

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5457737号
(P5457737)

(45) 発行日 平成26年4月2日(2014.4.2)

(24) 登録日 平成26年1月17日(2014.1.17)

(51) Int. Cl. F I
G05B 23/02 (2006.01) G05B 23/02 P
 G05B 23/02 301X

請求項の数 8 (全 42 頁)

(21) 出願番号	特願2009-151745 (P2009-151745)	(73) 特許権者	504132272 国立大学法人京都大学
(22) 出願日	平成21年6月26日 (2009. 6. 26)		京都府京都市左京区吉田本町36番地1
(65) 公開番号	特開2011-8562 (P2011-8562A)	(74) 代理人	100099933 弁理士 清水 敏
(43) 公開日	平成23年1月13日 (2011. 1. 13)	(72) 発明者	加納 学 京都府京都市西京区京都大学桂 国立大学 法人京都大学工学研究科内
審査請求日	平成24年6月20日 (2012. 6. 20)	(72) 発明者	藤原 幸一 京都府京都市西京区京都大学桂 国立大学 法人京都大学工学研究科内
		審査官	青山 純

最終頁に続く

(54) 【発明の名称】 プラント制御情報生成装置及び方法、並びにそのためのコンピュータプログラム

(57) 【特許請求の範囲】

【請求項1】

プラントの稼動条件と、当該稼動条件での前記プラントの過去の稼動結果から得られるデータ項目の実測値との組合せからなる実測サンプルを複数個記憶するためのデータベースと、

前記データベースに記憶された前記複数個の実測サンプルを、複数個のクラスにクラスタリングするためのクラスタリング手段と、

前記クラスタリング手段によりクラスタリングされた各クラスに対して、前記プラントにおける稼動条件と前記データ項目の値との関係を示す統計的モデルを構築するためのモデル構築手段と、

新たな稼動条件が与えられたことに応答して、前記統計的モデルと、前記新たな稼動条件とに基づいて、前記プラントの制御に関連する情報を算出するための制御情報算出手段とを含む、プラント制御情報生成装置であって、

前記クラスタリング手段は、

前記複数個の実測サンプルの任意の2個の間の相関の程度を成分とする行列を、前記データベースに記憶された前記実測サンプルから算出するための行列算出手段と、

前記行列算出手段により算出された前記行列を類似度行列として、前記複数個の実測サンプルに対するスペクトル・クラスタリングを行なって、前記複数個の実測サンプルを前記複数個のクラスにクラスタリングするための手段とを含み、

前記行列算出手段は、

10

20

前記プラントの稼動条件と前記データ項目とにより定義された実測ベクトルのベクトル空間において、前記複数個の実測サンプルにより定められる複数個の実測ベクトルのうち、基準となる実測サンプルに対応する実測ベクトルを他の実測ベクトルから減算するための減算手段と、

前記減算手段により減算された後の前記他の実測ベクトルの任意の2個について、共通の線形部分空間上に存在するか否かを基準として、互いの相関を求めるための相関算出手段と、

前記減算手段による減算と、前記相関算出手段による相関の算出とを、前記基準となる実測サンプルを変えながら前記複数個の実測サンプルの全てに対して行なうための繰返算出手段と、

前記繰返算出手段により算出された、前記複数個の実測サンプルの任意の2個の間で算出される複数個の相関に基づいて、前記行列を決定するための手段とを含む、プラント制御情報生成装置。

【請求項2】

前記行列を決定するための手段は、前記繰返算出手段により、前記複数個の実測サンプルの任意の2個の間で算出された前記複数個の相関を合計した値を成分とする行列を前記行列として出力するための手段を含む、請求項1に記載のプラント制御情報生成装置。

【請求項3】

前記相関算出手段は、

前記減算手段により減算された後の前記他の実測ベクトルの任意の2個について、共通の線形部分空間上に存在するか否かを基準として、互いの相関係数を求めるための相関係数算出手段と、

前記相関係数算出手段により算出された相関係数が所定のしきい値より大きいか否かにしたがって、前記相関の値を2値化するための手段とを含む、請求項1又は請求項2に記載のプラント制御情報生成装置。

【請求項4】

前記制御情報算出手段は、

新たな稼動条件が与えられたことに応答して、当該稼動条件が前記複数個のクラスのいずれに属するかを判定するための手段と、

当該クラスに対応する統計的モデルを、前記新たな稼動条件に適用することにより、前記新たな稼動条件の下での前記プラントにより生成される物質の属性を予測し、予測情報を出力するための手段とを含む、請求項1～請求項3のいずれかに記載のプラント制御情報生成装置。

【請求項5】

前記プラント制御情報生成装置は前記プラントの稼動状態を測定するための測定手段をさらに含み、

前記制御情報算出手段は、前記測定手段による測定結果が新たな稼動条件として与えられたことに応答して、当該稼動条件と前記複数個のクラスとを比較することにより、前記プラントの状態が異常か否かを判定するための異常判定手段を含む、請求項1～請求項3のいずれかに記載のプラント制御情報生成装置。

【請求項6】

前記異常判定手段は、

前記測定手段による測定結果が新たな稼動条件として与えられたことに応答して、前記複数個のクラスの統計的モデルのうち、当該稼動条件が適合するものを判定するための適合モデル判定手段と、

前記適合モデル判定手段により、前記新たな稼動条件が前記複数個のクラスの統計的モデルのいずれにも適合しないと判定されたことに応答して、前記プラントの異常を示す信号を生成し出力するための手段とを含む、請求項5に記載のプラント制御情報生成装置。

【請求項7】

コンピュータにより実行されると、当該コンピュータを、

10

20

30

40

50

プラントの稼働条件と、当該稼働条件での前記プラントの過去の稼働結果から得られるデータ項目の実測値との組合せからなる実測サンプルを複数個記憶するためのデータベースと、

前記データベースに記憶された前記複数個の実測サンプルを、複数個のクラスにクラスタリングするためのクラスタリング手段と、

前記クラスタリング手段によりクラスタリングされた各クラスに対して、前記プラントにおける稼働条件と前記データ項目の値との関係を示す統計的モデルを構築するためのモデル構築手段と、

新たな稼働条件が与えられたことに応答して、前記統計的モデルと、前記新たな稼働条件とに基づいて、前記プラントの制御に関連する情報を算出するための制御情報算出手段として機能させるコンピュータプログラムであって、

10

前記クラスタリング手段は、

前記複数個の実測サンプルの任意の2個の間の相関の程度を成分とする行列を、前記データベースに記憶された前記実測サンプルから算出するための行列算出手段と、

前記行列算出手段により算出された前記行列を類似度行列として、前記複数個の実測サンプルに対するスペクトル・クラスタリングを行なって、前記複数個の実測サンプルを前記複数個のクラスにクラスタリングするための手段とを含み、

前記行列算出手段は、

前記プラントの稼働条件と前記データ項目とにより定義された実測ベクトルのベクトル空間において、前記複数個の実測サンプルにより定められる複数個の実測ベクトルのうち、基準となる実測サンプルに対応する実測ベクトルを他の実測ベクトルから減算するための減算手段と、

20

前記減算手段により減算された後の前記他の実測ベクトルの任意の2個について、共通の線形部分空間上に存在するか否かを基準として、互いの相関を求めるための相関算出手段と、

前記減算手段による減算と、前記相関算出手段による相関の算出とを、前記基準となる実測サンプルを変えながら前記複数個の実測サンプルの全てに対して行なうための繰返算出手段と、

前記繰返算出手段により算出された、前記複数個の実測サンプルの任意の2個の間で算出される複数個の相関に基づいて、前記行列を決定するための手段とを含む、コンピュータプログラム。

30

【請求項8】

プラントの稼働条件と、当該稼働条件での前記プラントの過去の稼働結果から得られるデータ項目の実測値との組合せからなる実測サンプルを複数個記憶するためのデータベースと、

前記データベースに記憶された前記複数個の実測サンプルを、複数個のクラスにクラスタリングするためのクラスタリング手段と、

前記クラスタリング手段によりクラスタリングされた各クラスに対して、前記プラントにおける稼働条件と前記データ項目の値との関係を示す統計的モデルを構築するためのモデル構築手段と、

40

新たな稼働条件が与えられたことに応答して、前記統計的モデルと、前記新たな稼働条件とに基づいて、前記プラントの制御に関連する情報を算出するための制御情報算出手段とを含む、プラント制御情報生成装置におけるプラント制御情報生成方法であって、

前記クラスタリング手段は、

前記複数個の実測サンプルの任意の2個の間の相関の程度を成分とする行列を、前記データベースに記憶された前記実測サンプルから算出するための行列算出手段と、

前記行列算出手段により算出された前記行列を類似度行列として、前記複数個の実測サンプルに対するスペクトル・クラスタリングを行なって、前記複数個の実測サンプルを前記複数個のクラスにクラスタリングするための手段とを含み、

前記行列算出手段は、

50

前記プラントの稼動条件と前記データ項目とにより定義された実測ベクトルのベクトル空間において、前記複数個の実測サンプルにより定められる複数個の実測ベクトルのうち、基準となる実測サンプルに対応する実測ベクトルを他の実測ベクトルから減算するための減算手段と、

前記減算手段により減算された後の前記他の実測ベクトルの任意の2個について、共通の線形部分空間上に存在するか否かを基準として、互いの相関を求めるための相関算出手段と、

前記減算手段による減算と、前記相関算出手段による相関の算出とを、前記基準となる実測サンプルを変えながら前記複数個の実測サンプルの全てに対して行なうための繰返算出手段と、

10

前記繰返算出手段により算出された、前記複数個の実測サンプルの任意の2個の間で算出される複数個の相関に基づいて、前記行列を決定するための手段とを含み、

前記プラント制御情報生成方法は、

前記クラスタリング手段が、前記データベースに記憶された前記複数個の実測サンプルを、複数個のクラスにクラスタリングするステップと、

前記モデル構築手段が、前記クラスタリングするステップにおいてクラスタリングされた各クラスに対して、前記プラントにおける稼動条件と前記データ項目の値との関係を示す統計的モデルを構築するステップと、

前記制御情報算出手段が新たな稼動条件が与えられたことに応答して、前記統計的モデルと、前記新たな稼動条件とに基づいて、前記プラントの制御に関連する情報を算出するステップとを含み、

20

前記クラスタリングするステップは、

前記行列算出手段が、前記複数個の実測サンプルの任意の2個の間の相関の程度を成分とする行列を、前記データベースに記憶された前記実測サンプルから算出するステップと

、
前記クラスタリングするための手段が、前記算出するステップにおいて算出された前記行列を類似度行列として、前記複数個の実測サンプルに対するスペクトル・クラスタリングを行なって、前記複数個の実測サンプルを前記複数個のクラスにクラスタリングするステップとを含み、

前記算出するステップは、

30

前記減算手段が、前記プラントの稼動条件と前記データ項目とにより定義された実測ベクトルのベクトル空間において、前記複数個の実測サンプルにより定められる複数個の実測ベクトルのうち、基準となる実測サンプルに対応する実測ベクトルを他の実測ベクトルから減算するステップと、

前記相関算出手段が、前記減算するステップにおいて減算された後の前記他の実測ベクトルの任意の2個について、共通の線形部分空間上に存在するか否かを基準として、互いの相関を求めるステップと、

前記繰返算出手段が、前記減算するステップにおける減算と、前記相関を求めるステップにおける相関の算出とを、前記基準となる実測サンプルを変えながら前記複数個の実測サンプルの全てに対して行なう繰返算出ステップと、

40

前記繰返算出ステップにおいて算出された、前記複数個の実測サンプルの任意の2個の間で算出される複数個の相関に基づいて、前記行列を決定するステップとを含む、プラント制御情報生成方法。

【発明の詳細な説明】

【技術分野】

【0001】

この発明は、化学プラント、半導体製造設備、製鉄プラント等の制御技術に関し、特に、過去の複数の稼動時の稼動条件とその結果とに基づいて、プラントの制御の精度を高めるための技術に関する。

【背景技術】

50

【0002】

化学プラントでは、複数の装置（反応炉等）を使用して同一の製品を生産することが多い。この場合、生産される製品が均一となるように、各装置の稼働条件を適切に設定することが必要である。しかし、装置内におけるプロセスの進行とともにプロセスの特性が変化すること、及び複数の装置を並列で運転している場合に、それぞれの装置のカタログスペックが同一でも実際には個体差があること等により、生産される製品が均一となるような条件を装置ごとに精度よく決定したり、生産される製品の組成を稼働条件から精度よく予測したりすることはむずかしい。同様の問題は、化学プラントに限らず、半導体製造設備、金属精錬プラント、水処理施設、及びガスプラント等でも生じる。また、何らかのものを製造する施設だけでなく、発電・変電施設のように非有体物の特性を変化させるための施設でもこうした問題は生じ得る。以下、本明細書では、何らかのものを製造する施設だけでなく、非有体物の特性を変化させる施設をも含めて「プラント」と呼ぶ。また、複合的な施設だけでなく、単独の設備であって、稼働条件によって異なる挙動を示すものもここでいう「プラント」に含めるものとする。

10

【0003】

これまで、本願発明者は、プロセスの特性変化が測定パラメータ間の相関関係の違いとして表現されることに着目し、パラメータ間の相関関係を考慮したモデリングがソフトセンサ設計に有効であることを示してきた。本願発明者はさらに、装置間の個体差に対応できる手法として、パラメータ間の相関関係の類似度を考慮して、過去の装置の稼働パラメータをクラスタリングする手法（向井法）と、相関識別法（Nearest Correlation Method；NC法）とを開発した（非特許文献1、2）。

20

【先行技術文献】

【非特許文献】

【0004】

【非特許文献1】向井、藤原、加納、長谷部：変数間の相関に着目したクラスタリング手法の開発およびその多変量統計モデリングへの利用、計測自動制御学会第8回制御部門大会、京都大学（2008）

【非特許文献2】藤原、加納、長谷部：相関関係を考慮したパターン認識手法の開発とソフトセンサ設計への適用、計測自動制御学会第9回制御部門大会、広島大学（2009）

30

【発明の概要】

【発明が解決しようとする課題】

【0005】

向井法は、パラメータ間の相関関係を指標としてデータをクラスタリングできるが、繰返計算を行なう必要がある。一方、NC法は、線形空間におけるクエリとサンプルの幾何学的な関係に着目した手法であり、教師信号を用いることなくクエリと類似の相関関係を有するサンプルを検出できるという特徴を持つが、クラスタリングに用いることはできなかった。

【0006】

ここでいうパラメータ間の相関関係とは、以下のようなものをいう。以下、説明を理解しやすくするために、2つのパラメータのみからなるサンプルを扱う。図1を参照して、互いに相関関係を持つサンプルの集合が3つあるものとする。図1においては、第1の集合は黒丸で、第2の集合は白丸で、第3の集合は「+」記号で、それぞれ表されている。ここでは、各集合を「クラス」と呼ぶ。各クラスに属するサンプルの間の相関関係はクラスごとに異なっている。あるクラスについて新たなサンプルが得られれば、新たなサンプルを使用して、そのクラスに適した制御を行なうことができる。したがって、ある新たなサンプルが得られたときに、そのサンプルを的確なクラスに分類することが必要である。

40

【0007】

通常、このような場合には、予めサンプルの集合をクラスに分類しておき、各クラスとサンプルとの間の「距離」を定義しておく手法が用いられる。新たなサンプルが得られると、そのサンプルと各クラスとの間の「距離」を算出して、最も距離が小さなクラスに新

50

たなサンプルを分類する。

【0008】

得られたサンプルの集合がどのクラスに属するかが明確であれば、上記した手法を用いることができる。しかし実際のプラントの稼動時に得られる実測サンプルがどのクラスに属するかは明らかでない。つまり、実際のプラントから得られるサンプルのクラスを図1のように3つのクラスに分類(クラスタリング)する手法が必要である。

【0009】

クラスタリング手法として最も頻繁に使用されるものにk-平均法がある。図1に示されるサンプルの集合をk-平均法を用いてクラスタリングした結果を図2に示す。図2では、クラスタリングによりサンプルを3つのクラスに分類している。各クラスに分類されたサンプルをそれぞれ黒丸、白丸及び「+」記号で表している。

10

【0010】

図2により明らかのように、k-平均法を用いた場合、サンプルを正確にクラスタリングすることはできない。これはk-平均法が基本的にサンプルの間の距離に基づいてクラスタリングを行なっているためである。すなわち、各クラスのサンプルが集中する中央部付近ではサンプルが全て同じクラスに分類され、その両側に散在するサンプルはそれぞれ別々のクラスに分類される。

【0011】

k-平均法を用いたクラスタリングとは別のクラスタリング手法も存在する。図3は、スペクトラル・クラスタリングと呼ばれるクラスタリング手法を用いて図1に示されるサンプルをクラスタリングした結果を示す。このスペクトラル・クラスタリングの手法についてはその詳細を後述する。図3において、各クラスに属するサンプルの表示方法は図1及び図2と同様である。このスペクトラル・クラスタリングは、後述するようにグラフ理論に基づくものであって、k-平均法と比較すると単純に距離に基づいてクラスタリングするものではない。したがって、クラスタリングをより正確に行なえる可能性がある。しかし、図3に示される例ではクラスタリングが正しく行なわれていない。

20

【0012】

過去の実測データを統計処理することにより、プラントの制御を行なう場合、実測データを正しくクラスタリングできなければ、制御の前提が失われることになり、正しい制御を実現できない。したがって、実測データを正しくクラスタリングすることができる技術が必要である。

30

【0013】

したがって本発明の目的は、サンプル間の相関関係を用いて、スペクトラル・クラスタリング手法によりクラスタリングしたサンプルの集合を用いて、精度よくプラントの制御を行なうことができるプラント制御情報生成装置及び方法、並びにそのためにコンピュータプログラムを提供することである。

【0014】

本発明に係るプラント制御情報生成装置は、プラントの稼動条件と、当該稼動条件でのプラントの過去の稼動結果から得られるデータ項目の実測値との組合せからなる実測サンプルを複数個記憶するためのデータベースと、データベースに記憶された複数個の実測サンプルを、複数個のクラスにクラスタリングするためのクラスタリング手段と、クラスタリング手段によりクラスタリングされた各クラスに対して、プラントにおける稼動条件とデータ項目の値との関係を示す統計的モデルを構築するためのモデル構築手段と、新たな稼動条件が与えられたことに応答して、統計的モデルと、新たな稼動条件とに基づいて、プラントに関する制御情報を算出するための制御情報算出手段とを含む。クラスタリング手段は、複数個の実測サンプルの任意の2個の間の相関の程度を成分とする行列を、データベースに記憶された実測サンプルから算出するための行列算出手段と、行列算出手段により算出された行列を類似度行列として、複数個の実測サンプルに対するスペクトル・クラスタリングを行なって、複数個の実測サンプルを複数個のクラスにクラスタリングするための手段とを含む。

40

50

【 0 0 1 5 】

このプラント制御情報生成装置においては、クラスタリング手段が、データベースに記憶された複数個の実測サンプルを複数個のクラスにクラスタリングし、クラスタリングされた後のクラスごとに、統計的モデルが算出される。クラスタリングの際、複数個の実測サンプルの任意の2個の間の相関の程度を成分とする行列を類似度行列として、スペクトラル・クラスタリングが行なわれる。サンプルの間の相関の相違に基づいて複数個のサンプルがクラスタリングされるため、単純にサンプル間の距離に基づいてクラスタリングされる場合と比較して、異なる相関関係にあるサンプルが正しく別々のクラスにクラスタリングされる確率が高くなる。その結果、正しいクラスタリングが可能になり、それらクラスから得られる統計的モデルは正確なものとなる。こうした処理は、新たな稼動条件が与えられ、新たな制御情報の生成が要求されるまでに行なっておけばよく、複雑な挙動を示すプラントについても、稼動時に、従来よりも正確な制御情報を生成することが可能になる。その結果、サンプル間の相関関係を用いて、スペクトラル・クラスタリング手法によりクラスタリングしたサンプルの集合を用いて、精度よくプラントの制御を行なうことができるプラント制御情報生成装置を提供することができる。

10

【 0 0 1 6 】

好ましくは、行列算出手段は、プラントの稼動条件とデータ項目とにより定義された実測ベクトルのベクトル空間において、複数個の実測サンプルにより定められる複数個の実測ベクトルのうち、基準となる実測サンプルに対応する実測ベクトルを他の実測ベクトルから減算するための減算手段と、減算手段により減算された後の他の実測ベクトルの任意の2個について、共通の線形部分空間上に存在するか否かを基準として、互いの相関を求めるための相関算出手段と、減算手段による減算と、相関算出手段による相関の算出とを、基準となる実測サンプルを変えながら複数個の実測サンプルの全てに対して行なうための繰返算出手段と、繰返算出手段により算出された、複数個の実測サンプルの任意の2個の間で算出される複数個の相関に基づいて、行列を決定するための手段とを含む。

20

【 0 0 1 7 】

このように、実測サンプルの各々を基準ベクトルとして他の実測ベクトルから減算することにより、そのベクトルは実測ベクトル空間の原点に移動し、それに伴って他の実測ベクトルも平行移動する。その結果、基準実測ベクトルを通る平面は実測ベクトル空間の線形部分空間となり、その上に存在する実測ベクトルが互いにある相関関係を持つと判定できる。こうした判定を全ての実測ベクトルに対して行ない、さらに基準ベクトルとして全ての実測ベクトルを用いてこの処理を繰返すことで、実測ベクトルの間の相関関係がより明確となる。その結果、そうした相関に基づいて算出される類似度行列が、実測ベクトルの間の相関関係の相違を明確に表すものとなって、スペクトラル・クラスタリングが正確に行なえる。

30

【 0 0 1 8 】

より好ましくは、行列を決定するための手段は、繰返算出手段により、複数個の実測サンプルの任意の2個の間で算出された複数個の相関を合計した値を成分とする行列を行列として出力するための手段を含む。

【 0 0 1 9 】

行列の各成分について、単純に複数個の相関を加算することで、複数個の実測サンプルの任意の2個の間の相関の程度を行列として表現できる。必要な計算量が少なく済み、処理を早めてリアルタイムで実行することができる。

40

【 0 0 2 0 】

相関算出手段は、減算手段により減算された後の他の実測ベクトルの任意の2個について、共通の線形部分空間上に存在するか否かを基準として、互いの相関係数を求めるための相関係数算出手段と、相関係数算出手段により算出された相関係数が所定のしきい値より大きいか否かにしたがって、相関の値を2値化するための手段とを含んでもよい。

【 0 0 2 1 】

2値化することによって、加算処理がより単純なものとなり、計算量をより削減するこ

50

とができる。

【0022】

本発明のある実施の形態では、制御情報算出手段は、新たな稼動条件が与えられたことに応答して、当該稼動条件が複数個のクラスのいずれに属するかを判定するための手段と、当該クラスに対応する統計的モデルを、新たな稼動条件に適用することにより、新たな稼動条件の下でのプラントにより生成される物質の属性を予測し、予測情報を出力するための手段とを含む。

【0023】

本発明の別の実施の形態では、プラント制御情報生成装置はプラントの稼動状態を測定するための測定手段をさらに含む。制御情報算出手段は、測定手段による測定結果が新たな稼動条件として与えられたことに応答して、当該稼動条件と複数個のクラスとを比較することにより、プラントの状態が異常か否かを判定するための異常判定手段を含む。

【0024】

好ましくは、異常判定手段は、測定手段による測定結果が新たな稼動条件として与えられたことに応答して、複数個のクラスの統計的モデルのうち、当該稼動条件が適合するものを判定するための適合モデル判定手段と、適合モデル判定手段により、新たな稼動条件が複数個のクラスの統計的モデルのいずれにも適合しないと判定されたことに応答して、プラントの異常を示す信号を生成し出力するための手段とを含む。

【0025】

本発明の第2の局面に係るコンピュータプログラムは、コンピュータにより実行されると、当該コンピュータを、プラントの稼動条件と、当該稼動条件でのプラントの過去の稼動結果から得られるデータ項目の実測値との組合せからなる実測サンプルを複数個記憶するためのデータベースと、データベースに記憶された複数個の実測サンプルを、複数個のクラスにクラスタリングするためのクラスタリング手段と、クラスタリング手段によりクラスタリングされた各クラスに対して、プラントにおける稼動条件とデータ項目の値との関係を示す統計的モデルを構築するためのモデル構築手段と、新たな稼動条件が与えられたことに応答して、統計的モデルと、新たな稼動条件とに基づいて、プラントに関する制御情報を算出するための制御情報算出手段として機能させる。このプログラムは、クラスタリング手段として、コンピュータを、複数個の実測サンプルの任意の2個の間の相関の程度を成分とする行列を、データベースに記憶された実測サンプルから算出するための行列算出手段と、行列算出手段により算出された行列を類似度行列として、複数個の実測サンプルに対するスペクトル・クラスタリングを行なって、複数個の実測サンプルを複数個のクラスにクラスタリングするための手段として機能させる。

【0026】

本発明の第3の局面に係るプラント制御情報生成方法は、プラントの稼動条件と、当該稼動条件でのプラントの過去の稼動結果から得られるデータ項目の実測値との組合せからなる実測サンプルを複数個記憶するためのデータベースと、データベースに記憶された複数個の実測サンプルを、複数個のクラスにクラスタリングするためのクラスタリング手段と、クラスタリング手段によりクラスタリングされた各クラスに対して、プラントにおける稼動条件とデータ項目の値との関係を示す統計的モデルを構築するためのモデル構築手段と、新たな稼動条件が与えられたことに応答して、統計的モデルと、新たな稼動条件とに基づいて、プラントに関する制御情報を算出するための制御情報算出手段とを含む、プラント制御情報生成装置におけるプラント制御情報生成方法である。クラスタリング手段は、複数個の実測サンプルの任意の2個の間の相関の程度を成分とする行列を、データベースに記憶された実測サンプルから算出するための行列算出手段と、行列算出手段により算出された行列を類似度行列として、複数個の実測サンプルに対するスペクトル・クラスタリングを行なって、複数個の実測サンプルを複数個のクラスにクラスタリングするための手段とを含む。このプラント制御情報生成方法は、クラスタリング手段が、データベースに記憶された複数個の実測サンプルを、複数個のクラスにクラスタリングするステップと、モデル構築手段が、クラスタリングするステップにおいてクラスタリングされた各ク

10

20

30

40

50

ラスに対して、プラントにおける稼動条件とデータ項目の値との関係を示す統計的モデルを構築するステップと、制御情報算出手段が新たな稼動条件が与えられたことに応答して、統計的モデルと、新たな稼動条件とに基づいて、プラントに関する制御情報を算出するステップとを含む。クラスタリングするステップは、行列算出手段が、複数個の実測サンプルの任意の2個の間の相関の程度を成分とする行列を、データベースに記憶された実測サンプルから算出するステップと、クラスタリングするための手段が、算出するステップにおいて算出された行列を類似度行列として、複数個の実測サンプルに対するスペクトル・クラスタリングを行なって、複数個の実測サンプルを複数個のクラスにクラスタリングするステップとを含む。

【図面の簡単な説明】

10

【0027】

【図1】図1は、相関関係の違いに基づき実測サンプルを真のクラスにクラスタリングした結果を示すグラフである。

【図2】図2は、図1に示す実測サンプルを対象にk-平均法によってクラスタリングした結果を示すグラフである。

【図3】図3は、図1に示す実測サンプルを対象に従来のスペクトル・クラスタリング法によってクラスタリングした結果を示すグラフである。

【図4】図4は、8個のノードを持つ無向グラフの例を示す図である。

【図5】図5は、アフィン部分空間P内におけるサンプル間の相関関係を説明するための模式図である。

20

【図6】図6は、図5に示す変数群を、サンプルx1が原点と一致するように平行移動した後のサンプルの分布を説明するための模式図である。

【図7】図7は、図5に示す変数群を、サンプルx3が原点と一致するように平行移動した後のサンプルの分布を説明するための模式図である。

【図8】変数間の相関関係を指標として、スペクトラル・クラスタリングのための類似度行列を求める、本発明の位置実施の形態に係る装置で採用したアルゴリズムを説明するためのフローチャートである。

【図9】図9は、図1に示す測定サンプルを対象に本願発明の一実施の形態によってクラスタリングした結果を示すグラフである。

【図10】図10は、本発明の一実施の形態に係るバッチプロセスの概略構成図を示す模式図である。

30

【図11】図11は、図10に示すバッチプロセスにおいて、製品の分子量を推定する、本願発明の一実施の形態に係るソフトセンサを実現するためのコンピュータのブロック図である。

【図12】図12は、本願発明の一実施の形態において、実測サンプルをクラスタリングするためのプログラムの制御構造を示すフローチャートである。

【図13】図13は、本願発明の一実施の形態において、新たなサンプルにしたがって、バッチプロセスで産生される製品の分子量を推定するプログラムの制御構造を示すフローチャートである。

【図14】図14は、バッチプロセスにおける反応炉温度とジャケット温度との時系列変化を示すグラフである。

40

【図15】図15は、k-平均法によるクラスタリング結果を示すグラフである。

【図16】図16は、本発明の一実施の形態に係るシステムにおけるクラスタリング結果を示すグラフである。

【図17】図17は、k-平均法及び本発明の第1の実施の形態に係るソフトセンサの推定性能を示すグラフである。

【図18】図18は、本願発明の第2の実施の形態に係るプロセス異常の検出装置における準備処理を実現するプログラムの制御構造を示すフローチャートである。

【図19】図19は、本願発明の第2の実施の形態に係るプロセス異常の検出装置における、プロセス異常を検出するためのプログラムの制御構造を示すフローチャートである。

50

【発明を実施するための形態】

【0028】

1. 最初に

まず、スペクトラル・クラスタリングの基礎であるグラフ理論を簡単に紹介し、次にスペクトラル・クラスタリングの代表的なアルゴリズムを述べる。

【0029】

1.1 グラフ理論

電車での乗換について考える。乗換案内図を見ると、駅間の距離及び配置、並びに路線の形状は、実際の地理とは異なって表現されている。これは、乗換を考える場合、駅と駅とがどのように路線で結ばれているかが問題であって、線路が具体的にどこを通っているか、又はその形状がどのようなものであるかは本質的ではないためである。つまり、電車の乗換では、駅と駅とのつながり方が重要であって、その経路がどのような形状かは問題ではない。その結果、乗換案内図では、駅は点で、駅間の路線は簡略化された単なる線として描かれることが多い。このように、点のつながり方を、点とそれらを結ぶ線とからなる構造で抽象化したものをグラフと呼ぶ。グラフが持つ様々な性質を探求するのがグラフ理論である。

10

【0030】

グラフ理論の始まりは一筆書きの問題にある。オイラーはあるとき、ケーニヒスベルグという町で「町の中心を流れるプレーゲル川に架かっている7つの橋を2度通らずに、全て渡って元の所に帰ってくるができるか。ただし、どこから出発してもよい。」という問題を考えた。これをケーニヒスベルグの問題という。オイラーはこの問題を解決するために、川と橋との位置関係をグラフとして表現し、このグラフが一筆書き可能であれば、ケーニヒスベルグの橋を全て1度ずつ通って戻ってくるルートが存在することになると考えた。そして彼は、このグラフが一筆書きできないことを証明し、ケーニヒスベルグの問題を否定的に解決した。一筆書き可能なグラフをオイラーグラフと呼ぶ。

20

【0031】

数学的なグラフの定義を与える。集合 V 、 E と、 E の元に2つの V の元を対応させる写像 $f: E \rightarrow V \times V$ があるとき、 $G = (V, E)$ を有向グラフという。 V の元を G のノード(頂点)、 E の元をエッジ(辺)と呼ぶ。一方、 $P(V)$ を V のべき集合とし、 E の元に $P(V)$ の部分集合を対応させる写像 $g: E \rightarrow P(V)$ があって、 $e \in E$ に対し $g(e) = \{v_1, v_2\}$ ($v_1, v_2 \in V$) であるとき、 $G = (V, E)$ を無向グラフと呼ぶ。 $g(e) = \{v_1, v_2\}$ を満たす v_1 と v_2 のペアを「互いに隣接している」といい、 $v_1 \sim v_2$ で表す。グラフによっては、エッジに重みがついている場合がある。これを重み付きグラフという。8個のノードを持つ無向グラフの例を、図4に示す。

30

【0032】

図4に示すグラフでは、各ノードに1~8のインデックスを付してあり、各ノードをインデックスで呼ぶものとする。すなわち、 $V = \{1, 2, 3, 4, 5, 6, 7, 8\}$ 、 $E = \{\{1, 2\}, \{1, 5\}, \{2, 5\}, \{2, 3\}, \{3, 4\}, \{4, 5\}, \{4, 6\}, \{6, 7\}, \{6, 8\}, \{7, 8\}\}$ と表すことができる。

【0033】

さらに、2つのグラフ $G = (V, E)$ と $G' = (V', E')$ とに対して、全単射 $h: V \rightarrow V'$ で $v_1 \sim v_2 \iff h(v_1) \sim h(v_2)$ を満たすものが存在するとき、 G 、 G' は同型であるという。そのような写像 h を同型写像と呼ぶ。同型写像の例としては、ノードのインデックスのみを入れ替える写像がある。

40

【0034】

グラフの表現には行列表現が便利である。特に、グラフは「隣接行列」と呼ばれる行列で表現することが多い。無向グラフ $G = (V, E)$ のノードに重複のないインデックスが与えられているとき、次の式によってこのグラフ G の隣接行列 A を定義する。

【0035】

【数 1】

$$(A)_{v,w} = \begin{cases} 1, & \text{if } v_1 \sim v_2 \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

この定義から明らかなように、無向グラフの場合には隣接行列は対称行列である。有向グラフの場合、ノード v_1 から v_2 に向かうエッジが存在するときに行列の $(A)_{v_1, v_2} = 1$ 、そうでないときは $(A)_{v_1, v_2} = 0$ とすればよい。また、重み付きグラフの場合は $(A)_{v_1, v_2}$ にノード v_1 から v_2 に向かうエッジに与えられた重みを代入する。なお、隣接行列の対角成分には 0 を割り当てることが多い。たとえば、図 4 で表現される無向グラフの隣接行列 $A \in \mathbb{R}^{8 \times 8}$ は、次式のように表現できる。

【0036】

【数 2】

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \quad (2)$$

10

20

無向グラフの隣接行列は対称行列であるからその固有値は全て実数であり、固有ベクトルは互いに直交する。無向グラフ G の隣接行列の固有値を $\lambda_1 < \lambda_2 < \dots < \lambda_N$ とし、それらの重複度を m_1, \dots, m_N として得られる以下の表をグラフ G のスペクトルと呼ぶ。

【0037】

【数 3】

$$\begin{array}{cccc} \lambda_1 & \lambda_2 & \dots & \lambda_N \\ m_1 & m_2 & \dots & m_N \end{array} \quad (3) \quad 30$$

同型なグラフのスペクトルは一致するという性質があり（ただし、逆は成り立たない）、スペクトルはグラフの性質の分類に使われる。

【0038】

グラフで表現できる対象は、乗換案内図及び一筆書きだけではなく数多く存在する。たとえば、電気回路及び Web のハイパーリンクの構造はグラフで表現できる。計算機の手帳構造は木構造（ツリー）で表現されるが、ツリーもグラフの一種である。巡回セールスマン問題、及び四色問題に代表される彩色問題等、グラフ理論では組合せ最適化問題として定式化される問題が多く、アルゴリズム論の分野で活発に研究されている。

【0039】

1.2 Max - Min cut 法

スペクトラル・クラスタリングは、重み付きグラフのグラフ分割法であり、いくつかのエッジをカットしてグラフをサブグラフに分割することで、ノードのクラスタリングを行なう。たとえば図 4 のグラフの場合、 $v_4 - v_6$ 間のエッジをカットすることで、 $G_1 = \{v_1, \dots, v_5\}$ 、 $G_2 = \{v_6, v_7, v_8\}$ の 2 つの部分グラフに分割でき、 $V = \{v_1, \dots, v_8\}$ が $V_1 = \{v_1, \dots, v_5\}$ と $V_2 = \{v_6, v_7, v_8\}$ とにクラスタリングされる。スペクトラル・クラスタリングにはいくつかのアルゴリズムがあるが、ここではグラフを 2 つのサブグラフに分割する方法である Max - Min Cut (Mcut) 法について述べる。

【0040】

40

50

いま、重み付きグラフ $G(V, E)$ とその隣接行列 W が与えられたとし、グラフ G を 2 つのサブグラフ A, B に分割する問題を考える。サブグラフ A, B 間の類似度 $\text{cut}(A, B)$ を次の式 (4) により定義する。

【0041】
【数4】

$$\text{cut}(A, B) \equiv W(A, B) \quad (4)$$

ただし

【0042】
【数5】

$$W(A, B) = \sum_{u \in A, v \in B} W_{u,v}, \quad W(A) \equiv W(A, A) \quad (5)$$

である。すなわち、サブグラフ間の類似度 $\text{cut}(A, B)$ とは、 A, B 間のエッジの重みの総和である。Mcut法は $\text{cut}(A, B)$ を最小、かつ $W(A)$ 及び $W(B)$ を最大化するようなサブグラフ A, B を探索する方法である。Mcut法の目的関数 J は次式で定義される。

【0043】
【数6】

$$\min J = \frac{\text{cut}(A, B)}{W(A)} + \frac{\text{cut}(A, B)}{W(B)} \quad (6)$$

グラフ $G(V, E)$ のノードのインデックスの与え方は任意であるから、隣接行列 W を以下のように定義しても一般性は失わない。

【0044】
【数7】

$$W = \begin{bmatrix} W_A & W_{A,B} \\ W_{B,A} & W_B \end{bmatrix} \quad (7)$$

ここで、 W_A, W_B はサブグラフ A, B 内の隣接行列であり、 $W_{A,B}, W_{B,A}$ は A, B 間の隣接行列である。隣接行列の定義より、 $W^T = W_{B,A}, W_{A,B}$ である。さらに、サブクラスへの分割を表現するベクトル $x = [1, \dots, 1, 0, \dots, 0]$ 、 $y = [0, \dots, 0, 1, \dots, 1]$ が次式を満たすようにこれらを定義する。

【0045】
【数8】

$$W(A, B) = x^T(D - W)x = y^T(D - W)y \quad (8)$$

$$W(A) = x^T W x \quad (9)$$

$$W(B) = y^T W y \quad (10)$$

ただし、 $D = \text{diag}(W e)$ 、 $e = [1, \dots, 1]^T$ である。これらの式を用いて、目的関数 (6) 式を書き換えると

【0046】
【数9】

$$\min_{x,y} J = \frac{x^T(D - W)x}{x^T W x} + \frac{y^T(D - W)y}{y^T W y} \quad (11)$$

が得られる。また、 $x^T D y = 0$ 、 $x^T W x > 0$ 、 $y^T W y > 0$ であることがわかる。 J の各項は x, y について対称であるから、以降、第1項のみについて考える。

【0047】

10

20

30

40

50

目的関数(11)式の目的変数 x, y はバイナリ変数であるから、この最小化問題はNP困難である。そこで、(11)式のバイナリ変数を緩和する。ここでは、ノードのインデックスに対応する指標ベクトル q を導入する。 q の、ノード u に対応する要素を $q_u = \{a, -b\}$ ($a, b > 0$)とする。 $q_u = a$ のとき $u \in A$ 、 $q_u = -b$ のとき $u \in B$ である。目的関数は次の(12)式のように書き直される。

【0048】
【数10】

$$\min_q J_q = \frac{q^T(D-W)q}{q^T W q} \quad (12)$$

10

さらに(12)式を次のようにスケールリングする。

【0049】
【数11】

$$\min_{\tilde{q}} \tilde{J}_q = \frac{\tilde{q}^T(I-\tilde{W})\tilde{q}}{\tilde{q}^T \tilde{q}} \quad (13)$$

ここで、 $\tilde{W} = D^{-1/2} W D^{-1/2}$ 、 $\tilde{q} = D^{1/2} q / |D|^{1/2}$ である(各変数の前の「 $\tilde{\cdot}$ 」は、式中では変数の上に記載されている)。また、 $\tilde{q}^T \tilde{W} \tilde{q} > 0$ である。 \tilde{W} の要素 $(\tilde{W})_{i,j}$ は $1 - (\tilde{W})_{i,j} \geq 0$ 、各列の和は1であるから、 \tilde{W} は確率推移行列とみなすことができる。

20

【0050】

さらに目的関数 \tilde{J}_q において、 $P = I - \tilde{W}$ とおくと \tilde{J}_q はPのRayleigh商で表現できることがわかる。Rayleighの原理によると、行列Aの商 $R(x)$ は最小固有値 λ_1 に対応する固有ベクトル x_1 によって最小となり、その最小値は λ_1 である(本明細書では、固有値は $\lambda_1 \dots \lambda_N$ のように並べるものとする)。よって、緩和した最小化問題はPの固有値問題に帰着することがわかる。Pのある固有ベクトルを $D^{1/2} e$ と仮定する。Wの第 i 行を w_i とすると、 $(D)_{i,i} = d_i = w_i e$ であるから

【0051】
【数12】

$$\begin{aligned} P D^{1/2} e &= I D^{1/2} e - \tilde{W} D^{1/2} e \\ &= D^{1/2} e - D^{-1/2} W D^{-1/2} D^{1/2} e \\ &= D^{1/2} e - D^{-1/2} W e \\ &= D^{1/2} e - D^{-1/2} D e \\ &= 0 \end{aligned} \quad (14)$$

30

であり、確かにPのある固有ベクトルは $D^{1/2} e$ で、その固有値は0であることが示される。ところで、 \tilde{W} は確率推移行列とみなすことができたから、その最大固有値は1である。よって、 $x^T P x = x^T (I - \tilde{W}) x \geq 0$ であるから、Pは半正定値行列である。これより、固有ベクトル $D^{1/2} e$ はPの最小固有値 $\lambda_1 = 0$ に対応する固有ベクトル z_1 であることがわかる。

40

【0052】

したがって、Rayleighの原理より最小固有値に対応する固有ベクトル z_1 が(13)式を最小とする解である。しかし、この解は全ての要素が正であり、求める q の要素が $q_u = \{a, -b\}$ であるという仮定に反する。したがってこの解は採用できない。次に小さな解はRayleigh商のmax-min定理より、2番目に小さな固有値 λ_2 である。 λ_2 に対応する固有ベクトル z_2 は、Pの対称性より $z_1^T z_2 = 0$ であるから、必ず正負の要素を持つことがわかる。すなわち、この解は最初の仮定に反しない。し

50

たがって、 z_2 に対応する固有ベクトル z_2 が求める最適な \tilde{q}^* である。

【0053】

以上の考察より、最終的な定式化は

【数13】

$$\min_{\tilde{q}} \tilde{J}_q = \frac{\tilde{q}^T (I - D^{-1/2} W D^{-1/2}) \tilde{q}}{\tilde{q}^T \tilde{q}} \quad (15)$$

において、固有値問題

【数14】

$$(I - D^{-1/2} W D^{-1/2}) z = \lambda z \quad (16)$$

10

の2番目の大きさの固有値 λ_2 に対応する固有ベクトル z_2 を選択すればよいことになり、求める q^* は $q^* = D^{-1/2} z_2$ で表される。なお、この λ_2 、 z_2 をそれぞれ F i e l d e r 値、F i e l d e r ベクトルと呼ぶ。

【0054】

M c u t 法は任意のグラフを2つのサブグラフに分割する方法である。したがって、グラフを任意の個数のサブグラフに分割したいときは、分割後のサブグラフに対して上記手順を繰返せばよい。

【0055】

20

理論的な考察により、緩和問題の目的関数が最小となる時、(6)式においてほぼ $W(A) = W(B)$ が成立することが示されている。すなわち、M c u t 法ではサブグラフ内の類似度がバランスすると期待できる。他のスペクトラル・クラスタリングのアルゴリズムでは必ずしもサブグラフ内の類似度がバランスしないため、このことが M c u t 法の利点とされる。

【0056】

1.3 その他のアルゴリズム

前節で紹介した M c u t 法に類似の手法として、N o r m a l i z e d C u t (N c u t) 法がある。N c u t 法の評価関数は次の(17)式で表される。

【0057】

30

【数15】

$$Ncut = \frac{Cut(A, B)}{deg(A, V)} + \frac{Cut(A, B)}{deg(B, V)} \quad (17)$$

ここで、 $V = A \cup B$ 、 $deg(A, V) = \sum_u \sum_v W_{u, v}$ であり、 $deg(B, V)$ も同様に定義される。解析の詳細は M c u t 法とほぼ同等である。

【0058】

M c u t / N c u t 法では類似度の与え方は任意であり、類似度行列が無向グラフの隣接行列としての性質を満たしていればよい。

【0059】

40

一方、類似度行列の構成に G a u s s i a n カーネルを用いることが従来の一手法として提案されている。この手法では、サンプル s_i, s_j の類似度 $(W)_{i, j}$ として

【0060】

【数16】

$$(W)_{i, j} = \exp\left(\frac{-d^2(s_i - s_j)}{2\sigma^2}\right) \quad (18)$$

を用いる。ここで $d(\cdot, \cdot)$ は距離関数、 σ は調整パラメータである。この手法に関連してさらに、グラフの2分割を繰返す M c u t 法と異なり、スケール後の類似度行列の固有空間にてサンプルの集合を直接任意の個数のクラスタに分割する手法が提案されて

50

いる。この場合、カーネルを用いて非線形写像を行なった後に固有値問題を解いていることになるから、カーネル主成分分析（主成分分析を以下「PCA」と呼ぶ。）と類似の手法であるとされる。この手法をNg - Jordan - Weiss (NJW) アルゴリズムと呼ぶ。図3に示したのはNJWアルゴリズムを用いたスペクトル・クラスタリングの結果である。

【0061】

NJWアルゴリズムの(18)式の決め方において、ローカルスケーリングという方法が提案されている。 $(W)_{i,j}$ は次のように書き換えられる。

【0062】

【数17】

$$(\hat{W})_{i,j} = \exp\left(\frac{-d^2(s_i - s_j)}{\sigma_i \sigma_j}\right) \quad (19)$$

ここで、 $d_i = d(S_i, S_K)$ であり、 S_K は S_i の K 番目に近傍のサンプルを示す。 K は調整パラメータであるが、数値実験では $K = 7$ で概ね良好な結果が得られたと報告されている。

【0063】

1.4 アプリケーション

既に述べたように、グラフとして表現できる構造は多いため、スペクトラル・クラスタリングも応用は広い。たとえば、Webのハイパーリンク及びソーシャルネットワークの構造解析、CAD等のメッシュ分割、タンパク質ネットワークの解析、並びにコンピュータビジョン等にスペクトル・クラスタリングが用いられている。各問題に適した類似度の与え方に工夫がある。

【0064】

2 変数間の相関関係を指標とした類似度行列

前節で紹介したCut法では、類似度の定義は任意であった。一方、NJWアルゴリズムでは、類似度として、サンプル間の距離をカーネルによって写像した値を用いている。しかし、いずれにせよ従来法では変数間の相関関係を指標としてクラスタリングを行うことは考慮されていない。そこで本実施の形態では、スペクトラル・クラスタリングにおいて変数間の相関関係を考慮した類似度行列の構成法を実現する。本実施の形態に係る方法は、変数間の相関関係を考慮した類似度行列を構築し、スペクトラル・クラスタリングによってクラスタリングを実施する方法である。この方法を本明細書ではNC - スペクトラル・クラスタリングと呼ぶことにする。

【0065】

本実施の形態で採用したクラスタリング方法のコンセプトを述べる。なお、以下では、説明を理解しやすくするために、サンプルが3次元空間内に存在するものとする。図5を参照して、この3次元空間内のアフィン部分空間 P が変数間の相関関係を表しており、 P 上のサンプルは全て同一の相関関係にしたがっているものとする。すなわち、 P 上のサンプル x_1, \dots, x_5 は同一の相関関係にしたがうサンプルである。一方、 P 上には存在しないサンプル x_6 及び x_7 は異なった相関関係を有している。

【0066】

いま、サンプル x_1 に着目する。まず、サンプル x_1 が原点となるように空間全体を平行移動させる。つまり、全サンプル x_i ($i = 1, \dots, 7$) から x_1 を減算する。この操作によってアフィン部分空間 P は原点を含むことになるため線形部分空間の定義を満たす。この線形部分空間を V とする。次に図6に示すように、任意のサンプルと原点を結ぶ直線を引き、この直線上で別のサンプルが発見できたとする。図6に示す例では、サンプル x_2 及び x_5 並びにサンプル x_3 及び x_4 がこのような関係を満たしている。このとき、これらのサンプルのペアの相関係数の絶対値は1である。このような関係が成り立つとき、「 $x_2 \sim x_5$ 」のように書く。一方、部分空間 V の要素ではないサンプル x_6 及び x_7 の相関係数の絶対値は1未満である。それゆえ、相関係数が ± 1 であるペアのサンプルは

10

20

30

40

50

、同一の相関関係を有していると判定できる。

【0067】

実際は相関係数が厳密に ± 1 になるペアは存在しないため、しきい値 ($0 < \quad 1$) を用いて相関関係を2値化して、同一の相関関係を有しているかを判定する。すなわち、相関係数の絶対値がしきい値以上となるペアを、同一の相関関係を有していると判定する。本実施の形態では、サンプルの幾何学的な配置のみを考慮したここまでの手続きをNC法と呼ぶ。NC法自体は前述したとおり公知である。

【0068】

ここで、各サンプルをノードとし、サンプル間が同一の相関関係を有しているか否かをそれらのサンプルに対応するノードが隣接しているかとして表現した無向グラフ及びその隣接行列 S_1 $R^{7 \times 7}$ を考える。この例では、 $x_2 \sim x_5$ 及び $x_3 \sim x_4$ であるから、隣接行列 S_1 は、 $(S_1)_{2,5} = (S_1)_{5,2} = 1$ 及び $(S_1)_{3,4} = (S_1)_{4,3} = 1$ とし、それ以外の要素を0とする行列として表現できる。

【0069】

次に、サンプル x_2 が空間の原点と一致するように平面を平行移動させて同様の手順を行ない、隣接行列 S_2 を求める。これを全てのサンプルについて繰返し、最終的に S_1, \dots, S_7 を求める。

【0070】

この手続きによって隣接行列が7つ得られる。これら7つのグラフは、同一のノードの集合を持ち、エッジの集合の部分集合が異なる7つの無向グラフと考えることができる。すなわち、 $G_1(V, E(G_1)), \dots, G_7(V, E(G_7))$ である。ここで、これらのグラフの和のグラフ $G = G_1 \dots G_7$ を考えると、グラフ G のエッジの集合はグラフ G_1 から G_7 のエッジの和集合 $E(G_1) \dots E(G_7)$ であることがわかる。これは隣接行列 S_i ($i = 1, \dots, 7$) のブール和 (各行列の対応する要素の論理和を要素とする行列) を求めることに等しいから、グラフ G の隣接行列は $S = S_1 + \dots + S_7$ となる。ただし、ここで+は行列のブール和を表す。以上の手続きより、サンプル間の相関関係を表現したグラフを得ることができる。

【0071】

しかし、サンプルの分布と原点の選び方によっては、異なった相関関係を持つサンプルのペアであっても、偶々同一の相関関係を有すると判断されてしまうことがある。たとえば図5及び図6に示した例について考える。図7のようにサンプル x_3 が原点と一致するように平面を平行移動させると、 $x_1 \sim x_3$ 以外に $x_6 \sim x_7$ と判定されてしまう (サンプル x_5 と原点とを結んだ直線上にサンプル x_7 が存在する。)。つまり、グラフ G では真に同一の相関関係を有するサンプル間のエッジと、偶々同一の相関関係を有すると判断されてしまったサンプル間のエッジとが区別できない。

【0072】

そこで、上記手続きにおいて、それぞれのサンプルのペアで相関関係を有すると判定された回数をカウントし、それに応じてエッジに重み付けを行なう。このグラフを G とする。つまり、隣接行列のブール和をとるのではなく、通常の和を用いてサンプル間の相関関係の大きさを決定し、 G の隣接行列 S を求める。

【0073】

【数18】

$$S = \sum_{i=1}^7 S_i \quad (20)$$

隣接行列 S_i は対称行列であったからその和も対称行列である。したがって、 S を類似度行列としてスペクトラル・クラスタリングを行なうことで、変数間の相関関係を指標としたクラスタリングを実施できることになる。

【0074】

図8を参照して、本実施の形態に係るNC-スペクトル・クラスタリングを実現するコ

10

20

30

40

50

コンピュータプログラムの制御構造について説明する。いま、サンプル $x_n \in R^M$ ($n = 1, \dots, N$) がデータベースに保存されており、これより変数間の相関関係を指標として、スペクトラル・クラスタリングのための類似度行列を求めたいとする。

【0075】

図8を参照して、このプログラムは、以下の繰返し処理のため繰返し制御変数 L に1を代入するステップ30と、類似度行列 S のための2次元配列 $S[N, N]$ の領域を確保するステップ32と、配列 $S[N, N]$ の全ての要素をゼロでクリアするステップ34とを含む。

【0076】

このプログラムはさらに、 L 番目 ($1 \leq L \leq N$) のサンプルについて、全サンプルとの間の隣接行列 S_L を保存するための2次元配列 $S_L[N, N]$ の領域を確保するステップ36と、配列 $S_L[N, N]$ の各要素をゼロでクリアするステップ38と、 $L \leq n \leq N$ を満たす全ての n についてサンプル x_n からサンプル x_L を減算して新たなサンプル x_n を算出するステップ42と、ステップ42で算出されたサンプル x_n に関し、全ての x_k と x_1 との組合せ ($1 \leq k, l \leq N$ 、ただし $k \neq l$) について、相関係数 $C_{k, l}$ を算出するステップ46と、ステップ48で算出された相関係数 $C_{k, l}$ のうちで $|C_{k, l}| \geq \theta$ (θ は予め定めるしきい値) を満たす k 及び l を求めるステップ48と、 $S_L[k, l]$ と $S_L[l, k]$ とに1を代入するステップ50と、配列 S の各要素に、配列 S_L の対応する要素の値を加算するステップ52とを含む。

【0077】

このプログラムはさらに、変数 L の値が $N - 1$ より小さいか否かを判定し、判定結果に応じて制御の流れを分岐させるステップ54と、ステップ54における判定結果が YES のときに実行され、変数 L に1を加算して制御をステップ36に戻すステップ56と、ステップ54における判定結果が NO のときに実行され、類似度行列を表す配列として配列 S を出力し、処理を終了するステップ58とを含む。

【0078】

このプログラムは類似度行列を算出するためのものであり、スペクトラル・クラスタリング自体のアルゴリズムは任意である。本実施の形態では NJW アルゴリズムを用いる。

【0079】

3 クラスタリングのシミュレーション例

シミュレーションにより、従来法と本実施の形態に係るスペクトル・クラスタリングによるクラスタリングの識別性能の比較を行なった。本シミュレーション例では、2次元及び5次元のデータをそれぞれ用いて2通りの比較を行なった。

【0080】

3.1 2次元の場合

それぞれ互いに異なる相関関係を有する3クラスのデータを用い、クラスタリングを実施する。各クラスに属するサンプルは次式を用いて生成した。

【0081】

【数19】

$$x_i = sa_i + n \quad (i = 1, 2, 3) \quad (21)$$

ここで $S \sim N(0, 10)$ であり、 $a_i \in R^2$ は係数行列である。また、 $n = [n_1 \ n_2]^T$ 、 $n_i \sim N(0, 0.1)$ である。ただし、 $N(m, \sigma)$ は平均 m 、標準偏差 σ の正規分布にしたがう乱数を表す。係数ベクトルはそれぞれ、 $a_1 = [1 \ 2]^T$ 、 $a_2 = [2 \ 2]^T$ 、 $a_3 = [2 \ 1]^T$ とし、各クラスにおいて100サンプル生成した。

【0082】

本シミュレーション例において、本実施の形態のステップ48で用いるパラメータ (しきい値) θ は、0.999とした。生成したサンプル及び k -平均法、従来のスペクトラル・クラスタリングでのクラスタリング結果が、それぞれ図1～図3に示したものである。本願実施の形態にしたがったクラスタリング結果を図9に示す。

10

20

30

40

50

【 0 0 8 3 】

図 1 ~ 図 3 と、図 9 とを参照して、k - 平均法でクラスタリングした結果 (図 2) 及び従来のスペクトル・クラスタリングでクラスタリングした結果 (図 3) は、いずれも図 1 と比較すると正しいクラスタリングとはいえないことがわかる。一方、本実施の形態によるクラスタリング結果 (図 9) では、原点付近のサンプルの属するクラスは正しく識別できていないが、それ以外の領域では正しく識別できていることがわかる。

【 0 0 8 4 】

3.2 5次元の場合

前節と同様にそれぞれ互いに異なる相関関係を有する 3 クラスのデータを用い、クラスタリングを実施する。各クラスに属するサンプルは次式を用いて生成した。識別率は次式で定義される。

【 0 0 8 5 】

【数 2 0】

$$\text{識別率 [\%]} = \frac{L}{K} \times 100 \quad (22)$$

ここで K は識別すべきクラスのサンプルの数であり、L は K 個のサンプルのうち実際に対象のクラスと同一であったサンプルの数である。各クラスに属するサンプルは次式を用いて生成した。

【 0 0 8 6 】

【数 2 1】

$$x_i = A_i s + n \quad (i = 1, 2, 3) \quad (23)$$

ここで A_i $R^{2 \times 5}$ は係数行列であり、 $S = [s_1 \quad s_2]$ 、 $S_i \sim N(0, 10)$ である。また、 $n = [n_1 \quad n_2 \quad n_3 \quad n_4 \quad n_5]^T$ 、 $n_i \sim N(0, 0.1)$ である。

係数行列 A_i は

【 0 0 8 7 】

【数 2 2】

$$A_1 = \begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 1 \\ 2 & 3 \\ 1 & 3 \end{bmatrix} \quad A_2 = \begin{bmatrix} 3 & 3 \\ 2 & 1 \\ 3 & 1 \\ 3 & 2 \\ 2 & 0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 2 & 1 \\ 3 & 4 \\ 1 & 3 \\ 0 & 4 \\ 3 & 1 \end{bmatrix} \quad (24)$$

とし、各クラスについて 100 サンプル生成した。すなわち、 $K = 100$ である。本シミュレーション例において、本実施の形態のステップ 48 で用いるパラメータ (しきい値) は、0.999 とした。

【 0 0 8 8 】

生成したサンプルに対し、k - 平均法 (KM)、スペクトラル・クラスタリング (SC)、及び本実施の形態に係る NC - スペクトラル・クラスタリング (NC) のクラスタリングにより各クラスについてどの程度の精度で識別できたかをテーブル 1 に示す。テーブル 1 を参照して、NC - スペクトル・クラスタリングによれば、k - 平均法及び従来のスペクトル・クラスタリングのいずれと比較しても、どのクラスについてもより高い精度でクラスを識別できていることが分かる。以上の結果より、本実施の形態に係る NC - スペクトラル・クラスタリングの有効性が示されたといえる。

【 0 0 8 9 】

10

20

30

40

【表 1】

Table 1: Discrimination performance of the 5-dimensional samples

	Discrimination rate [%]		
	Class 1	Class 2	Class 3
KM	43	19	75
SC	40	35	31
NC	85	84	93

4. 第 1 の実施の形態

4.1 概略構成

本実施の形態は、並列化されたバッチプロセスの運転データのクラスタリングを対象とし、クラスタリング結果に基づいてソフトセンサ設計を行なうものである。この実施の形態のシミュレーションで用いたバッチプロセスモデルの詳細を最後に付録として示す。

【0090】

図 10 を参照して、本実施の形態に係るプラント 80 は、各々、内部で逐次並列反応

【0091】

【数 23】



(ただし、A, B は原料であり、C は製品、D は副製品) を行なう 5 台の反応炉 82, 84, ..., 86 と、これら反応炉、及びこれら反応路のジャケットの温度をそれぞれ測定するための図示しない複数個のセンサと、これら複数個のセンサの出力にしたがって、所定のプロセス制御を反応炉ごとに行なう複数のコントローラ 92, 94, ..., 96 と、これら反応炉で過去に行なわれたバッチプロセスの結果から得られた生成物に関する情報、及び各プロセスを開始するときの各反応炉に関する所定のプロセスパラメータ(ここではジャケット温度 T_j と反応炉内温度 T_r とする。)に基づき、新たなバッチプロセスで得られる製品 C の分子量 M_c を推定するための、コンピュータからなる製品分子量推定装置 98 とを含む。本プラントでは、目的生成物は製品 C である。プロセスの目的は副製品 D の生成を抑制し、できるだけ多くの製品 C を得ることである。バッチプロセス運転開始前に原料 A 及び B が室温で反応炉に仕込まれる。

【0092】

反応 1 と反応 2 との進行は、いずれも温度によって変化するものとする。条件によっては、反応 1 により生成した製品 C が、反応 2 によって消費されて副製品 D が精製してしまう。したがって、反応炉内の温度を調整する必要がある。プラント 80 はその目的のため、冷媒と熱媒とを用いたジャケット温度調節部 100 を有している。

【0093】

製品分子量推定装置 98 は、実質的にコンピュータからなり、反応炉 82, 84, ..., 86 による過去のバッチプロセスから得られたデータ(反応炉温度とジャケット温度、及び最終的に得られた製品 C の分子量 M_c) を記憶し、これらをそれぞれ適切なクラスに分類し、各クラス別に、各バッチの開始時のデータと、バッチプロセスにより得られた製品 C の分子量とに関する統計的モデルをパラメータの値から構築する機能を持つ。製品分子量推定装置 98 はさらに、反応炉温度とジャケット温度とからなる新たなデータが与えられると、それを適切なクラスに分類し、そのクラスの統計的モデルを用いて製品 C の分子量を推定する。

【0094】

4.2 コンピュータによる実現

上述の実施の形態は、コンピュータシステムハードウェア及びコンピュータシステム上で実行されるプログラムによって実現され得る。図 11 はこの実施の形態で用いられる製品分子量推定装置 98 のブロック図である。なお、ここで示す製品分子量推定装置 98 は

単なる例であって、他の構成も利用可能である。

【0095】

図11を参照して、製品分子量推定装置98は、コンピュータ120と、全てコンピュータ120に接続された、モニタ122と、キーボード126と、マウス128と、プリンタ124と、を含む。さらに、コンピュータ120はDVD-ROM(Digital Versatile Disk Read-Only-Memory: デジタル多用途ディスク読出専用メモリ)ドライブ150と、半導体メモリポート152とを含む。

【0096】

図11を参照して、コンピュータ120はさらに、DVD-ROMドライブ150と半導体メモリポート152とに接続されたバス142と、全てバス142に接続された、CPU140と、コンピュータ120のブートアッププログラムを記憶するROM144と、CPU140によって使用される作業領域を提供するとともにCPU140によって実行されるプログラムのための記憶領域となるRAM146と、音声データ、音響モデル、言語モデル、レキシコン、及びマッピングテーブルを記憶するためのハードディスクドライブ148と、ネットワーク164への接続を提供するネットワークインターフェイス154とを含む。

【0097】

上述の実施の形態に係る製品分子量推定装置98を実現するソフトウェアは、DVD-ROM162又は半導体メモリ160等の媒体に記録されたオブジェクトコード又はスクリプトの形で流通し、DVD-ROMドライブ150又は半導体メモリポート152等の読出装置を介してコンピュータ120に提供され、ハードディスクドライブ148に記憶される。CPU140がプログラムを実行する際には、プログラムはハードディスクドライブ148から読出されてRAM146にロードされる。図示しないプログラムカウンタによって指定されたアドレスから命令がフェッチされ、その命令が実行される。CPU140はハードディスクドライブ148から処理すべきデータを読出し、処理の結果をこれもまたハードディスクドライブ148に記憶する。推定された製品分子量のハードコピーが、プリンタ124により出力される。

【0098】

コンピュータ120の一般的な動作は周知であるので、詳細な説明は省略する。

【0099】

ソフトウェアの流通の方法に関して、ソフトウェアは必ずしも記憶媒体上に固定されたものでなくてもよい。たとえば、ソフトウェアはネットワークに接続された別のコンピュータから分配されてもよい。ソフトウェアの一部がハードディスクドライブ148に記憶され、ソフトウェアの残りの部分についてはネットワークを介してハードディスクドライブ148に取込み、実行の際に統合する様にしてもよい。

【0100】

また、ソフトウェアの流通形態はオブジェクトコードには限らない。前述したようにスクリプト形式でもよいし、ソースプログラムの形で供給され、コンピュータ120にインストールされた適切なコンパイラでオブジェクトコードに変換されるという流通形態もあり得る。

【0101】

典型的には、現代のコンピュータはコンピュータのオペレーティングシステム(OS)によって提供される一般的な機能を利用し、所望の目的にしたがって制御された態様で機能を達成する。したがって、OS又はサードパーティから提供されうる一般的な機能を含まず、一般的な機能の実行順序の組合せのみを指定したプログラムであっても、そのプログラムが全体として所望の目的を達成する制御構造を有する限り、そのプログラムがこの発明の範囲に包含されることは明らかである。

【0102】

図12及び図13を参照して、コンピュータ120により実行され、製品Cの分子量M_Cを推定するためのプログラムの制御構造について説明する。このプログラムは、バッチ

10

20

30

40

50

プロセスが終了すると、そのバッチプロセスによって得られた実測データを過去の実測データとともにデータベースに保存し、将来の製品分子量推定に備えてクラス別の統計的モデルを構築するための統計的モデル構築プログラム（図 1 2）と、新たなデータが与えられると、構築済のクラス別の統計的モデルのうちで最適な統計的モデルを選択し、その統計的モデルを新たなデータに適用することにより、最終的な製品分子量 M_c を推定するためのプログラム（図 1 3）とを含む。

【 0 1 0 3 】

図 1 2 を参照して、統計的モデル構築プログラムは、バッチプロセスが終了し、そのバッチプロセスによって得られた実測データ（プロセス開始時の反応炉温度、ジャケット温度その他の関連パラメータ、並びに最終的に得られた製品 C の平均分子量）が与えられると、その実測データをコンピュータ 1 2 0（図 1 1）内に構築されたデータベース（以下単に「DB」と呼ぶ。）に追加するステップ 1 8 0 と、ステップ 1 8 0 に続き、DB 内の実測データに対して主成分分析を行なってクラスタリングのための次元数を削減するステップ 1 8 2 とを含む。本実施の形態では、図面を理解しやすくするため、主成分分析により、実測データの各々を第 1 主成分及び第 2 主成分からなる 2 次元のデータに変換するものとする。

【 0 1 0 4 】

このプログラムはさらに、ステップ 1 8 2 に続き、ステップ 1 8 2 により得られた実測データに対して、図 8 を参照して説明した NC - スペクトラル・クラスタリングのためのプログラムを適用することにより、実測データを複数のクラスに分類するステップ 1 8 4 と、ステップ 1 8 4 により得られたクラスごとに、予め選択された構築方法により統計的モデルを構築するステップ 1 8 6 と、ステップ 1 8 6 で得られた各クラスの統計的モデルを HDD 1 4 8 に保存して処理を終了するステップ 1 8 8 とを含む。

【 0 1 0 5 】

一方、図 1 3 を参照して、新たなサンプル（新たなバッチプロセス開始時の反応炉温度、ジャケット温度その他の関連パラメータ等）が与えられたことに応答して、図 1 2 のステップ 1 8 2 での主成分分析の結果にしたがって主成分スコアを算出するステップ 2 0 0 と、この主成分スコアに基づき、新たなサンプルがどのクラスに属するかを推定するステップ 2 0 2 と、ステップ 2 0 2 で推定されたクラスに対する統計的モデルを HDD 1 4 8 から読み出し、新たなサンプルを当該統計的モデルに適用することで新たなバッチプロセスにより生成する製品 C の分子量を推定して処理を終了するステップ 2 0 4 とを含む。

【 0 1 0 6 】

ステップ 2 0 2 では、たとえば各クラスとの間で求めた Q 統計量を指標としてクラスを推定することが考えられる。

【 0 1 0 7 】

4.3 シミュレーション

本実施の形態について、以下のような条件でその性能についての数値シミュレーションを行ない、従来の方法による結果と比較した。以下の数値シミュレーションでは、バッチプロセス運転開始前に原料 A 及び B が室温 2 0 で反応炉に仕込まれるものとする。このシミュレーションでは、原料 A 及び B の仕込み量はそれぞれ $N(20, 0.1)$ にしたがう乱数として変化するものとする。

【 0 1 0 8 】

反応 1 は 9 0 以上で進行し、反応炉温度が 1 0 0 を超えると反応 2 が急速に進行するものとする。反応 1 を促進するため、バッチ開始後は速やかに反応炉を室温から 9 0 まで加熱する必要がある。反応が進むと反応熱によって反応炉内温度がさらに上昇するが、反応炉内温度が 1 0 0 を超えると反応 2 が急速に進行し、製品 C が消費され、副製品 D が生成してしまう。したがって、反応炉内温度が 9 0 付近まで上昇した後は、ジャケット温度調節部 1 0 0 により、反応炉を冷却し反応炉内温度の上昇を抑制する。シミュレーションでは、反応炉内温度制御には、デュアルモード制御が用いられるものとした。

【 0 1 0 9 】

本プロセスにおいて運転中に測定されるのは、ジャケット温度 T_j と反応炉内温度 T_r であり、それぞれサンプリング周期は1分である。バッチ終了時間は120分で固定する。あるバッチの運転データを図14に示す。

【0110】

本シミュレーションでは5台の反応炉82, 84, ..., 86が並列に運転されている。バッチ終了時にはこれら反応炉のメンテナンスが行なわれるため、バッチ毎に反応炉の伝熱係数が変化する。反応炉それぞれの伝熱係数 U_i ($i = 1, \dots, 5$) は、 $U_1 = U_2 = U(40.60, 40.62)$, $U_3 = U_4 = U(40.57, 40.59)$, $U_5 = U(40.54, 40.56)$ として。ただし、 $U(a, b)$ は閉区間 $[a, b]$ 上の一様乱数であることを示す。すなわち、本プロセスでは5台の反応炉が存在するが、それらの個体差としては3種である。本シミュレーションでは、各反応炉についてプロセス20バッチ、計100バッチのデータがDBに保存されているとし、これらのデータより製品の分子量推定を行なった。

10

【0111】

4.4 クラスタリング

データベースに保存されている100バッチ分の運転データのクラスタリングを、従来法である k -平均法と、本実施の形態に係るNC-スペクトラル・クラスタリングを用いて実施した。シミュレーションでは、図12のステップ182から処理を開始した。すなわち、100データに対してPCAを行なって各データの次元を圧縮した。ここで採用する主成分数は2、主成分スコア t_1 , t_2 を入力としてクラスタリングを実施した。クラス数は3とし、NC-スペクトラル・クラスタリングにおけるしきい値 $\theta = 0.99$ とした。

20

【0112】

クラスタリング結果を図15及び図16に示す。ここでは $t_1 - t_2$ 平面上でのサンプルの分布を示している。 k -平均法ではさらに、各クラスを中心を印で示している。この結果より、 k -平均法ではクラス中心からの距離に基づいてサンプルを分類しているが、本実施の形態では各クラスは相関関係の違いを表現していることがわかる。

【0113】

4.5 ソフトセンサ設計

製品Cの分子量 M_c [kmol] を推定するソフトセンサを設計する。設計手順は図12及び図13に示したとおりであり、図13のステップ204で新たなプロセスでの製品の分子量を推定した。なお、対比のために、 k -平均法を用いたクラスタリングも行ない、その結果を用いて図13に示す方法で製品の分子量を推定した。クラスタリング時のクラス数は3である。

30

【0114】

以下のように、3クラスそれぞれで統計的モデルを構築した。ここで、統計的モデルの入力は反応炉温度 $T_r(k)$ 及びジャケット温度 $T_j(k)$ ($k = 0, \dots, 120$)、出力は製品Cの分子量 M_c である。モデル構築にはPLS(Partial Least Square)回帰を用いた。採用した潜在変数の数はともに3である。新たに測定されたサンプルの識別には、 k -平均法ではクラス中心からの距離、本実施の形態では前述の通り各クラスとの間で求めたQ統計量を指標として用いた。

40

【0115】

本シミュレーションでは、新たな反応炉 R_6 の M_c を推定した。 R_6 の伝熱係数は $U_6 = U(40.60, 40.62)$ とした。検証には20サンプルを用い、 k -平均法及び本実施の形態それぞれを用いたソフトセンサの推定性能を比較した。

【0116】

結果を図17(A)(B)にそれぞれ示す。図中、グラフの横軸が真値、縦軸が予測値を示している。「RMSE」は根平均2乗誤差、「R」は真値と予測値との相関係数である。この結果より、 k -平均法を用いた場合は精度の良い推定ができていないが、本実施の形態を用いた場合は高い推定性能を達成していることが分かる。たとえば、 k -平均法

50

を用いた場合のR M S Eが0.007であるのと比較して、本実施の形態ではR M S Eは0.004となり、R M S Eにして43%改善した。このことより、本実施の形態によれば、k-平均法を用いた場合よりも装置の個体差を正しくクラスタリングできていることがわかる。

【0117】

5. 第2の実施の形態

5.1 構成

第1の実施の形態と同様のシステムハードウェア構成で、プロセス異常を検出する実施の形態について、説明する。

【0118】

図18に、この実施の形態に係るプロセス異常の検出装置を実現するためのプログラムの制御構造をフローチャート形式で示す。このプログラムは、図11に示すCPU140によって実行され、それによって以下に説明するようなプロセス異常の検出装置を実現することができる。

【0119】

図18を参照して、本実施の形態では、上記したNC-スペクトラル・クラスタリングと多変量統計的プロセス管理手法(Multivariate Statistical Process Control; MSPC)を組合せた異常検出を行なう。なお、比較のためにNC-スペクトラル・クラスタリングに代えてk-平均法を用いた場合の結果についても後に示す。以下、k-平均法又はNC-スペクトル・クラスタリングと、MSPCとを組合せた異常検出法をそれぞれKM-MSPC、NC-MSPCと呼ぶことにする。

【0120】

図18を参照して、本実施の形態に係るプロセス異常の検出装置を実現するためのプログラムの制御構造は以下のとおりである。すなわち、このプログラムは、これまでにDBに記憶されていた全サンプルに対してNC-スペクトル・クラスタリングを用いたクラスタリングを実施するステップ230と、ステップ230に続き、全クラスについて T^2 統計量とQ統計量との管理限界 T^2_S 、 Q_S ($S=1, 2, \dots, S$)(各記号の前の「」は、通常は各記号の直上に記載されるべきものである。)とを設定するステップ232とを含む。ただし、Sはクラス数である。

【0121】

さらに、本実施の形態はバッチプロセスを対象とし、サンプルはジャケット温度 T_j と反応炉内温度 T_r である。したがって、得られたデータは時系列データとなる。得られた時系列データを一定のインターバルで分割し、インターバル終了ごとにプロセスに異常が発生しているかを判定する。以下、そうした判定を行なうプログラムの制御構造について説明する。

【0122】

図19を参照して、このプログラムは、所定のインターバル(たとえば10分)が経過するまで待機するステップ250と、10分が経過するごとに、対象となる反応炉の反応炉温度、ジャケット温度等の新たな測定値 x_q をサンプリングするステップ252と、ステップ252でサンプリングされた測定値 x_q と、新たなサンプル x_q と、全てのクラスとの間で T^2 統計量とQ統計量 T_q^2 、 Q_q とを求めるステップ254と、全てのクラスに対して第1の判定式 $T_q^2 < T^2_S$ が成立しているか否かを判定し、判定結果に応じて制御の流れを分岐させるステップ256と、ステップ256で第1の判定式が成立していないと判定されたときに、全てのクラスに対して第2の判定式 $Q_q < Q_S$ が成立しているか否かを判定し、判定結果に応じて制御の流れを分岐させるステップ258とを含む。

【0123】

ステップ256又は258における判定結果がYESのとき、すなわち新たにサンプリングされた測定値 x_q がいずれのクラスにも属していないと判定されたときには、ステッ

10

20

30

40

50

プ 2 6 2 でプロセス異常が発生したと判定され、必要な処理が実行される。ステップ 2 5 6 及び 2 5 8 の双方において判定結果が N O のときには、ステップ 2 6 0 でプロセスは正常であると判定される。この後、制御の流れはステップ 2 5 0 に戻り、さらに 1 0 分が経過するまで待機した後、上記した処理を再び実行する。

【 0 1 2 4 】

5 . 2 シミュレーション

本実施の形態について、プロセスの運転時間を 1 0 分間隔にて 1 2 分割し、1 0 分ごとに異常の判定を実施するシミュレーションを行なった。

【 0 1 2 5 】

本ケーススタディでは、対象とする異常として、ジャケット温度測定センサのドリフトを考慮した。すなわち時間と共にセンサのゼロ点が移動し、センサにて測定される測定値は $T_j(k) = T_{j_0}(k) + T \times (k - k_0)$ のように変化する。ここで T_{j_0} はゼロ点が固定されているときの測定値、 k は測定時刻、 k_0 はドリフト開始時間、 T は単位時間当たりのドリフト量 (K / 分) である。

いま、新たな反応炉 R_6 を対象にドリフト開始時間 $k_0 = 20$ 分とし、 T を変化させ異常を発生させた。対象とするデータは後掲のテーブルに示す 6 ケースであるが、このうちケース 1、ケース 2 は正常データ ($T = 0$) である。本実施の形態に係る装置の目的は、異常発生後できるだけ早く、正確に異常を検出することである。なお、 T^2 統計量、 Q 統計量の管理限界はそれぞれ信頼区間 9 5 % として設定し、採用した主成分数は 5 である。

【 0 1 2 6 】

K M - M S P C 及び提案する N C - M S P C を用いた異常検出結果を表 2 - 表 1 7 に示す。表では、1 1 分 ~ 2 0 分 (第 2 インターバル)、2 1 分 ~ 3 0 分 (第 3 インターバル)、3 1 分 ~ 4 0 分 (第 4 インターバル)、4 1 分 ~ 5 0 分 (第 5 インターバル) におけるそれぞれの管理限界 (C L) と各ケースでの T^2 統計量、 Q 統計量を示す。

【 0 1 2 7 】

この結果より、K M - M S P C、N C - M S P C 双方ともに、未だ異常の発生していない第 2 インターバルの全てのケースと、正常データであるケース 1、ケース 2 については、その後も正しく正常であると判定している。しかし、K M - M S P C では第 3 インターバル、第 4 インターバルでケース 4 の異常を検出できていない。一方、N C - M S P C では第 3 インターバルでケース 4 の異常を正しく検出できている。以上より、N C - M S P C を採用した本実施の形態が K M - M S P C より素早く異常を検出することができる。

【 0 1 2 8 】

10

20

30

【表 2】

Table 2: The T^2 statistics and its control limits of each class with KM-MSPC (Interval 2)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	5.87		0.06	2.23	3.99	17.96	19.57	2.16
Class2	5.39		8.29	11.11	8.34	17.16	14.16	11.06
Class3	3.71		9.32	21.13	1.81	1.85	0.96	20.48
Detection			-	-	-	-	-	-

Table 3: The Q statistics and its control limits of each class with KM-MSPC (Interval 2)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	0.68		0.05	0.00	0.03	0.65	0.52	0.15
Class2	0.35		0.07	0.03	0.00	0.23	0.27	0.02
Class3	0.22		0.00	0.04	0.01	0.08	0.42	0.24
Detection			-	-	-	-	-	-

【 0 1 2 9 】

【表 3】

Table 4: The T^2 statistics and its control limits of each class with KM-MSPC (Interval 3)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	4.07		3.37	800	4477	3141	9916	62.53
Class2	10.05		7.05	2.19	16.34	0.97	2.64	13.36
Class3	5.19		8.97	1.45	0.91	0.82	12.04	1.69
Detection			-	-	-	-	-	-

Table 5: The Q statistics and its control limits of each class with KM-MSPC (Interval 3)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	6.43		1.69	66299	63208	145396	237923	47732
Class2	0.57		369	239	141	525	552	427
Class3	22.77		71.50	1.71	440	20.67	378	129
Detection			-	-	○	-	○	○

【 0 1 3 0 】

10

20

30

【表 4】

Table 6: The T^2 statistics and its control limits of each class with KM-MSPC (Interval 4)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	2.49		6.63	13.36	0.74	0.48	6.90	6.45
Class2	5.72		2.50	6.93	15.18	6.53	5.34	16.89
Class3	6.34		10.57	3.62	44.52	31.45	22.98	40.49
Detection			—	—	—	—	—	○

Table 7: The Q statistics and its control limits of each class with KM-MSPC (Interval 4)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	9.59		30.20	46.95	15.17	7.47	21.56	21.93
Class2	26.46		4.64	15.02	232	91.75	182	494
Class3	24.52		10.89	8.60	376	138	187	720
Detection			—	—	○	—	○	○

【 0 1 3 1 】

【表 5】

Table 8: The T^2 statistics and its control limits of each class with KM-MSPC (Interval 5)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	3.33		0.91	0.74	1.91	8.75	47.85	39.37
Class2	4.62		10.40	15.00	0.25	8.17	5.45	11.33
Class3	6.60		5.36	1.48	17.63	2.50	42.68	22.13
Detection			—	—	—	—	○	○

Table 9: The Q statistics and its control limits of each class with KM-MSPC (Interval 5)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	19.96		1.13	4.23	47.77	39.36	290	327
Class2	41.81		77.71	73.36	223	142	2286	1183
Class3	28.06		170	133	352	93.66	2903	1680
Detection			—	—	○	○	○	○

【 0 1 3 2 】

10

20

30

【表 6】

Table 10: The T^2 statistics and its control limits of each class with NC-MSPC (Interval 2)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	5.32		14.48	13.66	13.24	18.17	12.74	13.95
Class2	4.83		8.38	11.28	9.17	20.25	16.99	11.19
Class3	5.28		0.98	3.59	1.65	3.28	4.48	3.11
Detection			-	-	-	-	-	-

Table 11: The Q statistics and its control limits of each class with NC-MSPC (Interval 2)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	0.40		0.52	0.30	0.13	0.02	1.15	0.03
Class2	0.27		0.10	0.06	0.00	0.20	0.29	0.00
Class3	0.94		0.00	0.36	0.37	1.21	0.53	1.61
Detection			-	-	-	-	-	-

【 0 1 3 3 】

【表 7】

Table 12: The T^2 statistics and its control limits of each class with NC-MSPC (Interval 3)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	8.79		0.48	0.38	11.30	4.88	12.25	0.57
Class2	5.78		1545	5553	11550	9982	19611	3320
Class3	5.64		8.94	0.70	20.95	0.93	13.60	1.12
Detection			-	-	○	-	○	-

Table 13: The Q statistics and its control limits of each class with NC-MSPC (Interval 3)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	6.80		3.47	182	260	184	372	183
Class2	9.55		364314	331046	483353	204054	252919	357002
Class3	9.43		48.25	5.30	197	19.97	154	77.50
Detection			-	-	○	○	○	○

【 0 1 3 4 】

10

20

30

【表 8】

Table 14: The T^2 statistics and its control limits of each class with NC-MSPC (Interval 4)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	1.35		83.42	3.30	1808	250	843	933
Class2	7.12		0.32	1.80	7.88	3.47	3.04	7.41
Class3	8.05		1.83	5.86	78.65	33.68	30.75	100
Detection			—	—	○	—	○	—

10

Table 15: The Q statistics and its control limits of each class with NC-MSPC (Interval 4)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	0.01		34.83	34.17	28.49	19.38	75.76	32.63
Class2	13.02		5.30	5.12	40.42	19.81	70.64	144
Class3	10.14		153	100	473	315	376	1305
Detection			—	—	○	○	○	○

【 0 1 3 5 】

【表 9】

20

Table 16: The T^2 statistics and its control limits of each class with NC-MSPC (Interval 5)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	8.07		5.04	6.00	2.76	0.30	30.13	19.61
Class2	3.86		1.30	0.79	1.03	4.18	31.11	19.68
Class3	9.63		1.58	0.18	18.72	0.69	69.31	0.04
Detection			—	—	○	—	○	—

Table 17: The Q statistics and its control limits of each class with NC-MSPC (Interval 5)

	CL	δT	Case1	Case2	Case3	Case4	Case5	Case6
			0	0	0.1	-0.1	0.5	-0.5
Class1	32.86		35.81	33.24	108	95.22	1078	545
Class2	23.23		1.50	6.65	53.52	27.93	278	245
Class3	26.33		201	161	324	143	2958	2362
Detection			—	—	○	○	○	○

30

今回開示された実施の形態は単に例示であって、本発明が上記した実施の形態のみに制限されるわけではない。本発明の範囲は、発明の詳細な説明の記載を参酌した上で、特許請求の範囲の各請求項によって示され、そこに記載された文言と均等の意味及び範囲内での全ての変更を含む。

40

【 0 1 3 6 】

Appendix A

バッチプロセスモデル

シミュレーションにて用いたバッチプロセスのモデルを示す。プロセスにおける収支式等は以下の通りである。

【 0 1 3 7 】

【数 2 4】

$$\frac{dM_A}{dt} = -R_1 - R_2 \quad (26)$$

$$\frac{dM_B}{dt} = -R_1 \quad (27)$$

$$\frac{dM_C}{dt} = R_1 - R_2 \quad (28)$$

$$\frac{dM_D}{dt} = R_2 \quad (29)$$

10

$$R_1 = k_1 M_A M_B \quad (30)$$

$$R_2 = k_2 M_A M_C \quad (31)$$

$$k_1 = A_1 \exp(-E_1/RT'_r) \quad (32)$$

20

$$k_2 = A_2 \exp(-E_2/RT'_r) \quad (33)$$

$$W = Mw_A M_A + Mw_B M_B + Mw_C M_C + Mw_D M_D \quad (34)$$

【0 1 3 8】

【数 2 5】

$$M_r = M_A + M_B + M_C + M_D \quad (35)$$

$$C_{PR} = (C_{pA} M_A + C_{pB} M_B + C_{pC} M_C + C_{pD} M_D) / M_r \quad (36)$$

30

【0 1 3 9】

【数 2 6】

$$V = W/\rho \quad (37)$$

$$Ar = V/D \quad (38)$$

$$Q_j = UA(T'_j - T'_r) \quad (39)$$

$$Q_r = -\Delta H_1 R_1 - \Delta H_2 R_2 \quad (40)$$

$$\frac{T'_r}{dt} = \frac{Q_r + Q_j}{M_r C_{pr}} \quad (41)$$

$$\frac{T'_j}{dt} = \frac{F_j \rho_j C_{pj} (T_j^{sp} - T'_j) - Q_j}{V_j \rho_j C_{pj}} \quad (42)$$

$$T'_r = T'_r + a^{(k)} - 0.866a^{(k-1)} \quad (43)$$

$$T'_j = T'_j + a^{(k)} - 0.866a^{(k-1)} \quad (44)$$

ここで、 $a(k) \sim N(0, 0.1)$ であり、 k は測定ステップ数を表す。ただし、 $N(m, \sigma)$ は平均 m 標準偏差 σ の正規分布にしたがう乱数である。温度コントローラからジャケット温度への動特性は、(45)式で与えられる。

【0 1 4 0】

【数 2 7】

$$P_V(s) = \frac{K}{\tau s + 1} \quad (45)$$

$K = 1/16$ 、 $\tau = 2$ 分である。モデルのパラメータをテーブル18に、初期状態をテーブル19に、それぞれ示す。

【0 1 4 1】

【表 10】

Table 18 : Model parameters for the batch process

$M_{wA} = 30 \text{ kg/kmol}$	$C_{pA} = 75.31 \text{ kJ/kmol} \cdot \text{K}$
$M_{wB} = 100 \text{ kg/kmol}$	$C_{pB} = 167.31 \text{ kJ/kmol} \cdot \text{K}$
$M_{wC} = 130 \text{ kg/kmol}$	$C_{pC} = 217.57 \text{ kJ/kmol} \cdot \text{K}$
$M_{wD} = 160 \text{ kg/kmol}$	$C_{pD} = 334.73 \text{ kJ/kmol} \cdot \text{K}$
$k_1^0 = 1.20 \times 10^9$	$\Delta H_1 = -41,840 \text{ kJ/kmol}$
$k_2^0 = 7.88 \times 10^{16}$	$\Delta H_2 = -25,105 \text{ kJ/kmol}$
$E_1/R = 1.00 \times 10^4$	$\rho = 1.00 \times 10^3 \text{ kg/m}^3$
$E_2/R = 1.70 \times 10^4$	$\rho_j = 1.00 \times 10^3 \text{ kg/m}^3$
$r = 0.50 \text{ m}$	$C_{pj} = 1.88 \text{ kJ/kg} \cdot \text{K}$
$F_j = 0.38 \text{ m}^3/\text{min}$	$V_j = 0.69 \text{ m}^3$
$U = 40.842 \text{ kJ/min} \cdot \text{m}^2 \cdot \text{K}$	

10

Table 19: Initial operating condition for the batch process

$M_{A,0} = 12 \text{ kmol}$	$M_{B,0} = 12 \text{ kmol}$
$M_{C,0} = 0 \text{ kmol}$	$M_{D,0} = 0 \text{ kmol}$
$T_{r,0} = 293.15 \text{ K}$	$T_{j,0} = 293.15 \text{ K}$

20

【符号の説明】

【 0 1 4 2 】

8 0 プラント

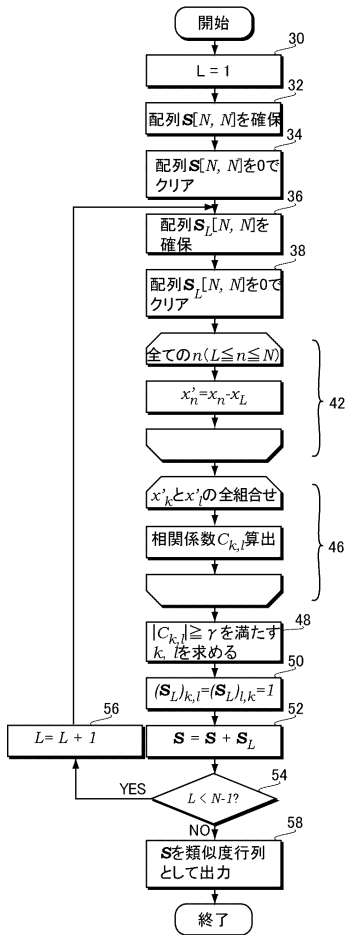
8 2 , 8 4 , 8 6 反応炉

9 2 , 9 4 , 9 6 コントローラ

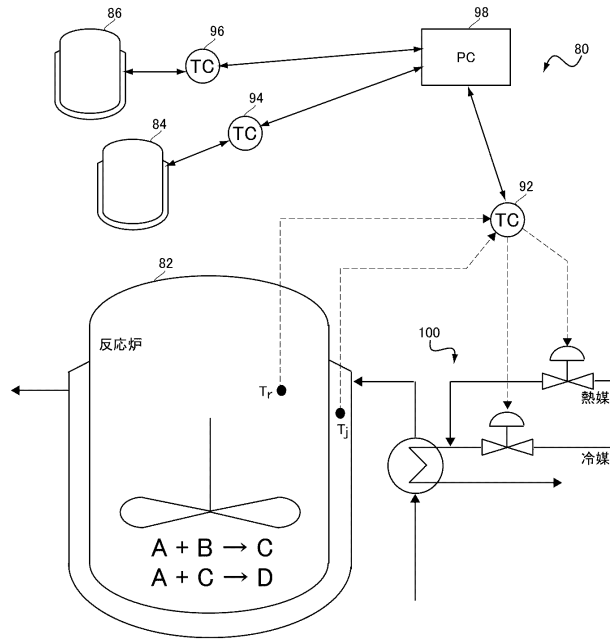
9 8 製品分子量推定装置

1 4 0 C P U

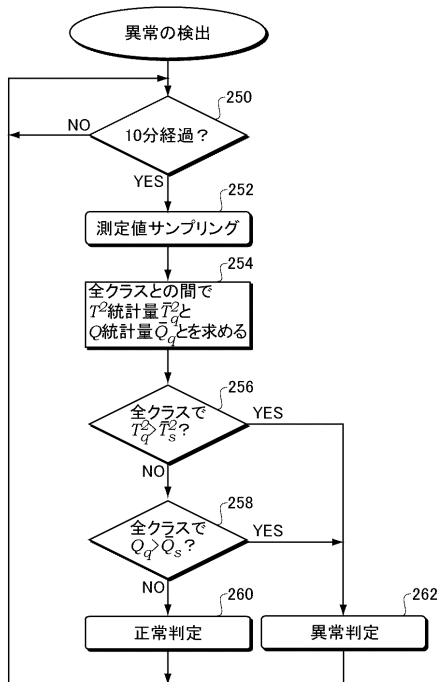
【図 8】



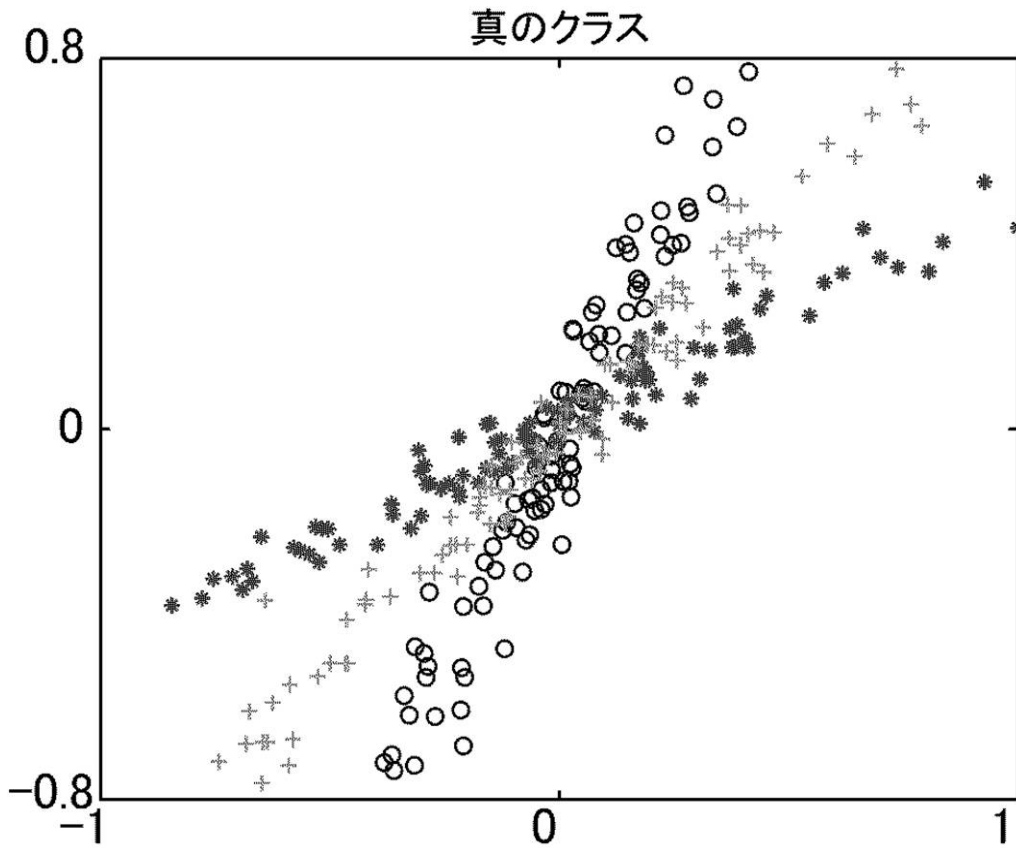
【図 10】



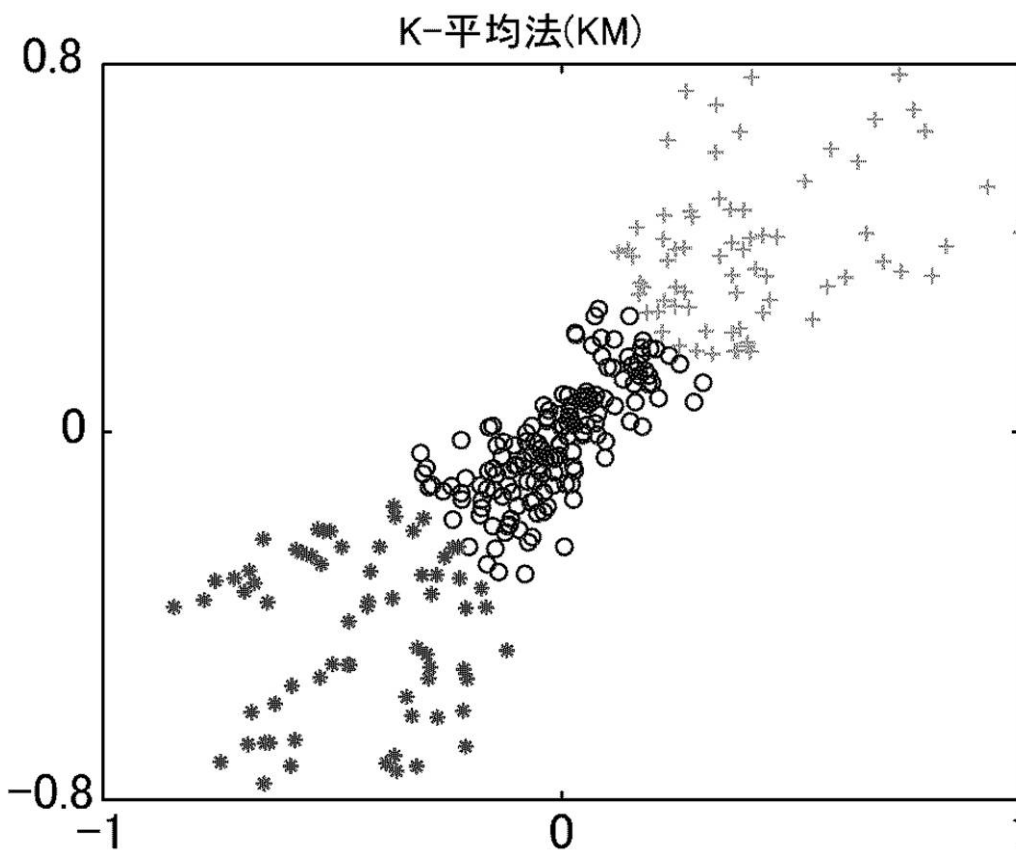
【図 19】



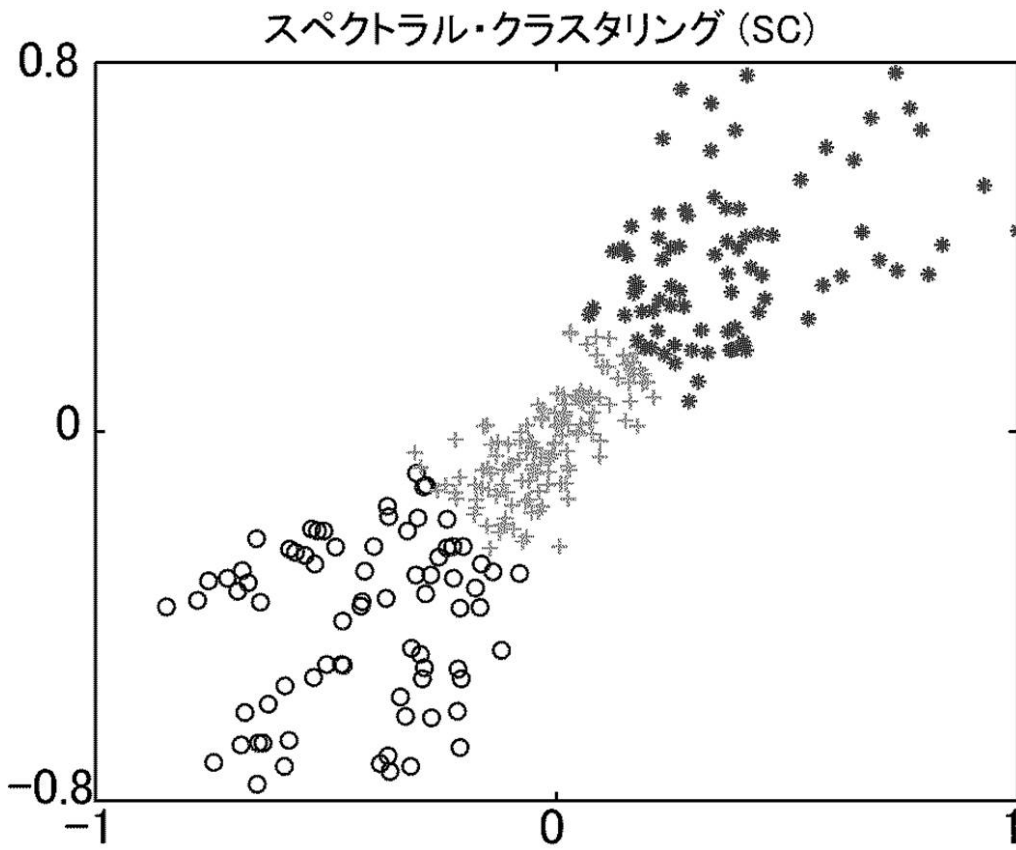
【図1】



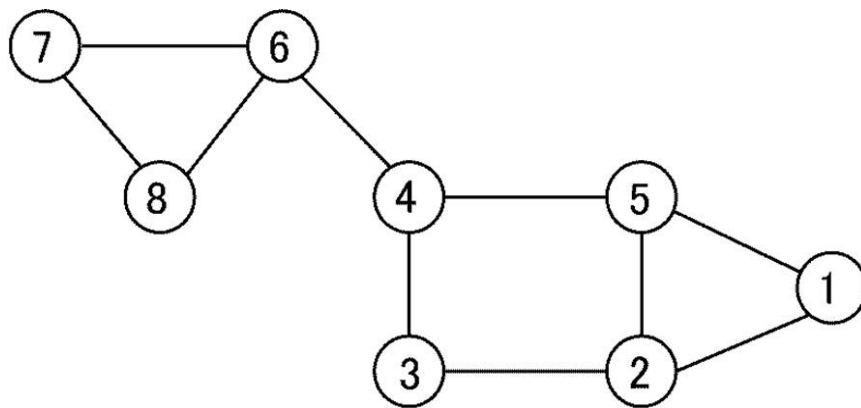
【図2】



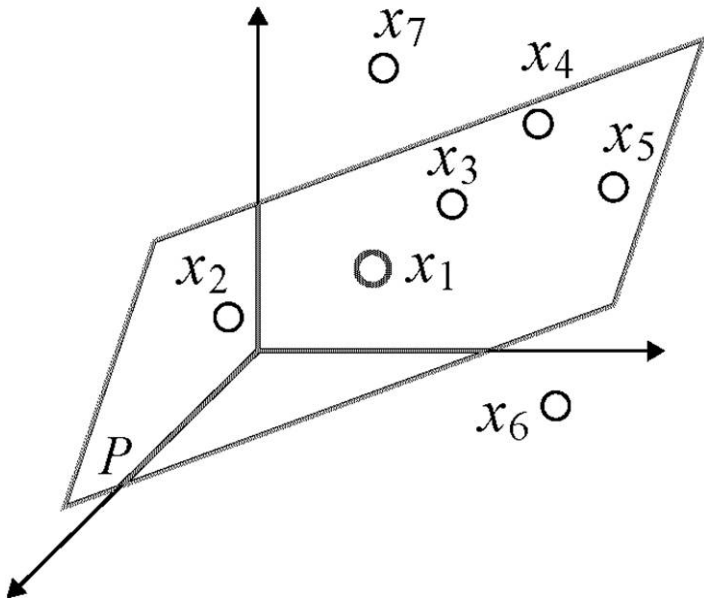
【図3】



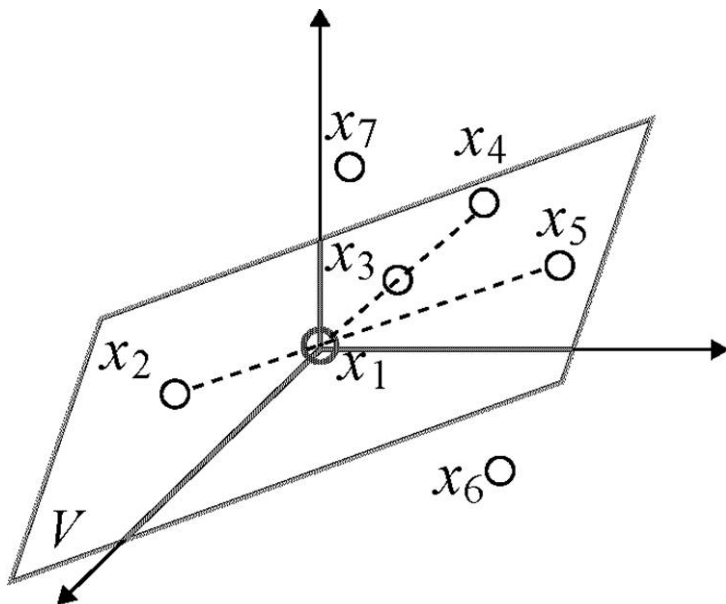
【図4】



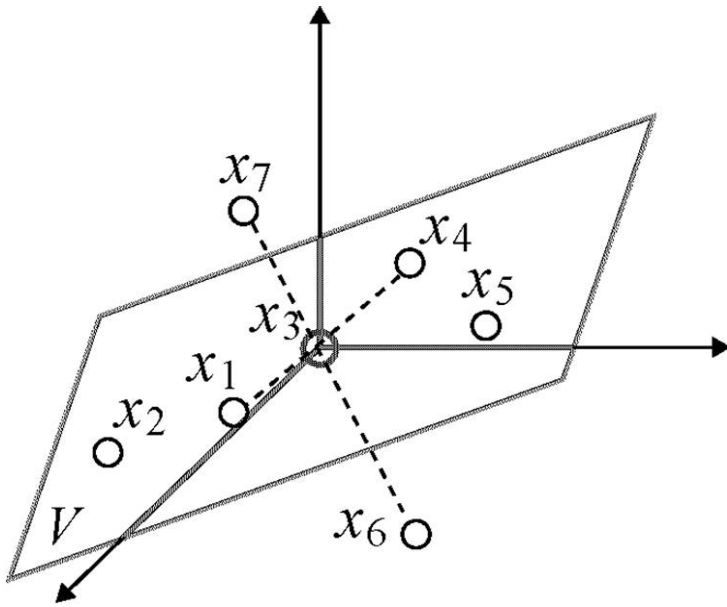
【 図 5 】



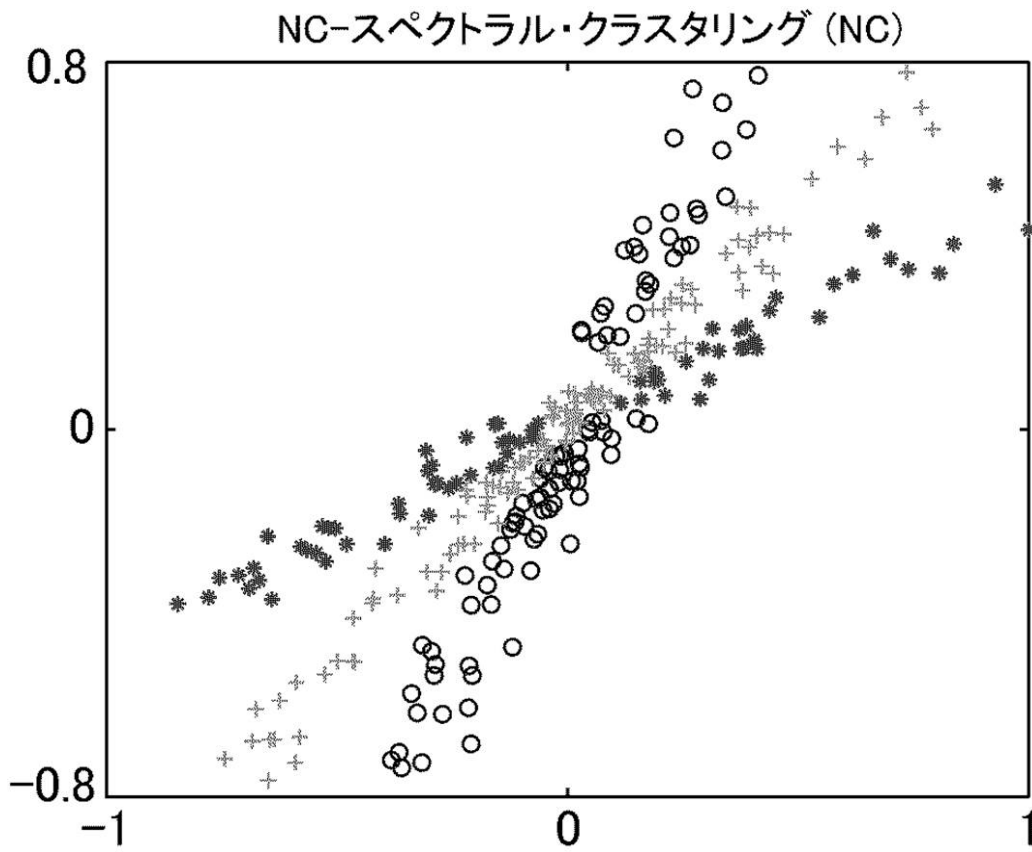
【 図 6 】



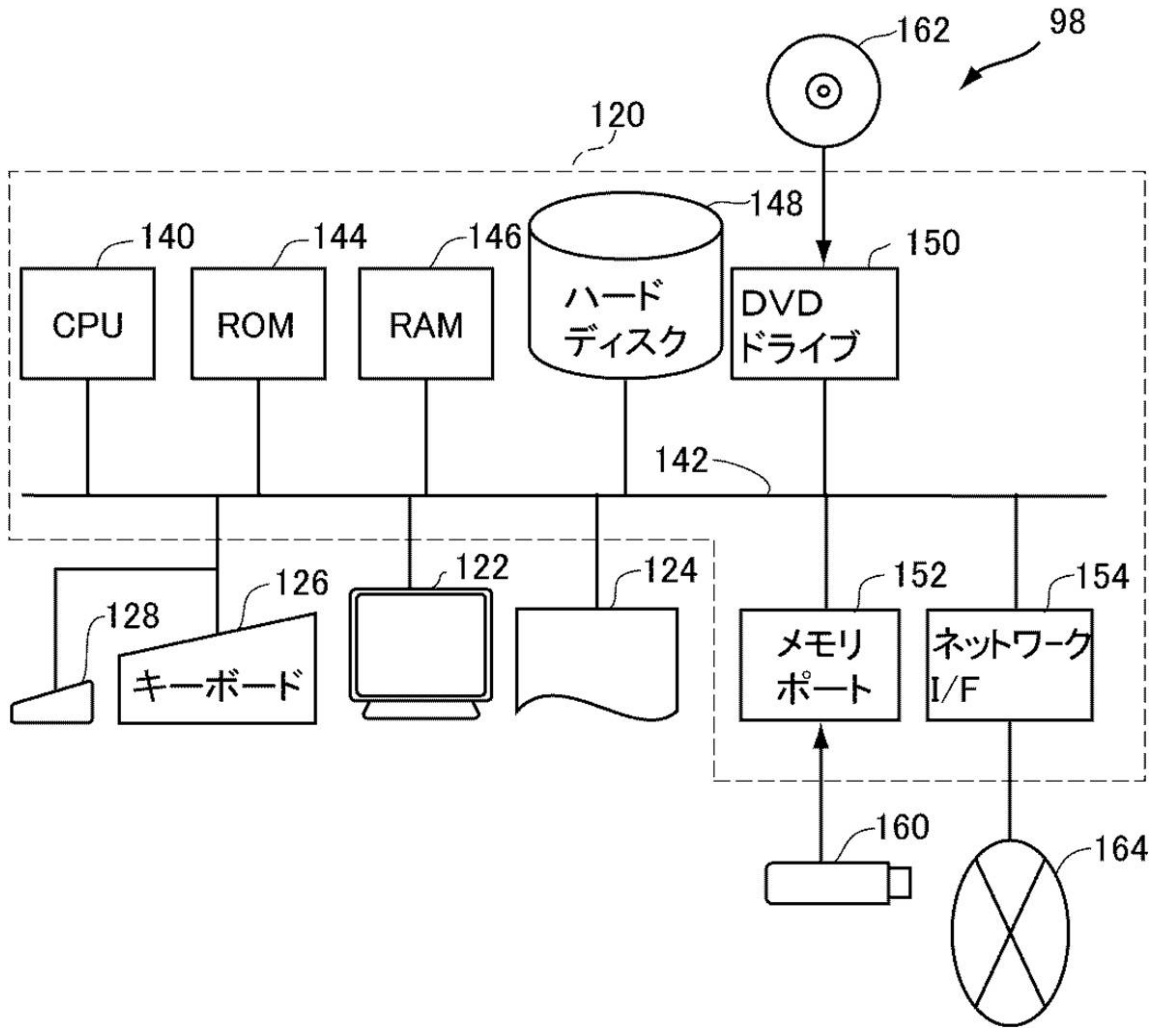
【図7】



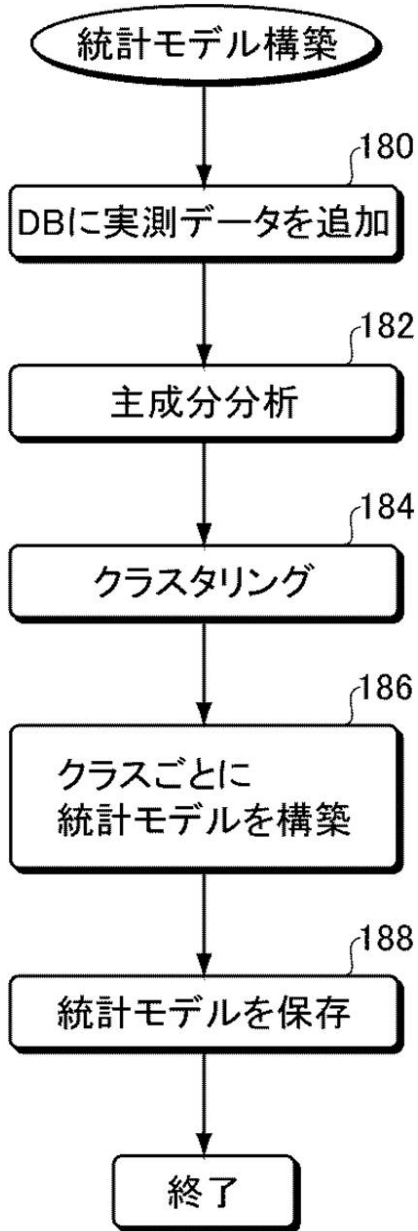
【図9】



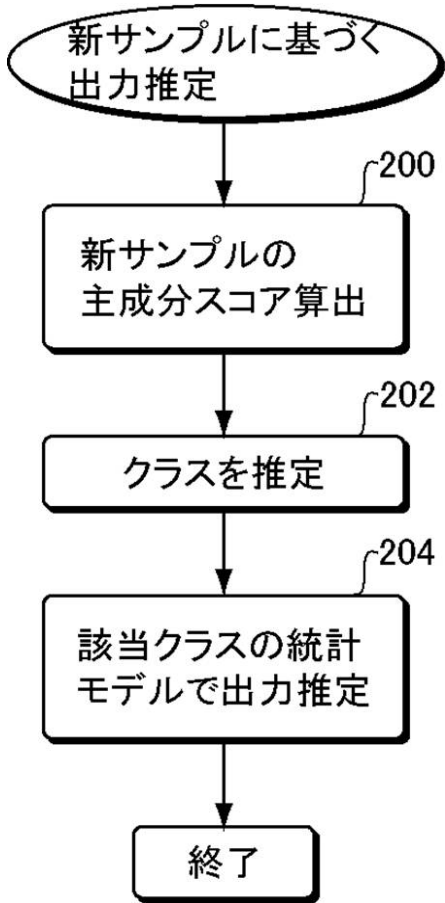
【図11】



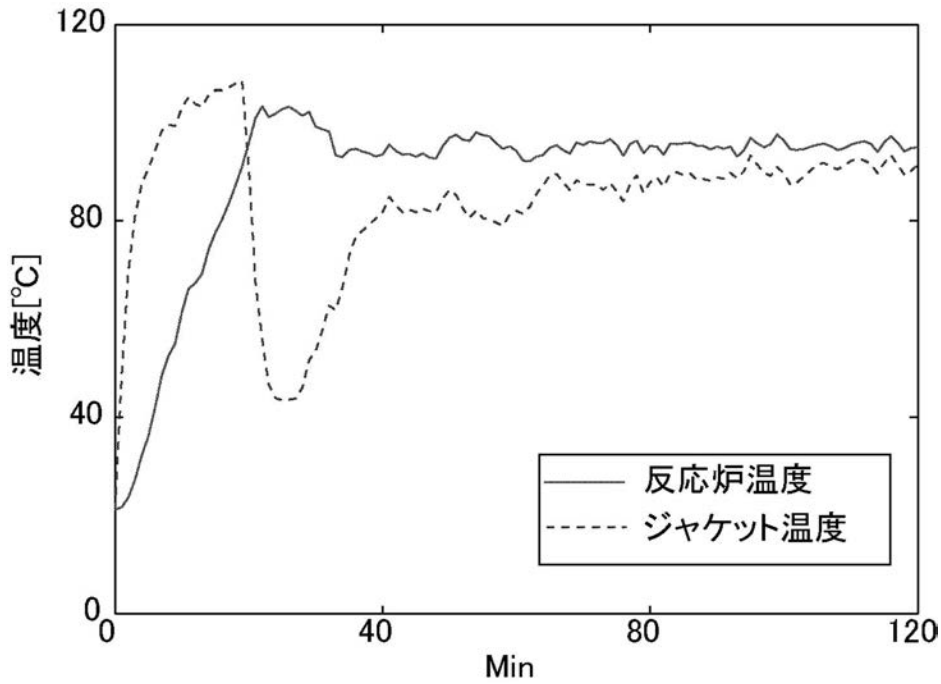
【図12】



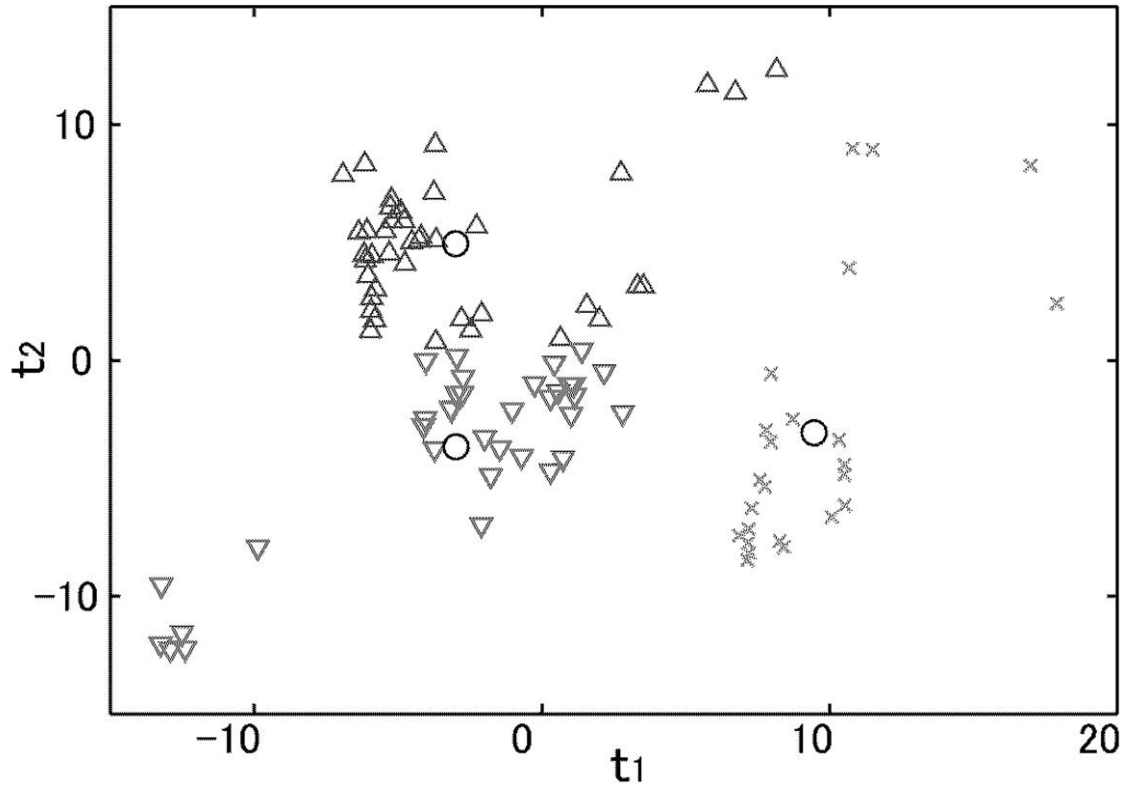
【図13】



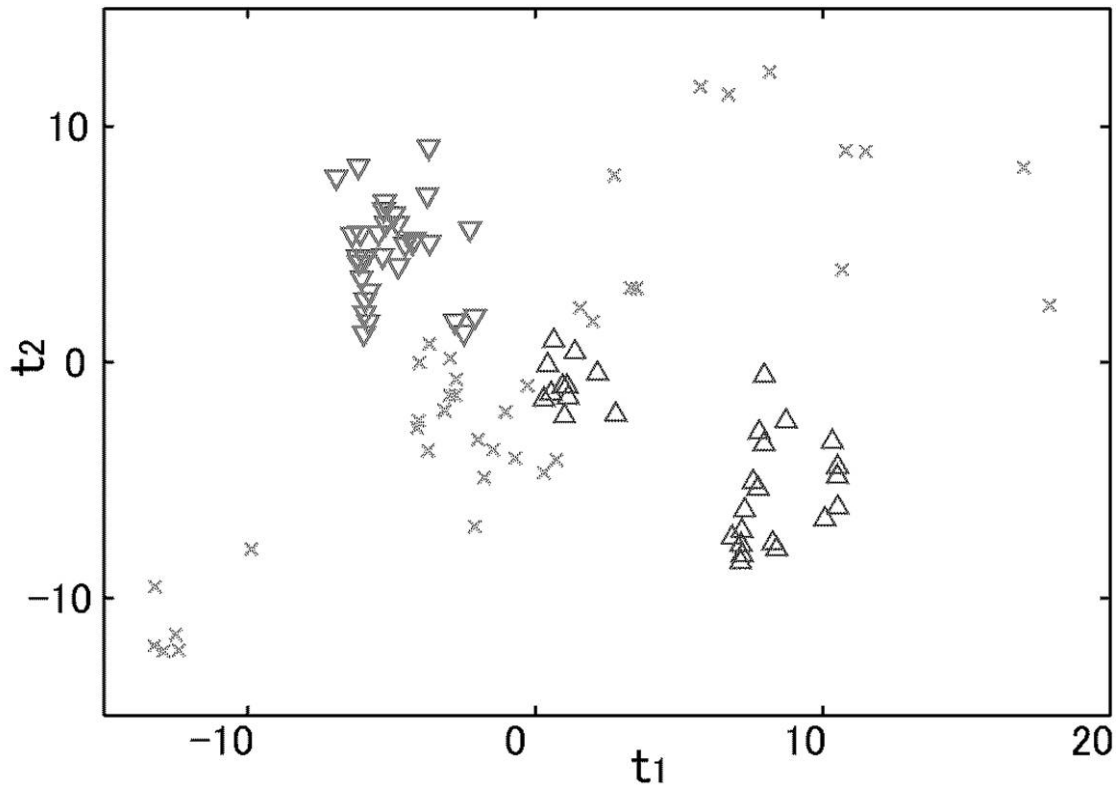
【図14】



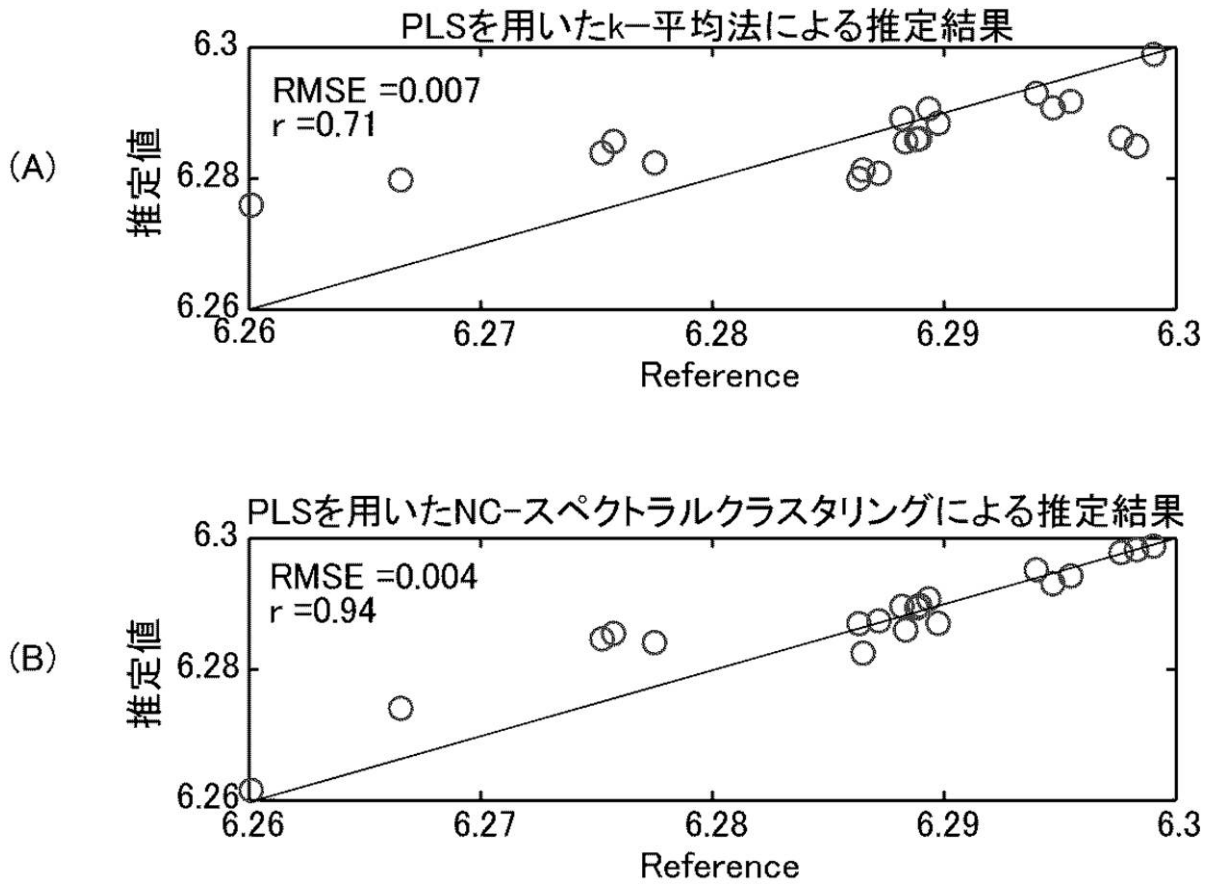
【 図 1 5 】



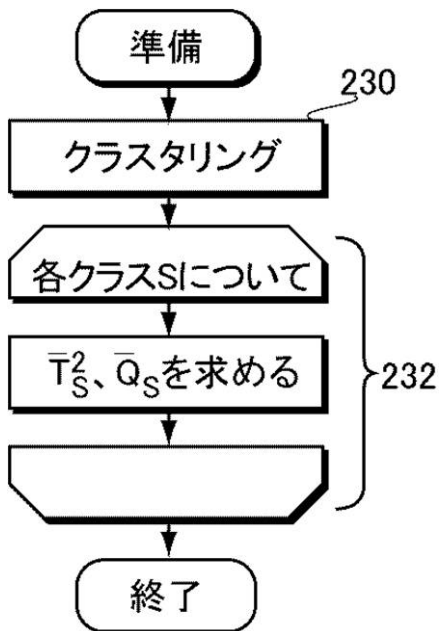
【 図 1 6 】



【図17】



【図18】



フロントページの続き

(56)参考文献 特開2007-272897(JP,A)

特開2006-276924(JP,A)

向井洋介, 変量間の相関に着目したクラスタリング手法およびその多変量統計モデリングへの利用, 化学工学会 研究発表講演要旨集 化学工学会第40回秋季大会, 日本, 社団法人 化学工学会, 2008年 8月24日, K302頁, [平成25年7月4日検索], インターネット<URL : https://www.jstage.jst.go.jp/article/scej/2008f/0/2008f_0_499/_pdf>

(58)調査した分野(Int.Cl., DB名)

G05B 23/02