

ヒットチャートと楽譜データに基づく日本の年代別流行曲のクラスタリング

Clustering Japanese Historical Popular Music from HIT Charts and Musical Score Data

宮田香月 赤木茅 江草遼平 寺野隆雄

* Kazuki Miyata, Kaya Akagi, Ryohei Egusa, and Takao Terano

千葉商科大学

Chiba university of Commerce

Abstract:

In this study, we analyzed the characteristics of hit songs from 1970 to 2022 in Japan in order to clarify how popular music has changed over time. The characteristics of the data were classified into seven categories: year, decade, BPM (Beat Per Minute), key, gender, number of transpositions, and the number of notes in the entire song. The results of principal component analysis and clustering are as follows. Principal component analysis and clustering revealed that the characteristics of hit songs were divided into four groups: sad songs, high-tempo women, mid-tempo men, and ballads & bands.

1. はじめに

近年、流行の音楽の消費速度が加速しているといわれている。日本において流行する曲の特徴には、不明点が多い。越川（2021）は、18~23歳の大学生1265名（男885人、女380人）を対象に、「理由はわからないが好きな曲」の調査を行った。複数の回答者から挙げられた80曲について、クラスター分析を行った結果、広告媒体として若者に好まれる音楽特徴を明らかにした[1]。先行研究では、曲の選定がアンケート回答者の主観に基づいており、また対象の属性も限定的である。本研究では、日本の1970年から2022年までのヒット曲を分析し、その特徴を明らかにすることで、日本のヒット曲の特徴とその年代ごとの推移を明らかにする。

2. 方法

2.1. データ

本研究では楽曲を、西暦、BPM(Beat Per Minute)、調、性別、転調回数、曲全体の音の開きの6項目に分類し

たデータを使用した。カラオケDAMの公式Webページ[2]から楽曲、DAMランキングのデータを取得し、1位の楽曲をヒット曲と定義した。楽曲は1970年から2022年までの全53曲である。この53曲についてDAMのWebページから楽曲名、西暦、性別のデータを収集した。さらにBPM、調、転調回数、曲全体の音の開きのデータはヤマハの「ぷりんと楽譜」[3]から配信されている楽譜に基づいて取得した。

各特長の概要は表1の通りである。

表1 曲の特徴

特徴	概要
西暦	楽曲の発表年
性別	男性または女性
BPM	1分あたりの拍数
調	長調または短調
転調回数	調が変わる回数
音の開き	楽曲中で最大となる音階の差

2.2. 分析手法

本稿の目的は年代ごとの曲の特徴の推移及び、その特徴である。そこで、性別、調、BPM、転調回数、曲全体の音の開きの5変数から主成分分析を行い第1,2

主成分を特徴を表す変数とした。

曲の特徴を把握するために、第一主成分と第二主成分を変数として k-means によってクラスタリングを行った。クラスタ数は、2 から 7 まで試し、主観的に理解しやすい結果が得られた 4 とした。

実施にあたっては Python ライブラリの scikit-learn の cluster.Kmeans (ver 0.14.0) [4]を利用した。

3. 分析

図 1 は、各変数の組み合わせごとの散布図である。自交点にはその変数のヒストグラムが表示されている。西暦と曲全体の音の開きの間に相関が見られるが、それ以外の項目単体によって、年代別の推移を説明することは困難であることが読み取れる。

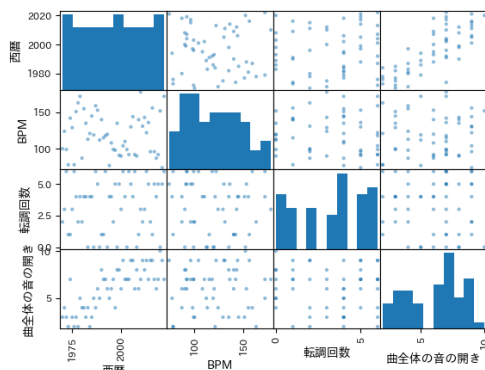


図 1 各変数のクロスプロット

表 2 は、主成分分析における各主成分の寄与率を表したものである。第一主成分の寄与率は 0.270177、第二主成分の寄与率は 0.252224 であった。第一主成分と第二主成分の寄与率の合計で 50%を超えているので、本分析はこれを採用する。

表 2 各主成分の寄与率

主成分	寄与率
第 1 主成分	0.270177
第 2 主成分	0.252224
第 3 主成分	0.198643
第 4 主成分	0.150510

図 2 は、主成分のプロットに各変数の情報をカラスケールとして加えている。性別を見ると、女性が右上、男性が左下に固まっている。調は短調が左上に固まっている。転調回数は、左から右に向かって上昇している。音の開きは上に行くほど大きくなっている。西暦に関しては古い年代の楽曲は真ん中から左下側に多くプロットされており、新しい年代の楽

曲は右上側に多い傾向が見て取れる。主成分に西暦は入っていないが、一定程度の年代別の曲の特徴の推移が存在していることが分かる。

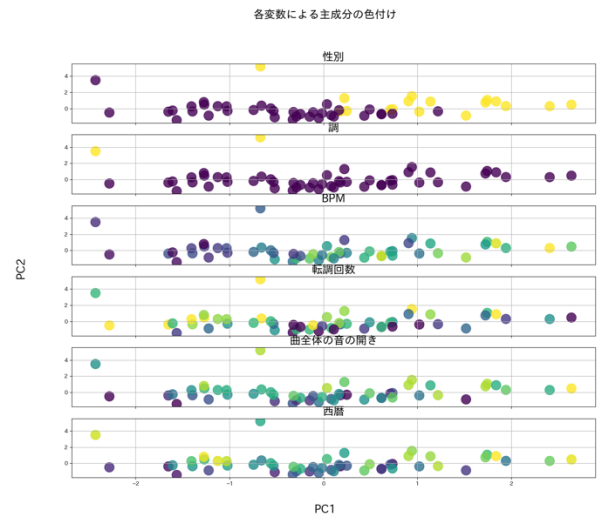


図 2 主成分プロット (カラスケール=年代)

k-means によるクラスタリング結果は紙幅の都合により、付録に提示した。クラスタ数は 4 に設定され、クラスター毎に異なる色で着色されている。各点に楽曲名の情報を追加した。見やすくするために、各クラスターを円で囲み、それぞれのクラスターを代表する名前をつけた。

命名の理由と、クラスターごとの特徴は以下のとおりである。悲しい曲クラスターは短調の曲のみで占められている。ハイテンポ女性アイドルクラスターは BPM と曲の開きが大きく、ボーカルは女性アイドルが多い。男性ミドルテンポクラスターは男性ボーカルのミドルテンポが大半である。バラード&バンドクラスターはテンポが低く、歌詞が恋愛をテーマとしている特徴があった。

4. 考察

クラスタリングを行った結果、選定した 53 曲を 4 グループに分けることが出来た。各グループを構成する楽曲から、それぞれのグループを特徴づける年代、特性が示唆された。曲の特徴から、年代ごとのヒット傾向が分析可能であることがわかった。また、各グループの特性から日本におけるヒット曲の要件に関しても示唆が得られた。

もの悲しいクラスターは短調の曲で構成されている。しかし、もの悲しいクラスターの曲数は極めて少なかった。このことから、もの悲しい楽曲は比較的ヒッ

ハイテンポ女性クラスタは、ほとんどの曲が 2000 年以降の新しい年代の楽曲で構成されており、近年のヒット曲の特徴は、高い BPM と、広い曲の開きで構成されていることが分かる。

バラード&バンドクラスタは、1970年代から2020年代と幅広い年代の楽曲で構成されていた。このクラスタのほとんどテンポが低く、歌詞は恋愛をテーマとしていた。この結果から、恋愛曲は時代に関係なくヒットする傾向があると示唆される。

5. 謝辭

6. 参考文献

- ## 7. 付録

