61**, 186-194.

Bellman, R. (1970). *Introduction to Matrix Analysis*, 2nd ed., McGraw-Hill, New York.

Cramér, H. (1946). *Mathematical Methods of Statistics*, Princeton University Press, New Jersey.

Edgeworth, F.Y. (1908). On the probable errors of frequency-constants II, *J. Roy. Statist. Soc.*, **71**, 651-678.

Elderton, W.P. and Johnson, N.L. (1969). *Systems of Frequency Curves*, Cambridge University Press, Massachusetts.

Fisher, R.A. (1922). On the mathematical foundation of theoretical statistics, *Philos. Trans. Roy. Soc. London Ser. A*, **222**, 309-368.

Fisher, R.A. (1936). Uncertain inference, *Proceedings of American Academy of Arts and Science,* **71**, 245-258.

Gauss, C.F. (1809). *Theoria Motus Corporum Coelestium*, Perthes & Besser (English translation by C.H. Davis (1857), Little Brown, Boston).

Heisenberg, W. (1927). Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik, *Z. Phys.*, **43**, 172-198.

Heisenberg, W. (1930). *Die Physikalischen Prinzipien der Quantentheorie*, Hirzel, Leipzig.

Herz, C.S. (1955). Bessel functions of matrix argument, *Ann. of Math.*, **61**, 474-523.

Kagan, A.M., Linnik, Yu. V. and Rao, C.R. (1973). *Characterization Problems in Mathematical Statistics*, Wiley, New York.

Keynes, J.M. (1911). The principal averages and the laws of error which lead to them, *J. Roy. Statist. Soc.*, **24**, 322-331.

Keynes, J.M. (1962). *A Treaties on Probability*, 194-205, Harper & Row, New York.

Matsunawa, T. (1992). Parametric statistical uncertainty relation, statistical fundamental equation and specification of model distributions, Research Memo., No. 465, The Institute of Statistical Mathematics, Tokyo.

Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory*, Wiley, New York.

Ord, J.K. (1967). On a system of discrete distributions, *Biometrika*, **54**, 649-656.

Pearson, K. (1895). Contribution to the mathematical theory of evolution II. Skew variation in homogeneous material, *Philos. Trans. Roy. Soc. London Ser. A*, **186**, 343-414.

Pearson, K. (1916). Contribution to the mathematical theory of evolution XIX, Second supplement to a memoir on Skew variation, *Philos. Trans. Roy. Soc. London Ser. A*, **216**, 429-457.

Poincaré, H. (1912). *Calcul des Probabilités*, 2nd ed., Gauthier-Villars, Paris.

Rao, C.R. (1945). Information and the accuracy attainable in the estimation of statistical parameters, *Bull. Calcutta Math. Soc.*, **37**, 81-91.

Schrödinger, E. (1926). Quantisierung als Eigenwertproblem, *Annalen der Physik*, **81**, 109-139.

Origin of Distributions
—— Nonparametric Statistical Uncertainty Relation and a Statistical Fundamental Equation ——

Tadashi Matsunawa

(The Institute of Statistical Mathematics)

A nonparametric statistical uncertainty relation is proved to specify nonparametric multivariate statistical models. The relation is represented by an inequality which is closely related to Love's inequality introduced by Edgeworth in 1908. It is shown that we can not simultaneously correctly describe both of a function of observation objects and a nonparametric statistical model. However, as a special case of the relation, we can get a nonparametric statistical fundamental equation, which is obtained as the condition of attainment of the equality sign in the relation. Making use of the result a generalized multivariate exponential family is derived as a family of nonparametric statistical uncertainty distributions. Some important multivariate distributions are constructed.

Key words: Nonparametric statistical uncertainty relation, nonparametric statistical fundamental equation, specification of model distribution, nonparametric specific intensity of model performance, nonparametric minimum statistical uncertainty distribution, multivariate distributions.