

京都パープルサンガ主催試合の観戦に 対する需要の予測

——神経回路網モデルと遺伝アルゴリズムを用いて——

鈴木 智 也*

目 次

- | | |
|------------------|--------------------|
| 1 はじめに | 4 パラメータ推定：遺伝アルゴリズム |
| 2 先行研究 | 5 推定結果 |
| 3 モデル構築：神経回路網モデル | 6 結 論 |

1 は じ め に

地域経済に対し、プロスポーツはしばしば大きな影響を及ぼす。ある地域をスポーツのチームが本拠地とすることで、その地域の経済が活性化することは珍しくない。その一方で、財政上の基盤が弱いチームに対して、自治体が金銭的援助を行うケースもある。このように、スポーツチームが存在する地域では、チームの財政面や人気面での浮沈と地元経済振興との間に深い関係が存在する。それゆえ、スポーツ観戦に対する需要は以前から経済学者の研究対象となってきた。

多くの先行研究では、観客数を被説明変数とする線形回帰モデルが用いられている。線形回帰モデルは、パラメータ推定値の計算を容易にするという利点を持つ。しかしながら、そのような実証研究では、観客数の変動を必ずしもうまく説明できていないのも実情である。線形回帰モデルを用いるということは、スタジアムの入場券購入に関する人々の意思決定を線形関数で近似するということである。本稿は、人々の意思決定の線形近似が観客数の変動をうまく説明できない原因になっているのではないかと考える。

線形近似は、喩えていうならば、地図を作る際に「地球は平坦である」と仮定することに似ている。例えば、伏見区といったように狭い地域の地図を作るならば、そのような仮定に問題はないうであろう。しかし、世界地図を作るとなると、その仮定は地図に大きなズレをもたらしてしまう。同様に、入場券購入に関する人々の意思決定を線形近似することは、パラメータ推定値の計算を容易にする一方で、モデルが人々の意思決定から乖離したものとなる可能性をもたらす。その乖離が無視できないほど大きければ、回帰モデルの説明力は弱くなる。そこで、本稿は、線形回帰モデルを非線形回帰モデルに代えて、京都パープルサンガ主催試合に対する観戦需要を推定する。

* 本稿に対して匿名のリフェリーより、有益なコメントを頂きました。ここに記して、感謝いたします。

非線形回帰モデルには様々な形態が存在するが、近年、注目を浴びているモデルの一つに“Neural Network Model”がある。これは「神経回路網モデル」と訳されたり、そのまま「ニューラルネットワークモデル」とカナ表記されたりする。このモデルは、名前が示唆するとおり、脳の神経回路網上で行われる情報処理を模倣している。本稿の文脈でいえば、潜在的な観客は、その時々を経済状況や気象条件、あるいは対戦相手など、様々な要因を考慮して、スタジアムへの入場券を購入するかどうかという情報処理を行う。このような人々の情報処理に倣ったモデルを利用すれば、観客数の変動をうまく説明できるのではないだろうか。そこで、本稿では、京都パープルサンガの主催試合に対する観戦需要を分析するために、神経回路網モデルを選択する。

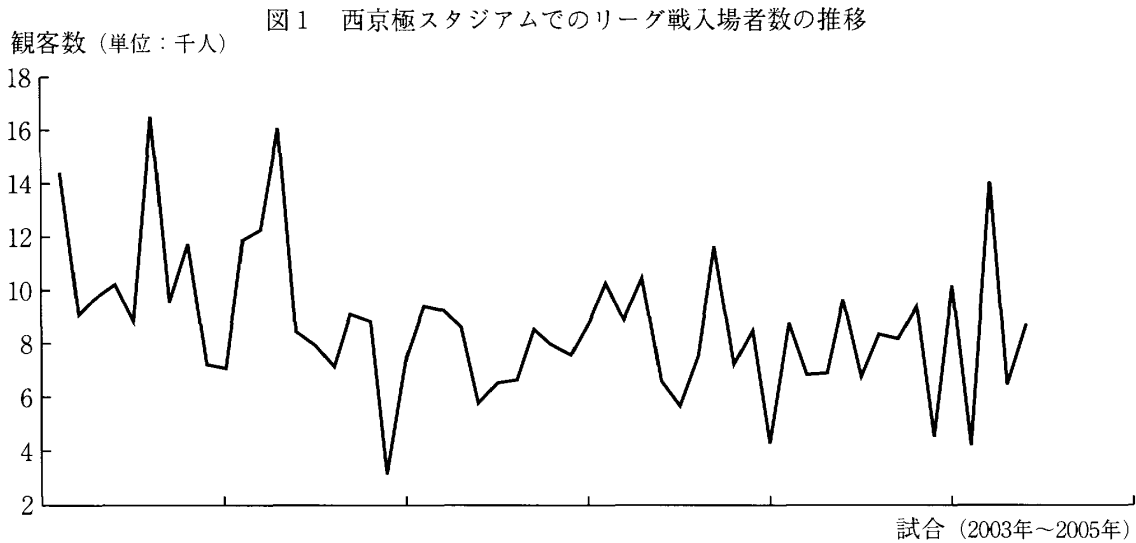
非線形回帰モデルを用いる際に問題となるのは、パラメータ推定値の計算である。非線形回帰モデルの場合、線形回帰モデルと異なり、最小二乗法によって推定値を求める公式が存在しない。後述するが、非線形回帰モデルでは、残差の二乗和を最小化するようにパラメータの値を求める過程で、局所解に陥ってしまう可能性がある。しかも、現段階では、どのような計算方法を使ったとしても、局所解が推定値となる可能性を完全には排除できない。とはいえ、数値演算の分野では、そのような局所解に陥る可能性を低くする解法の研究も進んでいる。その研究成果の一つとして、“Genetic Algorithm”と呼ばれる解法がある。これは「遺伝アルゴリズム」と訳されたり、GAと略されたりする。経済学の文献ではまだあまり使われていないが、本稿は遺伝アルゴリズムで神経回路網モデルを推定する。

2 先行研究

スポーツ観戦は、観光などと同様に、余暇活動の一つである。しかし、スポーツ観戦には一般的な観光と大きく異なる点がある。観光の場合、来場者数は規則的な季節変動を示すので、予測は比較的容易である¹⁾。それに対し、スポーツ観戦の場合、観客数が試合毎に大きく不規則に変動するので、その予測は難しくなる。これは京都パープルサンガの主催試合においても例外ではない。図1は、2003年から2005年までの三年間における西京極スタジアムでのリーグ戦54試合について、入場者数の推移を描いたものである。これら54試合のうち、観客数の最高値が16,614人(2003年8月2日 J1 第15節)、最低値が3,124人(2004年5月19日 J2 第13節)と、実にその差は一万三千人以上となっている。また、平均値が8,745人なのに対し、標準偏差は2,703である。これらの数値から明らかなように、京都パープルサンガが西京極スタジアムで主催するリーグ戦においても、観客数は試合毎に大きく不規則に変動している。

変数の変動が不規則でかつ大きいということは、予測の困難さを意味する。スポーツ観戦に対する需要の実証分析では、観客数を被説明変数とする線形回帰モデルが用いられることが多いが、残念ながら、それらの分析結果では決定係数が必ずしも高くない。海外の文献を例にとると、対

1) 京都を訪れる観光客数の予測に関する文献として、鈴木(2007)を参照のこと。



象とするスポーツや説明変数の選択などに違いはあるものの, Bhattachaya and Smyth (2003), Borland and Lye (1992), Falter and Pérignon (2000), Peel and Thomas (1988) や Wilson and Sim (1995) など, 自由度修正済み決定係数は0.5前後もしくはそれを下回っている。つまり, 線形回帰モデルを用いたスポーツ経済の実証研究の多くでは, 観客数の変動のうち半分程度の割合しか説明できていないのが現状である。

京都パープルサンガを対象とした鈴木 (2006) でも, 同様に, 観客数を被説明変数とする線形回帰モデルが推定されている²⁾。合計九通りのモデルが推定され, 最終的に選ばれたモデルは次のようなものであった。

$$\begin{aligned} \log attend_t = & \beta_0 + \beta_1 \log price_t + \beta_2 \log output_t + \beta_3 tv_t \\ & + \beta_4 \log opponent_t + \beta_5 rain_t + \beta_6 weekday_t + \beta_7 kansai_t \\ & + \beta_8 special_t + \beta_9 goldweek + u_t. \end{aligned} \quad (1)$$

なお, u は攪乱項であり, 通常の仮定を満たすものとする。左辺の $attend$ はリーグ戦各試合の観客数である。右辺の $price$ は入場券の実質価格, $output$ は実質所得, tv はテレビ中継のダミー変数, $opponent$ は対戦相手の集客力, $rain$ は雨のダミー変数, $weekend$ は平日開催試合のダミー変数, $kansai$ は在阪チームとの対戦のダミー変数, $special$ は下部リーグへの降格や優勝がかかった特別な試合のダミー変数, $goldweek$ はゴールデンウィーク期間中 (ただし連休最終日は除く) の試合のダミー変数である。それぞれの変数に対してどのデータを用いたかについては, 付論の表2にまとめられている。2003年から2005年のリーグ戦をサンプルとして, (1)式の線形回帰モデルを最小二乗法で推定したところ, 自由度修正済み決定係数は0.565であった。これは, 先行研究と同様に, 観客数の変動のうち半分近くをモデルで説明できていないということである。

先行研究において, 観客数の変動をうまく説明できていない理由として, 重要な説明変数が含

2) 本来, 需要分析には供給面も考慮した同時方程式モデルを用いるべきである。しかしながら, スタジアムが満員にならない限り, 入場者数が需要量を測ると解釈できる。京都パープルサンガの場合も, サンプルにおいて, 入場者数がスタジアムの収容人数に達した試合はない。

まれていない可能性は残る。しかしながら、スポーツ経済の先行研究では、データの利用可能性と自由度が許す限り、既に異なる多くの変数が説明変数として採択されている。したがって、これ以上、説明変数をあれこれ試して増やしていくことは生産的でないであろう。本稿はむしろモデルの線形性に問題があるのではないかと考える。次節では、線形回帰モデルに代わる非線形回帰モデルが構築される。

3 モデル構築：神経回路網モデル

線形回帰モデルを用いて京都パープルサンガ主催試合の観客数を説明した先行研究として、上述の鈴木（2006）がある。それに対し、本稿は線形回帰モデルを非線形回帰モデルに代えて、モデルの予測精度が向上するの否かを検討する。そのため、説明変数は（1）式のものとまったく同じものを用いることにする。変数とデータの出典は、付論の表2にまとめられている通りである。被説明変数は観客数であり、定数項と九個の説明変数から成る非線形回帰モデルで、京都パープルサンガ主催試合の観客数変動を説明する。

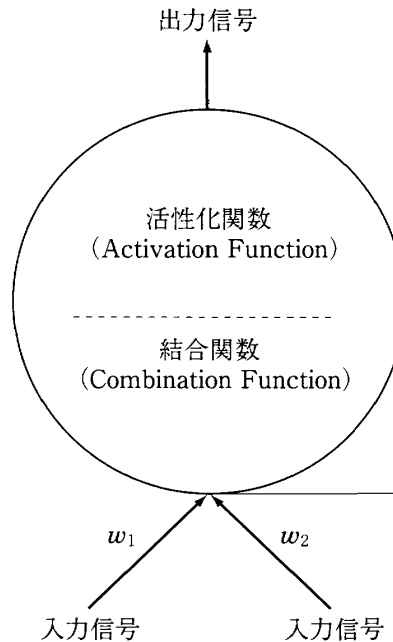
近年、経済学を含む幅広い分野で注目されている非線形モデルの一つに、神経回路網モデルがある³⁾。神経回路網とは、社会科学では耳慣れないが、自然科学では生物の高等学校用教科書に載っているような広く知られた言葉である。人間の脳には「ニューロン」と呼ばれる莫大な数の神経細胞が存在し、ニューロン同士の結びつきによって形成されるのが神経回路網である。人間の脳は、神経回路網上のニューロン間で互いに情報をやり取りし、高度な情報処理を実現している。端的に言えば、神経回路網上で行われる情報処理機能に倣ったモデルが、神経回路網モデルと総称されるのである。

では、人間の脳において、どのように情報処理が行われるのであろうか。上述のように、情報処理の基本単位はニューロンである。目や耳などの刺激受容器が外部からの刺激を感知した場合、その刺激はニューロンへ信号として伝えられる。刺激が一定の強さ（「閾刺激」）以上の場合、興奮が引き起こされ、ニューロンは「発火」する。その興奮はさらに隣接するニューロンへ伝えられる。つまり、各ニューロンは刺激の強さが閾値に達しているか否かで、その情報を次のニューロンへ伝えるか否かを選択する。また、それぞれのニューロンは複数のニューロンから受けた情報を統合して次のニューロンへ伝える。このようにして、知覚が統合され、情報処理が行われるのである。

神経回路網上での情報処理をモデル化する際に、注意したいポイントが二点ある。第一に、ニューロンの発火は閾刺激以上の強さの刺激を要するが、刺激が閾刺激より強すぎてもニューロンは発火しないという点である。このような規則性は「全か無かの法則」と呼ばれる。第二に、

3) 神経回路網モデルは1970年代頃から主に自然科学の分野で研究されてきた。たとえば、Rumelhart and McClelland (1986) を参照のこと。経済学において、比較的早い段階で神経回路網モデルの有用性を唱えた代表例は、Sargent (1993, 1999) である。

図2 個別ニューロンの情報処理モデル



ニューロンからニューロンへの興奮の伝達は一方向にしか起こらないという点である。これら二点を考慮して、神経回路網モデルが定式化される。

Kendrick *et al.* (2006) によれば、神経回路網モデルの定式化には三つの段階がある。第一段階は、個別ニューロンにおける情報処理の定式化である。個別ニューロンの情報処理は、図2のように二つのパートから成る。それは別のニューロンから受けた刺激を一つに結合することから始まる。これを数式で表現したものが「結合関数 (“Combination Function”)」である。通常、結合関数は刺激信号の加重和として定義される。第 i 番目のニューロンへの刺激が q_i 個あるとし、そのニューロンで結合された刺激信号を z_i で表すと、

$$z_i = \omega_0 + \sum_{i=1}^{q_i} \omega_i x_i \quad (2)$$

となる。本稿の場合、鈴木 (2006) の線形回帰モデルが比較対象なので、 x は(1)式の右辺にある九個の説明変数である。次に、ニューロンは結合された信号 z_i を圧縮し、それが閾刺激を超える場合に、別のニューロンへ興奮を伝える。この働きを表現した数式は「活性化関数 (“Activation Function”)」あるいは、結合信号を圧縮することから “Squasher” とも呼ばれる。第 i 番目のニューロンが次のニューロンへ伝える信号を a_i で表すと、

$$a_i = S(z_i), \quad (3)$$

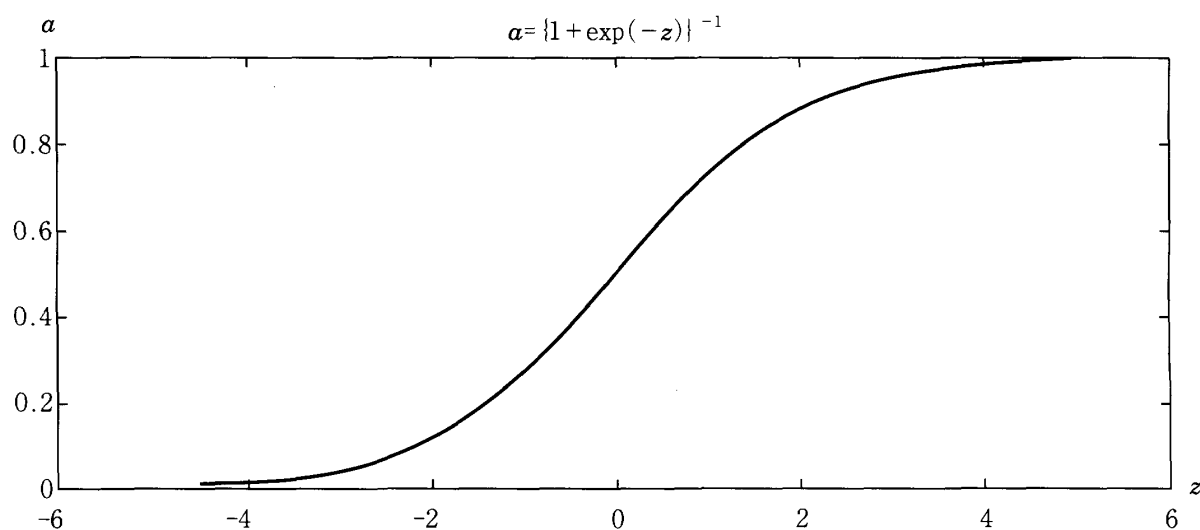
となる。ここで、 S は非線形の活性化関数である。もっとも代表的な活性化関数は次のようなロジスティック関数である⁴⁾。

$$S(z) = \{1 + \exp(-z)\}^{-1} \quad (4)$$

ロジスティック関数は、図3のように、0 から 1 の間に情報を圧縮する。この単調非減少関数は、

4) 活性化関数として用いられる関数については、たとえば、Campbell *et al.* (1997) の第12章、Judd (1998) の第6章、Kendrick *et al.* (2006) の第2章、McNelis (2005) の第2章などを参照のこと。

図3 代表的な活性化関数であるロジスティック関数



$z=0$ の周辺で大きく値が変わるので、上述の「全か無かの法則」をうまく近似しているといえる。本稿では、ロジスティック関数を活性化関数として採択する。

神経回路網モデルの定式化における第二段階は、ニューロン同士の結合をモデル化することである。上述のように、ニューロンからニューロンへの情報伝達は一方向である。この点を踏まえ、情報は前方へ伝播されると仮定する。代表的なモデルは、図4のような、階層型ネットワークである。図中のネットワークは入力層、中間層、出力層の計三層から構成されている⁵⁾。入力層では、一つのニューロンが種類の刺激を受け、それを中間層へ伝達する。中間層のニューロンは、入力層からの信号を加重和して圧縮し、出力層へ伝える。出力層ニューロンは、中間層から受け取った信号を結合する。中間層のニューロン数を q とし、出力層のニューロン数を 1 とすると、出力層の結合は、一般的に、次のような線形関数で表現される。

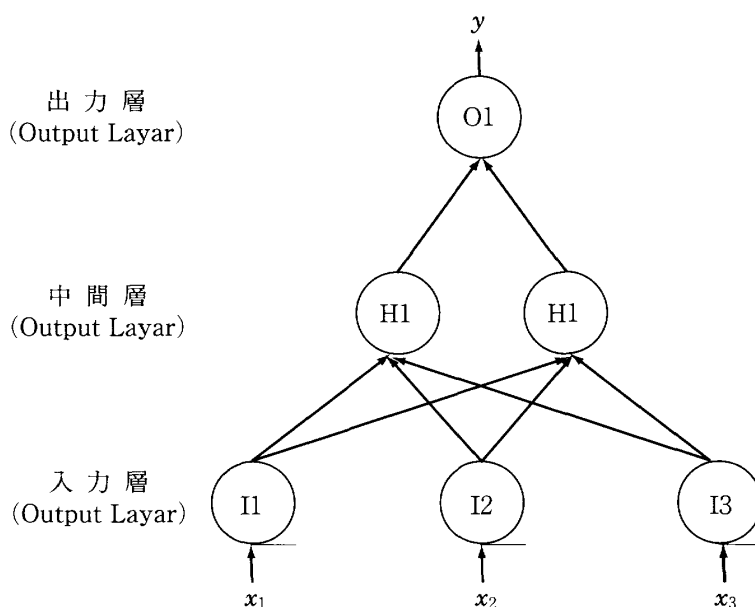
$$y = \theta_0 + \sum_{i=1}^q \theta_i a_i. \quad (5)$$

本稿の場合、鈴木 (2006) の線形回帰モデルが比較対象であるので、左辺の y は (1) 式の左辺と同じになる。中間層ニューロンの個数については、恣意的ではあるが、 $q=2$ とする。本稿のモデルには定数項と九個の説明変数があるので、中間層ニューロンを一個増やすだけで、推定すべきパラメータが 11 個加わり、自由度が 11 も減ってしまう。サンプル数が 54 なので、中間層ニューロンの個数を大きく増やすわけにはいかないのである。そこで、本稿は、中間ニューロン層の個数を二個に抑え、三層構造の前方伝播型ネットワークを採択する。この場合、モデルは (2) 式から (5) 式で記述される。

神経回路網モデルの定式化における第三段階は、パラメータ推定の基準を決めることである。本稿の場合、推定されるパラメータは (2) 式の ω と (5) 式の θ である。(2) 式と (5) 式を個別

5) 図4はあくまで一例である。たとえば、中間層は何個あっても、逆に無くてもよい。また、各層のニューロン数も図の通り限定されるわけではない。さらに、前方伝播型以外のモデルも存在し、それらについては、Campbell *et al.* (1997) の第12章や McNelis (2005) の第2章を参照のこと。

図4 階層型ニューラルネットの一例



にみると、それぞれ線形回帰モデルにみえるが、実際にはそうではない。たとえば、(2)式では左辺の変数 z_i が観察不可能であり、(5)式では右辺の変数 a_i が観察不可能である。そして、その z_i と a_i は(3)式によって結び付けられる。したがって、それぞれの式のパラメータを別々に推定するのではなく、すべての式のパラメータを一括して推定しなければならない。そのパラメータ推定には、何らかの基準を置く必要がある。代表的な基準は残差の二乗和を最小化することである。モデルから予測される被説明変数 y の値を $\hat{y}$ 、パラメータ推定値のベクトルを $\hat{\theta}$ と $\hat{\omega}$ で表すと、

$$\min_{\hat{\theta}, \hat{\omega}} \sum (y - \hat{y})^2 \quad (6)$$

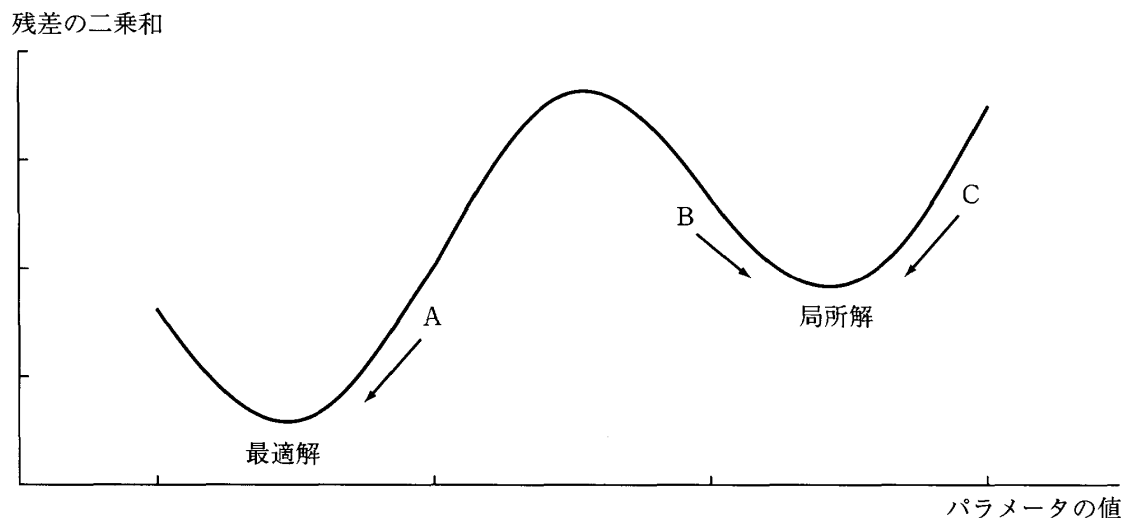
が推定値を求める基準となる。本稿は、残差二乗和の最小化を基準として、パラメータを推定する。

4 パラメータ推定：遺伝アルゴリズム

残差二乗和の最小化をパラメータ推定の規準とするにしても、非線形回帰モデルの場合、線形回帰モデルとは異なり、推定値を求める公式が存在しない。そこで、ニュートン法などの数値演算に頼ることになる。その方法に関して、図5を用いて、ごく大雑把に説明しておく⁶⁾。パラメータが一つしかない事例が図に描かれているものの、パラメータ数を増やしても推定の手順は基本的に同じである。数値演算による解法では、まず適当にパラメータの値を与える。その初期点がA点であったとする。次に、A点から残差の二乗和を小さくする方向を探索する。その方向が見つかったら、パラメータの値を少し変えて、A点から移る。残差二乗和がある一定以上

6) 数値演算によるパラメータ推定の詳細については、Davidson and MacKinnon (2004) の第6章など、中級以上の計量経済学に関する書籍を参照のこと。

図5 非線形モデルにおける残差二乗和とパラメータの値



には小さくなくなるまで、探索を反復する。反復探索の結果、パラメータ推定値が得られる。注意したいのは、A点から探索を始めると最適解に辿り着くが、B点やC点からでは局所解に辿り着いてしまうことである。このように、非線形回帰モデルの場合、初期点の選び方次第で推定結果が異なってしまう。

上述のような局所解の存在は、図5が与える印象よりも、はるかに厄介な問題となる。第一に、図5と異なり、モデルによっては幾つもの局所解が存在し得る。局所解の数が多いほど、無作為に初期点を選んだ結果として、局所解に陥る可能性が高くなる。第二に、最適解に関して、我々はその位置を知らない。もしも最適解の位置があらかじめ判明しているならば、その近傍で初期点を選ぶことで、我々は局所解に陥る可能性を低くできるであろう。しかしながら、我々は最適解の位置を知らないため、初期点を無作為に選ばざるを得ず、その結果、局所解に陥る可能性が残ってしまう。このように、非線形回帰モデルのパラメータ推定には、局所解の問題がつきまとうことになる。局所解に陥る可能性を低くするには、できるだけ多くの初期値を試すしかない。たとえば、100組のパラメータ初期値を試せば、最大で100組のパラメータ推定値が得られる。そこで、それらの推定値の中から、もっとも残差の二乗和を小さくするものを最適解として選ぶのである。理論的には、ありとあらゆる初期値を試せば、局所解に陥ることはなくなる。しかしながら、現実には、初期値を無限に試すことはできない。さらに、図5の場合とは異なって、モデルが多数のパラメータを持つならば、十分に多様な初期値を試すことは困難になる。

パラメータが多数存在する場合に効率良く解を探索する方法として、遺伝アルゴリズムが開発されてきた。このアルゴリズムは、その名が示すように、生物の遺伝と自然淘汰に倣ったものである。ニュートン法など従来の解法が一組の初期値から探索を開始するのに対し、遺伝アルゴリズムは複数組の初期値から探索を開始する。それぞれのパラメータ初期値は、親世代の染色体や遺伝子に見立てられ、後述の「選択 (“selection”)」「交叉 (“crossover”)」「変異 (“mutation”)」といった操作を加えられる。結果として得られるパラメータの値は、子世代の染色体や遺伝子に

見立てられる。このような操作を何世代も繰り返した後、本稿の文脈でいえば、残差の二乗和を最小化するようなパラメータの組み合わせを推定値として選ぶ。以上が基本的な遺伝アルゴリズムの手順である。本稿は、Kendrick *et al.* (2006) の第11章と第12章、および McNelis (2005) の第3章を参考に、以下のような遺伝アルゴリズムでパラメータ推定を行う。

遺伝アルゴリズムの第一歩は、複数の組み合わせのパラメータ初期値を用意することである。本稿の文脈でいえば、 ω や θ の初期値を複数通り用意するということである。それぞれの組み合わせは個人の染色体に喩えられる。九個の説明変数と二個の中間層ニューロンから成る本稿のモデルについていえば、各人の染色体は次のようにベクトル表示される。

$$\Omega(n) = [\omega_{0,1}(n) \cdots \omega_{9,1}(n) \omega_{0,2}(n) \cdots \omega_{9,2}(n) \theta_0(n) \theta_1(n) \theta_2(n)] \quad \text{for } n=1, \dots, N.$$

なお、染色体 Ω の各要素は遺伝子に相当する。遺伝子の初期値の生成にあたって、本稿はパソコンで乱数を近似的に生成する⁷⁾。その際、乱数発生に対して確率分布を指定しなければならない。本稿の場合、恣意的ではあるが、平均が0で分散が30の独立な正規分布に従って、各遺伝子の値が実現すると仮定した。また、本稿では $N=200$ とし、200人分の染色体を用意する。この200人を親世代と呼ぶ。

次のステップは選択である。生物の世界では、強い個体ほど、子を残す確率が高い。これを反映して、親世代の中から二人の個体を親として選び出す。ここでは、親世代の各人のパラメータを元に、残差の二乗和をそれぞれ計算する。残差の二乗和が小さい個体ほど、優れていると考える。ごく簡単な選択方法としては、残差の二乗和がもっとも小さくなる個体と次に小さくなる個体を親として選び出すことが考えられる。本稿では、それよりもランダム性の高い選択方法を採用する。残差の二乗和が小さい個体ほど選ばれやすくするため、親として選ばれる確率を残差の二乗和の逆数に比例して設定する。こうして、二人の個体を200人の親世代の中から選択する。

生物が子を産むとき、親の染色体の一部は、配偶者の染色体の対応する部分と交換される。これは交叉と呼ばれる。本稿では、染色体の一点においてのみ交叉が起こると単純化する。また、本稿では、染色体上で交叉の起こる点を無作為に選ぶ。親として選ばれた個体を l と m で表そう。本稿のモデルの場合、推定すべきパラメータが全部で23個なので、各個体は23個の遺伝子を持つ。ここで、1から22までの数字を一様分布から選び出す。仮に選ばれた数字が21だったとすると、21番目の遺伝子と22番目の遺伝子を境に二人の遺伝子が交換され、次のような一組の染色体が新たに生成される。

$$\Omega(l)' = [\omega_{0,1} \cdots \omega_{9,1}(l) \omega_{0,2}(l) \cdots \omega_{9,2}(l) \theta_0(l) \theta_1(m) \theta_2(m)]$$

$$\Omega(m)' = [\omega_{0,1} \cdots \omega_{9,1}(m) \omega_{0,2}(m) \cdots \omega_{9,2}(m) \theta_0(m) \theta_1(l) \theta_2(l)]$$

これらの染色体は親世代の遺伝子からのみ構成されていることに注意されたい。

生物の世界において、子の形質は親の形質を正確に引き継ぐわけではない。一部の遺伝子は変異を起こす。これを反映して、新たに生成された一組の遺伝子の一部をランダムに変異させ、子

7) 本稿の分析では、遺伝アルゴリズムのためのコードを、MATLAB で独自に書いた。参考にしたのは、Kendrick *et al.* (2006) の第11章および第12章で公開されているコードである。

世代の染色体を生成する。変異はランダムに起こるので、計算上、変異の起こる可能性を指定する必要がある。本稿では、変異の起こる可能性を一様分布から選び出す。また、染色体上のどの遺伝子に変異が起こるかは事前にわからないので、変異の起こる遺伝子もランダムに選ぶことにする。さらに、変異の大きさもランダムに選ばれるとする。選択された親から交叉によって生成された一組の染色体に対して、このような変異の操作を施し、200人分の染色体を生成する。これが子世代の染色体となる。

200人から成る子世代の染色体を元に、上記の手順によって、200人分の孫世代の染色体を生成する。さらに、孫世代の染色体から、同様に、200人分のひ孫世代の染色体を生成する。これを繰り返して、各世代200人で100世代の染色体を生成する。その際、各世代において、残差の二乗和を最小化する染色体を探し出しておく。最後に、それらの染色体の中でも、全世代を通じて残差の二乗和を最小化する染色体を選ぶ。これが本稿で行う遺伝アルゴリズムである。さらに、本稿は、最終的に選ばれた染色体のパラメータ値を初期値として従来の数値演算解法を用い、パラメータを推定する⁸⁾。

本稿は以上のような遺伝アルゴリズムで神経回路網モデルを推定するが、その際に注意すべき点がある。それは、本稿の変数が各々単位の異なるデータで測られているということである。本稿は、McNelis (2005) の推奨に従い、次の式によって各変数のデータを0から1に圧縮する。

$$x_{i,t}^* = \frac{x_{i,t} - \min(x_i)}{\max(x_i) - \min(x_i)}, \quad \text{for } i=1, \dots, 5. \quad (7)$$

ここで添え字 t は観測値の番号を表し、 $\min(x_i)$ と $\max(x_i)$ はそれぞれ変数 x_i の最小値と最大値を表す。(7)式は、変数 x_i の観測値を0から1の間に圧縮するもので、入力層ニューロンの働きを記述していると解釈してよい。なお、この式による変形は、説明変数のデータだけでなく、被説明変数のデータにも適用される。

5 推定結果

京都パープルサンガの観客数変動を説明するために、第3節で構築した神経回路網モデルを用いて、第4節で説明した遺伝アルゴリズムで、パラメータ推定を行った。残念ながら、Brooks (2005) の第12章などで指摘されているように、神経回路網モデルにおいては、個々のパラメータが経済学的な意味を持たない⁹⁾。したがって、推定結果を式で表示しても仕方がない。神経回路

8) 遺伝アルゴリズムよりも複雑な推定方法として、Granger and Jeon (2003) の“Thick Modeling”がある。遺伝アルゴリズムや従来の方法などで、パラメータを何度も推定し、それらの平均を取るという方法である。本研究では、初期原稿の段階でこの方法を試したが、結果はおもわしくなかった。うまくいかない原因が、用いたコードに誤りがあるからなのか、この手法が本稿のモデルにそぐわないからなのか、現段階では不明である。Thick Modeling による推定は今後の課題としたい。

9) 神経回路網モデルは“Black Box”と批判されることが多い。したがって、経済学の文献では、神経回路網モデルの主たる用途は予測に限られている。たとえば、ファイナンスでは Franses and van Dijk (2000) や McNelis (2005)、マーケティングでは Smith and Gupta (2002) などが予測のために神経回

図6 線形回帰モデルと神経回路網モデルの事後的予測力

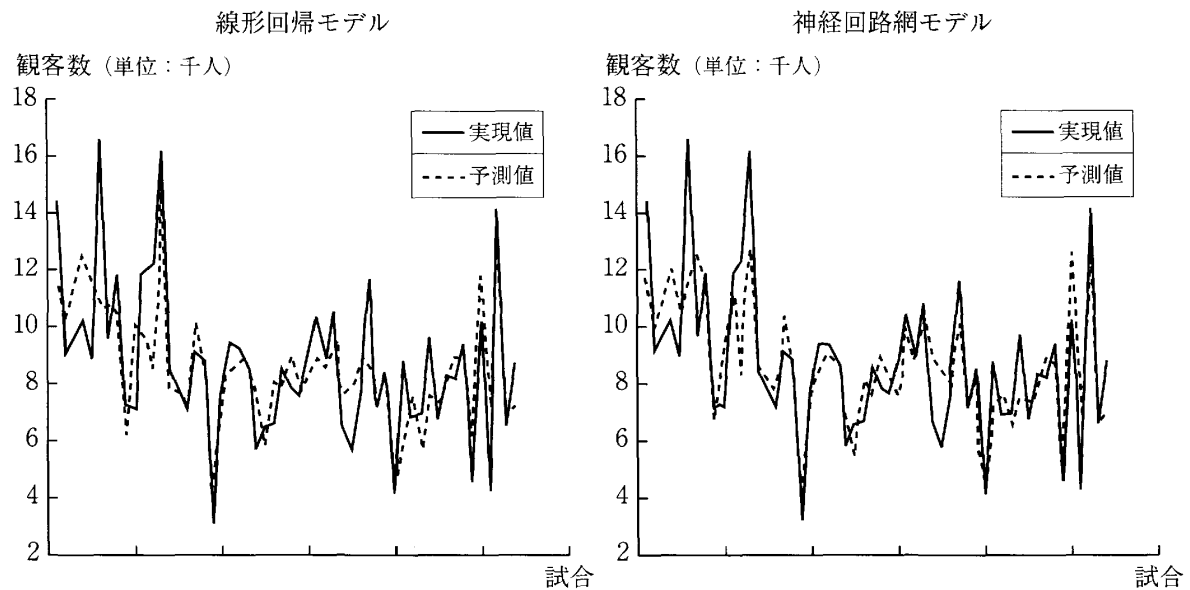


表1 モデルの事後的な説明力の比較

モデル	決定係数	HQ 情報量
線形回帰モデル	0.616	15.461
神経回路網モデル	0.618	15.338

注：HQ 情報量は小さいほど望ましい。

網モデルを用いた文献では、サンプル期間内での事後的な予測を表示するのが通例である。図6には、鈴木(2006)で用いられた線形回帰モデルと予測力を比較する目的で、神経回路網モデルによる事後的な予測が描かれている。なお、線形回帰モデルの場合、(1)式のように被説明変数は観客数の対数値であったが、指数関数を用いて、単位を人数に直してある。同様に、神経回路網モデルにおいては、まず(7)式から対数の値を逆算し、さらに指数関数を用いて単位を人数に戻してある。

図6を一瞥しただけでは、観客数の変動に関して、どちらのモデルがより高い説明力を持つのか明らかでない。そこで、モデル間の説明力の差を統計学的に検証する必要がある。本稿の場合、二つのモデルで優劣を競おうとすると、仮説が非入れ子型になるので、尤度比検定といった標準的な検定方法は使えない。幸い、二つのモデルは共通の被説明変数と説明変数から成るので、決定係数でモデルの説明力を比較する。なお、決定係数を算出するにあたって、どちらのモデルにおいても、被説明変数の単位は人数に戻してある。表1にまとめられているように、決定係数を比較すると、神経回路網モデルの方が線形回帰モデルよりも高いという結果になった。したがって、京都パープルサンガの観客数変動に関して、神経回路網モデルの方が線形回帰モデルよりも高い説明力を持つといえる。

決定係数の比較では神経回路網モデルに軍配が上がったが、他の基準ではどうであろうか。二つのモデルで説明変数は共通であるが、推定すべきパラメータ数は神経回路網モデルの方が多い。

、路網モデルを用いている。

パラメータ数の増加は自由度の減少を意味する。そこで、自由度の減少分にあわせて、パラメータ数の増加を罰するような情報量基準を使う。例えば、Hannan and Quinn (1979) は次のような情報量による基準（以下「HQ 情報量基準」）を提唱している。

$$HQ = \log \left\{ \frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2 \right\} + \frac{k}{T} \log \{ \log(T) \}. \quad (8)$$

ここで、 y と $\hat{y}$ はそれぞれ被説明変数の実現値と予測値であり、 k は推定すべきパラメータ数、 T は推定に用いた観測値の数である。(8)式の右辺第一項は、残差の二乗和に依存している。他方、右辺第二項は、自由度減少に対する罰則であり、パラメータ数の増加に伴って上昇する。つまり、HQ 情報量基準では、(8)式で表現される情報量が小さいほど望ましいということになる。なお、HQ 情報量を算出するにあたって、どちらのモデルにおいても、被説明変数の単位は人数に戻してある。HQ 情報量基準で二つのモデルを比較すると、表1にまとめられているように、神経回路網モデルにおいて情報量が小さくなった。したがって、自由度の減少分を差し引いても、神経回路網モデルの方が線形回帰モデルよりも高い説明力を有しているといえる。

以上のように、非線形な神経回路網モデルを使うことで、線形回帰モデルよりも、観客数の変動をうまく説明できると判明した。しかしながら、別の観点から、非線形性の問題を論じてみるのも有用であろう。言うまでもなく、線形回帰モデルでは、非線形性が無視されている。これがどのくらい致命的な誤りであるのかを検定する。Lee *et al.* (1993) は、ラグランジュ乗数検定の枠組みで、検定方法（以下「LWG 検定」）を考案した。制約なしのモデルは神経回路網モデルと線形回帰モデルを組み合わせたものであり、制約ありのモデルは線形回帰モデルである。LWG 検定では、まず線形回帰モデルを推定し、残差を得る。次に、非線形回帰モデルの推定に関わる煩雑さを避けるため、結合関数のパラメータを無作為に与えて、ニューロンのデータを作る¹⁰⁾。そして、残差をニューロンに回帰し、ニューロンのすべての係数がゼロであるという仮説を F -検定する。この手順を千回繰り返し、5%以上の頻度で F -値が有意なら、有意水準5%で「線形回帰モデルで無視された非線形性は重要でない」という帰無仮説を棄却する。本稿のデータを用いてLWG 検定をやってみたところ、残念ながら、有意水準5%で上記の仮説を棄却できなかった。つまり、LWG 検定の結果は「線形回帰モデルで無視された非線形性は重要とはいえない」というものになった。

LWG 検定の結果を解釈するには、留意すべき点が二つある。上述のように、LWG 検定では、ニューロン生成の際、結合関数のパラメータは推定されず、無作為に値が与えられる。任意に与えたパラメータ値が見当外れであれば、たとえ無視できない非線形性が存在しても、残差のニューロンへの回帰は有意とならないであろう。このような LWG 検定の検出力の弱さは、留意すべき点の一つである。さらに、LWG 検定における制約なしのモデルは、神経回路網モデルそのものではなく、神経回路網モデルと線形回帰モデルの組み合わせである。したがって、制約なしのモデルと制約ありのモデル（線形回帰モデル）を比較して後者が選ばれたとしても、それ

10) Lee *et al.* (1993) の検定では、(2)式と異なり、結合関数から定数項が除かれている。

が直ちに神経回路網モデルの否定に繋がらない。この二点から、LWG 検定の結果は線形回帰モデルの神経回路網モデルに対する優位性の証明にはならない。とはいえ、LWG 検定で非線形性の重要性を検出できなかったことは、神経回路網モデルの高い説明力を割り引いて受け止める必要性を示唆しているといえよう。

6 結 論

スポーツの観客数は、試合毎に大きく変動し、予測するのが困難である。スポーツ経済の先行研究において、線形回帰モデルでは、観客数の変動をあまりうまく説明できていない。線形回帰モデルは、潜在的な観客の入場券購入に関する意思決定を、簡単な線形関数で近似したものである。本稿は、先行研究において観客数の変動がうまく説明されない一因をモデルの線形性に求め、非線形回帰モデルを用いることにした。非線形回帰モデルのなかでも、人間の脳で行われる情報処理に倣った神経回路網モデルに着目した。

本稿は、西京極スタジアムでの京都パープルサンガ主催試合の観客数に関するデータを用いて、線形回帰モデルと神経回路網モデルとで、サンプル期間内の事後的な予測力を比較した。その結果、決定係数でみてもHQ 情報量基準でみても、神経回路網モデルの方が線形回帰モデルよりも優れていた。ただし、LWG 検定で、強い非線形性を検出できなかったので、神経回路網モデルの線形回帰モデルに対する優位性は若干割り引いて受け止める必要があろう。

実際、たまたま神経回路網モデルの説明力が高いという結果が本稿の分析で出ただけという可能性は完全には否定できない。よく知られているように、非線形回帰モデルでは、推定結果がパラメータの初期値をどう選ぶかで変わってくる。非線形回帰モデルのパラメータ推定は乱数の生成を要するが、コンピュータ上で乱数を発生させるには乱数の種 (“seed”) を用意する。その乱数の種の値によって、生成される乱数が異なり、推定結果が異なってくるのである。したがって、本当に神経回路網モデルが高い説明力を持つのか確かめるためには、乱数の種を数多く試してそれらの結果を比較するという、モンテカルロ的な分析が必要となる。この点については、今後の研究課題としたい。

付 論

本文で用いた変数の記号とデータの出典を表2にまとめた。「京都サンガ」は京都パープルサンガ公式ホームページ、「J リーグ」は J リーグ公式ホームページ、「京都新聞」は京都新聞縮刷版 CD-ROM、統計京都は京都府発行の『統計京都』各号をそれぞれ指す。

表2 変数の説明とデータの出所

記 号	変 数 の 説 明	データの出典
<i>attend</i>	西京極スタジアム入場者数	京都サンガ
<i>price</i>	S 席メイン当日券の実質価格	

	S席メインの当日券価格 消費者物価指数（京都市）	京都新聞 統計京都
<i>output</i>	鉱工業生産指数（京都府）	統計京都
<i>tv</i>	地上波によるテレビ生中継で1を取るダミー変数	京都新聞
<i>opponent</i>	対戦相手の平均入場者数	Jリーグ
<i>rain</i>	雨天の試合で1を取るダミー変数	京都サンガ
<i>weekday</i>	平日の試合で1を取るダミー変数	京都サンガ
<i>kansai</i>	在阪チームとの対戦で1を取るダミー変数	京都サンガ
<i>special</i>	降格・昇格・優勝に絡む試合で1を取るダミー変数	京都新聞
<i>goldweek</i>	ゴールデンウィーク中の試合で1を取るダミー変数	

参 考 文 献

- [1] Bhattacharya, M., and Smyth, R. (2003) "The Game is not the Same : the Demand for Test Match Cricket in Australia," *Australian Economic Papers*, Vol. 42, pp. 77-90.
- [2] Borland, J., and Lye, J. (1992) "Attendance at Australian Rules Football: a Panel Study," *Applied Economics*, Vol. 24, pp. 1053-1058.
- [3] Brooks, C. (2002) *Introductory Econometrics for Finance*. Cambridge, UK : Cambridge University Press.
- [4] Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997) *The Econometrics of Financial Markets*. Princeton, NJ : Princeton University Press.
- [5] Davidson, R., and MacKinnon, J. G. (2004) *Econometric Theory and Method*. New York, NY : Oxford University Press.
- [6] Falter, J. M., and Pérignon, C. (2000) "Demand for Football and Intramatch Winning Probability : An Essay on the Glorious Uncertainty of Sports," *Applied Economics*, Vol. 32, pp. 1757-1765.
- [7] Franses, P. H., and van Dijk, D. (2000) *Non-Linear Time Series Models in Empirical Finance*. Cambridge, UK : Cambridge University Press.
- [8] Granger, C. W. J., and Jeon, Y. (2003) "Thick Modeling," *Economic Modelling*, Vol. 21, pp. 323-343.
- [9] Hannan, E. J., and Quinn, B. G. (1979) "The Determination of the Order of an Autoregression," *Journal of the Royal Statistical Society, Series B*, Vol. 41, pp. 190-195.
- [10] Judd, K. L. (1998) *Numerical Methods in Economics*. Cambridge, MA : MIT Press.
- [11] Kendrick, D. A., Mercado, P. R., and Amman, H. M. (2006) *Computational Economics*. Princeton, NJ : Princeton University Press.
- [12] Lee, T. H., White, H., and Granger, C. W. J. (1993) "Testing for Neglected Nonlinearity in Time Series Models," *Journal of Econometrics*, Vol. 56, pp. 269-290.
- [13] McNelis, P. D. (2005) *Neural Networks in Finance*. Burlington, MA : Elsevier Academic Press.
- [14] Peel, D., and Thomas, D. (1988) "Outcome Uncertainty and the Demand for Football : an Analysis of Match Attendances in the English Football League," *Scottish Journal of Economics*, Vol. 35, pp. 242-249.
- [15] Rumelhart, D. E., and McClelland, J. L. (1986) *Parallel Distributed Processing*. Cambridge, MA : MIT Press (甘利俊一監訳『PDP モデル』産業図書, 1989年).

- [16] Sargent, T. J. (1993) *Bounded Rationality in Macroeconomics*. Oxford, UK : Oxford University Press.
- [17] ——— (1999) *The Conquest of American Inflation*. Princeton, NJ : Princeton University Press.
- [18] Smith, K., and Gupta, J. (2002) *Neural Networks in Business*. Hershey, PA : IRM Press.
- [19] Wilson, P., and Sim, B. (1995) “The Demand for Semi-Pro League Football in Malaysia 1989-91 : a Panel Data Approach,” *Applied Economics*, Vol. 27, pp. 131-138.
- [20] 鈴木智也 (2006) 「西京極スタジアムにおける京都パープルサンガホームゲーム観戦に対する需要の分析」『龍谷大学経済学論集』第46巻第1・2号, 15～36頁。
- [21] ——— (2007) 「ニューラルネットモデルによる京都への観光需要の予測」『社会科学研究年報』龍谷大学社会科学研究所, 第37号, 69～81頁。

(受付 2007年8月10日)