

Title	転調判定を用いたHMMによる和声機能同定
Author(s)	上原, 由衣
Citation	
Issue Date	2019-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/15924
Rights	
Description	Supervisor:東条 敏, 先端科学技術研究科, 修士(情報科学)

修士論文

転調判定を用いた HMM による和声機能同定

1730003 上原由衣

主指導教員 東条敏 教授
審査委員主査 東条敏 教授
審査委員 緒方和博 教授
鵜木祐史 教授
岡田将吾 准教授

北陸先端科学技術大学院大学先端科学技術研究科 (情報科学)

平成 31 年 2 月

Abstract

Along with the development of computer science, a large variety of research attempts to externalize human intuition, such as natural language understanding. Since music has similar aspects to natural language, we should consider applying the approaches of natural language processing (NLP) also to music, however, we need to take care of unique features that are different from language. Music is made from multiple normative structures, e.g., harmonic structure, metrical structure, and so on. Among existing structures, we focus on the theory of functional harmony that regulates the role of chords, and decides the key. Especially, we expect that the harmonic analysis and the key determination can be processed simultaneously as one and the same effort.

In the area of computational music research, the key detection and the function of chords acquisition have been researched independently. The key detection algorithms are mainly based on histogram of the pitch classes which referred as the key-profile, representing the importance of each pitch class in a key. The key-profile is originally obtained by a psychological experiment. Recent studies show that it can be learned statistically. Although the histogram based key detection algorithms are widely used, they need to specify the scope on the score for applying the algorithm.

An existing previous work employs the distance between chords such as Tonal Pitch Space (TPS) rather than the pitch classes. The key is detected by the Viterbi algorithm, not requiring a fixed scope. TPS defines the distance between chords in a uniform way, and thus, cannot consider the difference of music styles.

Statistical learning is a solution to consider the difference of music styles. Previous research found the function of chords, achieved statistically. There are two mainstreams, the classification of chords and generative modeling. Generative models are more advantageous because of their predictive power, and can be applied to many applications.

Generative models are, in general, difficult to fix the appropriate size, such as the number of hidden states of the Hidden Markov Model (HMM). Although the perplexity of the output probability is prone to decrease when

the number of hidden states of HMM increases, a larger number of them may be not always meaningful.

This research presents a new approach. We consider the performance of the key detection when obtaining the model size. In this research, we employ two algorithms: the generative modeling of the function of chords, and the key detection using the set partitioning problem. We use the choral pieces written by J.S.Bach as our dataset.

We modeled the function of chords of major key and minor key by HMM, and investigated the number of hidden states among 2 and 12. To learn the models, we transposed all the major keys to C-major, while all the minor ones to a-minor. A thousand different initial values were tested to learn the parameters, and the optimal values are ranked by the perplexity.

The result shows that the minimum perplexity of output probabilities decreases along with the increase of the number of hidden states. While the distances of the optimal values also increase along with the number of hidden states. This suggests that there would be many optimal values in the parameter space and it is difficult to fix a robust solution when the number of hidden states is large.

We apply the proposed key detection algorithm to obtain the appropriate number of hidden states. To this purpose, we select a key to maximize the log output probability, calculated by the obtained HMM. The pitch of target chord sequence is transposed to a possible set of candidate keys, among which one that is the nearest either to C-major or to a-minor would be chosen.

The algorithm above can select one optimal key to one target chord sequence. Using the set partitioning algorithm is a solution to detect modulations, so that we can obtain the optimal key blocks to one target chord sequence. The function of chords should work well to detect a key, especially when there are modulations in the target sequences. To investigate the appropriate model size, the sum of log probabilities of detected key blocks of all pieces in dataset is used. The obtained result shows that the model with 6 hidden states is best for both of major and minor keys.

The major key parameters with 3 hidden states are very similar to the

knowledge of the functional harmony theory, resulting the hidden state of Tonic, that of Dominant, and that of Subdominant. However, the chords are classified into finer-grained functions, up to 6 hidden states. When the number of hidden states is 4, the state for Subdominant is separated into two: $\{\mathbf{IV}, \mathbf{ii}\}$ and $\{\mathbf{vi}\}$, as well as the that of Dominant is separated into $\{\mathbf{V}\}$ and $\{\mathbf{vii}^\circ\}$ with 5 hidden states. Finally, when the number of hidden states is 6, each chord is assigned to each hidden state.

The results of minor-scale are significantly different from that of major-scale, hidden states correspond to Tonic and Dominant of the relative major key are obtained when the number of hidden state is larger than 4. With 6 hidden states, in addition to Tonic, Dominant and Subdominant, the Tonic of relative major and that of Dominant of relative major are obtained. This result reflects the feature of choral, whose melody was composed in the middle ages in Gregorian modes instead of modern tonalities, prior to the harmonization by J.S.Bach, because the relative keys share the common pitch class in Gregorian modes.

This research showed the appropriate model size of HMM by considering the performance of the key detection. In contrast to the evaluation with perplexity, the selected number of hidden states is not large, so that obtained parameters are meaningful. Since the parameters also reflect the musical style of the corpus pieces, our approach has a potential to be applied different music styles, such as music in post-romanticism. In addition, we present the new algorithm, which can detect modulation automatically without giving fixed scope and works well since the distance between chords have been learned statistically using the dataset.

Although we have realized the efficient performance in our research, we need to admit that our approach could include several mistakes in key detection, which sometimes treats small groups of chords as a modulation. This is because HMM cannot represent more external structures such as cadence. In the future, we will consider more sophisticated approaches, for example, introducing dependency of music structures into model.

目次

第 1 章	序論	1
1.1	研究の背景	1
1.2	研究の目的	3
第 2 章	基礎理論	4
2.1	音楽理論	4
2.1.1	音階	4
2.1.2	音名	5
2.1.3	音程	6
2.1.4	調	7
2.1.5	和音の種類	9
2.1.6	和声機能とカデンツ	12
2.2	HMM	14
2.2.1	Hidden Markov Model	14
2.2.2	前向き・後ろ向きアルゴリズム	15
2.2.3	Viterbi アルゴリズム	16
2.2.4	Baum-Welch アルゴリズム	17
2.3	集合分割問題	20
2.3.1	数理最適化の基礎	20
2.3.2	集合分割問題	20
第 3 章	先行研究	22
3.1	和声機能学習の先行研究	22
3.1.1	クラスタリングを用いた和声機能の教師なし分類	22
3.1.2	HMM を用いた和声機能の教師なし学習	23

3.2	調判定の先行研究	23
3.2.1	ヒストグラムベースの調判定アルゴリズム	23
3.2.2	和音間距離の最短経路探索に基づく調判定アルゴリズム	24
第 4 章	転調判定を用いた和声進行生成モデルのモデルサイズ選択	26
4.1	概要	26
4.2	和声機能の教師なし学習	26
4.2.1	HMM による和声進行の生成モデル	26
4.2.2	初期値の設定	28
4.2.3	局所解の評価	28
4.3	転調解析	28
4.3.1	HMM を利用した調判定アルゴリズム	28
4.3.2	集合分割問題による転調判定	29
4.4	HMM 隠れ状態数の選択	30
第 5 章	実験と結果	31
5.1	データセット	31
5.2	和声機能の教師なし学習	33
5.2.1	初期値の設定と局所解の評価	33
5.2.2	隠れ状態数 3 における機能 and 声理論との類似性	34
5.3	調判定結果による隠れ状態数選択	35
5.3.1	隠れ状態数と調判定結果の対数尤度合計	35
5.3.2	選択された隠れ状態数 6 の性質 (長調)	36
5.3.3	選択された隠れ状態数 6 の性質 (短調)	38
5.4	集合分割問題を用いた調判定	40
5.5	実験結果のまとめ	50
第 6 章	終わりに	54
参考文献		56

目次

2.1	半音階	4
2.2	長音階および(自然)短音階	5
2.3	和声的短音階および旋律的短音階	6
2.4	音程	7
2.5	長調の例: D major	8
2.6	最も調号が少ない場合	8
2.7	3和音の原型	9
2.8	7の和音の原型	11
2.9	音階上の3和音および7の和音	12
2.10	和声の機能進行	13
2.11	HMM	14
4.1	HMMによる和声進行モデル	27
5.1	隠れ状態数とパープレキシティ	34
5.2	隠れ状態数とパラメータの類似度	34
5.3	隠れ状態数3のパラメータ(長調)	35
5.4	隠れ状態数と対数尤度合計	36
5.5	隠れ状態数6のパラメータ(長調)	37
5.6	隠れ状態数3-6の和音出力確率(長調)	38
5.7	隠れ状態数6のパラメータ(短調)	39
5.8	BWV281の調判定結果(楽曲全体)	41
5.9	BWV281の調判定結果(フレーズの区切りを与えた場合)	42
5.10	BWV255の調判定結果(楽曲全体)	44
5.11	BWV255の調判定結果(フレーズの区切りを与えた場合)	45

5.12	BWV266 の調判定結果（楽曲全体）	46
5.13	BWV266 の調判定結果（フレーズの区切りを与えた場合）	47
5.14	BWV271 の調判定結果（楽曲全体）	48
5.15	BWV271 の調判定結果（フレーズの区切りを与えた場合）	49
5.16	BWV292 の調判定結果（楽曲全体）	51
5.17	BWV292 の調判定結果（フレーズの区切りを与えた場合）	52

表目次

2.1	調性一覧	10
2.2	3 和音の種類と名称	10
2.3	7 の和音の種類と名称	11
2.4	和音機能と分類	13
3.1	Krumhansl-Schmuckler アルゴリズム Major Key-profile	24
3.2	Krumhansl-Schmuckler アルゴリズム Minor Key-profile	24
3.3	ピッチクラスのレベルの深さ	25
4.1	長調の構成和音	27
4.2	短調の構成和音	27
5.1	出力記号の構成比率 (長調)	32
5.2	出力記号の構成比率 (短調)	32

第 1 章

序論

1.1 研究の背景

統計的機械学習の発展に伴い，従来人間の経験知によって設計されてきた知識を統計的手法により獲得する取り組みが様々な分野で活発になっている．音楽学は言語との共通点がしばしば指摘されており，自然言語処理研究の豊富な実績を受けて，統計的手法による知識の獲得が期待されている分野の一つである．一方，音楽の構造には音楽に特有の部分もあり，言語とは異なる点も多い．

楽曲にある種の構造が内在していることは周知の事実であるが，それらを研究し体系化するのが音楽学である．しかしながら，その構造は全体としては非常に複雑なものであり，いくつかの視点に分解され，その中で理論が整理されてきた．中でも和声学と呼ばれる分野は，比較的明瞭な法則性が見られ，音楽分析における必須の要素技術となっている．和声学とは狭義には機能และ声理論のことを言う．機能และ声理論の主張は，調性を確立するという働きにおいて，表層の和音がいくつかの機能に分類されるというものである．和音の機能すなわち和声機能とは，調性を確立する働きのことであるから，和声解析と調判定は切り離せない関係がある．すなわち和声機能の知識が得られたならば，和声解析とは，その知識を用いて曲中の局所的な調性すなわち転調を判別することと基本的に同じである．

音楽情報科学の分野では，調判定と和声機能の獲得は独立に研究されてきた．調判定アルゴリズムは，Krumhansl-Schmuckler[6]に代表されるようにピッチクラスのカスタムヒストグラムに基づく手法が主流である．このピッチクラスのカスタムヒストグラムは Key-profile と呼ばれ，元々は心理学的実験によって獲得されたが [6]，近年の研究では統計的に学習可能であることが報告されている [13]．ヒストグラムベースの調判定は広く用いられているが，あらかじめ判定対象範囲を定める必要があるという制約がある．

一方、ピッチクラスではなく和音間の距離を用いた手法も存在する。Sakamoto ら [7] は Tonal Pitch Spaces (TPS)[8] に基づく和音間距離をコストとした最短経路探索による手法を提案したが、これはヒストグラムに基づく方法とは異なり固定の判定対象範囲を定める必要がない。Tonal Pitch Space[8] は F.Lerdhahl の考案した理論で、和音の相対的な重要度を考慮し和音間距離を統一的な手法により算出するものである。しかしながら、和音間距離は時代や楽曲様式によって異なることを考えると、良好に当てはまるのが古典ロマン期の楽曲に限られるという問題点がある。

統計的学習は楽曲様式を考慮するための解決策の一つである。先行研究において、機能と声理論の知見が統計的手法により獲得できることが報告されている。和声機能の統計的学習には大きく分けて、クラスタリングと生成モデルの2種類がある。クラスタリングベースの手法では、Rohrmeier ら [14] や Jacoby ら [15] の研究があり、基本的に和声機能の分類を目的としたものと言える。いずれも J.S.Bach のコラール作品を用い、機能と声理論の知見に近い知識をデータから統計的に獲得した。一方、Tsushima ら [9] は生成モデルとして Hidden Markov Model (HMM) および Probabilistic Context-free Grammar (PCFG) を用い、コードネームの与えられたポピュラーおよびジャズの楽曲について、機能と声理論の知見に近い知識が得られることを報告している。

得られた和声機能の知識を調判定やメロディへの和声付けなどの応用に活かすことを考えると、クラスタリングよりも生成モデルによる手法が有効である。Tsushima らの研究 [9][10] によりその可能性が示されたが、ポピュラー音楽は一般にクラシック音楽に比べ和声進行が単純であるものが多く、より複雑な和声進行を含むクラシック音楽において同様の手法が有効であるかどうか検討の余地がある。クラスタリングによる手法の例として紹介した Jacoby らの研究では、補足研究として、J.S.Bach のコラール作品のうち長調のデータセットにおいて、生成モデルである HMM を用い、隠れ状態数を 3 とし出力記号として 7 つの音階構成音を用いた場合には機能と声理論の知見が得られた一方、転回形を考慮して出力記号を増やすと機能と声理論の知見に合致する結果をもたらさなかったと報告された。[15]。しかしながら Jacoby らの研究における HMM は補足研究の位置付けであり、HMM の学習結果に影響を及ぼす初期値の扱いや隠れ状態数の選択について検討が必要である。

一般に、生成モデルにおいて、適切なモデルサイズを選択することは難しい問題である。モデルサイズとは例えば、HMM における隠れ状態の数などである。Tsushima ら [9] の研究では、隠れ状態の数を増やすほどパープレキシティは減少しすなわち尤度は増加する傾向があることが示されたが、大きい隠れ状態数のモデルが必ずしも音楽的に本質的な意味があるとは限らないという問題点がある。

1.2 研究の目的

本研究の目的は、先行研究において HMM による統計的学習により特定の少ない隠れ状態数で機能と声理論の知見が得られることが知られていたものの、尤度の観点からは隠れ状態数が多いほど良く、最適な隠れ状態数を決定する方法が知られていなかった点に対し、得られた HMM を転調判定に用い、その結果得られる尤度の合計によって最適な隠れ状態数を決定することを提案・実装することである。データセットとして、コードネームの振られていないクラシックの楽曲である J.S.Bach のコラール作品を用いる。そして提案手法により選ばれたパラメータを機能と声理論の知見と照らし合わせ、その内容を定性的に分析する。

本研究では、以下の 2 つのアルゴリズムを用いる。

1. HMM による和声進行のモデリングと統計的学習
2. 集合分割問題を用いた転調判定

HMM の学習においては、隠れ状態数 2–12 の場合について 1000 通りの初期値からそれぞれ学習を行う。学習にあたっては、調性が調号から判断できる冒頭および終端のフレーズを用い、長調を C major に、短調を a minor に移調する。転調判定については、学習で得られたパラメータを用いた HMM により対数尤度を計算し、これをスコアとした集合分割問題により転調判定を行う。ここで学習された HMM の隠れ状態のパラメータは和声機能に相当した内容となることを期待しており、この転調判定は和声機能が調性を確立するという考え方をストレートに実現するアルゴリズムである。

そして以下の 2 つの観点から HMM による和声進行モデルを評価し選択する。

1. 同じ隠れ状態数の HMM に対し、初期値により局所解が異なるため、Tsushima ら [9] の研究にならいパープレキシティにより最も良い局所解を選択
2. 異なる隠れ状態数の HMM に対し、最適な隠れ状態数を見出すため、転調判定を行い同じ調に属する部分系列の対数尤度の和を持って、その状態数を選択

以上のように、機能と声の調性を確立する働きに着目し、和声機能の統計的学習と転調判定アルゴリズムの構築を一体の研究として行うことによって、HMM による和声進行モデルに対する最適な隠れ状態数を選択し、選ばれたパラメータの示す知見を明らかにする。また、以上に付随して局所的転調に対応可能な転調判定システムを併せて獲得する。

第 2 章

基礎理論

2.1 音楽理論

2.1.1 音階

音階とは、1 オクターブ中の構成音を高さの順に並べたものである。様々な時代や地域の音楽を含めると音階には多様な種類があるが、ここでは、17 世紀以降のクラシック音楽で基本的に用いられる長音階と短音階、およびその導入として半音階を紹介する。

まずはじめに半音階を図 2.1 に示す。これは、オクターブを 12 等分した音を順に並べた音階になっている。オクターブはある音とその音に対し周波数が 2 倍である音との関係を示す用語である。隣り合う音の周波数比は全て等しく $\sqrt[12]{2}$ であり、これを半音と呼ぶ。ここで、この音階に多数の \sharp 記号 (音程を半音上げることを表す臨時記号) が見られるが、これには歴史的文化的な事情があり、後に述べる長音階、短音階のような全音階のほうが標準であることに由来する。なお、半音階の書き表し方は一通りではなく、図 2.1 の上段のようも下段のようにも書ける。これは後に述べる異名同音の問題があるからである。

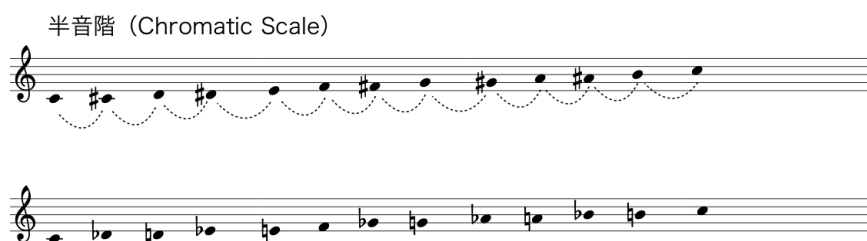


図2.1 半音階



図2.2 長音階および(自然)短音階

長音階 (Major Scale), 短音階 (Minor Scale) は図 2.2 の通りである。長音階, 短音階は全音階と呼ばれる音階群の一種である。全音階は 1 オクターブ中の構成音が 7 音である 7 音音階の一種であり, 全音 5 つと半音 2 つを含む。ここで, 全音とは半音 2 つ分の音程をいう。歴史的には全音は必ずしも半音 2 つ分とは限らないが, ここでは詳細には立ち入らずこの定義を用いる。図 2.2 において実線で表している音程が全音, 点線で表している音程が半音である。先に述べたように半音階では多数の臨時記号が必要なのに対し, 長音階, 短音階では臨時記号は現れない。このことは, 5 線譜の 1 ステップは全音階の隣り合う音分の増減を表しており, つまり 1 ステップが全音の場合と半音の場合があって周波数比の意味で均等ではないことを意味する。また, これらの全音階はピアノでは全て白鍵に相当する。

音階の起点および終点となる音を 主音と呼ぶ。主音から主音までが 1 オクターブである。主音を I 度と数え, 順に II 度, III 度, … VII 度と数える。主音に加え, V 度は属音, IV 度は下屬音という特別な名称を持つ。また, 主音からみて半音下の音を導音と呼び, 導音は主音へ強く進行する性質を持つ。図 2.2 に示した長音階では第 VII 音が主音から見て半音下の関係にあるのに対し, 短音階では全音下であり導音になっていない。そのため実際の楽曲では図 2.3 に示すように, 短音階の VII 度が半音高くなる和声的短音階が用いられることが多い。さらに旋律においては, 滑らかさを考慮した旋律的短音階がしばしば用いられる。

2.1.2 音名

音名とは, 普通には鍵盤楽器の各鍵の名前であるとするのが自然である。絶対音高を表すとされることもあるが, 正確にはある音名に対する絶対音高の物理的定義があるわけではない。鍵名としての音名では, 例えばピアノであれば 88 鍵のそれぞれを識別可能な名称にする必要があるが, オクターブ違いの音が人間にとって同質のものに聞こえるという性質から, オクターブの違いを同一視したピッチクラスとしての音名を用いることが普通である。言い換えるとピッチクラスとは同じ音名で表される音の集合である。

和声的短音階 (Harmonic Minor Scale)



旋律的短音階 (Melodic Minor Scale)



図2.3 和声的短音階および旋律的短音階

また、全音階が基準となっていることを背景に、音名も全音階の構成要素に対してつけられるのが普通である。全音階は7音音階であるから、音名は7つあれば良い。具体的な名は国や地域によって異なるが、本稿では英米方式の {A,B,C,D,E,F,G} を用いる。図 2.2の短音階が左から順に ABCDEFGA に相当し、一方長音階は CDEFGABC である。これらは、先に述べたようにピアノでは全て白鍵に相当する音である。では、黒鍵に相当する音はどのように表せば良いのだろうか。英米方式では記譜と全く同様に音名に♯やbをつけて表せば良い。ところが、例えばCとDの間の黒鍵の音には2通りの表し方がある。Cを半音高くしたものと考えればC♯であるし、Dを半音低くしたものと考えればD♭である。どちらを用いるかは一般に調の解釈によって異なり、このように同じ音が別の名称を持つことを異名同音という。

2.1.3 音程

音程は隣り合う音のピッチの差を表す用語である。音程は数字を使って表すのだが、前節で述べたように記譜法および音名が全音階を基準としていることを背景としてやや複雑な定義になっている。さらに用いられる数字は数学的な意味を持たず、したがって演算不能なことに注意しなければならない。

音程では、同一の音が1度である(0度とは表現しない)。そして、全音階上で隣り合う2音の関係が2度である。全音階は7音音階であるから、音階内には1度から7度までの音程が存在することになる。また、1オクターブ上の音との関係は8度という(ゆえにオクターブという)。ここで、全音階上で隣り合う音の関係には全音と半音があったことを思い出すと、それらを同じ2度と表すのは不都合である。そこで隣り合う全音は長2度、隣り合う半音は短2度と表す。3度も同様に含まれる音程の集合が「全音2つ」の場合と

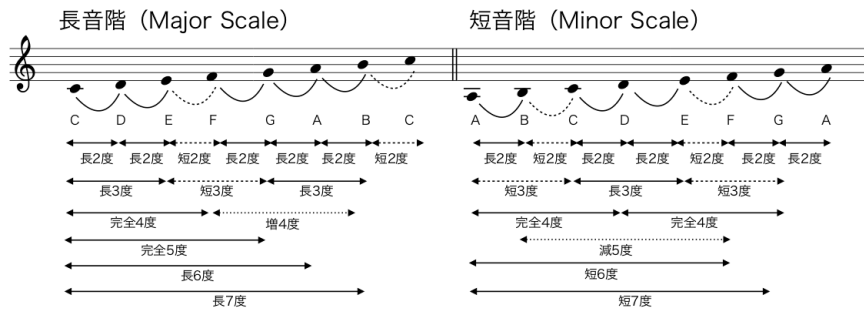


図2.4 音程

「全音1つと半音1つ」である場合の2種類があり、それぞれ長3度、短3度と表す。6度、7度も同様であり、長6度は「全音4つと半音1つ」に対し短6度は「全音3つと半音2つ」、長7度は「全音5つと半音1つ」に対し短7度は「全音6つと半音2つ」である。4度、5度は若干事情が異なる。4度は「全音2つと半音1つ」の場合と「全音3つ」の場合があり、5度は「全音3つと半音1つ」の場合と「全音2つと半音2つ」の場合がある。ここで、「全音3つ」と「全音2つと半音2つ」は半音2つで全音であることから実は音響的には全く同じである。これは古くから3全音と呼ばれ、不快な響きとされてきた音程である。3全音は全音階上ではFとB（またはBとF）の関係に対してのみ生じ、他の4度および5度は全てもう一方の音程である。このことから、3全音ではない4度および5度をそれぞれ完全4度、完全5度と表す。一方、3全音に相当する4度を増4度、3全音に相当する5度を減5度と表す。これは3全音は完全4度より半音広く完全5度より半音狭いためである。以上に述べた音程の一例を図2.4に示す。

2.1.4 調

調は音階の種類（長音階か短音階か）と主音の組み合わせによって決まる。例えば図2.2に示した長音階は、主音がCであるのでC majorと呼ばれる。一方同じ図に示した短音階は、主音がAであるのでa minorである。長音階に属する調を長調といい、短音階に属する調を短調という。一般に長調の場合は主音を大文字で書き表し、短調の場合は主音を小文字で書き表す。

まず、長調の一例を示す。先に図2.2で示した長音階の主音はCであったが、主音をDとした長音階D majorは図2.5のようになる。長音階の全音・半音の並びを保つために必要な臨時記号はこの場合2つ（FとC）である。D majorにおいては常にFとCに♯が



図2.5 長調の例: D major

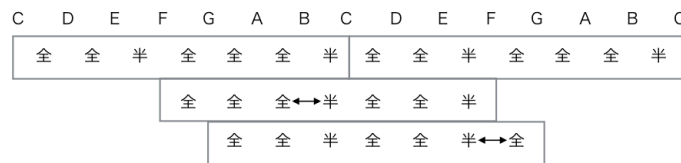


図2.6 最も調号が少ない場合

必要であるため毎回書くのは面倒である。そこで、一般に右のように音部記号の横に調号を配置して、以後調号が示す変化記号が常に適用されるものとするのが普通である。

さて、このように各主音に対して必要な調号を調べていたのでは効率が悪い。そこでまずは、調号なしの場合を除いて最も調号が少なくすむ主音は何かを考える。ところで、長音階とは全音と半音の並びが「全全半全全全半」であった。主音を C から DEF... などと 1 音ずつずらしていくことを考えるとき、この並びは 1 オクターブ上または下の C に移動するまで決して重ならないことに注意したい。まさにそのことによって、物理的な音高がわからなくとも相対的に音階が識別できるのである。主音を移動したとき、この全音と半音の並びに対し差異が最も小さくなるものが 2 つある。図 2.6 に示すように主音が F の場合と G の場合である。ここで、矢印で示した全音と半音の関係を入れ替えれば、長音階の並びになることがわかる。それを行うには、主音が F の場合は B に \flat をつければ良く、主音が G の場合は F に \sharp をつければ良い。G は C に対し完全 5 度上の音であるから、以降同じことを繰り返せば良い。すなわち G に対し完全 5 度上の D を主音とする場合、G major に対してつけた調号 $F\sharp$ に加えて $F\sharp$ の完全 5 度上の $C\sharp$ も調号に追加すれば良い。こうして、図 2.5 に示したように D major の調号が $F\sharp$ および $C\sharp$ となるのである。F に続く場合も考え方は同様であるが、F は C に対し完全 4 度上の音である一方、1 オクターブ上の C からみて完全 5 度下の音ともいえる点に注意したい。すなわち完全 5 度下に向かって順に主音を移動し調号を加えていけばよい。

この操作をどこまで続けることができるだろうか。これには、2 つの考え方がある。1 つは、7 音音階の全ての音を変化させた場合すなわち調号 7 つまでという考え方である。も

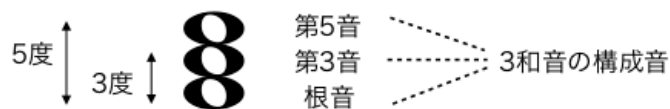


図2.7 3和音の原型

う一つは、この5度の操作を続けていくと $\sharp 6$ つと $b 6$ つの調が異名同音すなわち主音 $F\sharp$ と Gb の調となるので、これ以降はもう一方がより少ない調号で音響的に同じ調を表現できるため調号6つまでという考え方である。なおこのように6回で異名同音に達するのは、5度は半音7つに相当するので1オクターブの半音数12と互いに素であり、12回たどるまで元の音名に戻らないが、6がちょうどその中間点に当たることに起因する。ところが実際には、長調では \sharp の調号は7つまで、 b の調号は6つまでが使われることが多い。その理由は $\sharp 7$ つの $C\sharp$ major は $\sharp 4$ つの $C\sharp$ minor という主音が同じ短調(同主調)が存在するのに対し、 $b 7$ つの Cb major は同主調が存在しない(Cb は B の異名同音なので B minor が用いられる)ためと考えられる。同様の理由で短調では \sharp の調号は6つまで、 b の調号は7つまで用いられることが多い。

なお、このように長調と短調では基本的に同じ調号をもつ調が1組あり、互いに平行調と呼ばれる。平行調は、例えば G major と e minor の関係がこれにあたる。一方、主音が同じ長調と短調の関係を同主調という。同主調は、例えば C major と C minor の関係がこれにあたる。

本節の最後に、調性の一覧を表 2.1 に示す。

2.1.5 和音の種類

和音とは、秩序を持って組み合わせられた音の集合体のことである。その要素となる音のことを構成音という。構成音の数と種類に応じて種々の形体があるが、17世紀以降のクラシック音楽で最も基本的なものが3和音形体である。和音の原型は図 2.7 に示す通りである。原型において最も低い音を根音といい、根音に対し3度上の構成音を第3音、根音に対し5度上の構成音を第5音という。

さらに3和音は構成音の音度によって以下表 2.2 のように分類される。

3和音の原型の上に根音に対し7度上の構成音である第7音を追加すると、図 2.8 に示すように7の和音の原型ができる。7の和音は基礎となる3和音の種類と第5音と第7音の3度音程の種類によって以下表 2.3 のように分類される。なおこの表にない組み合わせ

長調 major	短調 minor	調号
C major	a minor	なし
G major	e minor	$\sharp\{F\}$
D major	b minor	$\sharp\{FC\}$
A major	f \sharp minor	$\sharp\{FCG\}$
E major	c \sharp minor	$\sharp\{FCGD\}$
B major	g \sharp minor	$\sharp\{FCGDA\}$
F \sharp major	d \sharp minor	$\sharp\{FCGDAE\}$
C \sharp major	a \sharp minor	$\sharp\{FCGDAEB\}$
F major	d minor	$\flat\{B\}$
B \flat major	g minor	$\flat\{BE\}$
E \flat major	c minor	$\flat\{BEA\}$
A \flat major	f minor	$\flat\{BEAD\}$
D \flat major	b \flat minor	$\flat\{BEADG\}$
G \flat major	e \flat minor	$\flat\{BEADGC\}$
C \flat major	a \flat minor	$\flat\{BEADGCF\}$

表2.1 調性一覧

根音と第3度のなす音程	根音と第5音のなす音程	名称
長3度	完全5度	長3和音
短3度	完全5度	短3和音
短3度	減5度	減3和音
長3度	増5度	増3和音

表2.2 3和音の種類と名称

は、次に述べる音階上の7の和音に現れない。

音階の各音を根音として3和音および7の和音を構成すると図 2.9のようになる。ここで短調で3和音および7の和音を構成する際は、通常2.1.1節で述べた和声的短音階が用いられる。長調と短調では各音度上の和音の種類が異なるが、和声的短音階を用いることによって、V度上の3和音とVII度上の3和音が長調と同じ種類になっていること、さらにV度上の7の和音がともに属7となることに注意したい。この属7の和音はVの上には現れず、調を決定づける鍵となる。

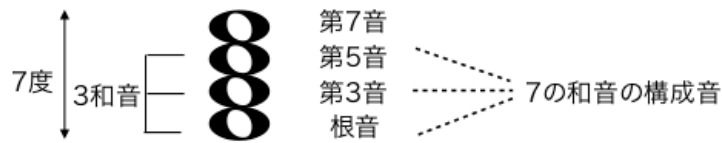


図2.8 7の和音の原型

基礎となる3和音	第5音と第7音のなす音程	名称
長3和音	短3度	属7の和音
長3和音	長3度	長7の和音
短3和音	長3度	名称なし(短3和音+長3度)
短3和音	短3度	短7の和音
減3和音	短3度	減7の和音
減3和音	長3度	導7の和音
増3和音	短3度	増7の和音

表2.3 7の和音の種類と名称

ここで、音階上の和音記号の書き方には決まった書き方があるわけではなく複数の慣例があるが、その一例について述べる。長3和音は根音の音度の大文字によって表し、短3和音は根音の音度の小文字によって表す。なお、このように音階上の和音記号を根音の音度によって表すことは、基本的にその和音が含まれる部分の調が与えられていることを前提とする。なお、音階上にない和音で別の調の音階上に存在する和音は、他の調から和音を借りるという意味で借用和音と呼ばれることがある。同様に長7和音は長3和音の表し方に7を伴い、短3和音は短3和音の表し方に7を伴う。減3和音は音度の小文字に○をつけて表し、減7の和音はこれに7を伴う。増3和音は音度の小文字に+をつけて表し、増7の和音はこれに7を伴う。導7の和音は音度の小文字に◎と7をつけて表す。属7の和音の表し方は長7と同様であるが、実際の音程は異なるので注意したい。

以上和音の種類について述べてきたが、最後に和音の転回型について概要を述べる。図2.7や図2.8のように和音の構成音を根音から順に積み上げたものを和音の基本型という。一方、構成音の順序を入れ替えて用いることもある。3和音において第3音を最も低音に配置した場合は第1転回型、第5音を最も低音に配置した場合は第2転回型と呼ぶ。7の和音においても同様で、第1転回型、第2転回型に加えて第7音を最も低音に配置す

図2.9 音階上の3和音および7の和音

る第3転回型までである。これらの転回形を区別するために複数の流儀による記法があるが、本項ではそれらを区別することが基本的に不要であるため詳細には立ち入らず、転回形は基本型に直して識別するものとする。

2.1.6 和声機能とカデンツ

本節では和声機能とカデンツについて述べる。前節までに述べた事項は定義であり記譜法と不可分であるのに対し、本節で述べる事項は楽曲のスタイルや内容に関わるものであり、必ずしも全ての楽曲に当てはまるものではない。まさにそのため、本研究では後に和声機能の統計的学習を行うのである。しかしここでは、機能と声理論における標準的な考え方や用語について説明する。

2.1.4 節で述べたように調は音階の種類と主音の高さによって決まる。音階の種類が同じであれば、全音と半音の並びは主音が変わっても保たれる(ように調を構成した)。同様に、2.1.5 節で示した音階上の音度と和音の種類の関係も、主音が変わっても保たれる。このような調と和音の体系を基礎とした音楽を一般に調性音楽という。調性音楽は17世紀までに支配的になり、その後の時代を通じて拡張と脱却が図られてきたものの、今日のポピュラー音楽に到るまで広く浸透している。

調の違いは主音の違いでしかないから、調性音楽では主音あるいは主音上の3和音(主和音という)を印象付けることが重要となる。各和音が調性を確立しようとする働きを和音の機能といい、次の3つに分類される。

トニック (Tonic)

主和音に代表される働きで、調の中心となる。

ドミナント (Dominant)

属和音に代表される働きで、主音に強く進行し調性を確立する。

サブドミナント (Subdominant)

属和音の機能を補助したり、主音に穏やかに進行する働きをもつ。

音階上の各音は一般にこれら 3 つの機能に以下のように分類される。以後それぞれの頭文字をとって T, D, S のように表すこともある。なお、ここでは長調短調の区別なく同じ分類が用いられる。

機能	主要 3 和音	副 3 和音
トニック (T)	I	VI
ドミナント (D)	V	VII
サブドミナント (S)	IV	II

表2.4 和音機能と分類

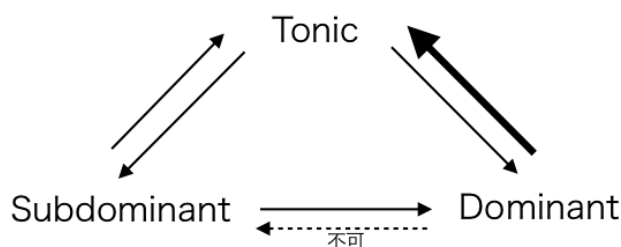


図2.10 和声の機能進行

機能の進行は図 2.10 の原則に従うとされている。トニックは他の 2 つの機能に自由に進行でき、サブドミナントも他の 2 つの機能に進行できるが、ドミナントへ進行する方が多い。一方、ドミナントはトニックにしか進行できない。これが調性を決定づける役割の所以である。なお、ここでは矢印を示していないが同じ機能に留まる進行ももちろん可能である。その際、一般に同じ機能内での主要 3 和音から副 3 和音への進行は可能であるのに対し、副 3 和音から主要 3 和音への進行は用いられない。

トニックから出発してトニックに戻る進行の過程をまとめると次の 3 つの型に分類される。

第1型 $T \rightarrow D \rightarrow T$

第2型 $T \rightarrow S \rightarrow T$

第3型 $T \rightarrow S \rightarrow D \rightarrow T$

これら3つの進行型を一般にカデンツというが、フレーズの終止部分である場合や、明確な低音部の動きを伴う場合にのみカデンツと呼ぶこともある。

2.2 HMM

2.2.1 Hidden Markov Model

マルコフ過程とは、現在の状態が1時刻前の状態のみによって決まるような確率過程のことであり、隠れマルコフモデルはこのような隠れ状態の確率的な状態遷移（この確率を遷移確率という）と、現在の状態からの出力記号列の生起確率を持つモデルである。すなわち、隠れ状態を s 、出力記号を x と表すと、一般に隠れマルコフモデルは図 2.11 のようになる。

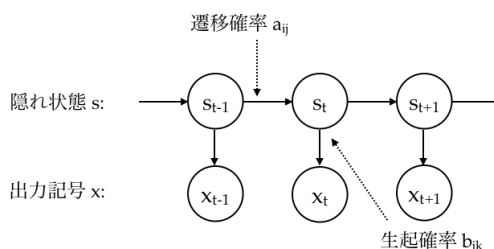


図2.11 HMM

隠れマルコフモデルについて、3つの問題設定ができる。1つ目は、遷移確率、生起確率を既知としたとき、出力記号列としての観測結果 \mathbf{x} が得られる確率を求める問題（評価）である。2つ目は、遷移確率、生起確率が既知としたとき、出力記号列 \mathbf{x} を生じる最も可能性の高い隠れ状態の系列を求める問題（復号）である。3つ目は、出力記号列 \mathbf{x} しか与えられていない場合に、未知パラメータ（遷移確率および生起確率）を求める問題（推定）である。

続いて、2.2.2 節、2.2.3 節、2.2.4 節でこれらの問題の解法の基礎となるアルゴリズムについて説明する、

2.2.2 前向き・後ろ向きアルゴリズム

本節では、2.2.1 節で述べた 1 つ目の問題に対する解法である、前向き・後ろ向きアルゴリズムについて説明する。まず、出力記号列を \mathbf{x} 、隠れ状態の系列を \mathbf{s} としたとき、遷移確率パラメータを a 、生起確率パラメータを b として、同時確率 $P(\mathbf{x}, \mathbf{s})$ は以下で表される。

$$P(\mathbf{x}, \mathbf{s}) = \prod_{t=1}^n a(s_{t-1}, s_t) b(s_t, x_t) \quad (2.1)$$

もしも観測系列 \mathbf{x} に加えて隠れ状態 \mathbf{s} も既知であるならば、(2.2) にしたがってこの \mathbf{x} と \mathbf{s} の同時確率を計算すれば良いので簡単である。一方、 \mathbf{s} が未知であり、出力系列の確率 $P(\mathbf{x})$ を求めたいならば、以下の周辺化が必要である。

$$P(\mathbf{x}) = \sum_{\mathbf{s}} P(\mathbf{x}, \mathbf{s}) \quad (2.2)$$

このような場合、全ての可能な \mathbf{s} についての和を単純な方法で計算しようとする、膨大な計算量となる。そこで、 $P(\mathbf{x})$ の効率的な計算方法として、以下に説明する前向き・後ろ向きアルゴリズムが知られている。

出力記号列 \mathbf{x} に対し、時刻 t の隠れ状態が ω_i である確率 $P(\mathbf{x}, s_t = \omega_i)$ を考えると、

$$P(\mathbf{x}, s_t = \omega_i) = P(x_1 x_2 \cdots x_t, s_t = \omega_i) \cdot P(x_{t+1} x_{t+2} \cdots x_n | s_t = \omega_i) \quad (2.3)$$

$$= \alpha_t(i) \cdot \beta_t(i) \quad (2.4)$$

と分解できる。ここで、 $\alpha_t(i)$ 、 $\beta_t(i)$ を以下のように定義する。

$$\alpha_t(i) \stackrel{\text{def}}{=} P(x_1 x_2 \cdots x_t, s_t = \omega_i) \quad (2.5)$$

$$\beta_t(i) \stackrel{\text{def}}{=} P(x_{t+1} x_{t+2} \cdots x_n | s_t = \omega_i) \quad (2.6)$$

$\alpha_t(i)$ は再帰的に求めることができ、前向きアルゴリズムと呼ばれている。その具体的な手順は以下となる。ここで、初期パラメータを ρ 、遷移確率パラメータを a 、生起確率パラメータを b とする。また、隠れ変数の状態数は c であるとする。

前向きアルゴリズム

step1 初期化

$$\alpha_1(i) = \rho_i b(\omega_i, x_1) \quad (i = 1, 2, \dots, c)$$

step2 再帰計算

$$a_t(j) = \left[\sum_{i=1}^c \alpha_{t-1}(i) a_{ij} \right] b(\omega_j, x_t)$$

step3 確率計算

$$P(\mathbf{x}) = \sum_{i=1}^c \alpha_n(i)$$

ここで、Step2 を見るとわかるように、 $\alpha_t(i)$ は 1 時刻前の値 $\alpha_{t-1}(i)$ を用いて、つまり観測の時系列と同方向に再帰的に求められることから、前向きアルゴリズムと呼ばれる。

一方、 $\beta_t(i)$ も同様に再帰的な方法で求めることができる。ただし前向きアルゴリズムとは逆に観測時系列を遡る方向に再帰的に計算することから、後ろ向きアルゴリズムと呼ばれる。以下がその手順となる。

後向きアルゴリズム

step1 初期化

$$\beta_n(i) = 1 \quad (i = 1, 2, \dots, c)$$

step2 再帰計算

$$\beta_t(i) = \sum_{j=1}^c a_{ij} b(\omega_j, x_{t+1}) \beta_{t+1}(j) \\ (t = 1, 2, \dots, (n-1)) \quad (i = 1, 2, \dots, c)$$

2.2.3 Viterbi アルゴリズム

本節では2.2.1 節で述べた 2 つ目の問題に対する解法である、Viterbi アルゴリズムについて説明する。2 つ目の問題は、遷移確率パラメータ、生起確率パラメータとともに観測系列 \mathbf{x} が得られたときに最も可能性の高い隠れ状態の系列を求めるというものであった。可能な全ての $P(\mathbf{x}, \mathbf{s})$ を求め、その中から $P(\mathbf{x}, \mathbf{s})$ が最大になる \mathbf{s} を得られれば問題は解けるが、2.2.2 節で述べたのと同様に、このような方法は計算量が膨大になり効率が悪い。そこで、このような問題に対する効率的な解法として、動的計画法に基づく Viterbi アルゴリズムが知られている。

これまで述べたことから推察されるように、Viterbi アルゴリズムは前向きアル

ゴリズムに類似している。時刻 t の隠れ状態が ω_j である確率の最大値 $\psi_t(j) = \max_{s_1, \dots, s_{t-1}} P(x_1 \cdots x_t, s_1 \cdots s_{t-1}, s_t = \omega_j)$ を考えると、 $\psi_t(j)$ は再帰的に計算できる。また、(2.7)のように最適な状態系列を復元するための記憶領域 $\Psi_t(j)$ を定義する。

$$\Psi_t(j) = k = \arg \max_i \{\psi_{t-1}(i) a_{ij}\} \quad (2.7)$$

以上から、Viterbi アルゴリズムの手順は以下となる。

Viterbi アルゴリズム

Step1 初期化

$$\psi_1(i) = \rho_i b(\omega_i, x_1)$$

$$\Psi_1(i) = 0$$

$$(i = 1, 2, \dots, c)$$

Step2 再帰計算

$$\psi_t(j) = \max_i \{\psi_{t-1}(i) a_{ij}\} b(\omega_j, x_t)$$

$$\Psi_t(j) = \arg \max_i \{\psi_{t-1}(i) a_{ij}\}$$

$$(t = 2, 3, \dots, n)(i, j = 1, 2, \dots, c)$$

2.2.4 Baum-Welch アルゴリズム

本節では2.2.1節で述べた3つ目の問題に対する解法である、Baum-Welch アルゴリズムについて説明する。すなわち、観測結果の系列すなわち出力記号列から HMM の隠れ状態の遷移確率パラメータと隠れ状態からの出力記号の生起確率パラメータを教師なし学習するアルゴリズムである。 θ を推定すべきパラメータの集合として、対数尤度関数 $J(\theta) = \ln \Sigma(\mathbf{x}, \mathbf{s}; \theta)$ を最大にする θ を求める最尤推定問題を考える。この関数は一般に非線形で解析解を求められないため EM アルゴリズムによる逐次解法を用いる。EM アルゴリズムは、 $J(\theta)$ のような log-sum 型の関数に対し sum-log 型の下限値を導出しその最大化を行う手法であり、目的関数は $Q(\theta, \theta^{prev}) = \Sigma P(\mathbf{s} | \mathbf{x}; \theta^{prev}) \ln P(\mathbf{x}, \mathbf{s}; \theta)$ となる。推定すべきパラメータ (遷移確率 $\mathbf{A} = \{a_{ij}\}$, 出現確率 $\mathbf{B} = \{b_{jk}\}$, 初期確率 $\boldsymbol{\rho} = \{\rho_i\}$)

は,

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{prev}) = \sum_{\mathbf{s}} P(\mathbf{s}|\mathbf{x}; \boldsymbol{\theta}^{prev}) \ln P(s_1) \quad (2.8)$$

$$+ \sum_{\mathbf{s}} P(\mathbf{s}|\mathbf{x}; \boldsymbol{\theta}^{prev}) \sum_{t=1}^{n-1} \ln a(s_t, s_{t+1})$$

$$+ \sum_{\mathbf{s}} P(\mathbf{s}|\mathbf{x}; \boldsymbol{\theta}^{prev}) \sum_{t=1}^n \ln b(s_t, x_t) \quad (2.9)$$

$$= Q(\mathbf{A}, \boldsymbol{\theta}^{prev}) + Q(\mathbf{B}, \boldsymbol{\theta}^{prev}) + Q(\boldsymbol{\rho}, \boldsymbol{\theta}^{prev}) \quad (2.10)$$

のように分解され個別に最大化できる。各ステップで \mathbf{A} , \mathbf{B} , $\boldsymbol{\rho}$ は以前のステップで求めたパラメータ $\boldsymbol{\theta}^{prev}$ から算出可能であり, その結果を次のステップの計算に用いることで逐次最大化を行う。

ここで, $Q(\mathbf{A}, \boldsymbol{\theta}^{prev})$, $Q(\mathbf{B}, \boldsymbol{\theta}^{prev})$, $Q(\boldsymbol{\rho}, \boldsymbol{\theta}^{prev})$ を最大化するパラメータを求めるための準備として, 2.2.2 節で導入した $\alpha_t(i)$, $\beta_t(i)$ を再び用いる。これらに加え, 以下の $\xi_t(i, j)$ を導入する。これは, 出力記号列 \mathbf{x} の条件のもとで, 時刻 t の隠れ状態が ω_i かつ時刻 $t+1$ の隠れ状態が ω_j である確率を表す。

$$\xi_t(i, j) \stackrel{\text{def}}{=} P(s_t = \omega_i, s_{t+1} = \omega_j | \mathbf{x}) \quad (2.11)$$

一方, 出力記号列 \mathbf{x} と, 時刻 t の隠れ状態が ω_i かつ時刻 $t+1$ の隠れ状態が ω_j である同時確率は,

$$P(s_t = \omega_i, s_{t+1} = \omega_j, \mathbf{x}) = P(x_1 \cdots x_t, s_t = \omega_i) \cdot P(s_{t+1} = \omega_j | s_t = \omega_i) \quad (2.12)$$

$$\cdot P(x_{t+1} | s_{t+1} = \omega_j) \cdot P(x_{t+2} \cdots x_n | s_{t+1} = \omega_j)$$

$$= \alpha_t(i) a_{ij} b(\omega_j, x_{t+1}) \beta_{t+1}(j) \quad (2.13)$$

であるから, $P(s_t = \omega_i, s_{t+1} = \omega_j | \mathbf{x}) P(\mathbf{x}) = P(s_t = \omega_i, s_{t+1} = \omega_j, \mathbf{x})$ の関係より,

$$\xi_t(i, j) = \frac{\alpha_t(i) a_{ij} b(\omega_j, x_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^c \alpha_n(i)} \quad (2.14)$$

さらに, 時刻 t で状態 ω_i にいる確率 $\gamma_t(i)$ を導入すると, 明らかに,

$$\gamma_t(i) = \sum_{j=1}^c \xi_t(i, j) \quad (2.15)$$

である。

これらを用いて, $Q(\mathbf{A}, \boldsymbol{\theta}^{prev})$ を条件 $\sum_{j=1}^c a_{ij} = 1$ のもとで最大にする a_{ij} , $Q(\mathbf{B}, \boldsymbol{\theta}^{prev})$ を $\sum_{k=1}^m b_{jk} = 1$ のもとで最大にする b_{jk} , $Q(\boldsymbol{\rho}, \boldsymbol{\theta}^{prev})$ を $\sum_{i=1}^c \rho_i = 1$ のもとで最大にする ρ_i は, それぞれ以下となる.

$$a_{ij} = \frac{\sum_{t=1}^{n-1} \xi_t(i, j)}{\sum_{t=1}^{n-1} \gamma_t(i)} \quad (2.16)$$

$$= \frac{\sum_{t=1}^{n-1} \alpha_t(i) a_{ij}^{prev} b_{jk}^{prev}(\omega_j, x_{t+1}) \beta_{t+1}(j)}{\sum_{t=1}^{n-1} \alpha_t(i) \beta_t(i)} \quad (2.17)$$

$$b_{jk} = \frac{\sum_{t=1}^n \delta(x_t, v_k) \gamma_t(j)}{\sum_{t=1}^n \gamma_t(j)} \quad (2.18)$$

$$= \frac{\sum_{t=1}^n \delta(x_t, v_k) \alpha_t(j) \beta_t(j)}{\sum_{t=1}^n \alpha_t(j) \beta_t(j)} \quad (2.19)$$

$$\rho_i = \frac{\gamma_1(i)}{\sum_{i=1}^c \gamma_1(i)} \quad (2.20)$$

$$= \frac{\alpha_1(i) \beta_1(i)}{\sum_{j=1}^c \alpha_n(j)} \quad (2.21)$$

以上により, Baum-Welch アルゴリズムによりパラメータを求める手順は以下となる.

Baum-Welch アルゴリズム

step1 初期化

a_{ij}, b_{jk}, ρ_i の初期化

step2 再帰計算

a_{ij}, b_{jk}, ρ_i の値をそれぞれ(2.17), (2.19), (2.20)にしたがって求める. ここで, $\alpha_t(i), \beta_t(i)$ はそれぞれ前向きアルゴリズム, 後ろ向きアルゴリズムにより求める.

step3 更新

Step2 で得た値でパラメータを $a_{ij}^{prev} = a_{ij}, b_{jk}^{prev} = b_{jk}, \rho_i^{prev} = \rho_i$ により更新する.

step4 判定

対数尤度 $\ln P(\mathbf{x}|\boldsymbol{\theta}) = \ln \sum_{i=1}^c \alpha_n(i)$ が更新前と比べ予め定めた閾値より増分が小さければ収束とみなし, そうでなければ Step2 に戻る.

Baum-Welch アルゴリズムは HMM の教師なし学習の基本的解法であるが, ここで用

いられる EM アルゴリズムは初期値依存性が高いことが知られている。局所解の問題を回避する方法として、ギブスサンプリング^{*1}を用いた BayesianHMM^{*2}がある。HMM は自然言語処理の形態素解析の分野でしばしば用いられており、この分野では EM アルゴリズムが初期値の影響で良い解をもたらさなかった一方 [11]、ギブスサンプリングによる BayesianHMM が局所解の問題を回避し、EM アルゴリズムに比べて性能の良い解を達成したことが報告されている [12]。ただし、BayesianHMM では事前分布を与える必要があること、局所最適解は保証されないことに注意する必要がある。

2.3 集合分割問題

2.3.1 数理最適化の基礎

数理最適化とは、与えられた制約条件をすべて満足し、目的関数 $f(x)$ の値を最大または最小にする決定変数 x の値を見つける問題である。このような x を最適解と呼び、最適解は唯一の場合もあれば複数または無限に存在する場合もある。このうち整数計画問題は、全ての変数が整数であるものを指す。次節で述べる集合分割問題は整数計画問題の一種である。

2.3.2 集合分割問題

集合分割問題は、複数の要素とその要素集合族および各集合のスコアまたはコストが与えられた時に、全ての要素が集合のどれかに1回だけ含まれるという条件で、選ばれた集合のコストの総和を最小またはスコアの総和を最大にする問題である。

ここで要素の集合を $M = \{1, 2, \dots, m\}$ 、部分集合の集合を $N = \{1, 2, \dots, n\}$ とする。また、 $i \in S_j$ ならば $a_{ij} = 1$ 、そうでなければ $a_{ij} = 0$ となる a_{ij} を要素にもつ行列 $A = (a_{ij})$ を導入する。さらに、 $\mathbf{x}_j = (x_1, \dots, x_n)^T$ を $j \in X$ ならば $x_j = 1$ 、そうでなければ $x_j = 0$ となる決定変数とすると、集合分割問題は0-1整数計画問題として定式化できる。

*1 直接サンプルを得ることが難しい複雑な関数についてその関数を再現するようにサンプリングを行うマルコフチェーンモンテカルロ法の一つ

*2 パラメータの事前確率を導入した HMM

集合分割問題

目的変数: $\sum_{j=1}^n c_j x_j \rightarrow$ 最小 (c_j がコストの場合), 最大 (c_j がスコアの場合)

制約条件: $\sum_{j=1}^n a_{ij} x_j = 1, i \in M, x_j \in \{0, 1\}, j \in N$

第 3 章

先行研究

3.1 和声機能学習の先行研究

本節では和声機能学習の先行研究について述べる。2.1.6 節で述べたように、和声機能は定義ではなく楽曲のスタイルや内容に関わるものである。和声機能の統計的学習には大きく分けてクラスタリングと生成モデルの 2 種類がある。

3.1.1 クラスタリングを用いた和声機能の教師なし分類

Rohrmeier ら [14] は J.S.Bach の 4 声コーラルを対象とし、長調と短調を別個に、ピッチクラスセットの進行（連続する 2 つのピッチクラスの関係）に対する階層クラスタ分析を行った。その結果、長調短調ともに最も顕著なクラスタはドミナント進行を示唆することが報告された。一方短調では、平行長調に相当するクラスタが得られることも報告された。

Jacoby ら [15] にもクラスタリングを用いた和声機能の分類を行ったが、情報理論に基づき、クラスタリングを行うだけでなくモデルを正確さと複雑さの 2 点から評価することを提案している。教師なし分類を行う前に、彼らは既存の音楽理論をベースとしたいくつかのモデルにこの評価手法を適用した。モデルの複雑さが同程度であれば正確さがより高い方がよいという観点から最適曲線を描くと、表層記号を転回形を無視した 7 つの音階構成和音とした場合、和声機能 TDS (Tonic, Dominant, Sub-dominant) による分類は最適曲線上に位置し、一方で 3 和音の種類（長 3 和音、短 3 和音、減 3 和音）による分類は最適曲線をかなり下回ることが分かった。また、教師なし分類においても、カテゴリ数 3 の分類結果は TDS による分類に極めて近いことが報告された。

3.1.2 HMM を用いた和声機能の教師なし学習

前節で述べた2つの研究は、Jacoby らの論文 [15] で言及されているように、和声機能の分類を目的としたものであって生成モデルではない点が本研究とは基本的に異なる。一方、Tsushima らの研究 [9] では生成モデルの観点から和声機能の教師なし学習を行っている。生成モデルとして HMM および PCFG を用い、学習の結果、コードネームの与えられたポピュラーおよびジャズの楽曲について機能と声理論の知見が得られることを報告している。

2.2.4 節で紹介したように、HMM を用いた形態素解析の分野では、局所解の問題を改善するために BayesianHMM を用いた方法が EM アルゴリズムによる方法よりもよい精度を達成したことが報告されている。そこで、Tsushima らはガウシアンサンプリングを用いた BayesianHMM による方法と EM アルゴリズムの双方で学習を行いその結果を比較したが、ここでは、BayesianHMM と EM アルゴリズムで学習結果の性能に顕著な差異はなかったことが報告されている。

また、彼らは HMM だけでなく PCFG モデルについても実験を行っている。生成モデルとして和声機能学習を行うことは、メロディへの和声付けといった幅広い応用を考慮してのことであり、そのような観点から PCFG モデルは大局的な構造を考慮できない HMM よりも優れていると考えられるためである。しかしながら和声機能の学習に関しては、PCFG はテストデータのパープレキシティおよびエラー率の観点で、最高の性能を達成した設定では HMM を上回ったものの、多くの場合 HMM と同程度の性能であり、より悪くなる場合もあったと報告されている。また PCFG のパラメータの初期化については、HMM で学習したパラメータによる初期化に顕著な効果があることがわかった。一方、学習した HMM および PCFG モデルによるメロディへの和声付けでは、PCFG が HMM を上回る性能を発揮したと報告されている [10]。

3.2 調判定の先行研究

3.2.1 ヒストグラムベースの調判定アルゴリズム

調判定アルゴリズムの中で最も典型的な手法は、ピッチクラスヒストグラムに基づくものである。ヒストグラムベースの調判定は、楽曲からピッチクラスヒストグラムを取得し、この情報を元に調の判定を行う。ヒストグラムベースの調判定は古くから行われており、古典的で最もよく知られたものに Krumhansl-Schmuckler アルゴリズム [6] がある。

Krumhansl-Schmuckler アルゴリズムでは心理実験の結果により設計された Key-profile と楽曲から得られたヒストグラムの相関係数により調を判定する。その実験とは、被験者にいくつかの和音列を聞かせ、試験音があらかじめ聞かせた音列とどれだけ調和しているかを答えさせるというものである。こうして得られた長調の Key-profile を表 3.1 に、短調の Key-profile を表 3.2 に示す。表においては C major および a minor に代表させているが、この値を各調の主音に基づき移調して用いる。

C	C \sharp	D	D \sharp	E	F	F \sharp	G	G \sharp	A	A \sharp	B
6.35	2.33	3.48	2.33	4.38	4.09	2.52	5.19	2.39	3.66	2.29	2.88

表3.1 Krumhansl-Schmuckler アルゴリズム Major Key-profile

A	A \sharp	B	C	C \sharp	D	D \sharp	E	F	F \sharp	G	G \sharp
6.33	2.68	3.52	5.38	2.6	3.53	2.54	4.75	3.98	2.69	3.34	3.17

表3.2 Krumhansl-Schmuckler アルゴリズム Minor Key-profile

この方法では、Key-profile があらかじめ設計されており異なるスタイルの楽曲に適応しづらい課題がある。そこで、Hu ら [13] は Latent Dirichlet Allocation(LDA) に基づく Key-profile の学習を提案している。しかしながら、これらヒストグラムベースの手法ではあらかじめ分析対象範囲を決める必要があるため、局所的な転調をピンポイントで捉えることはできない。

3.2.2 和音間距離の最短経路探索に基づく調判定アルゴリズム

Sakamoto ら [7] は予めコードが与えられている場合について、F. Lerdhahl の構成した Tonal Pitch Space (TPS)[8] に基づく和音間距離をコストとした最短経路探索による調判定を提案した。この方法は、音高ではなく和音を単位としている点、局所的な転調判定が可能である点において、ヒストグラムベースの方法とは異なるものである。

TPS は和音間距離を定量化するための理論である。TPS ではまず、以下のように音階内の各音の重要度を 5 段階のレベルに分ける。レベルの深さが小さいほど重要度の高い音である。

Level a = 0 C (主音) \dots octave level

- Level b = 1 G (属音) … fifth level
- Level c = 2 E (主和音構成音) … triadic level
- Level d = 3 全音階構成音 … diatonic level
- Level e = 4 半音階構成音 … chromatic level

表 3.3 に各音のレベルの深さを示す.

(a)	C											
(b)	C						G					
(c)	C			E			G					
(d)	C		D		E	F		G		A		B
(e)	C	C \sharp /D \flat	D	D \sharp /E \flat	E	F	F \sharp /G \flat	G	G \sharp /A \flat	A	A \sharp /B \flat	B
p/c	0	1	2	3	4	5	6	7	8	9	10	11

表3.3 ピッチクラスのレベルの深さ

一方, 主音を別の音に変えることは, 表 3.3 に示した深さの形状をシフトさせることに対応するが, シフト後と元の形状を比較した時, 完全 5 度上への移動すなわち主和音から属和音への移動が最も形状の差異が少なくなる. そこで, 和音間の距離を 5 度の移動を何回繰り返したか + シフト前後の形状の差異 (シフト前に比べてレベルの深さが減ったものの数) と定義する.

以上は同じ調の内部における和音間距離であるが, 調間にまたがる和音の移動については調号の数の変化を距離に加算する. 調号の数の変化は, 2.1.4 節で述べたように調号 1 つの増減が主音の 5 度の移動に対応している. さらに, 調号の差異が大きい場合に対しては, ここでは詳細を述べないが, 主和音同士の距離を手がかりに調間距離を定義する.

このようにして定義された TPS の和音間距離をもとに, Sakamoto らは与えられたコードネーム列に対し, 調 \times コードネームの組み合わせにより候補ノードを生成し, コードネーム列全体の和音間距離の和が最も少なくなる調の系列を発見する最短経路探索問題として, Viterbi アルゴリズムによる解析方法を提案した. したがってこの手法においては, 調 \times コードネームの組み合わせ数が状態の数となり, 状態数が多い分計算量が多い. また, 本質的に和音間距離があらかじめ設計されたものであるため, 5 度上での移動の重要性など音楽学の知識を積極的に取り入れることができるメリットがある反面, 知識が固定されていることにより楽曲や時代による差異への対応を行うには, それらに適した新たな事前知識とそれによる和音間距離の定義変更を必要とする.

第 4 章

転調判定を用いた和声進行生成モデルのモデルサイズ選択

4.1 概要

本研究では Hidden Markov Model (HMM) により和声進行をモデリングし、その学習を行うとともに、得られた HMM パラメータと集合分割問題を用いた転調判定システムを構築して実行し、その結果得られる対数尤度の合計によって HMM の最適な隠れ状態数を決定する。本研究は以下の 3 つのステップからなる。

1. HMM による和声進行のモデリングと統計的学習
2. 集合分割問題を用いた転調判定
3. 転調判定の結果得られる対数尤度の合計を異なる隠れ状態数で比較し、最も良い隠れ状態数を選択

HMM と集合分割問題の基礎理論については [2.2.1 節](#) および [2.3 節](#) で述べたが、以下の節ではこれらの理論を応用した本研究の具体的な手法について説明する。

4.2 和声機能の教師なし学習

4.2.1 HMM による和声進行の生成モデル

本研究では、[図 4.1](#) に示すように和声機能 (Tonic, Dominant 等) を期待した隠れ状態 s_t から個別の和音 x_t が確率的に出力され観測されるという HMM を用いて、観測結果で

ある和音列から和声機能の遷移確率パラメータおよび和声機能からの和音の出力確率パラメータを Baum-Welch アルゴリズムで推定する．ここで「和声機能を期待した隠れ状態」

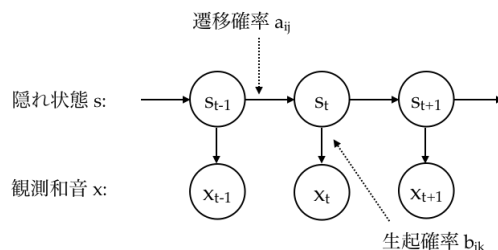


図4.1 HMM による和声進行モデル

と述べたのは，本モデルが機能と声モデルをよく近似し，学習される隠れ状態のパラメータが和声機能を表すことを期待するためである．しかしながら，この隠れ状態は事前に何ら音楽的知識を仮定するものではない．

隠れ状態数は事前に指定する必要があるため，本研究では隠れ状態数 2–12 の場合について実験を行なう．すなわち，例えば隠れ状態数が 4 の場合は，隠れ状態 $s_t \in \{\omega_1, \omega_2, \omega_3, \omega_4\}$ である．一方，観測 x については，本研究では調判定への応用をふまえ Diatonic Scale 上の 3 和音を構成要素とする．すなわち長調においては表 4.1 に示すように構成和音は 7 つであるが，短調においては 2.1.1 節で述べたように第 VII 音を半音上げる場合とそうでない場合の両方が用いられるため，その両方によって生じうる和音を網羅し表 4.2 に示す 10 和音を構成要素とした．

k	1	2	3	4	5	6	7
name	C major	D minor	E minor	F major	G major	A minor	B diminished
symbol	I	ii^o	iii	IV	V	vi	vii^o

表4.1 長調の構成和音

k	1	2	3	4	5	6	7	8	9	10
name	A minor	B diminished	C major	C augmented	D minor	E minor	E major	F major	G major	G [#] diminished
symbol	i	ii^o	III	III⁺	iv	v	V	VI	VII	vii^o

表4.2 短調の構成和音

4.2.2 初期値の設定

2.2.4 節で触れたように EM アルゴリズムの解は局所最適解であり，初期値の影響を受ける．そのため Tsushima らの先行研究 [9] では BayesianHMM により局所最適解を避けるアプローチと EM アルゴリズムの両方を実験しているが，得られた解の性能に顕著な差異はなかったと報告している．これを念頭に，本研究では EM アルゴリズム（Baum-Welch アルゴリズム）を用いて N_{init} 個の初期値から解を生成し，その中から後に述べる方法により最適な局所解を選択する．ただし和声機能の遷移確率はランダムに初期化し，和声機能からの和音の出現確率は一律に初期化するものとする．

4.2.3 局所解の評価

学習を行い得られた N_{init} 組のパラメータについて，同じ隠れ状態数の異なる局所解に対しては，Tsushima らの先行研究 [9] にならい，その性能を系列のパープレキシティによって評価する．系列のパープレキシティ \mathcal{P} は系列の対数尤度 $\ln P(\mathbf{x}|\boldsymbol{\theta})$ から(4.1)に従って計算される．したがって，対数尤度が大きいほどパープレキシティが小さいという関係性がある．ここで， \mathbf{x} は対象の系列， $\boldsymbol{\theta}$ は学習されたパラメータである．

$$\mathcal{P} = \exp\left(-\frac{1}{|\mathbf{x}|} \ln P(\mathbf{x}|\boldsymbol{\theta})\right) \quad (4.1)$$

コーパスサイズが小さいため，データを学習用と評価用に分けることができないが，学習時にはランダムな曲順で全曲を 1 つの系列に連結して用いるのに対し，評価時は 1 曲毎にパープレキシティを計算しその平均をもって評価した．

4.3 転調解析

4.3.1 HMM を利用した調判定アルゴリズム

音楽学において機能และ声解析とは，和音の進行を調性や和音の機能に着目して分析する方法のことである．HMM による機能และ声モデルの模倣と学習を行なうが，これによりある調（例えば G major）を仮定した時に，ある和声進行についてもっとも対数尤度の高い機能の系列を具体的に計算できる．逆に調性を判定するには対数尤度の最も高くなるモデルを選択すれば良い．すなわち，具体的には以下のステップで調判定を行う．

1. 調判定を行いたい楽曲の調は未知であるので，仮の調（例えば G major）を仮定し

た上で C major に転調する (G から C の差分だけ楽曲全体のピッチを移行する)。この処理を全ての調に対して行う。ただし、長調と短調ではモデルが別個であるため、長調同士、短調同士について移調を行う。

2. 1. で生成した調の候補全てに対し、HMM により系列の対数尤度を計算し、最も対数尤度の高い調をその曲の調であると判定する。例えば、G major を仮定して C major に移調した系列の対数尤度が最も高かった場合、その曲は G major であると判定する。

ここで、特に本来の調と異なる調を仮定して C major への移調を行うと、構成和音に含まれない和音が多く含まれる。したがって、4.2.1 節で述べたように観測 x の集合は Diatonic Scale 上の和音としているが、これに含まれない観測が生じることとなる。構成和音に含まれない和音は除いて対数尤度を算出するが、3 和音系 (Major, Minor, Diminish, Augmented) の非構成音に関しては出現回数分ペナルティを加算する。その理由は、これらの和音が他の調では構成和音となりうるためである。なお、3 和音上の 7 の和音もこれに含まれる。一方、3 和音系以外の非構成和音に関しては、同じ調の影響下で経過的に生じる和音である場合が多く、ペナルティは加算しない。

4.3.2 集合分割問題による転調判定

和音系列をブロック分割し、各ブロックで対数尤度最大となる調をそのブロックの調とする仮定のもと、各ブロックの対数尤度の和が最大になるようなブロックの分割方法を求める問題として、転調判定をモデル化する。これは部分集合の候補が部分系列 (ブロック) であるような、集合分割問題の特別な形とみなせる。つまり、観測系列 $x_1x_2x_3x_4x_5\dots x_n$ としたとき、部分集合の候補は(4.2)となる。ここでブロックの最小の長さは 4 とする。

$$\begin{aligned}
 & \{x_1, x_2, x_3, x_4\}, \{x_2, x_3, x_4, x_5\}, \dots, \{x_{n-3}, x_{n-2}, x_{n-1}, x_n\}, \\
 & \{x_1, x_2, x_3, x_4, x_5\}, \{x_2, x_3, x_4, x_5, x_6\}, \dots, \{x_{n-4}, x_{n-3}, x_{n-2}, x_{n-1}, x_n\}, \\
 & \dots \\
 & \{x_1, x_2, \dots, x_n\}
 \end{aligned} \tag{4.2}$$

これらの中で、さらに各出力記号が 1 つのブロックにのみ属する組み合わせが転調判定におけるブロック分割の候補となる。つまりここでは、ある出力記号すなわち和音が複数のブロックに重複して含まれることを許可しない。これは各出力記号について、唯一の調が割り当てられることを意味する。したがって、ブロック分割の候補は例えば以下のような

ものである。

$$\{x_1, x_2, x_3, x_4\}, \{x_4, x_6, x_7, x_8, x_9, x_{10}, x_{11}\}, \{x_{12}, x_{13}, x_{14}, \dots, x_n\}$$

ここで、各ブロックの出力記号列（すなわち上の例では $x_4x_5x_6x_7x_8x_9x_{10}x_{11}$ など）について HMM を用いて対数尤度を求める。このとき、この出力記号列を 4.3.1 節で述べた方法で転調を行い、最も対数尤度が高い調をこのブロックの調として選択し、同時にその対数尤度を得る。こうして各ブロックの対数尤度と調が得られた上で、集合分割問題は 2.3.2 節に述べた整数計画問題として、ブロックの対数尤度の和が最も大きくなるようなブロック分割を得ることに相当する。

4.4 HMM 隠れ状態数の選択

こうして 4.2 節に示したように個別の隠れ状態数における和声進行モデルのパラメータ学習とパープレキシティ評価による最良の局所解選択が可能になり、4.3 節に示した方法で学習した HMM を用いた転調判定が可能となった。そこで、各隠れ状態において最小のパープレキシティとなったパラメータを用い、学習データ全体に対し転調判定を行なった結果の対数尤度の和を計算する。すなわち、対数尤度の和が最も小さくなる隠れ状態数を最良のモデルとして選択する。

第 5 章

実験と結果

5.1 データセット

データとして J.S.Bach の作品番号 BWV253-438 に含まれる 4 声コラールのうち，長調の楽曲 94 曲，短調の楽曲 68 曲を用いる．ただし，和声機能パラメータ学習では以下に述べる理由により，長調の楽曲 83 曲，短調の楽曲 47 曲を用いた．

和声機能学習では，曲中には一般に転調が含まれるため，一般に調号にしたがった調性を仮定できる冒頭フレーズと終端フレーズのみを用い，長調および短調のグループごとに全ての曲についてこれを連結して和音系列を作る．フレーズは，コラールにおいてフレーズの区切りを示すフェルマータと呼ばれる記号が付与されていることから，これを用いてフレーズ区切りを得る．曲順による学習結果への影響を考慮し，系列のつなぎ方はランダムな曲順で複数作成し，学習データとして用いる．

コラールは聖歌を旋律に和声付けをした楽曲であり，古いものでは中世の旋律も含まれるため，一部に古い時代の調性システムの影響を受けた調号の指定が見られる．そこで，学習データにおいては開始和音と終了和音が，長調ではともに C major，短調ではともに A minor であるものに限定した．さらに，コラールにおいてはピカルディ終止と呼ばれる特殊な終止が用いられることが普通である．これは短調の楽曲において終結部分の主和音が本来短 3 和音であるところ長 3 和音に転じるものである．和声機能の学習データにおいてはこのピカルディ終止を短 3 和音に置き換える処理を行った．短調の学習データはこの置き換えを考慮した上で開始および終結がともに A minor であるものを全て用いた．

和声機能の学習データにおける出力記号の構成比率は，長調が表 5.1，短調が表 5.2 であった．短調では導音の導入に伴う全ての組み合わせを出力記号として用いたにも関わらず，長調よりも多くの例外和音があることがわかる．また，短調では平行調の主和音で

ある C major の出現比率が高いことも注目すべきである。開始和音と終了和音がともに A minor であるものに学習データを限ったにも関わらず、フレーズ内での平行調への遷移が含まれていることが示唆される。

順位	和音名	比率
1	C major(I)	0.3050
2	G major(V)	0.1956
3	D minor(ii)	0.1235
4	A minor(vi)	0.1076
5	F major(IV)	0.0957
6	B diminished(vii^o)	0.0544
7	E minor(iii)	0.0437
-	Others	0.0745

表5.1 出力記号の構成比率 (長調)

順位	和音名	比率
1	A minor(i)	0.2859
2	E major(V)	0.1494
3	C major(III)	0.0791
4	B diminished(ii^o)	0.0722
5	D minor(IV)	0.0623
6	G major(VII)	0.0613
7	G \sharp diminished(vii^o)	0.0524
8	F major(VI)	0.0514
9	E minor(v)	0.0346
10	C augmented(III⁺)	0.0208
-	Others	0.1306

表5.2 出力記号の構成比率 (短調)

5.2 和声機能の教師なし学習

5.2.1 初期値の設定と局所解の評価

HMM の隠れ状態数 2–12 の場合について、各隠れ状態数に対し 4.2.2 節で述べた方法に従い $N^{init} = 1000$ すなわち 1000 の異なる初期値から Baum-Welch アルゴリズムにより最適解を求めた。得られた N^{init} 個のパラメータに対し、4.2.3 節に述べたように系列のパープレキシティを計算し、パープレキシティが小さい順に順位付けを行なった。

さらに得られた局所解のばらつきを調べるために、パープレキシティにより順位付けされた上位 20 件のパラメータに対し、各隠れ状態に対する出力記号の生起確率を 1 つのサンプルベクトルとみなし、K-Means クラスタリングを行った。セントロイドは、パープレキシティが最も小さい局所解の生起確率パラメータ $\mathbf{B}^1 = \mathbf{b}_{jk}^1$ を用い、以下の $\mathbf{b}_0^c, \dots, \mathbf{b}_j^c$ のように隠れ状態の数だけ設定した。ここで j は隠れ状態番号、 k は出力記号番号を表す。

$$\begin{aligned} \mathbf{b}_0^c &= (b_{01}^1, b_{02}^1, \dots, b_{0k}^1) \\ \mathbf{b}_1^c &= (b_{11}^1, b_{12}^1, \dots, b_{1k}^1) \\ &\dots \\ \mathbf{b}_j^c &= (b_{j1}^1, b_{j2}^1, \dots, b_{jk}^1) \end{aligned} \tag{5.1}$$

そして他の局所解も同様に、各隠れ状態に対応する生起確率行列の行をベクトルとみなして、先に定義したセントロイドに対してクラスタリングを行った。こうしてパープレキシティを用いて順位付けした局所解のうち上位 20 件に対しクラスタリングを行い、セントロイドに対する二乗平均距離を算出した。遷移確率行列に対しても上に述べたのと同様の処理を行った。

得られた結果は、以下の傾向を示した。

- 隠れ状態数が多いほどパープレキシティは小さい (図 5.1)
- 隠れ状態数が多いほど得られたパラメータのばらつきは大きい (図 5.2)

すなわち隠れ状態数が大きすぎると安定して同じ局所解が得られなくなり、パープレキシティのみによって隠れ状態数を多いほど良いとすることに問題があることが示唆された。このように全体としては隠れ状態数が大きいと平方距離の平均は増加し、上位 20 件のパラメータの類似度が小さくなっているが、平方距離平均の増加は単調ではなく、例えば長調の隠れ状態数 6 は比較的小さい値を示している。

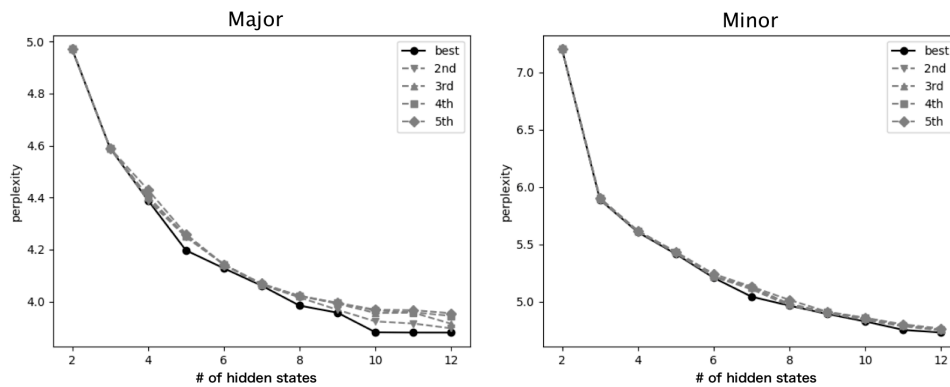


図5.1 隠れ状態数とパープレキシティ

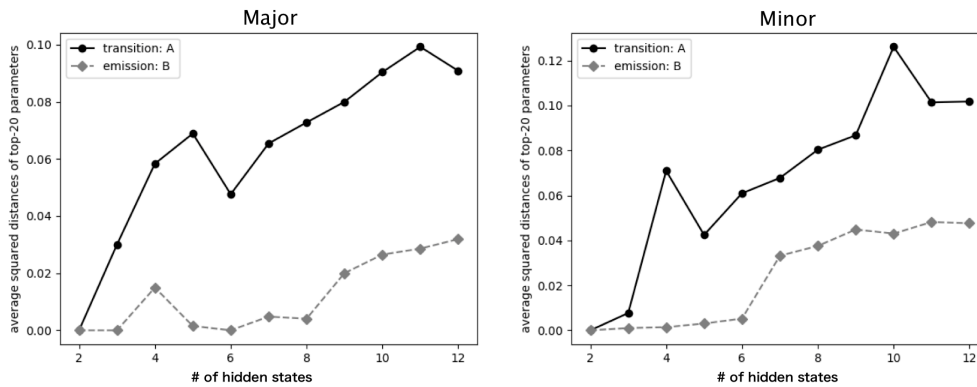


図5.2 隠れ状態数とパラメータの類似度

5.2.2 隠れ状態数 3 における機能と声理論との類似性

ところで、機能と声理論においては、2.1.6 節で述べたように和音の機能をトニック (T)、ドミナント (D)、サブドミナント (S) の 3 つに分類している。長調の隠れ状態数 3 の結果のうち最もパープレキシティの小さいものは、図 5.3 のようになった。各隠れ状態の出力記号出力確率を見ると、隠れ状態 0 では **I**、隠れ状態 1 では **V**、隠れ状態 2 では **ii** が最も大きくなっている。従って、隠れ状態 0 はトニック、隠れ状態 1 はドミナント、隠れ状態 2 はサブドミナントに近いものになっている。隠れ状態の遷移確率においても、ドミナントがトニックに進行しやすい点、サブドミナントがドミナントに進行しやすい点など機能と声理論との類似点が見られる。隠れ状態の遷移確率において同じ状態間への遷移がきわめて少ないのは、前処理において同じ出力記号が連続する場合は一体のものとして扱ったことが影響していると考えられる。

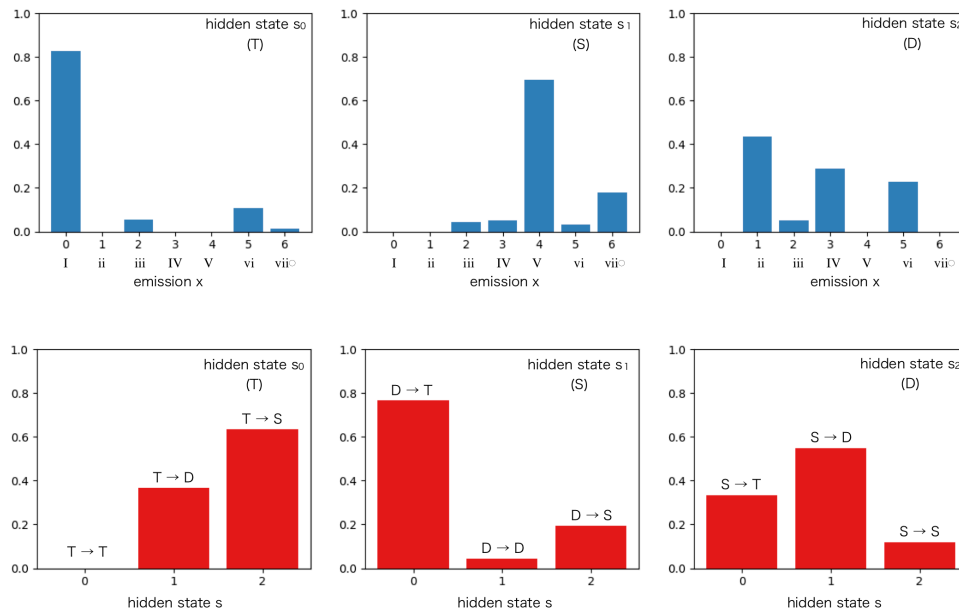


図5.3 隠れ状態数3のパラメータ（長調）

一方で機能と声理論では禁止されているドミナントからサブドミナントへに相当する確率が、少ないながらも存在していることが見られる。これは **VI** がトニックよりもサブドミナントの方に多く出現していることと関係がある可能性がある。機能と声理論においては終止における **VI** はトニックとみなされるが、サブドミナントやドミナントへ進行する経過の中ではサブドミナントとみなすことが多い。

5.3 調判定結果による隠れ状態数選択

5.3.1 隠れ状態数と調判定結果の対数尤度合計

調判定で得られた結果の対数尤度の全ての楽曲の合計は、図 5.4 に示すように長調の楽曲の合計、短調の楽曲の合計ともに隠れ状態数 6 が最大であった。ここで、調判定は楽曲全体に対して行うことが理想であるが、楽曲全体に対応する長い系列に対し集合分割問題を行うには計算時間を要するため、2-12 の異なる隠れ状態数かつ学習データに含まれる楽曲全てについて計算を行う必要上、楽曲を 5.1 節に説明したフェルマータ記号によりフレーズに区切って処理を行なった。長調と短調で絶対値に差があるのはデータセットにおける長調と短調の楽曲数に違いがあるためであり、隠れ状態数に対する相対的な値の大小のみに注目されたい。この結果から、隠れ状態 6 の最も良いパープレキシティを得たパラ

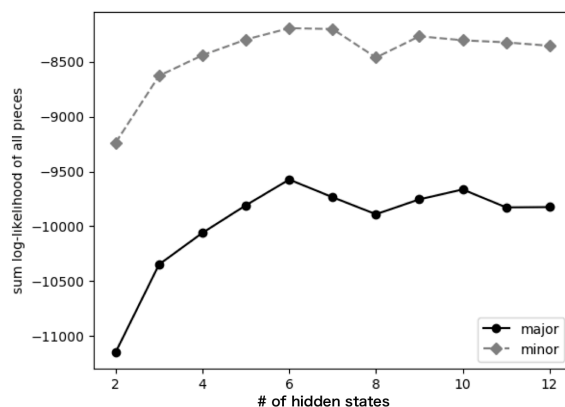


図5.4 隠れ状態数と対数尤度合計

メータが対数尤度の観点で最も良い性能を示したと言える。このように調判定の性能に関しては、隠れ状態を増やせば増やすほど減少したパープレキシティとは異なり、現実的な隠れ状態数でピークを迎えることがわかった。

5.3.2 選択された隠れ状態数 6 の性質（長調）

調判定で最も良い性能を示した隠れ状態数 6 の場合を見ると、長調は図 5.5 の結果となった。この結果は以下の傾向を示している（図 5.5）。

- **iii** と **vi** が 1 つの状態であることを除き、全体的にほぼ 1 和音が 1 機能に割り当てられている。
- **V** に対応する隠れ状態 s_1 と **vii^o** に対応する隠れ状態 s_5 は **I** に対応する隠れ状態 s_3 に最も進行しやすく、いずれもドミナントの性質を示しているが、**vii^o** は **V** に比べて **iii** と **vi** に対応する隠れ状態 s_2 に進行しにくいなどの差異も見られる。
- **ii** に対応する隠れ状態 s_0 は **IV** に対応する隠れ状態 s_4 よりもドミナントの性質をもつ隠れ状態への遷移確率が高い。

1 和音 1 機能の様相を示しているのは、2.1.6 節に述べたように和声の機能が主要 3 和音と副 3 和音に分けられるように、同じ機能の中でも個別の和音には若干異なる性質があるためと考えられる。その様子は、隠れ状態数 6 までのパラメータを隠れ状態数を一つずつ増やしながらか段階的に見ることによって明らかになる。図 5.6 は隠れ状態数 3 から 6 までの和音の出力確率パラメータを、わかりやすさのため、隠れ状態を性質に近い順に並び替えて示したものである。この結果は、以下の性質を示している。

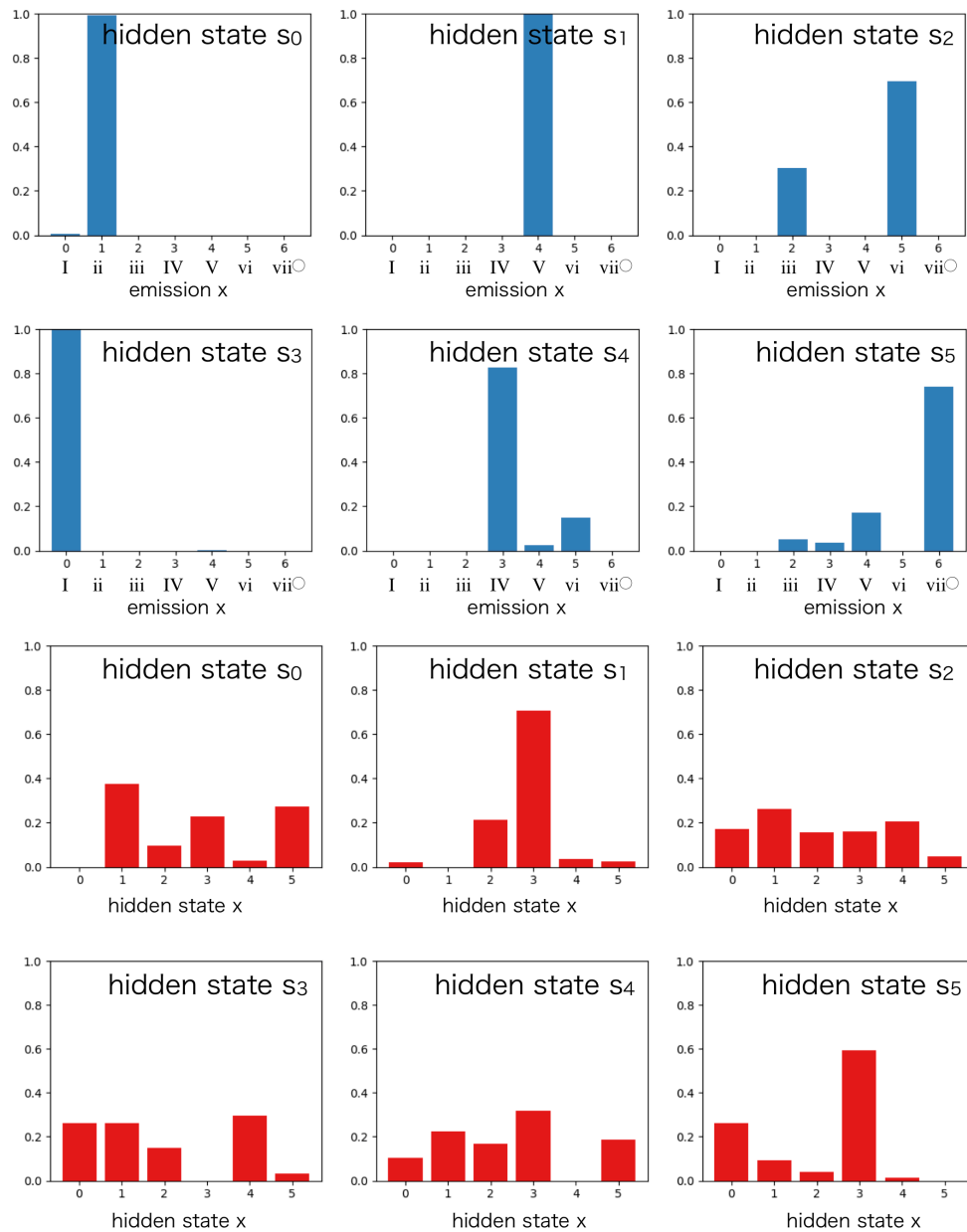


図5.5 隠れ状態数 6 のパラメータ (長調)

- 隠れ状態数 3 では、左から順にトニック、サブドミナント、ドミナントに対応するグループが形成される。
- 隠れ状態数 4 では、サブドミナントが **vi** を中心とするグループと、**ii** および **IV** のグループに分割される。
- 隠れ状態数 5 では、ドミナントが **V** のグループと、**vii°** のグループに分割される。

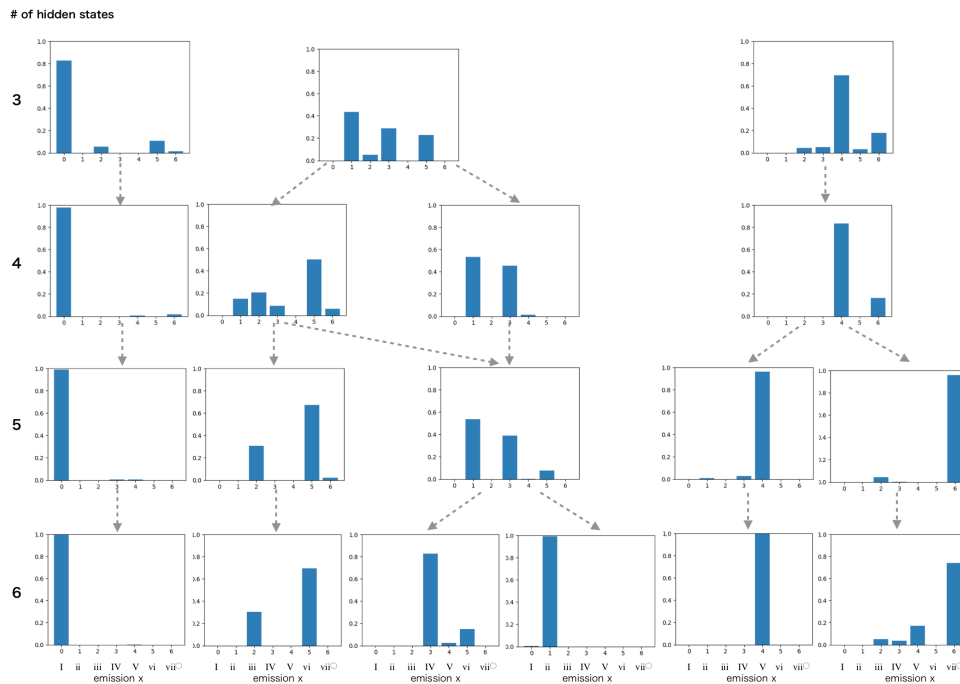


図5.6 隠れ状態数 3-6 の和音出力確率（長調）

また、**vi** が **iii** を含むトニック的なグループと、サブドミナントの **ii** および **IV** のグループに分配される。

- 隠れ状態数 6 では、サブドミナントが **IV** のグループと、**ii** のグループに分割される。

5.3.3 選択された隠れ状態数 6 の性質（短調）

一方、隠れ状態数 6 の短調の結果（図 5.7）は、以下に示すように 5.1 節で示唆した平行調への遷移を含む様相を示している。

- 隠れ状態 0 は、トニック **i** に対応するグループである。
- 隠れ状態 1 は、**V** と **vii°** に対応するドミナントのグループである。
- 隠れ状態 2 は、**ii°** および **VII** に対応するが、これは平行調に置き換えれば **vii°** および **V** であり、したがって平行調のドミナントに対応するグループである。
- 隠れ状態 3 は、**iv** と **VI** のサブドミナントのグループである。ここで、ドミナント化していない短 3 和音の **v** がこのグループに含まれる。
- 隠れ状態 4 は **ii°** を中心としたサブドミナントのグループである。

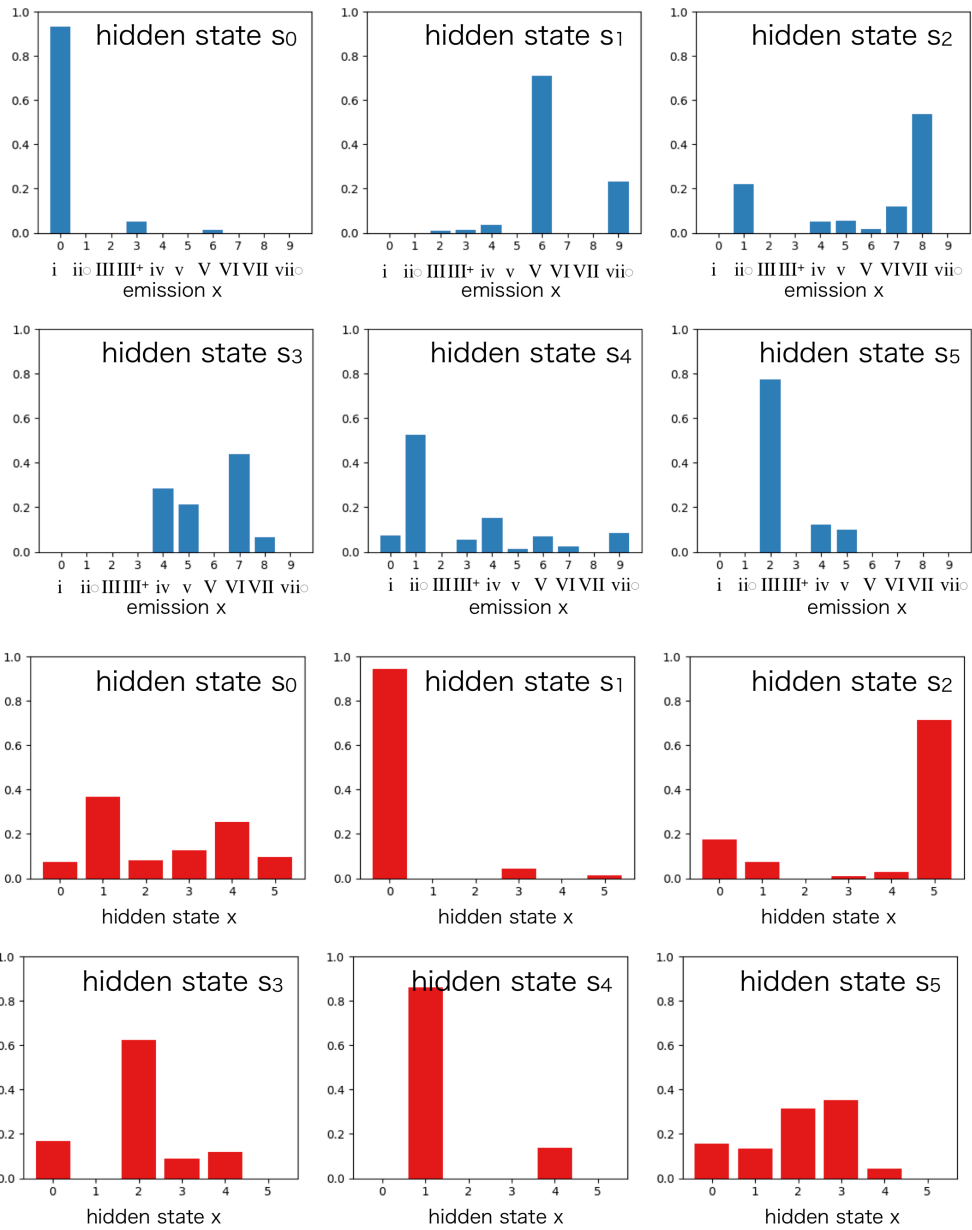


図5.7 隠れ状態数6のパラメータ（短調）

- 隠れ状態5は **III** すなわち平行調における **I** を中心としたグループである。

この短調の結果は、J.S.Bach のコラール作品はより古い時代の聖歌に和声付けをしたものであり、5.1 節で示したように和音の構成比率からも示唆された旋法的特徴を反映した結果と考えられる。

5.4 集合分割問題を用いた調判定

5.2 節において、各隠れ状態数内でのパープレキシティによる局所解のランク付け、および調判定の結果を用いた隠れ状態数の選択により、隠れ状態数 6 における最もパープレキシティの小さい HMM パラメータが選択された。ところで、調判定の先行研究においては 3.2 節に述べたように以下の 2 つが課題となっていた。

- ヒストグラムベースの調判定手法では、あらかじめ分析対象範囲を決める必要があるため、局所的な転調をピンポイントで捉えることはできない
- TPS による和音間距離をコストとした最短経路探索による方法では、和音間距離が固定のため、楽曲のスタイルを反映できない

本節では、以上の課題に対する、5.2 節で述べた提案手法によって得られた隠れ状態数 6 の HMM を用いた集合分割問題による調判定手法の効果を、具体的に楽曲例をいくつか提示することにより定性的に考察する。ここで、例示する楽曲の五線譜下部の記号は、以下の内容を示す。

- 1 段目 通し番号
- 2 段目 和音の根音とその展開数 (0 は基本形, 1 は第 1 転回等)
- 3 段目 調判定結果の主音 (大文字は長調, 小文字は短調を示す)
- 4 段目 隠れ状態番号 (図 5.5 または図 5.7 に対応する)

主調と属調を含む長調の解析例 1 (BWV281)

まず、長調かつ主調と属調のみを含む長調の典型的な楽曲の解析例について示す。この曲は主調と属調のみを含み、各調の内部では他の調からの借用和音もなく、比較的単純な進行と言える。図 5.8 に示す解析結果は、通し番号 20–31 において主調 F major から属調 C major への転調を検出することに成功している。しかしながら、転調の区切りが 1 音ずれてしまっている (本来通し番号 32 まで C major となるべきである) 点と、通し番号 8–11 で B♭ major を誤検出している点に課題がある。

一方、5.1 節で述べたように、対象楽曲のコラールはフレーズの区切りにフェルマータ記号が付与されていることから、これを用いてあらかじめフレーズに区切った上で区切り内の調判定を行うことにより上記のような問題が解消する場合がある。図 5.9 は同じ

bwv281.mxl

The image displays a musical score for BWV 281, consisting of four systems of music. Each system includes a treble clef staff, a bass clef staff, and a set of annotations below. The annotations are organized into rows: 'No.' (measure numbers), 'name' (chord names), 'key' (key signatures), and 'state' (fingerings). The first system covers measures 0-19, and the second system covers measures 20-40.

No.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
name	F0	F0	C1	F3	B-1	C0	Dm0	C1	C0	F0	B-0	F1	Gm0	Edim1	F0	F1	Fother	C0	F0	F1
key	F0	F0	F0	F0	F1	F1	F1	F1	B-2	B-2	B-2	B-2	F3	F3	F3	F3	F4	F4	F4	F4
state	s3	s3	s1	s3	s4	s1	s2	s1		s1	s3	s1	s0	s5	s3	s3		s1	s3	s3

No.	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Bdim0	G1	C0	C0	C0	Dm0	Dm1	Am0	Am3		Dm1	Dother	G0	C0	F0	Gm0	Edim1	F1	Gm1	C0	C0	F0
C5	C5	C5	C5	C6	C6	C6	C6	C7		C7	C7	C7	F8	F8	F8	F8	F8	F9	F9	F9	F9
s5	s1	s3	s3	s0	s0	s0	s2	s0		s0	s1	s1	s3	s0	s5	s3	s0	s1	s1	s3	s1

図5.8 BWV281 の調判定結果（楽曲全体）

bwv281.mxl

No. 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
name F0 F0 C1 F3 B-1 C0 Dm0 C1 C0 F0 B-0 F1 Gm0 Edim1 F0 F1 Fother C0 F0 F1
key F0 F0 F0 F0 F0 F0 F1 F1 F1 F1 F0 F0 F0 F0 F1 F1 F1 F1 F1 C0
state s3 s3 s1 s3 s4 s1 s2 s1 s1 s3 s4 s3 s0 s5 s3 s3 s1 s3 s4
6

20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39
Bdim0 G1 C0 C0 C0 Dm0 Dm1 Am0 Am3 Dm1 Dother G0 C0 F0 Gm0 Edim1 F1 Gm1 C0 C0
C0 C0 C0 C0 C0 C0 C0 C0 C0 C0 C0 F0 F0 F0 F0 F0 F0 F0
s5 s1 s3 s3 s3 s0 s0 s2 s2 s0 s1 s3 s3 s0 s5 s3 s0 s1 s1
9

40
F0
F0 1
s3 1

図5.9 BWV281 の調判定結果（フレーズの区切りを与えた場合）

BWV281 についてフレーズ区切りを与えて調判定を行なった結果であるが、上記の問題点がなく正しい解析結果を得られている。

主調と属調を含む長調の解析例 2 (BWV255)

この曲も主調と属調のみを含むが、先に示した BWV281 よりもやや複雑な和声進行となっている。図 5.10 に示した結果はその推移を概ね良好に捉えている。しかし通し番号 22 番目の和音に対応するは本来 G major であるべきところ、これに続く C major のグループに組み込まれてしまっているなど調の区切りの正確さに欠ける点と、冒頭から通し番号 4 までは G major というよりは C major における半終止であると考えられる点に課題がある。

一方、図 5.11 には同じ BWV255 の楽曲についてフレーズ区切りを与えて調判定を行なった結果を示す。この曲における最後のフレーズ（通し番号 33-43）は、本来 C major 内で vi から IV に向かう一連の流れが転調とみなされてしまっている。

主調と平行調を含む短調の解析例 (BWV266)

BWV266 は短調の主調 e minor と平行調 G major を含む場合である。和音の構成比率および和声機能パラメータの学習結果が示唆するように、コラール作品で短調に分類される楽曲においては、平行調と共通の和音あるいは平行調への転調が多く含まれる。

BWV266 は通し番号 16-33 において明らかに G major に転調しているが、楽曲全体の解析結果 (図 5.12) はこの転調を検出することに成功している。ただしこのフレーズの冒頭部分では通し番号 17 の借用和音 C \sharp diminished に引きずられて D major を検出してしまっている。さらに、続くフレーズ 34-70 で主調の e minor への復帰も正しく検出できている。

あらかじめフレーズの区切りを与えた場合 (図 5.13) は、フレーズ内の調がより一貫している。ここではフレーズ 34-70 の解析結果は e minor で一貫しているが、先に述べたフレーズに区切らない場合の検出結果も間違いではない。このように転調とみなすべきか、または転調の分割点をどこにするかは、一意に定まらない場合がある。

長調と短調を含む複雑な場合 (BWV271)

BWV271 は D major に始まり、b minor の半終止で終わる、旋法的な性格を持った曲である。楽曲全体の解析結果 (図 5.14) は冒頭は G major に揺れているものの開始と終

bwv255.mxl

No.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
name	C0	G0	D0	D0	G0	C0	Gm5	F0	Am2	Bdim1	C0	G0	Am0	F#dim1	G1	G5	D0	G0	Cother
key	G0	G0	G0	G0	C1	C1	C1	C1	C2	C2	C2	C2	C2	C2	G3	G3	G3	G3	G4
state	s4	s3	s1	s1	s1	s3	s4	s2	s5	s3	s1	s2	s3	s3	s1	s3			

	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
Am1	D0	D0	G0	C1	Bdim1	C0	G0	G0	C0	C3	D2	F#dim1	G0	C1	D1	G#dim0	Am0	Dm2	C2	F0	
G4	G4	G4	C5	C5	C5	C5	G6	G6	G6	G7	G7	G7	G7	G7	G7	G7	C8	C8	C8	C8	
s0	s1	s1	s1	s3	s5	s3	s1	s4	s4	s4	s1	s5	s3	s4	s5		s2	s0	s3	s4	

	40	41	42	43
Coother	C9	G0	G0	C0
	C9	C9	C9	C9
	s1	s1	s3	s1

図5.10 BWV255 の調判定結果（楽曲全体）

bwv255.mxl

No. 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
 name C0 G0 D0 D0 G0 C0 Gm5 F0 Am2Bdim1 C0 G0 Am0F#dim1 G1 G5 D0 G0 Cother
 key C0 C0 C0 C0 C0 C0 C0 C0 C0 C0 C0 G0 G0 G0 G0 G0 G0 G1
 state s3 s4 s1 s3 s4 s2 s5 s3 s3 s0 s5 s3 s3 s1 s3
 5
 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33
 Am1 D0 D0 G0 C1 Bdim1 C0 G0 G0 C0 C3 D2 F#dim1 G0 C1
 G1 G1 G1 G1 C0 C0 C0 C0 G0 G0 G0 G0 G0 G1 F0
 s0 s1 s1 s3 s3 s5 s3 s1 s3 s4 s4 s1 s5 s3 s1
 9
 34 35 36 37 38 39 40 41 42 43
 D1 G#dim0 Am0 Dm2 C2 F0 Cother G0 G0 C0
 F0 F0 F0 F0 F0 F0 C1 C1 C1 C1
 s2 s2 s1 s3 s1 s1 s3

図5.11 BWV255 の調判定結果（フレーズの区切りを与えた場合）

bwv266.mxl

No.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
name	Em0	Em1	Em0	B1	B0	Em0	A1	B1	Em0	Aother	Bm1	Am1	Em2	D#dim3	Am1	B0	Em0	C#dim1	D0	D3	G1	Am0	F#dim1
key	e0	e0	e0	e0	e1	e1	e1	e1	e1	G2	G2	G2	G2	e3	e3	e3	e3	D4	D4	D4	D4	G5	G5
state	s0	s0	s0	s1	s0	s0	s1	s0	s0	G2	s2	s0	s2	s4	s4	s1	s0	s5	s3	s3	s4	s0	s5

	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0
G5	G5	G5	G6	G6	G6	G6	G6	G7	G7	G7	G7	G7	G8	G8	G8	G8	G8	e9	e9	e9	e9	e9	e9	e9	e9	e9
s3	s3	s1	s2	s2	s4	s0	s0	s1	s3	s1	s3	s3	s3	s4	s4	s4	s4	s1	s1	s0	s0	s0	s0	s0	e9	s1

	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70			
C0	C0	F#dim1	B0	Em0	Em0	A1	D0	D3	F#dim5	F#dim1	G0	Em1	Eother	Eother	B1	Em0	Em0	Gaug3	Em1	Em5	Eother	B0			
e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9
s3	s4	s1	s0	s0	s0	e9	s3	s3	s2	s2	s5	s0	e9	e9	s1	s0	s0	s1	s0	s0	e9	e9	e9	e9	s1

	71	72	73	74	75	76	77	78	79	80	81	82	83
Em0	Em0	Em0	Aother	B1	B1	Em0	Bm1	C0	Gaug1	F#dim2	Am1	B0	
e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9	e9
s0	s0	s0	s1	s1	s0	s3	s2	s0	s4	s4	s1	s1	

図5.12 BWV266 の調判定結果（楽曲全体）

bwv266.mxl

The image shows a musical score for BWV 266, divided into three systems. Each system consists of a treble clef staff, a middle staff, and a bass clef staff. Below each system is a table of chord names and states for each measure.

No.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
name	Em0	Em1	Em0	B1	B0	Em0	A1	B1	Em0	Aother	Bm1	Am1	Em2	D#dim3	Am1	B0	Em0	C#dim1	D0	D3	G1	Am0	F#dim1
key	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0
state	s0	s0	s0	s1	s1	s0	s1	s0	s0	e0	s5	s3	s0	s4	s4	s1	s2	s1	s1	s3	s0	s0	s5

No.	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
name	G0	G0	G0	G0	G0	C0	Coher	Am0	Am0	D0	G0	D0	G0	G3	Em0	Em0	F#dim3	Am2	B1	B0	Em0	Em0	Em1	Em0	Eother	B0
key	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	D0	G0	G3	Em0	Em0	F#dim3	Am2	B1	B0	Em0	Em0	Em1	Em0	Eother	B0
state	s3	s1	s2	s2	s4	s0	s0	s1	s3	s1	s2	s5	s5	s0	s0	s4	s4	s1	s1	s0	s0	s0	s0	e0	e0	e1

No.	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70			
name	C0	F#dim1	B0	Em0	Em0	A1	D0	D3	F#dim5	F#dim1	G0	Em1	Eother	Eother	B1	Em0	Em0	Gaug3	Em1	Em5	Eother	B0			
key	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0	e0			
state	s4	s4	s1	s0	s0	e0	s3	s3	s2	s2	s5	s0	e0	e0	e0	s1	s0	s0	s1	s0	s0	e0	e0	e1	s1

No.	71	72	73	74	75	76	77	78	79	80	81	82	83
name	Em0	Em0	Em0	Aother	B1	B1	Em0	Bm1	C0	Gaug1	F#dim2	Am1	B0
key	e0	e0	e0	e0	e0	e0	e0	e1	e1	e1	e1	e1	e1
state	s0	s0	s0	s1	s1	s0	s3	s2	s0	s4	s4	s1	s1

図5.13 BWV266 の調判定結果（フレーズの区切りを与えた場合）

bwv271.mxl

The image shows a musical score for BWV 271 in three systems. Each system consists of four staves: two treble clefs and two bass clefs. Below each system is a table of chord analysis data.

No.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
name	D0	G0	D1	Em0	C#dim1	D0	Em1	A0	D0	F#1	Bm0	D3	Bm1	Bm5	Bother	Bother	F#0	F#0	Bm0	Bm0	Bm3
key	G0	G0	G0	G0	D1	D1	D1	D1	D1	b2	b2	b2	b2	b2	F#3	F#3	F#3	F#3	b4	b4	b4
state	s1	s3	s1	s2	s5	s3	s0	s1	s3	s1	s0	s5	s0	s0	s0	s3	s3	s0	s0	s0	s0

	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
name	C#dim1	Em0	D1	F#m0	G0	Bm2	C#dim1	A2	G2	D0	Em3	D0	D0	G0	G3	F#dim3	D#dim0	Em0	G3	Em1	Em0	B0	E1
key	b4	D5	D5	D5	D5	D5	D6	D6	D6	D6	D7	D7	D7	D7	e8	e8	e8	e8	e8	e9	e9	e9	e9
state	s4	s0	s3	s2	s4	s2	s5	s1	s4	s3	s0	s3	s3	s5	s5	s4	s1	s0	s5	s0	s0	s1	s1

	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68
name	A0	Bm3	A0	A0	Bm0	G#dim1	A1	Bm1	E0	A0	D0	D0	Em0	C#dim1	D1	G0	A0	Em1	F#m0	C#dim1	D0	D0	Bm4	Bm0	F#0
key	A10	A10	A10	A11	A11	A11	A11	A11	A12	A12	A12	A12	D13	D13	D13	D13	b14	b14	b14	b14	b14	b15	b15	b15	b15
state	s3	s0	s3	s0	s0	s5	s3	s0	s1	s3	s4	s4	s0	s5	s3	s4	s2	s5	s3	s2	s5	s0	s0	s0	s11

図5.14 BWV271 の調判定結果（楽曲全体）

bwv271.mxl

The image displays a musical score for BWV 271 in G major, consisting of four systems of music. Each system includes a treble clef, a bass clef, and a grand staff. Below each system, there are two rows of text: 'name' and 'state', which correspond to the chord names and state codes listed in the table below.

No.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
name	D0	G0	D1	Em0C#dim1	D0	Em1	A0	D0	F#1	Bm0	D3	Bm1	Bm5	Bother	Bother	F#0	F#0	Bm0	Bm0	Bm3	Bm3	
key	G0	G0	G0	G0	D1	D1	D1	D1	D1	b0	b0	b0	b0	b0	b0	b1	b1	b1	b1	b1	D0	D0
state	s1	s3	s1	s2	s5	s3	s0	s1	s3	s1	s0	s5	s0	s0	s0	s1	s1	s0	s0	s2	s2	s2

	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43
C#dim1	Em0	D1	F#m0	G0	Bm2C#dim1	A2	G2	D0	Em3	D0	D0	D0	G0	G3	F#dim3	D#dim0	Em0	G3	Em1	Em0	B0	E1	A0
D0	D0	D0	D0	D0	D0	D0	D0	D0	D0	D0	D0	D0	e0	e0	e0	e0	e0	e0	e1	e1	e1	e1	A0
s5	s0	s3	s2	s4	s2	s5	s1	s4	s3	s0	s3	s2	s2	s5	s5	s4	s1	s0	s5	s0	s0	s1	s1

	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68
A0	Bm3	A0	A0	Bm0G#dim1	A1	Bm1	E0	A0	D0	D0	Em0C#dim1	D1	G0	A0	Em1	F#m0C#dim1	D0	D0	Bm4	Bm0	F#0	F#0	F#0	F#0	
A0	A0	A0	A0	A0	A0	A1	A1	A1	A1	b0	b0	b0	b0	b0	b0	b0	b0	b0	b0	b0	b0	b0	b0	b0	b0
s3	s0	s3	s3	s0	s5	s3	s0	s1	s3	s1	s5	s5	s3	s2	s5	s3	s2	s5	s3	s2	s5	s5	s0	s0	s1

図5.15 BWV271 の調判定結果（フレーズの区切りを与えた場合）

了の調を正しく判定できている。D major と b minor は主調と平行調の関係にあり、先に述べたようにこの調遷移はコラールにおいては比較的出現しやすいと言える。それだけでなく、通し番号 33–42 では D major から e minor への転調があるが、こちらも正しく検出できている。続く通し番号 43–53 の A major への転調もまた、正しく検出されている。楽曲をセクションに区切った場合 (図 5.15) も、解析結果はほぼ同様である。

複雑な借用和音を含む場合 (BWV292)

最後に、冒頭に複雑な進行を含む場合を示す。BWV292 の冒頭のフレーズは V 度のまわりを F \sharp が、IV 度のまわりを B \flat が修飾していることによって、それぞれ G major, F major と検出されているが、一貫して C major とするのが最も自然である。楽曲全体の解析結果 (図 5.16) は通し番号 27–30, 71–75 に局所的な和音の中心を転調と誤検出しているが、そのほかは調の遷移を良好に解析できている。フレーズをあらかじめ与えた場合 (図 5.17) は、冒頭を除いてほぼ良好に解析できている。

5.5 実験結果のまとめ

本研究では、HMM による和声進行の生成モデル学習を行い、最適な隠れ状態数を転調判定の対数尤度のデータセットに対する合計により求めた。その結果、隠れ状態数は同一調性内では多ければ多いほどパープレキシティが小さいという先行研究と同様の結果を得たが、転調判定の性能からは妥当な状態数に落ち着くことを示した。

長調のモデルでは、隠れ状態数 6 に対し得られた隠れ状態のパラメータはほぼ 1 和音 1 状態の様相であり、{ii}, {V}, {iii, vi}, {I}, {IV}, {vii $^\circ$ } であった。隠れ状態数 3 では、得られた隠れ状態のパラメータは機能と声理論の知見に非常に近いものであり、トニック、ドミナント、サブドミナントに相当する隠れ状態が得られた。そして、隠れ状態数を 6 まで増やしていくにつれて、基本的に同じ機能 (ドミナント等) が分裂し、より詳細な隠れ状態が得られることがわかった。例えば隠れ状態数 4 では、図 5.6 に示したようにサブドミナントが IV および ii のグループと vi のグループに分かれる。さらに隠れ状態数 5 では、ドミナントが V と vii $^\circ$ に分けられるとともに、VI がトニックとサブドミナントの 2 つの機能に分割される。そして、隠れ状態数 6 では、ii と IV が独立したグループに分かれる。

短調のモデルに対し得られたパラメータは、長調とはかなり異なっている。短調では、隠れ状態数 4 以上の場合に平行長調に相当するグループが形成された。これは学習データ

bwv292.mxl

No.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
name	C0	Coher	G1	Am0	F#dim1	G0	G3	C1	C1	F0	F0her	C0	C0	C0	D3	G1	Bm0	C3	F#dim1	G0	Am0	G1	Am1	Am0	G0her	D0	D0	G0
key	G0	G0	G0	G0	G1	G1	G1	G1	F2	F2	F2	F2	F2	F2	F2	G3	G3	G3	G3	G4	G4	G4	G4	D5	D5	D5	D5	e6
state	s4	s3	s0	s5	s3	s3	s4	s3	s3	s4	s4	s4	s4	s4	s4	s3	s2	s4	s5	s3	s0	s3	s0	s0	s3	s3	s3	s5

No.	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56
name	Em0	Em0	Em0	Em3	Am1	F#dim2	G1	Am0	F1	G1	Em2	G1	C0	F0	F0	G3	C1	Em0	F3	Bdim1	C0	Dm0	C1	Dm1	Dm0	Coher	G0	G0	C0
key	e6	e6	e6	G7	G7	G7	G7	C8	C8	C8	C8	C8	C8	C8	G9	G9	G9	G9	C10	C10	C10	C10	C10	C10	G11	G11	G11	G11	C12
state	s0	s0	s0	s0	s0	s5	s3	s2	s4	s1	s2	s1	s3	s4	s3	s3	s4	s2	s4	s5	s3	s0	s3	s0	s0	s3	s3	s3	s3

No.	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	
name	C0	Coher	G1	C1	F#dim3	G2	G2	G2	Am1	Am1	D0	D0	D0	G0	C0	Dm0	C1	F0	F3	Am2	Bdim1	F0her	C0	Coher	G1	G1	C0	
key	C12	C12	G13	G13	G13	G13	G13	G14	G14	G14	G14	G14	G14	G14	F15	F15	F15	F15	F15	C16	C16	C16	C16	C16	C17	C17	C17	C17
state	s3	s1	s4	s5	s3	s3	s3	s0	s0	s1	s1	s1	s3	s1	s2	s1	s3	s3	s2	s5	s3	s3	s3	s1	s1	s1	s3	s3

No.	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	
name	F0	F0	F0her	C0	C0	Dm0	Dm0	Am0	Am3	Dm1	Bdim2	Dm1	G0	G0	G0	C0	
key	F18	F18	F18	F18	C19	C19	C19	C19	C20	C20	C20	C20	C21	C21	C21	C21	
state	s3	s3	s3	s1	s0	s0	s0	s0	s2	s0	s0	s5	s0	s1	s1	s1	s3

図5.16 BWV292 の調判定結果（楽曲全体）

bwv292.mxl

The image displays a musical score for BWV 292, consisting of four systems of music. Each system includes a table of chord names and states corresponding to the measures of the score. The chord names are listed in the 'name' row, and the states are listed in the 'state' row. The measures are numbered from 0 to 99.

No.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
name	C0	Coher	G1	Am0	F#dim1	G0	G3	C1	C1	F0	Foher	C0	C0	C0	D3	G1	Bm0	C3	F#dim1	G0	Am0	G1	Am1	Am0	Coher	D0	D0	G0
key	G0	G0	G0	G0	G0	G0	F1	F1	F1	F1	F1	C0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0
state	s4	s3	s0	s5	s3	s3	s1	s1	s1	s3	s1	s1	s4	s4	s1	s3	s2	s4	s5	s3	s0	s3	s0	s0	s0	s1	s1	s3

No.	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56
name	Em0	Em0	Em0	Em3	Am1	F#dim2	G1	Am0	F1	G1	Em2	G1	C0	F0	F0	G3	C1	Em0	F3	Bdim1	C0	Dm0	C1	Dm1	Dm0	Coher	G0	G0	C0
key	G0	G0	G0	G0	G0	G0	C1	C1	C1	C1	C1	C1	C1	C0	C0	C0	C0	C0	C0	C0	C0	C0	C0	C0	C0	C0	C0	C0	C0
state	s2	s2	s2	s2	s0	s5	s1	s2	s4	s1	s2	s1	s3	s1	s4	s4	s1	s3	s2	s4	s5	s3	s0	s3	s0	s0	s1	s1	s3

No.	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83
name	C0	Coher	G1	C1	F#dim3	G2	G2	G2	Am1	Am1	D0	D0	D0	G0	C0	Dm0	C1	F0	F3	Am2	Bdim1	Foher	C0	Coher	G1	G1	C0
key	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	G0	C0	C0	C0	C0	C0	C0	C0	C1	C1	C1	C1	C1	C1
state	s4	s4	s3	s4	s5	s3	s3	s3	s0	s0	s1	s1	s1	s3	s1	s3	s0	s3	s4	s4	s2	s5	s3	s0	s1	s1	s3

No.	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
name	F0	F0	Foher	C0	C0	Dm0	Dm0	Am0	Am3	Dm1	Bdim2	Dm1	G0	G0	G0	C0
key	C0	C0	C0	C0	C0	C0	C0	C0	C0	C0	C0	C0	C0	C0	C0	C0
state	s4	s4	s0	s3	s3	s0	s0	s2	s2	s0	s5	s0	s1	s1	s1	s3

図5.17 BWV292 の調判定結果（フレーズの区切りを与えた場合）

として用いたコラールの特徴が反映された結果と考えられる。学習データのコラールは J.S.Bach が和声付けをしたものであるが、元のメロディは中世の聖歌によるものであり、教会旋法が用いられているからである。

状態数 6 の短調の隠れ状態は、 $\{i\}$, $\{V, vii^\circ\}$, $\{ii^\circ, VII\}$, $\{iv, v, VI\}$, $\{ii^\circ\}$, $\{III\}$ のようになった。これは一見複雑に見えるが、次のように解釈できる。すなわち、 $\{i\}$ はトニックであり、 $\{V, vii^\circ\}$ はドミナントであり、 $\{iv, v, VI\}$ と $\{ii^\circ\}$ はサブドミナントである。そして、 $\{III\}$ は平行長調 C major におけるトニックであり、同様に $\{ii^\circ, VII\}$ は平行長調におけるドミナントである。 $\{ii^\circ\}$ が 2 つのグループに分割されたのは興味深い。一方のグループは通常のサブドミナントとしてドミナントに相当するグループに進行しやすいものであり、もう一方は平行長調 C major のトニックに進行しやすいグループである。

本研究で短調の場合に和声的短音階を用いるのではなく、導音の導入により生じる和音のすべての組み合わせを用いた理由は、短調の楽曲上に、和声的短音階上の 3 和音では含まれない C major が多く生じていたためである。この事実および得られた隠れ状態の結果は、J.S.Bach のコラール作品中の短調の楽曲では、短調と平行調が明確な区分なく移り変わり、したがって今日の視点では旋法的な性格を有していることを示唆している。

第 6 章

終わりに

本研究の成果は次の 3 点である。

1. 先行研究において HMM による統計的学習により特定の少ない隠れ状態数で機能
和声理論の知見が得られることが知られていたものの、尤度の観点からは隠れ状態
数が多ければ多いほど良く、最適な隠れ状態数を決定する方法が知られていなかっ
た点に対し、得られた HMM を転調判定に用い、その結果得られる尤度の合計に
よって最適な隠れ状態数を選択できることを示した。
2. 選ばれたパラメータの示す内容を音楽学の知見に照らして分析し、機能
和声理論の知見にかなう和声機能を示しているとともに楽曲特有の特徴を反映していることを
定性的に明らかにした。
3. 本研究に付随して構築した調判定アルゴリズムは、判定対象範囲の指定を必要とせ
ずに転調検出が可能であり、かつ楽曲様式を反映することができる点で先行研究に
ない利点を持つ。

このように、HMM による和声機能の統計的学習に対し、提案手法を用いて最適なモデル
サイズおよび局所解を選択し、機能
和声理論の知見にかなうとともに楽曲様式を反映した
和声機能が獲得できることが明らかになった。これにより、既存の和声理論が最もよく
当てはまる古典派の楽曲に限られない知識を獲得し、このモデルを多様な楽曲に展開でき
る可能性を示した。

例えば、本研究により示唆された J.S.Bach の短調のコラールに見られる短調と長調が
明確な区分なく流れるように移り変わる和声進行は、機能
和声理論や TPS[8] が最もよく
当てはまる古典派の時代にはあまり用いられなくなったが、ロマン派以降再びその魅力が
再発見され、G.Fauré らの旋法的な和声進行へとつながっている。後期ロマン派以降の楽

曲では、調性音楽の基本的枠組みを保ちつつも、和声の解釈は拡大され、作曲家によって多様で個性的な進行が用いられるようになった。

しかしながら、HMMは生成モデルとして最もシンプルなものの一つであり、HMMでは表現できない音楽的現象が存在すると考えられる。例えば、HMMは直前の和声機能とのつながりしかモデルに含まれないため、カデンツのように、直前の和音ではなく、より以前にある和音が現在の和音を強く予期するといった、遠隔の係り受け構造を考慮できないことなどである。さらに、和音だけでなく調にも近親調への転調や復調といった構造があるにも関わらず、その点を考慮することができない。これらの理由により提案手法の転調判定においても、局所的に支配的な和音を転調と判定する誤りが見受けられた。このようにHMMには課題もあるが、本研究で明らかになったようにHMMの隠れ状態は和声機能をよく表現しており、局所的な構造に対し表層の和音から和声機能の解析が可能であるため、自然言語処理において形態素解析が構文解析の前提となっているように、より複雑なモデルを学習するための前提として有効に機能すると考えられる。今後、この点を活かしつつ、遠隔の係り受け構造を考慮したより複雑なモデルによる楽曲解析手法の実現が望まれる。

参考文献

- [1] L. Bernstein. *The Unanswered Question: Six Talks at Harvard*. Cambridge, Harvard University Press (1976)
- [2] H. Schenker. *Beethovens Fünfte Sinfonie*, Universal Edition (1925)
- [3] A. Cadwallader and D.Gagné. *Analysis of Tonal Music A SCHENKERIAN APPROACH*, Oxford University Press (1998)
- [4] Fred Lerdahl and Ray Jackendoff. *A Generative Theory of Tonal Music*, The MIT Press (1983)
- [5] Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo. ATTA: Automatic Time-span Tree Analyzer Based on Extended GTTM, *In Proc. of the 6th International Conference on Music Information Retrieval Conference (ISMIR)*, pp. 358-365 (2005)
- [6] C. Krumhansl and E.J. Kessler. Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys, *Psychological Review* , 89:334-68 (1982)
- [7] Shoki Sakamoto, Sean Arn, Masaki Matsubara, and Satoshi Tojo. Harmonic analysis based on Tonal Pitch Space, *Eighth International Conference on Knowledge and Systems Engineering (KSE)*, pp. 230-233 (2016)
- [8] Fred Lerdahl. *Tonal Pitch Space*, Oxford University Press (2001)
- [9] Hiroaki Tsushima, Eita Nakamura, Katsutoshi Itoyama and Kazuyoshi Yoshii. Generative Statistical Models with Self-Emergent Grammar of Chord Sequences, *ournal of New Music Research*, vol.47, no.3, pp. 226-248 (2018)
- [10] Hiroaki Tsushima, Eita Nakamura, Katsutoshi Itoyama, and Kazuyoshi Yoshii. Function and Rhythm Aware Melody Harmonization Based on Tree-Structured Parsing and Split-Merge Sampling of Chord Sequences, in *Proc. 18th Interna-*

- tional Society for Music Information Retrieval Conference (ISMIR)*, pp. 502-508 (2017)
- [11] M. Johnson. Why Doesn't EM Find Good HMM POS-Taggers?, in *Proc. EMNLP-CoNLL*, pp. 296-305 (2007)
- [12] S. Goldwater and T. Griffiths. A Fully Bayesian Approach to Unsupervised Part-of-Speech Tagging, in *Proc. ACL*, pp. 744-751 (2007)
- [13] D. Hu and L. K. Saul. A probabilistic topic model for unsupervised learning of musical key-profiles, in *Proc. 10th International Society for Music Information Retrieval Conference (ISMIR)*, pp. 441-446 (2009)
- [14] M. Rohrmeier and I. Cross, "Statistical Properties of Tonal Harmony in Bach's Chorales," in *Proc. International Conference on Music Perception and Cognition (ICMPC)*, pp. 619-627 (2008)
- [15] N. Jacoby, N. Tishby, and D. Tymoczko, An Information Theoretic Approach to Chord Categorization and Functional Harmony, *Journal of New Music Research*, 44(3), pp. 219-244 (2015)
- [16] 東条敏, 平田圭二. 音楽・数学・言語-情報科学が拓く音楽の地平, 近代科学社 (2017)
- [17] 東川清一. 音楽理論入門, 筑摩書房 (2017)
- [18] 石井健一郎、上田修功. 続・わかりやすいパターン認識 教師なし学習入門, オーム社 (2014)
- [19] 藤澤克樹、梅谷俊治. 応用に役立つ 50 の最適化問題, 朝倉書店 (2009)

謝辞

本研究を進めるにあたり，多大なご指導とご助言を賜り，なかなか研究が進まない時も励ましてくださいました，東条敏教授に深く感謝申し上げます。また，筑波大学の松原正樹助教には，和声理論について有益な議論をいただきましたこと，感謝申し上げます。京都大学研究員の中村栄太氏には，和声機能発見の先行研究について詳細なご解説をいただき，また統計的学習に関する有益なアドバイスをいただきました。ここに感謝申し上げます。