

# 人工知能アルゴリズム探検隊

## 第38回 統計解析ソフトを使ったアンケートの因子分析

牧野 浩二

人の行動データやアンケート結果から共通項を抽出するのが得意な「因子分析」(図1)について解説します。前回(第37回, 2020年5月号)はアルゴリズムを紹介しました。今回は本格的な応用ができるように、「R」という統計解析ソフトウェアを使った実例を示します。

### 統計解析環境 R の準備

#### ● インストール

Rのインストールは下記URLから行います。

<https://www.r-project.org/>

ダウンロード・リンクが分かりづらいため少し解説します。まず、ダウンロードするためのリンクは英語の文章中に太字の青文字で書かれた「download R」をクリックします(図2)。その後、ダウンロード・サイトの一覧が表示されるので、JAPANの下にある2つあるリンクのどちらかをクリックします。これは、お住まいの地域に近い方が速くダウンロードできるといわれています。

クリック後に表示されたページからWindowsの場合は「Download R for Windows」をクリックします。「install R for the first time」をクリック後「Download R 3.6.1 for Windows」をクリックすることでダウンロードが始まります。ダウンロードが終了したら、ダウン

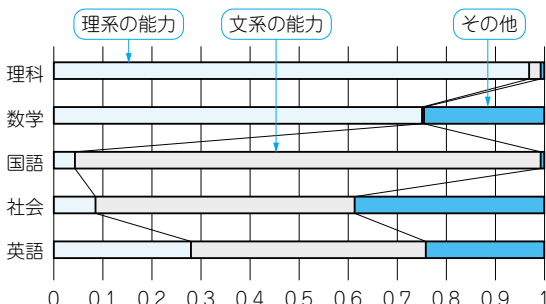


図1 因子分析は観測可能なデータから観測不可能なものへの傾向をあぶり出す

ロードしたR-3.6.1-win.exeを実行するとインストールが始まります。

#### ● データを準備する

今回も、前回使用した英語、国語、社会などの成績表データを使います。データ study\_esk.csvは本誌ウェブ・ページから入手できます。

<https://www.cqpub.co.jp/interface/download/contents.htm>

因子分析を始める前に、study\_esk.csvをドキュメント・フォルダに移動します。なお、他のフォルダに置いたデータを使いたいときには「ファイルメニューからディレクトリの変更」を選択することで行ってください。

#### ● Rのコマンドに触れてみる…データの読み出し / 表示と共分散を求める

Rでstudy\_esk.csvのデータを読み出すにはread.csvコマンドを用います。このコマンドは、dataだけ入力することで読み込まれたデータを表示させることができます。コマンドを以下に示します。

```
>data = read.csv("study_esk.csv",
header=T)
>data
```

実行結果は、リスト1に示します。

次に、共分散を以下のcorコマンドで求めてみます。

```
>cor(data)
```

実行結果をリスト2に示します。この結果を見ると、

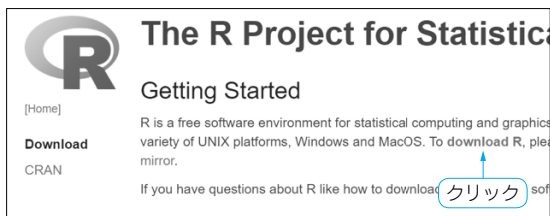


図2 統計解析環境 R をインストールする