(19)日本国特許庁(JP)　(12)**公開特許公報(A)**　(11)特許出願公開番号

特開2011-154554

(P2011-154554A)

(43)公開日　**平成23年8月11日(2011.8.11)**

(51)Int.Cl.		F I			テーマコード（参考）
G06Q 10/00	***(2006.01)***	G06F	19/00	１００	
G06Q 50/00	***(2006.01)***	G06F	17/60	１３８	

審査請求 未請求　請求項の数 9　ＯＬ　（全 16 頁）

(21)出願番号　特願2010-15910(P2010-15910)

(22)出願日　平成22年1月27日(2010.1.27)

(71)出願人　000004237
日本電気株式会社
東京都港区芝五丁目７番１号

(74)代理人　100103090
弁理士　岩壁　冬樹

(74)代理人　100124501
弁理士　塩川　誠人

(72)発明者　小阪　勇気
東京都港区芝五丁目７番１号　日本電気株式会社内

(72)発明者　中田　貴之
東京都港区芝五丁目７番１号　日本電気株式会社内

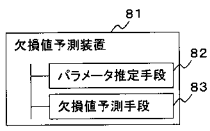
(54)【発明の名称】欠損値予測装置、欠損値予測方法及び欠損値予測プログラム

(57)【要約】

【課題】行列形データにおける欠損値の予測精度を向上できる欠損値予測装置を提供する。

【解決手段】パラメータ推定手段８２は、２つの因子を含む行列形データにおける一方の因子の因子要素ごとに定義される行列であって、その行列形データにおけるもう一方の因子の各因子要素の特徴を表す行列である因子行列を変換する関数のパラメータのうち、その関数が因子行列を変換したときのデータが行列形データである尤もらしさを最大にするパラメータを推定する。欠損値予測手段８３は、パラメータ推定手段８２が推定したパラメータ及び行列形データにおける既知の行列要素の値を用いて、行列形データにおける行列要素の欠損値を予測する。

【選択図】図８



【特許請求の範囲】

【請求項１】

２つの因子を含む行列形データにおける一方の因子の因子要素ごとに定義される行列であって、当該行列形データにおけるもう一方の因子の各因子要素の特徴を表す行列である因子行列を変換する関数のパラメータのうち、当該関数が因子行列を変換したときのデータが前記行列形データである尤もらしさを最大にするパラメータを推定するパラメータ推定手段と、

前記パラメータ推定手段が推定したパラメータ及び前記行列形データにおける既知の行列要素の値を用いて、前記行列形データにおける行列要素の欠損値を予測する欠損値予測手段とを備えた

ことを特徴とする欠損値予測装置。

【請求項２】

パラメータ推定手段は、ガウス分布の共分散行列を用いて定義された関数の因子行列における確率分布に基づいて、当該関数が因子行列を変換したときのデータが行列形データである尤もらしさを最大にする関数のパラメータを推定する

請求項１記載の欠損値予測装置。

【請求項３】

パラメータ推定手段は、ガウス分布の共分散行列を共分散を表す２つの行列のクロネッカー積として表現した確率分布に基づいて、当該関数が因子行列を変換したときのデータが行列形データである尤もらしさを最大にする関数のパラメータを推定する

請求項２記載の欠損値予測装置。

【請求項４】

パラメータ推定手段は、関数の因子行列における確率分布と、当該関数及び因子行列における行列形データの確率分布とにより決定される周辺尤度の値を最大化する関数のパラメータを推定する

請求項２または請求項３記載の欠損値予測装置。

【請求項５】

パラメータ推定手段は、行列形データが２つの因子として製品及び当該製品の不具合を示す情報を含む場合に、当該行列形データにおける一方の因子である製品を示す情報ごとに定義される行列であって、当該行列形データにおけるもう一方の因子である当該製品の不具合を示す情報の特徴を表す因子行列を変換する関数のパラメータのうち、当該関数が因子行列を変換したときのデータが前記行列形データである尤もらしさを最大にするパラメータを推定し、

欠損値予測手段は、推定されたパラメータ及び前記行列形データにおける既知の不具合箇所の情報を用いて、前記製品の未発生の不具合箇所の情報を予測する

請求項１から請求項４のうちのいずれか１項に記載の欠損値予測装置。

【請求項６】

２つの因子を含む行列形データにおける一方の因子の因子要素ごとに定義される行列であって、当該行列形データにおけるもう一方の因子の各因子要素の特徴を表す行列である因子行列を変換する関数のパラメータのうち、当該関数が因子行列を変換したときのデータが前記行列形データである尤もらしさを最大にするパラメータを推定し、

推定されたパラメータ及び前記行列形データにおける既知の行列要素の値を用いて、前記行列形データにおける行列要素の欠損値を予測する

ことを特徴とする欠損値予測方法。

【請求項７】

ガウス分布の共分散行列を用いて定義された関数の因子行列における確率分布に基づいて、当該関数が因子行列を変換したときのデータが行列形データである尤もらしさを最大にする関数のパラメータを推定する

請求項６記載の欠損値予測方法。

【請求項８】

　コンピュータに、
　２つの因子を含む行列形データにおける一方の因子の因子要素ごとに定義される行列であって、当該行列形データにおけるもう一方の因子の各因子要素の特徴を表す行列である因子行列を変換する関数のパラメータのうち、当該関数が因子行列を変換したときのデータが前記行列形データである尤もらしさを最大にするパラメータを推定するパラメータ推定処理、および、
　前記パラメータ推定処理で推定されたパラメータ及び前記行列形データにおける既知の行列要素の値を用いて、前記行列形データにおける行列要素の欠損値を予測する欠損値予測処理
　を実行させるための欠損値予測プログラム。
【請求項９】
　コンピュータに、
　パラメータ推定処理で、ガウス分布の共分散行列を用いて定義された関数の因子行列における確率分布に基づいて、当該関数が因子行列を変換したときのデータが行列形データである尤もらしさを最大にする関数のパラメータを推定させる
　請求項８記載の欠損値予測プログラム。
【発明の詳細な説明】
【技術分野】
【０００１】
　本発明は、欠損値予測システム、欠損値予測方法及び欠損値予測プログラムに関し、特に、２つの因子を含んだ行列形データにおける未知の行列要素の値を予測する欠損値予測システム、欠損値予測方法及び欠損値予測プログラムに関する。
【背景技術】
【０００２】
　未知の行列要素を含む行列形データを扱う際、観測した既知の部分から未知の行列要素の値（以下、欠損値と記すこともある。）を精度良く予測することが要求される。例えば、複数のユーザによる商品の評点を集めた行列データに基づく商品を推薦する協調フィルタリングシステムでは、ユーザによって評価されていない商品の未知評点を予測して、予測した評点の高い商品をユーザに推薦する。
【０００３】
　このようなシステムでは、行列データの欠損値を予測する一般的な方法として、欠損値を行列要素に含む行列形データを確率モデルを用いてサイズの小さい複数の行列データに分解し、その分解した行列データから元の行列データの欠損値を予測する方法が用いられる。このように欠損値を予測する方法が、非特許文献１及び非特許文献２に記載されている。
【０００４】
　非特許文献１に記載された方法では、欠損値を行列要素に含む行列形データの線形な構造を確率モデルにより抽出し、行列を構成する因子ごとにサイズの小さい行列データに分解する。そして、分解した各行列データの線形結合を用いて元の行列形データに近似する行列データを予測する。
【０００５】
　また、非特許文献２に記載された方法では、欠損値を行列要素に含む行列形データの非線形な構造を確率モデルにより抽出し、行列を構成する因子ごとにサイズの小さい行列データに分解する。そして、分解した行列データを用いて元の行列形データに近似する行列データを予測する。すなわち、非特許文献２に記載された方法は、非特許文献１に記載された方法を拡張した方法であると言える。
【先行技術文献】
【非特許文献】
【０００６】
【非特許文献１】Ruslan Salakhutdinov and Andriy Mnih,"Probabilistic Matrix Facto

rization" In Neural Information Processing Systems (NIPS), 2007.
【非特許文献２】Neil D. Lawrence and Raquel Urtasun, "Non-linear Matrix Factorization with Gaussian Processes", Proceedings of the 26th International Conference on Machine Learning (ICML), pp.601-608, Montreal, Canada, 2009.

【発明の概要】

【発明が解決しようとする課題】

【０００７】

非特許文献１及び非特許文献２に記載された方法では、因子行列が与えられた場合、元の行列データが条件付独立であるため、元の行列データに対して条件付独立の制約がない場合に比べると、予測精度が低いという問題がある。

【０００８】

例えば、非特許文献１に記載された方法では、元の行列形データに含まれる各因子の因子行列が与えられた場合、元の行列形データの各行列要素は条件付独立である。また、非特許文献１に記載された方法では、線形な構造を抽出する方法であるため、非線形で複雑な構造を抽出できないという課題がある。

【０００９】

一方、非特許文献２に記載された方法では、非線形な構造を抽出できるため、非特許文献１の欠損値予測システムに比べて、予測精度は高い。具体的には、非特許文献２に記載された方法では、一方の因子に事前分布を仮定して積分消去（周辺化）し、元のデータに与えるその因子の影響を削除することで非線形な構造を抽出する。しかし、非特許文献２に記載された方法においても、元の行列形データに含まれる２つの因子のうち、削除しなかった因子の因子行列が与えられた場合、元の行列形データの削除した方の因子要素は条件付独立である。

【００１０】

以上のように、非特許文献１及び非特許文献２に記載された方法では、因子行列が与えられた場合、元の行列データが条件付独立であるという性質を有する。そのため、元の行列形データに対して条件付独立の制約がなく、また、元の行列形データの各行列要素間の相関関係を組み込んだモデルの予測精度と比べると、精度が低下してしまうという課題がある。

【００１１】

そこで、本発明は、行列形データにおける欠損値の予測精度を向上できる欠損値予測装置、欠損値予測方法及び欠損値予測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】

【００１２】

本発明による欠損値予測装置は、２つの因子を含む行列形データにおける一方の因子の因子要素ごとに定義される行列であって、その行列形データにおけるもう一方の因子の各因子要素の特徴を表す行列である因子行列を変換する関数のパラメータのうち、その関数が因子行列を変換したときのデータが行列形データである尤もらしさを最大にするパラメータを推定するパラメータ推定手段と、パラメータ推定手段が推定したパラメータ及び行列形データにおける既知の行列要素の値を用いて、行列形データにおける行列要素の欠損値を予測する欠損値予測手段とを備えたことを特徴とする。

【００１３】

本発明による欠損値予測方法は、２つの因子を含む行列形データにおける一方の因子の因子要素ごとに定義される行列であって、その行列形データにおけるもう一方の因子の各因子要素の特徴を表す行列である因子行列を変換する関数のうち、その因子行列を変換したときのデータが行列形データである尤もらしさを最大にする関数を推定し、関数を推定する際に用いられたパラメータ及び行列形データにおける既知の行列要素の値を用いて、行列形データにおける行列要素の欠損値を予測することを特徴とする。

【００１４】

本発明による欠損値予測プログラムは、コンピュータに、２つの因子を含む行列形デー

タにおける一方の因子の因子要素ごとに定義される行列であって、その行列形データにおけるもう一方の因子の各因子要素の特徴を表す行列である因子行列を変換する関数のパラメータのうち、その関数が因子行列を変換したときのデータが行列形データである尤もらしさを最大にするパラメータを推定するパラメータ推定処理、および、パラメータ推定処理で推定されたパラメータ及び行列形データにおける既知の行列要素の値を用いて、行列形データにおける行列要素の欠損値を予測する欠損値予測処理を実行させることを特徴とする。

【発明の効果】

【００１５】

本発明によれば、行列形データにおける欠損値の予測精度を向上できる。

【図面の簡単な説明】

【００１６】

【図１】本発明による欠損値予測装置の一実施形態を示すブロック図である。

【図２】本発明の実施形態における動作の例を示すフローチャートである。

【図３】行列形データの例を示す説明図である。

【図４】本発明の実施例における不具合予測システムの例を示す説明図である。

【図５】本発明の実施例における動作の例を示すシーケンス図である。

【図６】送信される行列形データの例を示す説明図である。

【図７】予測した値で欠損値を埋めた予測結果の例を示す説明図である。

【図８】本発明による欠損値予測装置の最小構成の例を示すブロック図である。

【発明を実施するための形態】

【００１７】

以下、本発明の実施形態を図面を参照して説明する。

【００１８】

図１は、本発明による欠損値予測装置の一実施形態を示すブロック図である。本実施形態における欠損値予測装置１０１は、入力手段１０２と、近似手段１０３と、予測手段１０４と、出力手段１０５とを備えている。

【００１９】

入力手段１０２は、例えば、ユーザにより入力される行列形データを近似手段１０３に通知する。ここで、行列形データは、因子数が２の行列であり、例えば、行列の縦軸と横軸がそれぞれ因子になるデータである。なお、以下の説明では、入力手段１０２に入力される行列形データをＹと記す。

【００２０】

近似手段１０３は、入力された行列形データＹにおける一方の因子の因子要素ごとに、入力された行列形データＹにおけるもう一方の因子の各因子要素の特徴を表す行列形データ（以下、因子行列データと記す。）を定義する。そして、近似手段１０３は、因子行列データを変換する関数のパラメータのうち、その関数が因子行列データを変換したときのデータ入力された行列形データＹである尤もらしさを最大にするパラメータを推定する。なお、以下の説明では、一方の因子における各因子要素の特徴を表す因子行列データをＸと記し、因子行列データＸを変換する関数をＦと記す。

【００２１】

具体的には、近似手段１０３は、入力手段１０２に入力された行列形データＹを近似することを目的として、入力された行列形データＹに含まれる一方の因子の因子要素ごとに、入力された行列形データＹに含まれるもう一方の因子の各因子要素の特徴を表す行列形データを引数とする関数のパラメータを推定する。

【００２２】

予測手段１０４は、近似手段１０３が推定した結果と行列形データの既知の行列要素の値から、未知の行列要素の値（すなわち、欠損値）を予測する。例えば、予測手段１０４は、近似手段１０３が推定したパラメータ及び既知の行列要素の値を用いて、欠損値を予測する。

【００２３】
出力手段１０５は、予測手段１０４が予測した欠損値の予測結果を出力する。出力手段１０５は、例えば、ディスプレイなどの表示装置（図示せず）に予測結果を表示させてもよい。もしくは、出力手段１０５は、欠損値予測装置１０１が備えている、もしくは、欠損値予測装置１０１に接続されている記憶装置（図示せず）に予測結果を記憶させてもよい。さらに、出力手段１０５は、他のシステム（図示せず）に予測結果を通知してもよい。

【００２４】
近似手段１０３と、予測手段１０４とは、プログラム（欠損値予測プログラム）に従って動作するコンピュータのＣＰＵによって実現される。例えば、プログラムは、欠損値予測装置１０１の記憶部（図示せず）に記憶され、ＣＰＵは、そのプログラムを読み込み、プログラムに従って、近似手段１０３及び予測手段１０４として動作してもよい。また、近似手段１０３と、予測手段１０４とは、それぞれが専用のハードウェアで実現されていてもよい。

【００２５】
次に、動作について説明する。図２は、本実施形態における動作の例を示すフローチャートである。また、図３は、行列形データの例を示す説明図である。図３に例示する行列形データは、各軸が行列形データの因子を表し、縦軸が因子「映画」を、横軸が因子「ユーザ」をそれぞれ表す。すなわち、行列形データＹは、縦軸と横軸それぞれが意味する内容が因子であり、因子数が２のデータである。また、図３に例示する行列形データの行列要素は、ユーザが各映画を１～５の範囲で評価した点数であり、空白の部分（空白の行列要素）が、行列形データの欠損値を表す。

【００２６】
以下の説明では、２変量、１シーケンスデータの行列形データが入力されるものとする。また、入力される行列形データをＹとし、行列形データＹは、Ｍ×Ｎ次元の行列形データとする。図３に例示する行列形データは、映画の種類（すなわち、因子「映画」の因子要素）がＭ個（具体的には、Ａ～Ｅの５個）、ユーザの数（すなわち、因子「ユーザ」の因子要素）がＮ個（具体的には、１～３０の３０人）の行列形データである。

【００２７】
まず、行列形データＹが入力されると、入力手段１０２は入力された行列形データＹを近似手段１０３に通知する（ステップＳ２０１）。近似手段１０３は、入力された行列形データＹを近似することを目的として、入力された行列形データに含まれる一方の因子の因子要素ごとに、その行列形データに含まれるもう一方の因子の各因子要素の特徴を表す行列形データを引数とする関数のパラメータを推定する（ステップＳ２０２）。

【００２８】
具体的には、まず、近似手段１０３は、欠損値を行列要素に含む行列形データＹを、以下のような（Ｍ×Ｎ）×１次元のベクトルデータ（以下、ベクトルＹと記す。）に変形する。なお、Ｍ及びＮは、因子要素の数である。

【００２９】
$$
\begin{aligned}
Y &= [Y_1 \quad Y_2 \quad \cdots \quad Y_N] \\
&= [y_{11}, y_{12}, \cdots, y_{1M}, y_{21}, y_{22}, \cdots, y_{2M}, \cdots, y_{N1}, y_{N2}, \cdots, y_{NM}]^T
\end{aligned}
$$

【００３０】
ここで、Y_Nは、ベクトルＹの行ベクトルを示し、y_{NM}は、ベクトルＹの行列要素を示す。次に、近似手段１０３は、ベクトルＹを以下の式１のように定義する。

【００３１】
$$Y = F(X) + \varepsilon \quad （式１）$$

【００３２】
Ｘは、因子行列であり、また、関数Ｆは、（Ｍ×Ｎ）×１次元のベクトルである。具体的には、関数Ｆは、以下のように定義される。

【００３３】

$F = [F_1 \quad F_2 \quad \cdots \quad F_N]$
$= [f_{11}, f_{12}, \cdots, f_{1M}, f_{21}, f_{22}, \cdots, f_{2M}, \cdots, f_{N1}, f_{N2}, \cdots, f_{NM}]^T$

【００３４】
ここで、F_Nは、関数Ｆの行ベクトルを表す。また、式１におけるεは、以下のように定義される。

【００３５】
$\varepsilon = N(0, \sigma^2 I)$

【００３６】
ここで、Ｎ（）は、ガウス分布を表す。また、σ^2は分散を、Ｉは、単位行列をそれぞれ表す。次に、各因子要素の特徴を表す因子行列として、以下に例示する因子行列Ｘ’を定義する。

【００３７】
【数１】

$$X' = \begin{bmatrix} X'_1 \\ X'_2 \\ \vdots \\ X'_{N-1} \\ X'_N \end{bmatrix} = \begin{bmatrix} x'_{1,11} & x'_{1,21} & \cdots & x'_{1,q-11} & x'_{1,q1} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ x'_{1,1M} & x'_{1,2M} & \cdots & x'_{1,q-1M} & x'_{1,qM} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ x'_{N,11} & x'_{N,21} & \cdots & x'_{N,q-11} & x'_{N,q1} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ x'_{N,1M} & x'_{N,2M} & \cdots & x'_{N,q-1M} & x'_{N,qM} \end{bmatrix}$$

【００３８】
上述の通り、Ｘ’は、元の行列形データＹの横軸の因子の特徴を表すＭＮ×ｑの行列である。ＭＮは、因子行列Ｘ’の特徴の個数を表す。ここで、行列形データＹの縦軸の因子の各因子要素の特徴は横軸の各因子要素で表わされ、その特徴の次元はＭである。すなわち、行列形データＹの１行目（すなわち、縦軸の因子の１番目の因子要素）の特徴は、行列要素の（１行目、１列目）、（１行目、２列名）、・・・、（１行目、Ｍ列目）で表される。すなわち、縦軸の因子要素数はＮであることから、因子行列Ｘ’は、縦軸の因子の因子要素ごとに、横軸の因子の各因子要素の特徴を表している行列形データであると言うことができる。

【００３９】
また、ｑは、因子行列Ｘ’の特徴の次元を表すパラメータである。ｑの値は、ユーザ等により予め与えられる。

【００４０】
ここで、X_1’は、因子行列Ｘ’の１行目の行ベクトルを表す。したがって、例えば、行列形データＹの１行目の行ベクトルY_1は、以下のように表すことができる。ここで、εは、式１におけるεと同様である。

【００４１】
$Y_1 = F_1(X_1') + \varepsilon$

【００４２】
次に、近似手段１０３は、ベクトルＹの確率分布を、以下に例示する式２のように定義する。

【００４３】
$P(Y \mid F(X')) = N(Y \mid F(X'), \sigma^2 I)$　（式２）

【００４４】
ここで、関数Ｆの確率分布をガウシアンプロセスを用いて以下に例示する式３のように

定義する場合について説明する。
【００４５】
　Ｐ（Ｆ｜Ｘ）＝ＧＰ（Ｆ｜０，K^{WX}）＝Ｎ（Ｆ｜０，K^{WX}）　（式３）
【００４６】
　ガウシアンプロセスは、非線形の入出力関係を正規確率過程から得られたものとして捉える確率過程であり、行列演算のみで確率過程を記述できることを特徴とする方法である。また、K^{WX}は、ガウス分布の共分散行列を表す。共分散行列の各要素k^{WX}_{ij}は、以下のように表すことができる。
【００４７】
　k^{WX}_{ij}＝＜X_i’，X_j’＞
【００４８】
　なお、＜＞は、内積を表す。すなわち、k^{WX}_{ij}は、X_i’とX_j’の内積として表すことができる。
【００４９】
　共分散行列K^{WX}は、推定が必要なパラメータであるが、サイズが大きいＮＭ×ＮＭの行列データである。そのため、共分散行列K^{WX}を推定しようとすると、ＮＭ×ＮＭ個の行列要素全てを推定する必要があり、計算コストが大きくなってしまう。そこで、近似手段１０３は、式３を変更し、以下に例示する式４のように関数Ｆの確率分布を定義する。
【００５０】
【数２】

$$P(F \mid X) = GP(F \mid 0, K^W \otimes K^X) = N(F \mid 0, K^W \otimes K^X) \qquad \text{(式4)}$$

ここで、

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ \vdots \\ X_M \end{bmatrix} = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{q-1,1} & x_{q1} \\ x_{12} & x_{22} & \cdots & x_{q-1,2} & x_{q2} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ x_{1M-1} & x_{2M-1} & \cdots & x_{q-1,M-1} & x_{qM-1} \\ x_{1M} & x_{1M} & \cdots & x_{q-1,M} & x_{qM} \end{bmatrix}$$

$$W = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ \vdots \\ W_N \end{bmatrix} = \begin{bmatrix} w_{11} & w_{21} & \cdots & w_{r-1,1} & w_{r1} \\ w_{12} & w_{22} & \cdots & w_{r-1,2} & w_{r2} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ w_{1N-1} & w_{2N-1} & \cdots & w_{r-1,N-1} & w_{rN-1} \\ w_{1N} & w_{2N} & \cdots & w_{r-1,N} & w_{rN} \end{bmatrix}$$

【００５１】
　ここで、ＸはＭ×ｑの行列形データであり、ＷはＮ×ｒの関数を表す行列形データである。なお、ｒは、関数Ｗの特徴の次元を表すパラメータであり、ｒの値は、ユーザ等により予め与えられる。すなわち、式４は、式３における共分散行列K^{WX}を、以下に示す内容（以下、推定対象パラメータと記す。）に置き換えたものである。
【００５２】
【数３】

$$K^W \otimes K^X$$

【００５３】
　ここで、以下の記号は、クロネッカー積を表す。

【００５４】

【数４】

$$\otimes$$

【００５５】

また、K^{W}は、マーセルの定理を満たすカーネルと呼ばれるＮ×Ｎの行列データである。同様に、K^{X}は、マーセルの定理を満たす（すなわち、カーネルと呼ばれる）Ｍ×Ｍの行列データである。

【００５６】

ここで、行列K^{W}の行列要素k^{W}_{ij}、及び、行列K^{X}の行列要素k^{X}_{ij}は、それぞれ以下のように表すことができる。

【００５７】

$k^{W}_{ij}=k^{W}(W_i, W_j)=\langle W_i, W_j\rangle$

$k^{X}_{ij}=k^{X}(X_i, X_j)=\langle X_i, X_j\rangle$

【００５８】

なお、＜＞は、内積を表す。すなわち、k^{W}_{ij}は、W_iとW_jの内積として表すことができ、k^{X}_{ij}は、X_iとX_jの内積として表すことができる。具体的には、行列K^{W}の行列要素$k^{W}(W_i, W_j)$は、W_iとW_jとを非線形に写像した高次元ベクトル空間上におけるW_iとW_jとの距離を表す。行列K^{X}の行列要素$k^{X}(X_i, X_j)$についても同様である。

【００５９】

上述の推定対象パラメータは、共分散行列K^{WX}と同様、推定が必要なパラメータである。推定対象パラメータの内容を推定する場合、近似手段１０３は、Ｎ×Ｎ＋Ｍ×Ｍ個の行列要素を推定することになる。すなわち、共分散行列K^{WX}を推定する場合には、ＮＭ×ＮＭ個の行列要素全てを推定する必要があるが、推定対象パラメータを推定する場合、Ｎ×Ｎ＋Ｍ×Ｍ個の行列要素を推定すればよい。したがって、共分散行列K^{WX}を推定する場合に比べ、計算コストを小さくすることができる。

【００６０】

次に、パラメータを推定する際の計算オーダを低減させる方法について説明する。上述の通り、式４に例示するモデルを用いることで、式３に例示するモデルよりも推定するパラメータを低減させることができる。さらに、式３に例示するモデルよりも計算オーダを低減させるため、まず、式４を以下の式５のように変形する。

【００６１】

【数５】

$$P(Y\mid X,W,S)=\int P(Y\mid F^{X})\int P(F^{X}\mid X,F^{W})P(F^{W}\mid W)dF^{X}dF^{W} \qquad \text{(式５)}$$

【００６２】

なお、F^{X}は、Ｘを引数とする関数である。また、F^{W}は、Ｗを引数とする関数である。ここで、式４及び式５に例示するモデルにおける計算オーダは、$O(M^{3}N^{3})$である。そのため、式５を近似して、以下に例示する式６を定義する。式６に例示するモデルにおける計算オーダは、$O(MNN^{2})$になる。

【００６３】

【数６】

$$P(Y\mid X,W,S)\approx\int P(Y\mid F^{X})\int\left(\prod_{n=1}^{N}P(F_n^{X}\mid X,F^{W})\right)P(F^{W}\mid W)dF^{W}dF^{X} \qquad \text{(式６)}$$

【００６４】

F_n^{X}は、各ｎについて独立であるという仮定のもとに定義される関数である。すなわ

ち、Ｆ$_{n}{}^{X}$を定義することは、式４において因子要素数Ｎの各因子要素の独立性を仮定しているとも言える。しかし、ここで仮定する独立性は、例えば、非特許文献２に記載された方法で用いられる独立性とは異なる。例えば、式６において、独立性を仮定したと言える部分は、以下の部分である。

【００６５】

【数７】

$$\prod_{n=1}^{N} P(F_n{}^X \mid X, F^W)$$

【００６６】

一方、非特許文献２に記載された方法では、独立性を仮定した部分は、以下の式７のように表わされる部分である。

【００６７】

【数８】

$$\prod_{n=1}^{N} P(F_n{}^X \mid X) \qquad \text{（式７）}$$

【００６８】

両者を比較すると、式６において独立性を仮定したと言える部分がＦ^{W}を含む点で式７と異なる。このように、式７における独立性の仮定とは異なり、式６では、独立性の仮定において欠落する情報をＦ^{W}で補っている。

【００６９】

なお、実問題においては、因子ごとの特徴を表すデータが新たに与えられることがある。例えば、「映画」という因子の特徴として、映画に出演した俳優の情報や、映画の発表年月日、映画製作会社、映画配給会社等の情報が与えられる。また、例えば、「ユーザ」という因子の特徴として、性別、年代、地域などの情報が与えられる。このように与えられるデータをメタデータと呼ぶ。これらのメタデータを利用し、式６に例示するモデルを、以下に例示する式８のように定義し、さらに式８を用いて式９のように定義してもよい。

【００７０】

【数９】

$$P(Y \mid X, S) \approx \int P(Y \mid F^X) \int (\prod_{n=1}^{N} P(F_n{}^X \mid X, F^W)) \int P(F^W \mid W) P(W \mid S) dW dF^W dF^X \qquad \text{（式８）}$$

【００７１】

【数１０】

$$P(Y \mid R, S) = \int P(Y \mid X, S) P(X \mid R) dX \qquad \text{（式９）}$$

【００７２】

ここで、Ｒは、行列形データＹの縦軸を表す因子（すなわち、因子要素数がＭの因子）のメタデータを表す。また、Ｓは、行列形データＹの横軸を表す因子（すなわち、因子要素数がＮの因子）のメタデータを表す。なお、メタデータを導入して拡張された式９は、式６の階層モデルと呼ばれる。

【００７３】

次に、パラメータを推定する具体的な方法について説明する。推定が必要なパラメータは、Ｋ^{W}、Ｋ^{X}及びσである。そこで、近似手段１０３は、関数Ｆが因子行列Ｘを変換したときの行列が行列形データＹである尤もらしさを最大にするパラメータを推定する。例えば、近似手段１０３は、以下の式１０に例示する周辺尤度ｌｏｇＰ（Ｙ｜Ｘ）を最大化

するパラメータＫ ^{W}、Ｋ ^{X} 及びσを、勾配法を用いて推定してもよい。

【００７４】

【数１１】

$$\log P(Y \mid X) = \log \int P(Y \mid F, X, \sigma^2) P(F \mid X) dF \qquad \text{(式１０)}$$

【００７５】

このように導出されたモデル（欠損値を予測するモデル）は、以下のような特徴を持つ。すなわち、欠損値を行列要素に含む行列形データＹが確率モデルにより関数Ｆと因子行列Ｘでモデル化され、さらに、関数Ｆの事前確率（例えば、上述の式３）がガウシアンプロセスで定義されているため、非線形な相関構造を考慮可能なモデルになっている。

【００７６】

さらに、本モデルは、元の行列形データＹが関数Ｆと因子行列Ｘとを用いて表されているため、元の行列形データＹに対して条件付独立の制約がない。具体的には、本モデルには、元の行列形データＹの各行列要素間の相関関係が組み込まれている。そのため、予測精度が低下することを抑制できる。

【００７７】

さらに、本モデルでは、関数Ｆの共分散行列の構造に着目し、サイズの大きい共分散行列を、共分散を表すサイズの小さい２つの行列のクロネッカー積で表現している。そのため、推定が必要なパラメータ数を減らして、計算量を削減することを可能にしている。具体的には、元の行列形データＹに含まれる一方の因子の各要素の特徴を表す行列形データ（すなわち、因子行列）Ｘに関しては、全行列要素数（特徴の次元×要素数）のパラメータを推定し、もう一方の因子に関しては、行列形データの要素数の約２乗個のパラメータを推定すればよい。

【００７８】

次に、予測手段１０４は、近似手段１０３が推定した結果と行列形データの既知の行列要素の値から、未知の行列要素の値を予測する（ステップＳ２０３）。予測手段１０４は、例えば、近似手段１０３が推定したパラメータであるＫ ^{W}、Ｋ ^{X}（すなわち、W及びX）及び、行列形データの既知の行列要素の値を用いて、以下に例示する式１１により未知の行列要素（欠損値）を予測する。

【００７９】

【数１２】

$$y_{ij} = \hat{f}_{ij} = (k^{w}_{i,:} \otimes k^{X}_{j,:})^T (K^W \otimes K^X + DI)^{-1} Y \qquad \text{(式１１)}$$

【００８０】

ここで、Ｄは、対角成分がσの二乗であるＮＭ×ＮＭの行列データである。また、Ｙは、元の行列形データＹを変形したベクトルデータである。また、$k^{W}_{i,:}$ は、Ｋ ^{W} のｉ行目を列ベクトルにしたＮ次元のベクトルを表し、$k^{X}_{i,:}$ は、Ｋ ^{X} のｉ行目における列ベクトルを表す。例えば、ベクトルＹの要素 y_{1M} の値が欠損している場合、予測手段１０４は、式１１を用いた以下に例示する式１２により未知の行列要素（欠損値）を予測すればよい。

【００８１】

【数１３】

$$y_{1M} = \hat{f}_{1M} = (k^{w}_{1,:} \otimes k^{x}_{M,:})^T (K^W \otimes K^X + DI)^{-1} Y \qquad \text{(式１２)}$$

【００８２】

以下、行列形データＹの全ての欠損箇所を予測したデータ（すなわち、行列形データの欠損値を予測した値で埋めたデータ）をＹ’とする。最後に、出力手段１０５は、欠損値

を予測した値で埋めた行列データＹ’を出力する（ステップＳ２０４）。
【００８３】
以上のように、本実施形態によれば、まず、２つの因子を含む行列形データＹの縦軸方向の因子の因子要素ごとに、その行列形データＹにおける横軸方向の因子の各因子要素の特徴を表す因子行列Ｘが定義される。次に、近似手段１０３が、因子行列Ｘを変換する関数Ｆのパラメータのうち、その関数が因子行列Ｘを変換したときのデータが行列形データＹである尤もらしさ（周辺尤度）を最大にするパラメータを推定する。そして、予測手段１０４が、推定されたパラメータK^W、K^X（すなわち、Ｗ及びＸ）及び行列形データＹにおける既知の行列要素の値を用いて、行列形データの行列要素の欠損値を予測する。このようにすることで、行列形データにおける欠損値の予測精度を向上させることができる。
【００８４】
また、例えば、非特許文献１に記載された方法では、元の行列形データを全ての因子ごとにサイズの小さな行列形データに分解し、分解した因子行列データの全要素数を推定する必要があるため、計算コストは大きい。また、非特許文献２に記載された方法では、片方の因子行列データの全要素数だけ推定すればよいが、それでも、非特許文献１の計算コストの半分であり、元の行列データが条件付き独立であるという性質は有したままである。
【００８５】
しかし、本実施形態によれば、サイズの大きい共分散行列を、共分散を表すサイズの小さい２つの行列のクロネッカー積で表現することで推定が必要なパラメータ数を減らしているため、計算コストを抑えることが出来る。
【実施例】
【００８６】
以下、具体的な実施例により本発明を説明するが、本発明の範囲は以下に説明する内容に限定されない。以下の実施例では、会社や工場から製品の不具合情報を受信し、その不具合情報から将来発生する不具合を事前に予測するシステム（以下、不具合予測システムと記す。）に本発明を適用する場合について説明する。
【００８７】
図４は、本実施例における不具合予測システムの例を示す説明図である。なお、上記実施形態と同様の構成については、図１と同一の符号を付し、説明を省略する。本実施例における不具合予測システムは、クライアントシステム４０１と、サーバシステム４０３とを備えている。クライアントシステム４０１と、サーバシステム４０３とは、通信ネットワーク４０２を介して相互に接続される。
【００８８】
初めに、本実施例の概要を説明する。まず、テレビジョンや冷蔵庫などの家電製品や半導体などの部品製品を製作する各会社で運用されているクライアントシステム、もしくは、製品を製作する各工場で運用されているクライアントシステムから、製品の不具合情報をサーバシステムに送信する。サーバシステムでは、送信された過去の製品の不具合情報（既知の不具合情報）から、現時点で未発生であるが将来発生する不具合を事前に予測して、未発生の不具合が起こるか否かを各クライアントシステムへ送信する。クライアントシステム側では、このように送信された予測情報を使うことで、不具合の発生を未然に防ぐことができるようになる。また、クライアントシステム側では、不具合が起こる前から不具合に備えることができるようになる。
【００８９】
以下、各構成要素について説明する。クライアントシステム４０１は、製品の故障情報等を管理する。具体的には、クライアントシステム４０１は、過去に発生した製品の故障情報を記憶し、その故障情報を通信ネットワーク４０２を介してサーバシステム４０３に送信する。
【００９０】

サーバシステム４０３は、欠損値予測装置１０１と、予測結果記憶部４０４とを備えている。サーバシステム４０３は、クライアントシステム４０１の情報を管理する。具体的には、サーバシステム４０３は、クライアントシステム４０１から製品の故障情報を受信すると、その故障情報をもとに将来発生する不具合を予測して、その予測情報をクライアントシステム４０１に送信する。

【００９１】

欠損値予測装置１０１は、上述の実施形態に記載した欠損値予測装置１０１と同様である。ここで、欠損値予測装置１０１は、入力された行列形データから、製品の未発生の不具合箇所の情報を予測する。また、予測結果記憶部４０４は、欠損値予測装置１０１が出力した欠損値予測済みの行列形データを記憶する。

【００９２】

次に、動作について説明する。図５は、本実施例における動作の例を示すシーケンス図である。

【００９３】

まず、各クライアントシステム４０１が、通信ネットワーク４０２を介して、サーバシステム４０３に、各製品の不具合情報を送信する（ステップＳ５０１）。図６は、本実施例において各クライアントシステム４０１から送信される行列形データの例を示す説明図である。図６に例示する行列形データの各軸は、行列形データの因子を表す。行列形データは、縦軸と横軸を持つデータであるため、因子の数は2になる。また、図６に例示する行列形データは、縦軸が「不具合」の内容を表し、横軸が不具合が発生しうる「製品」を表す。

【００９４】

また、以下の説明では、行列形データの行列要素が製品の不具合発生件数である場合について説明する。すなわち、図６に例示する行列要素は、製品の不具合件数を示し、空白の行列要素が欠損値（すなわち、不具合未発生）を示す。ただし、行列要素は、不具合件数に限られない。行列要素は、例えば、不具合が発生する確率であってもよく、不具合が発生したか否かを表すバイナリ値であってもよい。なお、行列要素が製品の不具合発生件数の場合、欠損箇所の予測値は、将来起こりうる不具合の件数と言える。

【００９５】

例えば、各クライアントシステム４０１は、図６に例示する行列形データをサーバシステム４０３に送信する。なお、行列形データにおける欠損部分を「（製品、不具合）」という形式で表現する場合、図６に例示する行列データの欠損部分は、（１、Ｄ）、（１、Ｅ）、（２、Ｄ）、（３、Ｅ）、（４、Ａ）、（４、Ｃ）、（５、Ｅ）、（３０、Ａ）及び（３０、Ｂ）である。

【００９６】

サーバシステム４０３は、製品の不具合情報を受信すると、その製品の不具合情報を行列形データとして欠損値予測装置１０１に入力する（ステップＳ５０２）。以下、この入力形データをＹとする。欠損値予測装置１０１（より具体的には、図１における近似手段１０３及び予測手段１０４）は、入力された行列形データＹをもとに欠損値を予測する（ステップＳ５０３）。

【００９７】

具体的には、上記実施形態に示したように、まず、製品を示す情報ごとに製品の不具合を示す情報の特徴を表す因子行列が定義される。近似手段１０３は、その因子行列を変換する関数のパラメータのうち、その関数が因子行列を変換したときのデータが入力された行列形データである尤もらしさを最大にするパラメータを推定する。そして、予測手段１０４が、推定されたパラメータ及び行列形データにおける既知の不具合箇所の情報を用いて、製品の未発生の不具合箇所の情報を予測する。

【００９８】

その後、欠損値予測装置１０１（より具体的には、図１における出力手段１０５）は、予測した値で欠損値を埋めた行列形データ（以下、Ｙ’と記す。）を出力する。具体的に

は、欠損値予測装置１０１は、予測結果を予測結果記憶部４０４に記憶させる（ステップＳ５０４）。

【００９９】

図７は、本実施例において予測した値で欠損値を埋めた予測結果（すなわち、行列形データＹ’）の例を示す説明図である。なお、行列形データＹにおける欠損部分を「（製品、不具合、予測値）」という形式で表現する場合、図７に例示する行列形データＹ’の欠損部分は、（１、Ｄ、１．０１）、（１、Ｅ、０．３）、（２、Ｄ、２）、（３、Ｅ、２．０）、（４、Ａ、０．５）、（４、Ｃ、２．５）、（５、Ｅ、５）、（３０、Ａ、２．０）及び（３０、Ｂ、２．０）である。

【０１００】

次に、欠損値予測装置１０１（より具体的には、図１における出力手段１０５）は、予測結果（すなわち、行列形データＹ’）をクライアントシステム４０１に送信し（ステップＳ５０５）、クライアントシステム４０１は、その予測結果（すなわち、行列形データＹ’）を受信する（ステップＳ５０６）。

【０１０１】

以上のことから、例えば、テレビジョンや冷蔵庫などの家電製品や半導体などの部品製品を製作する各会社や製品を製作する各工場で運用されているクライアントシステム４０１側では、将来起こるかもしれない不具合の予測情報を使用することで、不具合の発生を未然に防ぐことができる。また、上記各会社や各工場では、不具合発生前から不具合に備えることができる。

【０１０２】

次に、本発明による欠損値予測装置の最小構成の例を説明する。図８は、本発明による欠損値予測装置の最小構成の例を示すブロック図である。本発明による欠損値予測装置８１は、２つの因子（例えば、縦軸と横軸の因子）を含む行列形データ（例えば、行列形データＹ）における一方の因子の因子要素（例えば、縦軸の因子の因子要素）ごとに定義される行列であって、その行列形データにおけるもう一方の因子の各因子要素（例えば、横軸の因子の各因子要素）の特徴を表す行列である因子行列（例えば、因子行列Ｘ）を変換する関数（例えば、関数Ｆ）のパラメータのうち、その関数が因子行列を変換したときのデータが行列形データである尤もらしさを最大にするパラメータを推定する（例えば、式５を用いて推定する）パラメータ推定手段８２と、パラメータ推定手段８２が推定したパラメータ（例えば、Ｗ、Ｘ）及び行列形データにおける既知の行列要素の値を用いて、行列形データにおける行列要素（例えば、式１１におけるy_{ij}）の欠損値を予測する（例えば、式１１により欠損値を予測する）欠損値予測手段８３とを備えている。

【０１０３】

そのような構成により、行列形データにおける欠損値の予測精度を向上できる。具体的には、元の行列形データが関数と因子行列とを用いて表されるため、元の行列形データに対して条件付独立の制約がなくなるため、予測精度が向上させることができる。

【０１０４】

なお、少なくとも以下に示すような欠損値予測装置も、上記に示すいずれかの実施形態に開示されている。

【０１０５】

（１）２つの因子（例えば、縦軸と横軸の因子）を含む行列形データ（例えば、行列形データＹ）における一方の因子の因子要素（例えば、縦軸の因子の因子要素）ごとに定義される行列であって、その行列形データにおけるもう一方の因子の各因子要素（例えば、横軸の因子の各因子要素）の特徴を表す行列である因子行列（例えば、因子行列Ｘ）を変換する関数（例えば、関数Ｆ）のパラメータのうち、その関数が因子行列を変換したときのデータが行列形データである尤もらしさを最大にするパラメータを推定する（例えば、式１０を用いて推定する）パラメータ推定手段と、パラメータ推定手段が推定したパラメータ（例えば、Ｗ、Ｘ）及び行列形データにおける既知の行列要素の値を用いて、行列形データにおける行列要素（例えば、式１１におけるy_{ij}）の欠損値を予測する（例えば、

式１１により欠損値を予測する）欠損値予測手段とを備えた欠損値予測装置。

【０１０６】

（２）パラメータ推定手段が、ガウス分布の共分散行列（例えば、K^{WX}）を用いて定義された関数の因子行列における確率分布（例えば、式３）に基づいて、その関数が因子行列を変換したときのデータが行列形データである尤もらしさを最大にする関数のパラメータを推定する欠損値予測装置。

【０１０７】

（３）パラメータ推定手段が、ガウス分布の共分散行列を共分散を表す２つの行列のクロネッカー積（例えば、推定対象パラメータ）として表現した確率分布（例えば、式４）に基づいて、その関数が因子行列を変換したときのデータが行列形データである尤もらしさを最大にする関数のパラメータを推定する欠損値予測装置。

【０１０８】

（４）パラメータ推定手段が、関数の因子行列における確率分布と、関数及び因子行列における行列形データの確率分布とにより決定される周辺尤度の値を最大化する関数のパラメータを推定する欠損値予測装置。

【０１０９】

（５）パラメータ推定手段が、行列形データが２つの因子として製品及びその製品の不具合を示す情報を含む場合に、その行列形データにおける一方の因子である製品を示す情報ごとに定義される行列であって、その行列形データにおけるもう一方の因子であるその製品の不具合を示す情報の特徴を表す因子行列を変換する関数のパラメータうち、その因子行列を変換したときのデータが行列形データである尤もらしさを最大にするパラメータを推定し、欠損値予測手段が、推定されたパラメータ及び行列形データにおける既知の不具合箇所の情報を用いて、製品の未発生の不具合箇所の情報を予測する欠損値予測装置。

【産業上の利用可能性】

【０１１０】

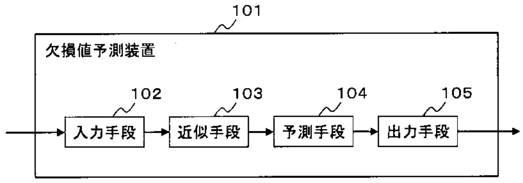
本発明は、２つの因子を含んだ行列形データにおける未知の行列要素の値を予測する欠損値予測システムに好適に適用される。

【符号の説明】

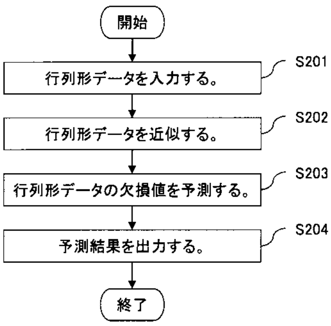
【０１１１】

１０１　欠損値予測装置
１０２　入力手段
１０３　近似手段
１０４　予測手段
１０５　出力手段
４０１　クライアントシステム
４０２　通信ネットワーク
４０３　サーバシステム
４０４　予測結果記憶部

【図１】



【図２】



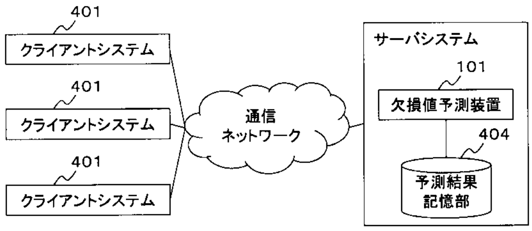
【図３】

ユーザ

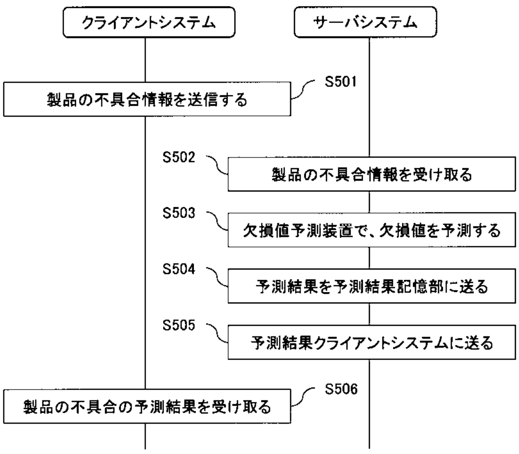
映画

	1	2	3	4	5	・・・	28	29	30
A	1	2	1	4	1	・・・	5	3	
B	3	1	5	2	2	・・・	5	5	
C	4	3	4		0	・・・		3	1
D	1		3		1	・・・	1	0	5
E				0		・・・	5	0	

【図４】



【図５】



【図６】

製品

不具合

	1	2	3	4	5	・・・	28	29	30
A	1	2	1		1	・・・	1	3	
B	3	1	5	2	2	・・・	5	5	
C	4	3	4		0	・・・	1	3	1
D			3	1	1	・・・	1	1	5
E		5		5		・・・	5	1	1

【図７】

製品

不具合

	1	2	3	4	5	・・・	28	29	30
A	1	2	1	<u>0.5</u>	1	・・・	1	3	<u>2</u>
B	3	1	5	2	2	・・・	5	5	<u>2</u>
C	4	3	4	<u>2.5</u>	0	・・・	1	3	1
D	<u>1.01</u>	<u>2</u>	3	1	1	・・・	1	1	5
E	<u>0.3</u>	5	<u>2.0</u>	5	<u>5</u>	・・・	5	1	1

【図８】