Modeling Musical Influence with Topic Models

Anonymous Author(s)
Affiliation
Address
email

Abstract

The role of musical influence has long been debated by scholars and critics in the humanities, but was not studied systematically in a data-driven way. To address the question we apply topic-modeling tools [1, 2] to a dataset of 24941 songs by 9222 artists, from the years 1922 to 2010. We find that the learned models are significantly correlated with a human-curated influence measure, and clearly outperform a baseline method. We find that musical influence and musical innovation are not monotonically correlated. However, we do find that the most influential songs were more innovative during two time periods: the early 1970's and the mid 1990's.

1 Introduction

The growing availability to researchers of music tracks and of methods to capture their acoustic signatures opens new possibilities to study the structure of an entire corpus of music. This paper provides a quantitative modeling approach based on topic models to study *musical influence*. Musical influence is often discussed, but has never been studied quantitatively at a large scale before. The challenge in modeling a whole musical corpus is two-fold: The audio signal itself is a complex continuous signal, with meaningful structure on multiple time-scales; and there exist intricate and evolving relations between artists, songs, and genres. Probabilistic topic models offer a unique way to rigorously unravel these

We model musical influence using *Dynamic Topic Models* (DTM) [1] and the *Document Influence Model* (DIM) [2]. These models were originally developed in the context of analyzing how the language of scientific papers evolves. Under the DIM, an influential scientific paper is one whose language is adopted by its successors in its scientific field. In our case, the audio content of songs replaces the text of a scientific paper, and we consider a song to be influential if its "musical-language" (or sound-content) has been adopted by later songs in related genres. We find that the DIM captures known historical dynamics of popular music, as validated by manually curated data. For example, it clearly shows the lineage leading from Reggae, Disco and Funk to modern electronic musical genres on one hand, and Hip Hop and Rap on the other. The model also agrees with a measure of musical influence inferred from a large human curated musical website, `allmusic.com`. Finally, it reveals interesting connections between influence and innovation.

2 The Problem Setup

Influence relations in the corpus of popular music have complex structure. Musical influence can be modeled at a hierarchy of levels, ranging from a sound segment – like an electronic distortion, to individual songs, to albums, to artists and musical bands. The relation between these levels is "soft": many songs are created in collaboration by several artists, many artists take part in several bands, and many songs were published in several versions, sometimes spanning a few decades. Finally, a well known thorny issue is that there exist no consistent metadata system which contains the above information for all music, and mapping music across metadata systems is hard. With these considerations in mind, we chose to model influence on the basis of individual songs, since a song is typically a clearly delineated unit in terms of its acoustic data and metadata.

A second critical design choice is about the scope of influence. An artist may be influenced by an artist, or by an individual song. A single song may influence many artists, or even originate a musical style. Here we model influence as a process where one song affects the "musical language" of a musical stream, or "topic". Such an approach was previously taken for modeling how one text document may influence an entire topic [2]. This song-to-topic approach is expected to generalize better than direct song-to-song modeling, since it allows to control the model complexity by the number of topics.

This idea of song-to-topic influence hinges on the basic idea of topic modeling: each song has a distribution across a set of genres, and influences an entire topic (i.e. genre), in proportion to its membership in that topic. The goal of the model is to assign this song-level topic-influence score, and is described in detail in Section 4. In our model we use only the acoustic data of a song, along with its year of release. We do not use any metadata such as genre, leaving this kind of information for validating out model.

3 Data and Features

We use the *Million Songs Dataset* [3] (MSD), a large scale, diverse and epoch-spanning dataset of songs. For this work we sampled 25K songs by 9222 artists. We biased our sample to include a larger portion of earlier songs since our model revolves around modeling historical trends, and since the dataset itself skewes heavily towards later years. Songs were divided into 28 time epochs, with all songs of the same epoch treated as concurrent.

Topic models were originally conceived for textual data, where each document is represented as a bag-of-words [4]. Music however, is naturally represented as a single continuous variable, with structure on multiple time scales from less than a millisecond to the entire song length. To convert the continuous acoustic signals into a dictionary of discrete musical-signature, we applied a widely-used two-stage procedure: First, we extract short time-scale features on the scale of 0.25-1 seconds; then we quantize them using K-means into 5000 groups. The clusters formed by K-means are treated as *musical-words*, and the histogram of their occurrence in a song gives us a bag-of-words representation. We also added long time-scale features such as tempo and key. All the raw audio features are provided by The Echonest. Overall each song is described by a bag-of-words with a vocabulary size of 5033.

4 Modeling Influence

Contemporary music has a strong "topic-like" structure in the form of musical genres (like hip hop or electronic), but at the same time, it exhibits nearly endless mixtures and interactions between genres. There is a clear sense of temporal evolution within and between these genres, which is fundamental to modeling influence [5, 6].

To capture these structures we use the *Dynamic Topic Model* (DTM) [1] and the *Document Influence Model* (DIM) [2]. The model consists of three interacting layers, with inference performed jointly. First is a classical topic layer applied to each time epoch separately. Second is a *time-dependent* layer: Each topic evolves with time, tying different epochs together. Finally there is the topic-dependent *influence* factor. Each song is seen as trying to "pull" future songs of its topic in its direction.

We treat each song d as comprised of a set of N_d *musical words*, $w_1^d, \dots, w_{N_d}^d$ taken from a vocabulary of size W. These words reflect both local and global audio structure, and are discussed in Section 3. Each song belongs to one of T time epochs, and we posit K topics. The model assigns a single topic k from $1 \dots K$ to the word w_n^d, as indicated by the variable $z_{n,k}^d$. In addition, the model assigns to each song a scalar normally distributed topic-influence score l_k^d controlling how much the topic k should later drift in direction of song d. The following relations define the probabilistic model that we use: For each epoch t and topic k the probability distribution of the words is governed by a W-dimensional parameter vector $\beta_{k,t}$. The probability distribution is: $p(w|\beta_{k,t}(w)) \propto exp(\beta_{k,t}(w))$. The temporal evolution of the topic-word distribution vectors β_t^k is given by $\beta_{k,t+1}|\beta_{k,t} \sim \mathcal{N}\left(\mu_{k,t}, \sigma^2 I\right)$, where σ^2 controls the rate of the topics' evolution, and : $\mu_{k,t} = \beta_{k,t} + exp(-\beta_{k,t}) \sum_d l_k^d \cdot \kappa(t, \tau(d)) \sum_n w_n^d z_{n,k}^d$, where l_k^d is each song's **topic-influence score**, $\tau(d)$ is the time of song d and $\kappa(t, \tau(d))$ is a kernel function controlling the time-decay of the influence scores. Each epoch evolves from a starting point that is the sum of two components: the topic's distribution in the previous time-epoch, and the sum of the songs in the previous epochs, scaled by their influence score and a time-delay kernel.

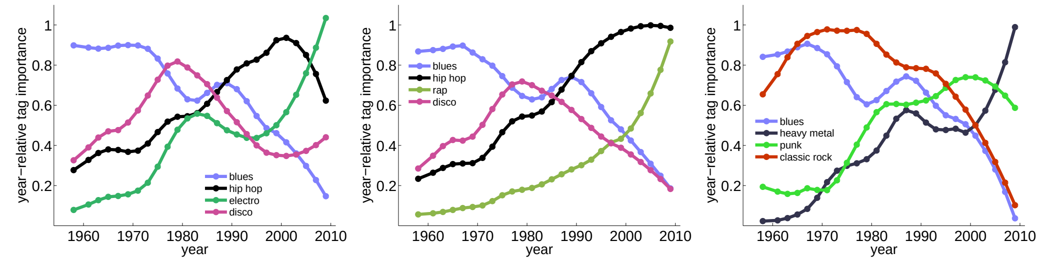

Figure 1: Top genre tags for 3 topics from the 20-topic model, for the years 1957-2010. The topics were chosen to reflect several different genres. The genres come from artist metadata partly available in the Million Songs Dataset, and were not used in training or selecting the model

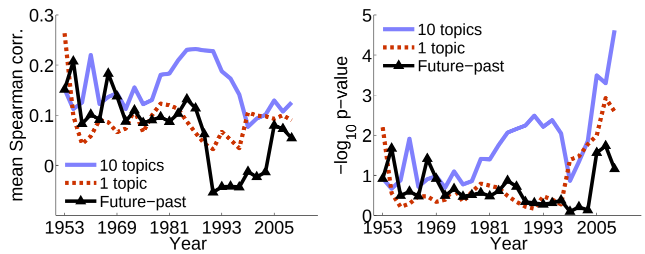

Figure 2: (left) The Spearman correlation and (right) the negative log p-value of the Spearman correlation across different epochs with `allmusic.com`'s influence rank for 10- and 1- topic models, and the future-past baseline described in Section 5. The highly significant p-values for the later years are possible because of the much larger number of songs available for those years.

Computing the posterior distribution of the topics and influence scores for this model is intractable. Gerrish & Blei [2] present a variational approximation and derive an algorithm for maximizing a lower bound on the marginal probability of the observed data. We used their code available at `code.google.com/p/princeton-statistical-learning`.

The variables of interest are the topic-influence scores l_k^d, and the topic assignments of each song. Together, they define the topic mixture of a song and how much it influences each of the topics. The influence of each song is defined as $l^d \equiv max_k l_k^d$ (using the mean across topics gives similar results). We set the time-kernel κ to a log-normal distribution.

5 Results

We applied 1, 5, 10 and 20-topic models to the songs data. We first looked at the matching of the topics with known musical genres. Our data includes 4803 genre tags, with a median of 36 tags per artist. The genre-tags are weighted to indicate the strength of the genre-artist association. For each learned topic we summed up the artist-genre scores weighted by the topic proportions of each song. We found that the topics broadly match widely accepted genres such as metal, electronic & hip hop, especially for the later years where the dataset is larger and more varied. We then investigated the temporal evolution of the topics. Figure 5 shows the genre tags associated with 3 of the topics of a 20-topic model, from 1957-2009. Note that these genre tags were in no way used in training or selecting the model.

A few examples of songs and artists found influential by the model include: Bob Dylan's famous songs "Rainy Day Women #12 & 35", and "Like a Rolling Stone". They also include the Beastie Boys' song "Paul Revere" and Run D.M.C.'s "Is It Live" from 1986, both considered very influential early hip hop / rap groups, as well as songs by Juan Atkins (Model 500), whose work is regarded by the musical encyclopedia `allmusic.com` as "perhaps the most influential body of work in the field of techno". More examples are available in the full paper and on our website.

The model we learned is unsupervised, and predicts the influence each song has on songs that followed it. To assess the validity and quantify the performance of our model, we compared it with the database of `allmusic.com`. This site includes a graph-like structure indicating artist-to-artist influence determined by human editors, covering 5000 of the ¿9000 artists in our data.

To bring `allmusic.com`'s artist-to-artist relation and our song influence measures in line, we averaged the influence scores of each artist ccording to the model, and compared it with sum of influenced artist as given by `allmusic.com`'s influence graph. We found that according to the `allmusic.com` influence measure, earlier artists tend to be much more influential than later artists, making overall comparisons of influence mostly time related. We thus evaluated our influence mea-

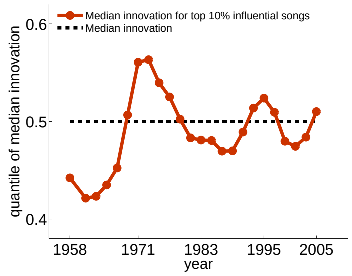

Figure 3: Median innovation of top 10% most influential songs, 1958-2005. In each epoch we standardize the median innovation to 0.5 (dashed black line). Median innovation for the most influential songs is below median at earlier years, and above median at the early 70's and mid 90's.

sure separately for each time-epoch, and averaged the results. The mean Spearman correlation across epochs between the scores of a 10-topic DIM and the `allmusic.com` data is $0.15 \ (p < 0.05)$. Figure 2 plots the per-epoch Spearman correlations and their respective negative log p-values for the 10- and 1-topic DIM, and a baseline method explained below.

As a baseline for comparing the DIM performance, Gerrish & Blei [2] proposed an easy to calculate heuristic influence measure. In this baseline, each word is assigned a weight for each epoch by: $w_t = \frac{\text{Frequency of } w \text{ in } [t, t+f]}{\text{Frequency of } w \text{ in } [t-b, t]}$, where f and b denote the time windows into the future and past respectively. The influence of each song is then the mean over w_t. The mean Spearman correlation with `allmusic.com`'s dataset is $r = 0.07, p > 0.05$, maximized over f and b; see Fig. 2.

6 Musical Innovation and Musical Influence

Having established a valid computational model of musical influence, we looked into the relation between influence and innovation. We used DIM to obain a measure of innovation, defined as the likelihood score of each song, since more innovative songs will be harder to account for by the model, and thus assigned a lower likelihood. Innovation is relative to the past, so to measure the innovation of a song from 1960, we use a model fitted using only songs up to 1960. We call this measure *time-restricted likelihood*.

To validate that *time-restricted likelihood* correlates with innovativeness, we surveyed the *least* likely songs from each time-epoch, as well as a comparable random selection of songs from the dataset. We found that 17 of 27 least likely songs are from artists or albums described in music literature as innovative or "experimental" during the relevant period. For example, Grandmaster Flash, considered by `allmusic.com` to be "Hip-hop's greatest innovator". For random songs, only 8 out of 27 were considered innovative, 6 of them from the earlier periods of the dataset up to 1970. Time-restricted likelihood also correlates well with other measures of innovation such as the use of rarer musical-words relative to the epoch. We thus refer to time-restricted likelihood as innovation score.

We further controlled for two trends: 1. Influence scores decline over the years. Since later songs have not yet had the chance to manifest their influence. 2. Overall innovation scores increase with time. This is a result of the dataset including more songs and more diverse songs in later years. We address these two trends by standardizing both the influence and innovation scores per each epoch, using order statistics.

We find that overall there is no monotonic correlation between the influence and innovation scores (Spearman $r = -0.019, p > 0.05$). However, we see that the relation fluctuates over the years. Fig. 3 shows the median innovation score for the 10% most influential songs in each epoch. We measure how innovative were the most influential songs, with innovation and influence both measured relative to the period. We see two periods when influential songs tended to be more innovative: the early 70's, and the mid-90's. The rise at the mid-90's stems mainly from electronic and hip-hop artists who were given both high innovative *and* high influence scores; examples include Cypress Hill, Outkast, Tricky & Mad Professor. All are considered both original and influential by critics.

References

[1] D.M. Blei and J.D. Lafferty. Dynamic topic models. In *Proceedings of the 23rd international conference on Machine learning (ICML)*, pages 113–120. ACM, 2006.

[2] S. Gerrish and D.M. Blei. A language-based approach to measuring scholarly impact. In *ICML*, 2010.

[3] T. Bertin-Mahieux, D.P.W. Ellis, B. Whitman, and P. Lamere. The million song dataset. In *ISMIR*, 2011.

[4] D.M. Blei, A.Y. Ng, and M.I. Jordan. Latent dirichlet allocation. *JMLR*, 3:993–1022, 2003.

[5] F. Holt. *Genre in popular music*. University of Chicago Press, 2007.

[6] F. Fabbri. A theory of musical genres: Two applications. *Popular Music Perspectives*, 1:52–81, 1982.