

Title	Studies on Functional Data Analysis Methods and Their Applications
Author(s)	Dou, Xiaoling
Citation	大阪大学, 2008, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/48834
rights	
Note	著者からインターネット公開の許諾が得られていないため、論文の要旨のみを公開しています。全文のご利用をご希望の場合は、 〈a href="https://www.library.osaka-u.ac.jp/thesis/#closed"〉 大阪大学の博士論文について 〈/a〉 をご参照ください。

Osaka University Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

Osaka University

氏 名 ^{ドゥ}竇 ^{シヤオ} 曉 ^{リン} 玲

博士の専攻分野の名称 博 士 (工 学)

学 位 記 番 号 第 2 2 1 1 6 号

学 位 授 与 年 月 日 平成 20 年 3 月 25 日

学 位 授 与 の 要 件 学位規則第 4 条第 1 項該当

基礎工学研究科システム創成専攻

学 位 論 文 名 Studies on Functional Data Analysis Methods and Their Applications
(関数データ解析の手法とその応用に関する研究)

論 文 審 査 委 員 (主査)

教 授 白旗 慎吾

(副査)

教 授 狩野 裕 教 授 長井 英生

論 文 内 容 の 要 旨

毎日の気温データや人間の身長などのような関数データは本来は関数である対象を不連続時点で観測して得られる。関数データ解析の目的はこれらのデータを滑らかな関数として表現し、回帰分析、主成分分析、判別分析を行うことである。

まず関数データの概略を述べた後に論文の前半では回帰関数及びその導関数の推定について考察した。どのようにして関数データの導関数を推定するのかは人間の成長関数速度、加速度などの研究分野ではよくある質問である。罰則付 **B-spline** 基底関数法や **kernel** 関数法など様々な手法があり、その中から適切な手法を選ぶのは必要である。私たちはこれらの関数データに対して回帰関数とその導関数の推定手法を比較した。シミュレーションの結果として、**B-spline** に基づく手法は **kernel** 関数法より、特に関数の定義域の中央部分では良い推定値を与えることが分かった。さらに、罰則付 **spline** 関数には関数の滑らかさを制御する平滑化パラメータを用いるが、各次数の導関数に対して平滑化パラメータを個別に選ぶ必要があることが分かった。

論文の後半では関数データの判別分析について考察した。関数データを基底関数で展開し、データを平滑化することによって、次元の縮減ができ、安定した結果が得られる。さらに、これにより、不規則時点で観測されたデータや欠測値を含んでいるデータを処理できるようになる。この考え方により、私たちは関数データに対して二つのノンパラメトリックな判別分析の手法関数部分空間法を提案した。それらを既存の手法と実際のデータで比べてみた結果、関数部分空間法は良い結果を与えている。特に、教師データのサイズが小さい場合には、伝統的な手法より安定した結果を与えることができる。基底関数展開によって、私たちは一種の関数データの線形判別分析法も開発して、フィルタリング法と比べた。**B-spline** 基底関数やほかの正規直交基底を使うときに、これら二つの手法は同じ結果を与える。

論文審査の結果の要旨

関数データとは、本来は関数である、もしくは関数（回帰関数）と考えた方が解釈しやすい場合で、観測の都合により離散点で観測されたデータを言う。日時観測データ（各地の気温、多くの企業の株価等）や人間の成長データなどの例がある。このようなデータは一つの対象に対しては不連続時点で観測される。関数データ解析の目的はこれらのデータを滑らかな関数として表現し、回帰分析、主成分分析、判別分析を行うことである。

本論文では、まず1章・2章で関数データ解析の特徴や用いる概念・基礎的な解析法について述べた後、3章ではもとの関数、その微分・2回微分の推定法の選択について実証的に考察している。回帰関数の推定に関しては膨大な研究がなされているが、その微分・2回微分の推定についての比較研究はほとんど行われていない。関数の推定については、代表的なB-スプライン関数による推定法を2種類（一つは基底関数のパラメータを別々に取る方法、もう一つは共通に取る方法）および核関数による推定法を2種類（一つは局所多項式回帰の一般化、もう一つは通常の核関数平滑化法）考え、乱雑度の様相の異なる4種類の関数に対しコンピュータ・シミュレーションにより精密に比較した。その結果、B-スプライン関数で、関数ごとに平滑化パラメータを選択する方法が最良である、との結論を得た。ただし、どの方法でも関数の定義域の端の方での当てはめが良くないため、その調整を行うようスプライン関数の節点を調整する方法を提案し、良好な成績を収めた。4章では複数の関数データグループに新しいデータを分類する判別方式を開発している。従来は線形手法が用いられていたが、データ数が少ないとき等に性能が低くデータ数が少ないときには適用で困難であった。そこでB-スプライン関数の空間にデータを射影し、そこから得られる距離を元に判別する方法を提案し、実際のデータとシミュレーションにより比較を行った。その結果、新しい方法は安定した結果を導くことが判明した。5章では対象間の距離を定義し、関数データに対するクラスター分析について考察している。

以上のように本論文は関数データ解析における新しい方法の提案と実証研究に寄与するものであり、博士（工学）の学位論文として価値あるものと認める。