A PROBABILISTIC TOPIC MODEL FOR UNSUPERVISED LEARNING OF MUSICAL KEY-PROFILES

Diane J. Hu and Lawrence K. Saul
University of California, San Diego

01 INTRODUCTION

Harmonic analysis studies the melodies and harmonies of musical compositions. Two basic tasks involve: 1) determining the musical key of a piece and 2) tracking its modulations. Our goal is to automate these two tasks for both MIDI and WAVE music files.

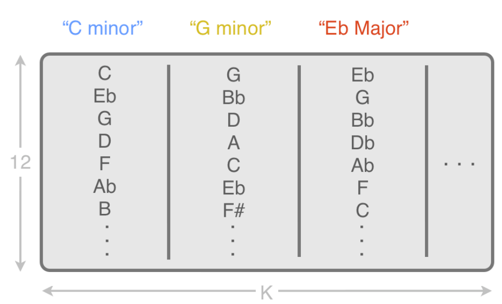
Our approach is based on Latent Dirichlet Allocation (LDA) [1], a probabilistic model that learns in an unsupervised setting. Our model learns each musical key as a key-profile – a distribution over pitch classes, and each song is modeled as a random mixture of key-profiles. Advantages are: (1) no need for labeled data, (2) ability to discover unknown correlations, (3) ability to generalize to different tonal systems.

02 OUR MODEL

02.1 Model Overview

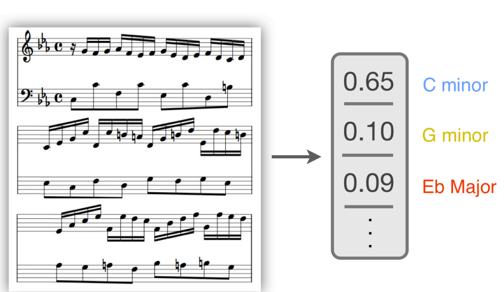
Musical Keys as "Topics"

Songs in the same key use similar sets of pitches. We look for commonly co-occurring notes in songs, to learn "key-profiles," distributions over pitches for each musical key. An example of most likely pitches are shown for each key below:



Songs as "Documents"

Songs modulate between multiple keys. Using the learned key-profiles, we can characterize each song as a distribution over musical keys.

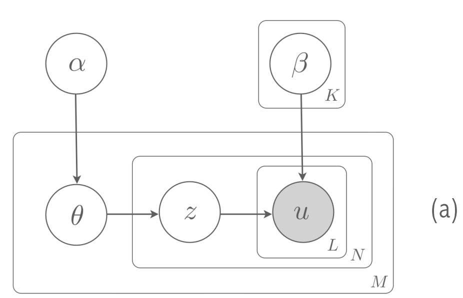


02.2 Modeling Symbolic Music

Input Data: Use MIDI files to obtain note counts for each measure.

Generative Process: Generates notes in each measure, and measures in each song:

1. Draw weight vector $\theta \sim Dirichlet(\alpha)$
2. For each measure $\mathbf{u}_n$ in the song:
   a. Choose a key $z_n \sim Multinomial(\theta)$
   b. For each note u_{nl} in the measure:
      Choose a pitch $\in \{A, A\sharp, \ldots, G, G\sharp\}$ from the distribution $p(u_{nl} = i | z_n = j, \beta) = \beta_{ij}$



Figures (a, b): Graphical representations of models using "plate notation." Nodes in boxes are replicated by number in corner.

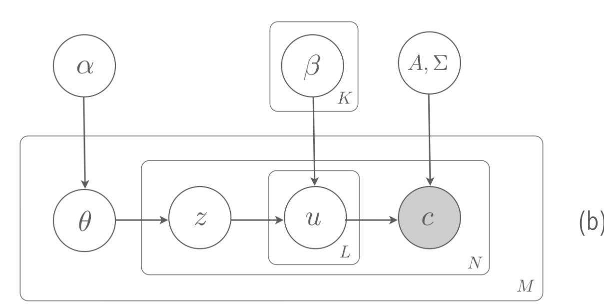
02.3 Modeling Audio Music (Work In Progress)

Input Data: Use MIDI files to synthesize WAVE files and compute 12-bin chroma vectors for each measure.

Generative Process: Same as MIDI, but with additional step: Once all notes $\{u_{n1}, \ldots, u_{nL}\}$ for measure $\mathbf{u}_n$ have been generated, draw chroma-vector c_n from the probability distribution:

$$p(c_n|\mathbf{u}_n, A) = \frac{1}{|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(c_n - A\mathbf{u}_n)^T \Sigma^{-1} (c_n - A\mathbf{u}_n)\right\}$$

where A and Σ are additional parameters to learn.



Figures (a, b): Overview of variables: M = # of songs, N = # of measures in song, L = # notes in measure, K = # latent classes

02.4 Incorporating Prior Knowledge

K, # of topics/keys: Set K = 24 to look for the 24 major and minor keys of tonal, western music.

β, key-profiles: Elements of β are tied cyclically, so that all major/minor key-profiles are related by simple transposition.

REFERENCES

[1] D. Blei, et al: "Latent Dirichlet allocation," 2003.

[2] T. Eerola, et. al: "MIDI Toolbox: MATLAB Tools for Music Research," 2004.

[3] C. Krumhansl: Cognitive Foundations of Musical Pitch, 1990.

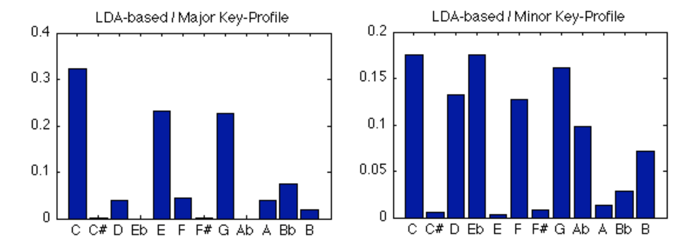
[4] D. Sleator, et al: "The Melisma Music Analyzer," 2001.

[5] D. Temperley: "A Bayesian approach to key-finding," 2002.

03 APPLICATIONS

Analyzing Key-profiles

The C major (left) and minor (right) key-profiles learned by our model.



Learned key-profiles are consistent with music theory principals. In both major and minor modes, weights are given in descending order to degrees of the triad, diatonic, and finally chromatic scales.

Key Finding (MIDI Input)

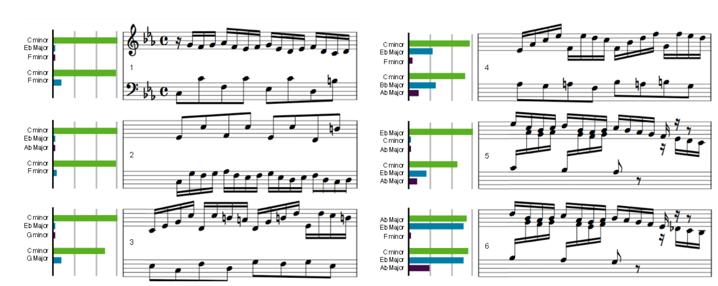
Classify the overall key of a song to be $k = argmax_k\, \theta^k$.

Dataset 1 (235 songs)		Dataset 2 (107 songs)			
LDA	$KS_{[2,3]}$	LDA	$KS_{[3,4]}$	$CBMS_{[4,5]}$	$BAYES_{[4,5]}$
86%	80%	79%	62%	66%	67%

Experiments show % of whole songs classified correctly for overall key finding. Two collections of classical midi files were used including works by Bach, Vivaldi, Mozart, Beethoven, Chopin, and Rachmaninoff. We test our model against several popular key-finding algorithms [3,5].

Modulation Tracking (MIDI Input)

Classify the key of each measure $\mathbf{u}_n$ to be $k = argmax_k\, p(z_n = k | \mathbf{u}_n)$.



An example of how our model analyzes the first six measures of Bach's Prelude in C minor from Book II of the Well-Tempered Clavier. Results are compared to annotations by a music theory expert, as recorded in [3].