

# Python による音源分離

- 情報数理体験講座 2021 の報告 -

## Sound source separation using software Python

-Report on information mathematics workshop 2021 -

清水 優祐\*

SHIMIZU, Yusuke\*

概要：プログラミング言語“Python”は、統計的機械学習に関するパッケージが充実しており、大量のデータを必要とする Web アプリケーションの開発や統計処理など、様々な分野で用いられている。そこで本報告書では、城西大学理学部数学科が主催している「高校生向け情報数理体験講座 2021」において扱った、統計的機械学習を利用した音源分離ライブラリ“spleeter”に関する講演内容を紹介する。Google Colaboratory を使用することで、Google アカウントに紐づいた個々の実行環境での操作が可能となり、ソースファイルや音声ファイルの共有が容易である。体験講座の参加者のアンケート結果も報告する。

### 1. はじめに

内閣府、文部科学省および経済産業省は、文理を問わず全ての大学・高専生が、正規課程にてリテラシーレベルの数理・データサイエンス・AI を取得することを目標としており、大学等の卒業単位として認められる数理・データサイエンス・AI 教育プログラムのうち、優れたプログラムを認定する制度を創設した（注1）。こういった背景もあり、データサイエンスを専門学べるような学部・学科・コースが、ここ数年の間に数多く誕生した。城西大学も、2020年4月に経済学部にてデータサイエンスコース、2021年4月に数理・データサイエンスセンターが発足し、データサイエンス教育プログラムがスタートしている。そのカリキュラムは各教育機関によって様々であるが、データ解析の手法を学ぶ統計教育は、数理・データサイエンス・AI の根幹をなす必要不可欠なものであるため、内容をより一層充実させていくべきである。その一環として、統計解析ソフトウェアを用いたデータ解析およびプログラミングに関する授業内容は、方法論を体系的に学ぶ実践的な技能を身につける上

\* 城西大学理学部数学科

(注1) [https://www.kantei.go.jp/jp/singi/ai\\_senryaku/suuri\\_datascience\\_ai/pdf/ninteisousetu.pdf](https://www.kantei.go.jp/jp/singi/ai_senryaku/suuri_datascience_ai/pdf/ninteisousetu.pdf)

でも極めて重要であると思われる。統計解析ソフトウェアには様々なものがあるが、近年では“Python”が人気であり、複数の検索エンジンから、対象となるプログラミング言語がどれだけ話題になっているかをインデックス化した PCI

(Programming Community Index) では、2021年10月にC言語とJava言語を押し退けて1位となった(注2)。Pythonは、YouTube、Instagram、Dropboxといった様々なアプリケーションやWebサービスで使用されている動的プログラミング言語であり、高度に最適化されたコンパイラとライブラリにより、十分な速度で実行することができる。城西大学では、坂戸・紀尾井町の両キャンパスにPython 3系(2022年1月現在の最新)がインストールされており、授業等で使用することが可能である。そこで本報告書では、城西大学理学部数学科が主催している、「高校生向け情報数理体験講座2021」において、Pythonに関する講演を行った際の内容を報告する。2章では、講座の実習で使ったPythonのプログラミングコードと、参加者のアンケート結果を掲載している。まとめと今後の展望は3章で述べる。

## 2. 体験講座の報告

2021年11月28日(日)の13:00-14:50に、城西大学紀尾井町キャンパスで開催された、城西大学理学部数学科主催の「高校生向け情報数理体験講座2021」で、「AIで楽曲から歌声を取り出そう」というタイトルで実習形式の講演を行った(図1)。本体験講座は、数学の理論がどのように実社会で応用されているのかを、プログラミングコードの実装を通して体験してもらうことを主眼としている。タイトルにある通り、今回は楽曲から歌声を抽出する音源分離という手法に焦点を当てた。音源分離により、音声データ内の不要な雑音の除去や、複数の音源から生じた音声から目的の音声のみを抽出することなどが可能となる。また、深層学習を用いることにより、分離精度を上げることができ、近年盛んに研究されている。音源分離は、SiriやGoogleアシスタントに代表されるvoice user interfaceや、オンラインミーティングソフトのノイズカット機能、イヤフォン補聴器、音楽や動画の編集・加工などで幅広く用いられている。ここで、音源分離のメカニズムについて簡単に説明する(注3)。音は波(音波)で伝わるため、振幅と周波数(周期の逆数)で定量化される。振幅は音の大きさを表し、単位はdB(デシベル)が使われる。人の日常会話はおおよそ50dBから60dBである。周波数は音の高さを表し、単位はHz(ヘル

(注2) <https://www.tiobe.com/tiobe-index/>

(注3) 体験講座での平易な説明と差異がないよう、ここでは非常にラフな記述になっていることに注意されたい。

ツ) が使われる。人の声の周波数はおおよそ 100Hz から 1000Hz であり、ギターはおおよそ 80Hz から 5000Hz, ピアノはおおよそ 27.5Hz から 15000Hz, ベースはおおよそ 40Hz から 400Hz, ドラムはおおよそ 20Hz から 15000Hz である。通常、音は色々な音波が重なって観測される。そこで音波を三角関数で表すと、三角関数の和を考える必要がある。図 2 は、 $\sin x$ ,  $2 \sin 2x$ ,  $\frac{1}{2} \sin 3x$  と、この 3 つの波形を重ね合わせた波形である。重なり合った波形を個々の波形に分けるためには、フーリエ変換により角周波数が異なる三角関数の和に分解する必要がある。詳細は省略するが、これを行うための主な手法として、ブラインド音声分離法とニューラルネットワークを用いる方法がある。特に後者は、音声認識にも応用される方法である。音源分離を Python で行う際には、音源分離ツール “spleeter” を用いる。以下は体験講座の実習で使用した Python のプログラミングコードである。コードの実装には、Google Colaboratoty を用いた。Google Colaboratoty は、ブラウザから Python を記述・実行できるサービスであり、Google アカウントに紐づいた環境であるため、煩雑な環境構築が不要というメリットがある。ソースファイルを共有することで、例えば自宅の PC での作業も可能であり、授業でも導入しやすいと思われる。ただし、GUI は作成できないため、その場合は、例えば Python プログラミングでよく使われる開発用ソフトである、Jupyter Notebook から実行する必要がある。体験講座の参加者にはあらかじめ Google アカウントを取得してもらい、Google drive に公開したソースファイルにアクセスし、Google Chrome 上で実行させた。なお、体験講座で使用した楽曲ファイルは、フリー楽曲サイト <https://dova-s.jp> のものを使用した。

1. 動画・音声ファイルを、再生・変換・保存する ffmpeg をインストール

```
!apt install ffmpeg
```

2. 音源分離ライブラリ spleeter をインストール

```
pip install spleeter
```

3. 音声を再生する Audio 関数をインポート

```
from IPython.display import Audio
```

4. 楽曲ファイル ○○.mp4 (○○はファイル名) を Google Colaboratoty 上にアップロード

5. 元の楽曲を聴いてみよう

```
Audio('〇〇.mp4')
```

6. 音源分離を実行（デフォルトではボーカルと伴奏の2音源に分離）

```
!spleeter separate -o output/ 〇〇.mp4
```

7. ボーカルを聴いてみよう

```
Audio('output/〇〇/vocals.wav')
```

8. 伴奏を聴いてみよう

```
Audio('output/〇〇/accompaniment.wav')
```

9. さらに多数の音源に分離するため Separator モジュールをインポート

```
from spleeter.separator import Separator
```

10. ボーカル・ドラム・ベース・その他の4音源に分離する

```
separator_4stem = Separator('spleeter:4stems')  
separator_4stem.separate_to_file("〇〇.mp4", "./output/4stems")
```

11. ドラムとベースとその他を聴いてみよう

```
Audio('/content/output/4stems/〇〇/drums.wav') #ドラム  
Audio('/content/output/4stems/〇〇/bass.wav') #ベース  
Audio('/content/output/4stems/〇〇/other.wav') #その他
```

12. ボーカル・ドラム・ベース・ピアノ・その他の5音源に分離する

```
separator_5stem = Separator('spleeter:5stems')  
separator_5stem.separate_to_file("〇〇.mp4", "./output/5stems")
```

13. ピアノとその他を聴いてみよう

```
Audio('/content/output/5stems/〇〇/piano.wav') #ピアノ  
Audio('/content/output/5stems/〇〇/other.wav') #その他
```

城西大学  
東京紀尾井町キャンパス  
高校生向け  
情報数理  
体験講座 2021

# 数学とコンピュータを 組み合わせてできること



日時 2021年 **11月28日(日)**  
13:00~14:50 (12:30 受付開始)

場所 **城西大学 東京紀尾井町キャンパス**  
3号棟 3401教室 東京都千代田区平河町 2-3-20

定員 **先着30名** **入場無料**

参加をご希望の方はお早めに、QRコードよりお申し込みください。

新型コロナウイルス感染症予防対策を十分とった上で実施いたします。  
当日必ずマスクの着用をお願いします。発熱等症状のある場合は、参加をお控えください。



## ACCESS MAP

- 東京メトロ半蔵門線 半蔵門駅1番出口より徒歩3分
- 東京メトロ有楽町線 麹町駅1番出口より徒歩5分
- 東京メトロ半蔵門線・南北線 永田町駅9番出口より徒歩7分



お問い合わせ先

城西大学 東京紀尾井町キャンパス入試・広報センター  
TEL 03-6238-1111 E-mail nyushik4@ju.ac.jp

図 1: 情報数理体験講座 2021 のポスター

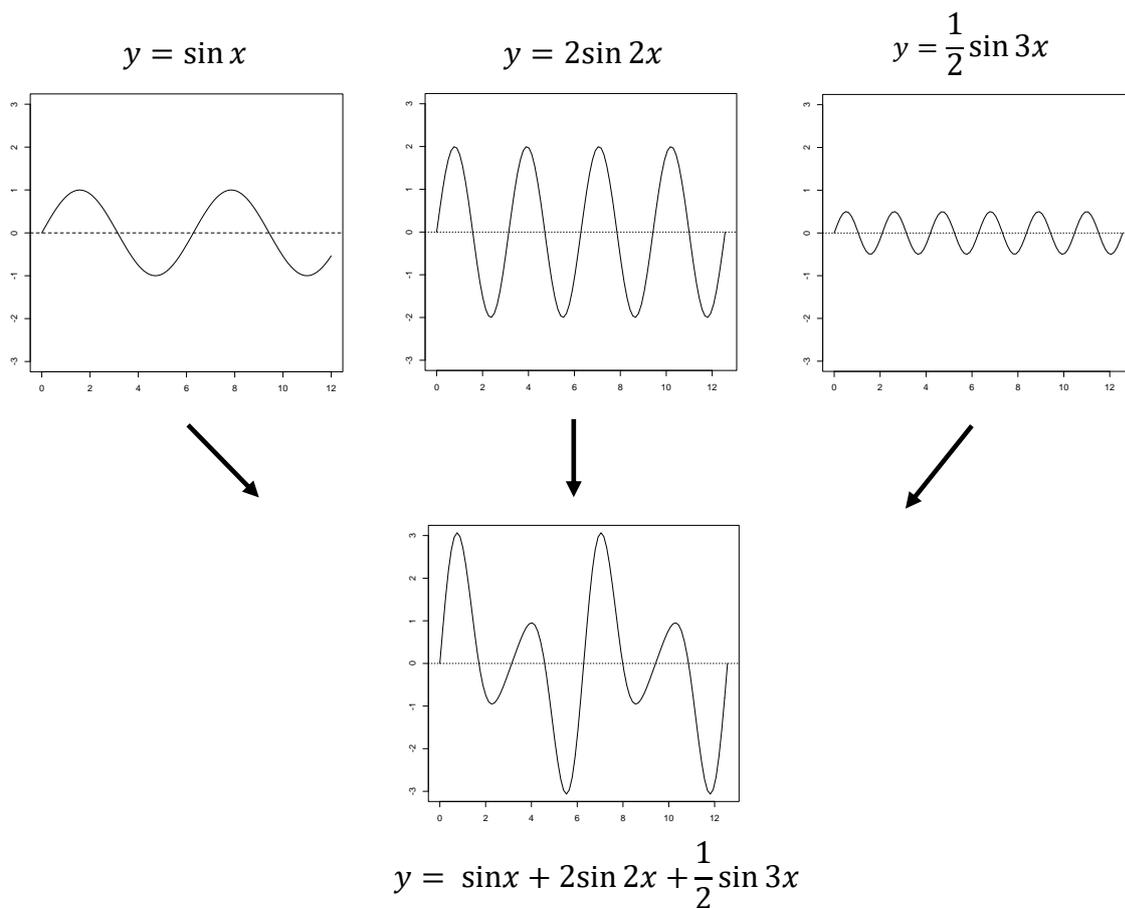


図 2: 音波の重ね合わせを表すイメージ図。上段が個別の 3 音源, 下段が観測音を表している。

アンケートの回答者は 5 名おり, “面白かったかどうかを 5 点満点で評価してください” の問いには,

- 4 点 : 結構綺麗に音源を取り出せせてすごいと思った。ボカロを入れたらどうなるのかが気になった。
- 4 点 : 面白い内容で、自分でもやってみたいと思った。
- 5 点 : 楽器ごとに音声を分離できると知らなかったのが非常に興味深い内容でした。もし自分が軽音部だったらそのアプリケーションを使って練習しやすくできるなど、幅広い分野に応用できそうで凄いな、と思いました。貴重な体験ありがとうございました！
- 5 点 : スマホの編集アプリで歌声を取り出すものがあるが、仕組みを知らなかったのが、今回の講演で少しわかった気がして楽しかった

- ・4点：私は昔から音楽に触れて生活してきたので、今回の講座で音楽についてより知れたのでとても良かったと感じます。本日はどうもありがとうございました。

といった意見があり、参加者の興味を惹くことができたと考えられる。なお、上記のアンケート結果は原文のまま掲載している。

### 3. まとめ・今後の展望

理学部数学科が主催している、「高校生向け情報数理体験講座 2021」で、「AIで楽曲から歌声を取り出そう」というタイトルで実習形式の講演を行った際の、プログラミング言語 Python によるコード実装の内容を、ソースコードとともに紹介した。また、三角関数の重ね合わせを例にして、音波の解析について簡単に説明した。参加者のアンケート結果より、Python を用いたプログラミング実習は、学生の興味を惹く内容であることが分かった。特に、楽曲の音源分離に関するテーマは、スマートフォンなどで手軽に音楽を聴くことができることもあり、実用例も豊富で意欲的に取り組みやすい題材であると言える。導入が容易である Google Colaboratoty を用いて教材を作成することで、統計解析ソフトウェアを用いたデータ解析およびプログラミングに関する授業内容として、初年次からでも取り入れることができるであろう。今後も数理・データサイエンス・AI 教育の一環として、Python に関するより質の高い講演や講義を行っていきたい。

#### 謝辞

体験講座の運営でお世話になりました高山晴子先生、同じ講演者の立場としてアドバイスをいただきました藤田昌大先生、実習補助の梅宮マオさん、川内陽平さん、重原多久斗さん、山田祥子さんには、心から感謝いたします。

#### 参考文献

- [1] 篠田浩一, 機械学習プロフェッショナルシリーズ:音声認識, 講談社, 2017
- [2] 高橋遼一, 機械学習実践シリーズ:Python で学ぶ音声認識, インプレス, 2021
- [3] 戸上真人, 機械学習実践シリーズ:Python で学ぶ音源分離, インプレス, 2020