

(19)日本国特許庁(JP)

(12)特許公報(B1)

(11)特許番号

特許第7381150号
(P7381150)

(45)発行日 令和5年11月15日(2023. 11. 15)

(24)登録日 令和5年11月7日(2023. 11. 7)

(51)Int. Cl. F I
 G 0 6 F 18/2133 (2023. 01) G 0 6 F 18/2133
 G 0 6 F 17/10 (2006. 01) G 0 6 F 17/10 Z

請求項の数 7 (全 24 頁)

(21)出願番号	特願2023-79260(P2023-79260)	(73)特許権者	322002104 株式会社エリアビジョン 山梨県甲府市飯田2-6-6
(22)出願日	令和5年5月12日(2023. 5. 12)	(74)代理人	110001830 弁理士法人東京U I T国際特許
審査請求日	令和5年5月12日(2023. 5. 12)	(72)発明者	新藤 久和 山梨県甲府市飯田2-6-6 株式会社エ リアビジョン内
早期審査対象出願		審査官	北川 純次

最終頁に続く

(54)【発明の名称】 数量化親和図法実行装置およびプログラム

(57)【特許請求の範囲】

【請求項 1】

n (n は 3 以上の整数) 個の言語データの入力を受け付ける言語データ入力手段 ,
 受け付けられた n 個の言語データを用いて , 互いに異なる 2 つずつの言語データ i , j
 (i j) のすべての組み合わせを表示する表示手段 ,

表示された 2 つの言語データ i , j の組み合わせのそれぞれについての親和度 k を表す
 数値の入力を受け付ける親和度入力手段 ,

n 行 n 列の n 次正方対称行列であって , 上記親和度入力手段によって入力された言語デ
 ータ i , j の組み合わせの親和度 k を上記 n 次正方対称行列の対応する成分 i , j の要素
 とし , 対角要素を互いに異なる 2 つの言語データ i , j の組み合わせについて受け付けら
 れる親和度 k を表す数値のうちの最大値とし , 行項目および列項目を同じとする n 次正方
 対称行列を表すデータを作成する行列作成手段 , ならびに

λ^2 を固有値 , P を上記行列作成手段によって作成された n 次正方対称行列 , P^T を n
 次正方対称行列 P の転置行列 , I を単位行列 , p_x を n 次正方対称行列 P の行和ベクトル
 の要素を対角要素とする行列 , x をスコアベクトルとしたときに , 以下のスコアベクトル
 算出式

$$\left(p_x^{-1/2} P^T p_x^{-1} P p_x^{-1/2} - \lambda^2 I \right) p_x^{1/2} x = 0$$

に基づいて , 上記行列作成手段によって作成された n 次正方対称行列から算出される複

数の固有値 λ^2 のそれぞれごとに、 n 個の言語データのそれぞれについてのスコアベクトル x を算出するスコアベクトル算出手段を備えている、

数量化親和図法実行装置。

【請求項 2】

上記スコアベクトル算出手段は、

上記行列作成手段によって作成される n 次正方対称行列から算出される複数の固有値 λ^2 のうち、固有値 $\lambda^2 = 1$ を除く複数の固有値 λ^2 のそれぞれごとに、 n 個の言語データのそれぞれについてのスコアベクトル x を算出するものである、

請求項 1 に記載の数量化親和図法実行装置。

【請求項 3】

上記スコアベクトル算出手段は、

上記行列作成手段によって作成される n 次正方対称行列から算出される複数の固有値 λ^2 のうち、固有値 $\lambda^2 = 1$ を除き、かつ複数の固有値 λ^2 のうち値の大きなものから順に m 個 ($m \leq n - 1$) の固有値 λ^2 のそれぞれを用いて、 n 個の言語データのそれぞれについてのスコアベクトル x を算出するものである、

請求項 1 に記載の数量化親和図法実行装置。

【請求項 4】

2 つまたは 3 つの固有値 λ^2 のそれぞれを用いて言語データのそれぞれについて算出されるスコアベクトル x を、互いに直交する 2 軸を備える 2 次元グラフまたは互いに直交する 3 軸を備える 3 次元グラフ上にプロットすることによって表すグラフ表示データ作成手段をさらに備えている、

請求項 3 に記載の数量化親和図法実行装置。

【請求項 5】

上記プロットを含む 3 次元グラフを表示画面に回転表示させる回転表示手段を備えている、

請求項 4 に記載の数量化親和図法実行装置。

【請求項 6】

上記グラフ表示データ作成手段は、

上記プロットにそれぞれに対応させて上記言語データを表示するものである、

請求項 4 に記載の数量化親和図法実行装置。

【請求項 7】

受け付けられた n (n は 3 以上の整数) 個の言語データを用いて、互いに異なる 2 つずつの言語データ i, j ($i \neq j$) のすべての組み合わせを表示画面に表示し、表示された 2 つの言語データ i, j の組み合わせのそれぞれについての親和度 k を表す数値の入力を入力装置から受け付ける処理をコンピュータに実行させる親和度入力プログラム、

n 行 n 列の n 次正方対称行列であって、受け付けられた言語データ i, j の組み合わせの親和度 k を上記 n 次正方対称行列の対応する成分 i, j の要素とし、対角要素を互いに異なる 2 つの言語データ i, j の組み合わせについて受け付けられる親和度 k を表す数値のうちの最大値とし、行項目および列項目を同じとする n 次正方対称行列を表すデータを作成する処理をコンピュータに実行させる行列作成プログラム、ならびに

λ^2 を固有値、 P を上記行列作成プログラムによって作成された n 次正方対称行列、 P^T を n 次正方対称行列 P の転置行列、 I を単位行列、 p_x を n 次正方対称行列 P の行和ベクトルの要素を対角要素とする行列、 x をスコアベクトルとしたときに、以下のスコアベクトル算出式

$$\left(p_x^{-1/2} P^T p_x^{-1} P p_x^{-1/2} - \lambda^2 I \right) p_x^{1/2} x = 0$$

に基づいて、上記行列作成プログラムによって作成された n 次正方対称行列から算出される複数の固有値 λ^2 のそれぞれごとに、 n 個の言語データのそれぞれについてのスコアベクトル x を算出する処理をコンピュータに実行させるスコアベクトル算出プログラムを

10

20

30

40

含む，

数量化親和図法実行プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

この発明は数量化親和図法実行装置および数量化親和図法実行プログラムに関する。

【背景技術】

【0002】

親和図法は、未知・未経験の分野、あるいは未来・将来の問題など、モヤモヤとしてハッキリしていない、つまり混沌とした状態の中から、事実あるいは推定、予測、発想、意見などを言語データとしてとらえ、それらの言語データを相互の親和性によってまとめあげ、構造を明らかにする方法である（非特許文献参照）。親和図法は漠然とした課題をテーマとして設定し、設定されたテーマに関して作成される統合図から課題を明確化することを狙いとする手法であり、未知の課題や未経験の課題など、不明確な課題に対する解決策を導き出すために用いられる。クオリティマネジメント用語辞典（吉澤正編集委員長，2004年，日本規格協会）では、親和図法を「未来・将来の問題，未知・未経験の問題など，モヤモヤしてははっきりしていない問題について，事実，意見，発想を言語データでとらえ，それらの相互の親和性によって統合した図を作成することにより，解決すべき問題の所在，形態を明らかにしていく方法」とも定義する。

10

【0003】

親和図法は一般には次のように進められる。

- (1) テーマの設定
- (2) テーマに関連する事実，意見，発想等の言語データ（テキストデータ）の収集
- (3) 言語データを記載したカードの作成
- (4) 親和性のある（関連している）言語データが記載されたカード同士のグループ化（一次グループの作成）（カード寄せ）
- (5) 親和性のある一次グループ同士のグループ化（二次グループの作成）
- (6) 一次グループおよび二次グループの配置（関係があるものを近い位置に配置するなど）
- (7) グループ同士の関係性（因果関係，相関，相反）の明示（一般には実線，片側矢印線，両側矢印線が用いられる）

20

30

【0004】

しかしながら，言語データに複合的な内容のもの（二つ以上の意味を持つようなもの）が含まれているような場合に特に，たとえばグループ化が親和図法の実施者によって異なることがあり，そうすると当然に異なる結果（異なる解決すべき問題の所在，形態）が導き出されることになる。

【先行技術文献】

【特許文献】

【0005】

【非特許文献1】納谷嘉信編，「おはなし新QC七つ道具」，第1版，財団法人日本規格協会，1987年4月20日，pp.29 - 47

40

【発明の概要】

【発明が解決しようとする課題】

【0006】

この発明は，親和図法の結果が実施者によって異なるようにすることを目的とする。

【0007】

この発明はまた，潜在的な事象（事柄）の検討も含めて問題の考察を支援することを目的とする。

【課題を解決するための手段】

50

【0008】

この発明による数量化親和図法実行装置は、 n (n は3以上の整数)個の言語データの
 入力を受け付ける言語データ入力手段、受け付けられた n 個の言語データを用いて、互い
 に異なる2つずつの言語データ i, j ($i \neq j$)のすべての組み合わせを表示する表示手
 段、表示された2つの言語データ i, j の組み合わせのそれぞれについての親和度 k を表
 す数値の入力を受け付ける親和度入力手段、 n 行 n 列の n 次正方対称行列であって、上記
 親和度入力手段によって入力された言語データ i, j の組み合わせの親和度 k を上記 n 次
 正方対称行列の対応する成分 i, j の要素とし、対角要素を上記親和度の最大値とし、行
 項目および列項目を同じとする n 次正方対称行列を表すデータを作成する行列作成手段、
 ならびに λ^2 を固有値、 P を上記行列作成手段によって作成された n 次正方対称行列、 P
 T を n 次正方対称行列 P の転置行列、 I を単位行列、 p_x を n 次正方対称行列 P の行和ベ
 クトルの要素を対角要素とする行列、 x をスコアベクトルとしたときに、以下のスコアベ
 クトル算出式

10

$$\left(p_x^{-1/2} P^T p_x^{-1} P p_x^{-1/2} - \lambda^2 I \right) p_x^{1/2} x = 0$$

に基づいて、上記行列作成手段によって作成された n 次正方対称行列から算出される複数
 の固有値 λ^2 のそれぞれごとに、 n 個の言語データのそれぞれについてのスコアベクトル
 x を算出するスコアベクトル算出手段を備えている。

【0009】

複数の言語データのうちの2つずつの組み合わせについてそれらの親和度 k (親和性、
 類似度、関連度と言ってもよい)が数値によって入力される。入力される親和度は、与え
 られる複数の言語データのうちの2つずつに対して、かつすべての組み合わせ (順序は問
 わない)に対して入力される。なお、後述するように、同一の言語データ同士の組み合わ
 せについては数値の入力を受け付ける必要は必ずしもない (言語データ i, j について、
 $i \neq j$)。同一の言語データ同士についての親和度は最高 (最大) の親和度とされるべき
 だからである。

20

【0010】

親和度の数値は、最も簡単には、1 (親和度有) および0 (親和度無) のいずれかとさ
 れる。もちろん、0, 1, 2, 3など、有無に加えて程度を表すことができる数値を用い
 て親和度を表す数値の入力を受け付けてもよい。

30

【0011】

入力された言語データ i, j の組み合わせの親和度 k が用いられて、受け付けられた言
 語データの数 n と等しい行数および列数を持つ n 次元の正方行列が作成される。作成され
 る n 次元正方行列は列項目と行項目を同じとするもので、入力された言語データ i, j の
 組み合わせの親和度 k を上記 n 次正方行列の対応する成分 i, j の要素とするので、対称
 行列 (n 次正方対称行列) となる。また、後述するように、対角要素 ($i = j$ に対応) は
 上記親和度の最大値とすることができる。対角要素は、同一の言語データ同士について
 の親和度に対応し、上述のように、同一の言語データ同士についての親和度は最高 (最大)
 の親和度とするべきだからである。

【0012】

この発明によると、作成された n 次正方対称行列から算出される複数の固有値 λ^2 が用
 いられて、固有値 λ^2 ごとに、 n 個の言語データのそれぞれについてのスコアベクトル x
 がスコアベクトル算出式にしたがって算出される。

40

【0013】

n 次正方対称行列からは n 個の固有値 λ^2 が算出される。ただし、数量化親和図法では
 必ず、固有値 $\lambda^2 = 1$ が得られ、これに対応するスコアベクトルの要素の値はすべて等し
 くなってしまうスコアベクトルとしての意味がなくなるので除外してもよい。この場合
 には、固有値 $\lambda^2 = 1$ を除く複数の固有値 λ^2 をスコアベクトル算出式に代入することによ
 ってスコアベクトルは求められることになり、 $n - 1$ 個の固有値 λ^2 のそれぞれごとに、

スコアベクトル x は求められる。もっとも、後述するように、 $n - 1$ 個の固有値 λ^2 のすべてを用いてスコアベクトル x を求める必要は必ずしもない。

【0014】

スコアベクトル x は、 n 次正方対称行列 P （およびその転置行列 P^T ）をスコアベクトル算出式に代入することによって求められるので、求められるスコアベクトル x は n 個の要素（成分）を含む。この n 個の要素が上述の n 個の言語データのそれぞれについてのスコアベクトルを構成する。

【0015】

固有値 $\lambda^2 = 1$ を除く $n - 1$ 個の固有値 λ^2 については、その値が大きければ大きいほど、 n 次正方対称行列の相関係数が大きいものとなる。すなわち、処理対象の n 個の言語データのそれぞれの全体における位置づけ（いずれの言語データと相関しており、いずれの言語データと相関していないか、その程度）を把握しやすいものとなる。

【0016】

一実施態様では、上記スコアベクトル算出手段は、上記行列作成手段によって作成される n 次正方対称行列から算出される複数の固有値 λ^2 のうち、固有値 $\lambda^2 = 1$ を除き、かつ上記複数の固有値 λ^2 のうち値の大きなものから順に m 個（ $m \leq n - 1$ ）の固有値 λ^2 のそれぞれを用いて、 n 個の言語データのそれぞれについてのスコアベクトル x を算出する。固有値 λ^2 が小さくなればなるほど相関係数が小さくなり、二元表の行と列の関係を説明する度合が小さくなる。それで、 λ^2 は寄与率とも呼ばれる。したがって、二元表の行と列の関係を解析するにあたって、寄与率を70%程度確保したいとすれば、累積寄与率（固有値の大きい方から対応する寄与率の累積和）が70%となるまでの固有値に属するスコアを利用すればよいことになる。たとえば、処理対象の正方対称行列が31次のものである場合、31個の固有値 λ^2 が求められる。固有値 $\lambda^2 = 1$ を除いた30個の固有値 λ^2 のうち、大きいものから累積寄与率が所望の大きさになるまでのスコアを用いればよいので、それより高次の固有値のスコアは使わないため計算を節約でき、処理速度を向上することができる。

【0017】

数量化親和図法実行装置は、一実施態様では2つまたは3つの固有値 λ^2 のそれぞれを用いて言語データのそれぞれについて算出されるスコアベクトル x を、互いに直交する2軸を備える2次元グラフまたは互いに直交する3軸を備える3次元グラフ上にプロットすることによって表すグラフ表示データ作成手段をさらに備えている。 n 個の言語データの全体における位置づけ、さらには類似する言語データのグループ（クラスタ）をプロットによって視覚的に把握することができる。

【0018】

2次元グラフや3次元グラフに加えて、複数の固有値 λ^2 を用いて算出されるスコアベクトル x を用いて樹形図やヒートマップを作成してもよい。

【0019】

この発明によると、与えられる言語データ同士の関係を、複数の観点（複数の固有値）のそれぞれから把握することができ、親和性のある（似ている）言語データ同士のグループを、2次元グラフ、3次元グラフ、樹形図、ヒートマップ等を用いて把握することができ、解決すべき問題における傾向や、問題を解決するための対策を見つけやすくすることができる。

【0020】

2次元グラフ、3次元グラフ等は、複数の言語データの2つずつの組み合わせについての親和度 k が同じであれば同じ結果となり、実施者によって異なることはない。また、固有値 λ^2 ごとに算出されるスコアベクトル x を検討することによって、複数の言語データ（複数の言語データから作成される行列）に潜在している事象（事柄）を検討することも可能である。

【0021】

この発明は、数量化親和図法をコンピュータに実行させるためのプログラムも提供する

。

【図面の簡単な説明】

【0022】

【図1】親和図法を数量化して行う方法を実行するコンピュータシステムのブロック図である。

【図2】3つ組データ作成プログラムの処理を示すフローチャートである。

【図3】3つ組データ作成プログラムの処理を示すフローチャートである。

【図4】子供の貧困をテーマにしたときの事象を表す文章（テキストデータ）例を示す。

【図5】3つ組データテーブル・ファイルの構造を示す。

【図6】親和度入力画面を示す。

10

【図7】2元表変換プログラムの処理を示すフローチャートである。

【図8】2元表変換プログラムによって作成される2元表（行列P）を示す。

【図9】数量化3類適用プログラムの処理を示すフローチャートである。

【図10】数量化3類適用プログラムの処理を示すフローチャートである。

【図11】第1固有値を用いて算出されるスコアベクトルと第2固有値を用いて算出されるスコアベクトルを、模式的に示している。

【図12】図8に示す2元表から算出される固有値，寄与率および累積寄与率を示す。

【図13】第1固有値～第7固有値のそれぞれを用いて算出される，言語データごとのスコアを示す。

【図14】2次元散布図を示す。

20

【図15A】3次元散布図を示す。

【図15B】3次元散布図を示す。

【図16A】回転表示される3次元散布図を示す。

【図16B】回転表示される3次元散布図を示す。

【図16C】回転表示される3次元散布図を示す。

【図16D】回転表示される3次元散布図を示す。

【図17】樹形図を示す。

【図18】ヒートマップを示す。

【発明を実施するための形態】

【0023】

30

図1は数量化親和図法を実行するコンピュータシステム（以下，「数量化親和図法実行装置」という）のブロック図である。

【0024】

数量化親和図法実行装置1は，システム全体を統括的に制御するCPU（Central Processing Unit）11を備えている。CPU11には，バスを介して，BIOS（Basic Input/Output System）等を記憶する不揮発性メモリから構成されるROM（Read Only Memory）12，プログラムおよびデータを一時的に記憶する揮発性メモリから構成されるRAM（Random Access Memory）13，キーボード，マウス等を含み各種の指示等を入力する入力部14，処理結果を表示するディスプレイ等を含む表示部15，ハードディスクやSSD（Solid State Drive）等から構成され，後述する被分析データ16-1，数量化親和図法実行プログラム16-2を記憶する記憶部16を備えている。

40

【0025】

数量化親和図法実行装置1は，社会問題を生じさせているさまざまな事象（事柄）を整理し，潜在的な事象の抽出を含めて，社会問題の構造を考察するために用いられる。後述するように数量化親和図法実行装置1によって，因果事象がたとえば散布図として表される。この散布図を検討することによって複数の解決策が導かれる。

【0026】

記憶部16に記憶される数量化親和図法実行プログラム16-2は，後述する（1）3つ組データ作成プログラム（関数），（2）2元表変換プログラム（関数），および（3）数量化3類適用プログラム（関数）を含む。これらの3つのプログラムを以下詳細に説

50

明する。パラメータを指定して呼び出すことによってこれらの3つのプログラムをそれぞれCPU11において実行することができる。

【0027】

(1) 3つ組データ作成プログラム

図2および図3は3つ組データ作成プログラムの処理を示すフローチャートである。

【0028】

3つ組データ作成プログラムは、あらかじめ用意される複数の言語データから選ばれる2つずつの言語データの組(i, j)のそれぞれについて、その親和性の度合い(親和度)を表すデータ k を関連付けたデータ(これを「3つ組データ(i, j, k)」という)を作成するプログラムである。

10

【0029】

はじめに処理対象の言語データがRAM13に読み込まれる(ステップS10)。言語データはテキスト形式データによってあらかじめ準備される。ここでは言語データの総数を n とする。

【0030】

図4は言語データ(言語によって表されるデータ(テキストデータ))の具体例を示すもので、 $n = 31$ (31個の言語データ)の例を示している。図4に示す31個の言語データ2は、子供の貧困に係る調査においてヒヤリングによって得られた情報を示している。社会的課題(この例では「子供の貧困を無くす」)を解決しようとする場合様々な観点からの情報が集積され、これらに数量化親和図法を適用することにより、課題を多角的に検討して明確化でき、解決策の導出を支援することができる。

20

【0031】

言語データのそれぞれには言語データを特定するためのID(データID)が関連付けられる。図4に示す例では1~31の互いに異なる数字(整数)がデータIDとして用いられている。データIDによって31個の言語データのそれぞれが識別(特定)される。

【0032】

図2に戻って、新規に(初めて)3つ組データを作成するか、作成途中の3つ組データの作成を再開するかが判断される(ステップS11)。

【0033】

図5は作成される3つ組データが保存される、3つ組データテーブル・ファイル31の構造の一例を示している。

30

【0034】

以下に詳述するように、3つ組データが作成されると、作成された3つ組データは3つ組データテーブル・ファイル31に保存される。新たに保存される3つ組データは、3つ組データテーブル・ファイル31において最下行に保存され、3つ組データが作成(保存)されるたびにファイル31の行数が増えていく(行データが下に積み重ねられる)。3つ組データは上述したようにデータID i 、およびデータID j によって特定される2つの言語データの組に親和度 k を関連付けたデータである。

【0035】

新規に3つ組データを作成する場合(図2のステップS11でYES)、3つ組データを保存するための3つ組データテーブル・ファイル31(図5)が新規に作成されてRAM13に記憶される。また、2つのカウンタ i および j がRAM13に記憶される。カウンタ i の初期値 i_s として「1」が、カウンタ j の初期値 j_s として、カウンタ i の初期値 i_s に1を加えた数、ここでは「2」が、それぞれ設定される(ステップS12)。

40

【0036】

作成途中の3つ組データの作成を再開する場合の処理(ステップS11でNO、ステップS13)については後述する。

【0037】

カウンタ i の初期値 i_s (=1)から、言語データの総数 n から1を減算した値($n-1$)まで、以下の処理が繰り返される(以下、 i ループという)(ステップS14)。こ

50

ここでは言語データの総数 n が 31 個であるから、カウンタ i には 1 ~ 30 の数字がそれぞれ順次入力される。

【0038】

また、カウンタ j の初期値 (= 2) から言語データの総数 n まで、以下の処理が繰り返される(以下、 j ループという)(ステップ S15)。カウンタ j には 2 ~ 31 の数字がそれぞれ順次入力される。

【0039】

以下に説明するように、はじめにカウンタ j のみがインクリメントされてカウンタ j が n (= 31) に到達すると、次にカウンタ i がインクリメントされて、その状態で再びカウンタ j のみがインクリメントされることによって、複数の言語データから選ばれる 2 つずつの言語データの組のすべてについて、以下に説明するように親和度 k が関連付けられる。

10

【0040】

カウンタ i 、 j の値(データ ID)にしたがって選ばれる、複数の言語データのうちの 2 つの言語データの組(i 、 j)が提示(表示画面に表示)される(ステップ S16)。

【0041】

たとえば、新規処理であれば、処理開始直後は、上述のカウンタ i の初期値 i_s が「1」、カウンタ j の初期値 j_s が「2」であるから、カウンタ $i = 1$ 、カウンタ $j = 2$ となり、データ ID が「1」である言語データと、データ ID が「2」である言語データの組が表示画面に表示される。

20

【0042】

図 6 は、言語データ表示画面(親和度入力画面)40 の一例を示している。言語データ表示画面 40 には、2 つ(一組)の言語データ 41、42 と、入力のやり方を説明する文章 43 が表示される。

【0043】

言語データ表示画面 40 に表示される一組の言語データについての親和度が入力される(受け付けられる)(ステップ S17、S18)。図 6 に示す例では、データ ID 「1」の言語データ(「両親は心を開かない」とデータ ID 「2」の言語データ(「親は支援の種類や内容を知らない」)が上下に並んで表示されている。この上下に並んで表示される 2 つの言語データの親和度の有無が、入力部 14 を用いて数値によって入力される(ステップ S18 で「有」、ステップ S19 で「親和度入力」)。たとえば、表示される 2 つの言語データに親和性があれば数字の 1 が入力部 14 から入力され(ステップ S18 で「有」、ステップ S19 で「親和度入力」)、親和度がなければ改行(数字の 0 でもよい)が入力部 14 から入力される(ステップ S18 で「無」)。

30

【0044】

親和度の有無に代えて、親和度の程度を入力するにしてもよい。たとえば、親和性がある場合に、その親和度が強ければ数字の「3」を、親和度が中程度であれば数字の「2」を、親和度が弱ければ数字の「1」を、それぞれ入力するにしてもよい。表示される一組の言語データの親和度(数値)は一般的には専門家の判断により入力される。もちろん、対象となる社会的課題(この実施例では子供の貧困)に詳しい複数の人物の話し合いによって一組の言語データの親和度を決めてもよい。

40

【0045】

親和度の有無に加えて、処理中断も、言語データ表示画面 40 を用いて入力される。たとえば「-1」の数値を入力することによって処理中断が入力される(ステップ S18 で「中断」、ステップ S19 で「中断入力」)。この機能は、言語データが多くなると、親和度の入力の数が指数関数的に増加し、一気呵成に入力を行うことが困難になるために設けられたものである。

【0046】

親和度が入力された場合(ステップ S18 で「有」、ステップ S19 で「親和度入力」)、カウンタ i のデータ ID、カウンタ j のデータ ID および入力された親和度 k (この

50

実施例では $k = 1$ 。上述したように $k = 3$ または 2 または 1 としてもよい。) を含む 3 つ組データが、データ保存用テーブルファイルの最下行に追加される (ステップ S 2 0)。

【 0 0 4 7 】

j ループが終わりかどうか、ここではカウンタ j が「 3 1 」に到達しているかどうか判断される (ステップ S 2 1)。j ループが終わっていない場合 (ステップ S 2 1 で NO)、カウンタ j がインクリメントされ (ステップ S 2 4)、その後に j ループが繰り返される (ステップ S 1 5)。すなわち、カウンタ i が固定された状態でカウンタ j がインクリメントされて、次のループがスタートする。これによってデータ ID「 1 」とデータ ID「 2 」の組の処理が終わると、次にデータ ID「 1 」とデータ ID「 3 」の組の親和度 k が入力され、さらに次のループではデータ ID「 1 」とデータ ID「 4 」の組の親和度 k が入力され、最後にデータ ID「 1 」とデータ ID「 3 1 」の組の親和度 k が入力される。

10

【 0 0 4 8 】

j ループが終わると (カウンタ j が「 3 1 」に到達すると) (ステップ S 2 1 で YES)、カウンタ i (初期値 i_s) がインクリメントされ、インクリメントされた値が新たな初期値 i_s として用いられて (ステップ S 2 2)、ステップ S 1 4 に戻る (ステップ S 2 3 で NO)。カウンタ j には、その初期値 j_s として、上述のカウンタ i の初期値 i_s に 1 を加えた数が設定される (ステップ S 2 2)。たとえば、カウンタ i の初期値 i_s が「 2 」にインクリメントされると、カウンタ j の新たな初期値 j_s は「 3 」となる。データ ID「 2 」とデータ ID「 3 」の組の親和度 k が入力され、次にデータ ID「 2 」とデータ ID「 4 」の組の親和度 k が入力され、最後にデータ ID「 2 」とデータ ID「 3 1 」の組の親和度 k が入力される。

20

【 0 0 4 9 】

i ループが終わると (カウンタ i が「 3 0 」に到達すると)、それは互いに異なる 2 つの言語データの組み合わせのすべてについて親和度 k の入力を終えたことを意味する (ステップ S 2 3 で YES)。複数の 3 つ組データを含む 3 つ組データテーブル・ファイル 3 1 (図 5 参照) が完成し、記憶部 1 6 に保存される (ステップ S 2 7)。

【 0 0 5 0 】

親和度の入力のとときに親和度無し (改行) が入力されると、3 つ組データの作成処理および 3 つ組データテーブル・ファイル 31 への保存処理はスキップされる (ステップ S 1 8 で「無」)。

30

【 0 0 5 1 】

親和度 k が入力されずに処理中断が入力されると (ステップ S 1 9 で「中断入力」)、3 つ組データテーブル・ファイル 3 1 には、カウンタ i の値 (データ ID)、カウンタ j の値 (データ ID) および処理が中断されたことを表す値 (- 1) が 3 つ組データとして 3 つ組データテーブル・ファイル 3 1 の最下行に追加され (ステップ S 2 5)、3 つ組データテーブル・ファイル 31 が記憶部 1 6 に保存される (ステップ S 2 6)。作成途中の 3 つ組データの作成が再開される際には、記憶部 1 6 に保存されている 3 つ組データテーブル・ファイル 3 1 の読み込みが行われ (ステップ S 1 0)、カウンタ i、j には、読み込まれた 3 つ組データテーブル・ファイル 3 1 の最下行に記憶されている 3 つ組データの 2 つのデータ ID (i、j) がそれぞれ与えられることで処理が再開される。最下行のデータ (「 - 1 」が格納されている行データ) は削除される (ステップ S 1 1 で NO、ステップ S 1 3)。

40

【 0 0 5 2 】

(2) 2 元表変換プログラム

図 7 は 2 元表変換プログラムの処理を示すフローチャートである。

【 0 0 5 3 】

2 元表変換プログラムは、上述した言語データ (図 4) および 3 つ組データテーブル・ファイル 3 1 (図 5) を用いて、次に説明する数量化 3 類適用プログラムによって処理されるべき行列データを作成する処理を実行する。

50

【 0 0 5 4 】

はじめに言語データ 2 および 3 つ組データテーブル・ファイル 3 1 が記憶部 1 6 から読み出されて R A M 1 3 に読み込まれる (ステップ S 3 0)。

【 0 0 5 5 】

R A M 1 3 に読み込まれた言語データ 2 の数 n に応じた行数および列数を持ち、すべての成分を 0 とする正方行列 P が作成される (ステップ S 3 1)。言語データ 2 の数が 3 1 個であれば、すべての成分を 0 とする 3 1 行 3 1 列の正方行列が作成される。

【 0 0 5 6 】

3 つ組データテーブル・ファイル 3 1 (図 5) は、上述したように複数の 3 つ組データを含み、各 3 つ組データには、2 つの言語データの組について、言語データを特定するデータ ID (i, j) と、親和度 k (ここでは、分かりやすくするために $k = 1$ とする) を表すデータ (3 つ組データ) が含まれている。3 つ組データテーブル・ファイル 3 1 に含まれる 3 つ組データ (i, j, k) のそれぞれについて、以下の処理が行われる (ステップ S 3 2, ステップ S 3 4 で N O)。

10

【 0 0 5 7 】

作成された行列 P において、読み込まれた 3 つ組データ (i, j, k) が用いられて、行列 P の第 i 行、第 j 列の成分が「0」から親和度「 k 」に置き換えられ、第 j 行、第 i 列の成分も「0」から親和度「 k 」に置き換えられる (ステップ S 3 3)。

【 0 0 5 8 】

たとえば、図 5 の 3 つ組データテーブル・ファイル 3 1 を参照して、最初に処理される 3 つ組データ (i, j, k) は (1, 2, 1) である。親和度 k が「1」であるので、3 1 行 3 1 列の正方行列 P において、第 1 行、第 2 列の成分が「0」から「1」に置き換えられ、かつ第 2 行、第 1 列の成分が「0」から「1」に置き換えられる。次の 3 つ組データ (i, j, k) は (1, 3, 1) であるから、3 1 行 3 1 列の正方行列 P において第 1 行、第 3 列の成分が「0」から「1」に置き換えられ、かつ第 3 行、第 1 列の成分が「0」から「1」に置き換えられる。

20

【 0 0 5 9 】

図 5 に示す 3 つ組データテーブル・ファイル 31 では、 $i = 1, j = 4$ の親和度 k は格納されていない (データが存在しない)。これは、上述した 3 つ組データ作成プログラムによる処理において、データ ID 「1」の言語データとデータ ID 「4」の言語データについて親和度は無い旨が入力されたためである。このため、正方行列 P において第 1 行、第 4 列の成分は「0」のままとなり、第 4 行、第 1 列の成分も「0」のままになる。

30

【 0 0 6 0 】

ステップ 3 3 の処理を経ることによって、行列 P には、2 つの言語データの組について、その i と j の順序 (行と列) を入れ替えた組み合わせについても、同じ値の親和度 k が格納される。たとえば、データ ID 「1」の言語データとデータ ID 「2」の言語データの親和度 $k_{1,2}$ と (行列 P における第 1 行第 2 列の要素)、データ ID 「2」の言語データとデータ ID 「1」の言語データの親和度 $k_{2,1}$ (親和度 $k_{1,2}$ と同じ値) (行列 P における第 2 行第 1 列の要素) が格納される。

40

【 0 0 6 1 】

3 つ組データテーブル・ファイル 3 1 に含まれるすべての 3 つ組データについての上述の処理が行われた後 (ステップ S 3 4 で Y E S), 行列 P の対角要素が「0」から親和度 k の最大値 (ここでは「1」) に置き換えられる (ステップ S 3 5)。行列 P の対角要素は同一言語データ同士の組に対応し、親和度 k は最大になるはずだからである。行列 P において、第 1 行第 1 列、第 2 行第 2 列、第 3 行第 3 列、...、第 3 1 行第 3 1 列の成分が「0」から「1」に置き換えられる。これによって行列 P は対角要素 (対角成分) のすべてに親和度 k の最大値が格納された 3 1 次正方対称行列となる。

【 0 0 6 2 】

最後に、行列 P の各行の行名 (行タイトル) および各列の列名 (列タイトル) に、言語データが挿入 (格納) されて 2 元表 (行列 P) が完成し、完成した行列 P が記憶部 1 6 に

50

記憶される（ステップ S 3 6）。言語データのそれぞれが関連付けられているデータ ID（1～31の数字）（図 4 参照）にしたがって、たとえば第 1 行の行名（行タイトル）にはデータ ID「1」の言語データ（「両親は心を開かない」）が挿入され、第 1 列の列名（列タイトル）にもデータ ID「1」の言語データ（「両親は心を開かない」）が挿入される。

【 0 0 6 3 】

図 8 は完成した 2 元表（行列 P）を示している。図 8 では簡略化のために、行名および列名を言語データではなく TEXT 1, TEXT 2, . . . , TEXT 3 1 と表記している。また成分「0」は「・」で表している。たとえば、「TEXT 1」はデータ ID「1」の言語データ（「両親は心を開かない」）、TEXT 2 はデータ ID「2」の言語データ（「親は支援の種類や内容を知らない」）であることを理解されたい。また、2 元表は行項目と列項目が同一であることを理解されたい。

10

【 0 0 6 4 】

（ 3 ） 数 量 化 3 類 適 用 プ ロ グ ラ ム

上述した 31 個の言語データのうちの 2 組ずつについては上述した入力される親和度によってその関係性が把握されるが、31 個の言語データのそれぞれの相関の程度は定かではない。与えられる言語データの相関を様々な観点（ディメンション：Dim）から分析するために数量化 3 類適用プログラムが用いられる。

【 0 0 6 5 】

以下、具体的に数量化 3 類適用プログラムの処理を説明する。

20

【 0 0 6 6 】

図 9 および図 10 は数量化 3 類適用プログラムの処理を表すフローチャートである。

【 0 0 6 7 】

はじめに言語データを含む 2 元表を表す行列 P（図 8）が記憶部 16 から読み出されて RAM 13 に記憶される（ステップ S 40）。

【 0 0 6 8 】

RAM 13 に記憶された行列 P に基づいて固有値が計算され、計算された固有値が大きい順に並べられる（ステップ S 41）。さらに計算された固有値を用いて以下に説明する寄与率および累積寄与率が計算される（ステップ S 42）。

【 0 0 6 9 】

図 11 は、数量化 3 類適用プログラムによって算出される複数の固有値の関係を説明するための模式図である。

30

【 0 0 7 0 】

行列 P の固有値は、以下の固有方程式を、未知数である固有値 の方程式として解くことによって求められる。ここで、E は単位行列である。

【 0 0 7 1 】

$$\det(P - E) = 0$$

【 0 0 7 2 】

det() は行列式を表す。

【 0 0 7 3 】

行列 P は 31 次の正方行列であるから、一般には行列 P から 31 個の固有値が算出される。固有値のそれぞれに対応する固有ベクトルは、それぞれの固有値を以下の連立方程式に代入することによって算出することができる。

40

【 0 0 7 4 】

$$(P - E) x = 0$$

【 0 0 7 5 】

行列の固有値および固有ベクトルの組は、たとえばオープンソースフリーウェアである統計解析ソフト「R」（<https://cran.r-project.org/>）を用いて算出することができる。

【 0 0 7 6 】

50

図 8 に示す 3 1 次の行列 P を用いると，最大固有値 = 1 を含む 3 1 個の固有値が求められる。しかしながら 3 1 個の固有値のうち「 1 」となる固有値は以下の処理では除外される。これは「 1 」となる固有値を用いると，各言語データのそれぞれについて算出されるスコアベクトル（スコアベクトルについては後述する）が同値となり，言語データ間の相関を表すスコアベクトルとしての意味がないからである。

【 0 0 7 7 】

算出される複数の固有値は，その値が大きいほど，二元表の行と列の項目同士の相関関係をより強く反映していることを表している。逆に固有値が小さいほど二元表の行と列の項目同士の相関関係が弱いことを表しており，複数の言語データの相互の関係は分かりづらなものとなる。図 1 1 は，2 つの固有値（第 1 固有値 > 第 2 固有値の関係を持つとする）のそれぞれを用いて言語データのそれぞれに対して算出されるスコアベクトル（Dim 1 と Dim 2）を模式的にプロットしたものである。図 1 1 において横軸と縦軸は，それぞれ第 1 固有値および第 2 固有値に属するスコアベクトルに対応している。プロットの番号は言語データ（テキストデータ）の番号に対応している。このように，値の大きい固有値に対するスコアベクトルを用いて，2 次元平面に各言語データをプロットすると，それらが相互にどのような関係にあるかを把握しやすくなる。これを 3 次元まで拡張して言語データ同士の関係を検討することが可能である。しかし，4 次元以上では可視化することはできないため，後述するように工夫が必要になる。

10

【 0 0 7 8 】

図 1 2 は，行列 P（図 8）から実際に算出される 3 0 個の固有値（No. 1 ~ No. 3 0）（「 1 」を除く固有値）を，寄与率および累積寄与率とともに示している。「寄与率」は，3 0 個の固有値のそれぞれについて，当該固有値をすべての固有値の総和によって除算した値である。「累積寄与率」は，複数の固有値のそれぞれに対応して算出される寄与率を，値の大きな固有値のものから順に累積加算した値である。なお，固有値，寄与率および累積寄与率は，実際は，図 1 2 に示す小数点以下 7 桁以降の数値まで算出され，図 1 2 では小数点以下 8 桁の数値が四捨五入されている。

20

【 0 0 7 9 】

上述したように，固有値の値が大きければ大きいほど，スコアベクトルは言語データ同士の相関関係をより強く表しているのので，言語データ同士の関係を検討するのに，すべてのスコアベクトルを用いる必要は必ずしもない。このため，問題により少なくとも累積寄与率で 70 から 80 % 程度が確保できるスコアベクトルを利用して，言語データ同士の関係を検討し，全体像を把握することを試みるのが現実的である（ステップ S 4 3）。

30

【 0 0 8 0 】

図 1 3 は，行列 P（図 8）から算出される固有値（図 1 2）のうち値の大きい順に 7 つの固有値（図 1 2 の No. 1 ~ No. 7）のそれぞれを用いて，3 1 個の言語データ（図 4）のそれぞれについて算出したスコアベクトルを示している。図 1 3 において，Dim 1，Dim 2，... Dim 7 は，行列 P に潜在する変数（軸）の個数を表しており，行列 P から算出される 3 1 個の固有値のうち，「 1 」となる固有値を除いて大きい方から 7 つの固有値（その固有ベクトル）のそれぞれに対応する。

【 0 0 8 1 】

固有値を λ^2 とすると，3 1 個の言語データのそれぞれのスコアベクトル x は，以下の式（スコアベクトル算出式と呼ぶ）を解くことによって求められる。

40

【 0 0 8 2 】

$$\left(p_x^{-1/2} P^T p_x^{-1} P p_x^{-1/2} - \lambda^2 I \right) p_x^{1/2} x = 0$$

【 0 0 8 3 】

上述したスコアベクトル算出式において，P は処理対象の正方行列を， P^T は正方行列 P の転置行列を，I は単位行列を，それぞれ表す。 p_x は正方行列 P の行和ベクトルの要素を対角要素とする行列である。 x がスコアベクトルである。なお，正方行列 P は対称行列であるから $P = P^T$ となる。スコアベクトル x は言語データごとかつ $n - 1$ 個の固有値

50

ごとに算出することができる。

【0084】

上述のスコアベクトル算出式は、2個の互いに異なる項目（行項目と列項目が互いに異なるもの）を取り扱う一般的な数量化3類から得られる固有値問題を基礎にして、これに行項目と列項目とを同じとする（区別しない）ことを条件としたとき表される式であり、2元表の相関係数を最大にするようなスコアベクトルを求める式である。

【0085】

固有値ごとに31個の言語データのそれぞれについて算出されるスコアベクトルxについても、上述した統計解析ソフト「R」を用いて算出することができる。このソフトでは、数量化3類は対応分析と呼ばれている。

10

【0086】

以下の説明では、算出される複数の固有値（ $\lambda^2 = 1$ を除く）を、値の大きいものから順番に、第1固有値、第2固有値、第3固有値、...と呼ぶ。

【0087】

図14は、第1固有値を用いてスコアベクトル算出式に基づいて算出された各言語データのスコアベクトル（以下、「Dim1スコア」という）を横軸とし、第2固有値を用いてスコアベクトル算出式に基づいて算出された各言語データのスコアベクトル（以下、「Dim2スコア」という。以下に説明するDim3スコアも同様。）を縦軸としたグラフであり、31個の言語データのそれぞれについて第1固有値および第2固有値を用いて算出されたスコアベクトルにしたがって、各言語データをプロットした散布図を示している。第1固有値および第2固有値に対応する2つの軸を用いることで図14に示す2次元（平面）グラフにおいて各言語データ同士の相関を把握することができる。数量化3類適用プログラムは図14に示すような2次元の散布図を表示部15に表示することができる（ステップS44）。

20

【0088】

図14を参照して、Dim1スコア（横軸）に着目すると、31個の言語データのうちのたとえば「25.奨学金制度の充実が必要」に大きなスコアが算出され、「5.働く場所がない」に小さいスコアが算出されており、一例として、これらの2つの言語データのDim1スコア軸上の距離が大きいことを一見して把握することができる。

【0089】

他方、Dim2スコア（縦軸）に着目すると、31個の言語データのうちのたとえば「25.奨学金制度の充実が必要」に大きなスコアが算出され、「29.食べ物を応援できるのはいい」に小さいスコアが算出されており、一例として、これらの2つの言語データのDim2スコア軸上の距離が大きいことを一見して把握することができる。

30

【0090】

図14に示す散布図をみると、Dim1スコアおよびDim2スコアが互いに近い複数の言語データ（プロット）が集中している部分がある。この集中している言語データが、子供の貧困を無くすという社会的課題の解決につながる重要情報であると考察することができる可能性がある。別のDimスコアを用いると、子供の貧困問題について、たとえば原因事象と結果事象とにプロットが散布図上で分かれる可能性もあり、そのような散布図が得られた場合には、たとえば原因事象のみをピックアップして再度上述した処理を繰り返すことも考えられる。

40

【0091】

なお、図14に示す散布図は、たとえば横軸をDim1スコア、縦軸をDim3スコアとして表示することもできることは言うまでもない。

【0092】

図15Aおよび図15Bは、31個の言語データのそれぞれについて第1固有値、第2固有値、第3固有値のもとで算出されたスコアベクトルを、互いに直交する軸（Dim1スコア、Dim2スコア、Dim3スコア）を持つ3次元空間にプロットした散布図を示している。図15A、図15Bに示すように各言語データの相関関係は3次元空間で把握

50

することもできる。数量化3類適用プログラムは図15A, 図15Bに示すような3次元の散布図を表示部15に表示することもできる(ステップS44)。

【0093】

図16A, 図16B, 図16Cおよび図16Dは, 図15A, 図15Bに示す3次元の散布図を, 視点を変えて見るために, 回転表示させている様子を示している。図16A~図16Dに示す回転表示も数量化3類適用プログラムによって実行され, 31個の言語データの相関関係を分かりやすく表示部15に表示させることができる(ステップS44)。

【0094】

図17は, 31個の言語データのそれぞれについてのDim1スコアからDim7スコアまでのスコアを用いて, 7次元空間における各言語データ間のユークリッド距離を計算し, クラスタ分析のワード法を用いて7つのクラスタにまとめた結果を樹形図として表示している。より具体的には, 図14に示す多数のプロットを含む2次元平面において最もユークリッド距離の近いプロットを持つ言語データの組み合わせを求め, その組み合わせをペアにし, ペアにされた組み合わせについてその重心を新たなプロットとし, 新たなプロットを含めて上記処理を繰り返すことによって, 最終的に図17に示すような樹形図が作成される。図17において左右方向にのびる直線の長さがユークリッド距離に対応する長さによって表されている。相関の程度に応じて言語データ同士を, まとまりをもつように表現することができる(クラスタリング, グループング)。数量化3類適用プログラムは図17に示すような樹形図を表示部15に表示することもできる(ステップS45)。

【0095】

図18は, 31個の言語データのそれぞれについて算出されたDim1スコアの大きい順番に言語データを並べ, Dim1スコア, Dim2スコアおよびDim3スコアを, 言語データのそれぞれに対応させて表示し, さらにDim1スコア, Dim2スコアおよびDim3スコアに, スコアの値の大きさに応じた色を付して表示するヒートマップを示している。数量化3類適用プログラムは図18に示すようなヒートマップを表示部15に表示することもできる(ステップS46)。Dim1, Dim2, Dim3等の行列Pに潜在する変数の間の類似性(行列Pに潜在する事象(事柄))を検討するのに有用である。

【0096】

最後に, 上述したスコアベクトル算出式の導出の基礎とした, 2種類の互いに異なる項目(行項目と列項目が互いに異なるもの)を取り扱う一般的な数量化3類から得られる固有値問題を説明しておく。

【0097】

数量化3類(Quantification Method of Type 3(QM3))は, 以下に示すテーブル1のような2元表の形式で与えられたデータの内部構造を分析する手法である。ここでXおよびYは互いに相違する項目であることを意味する。 X_i はi番目のX項目を, Y_j はj番目のY項目を表す。そして, 各項目にスコア(数量)が対応づけられており, これを x_i, y_j としている。 p_{ij} は項目 X_i と項目 Y_j の関連度を表し, p_i および p_j はそれぞれi番目の行和とj番目の列和を表している。関連度 p_{ij} は項目 X_i と Y_j との間に関連があれば「1」, なければ「0」とされる。関連の度合を強中弱に対応させて3, 2, 1のようにしてもよい。ここでは, 簡単のために1と0とする。テーブル1では, 2元表の要素の合計で要素の値を除いて相対度数の形に変換して合計が1となるようにしている。

【0098】

テーブル1 2元表

10

20

30

40

items		Y_1	Y_2	...	Y_j	...	Y_n	
	scores	y_1	y_2	...	y_j	...	y_n	
X_1	x_1	p_{11}	p_{12}	...	p_{1j}	...	p_{1n}	$p_{1\cdot}$
X_2	x_2	p_{21}	p_{22}	...	p_{2j}	...	p_{2n}	$p_{2\cdot}$
...
X_i	x_i	p_{i1}	p_{i2}	...	p_{ij}	...	p_{in}	$p_{i\cdot}$
...
X_m	x_m	p_{m1}	p_{m2}	...	p_{mj}	...	p_{mn}	$p_{m\cdot}$
		$p_{\cdot 1}$	$p_{\cdot 2}$...	$p_{\cdot j}$...	$p_{\cdot n}$	1

【 0 0 9 9 】

上記 2 元表は、量とは関係ない 2 種類の項目 X_i および Y_j から構成されているが、数量化 3 類は項目 X_i および Y_j に数量 x_i および y_j を対応させること（数量化）により内部構造を分析するものである。

【 0 1 0 0 】

一般性を失うことなく、 x_i と y_j のそれぞれの平均を次のように 0 とおくことができる。

$$\sum_i x_i p_{i\cdot} = 0 \quad (1.1)$$

$$\sum_j y_j p_{\cdot j} = 0 \quad (1.2)$$

【 0 1 0 2 】

また、 x_i と y_j のそれぞれの分散も 1 とおくことができる。

$$\sum_i x_i^2 p_{i\cdot} = 1 \quad (1.3)$$

$$\sum_j y_j^2 p_{\cdot j} = 1 \quad (1.4)$$

【 0 1 0 4 】

相関係数 は次式のように表される。

$$\rho = \sum_i \sum_j x_i y_j p_{ij} \quad (1.5)$$

【 0 1 0 5 】

数量化 3 類では、式(1.3)と(1.4)を制約条件として、相関係数 が最大となるように数量 x_i および y_j を決める。そこで、2 つの未定乗数 λ_1 と λ_2 を用いて Lagrange の未定乗数法を用いる。次式の Q を最大にすればよい。

【 0 1 0 6 】

$$Q = \sum_i \sum_j x_i y_j p_{ij} - \frac{\lambda_1}{2} \left(\sum_i x_i^2 p_{i\cdot} - 1 \right) - \frac{\lambda_2}{2} \left(\sum_j y_j^2 p_{\cdot j} - 1 \right) \quad (1.6)$$

【 0 1 0 7 】

数量 x_i および y_j について偏微分して 0 とおくと以下の式が得られる。

【 0 1 0 8 】

$$\frac{\partial Q}{\partial x_i} = \sum_j y_j p_{ij} - \lambda_1 x_i p_{i\cdot} = 0 \quad (1.7)$$

$$\frac{\partial Q}{\partial y_j} = \sum_i x_i p_{ij} - \lambda_2 y_j p_{.j} = 0 \quad (1.8)$$

【 0 1 0 9 】

式 (1.7) に x_i を乗じ、添字 i について総和をとると次式が得られる。

【 0 1 1 0 】

$$\sum_i \sum_j x_i y_j p_{ij} - \lambda_1 \sum_i x_i^2 p_{.i} = 0 \quad (1.9)$$

【 0 1 1 1 】

同様に、式 (1.8) に y_j を乗じて添字 j について総和をとると次式が得られる。

【 0 1 1 2 】

$$\sum_i \sum_j x_i y_j p_{ij} - \lambda_2 \sum_j y_j^2 p_{.j} = 0 \quad (1.10)$$

【 0 1 1 3 】

式 (1.3) と (1.4)、および (1.9) と (1.10) から、 $\lambda_1 = \lambda_2$ となるので、 λ_1 および λ_2 を λ に改める。そうすると、式 (1.5) から $x_i = \frac{\sum_j y_j p_{ij}}{\lambda p_{.i}}$ となる。

【 0 1 1 4 】

式 (1.7) は次のように変形することができ、この変形式から数量 (スコア) x_i を求めることができる。

【 0 1 1 5 】

$$x_i = \frac{\sum_j y_j p_{ij}}{\lambda p_{.i}} \quad (1.11)$$

【 0 1 1 6 】

式 (1.11) を式 (1.8) に代入すると次式が得られる。

【 0 1 1 7 】

$$\sum_i \frac{\sum_k y_k p_{ik}}{p_{.i}} p_{ij} - \lambda^2 y_j p_{.j} = 0 \quad (1.12)$$

【 0 1 1 8 】

ここで、以下のような表記法を導入する。

【 0 1 1 9 】

$$\mathbf{x} = (x_1 \quad x_2 \quad \cdots \quad x_m)^T \quad (1.13)$$

$$\mathbf{y} = (y_1 \quad y_2 \quad \cdots \quad y_n)^T \quad (1.14)$$

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mn} \end{bmatrix} \quad (1.15)$$

$$\mathbf{P}_x = \begin{bmatrix} p_{1.} & 0 & \cdots & 0 \\ 0 & p_{2.} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{m.} \end{bmatrix}, \quad \mathbf{P}_x^k = \begin{bmatrix} p_{1.}^k & 0 & \cdots & 0 \\ 0 & p_{2.}^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{m.}^k \end{bmatrix} \quad (1.16)$$

10

20

30

$$P_y = \begin{bmatrix} p_{.1} & 0 & \cdots & 0 \\ 0 & p_{.2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{.n} \end{bmatrix}, \quad P_y^k = \begin{bmatrix} p_{.1}^k & 0 & \cdots & 0 \\ 0 & p_{.2}^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{.n}^k \end{bmatrix} \quad (1.17)$$

【 0 1 2 0 】

これらの表記法を用いると式 (1.12) は次のように表すことができる。

【 0 1 2 1 】

$$P^T P_x^{-1} P_y - \lambda^2 P_y y = 0 \quad (1.18)$$

【 0 1 2 2 】

式 (1.18) は次式のように整理することができ、これは λ^2 を固有値とする固有値問題を表している。

【 0 1 2 3 】

$$\left(P_y^{-1/2} P^T P_x^{-1} P P_y^{-1/2} - \lambda^2 I \right) P_y^{1/2} y = 0 \quad (1.19)$$

したがって、

【 0 1 2 4 】

$$v = P_y^{1/2} y \quad (1.20)$$

とおくと、以下の式となる。

【 0 1 2 5 】

$$\left(P_y^{-1/2} P^T P_x^{-1} P P_y^{-1/2} - \lambda^2 I \right) v = 0 \quad (1.21)$$

【 0 1 2 6 】

上述したスコアベクトル算出式は、上記式 (1.21) について x (行項目) と y (列項目) の区別を無くす (すなわち、項目 X と項目 Y を同じにする) ことによって導出することができる。

【符号の説明】

【 0 1 2 7 】

1 数量化親和図法実行装置

1 1 CPU

1 2 ROM

1 3 RAM

1 4 入力部

1 5 表示部

1 6 記憶部

1 6 - 1 言語データ

1 6 - 2 数量化親和図法実行プログラム

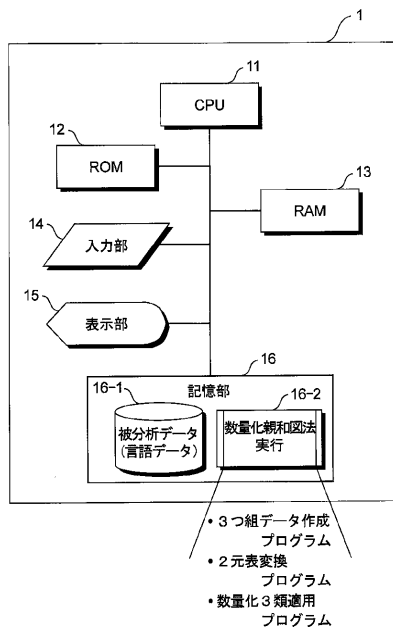
【要約】

【課題】親和図法の結果が実施者によって異なるようにする。

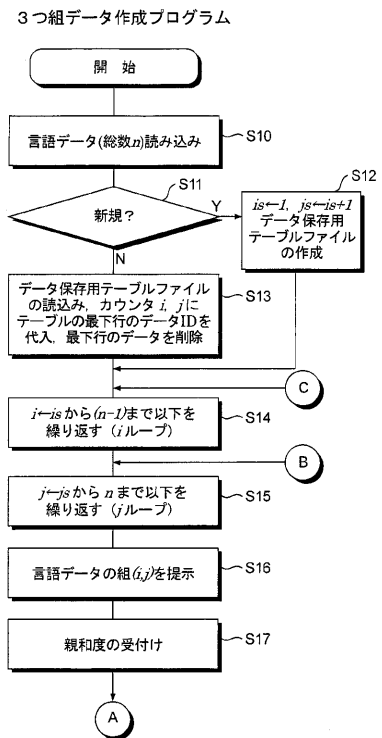
【解決手段】数量化親和図法実行装置は、 n 個の言語データ i, j を用いて、互いに異なる 2 つずつの言語データ i, j のすべての組み合わせについて数値によって入力される親和度 k を用いて、親和度 k を n 行 n 列の n 次正方対称行列の対応する成分 i, j の要素とし、対角要素を上記親和度の最大値とし、行項目および列項目を同じとする n 次正方対称行列を表すデータを処理対象とする。上記 n 次正方対称行列から固有値を求め (ステップ S 4 1)、固有値のそれぞれを用いてスコアベクトル算出式に基づいて、 n 個の言語データのそれぞれのスコアベクトル x が算出される (ステップ S 4 3)。

【選択図】図9

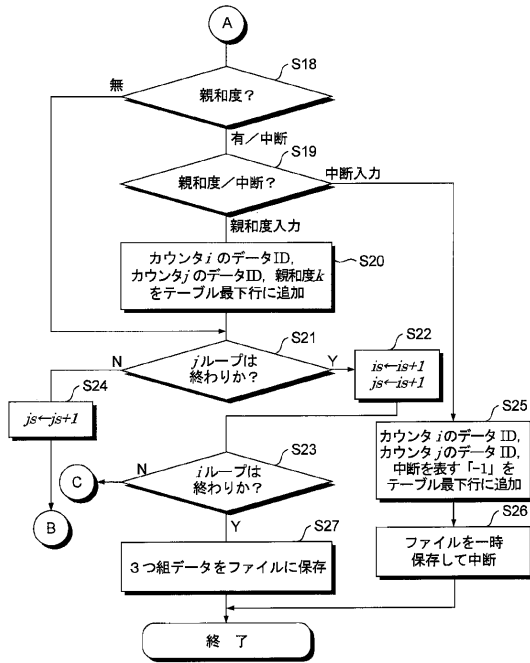
【図1】



【図2】



【 図 3 】



【 図 4 】

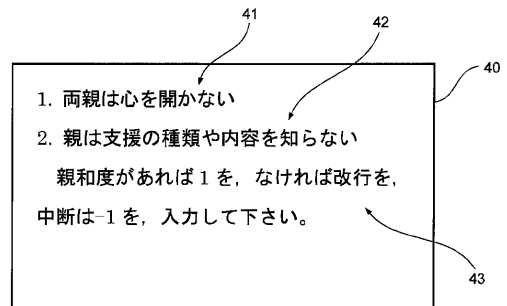
1. 両親は心を開かない
2. 親は支援の種類や内容を知らない
3. 両親は恥と支援を受けることに抵抗する
4. 親が適切な教育を受けずに育った影響がある
5. 働く場所がない
6. 仕事を探するとき、労働条件は不利です
7. 成人社会を知らずに親になる
8. 働いて収入を得て子どもを育てることができない
9. 日本語が苦手なため、正確な情報が得られない
10. 学校に行かずに兄弟たちの世話をする
11. 外国人の子どもは学習についていけない
12. 学外での支援が必要
13. 政府主導のトップダウン会議への参加に積極的
14. 自発的に仲間の集まりを作るように彼らを励ます
15. 政府は自発的に行われた集会を支援する
16. 政府は適切に対応しなければならない
17. 善意に頼らないで
18. 子ども委員会は、行政からの依頼がなければ動きにくい
19. 食事など学習支援を充実させる
20. 支援体制は整備されていない
21. 支援体制を整えるには外部からの機会が必要
22. 支援対象者の理解が必要
23. 経済的支援が最初に提供されるべきである
24. 制服などの物質の支援を検討する価値がある
25. 奨学金制度の充実が必要
26. 貧困家庭ではない子どもたちと一緒に学習支援が必要
27. 放課後学習支援を充実させる
28. 資格と仕事を得るといふ考えは両方とも必要である
29. 食べ物を応援できるのはいい
30. 商品券はレトルト食品より優れている
31. 医大と企業の連携が学習支援に優れている

【 図 5 】

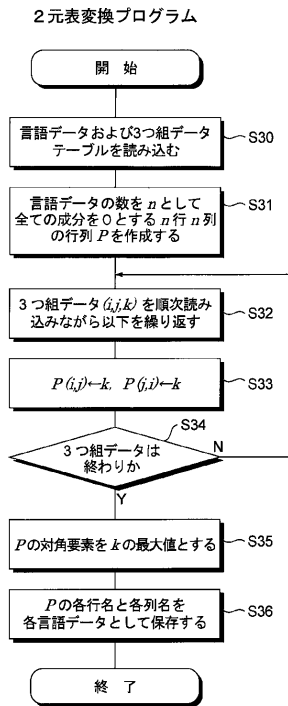
3つ組データテーブル・ファイル 31

言語データ i (データ ID)	言語データ j (データ ID)	親和度 k
1	2	1
1	3	1
1	13	1
⋮	⋮	⋮
2	3	1
2	7	1
2	9	1
2	20	1
⋮	⋮	⋮
23	25	1
24	30	1
26	27	1
26	31	1

【 図 6 】



【 図 7 】

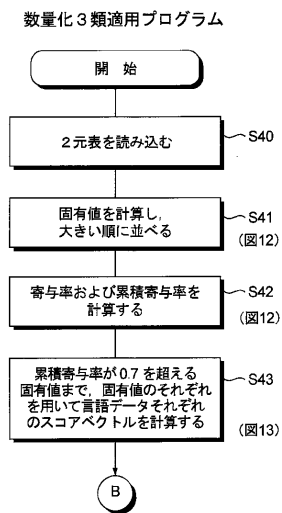


【 図 8 】

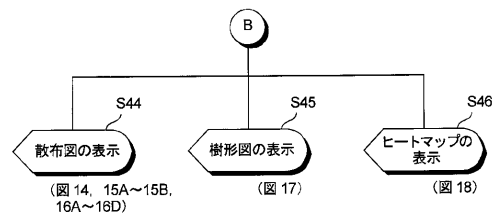
行列 P

	T E X T 1	T E X T 2	T E X T 3	T E X T 4	T E X T 5	T E X T 6	T E X T 7	T E X T 8	T E X T 9	T E X T 10	T E X T 11	T E X T 12	T E X T 13	T E X T 14	T E X T 15	T E X T 16	T E X T 17	T E X T 18	T E X T 19	T E X T 20	T E X T 21	T E X T 22	T E X T 23	T E X T 24	T E X T 25	T E X T 26	T E X T 27	T E X T 28	T E X T 29	T E X T 30	T E X T 31	
TEXT1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TEXT31	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

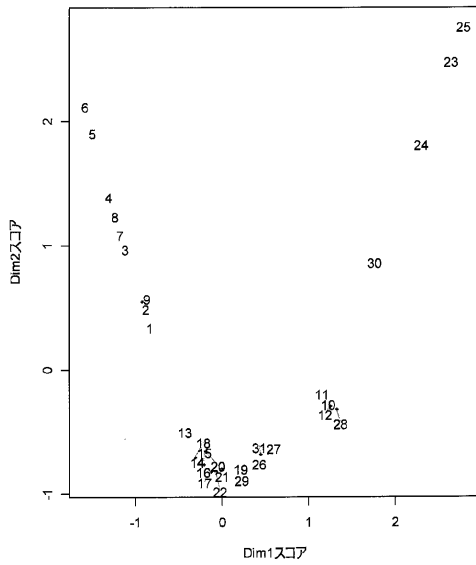
【 図 9 】



【 図 1 0 】



【 図 1 1 】



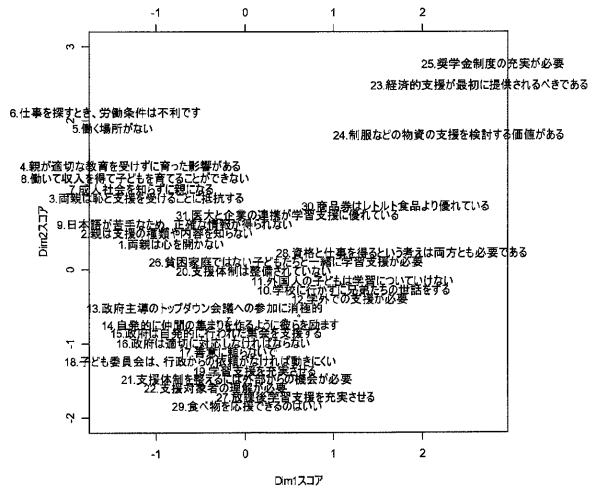
【 図 1 2 】

No.	固有値	寄与率	累積寄与率
1	0.948	0.142	0.142
2	0.900	0.135	0.277
3	0.827	0.124	0.401
4	0.696	0.104	0.506
5	0.625	0.094	0.599
6	0.441	0.066	0.666
7	0.419	0.063	0.728
8	0.368	0.055	0.784
9	0.304	0.046	0.829
10	0.219	0.033	0.862
11	0.154	0.023	0.885
12	0.131	0.020	0.905
13	0.115	0.017	0.922
14	0.095	0.014	0.937
15	0.087	0.013	0.950
16	0.059	0.009	0.959
17	0.051	0.008	0.966
18	0.044	0.007	0.973
19	0.042	0.006	0.979
20	0.031	0.005	0.984
21	0.027	0.004	0.988
22	0.024	0.004	0.991
23	0.017	0.003	0.994
24	0.015	0.002	0.996
25	0.010	0.001	0.998
26	0.008	0.001	0.999
27	0.005	0.001	0.999
28	0.002	0.000	1.000
29	0.001	0.000	1.000
30	0.000	0.000	1.000

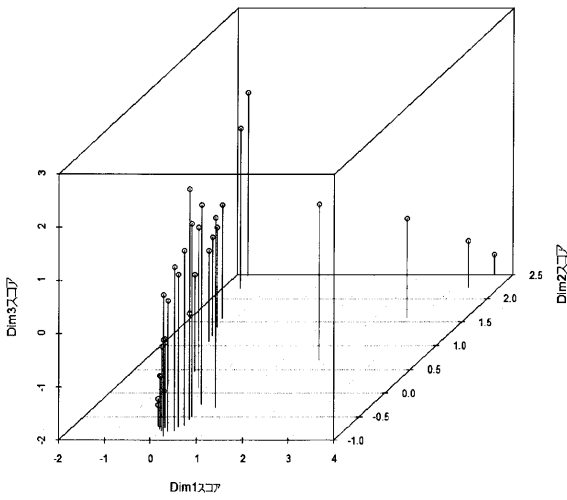
【 図 1 3 】

	Dim1	Dim2	Dim3	Dim4	Dim5	Dim6	Dim7
TEXT1	-0.84	0.35	-0.27	1.25	-0.22	0.47	-0.56
TEXT2	-0.88	0.50	-0.10	1.33	0.45	-0.18	-1.78
TEXT3	-1.12	0.98	0.18	1.14	0.51	0.05	0.71
TEXT4	-1.30	1.39	0.53	-0.02	0.32	-0.03	1.40
TEXT5	-1.50	1.91	1.05	-3.20	-0.91	0.01	-0.57
TEXT6	-1.59	2.12	1.28	-4.79	-1.57	0.04	-1.94
TEXT7	-1.17	1.10	0.28	1.06	0.66	-0.09	1.02
TEXT8	-1.24	1.24	0.38	0.93	0.69	-0.04	1.97
TEXT9	-0.92	0.55	-0.13	1.99	0.77	-0.54	-6.05
TEXT10	1.22	-0.27	2.32	0.59	-0.85	0.19	-0.02
TEXT11	1.09	-0.21	1.39	0.08	0.23	1.15	-0.07
TEXT12	1.25	-0.29	2.62	0.80	-1.40	-0.96	0.05
TEXT13	-0.42	-0.49	-0.78	0.45	-1.43	0.91	0.19
TEXT14	-0.28	-0.73	-0.87	0.08	-1.46	0.62	0.37
TEXT15	-0.21	-0.76	-0.76	-0.09	-0.91	0.12	0.19
TEXT16	-0.18	-0.80	-0.75	-0.17	-0.82	-0.01	0.21
TEXT17	-0.14	-0.81	-0.68	-0.28	-0.44	-0.39	0.08
TEXT18	-0.31	-0.70	-0.89	0.18	-1.64	0.91	0.42
TEXT19	0.21	-0.79	-0.02	-0.65	1.38	-0.51	0.04
TEXT20	-0.19	-0.65	-0.53	-0.11	-0.04	-0.42	-0.28
TEXT21	-0.01	-0.80	-0.44	-0.38	0.09	-0.29	0.09
TEXT22	-0.08	-0.82	-0.56	-0.39	0.03	-0.57	0.02
TEXT23	2.64	2.49	-1.94	-0.06	-0.14	-1.19	0.10
TEXT24	2.29	1.82	-0.99	0.00	0.05	2.44	-0.24
TEXT25	2.79	2.77	-2.37	-0.08	-0.24	-3.63	0.33
TEXT26	0.43	-0.75	0.38	-0.75	1.99	0.05	0.09
TEXT27	0.60	-0.63	0.66	-0.56	1.66	0.41	0.04
TEXT28	1.32	-0.32	3.20	1.19	-2.41	-2.93	0.16
TEXT29	0.22	-0.88	-0.03	-0.98	2.38	-1.55	0.13
TEXT30	1.76	0.87	0.24	0.06	0.20	3.62	-0.32
TEXT31	0.45	-0.68	0.37	-0.54	1.25	0.17	0.07

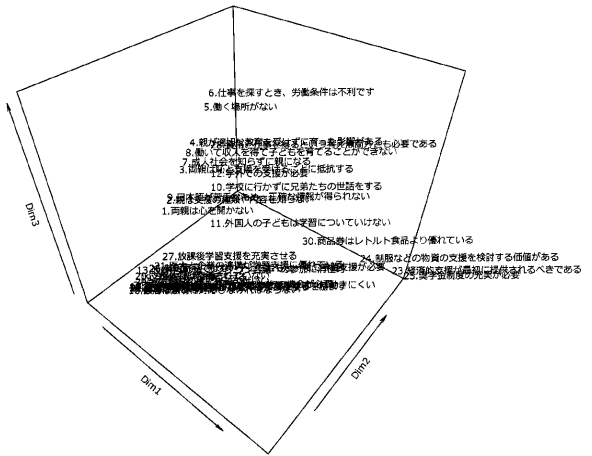
【 図 1 4 】



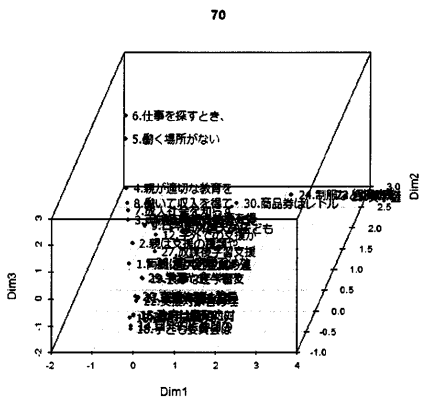
【図15A】



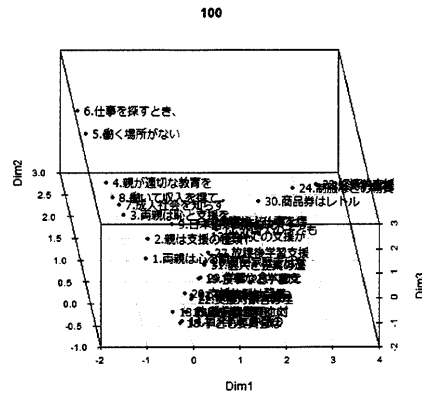
【図15B】



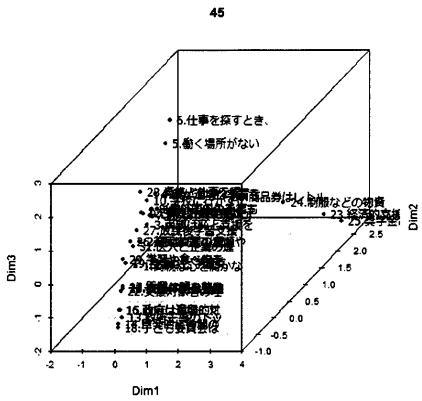
【図16A】



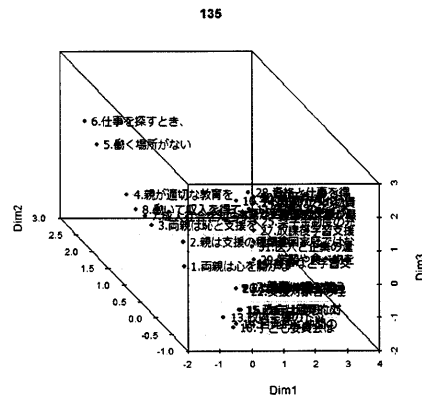
【図16C】



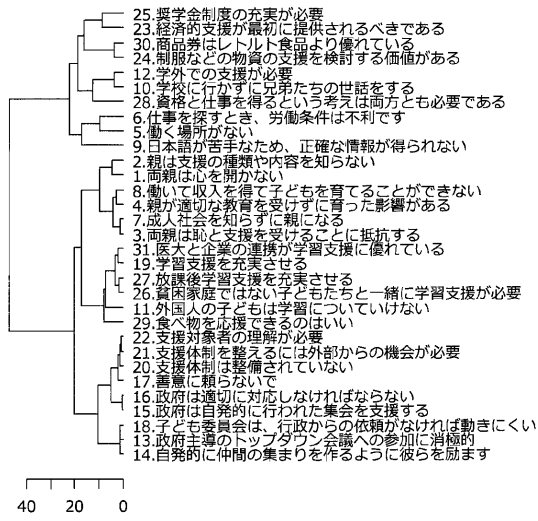
【図16B】



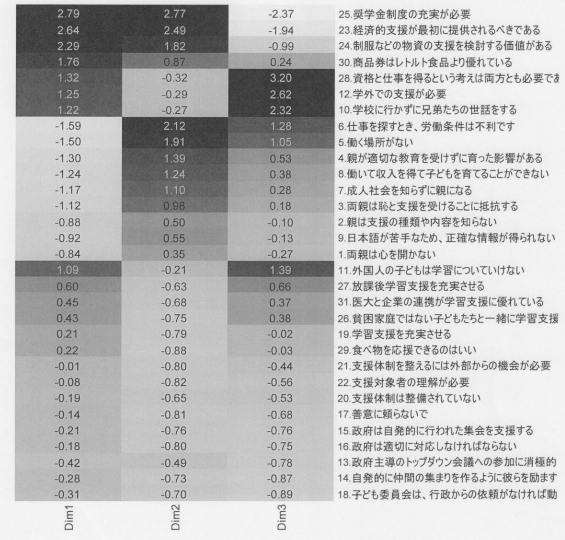
【図16D】



【 図 1 7 】



【 図 1 8 】



フロントページの続き

(56)参考文献 特開 2 0 0 1 - 1 5 5 0 6 3 (J P , A)

Koike, T. and Shindo, H. , Mining the latent information by the Affinity Diagram Method quantificated by QM3 (Correspondence Analysis) [online] , 2020年 , pp. 1 39 , [検索日 20 23.06.07], インターネット : <URL:https://web.archive.org/web/20200923102444/https://www.yafo.or.jp/wp/wp content/uploads/2020/03/01c9c442a37b8842fb94a64c23a0e878.pdf >
長畑 秀和 , R で学ぶデータサイエンス 初版 , 第1版 , 株式会社朝倉書店 , 2018年03月25日 , pp. 26 29 , ISBN 978 4 254 12227 5

(58)調査した分野(Int.Cl. , D B 名)

G 0 6 F 1 8 / 0 0 - 1 8 / 4 0

G 0 6 F 1 7 / 1 0

G 0 6 N 5 / 0 0 - 5 / 0 2 5