1996年度日本オペレーションズ・リサーチ学会
秋季研究発表会

不完全情報の主成分分析による学生の成績評価

01300450　日本大学　＊高橋磐郎　TAKAHASHI Iwaro
01011500　日本大学　大澤慶吉　OSAWA Keikichi
01404360　日本大学　西澤一友　NISHIZAWA Kazutomo

§1．はじめに

不完全情報下での主成分分析についてはすでに[2],[3]で述べてきた。

一般的には学生の成績評価を行う場合には、個々の学生の履修科目の単純平均点によってランク付けを行う場合が多い。しかし、このような単純平均点でランク付けを行うことは不適当に思われる。

そこで、単なる形式的な単純平均点よりは主成分分析を用いたスコア平均点で評価するほうが合理的と考えられるので実際のデータ（A大学B学科）を用いて解析を行った。

§2．不完全情報下での主成分分析のデータモデル

ここでは学生nが科目jを未履修である場合、そのデータを次のように仮定する。

(1)　$\alpha+\beta p_n+\gamma q_j$

p_n：n番目の学生の単純平均
q_j：j科目の単純平均
α, β, γ：未知パラメータ

次にδ記号を用いることで学生の全成績(n=1～N, j=1～p)を

(2)　$\alpha_{nj}=\delta_{nj}a_{nj}+(1-\delta_{nj})(\alpha+\beta p_n+\gamma q_j)$

$$\delta_{nj}\begin{cases}=1 & \text{（履修科目）}\\ =0 & \text{（未履修科目）}\end{cases}$$

と書ける。

(3)　$z_n=\sum_n \alpha_{nj}u_j$

となるが、z_nの分散$v_z=\Sigma(z_n-\bar{z})^2/N$，$\bar{z}=\sum_n z_n/N$を$u_1^2+\cdots+u_p^2=1$のもとで最大になるように$u_1,\cdots u_p$, α, β, γを決めようという考えが、ここで用いる原則である。

§3．解析および計算方式

上記の原則に基づく解の候補は、ラグランジュ関数

$L=v_z-\lambda(u_1^2+\cdots+u_p^2-1)$の停留点の中に求められる。つまり、

(4)　$\partial L/\partial u_i=0$　(i=1～p)

(5)　$\partial L/\partial\alpha=0$　、$\partial L/\partial\beta=0$
$\partial L/\partial\gamma=0$

の解となるが、このうち(4),(5)からそれぞれ次の式(6),(7)が得られる。

(6)　$Ru=\lambda u$,　$u^T=(u_1,\cdots u_p)$

$$(7)\begin{cases}(EE)\alpha+(PE)\beta+(QE)\gamma+(AE)=0\\(PE)\alpha+(PP)\beta+(QP)\gamma+(AP)=0\\(QE)\alpha+(QP)\beta+(QQ)\gamma+(AQ)=0\end{cases}$$

式(6)のRの中にはα、β、γが含まれているし、又、式(7)の(EE),(PE)・・・・・等にはu_1,・・・・・u_pが含まれているので、式(6),(7)を連立させて解くために、次のような逐次近似計算を用いる。

先ず、$u_1=\cdots\cdots=u_p=1/\sqrt{p}$を定め(EE),(PE),・・・・・等を計算して式(7)を解きα、β、γを求める。　これを用いて式(6)の主固有ベクトル（最大固有値に対する）$u_1,\cdots,u_p$を求める。これを新たな初期値として以上の手順を繰り返す。

§4．解析結果

ここでは、A大学B学科を卒業した92名の学生が選択した65科目の選択科目での成績評価を主成分分析で解析した。

全選択科目を履修した学生は2名であった。学生の単純平均による成績(p_n)と科目平均(q_j)の一部を表1に示す。

表1．学生の成績と科目平均

学生	成績	科目	平均
1	64.342	1	75.200
2	69.575	2	72.838
3	74.636	3	84.932
4	64.666	4	77.933
·	·	·	·
·	·	·	·
·	·	·	·
90	89.228	63	83.315
91	68.057	64	84.717
92	62.638	65	85.454

又、反復回数６回で収束し、固有値λ、パラメータβ、γ（ここではα＝０）は

$$\lambda = 24.4523$$
$$\beta = 0.64165$$
$$\gamma = 0.38952$$

と求まった。科目のウエイトの一部を表２に示す。

表２．科目のウエイト

科目	ウエイト(u_j)
1	0.040739
2	0.078096
3	0.057578
4	0.048644
5	0.022592
·	·
·	·
30	0.122660
31	0.227482
32	0.217303
33	0.087083
·	·
·	·
64	0.147421
65	0.296188

次に主成分スコア平均の一部を表３に又、単純平均とスコア平均によるランク付けの一部を表４に示す。

表３．主成分スコア平均

学生	スコア平均	（履修科目）
1	66.95	(35/65)
2	71.24	(33/65)
·	·	·
18	84.58	(36/65)
21	84.01	(30/65)
40	84.45	(29/65)
42	84.58	(31/65)
·	·	·
56	78.40	(30/65)
79	78.59	(30/65)
·	·	·
90	87.90	(35/65)
91	70.17	(35/65)
92	65.23	(36/65)

表４．単純平均とスコア平均によるランク

学生	単純平均(ﾗﾝｸ)	スコア平均(ﾗﾝｸ)
19	90.06 (1)	88.42 (1)
90	89.23 (2)	87.90 (2)
42	85.71 (3)	84.58 (3)
40	85.38 (4)	84.45 (5)
21	85.33 (5)	84.01 (6)
18	85.17 (6)	84.58 (3)
51	84.41 (7)	83.90 (7)
11	84.25 (8)	83.55 (8)
86	84.10 (9)	83.21 (9)
61	83.27 (10)	82.52 (11)
37	83.00 (11)	82.63 (10)
·	·	·
·	·	·
56	78.40 (23)	78.40 (24)
79	78.27 (24)	78.59 (23)
·	·	·
·	·	·
74	64.92 (79)	67.19 (81)
30	64.87 (80)	67.64 (79)
4	64.67 (81)	67.57 (80)
·	·	·
46	60.14 (92)	63.30 (92)

§５．まとめ

一般的な単純平均でランクが高い学生の場合、科目を多く履修していると主成分スコア平均でのランクはより高くなる。しかし、逆に単純平均でランクが低い学生の場合には科目を多く履修していることで主成分スコア平均でのランクは下がることが分かった。このように、主成分分析による成績評価では履修科目数等をも考慮した合理的なランク付けを行うことができる。

参考文献

[1] 奥野忠一他：「多変量解析法」日科技連、1971
[2] 高橋磐郎他：不完全情報下での主成分分析、1995年度日本ＯＲ学会秋季研究発表会アブストラクト集
[3] 高橋磐郎他：不均一構造・不完全情報下での主成分分析、1996年度日本ＯＲ学会春季研究発表会アブストラクト集