

# 主観的評価の構造分析

聴覚部・電子情報学科・情報工学専攻 小池 将貴

**要 旨：**論文・感想文のような複雑で多面的な作品の評価は、主観的にならざるを得ない。評価主体の主観によっては、評価結果が分かれることも生じる。それならば、主観による評価の多様化を逆手にとって、評価対象のみならず評価主体をも複数化することによって、主観的評価の構造を見ようと試みた。

主観的評価とは、評価主体が主観的に「良いと思う」順に評価対象を順序づけることと定め、まず、主観を十分汲み上げることのできる評価方法を提示した (Bradley-Terry モデル)。この評価法で4種類の評価主体に評価対象 (学生の作文10篇) を評価させたデータ  $\{\pi_i^{(k)}; k=1, \dots, 4; i=1, \dots, 10\}$  を主観的評価の典型とみなし、そのデータを素材にして、主観的評価を合理的に説明することのできる方法を提示した (主成分分析法)。

その結果、2つの主成分 (“努力志向” と “成人志向”) が見つかった。そして、作文  $i$  の主成分の値を  $(F_1^{(i)}, F_2^{(i)})$  とし、評価主体  $k$  の主成分に対する価値的加重付けを  $(\alpha_1^{(k)}, \alpha_2^{(k)})$  とし、この2つのベクトルの内積

$$\alpha_1^{(k)} F_1^{(i)} + \alpha_2^{(k)} F_2^{(i)}$$

の値により、作文を順序づけたところ、主観的評価 ( $\pi_i^{(k)}$  による順序づけ) 結果をほぼ再現することができた。  
キーワード：評価, 一対比較, 主成分分析

## 1. 問題提起

聴覚障害者の高等教育の一環として、学生に自己の考えを文章にまとめる演習を課している。出来上がった学生の論文・感想文は内容・体裁が多彩で、誰が評価しても結果は同じというわけにはいかない。この一見曖昧で恣意的に感じられる主観的評価を合理的に説明できないものだろうか。

## 2. 解決方針

ここに10篇の作文がある。これは、「米国留学の旭丘高校生射殺事件<sup>1)</sup>」をどのように考えるか、ワープロで1～2頁に纏めよ。」という演習課題を与えられて、筑波技術短期大学・聴覚部 (以降、短大と略称) の電子情報学科・情報工学専攻の3年生が、作文したものである。作文は1993年11月に作成され、その作成者氏名は伏せて、1～10の乱数によって番号を付けた。

これを素材にして、主観的評価の構造分析を試みる。なお、主観的評価とは、評価主体が主観的に「良いと思う」順に評価対象 (10篇の作文) を順序づけることとする。

主観的評価を合理的に説明する前に、まず主観そのものを十分に発揮させることのできる評価方法を提示する。その評価方法で得られた主観的評価結果を、合理的

に説明することによって問題が解決されたと考える。

## 3. 方法

### 3.1 評価対象の順序づけ — Bradley-Terry モデル

主観を十分に発揮させて順序付けるには、多くの評価対象を同時に評価させるよりは、そのうちの2つづつを比べてどちらが「良いと思う」かを尋ねる方が適切であろう。そこで、一対比較を利用する Bradley-Terry モデル<sup>2)</sup>を適用することにした。

この方法は、10篇の作文のうちの2つづつを一対比較すれば、10篇全体を一系列に順序付けてくれる。

理論的前提として、各作文  $i$  は、潜在的に「良さ」 ( $\pi_1, \pi_2, \dots, \pi_i, \dots, \pi_{10}$ ) を持ち、作文  $i$  が他の作文  $j$  と一対比較されたときに「より良い」との評価を得る確率が

$$P_{ij} = \frac{\pi_i}{\pi_i + \pi_j}$$

で与えられると仮定する。これを Bradley-Terry モデル (BT モデル) という。

BT モデルへのインプットデータは、“作文  $i$  と作文  $j$  との  $n_{ij}$  回の一対比較のうち、 $i$  は  $j$  より良いと評価されたのは  $x_{ij}$  回である”。という形式で与えるのだが、そのためには、評価主体に一対比較 (2つの作文の任意の組

み合わせ  $(i, j)$  に対して、どちらが良いかを判定すること)を十分な回数実施してもらう。

すると、独立に  $n_{ij}$  回試行して、そのうち  $x_{ij}$  回良い評価を得る確率は2項分布によって表現されるので、その積として尤度関数が、

$$\begin{aligned} \Pr(X_{ij}=x_{ij}; i \neq j; i, j=1, 2, \dots, 10) \\ = \prod_{i=1}^{10} \prod_{j>i}^{10} \left\{ \frac{n_{ij}!}{x_{ij}! x_{ji}!} \times \frac{\pi_i^{x_{ij}} \pi_j^{x_{ji}}}{(\pi_i + \pi_j)^{n_{ij}}} \right\} \\ = \prod_{i=1}^{10} \prod_{j>i}^{10} \left\{ \frac{n_{ij}!}{x_{ij}! x_{ji}!} \times \frac{1}{(\pi_i + \pi_j)^{n_{ij}}} \right\} \times \prod_{i=1}^{10} \pi_i^{t_i} \quad ; t_i = \sum_{j=1; j \neq i}^{10} x_{ij} \end{aligned}$$

のように得られる。

右辺の2番目の式により、「良さ」 $\pi_i$ は、一対比較  $(i, j)$  の比較件数  $n_{ij}$  と作文  $i$  の「良い」評価獲得総件数  $t_i$  とだけで決まることがわかる。この尤度関数を用いて、最尤法で未知母数

$$(\pi_1, \pi_2, \dots, \pi_i, \dots, \pi_{10})$$

を推定する。実際の計算では、対数尤度にラグランジュの未定乗数法を適用し、逐次近似式を導いて「良さ」 $(\pi_1, \pi_2, \dots, \pi_i, \dots, \pi_{10})$  の推定値をアウトプットした。作文の順序づけは、これらの「良さ」の推定値の大きさに順に依ればよい。

計算実行例を、表1と図1とによって説明する。

表1は、評価主体(④企業エンジニアに属する16名)の各々に数回づつ一対比較を試みていただいた結果を纏めたものである。例えば、作文8番は6番に対し、「より良い」との評価を3回( $x_{8,6}=3$ )獲得し、6番は8番「より良い」との評価を1回獲得している( $x_{6,8}=1$ )。従って、 $n_{6,8}=n_{8,6}=4$ 。

この表1のデータをインプットデータとして、BTモデルで計算した結果が、図1の④企業エンジニアに示すようなアウトプットである。それによると、例えば、作文8番の「良さ」は、 $\pi_8=49.33$ であり、6番の「良さ」は、 $\pi_6=17.06$ であることがわかる。

### 3.2 潜在構造の明示化 — 主成分分析法

主観的評価を合理的に説明するためには、まず、主観そのものを十分発揮させた評価結果を得ることが前提なので、BTモデルを採用した。そのインプットである一対比較を1人でこなすのは負担が重い。よって、同質の人を纏めて評価主体とし、そこでなされた一対比較は、あたかも一人でなされたとみなすことにした。

さて、評価主体の主観によって評価が分かれるならば、

評価主体をも複数グループ用意することによって、主観的評価の構造を見ようとした。そこで、次の評価主体を選定した。

- ①短大教官：短大教官の約半数の25名。
- ②短大学生：短大・情報工学専攻の1, 2, 3年生全員30名。ただし、作文を作成した3年生10名は、自分以外の作文を評価させた。
- ③私大学生：東京西郊にある私立の四年制総合大学のESSクラブの学生有志20名。短大卒業生が就職した場合に、そこで同僚の役割を果たす者として選定した。
- ④企業エンジニア：一部上場電気機器メーカーの湘南地区工場の若手エンジニア16名。四年制大学卒あるいは修士課程卒で入社し、五年のキャリアをもつ。短大卒業生の就職先で直属の指導者の役割を果たす者として選定した。

次に、こうして選定された評価グループ毎の10篇の作文の順序づけデータを作成した。実際には、アンケート調査により、各グループのメンバーに一対比較を数回繰り返してもらい、収集された一対比較データをBTモデルで計算し、各作文の「良さ」を求めた。図1の①~④参照。以降では、これを主観的評価の典型とみなし、その解明を問題解決とみなす。

図1に示した評価グループ①~④の作文の「良さ」の2元表データ

$$\{\pi_i^{(k)}; k=1, \dots, 4; i=1, \dots, 10\}$$

をインプットデータとして、主成分分析法<sup>3)</sup>を適用した。

計算結果の表2の〈固有値〉を見ると、固有値が1.0以上の主成分は、2つあり、それだけで累積寄与率が88.4%もある。そこで、この2つの濃縮主成分で以て、直交座標系を構成し、評価グループと評価対象をそれぞれ位置づけてみることにした。こうして、評価主体(①~④)を図示したものが図2の〈因子負荷量散布図〉である。評価対象(10篇の作文)を図示したものが、図3の〈主成分得点散布図〉である。

さらに、具体的な数値を表3の〈因子負荷量〉と、表4の〈主成分得点〉とに示す。

## 4. 主観的評価の構造解明

上述の「3. 方法」にて得られた図1~3や表1~4は、アンケートデータを数学的手法で要約したものであり、事実データの集約として客観性をもつ。しかし、これから行う主成分の抽出は、客観データに基づくとはいえず、

ひとつの解釈である。まず、図2、図3におけるx軸、y軸の意味を解釈してみる。

#### 4.1 x軸となる第1主成分の意味

図2の評価主体をx軸（主成分1）に正射影してみると、x座標値はすべて正の値である。つまり、主成分1は、すべての評価主体が共有する価値観を表わしている。

それが具体的に何を指すのかを明らかにするために、図3の各点（作文番号）を、x軸（主成分1）に正射影し、x座標値がプラスの作文（8番、9番、3番）と、x座標値がマイナスの作文（2番、6番、7番）とを対比させて見直してみた。すると、x座標値がプラスの作文は、長さが2頁にわたって細かい字で豊富に書き込んである。x座標値がマイナスの作文は、長さが半頁で、なるべく行間を空け、ことさらに大きな活字を使って紙面を早く埋めてしまおうとしているのが見てとれる。これでわかるのは、主成分1は、「とにかく課題が与えられた以上はしっかりと取り組んで努力する」という姿勢を評価しようとしているのである。そこで、主成分1を“努力志向”と命名する。

#### 4.2 y軸となる第2主成分の意味

図2の評価グループをy軸（主成分2）に正射影してみると、y座標値の符号がマイナス側に短大教官と短大学生の評価主体がある。プラス側に企業エンジニアと私大学生がある。つまり、主成分2は、評価主体が両極に別れてしまうような価値観を表している。

それが、何を表すのか明らかにするために、図3の各点（作文番号）のy座標値がマイナスの作文（7番、9番、8番）とプラスの作文（4番、3番、10番、5番）とを丹念に読み比べてみて、作文の作成者の性格に特徴があることに気づいた。即ち、y座標値がマイナスの方は、「生真面目に良く勉強し、従順である」。y座標値がプラスの作文は、「要領よく勉強し、世間的に大人びている」。そこで、主成分2は、“成人志向”と命名した。

#### 4.3 主観的評価の再構成

構造分析の素材とすべく、評価主体毎に、10篇の作文の「良さ」 $\{\pi_i^{(k)}; k=1, \dots, 4; i=1, \dots, 10\}$ をBTモデルによって計算した（図1）。その「良さ」のデータを、わかりやすく「順位」に書き換えて、表5の「主観的評価」の箇所に示す。これは、評価主体の主観を十分に汲み上げた主観的評価の典型とみなす。つまり、表5の「主観的評価」の順位データを構造分析探究のための素材として、その順位を別途に再現する仮説を提示することが、評価構造を解明することであるとす。それを以下に示す。

評価主体k毎に順序づけを再現する。まず、表3に載っている評価主体kの主成分1（努力志向）、主成分2（成人志向）の因子負荷量の値をそれぞれ $(\alpha_1^{(k)}, \alpha_2^{(k)})$ 代入し、表4に載っている評価対象・作文iの主成分1（努力志向）、主成分2（成人志向）の主成分得点の値をそれぞれ $(F_1^{(i)}, F_2^{(i)})$ 代入する。そして、この2つのベクトルの内積

$$\alpha_1^{(k)} F_1^{(i)} + \alpha_2^{(k)} F_2^{(i)}$$

を計算する。こうして得た「再構成評価」値を、あたかも作文iの「良さ」とみなして、作文を順序づける。すると、表5の「主観的評価」の順位データをほぼ再現することができた。

例えば、短大学生の評価した作文8の再構成された「良さ」は、

$$0.891 \times 1.850 + (-0.275) \times (-1.084) = 1.946$$

のように計算される。そして、この例題のような計算をすべての作文について行い、その値の大きさの順に並べ替えたものが表5の「再構成評価」である。

表5の「主観的評価」と「再構成評価」とを比較検討すると、すべての評価主体において、評価順位が完璧に近く再現されていることがわかる。従って、主観的評価

表1 BTモデルへのインプット（④企業エンジニアを例として）

i	j	1	2	3	4	5	6	7	8	9	10
1		0.	1.	0.	0.	0.	0.	2.	0.	0.	0.
2		0.	0.	0.	0.	0.	1.	1.	2.	1.	0.
3		1.	1.	0.	0.	0.	2.	2.	1.	1.	1.
4		1.	1.	1.	0.	0.	1.	1.	0.	1.	0.
5		1.	2.	0.	1.	0.	0.	1.	0.	0.	1.
6		2.	0.	0.	0.	1.	0.	2.	1.	0.	0.
7		1.	1.	0.	0.	1.	1.	0.	0.	0.	0.
8		1.	1.	1.	0.	1.	3.	2.	0.	0.	1.
9		1.	0.	0.	0.	0.	1.	0.	1.	0.	0.
10		1.	0.	1.	1.	0.	1.	1.	1.	1.	0.

表2 主成分分析のアウトプット〈固有値〉

No.	固有値	寄与率	累積寄与率
1	2.194	0.549	0.549
2	1.343	0.336	0.884
3	0.273	0.068	0.953
4	0.190	0.047	1.000

表3 主成分分析のアウトプット〈因子負荷量〉

変数名	主成分1	主成分2
短大教官	0.719	-0.628
短大学生	0.891	-0.275
私大学生	0.833	0.391
企業エンジニア	0.436	0.849

表4 主成分分析のアウトプット〈主成分得点〉

サンプル名	主成分1	主成分2
作文1	-0.457	-0.205
作文2	-1.123	-0.514
作文3	0.913	1.309
作文4	0.069	1.500
作文5	-0.499	0.552
作文6	-1.093	-0.406
作文7	-0.850	-1.008
作文8	1.850	-1.084
作文9	1.040	-1.088
作文10	0.150	0.943

表5-1 評価の再構成 (短大教官)

(表内の数字は作文番号)

	1位	2位	3位	4位	5位	6位	7位	8位	9位	10位
主観的評価	9	8	7	10	3	2	6	5	1	4
再構成評価	8	9	7	3	1	10	2	6	5	4

表5-2 評価の再構成 (短大学生)

(表内の数字は作文番号)

	1位	2位	3位	4位	5位	6位	7位	8位	9位	10位
主観的評価	8	3	9	10	1	7	4	6	5	2
再構成評価	8	9	3	10	1	4	7	5	6	2

表5-3 評価の再構成 (私大学生)

(表内の数字は作文番号)

	1位	2位	3位	4位	5位	6位	7位	8位	9位	10位
主観的評価	3	9	8	4	1	5	10	6	2	7
再構成評価	3	8	4	10	9	5	1	6	7	2

表5-4 評価の再構成 (企業エンジニア)

(表内の数字は作文番号)

	1位	2位	3位	4位	5位	6位	7位	8位	9位	10位
主観的評価	3	4	10	5	8	9	2	6	7	1
再構成評価	3	4	10	5	8	1	9	6	2	7

NO.	VALUE	.....I.....I.....I.....I.....I.....I.....I
1	24.29	+++++
2	38.06	+++++
3	38.52	+++++
4	21.42	+++++
5	25.98	+++++
6	26.28	+++++
7	51.20	+++++
8	111.29	+++++
9	123.60	+++++
10	39.36	+++++
		.....I.....I.....I.....I.....I.....I.....I

①短大教官

NO.	VALUE	.....I.....I.....I.....I.....I.....I.....I
1	36.55	+++++
2	8.62	+++
3	74.48	+++++
4	27.82	+++++
5	12.30	++++
6	16.74	++++
7	35.19	+++++
8	169.40	+++++
9	67.94	+++++
10	50.96	+++++
		.....I.....I.....I.....I.....I.....I.....I

②短大学生

NO.	VALUE	.....I.....I.....I.....I.....I.....I.....I
1	58.72	+++++
2	1.82	+
3	99.28	+++++
4	72.46	+++++
5	51.32	+++++
6	10.03	+++++
7	1.39	
8	76.36	+++++
9	80.89	+++++
10	47.73	+++++
		.....I.....I.....I.....I.....I.....I.....I

③私大学生

NO.	VALUE	.....I.....I.....I.....I.....I.....I.....I
1	7.80	++++
2	21.88	+++++
3	104.60	+++++
4	104.43	+++++
5	54.65	+++++
6	17.06	+++++
7	8.57	++++
8	49.33	+++++
9	28.76	+++++
10	102.93	+++++
		.....I.....I.....I.....I.....I.....I.....I

④企業エンジニア

図1 BTモデルのアウトプット (作文*i*の良さ $\pi_i$   $1 \leq i \leq 10$ )

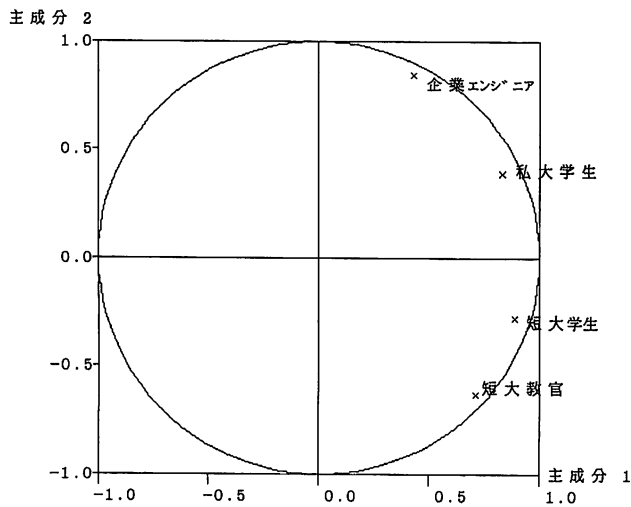


図2 主成分分析のアウトプット〈因子負荷量散布図〉

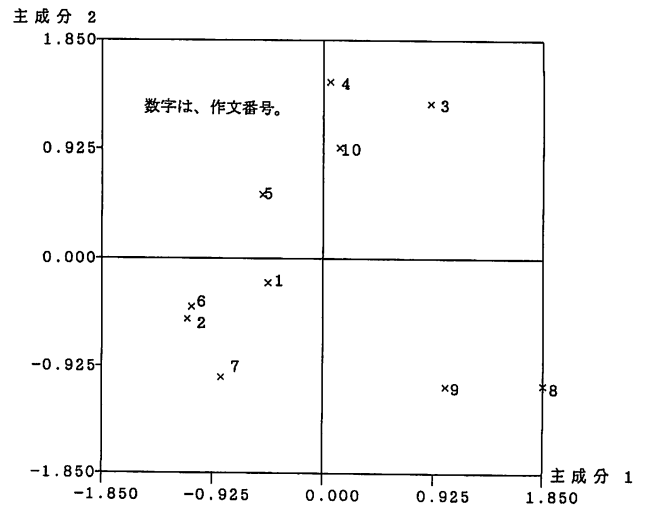


図3 主成分分析のアウトプット〈主成分得点散布図〉

の潜在構造が明示化されたと考える。

## 5. 結論

論文・感想文のような複雑で多面的な作品の評価が、評価対象の優劣を差別化するのは当然として、評価主体の主観によっても分かれることを、“努力志向”と“成人志向”と言う潜在特性を介在させることにより、説明することができた。

## 6. 参考文献

- 1) 朝日新聞全文記事情報 (1992年版), 米国での旭丘高校生射殺事件, 紀伊国屋書店, 東京.
- 2) 東京大学教養学部統計学教室編: 自然科学の統計学 (初判), 1992, 166-170, 東京大学出版会, 東京.
- 3) 柳井晴夫/高木広文編著: 多変量解析ハンドブック (初判), 1993, 70-83, 現代数学社, 京都.