

# 最大マージンのアプローチによる統計的な音楽情報の解析\*

米田隆一, 西本卓也, 嵯峨山茂樹 (東大情報理工)

## 1 はじめに

本稿では、MIDI、楽譜等のシンボリックな音楽情報を入力としてラベル(対旋律、和音、調等)を付与する汎用的な手法を提案する。縦の和音、横の声部進行等の2次元的な情報は、マルコフ連鎖のような1次元構造モデルでは不十分であり、より広いコンテキストの重視が必須である。このような背景のもと、広いコンテキストを柔軟に設定でき、かつ、我々の持つ音楽的知識を素性関数の設計という操作に還元することができる手法という観点から、最大マージン法に基づくアプローチを論じる。

## 2 音楽の確率定式化

### 2.1 一般化した音楽の数理構造

旋律、和音等を作成する際、縦の和音構成、横の声部進行共に考慮にいなければならない。さらに、音楽の持つ繰返し、模倣等の構造を考えると、より遠くのコンテキストまで影響が及んでいると考えられる。これらの特性を幾何学的に表現したものが図1である。各ボックスは音高、和音等のラベルを表し、灰色のボックスは現在付与しようとするラベルであり、矢印はボックス同士が関係を持っていることを表す。これは、MIDI、楽譜が持つ2次元的な情報によくマッチする。

### 2.2 対位法、和声学における文脈依存の例

対位法においては、平行5度、同一音程の4度以上連続の禁止等の制約がある。これは、旋律の1つ、あるいは数個前後の文脈により対旋律の決定が可能である。和声学においては、典型的な和声の終止定型が存在しバス旋律より決定が可能である(例: fa-so-so-do に対する II-I-V-I 曲末)。これは旋律の先読みにより和音の決定が可能である。これらの文脈依存性は、最大マージン法における素性関数により容易に設計可能である。

## 3 最大マージン原理

### 3.1 サポートベクトルマシン

最適ラベル系列の決定問題のような、周辺の文脈に依存する値をモデル化する際、サポートベクトルマシン(以下、SVM)を順次適用してラベルを順次決定していく手法は有用であると考えられる。SVMではモデルの複雑さは入力次元数に

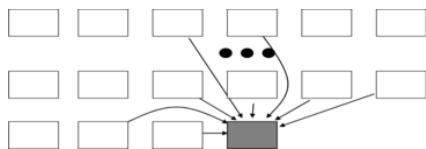


Fig. 1 音楽を一般化した数理的な構造

依存しない[1]ので、文脈として任意の素性を設計できるためである。

サポートベクトルマシン[1]は、素性空間においてマージン最大化の戦略をとる線型分類器であり、入力次元数に依存しない汎化能力をもつことが証明されている。原理としては、素性ベクトル  $x_i$  と正負の2値ラベル  $y_i$  のふたつ組み  $(x_i, y_i)$  で表現される訓練事例に対して、正負のラベルを正しく分離する平面のうちマージンが最大となる、式(1)のモデルを選択する。 $\alpha$  は各事例に対応し、学習によって求まるモデルパラメータである。最大エントロピーによるアプローチ[2]では素性に重みが付くのにに対し、本原理は分離に真に重要なサポートベクトルという事例に重みが付く点が異なる。

$$f(\mathbf{x}) = \text{sign}\left(\sum_{x_i}^{\#x} \alpha_i y_i (\mathbf{x}_i \cdot \mathbf{x}) + b\right) \quad (1)$$

また、素性ベクトルの各次元は、事例の有無をあらわす以下のような2値関数を考えることが多い。

$$x(\text{前音高}) \equiv \begin{cases} 1 & \text{if 前音高} = \text{so} \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

SVMは正負の2値のみを分類する機械であるが、和音等の複数のクラスを識別する手法として不十分である。SVMを多値分類に拡張する手法としてone-against-all法がある。これは、あるクラスかそれ以外か(図2では灰色のボックスに相当)の分類器をクラス数用意し(図2では3個の実線に相当)、クラスが曖昧な場合は分離平面から最も離れた分類器の結果を選択する手法である(図2では3クラスの識別境界は最終的に破線のようになる)。

### 3.2 SVMの順次適用

図3に、音高(階名)からカデンツラベルを決定する方法を示す。式(2)に示すとおり、前音高

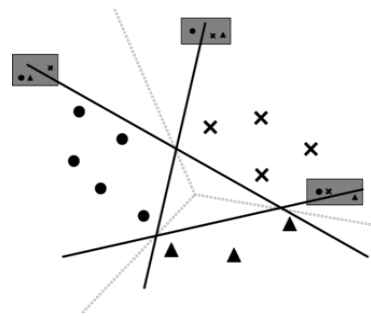


Fig. 2 SVM(one-against-all法)による多クラス分類

\*Max-Margin Approach to Music Modeling. by YONEDA, Ryuichi, NISHIMOTO, Takuya, SAGAYAMA, Shigeki (Graduate School of Information Science and Technology, The University of Tokyo)

などの事例の有無を素性ベクトルのひとつの次元と定義し、周辺のコンテキストから素性ベクトルを構成する。括弧の内容は和音、和音機能をあらわし、それが既知の場合は素性として使用してもかまわない。現ラベルが決定されるとコンテキストをシフトし、次ラベルを決定することができる。本稿ではこのようなSVMの順次適用法をラベル決定問題に使用する(調認識を除く)。

## 4 評価実験

### 4.1 実験方法

カデンツ同定と調認識の実験を除き、Humdrum Toolkit[3]に付属する和声ラベル付きのバッハのコラール16曲(humdrum-kern フォーマット)をすべて階名に変換したものを評価実験に用いた。音符を持たない弱拍部(2連続8分音符のソプラノ旋律に対応する4分音符のバス旋律の弱拍部)等には、特別のラベルを与えた。また、曲頭、曲末にもまた別の特別のラベルを与えた。また、全実験においてソフトマージンパラメータ<sup>1</sup>cを100に指定した。

### 4.2 対旋律付け、和声付け、和声解析

ソプラノとバスの旋律のペアを学習する対旋律付け実験、各声部の旋律と和音のペアを学習する和声付け実験において、テスト旋律(例えばベートーヴェン第9交響曲第4楽章の主旋律)に対する出力系列は妥当なものであると確認された。

前述のコラール16曲のうち15曲を学習データ、1曲(#054, BWV151)をテストデータとした和声解析結果は55事例中誤り14箇所(精度75%)であった(各4パートの現在音とその前後2音、および前2ラベルを素性とした)。

### 4.3 カデンツ同定

文献[4]のpp. 104-105のバス課題において、カデンツの決定にあたり、階名を入力とするK1-B, K1-I, K2-B, K2-I, K3-B, K3-I, O(K1, K2, K3はそれぞれT-D-T, T-S-D-T, T-S-Tを、B, I, OはそれぞれBegin, Inside, Otherをあらわす。)の7ラベルの付与を考えた。すなわち、各カデンツの開始部(ここでは各カデンツの先頭のTを除外したうえでの先頭とする。図4参照、曲末5音符が図3に相当)をB, 終了までをI, その他をOとラベル付けした。定量評価は、1曲とそれ以外に分割する交差検定で行った。現在音とその前後2音、前2ラベルを素性としたところ、総音符数229中、誤りは2箇所(精度99%)

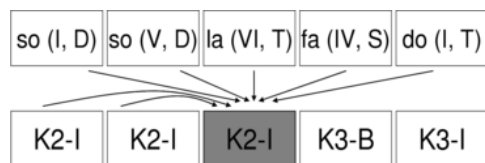


Fig. 3 現ラベル決定のためのコンテキストの例: 現音高+前後2個、前2ラベルをコンテキストとする。

<sup>1</sup>分離平面の選択時に事例のマージン領域への入り込みを許可する場合

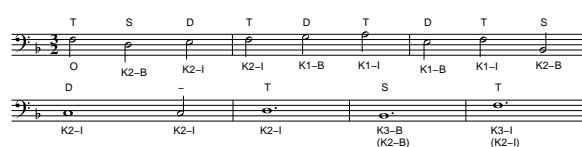


Fig. 4 K1, K2, K3, B, I, Oラベルの付与法

Table 1 実験の概観

	入力形式	定量的評価
対旋律付け	humdrum	60% ~ 70%
和声付け	humdrum	60% ~ 70%
カデンツ同定	独自形式	99%
和声解析	humdrum	75%
調認識	MIDI	63%

であった。誤り箇所は図4の曲末2音符であった(K3-B, K3-Iが正解)。

### 4.4 調認識

MIDIの音高(mod 12)の相対頻度を値とする12次元素性空間上でSVMにより調を決定した。すなわち、式(2)において素性関数を12個用意し、値を楽曲全体から得られる相対頻度とした。20~30秒程度の楽曲96曲(24調、各調4曲ずつ)からなるMIDIデータを、すべての調が1セット中に含まれるよう4分割し交差検定したところ、誤りは36曲(精度63%)であった。

## 5 まとめ

周辺の文脈に依存するラベル列のモデル化に適した最大マージン法(サポートベクトルマシン)を音楽情報の解析に適用し、シンボリックな音楽情報を入力とする種々のタスク(対旋律付け、和声付け、カデンツ同定、和声解析、調認識)を統一的な枠組みでできることを示した。今後の展開としては、音程等、音楽のモデル化に有効な素性を吟味したい。また、カデンツ同定から楽曲構造解析へ拡張したい。また、音高のヒストグラム情報のみによる調認識は改善の余地がある。

謝辞 実験には汎用の分類器<sup>2</sup>、チャンカー<sup>3</sup>を用いた。

## 参考文献

- [1] N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines and other kernel-based learning methods*, Cambridge University Press, 2000.
- [2] 陳, 米田, 西本, 嵯峨山, “マルコフ確率場に基づく統計的な音楽情報の解析,” 音講論(春), 2-2-10, 2006.
- [3] D. Huron, The Humdrum Toolkit: Software for Music Research, <http://dactyl.som.ohio-state.edu/Humdrum/>.
- [4] 島岡讓, 和声理論と実習I, 音楽之友社, 1964.

<sup>2</sup><http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

<sup>3</sup><http://chasen.org/~taku/software/yamcha/>