

論文内容の要旨

博士論文題目 Statistical Analysis of Optimization Algorithms
(最適化アルゴリズムの統計的解析)

氏名 船谷 浩之

(論文内容の要旨)

Mathematical algorithms have been performing important roles in the modern society; optimization to find the best solution to mathematical problems, classification to find groups in observations, and statistical analysis to find underlying regularities in various phenomena in the nature. In spite of the great results that these algorithms have been yielding, their mathematical mechanisms have yet to be so well studied.

In this dissertation, three subjects are analyzed: a genetic algorithm, an SVM with forgetting factor, and neural spike sequences. For genetic algorithms, we propose a new analytical method from a network point of view. First the genetic algorithm is formulated as a network with nodes representing generations connected by two genetic operations: crossover and mutation. Then, its characteristic path length is derived mathematically.

For support vector machines, a forgetting factor is proposed and the asymptotic generalization ability is derived for a simple linearly separable problem. The learning curve is mathematically derived for two types of forgetting factors: exponential and factorial forgetting factors. It is proved that the learning curve of factorial forgetting factors converges to zero in the asymptotic state.

At last, for the analysis on neuron spike sequences, information geometrical method is employed to classify neurons in cortical areas. The interspike intervals of a spike sequence of a neuron is modeled as a gamma process with a time-variant spike rate, and a semi-parametric estimation problem is formulated for the parameters and is solved using an information geometrical method to derive the optimal estimators from a statistical viewpoint.

氏名	船谷 浩之
----	-------

(論文審査結果の要旨)

情報爆発と言われる現在、情報を効率的に処理する数理アルゴリズムは技術的にも経済的にも重要な役割を果たしている。これらのアルゴリズムには、最適解を見つける数理計画法やデータを統計的性質に基づいて処理する機械学習などが含まれる。その重要性にもかかわらず、数理アルゴリズムにはそのメカニズムが明らかになっていないものが少なくない。本研究はそれらのうち三つのアルゴリズムに焦点を絞り、その性質を理論的に明らかにしたものである。

一つめは準最適化アルゴリズムの一種である遺伝的アルゴリズム (GA) を対象としている。GA は短時間で比較的性質のよい解を見つけることが知られているが、その理由は不明である。本研究では GA の状態遷移をネットワークとみなし、複雑ネットワークとしての性質を調べるという独特のアプローチにより、GA に含まれる二つの演算、交差と突然変異の影響を理論的に明らかにした。

二つめは非線形二値判別器であるサポートベクトルマシン (SVM) を対象としている。従来の SVM はすべての例題を同等に扱うが、例題が時間経過とともに与えられる場合も多く、その場合には適応フィルタにおける RLS アルゴリズムのような忘却係数を導入することが可能である。本研究は統計的漸近論を用い、忘却係数を持つ SVM の学習曲線を導出した。

三つめは神経スパイク列の情報幾何学に関するものである。神経スパイク列の特徴量の推定問題は、数学的にはセミパラメトリック推定として定式化され、その求解には推定関数法が用いられる。本研究は情報幾何学を用い、不応期を持つ神経細胞の特徴量を推定するための最適な推定関数を導出した。

以上をまとめると、本論文は、最適化あるいは機械学習における重要なアルゴリズムについて、高度な数学的手法を用いてその性質を数理的に明らかにしたものであり、数理アルゴリズムの理論的裏付けに応用に大きく寄与するものである。したがって、博士 (工学) の学位に値するものと認められる。