

多次元パラメータを用いた マルチエージェントシミュレーションの 結果に対する要因分析

*Causation Analysis for the Results of Multi-Agent
Simulation Implemented by Multi-dimensional Data*

東京大学大学院工学系研究科
塚本鋭

目次

- 背景
- 目的
- 手法
- 空間囚人のジレンマモデル
- 問題提起
- 数値実験
- 数値実験の検証
- 結果の可視化
- 結言

背景

- 単純な個体が複雑な動きを創発する現象を解明するために、マルチエージェントシミュレーション(MAS : *Multi-Agent Simulation*)が注目されている
- 現在のMAS研究は、シミュレーション内容をより現実問題近づけるためにモデル自体を改良することが目的とされていることが多い
- 一方で、MAS研究には、ある1つのパラメータが少し違うだけで結果が大きく異なってしまうという問題が存在する

背景

- この問題を解決するためには、可能な限り多くのパラメータセットを用いて実験を行い、結果を比較することが、結果の信頼性を高めるために有用な手法であると考えられることができる
- 今までには計算能力の限界により、この手法はあまり検討されてこなかった
 - シミュレーションの中で、パラメータ自身を変化させることで信頼性を確保する手法が主流だった
- しかし近年、計算機能力の飛躍的な向上により、比較的安価で大規模なシミュレーションが行えるようになった

目的

- MASにおける結果の信頼性を向上させるための新しい方法論に取り組む
- 大規模なパラメータスタディを行うことで、パラメータが結果に与える影響を考察する
- 大量の結果から有意な情報を取り出すための方法を考察する

多次元パラメータを用いる単純なモデルとして
空間囚人のジレンマ(SPD : *Spatial Prisoner's Dilemma*)
モデルを考える

空間囚人のジレンマとは

生物学や経済学、社会学などの様々な分野において、利己的な行動主体が、利他的行動(協調行動)を選択するメカニズムの解析に用いられている

囚人のジレンマ問題とは

利得行列

相手

自分

| | 協力(C) | 裏切(D) |
|-------|--------|--------|
| 協力(C) | R, R | S, T |
| 裏切(D) | T, S | P, P |

ただし

$$S < P < R < T$$

$$S + T < 2R$$

を満たす

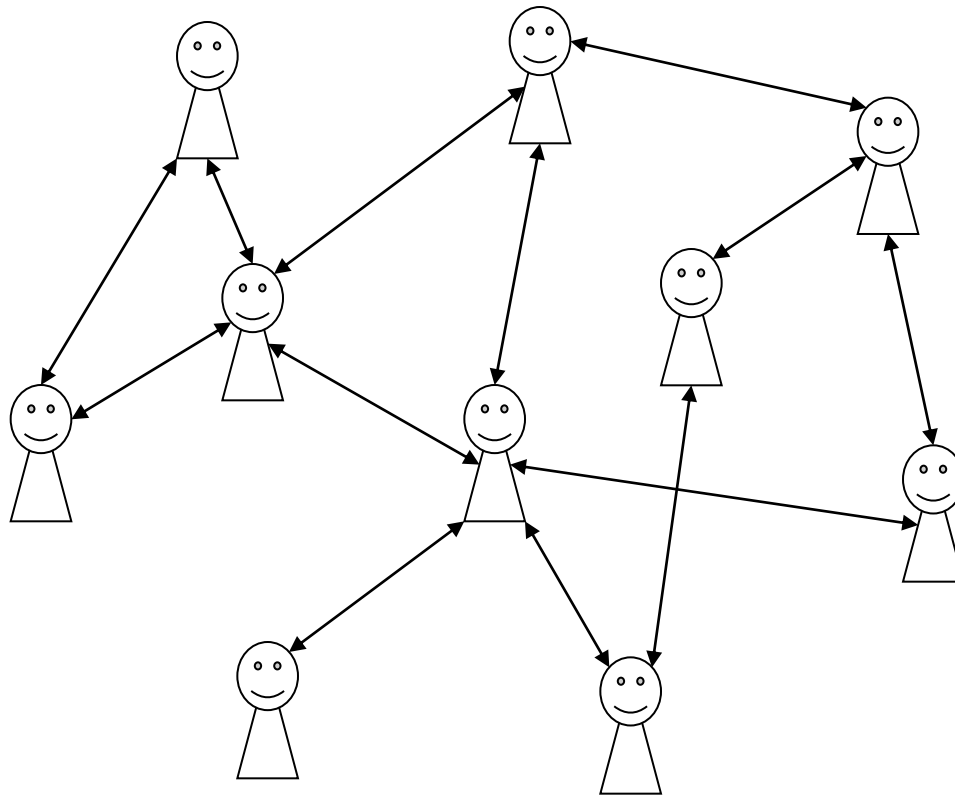
自分も相手も裏切を選択するのが合理的な行動である

限定された複数の相手と繰り返し対戦する(ネットワーク構造を導入する)ことで協調戦略を取る行動主体が現れる
これが空間囚人のジレンマ (SPD) モデル [Nowak 92]

SPDモデル

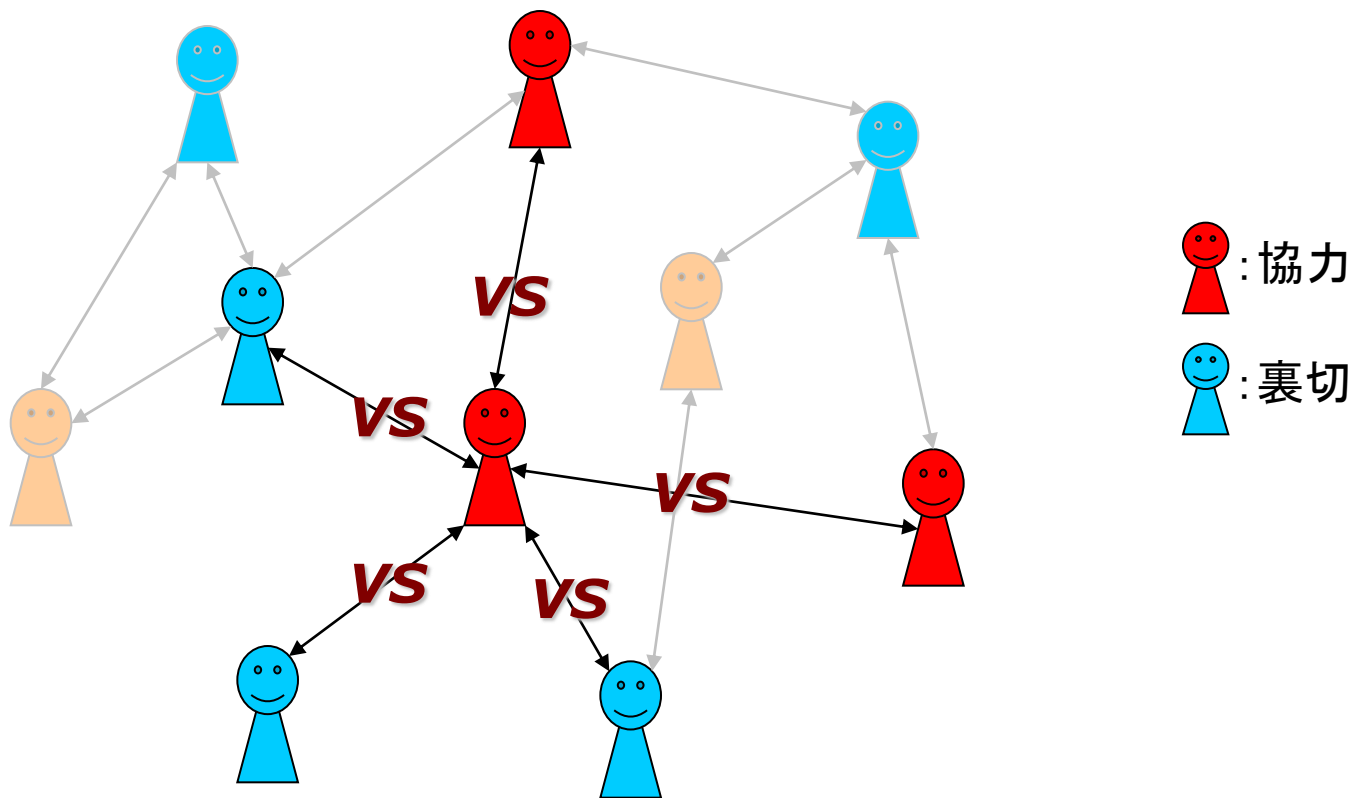
初期ネットワークを生成し、すべてのノード上にエージェントを配置する

本研究では、ネットワークは変化しないものとする



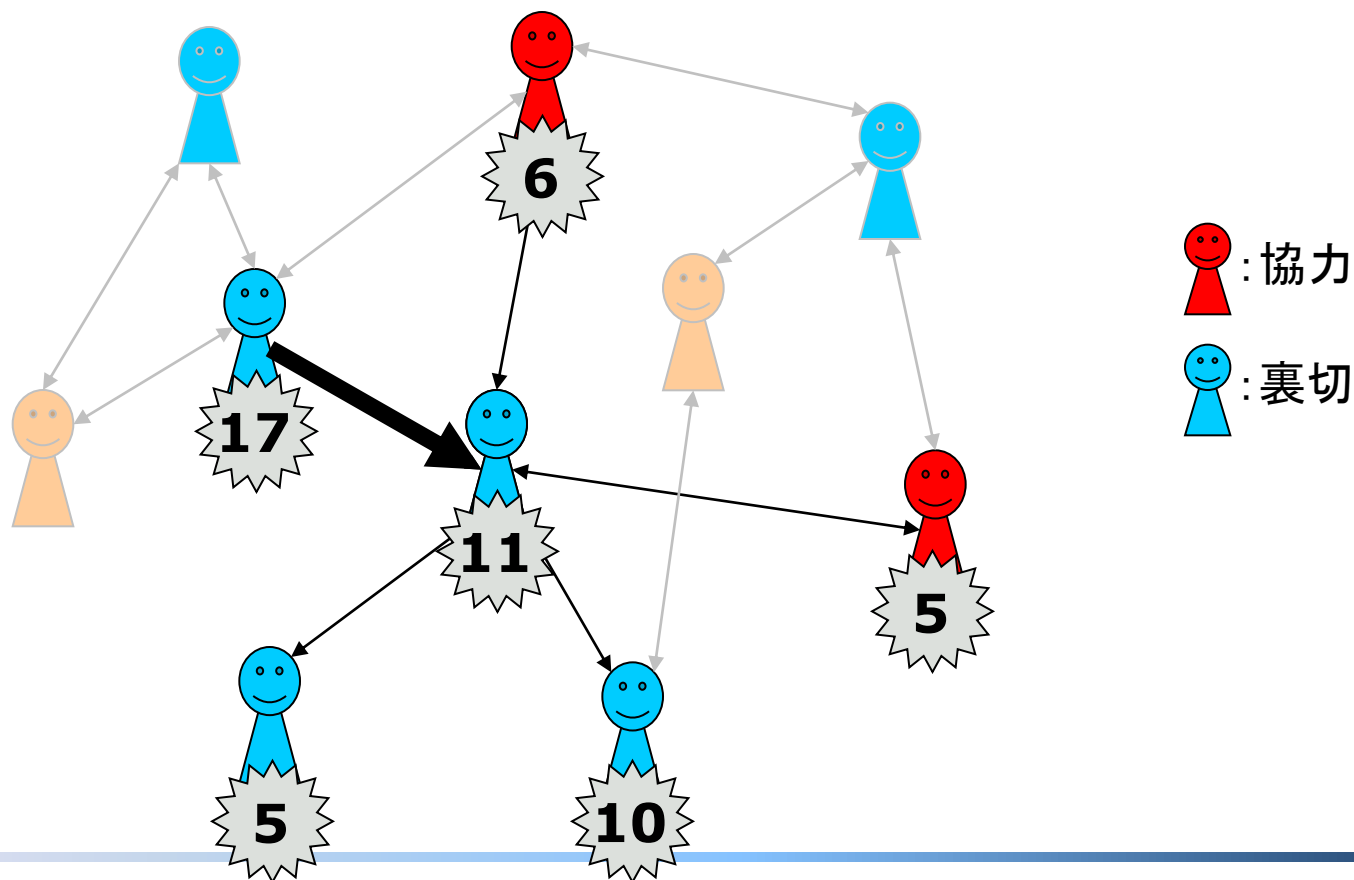
SPDモデル

エージェントは各ステップにおいて協力(C)か裏切(D)という戦略をとり、リンクでつながっているすべてのエージェントと囚人のジレンマゲームを行う(対戦)



SPDモデル

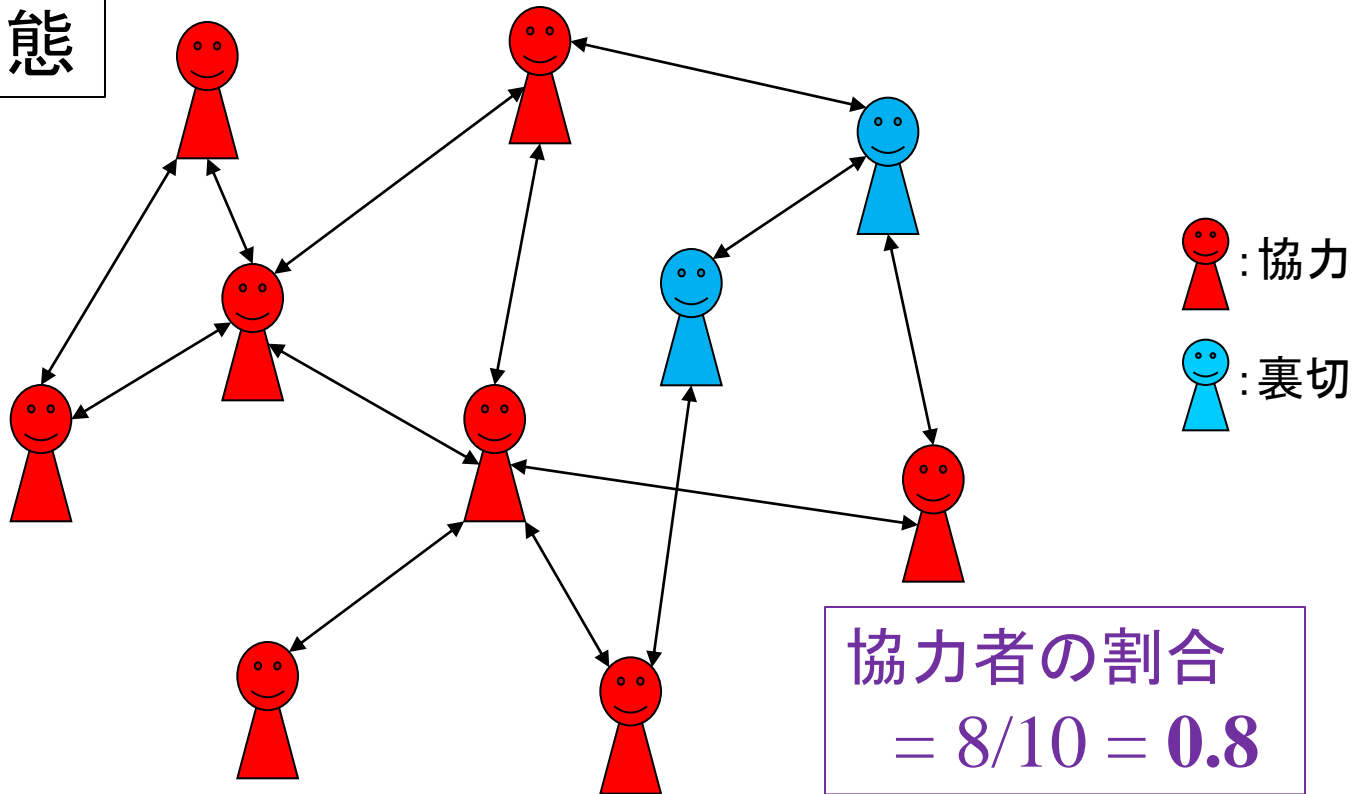
エージェントは自分と周りの報酬に応じて戦略を更新する
最高報酬を得たノードを模倣するというモデルを使用(模倣)



SPDモデル

対戦と模倣で1ステップとし、すべてのエージェントの戦略が安定する(終状態)まで、繰り返し行う
終状態での協力者の割合を結果とする

終状態



問題提起


SPDの既存研究には相反する2つの知見が存在している

Santos et al.(2005)

スケールフリー性(≡ Heterogeneity)を有するネットワークが協力者の割合を増加させることを示した

Fu et al.(2007)

Heterogeneityなネットワークは必ずしも協力者の数を増やさないことを示した



これらの相反する知見が存在する理由は、用いた初期パラメータの値が異なるためであると考えられる
知見を深めるためには、より系統立った網羅的な研究を行う必要性がある

数値実験

- スケールフリー性を有するネットワークを網羅的に作成し、それらのネットワークがどの程度協力者の割合を増やすか観測する
- ネットワークを生成するために4個の設計変数(べき乗指数※1、最小次数、最大次数、最大次数のノード数)を設定する
- またネットワークの重要な変数である2個の説明変数(ノード数、平均次数)を考慮する
- 計6個の変数が協力者の割合に与える影響を調べる

※1 べき乗指数とは、ネットワークの次数分布における「傾き」を示す Heterogeneityを示す1つの指標である

数値実験のパラメータ

ネットワーク生成のパラメータ

| | | | | | | | | | | | | | | | | | | | | |
|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ベキ乗指数 | 1.5 | 1.7 | 1.9 | 2.1 | 2.3 | 2.5 | 2.7 | 2.9 | 3.1 | 3.3 | 3.5 | 3.7 | 3.9 | 4.1 | 4.3 | 4.5 | 4.7 | 4.9 | 5.1 | 5.3 |
| 最小次数 | 1 | 2 | 3 | 4 | 5 | 6 | | | | | | | | | | | | | | |
| 最大次数 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 60 | 70 | 80 | 90 | 100 | | | | | | | |
| 最大次数のノード数 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 | |

このなかで計算時間の関係からノード数が60000以下である
10319個のパラメータセットに対して実験を行った

囚人のジレンマゲームの利得行列

| | | |
|----|---------|--------|
| | 協力 | 裏切 |
| 協力 | 1, 1 | 0, T |
| 裏切 | T , 0 | 0, 0 |

初期配置は協力と裏切を半
数ずつランダムに配置する

各施行5000ステップまで行い、 $T=1\sim 2$ で0.04刻みの26パ
ターン、各30施行(計780施行)の平均を1つのパラメータ
セットにおける結果(協力者の割合)とする

数値実験の結果

パラメータセットと結果の例

| 協力者の割合 | ベキ乗指数 | 最小次数 | 最大次数 | 最大次数のノード数 | ノード数 | 平均次数 |
|--------|-------|------|------|-----------|-------|-------|
| 0.275 | 1.7 | 6 | 20 | 20 | 846 | 10.32 |
| 0.517 | 5.1 | 6 | 15 | 60 | 12848 | 7.15 |
| 0.425 | 2.7 | 5 | 15 | 50 | 2964 | 7.25 |
| 0.519 | 4.5 | 4 | 20 | 3 | 7243 | 4.95 |
| 0.397 | 2.1 | 1 | 25 | 9 | 11894 | 2.18 |
| 0.477 | 3.3 | 6 | 20 | 10 | 1593 | 8.42 |
| 0.128 | 1.9 | 5 | 90 | 1 | 1376 | 15.21 |
| 0.593 | 4.5 | 5 | 35 | 2 | 25383 | 6.35 |
| 0.320 | 2.3 | 4 | 40 | 20 | 13839 | 8.29 |
| 0.501 | 2.1 | 2 | 15 | 4 | 608 | 3.84 |
| 0.411 | 1.9 | 2 | 25 | 9 | 2808 | 4.92 |
| 0.270 | 1.7 | 1 | 45 | 5 | 6312 | 3.93 |
| 0.372 | 2.3 | 5 | 20 | 6 | 548 | 8.31 |
| 0.453 | 4.9 | 4 | 15 | 7 | 7353 | 4.77 |
| 0.535 | 2.9 | 4 | 20 | 9 | 2459 | 6.16 |
| 0.530 | 4.7 | 6 | 15 | 50 | 7919 | 7.29 |

数値実験結果の検証

1. 設計変数4個が協力者の割合に“線形”に影響を与えていると仮定する



設計変数による重回帰分析

以降、数値実験の10319個の結果を、学習用データ5159個と検証用データ5160個にランダム関数を用いて分割したものを利用する

本研究では、モデルによって導かれた結果と実験結果の整合性を測る指標として、相関係数を用いる※2

※2 これは後のVMSにおける検証結果と一貫性を保つためであり、より厳密には結果の差の二乗和等の指標を用いることが望ましいと考えられる

重回帰分析

回帰式の係数の大きさによって各変数の影響度を確かめるために、各変数を[0,1]に正規化する

Ex) 最小次数

| | | | | | | |
|------|---|-----|-----|-----|-----|---|
| 最小次数 | 1 | 2 | 3 | 4 | 5 | 6 |
| 正規化後 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 1 |

$$(\text{正規化後の値}) = \frac{(\text{正規化前の値} - \text{最小値})}{(\text{最大値} - \text{最小値})}$$

Excelによる重回帰分析

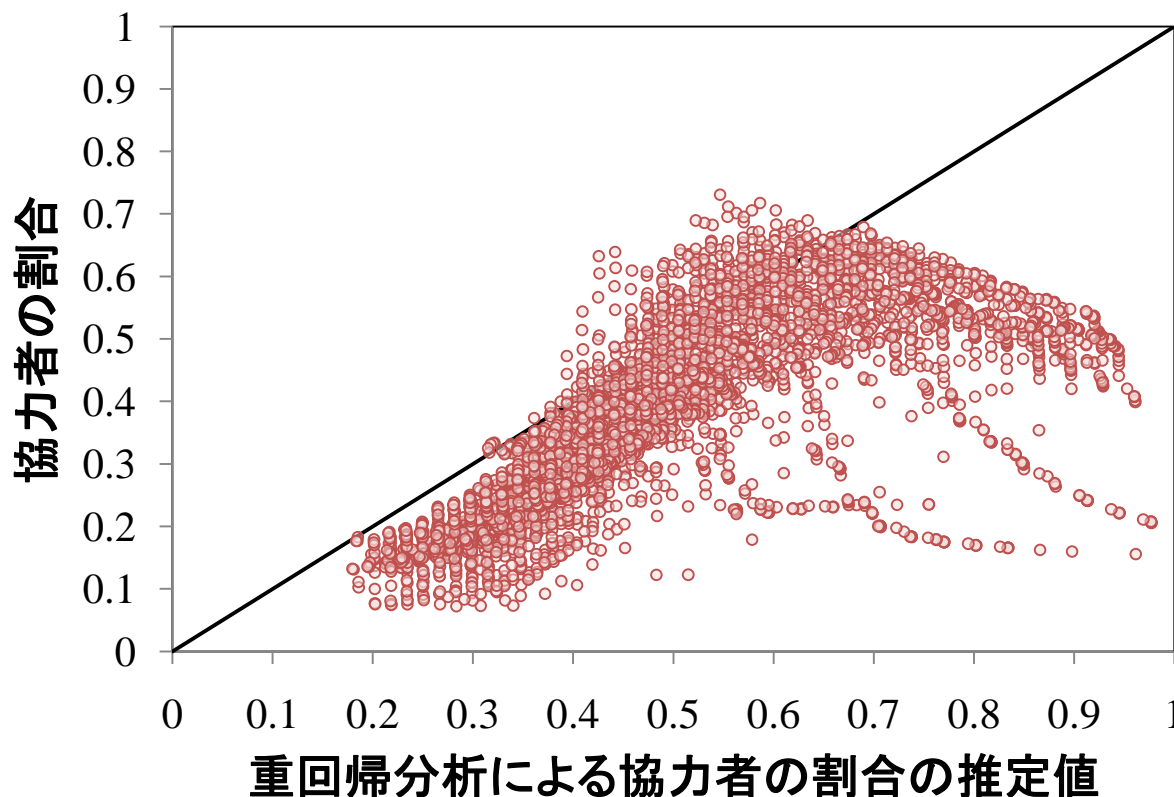
ベキ乗指数、最小次数、最大次数、最大次数のノード数をそれぞれ x_1 、 x_2 、 x_3 、 x_4 とする

回帰式

$$\text{協力者の割合} = 0.60x_1 + (-0.08x_2) + (-0.13x_3) + (-0.01x_4) + 0.40$$

重回帰分析

回帰式からの推定値と実験値の比較



相関係数
 $R = 0.743$

設計変数と協力者の割合には線形の関係性があるとは言えない

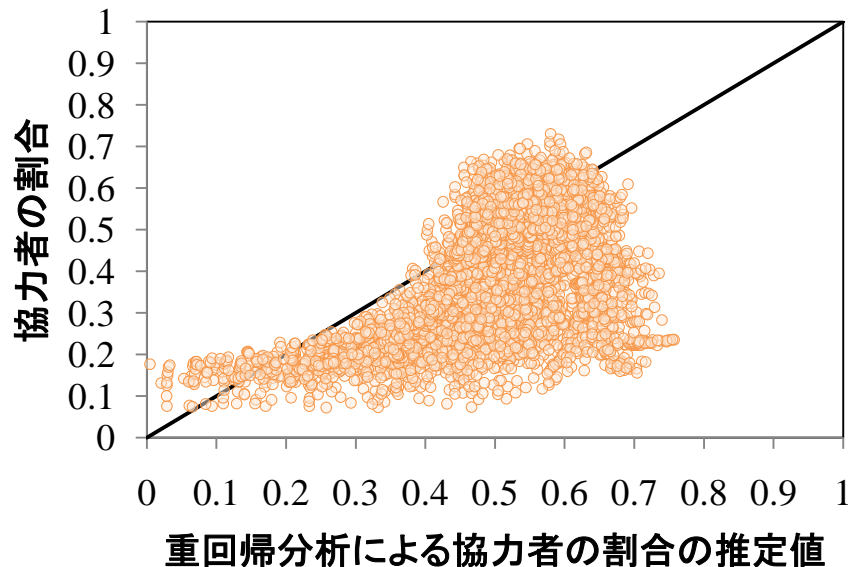
数値実験結果の検証

2. 説明変数2個が協力者の割合に“線形”に影響を与えていると仮定する

ノード数、平均次数を正規化し、それぞれ x_5 、 x_6 として重回帰分析を行う

回帰式

$$\text{協力者の割合} = 0.07x_5 + (-0.71x_6) + 0.70$$



相関係数

$$R = 0.565$$

変数と協力者の割合に線形な関係性があるとは言えない

数値実験結果の検証

3. 設計変数4個が協力者の割合に“非線形”に影響を与えていると仮定する



ニューラルネットワークを用いた分類分析

ニューラルネットワークとは

脳内の神経回路を模した情報処理機構であり、中間層を用いることで、高度で複雑なモデルを構築することができる

本研究では3層(中間層(隠れ層):1層)のニューラルネットワークモデルを用いるものとする

数値実験結果の検証

ニューラルネットワークを用いる際の問題点

ニューラルネットワークを用いて、設計変数の影響度を測定する際には感度分析が用いられる

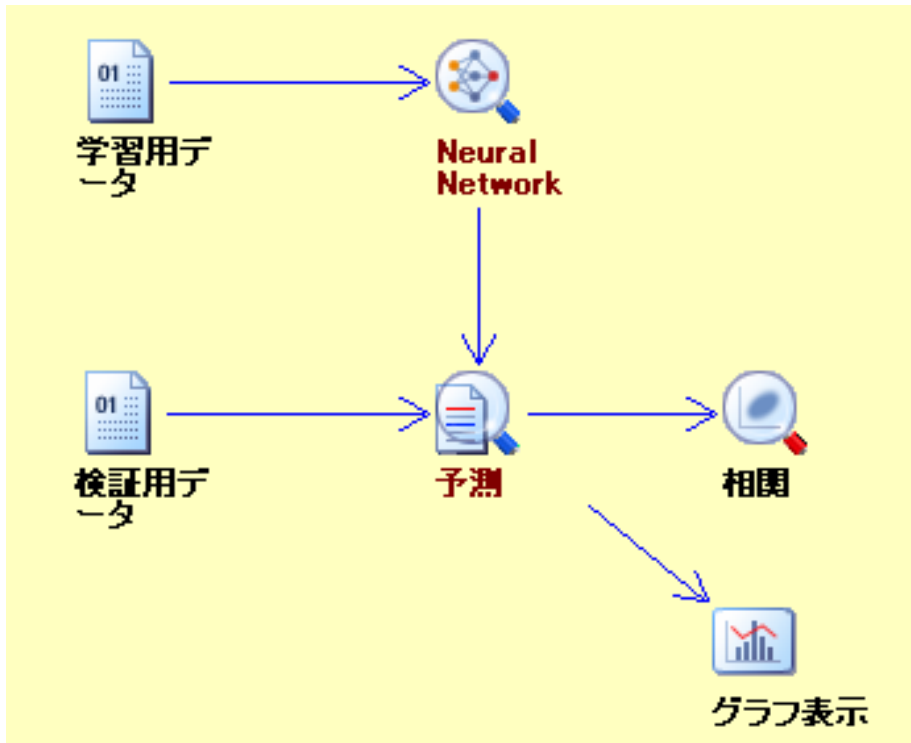
ニューラルネットワークにおける基本的な感度分析の方法は、測定変数以外の変数を平均値に固定し、測定変数を最小値から最大値まで変化させた時の、結果の範囲を調べるというものである
しかしこの方法は線形な関係性の時にしか使用できない



本研究では、測定変数を含むモデルと含まないモデルの予測精度の差に注目し、測定変数が結果に与えている影響を見積もる

ニューラルネットワーク

Visual Mining Studio(VMS)を用いてニューラルネットワークモデルを構築・検証するモデル作成した



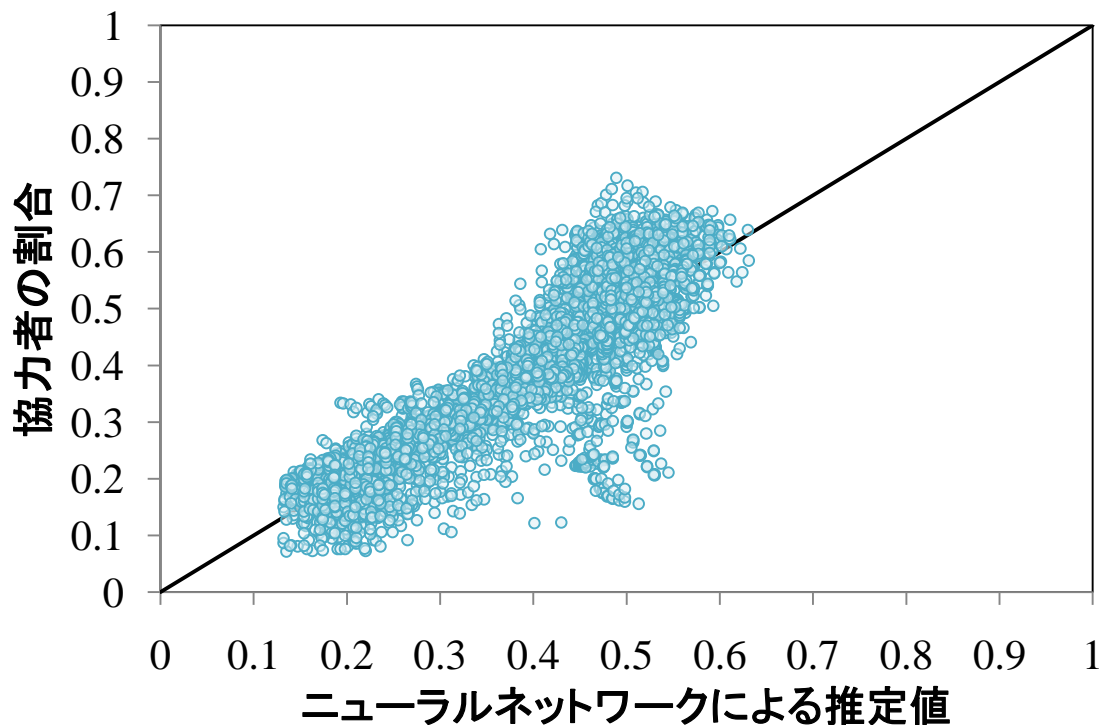
先ほどと、同様にデータを学習用と検証用に分割して用いている

VMSで実装されているファイル分割のコマンドを用いていないのは、各実験で同じデータを用いて比較するためである

ニューラルネットワーク

設計変数4個を用いたモデルの予測精度を求める予備実験より、最も予測精度が高かった6個のユニットの隠れ層のモデルを用いた

6個のユニットからなるモデルによる推定値と実験値の比較



目的関数 :

Liner-Sum of Square

繰り返し最大数 : 100

重みなし

相関係数

$R = 0.879$

非線形な関係性がモデルで表現されている

ニューラルネットワーク

測定変数をモデルに含めない時のモデルからの予測値と実験値の相関係数

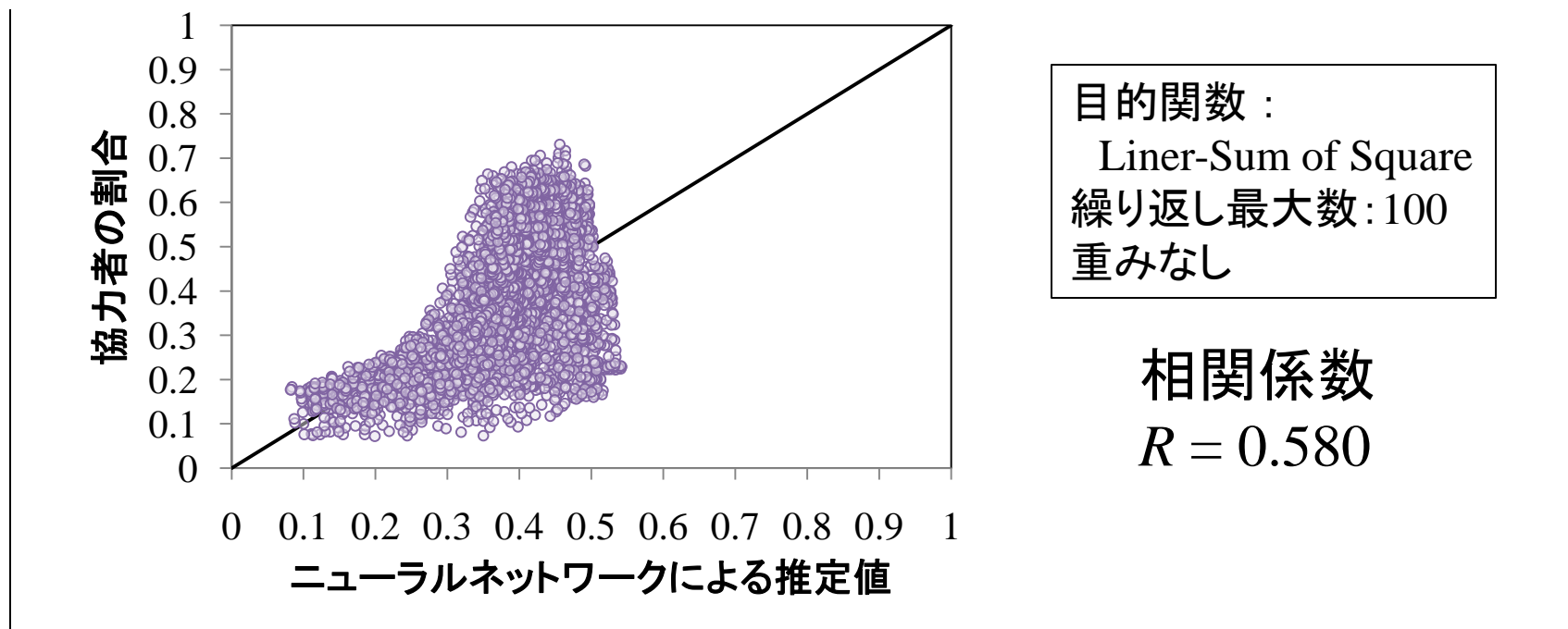
| モデル作成に用いた設計変数 | | | | 相関係数 |
|---------------|------|------|-----------|-------|
| ベキ乗指数 | 最小次数 | 最大次数 | 最大次数のノード数 | |
| ○ | ○ | ○ | ○ | 0.879 |
| | ○ | ○ | ○ | 0.478 |
| ○ | | ○ | ○ | 0.831 |
| ○ | ○ | | ○ | 0.842 |
| ○ | ○ | ○ | | 0.932 |

- ベキ乗指数が協力者の割合の予測に最も強い影響を与えている
- 次に影響を与えているのが最小次数
- 最大次数のノード数は逆に予測精度を下げている

数値実験結果の検証

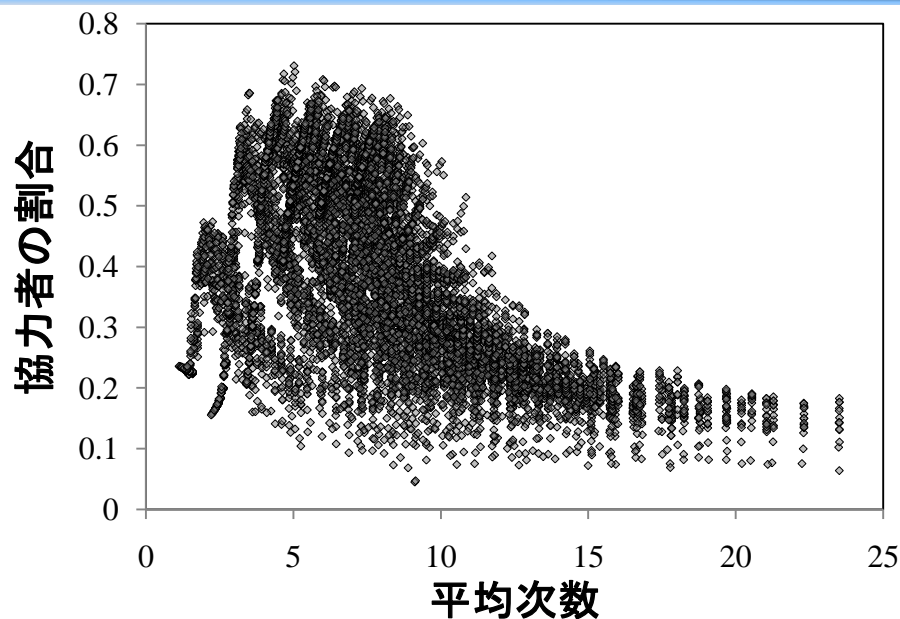
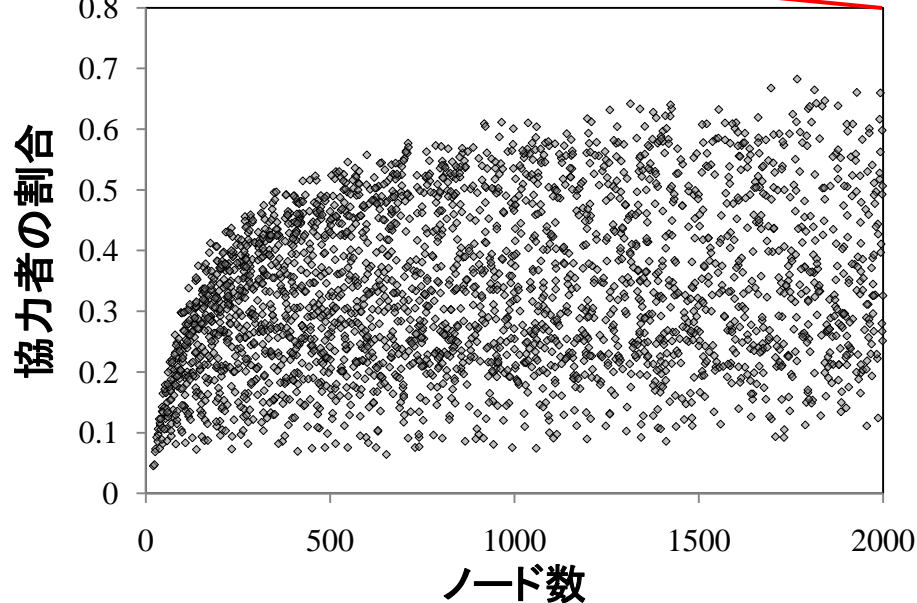
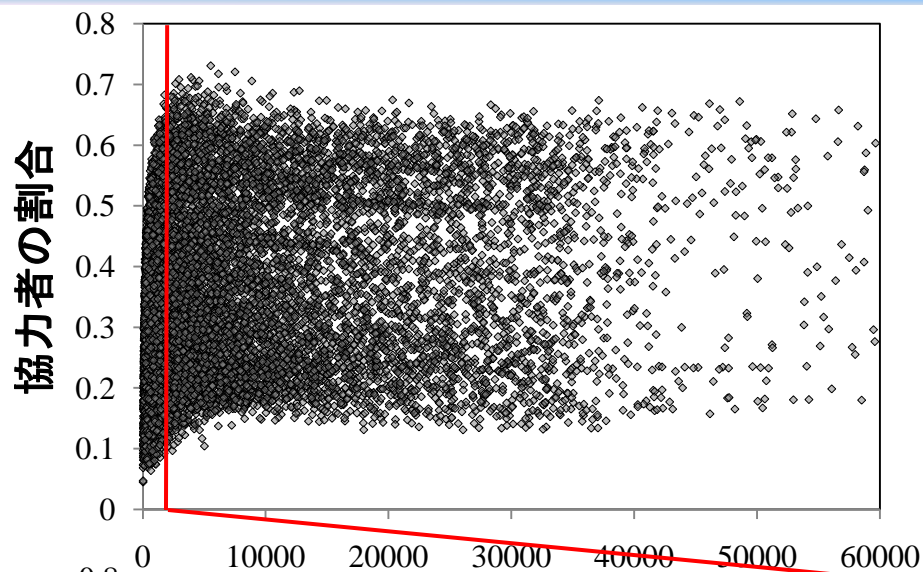
4. 説明変数2個が協力者の割合に“非線形”に影響を与えていると仮定する

6個のユニットの中間層のモデルによる推定値と実験値の比較



分類分析では2個の説明変数を与える影響は不明確

説明変数の可視化

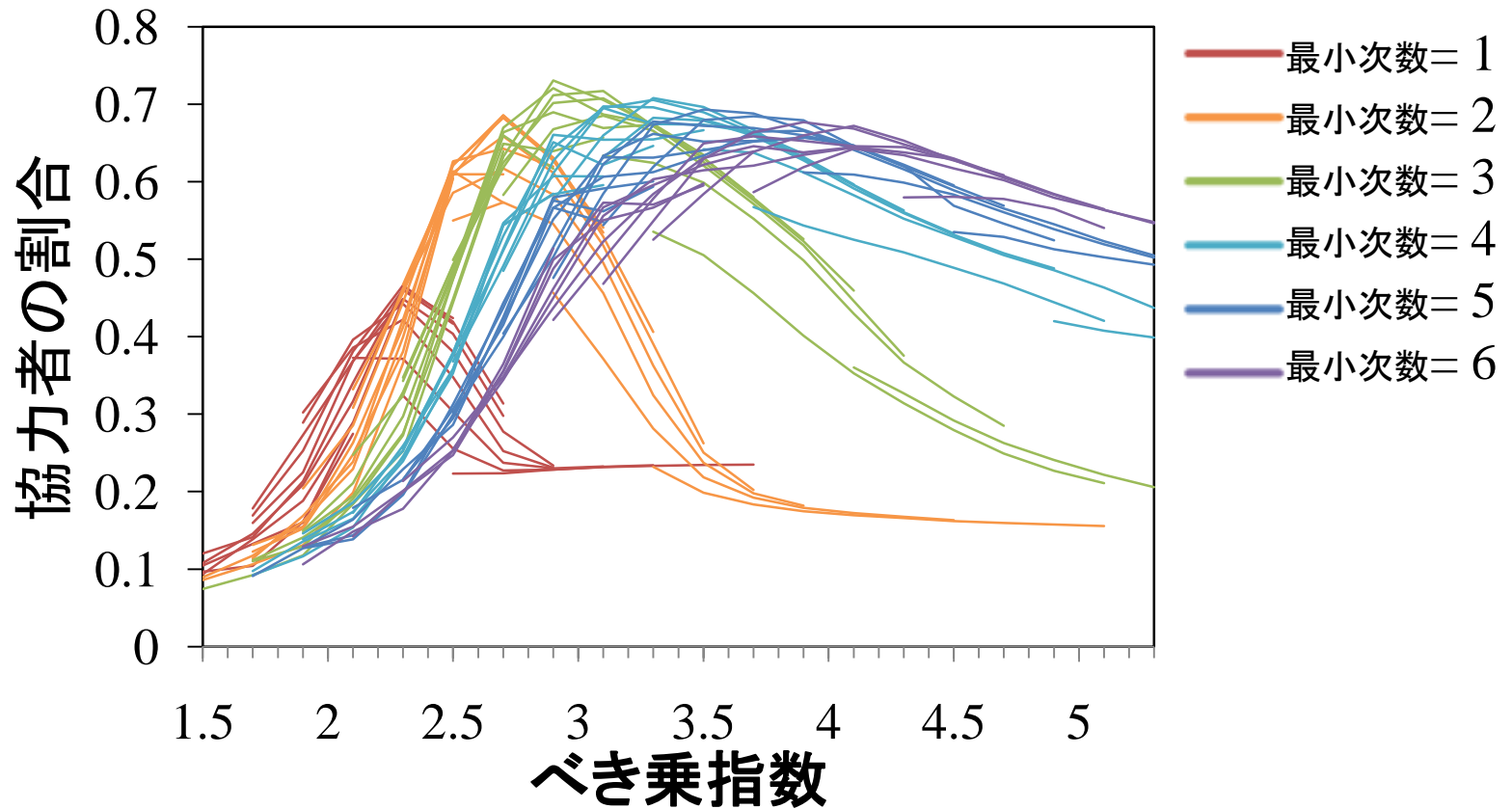


ノード数が小さすぎると
協力者の割合が減少する
ことがわかる

べき乗指数の影響の可視化

- ニューラルネットワークにおいて、協力者の割合に最も強い影響を与えていたべき乗指数と協力者の割合の関係を可視化する
- P.25の結果より、最大次数のノード数は固定(1とする)
- ノード数は1000以上のパラメータセットのみを考える
- 最小次数によって色分けを行い、同じ最大次数を持つパラメータセットを同じ系列とした

ベキ乗指数の影響の可視化



ベキ乗指数によって協力者の割合にピークが生じている
ピークは最小次数によって異なる

まとめ

- 多量のシミュレーション実験の結果から、べき乗指数が協力者の割合に非線形な強い関係性を示すことを導いた
- 空間囚人のジレンマにおける相反する知見 (P.15参照)に対して、網羅的なパラメータスタディを行ったことで、べき乗指数(Heterogeneity)が異なることで、協力者の割合は大きく変化するという知見が得られた

結言

- 多変数を用いた大規模なマルチエージェントシミュレーションを行い、実験結果から有意な情報を得るといふ新しい方法論を提案した
- 線形の解では解決できない問題に対し、ニューラルネットワークを用いた分析から、非線形な関係を可視化することができた

参考文献

- Nowak, M. and May, R.: Evolutionary games and spatial chaos, *Nature*, Vol.359, pp.826-829 (1992)
- A-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, Vol. 286, pp. 509–512 (1999)
- Santos, F. and Pacheco, J.: Scale-Free Networks Provide a Unifying Framework for the Emergence of Cooperation, *Physical Review Letter*, Vol. 95, No. 098104 (2005)
- Fu, F., Liu, L. H., and Wang, L.: Evolutionary Prisoner's Dilemma on heterogeneous Newman-Watts small-world network, *European Physical Journal B*, Vol. 56, pp. 367–372 (2007)
- 豊田秀樹：データマイニング入門, 東京図書 (2008)

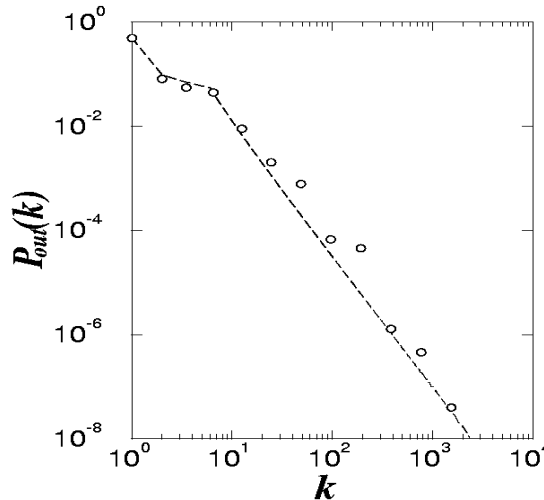


Appendix

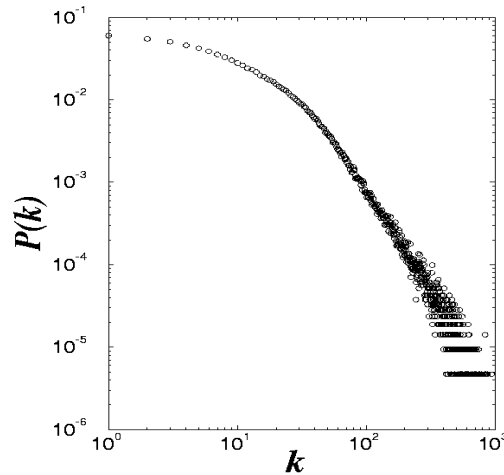
ネットワークについて

世の中には様々なネットワークが存在しているが、その中のいくつかはスケールフリー性を有していることがわかっている

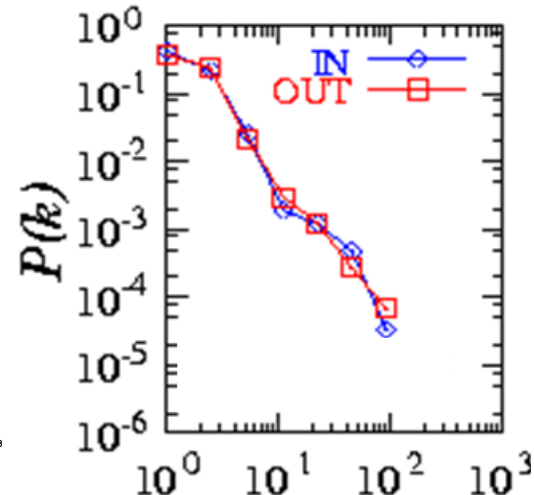
※スケールフリー性を有するネットワークとは、次数分布 $p(k)$ がべき乗則 $p(k) \sim k^{-\gamma}$ に従うネットワークのことである



WWW



Actor connectivity



Metabolic network

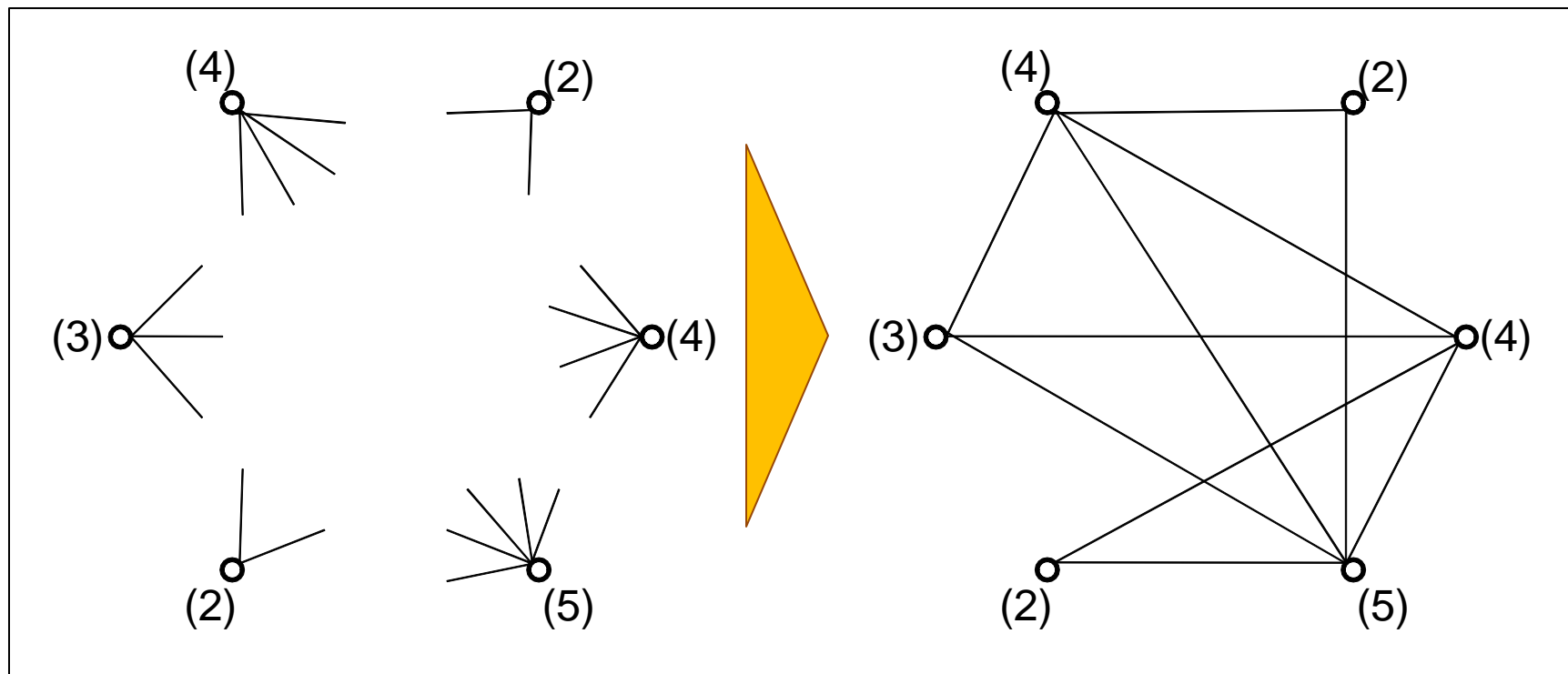
[Barabasi, 2004]

ネットワークの生成

コンフィギュレーションモデル

Newman(2003)により提唱された、任意の次数分布を持つネットワークを生成できるモデル

接続の仕方はランダムモデルに等しい



ネットワークの生成

べき乗則を持つ次数分布の生成

※スケールフリー性
 $p(k) \sim k^{-\gamma}$

スケールフリー性を有する
ネットワークの説明変数

ノード数: N

リンク数: M

平均次数: $\langle k \rangle$

べき乗指数: γ

最小次数: k_{min}

最大次数: k_{max}

最大次数のノード数: $n_{k_{max}}$

$$\begin{aligned} N &= n_{k_{max}} \int_{k_{min}}^{k_{max}} \frac{p(k)}{p(k_{max})} dk \\ &= \frac{n_{k_{max}} k_{max}^{\gamma}}{(-\gamma + 1)} (k_{max}^{-\gamma+1} - k_{min}^{-\gamma+1}) \end{aligned}$$

$$\begin{aligned} 2M &= n_{k_{max}} \int_{k_{min}}^{k_{max}} k \frac{p(k)}{p(k_{max})} dk \\ &= \frac{n_{k_{max}} k_{max}^{\gamma}}{(-\gamma + 2)} (k_{max}^{-\gamma+2} - k_{min}^{-\gamma+2}) \end{aligned}$$

ネットワークの生成

べき乗則を持つ次数分布の生成

平均次数 $\langle k \rangle$ ($=2M/N$)

$$\langle k \rangle = \frac{(-\gamma + 1) (k_{max}^{-\gamma+2} - k_{min}^{-\gamma+2})}{(-\gamma + 2) (k_{max}^{-\gamma+1} - k_{min}^{-\gamma+1})}$$

以上より

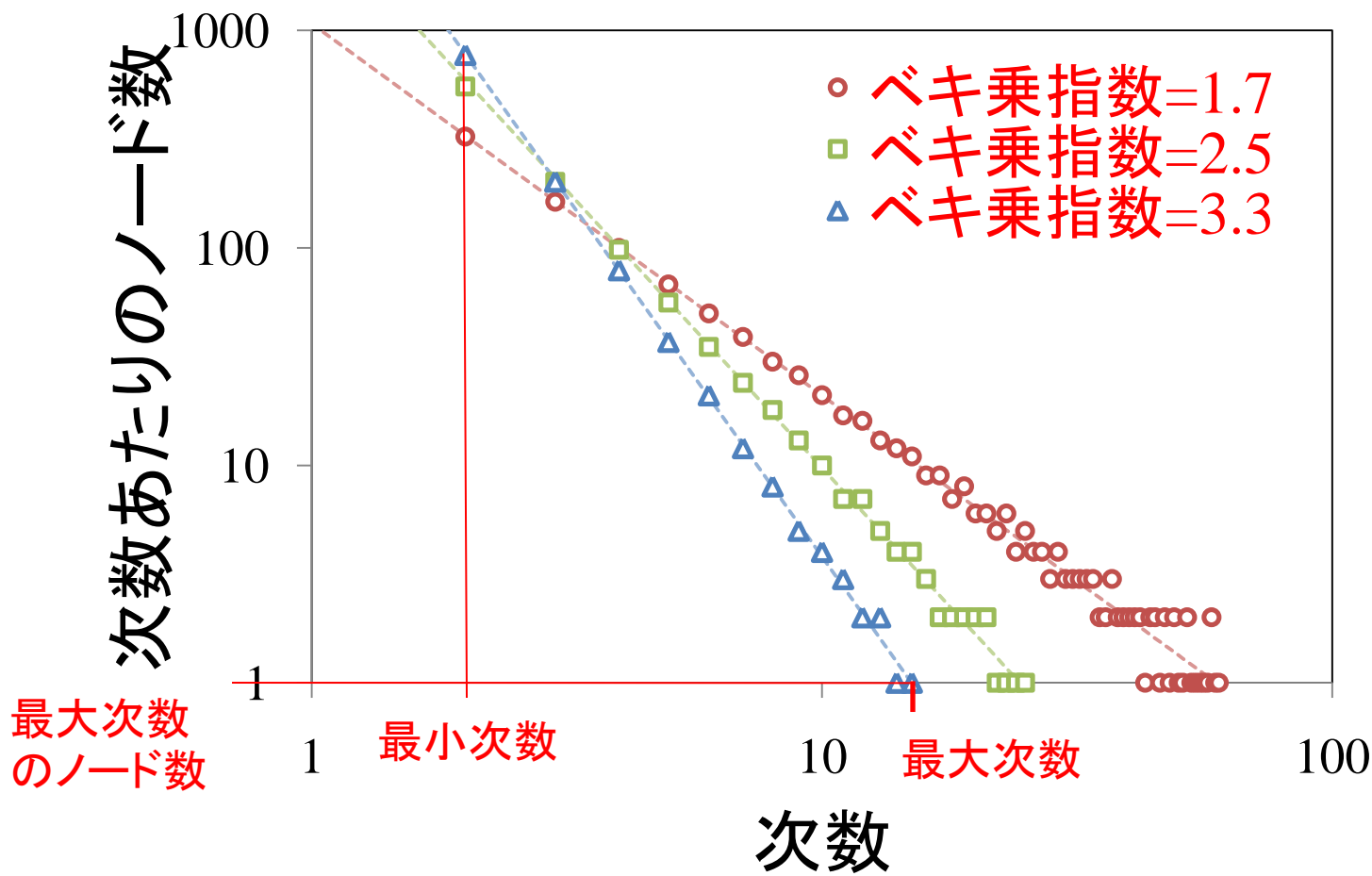
γ, k_{min}, k_{max} を決定 $\Rightarrow \langle k \rangle$ が求まる

$n_{k_{max}}$ を決定 $\Rightarrow N, M$ が求まる

$\gamma, k_{min}, k_{max}, n_{k_{max}}$ の4変数からネットワークを生成する
 N と $\langle k \rangle$ を説明変数とする

ネットワークの生成

スケールフリーネットワーク



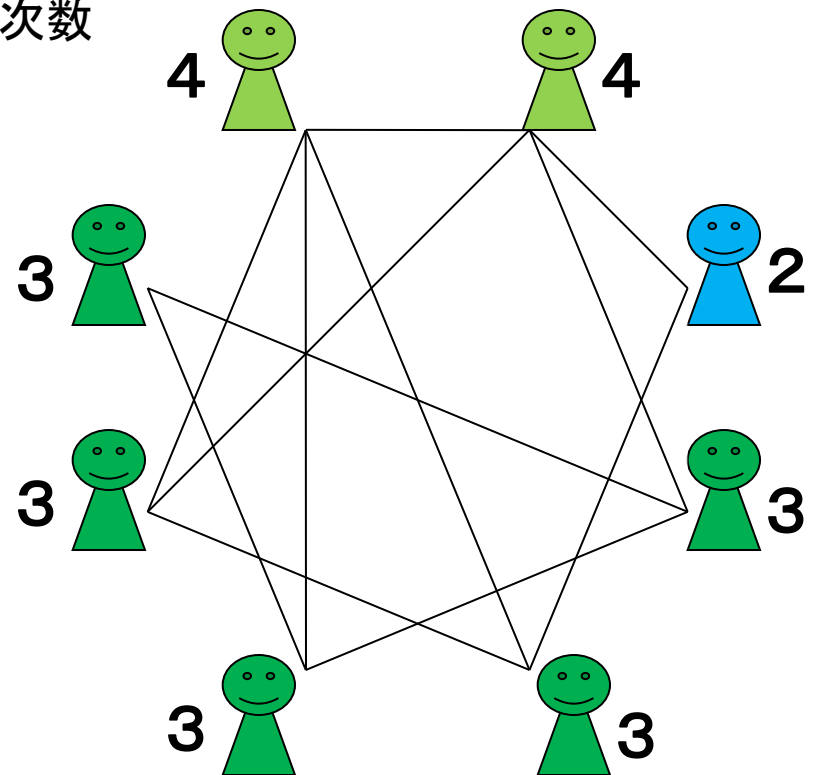
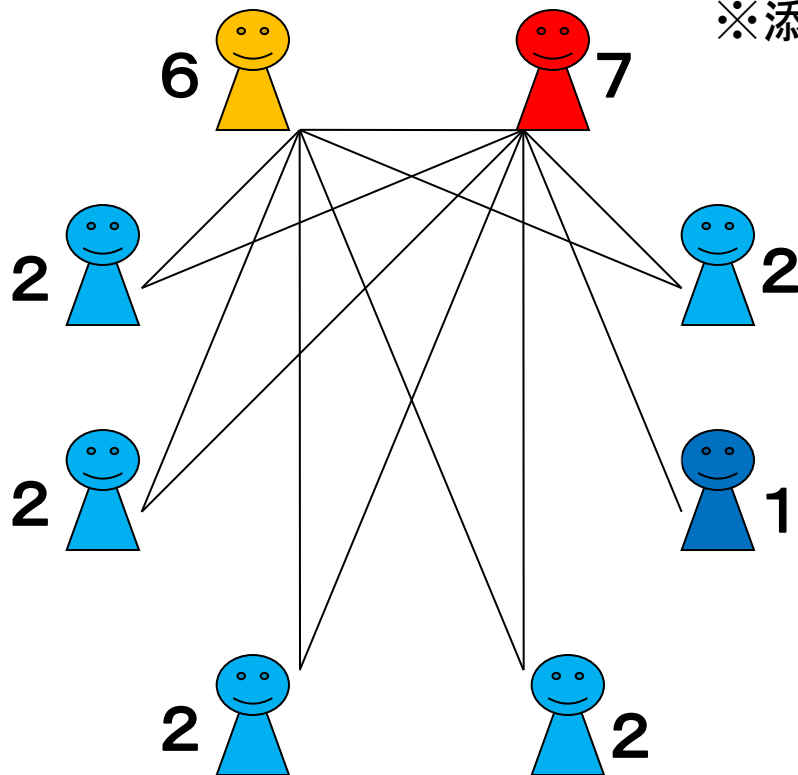
ベキ乗指数が小さいほど、Heterogeneityは大きくなる

べき乗指数による違い(イメージ)

小さい ← べき乗指数 → 大きい

大きい ← Heterogeneity → 小さい

※添え数字は次数



ニューラルネットワークの予備実験

設計変数4個を用いたモデルにおいて、中間層のユニット数を与える影響

| ユニット数 | 4 | 5 | 6 | 7 | 8 |
|-------|-------|-------|-------|-------|-------|
| 相関係数 | 0.843 | 0.839 | 0.879 | 0.861 | 0.842 |

ニューラルネットワークの予備実験

設計変数4個、説明変数2個から4パラメータより構築したモデルでの結果（中間層のユニット数は6）

| モデル作成に用いた設計変数 | | | | | | 相関係数 |
|---------------|------|------|-----------|------|------|-------|
| ベキ乗指数 | 最小次数 | 最大次数 | 最大次数のノード数 | ノード数 | 平均次数 | |
| | | ○ | ○ | ○ | ○ | 0.614 |
| | ○ | | ○ | ○ | ○ | 0.799 |
| | ○ | ○ | | ○ | ○ | 0.821 |
| | ○ | ○ | ○ | | ○ | 0.828 |
| | ○ | ○ | ○ | ○ | | 0.636 |
| ○ | | | ○ | ○ | ○ | 0.859 |
| ○ | | ○ | | ○ | ○ | 0.833 |
| ○ | | ○ | ○ | | ○ | 0.838 |
| ○ | | ○ | ○ | ○ | | 0.839 |
| ○ | ○ | | | ○ | ○ | 0.845 |
| ○ | ○ | | ○ | | ○ | 0.875 |
| ○ | ○ | | ○ | ○ | | 0.833 |
| ○ | ○ | ○ | | | ○ | 0.853 |
| ○ | ○ | ○ | | ○ | | 0.864 |
| ○ | ○ | ○ | ○ | | | 0.879 |