社会学の研究とコンピューターの利用

大阪大学人間科学部　西　田　春　彦

社会学でコンピューターを利用するのは，社会調査，既存統計資料の解析やシミレーションにおいてであるが，本命は社会調査における集計と解析であろう。

1. 社会階層の調査

日本の社会学でコンピューターを大巾に利用したのは，1952 年の六大都市の社会成層と移動についての日本社会学会の調査が多分最初であろう。① この調査は日本の段階的構造を成員の社会的地位の面から把えようとしたもので，社会的地位は本人と近親者の職業，収入，学歴，財産，生活程度などによって規定されると考えて，これらの変数に関係した調査項目を用いて，有効回収 1746 票の面接調査を行った。その集計は統計数理研究所の IBM によって，項目毎の単純集計，項目間のクロス集計および若干の相関係数を算出したものであった。

その後，1955 年に日本社会学会により同じ主題について全国規模の第 2 回標本調査（有効 3677 票の面接調査）がなされ，統計数理研究所で集計された。② 第 2 回調査の集計では，通常の集計のほかに，標本の主観的所属階層の高さ（S）を従属変数とし，標本の職業(X_1)収入（X_2），学歴（X_3)を独立変数とする回帰方程式を試案的に計算している。たとえば，岡山県宇垣地区では，　$S=1.166+0.656X_1+0.631X_2-0.687X_3$，岩手県長坂地区では，$S=0.389+0.667X_1+0.700X_2-0.501X_3$　になっている。

現在ならば，林の数量化理論第 I 類を適用するところであろうが，当時は変数における反応（観測値）を数量化して得点に直し計算している。これらの社会階層の調査は社会調査の数量的処理について一つの大きな刺激となり，複雑で大量な集計のスピード化，正確化に，又資料の統計解析にコンピューターを利用する気運を高めた。ただし，社会調査では調査項目によっては機械集計を行いにくい部分もあって，手集計を併用することが多いし，また一応の集計結果（解析を含めて）を見て標本を分類し直し，再集計するのが通常である。これらの作業の流れは図 1 を参照されたい。③

2. 再集計の例

表 1 は，大阪市愛隣地区調査における層のつくり方を示している。この標本調査は 1967 年 8 月に実施され，1967年 9 月 20 日単純集計終了，1967 年 12 月 9 日クロス集計終了，

1967年12月13日表1による8層で再集計終了となっている。[④] 集計は東洋コンピューターサービス，バロース 270による。これは住民の住居形式と家族類型を組み合せた層によって回答の傾向を見る方が，単純な集計や住居形式と他質問とのクロス集計，家族類型と他質問とのクロス集計よりも意味のある集計であるという判断によるもので，クラスター分析のような発想に基づくものである。

表1　住居形式と家族類型

ⓘは層の番号

	一軒屋・長屋	アパート	宿　屋	その他	計
単　身	③ 17	① 57	②216	③ 39	329
世帯持ち	④159	⑤141	⑥ 92	⑦ 89	481
その他	⑧ 37	⑧ 21	⑧ 18	⑧ 22	98

もう一つ再集計の例をあげよう。G新聞紙の購読の有無（G）を従属変数とし，標本の人生観（Y）と五大全国紙の標本の選好順位の型（Z）を説明変数とし，Gに対してYとZの寄与の程度を知りたい。潜在構造分析を用いて，各標本においてその人生観が地味な生活態度を持つとき1，消費型の生活態度を持つとき2で示す。新聞紙の選好順位の型は，1対比較法に準じた形で得た選好の度数行列に林の数量化理論を適用して，$a>b>c>d>e$, $e>d>c>b>a$,……のような選好順位をもつ人々に分類する。上位4つの型に90 %あまりの人々が所属する。これから表2の度数表を作り，カテゴリ的反応について重回帰分析に準じたL. A. Goodman 流の解析法を適用する。[⑤]

社会学で用いられる解析法には，いわゆる多変量解析法のほかグラフ理論の適用も試みられるようになった。ただし，社会学では実験は実施しにくいことが多いので，その難点を補うための path 分析も次第に用いられるようになっている。 path 分析は変数間の因果関係を求める重回帰分析の特別な形式といえるだろう。また，社会学の扱う資料には質的な変数が多い。男，女や都市，農村，投票した，しないなど，質的なものの処理が望まれるので，林の数量化理論，潜在構造分析などにも社会学者は親近感をもっている。

潜在構造分析は，対象とする集団が相互に排反的ないくつかの同質の潜在的部分集団より成り，その部分集団ごとに調査項目に対する顕在的反応が異なるという考えに基づくもので，1950年にその方法の内容が発表された。潜在構造分析に対するラザースフェルトの解法のプログラムを1959年三和銀行の IBM 650で同行と共同で作ったが，5項目2潜

表2

型	人生観 Y	選好順位 Z	G購読の有無	度数
1	1	1	1	34
2	1	1	2	183
3	1	2	1	9
4	1	2	2	58
5	1	3	1	29
6	1	3	2	25
7	1	4	1	5
8	1	4	2	35
9	1	5	1	6
10	1	5	2	30
11	2	1	1	20
12	2	1	2	104
13	2	2	1	5
14	2	2	2	29
15	2	3	1	12
16	2	3	2	11
17	2	4	1	2
18	2	4	2	16
19	2	5	1	4
20	2	5	2	13

在クラスの条件で，1450 ステップ，印刷時間を含め演算時間18分であった。[6]

3. 個体特性の計算

社会学にとって1つの課題は対象の個体特性の解明にあると考える。個体特性はその個体を特色づける成分（ないし因子）とその大きさ，成分間の相関関係で表現できるだろう。いま，日本の7大都市（東京，横浜，名古屋，京都，大阪，神戸，北九州）の人口を例にとって，特性の算出の手順を以下に示す。[7]

① 第i都市のn個の変数の資料から相関行列$\boldsymbol{R}_i$をつくる。$i=1, 2, \cdots\cdots, N_0$

② $\boldsymbol{R}_i$を式(2)のように分解する。

$$\boldsymbol{R}_i = \boldsymbol{A}_i \ \boldsymbol{A}'_i \qquad (i=1,2,\cdots\cdots,N_0) \tag{2}$$

この分解には色々の手法が可能であろうが，ここでは主成分分析法を用いた。$\boldsymbol{A}_i$を次のように扱う。

③
$$\begin{aligned}\boldsymbol{A} &= (\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots\cdots\cdots, \boldsymbol{A}_N) \\ &= \boldsymbol{A}_0 \ (\boldsymbol{Q}_1, \boldsymbol{Q}_2, \cdots\cdots\cdots, \boldsymbol{Q}_N)\end{aligned} \tag{3}$$

ただし，

$$\boldsymbol{A}_i = \boldsymbol{A}_0 \ \boldsymbol{Q}_i \qquad (i=1,2,\cdots\cdots,N) \tag{4}$$

$\boldsymbol{A}_i$のrankはnより小さいとしておく。$\boldsymbol{A}_i$のrankの最大をMとすれば，$\boldsymbol{A}_0$のrankはMである。この$\boldsymbol{A}_0$はN個の都市に共通した成分行列で，共通空間と呼ぶ。$\boldsymbol{Q}_i$はこの$\boldsymbol{A}_0$を第i都市の$\boldsymbol{A}_i$に変換する行列であって，第i都市の特別な性格によって定まると考える。

④ $\boldsymbol{Q}_i$の後から$\boldsymbol{Q}'_i$をかける。

$$\Phi_i = \boldsymbol{Q}_i \ \boldsymbol{Q}'_i \qquad (i=1,2,\cdots\cdots\cdots, N) \tag{5}$$

⑤ Φ_iの主対角要素の平方根をϕ_j（$j=1,2,\cdots\cdots$, M）とする。ϕ_j（$j=1,2,\cdots\cdots$ M）を要素とする対角行列$\boldsymbol{W}_i$をつくる。

$$\boldsymbol{W}_i = \begin{bmatrix} \phi_1 & & & \bigcirc \\ & \phi_2 & & \\ & & \ddots & \\ \bigcirc & & & \phi_M \end{bmatrix} \tag{6}$$

$i=1,2,\cdots\cdots\cdots, N_0$

$\boldsymbol{W}_i$は正則行列であるとし，$\boldsymbol{W}_i^{-1}$が存在するものとする。$\boldsymbol{W}_i$は第i都市の成分の大きさを意味する。

⑥
$$\boldsymbol{R}^*_i = \boldsymbol{W}_i \ \Phi_i \ \boldsymbol{W}_i^{-1} \qquad (i=1,2,\cdots\cdots\cdots, N) \tag{7}$$

$\boldsymbol{R}^*_i$は第i都市のM個の成分間の相関係数の行列である。

⑦ $\boldsymbol{E}_i = \boldsymbol{A}_0 \ \boldsymbol{W}_i$ 　　　　$(i = 1, 2, \cdots\cdots, \ N)$ 　　　　(8)

$\boldsymbol{E}_i$ は共通の成分行列を第 i 都市の成分の大きさ（つまり重み）で修正したもので標準化成分空間と呼んでおく。以上から，第 i 都市の個体特性はその都市の $\boldsymbol{W}_i$ と $\overset{*}{\boldsymbol{R}}_i$ で表現できると考える。なお，$\boldsymbol{R}_i$ は $\overset{*}{\boldsymbol{R}}_i$ と次の関係にある。

$$\begin{aligned} \boldsymbol{R}_i &= \boldsymbol{A}_0 \ \boldsymbol{W}_i \ \boldsymbol{W}_i^{-1} \ \Phi_i \ \boldsymbol{W}_i^{'-1} \boldsymbol{W}_i \ \boldsymbol{A}_0' \\ &= \boldsymbol{E}_i \ \overset{*}{\boldsymbol{R}}_i \ \boldsymbol{E}_i' \qquad (9) \\ & i = 1, 2, \cdots\cdots\cdots, \ \mathrm{N}_0 \end{aligned}$$

計算結果は別表のとおりであるが，第 1 成分は扶養，第 2 成分は職業の種類，第 3 成分は従業上の地位，第 4 成分は人口の集中を表わし，東京は 1970 年には 2 成分に近づき，横浜北九州も 2 成分の構造に近づいている。7 大都市の人口現象はより単純な構造へ動いているように見える。(この研究は昭和 4 8 年度文部省科研費による。計算は三菱電機 ,MELCOM 7700 による。)

日本の社会学でコンピューターを解析用に利用する傾向はまだ強くはない。社会学的な事象に計量的な方法を適用できる範囲は，現状では相当に限られていることにもその一因があろう。

(昭和 5 1 年 7 月)

参 考 文 献

1) 尾高邦雄，西平重喜；　わが国六大都市の社会的成層と移動，社会学評論　12，2－5 1，(1953).
2) 日本社会学調査委員会；　わが国における社会的移動，社会学評論　25，2－60,(1956).
日本社会学会調査委員会編；　日本社会の階層的構造，有斐閣，(1958).
3) 西田春彦，複合標本方式の適用，社会学評論　76，64－72,(1969).
4) 大阪社会学研究会；　愛隣地区綜合実態調査報告　4，(1968).
5) 平松　闊，山本剛郎；　グッドマンモデルの適用例，社会学評論　102，(1975).
6) 西田春彦，池田一貞；　潜在構造分析の理論と計算法，和歌山大学教育研究所，77－79,(1960).
7) 西田春彦；　計量社会学入門，森北出版，133－139,(1973).

図1 調 査 手 順 図

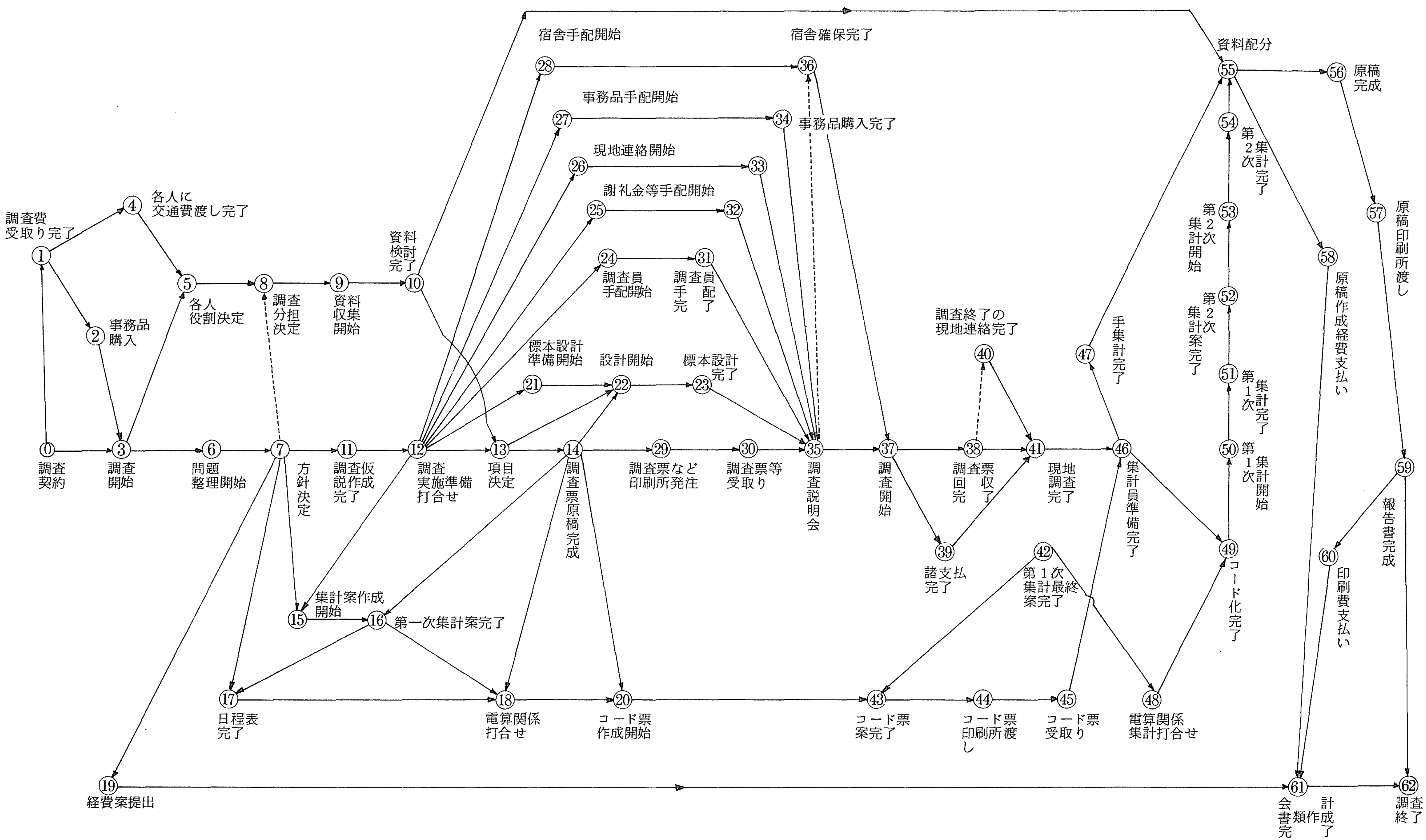
調査費受取り完了
各人に交通費渡し完了
事務品購入
調査契約
調査開始
各人役割決定
問題整理開始
調査分担決定
方針決定
資料収集開始
資料検討完了
調査仮説作成完了
調査実施準備打合せ
項目決定
集計案作成開始
第一次集計案完了
日程表完了
電算関係打合せ
コード票作成開始
経費案提出
宿舎手配開始
宿舎確保完了
事務品手配開始
事務品購入完了
現地連絡開始
謝礼金等手配開始
調査員手配開始
調査員手配完了
標本設計準備開始
設計開始
標本設計完了
調査票原稿完成
調査票など印刷所発注
調査票等受取り
調査説明会
調査開始
調査票回収完了
諸支払完了
調査終了の現地連絡完了
現地調査完了
第1次集計最終案完了
コード票案完了
コード票印刷所渡し
コード票受取り
集計員準備完了
電算関係集計打合せ
手集計完了
コード化完了
第1次集計開始
第1次集計完了
第2次集計案完了
第2次集計開始
第2次集計完了
資料配分
原稿完成
原稿作成経費支払い
原稿印刷所渡し
報告書完成
印刷費支払い
会計書類作成完了
調査終了

表3　7大都市が相互に独立した単位であるとした場合の W_i と R_i^*

W_i（成分の大きさ）	昭和35年 1	2	3	4	昭和40年 1	2	3	4	昭和45年 1	2	3
東　京	.378	.377	.468	.223	.408	.296	.497	.260	.394	.405	.125
横　浜	.409	.365	.239	.472	.378	.397	.174	.538	.387	.365	.334
名古屋	.414	.408	.403	.394	.366	.415	.463	.484	.401	.403	.463
京　都	.359	.407	.436	.264	.381	.399	.382	.113	.387	.393	.394
大　阪	.389	.356	.444	.255	.403	.405	.466	.154	.377	.379	.414
神　戸	.373	.363	.373	.464	.374	.354	.353	.396	.384	.395	.474
北九州	.310	.363	.181	.470	.326	.363	.135	.461	.303	.289	.323

R_i^*（成分間の相関関係）		昭和35年 1	2	3	4	昭和40年 1	2	3	4	昭和45年 1	2	3
東　京	2	－.085				.141				－.113		
	3	－.038	－.203			.025	.199			－.776	－.528	
	4	－.977	－.098	－.027		－.312	.805	－.236				
横　浜	2	.061				.071				.067		
	3	.033	.927			.350	－.878			.312	－.491	
	4	.111	－.242	.132		.028	.019	.234				
名古屋	2	－.029				.009				.045		
	3	－.077	－.082			－.012	.062			.012	－.053	
	4	.016	.038	.344		.052	.133	－.034				
京　都	2	.070				－.120				－.062		
	3	.117	－.059			.188	.226			.241	.188	
	4	.800	－.314	－.345		.375	－.865	.262				
大　阪	2	.063				－.048				.020		
	3	.037	－.190			－.053	－.009			－.011	.093	
	4	－.140	.922	.137		－.266	－.273	－.906				
神　戸	2	－.084				－.068				－.069		
	3	－.233	－.191			－.207	－.148			－.025	－.132	
	4	.030	.046	.144		.259	－.113	.888				
北九州	2	.003				.054				.209		
	3	.432	.534			－.377	.033			－.355	.832	
	4	－.060	－.084	－.794		－.126	－.125	－.869				