

AI 自習ドリル

第2回 新しい軸を追加してデータを多角的にみる「主成分分析」

牧野 浩二，足立 悠

主成分分析とは、たくさんの項目からなるデータ（アンケート・データや特徴データ）を人間に分かりやすい形で表示するための方法です。このドリルでは簡単な例題を用いて説明し、イメージをつかむところから始めます。

1 主成分分析でこんなことができる

主成分分析のできることの例を挙げてみます。

● 例1…選手や生徒の特性を抽出する

▶スポーツ選手

サッカー選手のデータとして、身長、体重の他にも持久走や短距離のタイム、パスやシュートの成功率などが蓄積されていたとします。これらを分類して特性を調べることができます。

▶学生の成績

成績は5教科で評価させることが多くありますし、場合によっては音楽や体育なども含めて評価することもあります。全部の評価を見ても、どの学生がどのような資質があるのか分かりにくいです。簡単な指標でみることで個々の学生の特徴をつかむことができます。

● 例2…ブランドや商品がどう思われているかを知る

▶企業ブランド

例えば腕時計を例に挙げると、セイコー、シチズン、カシオ、ロレックス、タグホイヤー、ノモス、スント、アップルなど多数のメーカーがあります。それぞれに使いやすさ、機能、価格、大きさ、重さ、歴史などのアンケートを取って、見やすい形にまとめると、ブランドの傾向がつかめます。

▶ドリンクの味わい

飲み心地や味などについて調べて分類することもできます。その結果から新しいジャンルの飲み物を検討することもできます。本稿でも紹介します。

● 例3…内閣府の調査でも使われている

内閣府の調査で以下のように活用されていました。

- 歳入・歳出構造の主成分分析⁽¹⁾
- 財政・金融における構造改革の進展⁽²⁾
- 仕事の性質と充実度に関する主成分分析結果⁽³⁾

ここでは歳入構造を分析するのに、どのように主成分分析を用いたかを簡単に説明します。この分析では図1-1に示すように、世界の幾つかの国（ドイツ、イタリア、フランス、英国、カナダ、オーストラリア、米国、フィンランド、スウェーデン、デンマーク）と日本の歳入構造の比較を行っています。

報告書をよく読むと、分析のために各国の個人所得税、法人所得税、社会保険料、財産税、間接税の5項目を点数化した5次元のデータを用いていることが分かりました。

主成分分析の結果、スウェーデン、フィンランド、デンマークからなる北欧式型、ドイツ、フランス、イタリアからなる大陸欧州型、米国、英国、カナダ、オーストラリア、日本からなる日米型に分けることができそうな分類結果が得られています。

そして、日本は法人税収と財産税収比率が他の国よりも高いことが分かる結果が得られました。確かに、結果の図を見ると納得してしまうような分布が得られています。なお、各国が点と線で表されているのは調査した年の違いを表しています。