

将棋の評価関数の非線形性

鈴木 豪 小谷 善行

東京農工大学大学院工学研究科

go@fairy.ei.tuat.ac.jp

kotani@cc.tuat.ac.jp

概 要

コンピュータ将棋における評価関数は、特徴量に対する線形和の形のもが多く使われている。しかし、一般に評価関数は非線形であると考えられる。本稿では将棋の非線形評価関数を学習し、その結果から非線形性と線形近似の妥当性を探った。将棋の非線形評価関数を得るためには3層の階層型ニューラルネットワークとTD(λ)法を使った。入力は単純なものから複雑なものまで3種類用意し、プロの棋譜から評価関数の学習を行った。そして得られた関数の出力を検証し、その非線形性を探った。実験の結果から、我々は将棋の評価関数の非線形性は小さく、特徴量に対する線形和の関数でも実用的であることを主張する。

Nonlinearity of an Evaluation Function in Shogi

Tsuyoshi SUZUKI

Yoshiyuki KOTANI

Tokyo University of Agriculture and Technology

go@fairy.ei.tuat.ac.jp

kotani@cc.tuat.ac.jp

Abstract

An evaluation function that has the form of the linear weighted sums of features in the position is used in the computer Shogi. But, generally it is thought that an evaluation function is nonlinear. In this paper, the nonlinear evaluation function of Shogi was learned by the neural network and TD(λ), and we investigated its nonlinearity. As the result, we insist that the nonlinearity of the evaluation function in Shogi is small and the form of linear weighted sum is practical.

1 はじめに

コンピュータ将棋における評価関数は、局面の生情報をそのまま使用せずに、駒価値や王の安全度などといった局面的特徴量と呼ばれる量に変換して、その値を変数とする関数で与えることが多い。評価関数はとくに次式のような線形和の形のもがよく使われている：

$$(\text{評価値}) = \sum_i w_i f_i(x) \quad x \in \text{局面}$$

ここで $f_i(x)$ は特徴 i を局面 x から求める関数であり、 w_i はその特徴量に対する重みである。しかし、一般に評価関数は非線形であると考えられる。そこで通常使われる線形和の評価関数は妥当であ

るだろうかと言う疑問がわいてくる。将棋においてもニューラルネットワークを使って非線形評価関数を学習させた研究がある[1,2]が、これらは妥当な出力を得る学習が行われたかということが議論の中心であり、得られた関数の非線形性についてほとんど議論されていない。本稿では TD(λ)法とニューラルネットワークを使って将棋の非線形評価関数の学習し、その結果得られた関数の出力に対し、非線形性を検証した。

学習に使用した TD(λ)法は Sutton により提案されたアルゴリズムである[3]。これは、過去の観測状態を考慮して効率の良い学習を行う。TD(λ)法とニューラルネットワークによるゲームプレイングプログラムの学習には様々なものがあるが Tesauero のバックギャモンプログラム TD-Gammon はその有効

性を最もよく示した例である[4]。TD-Gammon は自己対戦の学習から、世界チャンピオンと互角に戦えるまでになった。将棋において TD(λ)法を使った学習には[5,6,7] などがあるが、ここでは線形の評価関数が使われている。本稿では TD(λ)と3層の階層型ニューラルネットワークを使用した。学習にはプロの棋譜を使い、入力に使った特徴量は駒価値のみの単純なものからより多くの評価要素を含んだものまで3種類用意した。そしてそれぞれの関数に対し1要素のみを変化させた出力、2要素の全ての組合せに対する出力を調査した。

第2章ではニューラルネットワークと TD(λ)法を使った学習法について述べる。第3章では第2章の方法で将棋の非線形評価関数の学習実験について述べる。第4章では実験結果を示しその考察を行う。最後に、第5章で我々の結論を述べる。

2 TD 法とニューラルネットワークによる学習

本章では非線形評価関数の学習に利用したアルゴリズムについて述べる。

2.1 TD(λ)法

時刻 $1, \dots, m$ にける観測列を $x^1, x^2, \dots, x^{m-1}, x^m$ とし、最終的な結果を z とする。予言は観測データ x^t とパラメータ w を用いて $P^t = P(x^t, w)$ と表されるものとする。このとき TD(λ)法では、近い未来の予言との差をとり、更に過去への関係を考慮してパラメータ w を更新する:

$$w \leftarrow w + \sum_{t=1}^m \Delta w^t$$

$$\Delta w^t = \alpha (P^{t+1} - P^t) \sum_{k=1}^t \lambda^{t-k} \nabla_w P^k$$

ここで λ は過去の予言への依存関係を表す正定数で、 α は学習率と呼ばれる定数である。更新式はインクリメンタルに計算できるため計算コストは大きくないことが知られている。

2.2 ニューラルネットワークの学習

ニューラルネットワークの出力は次のように計算される。 $y_i^t (= P_k^t)$ をニューラルネットワークの i 番

目のユニットの時刻 t における出力 z_k を出力ユニット k に対応する出力 ($k \in O$) (理想的には z_k は各 $P_k^t (t=1, \dots, m)$ により予言される)、 m を観測ベクトル (x^t) の数、 $P_k^{m+1} \equiv z_k$ 、 w_{ij}^t をユニット i からユニット j へのコネクシオンの時刻 t における重み、 FO_j をユニット j からのコネクシオンを持つユニットのインデックスの集合とする。このときユニット j の出力は

$$s_j^t = \sum_{i \in PI_j} w_{ij}^t y_i^t$$

$$y_j^t = f(s_j^t) = \frac{1}{1 + e^{-s_j^t}}$$

となる。学習には TD(λ)法において $\lambda=0$ とおいた、TD(0)を使用した。このとき 2.1 節の更新式からネットワーク重みの更新は次のようになる:

$$w_{ij}^{t+1} = w_{ij}^t + \alpha \delta_j^t y_i^t$$

$$\delta_j^t = -\frac{\partial E^t}{\partial s_j^t} = \begin{cases} (P_i^{t+1} - P_i^t) y_i^t (1 - y_i^t) & \text{for } i \in O \\ \sum_{j \in FO_i} \delta_j^t w_{ij}^t y_i^t (1 - y_i^t) & \text{otherwise} \end{cases}$$

ここで α は学習率と呼ばれる定数である。

3 将棋の非線形評価関数の学習

3層の階層型ニューラルネットワークと TD(0)を使って将棋の評価関数の学習を行った実験について述べる。

3.1 学習の条件

(I) 観測ベクトル

局面の状態は観測ベクトル x^t を用いて表される。

観測ベクトルの要素としては次の3種類を用意した。これらは V1 から V3 になるに従って、より複雑なものになっている。

(V1) 先手と後手の盤上の駒および持ち駒の数の差(13要素):

{歩, 香, 桂, 銀, 金, 角, 飛,
と, 成香, 成桂, 成銀, 馬, 竜}

(V2) 先手と後手の盤上の駒数の差、持ち駒の数の差、手数(21要素):

{歩, 香, 桂, 銀, 金, 角, 飛,

と、成香、成桂、成銀、馬、竜(盤上の駒)、
歩、香、桂、銀、金、角、飛(持ち駒))

(V3) 先手と後手の盤上の駒の数・持ち駒数・王の
周りの利きの数・手数(53要素):

(歩、香、桂、銀、金、角、飛、
と、成香、成桂、成銀、馬、竜(先手の盤上の駒)、
歩、香、桂、銀、金、角、飛、
と、成香、成桂、成銀、馬、竜(後手の盤上の駒)、
歩、香、桂、銀、金、角、飛(先手の持ち駒)、
歩、香、桂、銀、金、角、飛(先手の持ち駒)、
先手陣にある後手の利き、後手陣にある先手の利き、
先手王の8近傍の先手の利き1、先手王の8近傍の
先手の利き2、先手王の8近傍の後手の利き1、先手
王の8近傍の後手の利き2、後手王の8近傍の後手の
利き1、後手王の8近傍の後手の利き2、後手王の8
近傍の先手の利き1、後手王の8近傍の先手の利き2、
先手王の自由度、後手王の自由度)

ここで、“利き1”とは升の一つでも利きがあれば1
そうでなければ0と数えたもので、“利き2”とは升に
複数の利きがあった場合はそれらをそのまま合計
したものである。

(II) 棋譜からの学習

観測ベクトルの作成は将棋の棋譜からとし、将棋
年鑑'90,'91,'92,'98,'99 年度版[8]の棋譜 2000
局を使用した。

(III) 教師値とその他パラメータ

観測ベクトル x に対する教師値 $T(x)$ を次式で定
義する:

$$T(x^i) = \begin{cases} P(x^{i+1}) & \text{if } i \neq m \\ z=1 & \text{if } i=m \text{ and 先手勝ち} \\ z=0 & \text{if } i=m \text{ and 後手勝ち} \end{cases}$$

(IV) ニューラルネットワーク

本実験では図1のような3層の階層型ニューラル
ネットワークを使った。3層のニューラルネットワ
ークはその中間ユニット数を増やせば、任意の連続
関数が表現できることが知られている。

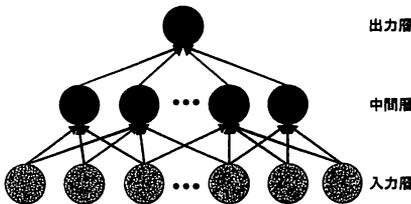


図1 3層の階層型ニューラルネットワーク

入力(I)における観測ベクトルであり、出力は入
力に対する勝敗の予測を示す実数である。すなわ
ち先手勝ちの可能性が高い場合には1に近く、後
手勝ちの可能性が高い場合には0に近い値となる。
学習率は $\alpha=0.8$ 固定とし、終了条件は全棋譜を繰
り返し 250 回学習したときとした。これは重みの更
新回数にすると約5800万回になる。隠れ層のユニ
ット数は V1,V2 に対しては 2, 4, 8, 16, 32, 64 とし、
V3 に対しては 2,4,8 とした。実験は、それぞれ重み
の初期値を変えて 12 回行った。

3.2 学習された関数の評価

学習で得られる評価関数は非線形の形をして
いるが、実際に得られたものは線形に近い出力を
示した。そこで得られた関数がどの程度線形に近
いかを評価するために、重回帰分析で用いられる
寄与率を利用した。

いま、実験で得られた非線形関数を $f(x)$ とし、
これを近似した線形関数を $a \cdot x + b$ とする。ここで
 $a = (a_1, \dots, a_n)$ はベクトル、 b はスカラーであり、 \cdot は
内積を表す。 $f(x)$ の線形近似関数は自乗誤差

$$\sum_{x \in \Omega} \{f(x^i) - (a \cdot x^i + b)\}^2$$

を最小にする超平面とする。ここで Ω はとりうる状
態の全ての集合をあらわす。自乗誤差を最小にし
る a と b は方程式

$$\begin{pmatrix} \sum_{i=1}^m x_1^i x_1^i & \dots & \sum_{i=1}^m x_1^i x_n^i & \sum_{i=1}^m x_1^i \\ \vdots & \ddots & \vdots & \vdots \\ \sum_{i=1}^m x_n^i x_1^i & \dots & \sum_{i=1}^m x_n^i x_n^i & \sum_{i=1}^m x_n^i \\ \sum_{i=1}^m x_1^i & \dots & \sum_{i=1}^m x_n^i & n \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m x_1^i f(x^i) \\ \vdots \\ \sum_{i=1}^m x_n^i f(x^i) \\ \sum_{i=1}^m f(x^i) \end{pmatrix}$$

を解くことによって得られる。可能な状態 Ω の数は
最も小さな観測ベクトル V1 を使っても数十億個あ
る。これらに対する出力値をすべて計算するのは
困難なので、本実験では Ω の部分集合を使って
 a と b を得た。評価は、この計算により得られた線
形近似の予測値 $a \cdot x^i + b$ と学習で得られた関数の
出力値 $f(x^i)$ を使って、線形近似した場合の適合
具合を寄与率で評価する。寄与率は次式で定義さ
れる:

$$(\text{寄与率}) = \frac{\sum_{i \in \Omega} (a \cdot x^i + b) - \overline{(a \cdot x + b)}^2}{\sum_{i \in \Omega} (f(x^i) - \overline{f(x)})^2}$$

ここで $\overline{(a \cdot x + b)}$ は予測値の平均、 $\overline{f(x)}$ は $f(x^i)$ の平均である。寄与率は決定係数とも呼ばれ、その平方根は重相関係数とよばれる。寄与率が1に近いほど線形関数での予測の精度が高く、逆に0に近いと予測が困難となる。

4 実験結果と考察

本章では学習で得られた評価関数の出力からその非線形性を考察する。各評価要素の単体および二つの評価要素の全ての組合せについて、その出力を調べた。以下で示すグラフは、それぞれ初期値を変えて12回行った出力値の平均である。以下の図において UNIT=XX は、中間層のユニット数が XX 個あることを意味する。

4.1 要素の単体での出力

ひとつの要素の入力値のみを変化させ、他の要素の入力値をすべて0に固定して学習で得られた関数の出力を調べた。

図2は観測ベクトルが V1 のときの先手からみた金の枚数の差によるネットワークの出力を表している。図2から中間ユニット数に関係なく、出力はほぼ同じ形で、得られた曲線も直線に近いことがわかる。V1の他の要素に対してもほぼ同様の傾向が見られた。

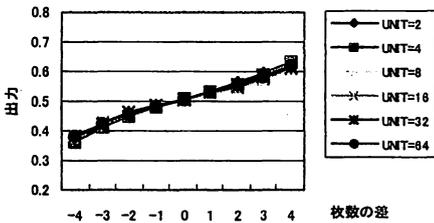


図2 金の枚数の差による出力

図2は見た目から、比較的直線に近いと考えられるが、これらを計量的に、判断するため3.2節の寄与率を計算して、その近さを計測した。表1は観測ベクトル V1 に対する評価要素の寄与率の結果の一部を表している。ほとんどの要素に対し、その値は1に近く、平均で0.92であった。このことから

V1 の各評価要素の単体での出力は線形関数で近似してもその精度が高いがわかる。

表1 観測ベクトル V1 の寄与率

| | UNIT=2 | UNIT=4 | UNIT=8 | UNIT=16 | UNIT=32 | UNIT=64 | 平均 |
|----|--------|--------|--------|---------|---------|---------|------|
| 歩 | 1.00 | 1.00 | 0.94 | 0.87 | 0.93 | 0.89 | 0.94 |
| 金 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 飛 | 1.00 | 1.00 | 1.00 | 0.95 | 0.73 | 0.83 | 0.92 |
| 成銀 | 1.00 | 1.00 | 1.00 | 0.97 | 0.99 | 0.99 | 0.99 |
| 馬 | 0.99 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 |

図3は観測ベクトルが V2 ときの盤上の歩の先手からみた枚数の差によるネットワークの出力を表している。ただし、その差は-4~4に限定した。これは例えば、枚数の差が18枚などといった、学習に現れない局面においてはその出力がでたらめなものとなり、評価の対象としてはふさわしくないためである。以下でも、他の要素に対して同様の制限をつけている。図3においても中間ユニット数に関係なくどの曲線も同様の傾向を示している。また、それぞれ直線に近いと考えられる(図3の出力値のスケールは図2のそれとは異なることに注意)。これは他の要素に対しても同様であった。

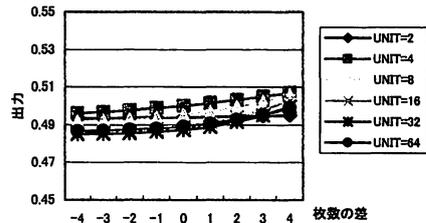


図3 歩の枚数の差による出力

表2に観測ベクトル V2 に対する持ち駒の歩・金・飛の寄与率を示す。どの寄与率も1に近く、平均で0.93であった。観測ベクトルベクトル V2 に対しても精度高く線形近似できることを示している。

表2 観測ベクトル V2 の寄与率

| | UNIT=2 | UNIT=4 | UNIT=8 | UNIT=16 | UNIT=32 | UNIT=64 | 平均 |
|------|--------|--------|--------|---------|---------|---------|------|
| 歩(持) | 0.97 | 0.94 | 0.96 | 0.98 | 0.99 | 0.99 | 0.97 |
| 金(持) | 1.00 | 0.99 | 1.00 | 1.00 | 0.97 | 0.95 | 0.98 |
| 飛(持) | 0.99 | 0.99 | 1.00 | 0.99 | 0.98 | 0.96 | 0.98 |

図4は観測ベクトルが V3 のときの先手陣にある後手の利きの数による出力である。図2、図3は駒価値に関係する出力であったが、駒価値とは異なる評価要素に対するこの場合も、ユニット数に関係なくいずれも直線に近い形で出力値が減少している。

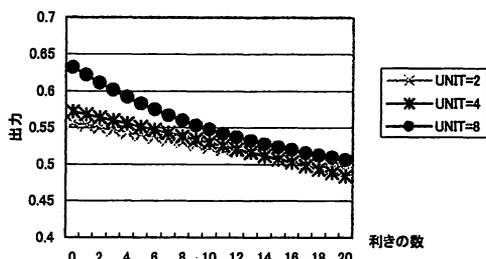


図4 先手陣にある後手の利きの数による出力

表3に観測ベクトル V3 の先手の利きの状態と王の周りの状態の寄与率を示す。いずれも 0.94 以上で、平均では 0.98 とこれらの評価要素も、非常に高い値を示した。王の周りの状態は非線形の度合いが高いと予想されたが、実験の結果は線形で評価してもよいことを示している。

表3 ベクトル V3 の寄与率

| | UNIT=2 | UNIT=4 | UNIT=8 | UNIT=16 | UNIT=32 | UNIT=64 | 平均 |
|--------|--------|--------|--------|---------|---------|---------|------|
| 先陣に後利 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.98 | 1.00 |
| 先王に先利1 | 1.00 | 1.00 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 |
| 先王に先利2 | 1.00 | 1.00 | 0.94 | 0.98 | 0.97 | 0.97 | 0.98 |
| 先王に後利1 | 1.00 | 1.00 | 1.00 | 1.00 | 0.99 | 0.99 | 1.00 |
| 先王に後利2 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 先王自由度 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |

一般に中間ユニット数が多いほど複雑な非線形関数を表すことができる。しかし、図2～図4が示すように、どの中間ユニット数の場合にも、ほぼ同様の出力傾向がみられた。すなわち、中間ユニット数はこの実験の最低の2個でも、より大きなユニット数を持ったネットワークと同様の出力をしている。他の要素に対しても同様の結果を得た。これらの結果から、中間ユニット数は少なくともよく、この意味において非線形性は小さいと考えられる。

4.2 手数と他要素との関係

駒価値は局面の進行度とともに変化すると考えられる。そこでここでは手数と駒の組合せによる出力の変化を示す。

観測ベクトル V2 に対し、中間ユニット数が8のときに、持ち駒に銀・金・角が1枚あると仮定し、手数を変化させたときの出力は図5となった。持ち駒は全体的に手数とともにその出力値が下がる傾向を示した。その中で、金の出力値だけは手数が増えても減少の幅が小さく、一方、飛と角の出力は大きく減少した。盤上の駒の出力は緩やかに増加する傾向を示した。

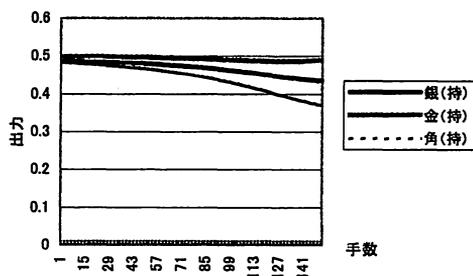


図5 手数と駒1枚の出力

観測ベクトル V2 に対し、手数と駒の枚数についての組合せと出力値で寄与率を計算した。その結果を表4に示す。寄与率が1に近いものもあるが、平均で 0.82 とそれほど大きな値にはならなかった。これは手数が増えると、枚数の差により、その出力値が緩やかに、ねじれるように変化している傾向を示したことによる。この様子を図6に示す。

表4 駒の枚数の差と手数の寄与率

| | UNIT=2 | UNIT=4 | UNIT=8 | UNIT=16 | UNIT=32 | UNIT=64 | 平均 |
|------|--------|--------|--------|---------|---------|---------|------|
| 歩 | 0.78 | 0.90 | 0.63 | 0.78 | 0.90 | 0.81 | 0.80 |
| 桂 | 0.82 | 0.88 | 0.88 | 0.84 | 0.76 | 0.76 | 0.82 |
| 成銀 | 0.78 | 0.82 | 0.90 | 0.85 | 0.76 | 0.88 | 0.83 |
| 馬 | 0.79 | 0.74 | 0.73 | 0.72 | 0.76 | 0.78 | 0.75 |
| 竜 | 0.79 | 0.87 | 0.85 | 0.82 | 0.92 | 0.91 | 0.86 |
| 香(持) | 0.79 | 0.96 | 0.92 | 0.91 | 0.91 | 0.89 | 0.90 |
| 角(持) | 0.79 | 0.88 | 0.89 | 0.90 | 0.94 | 0.94 | 0.89 |

図6は手数が少ないうちは枚数の差による出力値の変化は小さいが、手数が進むにつれてそれは大きくなっている。

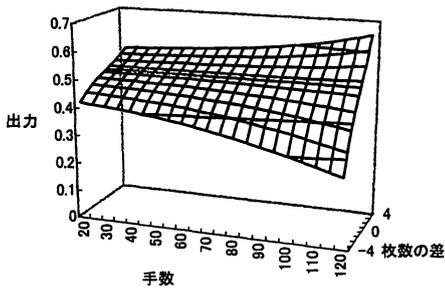


図6 手数と桂馬の枚数の差による出力

観測ベクトルV3を使って学習した手数と先手陣にある後手の利き・先手王の8近傍の状態の寄与率を表5に示す。これらの寄与率もほとんどが0.9以上で、平均0.92であった。このことから手数と自陣にある敵の利き・王の周りの状態も線形に評価可能であることを示している。

表5 先手王の8近傍状態と手数の寄与率

| | UNIT=2 | UNIT=4 | UNIT=8 | 平均 |
|-------------|--------|--------|--------|------|
| 先手陣にある後手の利き | 0.98 | 0.87 | 0.86 | 0.91 |
| 先手王に先手の利き1 | 0.99 | 0.84 | 0.90 | 0.91 |
| 先手王に後手の利き1 | 0.95 | 0.95 | 0.91 | 0.94 |
| 先手王の自由度 | 0.95 | 0.94 | 0.91 | 0.93 |

4.3 その他評価要素の関係

図7に、観測ベクトルをV3としたとき、先手の竜と後手王の自由度の出力を示す。実験前、王の周りの敵の利きの状態と持ち後は比較的大きな非線形性を持つと予想したが、得られた結果は大きな比較的平面に近いグラフとなった。後手から見た逆の出力に対しても同様の結果であった。

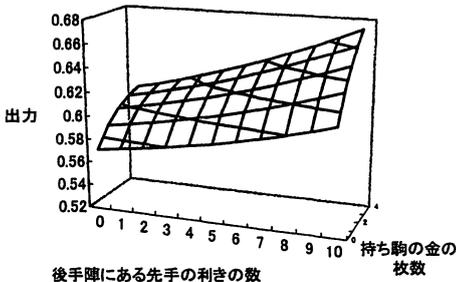


図7 先手竜-後手王の自由度の出力

表6に先手竜と利き・王の8近傍の状態の寄与率を示す。平均は、0.95であり、これらの関係についても線形で近似しても精度が高い結果となった。

表6 先手竜と利きの寄与率

| | UNIT=2 | UNIT=4 | UNIT=8 | 平均 |
|-----------------|--------|--------|--------|------|
| 先手竜 | | | | |
| -先手陣にある先手の利き | 0.99 | 0.97 | 0.84 | 0.93 |
| 先手竜 | | | | |
| -後手陣にある先手の利き | 0.99 | 0.98 | 0.97 | 0.98 |
| 先手竜 | | | | |
| -先手王の8近傍の先手の利き1 | 0.99 | 0.99 | 0.95 | 0.98 |
| 先手竜 | | | | |
| -先手王の8近傍の後手の利き1 | 0.97 | 0.88 | 0.69 | 0.85 |
| 先手竜-先手王の自由度 | 0.98 | 0.97 | 0.89 | 0.94 |
| 先手竜 | | | | |
| -後手王の8近傍の先手の利き1 | 0.99 | 0.98 | 0.93 | 0.97 |
| 先手竜 | | | | |
| -後手王の8近傍の後手の利き1 | 0.97 | 0.99 | 0.95 | 0.97 |
| 先手竜-後手王の自由度 | 0.97 | 0.98 | 0.96 | 0.97 |

表7に先手王の8近傍の状態と後手の先手陣にある利きの状態の寄与率を示す。これらの平均は0.97であった。そのほか異なる駒の種類の間についても寄与率は非常に1に近い値となった。

表7 王の8近傍の状態と利きの寄与率

| | UNIT=2 | UNIT=4 | UNIT=8 | 平均 |
|-------------|--------|--------|--------|------|
| 先手陣にある後手の利き | | | | |
| -先手王の先手の利き1 | 1.00 | 1.00 | 0.92 | 0.97 |
| 先手陣にある後手の利き | | | | |
| -先手王の後手の利き1 | 1.00 | 0.97 | 0.94 | 0.97 |
| 先手陣にある後手の利き | | | | |
| -先手王の自由度 | 1.00 | 1.00 | 0.94 | 0.98 |

全体としてまれにしか起こらない局面(あるいは学習時に1度も現れなかった局面)に対する入力ベクトルに対し、その出力がとびぬけた値を示すことがあった。これは中間ユニット数が多くなると増加する傾向にある。中間ユニットが増えると寄与率が若干下がる傾向となるが、これは関数の非線形性よりこのような値に引きずられた結果が主な原因と考えられる。寄与率が小さな要素に対し、入力ベ

クトルをより小さな範囲(より起こりそうな局面)に限定した場合にはこの値は1に近くなった。範囲を限定しても寄与率が小さいものもごく一部には存在したが、これはうまく学習できなかったことが原因である。他の棋譜500局面を使用した別の実験においては、このような場合にも高い値を示すことが確認できた。本実験で用いた観測ベクトル V1,V2,V3 を使った評価関数は各要素に対する出力が線形関数に近く、我々は高度に複雑な評価関数でなければ線形のものでもほぼ妥当な出力が得られると推測する。付録に二つの要素間の出力の変化を示すグラフをいくつか示す。

6 まとめ

TD(λ)法とニューラルネットワークを使って将棋の評価関数を学習し、その非線形性を探った。中間ユニットの数を変化させてもその出力は類似しており、その意味で非線形性は小さいと考えられる。また、二評価要素間の全ての組合せに対し、その寄与率から、線形近似の可能性を探った。結果は、本実験で使用した観測ベクトルを使った関数において、ほとんどの全ての寄与率が1に近く、線形近似でも高い精度が得られる結果となった。我々はこれらの結果から特徴量を使った将棋の評価関数の非線形性は小さく、一般に使われる線形和形でも十分実用的であると予想する。

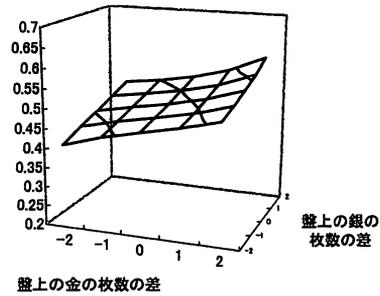
参考文献

- [1] 春日,鈴木,小谷. ニューラルネットワークを用いた詰め将棋評価実験. GPW'99, pp.101-104. 1999.
- [2] 徳田, 小谷. 将棋におけるニューラルネットワークを用いた評価関数の生成実験, 第 2 回 GPW, 1994.
- [3] Sutton, R.S. Learning to Predict by the Methods of Temporal Differences. Machine Learning, Vol. 3, pp.9-44, 1988.
- [4] Tesauro, G. Temporal Difference Learning and TD-Gammon. Communications of the ACM, Vol. 38, pp.58-68, 1995.
- [5] 薄井,鈴木,小谷. TD 法を用いた将棋の評価関数の学習. GPW'99, pp.31-18,1999.
- [6] 薄井,鈴木,小谷. プロの棋譜を用いた TD 法による将棋の評価関数の学習. TD 法を用いた将棋の評価関数の学習. ゲーム情報学研究会.

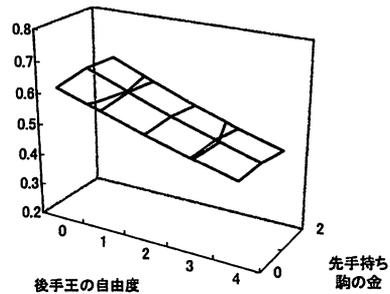
2000-GI-4, pp.17-21, 2000.

- [7] Suzuki,T. Inui,N. KOTANI,Y. TD(λ, μ):Temporal Difference Learning using Future Observation. ゲーム情報学研究会, 2000-GI-4, pp.9-16, 2000.
- [8] 日本将棋連盟. 平成 1,2,4,10,11 年度版将棋年鑑. 1990-1992,1998,1999.

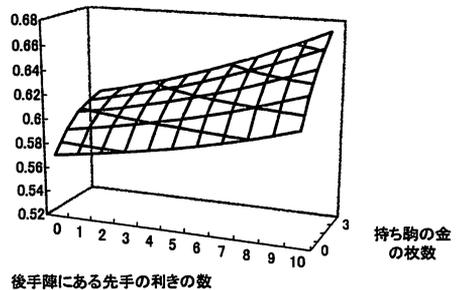
付録 二要素間の出力の変化



図A 盤上の金一盤上の銀 (V2, UNIT=32, 寄与率=0.92)



図B 後手王自由度一先手持駒金 (V3, UNIT=8, 寄与率=0.96)



図C 後手陣の先手利き一先手持駒金 (V3, UNIT=4, 寄与率=0.97)