

令和 4 年 5 月 13 日現在

機関番号：32663

研究種目：若手研究

研究期間：2018～2021

課題番号：18K13953

研究課題名（和文）ロバスト・スパース・モデリングに基づく実用的な異常検出法の開発

研究課題名（英文）Development of practical anomaly detection based on robust sparse modeling

研究代表者

大久保 豪人（Ohkubo, Masato）

東洋大学・経営学部・講師

研究者番号：40777976

交付決定額（研究期間全体）：（直接経費） 2,700,000円

研究成果の概要（和文）：センサー・データを対象とした異常検出問題を考える。センサー・データは人の手を介さずリアルタイムに自動で取得・蓄積されるため、学習データに大量の異常データが混入している可能性がある。そのような状況下では、従来の統計的モデリング法を適用しても適切な分析ができるとは限らず、異常検出の性能が著しく低下してしまう危険性がある。そこで、混入した異常データの影響を受けることなく統計モデルを推定できるロバスト・スパース・モデリングを応用した統計的異常検出法の開発に向けた理論研究を行う。

研究成果の学術的意義や社会的意義

統計的異常検出法は製造業における設備機器の状態監視保全の中核をなす技術であるだけでなく、様々な製品・サービスに応用され、安心・安全な社会システムの構築に重要な役割を担っている。本研究の成果である統計的異常検出法は、特にセンサーから自動で取得・蓄積されたデータを対象とした場合の異常検出性能を飛躍的に向上させるとともに、その原因の特定に有益な情報を同時に提供するものである。この研究成果により、統計的異常検出法の応用可能性が広がり、設備機器の故障予測や重篤な事故の未然防止等の様々な社会問題の解決につながることを期待される。

研究成果の概要（英文）：We consider the problems accompanied with the anomaly detection for sensor data. Since sensor data is automatically acquired and accumulated in real time, there is a possibility that a large amount of anomaly data is mixed in the learning data. Such a learning data can bring significant reduction of the performance of anomaly detection, even if the conventional statistical modeling method is applied. Therefore, we conduct theoretical research so that we apply robust and sparse modeling that can estimate statistical models without being affected by mixed anomaly data to our statistical anomaly detection procedures.

研究分野：複合領域

キーワード：異常検出 ロバスト統計 スパース・モデリング タグチメソッド MTシステム

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

近年、製造業や医療などの様々な分野において、統計的異常検出法を応用したサービスが次々と実用化され、安心・安全な社会システムの構築に重要な役割を担うようになっている。統計的異常検出法は未知の異常を検出できるという特徴をもつため、設備機器の故障予測や突発的な体調変化の検知などに有用であることが知られていた。研究開始当初はいわゆる「モノのインターネット (IoT)」技術の黎明期であり、センサーやスマートデバイスから保全データや生体データなどが容易に収集・蓄積できる環境が整いつつあった。そして、それらのデータを活用し、これまで検出できなかった異常の兆しをつかみ、設備機器の故障や重篤な事故の未然防止などにつなげるための研究開発が産学の各所で始まっていた。

しかしながら、センサーやスマートデバイスから収集・蓄積されたデータ(履歴データ)に基づいて異常の兆しを検出しようとする場合、統計的異常検出法における統計モデルの推定プロセスを改良する必要性が生じる。履歴データは人の手を介さずリアルタイムに自動で取得・蓄積されるため、学習データに大量の異常データが混入している可能性がある。そのような状況下では、従来の統計的モデリング法を適用しても適切な分析ができるとは限らず、異常検出の性能が著しく低下する危険性もあった。ゆえに、統計的モデリング法を改良し、混入した異常データの影響を受けることなく異常検出可能な手法の開発が急務となっていた。

2. 研究の目的

報告者はこれまでも統計的異常検出法の研究開発に従事し、国内外の学会で高い評価を得てきた。その成果の一つにスパース・モデリングを応用した統計的異常検出法があるが、この手法は異常の検出精度だけでなく、統計モデルの解釈容易性にも焦点を当てた実用的な異常検出法となっている。この手法を適用すれば、異常を検出するだけでなく、その原因の特定に有益な情報を同時に獲得できる。しかしながら、殊に履歴データを解析する場合、前述の問題が発生し得るため、異常検出法の改善が必要となっていた。

そこで本研究では、「統計モデルの解釈容易性」と「大量の異常データの混入に対するロバスト性」を両立する統計的モデリング法を応用した新たな異常検出法を開発することを目的とした。具体的には、Hirose et al. (2017)のロバスト・スパース・モデリングを異常検出に応用することを考える。ロバスト・スパース・モデリングは、Fujisawa & Eguchi (2008)のガンマダイバージェンスに基づくロバスト推定法を発展させた統計的モデリング法であり、統計モデルをロバストかつスパースに推定することが可能となる。

なお、当初の計画以上に進展したため、より発展的なテーマについての研究も実施している。具体的には、履歴データ解析法の実用性を高めるため、特に製造業における設備機器の状態監視保全に統計的異常検出法が用いられる状況を想定した開発を行っている。

3. 研究の方法

本研究では、主に報告者の所属する大学での研究と、共同研究者の所属する研究機関(早稲田大学、統計数理研究所など)や日本品質管理学会のテクノメトリクス研究会(年4回開催)での議論を中心として研究を進めた。ただし、実務的な動向を把握するため、報告者が招聘研究員を委嘱されている早稲田大学データサイエンス研究所のシンポジウムや国内外の研究発表会などの場を利用して、定期的の実務家とのディスカッションも行っている。具体的な研究の進め方については以下の通りとした(開発する手法ごとに次の三段階を繰り返す)。

- (1) 各種統計的方法に関する数理的な検討
- (2) 実データ解析による性能評価
- (3) モンテカルロ・シミュレーションによる再現性の確認

まず、(1)では、主に製造業における設備機器の状態監視保全への適用を想定したうえで、需要に合わせた各種統計的方法を選定もしくは開発し、その数理的な性質と異常検出問題への応用可能性について議論した。次に、(2)では、統計的異常検出法のプロトタイプを開発したうえで、その性能を評価した。特に、本研究の中核となる手法については、「統計モデルの解釈容易性」と「大量の異常データの混入に対するロバスト性」を定量的に評価した。最後に、(3)では、異なるデータに対しても(2)と同様の性能が発揮できるか評価するためにモンテカルロ・シミュレーションを通して、再現性を確認した。

4. 研究成果

本研究では、主にマハラノビス・タグチ(MT)法と呼ばれる統計的方法の改良を通して、所望の統計的異常検出法の開発を行った。MT法とは製造業を中心として広く活用されている実用的な異常検出法である。以下、科研費の助成を受けて実施された研究成果について報告する。

4.1. ロバスト・スパース・モデリングを応用した異常検出法に関連する成果

「2. 研究の目的」に示したロバスト・スパース・モデリングを応用した異常検出法に関連する研究成果を示す。ここで、三番目の研究成果はロバスト・スパース・モデリングを応用したものではないが、同種の問題において新たな視座を与えるものと捉えている。

- (1) 本テーマの実施に当たり、報告者らのこれまでの研究成果の一つであるガンマダイバージェンスに基づくロバスト推定法を応用した異常検出法を実務的な観点から再評価を行った。これにより、高汚染データを対象として、異常検出プロセスを開発する意義が明確になった。なお、本研究の成果は査読付き原著論文として国際誌に掲載されている。
- (2) Hirose et al. (2017)のロバスト・スパース・モデリング法を応用した異常検出プロセスを開発した。本研究により、これまでの報告者らの研究成果(大久保・永田 (2017))における課題であった大量の異常データ混入の問題を解決できた。すなわち、「統計モデルの解釈容易性」と「大量の異常データの混入に対するロバスト性」を両立する異常検出法を確立できた。なお、本研究の成果の一部を国際会議にて発表している。
- (3) ノイズを伴う履歴データに対して、「統計モデルの解釈容易性」を保ちながら異常を検出するため、大久保・永田 (2017)のガウシアン・グラフィカル・モデルを応用した新たな異常検出法を提案した。これまでの研究成果から得た知見は、データがノイズを伴って観測される場合においても有用であり、異常検出法の発展に寄与するものである。なお、本研究の成果の一部を国際会議にて発表している。また、本研究の最終的な成果は査読付き原著論文として国際誌に掲載されている。

4.2. 異常検出法全般に関連する成果

製造業における設備機器の状態監視保全への適用を想定した異常検出法を3つ開発した。一つ目は一般の統計的異常検出法に関する成果、二つ目以降はMT法に関する成果である。

- (1) Song et al. (2007)の条件付き異常検出法を、彼らとは異なる問題設定のもとで実現する条件付き異常検出法を提案した。提案した異常検出法は、環境を示す補助変数を利用することにより、環境変化に柔軟に対応できる文脈型の異常検出法である。なお、本研究の成果の一部を国際会議にて発表している。また、本研究の最終的な成果は査読付き原著論文として国際誌に掲載されている。
- (2) 運転状況などに応じた複数の正常モードをもつデータからの異常検出を可能とするフレームワークを開発した。このフレームワークのもとでは、学習データに対して、複数の正常モードに対応するラベルを付与する必要があるものの、これまでの研究成果を複数の正常モードをもつデータに対して容易に展開することが可能となる。なお、この成果の一部を国際会議にて発表し、Best Paper Awardを受賞した。また、本研究の最終的な成果は査読付き原著論文として国際誌に掲載されている。
- (3) 報告者らの研究成果の一つである高次元データからの異常検出をより適切に実行できる異常検出法を新たな観点から再評価した。ノイズ掃き出し法などの高次元主成分分析を応用した場合の異常検出法との比較研究を行うとともに、手法の拡張について一定の方向性を見出すことができた。なお、この成果の一部を国際会議にて発表している。
- (4) 異常検出法を統計的パターン認識の枠組みに則って実行する方法について議論を行った。その予備的な研究として、MT法の関連手法の一つであるタグチのT法を分類問題に応用する方法を開発した。この成果を今後、分類問題ではなく異常検出問題に適用可能な方法に発展させる予定である。なお、この成果の一部を国際会議にて発表している。

4.3. その他の成果

国際会議における企画セッションの招待を受け、講演を行うなど、国内外への研究成果の発信も積極的に行った。具体的には、MT法とその関連手法に関する総合的な研究成果の報告を国際会議における企画セッションにて行った。また、国内の学術雑誌からの招待を契機に、これまでの研究成果に関する総合的な報告を論文にまとめた結果、公刊されている。

【参考文献】

- Fujisawa, H., & Eguchi, S. (2008): Robust parameter estimation with a small bias against heavy contamination. *Journal of Multivariate Analysis*, **99**(9), 2053-2081.
- Hirose, K., Fujisawa, H., & Sese, J. (2017): Robust sparse Gaussian graphical modeling. *Journal of Multivariate Analysis*, **161**, 172-190.
- 大久保豪人・永田靖. (2017): グラフィカル・モデリングに基づくマハラノビス・タグチ法. *応用統計学*, **46**(1), 13-26.
- Song, X., Wu, M., Jermaine, C., & Ranka, S. (2007): Conditional anomaly detection. *IEEE Transactions on Knowledge and Data Engineering*, **19**(5), 631-645.

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 3件/うち国際共著 0件/うちオープンアクセス 2件）

1. 著者名 Ohkubo Masato and Nagata Yasushi	4. 巻 24
2. 論文標題 Anomaly Detection for Noisy Data with the Mahalanobis-Taguchi System	5. 発行年 2020年
3. 雑誌名 Quality Innovation Prosperity	6. 最初と最後の頁 75 ~ 75
掲載論文のDOI (デジタルオブジェクト識別子) 10.12776/qip.v24i2.1441	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Masato Ohkubo and Yasushi Nagata	4. 巻 -
2. 論文標題 Anomaly detection for unlabelled unit space using the Mahalanobis-Taguchi system	5. 発行年 2019年
3. 雑誌名 Total Quality Management & Business Excellence	6. 最初と最後の頁 pp.1-15
掲載論文のDOI (デジタルオブジェクト識別子) 10.1080/14783363.2019.1616542	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Masato Ohkubo and Yasushi Nagata	4. 巻 30(sup1)
2. 論文標題 Conditional anomaly detection based on a latent class model	5. 発行年 2019年
3. 雑誌名 Total Quality Management & Business Excellence	6. 最初と最後の頁 pp. S227-S239
掲載論文のDOI (デジタルオブジェクト識別子) 10.1080/14783363.2019.1665847	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 大久保豪人	4. 巻 13(2)
2. 論文標題 マハラノビス・タグチ・システムにおける高次元データ解析法の展開	5. 発行年 2019年
3. 雑誌名 横幹	6. 最初と最後の頁 pp. 117-122
掲載論文のDOI (デジタルオブジェクト識別子) 10.11487/trafst.13.2_117	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計9件（うち招待講演 3件 / うち国際学会 4件）

1. 発表者名 大久保豪人・永田靖
2. 発表標題 ロバスト・スパース・グラフィカル・モデリングに基づくマハラノビス・タグチ法
3. 学会等名 日本品質管理学会 第119回研究発表会
4. 発表年 2019年

1. 発表者名 Masato Ohkubo and Yasushi Nagata
2. 発表標題 Anomaly detection for noisy data with the Mahalanobis-Taguchi system
3. 学会等名 22th QMOD conference on Quality and Service Sciences ICQSS (国際学会)
4. 発表年 2019年

1. 発表者名 Masato Ohkubo
2. 発表標題 THE MAHALANOBIS-TAGUCHI SYSTEM BASED ON STATISTICAL MODELING
3. 学会等名 The Fifth International Conference on the Interface between Statistics and Engineering (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 大久保豪人・永田靖
2. 発表標題 潜在クラスモデルに基づくマハラノビス・タグチ法
3. 学会等名 日本品質管理学会 第116回研究発表会
4. 発表年 2018年

1. 発表者名 Masato Ohkubo and Yasushi Nagata
2. 発表標題 Conditional anomaly detection based on a latent class model
3. 学会等名 21th QMOD conference on Quality and Service Sciences ICQSS (国際学会)
4. 発表年 2018年

1. 発表者名 大久保豪人
2. 発表標題 グラフィカル・モデリングに基づくマハラノビス・タグチ法
3. 学会等名 2018年度統計関連学会連合大会 (招待講演)
4. 発表年 2018年

1. 発表者名 Masato Ohkubo, Hironori Fujisawa, and Yasushi Nagata
2. 発表標題 Anomaly detection with Mahalanobis-Taguchi method based on robust sparse graphical modeling
3. 学会等名 Asian Network for Quality (ANQ) Congress 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 大久保豪人
2. 発表標題 統計的モデリングに基づくマハラノビス・タグチ・システム
3. 学会等名 第9回横幹連合コンファレンス (招待講演)
4. 発表年 2018年

1. 発表者名 大久保豪人・永田靖
2. 発表標題 ベイズの主成分分析を応用したマハラノビス・タグチ法
3. 学会等名 日本品質管理学会 第48回年次大会研究発表会
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------