

直交計画を用いた複数モデルの 同時実験に関する研究

浮田 善文

- § 1. はじめに
- § 2. 準備
- § 3. 実験計画法
- § 4. 複数モデルの同時実験
- § 5. ベイズ理論による枠組み
- § 6. 実行例による同時実験の性能評価
- § 7. 考察
- § 8. まとめ

梗概

本稿では、類似する複数目的に対し、各目的に含まれる考慮すべき要因を全て調べる事が可能な同時実験を提案する。実験計画法のモデルについて考えると、これまで広く用いられている伝統的なモデルは、パラメータが実数値で各要因の効果が分かりやすいが、パラメータ間に制約がおかれているという特徴がある。同時実験を行う場合、従来に比べ因子数が非常に大きい場合を扱うため、プログラミングを考えると、パラメータ間の制約は無いほうが望ましい。近年、実験計画法のモデルは、全てのパラメータが独立となる直交基底関数モデルでも表現可能であることが示されている。そこで本稿では、扱う因子数およびパラメータの次元数が非常に大きい場合のプログラミングに、伝統的なモデルと直交基底関数モデルの両方を用いる新たな方法を提案する。本稿で提案する同時実験の性能評価であるが、個別に直交計画を実験する場合の実験回数の総和と、同時実験で直交計画を実験する場合の実験回数を比較し、同時実験を利用することで実験回数を減らすことが可能であることを、実行例により示す。さらに、ベイズ理論による枠組みでのパラメータ推定を考え、同時実験は個別実験よりも、事後平均が真のパラメータに近いことを示す。最後に、同時実験についての考察を行い、その特性を明らかにする。

キーワード：実験計画法，線形基底関数モデル，データ収集コスト，ベイズ線形回帰，事後平均，プログラミング

第1章 はじめに

実験計画法は、様々な要因が目標となる特性値にどのように影響を与えているか、またそれぞれの要因をどのような値にすれば特性値が最適になるかなどを明らかにする統計的実験手法である [1, 2, 3]. 実験計画法はもともとは農業実験から生まれたものであるが、現在では自然科学の多くの分野で広く活用されており、様々な応用先で不可欠な手法となっている.

ここで要因のうち、実験で取り上げ、水準を設定し比較される項目を因子と呼ぶ. 実験する水準組合せの集合は実験計画と呼ばれ、実験計画の中でも直交計画は、同じ大きさの実験計画の中で各効果の不偏推定量の分散を最小にすることが知られており [4, 5, 6], 実用上非常に重要である. また直交計画は応用先も多く、最適解の探索 [7] やソフトウェアテスト [8], 近年ではマテリアルズインフォマティクス [9] などでも利用されている.

実験計画法については、これまで国内外で多くの研究が行われているが、コスト面を考えた場合、特に医療・ヘルスケア分野を中心に、データ収集にかかるコストは依然、大きな問題である. 一例として、新薬開発を考えると、基礎研究で様々な候補物質が創成され、非臨床試験、臨床試験の順に、候補物質の有効性と安全性を確認することになる [10]. 各製薬会社が同種の新薬を開発する場合であっても全く同じとは限らないため、実験により特性値への影響を調べたい要因の集合も異なる. このため、調べたい要因の集合は比較的類似であるにも関わらず、実際には、各製薬会社ごとに実験に高額なコストをかけているのが実状である.

ここでもし実験計画法において、類似する複数目的に対して同時実験を行うことができれば、より少ない実験回数、コストでの実験の実施が期待できる. 一般に、複数目的に対する実験を行うのであれば因子数も大きくなることが考えられる.

本稿では、類似する複数目的に対し、各目的に含まれる考慮すべき要因を全て調べることが可能な同時実験を提案する. 実験計画法のモデルについて考えると、これまで広く用いられている伝統的なモデルは、パラメータが実数値で各要因の効果が分かりやすいが、パラメータ間に制約がおかれているという特徴がある. 同時実験を行う場合、従来に比べ因子数が非常に大きい場合を扱うため、プログラミングを考えると、パラメータ間の制約は無いほうが望ましい. 近年、実験計画法のモデルは、全てのパラメータが独立となる直交基底関数モデルでも表現可能であることが示されている [11]. さらに、伝統的なモデルと直交基底関数モデルが1対1で対応し、その線形変換の関係式も明らかとなっている [12].

そこで本稿では、扱う因子数およびパラメータの次元数が非常に大きい場合のプログラミングに、伝統的なモデルと直交基底関数モデルの両方を用いる新たな方法を提案する.

具体的には、各要因の効果を知りたいときやパラメータの事前確率分布を導入するときは伝統的なモデルを用い、途中のパラメータ推定やプログラミングには直交基底関数モデルを用いる。また直交基底関数モデルを使うことで、ベイズ線形回帰 [13] に関する結果も利用可能となる。本稿で提案する同時実験の性能評価であるが、個別に直交計画を実験する場合の実験回数の総和と、同時実験で直交計画を実験する場合の実験回数を比較し、同時実験を利用することで実験回数を減らすことが可能であることを、実行例により示す。さらに、ベイズ理論による枠組み [14] でのパラメータ推定を考え、同時実験は個別実験よりも、事後平均が真のパラメータに近いことを示す。最後に、同時実験についての考察を行い、その特性を明らかにする。

本論文は以下のように構成されている。2章では準備として、指標による関数の表現、ガロア体上の直交基底関数などについて述べる。3章では、実験計画法について概説する。必要となる用語を説明したあと、伝統的なモデルと直交基底関数モデルの2つについて説明する。4章では、本稿で提案する複数モデルの同時実験について説明する。また、データ収集のコストとパラメータの推定精度についても述べる。さらに同時実験のプログラミングのための2つのモデルの利用の流れについても説明する。5章で、ベイズ理論の枠組みについて述べる。6章では、同時実験の性能評価として、同時実験は個別実験よりも実験回数を減らすことができることを実行例により示す。さらに、同時実験は個別実験よりも、事後平均が真のパラメータに近いことを示す。7章では、実行例に対して考察を行い、8章でまとめと今後の課題を述べる。

第2章 準備

本稿では、実験計画において、直交基底関数により表現される線形基底関数モデルを考える。そこで、この章では3章以降の内容で必要となる、指標による関数の表現、ガロア体上の直交基底関数などについて説明する。

2.1 指標による関数の表現

G を（演算を掛け算の記号で表した）有限アーベル群であるとし、 S^1 を複素平面内の単位円周とする。 G 上の指標 [15] とは、複素数値関数 $\chi: G \rightarrow S^1$ で、次の条件を満たすもののことである：

$$\chi(x \cdot x') = \chi(x)\chi(x') \quad x, x' \in G \quad (1)$$

また、 G の指標群は $\{\chi_a(x) | a \in G\}$ で表すことができ、次式が成立する。

$$\frac{1}{|G|} \sum_{x \in G} \chi_a(x) \chi_b^*(x) = \begin{cases} 1, & a = b \\ 0, & a \neq b \end{cases} \quad (2)$$

ここで $\chi_b^*(x)$ は $\chi_b(x)$ の複素共役である。このとき、任意の関数 $f: G \rightarrow C$ (C は複素数

体) は, 指標の線形結合により,

$$f(\mathbf{x}) = \sum_{\mathbf{a} \in G} f_{\mathbf{a}} \chi_{\mathbf{a}}(\mathbf{x}) \quad (3)$$

で一意に表現可能で, 係数 $f_{\mathbf{a}}$ は次式で与えられる.

$$f_{\mathbf{a}} = \frac{1}{|G|} \sum_{\mathbf{x} \in G} f(\mathbf{x}) \chi_{\mathbf{a}}^*(\mathbf{x}) \quad (4)$$

2.2 ガロア体上の直交基底関数

q 個の元からなる体をガロア体 [16] といい, $GF(q)$ で表す. なお, ガロア体の位数 q は, $q = p^m$ でなければならない. ただし, p は素数である.

次に長さ n のベクトルを考え, $\mathbf{x} = (x_1, x_2, \dots, x_n)$, $\mathbf{a} = (a_1, a_2, \dots, a_n)$, $\mathbf{b} = (b_1, b_2, \dots, b_n)$ とする. ここで, $\mathbf{x}, \mathbf{a}, \mathbf{b} \in GF(q)^n$ としておく.

例 1. [11] $q = 3, n = 5$ の場合を考える. このとき, $GF(3)^5 = \{00000, 10000, \dots, 22222\}$ であり, $|GF(3)^5| = 243$ となる.

ガロア体上においても式 (2) は成立し, 次式で与えられる [11].

$$\frac{1}{q^n} \sum_{\mathbf{x} \in GF(q)^n} \chi_{\mathbf{a}}(\mathbf{x}) \chi_{\mathbf{b}}^*(\mathbf{x}) = \begin{cases} 1, & \mathbf{a} = \mathbf{b} \\ 0, & \mathbf{a} \neq \mathbf{b} \end{cases} \quad (5)$$

式 (3), 式 (4) についても同様に成立する.

なお, 詳しくは次章で説明するが, 本稿では, $\chi_{\mathbf{a}}(\mathbf{x})$ を直交基底関数とする直交基底関数モデルを扱う.

例 2. $q = 3$ の場合を考える. このとき, 直交基底関数 $\chi_{\mathbf{a}}(\mathbf{x})$ は $\chi_{\mathbf{a}}(\mathbf{x}) = e^{i2\pi \mathbf{a} \cdot \mathbf{x} / 3}$ と書くことができる [15].

2.3 ガウス分布

N 次元ベクトルの連続変数に対して定義されるガウス分布は

$$\begin{aligned} \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) &= \frac{1}{(2\pi)^{N/2} |\boldsymbol{\Sigma}|^{1/2}} \\ &\times \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} \end{aligned} \quad (6)$$

で与えられ, N 次元ベクトル $\boldsymbol{\mu}$ は平均, $N \times N$ 行列 $\boldsymbol{\Sigma}$ は共分散と呼ばれ, $|\boldsymbol{\Sigma}|$ は $\boldsymbol{\Sigma}$ の行列式を表す. また, T はベクトルまたは行列の転置を表す.

第3章 実験計画法

この章では、実験計画法の概要について説明する。まず必要な用語を与えた後、実験計画法の伝統的なモデルについて述べる。このモデルはパラメータが実数値で各要因の効果が分かりやすく広く使われているが、パラメータ間に制約がおかれているという特徴がある。次に、全てのパラメータが独立な直交基底関数モデルについて説明する。さらに、両モデル間の関係式を与える。最後に、実験形式と直交計画についても説明する。

3.1 必要となる用語の説明

はじめに、説明に必要な記号を定義しておく。長さ n のベクトル \mathbf{a} の 0 でない成分の数を \mathbf{a} のハミング重みといい、 $H_w(\mathbf{a})$ で表すものとする。

F_1, F_2, \dots, F_n をモデルに含まれる n 個の因子とする。本稿では、 n 個の因子は全て同じ水準数を持つものとし、その水準数を q としておく。このとき、各因子の水準は $GF(q)$ により表現され、水準組合せは、 $\mathbf{x} = (x_1, x_2, \dots, x_n) \in GF(q)^n$ で表現される。交互作用については、2 因子間の交互作用の一部のみが考えられ、3 因子以上の交互作用は含まれないものとする^{*1}。

ここで、集合 $A \subseteq \{0, 1\}^n$ はモデルに含まれる一般平均、因子、因子間の交互作用を表す集合とする。因子 F_i は $H_w(\mathbf{a}) = 1, a_i = 1$ となる \mathbf{a} により表され、因子 F_i と因子 F_j の交互作用は $H_w(\mathbf{a}) = 2, a_i = 1, a_j = 1$ となる \mathbf{a} により表されるものとする。また、一般平均は $H_w(\mathbf{a}) = 0$ となる \mathbf{a} により表されるものとする。

例えば、 $A = \{000, 100, 010, 001, 110\}$ の要素は順番に、一般平均、因子 F_1 、因子 F_2 、因子 F_3 、因子 F_1 と因子 F_2 の交互作用を示している。

3.2 伝統的なモデル [3]

集合 A に対し、交互作用のある 2 因子の番号対の集合を J_A とする。 $\mathbf{a} = (a_1, a_2, \dots, a_n) \in A$ とすると、 J_A は $J_A = \{\{i, j\} | a_i \neq 1, a_j \neq 1, H_w(\mathbf{a}) = 2, \mathbf{a} \in A\}$ で与えられる。

水準組合せ \mathbf{x} で実験を行ったときの出力である目標値を $t(\mathbf{x})$ で表し、以下のモデルを仮定する。

$$t(\mathbf{x}) = \mu + \sum_{i=1}^n \alpha_i(x_i) + \sum_{\{i,j\} \in J_A} \beta_{i,j}(x_i, x_j) + \epsilon \quad (7)$$

ここで

^{*1} ただし、本稿の結果を 3 因子以上の交互作用を含む場合へ拡張することは容易である。

- μ : 一般平均
- $\alpha_i(x_i)$: 因子 F_i を水準 x_i としたときの主効果
- $\beta_{i,j}(x_i, x_j)$: 因子 F_i を水準 x_i とし, 因子 F_j を水準 x_j としたときの交互作用効果
- ϵ : 平均 0, 分散 σ^2 のガウス確率変数

である.

主効果と交互作用効果については, 各因子の中での水準の効果の相対的な差を知ることが目的であるため, 一般に以下の仮定がおかれている [5].

$$\sum_{\varphi=0}^{q-1} \alpha_i(\varphi) = 0 \quad (8)$$

$$\sum_{\varphi=0}^{q-1} \beta_{i,j}(\varphi, \psi) = 0 \quad (9)$$

$$\sum_{\psi=0}^{q-1} \beta_{i,j}(\varphi, \psi) = 0 \quad (10)$$

ここで, 因子数 n , 水準数 q , 含まれる 2 因子交互作用の数を n_I とするときの独立なパラメータ数 K は, 上記の制約により, $K = 1 + (q-1) \times n + (q-1)^2 \times n_I$ で与えられる.

これらの独立なパラメータのみをまとめて列ベクトル \mathbf{u} で表すとする. このとき, 式 (7) の $\mu + \sum_{i=1}^n \alpha_i(x_i) + \sum_{\{i,j\} \in J_A} \beta_{i,j}(x_i, x_j)$ は, \mathbf{x} と \mathbf{u} の関数とみることができるため,

$$y(\mathbf{x}, \mathbf{u}) = \mu + \sum_{i=1}^n \alpha_i(x_i) + \sum_{\{i,j\} \in J_A} \beta_{i,j}(x_i, x_j) \quad (11)$$

とおく. これより, 式 (7) は \mathbf{u} を用い, 以下のように表現可能である.

$$t(\mathbf{x}) = y(\mathbf{x}, \mathbf{u}) + \epsilon \quad (12)$$

例 3. [12] $n = 2, q = 3, A = \{00, 10, 01, 11\}$ の場合を考える. このとき, パラメータは式 (7) より,

$$\mu, \alpha_1(0), \alpha_1(1), \alpha_1(2), \alpha_2(0), \alpha_2(1), \alpha_2(2), \beta_{1,2}(0, 0),$$

$$\beta_{1,2}(0, 1), \beta_{1,2}(0, 2), \beta_{1,2}(1, 0), \beta_{1,2}(2, 0), \beta_{1,2}(1, 1),$$

$\beta_{1,2}(1, 2), \beta_{1,2}(2, 1), \beta_{1,2}(2, 2)$ の合計 16 個である. ただし, 制約式 (8)-(10) により, 独立なパラメータ数は $K = 9$ である. このとき, 独立なパラメータのみをまとめた列ベクトル \mathbf{u} は次式で与えられる.

$$\mathbf{u} = [\mu, \alpha_1(0), \alpha_1(1), \alpha_2(0), \alpha_2(1), \beta_{1,2}(0, 0), \beta_{1,2}(0, 1), \beta_{1,2}(1, 0), \beta_{1,2}(1, 1)]^T \quad (13)$$

なお, \mathbf{u} の要素は実数であり, $\mathbf{u} \in \mathbb{R}^9$ である.

3.3 直交基底関数モデル [11]

集合 A に対し, $I_A = \{(b_1 a_1, b_2 a_2, \dots, b_b a_n) | \mathbf{a} \in A, b_i \in GF(q)\}$ と定義すると, 集合 A に含まれる一般平均は f_0 , 因子 F_i の主効果は $\{f_{\mathbf{a}} | H_w(\mathbf{a}) = 1, a_i \neq 0, \mathbf{a} \in I_A\}$, 因子 F_i と因子 F_j の交互作用効果は $\{f_{\mathbf{a}} | H_w(\mathbf{a}) = 2, a_i \neq 0, a_j \neq 0, \mathbf{a} \in I_A\}$ により表現される. このとき, 以下のモデルを仮定する.

$$t(\mathbf{x}) = \sum_{\mathbf{a} \in I_A} f_{\mathbf{a}} \chi_{\mathbf{a}}(\mathbf{x}) + \epsilon \quad (14)$$

ここで ϵ は平均 0, 分散 σ^2 のガウス確率変数である.

このとき, パラメータ数 K は $K = |I_A|$ であり, 全て独立なパラメータである. これらのパラメータ $f_{\mathbf{a}}, (\mathbf{a} \in I_A)$ をまとめて列ベクトル \mathbf{w} で表す.

式 (14) の $\sum_{\mathbf{a} \in I_A} f_{\mathbf{a}} \chi_{\mathbf{a}}(\mathbf{x})$ は \mathbf{x} と \mathbf{w} の関数とみることができ, $y(\mathbf{x}, \mathbf{w}) = \sum_{\mathbf{a} \in I_A} f_{\mathbf{a}} \chi_{\mathbf{a}}(\mathbf{x})$ とおく. これより, 式 (14) は \mathbf{w} を用い, 以下のように表現可能である.

$$t(\mathbf{x}) = y(\mathbf{x}, \mathbf{w}) + \epsilon \quad (15)$$

例 4. [12] $n = 2, q = 3, A = \{00, 10, 01, 11\}$ の場合を考える.

$I_A = \{00, 10, 20, 01, 02, 11, 12, 21, 22\}$ となる. このとき, パラメータ数は 9 であり, 全て独立なパラメータである. また, \mathbf{w} は次式で与えられる.

$$\mathbf{w} = [f_{00}, f_{10}, f_{20}, f_{01}, f_{02}, f_{11}, f_{12}, f_{21}, f_{22}]^T \quad (16)$$

なお, \mathbf{w} の要素は複素数であり, $\mathbf{w} \in \mathbb{C}^9$ である.

3.4 両モデル間の関係 [12]

ここまで 2 つのモデルを個別に見てきたが, 伝統的なモデルと直交基底関数モデルには 1 対 1 の対応があり, パラメータ間に以下の関係を満たす行列 \mathbf{M} を構成することができる.

$$\mathbf{u} = \mathbf{M}\mathbf{w} \quad (17)$$

また, \mathbf{M} の逆行列 \mathbf{M}^{-1} も存在し, 以下の関係を満たす行列 \mathbf{M}^{-1} も構成できる.

$$\mathbf{w} = \mathbf{M}^{-1}\mathbf{u} \quad (18)$$

一般の場合について, 行列 \mathbf{M} と \mathbf{M}^{-1} を付録に載せるが, 以下では前述の例を用い, 両モデルの関係を示す.

例 5. [12]

式 (13) の \mathbf{u} と式 (16) の \mathbf{w} に対し, $\mathbf{u} = \mathbf{M}\mathbf{w}$ を満たす行列 \mathbf{M} は以下で与えられる.

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \omega_3 & \omega_3^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_3 & \omega_3^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & \omega_3 & \omega_3^2 & \omega_3 & \omega_3^2 \\ 0 & 0 & 0 & 0 & 0 & \omega_3 & \omega_3 & \omega_3^2 & \omega_3^2 \\ 0 & 0 & 0 & 0 & 0 & \omega_3^2 & 1 & 1 & \omega_3 \end{bmatrix}$$

ここで, $\omega_3 = e^{2\pi i/3}$ である.

また, $\mathbf{w} = \mathbf{M}^{-1}\mathbf{u}$ を満たす行列 \mathbf{M}^{-1} は以下で与えられる.

$$\mathbf{M}^{-1} = \begin{bmatrix} 1 & & & 0 \\ & S_1 & & \\ & & S_1 & \\ 0 & & & S_2 \end{bmatrix} \quad (19)$$

ここで, 行列 S_1 と行列 S_2 はそれぞれ以下となる.

$$S_1 = \frac{1}{3} \begin{bmatrix} 1 - \omega_3 & \omega_3^2 - \omega_3 \\ 1 - \omega_3^2 & \omega_3 - \omega_3^2 \end{bmatrix} \quad (20)$$

$$S_2 = \frac{1}{3} \begin{bmatrix} -\omega_3 & \omega_3^2 & \omega_3^2 & -1 \\ 1 & -\omega_3^2 & -\omega_3 & 1 \\ 1 & -\omega_3 & -\omega_3^2 & 1 \\ -\omega_3^2 & \omega_3 & \omega_3 & -1 \end{bmatrix} \quad (21)$$

3.5 実験形式

本稿においては, 集合 A は事前に与えられるものとする. このもとで, まず実験する水準組合せの集合である実験計画 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ を決定する. その後, 実験を行い, 対応する目標値 $t(\mathbf{x}_1), \dots, t(\mathbf{x}_N)$ を得る. また, 目標値をまとめて列ベクトル $\mathbf{t} = [t(\mathbf{x}_1), \dots, t(\mathbf{x}_N)]^T$ で表す. 実験後には, 得られた N 個の入出力の対であるデータ集合 $\{(\mathbf{x}, t(\mathbf{x})) | \mathbf{x} \in \mathbf{X}\}$ から, 各効果の推定等を行う.

このため, 大きさ N の実験計画 \mathbf{X} をどのように決めるかが重要となる. 直交計画は, 同じ大きさの実験計画の中で各効果の不偏推定量の分散を最小にすることが知られている [4, 5, 6].

そこで, 本稿では, 実験計画を直交計画とした場合について話を進める. 次の 3.6 節で直交計画について説明する.

3.6 直交計画 [11]

$v(\mathbf{a}) = \{i | a_i \neq 0, 1 \leq i \leq n\}$ とする. 集合 A に対し, $k_A \times n$ 行列 G_A を

$$G_A = \begin{bmatrix} g_{1,1} & g_{1,2} & \cdots & g_{1,n} \\ g_{2,1} & g_{2,2} & \cdots & g_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{k_A,1} & g_{k_A,2} & \cdots & g_{k_A,n} \end{bmatrix} \quad (22)$$

$g_{i,j} \in GF(q) (1 \leq i \leq k_A, 1 \leq j \leq n)$ とし, 以下の二つの条件を満たす行列とする.

1. \mathbf{g}_j を G_A の j 番目の列とすると, 任意の $\mathbf{a}', \mathbf{a}'' \in A$ に対し, 集合 $\{\mathbf{g}_j | j \in v(\mathbf{a}' + \mathbf{a}'')\}^{*2}$ が $GF(q)$ 上で線形独立
2. \mathbf{g}_i を G_A の i 番目の行とすると, 集合 $\{\mathbf{g}_i | i \leq k_A\}$ が $GF(q)$ 上で線形独立

このとき, 集合 A に対する直交計画 \mathbf{X}_A は次式で定義される.

$$\mathbf{X}_A = \{\mathbf{x} | \mathbf{g} = \mathbf{r}G_A, \mathbf{r} \in GF(q)^{k_A}\} \quad (23)$$

このとき, 直交計画の大きさ, すなわち実験回数は $|\mathbf{X}_A| = q^{k_A}$ である.

第4章 複数モデルの同時実験

この章では, 複数のモデルに対し, 同時に実験を行うことで, データ収集のコストを削減することを考える^{*3}. なお, 実験計画には前章で説明した直交計画を用いる.

4.1 複数モデルの同時実験

はじめに, L 個の伝統的なモデル $\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_L$ が存在する場合について考える^{*4}. このとき, $\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_L$ に含まれる全ての因子と全ての交互作用を含む新たな伝統的なモデル \mathcal{M} が存在すると仮定する. ここで, \mathcal{M}_l に対する直交計画を \mathbf{X}_{A_l} とし, \mathcal{M} に対する直交計画を \mathbf{X}_A としておく. このとき, モデル \mathcal{M} に対する実験を, 複数モデルの同時実験と呼ぶ.

例として, $L = 6$ の場合について, モデルごとの実験を図1に, 複数モデルの同時実験を図2に示す. 図2より, 同時実験の実験回数 $|\mathbf{X}_A|$ は, 6つのモデルの類似度, 言い換

^{*2} $\mathbf{a}' = (a'_1, a'_2, \dots, a'_n), \mathbf{a}'' = (a''_1, a''_2, \dots, a''_n)$ とすると, ベクトル間の加法は, 排他的論理和 \oplus を用い, $\mathbf{a}' + \mathbf{a}'' = (a'_1 \oplus a''_1, a'_2 \oplus a''_2, \dots, a'_n \oplus a''_n)$ で定義される.

^{*3} 本章の内容は, 素案として発表 [17] したものを発展させたものである.

^{*4} このとき, 3.4節で述べたように伝統的なモデルと直交基底関数モデルは相互に変換が可能であるため, L 個の直交基底関数モデルが存在するということもできる.

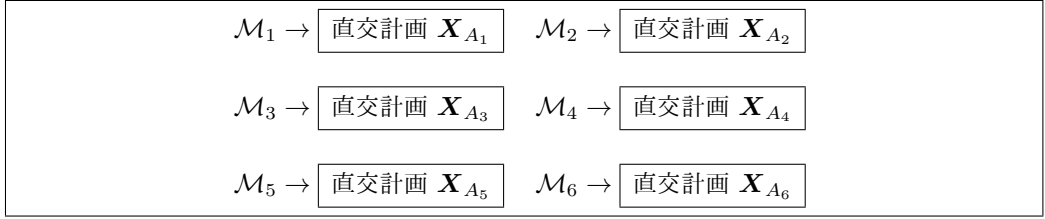


図1 モデルごとの実験 ($L = 6$)

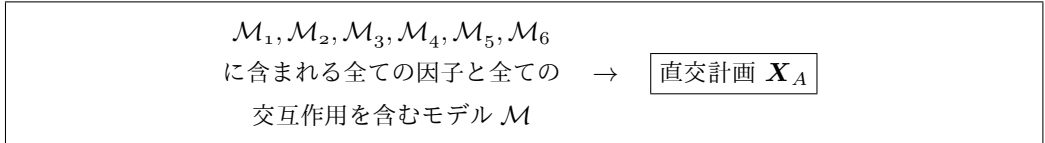


図2 複数モデルの同時実験 ($L = 6$)

えると、6つのモデルが2因子交互作用の重複をどのくらい持つか、に依存することが分かる。

4.2 複数モデルの同時実験が必要とされる状況

本節では、複数モデルの同時実験が必要とされる状況について説明する。まず、同時実験が計画できるということは、関心を持つ大きな交互作用のリストが確定していることを意味している。本稿では、モデル \mathcal{M} の存在を仮定しており、関心を持つ大きな交互作用のリストも確定している。このため、モデル \mathcal{M} に対する実験（複数モデルの同時実験）は可能である。

このとき、 L 個のモデル $\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_L$ はどのように考えられるかについて説明する。ここで重要となるのが交互作用であるが、文献 [2] では「どの2因子交互作用を検討に含め、どの2因子交互作用を検討からはずすのかを事前に明確にすることは必ずしも簡単ではない。しかし、そこは問題解決の初期段階だと割り切って、大胆に（しかし、慎重に）区別することによって、より少ない回数の実験で解析を進めることが可能になる」と記述されている。このため、検討から外すかどうかの判断が難しい交互作用が多いときには、関心を持つ交互作用のリストが大きくなる。また、この大きな交互作用リストには、重要度が高い（必ず検討に含める）交互作用から、それほど重要度が高くない交互作用まで、様々な交互作用が含まれることになる。ここでもし、実験のコストが低いのであれば、交互作用リストが大きな場合であっても、これまで通り、関心を持つ交互作用リストの全てを調べる（実験回数の多い）実験を実施することになる。しかし、本稿では、実験のコストが高い場合について考えており、実験が非常に高コストで実験回数を制限せざるを得ない場合、（関心を持つ交互作用リストの全てを調べることはできないが）重要度の高い交

表 1 実験回数の比較 [17]

モデルごとの実験回数の和	同時実験の実験回数
$\sum_{l=1}^L \mathbf{X}_{A_l} $	$ \mathbf{X}_A $

相互作用については調べることができる実験も意味を持つことになる．言い換えると，上記の文献 [2] の記述にある「大胆に区別することによって」得られるモデル（大きな交互作用リストの部分集合に限定したモデル）が \mathcal{M}_1 等に相当する．さらに，複数の目的が考えられる場合，部分集合の選び方も目的に応じて複数存在する．また，交互作用リストをどこまで絞り込むかについても実験者に依存するため，様々な部分集合が存在する．このような様々な部分集合に対応するモデルが $\mathcal{M}_1 \sim \mathcal{M}_L$ である．

以上より，実験が非常に高コストであり，かつ関心を持つ大きな交互作用リストを持つ場合に，より少ない実験回数，コストでの実験の実施が期待できる複数モデルの同時実験が必要とされる．

以下，次節以降では，データ収集のコストとパラメータ推定の精度に関して，モデルごとの実験と複数モデルの同時実験の違いについて説明する．

4.3 データ収集のコスト

本稿では，データ収集コストの評価基準として，実験回数を用いる． L 個のモデルが存在するときのモデルごとの実験回数の和と同時実験の実験回数を表 1 に載せる．表 1 より，もし $|\mathbf{X}_A| < \sum_{l=1}^L |\mathbf{X}_{A_l}|$ であれば，モデルごとに実験を実施するよりも，モデル \mathcal{M} に対して実験を実施することで，データ収集コストを削減可能であることが分かる．

ここで，それぞれの実験回数である $|\mathbf{X}_A|$ と $\sum_{l=1}^L |\mathbf{X}_{A_l}|$ はモデルに含まれる因子と交互作用が分かれば計算できるため，実際に実験を行う前に知ることが可能である．データ収集コストのより低い実験を行うには， $|\mathbf{X}_A| \leq \sum_{l=1}^L |\mathbf{X}_{A_l}|$ であれば同時実験を選択し，それ以外の場合はモデルごとの実験を選択すればよい．また， $\sum_{l=1}^L |\mathbf{X}_{A_l}| \leq |\mathbf{X}_A|$ であっても，次節で述べるように，モデルごとに実験を実施するよりも同時実験のほうがパラメータの推定精度の点で優れているため，同時実験を選択することも考えられる．このため，コストとパラメータの推定精度の両方を考慮し，同時実験を選択するかどうかを決めるのが良い．

4.4 パラメータの推定精度

本節では、最小二乗推定量の分散の値について、モデル \mathcal{M}_I とモデル \mathcal{M} の比較を行う。まず、 $f_{\mathbf{a}}$ の最小二乗推定量 $\hat{f}_{\mathbf{a}}$ は次式で与えられる [6].

$$\hat{f}_{\mathbf{a}} = \frac{1}{N} \sum_{i=1}^N \mathcal{X}_{\mathbf{a}}^*(\mathbf{x}_i) t(\mathbf{x}_i) \quad (24)$$

このとき、文献 [6] より、以下の結果が得られている。

1. 大きさ N のどのような計画 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ に対しても、 $f_{\mathbf{a}} (\mathbf{a} \in I_A)$ の最小二乗推定量 $\hat{f}_{\mathbf{a}}$ の分散 $V(\hat{f}_{\mathbf{a}})$ について、次式が成立する。

$$V(\hat{f}_{\mathbf{a}}) \geq \frac{\sigma^2}{N} \quad (25)$$

2. A に対する大きさ N の直交計画が存在すれば、これにより得られるデータに基づく、効果 $f_{\mathbf{a}}$ の最小二乗推定量 $\hat{f}_{\mathbf{a}}$ の分散 $V(\hat{f}_{\mathbf{a}})$ は、式 (25) の右辺に到達する。すなわち、

$$V(\hat{f}_{\mathbf{a}}) = \frac{\sigma^2}{N} \quad (26)$$

式 (26) は、直交計画の大きさ N を大きくすれば、効果の最小二乗推定量の分散を小さくできることを示している。モデル \mathcal{M}_I とモデル \mathcal{M} を比較する場合、両者の関係から、 $|\mathbf{X}_{A_I}| \leq |\mathbf{X}_A|$ であるため、同時実験を実施したほうが常に、効果の最小二乗推定量の分散を小さくできることがわかる。

4.5 プログラミングのための 2 つのモデルの利用の流れ

これまで、実験計画法に関する従来研究では全て、伝統的なモデルが用いられてきた。しかし、本稿では、複数モデルの同時実験を扱うため、一般に従来に比べて、扱う因子数およびパラメータの次元数が非常に大きくなってしまう。このため、本稿では従来研究と異なり、伝統的なモデルと直交基底関数モデルを組み合わせた新たな方法を提案する。この方法は、それぞれのモデルの利点を活かすものであり、2 つのモデルの利用の流れを図 3 に載せる。なお、両モデル間でのパラメータの変換は、式 (17), (18) を利用する。図 3 では最初に、伝統的なモデルが記載されている。これは、伝統的なモデルは、パラメータが実数値をとり、各要因の効果が分かりやすいためである。また、後述するベイズ的な扱いにおいて、パラメータの事前確率分布の導入が可能である。

次に記載されているのは、直交基底関数によるモデルである。本稿では、複数モデルの同時実験を扱うため、一般に因子数が非常に大きくなる。このため、プログラミングのための処理を考えると、パラメータに制約がないこのモデルが望ましい。また、このモデル

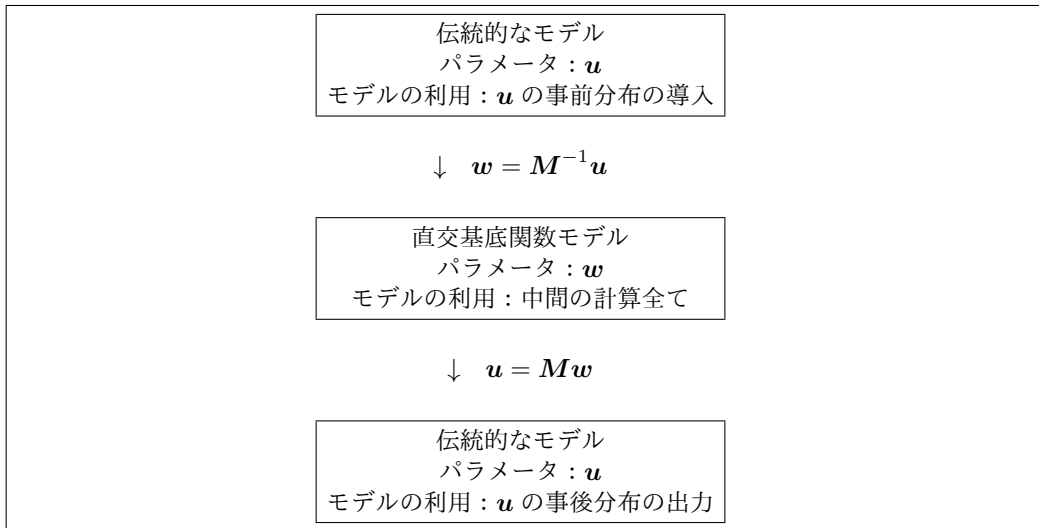


図3 2つのモデルの役割

は基底関数の線形結合で表現されているため、同様に表現されているこれまでの研究成果（ベイズ線形回帰に関する結果等）の利用が容易である。

最後に再び、伝統的なモデルが記載されている。直交基底関数によるモデルは途中の計算での利点はあるもの、複素空間上で計算されるため、パラメータの値は複素数をとってしまう。このため、最後に、パラメータが実数値をとる伝統的なモデルに変換している。このため、事後分布の出力が可能となる。

第5章 ベイズ理論による枠組み

この章では、ベイズ理論による枠組み [14] のために、3.2 節の伝統的なモデルと 3.3 節の直交基底関数モデルの両モデルを用いる。

まず伝統的なモデルであるが、パラメータが実数値のため、各効果が分かりやすく、パラメータの事前確率分布を導入することに適している。このため、伝統的なモデルに対し、パラメータ \mathbf{u} の事前確率分布を導入する。

一方、3.3 節の直交基底関数モデルは、全てのパラメータが独立であるため、因子数が多い場合のプログラミングに適しており、また、ベイズ線形回帰 [13] の結果を用いることで、事後分布を解析的に導出することが可能となる。

また、3 章で既に述べたように、両モデルには 1 対 1 の対応がある。両モデルのパラメータ間の関係である式 (17), (18) を使うことで、両モデルのパラメータ間の変換が可能である。以下で、両モデルを用いたベイズ的な取扱いについて話を進める。

5.1 尤度関数

3.3 節で説明した直交基底関数モデルについて考える．式 (15) はガウス分布を用い、

$$p(t|\mathbf{x}, \mathbf{w}, \sigma^2) = \mathcal{N}(t|y(\mathbf{x}, \mathbf{w}), \sigma^2) \quad (27)$$

と表すことができる．ただし、 $y(\mathbf{x}, \mathbf{w}) = \sum_{\mathbf{a} \in I_A} f_{\mathbf{a}} \mathcal{X}_{\mathbf{a}}(\mathbf{x})$ である．ここで N 個の入力 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ と対応する出力 $\mathbf{t} = [t(\mathbf{x}_1), \dots, t(\mathbf{x}_N)]^T$ が式 (27) から独立に生成されたと仮定すると、次の尤度関数が得られ、調整可能なパラメータ \mathbf{w} と σ^2 の関数となる．

$$p(\mathbf{t}|\mathbf{X}, \mathbf{w}, \sigma^2) = \mathcal{N}(\mathbf{t}|\Phi\mathbf{w}, \sigma^2\mathbf{I}) \quad (28)$$

ここで

$$\Phi = \begin{bmatrix} \mathcal{X}_{a_1}(\mathbf{x}_1) & \mathcal{X}_{a_2}(\mathbf{x}_1) & \dots & \mathcal{X}_{a_K}(\mathbf{x}_1) \\ \mathcal{X}_{a_1}(\mathbf{x}_2) & \mathcal{X}_{a_2}(\mathbf{x}_2) & \dots & \mathcal{X}_{a_K}(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{X}_{a_1}(\mathbf{x}_N) & \mathcal{X}_{a_2}(\mathbf{x}_N) & \dots & \mathcal{X}_{a_K}(\mathbf{x}_N) \end{bmatrix} \quad (29)$$

であり、 K はパラメータ数である．

この後、事前分布は \mathbf{u} について考えるため、 $\mathbf{M}^{-1}\mathbf{u} = \mathbf{w}$ を使い、式 (28) を変形した、以下の尤度関数も得ておく．

$$p(\mathbf{t}|\mathbf{X}, \mathbf{u}, \sigma^2) = \mathcal{N}(\mathbf{t}|\Phi\mathbf{M}^{-1}\mathbf{u}, \sigma^2\mathbf{I}) \quad (30)$$

5.2 事前分布と事後分布 [17]

この節ではモデルパラメータ \mathbf{u} の事前確率分布を導入し、ベイズ的な取り扱いについて説明する．式 (30) で定義される尤度関数に対応する共役事前分布は平均 \mathbf{m}_0 、共分散 \mathbf{S}_0 を持つガウス分布

$$p(\mathbf{u}) = \mathcal{N}(\mathbf{u}|\mathbf{m}_0, \mathbf{S}_0) \quad (31)$$

で与えられる [13]．このとき、実験により得られる、 N 個の入出力の対 \mathbf{X}, \mathbf{t} が与えられたもとでの事後分布は平均 \mathbf{m}_N 、共分散 \mathbf{S}_N を持つガウス分布となり、次式で与えられる．

$$p(\mathbf{u}|\mathbf{X}, \mathbf{t}, \sigma^2) = \mathcal{N}(\mathbf{u}|\mathbf{m}_N, \mathbf{S}_N) \quad (32)$$

ただし、

$$\mathbf{m}_N = \mathbf{S}_N \left(\frac{1}{\sigma^2} (\mathbf{M}^{-1})^* \Phi^* \mathbf{t} + \mathbf{S}_0^{-1} \mathbf{m}_0 \right) \quad (33)$$

$$\mathbf{S}_N^{-1} = \frac{1}{\sigma^2} (\mathbf{M}^{-1})^* \Phi^* \Phi \mathbf{M}^{-1} + \mathbf{S}_0^{-1} \quad (34)$$

であり, $*$ は共役転置, すなわち各要素の複素共役をとって転置を表す.

(式 (32) の証明) \mathbf{M} を加えることを除き, 式 (32) は文献 [13] と同様に証明される.

以下では, 議論を簡単にするために, 文献 [13] と同様に, 単一の精度パラメータ α で記述される平均が $\mathbf{0}$ の等方的ガウス分布

$$p(\mathbf{u}) = \mathcal{N}(\mathbf{u}|\mathbf{0}, \alpha^{-1}\mathbf{I}) \quad (35)$$

を考える. このとき, 事後分布に関する式 (33), (34) はそれぞれ, 以下となる.

$$\mathbf{m}_N = \mathbf{S}_N \left(\frac{1}{\sigma^2} (\mathbf{M}^{-1})^* \Phi^* \mathbf{t} \right) \quad (36)$$

$$\mathbf{S}_N^{-1} = \frac{1}{\sigma^2} (\mathbf{M}^{-1})^* \Phi^* \Phi \mathbf{M}^{-1} + \alpha \mathbf{I} \quad (37)$$

以後, 本稿では, 超パラメータ α, σ^2 を既知の定数として扱う. ただし, 完全にベイズ的に取り扱うために, α と分散の逆数 ($1/\sigma^2$) に対しても事前分布を導入し, エビデンス近似等で周辺化 [13] することも可能である.

第6章 実行例による同時実験の性能評価

ここでは, 真のパラメータを設定したもとの, 直交計画について実験を行い, ガウスノイズを加えたものを目標値とする人工データから, 真のパラメータを復元することを考える. 因子数は $n = 30$, 水準数は $q = 3$, モデル数は $L = 10$ とする. ノイズの分散 σ^2 は既知と仮定し, $\sigma^2 = 1$ を用いる. また, ベイズ理論の枠組みで考えるとき, パラメータ α は既知とし, $\alpha = 1$ を用いる.

次に人工データについて説明する. まず, 式 (35) のガウス分布から, \mathbf{u} を1つ発生させ, これを真のパラメータ \mathbf{u}^* とする. 直交計画を \mathbf{X} とすると, $\mathbf{x} \in \mathbf{X}$ に対し, 式 (12) に \mathbf{u}^* を代入した $t(\mathbf{x}) = y(\mathbf{x}, \mathbf{u}^*) + \epsilon$ により得られる t を目標値とする人工的な訓練データを生成する. なお, 使用したプログラミング言語は Python 3 であり, 実行環境は Google Colaboratory [18] を用いた.

6.1 10 個の伝統的なモデル

始めに, 10 個の伝統的なモデル $\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_{10}$ を考え, 各モデルは一般平均, 30 個の因子の主効果, 10 個の2因子交互作用効果からなるとする. なお, 2因子交互作用は, random モジュールによる乱数を使用し10個選択した. 各モデルに対し, 乱数により選択した10個の2因子交互作用を表2に示す. このとき, モデル $\mathcal{M}_1, \dots, \mathcal{M}_{10}$ は全て, パラメータ数 $K = 1 + 2 \times 30 + 4 \times 10 = 101$ である.

表 2 モデル \mathcal{M}_l に含まれる 2 因子交互作用

モデル	モデルに含まれる 2 因子交互作用
\mathcal{M}_1	[0, 28], [2, 24], [3, 8], [3, 15], [4, 18], [6, 25], [12, 20], [14, 15], [15, 24], [25, 27]
\mathcal{M}_2	[1, 19], [1, 27], [2, 11], [5, 13], [5, 26], [6, 19], [8, 9], [18, 21], [21, 27], [23, 25]
\mathcal{M}_3	[0, 29], [2, 19], [4, 17], [6, 7], [7, 18], [8, 17], [11, 29], [15, 19], [15, 26], [18, 20]
\mathcal{M}_4	[0, 2], [1, 7], [2, 4], [3, 23], [7, 9], [9, 29], [12, 15], [12, 17], [16, 17], [24, 25]
\mathcal{M}_5	[0, 16], [1, 20], [5, 28], [7, 24], [8, 19], [11, 23], [14, 26], [20, 29], [22, 25], [23, 26]
\mathcal{M}_6	[0, 4], [0, 8], [1, 8], [2, 26], [10, 24], [11, 23], [15, 24], [15, 25], [18, 21], [18, 25]
\mathcal{M}_7	[1, 2], [1, 6], [1, 18], [2, 7], [2, 13], [3, 11], [4, 10], [12, 20], [16, 29], [17, 26]
\mathcal{M}_8	[0, 20], [1, 2], [4, 7], [4, 12], [6, 12], [6, 22], [7, 11], [12, 14], [14, 15], [16, 25]
\mathcal{M}_9	[0, 10], [2, 10], [4, 5], [8, 11], [14, 16], [14, 19], [17, 29], [19, 22], [19, 28], [21, 27]
\mathcal{M}_{10}	[0, 18], [1, 5], [1, 18], [2, 10], [6, 14], [8, 26], [13, 15], [15, 16], [15, 26], [20, 25]

6.2 同時実験のためのモデル

6.1 節で与えた 10 個の伝統的なモデル $\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_{10}$ について同時実験を行うことを考える．表 2 に、各モデルに含まれる 10 個の 2 因子交互作用が示されている．ここで、表 2 に含まれる合計 100 個の 2 因子交互作用をすべて含む新たなモデルを考え、そのモデルを \mathcal{M} とする．モデル \mathcal{M} に含まれる 2 因子交互作用を表 3 で示す．表 2 を見ると、10 個の 2 因子交互作用 [1, 2], [1, 18], [2, 10], [11, 23], [12, 20], [14, 15], [15, 24], [15, 26], [18, 21], [21, 27] がそれぞれ 2 つのモデルに含まれていることが分かる．表 3 ではこのような重複は除かれているため、表 3 に含まれる 2 因子交互作用の数は 90 となっている．このとき、モデル \mathcal{M} のパラメータ数は $K = 1 + 2 \times 30 + 4 \times 90 = 421$ である．

表3 モデル \mathcal{M} に含まれる 2 因子交互作用

モデル	モデルに含まれる 2 因子交互作用
\mathcal{M}	$[0, 2], [0, 4], [0, 8], [0, 10], [0, 16], [0, 18], [0, 20],$ $[0, 28], [0, 29], [1, 2], [1, 5], [1, 6], [1, 7], [1, 8],$ $[1, 18], [1, 19], [1, 20], [1, 27], [2, 4], [2, 7], [2, 10],$ $[2, 11], [2, 13], [2, 19], [2, 24], [2, 26], [3, 8], [3, 11],$ $[3, 15], [3, 23], [4, 5], [4, 7], [4, 10], [4, 12], [4, 17],$ $[4, 18], [5, 13], [5, 26], [5, 28], [6, 7], [6, 12], [6, 14],$ $[6, 19], [6, 22], [6, 25], [7, 9], [7, 11], [7, 18], [7, 24],$ $[8, 9], [8, 11], [8, 17], [8, 19], [8, 26], [9, 29], [10, 24],$ $[11, 23], [11, 29], [12, 14], [12, 15], [12, 17], [12, 20],$ $[13, 15], [14, 15], [14, 16], [14, 19], [14, 26], [15, 16],$ $[15, 19], [15, 24], [15, 25], [15, 26], [16, 17], [16, 25],$ $[16, 29], [17, 26], [17, 29], [18, 20], [18, 21], [18, 25],$ $[19, 22], [19, 28], [20, 25], [20, 29], [21, 27], [22, 25],$ $[23, 25], [23, 26], [24, 25], [25, 27]$

0020010022210201000120110020002
100112100211210010222101001022
021010022010100021000112010000
000002110100000021002220102210
110000000001011000110000000101

図4 \mathcal{M}_1 に対し生成された G_A 行列

6.3 データ収集のコスト

はじめに、モデル \mathcal{M}_1 に対する直交計画を考える。このとき、 $k_A = 5$ の直交計画を構成可能であり、直交計画の大きさは $|\mathbf{X}_{A_1}| = 3^5 = 243$ である。モデル \mathcal{M}_1 に対し生成された G_A 行列を図4に示す。モデル \mathcal{M}_2 からモデル \mathcal{M}_{10} も同様に、 $k_A = 5$ の直交計画を構成可能で、 $|\mathbf{X}_{A_2}| = |\mathbf{X}_{A_3}| = \cdots = |\mathbf{X}_{A_{10}}| = 243$ となる。次に、モデル \mathcal{M} に対する直交計画を考える。このとき、 $k_A = 6$ の直交計画を構成可能であり、直交計画の大きさは $|\mathbf{X}_A| = 3^6 = 729$ である。モデル \mathcal{M} に対し生成された G_A 行列を図5に示す。以上をまとめると、モデルごとの実験回数の和と同時実験の実験回数は表4で与えられる。表4より、同時実験の実験回数はモデルごとの実験回数の総和の $1/3$ 以下であり、同時実験のほうがデータ収集コストが低いことが分かる。

0	0	1	0	0	0	2	0	1	2	2	0	2	0	0	0	1	0	1	0	0	1	1	2	2	0	0	1	2	0
1	0	0	2	0	2	0	2	0	0	0	0	0	1	2	0	2	1	0	1	0	1	1	1	2	0	0	0	0	0
0	0	0	0	0	2	1	1	0	1	2	0	0	1	2	1	0	0	1	1	2	2	2	2	0	0	0	2	0	0
0	1	0	2	0	0	1	0	0	2	1	2	0	1	2	0	0	2	0	0	2	0	0	0	1	0	1	0	2	1
0	0	0	1	0	0	0	1	0	2	0	1	1	1	0	1	1	0	1	2	1	0	0	0	1	0	0	1	1	
0	0	0	0	1	0	0	0	1	0	1	0	1	1	0	1	0	1	0	1	0	1	0	2	1	1	0	1	1	0

図5 \mathcal{M} に対し生成された G_A 行列

表4 実験回数の比較

モデルごとの実験回数の和	同時実験の実験回数
2430	729

表5 パラメータの推定精度の比較

直交計画 \mathbf{X}_{A_1} を用いた場合の 効果の最小二乗推定量の分散	直交計画 \mathbf{X}_A を用いた場合の 効果の最小二乗推定量の分散
$V(\hat{f}_a) = 1/243$	$V(\hat{f}_a) = 1/729$

6.4 パラメータ推定の精度

モデル \mathcal{M}_1 に対する直交計画とモデル \mathcal{M} に対する直交計画について、効果 f_a の最小二乗推定量 \hat{f}_a の分散 $V(\hat{f}_a)$ の比較を行う。式 (26) により与えられる分散の値を表5に示す。表5より、モデルごとに実験を行うよりも、モデル \mathcal{M} に対して同時実験をしたほうが、パラメータの推定精度が良いことが分かる。

ここまでは、モデル \mathcal{M}_1 に対する直交計画の実験回数とモデル \mathcal{M} に対する直交計画は実験回数が異なる場合を考えてきた。以下で同じ実験回数の729回とした場合についても考える。モデル \mathcal{M}_1 のパラメータ推定において、モデル \mathcal{M}_1 に対する243回の実験結果に加え、モデル $\mathcal{M}_2 \sim \mathcal{M}_{10}$ に対する実験結果のうち、任意の486回の実験結果を利用し、合計729回の実験結果からパラメータ推定を行うとする。このとき、(25)式より、 f_a の最小二乗推定量 \hat{f}_a の分散 $V(\hat{f}_a)$ について、 $V(\hat{f}_a) \geq 1/729$ となる。一方、直交計画による同時実験の場合、(26)式により、 \hat{f}_a の分散 $V(\hat{f}_a)$ の下限に到達することができ、 $V(\hat{f}_a) = 1/729$ となる。以上より、同じ実験回数であれば、最小二乗推定量の分散を評価基準とするときに、直交計画を用いた同時実験は最適な実験となることが分かる。これは直交計画の最適性 [4, 5] によるものである。

6.5 真のパラメータと事後平均

まず、モデル \mathcal{M}_1 に対する真のパラメータと事後平均を表 6 に載せる。

次に、モデル \mathcal{M} に対する真のパラメータと事後平均を表 7 に載せる。ただし、モデル \mathcal{M} のパラメータ数は全部で 421 であるが、紙面の制約上、表 6 との比較のため、モデル \mathcal{M}_1 に含まれる 101 個のパラメータのみ、表 7 に載せている。表 6 と表 7 より、 \mathcal{M}_1 と \mathcal{M} ともに事後平均が真のパラメータに近いことが分かる。ここで実験回数を考えると、 \mathcal{M}_1 に対する実験回数は 243 であり、 \mathcal{M} に対する同時実験の実験回数は 729 である。このとき、 \mathcal{M}_1 と \mathcal{M} に対する真のパラメータと事後平均の平均二乗誤差を表 8 に載せる。実験回数からも明らかであるが、表 8 より、各モデルで個別に実験を行うよりも、同時実験を行うことで平均二乗誤差を小さくすることができることが分かる。

次に、事前分布と事後分布についての比較を考える。事前分布の共分散行列 \mathbf{S}_0 のトレースと事後分布の共分散行列 \mathbf{S}_N のトレースをそれぞれ、 $tr(\mathbf{S}_0)$ 、 $tr(\mathbf{S}_N)$ と表しておく。このとき、モデル \mathcal{M}_1 とモデル \mathcal{M} それぞれについて、 $tr(\mathbf{S}_0)$ と $tr(\mathbf{S}_N)$ を表 9 に載せる。表 9 より、モデル \mathcal{M}_1 、 \mathcal{M} 共に、事後分布の共分散行列のトレースは非常に小さく、十分に不確定性の少ない事後分布が得られていることが分かる。これは、実験回数がそれぞれ、243, 729 と多いことによるものである。

第 7 章 考察

本稿では、類似する複数目的に対し、各目的に含まれる考慮すべき要因を全て調べることが可能な同時実験を考えてきたが、一般に因子数も多くなる。そこでまず、因子数が多い場合について考察する。

実際の実験では、因子数が少ない場合を扱うことが多いが、田口 [1] によると「因子の数は実験によって異なるが、実験計画法をうまく利用するという立場からいえば、少なくとも 7~8 個、一般にいくら多くてもよい」と記載されている。またこの文献では「たとえ 50 や 100 の因子があっても、数十の実験ですませることもできる」と指摘されている。このように因子数が多い場合も既に想定されているが、ここで大事なことは、多くの回数の実験を、ランダムな順序で実施することである。ランダム化の困難な場合の対応方法については文献 [1, 2] 等が詳しく、これらを利用することで、因子数が多い場合でも同時実験が実施できると考えられる。

次に 2 因子交互作用について考える。前章の実行例では 10 個のモデルそれぞれにおいて、2 因子交互作用をランダムに 10 個選択した。前章では、因子数が 30 の場合を考えたが、30 個から 2 個を選択する組合せ数は 435 となる。435 個からランダムに選ぶため、6.2 節でも述べたように、重複は少なく 10 個であった。しかし一般に、類似する複数目的を考

表 6 モデル \mathcal{M}_1 に対する真のパラメータと事後平均

	真の値	事後平均		真の値	事後平均
μ	1.62	1.7	$\alpha_{25}(0)$	-0.35	-0.31
$\alpha_0(0)$	-0.61	-0.57	$\alpha_{25}(1)$	-1.14	-1.09
$\alpha_0(1)$	-0.53	-0.54	$\alpha_{26}(0)$	-0.35	-0.29
$\alpha_1(0)$	-1.07	-1.18	$\alpha_{26}(1)$	-0.21	-0.23
$\alpha_1(1)$	0.87	0.91	$\alpha_{27}(0)$	0.59	0.54
$\alpha_2(0)$	-2.3	-2.23	$\alpha_{27}(1)$	0.84	0.89
$\alpha_2(1)$	1.74	1.6	$\alpha_{28}(0)$	0.93	0.94
$\alpha_3(0)$	-0.76	-0.66	$\alpha_{28}(1)$	0.29	0.36
$\alpha_3(1)$	0.32	0.42	$\alpha_{29}(0)$	0.89	0.91
$\alpha_4(0)$	-0.25	-0.39	$\alpha_{29}(1)$	-0.75	-0.77
$\alpha_4(1)$	1.46	1.48	$\beta_{0,28}(0,0)$	1.13	1.04
$\alpha_5(0)$	-2.06	-2.08	$\beta_{0,28}(0,1)$	1.2	1.09
$\alpha_5(1)$	-0.32	-0.2	$\beta_{0,28}(1,0)$	0.19	0.25
$\alpha_6(0)$	-0.38	-0.4	$\beta_{0,28}(1,1)$	-0.38	-0.3
$\alpha_6(1)$	1.13	1.09	$\beta_{2,24}(0,0)$	-0.12	0.02
$\alpha_7(0)$	-1.1	-1.18	$\beta_{2,24}(0,1)$	0.02	0.04
$\alpha_7(1)$	-0.17	-0.15	$\beta_{2,24}(1,0)$	-1.12	-1.11
$\alpha_8(0)$	-0.88	-0.84	$\beta_{2,24}(1,1)$	-0.52	-0.59
$\alpha_8(1)$	0.04	-0.02	$\beta_{3,8}(0,0)$	-0.17	-0.41
$\alpha_9(0)$	0.58	0.57	$\beta_{3,8}(0,1)$	0.99	1.21
$\alpha_9(1)$	-1.1	-1.15	$\beta_{3,8}(1,0)$	0.21	0.32
$\alpha_{10}(0)$	1.14	1.29	$\beta_{3,8}(1,1)$	2.19	1.97
$\alpha_{10}(1)$	0.9	0.86	$\beta_{3,15}(0,0)$	-0.25	-0.26
$\alpha_{11}(0)$	0.5	0.43	$\beta_{3,15}(0,1)$	0.04	0.19
$\alpha_{11}(1)$	0.9	0.85	$\beta_{3,15}(1,0)$	-0.23	-0.24
$\alpha_{12}(0)$	-0.68	-0.67	$\beta_{3,15}(1,1)$	1.33	1.3
$\alpha_{12}(1)$	-0.12	-0.07	$\beta_{4,18}(0,0)$	0.82	0.89
$\alpha_{13}(0)$	-0.94	-0.98	$\beta_{4,18}(0,1)$	-0.56	-0.81
$\alpha_{13}(1)$	-0.27	-0.17	$\beta_{4,18}(1,0)$	1.95	1.86
$\alpha_{14}(0)$	0.53	0.49	$\beta_{4,18}(1,1)$	-1.33	-1.15
$\alpha_{14}(1)$	-0.69	-0.71	$\beta_{6,25}(0,0)$	-0.67	-0.64
$\alpha_{15}(0)$	-0.4	-0.4	$\beta_{6,25}(0,1)$	-1.67	-1.51
$\alpha_{15}(1)$	-0.69	-0.73	$\beta_{6,25}(1,0)$	0.83	0.73
$\alpha_{16}(0)$	-0.85	-0.83	$\beta_{6,25}(1,1)$	-0.5	-0.54
$\alpha_{16}(1)$	-0.67	-0.75	$\beta_{12,20}(0,0)$	1.62	1.6
$\alpha_{17}(0)$	-0.01	0	$\beta_{12,20}(0,1)$	0.01	0.05
$\alpha_{17}(1)$	-1.12	-1.25	$\beta_{12,20}(1,0)$	-0.69	-0.94
$\alpha_{18}(0)$	0.23	0.19	$\beta_{12,20}(1,1)$	0.62	0.74
$\alpha_{18}(1)$	1.66	1.66	$\beta_{14,15}(0,0)$	-0.66	-0.81
$\alpha_{19}(0)$	0.74	0.71	$\beta_{14,15}(0,1)$	3.03	3.16
$\alpha_{19}(1)$	-0.19	-0.12	$\beta_{14,15}(1,0)$	0.82	0.94
$\alpha_{20}(0)$	-0.89	-0.96	$\beta_{14,15}(1,1)$	0.65	0.66
$\alpha_{20}(1)$	-0.75	-0.55	$\beta_{15,24}(0,0)$	-1.94	-1.76
$\alpha_{21}(0)$	1.69	1.75	$\beta_{15,24}(0,1)$	-2.51	-2.63
$\alpha_{21}(1)$	0.05	-0.02	$\beta_{15,24}(1,0)$	-2.11	-1.95
$\alpha_{22}(0)$	-0.64	-0.6	$\beta_{15,24}(1,1)$	-0.41	-0.52
$\alpha_{22}(1)$	0.19	0.09	$\beta_{25,27}(0,0)$	1.12	1.07
$\alpha_{23}(0)$	2.1	2.04	$\beta_{25,27}(0,1)$	0.31	0.52
$\alpha_{23}(1)$	0.12	0.06	$\beta_{25,27}(1,0)$	-2.09	-2.02
$\alpha_{24}(0)$	0.62	0.64	$\beta_{25,27}(1,1)$	-0.23	-0.4
$\alpha_{24}(1)$	0.3	0.23			

表7 モデル \mathcal{M} に対する真のパラメータと事後平均

	真の値	事後平均		真の値	事後平均
μ	1.62	1.66	$\alpha_{25}(0)$	-0.35	-0.34
$\alpha_0(0)$	-0.61	-0.66	$\alpha_{25}(1)$	-1.14	-1.18
$\alpha_0(1)$	-0.53	-0.5	$\alpha_{26}(0)$	-0.35	-0.3
$\alpha_1(0)$	-1.07	-1.09	$\alpha_{26}(1)$	-0.21	-0.29
$\alpha_1(1)$	0.87	0.9	$\alpha_{27}(0)$	0.59	0.61
$\alpha_2(0)$	-2.3	-2.36	$\alpha_{27}(1)$	0.84	0.89
$\alpha_2(1)$	1.74	1.79	$\alpha_{28}(0)$	0.93	0.91
$\alpha_3(0)$	-0.76	-0.81	$\alpha_{28}(1)$	0.29	0.27
$\alpha_3(1)$	0.32	0.33	$\alpha_{29}(0)$	0.89	0.85
$\alpha_4(0)$	-0.25	-0.21	$\alpha_{29}(1)$	-0.75	-0.75
$\alpha_4(1)$	1.46	1.45	$\beta_{0,28}(0,0)$	1.13	1.15
$\alpha_5(0)$	-2.06	-2.09	$\beta_{0,28}(0,1)$	1.2	1.21
$\alpha_5(1)$	-0.32	-0.3	$\beta_{0,28}(1,0)$	0.19	0.09
$\alpha_6(0)$	-0.38	-0.3	$\beta_{0,28}(1,1)$	-0.38	-0.26
$\alpha_6(1)$	1.13	1.09	$\beta_{2,24}(0,0)$	-0.12	-0.03
$\alpha_7(0)$	-1.1	-1.07	$\beta_{2,24}(0,1)$	0.02	-0.11
$\alpha_7(1)$	-0.17	-0.22	$\beta_{2,24}(1,0)$	-1.12	-1.13
$\alpha_8(0)$	-0.88	-0.96	$\beta_{2,24}(1,1)$	-0.52	-0.42
$\alpha_8(1)$	0.04	0.13	$\beta_{3,8}(0,0)$	-0.17	-0.17
$\alpha_9(0)$	0.58	0.6	$\beta_{3,8}(0,1)$	0.99	0.98
$\alpha_9(1)$	-1.1	-1.06	$\beta_{3,8}(1,0)$	0.21	0.27
$\alpha_{10}(0)$	1.14	1.13	$\beta_{3,8}(1,1)$	2.19	2.18
$\alpha_{10}(1)$	0.9	0.91	$\beta_{3,15}(0,0)$	-0.25	-0.11
$\alpha_{11}(0)$	0.5	0.51	$\beta_{3,15}(0,1)$	0.04	0.09
$\alpha_{11}(1)$	0.9	0.93	$\beta_{3,15}(1,0)$	-0.23	-0.35
$\alpha_{12}(0)$	-0.68	-0.72	$\beta_{3,15}(1,1)$	1.33	1.33
$\alpha_{12}(1)$	-0.12	-0.1	$\beta_{4,18}(0,0)$	0.82	0.85
$\alpha_{13}(0)$	-0.94	-0.94	$\beta_{4,18}(0,1)$	-0.56	-0.62
$\alpha_{13}(1)$	-0.27	-0.27	$\beta_{4,18}(1,0)$	1.95	1.89
$\alpha_{14}(0)$	0.53	0.54	$\beta_{4,18}(1,1)$	-1.33	-1.21
$\alpha_{14}(1)$	-0.69	-0.75	$\beta_{6,25}(0,0)$	-0.67	-0.62
$\alpha_{15}(0)$	-0.4	-0.31	$\beta_{6,25}(0,1)$	-1.67	-1.66
$\alpha_{15}(1)$	-0.69	-0.74	$\beta_{6,25}(1,0)$	0.83	0.81
$\alpha_{16}(0)$	-0.85	-0.82	$\beta_{6,25}(1,1)$	-0.5	-0.52
$\alpha_{16}(1)$	-0.67	-0.67	$\beta_{12,20}(0,0)$	1.62	1.69
$\alpha_{17}(0)$	-0.01	-0.04	$\beta_{12,20}(0,1)$	0.01	-0.06
$\alpha_{17}(1)$	-1.12	-1.07	$\beta_{12,20}(1,0)$	-0.69	-0.68
$\alpha_{18}(0)$	0.23	0.24	$\beta_{12,20}(1,1)$	0.62	0.6
$\alpha_{18}(1)$	1.66	1.74	$\beta_{14,15}(0,0)$	-0.66	-0.68
$\alpha_{19}(0)$	0.74	0.83	$\beta_{14,15}(0,1)$	3.03	3.02
$\alpha_{19}(1)$	-0.19	-0.26	$\beta_{14,15}(1,0)$	0.82	0.92
$\alpha_{20}(0)$	-0.89	-0.78	$\beta_{14,15}(1,1)$	0.65	0.59
$\alpha_{20}(1)$	-0.75	-0.77	$\beta_{15,24}(0,0)$	-1.94	-1.93
$\alpha_{21}(0)$	1.69	1.68	$\beta_{15,24}(0,1)$	-2.51	-2.5
$\alpha_{21}(1)$	0.05	0.14	$\beta_{15,24}(1,0)$	-2.11	-2
$\alpha_{22}(0)$	-0.64	-0.63	$\beta_{15,24}(1,1)$	-0.41	-0.41
$\alpha_{22}(1)$	0.19	0.15	$\beta_{25,27}(0,0)$	1.12	1.13
$\alpha_{23}(0)$	2.1	2.09	$\beta_{25,27}(0,1)$	0.31	0.24
$\alpha_{23}(1)$	0.12	0.13	$\beta_{25,27}(1,0)$	-2.09	-2.18
$\alpha_{24}(0)$	0.62	0.61	$\beta_{25,27}(1,1)$	-0.23	-0.18
$\alpha_{24}(1)$	0.3	0.25			

表 8 真の値と事後平均の平均二乗誤差

	\mathcal{M}_1	\mathcal{M}
平均二乗誤差	0.009738614	0.004464846

表 9 事前分布と事後分布の共分散行列のトレース

	$tr(\mathbf{S}_0)$	$tr(\mathbf{S}_N)$
モデル \mathcal{M}_1	101	1.1349317415890783
モデル \mathcal{M}	421	2.123970137972207

える場合、考慮すべき 2 因子交互作用の集合は類似していると考えられるため、実際には前章の実行例よりも、複数のモデルに含まれる 2 因子交互作用がより多く含まれると考えられる。複数モデル間での 2 因子交互作用の重複が多いほど、モデル \mathcal{M} に含まれる 2 因子交互作用の数は少なくなるため、モデル \mathcal{M} に対し、より少ない実験回数の直交計画を作ることが期待できる。これより、データ収集のコストの点で同時実験のメリットは大きいと考えられる。

最後に、直交基底関数モデルを用いる利点について考える。一つ目の利点は、既に述べてきたように、パラメータが全て独立となるため、プログラミングが容易になることである。また、二つ目の利点として、因子数が非常に多い場合や 3 因子以上の交互作用を含む場合、計算時間は増加するが、フーリエ変換アルゴリズムを利用可能であるため、計算の高速化が可能となることがあげられる。

第 8 章 まとめ

本稿では、類似する複数目的に対し、各目的に含まれる考慮すべき要因を全て調べることが可能な同時実験を提案した。次に、個別に直交計画を実験する場合の実験回数の総和と、同時実験で直交計画を実験する場合の実験回数を比較し、同時実験を利用することで実験回数を減らすことが可能であることを、実行例により示した。さらに、ベイズ理論による枠組みのもとでのパラメータ推定を考え、同時実験は個別実験よりも、事後平均が真のパラメータに近いことを示した。最後に、同時実験についての考察を行い、その特性を明らかにした。

最後に今後の課題を述べる。本研究では各因子の水準数が同じ場合を考えてきたが、1 つ目の課題として、各因子で水準数が異なる場合についての実験が考えられる。2 つ目として、直交計画についての実験後、さらに追加して実験を行う場合、次に実験すべき実験計画を考えることも考えられる。分布モデルについても、本研究では線形モデルのみ扱ってきたが、非線形性を持つ分布モデルに対する同時実験の特徴を明らかにすることも課題

である。また、実際には、類似する目的を持つ会社が同業のため実験内容を他社には知られたくない等の理由で、コストが高くなっても自社で実験を行い、同時実験に参加しない場合も考えられる。このような場合でも、同時実験ができるよう会社間の調整を担う第三者組織の検討も今後の課題である。

謝辞

本研究を進めるにあたり、有益な御討論や御指摘を頂いた早稲田大学の松嶋敏泰教授、湘南工科大学の齋藤友彦准教授に深く感謝致します。また、本研究は JSPS 科研費 JP17K00316 の助成を受けたものです。

参考文献

- [1] 田口玄一：新版 実験計画法 上, 丸善株式会社 (1962).
- [2] 永田靖：入門実験計画法, 日科技連 (2000).
- [3] 芳賀敏郎：医薬品開発のための統計解析, 第2部実験計画法改訂版, サイエンティスト社 (2014).
- [4] S. Moriguti: Optimality of Orthogonal Designs, Res. Stat. Appl. Res., JUSE, Vol.3 (1954).
- [5] 高橋磐郎：組合せ理論とその応用, 岩波全書 (1979).
- [6] 浮田善文, 松嶋敏泰：複素空間での直交計画の最適性に関する一考察, 情報処理学会第82回全国大会講演論文集, Vol.1, pp.171-172 (2020).
- [7] 田中秀俊：ランダムなサイズの直交計画近傍を用いた大局的最適解探索法, 情報処理学会論文誌数理モデル化と応用, Vol.44, No.8, pp.35-42 (2003).
- [8] 須田健二, 五味弘：有限射影幾何を用いたソフトウェアテスト向けの直交表自動生成プログラムの開発とその応用, 情報処理学会論文誌, Vol.57, No.8, pp.1800-1809 (2016).
- [9] 船津公人, 柴山翔二郎：実践マテリアルズインフォマティクス-Python による材料設計のための機械学習, 近代科学社 (2020).
- [10] 亀井淳三：治験薬学, 南江堂 (2012).
- [11] Ukita, Y., Saito, T., Matsushima, T. and Hirasawa, S.: A Note on a Sampling Theorem for Functions over $GF(q)^n$ Domain, *IEICE Trans. Fundamentals*, Vol.E93-A, No.6, pp.1024-1031 (2010).
- [12] Ukita, Y.: A Note on the Evidence Approximation in Bayesian Experimental Design Models on an Orthonormal System, *Journal of Signal Processing*, Vol.22, No.6, pp.307-314 (2018).
- [13] Bishop, C.M.: Pattern Recognition and Machine Learning, Springer (2006). 元田浩,

- 栗田多喜夫, 樋口知之, 松本裕治, 村田昇 (監訳): パターン認識と機械学習 上 - ベイズ理論による統計的予測, シュプリンガー・ジャパン (2007).
- [14] Chaloner, K. and Verdinelli, I.: Bayesian Experimental Design: A Review, *Statistical Science*, Vol.10, No.3, pp.273-304 (1995).
- [15] Stein, E. M. and Shakarchi, R.: Fourier Analysis, Princeton University Press (2003).
新井仁之, 杉本充, 高木啓行, 千原浩之 (訳): フーリエ解析入門, 日本評論社 (2007).
- [16] 平澤茂一, 西島利尚: 符号理論入門, 培風館 (1999).
- [17] Ukita, Y. and Matsushima, T.: A Study on Simultaneous Experiments for Related Linear Models based on an Orthonormal System (poster presentation), *Bayes on the Beach 2019*, <https://botb2019.files.wordpress.com/2019/12/ukita.pdf> (2019).
- [18] Google Colaboratory: <https://colab.research.google.com/> (2022).

付録 A 行列 M と M^{-1}

伝統的なモデルの独立なパラメータのみからなる列ベクトル \mathbf{u} と, 直交基底関数モデルのパラメータからなる列ベクトル \mathbf{w} に対し, 次式を満たす $k \times k$ 行列 M を構成可能である [12].

$$\mathbf{u} = M\mathbf{w} \quad (38)$$

ただし,

$$M = \begin{bmatrix} 1 & & & & & \\ & R_1 & & & & 0 \\ & & \ddots & & & \\ & & & R_1 & & \\ & & & & R_2 & \\ & 0 & & & & \ddots \\ & & & & & & R_2 \end{bmatrix}$$

であり, R_1 は次式で与えられる $(q-1) \times (q-1)$ 行列であり,

$$R_1 = \begin{bmatrix} \mathcal{X}_1(0) & \cdots & \mathcal{X}_{q-1}(0) \\ \mathcal{X}_1(1) & \cdots & \mathcal{X}_{q-1}(1) \\ \vdots & \ddots & \vdots \\ \mathcal{X}_1(q-2) & \cdots & \mathcal{X}_{q-1}(q-2) \end{bmatrix}$$

R_2 は次式で与えられる $(q-1)^2 \times (q-1)^2$ 行列である.

$$R_2 = \begin{bmatrix} \mathcal{X}_1(0)\mathcal{X}_1(0) & \cdots & \mathcal{X}_{q-1}(0)\mathcal{X}_{q-1}(0) \\ \mathcal{X}_1(1)\mathcal{X}_1(0) & \cdots & \mathcal{X}_{q-1}(1)\mathcal{X}_{q-1}(0) \\ \vdots & \ddots & \vdots \\ \mathcal{X}_1(q-2)\mathcal{X}_1(q-2) & \cdots & \mathcal{X}_{q-1}(q-2)\mathcal{X}_{q-1}(q-2) \end{bmatrix}$$

また、次式を満たす $k \times k$ 行列 M^{-1} を構成可能である.

$$w = M^{-1}u \quad (39)$$

ここで,

$$M^{-1} = \begin{bmatrix} 1 & & & & & \\ & S_1 & & & & 0 \\ & & \ddots & & & \\ & & & S_1 & & \\ & & & & S_2 & \\ 0 & & & & & \ddots & \\ & & & & & & S_2 \end{bmatrix} \quad (40)$$

である. $q' = q - 1$ とおくと, S_1 は $q' \times q'$ 行列であり, 次式で与えられ,

$$S_1 = \frac{1}{q} \begin{bmatrix} \mathcal{X}_1^*(0) - \mathcal{X}_1^*(q') & \cdots & \mathcal{X}_1^*(q-2) - \mathcal{X}_1^*(q') \\ \mathcal{X}_2^*(0) - \mathcal{X}_2^*(q') & \cdots & \mathcal{X}_2^*(q-2) - \mathcal{X}_2^*(q') \\ \vdots & \ddots & \vdots \\ \mathcal{X}_{q'}^*(0) - \mathcal{X}_{q'}^*(q') & \cdots & \mathcal{X}_{q'}^*(q-2) - \mathcal{X}_{q'}^*(q') \end{bmatrix}$$

S_2 は $q'^2 \times q'^2$ 行列であり, 以下で与えられる.

$$S_2 = \frac{1}{q^2} \begin{bmatrix} (\mathcal{X}_1^*(0) - \mathcal{X}_1^*(q'))(\mathcal{X}_1^*(0) - \mathcal{X}_1^*(q')) \\ (\mathcal{X}_2^*(0) - \mathcal{X}_2^*(q'))(\mathcal{X}_1^*(0) - \mathcal{X}_1^*(q')) \\ \vdots \\ (\mathcal{X}_{q'}^*(0) - \mathcal{X}_{q'}^*(q'))(\mathcal{X}_{q'}^*(0) - \mathcal{X}_{q'}^*(q')) \\ \cdots (\mathcal{X}_1^*(q-2) - \mathcal{X}_1^*(q'))(\mathcal{X}_1^*(q-2) - \mathcal{X}_1^*(q')) \\ \cdots (\mathcal{X}_2^*(q-2) - \mathcal{X}_2^*(q'))(\mathcal{X}_1^*(q-2) - \mathcal{X}_1^*(q')) \\ \vdots \\ \cdots (\mathcal{X}_{q'}^*(q-2) - \mathcal{X}_{q'}^*(q'))(\mathcal{X}_{q'}^*(q-2) - \mathcal{X}_{q'}^*(q')) \end{bmatrix}$$