

주성분분석을 통한 바둑 포석 분석

이병두, 박종욱

세한대학교 체육학부

blee026@korea.com, pjw@sehan.ac.kr

Applying Principal Component Analysis to Go Openings

Byung-Doo Lee, Park Jong-Wook

Division of Sports Science, Sehan University

요 약

바둑은 2,500년 이상의 역사를 갖고 있지만 바둑에 대한 이론적 연구는 여전히 미흡하다. 최근에 인공지능을 활용하여 연구를 하고 있으나 현저한 이론적 실체를 제공하고 있지 못하다. 본 연구는 세계 최고의 바둑기사인 이세돌 프로 9단의 기보를 중심으로 바둑의 초반전인 포석을 분석하기 위하여 주성분분석을 적용하였다. 분석 결과 361개의 고유벡터 중 48개의 가장 큰 고유벡터가 전체 분산의 99.9% 정보를 수록하고 있으며, 전체 분산의 90.5% 정보를 30개의 가장 큰 고유벡터로 처리할 수 있음이 밝혀졌다. 이 결과는 향후 프로기사의 포석에 대한 패턴인식을 연구하는데 상당한 기여가 있을 것이다.

ABSTRACT

Although the history of the game of Go is more than 2,500 years, the theoretical studies of Go are still insufficient. In recent years a lot of studies using Artificial Intelligent(AI) have been conducted, but they do not provide the prominent theoretical reality. We applied Principal Component Analysis(PCA) to the professional Go openings, which are the early stage in Go, to analyze them especially focused on the Go game records of the professional 9-dan player Lee Sedol who is the world's top professional Go player. The results showed that among the 361 eigenvectors the 48 most significant eigenvectors capture most of the variance (99.9%) and the 30 most significant eigenvectors enable to possess 90.5 percent of the total variance. This result would be expected to considerably contribute to pattern recognition research of the professional Go openings in the near future.

Keywords : Go openings(바둑포석), Principal Component Analysis(주성분분석)

Received: Feb. 12, 2013 Accepted: Mar. 26, 2013
Corresponding Author: Byung-Doo Lee(Sehan University)
E-mail: blee026@korea.com

© The Korea Game Society. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ISSN: 1598-4540 / eISSN: 2287-8211

1. 서론

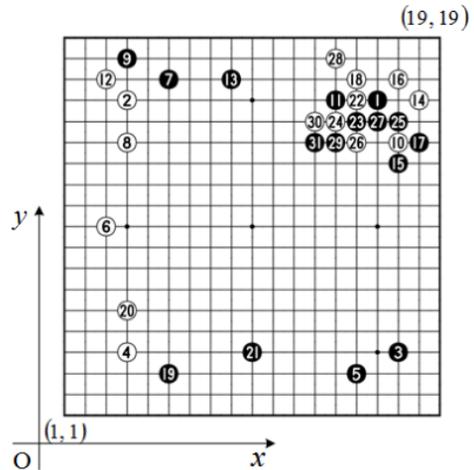
바둑은 인류 역사의 태동과 함께 대략 2,500년 전 중국에서 기원된 보드(board) 게임으로 추정하고 있다. 이 게임은 한국에서 ‘바둑(Baduk)’이라 불리고, 중국에서는 ‘웨이치(weiqi)’, 일본에서는 ‘이고(Igo)’라고 불리나, 전 세계적으로 통용되고 있는 용어는 ‘고(Go)’이다[1].

바둑 게임은 3단계, 즉 초반전, 중반전, 종반전으로 구분이 된다. 초반전은 30~50여 수로 된 포석이 되며 바둑의 골격을 만들어 가는 과정으로 전략(strategy) 개념에 대한 이해가 요구된다. 중반전은 적군의 돌을 공격하거나 아군의 돌을 수비하는 전투의 연속으로 전술(tactics) 개념의 이해가 필요하다. 마지막으로 끝내기로 불리는 종반전은 쌍방의 영역을 최종 정리하는 단계로 선·후수(sente and gote), 사활(life-and-death) 및 집의 크기 개념을 이해해야 한다.

이미 서구에서는 바둑을 흥미로운 신 연구영역으로 삼아 1990년대부터 본격적으로 연구를 진행하고 있으며, 주로 기계학습(machine learning), 복잡도이론(complexity theory), 조합게임이론(combinatorial game theory), 인지심리학(cognitive psychology) 분야에서 체계적으로 바둑을 다루고 있다[2]. 한 예로, 종반전인 끝내는 인공신경망(artificial neural network), 조합게임이론 등을 이용하여 이미 컴퓨터 바둑의 수준이 프로기사와 대등하게 되었으며[3,4], 중반전은 현재 본격적으로 연구되고 있는 몬테카를로 시뮬레이션(Monte-Carlo simulation)을 이용하여 상당한 아마 기력 수준까지 이르고 있다[5]. 한편 초반전인 포석에 대한 연구는 미진한 편이며, 주로 대용량으로 데이터베이스화된 패턴들로부터 패턴인식(pattern recognition)을 활용하여 다음 착점을 결정하는 초보적인 수준에 머물고 있다.

포석은 [Fig. 1]과 같이 보통 초반전의 50수 정도를 의미하며, 중반전인 전투 장면에 들어가기 전의 단계를 말한다. 포석은 바둑의 초석과 같은 것

으로 이를 잘 꾸미면 중반전 이후 바둑의 운영 및 전투에 있어 매우 유리해진다.



[Fig. 1] A numbered opening game(1~31)

[6]은 포석을 “대국 초반에 요소를 찾아 돌을 배치하는 일로서 3선 또는 4선에 돌을 집중 배치하여 실리와 세력의 조화를 꾀하는데 이는 중반 전투를 위해 대형을 갖추는 준비 과정”으로 정의를 하고 있다. 한편 [7]은 포석을 “영토의 기초를 형성하는 것과 관련된 활동을 포괄하는 과정”으로 보고 있다. 이와 같이 애매한 포석의 정의로 인하여 프로기사들도 포석을 정확히 설명하는데 매우 어려워하고 있다.

본 연구에서는 주성분분석(PCA)을 통해 평균포석에 대한 고유값(eigenvalue)과 고유벡터(eigenvector)의 특성과 주성분(principal component)의 개수를 살펴보고, 아울러 설문조사를 통해 평균포석에 대한 성향을 분석해 보았다.

2. 본론

2.1 실험 환경 및 자료 수집

2013년 현재 한국기원 소속 프로기사는 280명(남자 231명, 여자 49명)이며, 이 중 이세돌 9단은

지난 2012년 국내 랭킹 1위의 기사로 전 세계적으로 잘 알려진 천재 기사이다.

바둑 대국에 있어 첫 수는 무척 중요하다. 첫 수를 어디에 두느냐에 따라 한 판의 바둑에 있어 흑백간의 전략이 확연히 달라질 수가 있기 때문이다. 한 판의 바둑을 두기 전에 흑으로 첫 수를 놓을 것인가를 결정하기 위해 두 대국자는 둘 가르기를 한다. 이 후 흑을 켜 대국자는 첫 수로 자신이 구사하려는 포석을 전개하게 되는데, 이는 전략을 구가하는 포석의 성향을 분석하는데 중요한 역할을 한다. 이를 위하여 이세돌 9단이 2012년 국내의 공식 기전에서 치룬 92대국(한국기원 제공) 중에서 흑으로 대국을 치룬 50대국을 실험용 포석으로 선정하였다.

2.2 PCA 분석

[Fig. 1]과 같은 50개의 포석인 실험용 데이터는 주성분분석을 위해 좌표값 입력을 통해 i 번째 입력용 포석 P_i 로 변형이 되었으며, P_i 는 다음과 같은 $361(=19 \times 19)$ 개의 요소로 구성된 행벡터가 된다.

$$(p_{1,1}, p_{1,2}, \dots, p_{1,19}, p_{2,1}, \dots, p_{19,19})$$

여기서 특정좌표 (i, j) 에 흑돌이 있는 경우에는 $p_{i,j}$ 의 값이 1, 백돌이 있는 경우에는 -1, 둘이 없는 경우에는 0의 값으로 처리하였다.

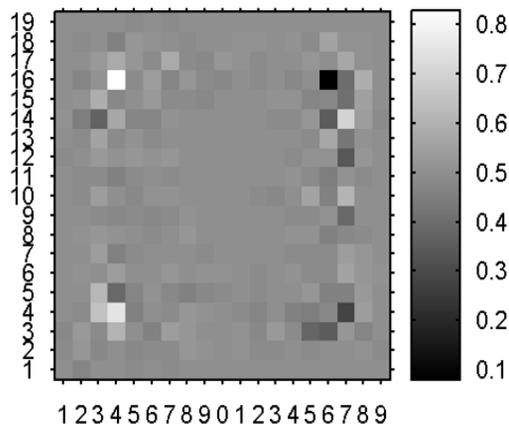
2.2.1 평균포석 분석

평균포석 μ 는 (eq. 1)과 같이 입력된 전체 포석 ($N=50$)에 대한 좌표값의 평균을 통해 구할 수 있다.

$$\mu = \frac{1}{N} \sum_{i=1}^N P_i \quad (\text{eq. 1})$$

여기서 P_i 는 입력된 i 번째의 포석이 된다.

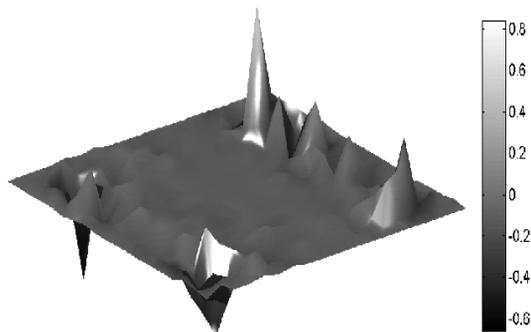
[Fig. 2]는 평균포석 μ 를 갖고 MATLAB의 imshow() 함수를 이용하여 흑백으로 그려낸 바둑판이 된다.



[Fig. 2] Grayscale Go board map of the mean opening

참고로 흑색이 짙을수록 흑 대국자가 더 선호하는 착점의 위치가 되며, 흰색이 짙을수록 백 대국자가 더 선호하는 착점의 위치가 된다.

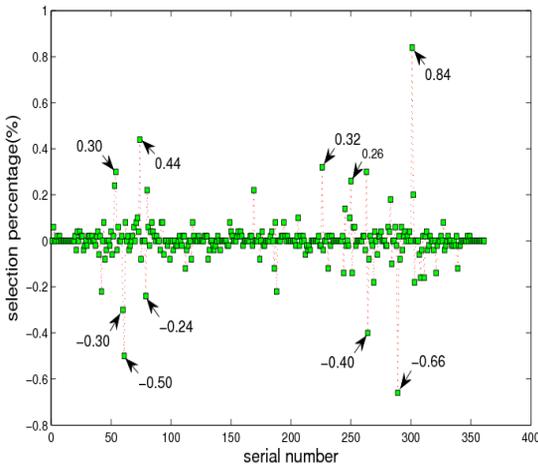
한편 [Fig. 3]은 평균포석 μ 를 갖고 MATLAB의 surf() 함수를 이용하여 3차원 공간상에 흑백으로 그려낸 바둑판이 된다.



[Fig. 3] 3-D shaded surface of the mean opening

[Fig. 3]을 살펴보면 흑 대국자는 우상귀, 우하귀, 우변의 순으로 착점을 많이 하는 것을 알 수 있으며, 반면에 백 대국자는 좌상귀와 좌하귀 순으로 착점을 주로 하는 것을 알 수 있다. 참고로 [Fig. 3]에서 우상귀에 있는 흑 착점의 최대값은 0.84이며, 좌상귀에 있는 백 착점의 최소값은 -0.66이 된다. 이를 달리 말하면 흑 대국자는 50판의 대

국 중에서 42판(=50×0.84)을 첫 수로 우상 화점 (16, 16)을 두고 있으며, 이에 대응하는 백대국자는 50판의 대국 중에서 33판(=50×0.66)을 좌상 화점(4, 16)을 두고 있는 것으로 확인이 되었다. 한편 특정좌표에 따른 착점의 선택률을 그려보면 [Fig. 4]와 같다.



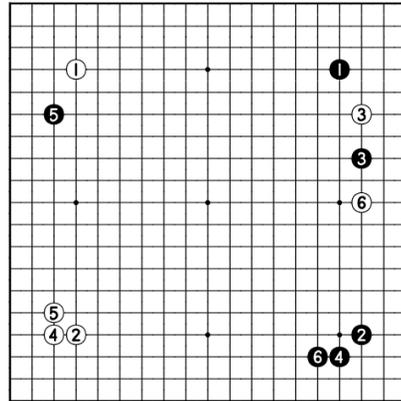
[Fig. 4] Selection percentage of the positions

[Fig. 4]에 있는 착점의 순위를 흑백으로 구분하여 표로 작성해보면 [Table 1]과 같다.

[Table 1] Selection percentage of the preferring positions ranked

player	rank	serial number	position (x,y)	selection percentage
black	1	301	(16,16)	84%
	2	74	(17, 4)	44%
	3	226	(17,12)	32%
	4	54	(16, 3)	30%
	5	250	(3,14)	26%
	6	53	(15, 3)	24%
white	1	289	(4,16)	66%
	2	61	(4, 4)	50%
	3	264	(17,14)	40%
	4	60	(3, 4)	30%
	5	79	(3, 5)	24%
	6	188	(17,10)	22%

[Table 1]에 있는 선호하는 착점을 기준으로 기보를 작성해보면 [Fig. 5]와 같다.



[Fig. 5] Preferring positions ranked for black and white players

[Fig. 5]를 살펴보면 이세돌 9단은 흑1, 2로 화점-소목포석1)을 선호하며, 이에 대적하는 백대국자는 백1, 2로 2연성 포석2)을 선호하는 것으로 나타났다. 이후 이세돌은 흑6과 같이 우하귀 소목으로 굳히기를 하여 실리를 추구하고 있으며, 이에 백대국자는 백6과 같이 갈라치기를 우선적으로 추구하는 예전의 방식이 아닌 백3으로 우선적으로 직접 걸쳐 들어가 전투를 유도하고 있다. 또한 백3의 걸침에 흑3과 같이 협공을 함으로서 공격을 선호하는 이세돌 9단의 기풍을 느낄 수가 있다. 아울러 흑5와 같이 넓은 쪽 그리고 자신의 진영이 아닌 상대방의 진영으로의 침투는 모양을 중시하는 바둑보다는 상대의 세력을 분할하는 바둑을 선호하고 있음을 알 수 있다.

결국 흑대국자인 이세돌 9단은 1) 화점-소목포석을 선호하고, 이에 대항하는 백대국자들은 2연성을 선호하며, 2) 이후 흑대국자는 소목을 활용하여 실리를 추구하고 있으며, 백대국자들은 화점에 있

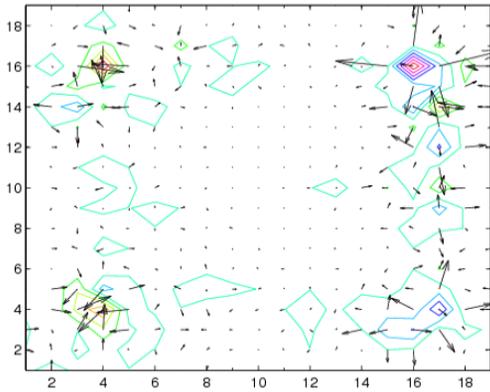
1) 화점-소목포석은 화점인 (16, 16)과 소목인 (17, 4)를 두는 포석을 칭함.

2) 2연성 포석은 두 개의 화점인 (4, 16)과 (4, 4)를 두는 포석을 말한다.

는 흑돌에 날일자 검침을 하여 적극적인 공격 성향을 보이고 있으며, 3) 이에 흑대국자는 한칸 협공으로 초반전에 두 대국자가 전투국면으로 몰입되는 것을 알 수 있다.

참고로 [Fig. 6]은 이변수함수인 평균포석 $F = \mu(x, y)$ 에 대해 (eq. 2)와 같은 MATLAB의 gradient() 함수와 quiver() 함수를 이용하여 그려낸 것이다.

$$\nabla F = \frac{\partial F}{\partial x} \hat{i} + \frac{\partial F}{\partial y} \hat{j} \quad (\text{eq. 2})$$



[Fig. 6] Gradient field of the function $F = \mu(x, y)$

바둑에 있어 하위지식(low-level knowledge)의 하나인 세력(influence)은 영토(territory)와 두터움(thickness)의 정도를 재는 척도로 사용된다. [Fig. 6]과 같은 경사도 지도는 향후 형세판단(board evaluation)을 위한 세력의 정도를 연구하는데 요긴하게 사용될 수 있다.

2.2.2 PCA를 통한 포석 분석

평균포석 μ 와 입력된 각 포석 P_i 간의 차인 $(P_i - \mu)$ 를 이용하여 361×361 차원인 공분산행렬(covariance matrix) C 를 (eq. 3)과 같이 구할 수가 있다.

$$C = \frac{1}{N} \sum_{i=1}^N (P_i - \mu)^T (P_i - \mu) \quad (\text{eq. 3})$$

여기서 $(P_i - \mu)^T$ 는 행렬 $(P_i - \mu)$ 의 전치행렬(transpose matrix)이다.

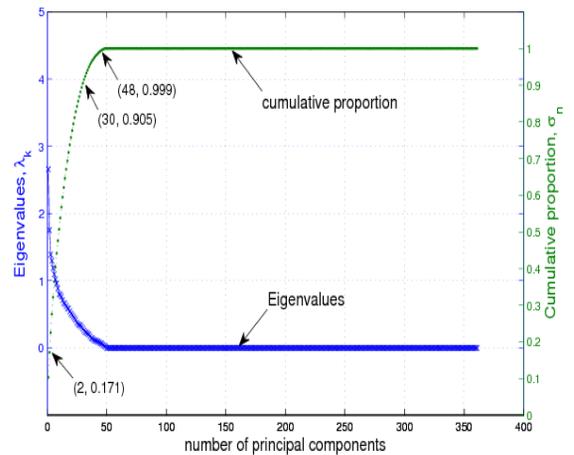
다음으로 수직고유벡터(orthogonal eigenvector) $E_k (k = 1, 2, \dots, 361)$, 공분산행렬 C 와 관련된 고유값 λ_k 를 (eq. 4)와 같이 구할 수가 있다.

$$CE_k = \lambda_k E_k, \quad \lambda_k \geq \lambda_{k+1} \quad (\text{eq. 4})$$

고유벡터는 공분산의 유의성(significance)의 순으로 정돈된 직교기저벡터의 집합이 된다. 고유값 λ_k 를 내림차순으로 정돈하였을 때 생성된 첫 번째 몇 개의 주성분(PC)들은 λ_k 가 큰 순서로 된 공분산행렬의 고유벡터가 된다. 이들은 입력된 포석의 대부분의 분산 정보를 포함하고 있다.

주성분분석의 목적인 입력 데이터의 차원을 줄이기 위해, 즉 c 차원을 ($n \ll c$)인 n 차원으로 줄이기 위해 고유값이 큰 첫 번째 몇 개의 PC를 선택할 수 있다[8,9]. 차원을 줄이기 위한 방법들이 있으나 대부분이 주먹구구식(rule of thumb)이다.

[Fig. 7]은 361개의 PC에 대한 k 번째 고유값 λ_k 와 (eq. 5)와 같은 전체 분산에 대한 누적 분산의 비인 σ_n 을 보여주고 있으며, 이는 고유벡터의 개수를 한정하는 지표로 사용된다.



[Fig. 7] Eigenvalues and cumulative proportion

$$\sigma_n = \frac{\sum_{i=1}^n \lambda_i}{\sum_{i=1}^c \lambda_i}, \text{ for } n = 1, 2, \dots, c \quad (\text{eq. 5})$$

여기서 $c (= 361)$ 는 전체 고유벡터의 수가 되며, n 은 선택된 PC들의 개수가 된다.

[Fig. 7]을 살펴보면 361개의 고유값 중 48개의 가장 큰 고유값으로 된 고유벡터가 전체 분산의 99.9% 정보를 수록하고 있음을 알 수 있으며, 전체 분산의 90.5%는 30개의 가장 큰 고유값으로 된 고유벡터로 처리할 수 있음을 알 수 있다. 아울러 두 개의 주축인 제1주축(PC1), 제2주축(PC2)으로 전체 분산의 17.1% 정보를 해결할 수도 알 수 있다. 결국 대략 30~40개의 가장 큰 고유값으로 된 고유벡터로 입력된 50개 포석의 전체 분산 대부분을 처리할 수 있음을 알 수 있다.

c 개로 된 원시 고유벡터 E 로부터 n 개의 주축을 선택하게 되면 다음과 같은 새로운 고유벡터 M 을 구할 수가 있다.

$$\begin{aligned} E &= (E_1, E_2, \dots, E_c) \\ &\quad \downarrow \\ M &= (E_1, E_2, \dots, E_n) \end{aligned}$$

(eq. 6)과 같이 새로운 고유벡터 M 을 이용하면 최종 포석 데이터 집합 F 를 구할 수가 있다.

$$F^T = M^T \bar{P}^T \quad (\text{eq. 6})$$

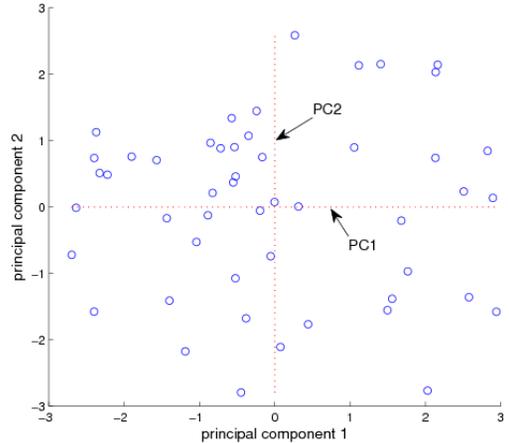
여기서 행렬 F^T , M^T 는 각각 F , M 의 전치행렬이 되며, \bar{P} 는 입력행렬 P 에서 평균벡터 μ 를 뺀 평균행렬이 된다.

한 예로 PC1, PC2로 된 고유벡터 $M = (E_1, E_2)$ 로 생성된 최종 포석 데이터 집합 F 를 갖고 2차원 평면상에 그려보면 [Fig. 8]과 같다.

2.3 평균포석에 대한 설문 조사

일반적으로 바둑기사들은 바둑판에 놓인 돌들을

살펴본 후, 기력에 비례하여 전체국면에 대한 형세, 수순, 큰 곳과 작은 곳, 급한 곳과 급하지 않은 곳 등을 찾아내는 탁월한 능력을 소유하고 있다.



[Fig. 8] 50 input Go openings along the two largest PCs

[Fig. 2]에 있는 평균포석의 배경 설명을 한 후, 평균포석에 대한 의견을 묻기 위해 [Table 2]와 같이 전문가 집단인 프로기사 3명(남자 3)과 아마기사 7명(남자 2, 여자 5)을 대상으로 15분간에 걸쳐 설문조사를 실시하였다. 참고로 프로기사의 기력은 4, 5단이었으며, 아마기사의 기력은 2~5단이였다.

[Table 2] General characteristics of surveyors

	type	case number	percent(%)
gender	male	5	50%
	female	5	50%
player	professional	3	30%
	amateur	7	70%
total		10	100%

설문내용은 4가지로 1) 포석기풍(세력/실리), 2) 포석형태(구형/신형), 3) 포석진행(안정적/공격적), 4) 포석흐름(매우빠름/매우느림) 등을 조사하였다.

설문내용에 대한 응답은 [Table 3]과 같이 의미

(semantic) 척도로 측정되었으며, 설문 분석을 위한 통계프로그램은 PSAW Statistics 18이 사용되었고, 신뢰도분석, 단순상관분석, 기술통계, 다중상관분석 등이 실시되었다.

[Table 3] Specified concepts corresponding to the elements

element	specified concept
pursuit	influence □□□□□ territory
shape	classic □□□□□ modern
progress	calm □□□□□ fighting
speed	very fast □□□□□ very slow

2.3.1 신뢰도분석

신뢰도분석은 조사 항목으로 된 그룹의 내부 일치성을 측정하는 것이다. 즉, 조사 대상을 유사한 측정도구로 여러 번 측정을 하거나, 한 가지 도구로 조사 대상을 반복 측정을 했을 때 질문에 대한 일관성 있는 응답을 하고 있는가를 조사하는 것이다. 주로 (eq. 7)과 같은 Cronbach's α 계수가 사용되며, 이상적인 알파값은 .70과 .90의 사이값이 된다[9].

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum_{i=1}^n \sigma_i^2}{\sigma_t^2} \right) \quad (\text{eq. 7})$$

여기서 n 은 항목수가 되고, σ_i^2 , σ_t^2 은 각각 항목 i 의 분산, 항목 전체의 분산이 된다.

신뢰도분석에서 변수 내부의 상관계수인 급내상관계수(intra-class correlation coefficient)는 여러 번의 측정에 따른 한 변수 내의 평균값에 있는 변화를 알려준다. [10,11]은 급내상관계수가 피어슨(Pearson) 상관계수와 비교할 때 더 적절한 재검사의 신뢰도의 지표라고 한다.

우리가 관심있는 것은 10명의 설문자들이 설문 항목에 대해 일관성있는 답변을 하였는가이다. 즉, 설문항목들에 대한 설문 값들의 차이가 없음을 나타내는 유의하지 않음을 찾는 것이다. 설문자들이

국내 바둑기사들 중에서 무작위로 추출되었기 때문에 모형은 이차원변량(two-way random)으로 선택하고, 설문항목에 대한 값이 비모수적데이터(non-parametric data)로 간주하여 Friedman 카이제곱(chi-square)검정을 실시하였다. 검정결과 [Table 4]와 같이 최종 연구 척도의 신뢰도는 Cronbach's α 계수가 .904로 설문자들의 답변이 매우 신뢰성이 높은 것으로 확인되었다.

[Table 4] Statistics for reliability

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	Number of items
.904	.889	10

한편 [Table 5]는 신뢰도를 높이기 위하여 설문자 10명에 대해 특정 설문자를 제거 시 발생하는 Cronbach's α 계수를 보여주고 있다. 참고로 첫 번째 열의 괄호에 있는 값은 설문자의 기력을 표시하고 있다. 즉 p는 프로그 기사를 나타내며, d는 아 마 기사를 나타내고 있다.

[Table 5] Item-total statistics

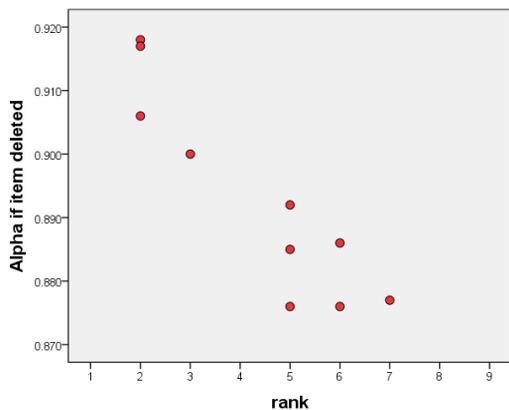
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Alpha if Item Deleted
item 1 (2d)	26.25	71.929	.164	.	.917
item 2 (5d)	25.50	54.286	.922	.	.876
item 3 (2d)	25.88	70.982	.187	.	.918
item 4 (4p)	25.75	55.643	.933	.	.876
item 5 (5p)	25.88	52.982	.894	.	.877
item 6 (3d)	26.00	63.714	.574	.	.900
item 7 (5d)	25.63	62.839	.740	.	.892
item 8 (2d)	26.00	66.857	.458	.	.906
item 9 (4p)	25.50	56.571	.791	.	.886
item 10 (5d)	25.25	54.500	.800	.	.885

2.3.2 단순상관분석

단순상관분석은 독립변수와 종속변수의 개념없이 한 변수와 또 다른 변수가 어떻게 변화하는가의 정도와 변화를 살펴보는 것이다. [Table 5]에 있는 기력과 (항목제거 후) 알파값과의 상관관계를

알아보기 위하여 단순상관분석을 실시하였다.

현재 국내에서는 아마기사와 프로기사 간에는 두 점의 접바둑으로 바둑 대국이 진행되기 때문에 프로기사에게는 2단계의 급수를 상향시켜 기력을 재조정하였다. 예를 들어 프로 5단(5p)은 아마 7단(7d)으로 급수를 상향시켰다. 이후 기력에 따른 알파값을 점도표로 그려보면 [Fig. 9]와 같이 음의 관계가 있음을 알 수 있다.



[Fig. 9] Alpha vs. rank plot

[Fig. 9]에서 보인 음의 관계가 과연 유의성을 갖고 있는가를 확인하기 위하여 단순상관분석을 실시하였으며, 분석 후 상관계수 보고서는 [Table 6]과 같았다.

[Table 6] Correlation between player's rank and its Alpha

		rank	Alpha if Item Deleted
rank	Pearson Correlation	1	-.921**
	Sig. (2-tailed)		.000
	N	10	10
Alpha if Item Deleted	Pearson Correlation	-.921**	1
	Sig. (2-tailed)	.000	
	N	10	10

** . Correlation is significant at the 0.01 level (2-tailed).

[Table 6]에 있는 상관계수 보고서를 살펴보면 뛰어난 상관계수가 $r = -.921$ 로 상당한 음의 관계

가 있음을 알 수 있다. 또한 양측검정용 유의확률값을 살펴보면 그 값이 $p = .000$ 으로 $p < .01$ 이 되어 '기력과 (항목제거 후) 알파값과는 강한 음의 상관관계가 있음'을 알 수 있다. 즉, [Fig. 2]의 평균포석에 대한 분석을 함에 있어 기력이 높은 프로는 매우 신뢰성있는 답을 하는 반면에, 기력이 낮은 아마는 비교적 신뢰성이 낮은 답을 하고 있다'라는 새로운 사실이 밝혀졌다.

2.3.3 기술통계

기술통계는 수치자료를 체계적으로 정리/요약해 주며, 주요 요약 내용으로는 빈도분포, 중앙경향치, 분산도 등이 된다.

[Fig. 2]의 평균포석에 대한 설문내용 4개에 대한 설문자 10명의 답을 요약해 보면 [Table 7]과 같다.

[Table 7] Descriptive statistics

	N	Mean	Std. Deviation
Q1: pursuit	10	3.50	.527
Q2: shape	10	3.40	.843
Q3: progress	10	3.50	.707
Q4: speed	10	2.70	.949
Valid N (listwise)	10		

[Table 7]에서 보듯이 1) Q1: 포석기풍(세력/실리)에 대해서는 $\mu = 3.50$ 으로 다소 실리를 추구하고 있으며, 2) Q2: 포석형태(구형/신형)에 있어서도 $\mu = 3.40$ 으로 다소 신형 포석을 선호하는 것을 조사되었다. 또한 3) Q3: 포석진행(안정적/공격적)에 있어서도 $\mu = 3.50$ 으로 다소 공격적인 성향을 나타내고 있음으로 조사되었으나, 4) Q4: 포석흐름(매우빠름/매우느림)에 있어서는 $\mu = 2.70$ 으로 다소 느린 포석을 구가하고 있음으로 조사되었다.

결국 이세돌 9단은 '다소 신형포석을 구가하며 다소 공격적이며 느리게 포석을 진행하면서 다소 실리를 챙기려는 의도'로 조사되었다.

2.3.4 다중상관분석

다중상관분석은 두 개의 변수가 아닌 여러 개의 변수간의 상관관계를 구하는 것이다.

설문자들이 답한 4개의 설문 항목값 간에 상관관계가 있는가를 알아보기 위하여 다중상관분석을 실시해 보았더니, 그 결과는 [Table 8]과 같았다.

[Table 8] Correlations among 4 variables

		Q1	Q2	Q3	Q4
Q1: pursuit	Pearson Correlation	1	-.500	-.447	-.111
	Sig. (2-tailed)		.141	.195	.760
	N	10	10	10	10
Q2: shape	Pearson Correlation	-.500	1	.745*	-.250
	Sig. (2-tailed)	.141		.013	.486
	N	10	10	10	10
Q3: progress	Pearson Correlation	-.447	.745*	1	-.414
	Sig. (2-tailed)	.195	.013		.234
	N	10	10	10	10
Q4: speed	Pearson Correlation	-.111	-.250	-.414	1
	Sig. (2-tailed)	.760	.486	.234	
	N	10	10	10	10

*. Correlation is significant at the 0.05 level (2-tailed).

[Table 8]에 있는 Q2: 포석형태(구형/신형)와 Q3: 포석진행(안정적/공격적)을 살펴보면, 유의확률값이 $p = .013$ 으로 $p < .05$ 이 되어 유의함을 알 수 있으며, 피어슨 상관계수가 $r = .745$ 로 강한 양의 관계를 이루고 있음을 알 수 있다.

결국 10명의 설문자를 통하여 [Fig. 2]에 있는 이세돌 9단의 평균포석을 분석해 보면 포석형태(구형/신형)와 포석진행(안정적/공격적)에는 강한 양의 관계가 있음이 밝혀졌다. 즉, '신형포석을 쓸수록 공격적인 성향을 보이며, 구형포석을 쓰면 안정적인 성향을 보인다'라는 바둑계에서 통용되는 일반적인 특성을 보여주고 있다.

3. 결론 및 제언

1990년대 이후부터 시작된 바둑에 대한 연구는

주로 컴퓨터 바둑을 제작을 위한 것이나, 아직까지 제작된 컴퓨터 바둑은 아마 수준에 머물고 있다. 국내에서는 이에 대한 연구가 거의 전무하며 그나마 유럽이나 미주에 있는 대학교 및 연구소에서 연구되고 있는 학문적인 연구는 주로 끝내기 및 형세판단에 머물고 있다. 이는 바둑에 대한 연구가 결코 쉽지 않다는 것을 반증하는 것이고, 향후 연구해야 할 대상과 폭이 크다는 것을 시사하고 있다.

본 연구에서는 전 세계적으로 거의 연구되고 있지 않은 바둑대국에서의 초반전인 포석에 대해 분석을 하였다. 실험대상으로는 2012년 국내 랭킹 1위의 기사로 전 세계적으로 잘 알려진 천재 기사인 이세돌 9단의 포석 중, 흑으로 둔 50개의 기보를 중심으로 이루어졌으며, 주성분분석 및 설문조사를 통해 평균포석에 대한 연구를 실시하였다.

3.1 연구결과

주성분분석 중의 산물인 평균포석을 살펴보면 1) 이세돌 9단은 화점-소목포석을 선호하고, 이에 대항하는 백대국자들은 2연성을 선호하며, 2) 이후 흑대국자는 소목을 활용하여 실리를 추구하고 있으며, 백대국자들은 날일자 검침을 통해 적극적인 공격 성향을 보이고 있으며, 3) 이에 흑대국자는 한 칸 협공으로 초반전에 전투국면으로 몰입되는 것을 알 수 있었다.

주성분분석을 위한 361개의 고유벡터 중 48개의 가장 큰 고유값으로 된 고유벡터가 전체 분산의 99.9% 정보를 수록하고 있음을 알 수 있었으며, 전체 분산의 90.5% 정보는 30개의 고유벡터로 처리할 수 있음도 알 수 있었다. 결국 대략 30~40개의 고유벡터로 50개 입력된 포석의 분산 대부분을 처리할 수 있음이 밝혀졌다.

또한 이세돌 9단의 평균포석을 분석함에 있어 다소 소수인 전문가 10명에 대한 신뢰도검사를 실시하였으며, 검사결과 α 계수가 .904로 설문자들의 답변이 매우 신뢰성이 높은 것으로 확인되었다.

기력과 α 계수와와의 단순상관분석에서는 프로그사가 아마기사보다 더 신뢰성있게 평균포석을 분석

하고 있다라는 새로운 사실이 밝혀졌다.

전반적으로 이세돌 9단의 평균포석 분석결과에 있어 이세돌은 다소 신형포석을 구가하며 다소 공격적이며 느리게 포석을 진행하여 다소 실리를 추구하는 것으로 최종 조사되었다.

마지막으로 설문항목값 간에 상관관계가 있는가를 알아보았더니, 신형포석을 쓸수록 공격적인 성향을 보이며, 구형포석을 쓰면 안정적인 포석진행을 한다는 일반적인 사실이 재확인되었다.

3.2 제언

일반적으로 주성분분석은 패턴인식을 위해 사용되는 하나의 방법이며, 본 연구에서 보여준 평균포석이나 방법론은 향후 포석에 대한 패턴인식을 위한 초석이 될 것이다. 또한 [Fig. 6]에서 보여준 경사도 지도는 향후 세력의 정도를 연구하는데 요긴하게 사용될 수 있다. 마지막으로 다소 소수로 구성된 10명의 전문가 집단에 대해 평균포석을 설문조사하였으며, 조사 결과 나타난 프로그시사와 아마기사의 시각차는 향후 바둑에 대한 설문조사를 실시함에 있어 차이를 두어 설문조사하여야 함을 간접적으로 시사하고 있는 것이다.

ACKNOWLEDGMENTS

This work was conducted under the research grant of Sehan University in 2013. In addition, the authors thank Yong-Sung Kim, the third year student in Sehan University, for his helping to create the input Go openings from the professional Go game records.

REFERENCES

- [1] Wikipedia, "Go", from [http://en.wikipedia.org/wiki/Go_\(game\)](http://en.wikipedia.org/wiki/Go_(game)), 2013.
- [2] Byung-Doo Lee, "Multi-strategic Learning, Reasoning and Searching in the game of Go", PhD thesis, Auckland University, pp. 16-22, 2005.
- [3] N. Sasaki, Y. Sawada, and J. Yoshimura, "A Neural Network Program of Tsume-Go", *Computer and Games*, pp. 167-182, 1999.
- [4] M. Müller, "Decomposition search: A combinatorial games approach to game tree search, with applications to solving Go endgames", *IJCAI-99*, Vol. 1, pp. 578-583, 1999.
- [5] T. Cazenave and B. Helmstetter, "Combining tactical search and monte-carlo in the game of go", *CIG'05*, pp. 171-175, 2005.
- [6] In-Man Kim, "Dictionary of Baduk Terms", Seolim Publishing Co., 1999.
- [7] Soo-Hyun Jeong, "Understanding of Modern Baduk", Nanam Publishing House, 2004.
- [8] B. D. Lee and Y. W. Keum, "Candidate First Moves for Solving Life-and-Death Problems in the Game of Go, using Kohonen Neural Network", *Journal of Korea Game Society*, Vol. 9, No. 1, pp. 105-114, 2009.
- [9] K. J. Chun, K. S. Oh, and S. H. Cho, "Photomosaics Using Principal Component Analysis", *Journal of Korea Game Society*, Vol. 11, No. 1, pp. 139-146, 2011.
- [10] N. Nikos, "A Step-by-Step Guide to SPSS for Sport and Exercise Studies", Routledge, 2006.
- [11] N. Walliman, "Statistics in Kinesiology", *Human Kinetics*, 2nd edition, 2001.



이 병 두 (Lee, Byung doo)

1982 한양대 원자력공학 학사
1991 서강대 정보처리학 석사
2005 Auckland University 컴퓨터공학 박사
2012-현재 세한대 체육학부 바둑학과 조교수

관심분야: 컴퓨터공학, 인공지능, 컴퓨터바둑



박 종 욱 (Park, Jong wook)

현재 세한대학교 체육학부 교수

관심분야: 운동생리학, 바둑게임
