

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2022年8月25日(25.08.2022)



(10) 国際公開番号

WO 2022/176014 A1

(51) 国際特許分類:
G06Q 10/00 (2012.01)

(21) 国際出願番号: PCT/JP2021/005698

(22) 国際出願日: 2021年2月16日(16.02.2021)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(71) 出願人: 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 Tokyo (JP).

(72) 発明者: 山本 太三 (YAMAMOTO Taizo); 〒1808585 東京都武蔵野市緑町3丁目9-1 NTT 知的財産センタ内 Tokyo (JP). 森

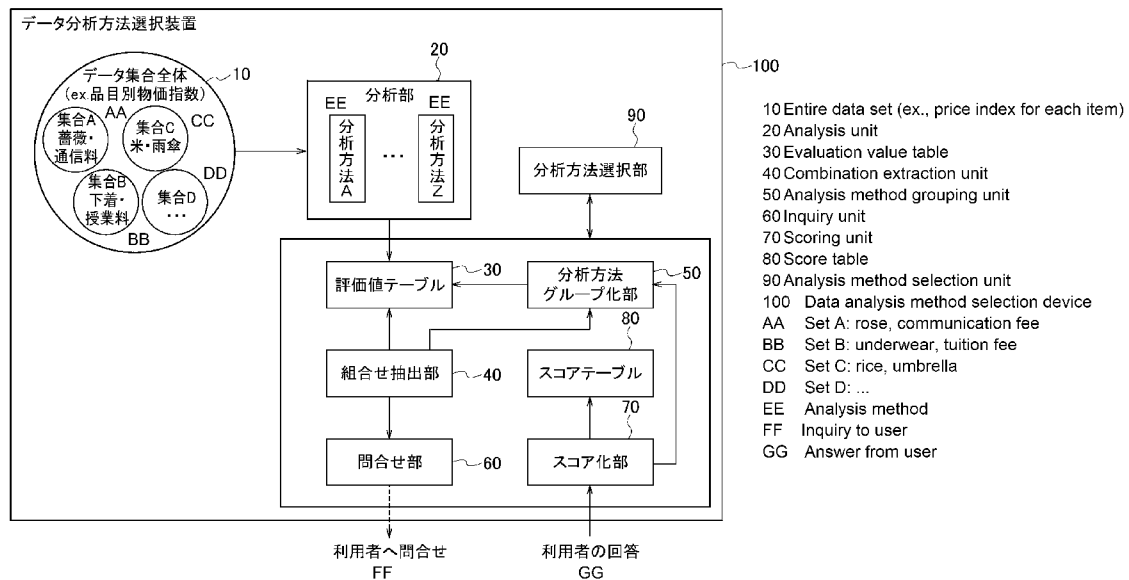
谷 高明(MORIYA Takaaki); 〒1808585 東京都武蔵野市緑町3丁目9-1 NTT 知的財産センタ内 Tokyo (JP). 西尾 学(NISHIO Manabu); 〒1808585 東京都武蔵野市緑町3丁目9-1 NTT 知的財産センタ内 Tokyo (JP). 三好 優(MIYOSHI Yu); 〒1808585 東京都武蔵野市緑町3丁目9-1 NTT 知的財産センタ内 Tokyo (JP).

(74) 代理人: 三好 秀和, 外(MIYOSHI Hidekazu et al.); 〒1050001 東京都港区虎ノ門1丁目2番8号 虎ノ門琴平タワー Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ,

(54) Title: DATA ANALYSIS METHOD SELECTION DEVICE, METHOD, AND PROGRAM

(54) 発明の名称: データ分析方法選択装置、方法及びプログラム



(57) Abstract: This data analysis method selection device is provided with: an analysis unit 10 that obtains evaluation values representing the relationship between two time series data for each set A, B, ..., using different analysis methods; a combination extraction unit 40 that extracts, for each analysis method, combinations of sets that differ from each other in the pattern of change of the evaluation value; an analysis method grouping unit 50 that, for each combination extracted by the combination extraction unit 40, classifies analysis methods into groups on the basis of whether the evaluation value is good or bad, and associates and records the classification results with the sets; an inquiry unit 60 that presents a user

EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告 (条約第21条(3))

with the time-series data of each set in a combination extracted by the combination extraction unit 40, and inquires of the user as to which set comprises similar time-series data; a scoring unit 70 that adds to the score of each analysis method belonging to the group for which the evaluation value of the set determined, in the user's answer, to be similar is better; and an analysis method selection unit 90 that repeats the processing for each functional component unit described above, and selects an analysis method, the score of which reaches a prescribed value.

(57) 要約: 2つの時系列データの関係性を表す評価値を集合A, B, ...ごとに異なる分析方法で求める分析部10と、分析方法に対応させて評価値の変化の傾向が異なる集合の組合せを抽出する組合せ抽出部40と、組合せ抽出部40で抽出された組合せのそれぞれについて、評価値の良否で分析方法をグループに分類し、該分類した結果を集合に対応させて記録する分析方法グループ化部50と、組合せ抽出部40が抽出した組合せのそれぞれの集合の時系列データを利用者に提示し、どちらの集合の時系列データが似ているかを利用者に問合せる問合せ部60と、利用者の回答で似ていると判定された集合の評価値が良い方のグループに属する分析方法のスコアを加点するスコア化部70と、上記の各機能構成部のそれぞれの処理を繰り返し、スコアが所定値になる分析方法を選択する分析方法選択部90とを備える。

明 細 書

発明の名称：データ分析方法選択装置、方法及びプログラム

技術分野

[0001] 本発明は、データ分析方法選択装置、方法及びプログラムに関する。

背景技術

[0002] データの集合を分析・評価するのに、データサイエンティスト（Data Scientist、以降DS）のサポートを受ける場合がある。DSは、さまざまな意向決定の局面において、データに基づいて合理的な判断を行えるように意向決定者をサポートする職務を行う。

[0003] DSは、各分野に精通した専門家ではあるが、ノウハウを持たない分野もある。よって、DSにノウハウのない場合は、適切なデータ分析ができない。

[0004] 一方、データ分析装置としては、時系列データ等のデータ集合に規則性を求め、各データの時間的変化量を示す指標値を算出し、時系列データをグラフ化するものが例えば特許文献1に開示されている。

先行技術文献

特許文献

[0005] 特許文献1：特許第6592411号公報

発明の概要

発明が解決しようとする課題

[0006] しかしながら、特許文献1に開示された技術は、求めた上記の指標値に従った順番で複数の時系列データをグラフ化したものを並べて表示するものである。よって、表示されたグラフが利用者の求めるもので無い場合もある。つまり、分析結果に利用者のフィードバックが効かないという課題がある。

[0007] このように従来は、完全な分析方法は存在しないという前提に基づいて複数の分析方法の結果を利用者に提示し、利用者がより良い分析方法を選択するという仕組みが存在しなかった。

[0008] 本発明は、この課題を鑑みてなされたものであり、ノウハウが無いような場合においても適切な分析手法を利用者のフィードバックを効かせて絞り込み、適切なデータ分析が選択できるデータ分析方法選択装置、方法及びプログラムを提供することを目的とする。

課題を解決するための手段

[0009] 本発明の一態様に係るデータ分析方法選択装置は、2つの時系列データをそれぞれ記録した集合を複数含むデータ集合と、前記2つの時系列データの関係性を表す評価値を前記集合ごとに異なる分析方法で求める分析部と、前記分析方法に対応させて前記評価値の変化の傾向が異なる前記集合の組合せを抽出する組合せ抽出部と、前記組合せ抽出部で抽出された前記組合せのそれぞれについて、前記評価値の良否で前記分析方法をグループに分類し、該分類した結果を前記集合に対応させて記録する分析方法グループ化部と、前記組合せ抽出部が抽出した前記組合せのそれぞれの前記集合の時系列データを利用者に提示し、どちらの前記集合の時系列データが似ているかを利用者に問合せる問合せ部と、前記利用者の回答で似ていると判定された前記集合の前記評価値が良い方の前記グループに属する前記分析方法のスコアを加点するスコア化部と、前記組合せ抽出部、前記分析方法グループ化部、前記問合せ部、及び前記スコア化部のそれぞれの処理を繰り返し、前記スコアが所定値になる前記分析方法を選択する分析方法選択部とを備えることを要旨とする。

[0010] また、本発明の一態様に係るデータ分析方法選択方法は、上記のデータ分析方法選択装置が行う方法であって、分析部は、2つの時系列データの関係性を表す評価値を前記時系列データがそれぞれ記録された集合ごとに異なる分析方法で求める分析ステップと、組合せ抽出部は、前記分析方法に対応させて前記評価値の変化の傾向が異なる前記集合の組合せを抽出する組合せ抽出ステップと、分析方法グループ化部は、前記組合せ抽出ステップで抽出された前記組合せのそれぞれについて、前記評価値の良否で前記分析方法をグループに分類し、該分類した結果を前記集合に対応させて記録する分析方法

グループ化ステップと、問合せ部は、前記組合せ抽出部が抽出した前記組合せのそれぞれの前記集合の時系列データを利用者に提示し、どちらの前記集合の時系列データが似ているかを利用者に問合せる問合せステップと、スコア化部は、前記利用者の回答で似ていると判定された前記集合の前記評価値が良い方の前記グループに属する前記分析方法のスコアを加点するスコア化ステップと、前記組合せ抽出ステップ、前記分析方法グループ化ステップ、前記問合せステップ、及び前記スコア化ステップのそれぞれの処理を繰り返し、前記スコアが所定値になる前記分析方法を選択する分析方法選択部とを行うことを要旨とする。

[0011] また、本発明の一態様に係るプログラムは、上記のデータ分析方法選択装置としてコンピュータを機能させるためのプログラムであることを要旨とする。

発明の効果

[0012] 本発明によれば、ノウハウが無いような場合においても適切な分析手法を利用者のフィードバックを効かせて絞り込み、適切なデータ分析方法を選択できるデータ分析方法選択装置、方法及びプログラムを提供することができる。

図面の簡単な説明

[0013] [図1]本発明の実施形態に係るデータ分析方法選択装置の構成例を示す図である。

[図2]ある集合の時系列データと、該時系列データを異なる分析方法で分析して求めた評価値の例を示す図である。

[図3]図1に示す評価値テーブルの一例を模式的に示す図である。

[図4]図1に示すスコアテーブルの一例を模式的に示す図である。

[図5]図1に示す分析方法選択部の作用を説明するための図である。

[図6]分析方法（１）を説明するための図である。

[図7]分析方法（２）を説明するための図である。

[図8]分析方法（３）を説明するための図である。

[図9]分析方法（４）を説明するための図である。

[図10]図１に示すデータ分析方法選択装置の処理手順を示すフローチャートである

[図11]汎用的なコンピュータシステムの構成例を示すブロック図である。

発明を実施するための形態

[0014] 以下、本発明の実施形態について図面を用いて説明する。複数の図面中同一のものには同じ参照符号を付し、説明は繰り返さない。

[0015] 図１は、本発明の実施形態に係るデータ分析方法選択装置の構成例を示す図である。図１に示すデータ分析方法選択装置１００は、適切な分析手法を利用者のフィードバックを効かせて絞り込み、適切なデータ分析方法を選択するものである。

[0016] データ分析方法選択装置１００は、データ集合１０、分析部２０、評価値テーブル３０、組合せ抽出部４０、分析方法グループ化部５０、問合せ部６０、スコア化部７０、スコアテーブル８０、及び分析方法選択部９０を備える。データ分析方法選択装置１００は、例えば、ＲＯＭ、ＲＡＭ、ＣＰＵ等からなるコンピュータで実現することができる。その場合、各機能構成部の処理内容はプログラムによって記述される。

[0017] データ集合１０は、２つの時系列データをそれぞれ記録した集合Ａ，Ｂ，Ｃ，Ｄ，…を複数含む。集合Ａは、例えば、切り花（バラ）と情報通信関係費のそれぞれの物価指数の推移を記録したものである。集合Ｂは、例えば、下着と授業料の物価指数の推移を記録したものである。

[0018] 分析部２０は、２つの時系列データの関係性を表す評価値を集合Ａ，Ｂ，…ごとに異なる分析方法で求める。分析方法は、例えばＤＳの頭の中にある複数の分析方法のことである。

[0019] 図２は、データ集合の時系列データと、該時系列データを異なる分析方法で分析して求めた評価値の例を示す図である。図２（ａ）は、切り花（バラ）と情報通信関係費のそれぞれの物価指数の時系列データを示す。図２（ｂ）は、例えば４つの分析方法（１）～（４）のそれぞれで分析した評価値を

示す。

[0020] 評価値は、例えば集合 A の 2 つの時系列データが似ていれば値が小さくなる数値である。評価値の具体的な計算方法については後述する。

[0021] 図 2 (c) は、下着（ブラジャー）と大学授業料（国立）のそれぞれの物価指数の時系列データを示す。図 2 (d) は、図 2 (c) に示す 2 つの時系列データを分析方法（1）～（4）のそれぞれで分析した評価値を示す。

[0022] 評価値テーブル 30 は、集合 A, B, …ごとに異なる分析方法で分析して求めた評価値のテーブルである。評価値テーブル 30 は、行を集合 A, B, …ごと、列を分析方法ごとに記録した表である。

[0023] 図 3 は、評価値テーブル 30 の例を示す図である。表の各行が集合 A, B, …に対応し、各列が分析方法に対応する。なお、図 3 の集合 A, B の評価値は、説明の都合により図 2 の集合 A, B と異なる。

[0024] 集合 A の分析方法（1）の評価値は 0.09、途中省略、分析方法（4）の評価値は -0.02 である。分析方法は（1）～（4）の 4 種類に限られない。

[0025] 組合せ抽出部 40 は、分析方法に対応させて評価値の変化の傾向が異なる集合の組合せを抽出する。組合せ抽出部 40 は、例えば集合 A と集合 B の組合せを抽出する。

[0026] 評価値の変化の傾向が異なるとは、図 3 の集合 A と B に示すように、分析方法（1）～（4）の評価値が例えば逆転している様な場合である。集合 A は、分析方法（1）の評価値が大きく、分析方法（2）～（3）の評価値が大きい。

[0027] 一方、集合 B は、分析方法（1）の評価値が小さく、分析方法（2）～（3）の評価値が大きい。この例の場合、組合せ抽出部 40 は、集合 A と集合 B の組合せを抽出する。

[0028] このように組合せ抽出部 40 は、評価値の傾向が反対、及び評価値の差の大きな集合の組合せを抽出する。

[0029] 分析方法グループ化部 50 は、組合せ抽出部 40 で抽出され組合せのそれぞれについて、評価値の良否で分類方法をグループに分類し、該分類した結

果を集合に対応させて記録する。評価値の良否とは、良を2つの時系列データが似ている場合の例えば小さい数値の評価値、否を2つの時系列データが似ている場合の例えば大きな数値の評価値とする。

[0030] 図3に示す集合Aの場合は、分析方法(1)を「否」、分析方法(2)～(4)を「良」にグループ分けする。図3に示す集合Bの場合は、分析方法(1)を「良」、分析方法(2)～(4)を「否」にグループ分けする。

[0031] なお、図3に示す評価値テーブルでは、分析方法の良否が陽として表されていない。良否は、例えば表のマス目に対応させた良否フラグで表わす様にしてもよい。

[0032] 問合せ部60は、組合せ抽出部40が抽出した組合せのそれぞれの集合の時系列データを利用者に提示し、どちらの集合の時系列データが似ているかを利用者に問い合わせる。問い合わせは、図示を省略している操作パネル等に、例えば「集合Aと集合Bのどちらが似ていますか？」等を表示して行う。

[0033] スコア化部70は、利用者の回答で似ていると判定された集合の評価値が良い方のグループに属する分析方法のスコアを加点する。利用者の回答は、例えばタッチパネルで構成される操作パネル(図示せず)に利用者がタッチすることで行われる。

[0034] 利用者の回答は、一方の集合の時系列データが似ている、他方の集合のデータ集合が似ている、及び分からない、の何れかである。これにより利用者(人)の感性を適切に取り込むことができる。

[0035] 図2に示した例において、利用者が集合Aの2つの時系列データの方が、集合Bよりも似ていると回答したと仮定する。この場合、スコア化部70は、集合Aの分析方法(1)にスコア1を加点する。

[0036] 図4は、スコア化部70がスコアを加点した結果を記録したスコアテーブルの例を示す図である。図4に示す例は、集合A-Bの組合せを7回利用者に問い合わせた場合を示す。また、集合C-Dの組合せを33回利用者に問い合わせた場合を示す。なお、集合A-Bにおいて7回の利用者は、それぞれ

異なる人である。

[0037] 集合Aは、図3に示したように分析方法(1)を「否」、分析方法(2)～(4)を「良」にグループ分けしているので、集合Aの方が似ていると判定された場合に分析方法の(2)～(4)マス目にスコア1が加点される。

[0038] なお、利用者は、分析方法(1)～(4)について関知しない。分析方法(1)～(4)及びそれぞれに対応する評価値は、データ分析方法選択装置100の内部の情報であり表に出ない。複数の分析方法及びそれぞれの評価値はブラックボックス化されている。

[0039] 分析方法選択部90は、組合せ抽出部40、分析方法グループ化部50、問合せ部60、及びスコア化部70のそれぞれの処理を繰り返し、スコアが所定値になる分析方法を選択する。

[0040] 分析方法選択部90の作用によって、問合せ部60は複数のデータ集合10の組合せを利用者に提示する。利用者に提示するデータ集合10の組合せの数PNは、データ集合10を構成する集合の数をNとすると次式で表せる。

[0041] [数1]

$$PN = N \times \frac{(N-1)}{2} \quad (1)$$

[0042] 例えば集合の数をA、B、Cの3つとすると、データ集合10の組合せは、A-B、B-C、C-Aの3つである。N=100とするとPN=4900である。

[0043] 図5は、N=3とした場合の分析方法選択部90の作用を説明するための図である。なお、集合A、B、Cの評価値の変化の傾向は、それぞれ異なるものとして説明する。

[0044] 問合せ部60は、最初に組合せA-Bのどちらの時系列データが似ているかを利用者に問い合わせる。例えば集合Aの方が似ていると回答した場合、図5に示すように分析方法(2)～(3)が評価値の良いグループに分類されているので、スコア化部70は分析方法(2)～(3)のそれぞれにスコア

1 を加点する。

[0045] この場合、図4に示す集合A-Bの行の方法(2)～(4)のそれぞれが加点され+1になる。図4の表記は異なっている。

[0046] 次に、問合せ部60は、組合せB-Cのどちらの時系列データが似ているかを利用者に問い合わせる。例えば集合Bの方が似ていると回答した場合、図5に示すように分析方法(1)(3)(4)のグループの評価値が良いので、スコア化部70は分析方法(1)(3)(4)のそれぞれにスコア1を加点する。

[0047] 次に、問合せ部60は、組合せC-Aのどちらの時系列データが似ているかを利用者に問い合わせる。例えば集合Cの方が似ていると回答した場合、図5に示すように分析方法(2)(3)(4)のグループの評価値が良いので、スコア化部70は分析方法(2)(3)(4)のそれぞれにスコア1を加点する。

[0048] 以上の処理の結果、スコアテーブルにおける各分析方法(1)～(4)のスコアは、分析方法(3)のスコアが3ポイントとなり最も多くなる。この場合、分析方法選択部90は分析方法(3)を選択する。

[0049] 実際は、利用者に提示するデータ集合10の組合せの数PNはもっと多数であり、分析方法を選択する所定値ももっと大きな数値になる。

[0050] 以上説明したように、本実施形態に係るデータ分析方法選択装置100は、2つの時系列データをそれぞれ記録した集合A, B, ...を複数含むデータ集合10と、2つの時系列データの関係性を表す評価値を集合A, B, ...ごとに異なる分析方法で求める分析部20と、分析方法に対応させて評価値の変化の傾向が異なる集合A, B, ...の組合せを抽出する組合せ抽出部40と、組合せ抽出部40で抽出された組合せ(A-B等)のそれぞれについて、評価値の良否で分析方法をグループに分類し、該分類した結果を集合に対応させて記録する分析方法グループ化部50と、組合せ抽出部40が抽出した組合せのそれぞれの集合(A-B等)の時系列データを利用者に提示し、どちらの集合A, Bの時系列データが似ているかを利用者に問合せる問合せ部60

と、利用者の回答で似ていると判定された集合の評価値が良い方のグループに属する分析方法のスコアを加点するスコア化部70と、組合せ抽出部40、分析方法グループ化部50、問合せ部60、及びスコア化部70のそれぞれの処理を繰り返し、スコアが所定値になる分析方法を選択する分析方法選択部90とを備える。これにより、ノウハウが無いような場合においても適切な分析手法を利用者のフィードバックを効かせて絞り込み、適切なデータ分析方法を選択できるデータ分析方法選択装置を提供することができる。

[0051] 本実施形態は、2つの時系列データの関係に着目し、その関係性を数値化し、利用者に2つの時系列データを画像化して提示して利用者の回答をフィードバックする。その結果、複数の分析方法から人（利用者）の感覚に近い分析方法を選択することができる。したがって、利用者に専門的な知識が無くても最適な分析方法の選択を可能にする。

[0052] つまり、本実施形態は、完全な分析方法は存在しないという前提に基づいて複数の分析方法の結果を利用者に提示し、利用者がより良い分析方法を選択するという仕組みを提供する。なお、分析方法を提示される利用者（後述する被験者）と、本実施形態に係るデータ分析方法選択装置100を利用する利用者とは基本的には別である。データ分析方法選択装置100を利用する人の方が多くなる。また、分析方法を提示される利用者は一人であってもよいし、複数であっても構わない。

[0053] なお、分析方法を提示される利用者が一人の場合、スコア化部70が加点するスコアは1である。また、データ分析方法選択装置100を利用する利用者が変わっても、ある一組の時系列データを分析する分析方法は最適なものが一つ選択される。

[0054] 次に分析方法の具体例について説明する。

[0055] （分析方法（1））

図6は、分析方法（1）を説明するための図である。図6は、2つの物価指数の時系列データを示す。図6の横軸は時間、縦軸は物価指数である。

[0056] 分析方法（1）は、一点鎖線と実線で示す比較する2つの物価指数につい

て、2つの時系列データの対応するデータの差分の累積値を、該累積したデータ数で除算する。なお、差分については符号付でもよいし、絶対値で扱ってもよい。なお、図6において破線で示すように、一方にしかデータがない場合は加算しない。

[0057] この分析方法（1）は、比較する2つの物価指数データの数が多く、季節変動のような時間ごとの変動が少ないものに好適である。

[0058] （分析方法（2））

図7は、分析方法（2）を説明するための図である。図7の横軸と縦軸の関係は図6と同じである。

[0059] 分析方法（2）は、2つの時系列データのそれぞれの変化量を求め、該変化量の差分の累積値を該累積したデータ数で除算する。図7に示す時間5の差分は $2 - (-2) = 4$ である。分析方法（1）と同様に、一方にしかデータがない場合は加算しない。

[0060] この分析方法（2）は、比較する2つの物価指数データの個数が多く、差分の絶対値は大きく且つ変動の形が似ているものに好適である。

[0061] （分析方法（3））

図8は、分析方法（3）を説明するための図である。図7の横軸と縦軸の関係は図6と同じである。

[0062] 分析方法（3）の計算方法は、上記の分析方法（2）と基本的に同じである。ただし、2つの時系列データの一方のデータしか無い場合は、他方の時系列データの変化量を該時系列データの変化量の平均値で補間する点である。なお、両方共にデータが無い区間については補間を行わない。

[0063] この分析方法（3）は、分析方法（2）と比べて、比較する2つの時系列データの一方にデータの無い区間が多いものに好適である。

[0064] （分析方法（4））

図9は、分析方法（4）を説明するための図である。図7の横軸と縦軸の関係は図6と同じである。

[0065] 分析方法（3）の計算方法は、上記の分析方法（2）と基本的に同じであ

る。ただし、上記の平均値は、時系列データが無くなる直前の複数の変化量の平均値である。平均するデータの個数及び平均時の重みづけについては変更してもよい。

[0066] この分析方法（４）は、上記の分析方法（１）が不適な季節変動の大きな時系列データの比較に好適である。

[0067] （データ分析方法選択方法）

図１０は、本実施形態に係るデータ分析方法選択装置１００が行うデータ分析方法選択方法の処理手順を示すフローチャートである。

[0068] データ分析方法選択装置１００は、２つの時系列データをそれぞれ記録した集合Ａ、Ｂ、…を複数含むデータ集合１０を備える。データ集合１０は予め用意される。集合…は適宜追加される。

[0069] データ分析方法選択装置１００の分析部２０は、２つの時系列データの関係性を表す評価値を集合Ａ、Ｂ、…ごとに異なる分析方法（例えば上記の（１）～（４））で算出する（ステップＳ１）。

[0070] 組合せ抽出部４０は、分析方法に対応させて評価値の変化の傾向が異なる集合の組合せを抽出する（ステップＳ２）。集合の組合せは、例えば、Ａ－Ｂ、Ｂ－Ｃ、Ｃ－Ａ等である。

[0071] 分析方法グループ化部５０は、組合せ抽出部４０で抽出された集合の組合せのそれぞれについて、評価値の良否で分析方法をグループに分類し、該分類した結果を集合に対応させて記録する（ステップＳ３）。

[0072] 問合せ部６０は、組合せ抽出部４０が抽出した組合せのそれぞれの集合の時系列データを利用者に提示し、どちらの集合の時系列データが似ているかを利用者に問い合わせる（ステップＳ４）。

[0073] 利用者は、どちらの集合の時系列データが似ているかを回答する（ステップＳ５）。回答は、例えば操作パネル（図示せず）等を利用者がタッチすることで行う。

[0074] スコア化部７０は、利用者の回答で似ていると判定された集合の評価値が良い方のグループに属する分析方法のスコアを加点する。例えば、集合Ａの

時系列データの方が似ていると判定された場合は、スコアテーブル（図4）の集合の例えばA-Bの方法（1）にスコアを加点する（ステップS6）。また、集合Bの時系列データの方が似ていると判定された場合は、スコアテーブル（図4）の集合の例えばA-Bの方法（2）（3）（4）にスコアを加点する（ステップS7）。

[0075] 分析方法選択部90は、組合せ抽出ステップ（ステップS2）、分析方法グループ化ステップ（ステップS3）、問合せステップ（ステップS4）、及びスコア化ステップ（ステップS5）のそれぞれの処理を繰り返し、スコアが所定値になる分析方法を選択する（ステップS8のYES）。なお、集合が追加された場合、処理の繰り返しは分析部20の処理（ステップS2）から繰り返す。

[0076] データ分析方法選択装置100は、図8に示す汎用的なコンピュータシステムで実現することができる。例えば、CPU90、メモリ91、ストレージ92、通信部93、入力部94、及び出力部95を備える汎用的なコンピュータシステムにおいて、CPU90がメモリ91上にロードされた所定のプログラムを実行することにより、データ分析方法選択装置100の各機能が実現される。所定のプログラムは、HDD、SSD、USBメモリ、CD-ROM、DVD-ROM、MOなどのコンピュータ読取り可能な記録媒体に記録することも、ネットワークを介して配信することもできる。

[0077] （評価実験）

本実施形態に係るデータ分析方法選択装置100で得られる効果を確認する目的で評価実験を行った。

[0078] 評価実験には、総省統計局が提供している消費者物価指数（品目別価格指数）から時系列データ380項目を使用した。380項目を組み合わせた約7万2千個の集合を用いて、評価値の算出方法が異なる分析方法の中から、最も適した分析方法を選択する実験を行った。

[0079] 分析方法は、上記の分析方法（1）～（4）の4つを用いた。集合の選択は、分析方法1種類につき20回実施した。その事前評価の結果、分析方法

(1) が最も被験者（利用者（人））の感覚に合うことが分かった。

[0080] その後、被験者4人に対して、ランダムに抽出した集合の10組について同様の評価を行った。その結果を表1に示す。

[0081] [表1]

被験者	評価回数	方法（1） との一致率	方法（2） との一致率	方法（3） との一致率	方法（4） との一致率
1	10	100%	75%	75%	88%
2	10	80%	60%	50%	60%
3	10	78%	67%	56%	67%
4	10	100%	67%	67%	78%
計	40	89%	67%	61%	72%

[0082] 表1に示すように、事前の評価で分析方法（1）が最も被験者に合うと判定された分析方法（1）の一致率が平均値89%と最も高く、データ分析方法選択装置100を用いることで、比較的少ない試行回数で分析方法が選択できることが分かった。

[0083] 本実施形態によれば、2つのデータの集合の時系列データの関係性に着目し、その関係性を数値化するだけでなく、視覚化して利用者に提示し、利用者から回答を得ることで複数の分析方法の中から人の感覚に近い分析方法を選択することができる。

[0084] つまり、DSと異なりノウハウのない利用者であっても適切なデータ分析方法を選択することが可能になる。

[0085] なお、上記の実施例では、分析方法を（1）～（4）の4種類で説明したが、本発明はこの例に限定されない。分析方法はn個（nは自然数）であってもよい。また、分析方法は上記の実施例に限定されない。また、集合A、Bは、物価指数の時系列データを例に示したが、他の時系列データであっても構わない。

[0086] このように本発明はここでは記載していない様々な実施形態等を含むことは勿論である。したがって、本発明の技術的範囲は上記の説明から妥当な特許請求の範囲に係る発明特定事項によってのみ定められるものである。

符号の説明

- [0087] 1 0 : データ集合
2 0 : 分析部
3 0 : 評価値テーブル
4 0 : 組合せ抽出部
5 0 : 分析方法グループ化部
6 0 : 問合せ部
7 0 : スコア化部
8 0 : スコアテーブル
9 0 : 分析方法選択部
1 0 0 : データ分析方法選択装置
A, B, C, D : 集合

請求の範囲

- [請求項1] 2つの時系列データをそれぞれ記録した集合を複数含むデータ集合と、
- 前記2つの時系列データの関係性を表す評価値を前記集合ごとに異なる分析方法で求める分析部と、
- 前記分析方法に対応させて前記評価値の変化の傾向が異なる前記集合の組合せを抽出する組合せ抽出部と、
- 前記組合せ抽出部で抽出された前記組合せのそれぞれについて、前記評価値の良否で前記分析方法をグループに分類し、該分類した結果を前記集合に対応させて記録する分析方法グループ化部と、
- 前記組合せ抽出部が抽出した前記組合せのそれぞれの前記集合の時系列データを利用者に提示し、どちらの前記集合の時系列データが似ているかを利用者に問合せる問合せ部と、
- 前記利用者の回答で似ていると判定された前記集合の前記評価値が良い方の前記グループに属する前記分析方法のスコアを加点するスコア化部と、
- 前記組合せ抽出部、前記分析方法グループ化部、前記問合せ部、及び前記スコア化部のそれぞれの処理を繰り返し、前記スコアが所定値になる前記分析方法を選択する分析方法選択部と
- を備えるデータ分析方法選択装置。
- [請求項2] 前記利用者の回答は、
- 一方の前記時系列データが似ている、他方の前記時系列データが似ている、及び分からない、の何れかである
- 請求項1に記載のデータ分析方法選択装置。
- [請求項3] 前記分析方法の1つは、
- 前記2つの時系列データの対応するデータの差分を累積した累積値を、該累積したデータ数で除算する
- 請求項1又は2に記載のデータ分析方法選択装置。

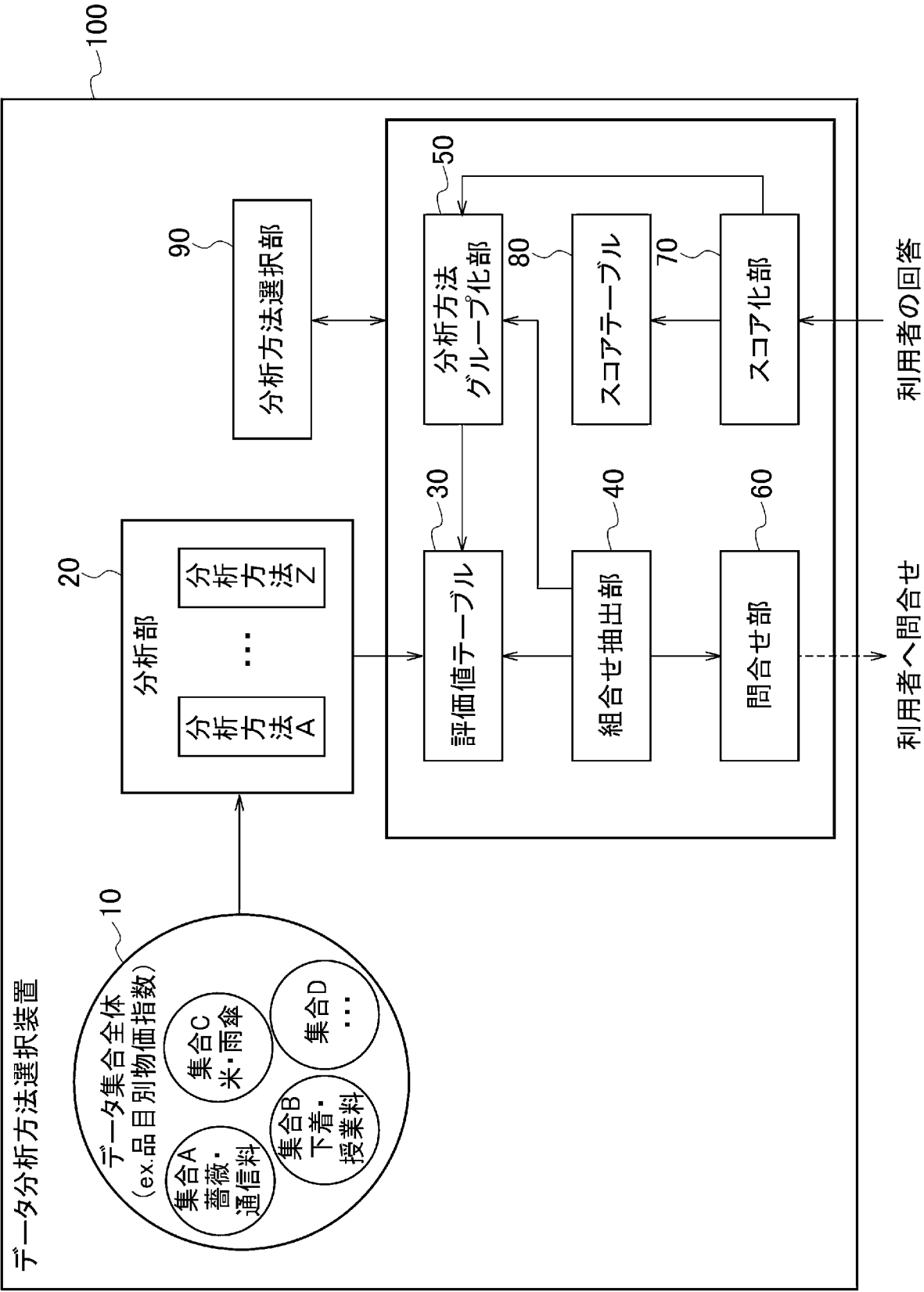
- [請求項4] 前記分析方法の1つは、
前記2つの時系列データのそれぞれの変化量を求め、該変化量の差分を累積した累積値を該累積したデータ数で除算する
請求項1又は2に記載のデータ分析方法選択装置。
- [請求項5] 前記分析方法の1つは、
前記2つの時系列データの一方しか無い場合は、他方の前記時系列データの前記変化量を該時系列データの前記変化量の平均値で補間する
請求項4に記載のデータ分析方法選択装置。
- [請求項6] 前記平均値は、
前記時系列データが無くなる直前の複数の前記変化量の平均値である
請求項5に記載のデータ分析方法選択装置。
- [請求項7] 分析部は、2つの時系列データの関係性を表す評価値を前記時系列データがそれぞれ記録された集合ごとに異なる分析方法で求める分析ステップと、
組合せ抽出部は、前記分析方法に対応させて前記評価値の変化の傾向が異なる前記集合の組合せを抽出する組合せ抽出ステップと、
分析方法グループ化部は、前記組合せ抽出ステップで抽出された前記組合せのそれぞれについて、前記評価値の良否で前記分析方法をグループに分類し、該分類した結果を前記集合に対応させて記録する分析方法グループ化ステップと、
問合せ部は、前記組合せ抽出部が抽出した前記組合せのそれぞれの前記集合の時系列データを利用者に提示し、どちらの前記集合の時系列データが似ているかを利用者に問合せる問合せステップと、
スコア化部は、前記利用者の回答で似ていると判定された前記集合の前記評価値が良い方の前記グループに属する前記分析方法のスコアを加点するスコア化ステップと、

前記組合せ抽出ステップ、前記分析方法グループ化ステップ、前記問合せステップ、及び前記スコア化ステップのそれぞれの処理を繰り返し、前記スコアが所定値になる前記分析方法を選択する分析方法選択部と

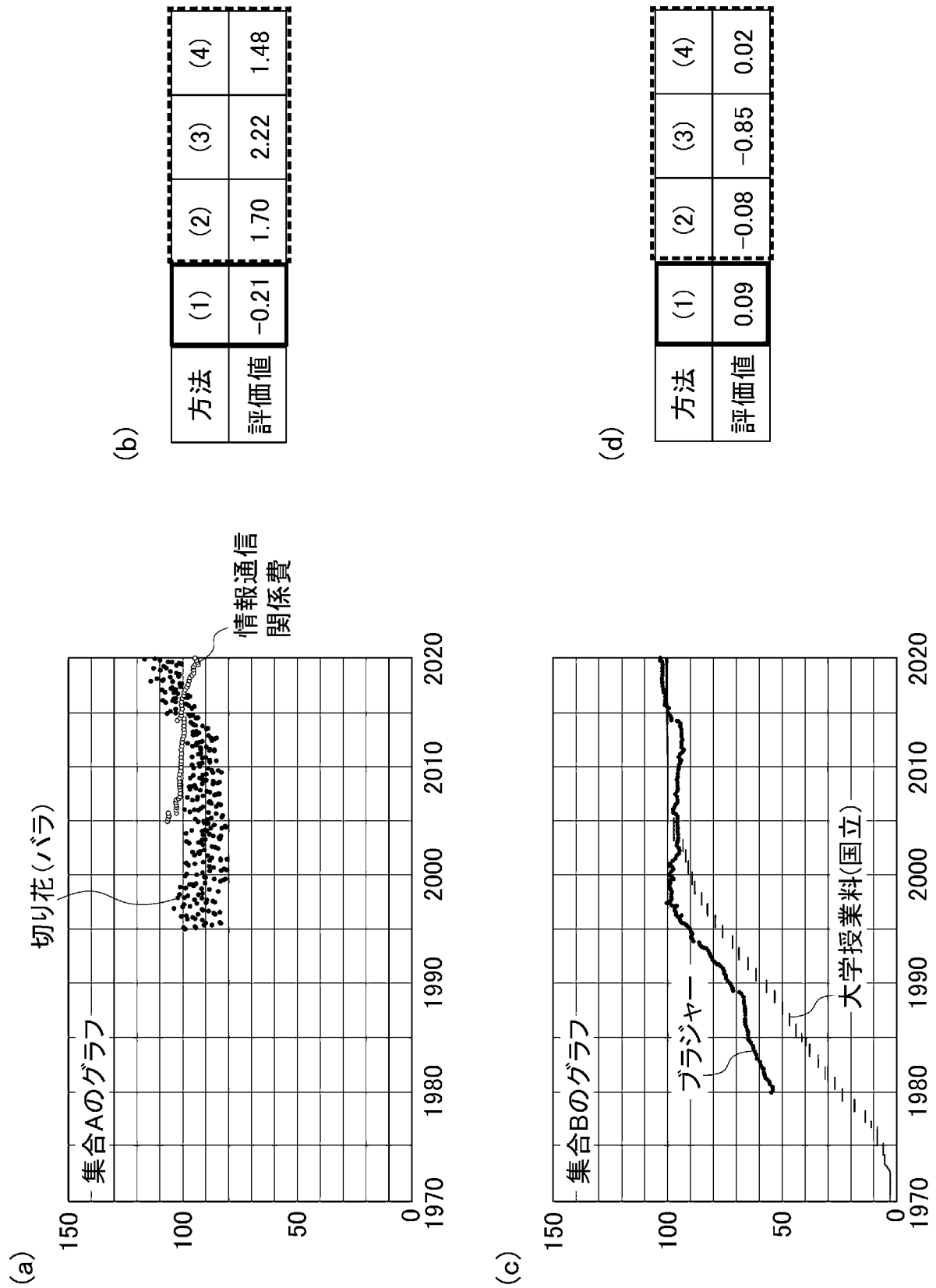
を行うデータ分析方法選択方法。

[請求項8] 請求項1乃至6の何れかに記載のデータ分析方法選択装置としてコンピュータを機能させるためのプログラム。

[図1]



[図2]



[図3]

評価値テーブル

データ集合	方法(1)	方法(2)	方法(3)	方法(4)	...
A	0.09	-0.08	-0.85	-0.02	...
B	-0.21	1.70	2.22	1.48	...
C	0.10	-1.10	2.03	-0.52	...
D	2.00	-2.88	3.25	2.91	...
...	...	...	...	...	...

[図4]

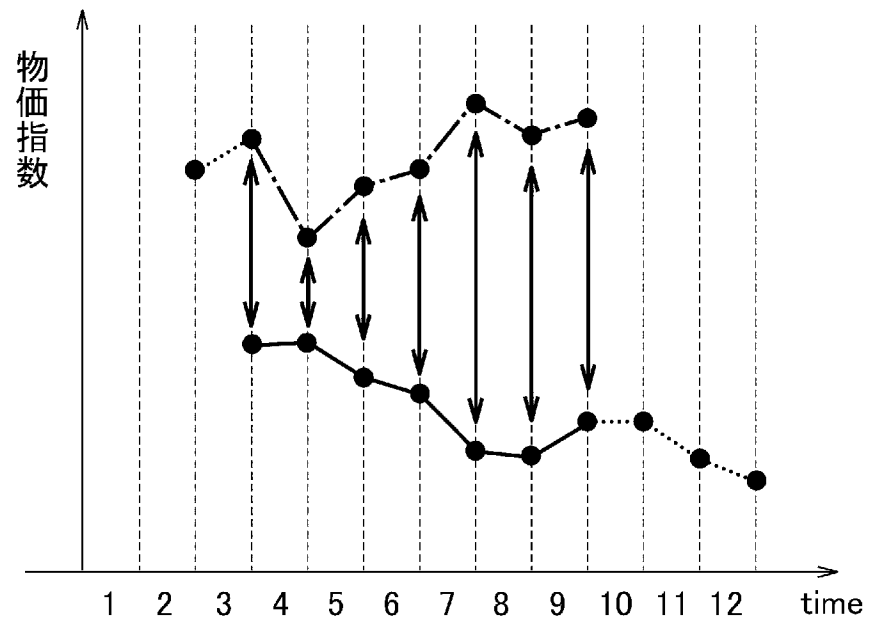
スコアテーブル

データ集合	方法(1)	方法(2)	方法(3)	方法(4)	...	分からない
A&B	+5	0	0	0	...	+2
C&D	+10	+2	+10	+2	...	+1
...	...	...	...	...	...	

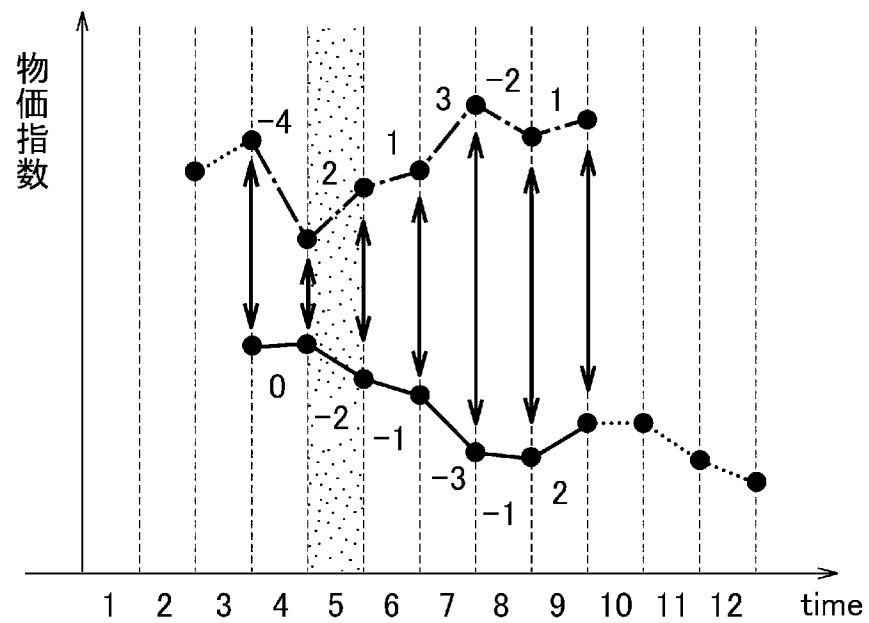
[図5]

	(1)	(2)	(3)	(4)
A	0.09	-0.08	-0.85	-0.02
B	-0.21	1.70	2.22	1.48
C	0.10	-1.10	2.03	-0.52
...				

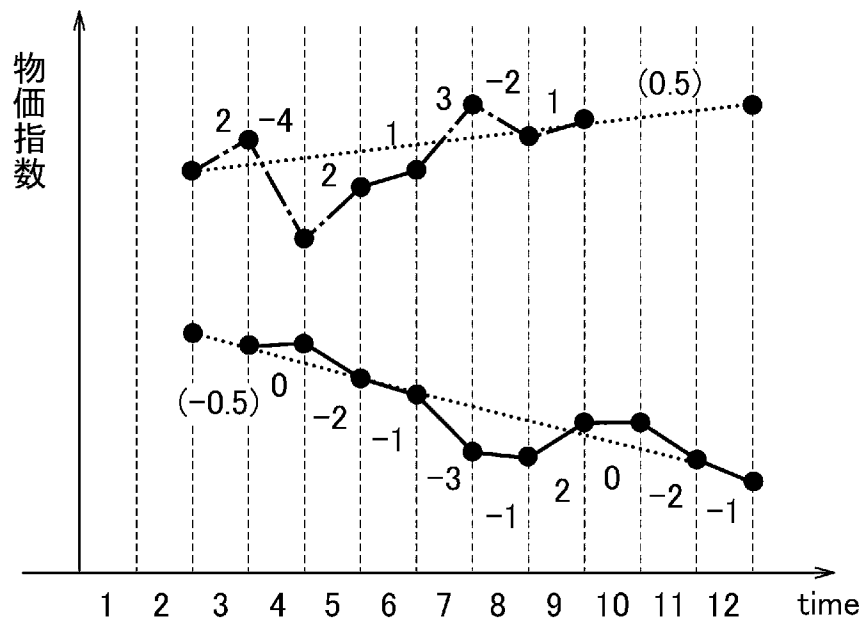
[図6]



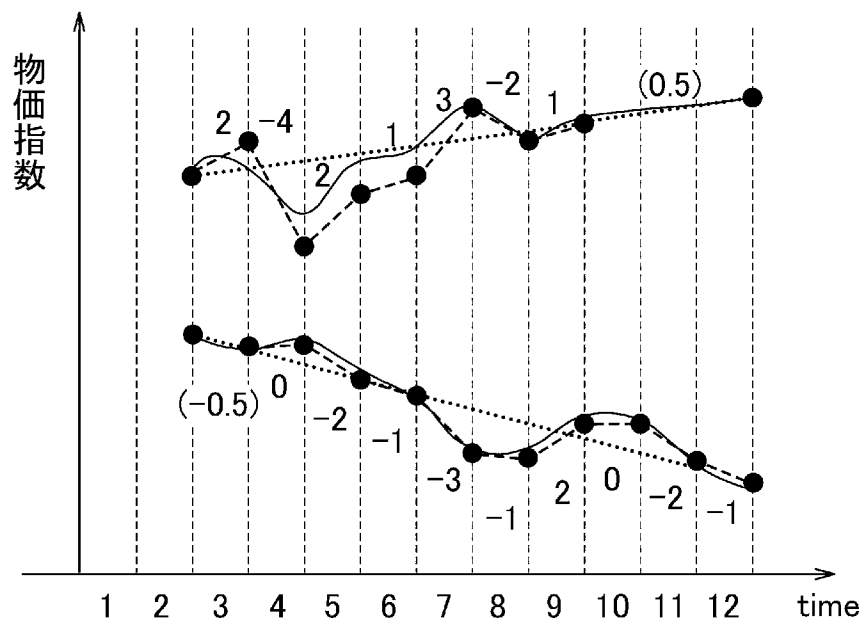
[図7]



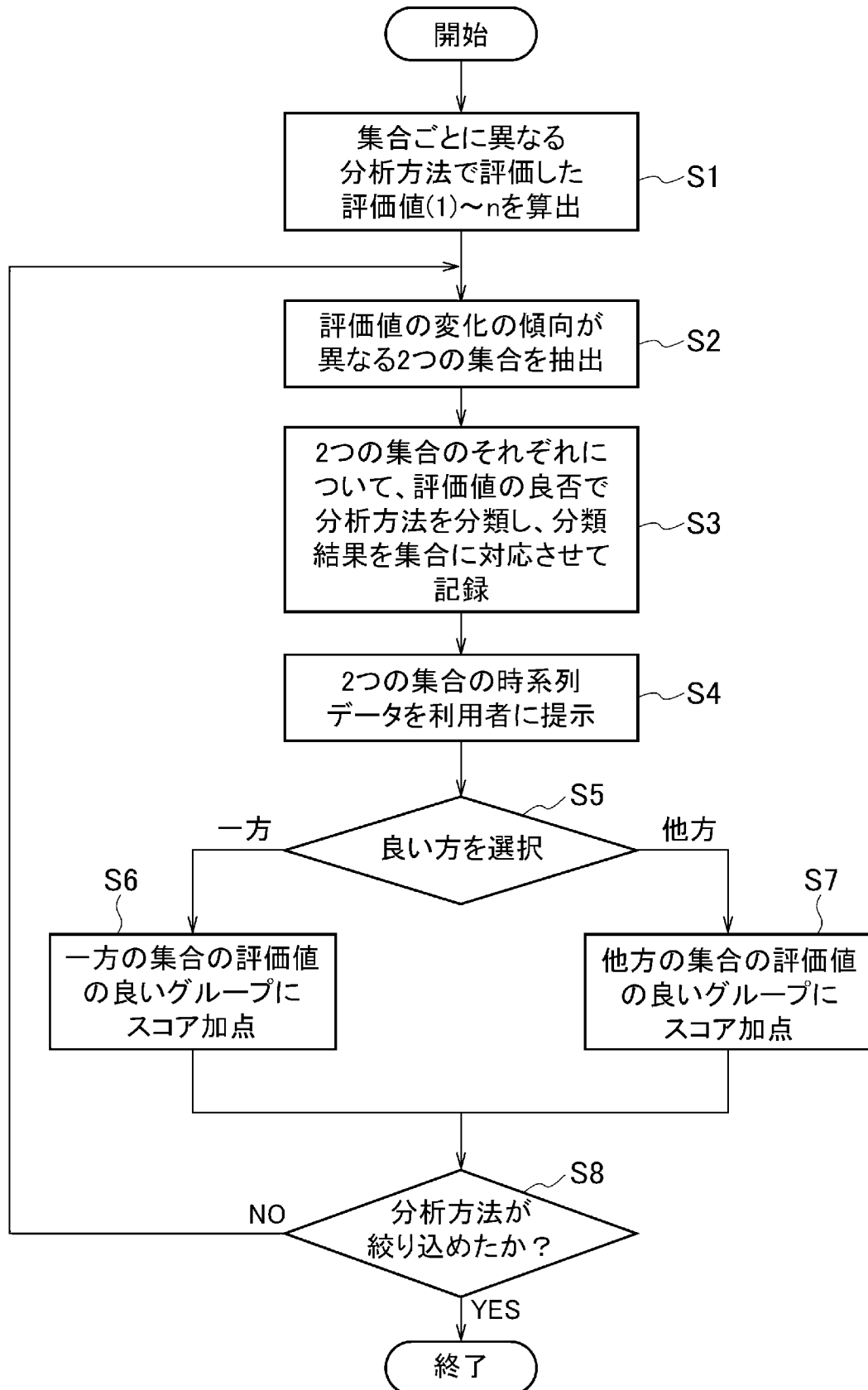
[図8]



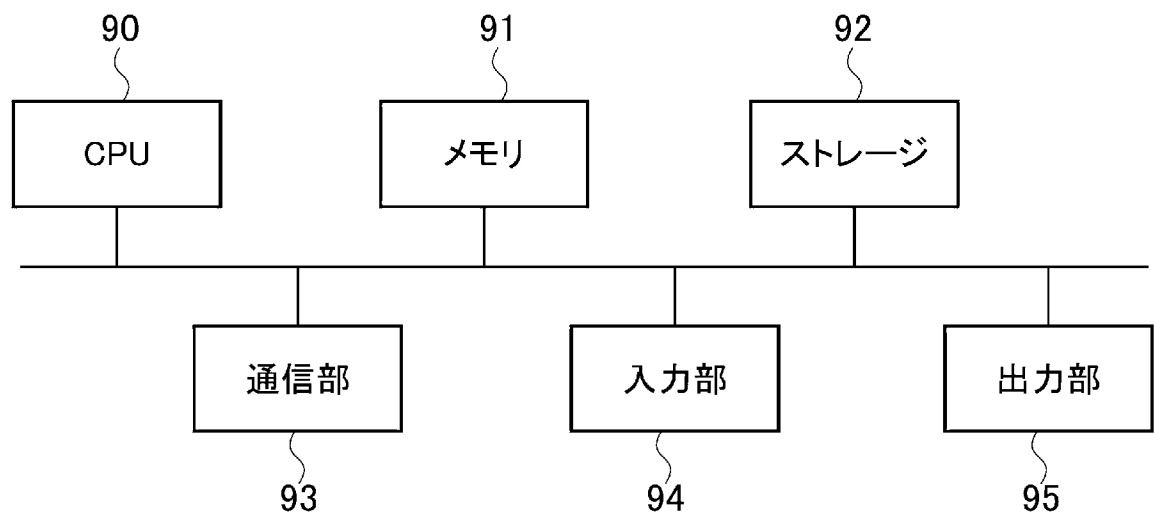
[図9]



[図10]



[図11]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2021/005698

A. CLASSIFICATION OF SUBJECT MATTER

G06Q 10/00(2012.01)i

FI: G06Q10/00 300

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06Q10/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2021

Registered utility model specifications of Japan 1996-2021

Published registered utility model applications of Japan 1994-2021

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2019/187012 A1 (MITSUBISHI ELECTRIC CORP) 03 October 2019 (2019-10-03) entire text, all drawings	1-8
A	JP 2005-157896 A (MITSUBISHI ELECTRIC CORP) 16 June 2005 (2005-06-16) entire text, all drawings	1-8
A	JP 2020-170371 A (MITSUBISHI ELECTRIC CORP) 15 October 2020 (2020-10-15) entire text, all drawings	1-8
A	JP 2019-105953 A (HITACHI LTD) 27 June 2019 (2019- 06-27) entire text, all drawings	1-8
A	JP 2010-205218 A (DAINIPPON PRINTING CO LTD) 16 September 2010 (2010-09-16) entire text, all drawings	1-8



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

28 April 2021 (28.04.2021)

Date of mailing of the international search report

18 May 2021 (18.05.2021)

Name and mailing address of the ISA/

Japan Patent Office

3-4-3, Kasumigaseki, Chiyoda-ku,

Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2021/005698

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2017/168967 A1 (MITSUBISHI ELECTRIC CORP) 05 October 2017 (2017-10-05) entire text, all drawings	1-8

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/JP2021/005698

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
WO 2019/187012 A1	03 Oct. 2019	TW 201942817 A	
JP 2005-157896 A	16 Jun. 2005	(Family: none)	
JP 2020-170371 A	15 Oct. 2020	(Family: none)	
JP 2019-105953 A	27 Jun. 2019	(Family: none)	
JP 2010-205218 A	16 Sep. 2010	(Family: none)	
WO 2017/168967 A1	05 Oct. 2017	CN 108885628 A	

A. 発明の属する分野の分類（国際特許分類（IPC）） G06Q 10/00(2012.01)i FI: G06Q10/00 300		
B. 調査を行った分野		
調査を行った最小限資料（国際特許分類（IPC）） G06Q10/00		
最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1922-1996年 日本国公開実用新案公報 1971-2021年 日本国実用新案登録公報 1996-2021年 日本国登録実用新案公報 1994-2021年		
国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）		
C. 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	WO 2019/187012 A1（三菱電機株式会社）03.10.2019（2019-10-03） 全文, 全図	1-8
A	JP 2005-157896 A（三菱電機株式会社）16.06.2005（2005-06-16） 全文, 全図	1-8
A	JP 2020-170371 A（三菱電機株式会社）15.10.2020（2020-10-15） 全文, 全図	1-8
A	JP 2019-105953 A（株式会社日立製作所）27.06.2019（2019-06-27） 全文, 全図	1-8
A	JP 2010-205218 A（大日本印刷株式会社）16.09.2010（2010-09-16） 全文, 全図	1-8
A	WO 2017/168967 A1（三菱電機株式会社）05.10.2017（2017-10-05） 全文, 全図	1-8
☐ C欄の続きにも文献が列挙されている。 ☒ パテントファミリーに関する別紙を参照。		
* 引用文献のカテゴリー “A” 特に関連のある文献ではなく、一般的技術水準を示すもの “E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの “L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す） “O” 口頭による開示、使用、展示等に言及する文献 “P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献 “T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの “X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの “Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの “&” 同一パテントファミリー文献		
国際調査を完了した日 28.04.2021		国際調査報告の発送日 18.05.2021
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号		権限のある職員（特許庁審査官） 小山 和俊 5L 9369 電話番号 03-3581-1101 内線 3562

国際調査報告
 パテントファミリーに関する情報

国際出願番号

PCT/JP2021/005698

引用文献			公表日	パテントファミリー文献			公表日
WO	2019/187012	A1	03.10.2019	TW	201942817	A	
JP	2005-157896	A	16.06.2005	(ファミリーなし)			
JP	2020-170371	A	15.10.2020	(ファミリーなし)			
JP	2019-105953	A	27.06.2019	(ファミリーなし)			
JP	2010-205218	A	16.09.2010	(ファミリーなし)			
WO	2017/168967	A1	05.10.2017	CN	108885628	A	