

# 感性情報処理

## ～中間レポート～

学籍番号: 26115065

氏名: 佐竹 寛弥

提出日: 2016 年 6 月 30 日

### 目次

<b>1</b>	<b>問題</b>	<b>2</b>
1.1	課題 1 . . . . .	3
1.2	課題 2 . . . . .	4
1.3	課題 3 . . . . .	4

## 1 問題

ある小学校でランダムに選んだ児童 11 名に、「次の野菜のうち、好きなものはどれですか?」と聞いたときの回答が、表 1 のようになった。

表 1: 小学校における好き嫌い調査（好きと答えた食べ物に○）

サンプル			カテゴリー								
			1	2	3	4	5	6	7	8	9
番号	学年	性別	にんじん	トマト	キュウリ	大根	いんげん	キャベツ	なす	タマネギ	レタス
1	1	女		○		○			○		○
2	1	女		○	○	○	○	○	○	○	
3	1	男			○	○			○		
4	1	男				○			○	○	
5	2	女		○	○			○	○		○
6	2	女		○		○			○	○	
7	2	男		○						○	○
8	3	女		○			○	○	○		○
9	3	女	○	○	○	○	○				○
10	3	女		○			○		○		
11	3	男	○	○	○	○		○		○	○

この表 1 の反応点について、サンプルとカテゴリーの相関が最大になるようにサンプル及びカテゴリーの数量化を行ったところ、相関値の高い順に表 2・3 のような結果が得られた。

表 2: 数量化の結果

サンプル	成分 1	成分 2
1	-0.37	0.30
2	-0.14	0.38
3	-1.04	-0.28
4	-2.40	0.02
5	0.73	0.12
6	-1.63	0.23
7	-0.27	-0.49
8	1.12	1.28
9	1.07	-1.00
10	0.56	2.38
11	0.38	-1.52

表 3: 数量化の結果

サンプル	成分 1	成分 2
1	1.44	-2.75
2	0.33	0.41
3	0.41	-1.00
4	-1.18	-0.58
5	1.30	1.66
6	1.06	0.14
7	-0.79	1.21
8	-1.63	-0.61
9	0.89	-0.48

### 1.1 課題 1

#### 課題

表 1 より表 2・3 の結果を得るためには, どのような処理が必要か, 文章で簡単に説明しなさい.

数量化 3 類の計算手順を元に説明する. 始めに, 各カテゴリーとサンプルに対して未定値  $x_i, y_i$  を割り当て, これら値の相関係数  $r_{xy}$  が最大になるように  $x_i, y_i$  を定める. ただし, 解を一意に定めるため,  $\bar{x} = \bar{y} = 0$  と仮定して計算する. さらに  $(x, y)$  を  $(v, w)$  と正規化し, "その平方和が 1 になる" という条件を加え, ラグランジュの未定乗数法を用いて相関係数を最大化するような  $x_i, y_i$  を求める. すると, 固有値問題に帰結する. 得られた固有値の大きいものから採用し, その固有値に対応する固有ベクトルが, 各カテゴリー, サンプルに割り当てべき数値である. ただし, 第 1 固有値に対する固有ベクトルは,  $\bar{x} = \bar{y} = 0$  という条件を満たさないで, 除外する. 第 2・3 固有値が, 成分 2・3 に対応する.

## 1.2 課題 2

### 課題

表 1 を解析しやすいように, 表 2・3 の成分 1 に従って並べなおしなさい.

表 2・3 の成分 1 の値を, それぞれカテゴリー, サンプルに割り当て, その値が降順になるようにソートする. ソートした結果を表 4 に示す.

表 4: 成分 1 によるソーティング後のデータ

				カテゴリー								
サンプル				1	5	6	9	3	2	7	4	8
スコア	番号	学年	性別	にんじん	いんげん	キャベツ	レタス	キュウリ	トマト	なす	大根	タマネギ
1.22	8	3	女		○	○	○		○	○		
1.07	9	3	女	○	○		○	○	○		○	
0.73	5	2	女			○	○	○	○	○		
0.56	10	3	女		○				○	○		
0.38	11	3	男	○		○	○	○	○		○	○
-0.14	2	1	女		○	○		○	○	○	○	○
-0.27	7	2	男				○		○			○
-0.37	1	1	女				○		○	○	○	
-1.04	3	1	男					○		○	○	
-1.63	6	2	女						○	○	○	○
-2.40	4	1	男							○	○	○
カテゴリースコア				1.44	1.30	1.06	0.89	0.41	0.33	-0.79	-1.18	-1.63

## 1.3 課題 3

### 課題

表 2・3 より, サンプル, カテゴリーそれぞれについての散布図を作成し, それぞれの散布図からどのようなことが読み取れるかまとめなさい.

図1・2に, サンプル, カテゴリーそれぞれの散布図を示す. なお, 散布図の作成には,R言語を使用した.

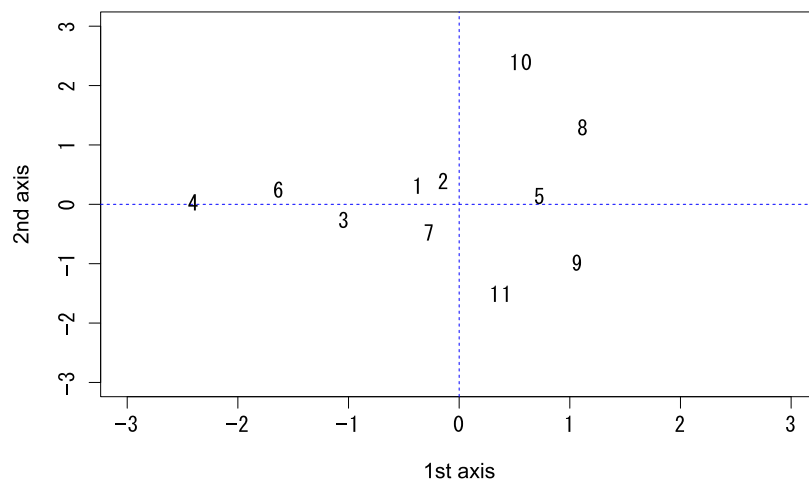


図 1: サンプルスコア散布図

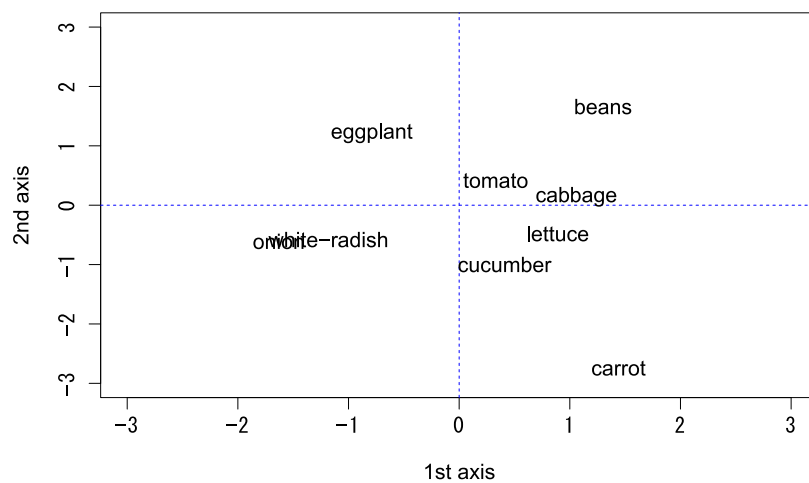


図 2: カテゴリースコア散布図

また, 成分 2 の意味を考察するために, 成分 2 によって表 1 をソートする. ソートした結果を, 表 5 に示す.

表 5: 成分 2 によるソーティング後の結果

			カテゴリー								
サンプル			5	7	2	6	9	4	8	3	1
番号	学年	性別	いんげん	なす	トマト	キャベツ	レタス	大根	タマネギ	キュウリ	にんじん
10	3	女	○	○	○						
8	3	女	○	○	○	○	○				
2	1	女	○	○	○	○		○	○	○	
1	1	女		○	○		○	○			
6	2	女		○	○			○	○		
5	2	女		○	○	○	○			○	
4	1	男		○				○	○		
3	1	男		○				○		○	
7	2	男			○		○		○		
9	3	女	○		○		○	○		○	○
11	3	男			○	○	○	○	○	○	○

始めに, これら散布図の軸の意味を, 図 1 のサンプルスコア散布図を元に考える. まず, 成分 1 のサンプルスコアによってソートされている表 4 からも分かるように, 第 1 軸の値が大きいほど学年が高く, 値が小さいほど学年が低い傾向が見られる. このことから, 第 1 軸は『学年の高さ』を表す軸であるといえる. 次に, 成分 2 のサンプルスコアによってソートしてみると, 第 2 軸の値が大きいほど女が, 値が小さいほど男が集中している. このことから, 第 2 軸は『性別:男女』を表す軸であるといえる.

次に, 先で定めた軸の意味を元に, 図 1 についての分析を行う. 図 1 のサンプルスコア散布図を見ると, 全体的に第 1 軸方向のばらつきは大きい, 第 2 軸方向のばらつきはほとんどない. つまり, 学年 (年齢) による好みの差が最も大きい, もしくは, 年齢の変化とともに好みも変化することが分かる. また, 表 4 を見ると, 反応点の分布は, 第 1 軸方向の値が大きいほど丸印が多くなっており, 上側三角形のようになっていた. つまり, 学年が上がるにつれて, 嫌いなものが減っている傾向があると考えられる. しかし, 1-7 は第 2 軸方向の値がほぼ 0 だが, 8-11 は比較的大きい値となっている. このことから, 学年が上が

るにつれて、男女の好みの差がはっきりしてくる傾向にあると考えられる。

さらに、図2についての分析を行う。図2のカテゴリースコア散布図を見ると、トマト・キャベツ・レタス・キュウリの4種の野菜は、原点からの距離が比較的小さい。つまり、これら4種の野菜は、小学生の中でも、ほとんど好き嫌いが分かれなような野菜であるといえる。特に、トマトはほぼ原点に分布していることから、年齢や学年を問わず好まれる傾向にあるといえる。また、キャベツ・レタス・キュウリは、近い位置に分布しており、比較的高学年の児童から好まれている。このことから、これら3種の野菜は、似たような性質によって好まれていると考えられる。たとえば、野菜のみずみずしさ・味の薄さなどが考えられる。

一方、大根・たまねぎ・なす・いんげん・にんじんの5種の野菜は、原点からの距離が比較的大きい。つまり、小学生の中で、好き嫌いが分かやすい野菜であるといえる。特に、いんげん・にんじんは特徴的に距離が大きいことから、この2種の野菜は、非常に好みが見られる野菜であるといえる。今回は、図2を見ると、いんげんは高学年女子に、にんじんは高学年男子によく好まれている。また、タマネギ・大根は、非常に近い位置に分布しており、比較的低学年の児童から好まれている。このことから、タマネギと大根は似た性質によって好まれていると考えられる。たとえば、みずみずしさ・味のしみこみやずさなどが考えられる。

他にも、第2軸方向の分布を見ると、比較的味の薄い野菜は値がほぼ0であり、味の濃い野菜は大きい値をとっている。つまり、いんげん・にんじんについて好みが大きく分かっている要因は、その味にあると考えられる。また例外的に、なすも大きな値をとっており、好みが見られる。これについては、味ではなく食感が要因になっていると考えられる。さらに、第1軸方向の分布では、学年が高い方向にいんげん・にんじんといった味の濃い野菜が集まっている。つまり、低学年の児童は、野菜の味をあまり好まないのではないかと考えられる。

## 参考文献

- [1] ライブラリ新数学大系: 多変量解析法入門・共著: 永田靖/棟近雅彦・出版: サイエンス社