

1-2. 自習の進め方（どこに何が書いてあるか。読者別のガイド）

本シリーズで扱うのは、あるデータ群とあるデータ群を同一のデータ群ではないと判断する方法（有意差検定）と多くの属性からなるデータ群を集約する技術（多変量解析）と AI の活用です。

高校レベルの数学をやったことがあるが、あまり得意ではない。大学の教養課程での講義は受けたことがないという人を読者に想定しています。具体的に言えば、微分・積分を習ったことはあるが、細かい解法は忘れてしまったという人です。いくつか説明の仕方がある場合には、可能な場合には、中学校レベルの数学でも理解できる説明を選びました。四則演算ができれば理解できるというのを理想にしました。とはいっても、大学の教養課程レベルで習う数学の内容が含まれますので、それについては、わかりやすい解説を（反対に言えば数学的に厳密ではない。）を付けました。項を分けてありますから、それらの説明に興味がない人、必要ではない人はそれらの項を読む必要はありません。この際、大学の教養課程レベルの数学をきちんと身に着けたい人はそれも読んでください。高等数学の入り口ぐらいは理解できます。図を使って作業の数式が意味する内容を具体的に説明しているので、大学の教養課程レベルの数学の一部が感覚的に身に付きます。反対に、そんなことはもう知っていて、ちょっとやり方を忘れてただけだという人には、計算手順の具体的な説明があります。そこだけ読んでください。想定しているのは、初めて学ぶ人なので、同じことを別の表現で何回も繰り返している部分があります、くどいという印象を受けるかもしれません。そこは我慢してください。

分散分析の最後に、単回帰についてのみ、章を設けて説明しました。単回帰の延長で重回帰を説明した方がいのかかもしれませんが、重回帰は多変量解析に入れました。単回帰分析を拡張する形で重回帰分析を説明すれば、線形代数学の知識は必要ありませんが、主成分分析や因子分析の理解には、初歩の線形代数学が必要になります。線形代数学を知っていることを前提に、主成分分析や因子分析の説明をすれば、ごちゃごちゃした感じのないすっきりした説明になります。それでもよいのですが、ベクトルや行列の基本的性質を知らなければ、様々な多変量解析を総合的に全体として理解することはできません。単回帰のときに、一対一に対応した説明変数と非説明変数のセットの関係を、一つの係数で表したように、 p 対 1 に対応した説明変数と非説明変数のセットの関係を、 p 個の係数で表そうとする。この時に、被説明変数を n 行の縦ベクトル、係数を p 行の縦ベクトルとして、行列やベクトルを一つの変数一つの係数のように塊として捉えて、 $n \times p$ 行列の割り算の代わりに変数行列の逆行列を被説明変数の行列にかけるというやりかた。あたかも四則演算の延長上に線形代数学を捉える感覚、その感覚に興味を持って貰いたいと考えたので、単回帰の延長上に重回帰を捉えるのではなくて、線形代数学を理解する糸口として、多変量解析の基礎的な作業として重回帰を例にするという説明にしました。そして、その後の説明でも、ところどころに初歩的な線形代数についての解説を入れました。また、偏相関解

析は重回帰と関連して解説するのが普通だと思います。その方が多重共線性の説明が楽です。しかしここでは、線形代数の空間的な捉え方の一例として偏相関を考えて、重回帰と切り離して、線形代数学の説明の直後に配置しました。最後に機械学習について、判別分析を教師あり学習、クラスター分析を教師なし学習の例として解説しました。判別分析は多変量分析のところにも解説があります。ここでは、プログラミングも必要になると考えて、Pythonによるプログラムも実装しました。これは今の流行に合わせたものです。筆者は従来Rを使っていたので、必要に応じてRのプログラムも例示していきたいと思えます。

それぞれの要求に応じた読み方は次の通りです。

分散分析

A. はじめて学ぶので、すみからすみまできちんと知っておきたい人

II章からIV章の全てを読む

ただし、III-2-7まで読んで、III-3を飛ばして、IV章を読んでから、飛ばした部分を丁寧に読んでもよい。

B. それぞれの分布の数式を導きだす数学的プロセスを知る必要はないが、各確率分布の特性と使い方を知っておきたい人

IV章

D. 各分布の導き方を知りたい人

III章

E. とりあえず分散分析の考え方だけを知りたい人

I章、II章

F. 式の形だけ知りたい人

付録-1式

多変量解析

A. 線形代数学の知識の無い人。

V章から読み始め、VI章を読んで、必要な分析方法の章を読む。ただし、主成分分析を知らないと、因子分析の章を読んでも理解できません。因子分析と主成分分析は、目的を異にする分析だということをしっかり理解してください。

B. 線形代数学の知識のある人（スペクトル分解ぐらいまで）

VI章を読んで、必要な分析法を読む。ただし、主成分分析を知らないと、因子分析の章を読んでも理解できません。

C. 線形代数学の知識はあるが、多次元の確率密度や母集団の分散共分散行列の推定など多変量解析の基礎的技術については知らない人

V-2, VI章を読んで、必要部分を読む。ただし、主成分分析を知らないと、因子分

析の章を読んでも理解できません。

- D. この際、教養として知識を身に着けたい人。ブログの主催者が何を考えているのか知りたい人。全部読んでください

機械学習

機械学習について学びたい人、Python によるプログラウを作りたい人。

VII から読んで、必要に応じて前の章を読んでください。

その他、様々な式については、索引のような表を作りました。式の細かいところを忘れてしまった人は、知識の確認に使ってください。