한국통계학회논문집
2010, 17권, 6호, 917–925

DOI: 10.5351/CKSS.2010.17.6.917

일반화 정준상관 행렬도와 프로크러스티즈 분석을 응용한 대한테니스협회 등록 선수의 체격요인, 체력요인 및 기초기술요인에 대한 분석연구

최태훈[a], 최용석[1,b]

[a]안동과학대학 스포츠레저과, [b]부산대학교 통계학과

요 약

일반적으로 정준상관 행렬도(canonical correlation biplot)는 정준상관분석에서 두 변수집단에 의해서 측정된 다변량 자료에서 변수 집단 간의 관계와 개체들의 관계를 탐색하기 위한 2차원 그림이다. 최근에 이를 활용하여 최태훈과 최용석 (2008)은 2006년도 한국여자골프협회(KLPGA) 선수에 대한 기술요인 변수군과 경기성적요인 변수군간의 관련성을 살펴보았고 최태훈 등 (2009)은 테니스 그랜드 슬램대회 선수특성요인과 경기요인에 대한 분석을 하였다. 더군다나 세 변수군 이상의 정준상관분석을 일반화 정준상관분석(generalized canonical correlation analysis)이라 하며 이와 관련하여 허명회 (1999, 6장)는 수량화 플롯을 제안하고 있다. 이를 행렬도의 의미에서 일반화 정준상관 행렬도(generalized canonical correlation biplot)라 하자. 본 연구에서는 대한테니스협회(KTA)에 등록된 남자선수들 중 상위50명의 체격요인, 체력요인 및 기초기술요인에 대한 분석을 일반화 정준상관 행렬도를 적용하여 살펴보고 프로크러스티즈 분석을 통하여 전체선수, 상위랭킹과 하위랭킹 선수간의 행렬도 형상비교를 시도 하였다.

주요용어: 일반화 정준상관 분석, 행렬도, 프로크러스티즈 분석.

1. 서론

일반화 정준상관 행렬도(canonical correlation biplot)는 정준상관분석(canonical correlation analysis)에서 세 변수군 이상에 의해 측정된 다변량 자료에서 변수군 간의 관계와 개체들의 관계를 탐색하기 위한 2차원 그림이다. 이는 일반적으로 최용석 (2006, 1장)의 한 변수군의 행렬자료에 대한 일반적인 행렬도를 세 변수 군의 행렬자료로 확장한 것으로 볼 수 있다. 국내에선 Park과 Huh (1996a, 1996b)가 정준상관분석에서 수량화 방법(quantification method)의 관점에서 2차원 그림을 제안하였고 이를 정준상관 행렬도라 하였으며, 세 변수 집단 이상인 경우까지 확장한 정준상관분석의 일반화를 시도하여 이를 허명회 (1999, 6장)에 수량화 플롯으로 잘 요약하고 있다.

일반적으로 행렬도는 복잡한 다변량 분석의 결과를 보다 쉽게 파악할 수 있기 때문에 최근 여러 분야에서 이에 대한 활발한 연구와 응용이 이루어지고 있다. 행렬도는 Gabriel (1971)에 의해서 개발되었고 국내에선 Choi (1991)가 이것의 저항버전을 개발하면서 처음으로 소개하였고, 허명회 (1993, 5장)가 국내에서 Biplot을 행렬도라 처음 불렀다. 더군다나 최용석 등 (2005a, 2005b)은 다변량 분산분석 모형에서 행렬도의 활용을 제안하고 응용의 예를 보였다. 최용석 등 (2005c), Choi 등 (2005a, 2005b) 그리고 Choi와 Kim (2008)은 군집분석을 이용하여 행렬도에 대한 활용의 폭을 넓혀왔다. 특히, 최근에 체육분야의 측정 평가에서 정준상관 행렬도의 응용을 시도하여 최태훈과 최용석 (2008)은 한국여자골프

이 논문은 2010년 교육과학기술부 교육역량강화사업비에 의한 연구과제 임.
[1] 교신저자: (609-735) 부산 금정구 장전동 산30, 부산대학교 통계학과, 교수. E-mail: yschoi@pusan.ac.kr

협회(KLPGA) 선수를 대상으로 기술요인 변수군과 경기성적요인 변수군 간의 관련성을 살폈고 최태훈 등 (2009)은 테니스 그랜드 슬램대회 선수특성요인과 경기요인에 대한 분석을 하였다.

본 연구에서는 2003년 3월부터 2004년 8월까지 전국대회 단체전 4강 이상 입상한 대학, 실업팀에 소속된 시니어 남자 테니스 선수들 중 2004년 대한테니스협회(KTA)에 등록된 랭킹 100위권 이내의 선수 50명을 대상으로 일반화 정준상관 행렬도를 통해 8개의 체격요인 변수군(신장, 체중, 상완장, 전완장, 수장, 상지장/신장 비율, 전완장/상지장 비율(브로컬지수), 피하지방 합(삼두근, 견갑골하단, 이두근, 장골능, 대퇴, 하퇴))과 8개의 체력요인 변수군(악력, 윗몸일으키기, 하버드스텝, 윗몸앞으로굽히기, 50m달리기, 제자리높이뛰기, 눈감고외발서기, 전신반응시간) 그리고 기초기술요인 변수군(그라운드 스트로크, 서비스, 발리)의 상호 연관성을 살펴보려 한다. 이를 위해서 2절에서 세 변수군에 대한 일반화 정준상관 행렬도와 행렬도 형상 비교를 위한 프로크러스티즈 분석을 소개하고 3절에서는 실제 분석의 예를 보이고자 한다. 4절에서는 프로크러스티즈 분석을 활용하여 전체선수와 상위랭킹, 하위랭킹별 행렬도의 형상 차이를 비교하고자 한다.

2. 일반화 정준상관 행렬도와 프로크러스티즈 분석

2.1. 일반화 정준상관 행렬도

이 절에서는 허명회 (1999, 6장)를 참고로 일반화 정준상관 분석과 관련된 수량화 플롯인 정준상관 행렬도의 대수적인 면을 간단히 요약하기로 하자. 일반화 정준상관분석은 여러 변수군 사이의 관계를 분석하는 다변량분석기법으로 여기서는 세 변수군에 한정하여 설명하기로 하자. 먼저 p_j개 변수로 이루어진 j번째 변수군에 의한 n개의 개체에 대해서 측정한 크기가 $n \times p_j$인 표준화된 자료행렬을 $\mathbf{X}_j,\ j = 1,2,3$라 하자. 이들에 의해서 이루어진 크기가 $n \times p$인 결합자료행렬을 $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$라 하면 $p = \sum_{j=1}^{3} p_j$이고 $n > \max(p_1, p_2, p_3)$이다.

크기가 $p_j \times 1$인 임의의 벡터 $\mathbf{b}_j,\ j = 1,2,3$에 의한 각 변수군의 선형결합을 $\mathbf{X}_1\mathbf{b}_1$, $\mathbf{X}_2\mathbf{b}_2$, $\mathbf{X}_3\mathbf{b}_3$이라 하면, 변수군 간의 상관을 최대화하기 위하여 이들이 서로 유사해야 한다. 이를 위해서 $\sum_{j=1}^{3}\sum_{k=1}^{3} \|\mathbf{X}_j\mathbf{b}_j - \mathbf{X}_k\mathbf{b}_k\|^2/n$을 최소화하는 계수벡터 $\mathbf{b}_j$를 찾으면 된다. 이는 제약조건 $\sum_{j=1}^{3}\mathbf{b}_j'\mathbf{X}_j'\mathbf{X}_j\mathbf{b}_j/n = 3$을 두고 $\sum_{j=1}^{3}\sum_{k\neq j}\mathbf{b}_j'\mathbf{X}_j'\mathbf{X}_k\mathbf{b}_k/n$을 최대화하는 계수벡터 $\mathbf{b}_j$를 찾는 것과 동일하다.

이를 위해서 라그랑지 승수법을 활용하면 일반화 정준계수 벡터와 일반화 정준상관을 대수적으로 더 쉽게 제공하는 식 (2.1)과 같은 스펙트럼 분해(spectral decomposition)

$$\frac{\mathbf{D}^{-\frac{1}{2}}(\mathbf{X}'\mathbf{X})\mathbf{D}^{-\frac{1}{2}}}{\sqrt{n}} = \mathbf{V}\mathbf{D}_{1+\mu}\mathbf{V}' \tag{2.1}$$

를 얻게 된다. 여기서 $\mathbf{D} = \mathrm{diag}(\mathbf{X}_1'\mathbf{X}_1/n, \mathbf{X}_2'\mathbf{X}_2/n, \mathbf{X}_3'\mathbf{X}_3/n)$는 크기가 $p \times p$인 블록 대각행렬이고, 크기가 $p \times r$인 고유벡터 $\mathbf{V}$는 직교행렬로 $\mathbf{V}'\mathbf{V} = \mathbf{V}\mathbf{V}' = \mathbf{I}_p$이며 $\mathbf{D}_{1+\mu} = \mathrm{diag}(1+\mu_1, \ldots, 1+\mu_r)$은 $1+\mu_1 \geq \cdots \geq 1+\mu_r$관계를 갖는 고유값을 대각원소로 하는 대각행렬이다.

일반화 정준상관 분석에서 특히, k번째 일반화 정준상관 $\mu_k/2$에 대응하는 일반화 정준계수 벡터는 $\mathbf{b}_{jk} = \sqrt{3}\mathbf{D}^{-1/2}\mathbf{v}_{jk},\ j = 1,2,3,\ k = 1,\ldots,r$이다. 이들에 의해서 s차원의 일반화 정준상관 행렬도를 위한 좌표는 각 변수군의 자료행렬 $\mathbf{X}_j,\ j = 1,2,3$에 대하여 행좌표 행렬과 열좌표 행렬은

$$\mathbf{R}_{\mathbf{X}_j} = \mathbf{X}_j\mathbf{B}_{j(s)}\mathbf{D}_{\mu(s)}^{\frac{1}{2}}, \qquad \mathbf{C}_{\mathbf{X}_j} = \mathbf{B}_{j(s)}\mathbf{D}_{\mu(s)}^{\frac{1}{2}}, \quad j = 1,2,3 \tag{2.2}$$

이고 여기서 $\mathbf{B}_{j(s)} = (\mathbf{b}_{j1}, \ldots, \mathbf{b}_{js})$는 크기가 $p_j \times s$인 일반화 정준계수 행렬이고, $\mathbf{D}_{\mu(s)} = \mathrm{diag}(\mu_1, \ldots, \mu_s)$는 $\mu_1 \geq \cdots \geq \mu_s$ 관계를 갖는 $\mu_k, k = 1,\ldots,s$를 대각원소로 하는 대각행렬이다.

식 (2.2)에서 제공되는 s차원의 일반화 정준상관 행렬도의 근사도는

$$\frac{\sum_{k=1}^{s}\lambda_k}{\sum_{k=1}^{r}\lambda_k}\times 100\%$$

이고 여기서 $\lambda_k = (\mu_k/2)^2,\ k = 1,\ldots,r$이다.

2.2. 프로크러스티즈 분석

프로크러스티즈 분석(procrustes analysis)이란 기하적 공간상에서 형상점(landmarks)에 의해서 나타낸 개체들의 형상을 측정하고, 기술하며 비교하는 형상분석(shape analysis)에서 개체간의 형상비교를 하기 위해 한 개체를 다른 개체 쪽으로 적합시키는 방법이다 (최용석과 현기홍, 2006; Choi 등, 2005a, 2005b; 최태훈 등, 2009). 이 분석에 대한 알고리즘을 최태훈 등 (2009)의 것을 재인용하기로 하자. X_r과 X_s를 개체간의 형상을 나타내는 크기가 $k\times m$인 중심화 형상좌표행렬이라 하면 벡터 t와 직교행렬 R에 의해서 X_r의 l번째 점 $\mathrm{x}_{r(l)}$의 좌표점을 변환한 $\mathrm{Rx}_{r(l)}+\mathrm{t}$을 고려하자. 그러면 X_s의 l번째 점 $\mathrm{x}_{s(l)}$과 $\mathrm{Rx}_{r(l)}+\mathrm{t}$간의 제곱거리합

$$\sum_{l=1}^{k}\left(\mathrm{x}_{s(l)}-\mathrm{Rx}_{r(l)}-\mathrm{t}\right)^T\left(\mathrm{x}_{s(l)}-\mathrm{Rx}_{r(l)}-\mathrm{t}\right) \tag{2.3}$$

을 생각할 수 있다. 식 (2.3)을 최소화 하는 R과 t를 찾는 것이 두 형상 X_r과 X_s가 잘 일치되도록 하게 하는 정보를 제공한다. 실제로 비정칙치분해(singular value decomposition) $\mathrm{X}_r{}^T\mathrm{X}_s = \mathrm{U}\Lambda\mathrm{V}^T$를 이용하여 $\widehat{\mathrm{R}} = \mathrm{VU}^T$과 $\hat{\mathrm{t}} = 0$를 제공받을 수 있고 이를 이용하여 두 형상의 일치성을 평가하는 측도(measure)인 프로크러스티즈 통계량

$$\mathrm{PS}\left(\mathrm{X}_r,\mathrm{X}_s\right) = \mathrm{tr}\left(\mathrm{X}_r^T\mathrm{X}_r\right)+\mathrm{tr}\left(\mathrm{X}_s^T\mathrm{X}_s\right)-2\mathrm{tr}\left(\Lambda\right) \tag{2.4}$$

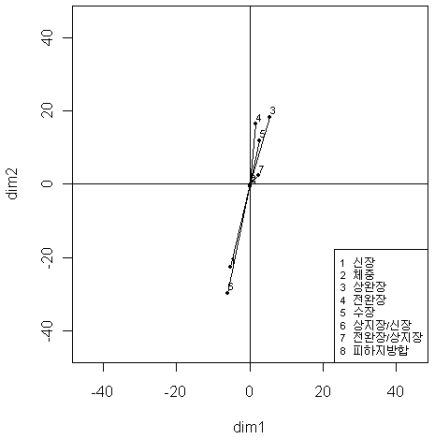
을 얻을 수 있다. 식 (2.4)의 프로크러스티즈 통계량 값이 0이면 두 형상이 일치한다고 평가할 수 있다.

3. 테니스 선수의 체격요인, 체력요인, 기초기술요인 분석

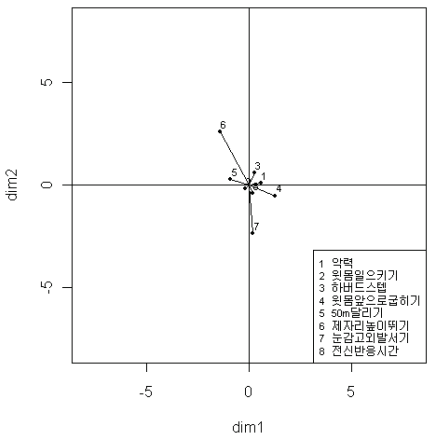
3.1. 전체자료 비교

2절에서 요약한 일반화 정준상관분석 행렬도를 응용하기 위한 자료는 2003년 3월부터 2004년 8월 15일전까지 전국대회 단체전 4강 이상 입상한 대학, 실업팀에 소속된 시니어 남자 테니스 선수들 중 2004년 대한테니스협회(KTA)에 등록된 랭킹 100위권 이내의 선수 50명에 대한 체격요인 변수군, 체력요인 변수군과 기초기술요인 변수군으로 이루어진 세 변수군에 의해 측정된 것이다. 이미 최태훈 (2004)은 이 자료에 대해서 요인분석과 경로분석을 통해 세 요인변수군 간에 관계를 살펴 본 바 있다. 그러나 이 연구는 세 요인변수군에 대하여 두 요인변수군 별로 짝을 지어 비교하고 있어 이들을 동시에 비교하는 방법이 요구 된다. 본 연구에서는 최근 2절에서 설명한 세 변수군 이상의 경우에 대한 일반화 정준상관 행렬도를 통해 시각적으로 세 요인변수군 간의 관계를 동시에 규명하는 새로운 방법을 시도하려 한다.

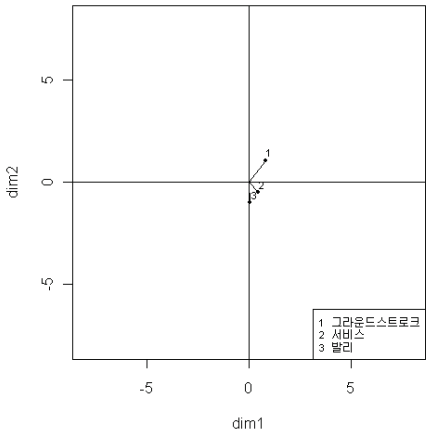
본 연구에서는 최태훈 (2004)의 분석을 통해 중요한 8개의 체격요인 변수군(신장, 체중, 상완장, 전완장, 수장, 상지장/신장 비율, 전완장/상지장 비율, 피하지방합)과 8개의 체력요인 변수군(악력, 윗몸일으키기, 하버드스텝, 윗몸앞으로굽히기, 50m달리기, 제자리높이뛰기, 눈감고외발서기, 전신반응시간), 그리고 기초기술요인 변수군(그라운드 스트로크, 서비스, 발리)인 세 변수군을 고려하기로 하자.



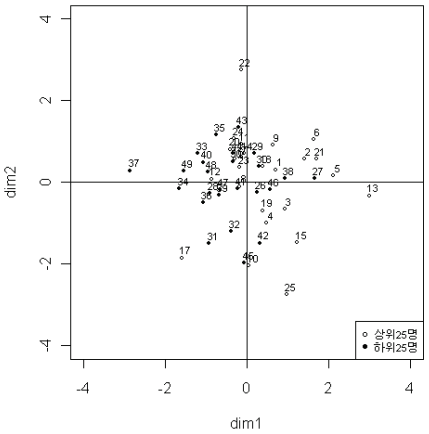
(a) 체격요인 변수군



(b) 체력요인 변수군



(c) 기초기술요인 변수군



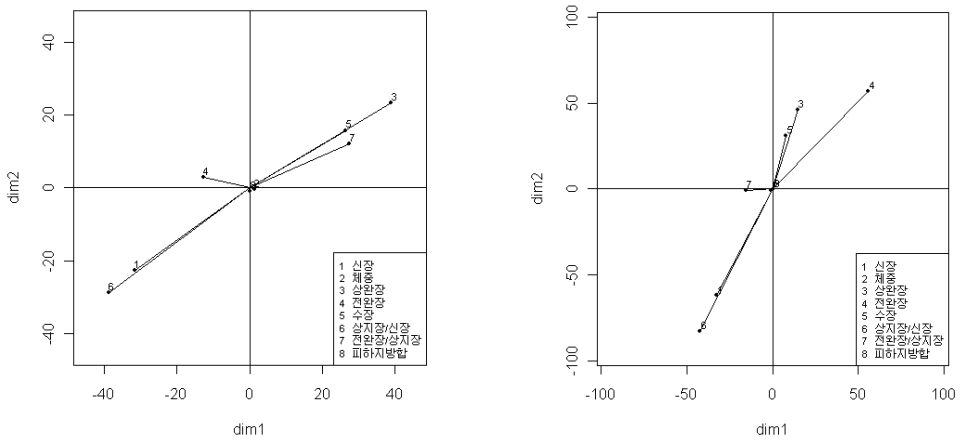
(d) 전체선수개체

그림 1: *전체자료에 대한 일반화 정준상관 행렬도*

이러한 체격, 체력, 기초기술의 측정항목들은 테니스선수들을 대상으로 가장 일반적으로 사용되는 측정항목들로 특정 스포츠종목에 적합한 이상적인 체격 등이 존재한다면 동일한 훈련, 생리적인 능력을 가진 선수들을 설명하는데 해당종목에 참가하는 선수의 체격 등의 그 자체로 설명될 수 있을 것이다.

먼저 그림 1의 (a), (b), (c)는 전체 선수자료에 대한 각각 체격요인 변수군, 체력요인 변수군, 기초기술요인 변수군에 대한 2차원 일반화정준상관 행렬도이고 (d)는 (c) 기초기술요인 변수군에 의한 전체 선수의 행렬도이다. 제 1, 2, 3 일반화 정준상관이 0.65, 0.47, 0.36으로 근사도에서 제 1과 제 2 정준축이 54%와 29%로 따라서 전체 적합도는 83%로 2차원 행렬도가 원자료를 잘 설명함을 나타내고 있다. 그림 1(d) 기초기술변수요인에 대한 전체선수개체에 대한 행렬도로 제 1 정준축(dim1)을 중심으로 오른편에 상위랭킹 선수들이 포진해 있고 왼편이 하위랭킹 선수들이 포진해 있다. 이러한 경향은 (a) 체격요인 변수군과 (b) 체력요인 변수군에 대해서도 같은 결과를 제공하므로 이들은 생략하였다.

일반적으로 정준상관 행렬도에서는 변수군간 비교를 통해 변수들의 연관성을 파악하는 것이 주된 것이나 개체군집과의 관련성을 살펴보는 것도 탐색적 분석관점에서 요구되는 점이다. 실제로 (d)



(a) 상위랭킹 선수자료 (b) 하위랭킹 선수자료

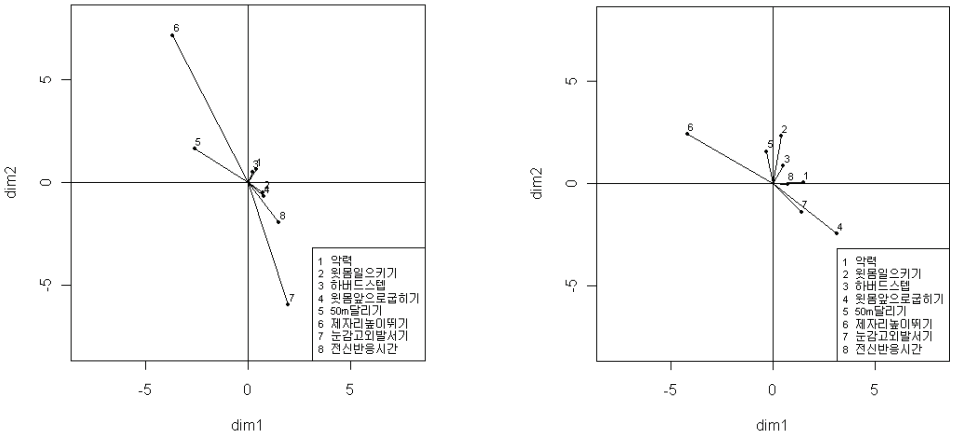
그림 2: *랭킹별 체격요인 변수군과 선수개체에 대한 일반화 정준상관 행렬도*

전체선수개체 행렬도와 변수군 행렬도 (a), (b), (c)를 서로 비교하여 상위랭킹과 하위랭킹 선수들의 특성을 파악하려 한다. 그림 1(a) 체격요인 변수군의 행렬도에서 제 1 정준축(dim1)은 왼편에는 신장(신장(1)과 상지장/신장(6))과 오른편에는 팔과 손의 길이(상완장(3), 전완장(4), 수장(5), 전완장/상지장(7))를 잘 구분하고 있다. 그림 1(b) 체력요인 변수군의 행렬도 제 1 정준축은 오른편에 전문체력(악력(1), 하버드스텝(3), 윗몸앞으로 굽히기(4), 눈감고외발서기(7), 전신반응시간(8))을 나타내는 변수와 왼편에는 순발력(제자리높이뛰기(6), 50m달리기(5), 윗몸일으키기(8))을 나타낸다. 그림 1(c) 기초기술요인 변수군의 행렬도에서 제 1 정준축의 오른편에 그라운드스트로크(1), 서비스(2), 발리(3) 세 가지 모두 있어 기초기술을 나타내는 축이다. 지금까지 살펴 본 전체 행렬도 (a), (b), (c), (d)를 비교 요약하자면 (a) 체격요인 변수군에서 팔과 손의 길이(상완장(3), 전완장(4), 수장(5), 전완장/상지장(7))는 (b) 체력요인 변수군에서 전문체력(악력(1), 하버드스텝(3), 윗몸앞으로 굽히기(4), 눈감고외발서기(7), 전신반응시간(8))을 나타내는 변수와 (c) 기초기술요인 변수군의 모두와 연관성이 높으며 특히, 이들은 (d) 전체선수에서 상위랭킹 선수들에게 요구되는 요인변수들임을 알 수 있다.

특히, 최태훈 (2004)에서는 상위랭킹과 하위랭킹 선수 간 차이를 잘 드러내는 영향력 있는 변수로 체격요인 변수군에서 상완장, 전완장, 상지장/신장 비율, 전완장/상지장 비율을, 체력요인 변수군에서는 악력, 하버드스텝, 윗몸앞으로 굽히기, 제자리높이뛰기, 눈감고외발서기를, 기초기술요인 변수군에서는 그라운드스로크, 서비스, 발리 세 변수 모두가 해당됨을 보여 주고 있다. 이는 일반화 정준상관 행렬도의 결과와 일치하며 더군다나 이들 모두는 상위랭킹에서 높은 수치를 나타내는 중요한 요인변수임을 보여 주고 있다.

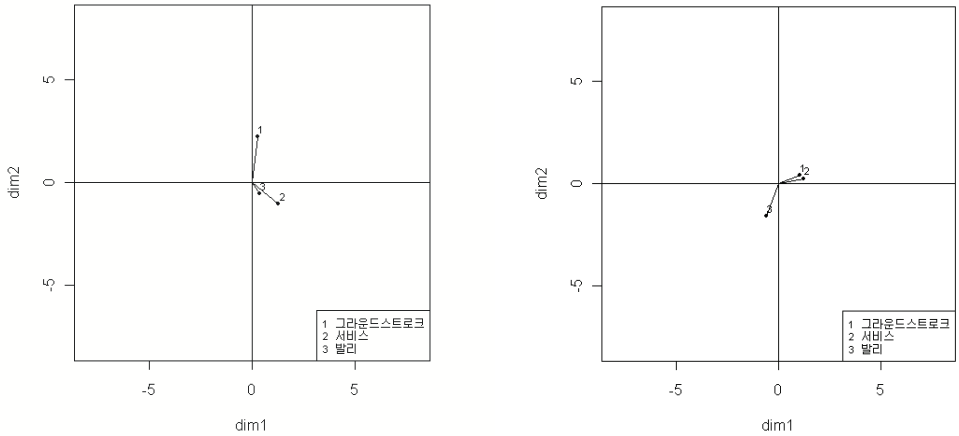
3.2. 랭킹별 비교

테니스 선수들의 상위랭킹(25명)과 하위랭킹(25명)에 대한 체격요인 변수군, 체력요인 변수군, 기초기술요인 변수군의 일반화 정준상관 행렬도를 통한 차이를 비교하기로 하자. 그림 2는 체격요인 변수군에 대한 (a) 상위랭킹 선수자료와 (b) 하위랭킹 선수자료의 일반화 정준상관 행렬도이다. 이들 2차원 행렬도에 대한 전체 적합도는 각각 77%와 81%로 원자료를 잘 근사하고 있고 이는 일반화 정준상관 행렬도의 특성상 그림 3 체력요인 변수군과 그림 4 기초기술요인 변수군의 행렬도의 적합도이기도 하다.



(a) 상위랭킹 선수자료 (b) 하위랭킹 선수자료

그림 3: *랭킹별 체력요인 변수군에 대한 일반화 정준상관 행렬도*



(a) 상위랭킹 선수자료 (b) 하위랭킹 선수자료

그림 4: *랭킹별 기초기술요인 변수군에 대한 일반화 정준상관 행렬도*

먼저 상위랭킹 선수자료와 하위랭킹 선수자료의 체격요인 변수군에 대한 일반화 정준상관 행렬도 그림 2의 (a)와 (b)에서 제 1 정준축에 대한 해석은 그림 1(a)와 대동소이하며 이축은 왼편에는 신장(신장(1)과 상지장/신장(6))과 오른편에는 팔과 손의 길이를 잘 구분하고 있다. 특히, 신장(1)과 상지장/신장(6), 상완장(3)과 수장(5)은 각각 랭킹에 관계없이 매우 상관이 높은 변수들이다. 다만, 상위랭킹에서 왼편의 전완장(4)이 하위랭킹에서 좌편의 전완장/상지장(7)이 영향을 미치는 중요변수로 여겨진다.

다음으로 상위랭킹 선수자료와 하위랭킹 선수자료의 체력요인 변수군에 대한 일반화 정준상관 행렬도 그림 3의 (a)와 (b)에서 제 1 정준축은 각각 그림 1(b)와 대동소이하여 오른편에 전문체력(악력(1), 하버드스텝(3), 윗몸앞으로 굽히기(4), 눈감고외발서기(7), 전신반응시간(8))을 나타내는 변수와 왼편에는 순발력(제자리높이뛰기(6), 50m달리기(5), 윗몸일으키기(8))을 나타낸다. (a) 상위랭킹에서는 주로 왼편의 순발력 보다는 오른편의 전문체력이 중요한 요인변수들이고 (b) 하위랭킹에서는 그 반대의 경우이다.

표 1: 행렬도의 형상비교를 위한 프로크러스티즈 통계량

	체격요인			체력요인			기초기술요인		
	전체	상위	하위	전체	상위	하위	전체	상위	하위
전체	0.00			0.00			0.00		
상위	52.05	0.00		2.06	0.00		3026.15	0.00	
하위	24.84	60.22	0.00	2.27	6.99	0.00	11697.65	11536.43	0.00

그림 4는 기초기술요인 변수군에 대한 (a) 상위랭킹 선수자료와 (b) 하위랭킹 선수자료의 일반화 정준상관 행렬도이다. (a)의 상위랭킹의 경우 제 1정준축이 그림 2(c)와 같이 오른편에 그라운드스트로크(1), 서비스(2), 발리(3) 세 가지 모두 있어 기초기술을 나타내는 축으로 해석이 된다. 그러나 (b) 하위랭킹의 경우 제 1 정준축은 왼편의 발리(3)와 오른편(그라운드스로크(1), 서비스(2))를 구분하는 축으로 해석이 가능하며 발리(3)가 하위랭킹 선수들에게는 특징적인 요인변수임을 보여 주고 있다. 특히, 상위랭킹에서는 서비스와 발리가 하위랭킹에서는 그라운드스트로크와 서비스가 서로 상관이 높음을 보여 주고 있다.

3.3. 행렬도의 형상비교를 위한 프로크러스티즈 분석

다음으로 표 1은 지금까지 살펴 본 체격요인 변수군, 체력요인 변수군, 기초기술요인 변수군 별 전체자료, 상위랭킹, 하위랭킹 행렬도의 형상을 비교하기 위한 프로크러스티즈 통계량을 제공하고 있다. 먼저 체격요인 변수군에서 전체자료와 하위랭킹의 행렬도가 통계량 24.84로 제일 변동(variability)이 적게 나타났는데 이는 그림 2(a) 상위랭킹의 경우 중요한 팔과 손의 길이 요인변수인 상완장(3)과 수장(5), 전완장/상지장(7)의 분산이 상대적으로 커서 그림 2(b) 하위랭킹보다 더 중요한 요인변수로 작용했기 때문이다. 체력요인 변수군의 행렬도 비교에서는 상위랭킹과 하위랭킹 간에 다소 변이가 인정되는데 이는 상위랭킹에 중요한 요인변수인 전문체력(악력(1), 하버드스텝(3), 윗몸앞으로 굽히기(4), 눈감고외발서기(7), 전신반응시간(8))의 경향과 분산이 하위랭킹에서 다소 다른 양상을 보여주고 있기 때문이다. 기초기술요인 변수군에서는 전체선수와 상위랭킹는 상대적으로 전체선수와 하위랭킹, 상위랭킹과 하위랭킹에 비해 변동이 적다. 이는 그림 1(c), 그림 4(a)와 (b)의 비교에서 살펴 본 내용과 일치한다.

4. 결론

2004년 대한테니스협회(KTA)에 등록된 랭킹 100위권 이내의 남자선수 50명을 대상으로 일반화 정준상관 행렬도를 통해 체격요인 변수군과 체력요인 변수군 그리고 기초기술요인 변수군의 일반화 정준상관 행렬도는 이들 세 변수군 간의 연관성을 동시에 그림으로 잘 보여 주었다. 특히, 랭킹별 행렬도 분석에서는 상위랭킹에서는 체격요인 변수군의 경우 상완장, 수장이 영향을 미치는 중요변수로 여겨지며, 체력요인 변수군에서는 전문체력(악력, 하버드스텝, 윗몸앞으로 굽히기, 눈감고외발서기, 전신반응시간)관련 변수가, 기초기술요인 변수군에서는 그라운드스트로크, 서비스와 발리가 매우 중요하며 특히, 서비스와 발리가 상관이 매우 높음을 보여 주고 있다. 이러한 행렬도상에서 살펴 본 각 변수군 연관성 형상 차이점을 프로크러스티즈 분석을 이용한 행렬도의 형상비교를 통해 유사한 경향을 살펴보았다.

참고 문헌

최용석 (2006). <행렬도 분석>, 기초과학 총서 2권, 부산대학교 기초과학연구원.

최용석, 강창완, 김경덕 (2005c). 명목형 다항반응 로지스틱회귀모형의 행렬도 분석, *Journal of The Korean Data Analysis Society,* **7**, 839–849.
최용석, 현기홍 (2006). <통계적 형상분석의 이해와 응용 -프로크러스티즈 분석의 저항성 버전 연구 및 개발->, 자유아카데미, 서울.
최용석, 현기홍,정수미 (2005a). 다변량 분산분석에서 추정된 모수행렬의 행렬도, *Journal of The Korean Data Analysis Society,* **7**, 851–858.
최용석, 현기홍, 정수미 (2005b). MANCOVA Biplot, <한국통계학논문집>, **12**, 705–712.
최태훈 (2004). 테니스선수의 체격, 체력, 기초기술과 랭킹간의 관계, 국민대학교, 박사학위논문.
최태훈, 최용석 (2008). 정준상관 행렬도와 군집분석을 응용한 KLPGA 선수의 기술과 경기성적요인에 대한 연관성분석. <응용통계연구>, **21**, 429–439.
최태훈, 최용석, 신상민 (2009). 테니스 그랜드슬램 대회의 선수 특성요인과 경기요인에 대한 분석연구-정준상관 행렬도와 프로크러스티즈 분석의 응용-, <응용통계연구>, **22**, 855–864.
허명회 (1993). <統計相談의 이해>, 자유아카데미, 서울.
허명회 (1999). <다변량수량화>, 자유아카데미, 서울.
Choi, Y. S. (1991). *Resistant Principal Component Analysis, Biplot and Correspondence Analysis,* Unpublished Ph.D. Dissertation, Department of Statistics, Korea University.
Choi, Y. S., Hyun, G. H. and Kim, J. G. (2005a). A numerical comparison of map variability in SCA Using the Procrustes Analysis, *Journal of the Korean Data Analysis Society,* **7**, 1531–1538.
Choi, Y. S., Hyun, G. H. and Yun, W. J. (2005b). Biplots' variability based on the Procrustes analysis, *Journal of the Korean Data Analysis Society,* **7**, 1925–1933.
Choi, Y. S. and Kim, H. Y. (2008). Applications of cluster analysis in biplots, *The Korean Communications in Statistics,* **15**, 65–76.
Gabriel, K. R. (1971). The biplot graphics display of matrices with applications to principal component analysis, *Biometrika,* **58**, 453–467.
Park, M. and Huh, M. H. (1996a). Canonical correlation biplot, *The Korea Communications in Statistics,* **3**(1), 11–19.
Park, M. and Huh, M. H. (1996b). Quantification plots for several sets of variables, *Journal 1973of the Korea Statistical Society,* **25**, 599–601.

2010년 10월 접수; 2010년 10월 채택

A Study on the Relationship between Physique, Physical Fitness and Basic Skill Factors of Tennis Players in the Korea Tennis Association Using the Generalized Canonical Correlation Biplot and Procrustes Analysis

Tae-Hoon Choi[a], Yong-Seok Choi[1,b]

[a]Department of Sport and Leisure Studies, Andong Science College
[b]Department of Statistics, Pusan National University

Abstract

The canonical correlation biplot is a 2-dimensional plot for graphically investigating the relationship between two sets of variables and the relationship between observations and variables in the canonical correlation analysis. Recently, Choi and Choi (2008) suggested a method for investigating the relationship between skill and competition score factors of KLPGA players using this biplot. Choi *et al.* (2010) used this biplot to analyze the player characteristic factors and competitive factors of tennis Grand Slam competition. Moreover, Huh (1999) provided a generalized canonical correlation analysis and biplot for more than three sets of variables. A Procrustes analysis is a useful tool for comparing shapes between configurations. This study will provide a method to investigate the relationship between physique, physical fitness and basic skill factors of tennis players in the Korea Tennis Association using a generalized canonical correlation biplot and Procrustes analysis.

Keywords: Biplot, generalized canonical correlation analysis, tennis players, Procrustes analysis.

This work was supported by the College Education Capacity Reinforcement Grant 2010 of Ministry of Education, Science and Technology.

[1] Corresponding Author: Professor, Department of Statistics, Pusan National University, Busan 609-735, Korea.
E-mail: yschoi@pusan.ac.kr