

まえがき

本書のテーマである「時系列解析」は過去の自身のデータから未来のデータを予測するために用いられる手法であるが、予測だけでなく、事象の分解・理解に強みを持つ手法でもある。現実世界で未来を予測することは非常に難しい問題であることは直観的に理解ができると思う。時系列解析手法を単に適用しただけでは予測がうまくいくことはほぼない。事象の観察 → モデリング → 不足している情報の考察（事象を再度観察） → 必要な情報をモデルに組み込む（再モデリング）という作業を繰り返すことでモデルを構築していく必要がある。本書では、簡単なデータを用いた簡単な課題を例にとり、基礎的なモデル構築の過程を段階的に体験できるように心がけた。また数式の提示は最小限にし、自学により応用範囲を広げてもらえるように、どの場面で、なぜその手法を使うのかを考えられるように説明することを心がけた。この点において、理論的な解説を最小限に留めたため、詳細を他書に譲ることも多くなっている点はご了承いただきたい。

データという観点から考えると、時系列データは多くのビジネス現場で発生するデータである。近年、IoTの普及により機械から吐き出されるセンサーデータの量が増えており、今後益々、時系列解析・信号解析の需要は増えていくと考えられる。しかし動的なデータとして時系列データを十分に活用できている現場は少ないのではないかと考えている。時系列データに対して時間を説明変数としたあまり適切ではない単回帰を適用しているような場面もしばしば見かける。時系列データは特殊なデータであり、時系列データに適した手法が存在することを本書を通してご理解いただきたい。

時系列解析手法はマーケティングや経済の領域だけでなくロケットの軌道推定、ロボットの制御、脳信号解析などでも古くから利用され、幅広い活用場面がある。

十分に時系列データを活用するには自らが立てた仮説に従ってデータを取得・加工する必要がある。仮説の実証には既存のツールをうまく活用しながらコードを書いて自らの考えを実現することが近道である。幸い、統計解析／機械学習／科学計算のための広範なツールを提供しているPython、R言語、Juliaといったプログラミング言語が現在存在する。データ解析プロセスを素早く回す上でも車輪の再開発をせずに既存ツールを上手に活用することは重要である。本書では、Pythonを使った時系列解析の基礎的な内容を提供する。Pythonで時系列データの解析を取り扱っている書籍は少ないため、本書がPythonユーザにおける時系列解析の裾野を広げる一助となれば幸いである。

本書の構成

本書は以下のように4つの章で構成されている。

- 第1章：時系列データの基本的な知識およびデータ加工方法，第2章以降で必要になってくる知識，四則演算でできるレベルの時系列データ解析方法を説明
- 第2章：経済・マーケティングの分野で多く用いられるARモデルに代表される自己回帰型の古典的なデータ解析手法について説明
- 第3章：工学分野の信号処理でも活躍の場面が多いカルマンフィルタに代表される状態空間モデルについて説明
- 第4章：IoT分野で活躍の場面が多い異常検知について説明

上の4つの章とは別に付録としてNumPy, Pandas, TensorFlowについての初歩的な操作方法について記載した。NumPyなどになじみがない方は参考にしてほしい。

サポートサイト

本書の事例で用いたコードや誤植などはサポートサイト

<https://www.kyoritsu-pub.co.jp/bookdetail/9784320044722>

から確認してほしい。

謝辞

福島真太郎氏には構成段階から多大な示唆をいただき，本書の内容をより正確なものにできた。本書の執筆・出版にあたりご協力いただいた多くの方々に感謝の意を表します。

2019年7月
島田直希