

현대시 구조의 사영탐색

안 주 선 1)

요약

현대시로부터 c^d -분할표를 생성하여 그의 구조적 특징을 사영탐색-플롯(Projection Pursuit-plot)을 이용하여 조사하는 방법을 소개하고, 여러 시집에서 자주 인용된 김소월 시와 서정주 시들에 적용하여 유사성을 비교한다.

주요용어: c^d -분할표, 사영탐색-플롯, 유사성

1. 서론

현대시의 구조적 특징은 "시에 율격구조(6.4조 또는 7.5조)를 갖는지, 작가가 선호한 음절이 짧은 것인지 또는 긴 것인지, 시가 민요조 시인지 또는 산문조 시인지, 표현이 은유적인지 또는 직설적인지" 등이 복합적으로 나타난 결과로 볼 수 있다[장만영, 박목월. (1979), 송하선. (1998) 참조].

이러한 현대시 구조적 특징을 사영탐색-플롯(Projection Pursuit-plot, 간단히 PP-plot)을 이용하여 탐색하는 방법을 소개한다. PP-plot은 Ahn et al(2003)에서 소개된 방법으로 Radon 변환을 c^d -분할표에 적용하여 자료를 탐색하는 그래프이다. c^d -분할표에서 c 가 소수이면 PP-자료를 만드는 변환행렬은 역행렬을 갖고, PP-자료들 사이의 거리제곱은 원 자료들 사이의 거리제곱의 상수배(c 와 d 에만 의존)인 성질을 갖는다. PP-자료는 분할표의 $\frac{c^d-1}{d-1}$ 방향에서 얻어진다.

PP-plot을 현대시에 적용하기 위해서는 먼저 각각의 시로부터 c^d -분할표 자료를 만들고, 그 분할표를 PP-자료로 변환해야 한다. 본 논문에서는 현대시로부터 c^d -분할표를 만드는 과정을 설명하고 PP-자료를 만들어 PP-플롯을 그려 어떤 시가 특별한 구조를 갖는지 또는 유사한 구조를 갖는지를 조사하는 방법을 보인다. 끝으로 위의 방법을 김정식(1977)과 서정주(1994)시집에서 몇 개씩을 선택하여 응용한다.

1) 강릉대학교 정보통계학과 교수, 강릉시 지변동 123

현대시 구조의 사영탐색

2. 자료생성

현대시의 구조적 특성을 탐색하기 위해 다음 세 단계로 분할표 자료를 생성한다.

1단계) 주어진 시에서 각 음절의 글자 수를 조사한다.

2단계) 범주의 수(c)를 2로 정하고 글자수가 3보다 작으면 짧은 음절로 보고 범주 0을 주며 글자수가 3보다 크거나 같으면 긴 음절로 보고 범주 1을 준다.

3단계) 시를 감상할 때 연(또는 문장)의 구별이 중요한 요소이므로, 한 연(또는 문장)에서 d음절씩 나누고, 분할한 d음절의 순서쌍들을 d차원 분할표의 셀 (x_1, x_2, \dots, x_d) 의 객체로 보고 c^d -분할 표를 만든다.

3. 사영탐색 플롯

소월시 중에서 음절수가 60미만 인 12개의 시 “1. 가는 길(ganeun), 2. 가시나무(gasi), 3. 개여울의 노래(gae), 4. 고적한 날(gojuk), 5. 먼 후일(mern), 6. 못 잊어(mots), 7. 봄밤(bombam), 8. 산유화(san), 9. 예전엔 미쳐 몰랐어요(yejun), 10. 왕십리(wang), 11. 진달래꽃(jindal), 12. 합장(habjang)”의 음절수와 짧은 음절의 비율, 3차원 분할표, PP-자료와 PP-플롯은 [표3.1], [표3.2], [표3.3]과 [그림 3.1]에 주어진다.

[표 3.1] 음절수와 짧은 음절의 비율

시	1	2	3	4	5	6	7	8	9	10	11	12	Mean
음절수	33	34	52	44	28	34	36	38	32	51	36	59	39.8
비율	0.52	0.41	0.44	0.57	0.32	0.35	0.42	0.82	0.50	0.49	0.56	0.54	0.49

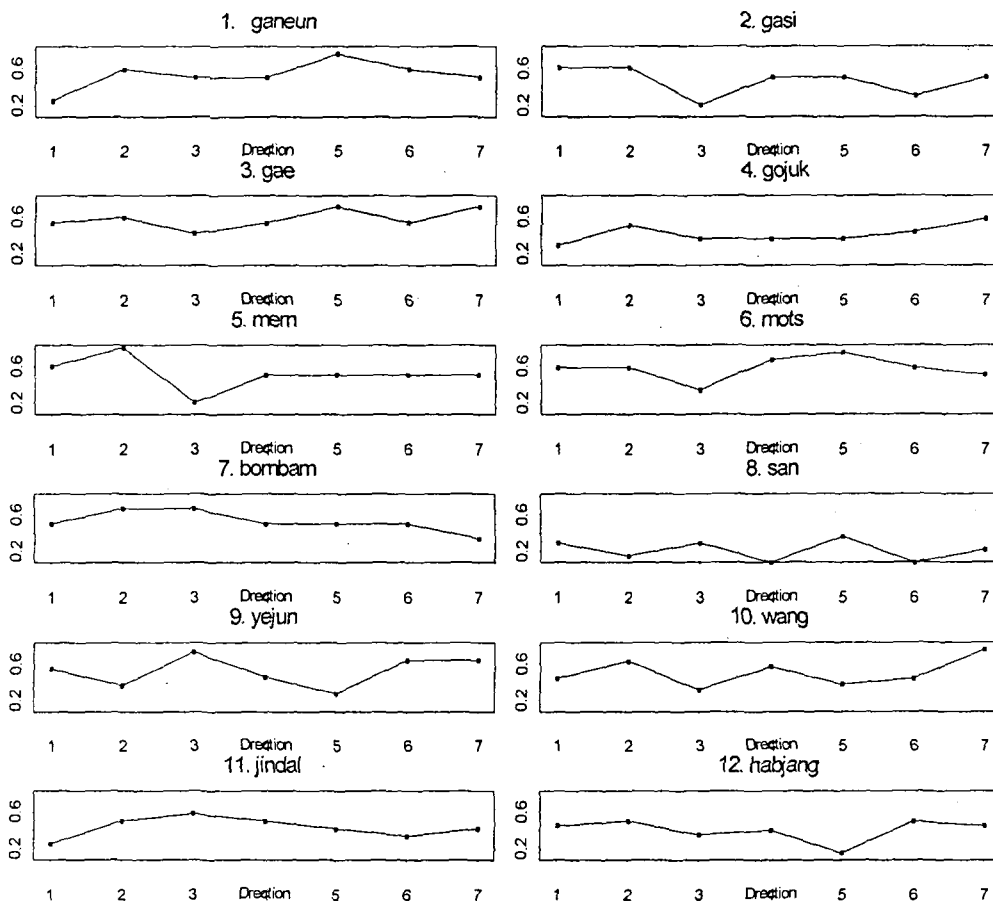
[표 3.2] 2^3 -분할표

셀\시	1	2	3	4	5	6	7	8	9	10	11	12
111	0	3	3	2	3	2	1	0	1	5	1	4
110	2	2	3	1	3	3	2	1	0	1	0	2
101	0	0	0	0	0	1	2	0	3	0	1	3
100	1	1	3	1	0	1	1	3	2	1	1	0
011	3	1	2	1	1	2	3	1	0	1	3	0
010	2	0	2	3	1	0	2	0	3	3	2	4
001	3	1	4	2	1	3	0	0	1	3	1	1
000	1	2	0	4	1	0	1	9	1	2	3	6

[표 3.3] [표 3.2]의 PP-자료

방향\시	1	2	3	4	5	6	7	8	9	10	11	12
1	0.25	0.60	0.53	0.29	0.60	0.58	0.50	0.29	0.55	0.44	0.25	0.45
2	0.58	0.60	0.59	0.50	0.80	0.58	0.67	0.14	0.36	0.63	0.50	0.50
3	0.50	0.20	0.41	0.36	0.20	0.33	0.67	0.29	0.73	0.31	0.58	0.35
4	0.50	0.50	0.53	0.36	0.50	0.67	0.50	0.07	0.45	0.56	0.50	0.40
5	0.75	0.50	0.71	0.36	0.50	0.75	0.50	0.36	0.27	0.38	0.42	0.15
6	0.58	0.30	0.53	0.43	0.50	0.58	0.50	0.07	0.64	0.44	0.33	0.50
7	0.50	0.50	0.71	0.57	0.50	0.50	0.33	0.21	0.64	0.75	0.42	0.45

[그림 3.1] [표 3.2]의 PP-플롯



산유화(8. san)는 PP-자료와 플롯에서 낮은 PP값을 가지며 유사한 것이 없음을 볼 수 있으나 봄밤(7. bombam)과 진달래꽃(11. jindal)은 꺾임이 완만한 유사한 플롯을 보이고 있다.

현대시 구조의 사영탐색

음절수가 60이상인 6개의 소월시 “ 1. 가을 아침에(so-ga), 2. 님의 노래(so-nim), 3. 벗 마을(so-buts), 4. 오는 봄(so-oneun), 5. 접동새(so-jub), 6. 초혼(so-cho)” 와 6개의 미당시 “7. 귀축도(mi-gyu), 8. 무등을 보며(mi-moo), 9. 자화상(mi-jawha), 10. 천사(mi-chu), 11. 학(mi-hag), 12. 화사(mi-whasa)”에 대한 4차원 분할표와 PP-플롯은 [표 3.4]와 [그림 3.2]에 주어진다.

[표 3.4] 6개의 소월시와 6개의 미당시의 PP-자료

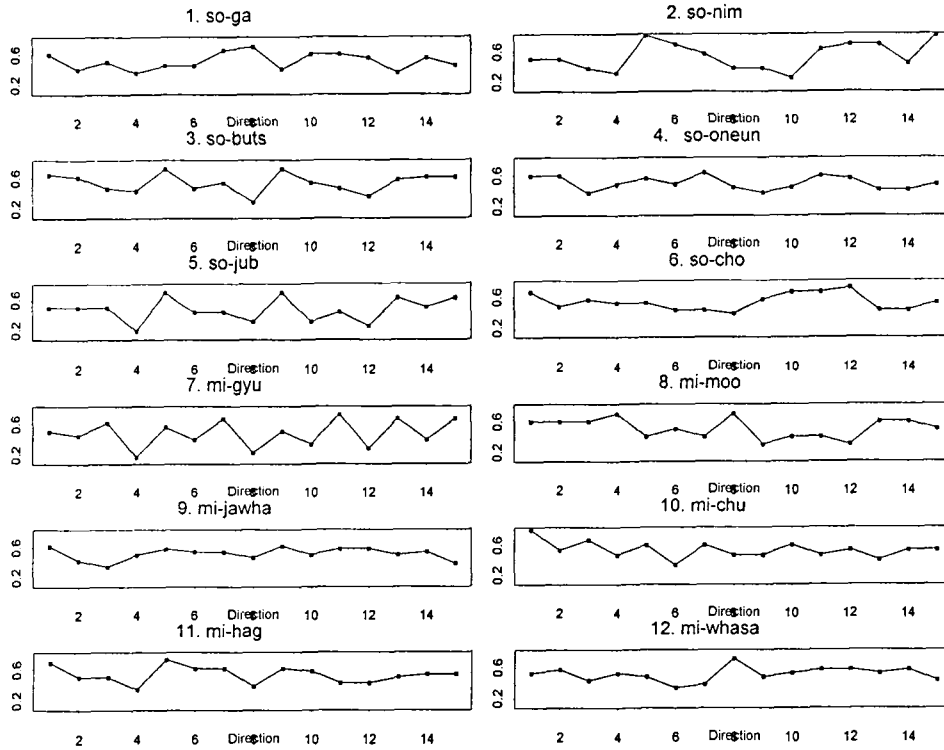
방향\시	1	2	3	4	5	6	7	8	9	10	11	12
1	0.60	0.50	0.64	0.59	0.50	0.67	0.50	0.60	0.62	0.80	0.69	0.52
2	0.40	0.50	0.61	0.59	0.50	0.48	0.44	0.60	0.42	0.53	0.50	0.57
3	0.50	0.38	0.46	0.37	0.50	0.57	0.61	0.60	0.35	0.67	0.50	0.43
4	0.35	0.31	0.43	0.48	0.19	0.52	0.17	0.70	0.50	0.47	0.35	0.52
5	0.45	0.81	0.71	0.56	0.69	0.52	0.56	0.40	0.58	0.60	0.73	0.48
6	0.45	0.69	0.46	0.48	0.44	0.43	0.39	0.50	0.54	0.33	0.62	0.33
7	0.65	0.56	0.54	0.63	0.44	0.43	0.67	0.40	0.54	0.60	0.62	0.38
8	0.70	0.38	0.29	0.44	0.31	0.38	0.22	0.70	0.46	0.47	0.38	0.71
9	0.40	0.38	0.71	0.37	0.69	0.57	0.50	0.30	0.62	0.47	0.62	0.48
10	0.60	0.25	0.54	0.44	0.31	0.67	0.33	0.40	0.50	0.60	0.58	0.52
11	0.60	0.63	0.46	0.59	0.44	0.67	0.72	0.40	0.58	0.47	0.42	0.57
12	0.55	0.69	0.36	0.56	0.25	0.71	0.28	0.30	0.58	0.53	0.42	0.57
13	0.35	0.69	0.57	0.41	0.63	0.43	0.67	0.60	0.50	0.40	0.50	0.52
14	0.55	0.44	0.61	0.41	0.50	0.43	0.39	0.60	0.54	0.53	0.54	0.57
15	0.45	0.81	0.61	0.48	0.63	0.52	0.67	0.50	0.38	0.53	0.54	0.43

소월시 벗마을(3. so-buts)과 미당시 학(11.-hag), 소월시 접동새(5. so-jub)와 미당시 귀축도(7. mi-gyu) 등은 매우 유사한 구조를 가지고 있음을 볼 수 있다. 다른 여러 시들의 유사성 비교는 독자들의 몫으로 돌리고자 한다.

4. 결론

PP-플롯으로부터 현대시 구조를 비교하고자할 때 수형도(Clustering Tree)를 참고하면 도움이 될 것이다. 짧은 음절과 긴 음절의 경계는 분석자의 몫일 것이나 인용된 12개 소월 시의 경우 음절수가 3미만인 것의 비율이 49%임을 볼 수 있으므로 2절에서와 같이 경계를 3으로 정하는 것이 좋을 것이다. 범주의 수(c)와 차원(d)은 각 셀의 평균적인 도수가 너무 적지 않게 정해야 할 것이다. 음절의 개수가 60미만인 짧은 시의 경우는 2^3 -분할표, 60이상인 시의 경우 2^4 -분할표를 만들어 조사하는 것이 좋을 것으로 생각된다.

[그림 3.2] [표 3.4]의 PP-플롯



참고문헌

김정식. (1977). < 소월시전집 >, 성공문화사.
 서정주.(1994). <미당 서정주 자선시집 민들레꽃>, 정우사
 송하선. (1998). < 한국명시해설 >, 국학자료원.
 장만영, 박목월. (1979). < 소월시감상 >, 박영사.
 Ahn, J.S., Hofmann, H. and Cook, D. (2003).
 A Projection Pursuit Method on the multidimensional squared
 Contingency Table. *Computational Statistics*. Vol. 18. 605-626.