

# カーネルPCAを用いた音楽のジャンル分析

丹澤 春香 (指導教員: 吉田 裕亮)

## 1 はじめに

音楽ストリーミングサービスの普及により、大規模な音楽データの解析と活用が研究課題となっている。特に、音楽ジャンルの自動識別や特徴分析は、音楽推薦システムの基盤となる重要な技術である。しかし、楽曲データから特徴抽出を行う際に、どの情報が各ジャンルを特徴付ける重要な要素であるかを明示的に判断することは難しい。この課題に対して、統計的機械学習法の一つである PCA (主成分分析) は、多変量を持つデータに対して本質的な特徴を抽出することが有効な手法となるであろう。

本研究では、Spotify API から取得した楽曲データに対してカーネル PCA を適用し、音楽ジャンルを特徴付ける本質的な変量の推定を試みる。これにより、音楽ジャンルの客観的な特徴づけと、より効果的な音楽推薦システムの実現に貢献することを目指す。

## 2 先行研究

音楽ジャンル識別における先行研究では、主に MFCC やスペクトル特徴量などの音響特徴を用いた機械学習による手法が提案されており、Tzanetakis らは [2] で、ティンバル・テクスチャ特徴量、リズム特徴量、ピッチ特徴量を組み合わせた手法を提案し、61%の識別精度を達成した。その後、Ali らは、[1] で MFCC と統計的の特徴量を組み合わせ、SVM を用いることで 77%まで精度を向上させた。これらの研究では、classic や jazz は高精度で識別できる一方、rock や blues など境界が曖昧なジャンルの識別は依然として課題となっている。

一方で、音楽ジャンル識別にカーネル PCA を適用した研究は、まだ報告されていないと思われる。カーネル PCA は非線形な特徴抽出が可能であり、音楽データのような複雑な構造を持つデータの分析に適している可能性がある。そこで本研究では、Spotify API から取得した楽曲データに対してカーネル PCA を適用し、各ジャンルの特徴を視覚化することで、ジャンル識別における有効性を検証する。また、先行研究で示された特徴量の中から、実際にジャンル識別に寄与する変量を明らかにすることを目指す。

## 3 カーネル PCA

### 3.1 PCA(主成分分析)

PCA とは、分散の大きい方向にデータを射影することで、多次元データの情報を、特性を保ちながらより低い次元に縮約させる方法である。

しかし、PCA は線形データ解析手法であるため、音楽データのような非線形な構造を持つデータに対しては、その特徴を適切に捉えることが困難である。この制限を克服するために、カーネル法を用いた拡張が必要となる。

### 3.2 カーネル法

一般にカーネル法では、非線形変換を介して、データ  $x$  のいろいろな特徴量を取り出している。

$\phi_1, \dots, \phi_d$  を特徴抽出するための非線形関数とし、特徴ベクトルを  $\vec{\Phi}(x) = (\phi_1(x), \dots, \phi_d(x))^T$  と書く。このとき、カーネル関数は特徴抽出の内積に基づき、以下のように定義できる。

$$k(x, x') = \vec{\Phi}(x)^T \vec{\Phi}(x') = \sum_{m=1}^d \phi_m(x) \phi_m(x')$$

したがって、非線形に写像した空間での  $\vec{\Phi}(x)$  と  $\vec{\Phi}(x')$  の内積が、入力特徴  $x$  と  $x'$  のみで計算でき、 $k(x, x')$  から非線形写像を構成することができる。このような関数  $k$  をカーネルと呼び、このように高次元に写像しながらカーネルの計算のみで最適な非線形関数を構成することを、一般に、カーネルトリックという。本研究では、ガウスカーネル

$$k(x, x') = \exp(-\beta \|x - x'\|^2)$$

を用いる。本研究では  $\beta$  をカーネルパラメータと呼ぶ。

### 3.3 カーネル PCA

カーネル PCA とは、高次元の特徴ベクトルに変換してから、通常の PCA を施し、低次元の線形部分空間を求める非線形解析手法である。すなわち、中心化されたデータ点の集合  $x^{(1)}, x^{(2)}, \dots, x^{(n)}$  からカーネル (Gram) 行列  $K = (k(x^{(i)}, x^{(j)}))$  を構成し、固有値問題  $K\alpha = \lambda\alpha$  を解き  $M$  個の優固有値  $\lambda_1, \dots, \lambda_M$  と、それらに付随する固有値ベクトル  $\alpha_1, \dots, \alpha_M$  を用いて、 $M$  次元 PCA プロットを行う手法である。本研究においては  $M = 2$  として、第 2 主成分までの PCA プロットを行った。

## 4 楽曲データ

実データとして、5 ジャンルの楽曲データ「jazz」「classic」「rock」「blues」「hip hop」の楽曲データを 100 曲ずつ計 500 曲用いる。各楽曲について、Spotify API から 14 次元の特徴ベクトルを取得する。特徴ベクトルの成分は [2] に基づいて、ビート関連特徴量 (「danceability」「energy」「liveness」「tempo」「time\_signature」)、ティンバル・テクスチャ特徴量 (「acousticness」「instrumentalness」「loudness」「speechiness」)、ピッチ関連特徴量 (「key」「mode」「valence」)、その他 (「length」「popularity」) に分類される。

## 5 分析方法

- (i) 取得したデータにカーネル PCA を施し、できる限りジャンルが識別できるように学習し、カーネルパラメータ  $\beta$  を調整することでジャンルの識別を示す PCA プロット図を求める。
- (ii) 学習に用いた特徴ベクトルにおいて、できる限り少ない変数によって (i) で得られた PCA プロット図 (またはそれとほぼ等価な図) が維持するように変数の削減を行う。

## 6 分析の実践と結果

### 6.1 実験 1

**目的** 5ジャンルのデータにカーネル PCA を施し、最適な 2次元の PCA プロットを行い、ジャンルの識別や特徴を考察する。

**手法と結果** 「jazz」・「classic」と「rock」・「blues」・「hip hop」の大きく 2つのグループに分かれることが確認された。また 14 変量から、「energy, tempo, valence」の 3 変量にまで削減することが可能であることが分かった。

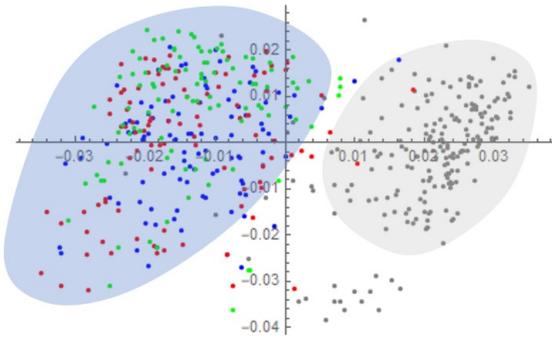


図 1: jazz・classic の識別結果。

**考察** 分析結果から、jazz と classic は他のジャンル（青色の領域）と分離され、図の灰色の領域に集中し分布していることが確認できた。この結果は、先行研究 [2] と両立するものである。すなわち、jazz と classic は他のジャンルから明瞭に分離される。特に選択された 3 つの変量 (tempo, acousticness, valence) は、これらのジャンルの特徴を効果的に捉えていると考えられる。実際に、jazz と classic は、比較的穏やかな tempo と安定した valence を示す傾向があり、これが図の緑の領域への分布として表れている。一方、rock (青)、blues (赤)、hip hop (緑) は混在して分布しており、これらのジャンル間の境界は不明瞭である。この結果より、これら 3 ジャンルの識別にはさらに別の識別機が必要であることが示唆された。

### 6.2 実験 2

**目的** 「rock」、「blues」、「hip hop」の 3 ジャンルの楽曲データにカーネル PCA を施し、2次元の PCA プロットによって識別を行い、その結果からデータの識別や特徴を考察する。

**手法と結果** 「rock」、「blues」、「hip hop」のジャンルについて楽曲データの変量「danceability, energy, speechiness, tempo, valence」を学習データに用いてパラメータ  $\beta$  を学習し最適な 2次元 PCA プロットを求めた結果が 2 である。また、この結果から「rock」、「blues」のみを抽出した図が 3 である。

**考察** hip hop (緑) と rock (青) については、図中の緑色の領域と青色の領域にそれぞれ分布が集中し、一部の外れ値を除いて比較的明瞭な分離が得られた。これは特に danceability と valence の違いが両ジャンルの特徴を効果的に表現できていることを示している。一方、blues (赤) は両者の中間に広く分布しており、明瞭な境界を見出すことができなかった。これは blues

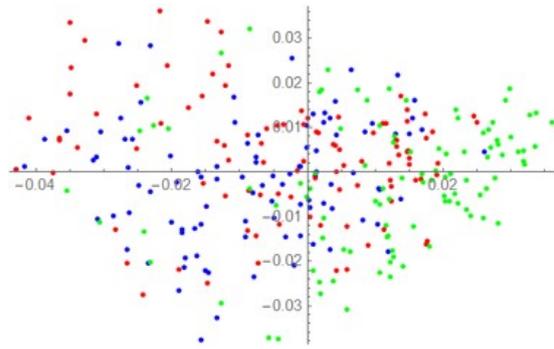


図 2: rock・blues・hip hop の PCA プロット

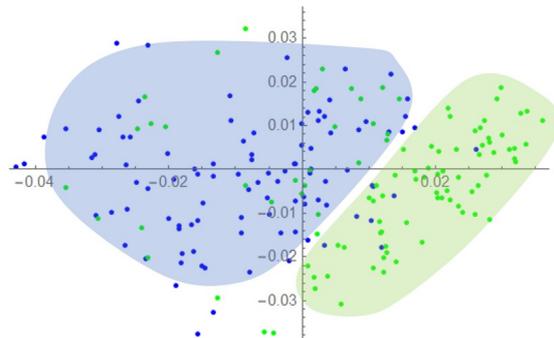


図 3: rock・hip hop の識別結果

が両ジャンルの音楽的要素を共有する過渡的な性質を持つことを示唆している。

## 7 まとめと今後の課題

本研究では、カーネル PCA を用いて音楽ジャンルの識別と特徴分析を、特に英語圏の楽曲について行った。具体的には、jazz・classic は他のジャンルから明瞭に分離でき、rock と hip hop についても適切な変量の選択により、高い精度での識別が可能であることが分かった。一方で、blues のような複数のジャンルの特徴を持つ音楽の識別には課題が残った。

今後の課題として、以下の二点を取り上げる。一点目として、分類精度の数値的検証を行うことだ。本研究ではカーネル PCA による音楽ジャンル識別の有効性を確認したが、今後は定量的な性能評価を実施する必要がある。分類精度の数値的な検証を行い、提案手法の有効性を客観的に評価することが必要だ。二点目として、日本の楽曲への応用が挙げられる。日本特有の音楽ジャンルについての分析は、これまでほとんど行われていないため、新たな知見が期待できる。

## 参考文献

- [1] Ali, M. R. and Siddiqui, K. T. A.: Automatic Music Genres Classification using Machine Learning, *International Journal of Advanced Computer Science and Applications*, Vol. 8, No. 8, pp. 337–344 (2017).
- [2] Tzanetakis, G. and Cook, P.: Musical Genre Classification of Audio Signals, *IEEE Transactions on Speech and Audio Processing*, Vol. 10, No. 5, pp. 293–302 (2002).