

## Study on analysis with partial least square path modeling using multiple factor analysis

Ri-Ra Park<sup>a</sup> · Eun-Kyung Lee<sup>a,1</sup>

<sup>a</sup>Department of Statistics, Ewha Womans University

(Received January 23, 2018; Revised March 1, 2018; Accepted April 3, 2018)

---

### Abstract

In this paper, we examine the methodology to predict consumer preferences using several groups of attributes of products and application to real data. In the food industry, studies are in progress to investigate the relationship between product attributes and consumer preferences; consequently, various methodologies are proposed. Among these methodologies, we consider multiple factor analysis (MFA). The result of the MFA enable the division of consumers into four clusters with similar liking and the defining of preference characteristics for each cluster. Also, using the results of multiple factor analysis, we find the partial least squares path model to predict consumer preferences through the characteristics of the product and the characteristics evaluated by consumers. We can understand the relationship between the cluster of consumers and the preferred/undesirable characteristics of products through the partial least squares path model applied to two clusters with different liking. When multiple factor analysis is used in the partial least squares path model, it is possible to investigate relationships between products and consumers by analyzing product characteristics and consumer preferences simultaneously. The results can be applied to product developments and sales which makes this methodology important and useful.

Keywords: multiple factor analysis, partial least square, path analysis, sensory evaluation

---

### 1. 서론

제품의 특성을 통해 소비자들의 기호도를 예측하는 것은 제조업 분야의 대표적인 관심사이다. 소비자들이 선호하는 특성을 파악하여 제품을 개발하고 홍보하는 것은 이윤으로 직결되기 때문이다. 이러한 문제는 식품 이외에도 뷰티, 영화 등 다양한 상품과 서비스에 적용될 수 있다. 제품의 특성과 소비자들의 기호도를 분석하기 위하여 다양한 통계적 방법들이 이용되고 있다. 본 논문에서는 이들 방법들을 살펴보고 소비자별로 다양한 취향을 갖는 향수에 대하여 수집된 자료에 적용시켜보고자 한다.

선행 연구에는 여러 가지 요인 분석 방법을 이용하여 제품의 관능적인 특성을 시각화하고 소비자의 기호도와 연관시키는 연구 (Jee 등, 2008), 소비자들의 기호도를 군집화하여 군집별 선호 특성을 살펴보는 연구 (Lee와 Lee, 2008) 등이 이루어졌다. 본 논문에서는 제품의 감각 속성과 기호도를 사용한 다중

---

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2015R1D1A1A01056790).

<sup>1</sup>Corresponding author: Department of Statistics, Ewha Womans University, 52, Ewhayeodae-gil, Seodaemun-gu, Seoul, 03760, Korea. E-mail: [lee.eunk@ewha.ac.kr](mailto:lee.eunk@ewha.ac.kr)

요인분석의 결과를 이용하여 부분 최소제곱 경로 모형을 적합하는 방법 (Pages와 Tenenhaus, 2001)을 살펴보고 이를 향수 자료에 적용시켜보고자 한다. Pages와 Tenenhaus는 오렌지 주스에 대하여 교육을 받은 패널이자 소비자인 식품영양학과 학생들로부터 측정된 감각속성자료와 소비자 기호도 자료, 그리고 실험실에서 측정된 오렌지 주스에 대한 화학적 특성자료를 이용하였다. 감각속성자료와 화학적 특성자료를 이용하여 다중요인분석을 실시하고, 이 결과를 바탕으로 소비자와 감각속성자료, 화학적 특성자료를 군집화한 후 이를 바탕으로 세운 최소제곱경로모형을 추정하여 소비자들의 기호를 예측하였다. 본 연구에서 이용하는 향수 자료는 전문가 패널들로부터 측정된 감각속성과 소비자들로부터 측정된 감각속성, 그리고 소비자 기호도 자료로 이루어져 있어 기존의 연구자료와는 달리 비슷한 감각속성에 대하여 두 집단으로부터 측정된 자료를 포함하고 있다. 본 연구에서는 이러한 자료의 특성을 고려하여 기존 연구의 다중요인분석을 이용한 부분 최소제곱 경로 모형을 자료에 맞게 변형하여 소비자를 군집화하고 추가적으로 소비자가 평가한 제품 특성을 바탕으로 기호도를 예측해보고자 한다.

2절에서는 다중요인분석과 부분 최소제곱 경로 모형에 대하여 살펴본다. 3절에서는 12가지 향수에 대하여 수집된 특성들과 소비자의 기호도를 모형에 적용, 분석하여 향수 제품의 감각 속성과 소비자가 지각하는 특성, 그리고 기호도 간의 관계에 대해 깊이 있게 살펴본다. 마지막 절에서는 분석 결과에 대하여 정리하고 논의를 하고자 한다.

## 2. 다중요인분석을 이용한 경로모형

### 2.1. 다중요인분석

다중요인분석(multiple factor analysis; MFA) (Pages, 2014)은 몇 개의 그룹으로 이루어진 변수들의 집합들을 이용하여 자료의 특성을 설명하는 분석 방법으로 각 변수들의 그룹은 정량적, 혹은 정성적인 변수들로 이루어져 있으며 하나의 변수 그룹은 같은 형태의 변수들로 이루어져 있다. 다중요인분석은 이들 변수 그룹들을 변수의 특성에 따라 차원축소방법을 적용하여 자료를 각 그룹을 대표하는 몇 개의 요인들로 요약하고 이들 요인들을 이용하여 자료를 살펴보는 과정으로 차원을 축소하여 고차원상의 관측들과 변수들 간의 관계를 낮은 차원에서 시각화하여 알아보기 쉽게 설명할 수 있다는 장점이 있다. 설명변수들이 모두 정량적 변수인 경우에는 주성분 분석(principal component analysis)의 확장으로 볼 수 있다.

$\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_G]$ 는 제품의 특성을 나타내는 표준화된 자료로  $g$ 번째 변수 그룹인  $I \times n_g$  행렬  $\mathbf{X}_g = [X_{g,1}, X_{g,2}, \dots, X_{g,n_g}]$ ,  $g = 1, \dots, G$ 로 구성되어 있다. 여기서  $X_{g,j}$ 는  $I$ 차원의 벡터로  $g$ 번째 변수 그룹 중  $j$ 번째 변수를 나타내며,  $I$ 는 고려하고 있는 제품의 수를, 그리고  $n_g$ 는  $g$ 번째 변수 그룹에서의 변수 개수를 나타낸다. 표준화된 변수 그룹들의 집합인  $\mathbf{X}$ 를 그대로 다중요인분석에 이용하는 경우 각 변수 그룹은 그룹 내 변수의 개수에 따라 전체 다중요인분석에 끼치는 영향력이 달라지므로 분석 전에 변수 그룹에 따라 적절한 가중치를 부여하여 각 변수 그룹의 중요도를 달리하여야 한다. 가장 일반적으로 쓰이는 방법은 각 그룹내의 첫 번째 요인의 분산이 1이 되도록 가중치를 부여하는 것이다. 즉,  $\lambda_{g,1}$ 을  $g$ 번째 변수 그룹의 가장 큰 특성값(eigenvalue)라고 할 때 각 변수그룹마다 가중치를  $1/\sqrt{\lambda_{g,1}}$ 로 하여 각 그룹내 요인의 가장 큰 분산을 1이 되도록 조정해준다.

$\mathbf{X}$ 와 임의의 벡터  $Z$ 와의 관계를 나타내는 값  $L(Z, \mathbf{X})$ 을 아래와 같이 정의하자.

$$L(Z, \mathbf{X}) = \sum_{g=1}^G \sum_{j=1}^{n_g} \text{corr} \left( Z, \frac{X_{g,j}}{\sqrt{\lambda_{g,1}}} \right)^2. \quad (2.1)$$

다중요인분석은 차원 축소를 위하여  $\mathbf{X}$ 와의 상관관계를 최대화 하는 벡터  $Z_1, \dots, Z_K$ 를 찾는 과정으로

$\text{var}(Z_k) = 1, k = 1, \dots, K, \text{cor}(Z_{k_1}, Z_{k_2}) = 0, k_1 \neq k_2$ 의 조건하에서 식 (2.1)을 최대화 하는 벡터를 찾는다. 이 과정을 통하여  $P = \sum_{g=1}^G n_g$  차원 공간상의 원자료를  $K$ 차원상에 표현할 수 있게 된다. 일반적으로 시각적으로 표현이 가능하고 간단하지만 자료의 많은 부분에 대한 설명이 가능한  $K = 2$  또는  $K = 3$ 을 이용한다. 식 (2.1)은 행렬을 이용하여 아래와 같이 정리될 수 있다.

$$\begin{aligned} L(Z, \mathbf{X}) &= \sum_{g=1}^G \frac{1}{I^2} Z^T \mathbf{X}_g \mathbf{M}_g \mathbf{X}_g^T Z \\ &= \frac{1}{I^2} Z^T \mathbf{X} \mathbf{M} \mathbf{X}^T Z, \end{aligned} \quad (2.2)$$

여기서  $\mathbf{M} = \text{diag}(\mathbf{M}_1, \dots, \mathbf{M}_G), \mathbf{M}_g = 1/\lambda_{g,1} \mathbf{I}$ 이다. 식 (2.2)로부터 다중요인분석은 각 변수 그룹의 가중치를 다르게 적용한 주성분 분석이라는 것을 파악할 수 있게 된다. 그러므로 모든 변수 그룹을 총괄하는 주성분을  $F_1, \dots, F_K$ 라고 할 때  $F_k = \sqrt{\lambda_k} Z_k$ 로 표현할 수 있다. 이 때  $\lambda_k$ 는  $\mathbf{X}^*$ 의 분산 공분산행렬의  $k$ 번째 특성값을 나타낸다.  $F_k$ 와  $Z_k$ 의 관계는 다음과 같이 정리될 수 있다.

$$\begin{aligned} Z_k &= \frac{1}{\lambda_k} \frac{1}{I} \mathbf{X} \mathbf{M} \mathbf{X}^T Z_k, \\ F_k &= \frac{1}{\lambda_k} \frac{1}{I} \mathbf{X} \mathbf{M} \mathbf{X}^T F_k \\ &= \frac{1}{\lambda_k} \frac{1}{I} \sum_{g=1}^G \frac{1}{\lambda_{g,1}} \mathbf{X}_g \mathbf{M}_g \mathbf{X}_g^T F_k. \end{aligned} \quad (2.3)$$

이를 이용하여  $g$ 번째 변수 그룹의  $k$ 번째 정준변수(canonical variable)  $F_{k,g}$ 를 아래와 같이 정의한다.

$$\begin{aligned} F_{k,g} &= \frac{1}{\lambda_k} \frac{1}{\lambda_{g,1}} \mathbf{X}_g \mathbf{X}_g^T F_k, \\ F_k &= \sum_{g=1}^G F_{k,g}. \end{aligned} \quad (2.4)$$

$F_{k,g} = [F_{k,g,1}, F_{k,g,2}, \dots, F_{k,g,I}]$ 라고 할 때 식 (2.4)로부터 아래의 관계를 파악할 수 있다.

$$F_{k,g,i} = \frac{1}{\sqrt{\lambda_k}} \frac{1}{\lambda_{g,1}} \sum_{j=1}^{n_g} \text{cor}(X_{g,j}, F_k) X_{g,j,i}. \quad (2.5)$$

이는  $g$ 번째 변수 그룹의  $k$ 번째 정준변수 중  $i$ 번째 관측에 대한 값은  $g$ 번째 변수 그룹의  $i$ 번째 관측에 대한 가중합으로  $k$ 번째 정준변수와 상관계수를 가중치로 이용하고 있는 것이다.  $K$ 차원 공간상의 상관원그림은 이들 관계를 이용한 것이다 (Figure 3.1 참고).

다중요인분석에서는 자료를 나타내는  $K$ 차원의 정준변수를 찾은 후 이들과 분석에 이용되지 않은 보충 변수(supplementary variable)들의 상관관계를 이용하여 같은 공간에 투영시켜 보충 변수들이 자료의 주요 차원들과 어떤 관계인지를 살펴볼 수 있게 된다. 관능검사분석을 위한 자료의 경우 전문 패널에 의해 평가되는 제품의 감각적 특성(sensory attribute) 뿐 아니라 실험으로부터 얻어지는 화학적 특성(physicochemical attribute) 등도 있다. 이들은 제품에 대한 객관적인 특성이라고 할 수 있다. 또한 소비자들로부터 얻어지는 주관적 기호도에 대한 특성(hedonic judgement) 자료도 있다. 다중요인분석은 제품에 대한 객관적인 특성들에 대한 자료를 이용하여 제품과 객관적 특성들 간의 관계를 잘 나타내는 요인을 찾아내고 소비자 기호도 자료를 요인들의 공간상에 투영시켜 제품의 객관적 특성과 소비자 기호 간의 관계를 파악할 수 있게 된다. 이는 추후 제품개발과 제품 홍보 등에 유용한 정보로 이용될 수 있다.

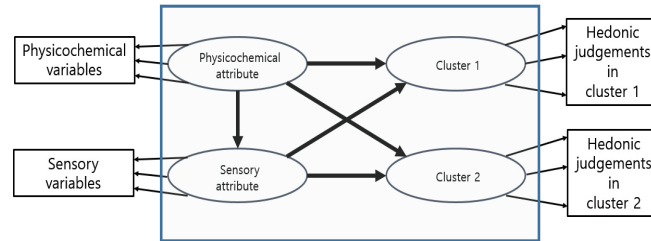


Figure 2.1. Partial least square path model (Pages and Tenenhaus, 2001).

## 2.2. 부분 최소제곱 경로모형

부분 최소제곱 경로모형(partial least square path model) (Lohmöller, 2013)은 구조방정식 모형 중 하나로 사회과학을 비롯한 다양한 분야에서 여러 개념간의 인과관계를 표현하기 위한 모형으로 이용된다. 이론과 논리적 타당성을 근거로 설정한 모형의 적합에 중점을 두는 일반적인 구조방정식 모형과는 달리 부분 최소제곱 경로모형은 이론적 연구가 부족한 상황에서 예측이나 설명을 목적으로 하는 탐색적 성격을 갖는 모형이다.

구조방정식 모형은 구조모형을 나타내는 내부 모형(inner model)과 측정모형을 나타내는 외부 모형(outer model)의 선형 결합 방정식 형태로 정의된다. 내부 모형은 관측되지 않은 잠재변수 간의 관계를 나타내며 식 (2.6)으로 표현된다. 이 때  $\eta$ 와  $\xi$ 는 잠재변수를 나타낸다.

$$\eta = \mathbf{B}\eta + \Gamma\xi + \zeta. \quad (2.6)$$

외부 모형은 잠재변수  $\eta$ 와  $\xi$ 와 측정변수들 간의 관계를 식 (2.7)와 같이 나타낸다. 여기서  $\mathbf{X}$ 는  $\eta$ 와 관련 있는 측정변수들을,  $\mathbf{Y}$ 는  $\xi$ 와 관계있는 측정변수들을, 그리고  $\zeta, \delta, \varepsilon$ 은 오차항을 나타낸다.

$$\begin{aligned} \mathbf{X} &= \Lambda_x \eta + \delta, \\ \mathbf{Y} &= \Lambda_y \xi + \varepsilon. \end{aligned} \quad (2.7)$$

부분 최소제곱 경로모형은 위와 같이 정의된 구조방정식 모형을 공분산과 가능도함수를 이용하여 추정하기보다는 가중치 벡터를 갖는 연속적인 회귀분석을 이용하여 추정한다. Figure 2.1은 Pages와 Tenenhaus (2001)에서 이용한 경로모형으로 제품에 대한 화학적 특성과 감각특성 등 객관적인 특성자료들을 잠재변수  $\eta$ 로 하고 상반된 특성을 나타내는 두 군집 cluster 1과 cluster 2의 기호도를 잠재변수  $\xi$ 로 하여 이들 간의 관계를 내부 모형으로 나타낸 경로모형이다. 외부모형을 위해 화학적 특성과 감각특성 등 객관적인 특성에 대한 변수들을  $\mathbf{X}$ 로, 두 군집의 기호도에 대한 변수를  $\mathbf{Y}$ 로 나타내었다. 그림에서 사각형 영역 내부의 잠재변수들 간의 관계를 내부 모형, 사각형 영역 외부에 표현된 관계를 외부모형이라고 한다.

본 연구에서는 객관적인 감각특성 자료 뿐 아니라 소비자들로부터 측정된 주관적 감각특성 자료를 활용하기 위하여 객관적인 감각특성을 이용하여 다중요인분석을 실시한 후 이를 이용하여 소비자를 4개의 군집으로 나누는 것뿐만 아니라 소비자들로부터 측정된 주관적 감각특성 변수들도 각 군집의 특성을 나타내는 변수들 군집으로 나누어 각각 잠재변수로 이용하였다 (Figure 3.5와 Figure 3.6). 다중요인분석과 부분최소제곱경로모형 추정을 위하여 R에서 제공하고 있는 FactoMineR (Lê 등, 2008), SensoMineR (Lê와 Husson, 2008), 그리고 plsmpm (Sanchez와 Trinchera, 2012) 패키지를 이용하였다.

**Table 3.1.** 12 sensory attributes in perfume data

Spicy	계피, 생강, 후추 등의 자극적인 강한 향
Heady	향의 강도
Fruity	과일 향기
Green	식물 향기
Vanilla	바닐라 향기
Floral	꽃 향기
Woody	나무 향기
Citrus	감귤류 과일 향기
Marine	바다 향기
Greedy	꿀, 초콜렛, 바닐라, 사탕 등의 식품을 연상시키는 향기
Oriental	동양에서 얻은 향료를 배합하여 얻은 향기
Wrapping	향수를 포장하고 있는 용기

### 3. 향수의 특성을 이용한 다중요인분석

#### 3.1. 향수 자료

향수 자료는 시중에서 판매하고 있는 12가지 향수에 대해 전문가와 소비자들이 평가한 자료로 프랑스 Argocampus Quest 대학의 석사 학생들로부터 얻어진 자료이다 (Lê와 Worch, 2014). 12명의 훈련된 패널리스트들은 두 세션에 걸쳐 12가지 감각 속성들을 0점부터 10점까지의 척도로 평가하였다. 패널들이 평가한 감각 속성에는 Spicy, Heady, Fruity, Green, Vanilla, Floral, Woody, Citrus, Marine, Greedy, Oriental, 그리고 Wrapping이 있다. Fruity, Green, Vanilla, Floral, Woody, Citrus, 그리고 Marine는 각각 과일, 식물, 바닐라, 꽃, 나무, 감귤류 과일, 그리고 바다 향을 의미하고, Oriental은 동양에서 서식하는 동물로부터 얻은 향료를 배합해서 얻은 오리엔탈 향이다. Spicy는 자극적인 강한 향으로 계피, 생강, 후추 등의 향을 의미하며 Heady는 향의 강도를 나타내고 Wrapping은 향수를 포장하고 있는 용기에 대한 평가이다. 마지막으로 Greedy는 프랑스어로 구르망(Gourmand)을 뜻하는 것으로 꿀, 초콜릿, 바닐라, 사탕 등과 같이 먹을 수 있는 식품을 연상시키는 향이라고 할 수 있다 (Table 3.1).

또한 103명의 소비자들은 향수의 21가지 감각특성을 0점부터 100점까지로 평가하고 추가적으로 개인적인 선호도도 1점부터 9점까지로 기록한다. 소비자들이 평가한 21가지 감각특성에는 intensity, freshness, jasmin, rose, camomille, fresh lemon, vanilla, citrus, anis, sweet fruit, honey, caramel, spicy, woody, leather, nutty, musk, animal, earthy, incense, 그리고 green이 있다. jasmin, rose는 꽃의 종류인 자스민, 장미의 향이고 camomille, anis는 식물의 종류인 캐모마일, 아니스 향이며 freshness, fresh lemon, sweet fruit는 상쾌한 향, 레몬, 달콤한 과일 향을 의미하여 과일 향 범주에 속한다. vanilla는 바닐라 향을 의미하고 caramel은 설탕과 바닐라 향료가 섞인 캐러멜 향, musk는 바닐라와 오리엔탈 향이 섞인 머스크 향이다. animal, leather는 동물, 가죽 향을 뜻해서 동물성 향료인 오리엔탈 향과 관련 있을 것으로 추측된다. 이들 감각특성은 전문가 패널에 의해 평가된 감각속성과 밀접한 관련이 있다.

향수 자료는 전문가 패널들이 평가한 12가지 감각 속성 자료(12개 제품 × 2번의 세션 × 12명 패널 × 12개 속성 변수)와 103명의 소비자들이 평가한 21가지 감각특성 자료(12개 제품 × 103명 소비자 × 21가지 특성), 그리고 소비자들의 기호도 자료(12개 제품 × 103명 소비자)로 구성되어 있으며 전문패널 평가 감각속성, 소비자 평가 특성, 그리고 소비자 기호도의 세가지 변수 그룹으로 정리된다. 전문패널이 평가한 감각속성 자료의 경우 12명의 패널로부터 2명의 세션에 의해 측정된 값의 평균을 취하여 각 향수 제품의 감각속성을 정리하였다. 전문 교육을 받은 전문가 패널들과는 달리 소비자들의 평가에는 개인적 취향 등의 외부 요인에 크게 영향을 받을 수 있다는 점을 고려하여 소비자들이 평가한 21가지

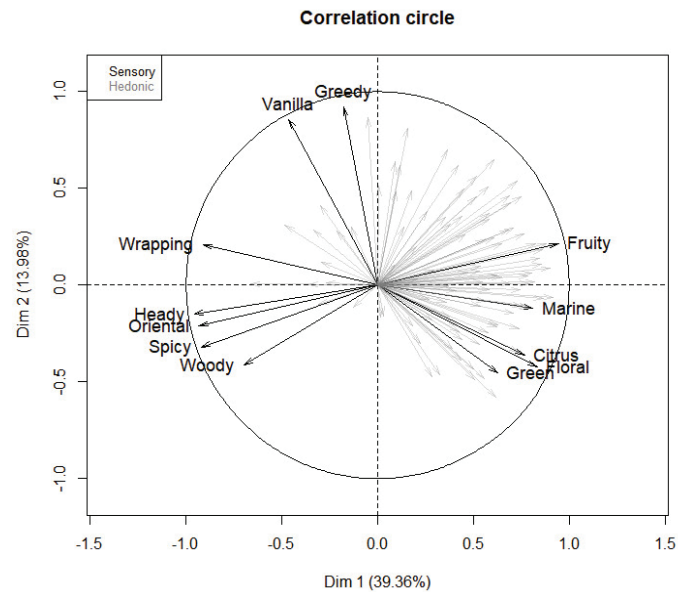


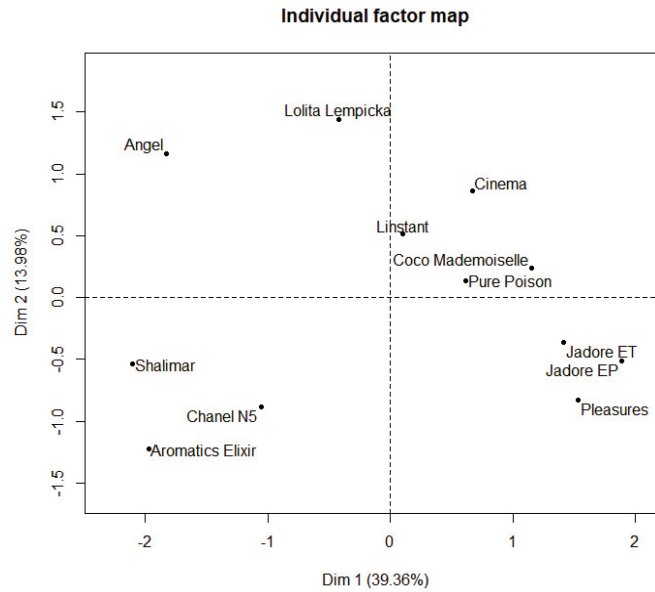
Figure 3.1. Representation of the sensory attributes evaluated by experts.

특성 자료는 103명에 의해 평가된 값들의 평균 대신 중앙값을 취하였으며 기호도 자료는 103명 소비자들의 12개 제품에 대한 평가를 그대로 사용하였다.

본 논문에서는 이들 세가지 변수 그룹의 향수 자료를 이용하여 향수의 특성을 파악하고 전문 패널 감각 속성, 소비자들이 평가하는 감각특성, 그리고 이들과 소비자들의 선호도와와의 관계 등을 다중요인분석으로 살펴보고자 한다. 전문패널과는 달리 소비자들은 같은 향수 제품을 시향하고 평가할 때 사람마다 향수의 특성을 다르게 지각할 수 있으므로 다중요인분석은 전문가 평가의 감각속성과 소비자 기호도만으로 시행하였다. 다중요인분석결과를 이용하여 소비자들을 특성별로 4개의 군집으로 분류하고 또한 소비자 평가 감각특성도 4개의 군집으로 분류하여 군집별로 선호하는 제품의 속성을 살펴보고 이를 이용하여 기호도를 예측해본다. 또한 각 군집별로 어떤 특성을 가지고 향수를 지각하는지 알아보고 이 특성을 사용하여 기호도를 예측하는 등 각 소비자 군별로 좀 더 자세히 살펴보기 위하여 경로분석을 이용하고자 한다.

### 3.2. 다중요인분석

향수 자료 중 전문패널 평가 감각속성과 소비자 기호도를 이용하여 다중요인분석을 실시하였다. Figure 3.1은 다중요인분석 결과 중 전문패널 평가 감각속성 그룹의 변수들만을 상관 원그림에 나타낸 것이고 Figure 3.2는 관측치인 향수들을 각 요인으로 변환된 공간에 나타낸 개별요인지도이다. Figure 3.1을 통해 첫 번째 요인과 두 번째 요인이 설명하는 변동의 비율은 각각 39.36%, 13.98%이고 두 요인이 전체 변동의 53.34%를 설명하고 있음을 알 수 있다. 첫 번째 요인과 양의 상관관계가 큰 변수들은 Fruity, Marine, Citrus, Floral, 그리고 Green이고 음의 상관관계가 큰 변수들은 Woody, Spicy, Oriental, Heady, 그리고 Wrapping이다. 또한 두 번째 요인과 양의 상관관계가 큰 변수들은 Vanilla와 Greedy이다. 이를 통해서 첫 번째 요인은 과일(Fruity), 바다(Marine), 감귤류 과일(Citrus), 꽃(Floral), 그리고 식물(Green) 향이 나는 특성이 있으면 양의 값을 갖고, 나무(Woody), 오리엔탈(Oriental) 향이 나거나



**Figure 3.2.** Individual factor map for perfume data.

향이 자극적이고(Spicy) 강하거나(Heady) 향수 용기가 예쁘면(Wrapping) 음의 값을 갖는다. 또한 두 번째 요인은 바닐라(Vanilla), 구르망(Greedy) 향이 나는 특성이 있으면 양의 값을 가진다.

Figure 3.2에서 비슷한 위치에 있는 향수들은 비슷한 특성을 갖는 경향이 있다고 할 수 있다. Cinema, Pure Poison, Coco Mademoiselle이 가까이에 위치하며 첫 번째와 두 번째 요인의 값이 모두 양의 값을 가진다. J'adore ET, J'adore EP, 그리고 Pleasures도 위의 향수들과 같이 첫 번째 요인이 양의 값을 가지나 두 번째 요인의 값은 음의 값을 가지고 있어 첫 번째 요인의 특성은 비슷하나 두 번째 요인의 특성은 반대가 됨을 알 수 있다. Shalimar, Aromatics Elixir, 그리고 Chanel N5는 첫 번째, 두 번째 요인 모두 음수이며, Angel은 첫 번째 요인은 음수이나 두 번째 요인은 양수임을 알 수 있다. Lolita Lempicka는 첫 번째 요인의 특성은 크지 않으나 두 번째 요인은 양수로 영향이 큰 것을 알 수 있다. L'Instant의 경우 두 요인 모두 영향이 크지 않고 평균적인 특성을 가지고 있음을 알 수 있다.

Figure 3.1과 Figure 3.2를 동시에 고려하면 각 함수의 감각속성을 파악할 수 있다. Lolita Lempicka와 Angel은 바닐라와 구르망 향이 많이 나는 것을 알 수 있다. 특히 Angel은 용기가 예쁜 특성이 있고, Woody, Oriental, Spicy, 그리고 Heady 특성이 강해서 그 향이 더욱 강하다는 사실을 알 수 있다. Cinema와 Coco Mademoiselle는 과일 향이 많이 나고, J'adore ET, J'adore EP, 그리고 Pleasures는 꽃 향과 감귤류 과일 향이 많이 나는 향수이다. 특히 J'adore EP와 Pleasures는 화살표의 길이가 더 길어서 그 특성이 더 강하고 바다 향이 가미되었고, Pleasures는 이에 식물 향이 더해진다. 마지막으로 Shalimar, Aromatics Elixir, 그리고 Chanel N5는 나무, 오리엔탈 향이 나고 그 향이 강하고 자극적이며 용기가 예쁜 특성이 있다.

Figure 3.3은 소비자 기호도 변수 그룹을 상관 원그림에 표현한 것이다. 향수 소비자의 기호가 첫 번째 요인의 양의 특성에 몰려 있다는 사실을 확인할 수 있다. Figure 3.2와 Figure 3.3을 동시에 고려하면 소비자들이 선호하는 향수의 종류를 파악할 수 있다. Lolita Lempicka, Cinema, Pure Poison, Coco Mademoiselle, J'adore ET, J'adore EP, 그리고 Pleasures 제품에 소비자 기호도의 대부분이 몰려 있

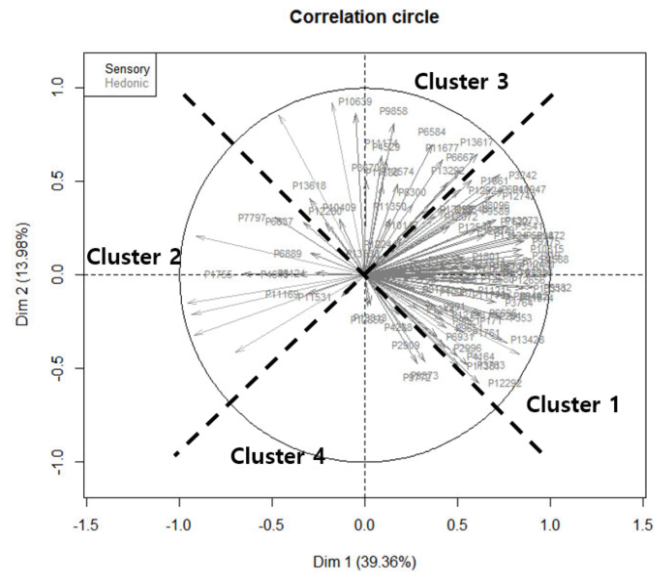


Figure 3.3. Representation of the sensory attributes evaluated by consumers.

는 것으로 보아 이들이 대중적으로 인기 있는 향수이고 그 외의 제품들은 소수의 마니아층이 있는 제품임을 알 수 있다. 따라서 많은 향수 소비자들이 과일, 꽃, 감귤류 과일이 포함된 향수 또는 바닐라와 그리디 향이 진하지 않고 은은하게 나는 향수를 선호한다는 것을 알 수 있다.

### 3.3. 군집화 과정

군집화 과정은 다중요인분석 결과를 이용하여 소비자들을 군집화하고 경로모형에 사용될 변수를 선택하는 과정으로 Figure 3.3의 기호도 변수 그룹을 상관 원그림에 나타내어 45°와 135°의 점선을 경계로 하여 4개의 군집으로 분할하는 것이다. 군집 1과 군집 2는 두 번째 요인의 특성보다는 첫 번째 요인의 특성에 더 관련 있는 군집으로 군집 1은 첫 번째 요인의 값이 두 번째 요인 값의 절대값보다 큰 양의 값을 갖는다. 군집 2도 첫 번째 요인의 값이 두 번째 요인 값의 절대값보다 큰 값을 갖지만 음의 값을 갖는다. 군집 3과 군집 4는 첫 번째보다는 두 번째 요인의 특성에 더 밀접하게 관련 있다고 할 수 있다. 이와 같이 나뉜 4개의 군집은 각각 69, 8, 20, 그리고 6명의 소비자가 있으며 대부분의 소비자들은 군집 1에 모여 있다. 좀 더 자세한 특성을 파악하기 위하여 군집 1과 군집 2를 하나의 경로 모형에, 그리고 군집 3과 4를 또 다른 경로 모형으로 구성하여 각 요인에 대해 반대의 특성을 지니는 두 개의 군집을 하나의 경로모형에서 함께 살펴본다.

각 경로모형에서 사용할 전문패널 평가 감각 속성 변수들을 선택하기 위해 Figure 3.1의 감각속성 상관 원그림을 살펴본다. 두 번째 요인보다는 첫 번째 요인에 더 큰 영향을 주는 감각속성 변수들은 양의 영향을 주는 Fruity, Marine, Citrus, 그리고 Floral와 음의 영향을 주는 Spicy, Oriental, Heady, 그리고 Wrapping 변수이므로 이들을 군집 1과 2를 위한 모형에 이용한다. 또한 첫 번째 축보다는 두 번째 축에 가까운, 즉 두 번째 요인에 양의 영향을 주는 Vanilla, Greedy 변수들은 군집 3과 4를 위한 모형에 이용한다. 두 번째 요인의 음의 값을 갖는 변수가 명확하지 않으므로 가장 유사한 Woody와 Green 변수를 두 번째 요인에 음의 영향을 주는 변수로 지정하고 군집 3과 4를 위한 모형 적용에 사용한다. 따라서



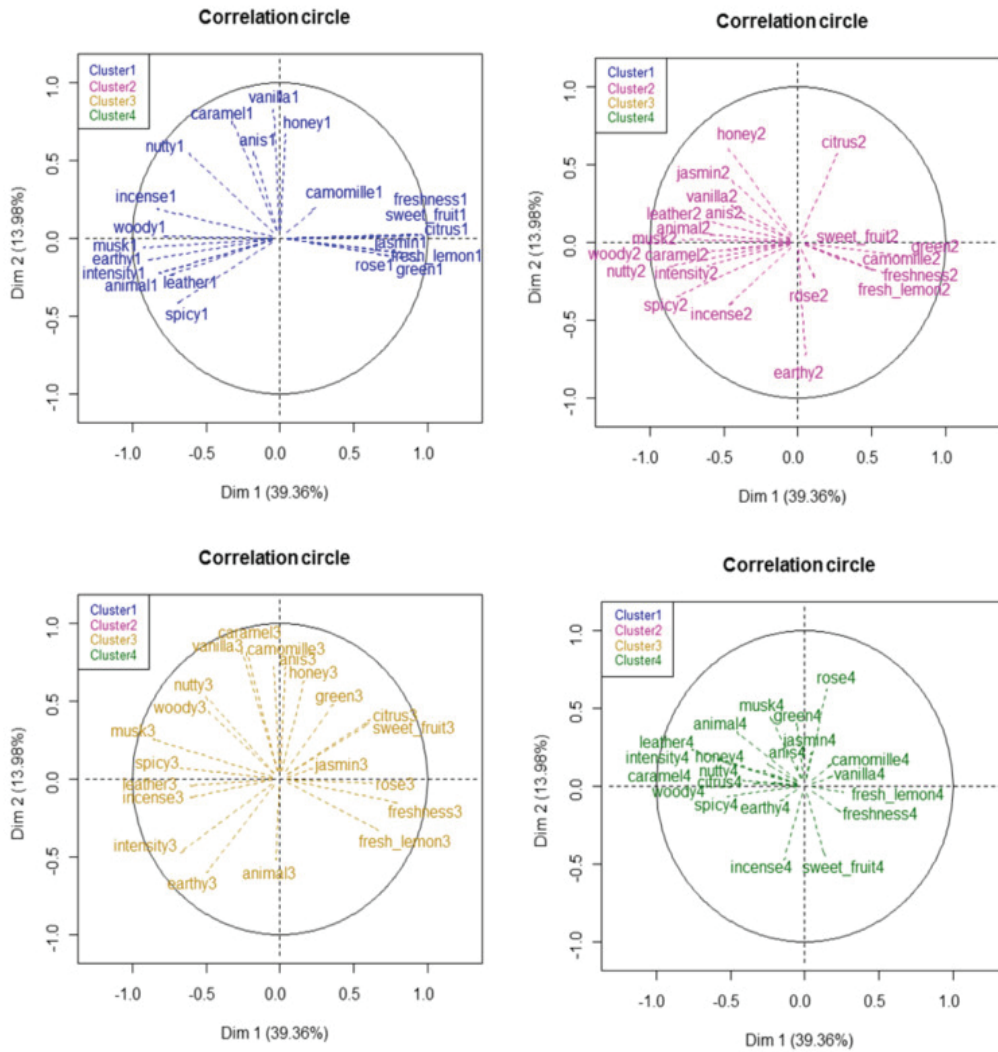


Figure 3.4. Representation of the sensory attributes evaluated by consumers in each cluster.

첫 번째 모형에는 8개, 두 번째 모형에는 4개의 변수가 제품의 감각 속성을 설명하는 관측 변수로 사용된다. 이와 같이 두 모형에 사용되는 상품 특성 변수들은 완전히 분리되고 각 모형에서 상품의 다른 특성을 바탕으로 소비자의 기호도를 예측한다.

각 군집별 소비자들이 어떤 특성을 가지고 향수를 지각하는지를 살펴보고 이들과 전문가 패널이 평가한 감각속성과의 관계, 그리고 제품과의 관계를 알아보기 위해 소비자들이 평가한 21가지 특성 변수의 군집별 중앙값을 계산하여 다중요인분석의 보충 변수로 사용하였다. 각 군집별로 계산된 21개 변수는 변수명 뒤에 해당 군집 번호를 붙여 구분하였다. Figure 3.4는 이를 상관원그림에 군집별로 표현한 것이다. 그리고 소비자들을 군집화 했던 기준을 활용하여 다음 모형 적합에 사용될 변수를 선택하였다. 예를 들어, 군집 1의 경우에는 두 번째 요인보다 첫 번째 요인에 더 가깝고 첫 번째 요인의 양의 값 범

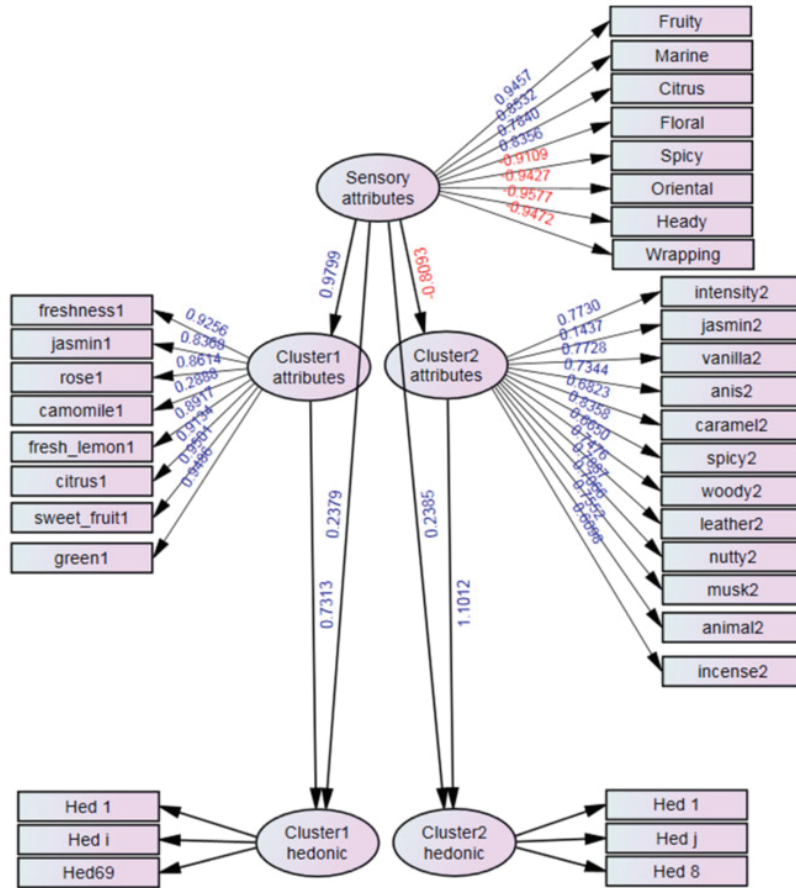


Figure 3.5. Path model for cluster 1 and cluster 2.

위에 위치한 변수들을, 군집 2의 경우에는 음의 값 범위에 위치한 변수들을 선택한다. 같은 방법으로 군집 3에서는 첫 번째 요인보다 두 번째 요인에 더 가깝고 두 번째 요인의 양의 값 범위에 위치한 변수들을, 군집 4는 음의 값 범위에 위치한 변수들을 선택할 수 있다. 군집 1에서 선택된 특성 변수로는 freshness1, jasmin1, rose1, camomille1, fresh lemon1, citrus1, sweet fruit1, 그리고 green1이 있고 군집 2에서는 intensity2, jasmin2, vanilla2, anis2, caramel2, spicy2, woody2, leather2, nutty2, musk2, animal2, 그리고 incense2이, 군집 3에서는 camomille3, vanilla3, anis3, honey3, caramel3, nutty3, 그리고 green3이, 군집 4에서는 sweet fruit4, incense4을 선택하였다. 각 군집에서 선택된 특성 변수는 군집별로 선호하는 제품의 속성과 큰 연관성이 있는 것을 확인할 수 있다.

3.4. 다중요인분석을 이용한 부분 최소제곱 경로 모형

다중요인분석 결과를 통해 소비자들을 군집화하고 두 개의 군집씩 묶어서 부분 최소제곱 경로모형 분석하기 위한 변수를 적절한 기준에 따라 선택하였다. 군집 1과 군집 2를 위한 모형 적합에 사용될 자료는 전문패널 평가 감각속성, 군집 1 소비자 69명의 기호도(Cluster1 hedonic), 군집 2 소비자 8명의

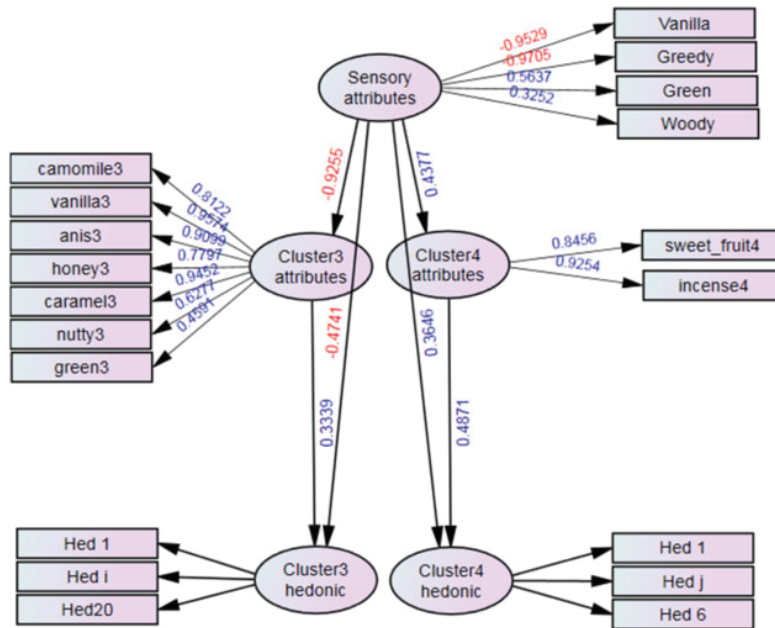


Figure 3.6. Path model for cluster 3 and cluster 4.

기호도(Cluster2 hedonic), 군집 1, 그리고 군집 2의 소비자들이 평가한 감각속성(Cluster1 attributes, Cluster2 attributes) 으로 구성되어 있다. 이 모형에서는, 제품의 전문패널 평가 감각속성을 이용해서 직접적으로 군집 1과 군집 2의 소비자의 기호도를 예측하는 경로와, 전문패널 평가 감각속성에서 군집 1과 군집 2의 소비자가 평가한 감각속성을 거쳐 해당 군집의 기호도를 예측하는 경로를 함께 고려한다. Figure 3.5는 군집 1과 군집 2의 부분 최소제곱 경로 모형을 적합한 결과를 나타낸다. 감각 속성의 잠재 변수는 과일, 바다, 감귤류 과일, 꽃 향이 나면 양의 값을 갖고, 향이 자극적이거나 강하고 오리엔탈 향이 나거나 용기가 예쁜 특성이 있으면 음의 값을 갖는다. 감각 속성과 군집별 특성 변수의 관계를 통하여, 군집 1의 소비자들은 감각 속성 잠재 변수의 양의 특성인 상쾌함, 자스민, 장미, 레몬, 감귤류 과일, 달콤한 과일, 식물 향을 잘 지각하고 있고 군집 2는 음의 특성인 강하거나 자극적인 향, 그리고 동물, 가죽, 머스크 향에 민감하게 반응한다고 할 수 있다.

내부 모형에서 각 군집의 기호도 잠재 변수에 영향을 미치는 설명 변수는 감각 속성과 해당 군집이 평가한 특성, 총 두 가지이다. 내부 모형은 다중 회귀 모형의 회귀 계수를 통해 경로계수를 생성한다. 따라서 감각 속성과 해당 군집이 지각하는 특성 중 어떤 것이 기호도에 더 큰 영향을 주는지 살펴보기 위해, 경로계수를 비교하였다. 이를 통하여 제품의 감각 속성이 직접적으로 기호도에 미치는 영향보다도 소비자가 평가한 특성이 더 큰 영향력을 가진다는 사실을 확인할 수 있다.

Figure 3.6은 군집 3과 군집 4의 모형 적합 결과를 그림으로 나타낸 것이다. 군집 3과 군집 4를 위한 모형도 앞서 설명한 것과 마찬가지로 해석할 수 있다. 감각 속성 잠재 변수는 바닐라와 구르망 향이 나면 음의 값을, 식물과 나무 향이 나면 양의 값을 갖는다. 군집 3의 소비자들은 감각 속성 잠재 변수의 음의 특성인 바닐라, 꿀, 캐러멜 향을 잘 지각하고 있다. 하지만 군집 4에서는 감각 속성과 소비자가 평가한 특성 사이에 큰 연관성이 없어서 군집 4의 소비자들이 어떤 특성에 더 민감하고 둔감한지를 파악하기 어

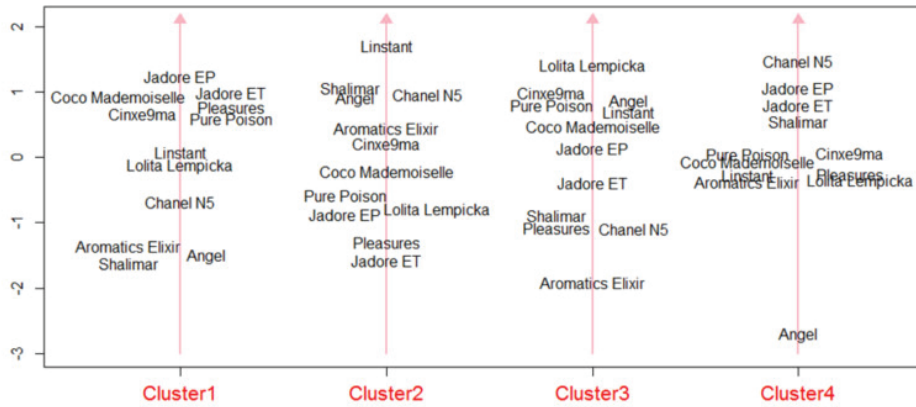


Figure 3.7. Latent variables summarising the four clusters of hedonic judgements.

럽다. 원인은 군집 4에 해당되는 소비자의 수가 적을 뿐만 아니라 해당 모형에 사용될 변수를 선택하는 과정에서 다중요인분석의 두 번째 요인 음의 특성을 설명하는 감각 속성 변수가 충분하지 않았기 때문으로 파악할 수 있다. 내부 모형의 경로계수를 통해, 군집에 따라 느끼는 향수의 특성은 매우 다르지만 이 특성이 기호도에 미치는 영향은 향수의 감각 속성이 직접 기호도에 미치는 영향보다 약하다는 사실을 알 수 있다. 따라서 군집 3과 군집 4의 모형에서는 군집 1과 군집 2의 모형과는 달리 제품의 감각 속성이 기호도에 더 큰 영향을 미치고, 군집 3은 바닐라와 구르망 향이 나는 향수를, 군집 4는 식물과 나무 향이 나는 향수를 선호한다는 것을 파악할 수 있다.

두 모형을 통해 얻은 군집별 기호도 잠재 변수를 사용하여 군집별로 선호하는 향수를 파악해보았다. Figure 3.7은 각 관측치를 제품명으로 표기하여 4가지 잠재 변수를 표현한 것으로 군집 1에는 소비자의 과반수가 포함되어 있어서, 대중적으로 인기 있는 향수를 선호하는 특성을 가지고 군집 2는 상대적으로 소수의 마니아층이 있는 향수를 선호하고 대중적인 향수를 선호하지 않는 것을 확인할 수 있다. 또한 군집 3은 바닐라와 구르망 향이 강한 향수를 선호하고 군집 4는 이와 반대의 특성을 갖고 Angel 제품을 기피하는 것으로 보인다.

#### 4. 결론 및 토의

본 논문에서는 다중요인분석을 이용한 부분 최소제곱 경로 모형을 통해 향수 소비자들을 비슷한 기호를 갖는 4개의 군집으로 군집화하고, 군집별로 선호하는 제품의 특성, 그리고 소비자들이 지각하는 향수의 특성을 파악하였다. 또한 경로 모형을 사용하여 제품의 감각 속성, 군집별 평가 특성, 그리고 기호도를 연결하고 기호도에 영향을 미치는 중요 변수도 살펴보았다.

군집별로 선호하는 향수의 감각 속성은 서로 상이하고, 군집 1과 군집 2는 소비자가 느끼는 특성이, 군집 3과 군집 4는 제품 자체의 특성이 기호도를 예측하는데 더 큰 영향을 주는 것을 알 수 있었다. 군집 1은 상쾌함, 자스민, 장미, 레몬, 감귤류 과일, 달콤한 과일, 식물 향에, 군집 2는 강하거나 자극적인 향 그리고 동물, 가죽, 머스크 향에 민감하고 해당 특성을 가진 제품을 선호한다는 사실을 알 수 있었다. 또한 대부분의 소비자들이 군집 1에 몰려 있어서 군집 1은 일반적인 향수 소비자들의 취향을 반영하고 있었다. 그리고 군집 3은 바닐라와 구르망 향이 나는 향수를 선호하고, 군집 4는 식물과 나무 향이 나는 향수를 선호하며 특정 향수 제품을 극도로 싫어하는 특성을 파악할 수 있었다. 그러나 군집 4의 소비자

들이 선호하는 제품 특성을 설명하는 변수가 충분하지 않아서 군집 4의 특성을 파악하고 기호도를 예측하는데 어려움이 따랐다.

## References

- Jee, J. H., Lee H. S., Lee, J. W., Suh, D. S., Kim, H. S., and Kim, K. O. (2008). Sensory characteristics and consumer liking of commercial soju marketed in Korea, *Korean Journal of Food Science and Technology*, **40**, 160–165.
- Lê, S. and Husson, F. (2008). SensoMineR: a package for sensory data analysis, *Journal of Sensory Studies*, **23**, 14–25.
- Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis, *Journal of Statistical Software*, **25**, 1–18.
- Lê, S. and Worch, T. (2014). *Analyzing Sensory Data with R*, CRC Press.
- Lee, S. J. and Lee, K. G. (2008). Understanding consumer preferences for rice wines using sensory data, *Journal of the Science of Food and Agriculture*, **88**, 690–698.
- Lohmöller, J. B. (2013). *Latent Variable Path Modeling with Partial Least Squares*, Springer Science & Business Media.
- Pages, J. (2014). *Multiple Factor Analysis by Example using R*, CRC Press.
- Pages, J. and Tenenhaus, M. (2001). Multiple factor analysis combined with PLS path modelling. Application to the analysis of relationships between physicochemical variables, sensory profiles and hedonic judgements, *Chemometrics and Intelligent Laboratory Systems*, **58**, 261–273.
- Sanchez, G. and Trinchera, L. (2012). plspm: Partial Least Squares data analysis methods. R package version 0.2-2, URL <http://CRAN.R-project.org/package=plspm>.

# 다중요인분석을 이용한 부분 최소제곱 경로 모형에 대한 고찰

박리라<sup>a</sup> · 이은경<sup>a,1</sup>

<sup>a</sup>이화여자대학교 통계학과

(2018년 1월 23일 접수, 2018년 3월 1일 수정, 2018년 4월 3일 채택)

---

## 요약

다중요인분석은 관능검사에서 주로 이용되는 분석으로 상품의 속성과 소비자들의 기호도에 대한 자료 분석에 주로 이용된다. 본 연구에서는 다중요인분석을 상품의 속성, 소비자들의 기호도 등의 자료에 이용하여 소비자들의 특성에 따라 몇 개의 군집으로 분류하고 이를 부분 최소제곱 경로모형을 이용하여 분류된 군집의 특성을 파악해 보고자 한다. 함수의 속성에 대한 자료와 소비자들이 파악한 함수의 속성, 그리고 그들의 기호도에 관한 실제 자료를 다중요인분석을 이용하여 살펴보고 이 결과를 이용하여 소비자들을 4개의 군집으로 분할한다. 분할한 군집별로 제품의 특성을 파악하고 이를 최소제곱경로모형에 적용하여 각 군집의 특성을 나타내는 잠재변수를 추정, 군집별로 소비자들이 선호하거나 기피하는 제품의 특성들, 그리고 각 군집별로 제품들을 어떻게 지각하는지 등을 파악한다. 다중요인분석을 활용한 부분 최소제곱 경로모형은 제품에 대한 특성과 소비자들의 기호도를 동시에 분석하여 이들의 관계를 규명하고 분석 결과를 제품 개발과 판매에 적용할 수 있다는 점에서 유용한 모형이라고 할 수 있다.

주요용어: 다중 요인 분석, 부분최소제곱, 경로모형, 관능검사

---

---

이 논문은 2015년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (2015R1D1A1A01056790).

<sup>1</sup>교신저자: (03760) 서울특별시 서대문구 이화여대길 52, 이화여자대학교 통계학과.

E-mail: lee.eunk@ewha.ac.kr