ゲーム情報学 7－8
（２００２．３．１５）

将棋におけるプロの指し手とそれ以外の手の判別分析

本堂　敦、鈴木　豪、小谷　善行
東京農工大　工学研究科
{kagyuu ,go ,kotani}@fairy.ei.tuat.ac.jp

要旨

局面評価関数が線形和であるときプロの指し手の特徴量ベクトルからプロが選ばなかった指し手の特徴量ベクトルを引いた差分ベクトルを評価関数で評価すると正の値になるはずである。すなわち、プロの指し手を最善手とする局面評価関数は、プロの指し手の特徴量ベクトルからプロの選ばなかった指し手の特徴量ベクトルを引いた差分ベクトルの線形和を正の値にするような線形結合である。このような線形結合は、判別分析により求めることが出来る。プロの棋譜 2000 対局分を判別分析した結果、86%の確率でプロの指し手の評価値がプロの選ばなかった指し手の評価値よりも高くなるような評価関数を得ることができた。また、得られた局面評価関数は、人間の調整した局面評価関数とほぼ同等な強さを持っていた。

Discriminate Analysis Distinguish Moves that Professional Shogi Player Elected from The Others

Atsushi HONDOH, Tsuyoshi SUZUKI, Yoshiyuki KOTANI
Tokyo Univ. of Agri. & Tech., 2-24-16 Nakamachi, Koganei,Tokyo, Japan
{kagyuu ,go ,kotani}@fairy.ei.tuat.ac.jp

Abstract

When the shape of the position evaluation function is liner. It becomes a plus that evaluated value of the result that the vector of the move whom a professional didn't choose from the vector of the professional move was subtracted. In other words, the position evaluation function that distinguish professional move from others is the same as the liner combination, which makes the difference vector plus. Discriminate Analysis can find such liner combination easily. We use discriminate analysis toward the record of the 2000 professional games to make the position evaluation function. That function made the evaluation value of the professional move lager than evaluation values of the others at 86% of rates. And that function is strong as the function that human experts adjusted.

1.はじめに

現状の多くの強い将棋プログラムは、人手によって局面評価関数が調整されている場合が多い。そして、その局面評価関数は学習アルゴリズムにより調整された局面評価関数よりも強い場合が多い。局面評価関数の調整は、他の将棋プログラムや自分自身と対戦させてみて、その棋譜を見ながら望みの駒が望むような動きをするように局面評価関数を微調整することでおこなわれる。たとえば図 1 において、プログラマは桂馬を右上に動かしたいのに、現在の局面評価関数を使った場合に将棋プログラムは左上に桂馬を動かして銀を取るような振る舞いをしたとする。この場合には、銀の評価値を下げ、歩の評価値を上げて、さらに玉自由度の評価値を上げるなどして、桂馬が右上へ動き歩を取り王手をかけるような手を選択するようにする。

この方法を繰り返せば、確かに的確に局面を判断できる局面評価関数を作ることが可能である。しかし、人手によって試行錯誤しながら局面評価関数を調整する方法には、莫大な時間がかかり、また将棋（研究対象のゲーム）に対する幅広い経験や知識が必要である。

そこで、これを自動化することを考える。我々は、本論文で、多くの強い将棋プログラムにようにプログラマの望むような手を指すように局面評価関数を調整する代わりに、プロ棋士の棋譜を用いてプロ棋士の指した手を指すように局面評価関数を調整する手法を提案する。

2.局面評価関数について

多くの将棋プログラムでは、局面評価関数はべき乗を使う多項式やニューラルネットなどの複雑な機構を使うのではなく、実装上高速に動作する整数型の積と和だけで表現できる線形和を利用している。また、「将棋の局面の静的評価関数の非線形性」（鈴木,2000）[1]において、ニューラルネットによる局面評価関数と特徴量の線形和による局面評価関数を比較した結果、局面評価関数は特徴量の線形和で十分表現できることを示された。そこで、本研究では一般に用いられている線形和による局面評価関数に関して、その特徴量の重みを調整することとする。

3.プロ棋士の指し手とその他の手の差分

いま、ある局面の特徴量ベクトルを p、重みベクトルを w とし、局面の評価関数を

$F(p) = {}^{t}w \cdot p$ ・・・ （式 1）

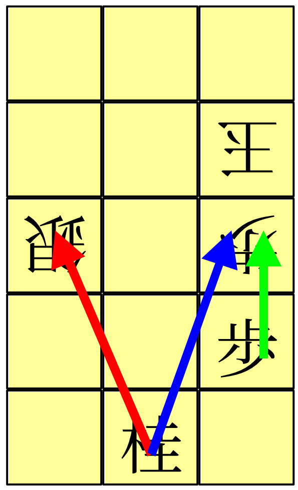


図 1:人手による調整の例

と表すとする。ここで、プロ棋士の指し手が合法手の中で手番側にとって最善手であると仮定すると、棋譜上のある局面の合法手についてプロ棋士の指し手 p_0 の評価値とプロ棋士の選ばなかった指し手 p_i (i= 1、2、3、…、n-1)の関係は（式 2）のようにあらわすことができる。

$F(p_0) > F(p_i)$

${}^t w \cdot p_0 > {}^t w \cdot p_i$ (i= 1、2、3、…、n-1) ・・・(式 2)

さらに、(式 2）の右辺を移項すると

${}^t w \cdot p_0 - {}^t w \cdot p_i > 0$ (i= 1、2、3、…、n-1) ・・・(式 3)

と表すことができる。ここで、評価関数 F(p) = ${}^t w \cdot p$ が線形結合であることを考慮に入れると(式 3)は

${}^t w \cdot (p_0 - p_i) > 0$ (i= 1、2、3、…、n-1) ・・・(式 4)

と変更することが出来る。すなわち、プロの指し手を最善手とするような評価関数 F(p)は、プロの選んだ指し手とプロの選ばなかった指し手の差分ベクトルの線形結合を正にするような線形結合であると見ることが出来る。

4. w の導出

判別分析は、二つの母集団において、同じ種類の変数が測定されている時に、それぞれの母集団からサンプルをひとずつ取り測定を行った場合、どちらのグループに属していたか高い確率で当てることを目的とするアルゴリズムである。具体的には、観測されたデータの線形結合がある値以下であればグループ A、以上であればグループ B となるように分ける線形結合を求める分析手法である。

本手法は、判別分析の手法を利用して、多数の棋譜から得られた事例に関して式４がなるべく多く満たされるような線形結合を求める。いま事例から大量に、プロ棋士の指し手とそれ以外の指し手の差分ベクトルを x_i を作ったとし、その線形結合(評価関数による評価) y_i がなるべく多く正になるような重みベクトル w を求めることを考える。

y_i ? ${}^t w \cdot x_i$ ・・・(式 5)

と表すとした場合、y の平均をなるべく大きく、分散をなるべく小さくするように w を調整すれば y_i がなるべく多く正になる（図 2)。すなわち 標準化距離

$$D(w) \; ? \; \frac{y\text{の平均}}{y\text{の標準偏差}} \quad \cdots(\text{式 } 6)$$

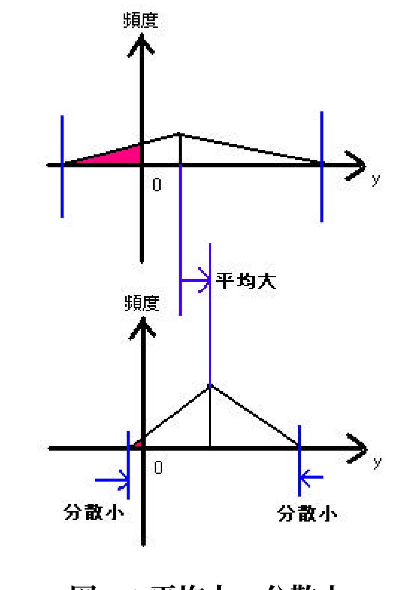


図 2：平均大・分散小

を最大化する w を求めればよい。

ここで、A_x を x の平均 ベクトル、S_x を x の共分散行列とすると

$$\begin{aligned} y\text{の平均} &?\,^t w \cdot A_x \\ y\text{の分散} &?\,^t w \cdot S_x \cdot w \end{aligned} \quad \cdots(\text{式 } 7)$$

だから

$$D(w) ? \frac{^t w \cdot A_x}{(^t w \cdot S_x \cdot w)^{1/2}} \quad \cdots(\text{式 } 8)$$

を最大化すればよい。w が D(w)を最大化するとき $D(w)^2$ も最大化するから

$$D(w)^2 \; ? \; \frac{(^t w \cdot A_x)^2}{^t w \cdot S_x \cdot w} \; ? \; \frac{^t w \cdot (A_x \cdot {}^t A_x) \cdot w}{w^t \cdot S_x \cdot w}$$

・・・(式 9)

を最大化しても良い。(二つ目の等号は、$^t w A_x$ がスカラ量であることから成立する。)

ここで、D(w) = D(kw) (k は定数)であることを利用して w のうち w= w_0($^t w_0$・S_x・w_0 = 1)なる w_0 を求めることとすると。(式 9)を最大化したいとき、$w^t{}_0$・S_x・w_0 = 1 を拘束条件とする未定係数法を利用して

$$L(w_0, ?) \; ?^t \, w_0 \cdot (A_x \cdot {}^t A_x) \cdot w_0 \; ? \; ?(^t w_0 \cdot S_x \cdot w_0 \; ?1)$$

・・・(式 10)

を最大化する w_0 を求めればよい。L(w_0、λ)を最大化する w_0 を求めるためには、w_0 で偏微分した結果が 0 になればよいから

$$\frac{?L(w_0, ?)}{?w_0} \; ? \; A_x \cdot {}^t A_x \cdot w_0 \; ? \; ? \cdot S_x \cdot w_0 \; ? \; 0$$

・・・(式 11)

を得る。

これを整理して

$$A_x \cdot {}^t A_x \cdot w_0 \; ? \; ? \cdot S_x \cdot w_0 \; ? \; 0$$

$$A_x \cdot {}^t A_x \cdot w_0 \; ? \; ? \cdot S_x \cdot w_0$$

$$S_x \cdot w_0 \; ? \; \frac{1}{?} A_x \cdot {}^t A_x \cdot w_0$$

$$w_0 \; ? \; \frac{1}{?} S_x^{\,?1} \cdot A_x \cdot {}^t A_x \cdot w_0$$

$^t A_x$・w_0 がスカラ量であることに注目すると、

$$w_0 \; ? \; \frac{^t A_x \cdot w_0}{?} \cdot S_x^{\,?1} \cdot A_x \; ? \; k \cdot S_x^{\,?1} \cdot A_x$$

・・・(式 12)

ここで、D(w) = D(kw) (k は定数)だから、D(w)を最大化する、すなわちプロ棋士の指し手の評価値をプロ棋士の指さなかった手の評価値よりもなるべく大きくする w_1 は

$$w_1 \; ? \; S_x^{\,?1} \cdot A_x \qquad \cdots(\text{式 } 13)$$

と表すことができる。

すなわち、求める重みベクトル列は、棋譜から得られた差分ベクトル x の共分散行列の逆行列と平均ベクトルをかけたものである。

5.実験

5.1.実験により調整する局面評価関数

判別分析による局面評価関数の調整では、一回の単純な計算で局面評価関数を求めることが可能である。すなわち、局面評価関数を複雑にすることができる。そこで、局面評価関数を調整することにする。

評価要素の項目に関して、まず思い浮かぶのが盤面上の駒の枚数と、駒台上の持ち駒の枚数である。盤面上の駒に関してこの他に、ある駒がどれだけのマスに一手で到達できるのかということも(潜在的に支配

している)評価関数で考慮に入れる。また、ある駒に対する味方側からの利き、敵側からの利きについても考慮に入れる。

これとは別に、玉を中心とした座標系で各マスに駒や壁や空白などがあることに関しても評価用途として取り入れる。駒については種類ごとに評価項目とする。ここで、座標は２次元として扱わずに１次元の距離を重ね合わせる事により平面上の位置を表現する。これは、実際に将棋プログラムに組み込んだ時に計算を早くするためと、学習に際して事例数を多くするという二つの意味がある。

座標系１は、近傍と呼ばれるものである。近傍の距離は

距離１ ＝ MAX(縦の変異、横の変異)
・・・(式 14)

であらわされる（図 3)。玉からのそれぞれの距離のマスについて、敵側の駒、味方側の駒、敵側利き数、空白、壁の個数を特徴量とする。

座標系２の距離は、

距離２ ＝ 縦の変異 ＋ 横の変異
・・・(式 15)

であらわされる（図 4)。距離 1 と同様に玉からのそれぞれの距離のマスについて、敵側の駒、味方側の駒、敵側利き数、空白、壁の個数を特徴量とする。

2	2	2
1	1	2
玉	1	2

図 3 距離 1

2	3	4
1	2	3
玉	1	2

図 4 距離 2

ここで、これらの特徴量には冗長な部分があるために共分散行列 S_x の逆行列を求めることが出来ない。そこで、一度項目を主成分に変換して、主成分に対して判別分析を行った後、主成分の重みから元の項目の重みを求める。

5.2.判別分析により局面評価関数を作る

プロ棋士の棋譜 2000 対局分から､プロ棋士の選んだ手とプロ棋士の選ばなかった手の特徴量ベクトルの差分ベクトルを抽出した結果 17,736,886 の事例を作ることが出来た。局面評価関数が、この事例をなるべく多くみたすようなに判別分析によって特徴量の重みを調整する実験を行った。

6.実験結果と考察

6.1.得られた評価関数の重みについて

得られた局面評価関数の重みについていくつか特徴的なものを述べる。なお局面評価関数は、持ち駒の歩一枚を 100 点に正規化してある。

表 1 は、駒台上の持ち駒の価値である。おおむね常識的な評価値の順である　歩＜香＜桂＜銀＜金＜角＜飛　の順に並んだ。ここで、王将の価値が 0 点であるが、これはプロの指し手とその他の手の差分ベクトルに関して王将の枚数に差がある事例が出てこないためである。実際の将棋プログラムに用いる場合には、玉の重みを大きい数

にするなどして対処する必要がある。

駒	歩兵	香車	桂馬	銀将	金将	角行	飛車	王将
価値	100	184	187	198	203	256	267	0

表 1.駒台上の持ち駒の価値

表２は、盤面上の駒の価値をまとめたものである。ここで、駒の価値はその可動性や玉との相対的な位置などで異なるが、ここでは玉との相対的な位置による駒の価値の平均についてまとめた。これを見ると、銀将の価値が他の駒に比べて低いもののその他の駒についてはおおむね常識的な評価値の順に並んでいることが分かる。

駒	歩兵	香車	桂馬	銀将	金将	角行	飛車	王将
価値	230	376	338	206	329	466	474	0
駒	と金	成香	成桂	成銀		龍馬	竜王	
価値	211	357	373	392		527	526	

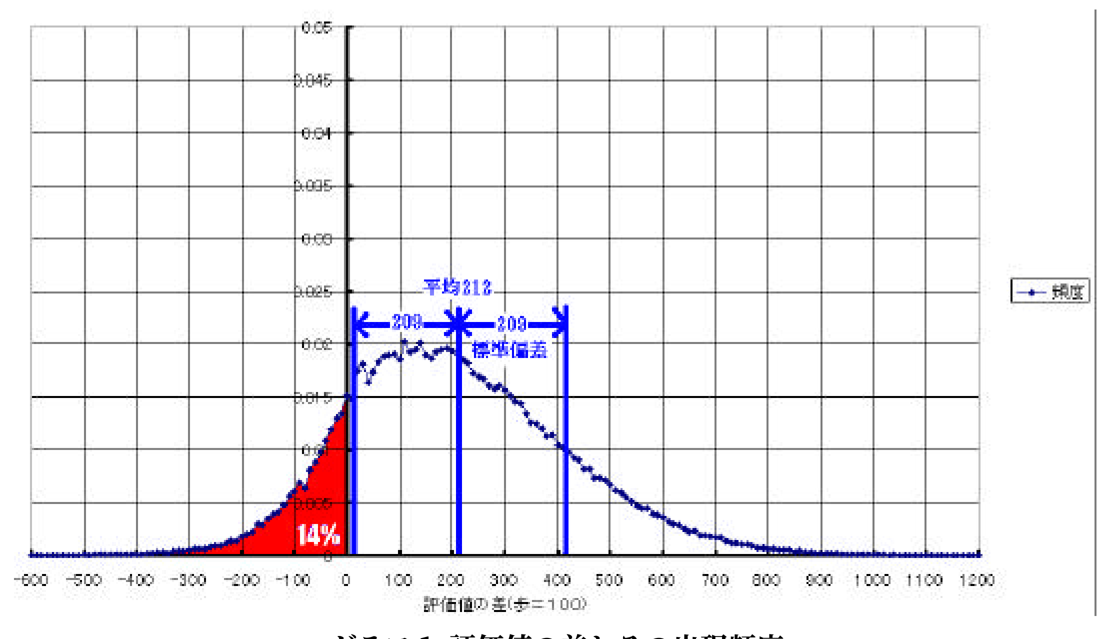
表 2.盤面上の駒の価値

6.2.得られた評価関数の性能について

判別分析によって求められた局面評価関数の重み列を使って、学習に使われていないプロ棋士の棋譜 91 対局を用いて局面評価関数の評価を行った。

グラフ１は、プロの指した手の特徴量ベクトルとその他の指し手の特徴量ベクトルの差分ベクトルを評価関数で評価した場合の値の出現頻度である。差分ベクトルの「評価値」の平均は 212 で標準偏差は 209 であり、標準偏差の範囲の中に入るような標準的な指し手についてはプロの指した手よりも評価値が低くなった。その一方で、差分ベクトルの「評価値」が負になる事例、すなわちプロの指した手よりもプロの選ばなかった指し手の得点のほうが高い事例が 14%あった。

なお、先ほどの 2000 対局分の棋譜から得られた 17,736,886 の事例のうちこの評価関数で評価した場合にプロの指し手の評価



グラフ１:評価値の差とその出現頻度

値との差が 212+209=412 点以上あるような事例を無視して判別分析を行って局面評価関数を求めたが、差分ベクトルの「評価値」が負の値になるような事例がやはり15%程残った。

ある局面の可能手を得られた評価関数によって評価して評価値によって降順に並べた場合、プロ棋士の選んだ指し手が何位に入っているのかをまとめたものが表 3 である。プロ棋士の選んだ指し手が最善手として選ばれる確率は 10%であったが、上位 10 手以内に這いいている確率は 68%、上位 20 手以内では 85%、上位 30 手以内には 94%の確率でプロ棋士の選んだ指し手が入っていた。可能手の中からプロの選んだ指し手を選び出すことは出来なかったが、多くの凡手の点数をプロの選んだ指し手の点数よりも低くすることには成功しているようである。

順位	1	10	20	30	40	50
プロ棋士の選んだ指し手の入っている確率	10%	68%	85%	94%	97%	98%

表 3:プロ棋士の指し手の入っている確率

6.3.人間の作った評価関数との対決

判別分析によって求められた局面評価関数と[1.はじめ]で述べたような人間が試行錯誤によって作り出した評価関数とを比較した。

比較には、東京農工大小谷研究室の将棋プログラムである Sexy-AI-Chan を利用して局面評価関数だけを本研究によって得られたものに入れ替えたものと元の局面評価関数を使用したもので先手後手を入れ替えて１０試合ずつ対戦させた。

その結果、どちらも５勝５敗となりえられた評価関数は人間が調整した評価関数と同等な強さを持っていることがわかった。

	A.元の評価関数	B.判別分析によって得られた評価関数
A 先手	５勝　５敗	５勝　５敗
B 先手	５勝　５敗	５勝　５敗
合計	１０勝１０敗	１０勝１０敗

表４:人間の作った評価関数との対戦

7.まとめ

本論文において、可能手の中からプロ棋士の選んだ指し手を選ぶような局面評価関数を作るために判別分析による局面評価関数の調整という手法を提案し、その実用性について実験・考察を行った。

本手法は、プロ棋士の選んだ指し手とプロ棋士の選ばなかった指し手の差分ベクトルに着目して、それらを局面の静的評価関数によって評価した時になるべくたくさん正の値になるようにすれば、局面的評価関数は棋譜上の局面についてプロ棋士の選んだ指し手を選ぶようになるという発想に基づくものである。

この発想に基づきプロ棋士の選んだ指し手を選ぶような局面評価関数を得るために、大量の棋譜からプロ棋士の選んだ指し手とその選ばなかった指し手の特徴量の差分ベクトルを得て、それらの線形和がなるべく多く正になるような線形結合を判別分析の方法を応用して得た。

得られた局面評価関数は、全てのプロ棋

士の選ばなかった指し手のうち 86％をプロ棋士が選んだ指し手より低く評価した。また、得られた評価関数が、プロの指し手を可能手の内最も高い指し手と評価する割合は 10%であったが、可能手を評価値の順に口授運に並べた場合にプロ棋士の選んだ指し手が、上位 10 手以内に入っている確率は 68%、上位 20 手以内では 85%、上位 30 手以内には 94%の確率でプロ棋士の選んだ指し手が入っていた。

そして、判別分析により得られた評価関数と人間の調整した Sexy-AI-Chan の評価関数を対戦させたところほぼ同等な強さを持っていることが分かった。

参考文献

[1]鈴木豪・小谷善行、 将棋の評価関数の非線形性 、 Game Programming Workshop in Japan '01、 2001、 pp124-131

[2]T.S.Anantharaman 、 Evaluation Tuning For Computer Chess ： Liner Discriminant Method、ICCA Journal Dec 1997、 1997

[3]D.F.Beal and M.C.Smith 、 Learning Piece Value Using Temporal Differences、ICCA Journal Vol.20 No.3 、 1997 、 pp147-151

[4]Tom E.Fawcett、Feature Discovery for Problem Solving Systems、 Doctor of Philosophy (Univ. of Massachusetts)、 1993

[5]堀田浩司・小林康幸、 重回帰分析による静的評価関数の計算 、 Game Programming Workshop in Japan '95、 1995、 pp187-196

[6]香山健太郎・井上博充、 将棋における差し手の絞込みの学習 、 Game Programming Workshop in Japan '96、 1996、 pp104-113