

実験計画法に基づくパラメータ探索を用いた 大規模マルチエージェントシミュレーションの分析

松島 裕康 *1*2 山下 倫央 *1*2 野田 五十樹 *1*2
Hiroyasu MATSUSHIMA Tomohisa YAMASHITA Itsuki NODA

*1産業技術総合研究所 サービス工学研究センター
Center for Service Research(CfSR), The National Institute of Advanced Industrial Science and Technology(AIST)

*2独立行政法人科学技術振興機構, CREST
Japan Science and Technology Agency, CREST

In this paper, in order to search sensitive parameters to affect the result of the simulation effectively, we propose the method based on design of experiment. The results of the preliminary experiment are revealed that searching parameters by using analysis of variance contribute to find sensitive combination of parameters.

1. はじめに

人間や組織の活動をエージェント単位で表現可能なマルチエージェントシミュレーションを応用した災害、交通、経済などの社会シミュレーションによる予測から、様々な政策決定や社会現象の解析など、社会課題解決のアプローチとして注目されている。大規模な社会シミュレーションともなると、扱うパラメータの増加、個々のエージェントの判断やその相互作用などで系が分散してしまうなど現象やパターンそのものも増大し複雑化していく。また、対象とする社会システムの挙動の構造分析やシステム同定を行うために網羅的な試行をすると、さらに扱うデータが爆発的に増大する。

このように大規模化・複雑化するシミュレーションに対し、京などに代表されるようなスーパーコンピュータの計算資源を活用し、網羅的な試行を効果的に行うためのシミュレーション実行フレームワークが必要とされてきている。近年では、PRACTIS[岡田 2012] や OpenMOLE[Reuillon 2010] などの HPC を活用してシミュレーションの様々なケースを網羅的に実行可能なアプリケーションの開発が進められている。しかしながら、パラメータを網羅的に実行する時、例えば計算資源が余っていたとしても実験する必要のない部分を計算資源に当てるよりも、(1) シミュレーションの結果を俯瞰してみるために十分なパラメータの組み合わせの実験を効率的に行い、(2) 特徴的な部分では集中的にまたは優先的に実験をする仕組みが必要となる。

そこで本研究では、実験計画法を応用し (1) 効率的に実験パラメータの組み合わせを網羅するシフト付繰り返し実験計画と、(2) 分散分析を応用し、結果に影響を与えるパラメータの組み合わせに関してより細かく実験していく方法について提案し検証する。

2. 実験計画法

実験計画法は、イギリスの農業試験場の技師として働いていた R. A. Fisher によって開発された効率的に実験を進めることを目的とした統計学の応用である。

連絡先: 松島 裕康, 産業技術総合研究所, 茨城県つくば市梅園 1-1-1 中央第二, Tel:029-861-2713, Fax:029-862-6548, E-mail: hiroyasu.matsushima@aist.go.jp

実験計画法では、検討する対象である因子と各因子に設定する段階を水準とし、各因子の直接的効果を主効果や複数の因子が互いに独立でない場合は因子間の交互作用を検討していく。特に、各因子の各水準が他のすべての因子の水準と組合せを効率的に実験する手法として代表的な直交表を用いた方法が挙げられる。具体的に例として、下記の 1 は $L_4(2^3)$ の直交表であり、 L はラテン方格の頭文字、添字の 4 は実験回数、括弧内の 2 は因子の水準数、右肩の 3 は因子の個数を示す。つまり、4 回の実験で 2 水準の因子が 3 つまで調べられる直交表を示す。また、第 3 列目は第 1 列目と第 2 列目の交互作用として割り付けることも可能である。

例えば、ある因子 A と B について実験を行う場合、 A と B は 2 のように第 1 列目と第 2 列目に割り付けられる。また、 A と B の交互作用の効果についても考慮したい場合は、 A と B を割り付けた第 1 列目と第 2 列目をそれぞれベクトル a , b とした時、第 3 列目は $a + b$ となるので、 A と B の交互作用を割り付けることができる。

表 1: $L_4(2^3)$ の直交表

| \ | 1 | 2 | 3 |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 1 | 0 | 1 | 1 |
| 2 | 1 | 0 | 1 |
| 3 | 1 | 1 | 0 |

表 2: 直交表 $L_4(2^3)$ を用いた因子の割り付け

| \ | 1 | 2 | 3 |
|---|-------|-------|------------------|
| 0 | A_1 | B_1 | $A_1 \times B_1$ |
| 1 | A_1 | B_2 | $A_1 \times B_2$ |
| 2 | A_2 | B_1 | $A_2 \times B_1$ |
| 3 | A_2 | B_2 | $A_2 \times B_2$ |

また他にも、 $L_8(2^7)$ や $L_{27}(3^13)$ などの直交表がある。

3. シフト付繰り返し実験計画

直交配列を用いた実験計画では、繰り返しが必要な場合、同じ直交配列の組み合わせで繰り返しを行う。しかし、同じ直交配列しか用いない場合、各要因の交互作用を発見することはできない。繰り返しを通じて交互作用の可能性を検証する場合には、異なる直交配列を用いて実験をすすめる必要がある。

そこで、本節では、異なる直交配列をシステムティックに作っていく事で、実験の直交性をできるだけ保ちつつ、計算資源が許す限り、要因の組み合わせを網羅していく方法を考える。

まず、2水準の直交配列 $L_{2^n}(2^{(2^n-1)})$ を以下の手順で構築していく。まず、ある整数 x を2進数で表した時、1の位より数えて k 桁目の位の数 ($\{0,1\}$) を $B^k(x)$ と表す。

1. $2^n - 1$ 個の要因を順番に並べ、1より番号を付与する。
2. $2^n - 1$ 個の $\{0,1\}$ 要素からなる配列を考え、各々の要素を上記の要因に順番に対応づける。
3. すべての要素が0である配列を初期配列 ($t=0$) とする。
4. t 番目の配列を作るときは、 $t-1$ 番目の配列より以下の手順で作成する。

- (a) $t \bmod 2^k = 1$ となる k を求める。
- (b) $\forall i \in \{1 \dots 2^{n-1} - 1\}$ について、 $B^k(i) = 1$ となる場合、 $t-1$ 番目の配列の第 i 要素の 0/1 を反転して t 番目の要素とする。 $B^k(i) = 0$ の場合には、 $t-1$ 番目の配列の第 i 要素をそのまま t 番目の要素とする。

例えば、 $L_{2^3}(2^{(2^3-1)})$ の場合、以下のように直交配列が作られる。

| | | | | | | | |
|-----------------|---|---|---|---|---|---|---|
| $t \setminus i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| 2 | 1 | 1 | 0 | 0 | 1 | 1 | 0 |
| 3 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
| 4 | 0 | 1 | 1 | 1 | 1 | 0 | 0 |
| 5 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| 6 | 1 | 0 | 1 | 1 | 0 | 1 | 0 |
| 7 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |

(1)

また、要因あるいは要素の順番 i を二進数で表した時に1つだけが1であるものを、配列のシグナチャ部と呼び、それ以外の部分を随伴部と呼ぶことにする。上記の $L_{2^3}(2^{(2^3-1)})$ の例では、1,2,4 番目の要素がシグナチャ部、それ以外が随伴部である。この時、 $L_{2^n}(2^{(2^n-1)})$ 中の各配列のシグナチャ部は、 $L_{2^n}(2^{(2^n-1)})$ でユニークとなる。

次に、上記の手順の中で、3の初期配列を異なるものに変更してできる配列を、 $L_{2^n}(2^{(2^n-1)})|z$ で表す。但し、 z はその初期配列とする。すなわち、もともとの直交配列 $L_{2^n}(2^{(2^n-1)})$ は、 $L_{2^n}(2^{(2^n-1)})|000\dots 0$ となる。この手順で作られた配列は、各々直交配列となる。例えば、 $L_{2^3}(2^{(2^3-1)})|0000101$ は以下のようなになる。

| | | | | | | | |
|-----------------|---|---|---|---|---|---|---|
| $t \setminus i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |
| 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| 2 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |
| 3 | 0 | 1 | 1 | 0 | 1 | 1 | 0 |
| 4 | 0 | 1 | 1 | 1 | 0 | 0 | 1 |
| 5 | 1 | 1 | 0 | 1 | 1 | 0 | 0 |
| 6 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| 7 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |

(2)

この $L_{2^n}(2^{(2^n-1)})|z$ について、すべて0の配列のうち、その随伴部の任意の要素を任意個1に変更したものを初期配列 z とした場合、それにより生成される各直交配列の各要素はす

べて異なる。さらに、随伴部は $2^n - 1 - n$ 桁あることから、このような初期配列の組み合わせは全てで 2^{2^n-1-n} である。これと、各 $L_{2^n}(2^{(2^n-1)})|z$ に含まれる 2^n 個の配列を組み合わせると、全てで 2^{2^n-1} 通りとなり、すべての要因の組み合わせの数と同じとなる。つまり、随伴部のみ変更した初期配列で生成する直交配列 $L_{2^n}(2^{(2^n-1)})|z$ に従って実験を進めれば、直交性をできるだけ維持しつつ、網羅的な要因の組み合わせ探索が行える。この方法を、「シフトつき繰り返し実験計画」と呼ぶことにする。

4. 分散分析を用いたパラメータ探索

シミュレーションのパラメータを因子として、実験計画法の直交表を用いて実験するパラメータの組み合わせを設定し、分散分析から各因子の主効果および交互作用の効果を判定するF検定により、パラメータの変化によるシミュレーション結果の変動を効率的に見つけ出す。ここでは、 n 人版繰り返し囚人のジレンマゲームを例題にとり、ノイズが混ざると各戦略をもつエージェントの分布が変化するパラメータの組み合わせを探索する手法を説明していく。

4.1 n 人版繰り返し囚人のジレンマゲーム

繰り返し囚人のジレンマゲームは、2人非ゼロ和ゲームの一種で、表3に代表される利得行列を用いて以下の手順で行われる。二人のエージェントは協調、または裏切りのどちらかの行動を選択し、選択した行動に応じて、表3の利得行列に従い両者は得点を得る。この対戦を繰り返し行い、合計得点を競うゲームである。 n 人版繰り返し囚人のジレンマゲームは、上

表 3: 囚人のジレンマゲームの利得表

| | | |
|--------|--------|--------|
| | A(協調) | A(裏切り) |
| B(協調) | (3, 3) | (0, 5) |
| B(裏切り) | (5, 0) | (1, 1) |

記の2人だけのゲームから複数人対戦となり、エージェントは各自がもつ戦略とそれに従った協調か裏切りのどちらかを選択する。さらに、多くの利得を獲得可能な戦略は徐々に増えていき、利得の少ない戦略は逆に淘汰されていくように拡張されている。1世代でエージェントの追加と淘汰が1度ずつ実行され、それを指定した世代数だけ実行すると、殆どの場合、ある戦略に収束していくのが一般的であるが、ある確率で選択した行動とは逆の行動を取らせるようにノイズを加えると、図1のように、各戦略をもつエージェント数の母集団に対する割合が、世代ごとに変化している様子が見られるようになる。この

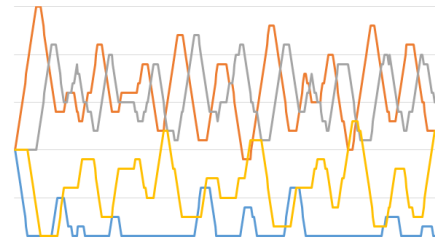


図 1: ノイズの影響を受けた n 人版囚人のジレンマゲームにおける各戦略の母集団の割合の推移

ノイズと繰り返しゲーム回数という2つのパラメータが母集

団における戦略の割合の変動に関してどのように影響を与えているかを例として、どのように分散分析を用いてパラメータを調査していくかを説明する。ここで母集団における戦略の割合の変動を、最も数の多い戦略が入れ替わった回数としてカウントとして置き換えることにより、1つの出力として簡単に扱えるようにする。直交表 $L_4(2^3)$ を用いた場合、ノイズと繰り返しゲーム回数をパラメータ A, B とし、の表2のように割り付けをすることができる。初期パラメータを(ノイズ, 繰り返しゲーム回数) = $([0.0, 0.05], [10, 100])$ とし、各組み合わせについて2回ずつ実験し、それらの結果による表4のような分散分析表を作成することができる。この時、 M, e, SS, f, V, F , はそれぞれ平均, 残差, 全変動, 自由度, 分散, 対応する要因を残差の V 値で割ったものに対応する。これと f 分布表を用いて f 検定を行う。例えば表4においては、残差に比べて有意な要因は**として表されている(有意水準 $\alpha = 0.01$)。 f 検定により判定された有意なパラメータについて図2のよう

表 4: 分散分析表の例

| | SS | f | V | F | 判定 |
|--------------|---------|---|---------|--------|----|
| M | 253.125 | 1 | 253.125 | | — |
| A | 171.125 | 1 | 171.125 | 1369.0 | ** |
| B | 171.125 | 1 | 171.125 | 1369.0 | ** |
| $A \times B$ | 171.125 | 1 | 171.125 | 1369.0 | ** |
| e | 0.5 | 4 | 0.125 | — | |

に分割点のパラメータの組み合わせを実験し4つの点で囲まれた領域について分散分析を再び行う。例えば、ノイズに関するのみ有意と判定された場合、ノイズパラメータの初期水準の区間を2つに分割し、図2の左下のように黒い線で囲まれた2つの領域について分散分析を行う(例。 $([0.0, 0.05], [10, 100])$ と $([0.05, 0.1], [10, 100])$ のそれぞれにおいて分散分析)。また、ノイズと繰り返しゲーム回数の両方または、交互作用について有意であった場合は、図2の右下のように両パラメータの区間を分割し各領域について分析することにより、有意なパラメータに関してより詳細な実験を実施することが可能となる。

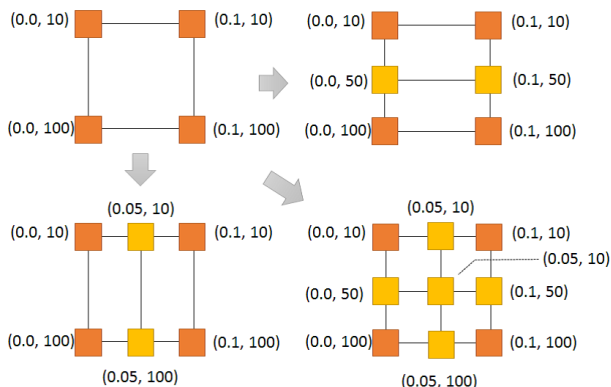


図 2: パラメータの2分割探索法

5. 実験

本論文では、シフトつき繰り返し実験計画と分散分析を用いたパラメータ探索について実験及び、その効果について検証していく。

5.1 シフトつき繰り返し実験計画の検証

提案したシフトつき繰り返し実験計画の各実験の直交性を確認するために、以下のような実験を行う。3.章に示した手順に従い、 $L_{2^n}(2^{(2^n-1)})|z$ を生成してゆき、その途中までの配列の各要素の共分散を求める。対象としては、各要素をランダムに0/1に決めた配列を生成するモンテカルロ法で同様の共分散を求める。

その結果を3に示す。このグラフでは、横軸が実験の回数(生成した配列の数)、縦軸が非対角成分の共分散の分布の標準偏差を示しており、赤線が提案手法、緑線がモンテカルロ法による結果を示している。また、要素の数としては、 $n = 4, 6, 8$ の場合の結果を示している。この結果からわかるように、提案手法は実験を重ねるごとに要素間の相関を急速に低減できることがわかる。これにより、実験の直交性を保ちながら、網羅的探索が可能となると考えられる。

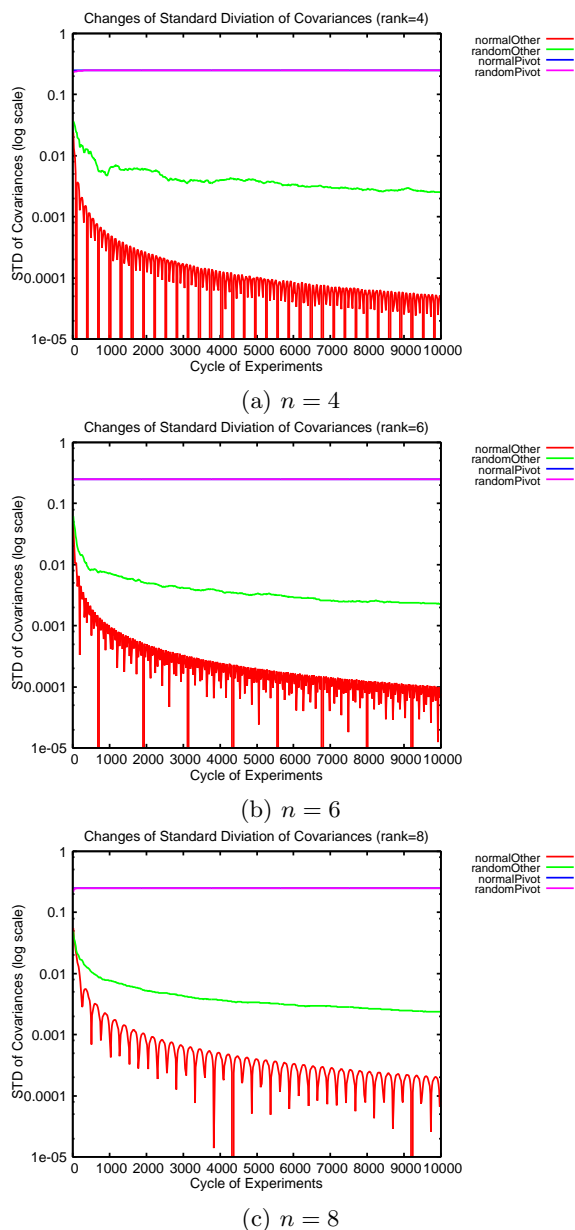


図 3: シフト付繰り返し実験計画の直交性検証

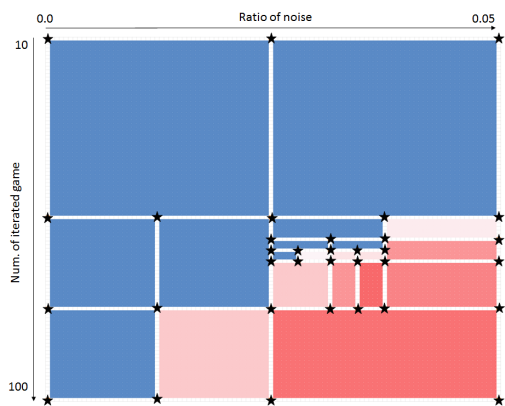


図 4: F 検定を用いたパラメータの効果の探索結果

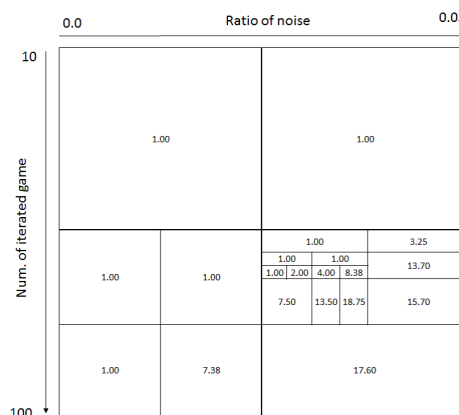


図 5: パラメータの組み合わせによる出力の平均

5.2 分散分析を用いたパラメータ探索

次に、提案した直交表によるパラメータ配置からの分散分析を用いたパラメータ探索について検証するための第一歩として、 n 人版繰り返し囚人のジレンマゲームを例題に実験する。ここでは、(1) 常に協調、(2) 常に裏切り、(3) しつぺ返し (最初のみ必ず協調を選択)、(4) ランダムに選択の 4 種の戦略を採用し、ノイズと繰り返しゲーム回数のパラメータを入力、最も多い戦略が入れ替わる回数を出力とし、各戦略をもつエージェントの人口変化が振動するパラメータの境界を探索していく。各入力の初期の水準は、(ノイズ, 繰り返しゲーム回数) = $([0.0, 0.05], [10, 100])$ と設定した。例題の他のパラメータとして、母集団を 100, 世代数を 300, 各戦略のエージェント数 25 体ずつとし、毎世代、最も利得の多い戦略のエージェントを 1 体追加し、最も利得の低いエージェントを 1 体淘汰する。 $L_8(2^7)$ の直交表を用いてパラメータを割り付けるが、1 の直交表を例にとると、ノイズを第 4 列に、繰り返しゲーム回数を第 6 列に、そして、これらパラメータの交互作用については第 2 列に割り付けて検証する。また、f 検定を実施する際の危険率 α の値を 0.01 として検定する。

実験結果である図 4 と 5 は縦軸に実験回数、横軸にノイズを示したパラメータ空間を表し、両図は対応している。図 4 における \star の点がパラメータの組み合わせを実験した箇所であり、それらで囲まれた青や赤の領域が分散分析を用いて F 検定を行った箇所である。図 5 の各領域における数値は各パラメータの組み合わせによる出力の平均値を示しており、図 4 と対応させて値が小さいほど、青色を示し、逆に値が高いほど濃い赤色を示している。例えば、一番左上の領域はノイズの 2 水準 (0.0, 0.25) と、繰り返しゲーム回数の 2 水準 (10, 50) の組み合わせを実験したところ、平均値は 1.0 であり F 検定を実施した結果、残差に比べて有意ではないと判断し、この水準の範囲内ではこれ以上パラメータを細かく分割して実験されない。結果をみると、青と赤の境界付近に近づくほど領域が細かく分割され、その付近から遠くなるほど領域は大きくなっていることが確認できる。このことから、本論文で設定したパラメータを用いた n 人版繰り返し囚人のジレンマゲームにおいてノイズと繰り返しゲーム回数が各戦略をもつエージェントの母集団に占める割合に影響を及ぼしている部分を探ることができた。

6. おわりに

実験計画法を応用し (1) 効率的に実験パラメータの組み合わせをシステマティックに網羅するシフト付繰り返し実験計画と、(2) 分散分析を応用し、結果に影響を与えるパラメータの組み合わせをより細かく探索する方法について提案しその有効性について検証した。(1) シフト付繰り返し実験計画はモンテカルロ法で生成した配列の組み合わせよりも、要素間の相関を急速に低減しつつ網羅的に探索できることを示した。(2) 分散分析を用いたパラメータ探索は、 n 人版囚人のジレンマゲーム上で母集団に対する戦略数の変動に影響するパラメータの組み合わせを効率的に探索できることを示した。

今後の課題としては、適用する問題をより規模を大きく、または複雑化した問題を通して検証する必要がある。また、より適切にパラメータの差異とその影響を確認していくためには、水準の値を細かくしてだけでなく広げて探索する方法も取り入れていくことが必要であると考えられる。

謝辞

本研究は、科学技術振興事業団「JST」の戦略的基礎研究推進事業「CREST」における研究領域「ポストペタスケール高性能計算に資するシステムソフトウェア技術の創出」の研究課題「超大並列計算機による社会現象シミュレーションの管理・実行フレームワーク」の支援を受けています。

参考文献

- [岡田 2012] 岡田 崇, 山下 倫央, 野田 五十樹: “網羅的シミュレーションによる群集誘導手法の検証”, 人工知能学会全国大会 (JSAI)2012, 3F2-OS-10-9, 2012.
- [Reuillon 2010] Reuillon, R., et al.: “Declarative Task Delegation in OpenMOLE”, International Conference on High Performance Computing and Simulation (HPCS) 2010, pp.55-62, 2010.