様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科研費 KAKENHI

科学研究費助成事業　　研究成果報告書

令和　５ 年　 ６ 月 １６ 日現在

機関番号：８２４０１
研究種目：基盤研究(B)（一般）
研究期間：2020～2022
課題番号：２０Ｈ０４２５０

研究課題名（和文）Machine learning and statistical methhods on infinite-dimensional manifolds

研究課題名（英文）Machine learning and statistical methhods on infinite-dimensional manifolds

研究代表者
Ｈａ　ＱｕａｎｇＭｉｎｈ（Ha, Quang Minh）

国立研究開発法人理化学研究所・革新知能統合研究センター・ユニットリーダー

研究者番号：９０８６８９２８

交付決定額（研究期間全体）：（直接経費）　　5,950,000円

研究成果の概要（和文）：私たちは、最適輸送と情報幾何学の枠組みにおける無限次元ガウス測度、ガウス過程、および無限次元正定演算子の幾何学に関して多くの結果を得ました。
(1) エントロピー正則化ワッサーシュタイン距離、正則化カルバック・ライブラーおよびレニー発散、正則化フィッシャー・ラオ距離など、関係する多くの関心量の明示的な数式。 これらは、機械学習や統計のアルゴリズムに簡単に使用できます。
(2) 特に、正規化された距離と発散の次元に依存しないサンプルの複雑さを示し、近似有限次元法の一貫性を証明する理論的分析。 正規化された距離と発散には多くの有利な理論的特性があり、実際のアルゴリズムに影響を与えます。

研究成果の学術的意義や社会的意義
Our results are the first in the setting of infinite-dimensional Gaussian measures and Gaussian processes. They (1) elucidate many theoretical properties of Optimal Transport; (2) have important consequences for the mathematical foundations of Gaussian process methods in machine learning.

研究成果の概要（英文）：We have obtained many results on the geometry of infinite-dimensional Gaussian measures, Gaussian processes, and infinite-dimensional positive definite operators in the framework of Optimal Transport and Information Geometry. These include
(1) Explicit mathematical formulas for many quantities of interest involved, including entropic regularized Wasserstein distance, regularized Kullback-Leibler and Renyi divergences, and regularized Fisher-Rao distance. These can readily be employed in algorithms in machine learning and statistics.
(2) Extensive theoretical analysis showing in particular dimension-independent sample complexities of the above regularized distances and divergences. These provide guarantees for the consistency of finite-dimensional methods to approximate them in practice. Moreover, we show explicitly that the regularized distances and divergences possess many favorable theoretical properties over exact ones, with implications for practical algorithms.

研究分野：Mathematical foundations of machine learning

キーワード：Gaussian measures　Gaussian processes　Optimal Transport　Information Geometry　Riemannian geometry　Divergences　Wasserstein distance　Entropic regularization

１．研究開始当初の背景

Geometrical methods have become increasingly important in Machine Learning and Statistics and other application domains, including Computer Vision, Signal Processing, and Brain Computer Interfaces. Many geometrical methods are based on the mathematical theories and techniques of Riemannian Geometry, Information Geometry, and Optimal Transport. Most of the current literature focuses on finite-dimensional methods. Among these are methods for the finite-dimensional manifolds of symmetric, positive definite (SPD matrices and Gaussian densities on Euclidean space, which play a central role in statistics, machine learning, and numerous practical applications.

２．研究の目的

This project aims to generalize the finite-dimensional geometrical methods for SPD matrices and Gaussian densities on Euclidean space to the setting of infinite-dimensional positive definite operators, infinite-dimensional Gaussian measures, and Gaussian processes.

３．研究の方法

There are several key techniques that we employ to obtain the mathematical formulations for the infinite-dimensional setting

(1) Regularization: this includes entropic regularization for the setting of Optimal Transport and covariance operator regularization for the setting of Information Geometry. In the Information Geometry setting, we show that regularization is key to proper, well-defined distances/divergences and the corresponding geometrical structures. In both settings, regularization leads to many favorable theoretical properties, including dimension-independent sample complexities for finite-dimensional approximation methods.

(2) For the application of the abstract framework of Gaussian measures on infinite-dimensional Hilbert spaces to the Gaussian process setting, the methodology of reproducing kernel Hilbert spaces (RKHS) plays a crucial role. In particular, we introduced the concept of cross-covariance RKHS operators, which is new in the literature (to the best of our knowledge).

４．研究成果

We have obtained many results on the geometry of infinite-dimensional Gaussian measures, Gaussian processes, and infinite-dimensional positive definite operators in the framework of Optimal Transport and Information Geometry. These include

(1) Explicit mathematical formulas for many quantities of interest involved, including entropic regularized Wasserstein distance, regularized Kullback-Leibler and Renyi divergences, and regularized Fisher-Rao distance. These can readily be employed in algorithms in machine learning and statistics.

(2) Extensive theoretical analysis showing in particular dimension-independent sample complexities of the above regularized distances and divergences. These provide guarantees for the consistency of finite-dimensional methods to approximate them in practice. Moreover, we show explicitly that the regularized distances and divergences possess many favorable theoretical properties over exact ones, with implications for practical algorithms.

The following papers have been accepted/published

(1) Ha Quang Minh. Finite sample approximations of exact and entropic Wasserstein distances between covariance operators and Gaussian processes, **SIAM/ASA Journal on Uncertainty Quantification**, volume 10, number 1, pages 96-124, 2022, https://epubs.siam.org/doi/abs/10.1137/21M1410488

(2) Ha Quang Minh. Convergence and finite sample approximations of entropic regularized Wasserstein distances in Gaussian and RKHS settings**, Analysis and Applications**, Vol. 21, No. 03, pp. 719-775 (2023)
https://www.worldscientific.com/doi/abs/10.1142/S0219530522500142

(3) Ha Quang Minh. Entropic regularization of Wasserstein distance between infinite-dimensional Gaussian measures and Gaussian processes, **Journal of Theoretical Probability**, 2022, https://link.springer.com/article/10.1007/s10959-022-01165-1

(4) Ha Quang Minh. Fisher-Rao Riemannian geometry of equivalent Gaussian measures on Hilbert space. **Geometric Science of Information (GSI) 2023**, to appear.

５．主な発表論文等

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

１．著者名 Ha Quang Minh	４．巻 10
２．論文標題 Finite Sample Approximations of Exact and Entropic Wasserstein Distances Between Covariance Operators and Gaussian Processes	５．発行年 2022年
３．雑誌名 SIAM/ASA Journal on Uncertainty Quantification	６．最初と最後の頁 96-124
掲載論文のDOI（デジタルオブジェクト識別子） 10.1137/21M1410488	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 該当する

１．著者名 Ha Quang Minh	４．巻 -
２．論文標題 Entropic Regularization of Wasserstein Distance Between Infinite-Dimensional Gaussian Measures and Gaussian Processes	５．発行年 2022年
３．雑誌名 Journal of Theoretical Probability	６．最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s10959-022-01165-1	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 該当する

１．著者名 Ha Quang Minh	４．巻 21
２．論文標題 Convergence and finite sample approximations of entropic regularized Wasserstein distances in Gaussian and RKHS settings	５．発行年 2023年
３．雑誌名 Analysis and Applications	６．最初と最後の頁 719-775
掲載論文のDOI（デジタルオブジェクト識別子） 10.1142/S0219530522500142	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 該当する

〔学会発表〕　計4件（うち招待講演　3件／うち国際学会　3件）

１．発表者名 Ha Quang Minh
２．発表標題 Riemannian distances between infinite-dimensional covariance operators and Gaussian processes
３．学会等名 4th International Conference on Econometrics and Statistics（招待講演）（国際学会）
４．発表年 2021年

1．発表者名 Ha Quang Minh
2．発表標題 Regularized information geometric and optimal transport distances between covariance operators and Gaussian processes
3．学会等名 Conference on Mathematics of Machine Learning（招待講演）（国際学会）
4．発表年 2021年

1．発表者名 Ha Quang Minh
2．発表標題 Fisher-Rao Riemannian geometry of equivalent Gaussian measures on Hilbert space
3．学会等名 Geometric Science of Information（国際学会）
4．発表年 2023年

1．発表者名 Ha Quang Minh
2．発表標題 Renyi divergences in RKHS and Gaussian process settings
3．学会等名 Information Geometry for Data Science（招待講演）
4．発表年 2022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