

メディア情報学実験・メディア分析 課題レポート

1510151 柳 裕太

2017 年 11 月 5 日

1 序論・仮説

今回の実験では、”きよしのズンドコ節”という曲の PV を扱った (PV16)。この PV では昭和の歌謡曲のエッセンスを入れたことによって、筆者のような 20 代前半にとっては個人的にはかなり印象深い PV であった。よって今回は、以下の仮説を立ててから解析に臨むことにした。

- PV を構成する任意の要素が洗練されていなくとも、好感度には影響しない
- 映像・メロディに迫力がなくとも、好感度には影響しない
- 聞き取りやすいメロディ・歌詞は好感度上昇に寄与する

2 調査結果分析

2.1 a-sum 分析 (調査全体)

2.1.1 主成分抜粋

主成分抜粋においては、累積寄与率と固有値の 2 つのデータを基準に足切りを行った。なお、PC5 以降は省略している。

表 1: 全体調査に対する主成分毎の累積寄与率と固有値

主成分番号	累積寄与率 (%)	固有値
PC1	43.83105671	8.327900774
PC2	68.24472267	4.638596532
PC3	83.05686358	2.814306774
PC4	90.02127311	1.32323781
PC5	93.34194823	0.630928273

講義内では、以下の条件で足切りすることが推奨されていた。

- 累積寄与率が 80% 以下の主成分
- 固有値が 1 以上の主成分

前者であれば PC2、後者であれば PC4 までとなるが、PC2 で足切りした場合、固有値が 2 を上

回る PC3 をも切り捨てる必要が生じることになる。よって、後者の広い基準を採用し、PC4 まで後の解析にかけることにした。

2.1.2 重回帰式による検証

PC1 から PC4 まで全ての主成分を対象に重回帰分析を行った。その結果は以下の通りである。

表 2: 重回帰分析結果: 各主成分の係数

主成分番号	偏回帰係数	標準誤差	t 値	P 値	標準化偏回帰係数	トレランス
PC1	-25.7	1.07	-23.9	2.29e-13	-0.941	1
PC2	6.67	1.44	4.65	0.000317	0.183	1
PC3	-10.5	1.84	-5.71	4.11e-05	-0.224	1
PC4	6.00	2.69	2.23	0.0413	0.0877	1
定数項	829	3.09	268	5.06e-29	NA	NA

表 3: 再重回帰分析結果: 重相関係数・自由度調整済重相関係数

重相関係数	重相関係数の 2 乗	自由度調整済重相関係数の 2 乗
0.988353434669365	0.97684251182273	0.970667181642125

表 4: 再重回帰分析結果: 分散分析

項目	平方和	自由度	平均平方	F 値	P 値
回帰	1.21e+05	4	3.02e+04	158	4.52e-12
残差	2.87e+03	15	191	NA	NA
全体	1.24e+05	19	6.51e+03	NA	NA

これを重回帰式にすると、以下の通りとなった。なお、目的変数は PV_{like} とした。

$$PV_{like} = -25.7PC_1 + 6.67PC_2 - 10.5PC_3 + 6.00PC_4 + 829 \quad (1)$$

この目的変数 PV_{like} に対して PC_1 から PC_4 までを説明変数として重回帰分析を行った結果、決定係数は $R^2 = .98$ であり、0.1% 水準で有意であった。 $(F(4, 15) = 158)$

2.1.3 目的関数に寄与する主成分の選定

主成分の選定において着目したのが P 値の列である。講義内では $p < 0.005$ が望ましいとされた。今回の解析では、表 2 の P 値の列によると、いずれの主成分もこの基準を満たしているため、全ての主成分を残すべきであることが適当であると判断した。

2.1.4 主成分を構成する質問・主成分命名

PV16 の主成分を構成する質問の選出には、各質問に対応した主成分負荷量を対象とし、それらの絶対値が大きい値を持つ質問を選抜することにした。

以下の表が、各主成分に対して主成分負荷量の絶対値が最も大きかった 5 つの質問とその負荷量を表している。

表 5: 各主成分に対する主成分負荷量絶対値 TOP5

(a) PC1		(b) PC2	
質問番号	主成分負荷量	質問番号	主成分負荷量
Q14	-0.935	Q7	0.874
Q17	-0.933	Q1	0.784
Q19	-0.902	Q9	0.781
Q3	-0.876	Q10	0.703
Q11	-0.857	Q13	0.659

(c) PC3		(d) PC4	
質問番号	主成分負荷量	質問番号	主成分負荷量
Q1	0.566	Q8	-0.615
Q4	0.533	Q2	-0.401
Q18	-0.499	Q18	-0.398
Q9	-0.495	Q14	0.296
Q15	0.460	Q6	0.281

このデータを元に、各主成分に命名した一覧は以下の通りである。

PC1 好感度・購買欲

- 各要素に対する好感度を問うた質問が主成分負荷量絶対値の上位の多くを占めたため
- それに加え、Q19:購買意欲を問うた質問も上位に入っていたため

PC2 映像・雰囲気のカオリティ

- 映像の洗練性と迫力、そして雰囲気のカオリティを問うた質問が主成分負荷量絶対値の上位 3 つであつたため
- その 2 問の主成分負荷量絶対値が以下の質問に対して差をあけていたため

PC3 映像・アーティストの洗練性

- 映像・アーティストの洗練性を問うた質問が主成分負荷量絶対値の上位 2 つを占めたため

PC4 映像の見やすさ

- 映像の見やすさを問うた Q8 が主成分負荷量絶対値の最上位だったため

- Q8 以降との主成分負荷量絶対値の差が 0.2 以上開いていたため

2.1.5 グループの類推

調査結果に対して、クラスタ解析を行った結果が以下の図 1 の通りである。

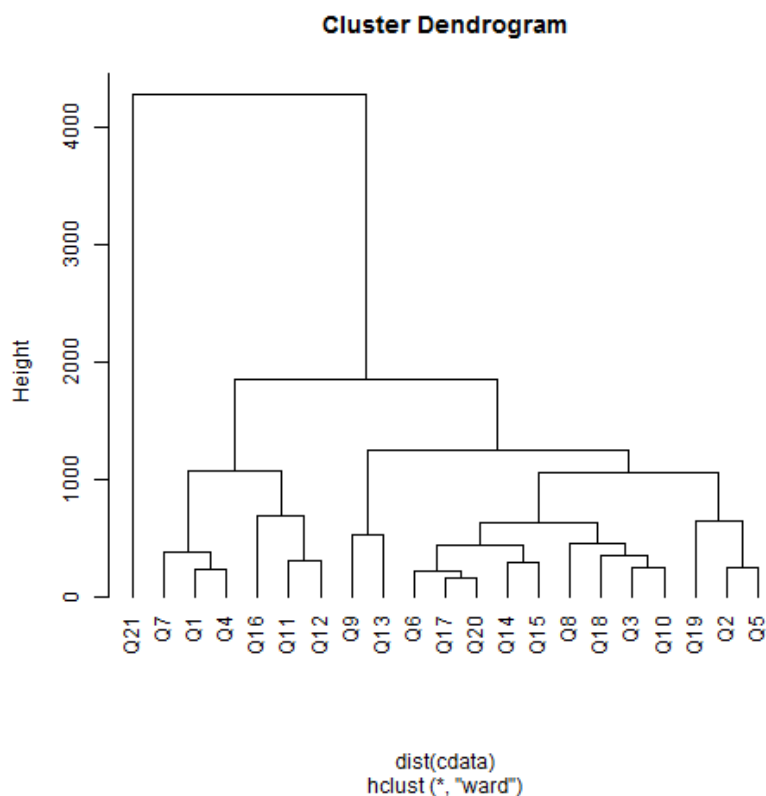


図 1: 全体調査結果に対するクラスタ分析結果

目的変数 Q20 に最も近い位置の質問が Q17 となった。その Q14 は、表 5 より PC1 を構成する質問の 1 つでもあった。また、図左部分に各要素の洗練性を問うた質問が集中している他、各主成分を構成していた質問が近い関係にある場合が多かった。

2.2 PV16 分析 (きよしのズンドコ節)

2.2.1 主成分抜粋

主成分抜粋においては、累積寄与率と固有値の 2 つのデータを基準に足切りを行った。なお、PC11 以降は省略している。

こちらの解析でも、講義内で推奨された基準を適用することにした。

前者であれば PC8、後者であれば PC4 までとなるが、両者のデータ共に値が著しく変化する境界があまり明瞭ではない。そこで、前者の広い基準を採用し、解析後の P 値等によって解析対象から外すことにした。

表 6: PV16 に対する主成分毎の累積寄与率と固有値

主成分番号	累積寄与率 (%)	固有値
PC1	38.38390791	7.29294250348674
PC2	48.55050945	1.93165429251404
PC3	55.79636735	1.37671299972187
PC4	61.89941496	1.15957904665
PC5	67.01570809	0.972095695622575
PC6	71.71468219	0.892805078337444
PC7	75.76539908	0.769636208492113
PC8	79.51202683	0.711859273699518
PC9	82.66178688	0.598454408935556
PC10	85.20054444	0.482363935634707

2.2.2 重回帰式による検証

PC1 から PC8 まで全ての主成分を対象に重回帰分析を行った。その結果は以下の通りである。

表 7: 重回帰分析結果: 各主成分の係数

主成分番号	偏回帰係数	標準誤差	t 値	P 値	標準化偏回帰係数	トレランス
PC1	-0.282	0.0119	-23.6	2.09e-64	-0.765	1
PC2	-0.171	0.0232	-7.38	2.66e-12	-0.239	1
PC3	-0.197	0.0275	-7.15	1.05e-11	-0.232	1
PC4	-0.0489	0.0300	-1.63	0.104	-0.0529	1
PC5	0.143	0.0327	4.39	1.72e-05	0.142	1
PC6	0.105	0.0341	3.07	0.00242	0.0994	1
PC7	-0.156	0.0368	-4.24	3.26e-05	-0.137	1
PC8	-0.00994	0.0382	-0.260	0.795	-0.00843	1
定数項	3.15	0.0323	97.7	2.64e-195	NA	NA

表 8: 再重回帰分析結果: 重相関係数・自由度調整済重相関係数

重相関係数	重相関係数の 2 乗	自由度調整済重相関係数の 2 乗
0.864590803140089	0.747517256874424	0.739101165436905

これを重回帰式にすると、以下の通りとなった。なお、目的変数は PV_{like} とした。

$$\begin{aligned}
 PV_{like} = & -0.282PC_1 - 0.171PC_2 - 0.197PC_3 - 0.0489PC_4 \\
 & + 0.143PC_5 + 0.105PC_6 - 0.156PC_7 - 0.00994PC_8 + 3.15
 \end{aligned} \tag{2}$$

表 9: 再重回帰分析結果: 分散分析

項目	平方和	自由度	平均平方	F 値	P 値
回帰	184	8	23.0	88.8	2.36e-67
残差	62.2	240	0.259	NA	NA
全体	246	248	0.993	NA	NA

この目的変数 PV_{like} に対して PC_1 から PC_8 までを説明変数として重回帰分析を行った結果、決定係数は $R^2 = .75$ であり、0.1% 水準で有意であった。 $(F(8, 240) = 88.8)$

2.2.3 目的関数に寄与する主成分の選定

第 2.1.3 節同様に、 $p < 0.005$ を基準に主成分を選定することにした。今回の解析では、表 7 の P 値の列によると、 PC_4 , PC_8 がこの基準を満たしていないため、除外して重回帰分析を再度行うことが適当であると判断した。

2.2.4 重回帰式による再検証

第 2.2.3 節より、 PC_4 , PC_8 を除外した状態でもう一度重回帰分析を行った。その結果得られたデータは次の表 10 の通りである。

表 10: 再重回帰分析結果: 各種成分の係数

主成分番号	偏回帰係数	標準誤差	t 値	P 値	標準化偏回帰係数	トレランス
PC1	-0.282	0.0120	-23.5	1.64E-64	-0.765	1
PC2	-0.171	0.0232	-7.36	2.79E-12	-0.240	1
PC3	-0.197	0.0275	-7.14	1.10E-11	-0.232	1
PC5	0.143	0.0328	4.38	1.77E-05	0.142	1
PC6	0.105	0.0342	3.06	0.00245	0.0994	1
PC7	-0.156	0.0368	-4.23	3.33E-05	-0.137	1
定数項	3.15	0.0323	97.6	2.45E-196	NA	NA

表 11: 再重回帰分析結果: 重相関係数・自由度調整済重相関係数

重相関係数	重相関係数の 2 乗	自由度調整済重相関係数の 2 乗
0.862928027094451	0.744644779945121	0.738313658786736

これを重回帰式にすると、以下の式 3 の通りとなった。

$$PV_{like} = -0.282PC_1 - 0.171PC_2 - 0.197PC_3 - 0.143PC_5 + 0.105PC_6 - 0.156PC_7 + 3.15 \quad (3)$$

表 12: 再重回帰分析結果: 分散分析

項目	平方和	自由度	平均平方	F 値	P 値
回帰	183	6	30.6	118	7.69e-69
残差	62.9	242	0.260	NA	NA
全体	246	248	0.993	NA	NA

この分析結果をまとめると、目的変数 PV_{like} に対して $PC_1, PC_2, PC_3, PC_5, PC_6, PC_7$ までを説明変数として重回帰分析を行った結果、決定係数は $R^2 = .74$ であり、0.1% 水準で有意であった。 $(F(6, 242) = 118)$

2.2.5 主成分を構成する質問・主成分命名

第 2.1.4 節と同様に、PV16 の主成分を構成する質問の選出には、各質問に対応した主成分負荷量を対象とし、それらの絶対値が大きい値を持つ質問を選抜することにした。

以下の表 13 が、各主成分に対して主成分負荷量の絶対値が最も大きかった 5 つの質問とその負荷量を表している。

表 13: 各主成分に対する主成分負荷量絶対値 TOP5

(a) PC1		(b) PC2		(c) PC3	
質問番号	主成分負荷量	質問番号	主成分負荷量	質問番号	主成分負荷量
Q14	-0.801	Q1	0.604	Q16	0.502
Q17	-0.756	Q7	0.548	Q12	0.478
Q3	-0.727	Q2	-0.418	Q9	-0.404
Q10	-0.724	Q11	0.414	Q5	0.347
Q6	-0.721	Q4	0.402	Q10	-0.345

(d) PC5		(e) PC6		(f) PC7	
質問番号	主成分負荷量	質問番号	主成分負荷量	質問番号	主成分負荷量
Q7	-0.366	Q8	0.568	Q18	-0.609
Q15	0.328	Q5	-0.438	Q2	0.250
Q11	0.319	Q4	-0.353	Q5	0.221
Q17	0.299	Q6	-0.338	Q15	0.211
Q9	-0.295	Q15	0.194	Q19	0.207

このデータを元に、各主成分に命名した一覧は以下の通りである。

PC1 好感度

- 各要素に対する好感度を問うた質問が主成分負荷量絶対値の上位を独占したため

- 好感度を問うた質問同士の主成分負荷量絶対値の差があまりなかったため

PC2 非映像洗練性

- 雰囲気・映像に対する洗練性を問うた質問が主成分負荷量絶対値の上位 2 つであったため
- その 2 問の主成分負荷量絶対値が以下の質問に対して差をあけていたため
- なお、PC2 の重回帰式上の係数が負であるため、洗練されていないほど目的変数上昇に寄与することを意味する

PC3 聴きやすさ

- 歌詞の分かりやすさ・メロディの聞き取りやすさが主成分負荷量絶対値の上位 2 つであったため
- その 2 問の主成分負荷量絶対値が以下の質問に対して差をあけていたため
- なお、PC3 の重回帰式上の係数が負であるため、洗練されていないほど目的変数上昇に寄与することを意味する

PC5 楽曲洗練性

- 楽曲の歌詞・メロディに対する洗練性を問うた質問が上位に入っていたため
- なお、歌詞・メロディのみならず、PC2 で採用された映像に対する洗練性を問うた質問も上位に入っていた。
 - － しかしこの質問の主成分負荷量が負だったため、洗練されていないほど目的変数上昇に寄与することを意味する。
 - － 楽曲の歌詞・メロディに対する洗練性を問うた質問の主成分負荷量は正であった

PC6 映像に対するアーティストの非親和性

- アーティストの親しみやすさと映像の見やすさを問うた質問が上位 2 つを占めたため
 - － 主成分負荷量の正負では、Q8:映像の見やすさが正、Q5:アーティストの親しみやすさが負であった。
- その 2 問の主成分負荷量絶対値が以下の質問に対して差をあけていたため

PC7 印象度

- PV の印象度を問うた質問が、他の質問の主成分負荷量絶対値に対して大差をつけて最上位となっていたため

2.2.6 グループの類推

調査結果に対して、クラスタ解析を行った結果が以下の図 2 の通りである。目的変数 Q20 に最も近い位置の質問が Q14 となった。その Q14 は、PC1 における主成分負荷量絶対値が最大だった質問でもあった。また、図左部分に各要素の洗練性を問うた質問が集中している他、各主成分を構成していた質問が近い関係にある場合が多かった。なお、PC7 を構成していた PV の印象度を問うた Q18 に最も近かったのは、メロディの聴きやすさを問うた Q12 であった。

3 結論

4 考察

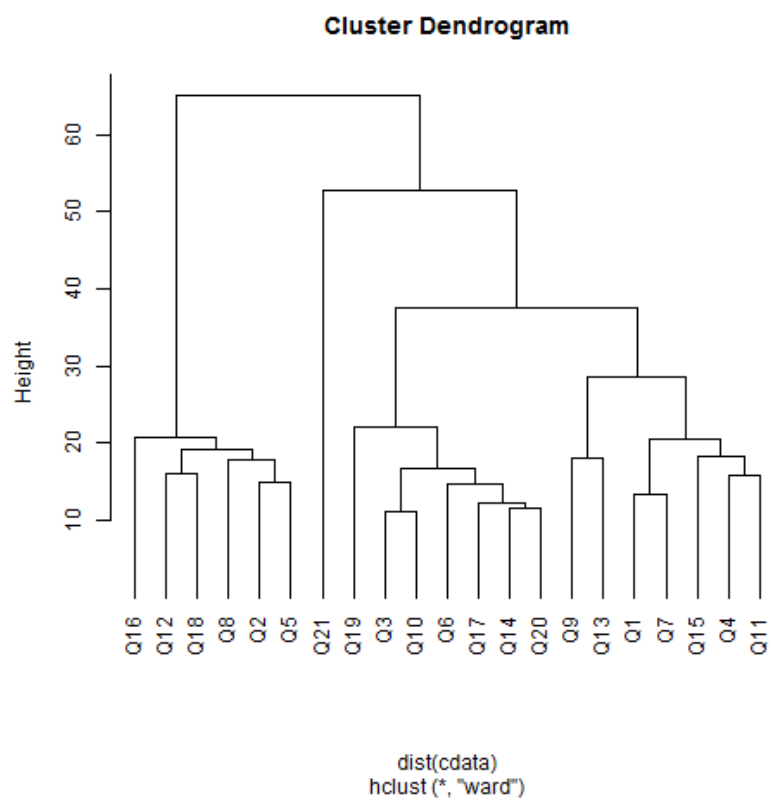


図 2: PV16 調査結果に対するクラスタ分析結果