

東京大学新領域創成科学研究科  
環境システム学専攻

令和3年度  
修士論文

異常状態検知のための  
機械学習による大学実験室における音の検知

2021年7月21日提出  
指導教員 大島 義人 教授

周 睿

# 目次

<b>第一章</b>	<b>緒言</b> .....	<b>4</b>
1.1	大学実験室の事故予防 .....	4
1.2	音からの異常状態の検知 .....	5
1.3	大学実験室の異常音検知の考え方 .....	6
1.4	異常音検知に関する研究例 .....	8
1.5	本研究の目的 .....	9
<b>第二章</b>	<b>研究方法</b> .....	<b>10</b>
2.1	音の取り方について .....	10
2.2	収録音の前処理 .....	11
2.2.1	音響信号のモノラル化 .....	11
2.2.2	プリエンファシス (pre-emphasis) .....	11
2.2.3	フレーム化処理と窓関数 .....	11
2.3	音響特徴量 MFCC .....	14
2.4	混合ガウスモデル (GMM) .....	16
2.4.1	混合ガウスモデル (GMM) の原理 .....	16
2.4.2	混合ガウスモデル (GMM) のパラメーター推定 .....	16
2.5	音のモデル化と外れ値検出 .....	19
2.5.1	音のモデル化 .....	19
2.5.2	外れ値検出 .....	20
2.5.3	聴覚評価 .....	21
<b>第三章</b>	<b>実験結果</b> .....	<b>23</b>
3.1	音のモデル化による音の検知 .....	23
3.1.1	個別の音の録音結果 .....	23
3.1.2	MFCC 特徴量の抽出結果 .....	25
3.1.3	GMMに基づく機械学習 .....	28
3.1.4	GMMモデルの学習効果の検証 .....	28
3.1.5	交差検証の結果 .....	33
3.1.6	まとめ .....	35
3.2	外れ値検出による音の検知 .....	36
3.2.1	全体の音の録音結果 .....	36
3.2.2	いつもの音のモデル化 .....	36

3.2.3	いつもと違う音の抽出結果.....	43
3.2.4	作られたいつもと違う音の検出結果.....	44
3.2.5	リアルタイムモニタリングによる異常音検出の試み .....	47
3.2.6	聴覚評価実験.....	49
3.2.7	まとめ .....	53
<b>第四章</b>	<b>結言 .....</b>	<b>54</b>
	<b>参考文献 .....</b>	<b>56</b>
	<b>謝辞.....</b>	<b>58</b>

# 第一章 緒言

## 1.1 大学実験室の事故予防

大学のような研究現場において、研究者の創造性と積極性を損なうことのないような安全・安心な実験作業環境が求められている。しかし、各大学の実験室をはじめとする様々な研究現場からの事故が年々絶えずに報告されており、事故による経済的損失や人的被害が出てしまっている。このような実情を踏まえ、大学の実験室における事故の発生を避け、または被害を軽減する対策を真剣に取り組む必要がある。

事故予防については、すでに産業界などを対象に、安全管理の方法論が立てられている。あるシステムにおいて、既に発生した故障を引き起こす全ての故障原因およびその経路としての故障モードを演繹的に分析する故障の木分析 (Fault Tree Analysis, FTA) という手法もあれば、作業や製造プロセスの中の潜在的な故障モードおよびその影響を帰納的に分析する故障モードと影響解析 (Failure Mode and Effects Analysis, FMEA) という手法もある。これらの手法は産業界において活用されており、事故予防に必要なものとして使われている。しかし、工場などの定まった生産・製造プロセスに沿って行われる定常作業に対し、大学実験室の場合は研究者が自ら実験設計をし、行われる実験も日々変わっていると考えられる。非定常作業の多い実験室において、起こりうる故障やその原因が日々変わっている可能性もあるため、産業界の安全管理手法をそのまま実験室に持ち込むことは現実的ではない。

現在の大学実験室の事故予防については、一般的に過去の事故報告書を参照し、各実験室における事故の起こりうるところを一つ一つ挙げるか、産業医現場巡視により事故事例を基に点検し対策を講ずることでリスクを排除しているか、または教員が過去の経験や実験マニュアルで学生である実験者に安全への重視をさせている。いずれも過去の事故から整理した事故原因から、その個別性の高い事故を予防することしかできない。大学実験室のような創造性が求められる研究現場において、今までにない新しいアイデアで考えられた実験プロセスや更新されていく新しい実験装置があるため、今までにないリスクが存在し、今までにない事故が起こる可能性があると考えられる。一方、実験室の実情を分かっているほどの人には正常化バイアスを持っている可能性が高いと考えられる。そのため、実験室のことをよく分かっている人だからといってリスクへの認知能力が高いとは限らない。これは例として、実験中に揮発性の薬品が漏れ出る危険性が低いことで、教員をはじめとする研究者が火災や爆発の危険性に比べて化学物質への暴露による有害性を軽視しがちだということが分かった<sup>[1]</sup>。

以上のことを踏まえ、大学実験室の事故を予防するために、過去の経験に依存しない、

特に未知のリスクや事故を予測・検知できるような一般的な手法が必要と考えられる。既に、産業界においては、事故防止の観点から異常検知の概念を採用し、画像や音の情報から、いつもと違う状態を検出し、事故や欠陥の排除などが行われている<sup>[2]</sup>。異常検知の手法では、過去に起こった異常を基に同じような異常を検知することができ、また、今まで起こったことのない異常への検知も可能であると示唆されている。これにより、異常検知の観点から、大学実験室の特徴を考慮した事故予防システムの開発の必要があると考えられる。

## 1.2 音からの異常状態の検知

大学実験室の異常を検知するにあたり、産業界の既存の異常検知手法を参考し、異常を検知するよう適合した情報を選択する必要がある。既に画像の情報から、火災や油漏れ等の異常を検知できるインテリジェント監視システムが活用されており、長時間の画像の時間変化量を学習して異常の兆候を判断し、事故を防止できる<sup>[3]</sup>。しかし、実験室の異常は画像だけでは検知されない場合が多い。よく事故につながるボンベなどの開閉状況、遠心分離機の異常振動などの異常状態は画像情報から検知されにくいであろう。ましてや、非定常作業の多い実験作業において、画像情報だけで何が異常状態なのかを簡単に定義することができないため、画像から正しい実験行動を学習しても誤った実験行動を検出することは難しいと考えられる。一方、実験者が実験をする際の実験作業、会話、歩行と、実験機器や換気装置稼働などによって様々な音が出て、実験室の音環境に構成している。このような音の情報から、画像では検知できない異常を検知できると期待される。

産業界に着目した研究によると、既に異常音を知覚することによって機械の保守・点検において重要であることが明らかになった。ここで例を示すと、寺島<sup>[4]</sup>は石油業において聴覚による異常発見件数が 20.9%の割合を占めており、五感による異常発見の中で聴覚が二位であることを示した。豊田ら<sup>[5]</sup>は、機械設備の点検において、特に軸受や歯車からの異常を検出する際、異常な振動と音響が特に優れたパラメーターであることを示した。中村<sup>[6]</sup>は、とある爆発火災の事故事例を分析し、爆発炎上の 14 分前に異音が確認されたことにより、音のような異常現象を早めに気づけば、重大事故を回避できる可能性があるという結論を示した。

実験室の場合、異常音検知による事故回避についての研究はまだされていないが、過去の事故事例を振り返ると、異常音と事故の関係が大きいと示されている。ここで例を示すと、田村<sup>[7]</sup>は、化合物の不安定さにより、減圧蒸留中に黒煙と高いポンプ音が発生し、すぐに加熱を止めたため事故に至らなかったが、普通以上の注意をしないと、爆発事故にな

るかもしれない。また、錯体のエバポレーターにおける水を除去する際に、副反応により温度が上がったため、大音量の爆発音とともに容器が粉碎され、反応物が天井まで飛ばされた。このような実験中における事故直前の音で、爆発音、轟音、警報音などの事故につながる音が少なくないことを事件事例から読み取れた。川泉<sup>8)</sup>は、過去の実験室の事故と危ない体験の概要を整理した。そこから、いくつかの実験事故から異常な音が発生したことが分かった。オイルバスでの加熱において、バシッという音が出たとともにガラス管が破裂して実験者の手を傷つけた。エタノールと硝酸の混合において、操作の過ちによって混合物の瓶が爆発音とともに破損した。アルケンの水素添加において、レギュレーターを開けているままで元栓を開けたとともに、キーンという甲高い音が出て水素が漏れた。他にも、熔融塩の突沸や有機蒸気の爆発などの事故も紹介されており、特に音について言及されていないが、いつもと違いそうな音が出るだろうと予想できる。

以上のような実験室における異常音を検知できれば、実験室における異常状態を検知することができる。それにより、多くの音に関係する事故の発生を予知し、時を移さず対策を実施すれば事故の発生を回避でき、または事故による被害を軽減できると考えられる。そのため、実験室の特徴に合わせた実験室における異常音を検知する方法を開発することが重要だと考えられる。

### 1.3 大学実験室の異常音検知の考え方

これまで実験室において多くの音に関係する事故が発生してしまってきたが、実験室の異常音に注目する研究は基本的になかった。ただし、廖<sup>9)</sup>は大学実験室においてどのような音があるかを明らかにされ、またその音の特徴をスペクトログラム分析という手法で解明され、異常音検知への応用の可能性を示した。スペクトログラム分析により音のスペクトログラムにおける特徴で、音の時間長短、特定周波数帯域、倍音、ピーク、スペクトログラム形状などの特徴が明らかになる。このような手法では、いちおう音を定量的に計算し、図に出力してはいるが、そのスペクトログラム特徴を明らかにするとき、基本的にスペクトログラム図を見ながら論述しているため、定性的な分析方法である。しかし、これからの実験室における異常音検知方法の開発においては、定性的な方法より、定量的な方法が望ましい。その理由は、異常音はよく出ない音のため、異常音が出る前に何の異常音が出るかは分からない状態で異常音検知をしていると考えられる。つまり、異常音のスペクトログラムにどのような特徴があるかを明らかにすることは難しい。また、大学実験室における異常音検知は、実験室に人がいなくても異常音を検知できることを目標とし、最終的に人間の聴覚の代わりに、リアルタイムモニタリングというような形で異常音を検知

する機能を持つべきであるため、人間にとっては定性的な特徴が分かりやすいかもしれないが、検知する側である機械などにとっては、定量的な特徴を使う方が一般的であろう。なので、大学実験室における音の定量的な特徴を明らかにできる方法を開発することがこれからの課題として研究する必要がある。

異常音検知の方法を検討する前に、何よりなのは異常音の定義を明らかにしなければならない。異常音の定義を明確にしなければ、その検知手法も定義によって変わってくるからである。一般的に、異常音は事故につながる音と考えられるが、異常状態を検知する場合、異常状態は必ずしも事故に行かないため、異常状態における音も必ずしも事故につながる音に聞こえないと考えられる。異常音は異常状態を反映するため、その異常状態に前提を設ける必要がある。音源の少ない、または鳴っている音の種類が限られている場合で、例えば装置の少なく、無人でリモート操作がよく行われている実験室では、異常が出るものは基本的に装置そのものであり、異常音もその装置が正しく稼働していない音である。この場合では、装置の正しく稼働している音と正しく稼働していない音を学習する。検知する際に、正常な音は正常と検知され、異常な音は異常と検知される。こういう手法の利点は、検知される異常音の種類まで分かる。つまり、その異常音から何の異常が出たかが分かる。しかし、その欠点は、普段異常音は出ないため、異常音の収集がとても困難で、収集不完全の場合は音が誤認識される可能性がある。一方、音源の多い、しかも鳴っている音が混じり合っている場合、上記の手法は難しい。例えば、一部の装置が正しく稼働しているにもかかわらず、他の一部の装置が正しく稼働していない場合、正常な音があるため、正常音として検知される可能性があるが、明らかに異常状態でありながら検知されないのは目的からずれてしまう。また、全ての装置が正しく稼働していても、人の行動などによって今までにない音が出て、それにより何か事故が起こるかは知らないが、「いつもと違う」状況にあるため、実験者や担当責任者が注意し、確認する必要があるため、異常状態として検知されるべきである。この場合では、装置などのモノが正しく稼働していない音を検知するではなく、この「いつもと違う」音を検知することである。この手法の利点は、異常音の収集を行う必要なく、いつもと違う音があれば異常音として検知できる。欠点としては、異常音が検知されてもそれは何の音なのかを知るために事後確認が必要である。また、異常のしきい値を設ける必要があるため、設けられたしきい値によって正常な音でも異常音として検知されることがある。

今まで産業界や医学界において異常音検知に関する研究がたくさんなされており、異常音検知の考え方について上記の両方も検討されている。大学実験室の場合、各研究室の研究テーマや使っている装置機器によって、各実験室の音環境も一概に論じることができない。本研究は両方の視点から、簡単な音環境と複雑な音環境の実験室における異常音検知

方法を検討する。

## 1.4 異常音検知に関する研究例

上野<sup>[10]</sup>らは、機械学習による機械の正常・異常音判定の有効性を確認した。異常音の収集が困難のため、機械の正常音だけでなく、異常の動作音も発生させた。収集したデータに高速フーリエ変換を行うことで正常音と異常音の波形を図にし、異常音の周波数成分の特徴を得た。その結論は、正常音と異常音に 100Hz を超えた周波数成分では著しい違いが見られた。

江原<sup>[11]</sup>らは、異常音を入手することが困難ということから、教師なし学習を用いた検出システムを提案した。音の特徴量を抽出し、それによって音の正常・異常を判断するシステムである。音の特徴量をメル周波数ケプストラム係数とし、機械学習モデルを One Class SVM とした。実験はベアリングを回転させ、正常音をとり、そしてベアリングに負荷を与えることで異常音を作った。正常音と異常音を機械学習モデルに入れたあと、外れ値として判断するパラメーターを調整することで異常音を外れ値として検出した。

大瀧<sup>[12]</sup>らは、車内における異常音の発生と流出を防止することを目標し、異常音を定量値として判断する技術とその機能を論述した。まずは異常音と正常音の違いについて、音の生波形だけで判断できないため、生波形の特定の帯域を二乗平均処理することで強調した波形におけるあるレベルでのしきい値を設けることで、それを超えた面積がある値に超えたら異常音として判断する。開発されたシステムを生産ラインに導入した結果、異常音に対する検出力が著しく上がったということである。

Yang<sup>[13]</sup>は、飛行機のエンジンからの異常音を認識する方法を開発した。飛行機エンジンの正常と異常な稼働音と事故中の音を取得し、飛行機エンジンの音響信号の MFCC と GFCC パラメーターを音響特徴量とし、GMM-UBM 機械学習モデルを使用し、音響特徴量の種類および次元数、GMM コンポーネントなどの指標が異常音検知結果への影響について論述した。また、ディープラーニングを用いた異常音検知の応用についても論述した。

Li<sup>[14]</sup>らは、人の聴覚に頼らないラウドスピーカーからの異常音を検知する方法を開発することを目標とし、短時間フーリエ変換という手法で取得したラウドスピーカーの異常音信号の時間周波数領域を若干の部分に分割し、それぞれの部分の時間周波数領域の行列を抽出した。抽出した行列が正常サンプルのラウドスピーカーの行列との距離によって異常音が存在するかどうかを確認した。この方法により、99%以上の不合格ラウドスピーカーが検出されたと示した。



## 1.5 本研究の目的

本研究は、大学実験室における異常状態を検知するために、その異常状態を反映するいつもと違う音を、異なる音環境における機械学習による定量的検出方法を検討し、最終的に大学実験室リアルタイムモニタリングという視点から異常音を検出する仕組みの開発に必要とする基礎的知見を検討する。具体的には、実験室における音の音響特徴量を抽出し、機械学習モデルを用いて音をモデル化するとともに、音のいつもの音から学習したモデルに対する対数尤度の外れ値検出によって実験室全体音からいつもと違う音の検出手法を検討する。また、外れ値検出に使われる評価指標と人の聴覚的評価との関係を明らかにし、今後の研究において活用される可能性を検討する。

## 第二章 研究方法

### 2.1 音の取り方について

本研究では、本学の化学工学系の実験室を対象としている。当実験室では、主に超臨界流体を使用し、高温高压に伴う合成・分解反応の実験が行われている。遠心分離機、超音波洗浄機、水素炎イオン化検出器（GC-FID）、高速液体クロマトグラフ（HPLC）、ヒュームフードなど複数の装置が日常的に稼働しており、実験用の流し台、薬品庫、冷蔵庫、作業機などの実験者の操作による音が出る箇所がある。なお、普段同時に実験を行っている実験者は2-3人であり、実験者による歩行音と会話音も存在する。

音を収録するためにマイクロフォン UC-59（リオン）を使用し、収録した音を波形収録装置 SA-A1（リオン）を通じてSDカードで保存した。マイクロフォンは、なるべく実験者のじゃまにならないように配慮し、実験台、もしくは収録する音源のそばに置いて録音した。

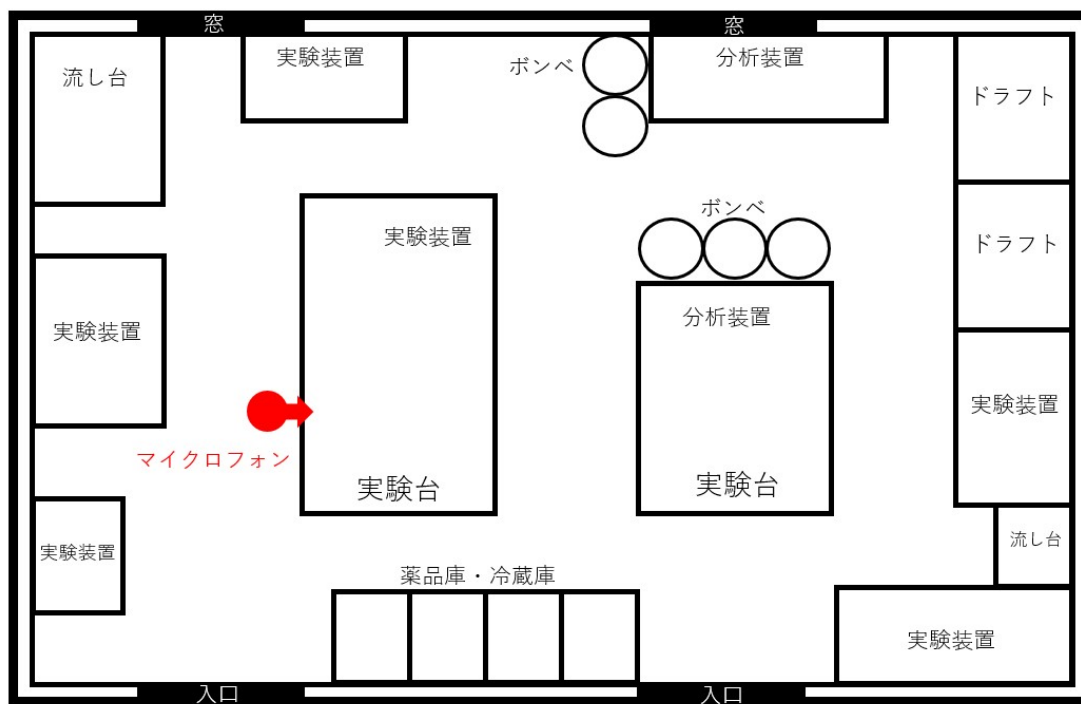


図1, 実験室の平面図およびマイクロフォンの置く場所

## 2.2 収録音の前処理

収録した音を定量的に分析するために、そのまま生の音データを利用するのではなく、あらかじめ処理することが必要である。それぞれの音の持続時間と振幅が違うため、前処理を行わないと、音が混在し、折り返し雑音や歪みによって区別できなくなる。それからの影響を小さくするため、音響特徴量の抽出の前一步として、前処理を行う必要がある。

### 2.2.1 音響信号のモノラル化

マイクロフォン UC-59 で収録した音響信号はステレオである。けれども、音響信号の特徴を計算するとき、モノラルの音響信号を用いるのが望ましい。そのため、ステレオ信号をモノラル信号に変えることが必要である。ここでは、Format Factory というソフトウェアでステレオの音響信号をモノラルにした。

### 2.2.2 プリエンファシス (pre-emphasis)

日常生活における音は主に低周波数と中周波数の帯域に存在しており、実験室のような回転装置の多い場所では特にそうである。音響信号が空気のような媒質の中に伝送するとき、周波数が高いほど、媒質による音のパワーの減衰が著しくなる。また、音響信号が伝送路のローパス特性による高周波数成分における減衰も避けることができない。そのため、音響信号の高周波数成分を強調する必要がある。

通常ではフィルタを用いてプリエンファシスを行う。プリエンファシスのインプットとアウトプットの関係は以下の式で示される。

$$signal_{out}(n) = signal_{in}(n) - a * signal_{in}(n - 1)$$

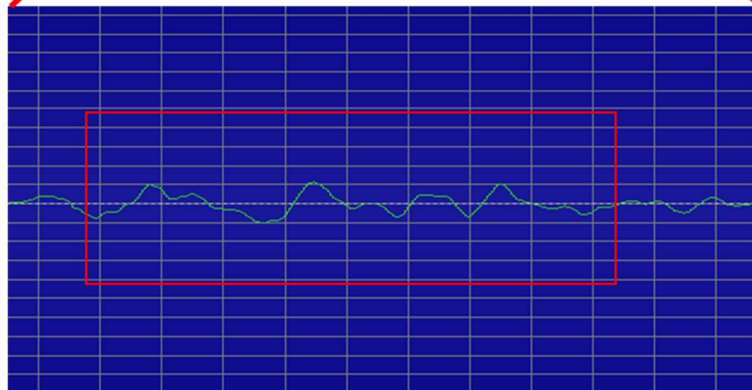
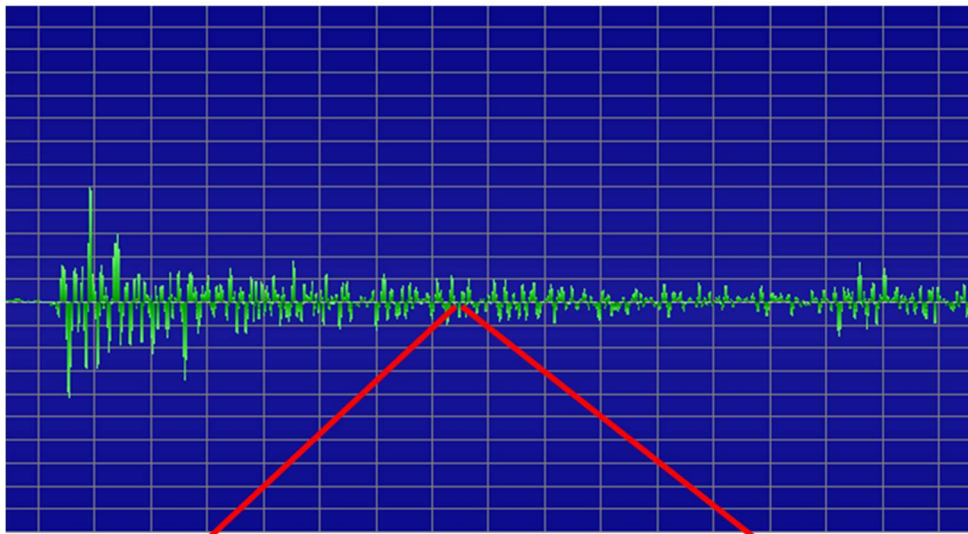
上式において、 $signal_{in}$ は入力信号、 $signal_{out}$ は出力信号、 $a$ はプリエンファシスの係数であり、よく 0.9–1.0 の値が取られている。本研究では 0.97 とした。

### 2.2.3 フレーム化処理と窓関数

音響信号を利用するために、信号の周波数領域における分布を明らかにすることが必要とされている。よく使われている方法は信号に対してフーリエ変換を行うことである。しかし、フーリエ変換は信号が安定することを前提としている。実験室において、定常に稼働している機器による定常な音は存在するが、多くの場合は非定常音の方が多いと考えられるため、長い時間領域における音響信号は不安定で、フーリエ変換を行っても意味のある結果が得られない。とはいえ、ある程度短い時間領域において、音響信号は安定すると考えられるため、短時間フーリエ変換を行うことで音響特徴量の抽出が可能になる。これ

を説明するために、下の図2に示されている。横軸は時間、縦軸は振幅である。

## 不安定な音



## 安定と見なせる

図2, 音の長いと短い時間領域における波形

この理由で、音響信号にフレーム化処理を行うことが必要である。フレーム化処理は、音響信号を一個一個の短い部分を分割し、短時間処理を行う。その短い部分の長さはよく10-30msに取られている。本研究では25msとした。

分割したフレーム信号に対してフーリエ変換を行う前に、数値解析を容易くするために、窓関数を乗じること。その目的はフレーム信号の幅が両端においてゼロに近づかせることでスペクトルの相互影響によるずれを最小限にできるが、その代わりに、信号の両端の部分が弱められてしまう。それによる影響を軽減するため、フレームを連続的に分割するのではなく、重ね合わせるようにすること。重ね合わせによる時間差は10msまたはフレ

ームの長さの半分にとられている。本研究では10msとした。

窓関数を選択するにあたって、両端の勾配を緩めにするのが重要である。よく使われている窓関数は、矩形窓、ハミング窓、ハンニング窓などがある。この中に、ハミング窓が最も窓の両端を緩めのできるため、本研究の窓関数はハミング窓を使用した。

## 2.3 音響特徴量 MFCC

生の音データを機械学習などに用いられるよう一般的に音響信号の特徴量を抽出してそれを学習している。音響信号には様々な特徴量が存在し、目的や明らかにしたい音響的特徴によって選択すべき特徴量も違ってくる。音の異常らしさを判断するのは最終的に人間の聴覚のため、なるべく人間の聴覚に一致する特徴量を使うことが望ましい。

人間の聴覚は周波数ベースの非線形システムである。音の周波数が 1000Hz 以下では、人の聴覚は周波数と線形関係であり、音の周波数が 1000Hz 以上では、人の聴覚は周波数と対数関係である。こういう聴覚特性を説明するために、Davies らは 1980 年メル周波数ケプストラム (Mel Frequency Cepstral Coefficients, MFCC) の概念を提示した。メル周波数ケプストラムはメル周波数尺度から抽出されるケプストラムパラメーターであり、メル尺度は人間の聴覚の非線形特性をあらわしている特徴量のため、本研究において適合した特徴量であると考えられる。メル周波数 (mel (f) /Hz) と周波数 (f/Hz) の関係は以下の式と図で示される。

$$mel(f) = 2595 * \lg\left(1 + \frac{f}{700}\right)$$

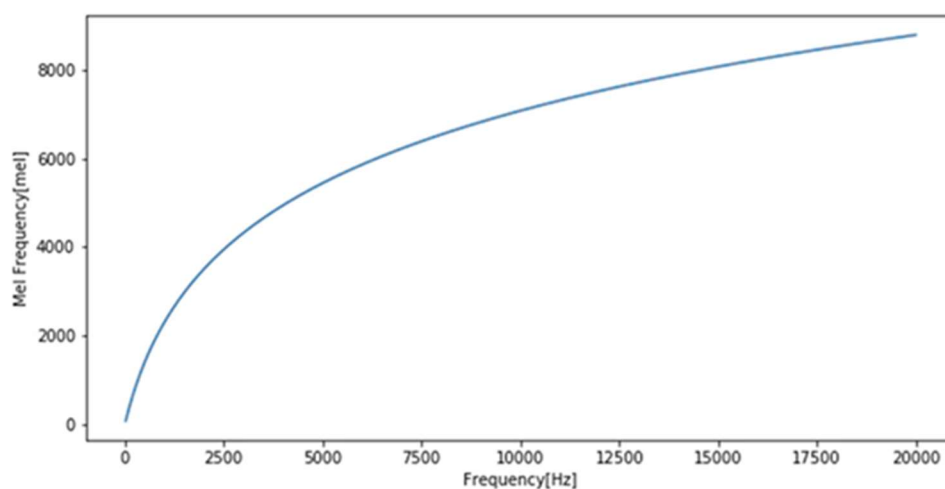


図 3, メル周波数と周波数の関係

メル周波数ケプストラム係数 (MFCC) は、よく音声認識において信頼性の高い評価指標として使われているが、他の音の検知のためにもよく使われている。伊藤<sup>[15]</sup>は、様々な異常音検出手法をまとめ、それぞれの検出手法において使われている特徴量を挙げた。この中に、家庭や医療現場におけるガラス音や転倒音、公共空間における爆発音、銃声および叫び声などへの検知のためにも MFCC を使われており、異常音の検知のためにも信頼

性の高い特徴量として考えられる。

抽出手順としては、音響信号に対して離散フーリエ変換を行い、得られたパワースペクトルにメル尺度の重ね合わさったフィルタバンクを通らせ、フィルタバンクから出力されるベクトルが得られ、離散コサイン変換を行うと 13 次元のメル周波数ケプストラム係数アレーが得られる。

以下の図において、MFCC の時間表現が示されている。横軸は時間（フレーム）であり、縦軸は 13 次元の MFCC 係数であり、1 フレームごとに 0.01 秒における 13 次元の MFCC アレーが可視化されている。柱が黄色いほど、その次元における MFCC 係数が正で、その絶対値が大きくなる。逆に柱が黒いほど、その次元における MFCC 係数が負で、その絶対値も大きくなる。

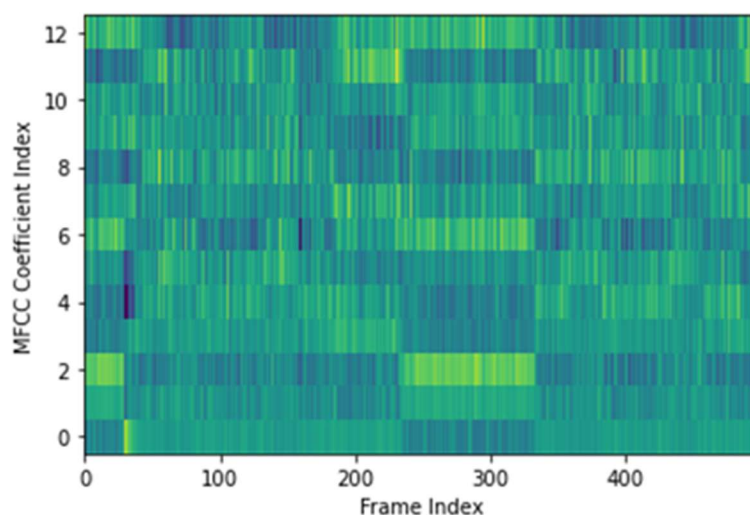


図 4, MFCC-時間表現

最後に、抽出された MFCC アレーを標準化 (standardization) する必要がある。データごとに平均を引き、標準偏差で割ると、データは平均が 0、分散が 1 の標準正規分布に従属する。データを標準化することで、データがある区間に限定される。次に機械学習を行うとき、データが使いやすくなる。ここで注意すべき点としては、学習データとテストデータを分割してから標準化すること。そうしないと、学習データとテストデータが互いに影響し合うため (data leak)、最終的にモデルには偏差が出てしまう。

そこで本研究では、メル周波数ケプストラム係数 (MFCC) を実験室の音の特徴量とした。MFCC の抽出は、Python3.8 で行った。

## 2.4 混合ガウスモデル (GMM)

### 2.4.1 混合ガウスモデル (GMM) の原理

前述した MFCC 音響特徴量を定量的に分析するには、機械学習を行うことで可能になる。一般的に、適合した機械学習モデルを選択する必要がある。音の異常検知において、混合ガウスモデル (Gaussian Mixed Model, GMM) がよく使われている。

混合ガウスモデル (GMM) は、複数のガウス分布の線形結合であり、その確率密度関数はM個のコンポーネント密度の重み付け和からなるものである。理論上では混合ガウスモデルは任意の分布をフィッティングすることができる。その式が以下のように説明される。

$$p(X|\lambda) = \sum_{i=1}^M \omega_i N(X|\mu_i, \Sigma_i)$$

$$N(X|\mu_i, \Sigma_i) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_i|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (X - \mu_i)^T \Sigma_i^{-1} (X - \mu_i) \right\}$$

中に、 $N(X|\mu_i, \Sigma_i)$  は混合モデルにおける i 番目のコンポーネントという。X は d 次元の特徴ベクトルである。 $\omega_i$  はその i 番目のコンポーネントの重み係数という。 $\mu_i, \Sigma_i$  は平均ベクトルと共分散行列である。ただし、

$$\sum_{i=1}^M \omega_i = 1 \quad 0 \leq \omega_i \leq 1$$

の条件を満たす。

以上のことから、混合ガウスモデルは重み係数 $\omega_i$ 、平均ベクトル $\mu_i$ と共分散行列 $\Sigma_i$ の三つのパラメーターで決定されることが分かる。故に、混合ガウスモデルは、

$$\{\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_M\}$$

$$\lambda = \{\omega_i, \mu_i, \Sigma_i\}, \quad i = 1, 2, 3, \dots, M$$

のパラメーターの集合とすることができる。

### 2.4.2 混合ガウスモデル (GMM) のパラメーター推定

あるデータがどのような分布に従属しているかを明らかにするには、上記のパラメーターを推定する必要がある。

GMMのパラメーター推定において、ある潜在変数 $Z$ を導入することでパラメーターを



推定する。潜在変数 $\mathbf{z}$ は以下の式に満たす。

$$p(z_k = 1) = \omega_k \quad 1 \leq k \leq M$$

$$z_k \in \{0,1\}$$

つまり、 $z_k = 1$ の時、 $k$ 番目の分布が選択された。逆に $z_k = 0$ の時、 $k$ 番目の分布が選択されなかったという仮定条件を設けられた(事前に何番目の分布が選択されたかは分からないため)。

$z_k$ が独立同分布であると仮定する場合、事前確率 $p(\mathbf{z})$ は以下の式に満たす。

$$p(\mathbf{z}) = p(z_1)p(z_2)p(z_3) \dots p(z_k) = \prod_{k=1}^M \pi_k^{z_k}$$

また、データは全てガウス分布に従属するため、

$$p(X|z_k = 1) = N(X|\mu_k, \Sigma_k)$$

なお、 $z_k$ もガウス分布に従属するため、

$$p(X|\mathbf{z}) = \prod_{k=1}^M N(X|\mu_k, \Sigma_k)^{z_k}$$

これにより、事前確率 $p(\mathbf{z})$ と潜在変数 $\mathbf{z}$ の条件付き確率(尤度)が説明されたため、事後確率 $p(\mathbf{z}|X)$ が求められる。ここで、 $k$ 番目の分布の事後確率を $\gamma(z_k)$ と表示する。

$$\begin{aligned} \gamma(z_k) &= p(z_k = 1|X) \\ &= \frac{p(z_k = 1)p(X|z_k = 1)}{\sum_{j=1}^M p(z_j = 1)p(X|z_j = 1)} \\ &= \frac{\omega_k N(X|\mu_k, \Sigma_k)}{\sum_{j=1}^M \omega_j N(X|\mu_j, \Sigma_j)} \end{aligned}$$

これにより、 $k$ 番目の分布の事後確率 $\gamma(z_k)$ が表示されたことで、アルゴリズムでパラメータ推定ができるようになる。

GMMのパラメータ推定において、EMアルゴリズムというのが使われている。複数の反復により最尤のパラメータの値を推定するアルゴリズムである。EMアルゴリズムには、EステップとMステップがあり、Eステップではパラメータの期待値を推定し、Mステップでは、Eステップで求められたパラメータの期待値を最大化し、パラメータが収束するかどうかを確認する。パラメータが収束するまでこのように反復するア

ルゴリズムが EM アルゴリズムまたは期待値最大化法 (Expectation Maximization Algorithm) と呼ばれている。

上記の求められた事後確率 $\gamma(z_k)$ を使用し、EM アルゴリズムによるパラメーターの推定は以下のようになる。

1. 分布の数を  $M$  個と仮定し、それぞれ  $k$  番目の分布の  $\omega_k$ 、 $\mu_k$ 、 $\Sigma_k$  の初期値を設定し、 $p(X) = \sum_{k=1}^M \omega_k N(X|\mu_k, \Sigma_k)$  の対数尤度関数を計算する。
2. E ステップ:  $\omega_k$ 、 $\mu_k$ 、 $\Sigma_k$  を基に、 $\gamma(z_k)$  を計算する。
3. M ステップ: E ステップで計算された  $\gamma(z_k)$  により、新しい  $\omega_{nk}$ 、 $\mu_{nk}$ 、 $\Sigma_{nk}$  の値を計算する

$$\mu_{nk} = \frac{\sum_{n=1}^N \gamma(z_k) X_n}{N_k} X_n$$

$$\Sigma_{nk} = \frac{1}{N_k} \sum_{n=1}^N \gamma(z_k) (X_n - \mu_{nk})$$

$$\omega_{nk} = \frac{N_k}{N}$$

ここでは、 $N$  はデータ点の数であり、 $N_k = \sum_{n=1}^N \gamma(z_k)$

4. 上記の新しいパラメーターを  $p(X)$  に代入し、対数尤度関数を計算する。
5. 対数尤度関数が収束するかどうかを確認する。もし収束しない場合、E ステップに戻り、反復を続ける。<sup>[16]</sup>

## 2.5 音のモデル化と外れ値検出

本研究では、混合ガウスモデル (GMM) を用いて音の検知をしているが、実は使われている実験室によって、その詳しい手順も変わってくる。

### 2.5.1 音のモデル化

音源が少なく、かつ鳴る音がほぼ一定である音環境において、異常音の種類も一定であり、簡単な音環境である。この場合では全ての音に対してGMMモデルを通じてモデル化し、次に音を検知する際に、ある種類の音が鳴ると、その音の分布がその種類の音から学習したGMMモデルの分布に従属していると検知される。ここで、例えばもし学習した音の種類が三つで、つまりGMMモデルが三つの場合、ある音をどの音かを検知したいとき、その音がその三つのGMMモデルそれぞれの対数尤度を計算し、尤度比検定を通じ、最も高い対数尤度を持つGMMモデルの音の種類が、検知したい音の種類と一致すると、検知が成功となる。

尤度比検定について、GMMモデルの尤度の高いほど、その音がそのモデルの分布に従属している可能性が高いと考えられる。尤度によって音がどの分布に従属するかを判断する。尤度比検定では、よく尤度の自然対数、対数尤度が使われている。尤度比検定を説明するために、Ayoub Malek<sup>[17]</sup>は男性と女性の音声認識において、対象となる音声は男性か女性からの音かを検知するのに尤度比検定の概念が使われた。ある音声は男性の音声である対数尤度は $p(X|H_m)$ で、女性の音声である対数尤度は $p(X|H_f)$ とした。音声の尤度比検定は以下の式である。

$$\frac{p(X|H_f)}{p(X|H_m)} = \begin{cases} \geq 1 & \text{accept } H_f \\ < 1 & \text{reject } H_m \end{cases}$$

以上の式により、二つの音声モデルから音声を検知できる。二つ以上の場合、最も高い対数尤度で音を検知すると考えられる。

簡単な音環境の実験室における音の検知の流れは以下になる。この流れにより、正常音と異常音に対してモデル化すれば、音を検知する際に、正常音か異常音を検知されるとともに、その音の種類まで分かるような仕組みである。

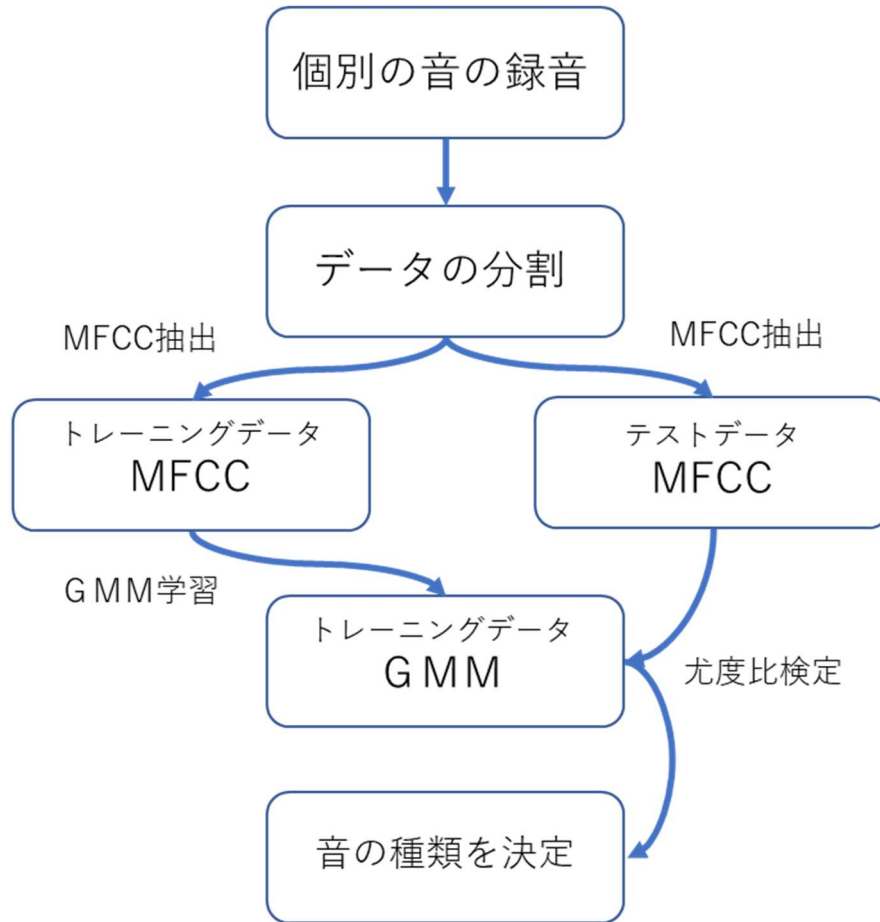


図5, 簡単な音環境の実験室における音の検知の流れ

### 2.5.2 外れ値検出

音源が多く、しかも混じり合っている場合、全ての音に対してモデル化することができない。たとえ一部の音をモデル化しても、モデル化されなかった音が出た場合、既に学習しているGMMモデルのいずれとして誤検知され、結果として意味がない。そもそも、異常音の収集が難しいということは一般的認識である。なので、上記のように事前に音データにタグを付ける、いわゆる教師有り学習の機械学習方法を、複雑な音環境では使用できない。

上記のため、異常音検知を目的として、何の音が異常音を事前に音データにタグを付けることができない。しかし、1.3で言及したとおり、簡単な音環境では装置などの正しくない稼働音を異常音と定義できるが、複雑な音環境では、「いつもと違う音」という視点で異常音を定義したほうが良い。この場合で、装置の正しくない稼働音といい、人の転倒音といい、または今までにない未知の音といい、全部あまり発生していない・発生したことがない音のため、データグループには入っていない、もしくは出現回数が少ない。そのため、機械的には他の音データの分布とは違うような分布になっている音として、「い

つもと違う音」として検知されると考えられる。

この考え方により、外れ値検出という手法が適合と考えられる。前述したとおり、簡単な音環境のある音に対してモデル化すると、その音のGMMモデルはその音の分布を学習する。尤度比検定の思想によって、その音はそのGMMモデルに対して最も高い対数尤度を持つが、他の音はそのGMMモデルに対しての対数尤度が低い。そのため、複雑な音環境において、実験室の全ての音を同時に長時間の録音を行えば、その音はその実験室のいつもの音として考えられ、装置の正しくない稼働音などのあまり発生していない音、または発生したことの無い音が出た場合、対数尤度が明らかに低いと考えられる。そこから外れ値が出ると考えられる。その時、適切な対数尤度のしきい値を設けることで、しきい値より低い対数尤度を持つ音が外れ値になり、いつもと違う音として検知される。たとえそれが正しくない装置の稼働音のような事故につながる異常音でなくても、その実験室にいる実験者や担当責任者には確認すべき音として未知の異常音の発現、もしくは注意を喚起するということにもなるため、実験室の安全管理の上、非常に有意義と考えられる。

### 2.5.3 聴覚評価

音の異常らしさを対数尤度の高さで定義しているが、実際にそれが人間の聴覚との関係はまだ不明である。人間が聞こえる異常音と機械が検出する異常音が一致しているのか、それを検証するために、対数尤度の高い音から低い音まで人間に聴かせ、その音がどれほどいつもと違いそうに、もしくはどれほどいつもと同じように聞こえるのかを明らかにするために、被験者実験を行う。その結果により、人間が聞こえる異常音と機械が検出する異常音が一致しているかを明らかにするとともに、複雑な音環境において、対数尤度を判断基準とした外れ値検出による異常音検知への可能性を検討する。

外れ値検出および聴覚評価実験により、複雑な音環境の実験室における音の検知の流れは以下になる。

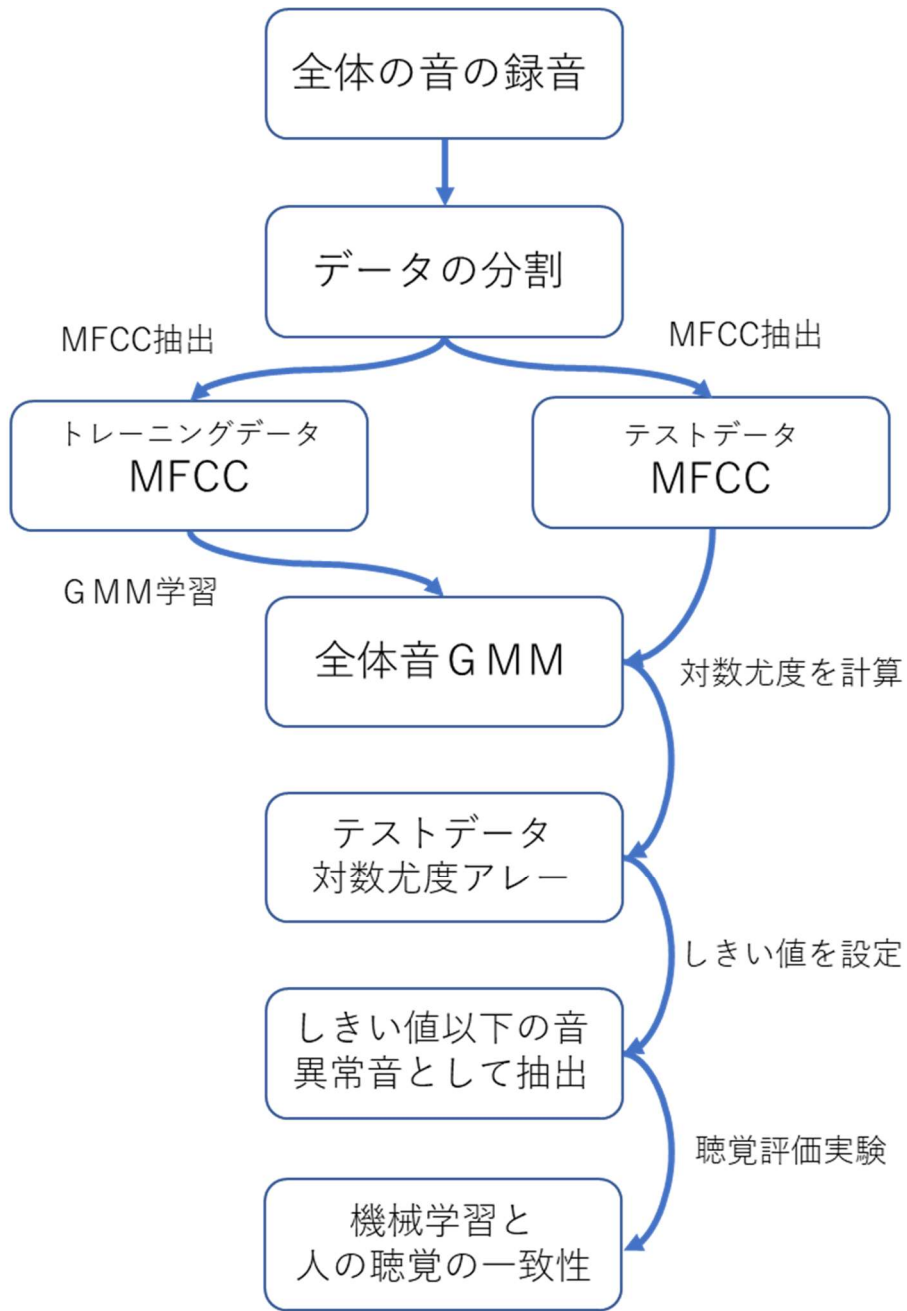


図6, 複雑な音環境の実験室における音の検知の流れ

## 第三章 実験結果

### 3.1 音のモデル化による音の検知

#### 3.1.1 個別の音の録音結果

簡単な実験室の音環境を予想しているため、音を個別に録音した。対象とする音は、簡単な音環境の実験室において主に存在する音、つまり装置の音とした。

本研究では、遠心分離機、水素炎イオン化検出器 (GC-FID)、高速液体クロマトグラフ (HPLC)、ヒュームフードを対象とした。録音する際に、なるべくマイクロフォンを対象の装置に向けて置いている。マイクロフォンの置き方と対象の装置は以下の図に示されている。



図7, 遠心分離機



図8，水素炎イオン化検出器（GC-FID）



図9，高速液体クロマトグラフ（HPLC）

録音は上記の装置の稼働音を録音した。ただし、遠心分離機の録音に際し、その稼働音だけでなく、稼働完了の提示音も録音した。他の装置の録音に際し、特にシナリオを考えないでそのまま録音した。録音された音の時間については以下の表になる。



表 1, 装置ごとの録音時間

音	時間/s
遠心分離機 (稼働音)	200
遠心分離機 (提示音)	80
水素炎イオン化検出器 (GC-FID)	1878
高速液体クロマトグラフ (HPLC)	825
ヒュームフード	606

上記の録音されたデータを学習する前に、データに分割する必要がある。まずは、データを5秒ごとのデータに分割する。それは、機械学習モデルの学習効果を評価する際に、複数のテストデータが必要だからである。分割した後、無作為で一部のデータをテストデータとして抽出し、残りのデータをトレーニングデータとした。トレーニングデータは比較的が多い方にすること。下の表で各装置音のトレーニングデータとテストデータの分配が示されている。

表 2, 装置ごとのトレーニングデータとテストデータ

音	トレーニングデータ	テストデータ
遠心分離機 (稼働音)	35	5
遠心分離機 (提示音)	11	5
水素炎イオン化検出器 (GC-FID)	356	20
高速液体クロマトグラフ (HPLC)	145	20
ヒュームフード	102	20

装置の都合により、遠心分離機の音がなかなか収集が難しかったため、データの量が少なかったが、今回使用するGMMモデルは、小標本でも認識精度が高いとされているため、今回は小標本で機械学習を行うこととした。

### 3.1.2 MFCC 特徴量の抽出結果

データが分割された後 MFCC 特徴量の抽出を行った。MFCC 特徴量は、0.01 秒につき一個の 13 次元のアーレーになっており、5 秒には 498 個 (窓関数をかけるときデータの両

端が重なっていないため) の MFCC 係数アレーがある。ここで例を示すと、以下の図には遠心分離機の提示音であり、0.01 秒の MFCC 係数アレーおよび標準化された後の MFCC 係数アレーが示されている。

```
[-0.50679051  0.58907611  1.78921994 -0.84249689 -1.49274864 -0.45653839
 1.01391968 -0.12568675 -1.1504449  0.32387664  0.36160204 -0.57676178
 1.75625541]
```

図 10, 13 次元の MFCC 係数アレー

```
[-0.83507756  0.97838063  1.31309649 -1.28415918 -1.3239942 -0.60103982
 0.82308214 -0.15076431 -1.13273406  0.32280161  0.44702362 -0.46308591
 1.35588281]
```

図 11, 標準化された 13 次元の MFCC 係数アレー

このような MFCC 係数アレーはデータ一つにつき 498 個があるため、MFCC-時間表現図にした。下の図には遠心分離機提示音および稼働音、水素炎イオン化検出器 (GC-FID)、高速液体クロマトグラフ (HPLC)、ヒュームフードの稼働音 MFCC-時間表現図が示されている。ただし、以下の MFCC-時間表現図は未標準化の MFCC 係数アレーからなっている。

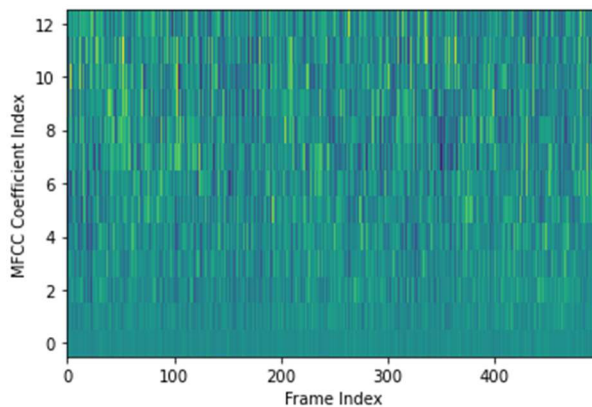


図 12, 遠心分離機稼働音 MFCC-時間表現

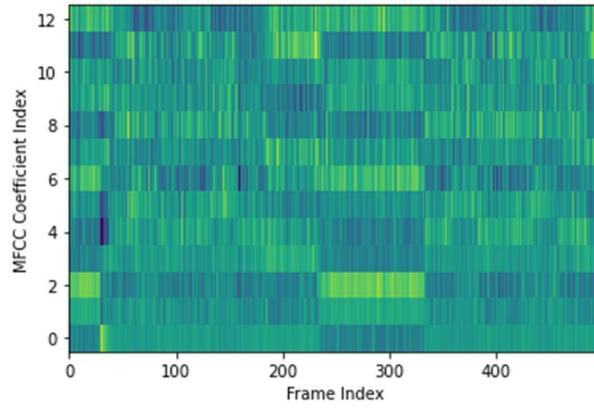


圖 13, 遠心分離機提示音 MFCC-時間表現

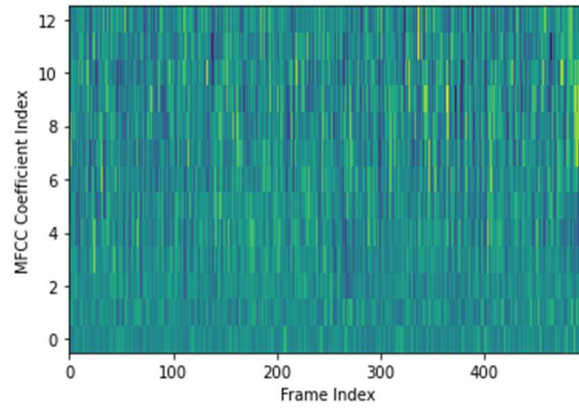


圖 14, GC-FID 稼働音 MFCC-時間表現

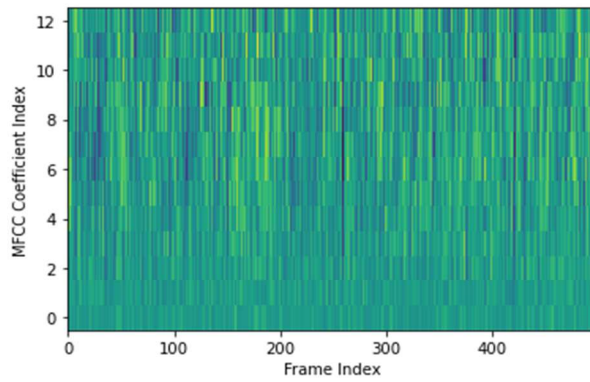


圖 15, HPLC 稼働音 MFCC-時間表現

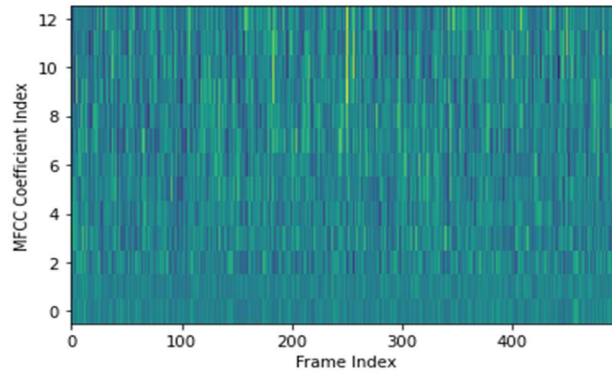


図 16, ヒュームフード稼働音 MFCC-時間表現

人間の聴覚でも聞き分けるように、遠心分離機の提示音の MFCC-時間表現図が他の稼働音の MFCC-時間表現図とは違っていることが分かる。それにもかかわらず、各装置の稼働音の MFCC-時間表現だけを見ていると、それぞれがどの音かを見分けることは難しい。そのため、機械学習モデルに定量的情報を学習させ、音を分類させることで音を検知することを期待している。

### 3.1.3 GMMに基づく機械学習

上記の MFCC 抽出と同様に、GMMに基づく機械学習は python3.8 より行った。表 1 で示されたトレーニングデータで GMM モデルに学習させ、GMM ファイル (.gmm) の形式で出力させた。

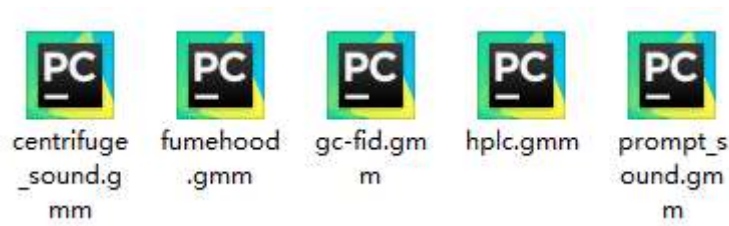


図 17, 学習した GMM モデルのファイル (.gmm)

### 3.1.4 GMM モデルの学習効果の検証

上記のトレーニングデータから学習した GMM モデルとテストデータにより、GMM モデルの学習効果を検証した。

前述したとおり、検証は尤度比検定で行った。それぞれの音のテストデータのそれぞれの GMM モデルに対する対数尤度を計算して比較する。対数尤度が高いほど、その音はそのモデルの分布に従属する可能性が高いと考えられる。最も高い対数尤度を持つ GMM モデルの音の種類が、検知したい音の種類と一致すれば検証できたとする。

以下の表で、遠心分離機提示音および稼働音、水素炎イオン化検出器 (GC-FID)、高速液体クロマトグラフ (HPLC)、ヒュームフードの稼働音のそれぞれのテストデータが、それぞれのGMMモデルに対する対数尤度が示されている。この中、FH (Fume Hood) がヒュームフードの稼働音、CS (Centrifuge Sound) が遠心分離機の稼働音を指し、PT (Prompt Tone) が遠心分離機の提示音を指している。

表 3, 遠心分離機稼働音——CS (Centrifuge Sound) テストデータの各モデルに対する対数尤度

GMMモデル テストデータ	FH	GC-FID	HPLC	CS	PT
1	-17.737	-17.164	-17.360	-16.494	-23.167
2	-17.722	-17.144	-17.467	-16.468	-23.25
3	-18.047	-17.322	-17.642	-16.717	-23.116
4	-17.734	-17.106	-17.401	-16.469	-22.867
5	-17.695	-17.136	-17.441	-16.549	-23.058

表 4, 遠心分離機提示音——PT (Prompt Tone) テストデータの各モデルに対する対数尤度

GMMモデル テストデータ	FH	GC-FID	HPLC	CS	PT
1	-22.435	-20.288	-20.774	-20.327	-12.474
2	-22.397	-20.504	-20.794	-20.343	-12.447
3	-22.462	-20.441	-20.763	-20.352	-12.607
4	-22.295	-20.186	-20.744	-20.250	-12.753
5	-22.698	-20.709	-20.967	-20.652	-12.303

表 5, 水素炎イオン化検出器—GC-FID テストデータの各モデルに対する対数尤度

GMMモデル テストデータ	FH	GC-FID	HPLC	CS	PT
1	-17.443	-16.104	-16.498	-16.775	-23.298
2	-18.252	-15.959	-16.262	-17.294	-23.117
3	-17.405	-16.241	-16.561	-16.679	-23.035
4	-17.619	-16.291	-16.621	-16.772	-23.283
5	-17.553	-16.078	-16.468	-17.131	-23.391
6	-17.534	-16.336	-16.737	-16.782	-23.013
7	-17.761	-16.408	-16.741	-16.913	-23.226
8	-17.561	-16.235	-16.608	-16.791	-23.213
9	-18.166	-16.466	-16.813	-17.498	-23.414
10	-17.827	-16.349	-16.665	-17.022	-23.156
11	-17.536	-16.235	-16.607	-16.826	-23.107
12	-17.642	-16.248	-16.622	-16.864	-23.187
13	-17.795	-16.446	-16.889	-17.338	-23.500
14	-17.975	-16.139	-16.557	-17.260	-22.948
15	-17.800	-16.336	-16.702	-17.026	-23.369
16	-18.041	-16.511	-16.817	-17.405	-23.633
17	-17.128	-15.991	-16.389	-16.855	-23.450
18	-17.254	-16.144	-16.560	-17.183	-23.151
19	-17.850	-15.911	-16.225	-17.307	-23.304
20	-17.318	-16.147	-16.606	-16.890	-23.541

表 5, 高速液体クロマトグラフ—HPLC テストデータの各モデルに対する対数尤度

GMMモデル テストデータ	FH	GC-FID	HPLC	CS	PT
1	-17.557	-16.362	-16.002	-16.937	-23.374
2	-17.655	-16.373	-16.106	-16.852	-23.438
3	-17.614	-16.285	-16.105	-16.922	-23.344
4	-17.506	-16.385	-16.107	-16.868	-23.355
5	-17.847	-16.583	-16.265	-17.294	-23.294
6	-18.214	-16.036	-15.694	-17.290	-22.945
7	-17.818	-16.483	-16.008	-16.910	-23.628
8	-17.971	-16.314	-16.004	-17.276	-23.482
9	-17.592	-16.414	-16.154	-17.139	-23.546
10	-17.556	-16.447	-15.918	-16.933	-23.493
11	-17.790	-16.373	-15.955	-16.936	-23.413
12	-17.560	-16.347	-15.903	-16.859	-23.415
13	-17.291	-16.163	-15.707	-16.650	-23.432
14	-17.604	-16.465	-15.881	-16.961	-23.492
15	-17.608	-16.518	-16.030	-16.991	-23.275
16	-17.680	-16.606	-16.409	-17.065	-23.463
17	-17.455	-16.273	-15.830	-16.927	-23.405
18	-17.477	-16.409	-16.168	-17.012	-23.499
19	-17.474	-16.394	-15.927	-17.040	-23.537
20	-17.398	-16.342	-16.007	-16.912	-23.228

表 5, ヒュームフード—FH (Fume Hood) テストデータの各モデルに対する対数尤度

GMMモデル テストデータ	FH	GC-FID	HPLC	CS	PT
1	-15.758	-16.953	-17.215	-16.918	-23.474
2	-15.784	-17.093	-17.351	-17.009	-23.596
3	-15.912	-17.220	-17.481	-17.143	-23.393
4	-15.939	-17.069	-17.318	-17.063	-23.461
5	-15.899	-17.216	-17.412	-17.064	-23.518
6	-15.677	-17.021	-17.192	-16.936	-23.592
7	-15.712	-16.869	-17.176	-16.815	-23.393
8	-15.856	-17.073	-17.437	-16.998	-23.433
9	-15.799	-16.960	-17.175	-16.896	-23.542
10	-15.909	-17.014	-17.227	-16.998	-23.534
11	-15.666	-16.925	-17.194	-16.905	-23.393
12	-15.748	-16.986	-17.290	-16.816	-23.665
13	-15.902	-17.027	-17.318	-16.965	-23.515
14	-15.887	-17.046	-17.301	-16.834	-23.414
15	-15.581	-16.819	-17.108	-16.751	-23.524
16	-15.677	-16.996	-17.254	-16.888	-23.538
17	-15.694	-17.069	-17.331	-16.989	-23.527
18	-15.911	-17.131	-17.404	-17.057	-23.636
19	-15.649	-16.874	-17.206	-16.733	-23.441
20	-15.709	-16.868	-17.212	-16.902	-23.338

上記の表で示されているように、遠心分離機稼働音の場合、全ての遠心分離機稼働音のテストデータが遠心分離機稼働音のトレーニングデータから学習した GMM モデルに対する対数尤度が最も高く、他の四種の音のトレーニングデータから学習した GMM モデルに対する対数尤度がより低くなっている。これは、遠心分離機提示音、水素炎イオン化検出器 (GC-FID)、高速液体クロマトグラフ (HPLC)、ヒュームフードの稼働音の場合でも同じである。つまり、全ての音がその音から学習した GMM モデルに対する対数尤度が最も高く、学習効果が良いと考えられる。



### 3.1.5 交差検証の結果

上記の結果により、GMM モデルの学習効果を検証した。しかし、それはそのデータの分割からの結果しか表現できない。ということは、このトレーニングデータから学習したモデルの学習効果は良いにもかかわらず、残りの他のデータをトレーニングデータとすると、必ずしも学習効果が良いとは言えない。一部のデータで良い結果が出ていても、残りのデータに良くない結果が出た場合、モデルの汎化性能が低いと認められ、過学習という。

以上のことから、交差検証 (Cross-validation) を行う。交差検証は、トレーニングデータとテストデータを複数回にランダムに分割し、目的はすべてのデータを学習し、過学習を防ぐ。交差検証には、よく K 分割交差検証 (K-fold cross-validation) が使われている。データを K 個の集合に分割し、その中の一つの集合をテストデータとし、残りの K-1 個の集合をトレーニングデータとして学習する。そして、そのテストデータを基に、K-1 のトレーニングデータから学習したモデルの学習効果 (認識率など) を評価する。こういう手順を K 回繰り返させ、最後に学習効果を評価するために、K 個の出力された認識率の平均を評価指標としている。それにより、すべてのデータはもれなくトレーニングデータとなるため、それでも学習効果が良いとすれば、過学習はしていないと考えられる。具体的な K 分割交差検証の手順としては、下の図に示されている。

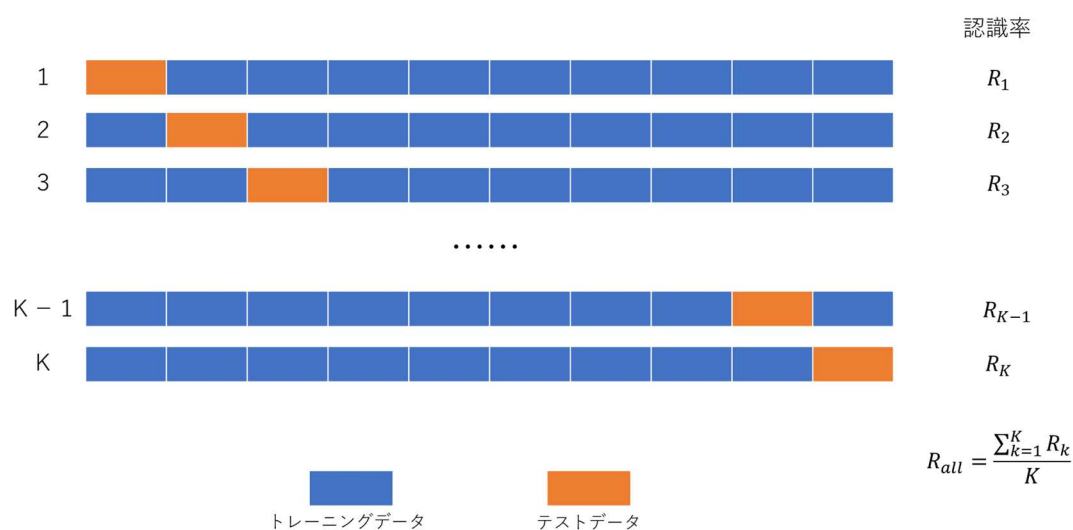


図 18, K 分割交差検証

今回はすべての音のデータに対して 10 分割交差検証を行った。その結果は、すべての音のすべてのテストデータが認識され、全体の認識率が 100% となっている。ここですべての音の 10 分割交差検証の結果を示さないが、水素炎イオン化検出器 (GC-FID) の稼働音の 10 分割交差検証の流れを示す。

1. K 分割交差検証アルゴリズムにおいて分割数を 10 に設定する。テストデータ集合

- の個数も 10 に設定する。
2. GC-FID の音を除き、他の音のデータをすべてトレーニングデータとして GMM モデルに学習させる。ただし、GC-FID の音は 10 個の集合に分割する。
  3. 10 個の集合から一個を引き、残りの 9 個を GC-FID GMM モデルに学習させる。
  4. 一個引いたデータをテストデータにし、GC-FID GMM モデルを含み、5 個の GMM モデルの対数尤度を計算し、比較する。GC-FID GMM モデルの対数尤度が最も高ければ検証できるとする。
  5. 上記の流れを 10 回繰り返す。最後に平均認識率を算出する。
  6. 上記の流れを他の音を対象に繰り返す。最後にそれぞれの音の平均認識率を算出する。

下の表で水素炎イオン化検出器 (GC-FID) の稼働音の 10 分割交差検証の結果が示されている。結果は、10 分割交差検証でもすべてのテストデータの GC-FID GMM モデルに対する対数尤度が最も高いことが分かった。つまり、すべてのデータがトレーニングデータに入っても学習効果が良く、過学習はないと考えられる。

他の音の場合でも同じような結果が示されたが、ここでは省略した。

表 6, 水素炎イオン化検出器—GC-FID 稼働音の 10 分割交差検証結果

GMMモデル 検証回数	FH	GC-FID	HPLC	CS	PT
1	-19.555	-16.195	-18.698	-16.989	-25.094
2	-19.465	-16.160	-18.597	-16.915	-25.051
3	-19.577	-16.221	-18.709	-16.982	-25.056
4	-19.525	-16.182	-18.675	-16.948	-25.053
5	-19.527	-16.197	-18.683	-16.964	-25.080
6	-19.637	-16.235	-18.742	-17.033	-25.115
7	-19.616	-16.211	-18.718	-16.984	-25.053
8	-19.613	-16.221	-18.717	-17.005	-25.113
9	-19.533	-16.181	-18.642	-16.938	-24.991
10	-19.567	-16.214	-18.717	-16.988	-25.126

### 3.1.6 まとめ

3.1 では簡単な音環境の実験室を前提とした実験室における音の検知方法を検討した。MFCC を音響特徴量とした GMM 機械学習モデルにより、小標本データを対象としても、装置の稼働音や提示音などの個別の音に対する検知性能が高いことが分かった。また、交差検証を行うことで過学習の可能性を排除した。以下が学習効果の検証結果になる。この実験室では事故などにつながる異常な音の収集が難しかったため、今回対象とした音は、その意味での異常音ではないが、異常音が収集できる、かつ音環境が簡単な実験室では、この手法のように異常音を含めて音をモデル化することで、異常音の検知が可能と示唆された。

表 7, GMM による音の検知の学習効果の検証

音	トレーニングデータ	テストデータ	認識率	交差検証
遠心分離機 (稼働音)	35	5	100%	100%
遠心分離機 (提示音)	11	5	100%	100%
水素炎イオン化検出器 (GC-FID)	356	20	100%	100%
高速液体クロマトグラ フ (HPLC)	145	20	100%	100%
ヒュームフード	102	20	100%	100%

## 3.2 外れ値検出による音の検知

### 3.2.1 全体の音の録音結果

複雑な音環境において、個別の音に対してモデル化することは難しいため、個別の音源に対して録音するのではなく、実験室全体の音を録音することとした。全体の音の録音に際し、マイクロフォンの置き方は下の図に示されており、特定の装置などに向いているのではなく、実験台に向いており、実験室の全体の音環境を録音した。録音は基本的に 2021 年前期に間欠に行われ、総計 38 時間 43 分の実験室全体の音を入手した。



図 19, 全体の音を録音している様子

### 3.2.2 いつもの音のモデル化

全体の音からいつもと違う音を検出するには、3.1 のように音を学習して GMM モデルによる機械学習を行う。3.1 と違うところは、個別の音に対してモデル化するのではなく、実験室全体のいつもの音に対してモデル化し、そのモデルに対する対数尤度の低い音を異常音としている。長時間の録音によりその音をいつもの音としているため、どれほど長い時間で録音すれば、それをいつもの音といえるかを明らかにしたい。

ここで、手元のデータから、8 時間 43 分のデータをテストデータとし、5 秒ごとに 6345 個データに分割した。ただし、この 8 時間 43 分のデータは連続的に録音されたわけではない。一方、残りの 30 時間のデータの中から、0.5 時間から 30 時間までのデータを抽出してトレーニングデータとして学習し、同じテストデータの GMM モデルに対する対数尤度の分布の違いを明らかにした。ここで、1% の最も対数尤度の低いデータを赤点にし

た。

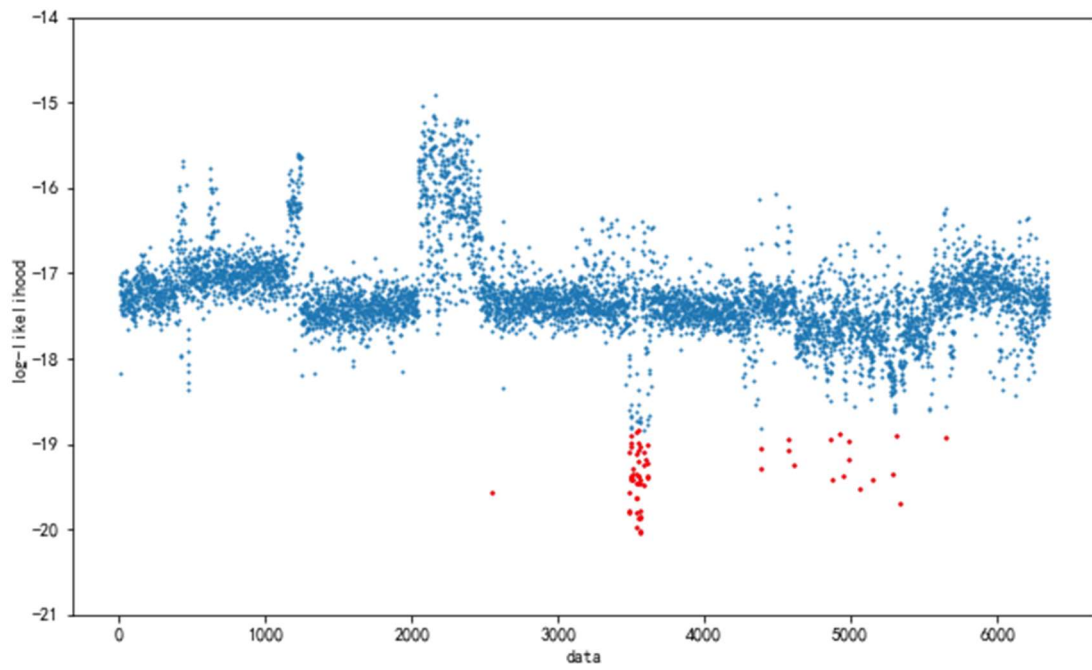


図 20, 0.5 時間のトレーニングデータから学習した GMM モデルの対数尤度

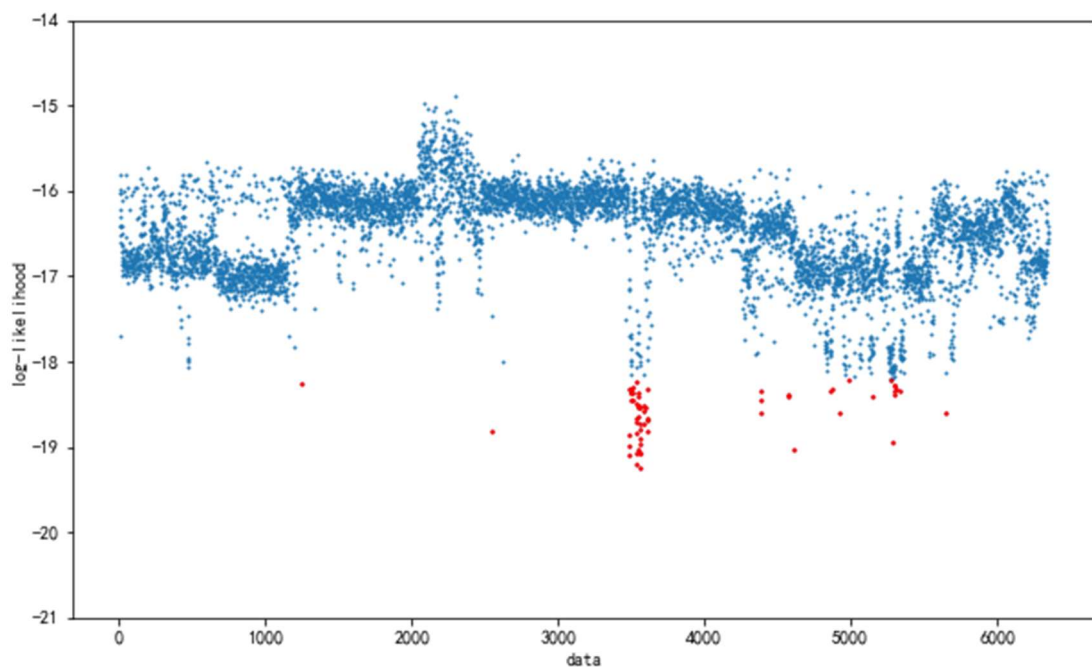


図 21, 1.0 時間のトレーニングデータから学習した GMM モデルの対数尤度

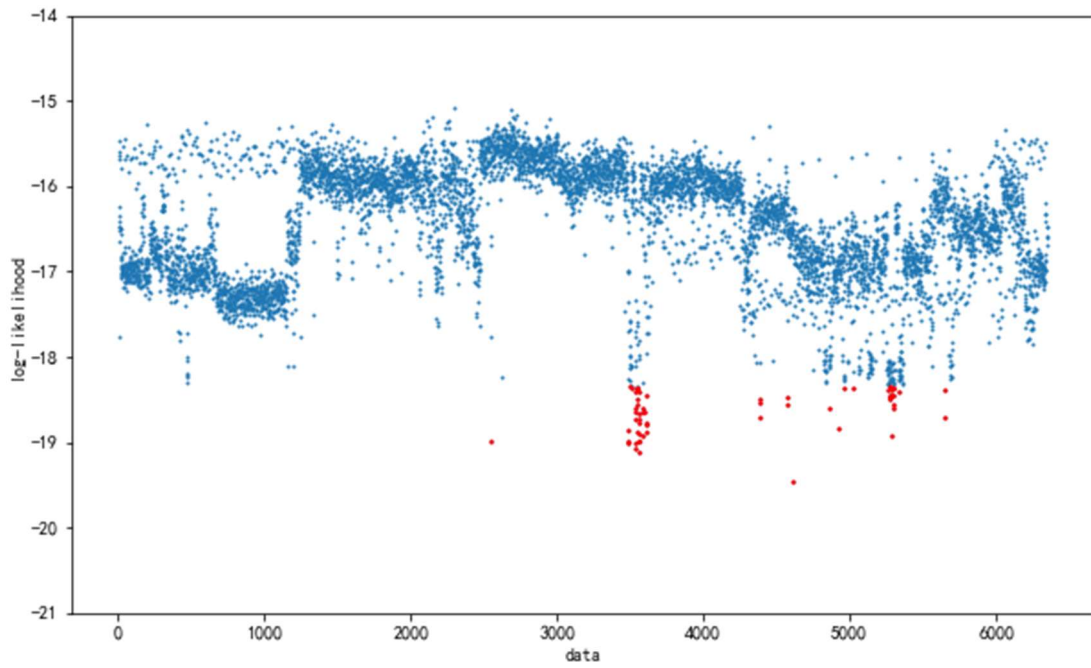


図 22, 2.0 時間のトレーニングデータから学習した GMM モデルの対数尤度

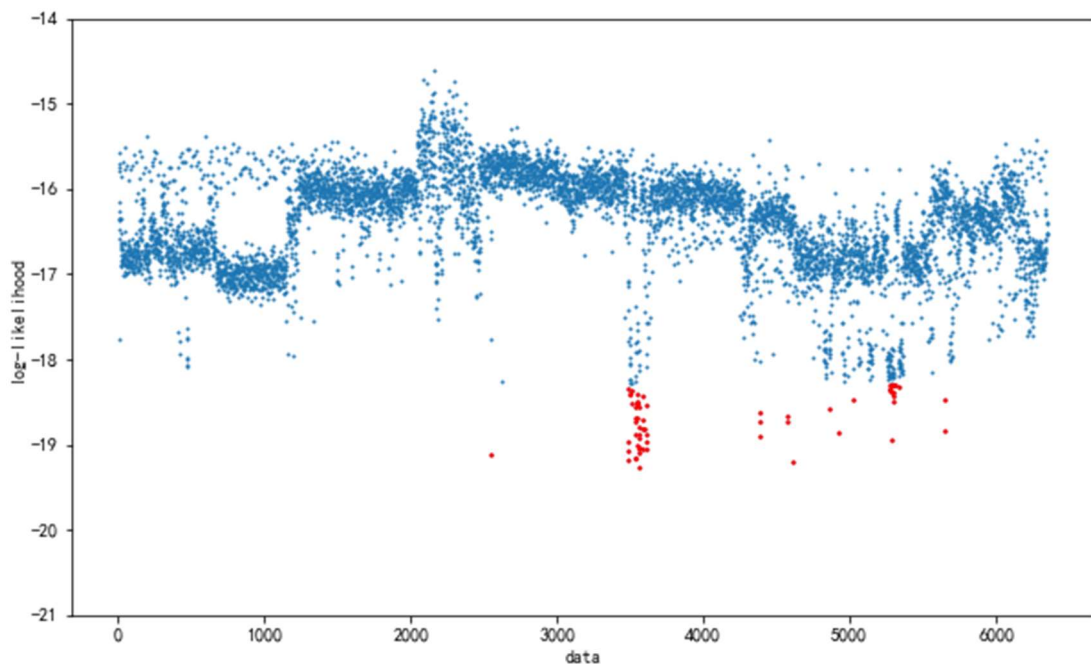


図 23, 3.0 時間のトレーニングデータから学習した GMM モデルの対数尤度



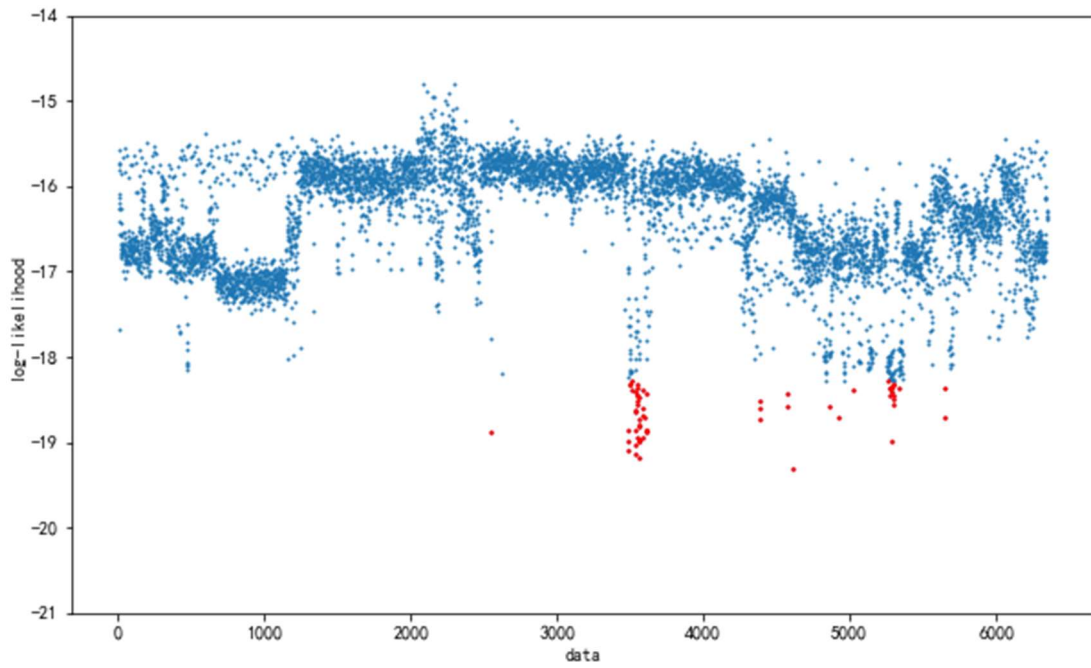


図 24, 5.0 時間のトレーニングデータから学習した GMM モデルの対数尤度

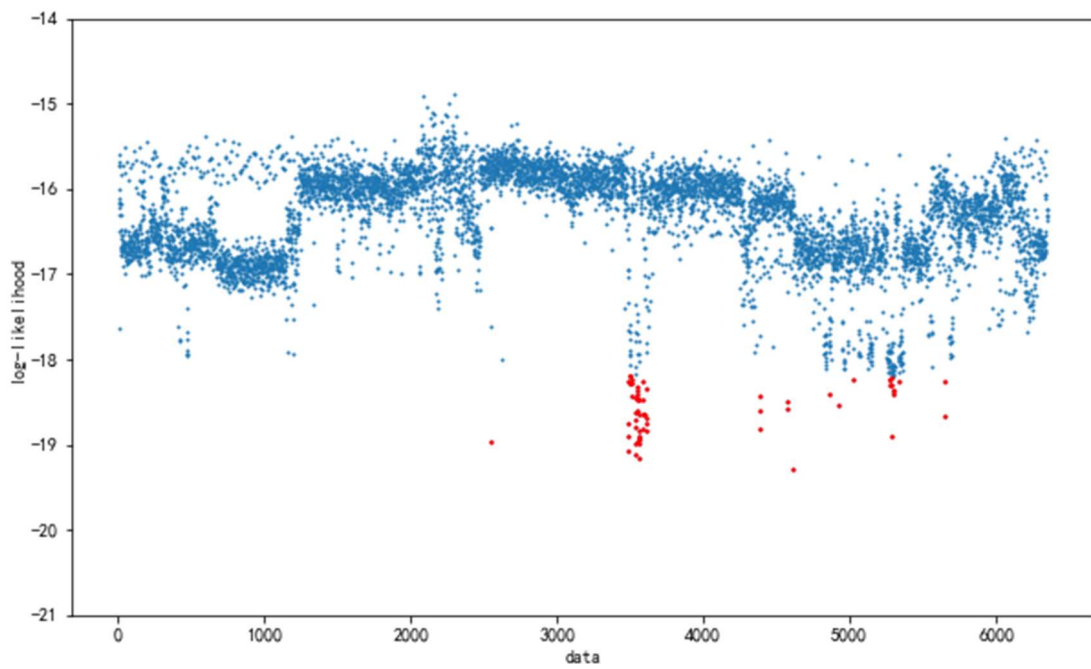


図 25, 10.0 時間のトレーニングデータから学習した GMM モデルの対数尤度

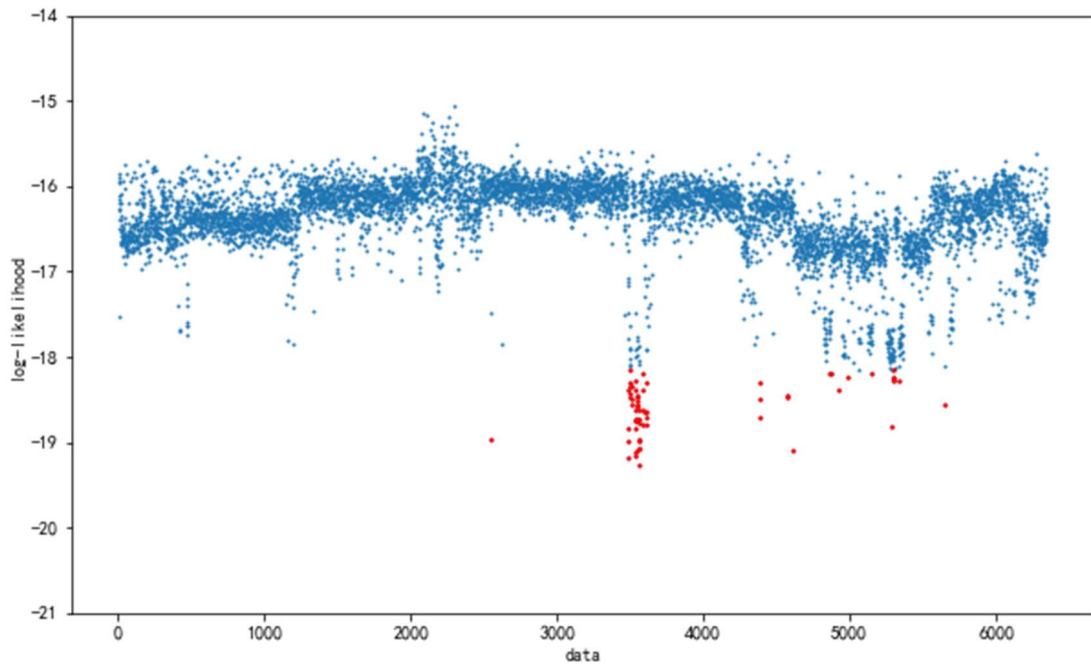


図 26, 20.0 時間のトレーニングデータから学習した GMM モデルの対数尤度

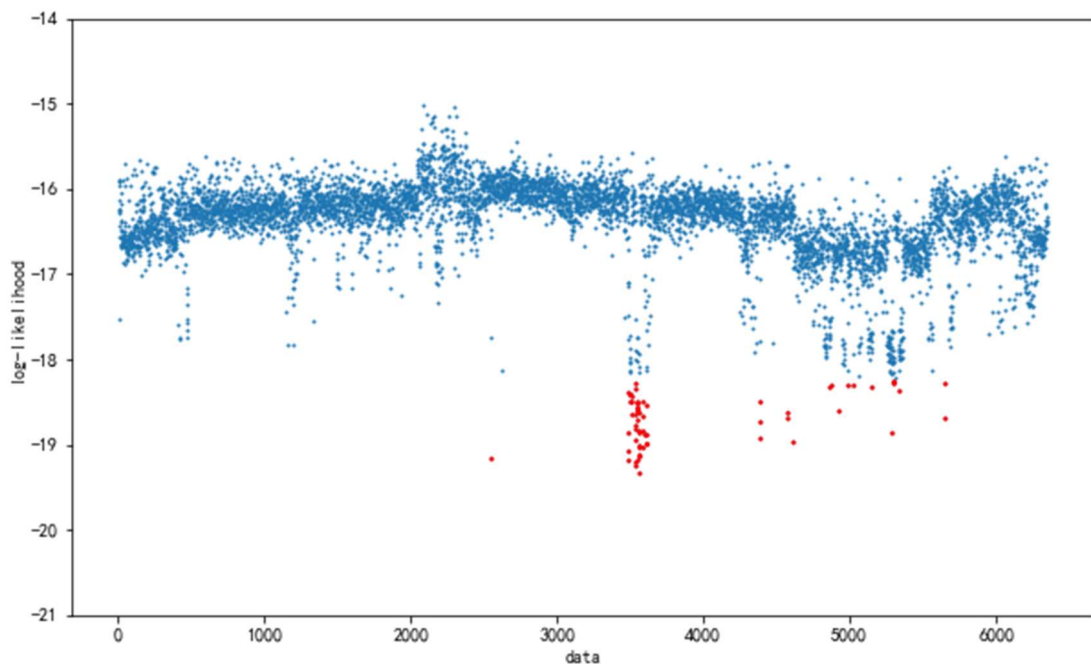


図 27, 30.0 時間のトレーニングデータから学習した GMM モデルの対数尤度

上の図では、横軸はテストデータの番号であり、縦軸はテストデータの異なる時間のトレーニングデータから学習した GMM モデルに対する対数尤度である。しきい値を 1% とする場合、つまり最も対数尤度の低い 1% のデータをいつもと違う音とする場合、いつ



もと違う音と検出されるテストデータを赤点と表示しており、他のテストデータを青点と表示している。

上の図に示されているように、0.5 時間のトレーニングデータから学習した GMM モデルの対数尤度の分布が他の分布と比較して全く違っており、0.5 時間だけでは足りないことが分かった。また、1 時間から 10 時間までのトレーニングデータから学習した GMM モデルの対数尤度の分布が、だんだん安定していき、1 時間から 10 時間まで増えたデータが同じような分布をしていることが分かった。ところが、10 時間と 20 時間の対数尤度の分布には、500 から 1000 までのテストデータの GMM モデルに対する対数尤度が著しく上がったことが分かったため、元の 10 時間のトレーニングデータには 500 から 1000 までのテストデータに含まれている音が少ないのに対し、増えた 10 時間のデータにはそのテストデータに含まれている音が多いため、対数尤度が上がったと考えられる。このことにより、20–30 時間の分布はあまり変わらないにもかかわらず、30 時間でいつもの音と認められかねることが分かった。音を収集する際に、なるべくもっと長い時間で録音しないと、音の分布がそれ以上に変化していく可能性があるため、いつもの音として認められない可能性がある。

また、大学実験室では行われる実験作業もよく変わっていると考えられ、行われる実験によってその日の音も昨日より変わる可能性があるため、いつもの音とする録音の時期が違くと、そのいつもの音も違う可能性がある。そのため、2019 年に「大学実験室における音響情報における解析」というテーマの研究のために、実験室の音の録音が残っているため、そこから同じく 30 時間のトレーニングデータと 8 時間 43 分のテストデータを抽出し、両方のテストデータにおける対数尤度の分布を明らかにし、トレーニングデータとテストデータが同じ年の場合と異なる場合における対数尤度を比較した下の図に示されている。

図 28 上と 29 上の場合、同じ 2019 年のテストデータの対数尤度において、2019 年の GMM モデルの方が対数尤度が高いことが分かった。また、図 28 下と 29 下の場合、同じ 2021 年のテストデータの対数尤度において、2021 年の GMM モデルの方が対数尤度が高いことが分かった。これにより、トレーニングデータとテストデータが同じ年の場合の方が、異なる年の場合よりテストデータの対数尤度が高いため、実験室の通常の状態を音から決定するには、短期間でサンプリングの方が、現状の通常の状態を表現している可能性が高い可能性があり、そのことに留意して今後検討していく必要があることが示された。

ここで結論としてまとめると、いつもの音を定義するために所要される時間は、30 時間以上の長時間の録音が必要と考えられ、また、録音する時期によりそのいつもの音も変

わる可能性が示唆され、短期間でのサンプリングの方が現状の通常のを表現する可能性が高いと示唆された。ただ、今のところ、テストデータにおいて異常音の検出された部位は、何時間のトレーニングデータを使われるのに関係なく、基本的に同じである。つまり、機械的にいつもと違う音と検知される音は、基本的にほとんどの場合でもいつもと違うと検知されると示唆された。

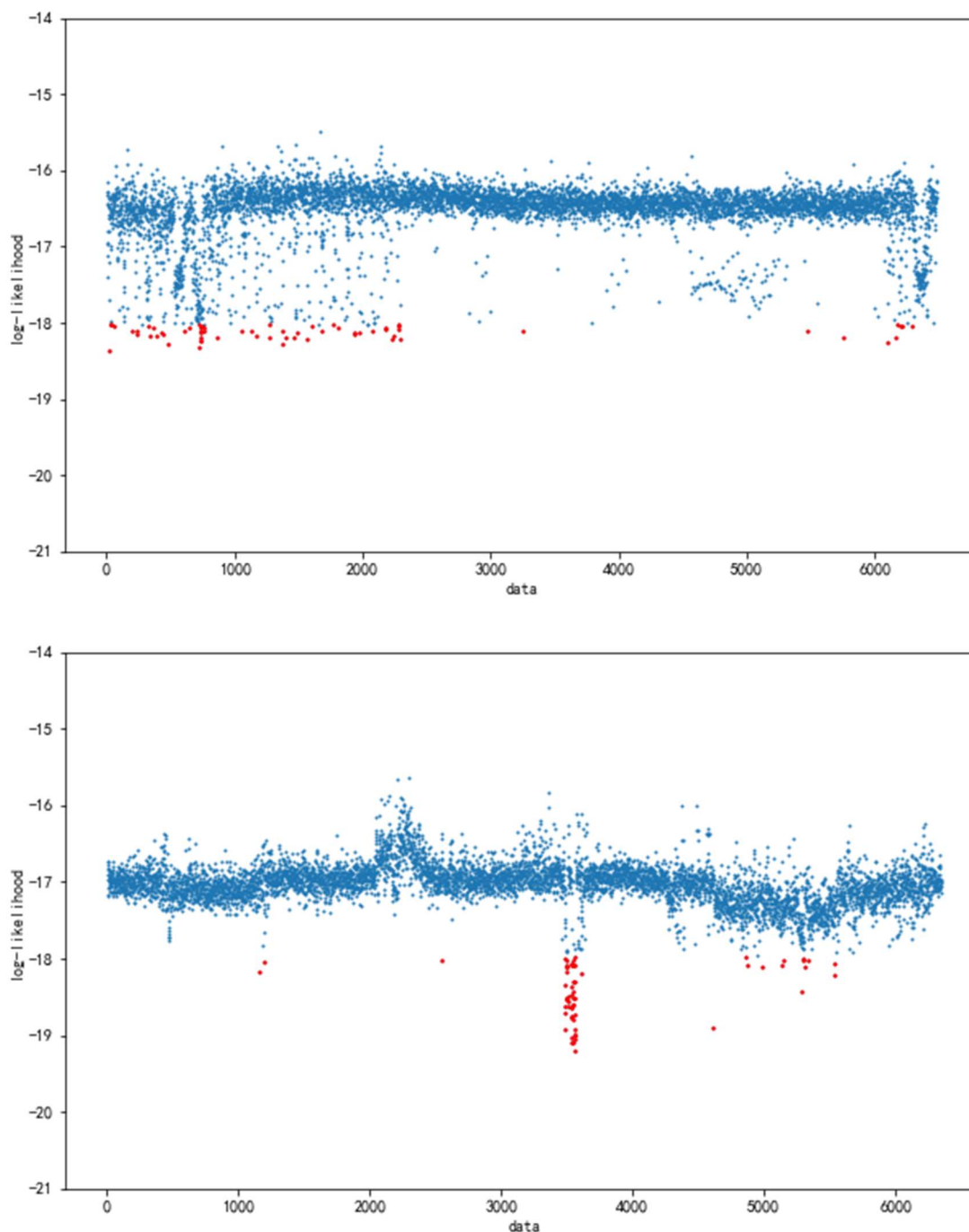


図 28, 2019 年の 30.0 時間のトレーニングデータから学習した GMM モデルの対数尤度  
(上 2019 年のテストデータ、下 2021 年のテストデータ)

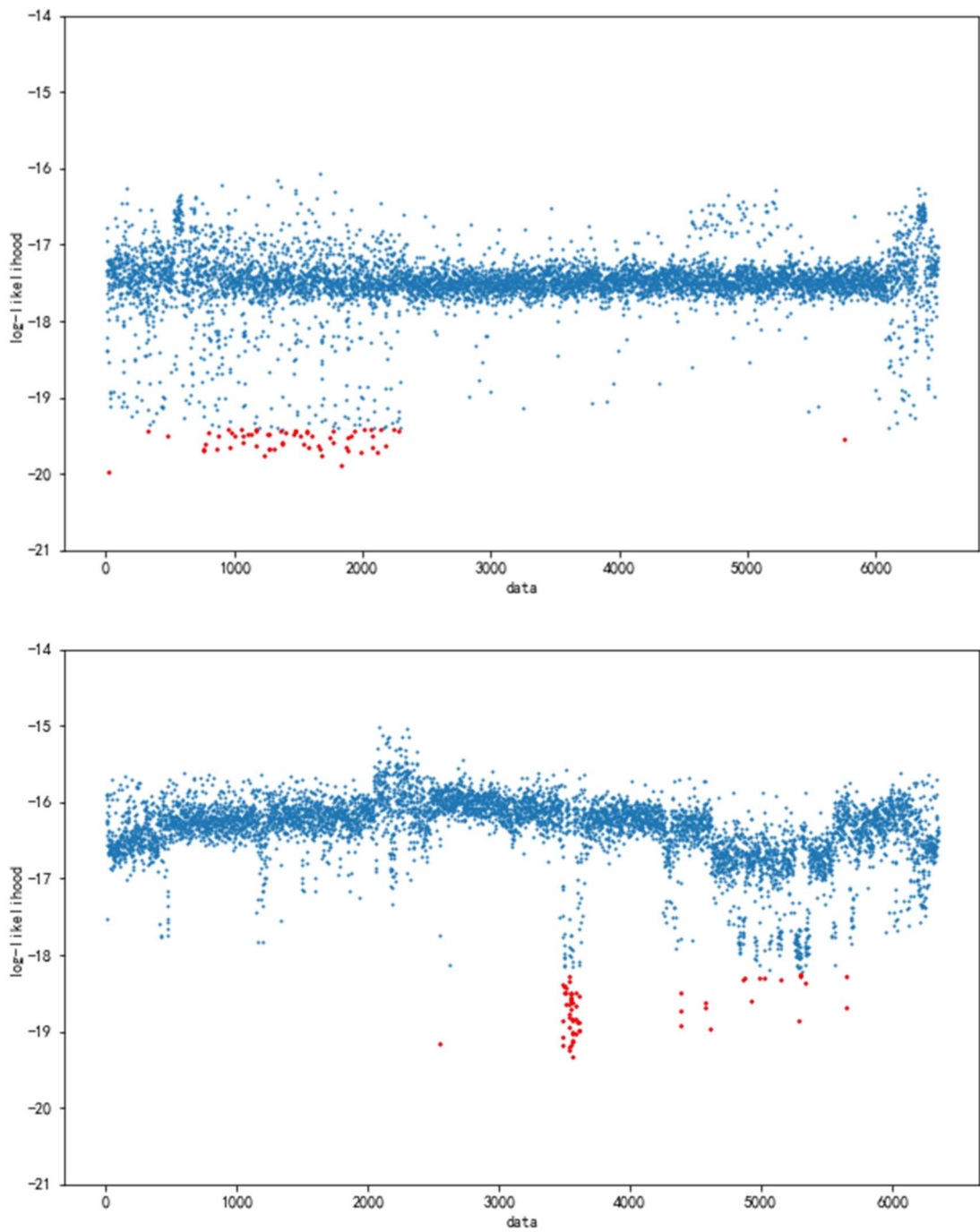


図 29, 2021 年の 30.0 時間のトレーニングデータから学習した GMM モデルの対数尤度  
(上 2019 年のテストデータ、下 2021 年のテストデータ)

### 3.2.3 いつもと違う音の抽出結果

上記の 8 時間 43 分 (6345 個、5 秒間につき) のテストデータを対象に、2021 年に録音された 30 時間の全体の音をいつもの音として GMM モデルに学習させ、その GMM モデルに対する最も対数尤度の低い 50 個の最も対数尤度いつもと違う音を抽出した。その

結果は以下の表に示されている。

表 8, 抽出されたいつもと違う音

音	出現回数/5s
ドライヤー	8
金属を磨く音	13
ガラスのぶつかる音	1
装置稼働音 1	2
装置稼働音 2	19
金属をたたく音	6
何の音もない	1

抽出されたいつもと違う音には、事故につながる異常音が見られなかった。その理由は事故につながる異常音はテストデータに存在しないか、あるいは事故につながる異常音は存在するが、検出されなかったと考えられる。どちらが理由かの検証は後で論述する。ここで抽出されたいつもと違う音は、テストデータの出現回数の少ない音として考えられるが、当実験室でよく実験を行われている学生にその音を聞かせると、異常ではないと答えられた。もし GMM モデルでは全体の音からいつもと違う音が検出できなければ、GMM モデルが適合しないという結論になるが、検証に通じてそうでない結論が出たら、ただそれ以上にいつもと違いそうな音はないとしか考えられない。そのため、GMM モデルが全体の音からいつもと違う音が検出できるかどうかについて検討する必要がある。

### 3.2.4 作られたいつもと違う音の検出結果

GMM モデルが全体の音からいつもと違う音が検出できるかどうかを検証するために、異常音が存在するかどうか分からないデータから検出できなかったため、いつもと違う音を作ってテストデータに入れた。作られたいつもと違う音は、遠心分離機の提示音と窒素がボンベから漏れ出る音である。遠心分離機の提示音は、事故につながる異常音ではないが、当実験室の実験者に聞くと、あまり出ない音のため、いつもと違う音と考えられ、その音をテストデータに入れた。また、窒素がボンベから漏れ出る音は、シューという音のため、明らかに異常な音であり、その大量に漏れ出ることによりリスクがあると考えられ

る。そのため、窒素がボンベから漏れ出らせ、その音を収集し、テストデータに入れた。下の図に示されているのは、窒素のボンベである。



図 30, 窒素のボンベ

それぞれの異常音を作ったあと、テストデータに入れ、テストデータが 6346 になり、そこから作られた異常音を対数尤度の低い音として検出されるかどうかを検証する。下の図 31、32、33 で示されているとおり、検出しきい値を 1% とすると、つまり最も対数尤度の低い 1% のデータを赤点にすると、二つの異常音を入れている場合では、その異常音が赤点として検出された (図 32、33 において矢印で標記されている)。その中、遠心分離機提示音も窒素がボンベから漏れ出る音も 1% の最も対数尤度の低い音に含まれていると見られた。遠心分離機提示音は 6345 個のテストデータの中の 1 位の対数尤度の低い音であり、窒素がボンベから漏れ出る音は 36 位の対数尤度の低い音であることが分かった。これにより、しきい値を 1% とした場合、30 時間のトレーニングデータから GMM モデルに学習させることにより、いつもと違う音は検出されることが可能と示唆された。また、テストデータから事故につながる異常音が抽出されなかったのは、存在する異常音が抽出されなかったのではなく、テストデータには異常音は存在しないからである。逆に異常音

が存在する場合、検出されると考えられる。設定されるしきい値により、いつもと違う音として検出される音は変わり、異常そうもない音も検出されるが、ここで検証したのは、いつもの音に含まれていない音や出現回数の少ない音は、いつもの音から学習した GMM モデルに対する対数尤度が低いことが示され、対数尤度は対象音のいつもの音に対する異常らしさを反映することが可能と示唆された。

表 9, 作られた異常音の対数尤度

異常音	対数尤度	対数尤度低い順位	対数尤度低い%
遠心分離機提示音	-20.334	1/6345	0.00%
窒素の漏れ出る音	-18.611	36/6345	0.57%

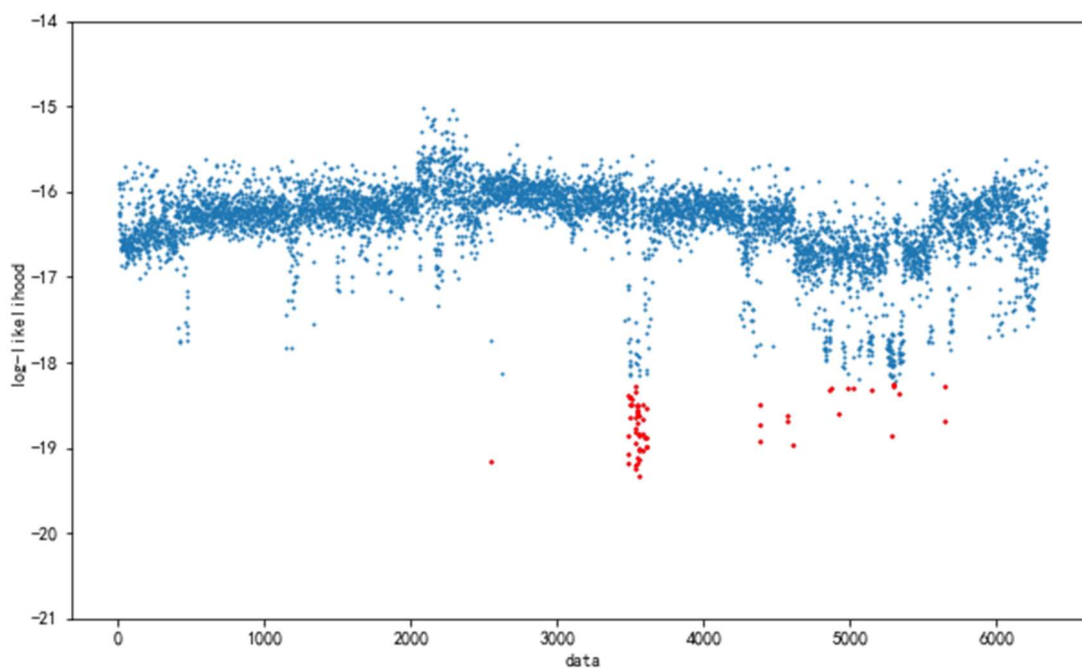


図 31, 異常音を入れていない場合



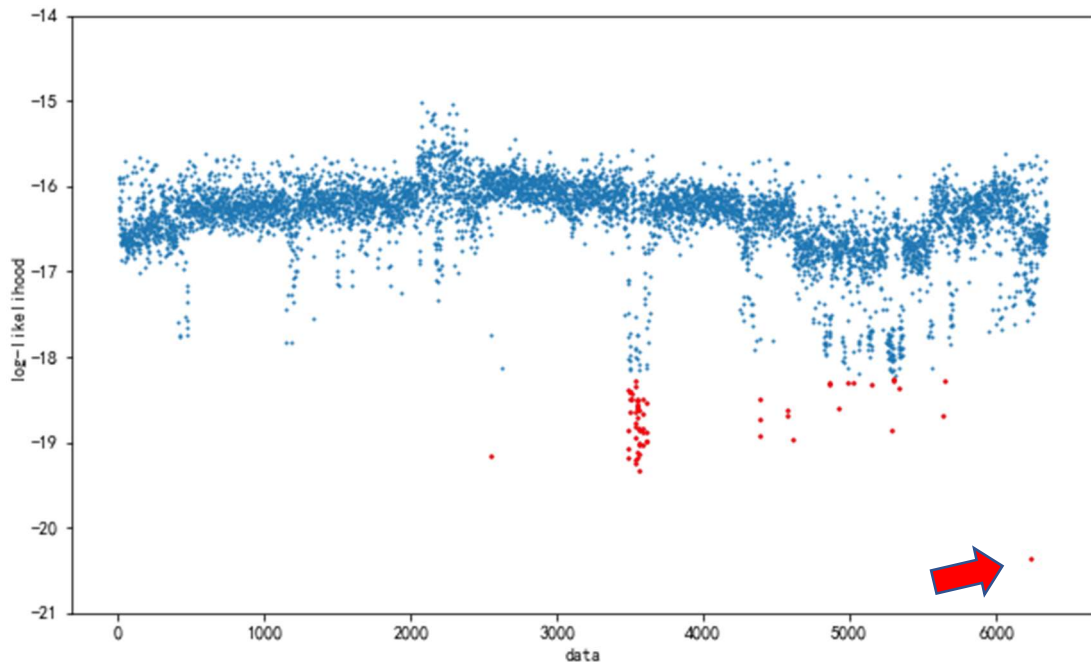


図 32, 異常音 (遠心分離機提示音) を入れている場合

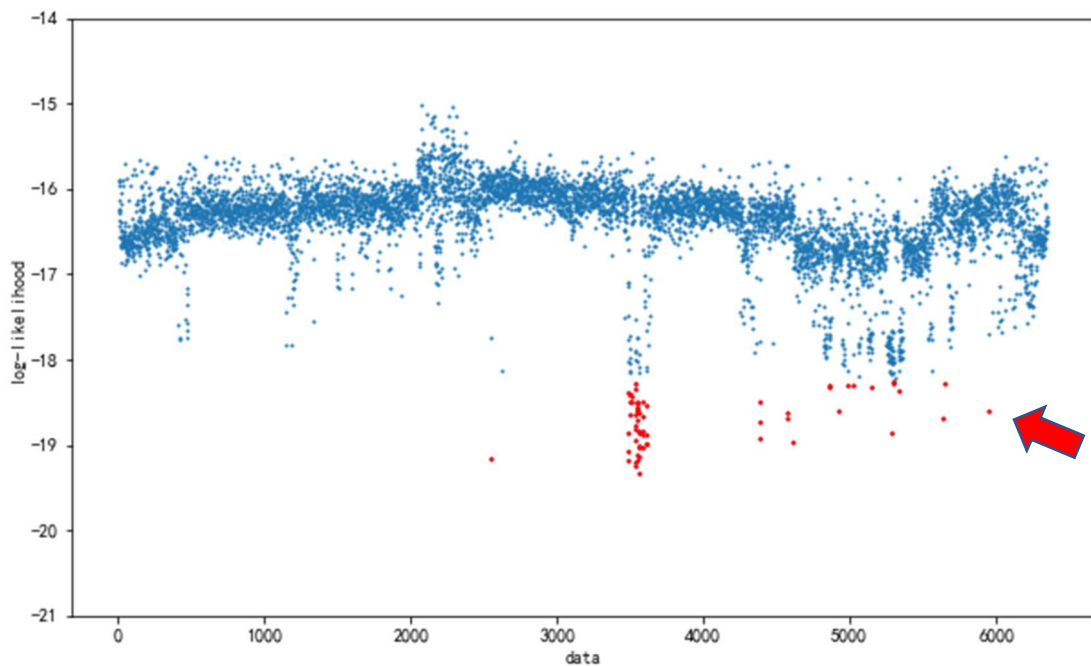


図 33, 異常音 (窒素がボンベから漏れ出る音) を入れている場合

### 3.2.5 リアルタイムモニタリングによる異常音検出の試み

3.2.4 により、いつもと違う音のいつもの音の GMM モデルに対する対数尤度が低いことが示されたことで、最終的に、GMM モデルの対数尤度による異常音検知方法により、大学実験室リアルタイムモニタリングにおける異常音検知への応用が期待されている。こ

ここでは、音をリアルタイムに録音し、その音の対数尤度をリアルタイムに計算して出力するプログラムがあれば、検出される異常音が時を移さず実験者に知らせることができる。そのため、録音と対数尤度の計算が同時に進むアルゴリズムを Python で構築した。録音は開始してから 5 秒後に終了し、そして自動的に繰り返すことができる。その録音が終了すると同時に、録音された音の用意された GMM モデルに対する対数尤度が即時に算出されて表示される。また、対象とする音は、事前に全体のいつもの音を用意するか、アルゴリズムにより自動的に音の録音ができ、GMM 学習も自動的に行われる。ここでしきい値を設定していれば、録音された音の対数尤度がしきい値より低くなると、警告などを出せば実験者に知らせると良いと考えられる。下の図に構築されたアルゴリズムにより録音と対数尤度の計算をしている様子を表示している。結果のところ、Start Recording が録音開始を示し、Recording Done が録音終了を示している。左の数字はデータの番号であり、右はそのデータの対数尤度を表示している。

```

Start Recording...
Recording Done...
0 [-17.45748084]
Start Recording...
Recording Done...
1 [-17.93063674]
Start Recording...
Recording Done...
2 [-17.89038944]
Start Recording...
Recording Done...
3 [-17.73238188]
Start Recording...
Recording Done...
4 [-17.81769157]
Start Recording...
Recording Done...
5 [-17.86038336]
Start Recording...
Recording Done...

```

図 34, アルゴリズムによるリアルタイム自動的録音と対数尤度の即時計算

ここでアルゴリズムの機能を検証するための模擬実験を行った。実験は、超音波洗浄機の音を 10 分間ほど録音し、それを超音波洗浄機のトレーニングデータとして GMM モデルに学習させる。また、検証実験として、超音波洗浄機の音を取りながら、金属器具を叩いて音を出し、その時の対数尤度が下がるかどうかを明らかにする。結果は以下の図に示されているとおり、トレーニングデータと同じ超音波洗浄機の稼働音は対数尤度が-18 以上で、比較的によく安定するが、金属をたたく時対数尤度が著しく-19.5 あたりに下がったことが示されている。ここで個別の音に対して実験を行い、アルゴリズムによりリアルタイムに音の対数尤度が計算され、異常音の検知が可能と示唆された。全体の音の場合、技術的問題があり、今までのデータを使用できなかったため検証できなかったが、今後の



実験による検証の必要がある。

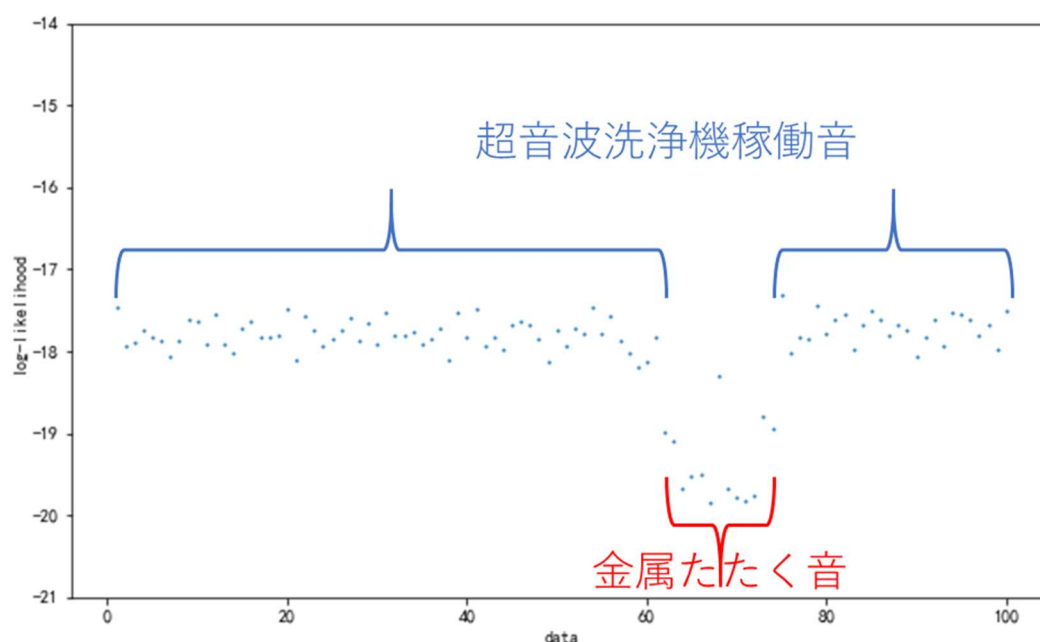


図 35, アルゴリズムによるリアルタイム異常音検知

### 3.2.6 聴覚評価実験

本研究は対数尤度の外れ値検出によりいつもと違う音を検出する方法を検討した。ここまでは機械的にいつもと違う音を検出できるかどうかを検討してきたが、最終的にその音がいつもと違うかどうかを判断するのは実験者の聴覚である。よって、ここで音の対数尤度と人間の聴覚との関係を明らかにし、今後の研究において活用される可能性を検討するため、実験室の音の聴覚評価実験を設計した。実験は、テストデータから対数尤度の高さの順で、違う対数尤度の音を抽出して実験のサンプルにした。対象とする被験者は、当実験室の実験者の学生と、同専攻の他の学生で、化学実験に参加したことのある学生になってもらった。実験は二回行われた。一回目は、対面での実験であり、被験者が筆者の説明と指示を受け、パソコンに入っている音を聴き、その音がどれほどいつもと同じそうか・違いそうかを筆者が配布した紙媒体のアンケートで答えた。二回目は、異なるテストデータから抽出された音を対象とした。また、二回目は対面ではなく、Google Formsでアンケートを作成し、被験者が自宅などでアンケートを答える形で被験者実験を行った。対象とする音を Google Drive に入れ、アンケートの各質問に音のリンクを入れた。特に、二回目では、当実験室の学生が自分の実験室の音と他の実験室の音を聞き分けるかを明らかにしたいため、当実験室において録音した音だけを対象としたのではなく、他の実験室でとられた装置の音も一個入れた。ただし、それを被験者に教えなかった。

二回の結果は以下の図で示されている。横軸は機械学習による音の対数尤度である。右

は対数尤度の高い音であり、左は対数尤度の低い音である。縦軸は被験者によるいつもと違う程度の判断であり、スコアと表示している。スコアが高ければいつもと違うとの結果であり、低ければいつもと同じとの結果である。また、黒点はその音のスコアの平均を表示しており、柱の両端がその音のスコアの標準偏差を表示している。

一回目では、当実験室の被験者（図 36）は対数尤度の高い音（-16 の近く）にいつもと同じと認識しており、対数尤度の低い音（-19 の近く）にもいつもと同じと認識しているが、スコアが対数尤度の高い音より高くなり、よりいつもと違いそうに認識していると考えられる。一方で、当実験室でない被験者（図 37）の場合では、対数尤度の高い音にいつもと同じと認識しており、対数尤度の低い音には、-19 付近ではいつもと同じと認識しているが、大体に対数尤度の低い音にはいつもと違うと認識している。ここでは、音の対数尤度の高さとの聴覚との相関が見られた。対数尤度の高い音は、どの実験室でも日常的に聞こえられる音と考えられ、両方の結果もいつもと同じと示された。対数尤度の低い音は、機械的にいつもと違うと認識されているにもかかわらず、テストデータには事故につながる異常音が存在しないため、その音でも当実験室の被験者にはいつもと違う音とは思わない。そのため当実験室でない被験者と違い、当実験室の被験者は対数尤度の低い音をいつもと同じ音と思っているからであると考えられる。

二回目では、他の実験室の音を入れており、その音が当実験室に存在しないため、対数尤度が約-20 であった。当実験室の被験者（図 38）の結果では、対数尤度の高さとの聴覚との明らかな相関が見られなかった。ただ、他の実験室の音に対し、いつもと違うという結果が見られた。これにより、当実験室の実験者は自分の実験室に存在しない音を聞き分けられることが分かった。その音は当実験室にとっていつもと違う音のため、対数尤度との相関がある可能性を示した。一方、当実験室でない被験者（図 39）の結果では、対数尤度の低い音のスコアが対数尤度の高い音のスコアよりやや少し高くなっており、弱い相関が見られた。ただ、他の実験室の音に対しても、対数尤度が低いにもかかわらず、いつもと同じという結果が見られた。つまり、他の実験室の正常音は、機械学習にも、当実験室の実験者にもいつもと違う音として検知されているが、その音の対数尤度が低いにも関わらず、正常な音のため、他の実験室の人はいつもと違うとは思わないと示唆された。

二回目の結果で、一回目の結果より音の対数尤度との聴覚との相関が弱く見られたのは、対面でない状況による意思伝達がうまくできなかったからか、または音を聞く環境が統一されていなかった（一回目では統一されていた）からだと考えられる。また、30 時間の音をいつもの音として学習させているため、それだけの音と当実験室の学生の思ういつもの音には差があるかもしれない。以上のことで、音の対数尤度との聴覚との相関があるという確固たる証拠を、今後さらに明らかにする必要がある。

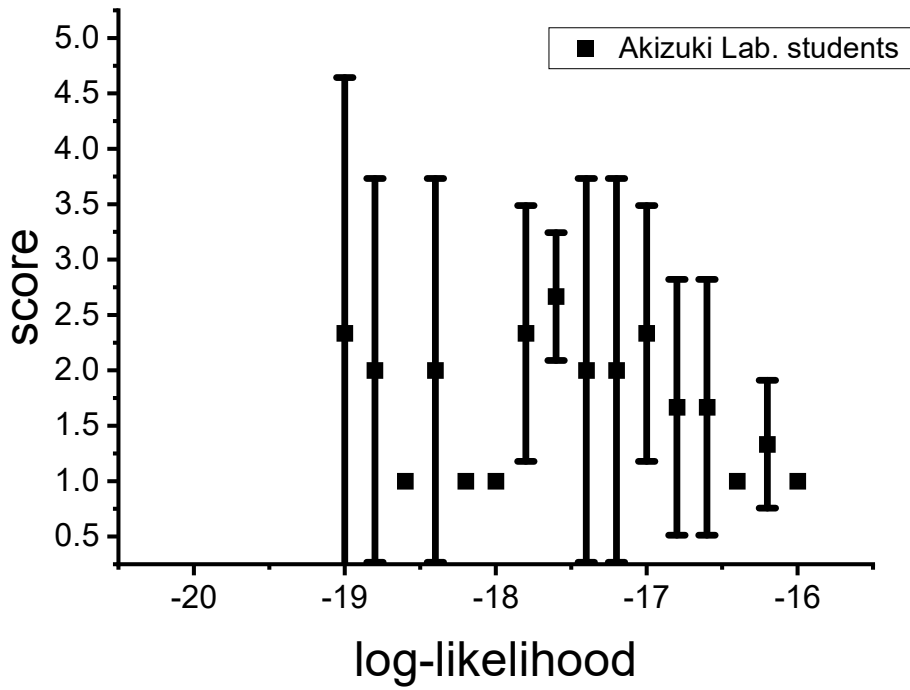


図 36, 一回目の当実験室被験者の結果

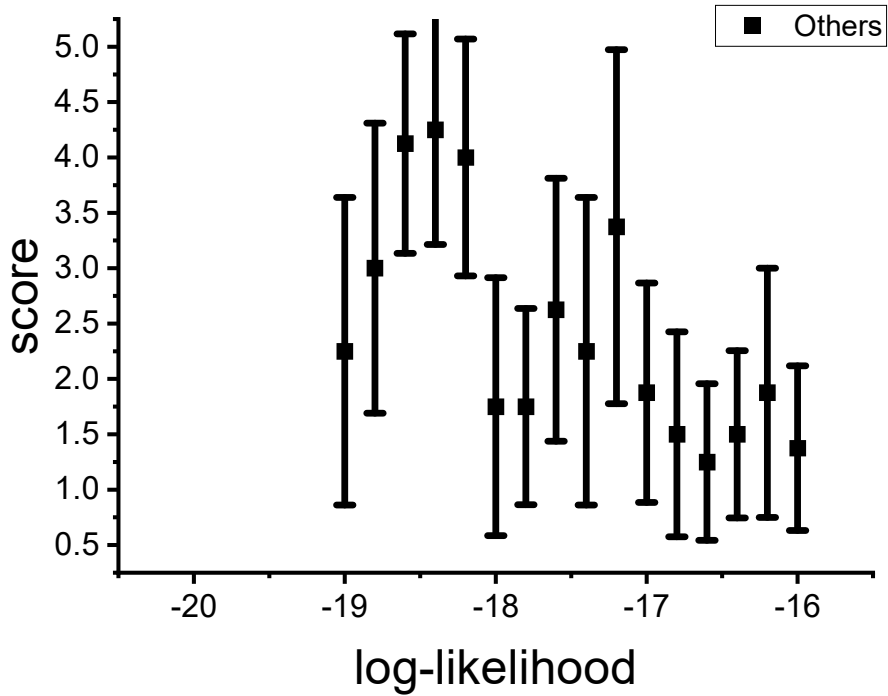


図 37, 一回目の当実験室でない被験者の結果

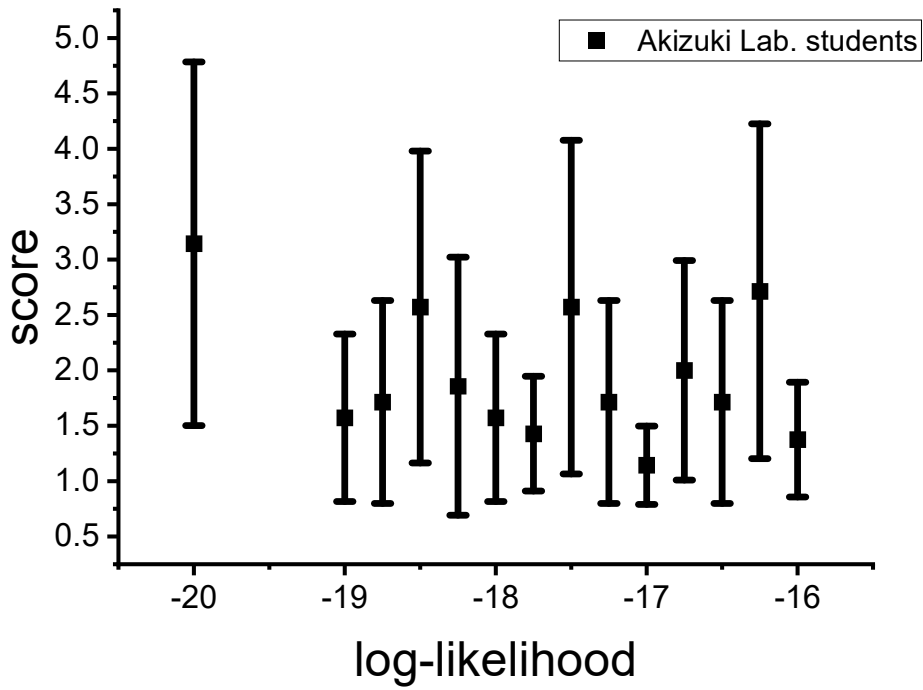


図 38, 二回目の当実験室被験者の結果

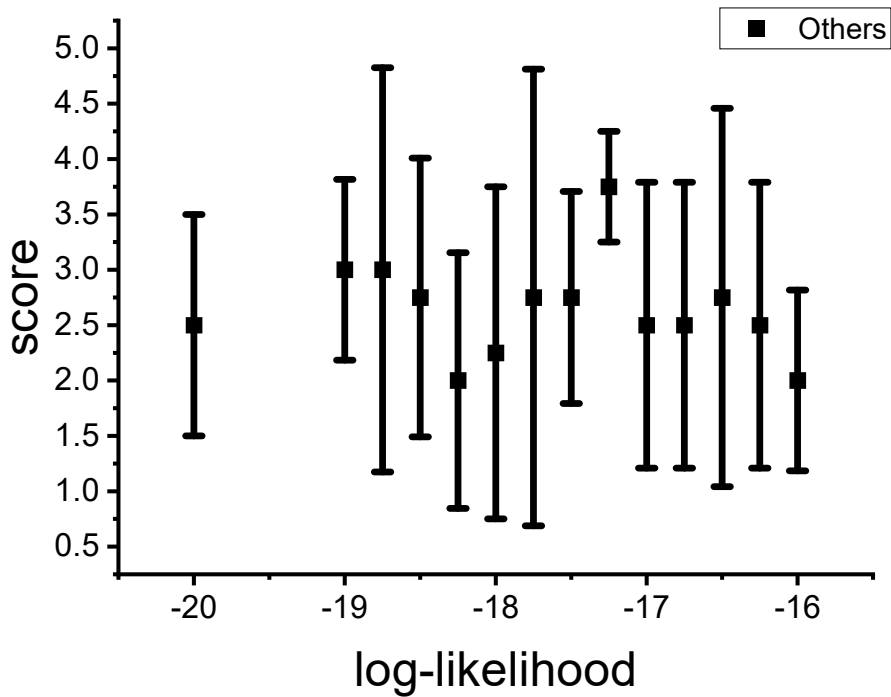


図 39, 二回目の当実験室でない被験者の結果

### 3.2.7 まとめ

3.2では、複雑な音環境からいつもと違う音を検出する手法を検討した。音がGMMモデルへの対数尤度の高さという視点を取り、対数尤度の低い音をいつもと違う音として検出するという手法を検討した。取られたデータから事故につながるいつもと違う音を検出されなかったにもかかわらず、異常音を作ってデータに入れて検出されたことにより、GMMモデルによる音の対数尤度の外れ値検出を使用し、適切なしきい値を設定していれば異常音の検出が可能と示唆された。ただし、適切なしきい値については、まだ検討する必要がある。外れ値検出による音の検出が可能との結果を基に、リアルタイムモニタリングに異常音を検出するアルゴリズムを開発し、そのアルゴリズムの効果を個別の音に対して検証できた。また、音の対数尤度の高さ与人間の聴覚との相関性について、被験者実験を行ったことにより、音の対数尤度の高さ与人間の聴覚には相関が存在する可能性を持つと示唆され、今後はさらに詳しく検討する必要がある。

## 第四章 結言

本研究では、大学実験室における異常状態を検知するために、大学実験室の音に着目し、それを定量的手法により簡単な音環境と複雑な音環境という二つ実験室シナリオにおける異常音を検知する手法を検討した。本研究は、人間の聴覚的尺度をあらわすメル周波数ケプストラム (MFCC) を音響特徴量とし、混合ガウスモデル (GMM) を音の機械学習モデルとした。それにより、音のモデル化と対数尤度の外れ値検出の二つの手法により、異なる音環境における異常音検知が可能と示唆された。

音のモデル化は、簡単な音環境の実験室を対象とし、その音環境に存在する、限りのある全ての音に対して録音し、録音された音データの MFCC 特徴量を抽出し、それを GMM 機械学習モデルに学習させることで個々の音がモデル化される。本研究では、その手法の有用性を検証するために、様々な装置の稼働音や提示音に対してモデル化し、テストデータのそのモデル化された音の GMM モデルに対する対数尤度を、尤度比検定により GMM モデルの認識率を検証した。結果はすべてのテストデータの音が、その音から学習した GMM モデルの対数尤度が最も高く、認識率がすべて 100%になっており、学習効果が良いと示した。また、過学習の恐れがあるため、データに対して 10 分割交差検証を行ったことで、過学習の可能性を排除した。これにより、音のモデル化という手法で、簡単な音環境における異常音検知へ応用できると示唆された。

外れ値検出は、複雑な音環境の実験室を対象とし、事前に音を分類する必要なく、ただいつもの音から学習した GMM モデルに対する明らかに対数尤度の低い音を検出することにより、いつもと違う音を検出される。取られた音をいつもの音とすると、30 時間以上の時間が必要であり、また録音する時期によりいつもの音も変わると示唆された。本研究では、30 時間のデータをいつもの音とし、検出しきい値を 1% に設定する場合、作られた異常音を検出されたことにより、外れ値検出の手法の有用性を検証した。ただし、設定されたしきい値によりいつもと同じ音も検証されることがあるため、適切なしきい値を設定することについては、さらに検討する必要がある。また、リアルタイムモニタリングによる異常音検知のアルゴリズムを構築し、個別の音に対してそのアルゴリズムの有用性を検証したが、今後は全体の音への有用性を検証する必要がある。最後に、評価指標とする対数尤度の高さ与人間の聴覚との関係について検討したが、対数尤度の高さ与人間の聴覚には相関が存在する可能性を持つことを示唆されたが、今後詳しく明らかにする必要がある。

最後に今後の展望として、異常音検知に用いられる音響特徴量について、MFCC の他、音の検知のためによく使われている色んな特徴量を考慮し、MFCC と同時に使うことで、

検知の効果が上がると考えられる。また、音の情報と画像情報、さらに気流や濃度などの情報と同時に異常検知に応用されれば、実験室における様々な異常を検知し、過去の経験や人間の知覚などに頼らずに事故の前兆を明らかにすることができる、そのような異常検知システムが実験室のみならず、様々な場面でも使われると安全管理において重要な意義を持つこととして期待される。

## 参考文献

- [1] 山本仁. 大学にフィットした合理的な化学物質リスクアセスメントのあり方[J]. 環境と安全, 2015, 6(3): 165-168.
- [2] 村川 正宏, 人工知能技術による異常検知システムとその産業応用, 日本原子力学会誌 ATOMOΣ 59(6), 335-339, 2017.
- [3] 宮本章広, 植草秀明, 後藤賢治. インテリジェント監視システム (AI・アドバンスト制御特集)[J]. 富士時報, 1998, 71(3): 167-171.
- [4] 寺島真介. 音響信号による設備監視[J]. 紙パ技協誌, 2003, 57(6): 870-877,024.
- [5] 豊田利夫, 中嶋智. 振動・音響的方法による設備診断[J]. 日本音響学会誌, 1983, 39(8): 550-555.
- [6] 中村将也. 石油コンビナート等特別防災区域内における 「未来の重大事故」 を無くすための提言について[J]. Safety & Tomorrow, 2013 (150): 18-26.
- [7] 化学実験における事故例と安全[M]. 株式会社 オーム社, 2013.
- [8] 川泉文男. 身近に見聞きした事故と" 危ない体験"(講座: 化学実験での事故防止のために-いくつかの事故例と安全教育 2)[J]. 化学と教育, 2005, 53(7): 394-397.
- [9] 廖 智聡, 大学実験室における音響情報の解析, 東京大学修士論文, 2020.
- [10] 上野圭介, 江原史朗. 機械学習のための機械動作音測定[C]//産業応用工学会全国大会講演論文集 2018. 一般社団法人 産業応用工学会, 2018: 51-52.
- [11] 江原史朗, 武藤義彦. One Class SVM を用いたベアリング異常音検知システム[J]. 産業応用工学会論文誌, 2021, 9(1): 31-37.
- [12] 大瀧淳司, 小菅太, 三譯勉, 等. 異音を定量的に検出する手法の開発[J]. 富士通テン技報, 2005, 23(2): 37-46.
- [13] Yang. Research on the Abnormal Sound Recognition Method of Aircraft Engine [D]. Nanchang Hangkong University, 2018.
- [14] Hong-Bin L I , Chu-Lin X U , Wen Z B . A STFT based method for detecting Rub & Buzz defects of loudspeaker and its application research[J]. Technical Acoustics, 2014.
- [15] 伊藤彰則. 環境音から異常を検知する統計的手法[J]. 日本音響学会誌, 2019, 75(9): 538-543.
- [16]  
[https://blog.csdn.net/jinping\\_shi/article/details/59613054?ops\\_request\\_misc=%257B%2522request%255Fid%2522%253A%2522162599530416780357261987%2522%252C%2522scm%2522%253A%25220140713.130102334..%2522%257D&request\\_id=162599](https://blog.csdn.net/jinping_shi/article/details/59613054?ops_request_misc=%257B%2522request%255Fid%2522%253A%2522162599530416780357261987%2522%252C%2522scm%2522%253A%25220140713.130102334..%2522%257D&request_id=162599)



[530416780357261987&biz\\_id=0&utm\\_medium=distribute.pc\\_search\\_result.none-task-blog-2~all~top\\_positive~default-1-59613054.first\\_rank\\_v2\\_pc\\_rank\\_v29\\_1&utm\\_term=%E9%AB%98%E6%96%AF%E6%B7%B7%E5%90%88%E6%A8%A1%E5%9E%8B&spm=1018.2226.3001.4187](https://www.google.com/search?q=voice+based+gender+recognition&rlz=1C1GCEWZ_C904416780357261987&biz_id=0&utm_medium=distribute.pc_search_result.none-task-blog-2~all~top_positive~default-1-59613054.first_rank_v2_pc_rank_v29_1&utm_term=%E9%AB%98%E6%96%AF%E6%B7%B7%E5%90%88%E6%A8%A1%E5%9E%8B&spm=1018.2226.3001.4187)

(最終閲覧 2021 年 7 月 21 日)

[17] <https://superkogito.github.io/blog/VoiceBasedGenderRecognition.html>

(最終閲覧 2021 年 7 月 21 日)

## 謝辞

本研究は東京大学大学院新領域創成科学研究科環境システム学専攻大島研究室において行われたものです。修士二年間の勉強と研究に渡り、多くの方々にお世話になりました。

指導教員である大島義人教授には、ゼミやディスカッションにおいて助言や啓発をいただき、研究の最初から最後まできちんとお世話いただきました。研究に対する考え方がいつも私にヒントをくださり、研究の視野を広げてくださりました。また、いつも励ましてくださっており、厚く御礼申し上げます。秋月信講師には、本研究の対象実験室の先生として、ゼミにて当実験室の方ならではの視点で、的確な知見とコメントをいただきました。特任研究員の根津友紀子さんには、日頃の実験の実施と設備の用意についてお世話になり、研究内容について私の分からないことに対して色々な助言や見解をいただきました。副査の徳永朋祥教授には、異なる専門でありながら、鋭い質問で研究の足りないところを意識させていただきました。皆様に心から感謝申し上げます。

大島秋月研究室の先輩と同期の方々には、研究や生活において大変お世話になりました。特に被験者実験に協力していただき、本研究の欠かせない部分に必要な結果を提供していただきました。毎週のゼミで、下手なゼミ発表にも関わらず、それでもいつも解ってくださり、質問やコメントを下さっている方にも、心から感謝申し上げます。

修士二年間、大島研究室で充実に楽しく過ごさせていただきました。ここに入るチャンスを提供してくださる大島先生、ここで一緒に研究・生活する研究室の皆さんには感謝いたします。そして何より、留学させて下さった、生活を支えてくださってきた両親には、ここで感謝の意を示します。

2021年7月21日

周睿