

## 科学研究費助成事業 研究成果報告書

平成 26 年 6 月 2 日現在

機関番号：15401

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23510171

研究課題名(和文) 動的システムに対する縮小推定アルゴリズムの開発とその実用化に関する研究

研究課題名(英文) A study on development and implementation of shrinkage estimation algorithms for dynamic systems

研究代表者

土肥 正(Dohi, Tadashi)

広島大学・工学(系)研究科(研究院)・教授

研究者番号：00243600

交付決定額(研究期間全体)：(直接経費) 4,300,000円、(間接経費) 1,290,000円

研究成果の概要(和文)：本研究では、信頼性理論や性能評価理論で頻出する動的確率システムの新しい統計的推定技法として縮小推定の概念に着目し、尤度情報に基づいた汎用的なパラメトリック推定や核推定法を用いたノンパラメトリック推定よりも、高精度かつ低い計算コストで推定を実行するための理論的枠組みを構築する。特にウェーブレット縮小推定について考察し、高次元の時間非定常確率過程に対するいくつかの推定アルゴリズムを提案する。さらに、ウェーブレットに基づいた統計的推定アルゴリズムを実装した汎用的な統計解析ツールを世界に先駆けて開発し、いくつかの個別研究領域における応用問題に適用した事例について詳細に検討する。

研究成果の概要(英文)：In this research we focus on the concept of shrinkage estimation as a novel statistical inference technique in dynamic stochastic systems, which is frequently used in dependability and performance evaluation. We develop a theoretical framework to perform the statistical estimation with more accuracy and cheaper computation cost, than the common approaches such as parametric estimation based on the likelihood information or non-parametric estimation via kernel functions. Especially, we consider the wavelet-based shrinkage estimation, and propose several statistical estimation algorithms for high-dimensional non-stationary stochastic processes. Further, we develop the statistical analysis tool to implement the wavelet-based statistical estimation algorithms, and investigate several application examples in individual research area.

研究分野：信頼性工学

科研費の分科・細目：社会・安全システム科学・社会システム工学・安全システム

キーワード：縮小推定 ウェーブレット 確率過程 統計的推論 推定アルゴリズム 非定常点過程 マルコフ過程  
信頼性評価

## 1. 研究開始当初の背景

確率過程の推定問題は理論統計学や計算統計学における中心的な課題であり、様々な確率過程に対して多種多様な統計的推定アルゴリズムが提案されている。それらは大きく頻度論的枠組みとベイズ論的枠組みに大別され、さらに確率過程を特徴づけるパラメータに関する情報からパラメトリック推定とノンパラメトリック推定に分類できる。頻度論的アプローチにおけるパラメトリック推定法の中でも特に、大標本を前提とした各種漸近的性質や豊富な適合性評価法(モデル選択法)の整備状況から、最尤法は最も合理的な統計的推定法であると考えられている。

しかしながら、観測されるデータ形式によっては最尤法がうまく機能しない(統計的一致性を持たないなど)ことが頻繁にあり、必ずしも最尤推定量が動的システムの推論に適しているわけではない。また尤度関数を最大にする問題は、多くの場合、多峰性を有する非線形関数の最適化問題に帰着され、準ニュートン法などの汎用的な高速アルゴリズムや大域的収束性を持つ EM (Expectation-Maximization) アルゴリズムを用いたとしても計算コストが膨大となり、常に実用的な求解が可能であるとは限らない。またカーネル法などのノンパラメトリック推定では、カーネル関数のウィンドウ幅を求めるためにクロスバリデーション法を用いるが、膨大な計算コストがかかるだけでなく、計算効率が推定精度に極端に敏感であり、かつデータへの適合性も一般的には低い。

一方、近年になって、学習理論の進展や大規模なデータマイニングの必要性から、確率過程のベイズ推定に注目が集まっており、マルコフ連鎖モンテカルロ(MCMC)などの計算統計論的技術を援用して高速にパラメータの事後分布を求める方法が整備さ

れてきた。同様に、ディレクレ過程やガンマ過程によって事前情報を表現するノンパラメトリックベイズ推定も、MCMC に基づいたシミュレーション技術の発展によりかなり実用化される目途がたってきた。しかしながら、依然として計算コストは高く、実時間内でのオンライン推定が必要な状況ではこれらの方法を適用することが困難である。

良く知られているように、ウェーブレット解析は画像処理や信号処理などの分野で頻繁に用いられており、近年、高速演算と高精度が要求される時系列解析においても適用され始めている。確率過程のノンパラメトリック推定を行うウェーブレット推定法は90年代初頭から徐々に整備され始め、非定常ポアソン過程などの比較的単純な構造を持つ点過程に対してはほぼ確立されたと言っても過言ではない。これに対して、非定常マルコフ過程、時間に依存した係数項を持つブラウン運動、複合ポアソン過程や拡散ジャンプ過程などのレヴィ過程に応用した例は少なく、今後様々な確率過程に対してウェーブレット縮小推定の理論を整備するとともに、その実務的有効性を検証する必要がある。ウェーブレット推定の基本原理は、ノイズを含んだ観測データをウェーブレット級数で表現した後にウェーブレット係数を抽出し、閾値法によってノイズ除去された係数の逆ウェーブレット変換をとることで推定を行うというものである。具体的に、ウェーブレットはスケーリング関数とウェーブレット関数の2種類の関数から生成される基底関数を持ち、解析対象関数を基底関数の線形結合で表現する。スケーリング関数から作られる基底関数はファザーウェーブレット、ウェーブレット関数から作られる基底関数はマザーウェーブレットと呼ばれ、それぞれ解析対象関数の滑らかさと細かさを表す。フーリエ変換で

用いられる三角関数と異なり，これらのウェーブレット基底関数は時間軸上で平行移動や拡大縮小するパラメータを持ち，解析対象関数の大域的な特性と局所的な細かい挙動を同時に捉えることができる．このため，閾値法によってノイズ除去された係数を線形結合式に代入することで，解析対象関数の背後に隠された真の関数を推定する．

非線形回帰モデルによって表現される非定常多次元確率過程のウェーブレット縮小推定に対して，David L. Donoho と Iain Johnstone は一連の研究の中で，閾値法の効果やそれらの統計的合理性に関するいくつかの数理的な特性を証明している．最近，申請者らの研究グループはハール (Harr) ウェーブレットを基本とし，Anscombe 変換と Fisz 変換を適用してポアソンデータを正規データに変換することで，非定常ポアソン過程の強度関数を推定するノンパラメトリック推定をソフトウェア信頼性評価に適用し，従来から用いられてきた最尤法や最小二乗法を凌駕する適合性を示すことを示した．これまでのウェーブレット縮小推定の応用事例は主に物理学で見受けられ，例えばバースト性をもつガンマ線の動的性質を調査するために利用されていた．その意味において，ソフトウェア信頼性工学においてウェーブレット縮小推定の絶大な効用が認められ，その他の工学的分野における応用可能性が期待できる段階がついに到来したと言える．

## 2. 研究の目的

本研究では，非定常マルコフ過程，時間非定常なドリフト項や拡散項を持つブラウン運動，複合ポアソン過程，拡散ジャンプ過程に対して適用可能なウェーブレット縮小推定の理論的枠組みを整備することを目的とし，種々のウェーブレット関数，閾値法（ハード閾値法とソフト閾値法の選択と閾値水準の決定法），データ変換法（対象

確率過程の分布族が既知である場合と未知の場合）の多角的な観点から，各確率過程とそのデータ構造のバリエーションに対して複数のウェーブレット縮小推定アルゴリズムを提案する．先にも述べたように，ポアソン過程以外のウェーブレット縮小推定アルゴリズムはまだほとんど整備されていない状況であり，信頼性，情報セキュリティ，性能評価，トラヒック理論，数理ファイナンスにおける諸分野に適用するためには，アルゴリズム自体の開発とさらなる理論的な解析が必要である．頻度論的アプローチに加え，ウェーブレット縮小推定をベイズ推論に適用することも可能であり，異なる統計パラダイム上でウェーブレット縮小推定アルゴリズムを提案する．

ウェーブレット縮小推定の実用化のためには，各種推定アルゴリズムを実装した汎用的な統計ツールを開発することが不可欠であるため，本研究の理論的側面におけるひとつの到達目標は，この汎用的なウェーブレット縮小推定ツール(フリーウェアとしてのウェブツール)を世界に先駆けて開発し，ウェーブレット縮小推定の発展と啓蒙に寄与することである．さらに，各応用分野（信頼性，情報セキュリティ，性能評価，トラヒック理論，数理ファイナンス）で設定されているベンチマークテストを通じて，推定アルゴリズムの実験的評価や他の推定法との包括的な比較を行うことが必要である．具体的には，ポアソン過程やマルコフ過程に基づいたソフトウェア信頼性モデルの推定と信頼性評価尺度の導出，コンピュータ端末のウィルス感染台数予測，複合ポアソン過程によるインターネットトラヒックデータの解析，CDMD 移動体通信網モデルの解析，株式オプションに代表される条件付き請求権の価格推定などの研究領域において，ウェーブレット縮小推定の有効性について検証を行う．

### 3. 研究の方法

本研究課題では、(1) 各種確率過程に対応したウェーブレット縮小推定アルゴリズムの開発と理論的な解析、(2) 信頼性、情報セキュリティ、性能評価、トラヒック理論、数理ファイナンスの各研究領域におけるウェーブレット縮小推定の応用とその評価、(3) ウェーブレット縮小推定のための汎用ツール（フリーウェアとしてのウェブアプリケーション）の開発、を3つの柱として研究を遂行した。理論的な解析においては、ハード閾値法とソフト閾値法に対応した閾値水準の決定法と、結果として生じるウェーブレット推定量の漸近的性質について調査した。また、Kolaczyk や Jansen によって考案されたベイズ閾値法を拡張し、事後分布を導出するエンピリカル・ベイズ法をポアソン過程以外の確率モデルに応用するための理論を新たに確立した。

(1) アルゴリズム設計と解析：2011年度は、非定常ポアソン過程、非定常マルコフ過程、複合ポアソン過程など主に点過程の推定問題に着目し、データ構造（時間データ、グループデータ）に応じた縮小推定アルゴリズムを開発・改良した。(2) 汎用推定ツールの開発：ツール開発において、データベースとの互換性検討やインタフェース設計を早期に行っておく必要があるため、2011年度から開発作業を開始した。(3) 信頼性・セキュリティ評価への応用：従前研究で世界に先駆けて開発されたソフトウェア信頼性評価におけるウェーブレット縮小推定アルゴリズムの拡張を行った。(4) アルゴリズム設計と解析：非定常ポアソン過程、非定常マルコフ過程、複合ポアソン過程に対するウェーブレットベイジアン縮小推定に関する理論的な解析を行った。また、点過程以外の確率過程として、ブラウン運動過程や拡散ジャンプ過程のウェーブレット縮小推定アルゴリズムを開発した。(5) 汎用推定ツールの開発：タスク(4)で開発されたアルゴリズムの実装を行い、タスク(2)で開発されたシステムの拡張を行った。システム開発の中でも主要な作業である推定カーネルの開発を行った。(6) 金融商品評価への応用：金融商品の価格付けはオンラ

インで行われなければならない、最尤法の代わりに一般化モーメント法が用いられることが多い。これに対して、パラメータを特定しないノンパラメトリック推定を高速に行うためにはウェーブレット縮小推定が非常に有効である。(7) 性能評価モデルの解析：自動車間通信システムの性能評価問題に着目し、ソフトハンドオーバーを伴うCDMD (Code Division Multiple Access) 移動体通信網のモデルである非定常マルコフ過程のオンライン推定について考察した。

### 4. 研究成果

アルゴリズム設計と解析において頻度論的アプローチのみに着目するため、ハード閾値法とソフト閾値法に対応した閾値水準の決定法を詳細に検討する必要がある、min-max 閾値、level-dependent 閾値、universal 閾値、leave-out-half クロスバリデーション閾値などの適用可能性について考察した。さらにハール系ウェーブレット以外にも、シャノンウェーブレット系やルマリエ-メイエウェーブレット系についても検討を加えた。縮小推定アルゴリズム自体は最適化アルゴリズムではないので、代数計算のみでアルゴリズムの記述が可能である。また、結果として得られる推定量に対する漸近的性質についても理論的な検討を加えた。性能評価理論の世界的権威である米国デューク大学の K. S. Trivedi 教授を訪問し、アルゴリズム設計と解析に関する討論を行った。また、ドリフト係数や拡散係数が時間の関数であるが未知の状況において、線形確率微分方程式で記述される動的システムのパラメータ推定問題について取り扱った。タスク(1)で求めた推定アルゴリズムと拡散極限（もしくは拡散近似）の関連を理論的に解析することで、ウェーブレット縮小推定を理解するための理論的枠組みが明らかになり、数

学的に興味深い結果が得られた。

汎用推定ツールの開発において、発言語は Java であり、システム全体としては 100 クラスで 40000 行程度の開発規模であった。推定エンジンにアルゴリズムを実装した後、各種適合性評価の計算値とデータの挙動を視覚的に表示する機能を標準装備した。このため、開発キットとして JDK1.5.0\_9 を利用し、チャートライブラリには JFreeChart を使用した。

信頼性・セキュリティ評価への応用では、大学院博士課程後期学生 X. Xiao との共同研究において、Anscombe と Fisz による正規変換を必要としない閾値法の下で一般的なアルゴリズム設計を行い、実際の開発プロジェクトで得られたフォールトデータに基づいてソフトウェアの信頼性評価を行った。次に、非定常ポアソン過程と非定常マルコフ過程によりウィルス感染台数予測を動的な系として記述し、感染初期の時期から将来における感染台数の予測や感染数爆発の確率を推定する問題を考えた。本研究では CodeRed などの過去のウィルス感染データを用いて統計解析を行い、インフルエンザウィルスなどの疫学的感染拡散現象の解析についても同様に考察することが可能であることがわかった。

金融商品評価への応用において、ウェブレット縮小推定の有効性を金融商品の価格付けに応用するためには、元資産に関する経済時系列データを解析する必要がある。数年単位の価格データを使った実証分析を行うことができた。

最後に性能評価モデルの解析では、パケット損失率・スルーット・遅延時間・回線使用率など、通信トラヒックの状態に関する変量の分布の裾野の振る舞いは通信品質・アプリケーション性能・ネットワーク設計などに大きな影響を与える。この問題にウェブレット縮小推定を適用すること

で、Gross-Error-Robustness などの性質を保障する推定方式を確立した。また研究代表者らの研究グループによって近年開発された、MAP と複合ポアソン過程に対する高速 EM アルゴリズムとの比較を通じて、計算効率と推定精度を比較した。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 21 件)

- (1) C. Luo, H. Okamura and T. Dohi, Performance evaluation of snapshot isolation in distributed database system under failure-prone environment, Journal of Supercomputing, 査読有, (in press).
- (2) X. Xiao and T. Dohi, Exponential-like software reliability models based on binomial process, International Journal of Reliability and Quality Performance, 査読有, (in press).
- (3) X. Xiao and T. Dohi, Interval estimation method for decision making in wavelet-based software reliability assessment, IEICE Transactions on Information & Systems (D), 査読有, (in press).
- (4) H. Okamura, R. Watanabe and T. Dohi, Variational Bayes for phase-type distribution, Communications in Statistics -Simulation and Computation, 査読有, vol. 43, no. 8, pp. 2031-2044, 2014. DOI:10.1080/03610918.2013.848895
- (5) X. Xiao and T. Dohi, Wavelet shrinkage estimation for NHPP-based software reliability models, IEEE Transactions on Reliability, 査読有, vol. 62, no. 1, pp. 211-225, 2013. DOI: 10.1109/TR.2013.2285055
- (6) X. Xiao and T. Dohi, Estimating software intensity function based on translation-invariant Poisson smoothing approach, IEEE Transactions on Reliability, 査読有, vol.62, no. 4, pp. 930-945, 2013. DOI:10.1109/TR.2013.2240897
- (7) X. Xiao and T. Dohi, A study on applying extreme value distribution to NHPP-based SRM, Information, 査読有, vol. 162, no. 1(B), pp. 575-580, 2013. <http://selab.hongik.ac.kr/down/journal/201301615-625.PDF>
- (8) X. Xiao and T. Dohi, A measurement-based approach for estimating error rate of a web server system, International Journal of Reliability and

Safety, 査読有, vol. 17, no. 1, pp. 1-16, 2013.  
DOI: 10.1504/IJRS.2013.055821  
(9) B. Zhou, H. Okamura and T. Dohi, Enhancing performance of random testing through Markov chain Monte Carlo methods, IEEE Transactions on Computers, 査読有, vol. 62, no. 1, pp. 186-192, 2013.  
DOI: 10.1109/TC.2011.208  
(10) H. Okamura and T. Dohi, Dynamic software rejuvenation policies in a transaction-based system under Markovian arrival processes, Performance Evaluation, 査読有, vol. 70, no. 3, pp. 197-211, 2013.  
<http://dx.doi.org/10.1016/j.peva.2012.07.004>  
(11) H. Okamura, T. Dohi and S. Osaki, Software reliability growth models with normal failure time distributions, Reliability Engineering and System Safety, 査読有, vol. 116, pp. 135-141, 2013.  
<http://dx.doi.org/10.1016/j.ress.2012.02.002>  
(12) H. Okamura, M. Tokuzane and T. Dohi, Quantitative security evaluation for software system from vulnerability database, Journal of Software Engineering and Applications, 査読有, vol. 6, no. 4A, pp. 15-23, 2013.  
DOI: 10.4236/jsea.2013.64A003  
(13) S. Hiroyama, T. Dohi and H. Okamura, Aperiodic checkpoint placement algorithms - survey and comparison -, Journal of Software Engineering and Applications, 査読有, vol. 6, no. 4A, pp. 41-53, 2013.  
DOI: 10.4236/jsea.2013.64A006  
(14) H. Okamura, T. Dohi and K. S. Trivedi, Improvement of expectation maximization algorithm for phase-type distributions with grouped and truncated data, Applied Stochastic Models in Business and Industry, 査読有, vol. 29, pp. 141-156, 2013.  
DOI: 10.1002/asmb.1919  
(15) S. Ukimoto and T. Dohi, A software cost model with reliability constraint under two operational scenarios, International Journal of Software Engineering and Its Applications, 査読有, vol. 7, no. 3, pp. 427-438, 2013.  
[http://onlinepresent.org/proceedings/vol19\\_2013/59.pdf](http://onlinepresent.org/proceedings/vol19_2013/59.pdf)  
(16) H. Okamura, T. Hirata and T. Dohi, Semi-parametric approach for software reliability evaluation using mixed gamma distributions, International Journal of Software Engineering and Its Applications, 査読有, vol. 7, no. 4, pp. 401-414, 2013.  
[http://www.sersc.org/journals/IJSEIA/vol7\\_no4\\_2013/35.pdf](http://www.sersc.org/journals/IJSEIA/vol7_no4_2013/35.pdf)

(17) H. Okamura and T. Dohi, Optimal trigger time of software rejuvenation under probabilistic opportunities, IEICE Transactions on Information & Systems (D), 査読有, vol. E96-D, no. 9, pp. 1933-1940, 2013.

[http://search.ieice.org/bin/summary.php?id=e96-d\\_9\\_1933](http://search.ieice.org/bin/summary.php?id=e96-d_9_1933)

(18) X. Xiao and T. Dohi, A comparative study of data transformations for wavelet shrinkage estimation with application to software reliability assessment, Advances in Software Engineering, 査読有, vol. 2012, pp. 1-9, 2012.

<http://dx.doi.org/10.1155/2012/524636>

(19) T. Dohi and T. Uemura, An adaptive mode control algorithm of a scalable intrusion tolerant architecture, Journal of Computer and System Sciences, 査読有, vol. 78, no. 6, pp. 1751-1774, 2012.

<http://dx.doi.org/10.1016/j.jcss.2011.10.022>

(20) X. Xiao and T. Dohi, Software failure time data analysis via wavelet-based approach, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences (A), 査読有, vol. E95-A, no. 9, pp. 1490-1497, 2012.

[http://search.ieice.org/bin/summary.php?id=e95-a\\_9\\_1490](http://search.ieice.org/bin/summary.php?id=e95-a_9_1490)

(21) B. Zhou, H. Okamura and T. Dohi, Application of Markov chain Monte Carlo random testing to test case prioritization in regression testing, IEICE Transactions on Information & Systems (D), 査読有, vol. E95-D, no. 9, pp. 2219-2226, 2012.

[http://search.ieice.org/bin/summary.php?id=e95-d\\_9\\_2219](http://search.ieice.org/bin/summary.php?id=e95-d_9_2219)

〔学会発表〕(計 57 件)

(1) X. Xiao and T. Dohi, Estimating software intensity function via multiscale analysis and its application to reliability assessment, The 17th IEEE Pacific Rim International Symposium on Dependable Computing (PRDC-2011), Los Angeles, USA, December 12-14, 2011.

## 6. 研究組織

### (1) 研究代表者

土肥 正 (DOHI Tadashi)  
広島大学・大学院工学研究科・教授  
研究者番号：00243600