

プロジェクト名： 機能と帰納

サブテーマ名： 予測とリスク解析

研究代表者

江口真透 [統計数理研究所]

鷺尾 隆 [大阪大学産業技術研究所・統計数理研究所]

足立 淳 [統計数理研究所]

椿 広計 [統計数理研究所]

1. 研究目標

(1) グラフマイニングを用いた因果ネットワーク発見

平成 19 年度は前年度の研究成果を受け、以下の項目を目標とした。

前年度に改良手法を更に高速な多頻度一般的部分 DAG 構造マイニング手法に拡張し、その拡張手法を計算機実装し、より高速な多頻度一般的部分 DAG 構造マイニングプログラム DIGDAG を開発する。上記拡張手法及び DIGDAG の性能を遺伝子ネットワークデータに適用し、大規模な埋め込み遺伝子ネットワーク構造を把握する。

(2) 遺伝子多様性の統計解析

バイオインフォマティクスの統計的課題である ‘ $p \gg n$ ’ 問題に挑戦する。バイオテクノロジーの急速な発展に伴い特徴次元（パラメータ数） p は膨大となる一方で、観測数 n は被験者のインフォームドコンセントの問題、さらには生命倫理の問題まで深刻に関連するので一定数より大きくすることは困難である。統計的な思考からありとあらゆるアプローチを尽くし、この問題の解決へ一歩でも近づくことを目指す。しかし、現実には一塩基多型、マイクロアレイ、プロテオームのデータが日々、世界的な規模で急速な勢いで得られている。理論と現実との整合性を保ちながら結果を出すことが私たちの急務である。ゲノムデータと表現形の相関研究について焦点を当てて研究をしている。ここでゲノムデータは遺伝子の多様性に関連する SNPs、マイクロアレイ、プロテオームを指し、表現形は臨床・病理医学的な知見で得られる疾病の進行度、治療の効果、薬剤の感受性、薬剤の奏功性に関する質的変量を指す。高次元のゲノムデータの中からどんな特徴的パターンが表現形の予測に関わっているかを探索するために有効な「統計的パターン認識」の方法を目指している。これは、私たちのグループで汎用な目的で開発されているブースティングの学習アルゴリズムをバイオインフォマティクスの文脈に沿った形にすることが肝要となる。

(3) 生物多様性の総合的理解を目指して

生物多様性を理解するためには、多様な生物の系統関係を正しくとらえることが必要である。しかしながら個々の生物の遺伝情報は進化の過程で様々な偏りが生じていることが最近になっ

て分かってきている。そのような偏りに対処し補正できる塩基，アミノ酸やコドンの新しい置換モデルと効率的な系統樹探索法を開発する。また，ゲノム情報等の大量データから系統樹を偏り無く推定するために，生物間で相同な遺伝子（オーソログ）を発見する方法も考える。その上で，哺乳類の系統進化を解析し生物多様性の理解を深めたい。

(4) リスク評価のための帰納推論と機能的統計解析のデザイン

リスクの定量的評価に関する分野横断型理論の構築と応用を，(1) 大量データからの知識発見を促す半自動的な帰納的推論と，(2) 調査・実験のコストが軽視できない領域での情報取得の機能的デザインの両面から，(3) 医薬品・食品，環境，金融・保険の3領域にわたり推進する。特に情報共有のためのリスク研究ネットワークの構築・活用によりプロジェクト型研究を担う人材を育成し，リスク評価に関する情報化時代の新たな科学的推論の形を提案する。

2. 年度研究計画

平成16年度(予備研究)

(1) グラフマイニングを用いた因果ネットワーク発見

平成16年度までに，グラフマイニング基礎手法の研究開発を行ってきた。また，gene ネットワークデータから因果ネットワークを導出する基礎理論に関する研究も行ってきた。生命の遺伝情報である遺伝子の塩基配列やアミノ酸配列から分子系統樹を推定する方法を開発してきた。しかし，より現実的なモデルの導入や，ゲノム時代に即した大量データの解析に対応する必要があった。

平成17年度

(1) グラフマイニングを用いた因果ネットワーク発見

gene ネットワークに代表される因果ネットワークにおいて，如何なる構造や基準に着目して不変な安定ネットワーク部分を抽出すべきかの検討を行う。そのために定量的強度付きネットワーク構造解析手法の検討，構造情報を直接扱うことが可能なグラフマイニングや木構造マイニングによる解析手法の検討を行う。

(2) 遺伝子多様性の統計解析

マイクロアレイデータから表現形への相関解析のためにアダプブーストを改良してグループ・ブーストを提案し，その予測性能について理論的考察と公開された幾つかの実データで検証した。

(3) 生物多様性の総合的理解を目指して

生物学上の問題解決をはかりながら，分子進化のモデリングと分子系統樹推定法の開発を進める。蛋白質コード領域の配列データから系統樹推定する際に，コドン単位の置換をモデル化したコドン置換モデルが有効であることを確かめる。具体的な生物学上の問題としては，長い間大陸から隔離され，独自の生物相を進化させたマダガスカルにおけるテンレック類，

原猿類，バオバブ類の多様化を系統進化の観点からとらえる研究を進める。

(4) リスク評価のための帰納推論と機能的統計解析のデザイン

リスク研究ネットワークを設立し，各分野で問題提起的なシンポジウムを開催。具体的には日本製薬工業協会との共催シンポジウム「医薬品の安全性情報を考える」の開催，ISM シンポジウム「統計科学と環境科学の新たな融合」開催，金融・保険における最新動向に関する国際ワークショップ「リスク管理における統計的技術」の開催。一般向け講演として「統計科学と保険の接点」と題したセミナーを開催。更に環境問題への直接的貢献窓口として NPO「環境統計統合機構」設立。PD 研究員を採用し，各分野でのプロジェクト研究の推進。

平成18年度

(1) グラフマイニングを用いた因果ネットワーク発見

前年度検討結果を受けて，遺伝子発現データにベイジアンネットワークを適用して得た gene ネットワークデータに，グラフマイニング及び木構造マイニング手法を適用し，導出される gene ネットワーク構造パターンの性質及び生物情報学的見地について詳細分析を行う。また，埋め込み木構造マイニング手法を DAG(Directed Acyclic Graph)で表される因果構造ネットワークデータに適用可能なように拡張し，同じく gene ネットワークデータに適用して詳細分析を行う。

(2) 遺伝子多様性の統計解析

SNP のハプロタイプブロックの特定のためのアルゴリズムを提案し，既存の方法と比較した。特に癌研究所で採られた SNP タイピングデータと臨床の情報を結合させ有用性を確認した。

(3) 生物多様性の総合的理解を目指して

真獣類はボレオ真獣類（北半球のローラシア大陸起源），アフリカ獣類（アフリカ起源），貧歯類（南アメリカ起源）の3つの主要なグループに分類でき，それらが大陸の分裂，移動と密接に関わっていることを解明する。また，コウモリは哺乳類のなかで特異な形態をもち，その起源については長い間なぞであったが，霊長目に近いという説が有力であった。ところが，霊長目とは類縁関係はなく，翼手目（コウモリ），奇蹄目（ウマ，サイ），食肉目（クマ，イヌ，ネコ）が進化的に1つのグループを構成していることが明らかにする。

(4) リスク評価のための帰納推論と機能的統計解析のデザイン

日本製薬工業会医薬品評価委員会と共同で種々の安全性データベース構築および活用に向けての検討を行なうとともに，降圧薬や経口抗菌薬などの使用成績調査データベース構築とその活用を進め，市販前の臨床試験のデータベースの設計に着手。研究会「保険と金融の統計学」を定期的に開催し，保険数学，マルチンゲール理論，極値理論に関する研究を実施。温室効果ガス観測技術衛星 GOSAT から二酸化炭素およびメタンのカラム濃度導出精度を評価する方法を研究。「地球温暖化現象とそのリスク予測」に関するシンポジウムを開催。

平成19年度

(1) グラフマイニングを用いた因果ネットワーク発見

埋め込み木構造マイニング手法を DAG(Directed Acyclic Graph)で表される因果構造ネットワークデータに適用可能にしたアルゴリズムを更に高速化, 広適用範囲化する。またその結果を受けて, gene ネットワークなどの因果ネットワークにおいて, 如何なる構造や基準に着目して特徴的ネットワーク部分を抽出すべきかの検討を行う。

(2) 遺伝子多様性の統計解析

質量スペクトルを測るプロテオームデータから疾病との相関を発見するために開発した「共通ピークアプローチ」を松浦氏の癌研究所・ゲノム解析センターの研究チームが行った実際の乳がん患者からのプロテオームデータに対して適用し, その有用性が拡大された。

(3) 生物多様性の総合的理解を目指して

ゲノムの大量データから系統樹推定を行う際には, サンプリング誤差は限りなく小さくなるが, 逆に推定の偏りに伴う誤差が顕著になる。ゲノムデータの解析を通じて真獣類の初期進化における系統関係を明らかにするために, 系統樹推定の偏りを少なくするためのさまざまなモデル化を試みる。また, ゲノム時代に即した大量データから哺乳類の系統進化を解析する。

(4) リスク評価のための帰納推論と機能的統計解析のデザイン

海外から専門家を招聘し, リスク解析に関する連続講義と共同研究を実施。クレジット・デフォルト・スワップ(CDS)取引データを用いた統計的モデリングによる, 信用リスクインデックスを作成。温室効果ガス観測技術衛星 GOSAT から二酸化炭素およびメタンのカラム濃度導出精度を評価するとともに, シンポジウムを開催。日本製薬工業会医薬品評価委員会と共同で種々の安全性データベース構築および活用に向けての検討を行い, 市販前の臨床試験のデータベースの設計に着手。

平成20年度

(1) グラフマイニングを用いた因果ネットワーク発見

前年度検討及び開発結果を受けて, gene ネットワークデータに開発したグラフマイニング及び木構造マイニング手法を適用し, 導出される gene ネットワーク構造パターンの性質について詳細分析を行う。

(2) 遺伝子多様性の統計解析

国立がんセンター研究所の乳腺・腫瘍内科グループとの共同研究によって乳がん患者グループの治療前, 治療初期, 治療後期の3時点のマイクロアレイを解析して, 時間変化によって予測因子と予後因子となる発現遺伝子を同定する。

(3) 生物多様性の総合的理解を目指して

近年の分子系統学においては, 全ゲノムデータを基盤とした解析が主流となりつつあるが, それでも哺乳類の中で解明されていない系統関係が複数残されている。本研究課題では, こ

これらの系統関係の解明を目的として、全ゲノムデータベースから大量の遺伝子配列を収集して統計的解析をおこなうことにより、全哺乳類の進化の道筋を明らかにする。

(4) リスク評価のための帰納推論と機能的統計解析のデザイン

クレジット・デフォルト・スワップ・インデックスの研究とウェブを通じた成果公開。中小企業信用リスク情報データベースが利用可能になり次第信用リスク推定に関するプロジェクトを開始。極値理論の金融リスク管理への応用に関するプロジェクトを開始。化学物質のリスクトレードオフに関する研究に着手。化学物質のリスク評価に関する ISM シンポジウムを開催。使用成績調査および臨床試験の大規模データのデータベースの構築を継続、統計的シグナル検出手法および定量的評価のための統計的手法の実務的な検討。

平成21年度

(1) グラフマイニングを用いた因果ネットワーク発見

これまでの開発及び検討結果を受けて、gene ネットワークなどの因果ネットワークデータから不変的及び特徴的な部分構造を発見する構造マイニング解析手法を体系化し、統一的な手法及びツールにまとめる。また、それをデータに適用して gene ネットワーク構造パターンの性質及び生物情報学的見地について詳細分析を行う。

(2) 遺伝子多様性の統計解析

国立循環器センター研究所の高血圧内科グループとの共同研究によって降圧剤の奏功性予測のための重要な関連を持つ SNP、及び SNP×SNP の発見をして予測キットを提案する。

(3) 生物多様性の総合的理解を目指して

分子系統樹を推定する上で、これまでに開発してきた様々な方法を、分子系統樹推定ソフトウェア MolPhy に実装し新たなバージョンをリリースする。また、最新のゲノムデータを使い、哺乳類以前の祖先生物の系統進化を解析し生物多様性の理解を深めたい。

(4) リスク評価のための帰納推論と機能的統計解析のデザイン

「保険と金融の統計学」に関する研究報告会を開催し、あわせて学術雑誌での特集号の編纂による研究成果公開を目指す。化学物質のリスクトレードオフに関する推論構築とそれを解くためのアルゴリズム開発の研究を推進する。また、これらに関係した ISM シンポジウムの開催を計画。使用成績調査および臨床試験の大規模データのデータベースの構築を継続するとともに、それらの大規模データベースに基づく統計的シグナル検出手法および定量的評価のための統計的手法についての研究成果を公開する。

平成22年度以降の展開

(1) グラフマイニングを用いた因果ネットワーク発見

本プロジェクトの成果として、gene ネットワークデータをはじめとして、DAG (Directed Acyclic Graph) で表される因果構造ネットワークデータから、事象間の広範な相互依存関係、因果関係に関する知識を発見する手法の確立の見通しが得られつつある。これを更に拡張し、科学、

学術，工学，社会全般に広く見られる DAG(Directed Acyclic Graph)で表される構造ネットワークデータから，機能をはじめとする知識を帰納的に導出する枠組みと手法，ツール開発を行っていく。

(2) 遺伝子多様性の統計解析

ゲノム・オミクスデータをバイオマーカーとして表現形との関連の高いものを同定し，オーダーメイド医療の実現に貢献したい。そのためにこれまでの成果を踏まえて機械学習，生物統計，バイオインフォマティクをまたがる分野を横断する新たなアプローチを開拓したい。

(3) 生物多様性の総合的理解を目指して

生命が持つゲノムは長い進化の歴史の産物であり，そこには突然変異に起因する進化の履歴が刻まれている。これまで生物間の相同な遺伝子の比較から系統関係を推定する分子系統学の発展により，生物の進化の道筋は系統樹として推定されてきたが，同時に幾つかの限界点も見えてきた。相同な遺伝子を比較するだけでは非常に短期間に種分化が起きた場合や，種分化の時に長い間にわたって種間交雑が続いた場合などでは，系統関係の推定が困難になってしまう。また系統関係がわかってもそれは進化の道筋が明確になっただけであり，進化のメカニズムについては何も知ることはできない。そこで，相同な遺伝子群の系統関係と，ゲノム上での互いの遺伝子の位置関係から，種間におけるゲノム構造の変異の歴史的順番を数値的最適化の手法を適用することによって再構築することを試みる。分子系統学とゲノム比較を高度に組み合わせることによって，ある遺伝子の進化の引き金となった突然変異が，ゲノム上で何時どのように起こったかを推定することができるようになる。こうして個々の突然変異が定着してきた歴史を解明することは，進化のメカニズムを知るための第一歩となるであろう。

(4) リスク評価のための帰納推論と機能的統計解析のデザイン

平成 21 年度までの研究成果として，異分野の交流からリスク解析の方法論での共通要素が抽出されることが期待される。そこで，医薬品・食品，環境，金融・保険という当初の応用分野主体の区分を見直し，方法論を軸とした研究班の構成にする。また，製品やサービスの質保証・信頼性という新たな切り口で定量的リスク解析への接近法を研究する。

3. 研究経費の推移

平成 17 年度実績： 23,700 千円

平成 18 年度実績： 34,290 千円

平成 19 年度見込： 34,860 千円

4. 平成 19 年度の研究実施体制

(1) グラフマイニングを用いた因果ネットワーク発見

[統計数理研究所] 樋口知之 Alexandre Termier

[情報学研究所] 佐藤 健 宇野毅明

[東京大学] 井元清哉
[大阪大学] 大原剛三

(2) 遺伝子多様性の統計解析

[江口真透・統計数理研究所]
[藤澤洋徳・統計数理研究所]
[伏木忠義・統計数理研究所]
[川喜田雅則・九州大学情報工学]
[松浦正明・(財) 癌研究会癌研究所 癌研究所・ゲノム解析センター] ★
[竹之内高志・奈良先端大学情報科学研究科] ★
[池田思朗・統計数理研究所]
[栗木 哲・統計数理研究所]
[南 美穂子・統計数理研究所]

(3) 生物多様性の総合的理解を目指して

[統計数理研究所 予測発見戦略研究センター] 曹 纓 佐々木 剛
[情報・システム研究機構 新領域融合センター] 松井 淳
[筑波大学 生物科学系] 橋本哲男
[東京工業大学 生命理工学研究科] 岡田典弘 西原秀典
[東京工業大学 情報理工学研究科] 下平英寿
[Fudan University, China] 長谷川政美 米澤隆弘

(4) リスク評価のための帰納推論と機能的統計解析のデザイン

[統計数理研究所] 椿 広計 藤田利治 金藤浩司 川崎能典 河村敏彦
柏木宣久 佐藤整尚 山下智志 志村隆彰 田野倉葉子 安藤雅和
[新領域融合研究センター] 河合研一 公文雅之 友定充弘
[東京大学] 国友直人 山本和夫
[京都大学] 佐藤俊哉
[岡山大学] 小野芳朗
[鹿児島大学] 高梨啓和 青木 敏
[(社)国際環境研究協会] 松本幸雄
[東京医療保健大学] 比江島欣慎

5. 平成19年度研究成果

(1) 成果物(知見・成果物・知的財産権等)

1. 拡張 DIGDAG アルゴリズム

埋め込み木構造マイニング手法を DAG (Directed Acyclic Graph) で表される因果構造ネットワークデータに適用可能なように拡張したアルゴリズム DIGDAG を, 更に高速化, 適用広範囲化し, より効率的にかつより複雑な一般飽和埋め込み有向非巡回グラフネットワークマイニングを行

なえるようにした。

2. 特徴的ネットワーク部分抽出

上記手法を gene ネットワークに関する因果ネットワークに適用し、特徴的ネットワーク部分の抽出を行なった。

3. クレジット・デフォルト・スワップ(CDS)取引データを用いた統計的モデリングによる、信用リスクインデックスを作成。

4. 温室効果ガス観測技術衛星 GOSAT から二酸化炭素およびメタンのカラム濃度導出精度を評価するとともに、シンポジウムを開催。

5. 日本製薬工業会医薬品評価委員会と共同で種々の安全性データベース構築および活用に向けての検討を行い、市販前の臨床試験のデータベースの設計に着手。

6. プロテオームの共通ピークパタン解析アルゴリズム

(2) 成果発表等

<論文発表>

[学術論文]

1. Alexandre Termier, Marie-Christine Rousset, Michele Sebag, Kouzou Ohara, Takashi Washio, and Hiroshi Motoda: DRYADEPARENT, An Efficient and Closed Attribute Tree Mining Algorithm, IEEE Transactions on Knowledge and Data Engineering, Vol. 20, No. 2 (February 2008)
2. Takashi Washio, Koutarou Nakanishi and Hiroshi Motoda: A Classification Method Based on Subspace Clustering and Association Rules, New Generation Computing, Vol.25, pp.235-245, Ohmsha, Ltd. and Springer (2007)
3. L. A. Pastene, M. Goto, A. N. Zerbini, K. Watanabe, Y. Bessho, D. Kerem, N. Kanda, M. Hasegawa, R. Nielsen, F. Larsen, and P. J. Palsboll (2007) Marine speciation during periods of global warming: timing of origin and radiation of the common minke whale, *Balaenoptera acutorostrata*. *Molecular Ecology* 16:1481-1500.
4. A. Matsui, F. Rakotondraparany, M. Hasegawa, and S. Horai (2007) Determination of a complete lemur mitochondrial genome from feces. *Mammal Study* 32: 7-16.
5. T. Yonezawa, M. Nikaido, N. Kohno, Y. Fukumoto, N. Okada, and M. Hasegawa (2007) Molecular phylogenetic study on the origin and evolution of Mustelidae. *Gene* 396:1-12.
6. Y. Zhang, N. Zheng, P. Nan, Y. Cao, M. Hasegawa, and Y. Zhong (2007) Computational

- simulation of interactions between SARS coronavirus spike mutants and host species-specific receptors. *Computational Biology and Chemistry* 31:134–137.
7. H. Nozaki, M. Iseki, M. Hasegawa, K. Misawa, T. Nakada, N. Sasaki, and M. Watanabe (2007) Phylogeny of primary photosynthetic eukaryotes as deduced from slowly evolving nuclear genes. *Mol. Biol. Evol.* 24:1592–1595.
 8. H. Nishihara, N. Okada, and M. Hasegawa (2007) Rooting the eutherian tree: the power and pitfalls of phylogenomics. *Genome Biology* 8:R199.
 9. M. Ushijima, S. Miyata, S. Eguchi, M. Kawakita, M. Yoshimoto, T. Iwase, F. Akiyama, G. Sakamoto, K. Nagasaki, Y. Miki, T. Noda, Y. Hoshikawa and M. Matsuura. Common peak approach using mass spectrometry data sets for predicting the effects of anticancer drugs on breast cancer. *Cancer Informatics*, 3 (2007) 285–293.
 10. M. Henmi, R. Yoshida and S. Eguchi. Importance sampling via the estimated sampler. *Biometrika* 94, 4, 985–991 (2007).
 11. H. Fujisawa, M. Isomura, S. Eguchi, M. Ushijima, S. Miyata, Y. Miki, M. Matsuura. Identifying haplotype block structure by using ancestor-derived model. *J. Human Genetics* 52 (9) (2007) 738–746.
 12. T. Kanamori, T. Takenouchi, S. Eguchi and N. Murata. Robust loss functions for boosting. *Neural Computation* 19 (2007) 2183–2244.
 13. M. Henmi, J. Copas and S. Eguchi. Confidence intervals and P-values for meta analysis with publication bias. *Biometrics* 63 (2007) 475–482.
 14. M. N. H. Mollah, M. Minami and S. Eguchi. Robust prewhitening for ICA by minimizing beta-divergence and its application to FastICA. *Neural Processing Letters* 25 (2007) 91–110.
 15. T. Takenouchi, M. Ushijima and S. Eguchi. GroupAdaBoost: accurate prediction and selection of important genes. *IPSI Transactions on Bioinformatics* (2007) 3, 1–8.
 16. Kumon, M. and A. Takemura, On a simple strategy weakly forcing the strong law of large numbers in the bounded forecasting game, *Annals of the Institute of Statistical*

Mathematics (to appear), DOI 10.1007/s10463-007-0125-5.

17. Kumon, M., Takemura, A. and Takeuchi, K., Game-theoretic versions of strong law of large numbers for unbounded variables, *Stochastics*, Vol.79, No.5, 449-468, 2007.
18. Kumon, M., Takemura, A. and Takeuchi, K., Capital process and optimality properties of a Bayesian Skeptic in coin-tossing games, *Stochastic Analysis and Applications* (accepted and to appear).
19. Ando, M. and Hodoshima, J., A note on bootstrapped White's test for heteroskedasticity in regression models, *Economics Letters*, 97, 46-51, 2007.
20. Hodoshima, J. and Ando, M., The finite-sample performance of White's test for heteroskedasticity under stochastic regressors, *Communications in Statistics: Simulation and Computation*, 36, 1201-1215, 2007.
21. 藤田 利治, 真山 武志, 降圧薬の使用成績調査データベース構築とその活用例, *日本統計学会誌*, 第36巻J-1号, 205-217, 2007。
22. Hayashi, K., Mizunuma, H., Fujita, T., Suzuki, S. Imazeki, S., Katanoda, K., and Hayashi, K., Design of the Japan Nurses' Health Study: A Prospective Occupational Cohort Study of Women's Health in Japan, *Industrial Health*, Vol. 45, 679-686, 2007.
23. 椿 広計, 統計科学の横断性と設計科学への寄与, *横幹*, Vol. 1, No. 1, 22-28, 2007年4月。
24. Seike, N., Kashiwagi, N. and Otani, T., PCDD/F contamination over time in Japanese paddy soils, *Environmental Science and Technology*, 41, 2210-2215, 2007.
25. Kajiwara, Y., Kashiwagi, N. and Kadokami, K., Nationwide study of dioxins in the freshwater fish *Carassius auratus (gibelio) langsdorfii* (crucian carp) in Japan: concentrations and estimation of source contribution ratios, *Chemosphere*, 69, 1177-1187, 2007.
26. Koga, T., Kuwano, K., Kito, G. and Kanefuji, K., Evaluation of QT interval using a linear model in individual cynomolgus monkeys, *Journal of Pharmacological and*

- Toxicological Methods, 55, 265-270, 2007.
27. 河村 敏彦, 岩瀬 晃盛, パラメータ設計における 2 ステップ最適性に関する統計的一考察, 統計数理, 55(2), 327-336, 2007。
 28. 敦賀 智裕, 山下 智志, デフォルト境界が不確実な場合の損失率: 優先劣後構造を持つ債権への応用, 金融研究, 26 巻別冊第 2 号, 79-103, 2007。
 29. 山下 智志, 吉羽 要直, 追加融資を考慮した信用リスク: 構造モデルによる EL と UL の解析解, 金融研究, 26 巻別冊第 2 号, 103-136, 2007。
 30. Kunitomo, N. and Kim, Y.-J., Effects of Stochastic Interest rates and Volatility on Contingent Claims, Japanese Economic Review, 88(1), 2007.

[会議録]

1. Kentarou Kido, Hiroshi Kuwajima and Takashi Washio: A Range Query Approach for High Dimensional Euclidean Space Based on EDM Estimation, Proc. of SDM2008: the 18th SIAM International Conference on Data Mining, (To appear)
2. Alexandre Termier, Yoshinori Tamada, Kazuyuki Numata, Seiya Imoto, Takashi Washio, Tomoyuki Higuchi: DIGDAG, a first algorithm to mine closed frequent embedded sub-DAGs, Proc. of MLG Workshop 2007, Mining and Learning with Graphs, pp.41-46 (2007)
3. Hiroshi Kuwajima and Takashi Washio: Large PSD Matrix Estimation from Partial Elements, Workingnotes of Seventh IEEE International Conference on Data Mining - Workshops: 0-7695-3019-2/07 © 2007 IEEE DOI 10.1109/ICDMW.2007.24, pp.337-342 (2007)
4. Hiroshi Kuwajima and Takashi Wasiho: Fast PSD Matrix Estimation by Column Reductions, ISM Report on Research and Education No. 25, The International Workshop on Data-Mining and Statistical Science (DMSS2007), pp.179-189 (2007)
5. Tsubaki, H., The Grammar of Technology Development-Needs for a New Paradigm of Statistical Thinking, The Grammar of Techonolgy Development, 15-22, 2007.
6. Kawasaki, Y., Koga, T. and Kanefuji, K., Common intervention analysis in multivariate

nonstationary time series, MODSIM 2007 International Congress on Modelling and Simulation, 2989-2995, 2007.

7. Kanefuji, K., Kawasaki, Y., Sato, S., Sumiya, T. and Ochi, Y., Statistical courseware with synchronized web-based statistical analysis system, Proceedings of 7th Hawaii International Conference on Statistics, Mathematics and Related Fields 7, 84-87, 2008.

[解説・総説]

1. 鷲尾 隆: 因果関係モデリングにおけるデータマイニング・グラフマイニング技術の活用, 日本化学会情報化学部会誌 Vol. 25 (2007), No. 3 October pp.76-80 (2007)
2. 江口 真透, 統計的パターン認識: 線型判別からアダブーストまで。日本化学会情報化学部会誌 25, 68 (2007)。
3. 江口 真透, リスク科学のための人工知能的方法と統計的思考。人工知能学会誌 22 巻 5 号 特集「定量的リスク科学」656-662 (2007 年 9 月)
4. 藤田 利治, ファーマコビジランスに不可欠な大規模データベース, 医学のあゆみ, 223, 855-860, 2007 年 12 月。

[研究ノート]

1. 前川功一, 河合研一, ボラティリティの長期依存性に関する一考察, 広島経済大学経済研究論集, 第 30 巻第 3・4 合併号, 平成 20 年 3 月。

[その他]

1. Takeuchi, K., Kumon, M. and Takemura, A., A new formulation of asset trading games in continuous time with essential forcing of variation exponent, The Institute of Statistical Mathematics Research memorandum, No.1042, August 7, 2007.

<会議発表等>

[招待講演]

1. Tree of Life Symposium (Beijing, China), 5 Jun 2007
"Rooting the eutherian tree ? Problems of phylogenomics"
M. Hasegawa

2. 江口真透。情報幾何学への招待。サマーセミナー2007, 2007年8月6日指宿市
3. 江口真透。表現形の相関研究：パターン認識の先に何が見えるか。春季統計学会, 2008年3月1日
4. Shinto Eguchi, Boosting methods for association studies in bioinformatics, International Conference on Multiple Decision Theory, Statistical Inference and Applications, タイペイ, 中華民国, 2007年12月28日
5. Shinto Eguchi, Boosting learning approach to association studies in bioinformatics, First joint meeting between Institute of Statistical Science, Academia Sinica, Taiwan and the Institute of Statistical Mathematics, November 29, 2007
6. 公文 雅之, 資産取引ゲームにおける最適戦略: Kelly criterion をめぐって, 応用経済時系列研究会, 統計数理研究所, 平成20年2月5日。

〔一般講演〕

1. 桑島 洋, 中西 耕太郎, 鷺尾 隆: 半正定性を満たす類似性尺度の高速推定手法, 人工知能学会2007年全国大会, 予稿集 CD, 1C1-4 (2007.6.22)
2. 城戸 健太郎, 桑島 洋, 中西 耕太郎, 鷺尾 隆: 半正定性に基づく高速ユークリッド距離推定手法, 人工知能学会2007年全国大会, 予稿集 CD, 1C1-5 (2007.6.22)
3. 大原 剛三, 鷺尾 隆: 大規模グラフにおける近似部分グラフ検索手法, 人工知能学会2007年全国大会, 予稿集 CD, 3F7-1 (2007.6.22)
4. 高林 健登, 原 昌弘, 大原 剛三, 元田 浩, 鷺尾 隆: 情報利得値の上界に着目した特徴的部分グラフの効率的なマイニング, 人工知能学会2007年全国大会, 予稿集 CD, 3F7-2 (2007.6.22)
5. 桑島 洋, 鷺尾 隆: 部分要素を用いた巨大半正定類似性尺度行列の高速推定, 第4回データマイニングと統計数理研究会 (SIG-DMSM), 人工知能学会, 講演報告集 CD-ROM (2007.7.26)
6. 城戸健太郎, 桑島洋, 鷺尾隆: ユークリッド距離の高速高精度推定と範囲問い合わせへの適用, 人工知能学会第79回知識ベースシステム研究会, 予稿集, pp.67-74 (2007.12.5)
7. Evolution 2007 Conference (Christchurch, New Zealand), 17 Jun 2007
MolPhy: new software for phylogenetic analyses using maximum likelihood.
J. Adachi

8. Evolution 2007 Conference (Christchurch, New Zealand), 18 Jun 2007
Molecular phylogeny and evolution of river dolphins: integration of mitochondrial DNA and SINE flanking sequences.
9. Y. Cao, H. Shimodaira, D. Wang, M. Hasegawa
10. Evolution 2007 Conference (Christchurch, New Zealand), 19 Jun 2007
Molecular phylogeny and evolution of prosimians based on complete sequences of mitochondrial DNAs.
11. A. Matsui, F. Rakotondraparany, I. Munechika, M. Hasegawa and S. Horai
第 23 回日本霊長類学会大会 2007 年 7 月 15 日
12. P-41 霊長類の糞を試料としたミトコンドリア DNA 全塩基配列決定。
松井淳, Felix Rakotondraparany, 長谷川政美, 宝来聡
13. 第 58 回 染色体学会・第 17 回 染色体コロキウム 2007 年合同年会 分科会 (Genomics & Epigenetics & Chromosome) 2007 年 11 月 28 日。G-02 ミトコンドリアゲノムによる原猿類の分子系統進化。松井淳, Felix Rakotondraparany, 宗近功, 長谷川政美, 宝来聡
14. 江口真透, 竹之内高志, マルチクラスの判別問題におけるミスラベルのモデリング, 統計関連学会, 2007 年 9 月 7 日,
15. 江口真透, 観察データの推測の限界—揺らぎモデルアプローチ—広島大学理学部, 科研費研究集会「統計的データ解析手法の評価と開発」2008 年 1 月 16 日
16. 公文 雅之, 竹内 啓, 竹村 彰通, 資産取引ゲームにおける最適戦略, 統計関連学会 連合大会, 神戸大学, 平成 19 年 9 月 7 日。
17. 竹村 彰通, 公文 雅之, 竹内 啓, 資産取引ゲームにおける最適戦略, 日本数学会 2007 年度年会, 東北大学, 平成 19 年 9 月 24 日。
18. 公文 雅之, Mutual information between betting games, Tutorial workshop on game-theoretic probability, 東京大学, 平成 20 年 2 月 28 日。
19. 河合 研一, 前川 功一, 実現ボラティリティの長期依存性に関する一考察, 科学研究費基盤研究(B)「高頻度データによる株価・為替レートの計量分析」(代表: 前川 功一) 研究集会, 広島経済大学, 平成 19 年 12 月 26 日。
20. 友定 充洋, 衛星画像の自動ミクセル分解手法の提案, 第 42 回日本リモートセンシング学会 学術講演会, 2007 年 5 月 10 日。
21. Tomosada, M., A proposition of automatic mixed pixel classification for remotely sensed multispectral image, International conference on Instrumentation, control and Information Technology, Kagawa University, Takamatsu, Japan, 2007 年 9 月 19 日。
22. 友定 充洋, 椿 広計, 次期地球環境観測衛星による物理量推定精度評価方法の提案, 第 2 回横幹連合コンファレンス, 京都大学, 2007 年 11 月 29 日。
23. Tomosada, M. and Tsubaki, H., An evaluation method of the retrieved physical quantity deriving from the satellite remote sensing using analysis of variance in experimental

- design, IFIP/IIASA/GAMM Workshop on Coping with Uncertainty (CwU) Robust Decisions, IIASA, Laxenburg, Austria, 2007年12月10日。
24. 佐藤 整尚, 陳 春航, 上下で異なったジャンプ構造を持つGARCHモデルについて, 統計関連学会連合大会, 神戸, 2007年9月8日。
 25. 松本 幸雄, 田村 憲治, 原沢 英夫, 佐々木 寛介, 椿 貴博, 温暖化にともなうオキシダント濃度の変化と健康影響の予測—東京, 大阪, 愛知周辺—, 大気環境学会, 岡山市, 2007年9月6日。
 26. 公文 雅之, 資産取引ゲームにおける最適戦略, 統計数理研究所 オープンハウス, 統計数理研究所, 平成19年7月19日。
 27. 河合 研一, 高頻度データに基づく GARCH-Jump モデルの推定, 統計数理研究所 オープンハウス, 統計数理研究所, 平成19年7月19日。
 28. 安藤 雅和, CDO(債務担保証券)の価格評価について, 融合研究若手クロストーク, 伊豆長岡温泉 KKR 千歳荘, 2007年11月26日。

<著書等>

Takashi Washio: Guest Editor of Special Issue, Applications eligible for data mining, Advanced Engineering Informatics, Vol.21, 2007, pp.241-242, pp.241-301 (2007)

Washio, T.; Zhou, Z.-H.; Huang, J.Z.; Hu, X. (.; Li, J.; Xie, C.; He, J.; Zou, D.; Li, K.-C.; Freire, M.M. (Eds.): Emerging Technologies in Knowledge Discovery and Data Mining, XIV, 675 p., Softcover, Lecture Notes in Artificial Intelligence (LNAI), Lecture Notes in Computer Science (LNCS), Vol. 4819, ISBN: 978-3-540-77016-9, Springer (2007)

椿 広計, 「リスク解析とは何か」, リスク学入門 1(橋木俊詔, 長谷部恭男, 今田高俊, 益永茂樹 編)第4章, 岩波書店, 東京, 2007年7月。[分担執筆]

川崎 能典, 「移動平均と時系列成分モデル」, 統計データ科学事典(杉山高一, 藤越康祝, 杉浦成昭, 国友直人 編著)所収, 朝倉書店, 東京, 566-567, 2007年6月。[分担執筆]

浅野 直人, 藤倉 良, 増沢 陽子, 石野 耕也, 小幡 雅男, 大塚 直, 柳 憲一郎, 高橋 滋, 小野 芳朗, 小瀬 知洋, 新美 育文, 牛嶋 仁, 山田 洋, 松村 弓彦, 岩間 徹, 「環境リスク管理と法」, 慈学社出版, 193-205, 2007年11月。[分担執筆]

統計データ科学事典, 小項目「情報幾何」分担執筆, 杉山高一・藤越康祝・杉浦成昭・

国友直人 編。(2007 年) 朝倉書店

<受賞>

1. 人工知能学会研究会優秀賞 (2007. 6. 21)
2. マーケティング分析コンテスト2007 佳作, マーケティング分析コンテスト事務局主催, 野村総合研究所, 宣伝会議協賛 (2007. 12. 19)
3. 日刊工業新聞社主催「第2回モノづくり大賞」特別賞, 受賞団体: 業種の異なる企業が集合した産学官連携「物づくりの機能性評価研究会」(河村敏彦助教がメンバーとして参加)
<http://www.nikkan.co.jp/sangakukan/>

(3) その他の成果発表

なし。