

続マトリックス・データ解析法による土木工学科学生の進路分析

米 長 泰* 菊 地 宗太郎**

Reasoning study on the civil engineering student courses after leaving the colleges by matrices data analytical method (No. 2)

Yasushi YONENAGA and Sotaro KIKUCHI

(平成2年8月21日受理)

Matrices data analytical method is known as a tool by which we understand the peculiarity of the phenomenon theoretically. We now attempted to apply this tool to reasoning study on the civil engineering student courses after leaving the colleges. We treated the comparison of each college as No.1 report. Then this report treats each annual comparison within one college.

1. 続編の趣旨

マトリックス・データ解析法は群をなす固体の個々の特長を抽出する手法として産業界に普及しつつある。今回は全国28の土木工学科を有する工業高専の進路を分析し、主成分の意味合いも、各校の散布図も、一応納得のいく解説が得られた。

そこで本文では第二弾とし、我が校18年間の卒業生の進路の特長把握を試みようとするものである。10年前は建設業界は大不況で先輩の教官は就職の世話が大変だったと伺っている。昨今は逆に超売手市場とやらで全く反対の現象となっている。また各時点々々では学生の人気や傾向といったものが存在する。これらを総合して、各年次の卒業生進路にどんな特長が暗示されるのか。丁度分析するのに適切な時期だと考え、続編として取り組んだ。

2. マトリックス・データ解析法とは

前編⁽¹⁾で詳しく説明しているから、要点だけを再説明する。文献⁽²⁾より抜粋した。

(1) インプット・データ

n個の固体とP個の変数によるn×P二元マトリックスである。マトリックスの要素はすべて数値で表わされなければならない。論理上の数値でもよい。

* 秋田工業高等専門学校土木工学科

** 同学生

(2) 相関係数行列

$$\begin{bmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1P} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{P1} & \gamma_{P2} & \dots & \gamma_{PP} \end{bmatrix}$$

特性値間の相関係数を行列に並べたもの。γ_{ij}とは特性値χ_iと特性値χ_jの相関係数である。対角要素が1の値をとる対称行列である。

(3) 固有値と固有ベクトル

下に示す行列方程式を解くことにより求められるものである。固有値λ_kは主成分Z_kの分散に等しい。また固有値λ_k (K=1, 2, ..., P)の合計はPになる。P組の固有ベクトルが存在する。

$$\begin{bmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1P} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{P1} & \gamma_{P2} & \dots & \gamma_{PP} \end{bmatrix} \begin{bmatrix} l_{K1} \\ l_{K2} \\ \vdots \\ l_{KP} \end{bmatrix} = \lambda_K \begin{bmatrix} l_{K1} \\ l_{K2} \\ \vdots \\ l_{KP} \end{bmatrix}$$

↑
固有地

↑
固有ベクトル

条件 (l_{K1}² + l_{K2}² + + l_{KP}² = 1)

λ_kを第K主成分の固有値と呼び、値が大きいほど重要な意味合いをもつ。値が1より小さな主成分は無視する。 λ₁ ≥ λ₂ ≥ λ₃ ≥ ...

(4) 累積寄与率

$$\frac{\lambda_1}{P}, \frac{\lambda_1 + \lambda_2}{P}, \frac{\lambda_1 + \lambda_2 + \lambda_3}{P}, \dots$$

を累積寄与率と呼び、事象全体の何%ぐらいが説明できているかを示す尺度である。通常80%前後に達

すれば十分と考えられる。

(5) 主成分

固有ベクトル l_{ki} と特性値 X_i の積算和を第 K 主成分 Z_k という。第 m 主成分迄求めるとき次式で与えられる。これらを順次第 1 主成分, 第 2 主成分, …… と呼ぶ。総合特性値である。上述のように $\lambda_m > 1$ の範囲にとどめるのがよい。

$$Z_1 = l_{11}X_1 + l_{12}X_2 + \dots + l_{1p}X_p$$

$$Z_2 = l_{21}X_1 + l_{22}X_2 + \dots + l_{2p}X_p$$

$$Z_k = l_{k1}X_1 + l_{k2}X_2 + \dots + l_{kp}X_p$$

$$Z_m = l_{m1}X_1 + l_{m2}X_2 + \dots + l_{mp}X_p$$

(6) 因子負荷量

主成分 Z_k と元の特性値 X_i を基準化した X_i との相関係数のことである。 a_{ik} と表わせば,

$$a_{ik} = \sqrt{\lambda_k} l_{ki}$$

により求められる。もちろん $-1 \leq a_{ik} \leq +1$ であるが、絶対値が 0.6 より大きな a_{ik} だけに着目し、その符号を加味しながら第 K 主成分の意味合いを解説することになる。

(7) 主成分スコア

特性値 X_i が n 個の固体から採取されたとき、最初のインプット・データは $n \times P$ 二元マトリックスとなる。第 K 主成分のスコアは固有ベクトル l_{ki} と基準化された特性値 X_i の積算和として求められる。即ち

$$Z'_k = l_{k1}X_1' + l_{k2}X_2' + \dots + l_{kp}X_p'$$

↑
固有ベクトル

実際には Z'_k が固体に応じて n 個存在している。

(8) 散布図

2つの主成分スコアを直角座標上にプロットしたもの。これまでの数値解析結果がビジュアルになりわかり易い。一例を図 1 に示す。これは平成元年度プロ野球セ・リーグ打撃 28 傑の散布図である。横軸は第 1 主成分「長距離砲」、縦軸は第 2 主成分「出塁盗塁」を表わしている。クロマティは両要素を備え MVP に相応しいといえる。また落合は三冠は逸したものの長距離打者としてはリーグ随一であることがわかる。本図を駆使すれば並の評論家以上の口がきけそうである。文献 (3) を参照されたい。

各固体 (選手一人々々) にデータ番号を記すと見やすくなる。データは中心部になるほど平均的になり、左右上下に離散するほど何か特長を有することになる。実際にはさきの因子負荷量と散布図とを見

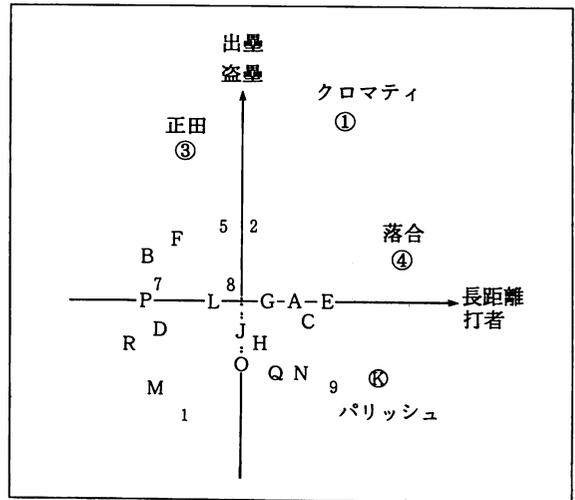


図 1. 散布図 (プロ野球打撃成績) の一例

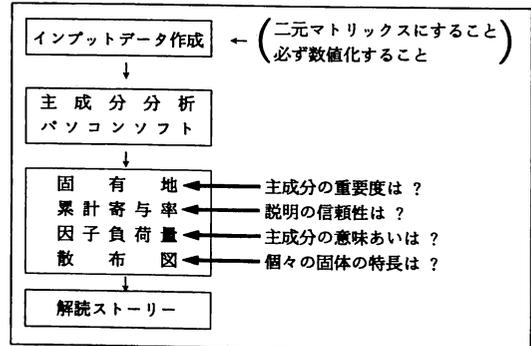


図 2. マトリックス・データ解析法の活用要領

比べながら、試行錯誤のうえ結論を得るのである。なお第 3 主成分迄取り上げる場合、第 1 VS 第 2 と第 1 VS 第 3 の散布図でよいと思いがちであるが、念のため第 2 VS 第 3 をプロットすると意外な新事実を発見することがある。これらは経験によって得られるノウハウである。

以上の道順をとりまとめると図 2 のようになる。

3. 学生進路分析

本校土木工学科は現 5 学年が第 18 期生である。目下殆どの学生が進路内定しているので、各期を 1 個体として、18 期全体を解析することにした。主成分の解釈は百人百様と言われるので、筆者と見解の相違あらば後日ご指摘いただきたい。

3・1 進路先の区分

前回の区分を参考としながらも、本文は同じ学内だということ、また筆者は就職担任として 300 人を

超える各社のイグゼクティブと面談させていただくなど若干の経験を積んだ。前回と全く同じ区分にしようかとも思ったが、結局区分(特性値)を大幅に変更することにした。要点は下記のとおり。

- (1) 卒業生総数は100分比でなく実数とする。入学から5年後の在籍数は重要である。
- (2) 資本金400億円以上を超大手とする。鹿島・清水建設は建設大手に入れる。また土木が全く傍系の大企業は加工・電工に入れる。
- (3) 建設業界を資本金100億円で大小に分ける。
- (4) 国鉄は公社、JRは超大手とする。
- (5) 基幹企業を網構造・コンクリート・電工に分ける。横河工事は網構造、化工機工事は電工・化工、三井造船は超大手とする。
- (6) 自営業は特殊企業と合せ別分野とする。

(7) 建設省・運輸省・文部省Ⅱ・Ⅲ種を合せて建設省とする。町役場は市役所とする。

3・2 インプット・データ

結局進路の区分は14となった。これを卒業生名簿にあてはめインプット・データを作成した。表1の通りである。入力サイズは前述の記号でいうと

$$\left. \begin{array}{l} \text{個体の数 } n = 18 \\ \text{特性値の数 } p = 14 \\ \text{主成分の数 } m = 5 \end{array} \right\} (18 \times 14 \text{ 二元マトリックス})$$

3・3 各主成分の出力データと考察

(1) 各主成分の重要度

第1～第5主成分の固有値・寄与率・累積寄与率を表2に示す。固有値は $\lambda_1=3.437$ と $\lambda_2=3.203$ が大きいから第3主成分以下は重要度が低く、参考程度といったところ。ただし累積寄与率80%に達する

表1. 各年次進路内訳

(18期は8月推定値) 単位:人

(期)年次	進路	1	2	3	4	5	6	7	8	9	10	11	12	13	14
		卒業生総数	建設省	公社	県庁市役所	建設業大手	〃 中小	コンサルタント	鋼構造	化工電子	超大手	進学	コンクリ舗装	ソフト	別分野
1	1	35	4	3	0	11	8	0	1	4	0	2	1	0	1
2	2	34	3	3	3	4	11	0	2	4	1	0	2	0	1
3	3	31	4	4	0	1	11	1	2	0	0	1	6	0	1
4	4	34	0	3	2	0	13	3	2	1	0	4	3	0	3
5	5	30	2	6	7	0	6	0	2	4	0	1	1	0	1
6	6	36	1	4	6	1	12	3	1	2	0	1	3	0	2
7	7	39	3	2	5	1	17	2	1	2	0	1	3	0	2
8	8	35	1	0	5	4	13	0	1	2	0	3	4	0	2
9	9	36	2	4	3	2	8	4	1	2	1	4	3	0	2
10	0	31	1	2	1	1	11	0	4	1	2	3	4	0	1
11	A	34	1	1	2	0	14	3	0	2	1	3	3	0	4
12	B	34	0	1	0	0	17	2	1	2	1	2	3	2	3
13	C	37	1	2	0	0	15	3	2	4	1	3	3	1	2
14	D	32	0	1	3	0	10	3	3	3	1	4	2	1	1
15	E	33	2	3	2	0	12	4	1	3	2	3	0	0	1
16	F	29	1	1	4	0	9	2	1	0	2	3	3	2	1
17	G	34	1	0	4	1	11	3	1	2	1	5	1	3	1
18	H	42	3	1	3	1	8	6	8	2	6	3	0	1	0

統マトリックス・データ解析法による土木工学科学生の進路分析

ためには第5主成分迄論及する必要がありそうだ。

(2) 各主成分の意味合い

各主成分と特性値の相関係数即ち因子負荷量を表3に示す。この中で絶対値が大きなもの(●印)に着目しながら主成分の意味合いを解説する。

〈第1主成分〉 コンサルタント、鋼構造、超大手、進学の4分野に人員が集中している。基幹産業の計画・施行业務希望である。好況時の現象であろう。

〈第2主成分〉 建設省が少なく建設業中小に大挙就職するのがプラスの主成分で別分野も多い。

〈第3主成分〉 卒業生総数と建設業中小への就職が比例する主成分である。ところが公務員との相関がない。大勢で卒業し大勢で建設業中小へといった集団ムードが感じられる。

〈第4主成分〉 地方公務員とコンクリート分野のどちらかへの偏りを示す。あまり重要度はない。

〈第5主成分〉 建設業大手とソフト産業に関する局部的な指摘である。

3・4 散布図と各年次の特長

マトリックス・データ解析法は表3因子負荷量と見比べながら図3と図4の散布図の特長を抽出するもので、本文のハイライトである。しかし、最終的には表1に示すソースデータがどうなっているのかヘフィードバックしなければならない。

(1) 第1主成分VS第2主成分(図3)

図3で目を引くのは④が極端に右寄りということ

表2. 各主成分の重要度

No.	固有値	寄与率	累積寄与率
1	3.437	0.246	0.246
2	3.203	0.229	0.474
3	1.666	0.119	0.593
4	1.479	0.106	0.699
5	1.350	0.096	0.795

である。即ちもっとも新しい世代の18期生は超大手、鋼構造、コンサルタント、進学に多数が押寄せ、この現象が開校以来の最大特長となっている。これに対し初期の1~3期生が真反対の左側に片寄っている。第1主成分では新旧が全く対称的な進路を採ったことを物語っている。

これに対し第2主成分では18期生と1期生が最下方にある。1期生は建設大手へ11人も就職しているからである。当時は鹿島建設はじめ大手建設会社への就職が多く、それが第2順位の特長として指摘されている。18期生も下方にあるのは建設業中小が少なく別分野がゼロだからである。

第2主成分の上端は④即ち12期生である。半数が建設業中小に就職し公務員がゼロであるが、これも第2主成分の特長といえる。不況だったら公務員の

表3. 因子負荷量

●印 要注意

特性値	主成分1	主成分2	主成分3	主成分4	主成分5
1. 卒業生総数	0.394	-0.232	0.791●	-0.042	-0.192
2. 建設省	-0.307	-0.731●	0.221	-0.354	0.079
3. 公団公社	-0.577	-0.426	-0.159	0.030	-0.480
4. 県庁市役所	-0.050	-0.148	-0.301	0.482	-0.519●
5. 建設業大手	-0.388	-0.485	0.290	0.037	0.612●
6. " 中小	-0.058	0.700●	0.573●	-0.094	-0.044
7. コンサルタント	0.782●	0.045	0.231	0.041	-0.370
8. 鋼構造	0.597●	0.515	-0.037	-0.401	-0.138
9. 化工電工	-0.157	0.459	0.384	0.641●	0.116
10. 超大手	0.826●	-0.365	0.008	-0.263	-0.050
11. 進学	0.643●	0.347	-0.179	0.195	0.256
12. コンクリ舗装	-0.451	0.546	-0.101	-0.613●	-0.046
13. ソフト	0.594	0.285	-0.253	0.192	0.398
14. 別分野	-0.234	0.760●	0.400	0.134	-0.180

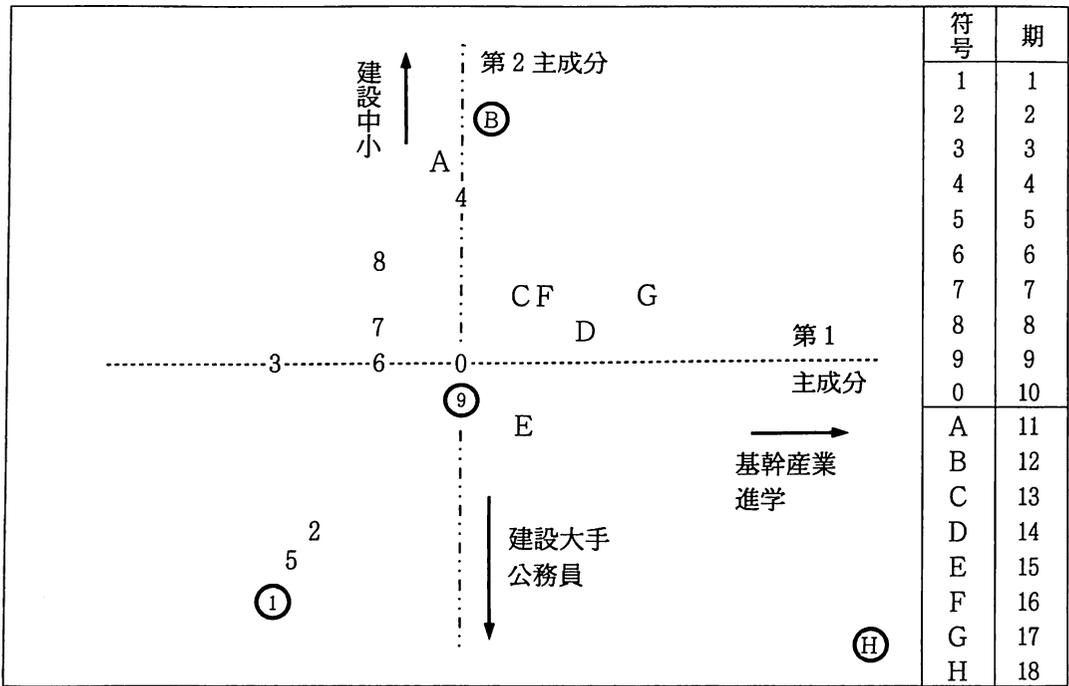


図3. 第1主成分 VS 第2主成分

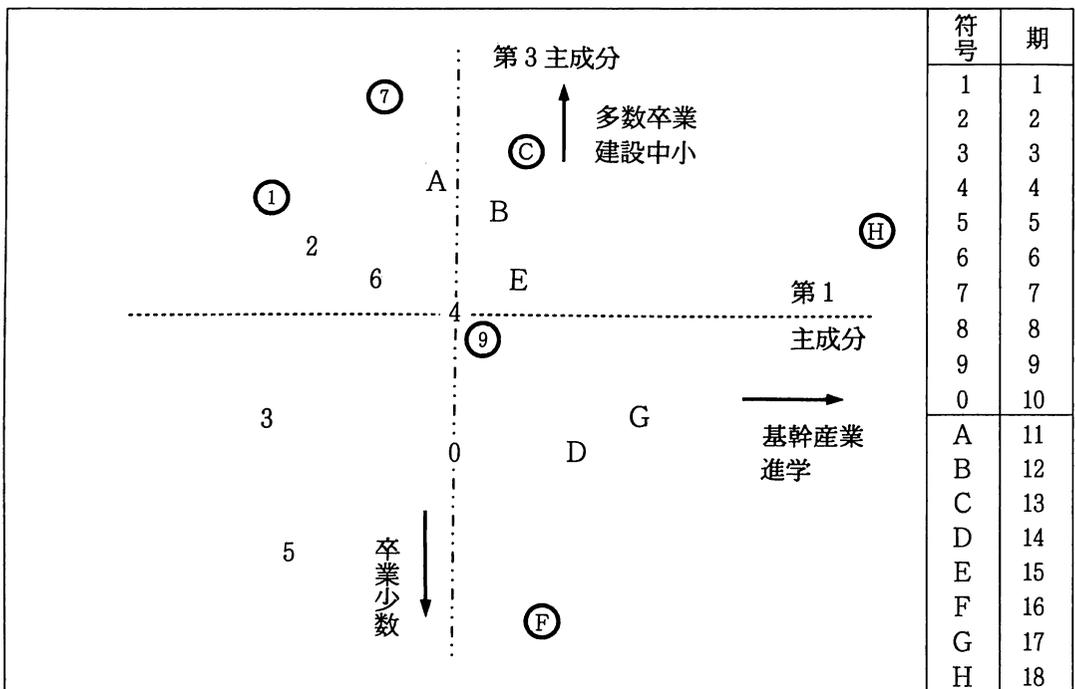


図4. 第1主成分 VS 第3主成分

続マトリックス・データ解析法による土木工学科学生の進路分析

数が増える筈であるが、12期生はもっとも勉強が嫌いなクラス?だったのであろうか。

(2) 第1主成分VS第3主成分(図4)

第3主成分は1・7・13期生が上方にある。卒業生数が多く、建設業中小と化工電工が正の相関にある。就職に苦労した当時の様子が推察できる。一方18期生⑩は在籍42名と最多を誇っているが、上述の相関が薄いため、主成分スコアは中程度で止まっている。

最下方は⑩即ち16期生である。卒業生総数29名と少なく建設業大手ゼロなど1期生と全く対称的な位置にある。9期生は常に原点付近にあり、全体の平均的存在であるといえる。

(3) その他の主成分

第4主成分では3期生と5期生の化工電工とコンクリートの分野が逆であることを指摘している。また第5主成分では建設業大手と地方公務員の負の関係を取上げ、能力のある学生がどちらを選ぶかともどいを示している。これらはすべて部分的な数値の特長で、全体に対するウェイトは低い。

(4) 総合評価

再び図3に目を通してみよう。散布図を見ると1~10期生は左半分に、11~18期生は右半分にあることがわかる。即ち第1主成分は18年間の世相の変遷を示す重要な評価メジャーだと言える。学生の進路は基幹産業、進学へと大きく指向している。

4. 所 感

今回は全国土木28高専に対し、また本文では本校土木に絞り18年間の学生進路を分析し、まずまずの

結論を得ることができた。各個体の特長を単純思考では不可能な次元で捉えることができたのは、マトリックス・データ解析法の効能だといえる。

18期生が大挙進級できたのは筆者の構造力学が甘かったためとのうわさがある。しかし昨年度は17名が単位を落としながらうまく1人5単位止まりとし、追加認定試験をクリアできたのは、全員が相当な潜在能力を有する証拠であろう。高専体育大会では剣道やバドミントンで全国大会へ代表選手を送るなど、文武両面で活達な所を見せている。また綱構造やコンサルタントへの就職が多いのは、もう一人構造力学で毎週深夜まで演習でしごく熱血教授がいたためであろう。この教授は平成元年度で定年退官された。ともかく18期生は学生進路の傾向に大きなインパクトを与えた。これを機に本校土木工学科学生の進路が大きく変貌しようというのか昔の傾向に戻ろうとするのか、引続き数年間の推移を見守りたい。

参 考 文 献

- 1) 米長泰他, マトリックス・データ解析法による土木工学科学生の進路分析 秋田高専研究紀要第25号P. 69~77 1990. 2.
- 2) 今村眞明, 殿守育子他 統計的手法活用マニュアル(主成分分析) MM06-861 三菱重工技術本部・非売品
- 3) 米長泰他 マトリックス・データ解析法による平成元年度プロ野球選手の業績評価 日本品質管理学会第37回研究発表会テキストP. 153~156 1990. 5.