最適計画基準の評価法に関する研究

M2016SS010 田川聖也
指導教員：松田眞一

1 はじめに

実験計画法は効率的にデータの収集を考える方法であり，実験での時間とコストを節約でき，開発速度を上げることができる．実験計画法の分野の 1 つに最適計画がある．最適計画はモデルを基に計画を評価する基準を定めて，その基準で最適化して求めた計画である．計算機の発達により最適計画は実務においても使われており，中でも D 最適計画が最も使われている．しかし最適計画はモデルに合うように収集されたデータを最適化するため，基準によって結果が異なる．そのため，どの最適基準が良いかを別の観点から評価する必要がある．そこで奥村 [4] は応答曲面の観点から最適基準 (A，D，I) の評価を行い，D 最適基準が最も良いという結論に至った．しかし奥村 [4] が行っていない視点から基準の評価を行うことで結論が変わる可能性もある．本研究では奥村 [4] の最適基準 (A，D，I) の評価法にまだ使われてない E，G 最適基準を評価指標とし，D 最適基準の有用性を再検討する．

2 実験計画法

実験計画法とは，取り上げる対象の結果とそれに影響を与えると思われる因子 (実験者が設定できるパラメータ) の関係を効率的に調べる統計的方法である．因子が 2 つ以上ある場合，交互作用が発生する．実験計画法の利点は少ない実験回数で因子と交互作用の影響を見ることができることである．(永田 [3] 参照)

3 応答曲面法

応答曲面法とは，応答と因子の関係について，実験計画にしたがってデータを収集し，それを解析することで応答と因子の関係を探索する一連の方法である．n は実験回数，$i = 1, \cdots, n$ に対して y_i は応答変数，p 個の因子を $\boldsymbol{x}_i = (1, x_{i1}, \cdots, x_{ip})^T$，$\boldsymbol{\beta} = (\beta_0,\beta_1, \cdots,\beta_p)^T$ は回帰係数，誤差を ε_i とし次のモデルを考える．

$$y_i = \boldsymbol{x}_i^T\boldsymbol{\beta} + \varepsilon_i$$

ここで $E(y_i) = \boldsymbol{x}_i^T\boldsymbol{\beta}$ とし，ε_i は互いに独立で $N(0,\sigma^2)$ に従うと仮定する．応答と因子の関係を表現する $E(y_i)$ について，$\boldsymbol{x}_i$ のどのような関数型も考えることが可能である．ここでは，$E(y_i)$ の 2 次式

$$E(y_i) = \beta_0 + \sum_{j=1}^{p}\beta_j x_{ij} + \sum_{j=1}^{p-1}\sum_{k=j+1}^{p}\beta_{jk}x_{ij}x_{ik} + \sum_{j=1}^{p}\beta_{jj}x_{ij}^2 \tag{1}$$

で表現できるものとして解析を行う．式 (1) は 2 次モデル，$E(y_i)$ は応答曲面と呼ばれる．$\beta_j x_{ij}$ は因子 x_{ij} の 1 次の効果，$\beta_{jj}x_{ij}^2$ は因子 x_{ij} の 2 次の効果，$\beta_{jk}x_{ij}x_{ik}$ は因子 x_{ij} と x_{ik} の「1 次 ×1 次の交互作用」を表す．因子や交互作用は，実用面では 2 次までで十分である．(Atkinson and Donev[1]，山田 [7] 参照)

4 最適計画

最適計画は，実験回数 n，計画領域，モデルを与え，計画の良さを測る基準を決め，その基準を最適化するように構成された計画である．最適計画は因子や水準数，実験回数など形式に制限がなく柔軟に実験計画を組める利点がある．しかし与えるモデルが妥当ではない場合，再現性のない結果となる．(Atkinson and Donev[1] 参照)

4.1 基本概念

$\boldsymbol{y}$ を $n \times 1$ の応答ベクトル，X を $n \times (p+1)$ の計画行列とする線形回帰モデルの最小二乗推定量 $\hat{\boldsymbol{\beta}}$ は残差平方和 $(\boldsymbol{y} - X\boldsymbol{\beta})^T(\boldsymbol{y} - X\boldsymbol{\beta})$ を最小にする

$$\hat{\boldsymbol{\beta}} = (X^TX)^{-1}X^T\boldsymbol{y} \tag{2}$$

である．$\hat{\boldsymbol{\beta}}$ の分散共分散行列は以下のようになる．

$$V(\hat{\boldsymbol{\beta}}) = \sigma^2(X^TX)^{-1} \tag{3}$$

$(X^TX)^{-1}$ が小さくなるように計画を組むことで，式 (3) の $V(\hat{\boldsymbol{\beta}})$ を小さくでき，信頼性のある統計的推測が可能となる．$V(\hat{\boldsymbol{\beta}})$ を総合評価するための最適基準がこれまで多く提案されている．有名な基準に A，D，I がある．X^TX を n で割ったものは情報行列と呼ばれている．

$$M = \frac{X^TX}{n} \tag{4}$$

最適計画は情報行列 M を最適性の基準で最適化した計画行列 X である．(Atkinson and Donev[1]，山田 [7] 参照)

5 最適基準の種類

5.1 A 最適基準

A 最適基準は情報行列の逆行列の対角成分の和 (trace) を最小化し，$V(\hat{\boldsymbol{\beta}})$ を小さくする立場からの基準である．

$$\min_X\left[tr(M^{-1})\right] = \min_X\left[tr\left\{(X^TX/n)^{-1}\right\}\right] \tag{5}$$

で表される．情報行列 M の固有値を用いて A 最適基準を定義することも可能である．M の固有値を $\lambda_1, \cdots,\lambda_p$ とするとき，M^{-1} の固有値は $1/\lambda_1, \cdots, 1/\lambda_p$ となり，M^{-1} の固有値の総和を最小化する

$$\min\sum_{i=1}^{p}\frac{1}{\lambda_i} \tag{6}$$

という基準として定義できる．(平尾ら [2] 参照)

5.2 D 最適基準

D 最適基準は最適計画が研究され始めた当初からよく用いられている．X^TX の行列式を最大化し $V(\hat{\boldsymbol{\beta}})$ を小さくする立場からの基準である．

$$\max_X|M| = \max_X|X^TX/n| \tag{7}$$

で表される．情報行列 M の固有値を用いて D 最適基準を定義することも可能である．M の固有値を $\lambda_1,\cdots,\lambda_p$ とするとき，その固有値の積を最大化する

$$\max \prod_{i=1}^{p} \lambda_i \tag{8}$$

という基準として定義できる．(Atkinson and Donev[1], Wit，Nobile and Khanin[6] 参照)

5.3 I 最適基準

I 最適基準は予測分散 $V(\boldsymbol{x}_i^T\hat{\boldsymbol{\beta}})=\sigma^2\boldsymbol{x}_i^T(X^TX)^{-1}\boldsymbol{x}_i$ を小さくする立場からの最適基準である．これらの立場の基準は，$V(\boldsymbol{x}_i^T\hat{\boldsymbol{\beta}})$ に σ^2，実験回数 n を用いて標準化を行った

$$d(\boldsymbol{x}_i) = n\frac{V(\boldsymbol{x}_i^T\hat{\boldsymbol{\beta}})}{\sigma^2} = \boldsymbol{x}_i^T M^{-1}\boldsymbol{x}_i \tag{9}$$

を用いて最適化を行う．I 最適基準は予測値の標準化分散 $d(\boldsymbol{x}_i)(i=1,\cdots,n)$ を平均した値を最小化する

$$\begin{aligned}\min_X \frac{1}{n}\sum_{i=1}^{n} d(\boldsymbol{x}_i) &= \min_X \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i^T M^{-1}\boldsymbol{x}_i \\ &= \min_X \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i^T (X^TX/n)^{-1}\boldsymbol{x}_i \end{aligned}\tag{10}$$

の基準として定義できる．(Atkinson and Donev[1]，Package ‘AlgDesign’ [5] 参照)

5.4 E 最適基準

E 最適基準は情報行列の最小固有値を最大化する最適基準である．$V(\hat{\boldsymbol{\beta}})$ を小さくする立場からの基準である．情報行列 M の最小固有値を λ_{min} とするとき，最適基準は

$$\max_X \lambda_{min}(M) = \max_X \lambda_{min}(X^TX/n) \tag{11}$$

として定義される．(Atkinson and Donev[1]，平尾ら [2] 参照)

5.5 G 最適基準

G 最適基準は予測値の分散 $V(\boldsymbol{x}_i^T\hat{\boldsymbol{\beta}})$ を小さくする立場からの基準である．標準化分散 $d(\boldsymbol{x}_i)(i=1,\cdots,n)$ のミニマックス基準

$$\begin{aligned}\min_X \max_i d(\boldsymbol{x}_i) &= \min_X \max_i \boldsymbol{x}_i^T M^{-1}\boldsymbol{x}_i \\ &= \min_X \max_i \boldsymbol{x}_i^T (X^TX/n)^{-1}\boldsymbol{x}_i \end{aligned}\tag{12}$$

として定義される．G 最適性の達成度合いを評価する指標がある．これは p 個の未知パラメータを含むモデル，計画領域，実験回数が与えられたときの G 最適計画に対して，同一モデル，計画領域，実験回数に関する計画 X がどの程度効率的に G 最適性を達成できているかを標準化分散 $d(\boldsymbol{x}_i)(i=1,\cdots,n)$ を用いて

$$G_{eff} = \frac{p}{\max_i d(\boldsymbol{x}_i)} \tag{13}$$

のように定義できる．この指標は $G\text{-}effciency$(G-効率) と呼ばれ 1 に近いほど G 最適性が達成できている．(Atkinson and Donev[1]，Package‘AlgDesign’[5]，山田 [7] 参照)

6 計画の評価

統計解析ソフト「R」のパッケージの 1 つである AlgDesign を導入し，A，D，I 基準でそれぞれの最適計画を生成できるようにした．今回は，AlgDesign で生成した最適計画 (A，D，I) を評価する指標に E，G 最適基準を用いる．E，G 基準は計画を生成するのに用いるのではなく評価としてのみ用いる．E，G 基準の観点から A，D，I 最適計画のうち，どの計画が最も良いのか判断する．

6.1 解析方法

モデルは 3 因子 2 次モデル ($p=10$)，水準数は 3 水準から 5 水準の範囲，各水準の値は 3 水準 (-1 0 1)，4 水準 (-3 -1 1 3)，5 水準 (-2 -1 0 1 2) としている．

6.2 解析手順

手順 1，2 は奥村 [4] を参照している．

手順 1 初期計画として完全要因計画のフル行列を生成し，それらを行方向にランダムに並び替える．

手順 2 初期計画から指定した計画数分の行ベクトル $\boldsymbol{x}_i$ をモデル，基準 (A，D，I) の最適性に合うように選択し，最適計画 (A，D，I) を生成する．

手順 3 生成した最適計画 (A，D，I) をそれぞれ情報行列 M^A，M^D，M^I に変換し，E，G 基準の評価値を求める．

手順 4 手順 1〜3 を 100 回繰り返す．繰り返しで求めた，E，G 基準の評価値の平均を出す．

6.3 E 基準での計画評価

生成した最適計画 (A，D，I) を E 最適基準で評価する．E 基準は情報行列の最小固有値を最大化する基準である．そのため各計画の評価値は A，D，I 基準でそれぞれ最適化した計画の情報行列 M の最小固有値

$$\lambda_{min}(M^A) \quad \lambda_{min}(M^D) \quad \lambda_{min}(M^I)$$

を求め，その値が大きい計画を E 基準の観点においてベストな計画とする．ベストの計画に 1 点，ベストの計画が 2 つある場合，同点とし両方に 1 点とする．3 つ全て同点である場合は引き分けに 1 点とする．値の差が 0.001 未満の場合，同点とする．水準が (4，4，4) の場合を以下に示す．

表 1 水準 (4，4，4) の E 基準平均値とベストな計画

計画数	A	D	I	ベスト
12	0.255	0.081	0.141	A
24	0.341	0.088	0.130	A
48	0.231	0.108	0.131	A
60	0.186	0.140	0.140	A

表 1 から A が 4 点で水準が (4，4，4) の場合 A が最も良くなる傾向がある．水準が (3，3，5) の場合を以下に示す．

表 2　水準 (3，3，5) の E 基準平均値とベストな計画

計画数	A	D	I	ベスト
12	0.123	0.070	0.078	A
22	0.156	0.086	0.087	A
33	0.125	0.081	0.097	A
40	0.112	0.092	0.092	A

表 2 から A は 4 点で水準が (3，3，5) の場合 A が最も良くなる傾向がある．評価した水準列，計画数を以下にまとめる．水準パターンは 3 つ全て等しい場合は 13 通り，3 つのうち 2 つが等しい場合は 25 通り，3 つ全て異なる場合は 4 通りの全 42 通りで評価している．

表 3　E 基準のベスト数の比較

水準	A	D	I	引き分け
3 つ全て等しい	12	0	0	1
2 つ等しい	25	0	0	0
3 つ全て異なる	4	0	0	0
総計	41	0	0	1

表 3 から総計で A が 41 点，引き分けが 1 点となり，E 基準の観点からは A 最適計画が最も良くなる．A は最小固有値が 3 つの基準の中で最も大きい基準である．表 3 から D，I は同点のため 2 つに絞った結果を以下に示す．

表 4　E 基準のベスト数の比較 (A を除く)

水準	D	I	引き分け
3 つ全て等しい	0	10	3
2 つ等しい	1	21	3
3 つ全て異なる	0	4	0
総計	1	35	6

表 4 から D よりも I の方が良い結果であることが分かる．最終的な E 基準の観点からの序列は $D<I<A$ となる．よって固有値は全て非負なので D は固有値が 0 に最も近く，逆行列を考える際に最も不安定さがあるといえる．

6.3.1　E 基準での A と D の固有値

A，D は基準の性質上，式 (6)，式 (8) のようにそれぞれ固有値の和と積で定義が可能である．水準が (4，4，4) 計画数を 24 に指定した場合の固有値に着目する．

表 5 から最小固有値は A が最も大きいが，最大固有値は D が最も大きいことが分かる．D は固有値の積を用いるため最小固有値が積に影響が出てしまうことを避け，和を用いる A よりも大きくとると思っていたが，実際は A のほうが大きくとっており予想と反した結果となった．その代わり，D は固有値の積を最大化する最適性を保つためにそれ以外の固有値，特に最大固有値を大きくとり，バランスを保っていることが分かった．A は固有値の逆数の和を最小化するように計画を考える．そのため他の計画よりも最小固有値を大きくとることで和の最小化に最も寄与しているといえる．注目した水準列以外でもこれらの傾向が見られた．

表 5　水準が (4，4，4)，計画数が 24 の固有値比較

	A	D	I
定義	固有値の和	固有値の積	定義なし
最小固有値	0.341	0.088	0.132
最大固有値	96.165	147.279	130.706

6.4　G 基準での計画評価

次に生成した最適計画 (A，D，I) を G 最適基準で評価する．G 基準は $d(\boldsymbol{x}_i)$ の最大値を最小化する基準である．評価値は A，D，I 基準でそれぞれ最適化した計画の情報行列 M を用いて

$$\max_i \boldsymbol{x}_i^T M^{A^{-1}} \boldsymbol{x}_i \quad \max_i \boldsymbol{x}_i^T M^{D^{-1}} \boldsymbol{x}_i \quad \max_i \boldsymbol{x}_i^T M^{I^{-1}} \boldsymbol{x}_i$$

となる ($i = 1, \cdots, n$)．その値が小さい計画を G 基準の観点においてベストな計画とする．しかし今回はパッケージ内の計算結果と合わせるため，式 (13) の G 効率値を評価値とする．G 効率値が最も大きい計画をベストな計画とする．水準が (5，5，4) の場合を以下に示す．

表 6　水準 (5，5，4) の G 基準平均値とベストな計画

計画数	A	D	I	ベスト
14	0.351	0.747	0.664	D
30	0.468	0.728	0.648	D
50	0.419	0.694	0.615	D
70	0.429	0.552	0.552	D I
90	0.438	0.491	0.491	D I

表 6 から D は 5 点，I は 2 点で水準が (5，5，4) の場合 D が最も良くなる傾向がある．水準が (3，4，5) の場合を以下に示す．

表 7　水準 (3，4，5) の G 基準平均値とベストな計画

計画数	A	D	I	ベスト
12	0.315	0.564	0.520	D
36	0.487	0.716	0.655	D
48	0.504	0.620	0.608	D
53	0.503	0.573	0.581	I

表 7 から D が 3 点，I が 1 点で水準が (3，4，5) の場合 D が最も良くなる傾向がある．評価した水準列，計画数を以下にまとめる．評価に用いた水準パターン，点数付けについては E 基準と同様である．

表 8 から総計では D が 37 点，I が 14 点，引き分けが 1 点となり，G 基準の観点からは D が最も良い結果となった．最終的な G 基準の観点からの序列は $A<I<D$ となる．結果を出す前は，G 基準と予測分散を小さくする立

表 8　G 基準のベスト数の比較

水準	A	D	I	引き分け
3 つ全て等しい	0	9	5	1
2 つ等しい	0	24	9	0
3 つ全て異なる	0	4	0	0
総計	0	37	14	1

場が同じである I 最適のほうが結果が良くなると予想していたが D の方が良い結果となった．D の予測分散の安定性はそれなりに高いのではと思われる．

6.5　基準ごとの選択する実験点の特徴

A，D，I 基準で選ばれた実験点の特徴を見る．実務において最良の実験点の位置を検討できない場合がほとんどであるため，バランス良く実験点を取ることが重要である．本研究は 3 因子で計画を構成しているため，立方体を考えることができる．実験点の特徴の把握しやすさを考え，3 つの水準が全て 3 水準の場合と 5 水準の場合を取り上げる (4 水準は中心となる点がないため省く)．3 水準の全ての実験点 3^3（27）点は，頂点 8，辺 12，面 6，中心 1 の位置に区分できる．5 水準の全ての実験点 5^3(125) 点は頂点 8，辺 36，面 54，中央 (中心付近)26，中心 1 の位置に区分できる．それぞれの水準の初期計画から計画数を指定すると基準ごとにどの位置の実験点を選択する傾向があるかを見ていく．

表 9　水準が (3，3，3)，計画数が 15，18，21

	15			18			21		
位置	A	D	I	A	D	I	A	D	I
頂点	8	8	7	7	8	8	8	8	8
辺	0	4	5	6	9	6	6	12	9
面	6	3	3	4	1	4	6	0	3
中心	1	0	0	1	0	0	1	1	1

表 9 から全て 3 水準の場合は，A は中心や面付近を取り，D，I は計画数がある程度多くならないと中心を取っていない傾向がある．D は頂点や辺を他の基準と比べて多く取る傾向がある．特に A との違いは顕著である．I は頂点や辺付近を取る傾向はあるが，面もそれなりに取っている．計画数が増えるにつれ，A は面，D は頂点，辺を先に全部取ろうとする傾向がある．

表 10　水準が (5，5，5)，計画数が 30，60，90

	30			60			90		
位置	A	D	I	A	D	I	A	D	I
頂点	8	8	8	8	8	8	8	8	8
辺	3	20	13	19	36	36	25	36	36
面	9	0	6	9	13	9	30	43	39
中央	9	1	2	23	2	6	26	2	6
中心	1	1	1	1	1	1	1	1	1

表 10 から全て 5 水準の場合は，A は中央付近が他の基準と比べて圧倒的に多い．A は 3 水準の場合は面を最も多く取る傾向があったが，5 水準で新たに中央という中心付近の位置が追加されたことで面ではなく，そちらを取る傾向に変わった．D は計画数が増えるにつれ，頂点，辺，面の位置の実験点を多く取る傾向がある．逆に中心付近の中央は計画数が増えてもほとんど取っていない．I は 5 水準の場合だと D の傾向がより強いと思われる．その理由として中央を A と比較するとほとんど取っていないため，I は D 寄りの傾向があるといえる．

表 9，10 から全体的にバランス良く実験点を取れていたのは，A と I 基準であった．また E と G の計画評価の結果を踏まえると中央付近が多いと固有値が大きく解析的に安定し，それ以外の外側付近が多いとモデルを用いた予測精度が上がるといえる．

7　まとめ

別の観点から計画の評価を行った結果，D は特別優れた基準ではないと思われる．最小固有値を小さくとる解析的な不安定さや実験点を外側中心に取る傾向から D だけに頼るのは良い実験計画を考える上で不十分である．今回どの観点からでも中立的な立ち位置であった I 基準が取り上げた 3 基準の中で最も万能であると思われる．I は基準の性質としては D に近いが A の解析的な安定さもある．また実験点を選択するときは，D と比べて部分的に集中しすぎず，実験点のバランスがとれている．これらを総合的に判断すると I が最も良い基準であるといえる．

8　おわりに

本研究を通して，それぞれの基準の性質を別の観点から評価できた．当初は評価指標にどの基準を用いるかで時間を使い，基準を変えて評価することも検討していたが，満足いく結果が出たと思う．しかし評価の観点はまだ残されているため．本研究とは異なる観点から評価を行うことでまた新規性のある結果が得られると思われる．

参考文献

[1] Atkinson，A. C. and Donev，A. N.：*Optimum Experimental Designs*，Oxford Science Publications，1992.

[2] 平尾将剛，澤正憲，神保雅一：『最適計画の構成法とその関連した話題』，数理解析研究所講究録，**1844**，23-38，2013.

[3] 永田靖：『入門実験計画法』，日科技連，2000.

[4] 奥村和也：『応答曲面の観点からみた最適計画の比較』，南山大学大学院数理情報研究科修士論文，2014.

[5] Package ‘ AlgDesign ’，https://cran.r-project.org/web/packages/AlgDesign/AlgDesign.pdf，2011.

[6] Wit，E.，Nobile，A. and Khanin，R.：Near-optimal designs for dual channel microarray studies，*Appl. Statist.*，**54**(5)，817-830，2005.

[7] 山田秀：『実験計画法-方法論-』，日科技連，2004.