

PLS 방법에 의한 “큰” 2원 교차표의 시각화

이용구^{1,a}, 최연임^a

^a중앙대학교 통계학과

요약

범주형 자료의 시각화에서 범주가 많지 않은 경우에는 기존의 Hayashi의 수량화 제3방법을 이용하여 두 변수의 범주들 사이의 연관성에 대한 시각화를 구할 수 있다. 그러나, Hayashi 방법은 큰 빈도의 범주들보다 작은 빈도의 범주들을 두드러지게 수량화하므로 결과가 불안정하다는 문제점이 있다 (허명희와 이용구, 2006). 이 연구의 목적은 범주수가 “큰” 두 범주형 변수 R 과 C 에 대하여 각 변수의 범주들 사이의 연관성을 살펴보기 위한 시각화 방법을 제안하는 데 있다. 이를 위하여 우리는 2개 변수군의 수치형 자료를 시각화하는 방법으로 제안된 허명희 등 (2007)의 PLS 시각화 방법을 범주형 자료에 적용하고자 한다. 즉, 범주형 변수 R 과 C 의 범주들 각각을 0/1로 더미 코드화하여 각각 R 개와 C 개의 범주군으로 변환한 다음 허명희 등 (2007)에서 제시한 PLS 시각화 방법을 적용하고자 한다. 이러한 방법은 Hayashi 수량화 방법의 문제점을 해결할 수 있을 뿐만 아니라 행변수와 열변수 각각이 여러 개의 범주형 변수들의 집합인 변수군의 경우에도 확대 적용 가능하다. 수치 예로서 German Credit 자료에서 10개 금융관련 변수의 34개 범주를 R 로 간주하고 10개 사회인구적 변수의 46개 범주를 C 로 간주하여 새 방법론을 적용해 보인다.

주요용어: Hayashi 수량화 제3방법, PLS 수량화, “큰” 2원 교차표.

1. 범주형 자료의 수량화

n 명의 응답자로부터 얻은 p 개 범주의 범주형 변수 R 과 q 개 범주의 범주형 변수 C 를 더미 코딩하여 $n \times p$ 행렬 $Z_R = \{z_{ij}^R\}$ 과 $n \times q$ 행렬 $Z_C = \{z_{ik}^C\}$ 가 만들어졌다고 하자. 즉, z_{ij}^R 은 응답자 i 가 R 의 j 번째 범주에 반응한 경우 1, 반응하지 않은 경우 0으로 코딩한다. 마찬가지로 z_{ik}^C 는 응답자 i 가 C 의 k 번째 범주에 반응한 경우 1, 반응하지 않은 경우 0으로 코딩한다. 따라서 R 과 C 가 택일형(single choice) 질문에서 나온 경우 $F = Z_R'Z_C$ 는 $p \times q$ 교차표(crosstab)가 된다.

Hayashi의 수량화 제3방법은 다음의 정식화로부터 R 의 p 개 범주에 대한 수량화 값 $p \times 1$ 벡터 x 와 C 의 q 개 범주에 대한 수량화 값 $q \times 1$ 벡터 y 를 구한다 (허명희, 1998; 허명희와 이용구, 2006).

$$\begin{aligned} & \text{maximize } x' \frac{Fy}{n}, \\ & \text{subject to } x' \left(\frac{D_r}{n} \right) x = 1, \quad y' \left(\frac{D_c}{n} \right) y = 1, \end{aligned} \quad (1.1)$$

여기에서 D_r 과 D_c 는 각각 행렬 F 의 $p \times 1$ 행 주변빈도, $q \times 1$ 열 주변빈도의 대각행렬이다. 즉, $D_r = \text{diag}(F1_q)$, $D_c = \text{diag}(F'1_p)$ 이다. 제2축의 행과 열 수량화는 각각 D_r 과 D_c 를 가중치로 제1축 수량화 벡터 x 와 y 에 직교하는 공간에서 찾는다. 이러한 Hayashi 수량화 제3방법은 1940년대 후반부터 1950년대 전반에 걸쳐서 일본의 C. Hayashi에 의해 창안된 교차표 분석 기법으로 사회조사 및 마케팅 조사자

이 논문은 2007년도 중앙대학교 학술연구비 지원에 의한 것임.

¹ 교신저자: (156-756) 서울시 동작구 흑석동 221 중앙대학교 통계학과, 교수. E-mail: leeyg@cau.ac.kr

료의 분석에서 매우 유용하다 (岩坪秀一, 1987; 駒澤勉, 1992). 또한, 이 방법은 프랑스에서 개발된 대응분석(correspondence analysis)과 기본 알고리즘을 공유하는 것으로 잘 알려져 있다. 그러나, 이 방법은 반응빈도가 작은 일부 범주들이 특이하게 큰 수량화 값을 갖게되는 문제가 있다. 이는 식 (1.1)의 제약식 $x'(D_r/n)x = 1$ 과 $y'(D_c/n)y = 1$ 하에서 행 범주 비율 벡터 R 의 i 번째 요소 r_i ($i = 1, \dots, p$) 또는 열 범주 비율 벡터 C 의 j 번째 요소 c_j ($j = 1, \dots, q$)가 작으면 해당하는 수량화 값 x_i 와 y_j 가 비교적 자유롭게 움직일 수 있기 때문이다 (허명희와 이용구, 2006). 이를 개선하기 위하여 허명희와 이용구 (2006)는 Hayashi 수량화 제3방법에 대한 축소 해(shrinkage solution) 방법을 제안하였다. Hayashi의 수량화 제3방법에 대한 구체적인 방법은 허명희 (1998, 1999), 岩坪秀一 (1987), 駒澤勉 (1992) 등을 참고할 수 있다.

한편, 허명희 등 (2007)은 q -변량의 수치형 Y 변수에 대한 회귀에서 p -변량의 수치형 X 변수가 다중 공선성의 문제를 갖는 경우에도 적용 가능한 PLS(partial least squares 또는 projected latent structure) 회귀에서 두 종류의 자료에 대한 시각화 방법을 제안한 바 있다. PLS 회귀는 설명변수의 수 p 가 관측치의 수 n 보다 큰 경우에도 기능할 수 있는 선형회귀적 방법이다. 이 방법은 1966년 Herman Wold에 의하여 계량경제적(econometric) 기법으로 창안되었는데 그 이후 여러 분석화학자들이 받아들여 계량화학적(chemometric) 기법으로 활용하고 있다. PLS 회귀에 대한 최근 20여 년에 걸친 연구는 수백 편에 이르고 있어 간단히 정리하기는 어렵다. 1~2편을 고른다면 Helland (2006)의 Encyclopedia of Statistical Sciences 소개 논문과 Rosipal과 Krämer (2006)의 개괄논문을 뽑을 수 있다. 국내 이론 연구자로는 2001년 이후 여러 연구결과를 낸 김중덕 교수가 있고 (김중덕, 2004; Kim, 2001, 2003a, 2003b, 2003c) 응용 연구로는 박성현 등 (1999), 전치혁 등 (2006)이 있다.

허명희 등 (2007)이 제안한 PLS 수량화는 선형계수들의 일정한 제약 하에서, p 개 X 변수의 선형결합과 q 개 Y 변수들의 선형결합 간 공분산을 최대화하는 선형계수를 찾는 것을 목적으로 한다. 즉,

$$\begin{aligned} & \text{maximize (wrt } b \text{ and } c) \quad \text{Cov}(Xb, Yc) & (1.2) \\ & b'b = 1, \quad c'c = 1, \end{aligned}$$

이다. 여기에서 $b : p \times 1$, $c : q \times 1$ 이다. 기본적으로 회귀모형과 실험계획 모형은 선형모형(Linear Model)이라는 하나의 모형으로 정리할 수 있다. 따라서, 위 PLS 회귀모형 방법은 자연스럽게 더미 코드화한 범주형 자료의 모형화에 적용할 수 있다. 즉, 위 모형을 두 범주형 변수 R 과 C 에 적용하면

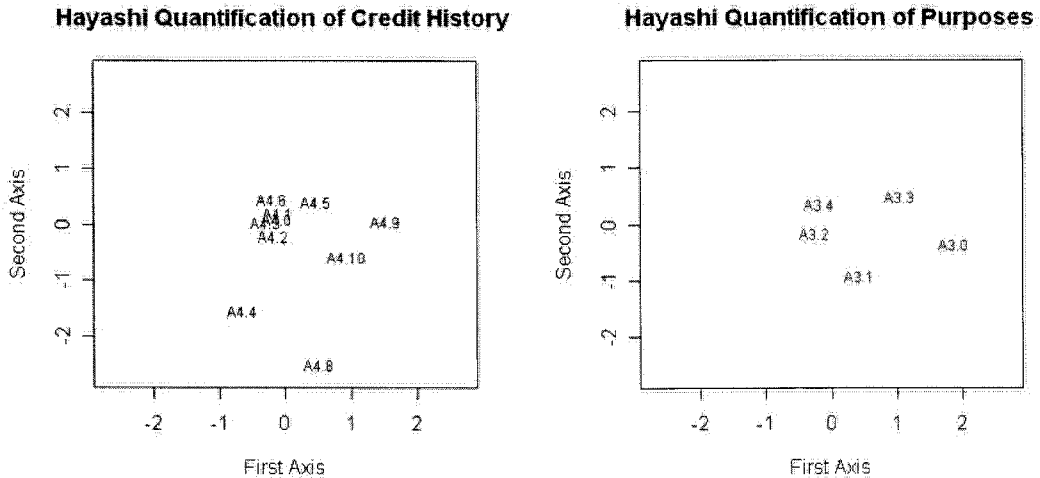
$$\begin{aligned} & \text{maximize (wrt } x \text{ and } y) \quad x' \frac{\bar{F}y}{n} & (1.3) \\ & \text{subject to } x'x = 1, \quad y'y = 1 \end{aligned}$$

와 같다. 여기서 \bar{F} 는 Z_R 과 Z_C 의 평균수정 행렬 간 교적(cross-product)이다. 즉 $\bar{F} = \bar{Z}'_R \bar{Z}'_C$, $\bar{Z}_R = Z_R - 1_n \bar{z}'_R$, $\bar{Z}_C = Z_C - 1_n \bar{z}'_C$, $\bar{z}'_R = 1'_n Z_R$, $\bar{z}'_C = 1'_n Z_C$ 이다.

제1축 수량화는 \bar{Z}_R 과 \bar{Z}_C 의 열들을 각각 $\bar{Z}_{R(x)} (= s)$ 와 $\bar{Z}_{C(y)} (= t)$ 에 사영하여 시각화한다 (허명희 등, 2007). 그리고, \bar{Z}_R 과 \bar{Z}_C 를 s 와 t 에 사영하여 남은 잔차로 대치하여 제2축 수량화에 들어간다. 이러한 PLS 수량화의 장점은 Hayashi 수량화 제3방법과는 달리 반응빈도가 작은 일부 범주들이 특이하게 나타나는 문제가 없다는 것이다. 그 이유는 앞의 PLS 정식화에서 제약식을 살펴보면 명확히 알 수 있다. 다음 절에서 Hayashi 수량화 제3방법과 PLS 수량화 방법에 의한 시각화 그림을 수치 예들 통해 비교해 보고자 한다.

2. German Credit 자료에서 두 범주형 변수 간 연관성

여기에서는 실제 자료를 이용하여 Hayashi 수량화 제3방법과 PLS 수량화에 의한 방법으로 시각



(a) 대출목적(A4. purpose)의 수량화: 신용이력 참조 (b) 신용이력(A3. credit history)의 수량화: 대출목적 참조

그림 1: Hayashi 수량화 제3방법에 의한 대출목적과 신용이력의 수량화

화를 실시하여 두 방법을 비교해 보기로 한다. 여기에서 활용하는 German Credit 자료는 신용대출자 1,000명에 대한 20개 인구사회 및 금융관련 변수와 이항형 결과(good/bad credit) 코드로 구성되어 있다(<http://mllearn.ics.uci.edu/MLSummary.html> 참조). 20개 인구사회 및 금융관련 변수 중 7개는 연속형이고 13개는 범주형이다. 여기에서는 각각의 범주의 수가 10개와 5개인 2개의 금융관련자료를 이용하여 Hayashi 수량화 제3방법과 PLS 수량화를 이용한 시각화를 실시하여 두 방법의 결과를 비교해 보기로 한다. 다음은 선택된 변수들과 변수값에 대한 설명이다.

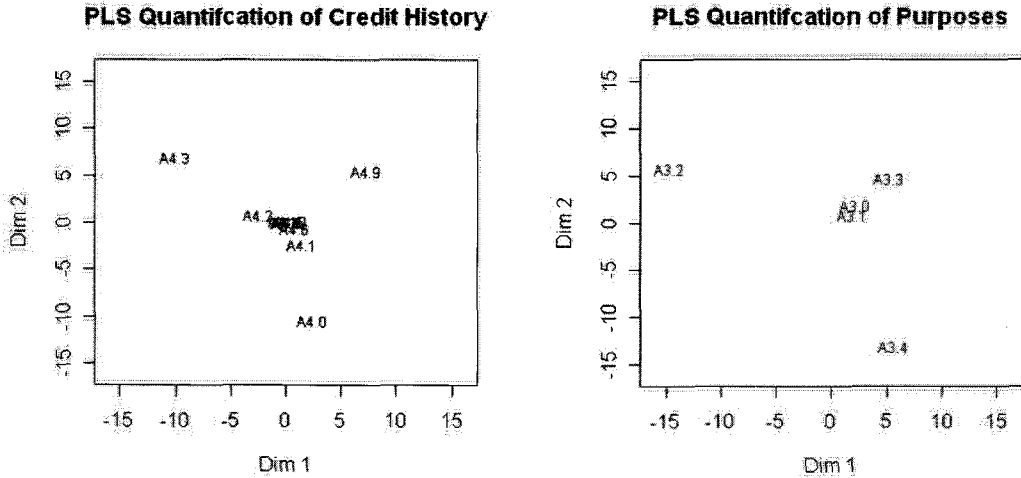
X 변수: A4 대출목적(purpose, 10개 범주)

- | | |
|---------------------------|------------------------|
| A4.0: new car | A4.1: used car |
| A4.2: furniture/equipment | A4.3: radio/television |
| A4.4: domestic appliances | A4.5: repairs |
| A4.6: education | A4.8: retraining |
| A4.9: business | A4.10: others |

Y 변수: A3 신용이력 (credit history, 5개 범주)

- A3.0: no credits taken, all credits paid duly
- A3.1: all credits at this bank paid back duly
- A3.2: existing credits paid back duly till now
- A3.3: delay in paying off in the past
- A3.4: critical account/other credits existing

그림 1은 대출목적 대 신용이력의 Hayashi 수량화 제3방법에 의한 수량화 그림이다. 그림 1(a)에서 범주 A4.4, A4.8과 A4.9가 원점에서 먼 곳에 타점되어 두드러져 보이는데, 확인해보면 A4.4와 A4.8의 빈도는 1.2%와 0.9%에 불과할 정도로 작다. 반면 범주 A4.3과 A4.0은 빈도가 28.0%와 23.4%로 크지만 원점 근처에 묻혀 거의 보이지 않는다. 그림 1(b)에서도 가장 두드러지게 보이는 범주 A3.0은 가장 작은 빈도(4.0%)를 갖는 범주이다. 범주 A3.2가 가장 큰 빈도(53.0%)를 갖지만 원점 근처에 있어 특



(a) 대출목적(A4. purpose)의 수량화: 신용이력 참조 (b) 신용이력(A3. credit history)의 수량화: 대출목적 참조

그림 2: PLS에 의한 대출목적과 신용이력의 수량화

표 1: 대출목적(A4)과 신용이력(A3) 간 교차표: 빈도 (행퍼센트)

	A3.0	A3.1	A3.2	A3.3	A3.4	Total
A4.0	7 (3.0)	12 (5.1)	120 (51.3)	17 (7.3)	78 (33.3)	234 (100)
A4.1	3 (2.9)	5 (4.9)	51 (49.5)	8 (7.8)	36 (35.0)	103 (100)
A4.2	7 (3.9)	8 (4.4)	106 (49.5)	10 (7.8)	50 (35.0)	181 (100)
A4.3	4 (1.4)	9 (3.2)	167 (59.6)	20 (7.1)	80 (28.6)	280 (100)
A4.4	0 (0.0)	1 (8.3)	10 (83.3)	0 (0.0)	1 (8.3)	12 (100)
A4.5	2 (9.1)	0 (0.0)	11 (50.0)	3 (13.6)	6 (27.3)	22 (100)
A4.6	0 (0.0)	3 (6.0)	23 (46.0)	5 (10.0)	19 (38.0)	50 (100)
A4.8	1 (11.1)	2 (22.2)	5 (55.6)	0 (0.0)	1 (11.1)	9 (100)
A4.9	15 (15.5)	7 (7.2)	33 (34.0)	23 (23.7)	19 (19.6)	97 (100)
A4.10	1 (8.3)	2 (16.7)	4 (33.3)	2 (16.7)	3 (25.0)	12 (100)
Total	40 (4.0)	49 (4.9)	530 (53.0)	88 (8.8)	293 (29.3)	1000 (100)

별한 의미를 부여하기 어렵다. 그림 2는 대출목적 대 신용이력의 PLS 수량화 그림이다. 그림 2(a)에서 범주 A4.0, A4.3, A4.9가 두드러지는데 이들 범주들은 각각 23.4%, 28.0%와 9.7%의 빈도를 갖는 큰 범주들이다. 그림 2(b)에서도 마찬가지로 가장 두드러지게 보이는 3개 범주 A3.2, A3.4, A3.3은 각각 53.0%, 29.3%, 8.8%를 갖는 큰 범주들이다.

A4의 Hayashi 수량화 그림 1(a)에서 두드러져 보이는 대출목적(A4)의 3개 범주 A4.4, A4.8, A4.9를 표 1에서 신용이력(A3)과의 연관성을 살펴보면 A4.4는 A3.2와, A4.8은 A3.1과, A4.9는 A3.0과 대응하고 있음을 알 수 있고, 이런 대응 관계는 신용이력(A3)의 Hayashi 수량화 그림 1(b)에서도 확인할 수 있다. 그러나, A4.9만 어느 정도 비중이 있는 범주일 뿐 A4.4와 A4.8은 워낙 비중이 작으므로 이러한 해석은 주목할 시사점이 되지 못한다. A4의 PLS 수량화 그림 2(a)에서 두드러져 보이는 대출목적(A4)의 3개 범주 A4.0, A4.3, A4.9를 표 1에서 신용이력(A3)과의 연관성을 살펴보면 A4.0은 A3.4와,

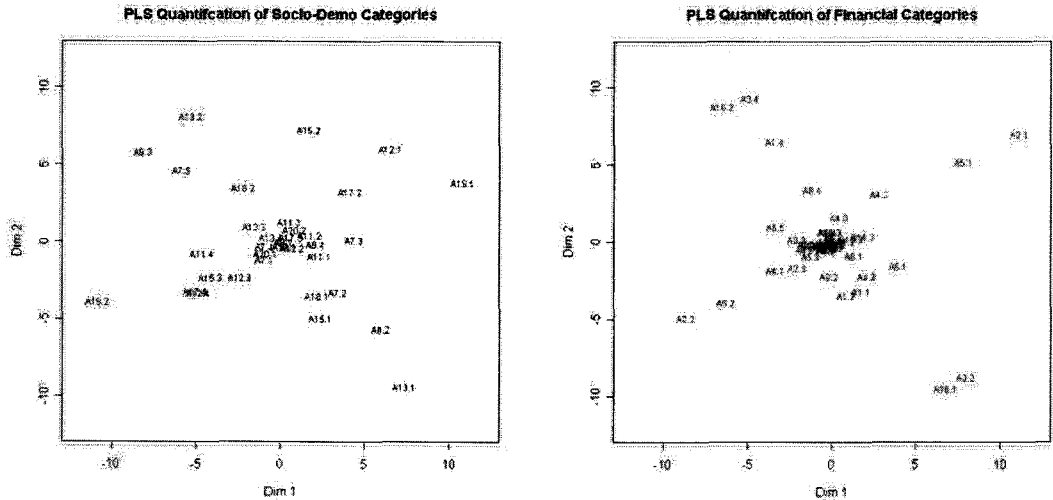
A4.3은 A3.2와, A4.9는 A3.3와 대응되어 있음을 알 수 있고, 이런 대응 관계는 신용이력(A3)의 PLS 수량화 그림 2(b)에서 확인할 수 있다. 그런데, 주요한 대응 범주 중 2개 범주 A3.2와 A3.4가 비중이 상당히 크고 나머지 한 범주 A3.3도 작지 않은 비중을 가지므로 PLS 수량화 결과가 안정성을 지니고 있는 것으로 기대할 수 있다.

3. German Credit 자료에서 인구사회적 범주군과 금융관련 범주군과의 연관성

여기에서는 앞 절에서 제시된 German Credit 자료에서 연속형 변수들도 모두 구간화하여 범주형 변수로 재코딩함으로써 모두 20개의 인구사회 및 금융관련 변수들을 분석할 것이다. 이 때 7개 연속형 변수 중 people은 2개 값만 가지므로 2개 항으로 바꾸고, 그 외 6개 연속형 변수(duration, amount, installment, residence, age, existing)는 각각 등구간 범주화 방법을 이용하여 4개 구간으로 범주화하였다. German Credit 자료의 20개 설명변수를 살펴보면 그 중 10개(A7, A9, A11, A12, A13, A15, A17, A18, A19, A20)는 인구사회적 변수이고 나머지 10개(A1, A2, A3, A4, A5, A6, A8, A10, A14, A16)는 금융관련 변수이다. 여기에서 인구사회적 변수 10개는 총 34개 범주로 되어 있고 금융관련 변수 10개는 총 46개 범주로 되어 있다. 따라서 여기에서는 10개의 인구사회적 변수들의 34개 범주들을 한 변수의 범주군으로 하고, 10개 금융관련 변수들의 46개 범주들을 또 하나의 변수의 범주군으로 하여 PLS 수량화를 이용한 연관성 분석을 실시하고자 한다. 이를 위하여 R과 C를 각각 34개 인구사회적 범주들과 46개 금융관련 범주들에 대한 응답자 반응으로 정의한다. R의 각각의 범주들을 더미 코딩하여 1000×34 행렬을 Z_R 로, C의 각각의 범주들을 더미 코딩하여 1000×46 행렬을 Z_C 로 정의한 후에 ($Z_R : Z_C$)에 대한 PLS 수량화를 실시 하였다. 이와같이 행변수(R)와 열변수(C) 각각의 범주가 34개와 46개로 일반적인 경우보다 상당히 “큰” 2원 교차표에 대한 시각화 작업이라고 할 수 있다. 이와같은 경우에 대한 PLS 시각화의 응용을 이용하여 인구사회적 속성과 금융관련 속성의 연관성을 살펴볼 수 있을 것으로 기대한다. 그림 3은 PLS 수량화 그림을 보여준다. 두 그림을 비교해 볼 때, 4분면마다 1개씩의 군집이 두드러지게 대응을 이루고 있음을 볼 수 있다.

- 제1사분면. 인구사회적 속성: A19.1 (no telephone)
 A12.1 (real estate)
- 금융관련 속성: A2.1 (shortest duration)
 A5.1 (smallest amount)
- 제2사분면. 인구사회적 속성: A13.2 (2nd age group)
 A9.3 (male single)
 A7.5 (employment ≥ 7)
- 금융관련 속성: A16.2 (two credit cards)
 A3.4 (other credits/critical account)
 A1.4 (no checking account)
- 제3사분면. 인구사회적 속성: A19.2 (telephone)
- 금융관련 속성: A2.2 (2nd shortest duration)
 A5.2 (2nd smallest amount)
- 제4사분면. 인구사회적 속성: A13.1 (youngest age)
 A9.2 (female non-single)
- 금융관련 속성: A3.2 (paid back duly)
 A16.1 (one credit card)

앞에 제시된 인구사회적 속성과 금융관련 속성간 연관성을 카이제곱 검증으로 평가한 결과는 다음과 같다.



(a) 인구사회적 속성의 수량화: 금융관련 속성 참조

(b) 금융관련 속성의 수량화: 인구사회적 속성 참조

그림 3: PLS에 의한 인구사회적 속성과 금융관련 속성의 수량화

- 제1사분면 대응군집. 카이제곱 77.4 (자유도 1, p -값 = 0.000)
- 제2사분면 대응군집. 카이제곱 3.6 (자유도 1, p -값 = 0.058)
- 제3사분면 대응군집. 카이제곱 10.5 (자유도 1, p -값 = 0.001)
- 제4사분면 대응군집. 카이제곱 20.4 (자유도 1, p -값 = 0.000)

앞에 제시된 특성들을 갖춘 응답자 그룹과 그렇지 않은 응답자 그룹 간 good credit 비율을 살펴보았다.

- 제1사분면 대응군집. good 비율 81.0% 대 67.9%. 카이제곱 10.6 (자유도 1, p -값 = 0.001)
- 제2사분면 대응군집. good 비율 100% 대 69.5%. 카이제곱 5.6 (자유도 1, p -값 = 0.018)
- 제3사분면 대응군집. good 비율 63.0% 대 70.7%. 카이제곱 0.79 (자유도 1, p -값 = 0.374)
- 제4사분면 대응군집. good 비율 60.7% 대 71.3%. 카이제곱 5.3 (자유도 1, p -값 = 0.022)

따라서 대체로 good credit 속성이 제1사분면과 제2사분면에 위치해있고 bad credit 속성이 제4사분면에 위치해 있음을 알 수 있다. 제3사분면에는 good credit 속성과 bad credit 속성이 혼재해 있다.

4. 맺음말

허명희 (1998)와 허명희와 이용구 (2006)에서 지적한 바와 같이 2원 교차표 자료에 대한 기존의 Hayashi 수량화 제3방법 또는 대응분석에서는 일부 작은 빈도 범주들이 두드러지게 보이는 현상이 발생하여 결과를 왜곡하는 문제점이 있었다. 반면, 이에 대한 대안적 방법으로 본 연구에서 범주형 자료의 시각화 기법으로 제안한 PLS 수량화는 1절에서 설명한 이론적 논리와 2절에서 제시한 실증적 결과에 의할 때, 큰 빈도 범주들에 의하여 주된 영향을 받으므로 안정적인 결과를 보여줄 수 있다. 또한 많은 수의 범주형 변수들이라도 2개 군으로 묶이는 경우에는 2원 교차표 형태로 나타낼 수 있으므로 PLS 수량화로 시각화하는 것이 가능하다. 이에 대한 예제로 본 연구에서는 German Credit 자료의 80개 범주를 외적 기준 없이 수량화한 결과를 제시하였다.

참고 문헌

- 김종덕 (2004). 고유벡터 기저를 이용한 회귀방법의 비교, <한국자료분석학회지>, **6**, 205–218.
- 박성현, 최엄문, 박창순 (1999). 편최소제곱 반응표면합수를 이용한 공정 최적화에 관한 연구, <품질경영학회지>, **27**, 237–250.
- 전치혁, 이해선, 이대원, 장창환 (2006). X-선 회절 데이터에 PLS 기법을 이용한 철광석의 환원을 예측, <한국통계학회 2006년 춘계학술발표회 논문집>, **30**.
- 허명희 (1998). <수량화 방법 I, II, III, IV>, 자유아카데미, 서울.
- 허명희 (1999). <다변량 수량화>, 자유아카데미, 서울.
- 허명희, 이용구 (2006). 수량화 제3방법의 축소 해, <응용통계연구>, **19**, 331–338.
- 허명희, 이용구, 이성근 (2007). PLS 기법에 의한 (X,Y)자료의 시각화, <응용통계연구>, **20**, 345–355.
- 岩坪秀一 (1987). <數量化法の基礎>, 朝倉書店, 동경.
- 駒澤勉 (1992). <數量化理論>, 放送大學教育振興會, 동경.
- Helland, I. (2006). *Partial least squares regression*, The Encyclopedia of Statistical Sciences, 2nd ed.(edited by Kotz), 5997–5962.
- Kim, J. D. (2001). A general weighting scheme of partial least squares regression, *Journal of the Korean Data Analysis Society*, **3**, 11–21.
- Kim, J. D. (2003a). Alternative expressions of regression vector for principal component regression and partial least squares regression, *Journal of the Korean Data Analysis Society*, **5**, 17–26.
- Kim, J. D. (2003b). Projection matrices for partial least squares regression and principal component regression, *Journal of the Korean Data Analysis Society*, **5**, 787–800.
- Kim, J. D. (2003c). Unified non-iterative algorithm for principal component regression, partial least squares and ordinary least squares, *Journal of the Korean Data and Information Science Society*, **14**, 355–366.
- Rosipal, R. and Krämer, N. (2006). Overview and recent advances in partial least squares, *Lecture Notes in Computer Science*, **3940**, 34–51, Springer-Verlag.

Visualizing Large Two-way Crosstabs by PLS Method

Yonggoo Lee^{1,a}, Younim Choi^a

^aDepartment of Statistics, Chung-Ang University

Abstract

On the visualization of categorical data, if the number of categories is small, we can consider Hayashi Quantification Method 3 for visualization of the categories of the variables. But it is known that the method is unstable because it quantifies more significantly for the small frequency categories rather than large frequency categories. The purpose of this research is to propose the visualization of large two-way crosstabulation data by PLS methods for checking the relationship between the categories of row and column variables. In this research, we utilize the PLS visualization methods (Huh *et al.*, 2007) that is proposed for visualization of the qualitative data to visualize the categories of the large categorical data. We also compared both methods by applying them to real data, and studied the results from PLS visualization method on the real categorized data with many categories.

Keywords: Hayashi Quantification III, PLS Quantification, large two-way cross-tabs.

This Research was supported by the Chung-Ang University Research Grants in 2007.

¹ Corresponding author: Professor, Department of Statistics, Chung-Ang University, 221 Heukseok-Dong Dongjak-Gu, Seoul 156-756, Korea. E-mail: leeyg@cau.ac.kr