

## AIMaP 研究集会等実施報告書

(Part 1/4) 名称・重点テーマ・キーワード等

項目	内容
名称	大規模複雑データの理論と方法論、及び、関連分野への応用
採択番号	2017A011
重点テーマ	大規模複雑データの数理と関連分野への応用
キーワード	高次元データ、ビッグデータ、機械学習、人工知能、ディープラーニング、 計算統計、生物統計、データマイニング
主催機関	筑波大学
運営責任者	青嶋 誠
開催日時(開始)	2017/12/1
開催日時(終了)	2017/12/3
開催場所	筑波大学自然系学系 D 棟 509 (筑波キャンパス内)

(Part 2/4) 最終プログラム・参加者数

項目	内容
最終 プログラム	<p>詳細プログラム：  <a href="http://www.math.tsukuba.ac.jp/~aoshima-lab/jp/symposium.html">http://www.math.tsukuba.ac.jp/~aoshima-lab/jp/symposium.html</a> (公式サイト)</p> <p>12月1日(金)</p> <p>14:00～14:05 開会</p> <p>14:05～14:45 石井 晶 (東京理科大・理工学部)  矢田 和善 (筑波大・数理物質系)  青嶋 誠 (筑波大・数理物質系)</p> <p>Equality tests of covariance matrices based on eigenstructures  in the high-dimensional context</p>

14:55~15:25 藤森 洸 (早稲田大・基幹理工学研究科)

西山 陽一 (早稲田大・国際教養学部)

The Dantzig selector for Cox's proportional hazards model

15:30~16:00 牧草 夏実 (島根大・大学院総合理工学研究科)

内藤 貫太 (島根大・大学院総合理工学研究科)

再生核ヒルベルト空間における正規性の検定

16:05~16:45 片山 翔太 (東京工業大・工学院経営工学系)

Support recovery of adaptive generalized lasso under high-dimensionality

16:55~17:35 筑瀬 靖子 (香川大)

High dimensional limit theorems on the stiefel manifold

## 12月2日(土)

9:40~10:20 今倉 暁 (筑波大・システム情報系, 人工知能科学センター)

複素モーメント型並列固有値解法とその応用

10:30~11:10 園田 翔 (早稲田大・先進理工学部)

深層学習のWasserstein 幾何学的解析にむけた取り組み

11:20~12:00 内藤 貫太 (島根大・大学院総合理工学研究科)

Spiridon Penev (University of New South Wales)

Locally robust methods and near-parametric asymptotics

13:15~13:55 柳本 武美 (統計数理研究所)

多項出現確率の推定法としての深層学習分類器

14:05~14:45 塩川 浩昭 (筑波大・計算科学研究センター)

大規模グラフクラスタリングの高速化

14:55～15:35 仲木 竜 (株式会社Rhelixa・代表取締役CEO)

全ゲノムシーケンシング時代を勝ち抜く計算科学技術

15:45～16:25 大川 英希 (筑波大・数理物質系, 宇宙史研究センター)

素粒子実験における多変量解析・機械学習・深層学習などのビッグデータ解析

－ LHC-ATLAS 実験を例に

16:35～17:15 小副川 健 (株式会社富士通研究所)

ディープラーニングのビジネス適用に向けた取り組み

## 12月3日(日)

9:20～9:50 川村 健太 (島根大・大学院総合理工学研究科)

内藤 貫太 (島根大・大学院総合理工学研究科)

Divergence に基づく局所密度推定

9:55～10:35 永井 勇 (中京大・国際教養学部)

分散共分散行列の逆行列における縮小推定法の提案

10:45～11:25 種市 信裕 (北海道教育大・札幌校)

関谷 祐里 (北海道教育大・釧路校)

外山 淳 (数学利用研究所)

多次元分割表の完全独立性検定における変換検定統計量の構築について

11:35～12:15 新村 秀一 (成蹊大学)

なぜ癌の遺伝子解析は30年以上成功しなかったのか？

12:15～12:20 閉会

参加者数

数学・数理科学:43人、諸科学:4人、産業界:3人、その他:0人

(Part 3/4) 論点・現状・今後の展開

項目	内容
当日の論点	<p>運営責任者の青嶋は、統計科学・機械学習・生物工学等の研究者が一堂に会する分野横断型の科研費シンポジウムを、2009年から毎年、筑波大学で開催してきた。今年も、大規模複雑データの理論と方法論について、関連分野への応用をテーマに産業界の技術者にまで幅広く講演を募り、統計科学を含む数学・数理科学から11件、諸科学から4件、産業界から2件の講演を選定した。数学・数理科学と諸科学・産業界との連携強化を意識し、諸科学・産業界からの問題提起と統計科学の最先端技法について、互いの情報を共有して、共同研究に発展する深い議論がなされた。特に、高次元データ解析とディープラーニングについて、素粒子実験やゲノム解析産業の現場からデータ解析における問題提起がなされ、統計科学から最先端の理論・方法論を提供すべく、非常に密な議論がなされた。本シンポジウムは、数学・数理科学と諸科学・産業界の双方にとって、大変に実りの多いものとなった。</p>
研究の現状と課題（既にできていること、できていないことの切り分け）	<p>大規模複雑データ解析に対して、統計的理論・方法論は目覚ましく発展している。しかし、諸科学・産業界に十分に認知されているとは言い難い。実際、ゲノム解析や人工知能など諸科学・産業界では非常に魅力的な事例が報告されているものの、数学的理論に裏打ちされた方法論を使用しているわけではなく、多くの場合、計算機の性能を鵜呑みにし、大規模複雑データに内在する巨大なノイズを甘く見積もったものになっている。それゆえ、データ解析に精度が保証されていない。最先端科学のデータ解析に、高精度で柔軟な統計的理論・方法論を提供することは急務である。</p>
新たに明らかになった課題、今後解決すべきこと	<p>素粒子実験、ゲノム解析、人工知能等の最先端科学におけるデータ解析の現場で、理論的な解釈の困難さなどが原因で、最先端の統計的方法論が十分には利用されていないことが明らかになった。大規模複雑データを扱う諸科学・産業界に、今後もより一層、最先端の統計的方法論を普及していく必要があるとともに、問題の共有と深い議論をするための機会を設けることは、今後も大事になるであろう。</p>
今後の展開・フォローアップ	<p>シンポジウム全体を通して参加者の反応は極めて良好で、数学・数理科学と諸科学・産業界からの分野融合の講演は、それらの研究・技術を公開し、アイデアが混ざり合い、まさしく共同研究へと繋がるきっかけとなった。実際、青嶋グループは、最先端ゲノムデータを大量に有する株式会社Rhelixaの仲木氏との共同研究と、また、素粒子論の大川氏との共同研究について、それぞれ準備を進めている。今後は共同研究の成果を形にして、当該分野での報告を予定している。</p>

(Part 4/4) 写真

項目	内容
添付写真 1	 A group of approximately 25 people, including men and women of various ages, are standing in a line in a classroom. They are positioned in front of two large projection screens. The room contains several rows of yellow desks and blue chairs. The lighting is bright, and the atmosphere appears to be formal or semi-formal.
添付写真 2	 This is a second group photo of the same group of people from the first photo, taken from a slightly different angle in the same classroom setting. The group is standing in front of the projection screens, and the room's layout with desks and chairs is visible.
添付写真 3	 A person is standing at the front of the classroom, presenting to an audience. The audience is seated at desks, seen from behind. Two projection screens are visible at the front. The screen on the right displays a slide with the title "NEW TEST PROCEDURE BY USING AN METHOD (LIVE)" and lists several items under "METHODOLOGY". The screen on the left shows a slide with a green background and some text.

(2017/08/24b)