第5章 音声を加工してみよう!

5.1 Mathematicaの離散フーリエ変換

この章では、コンピュータを使って離散フーリエ変換および離散フーリエ積分を計算し、波の 特徴を捉えたり、波 (音; 音声; 音波)の加工を行ってみましょう。なお、離散フーリエ変換および 離散フーリエ積分の計算には、ウルフラム・リサーチ社 (Wolfram Research, Inc.)の開発した数 式処理ソフトウェア「Mathematica」を使用します (以後、単に Mathematica と記載します)。

 $\text{Mathematica}\ \mathcal{O}$

離散フーリエ変換 Fourier[リスト
$$u_r$$
] (= リスト v_s)

および

離散フーリエ積分 InverseFourier[リスト
$$v_s$$
] (= リスト u_r)

のデフォルトの定義では、(時間領域の)長さ nのリスト (配列) ur に対して離散フーリエ変換

$$v_s = \frac{1}{\sqrt{n}} \sum_{r=1}^n u_r e^{2\pi i (r-1)(s-1)/n}$$
 (s = 1, 2, ..., n)

を計算し、(周波数領域の)長さ nのリスト (配列) vs に対して離散フーリエ積分

$$u_r = \frac{1}{\sqrt{n}} \sum_{s=1}^n v_s e^{-2\pi i (r-1)(s-1)/n} \qquad (r = 1, 2, \cdots, n)$$

を計算します¹。なお、離散フーリエ変換および離散フーリエ積分の定義では添え字は0から始まりますが、Mathematicaのリスト (配列)の添え字が1から始まることに注意しましょう (指数関数のべきを $\pm 2\pi i (r-1)(s-1)/n$ として調節してあります)。

¹離散フーリエ変換と記述されていますが、実際には、高速フーリエ変換のアルゴリズムを使って計算します。

このままでは本テキストの定義とも異なり使いづらいので、オプションを設定して信号処理に適した離散フーリエ変換および離散フーリエ積分の数式に設定しましょう。Mathematicaのオプション付きの離散フーリエ変換 *v_s* および離散フーリエ積分 *u_r* の計算式は、それぞれ

$$v_s = \frac{1}{n^{(1-a)/2}} \sum_{r=1}^n u_r e^{2\pi i b(r-1)(s-1)/n} \qquad (s = 1, 2, \cdots, n),$$
$$u_r = \frac{1}{n^{(1+a)/2}} \sum_{s=1}^n v_s e^{-2\pi i b(r-1)(s-1)/n} \qquad (r = 1, 2, \cdots, n)$$

となっており、オプション

$$(a,b) = (0,1)$$
 デフォルト,
 $(a,b) = (-1,1)$ データ解析,
 $(a,b) = (1,-1)$ 信号処理,

を設定することで、各用途に適した計算式となります。本テキストの定義に合ったオプションは、

(a,b) = (-1,-1)

で、離散フーリエ変換 vs および離散フーリエ積分 ur は、それぞれ

$$v_s = \frac{1}{n} \sum_{r=1}^n u_r e^{-2\pi i (r-1)(s-1)/n} \qquad (s = 1, 2, \cdots, n),$$
$$u_r = \sum_{r=1}^n v_r e^{2\pi i (r-1)(s-1)/n} \qquad (r = 1, 2, \cdots, n)$$

となります。この章では、改めて、この一対の式を離散フーリエ変換および離散フーリエ積分と 定義し直すことにしましょう。なお、上記の離散フーリエ変換 v_s の式は、第4章では離散周波数 スペクトル密度と呼ばれていた式ですが、この章では、離散フーリエ変換と離散周波数スペクト ル密度を同じ意味として扱うことにします。また、新しい式では、 T_0 を含みませんが、周波数分 解能 $\Delta f = \frac{1}{T_0}$ [Hz]を決める重要な数値であることに注意しましょう。Mathematicaでは、

離散フーリエ変換 Fourier[リスト *u_r*, FourierParameters -> {-1, -1}]

および

離散フーリエ積分 InverseFourier[リスト v_s, FourierParameters -> {-1, -1}]

のように入力します。

具体的な計算例については、ホームページにリンクされている 「テキスト関連プログラム 11 (Mathematica ノートブック)」を ご覧ください。

* Chop[]命令は、ゼロに近い近似実数を厳密な整数0に置き換えます。

5.2 サンプリング

離散フーリエ変換を用いた自然現象などの波の解析や加工にあたって、サンプリングは非常に 重要な要素となります。扱う波の特徴や性質をよく知っている必要があります。本テキストでは、 最も身近な波として、音(音声;音波)と呼ばれる波について扱って行きます。

人間が耳から聞くことのできる音の周波数は、20 [Hz] から 20k [Hz] と言われています²。なお、 低い周波数の音ほど低い音に聞こえ、高い周波数ほど高い音に聞こえます。人間が耳から聞くこ とのできる音の周波数を全てカバーしようとすると、サンプリング間隔は、

$$\Delta t = \frac{1}{f_s} = \frac{1}{2 \times f_{\text{max}}} = \frac{1}{2 \times 20000} = \frac{1}{40000} \ [\text{P}]$$

となり、実に、1秒あたり 40,000 回ものサンプリングが必要となってきます。現在の主なオーディ オメディアのサンプリング数を挙げてみると、

電話:	11k 回 ([Hz]),
AM ラジオ :	22k 回 ([Hz]),
FM ラジオ (テレビ):	32k 回 ([Hz]),
CD :	44k 回 ([Hz]),
DVD :	48k 回 ([Hz])

となっており、デジタルオーディオと呼ばれている CD や DVD は、人間が耳から聞くことので きる音の周波数を全てカバーしていることがわかります³。特に、DVD のサンプリング周波数は、

$$f_s = \frac{1}{\Delta t} = \frac{1}{\frac{1}{48000}} = 48,000 \text{ [Hz]}$$

となり、人間が耳から聞くことのできる音の周波数を超えていることがわかります。このような サンプリング周波数が採用されているのは、人間は耳から聞くだけでなく皮膚などを通して音を 感じ取っていることが、最近の研究でわかってきたためです (イヌやネコには聞こえていること でしょう)。

逆に、人間の口から発せられる音 (声)の周波数は、標準的な人で 300 [Hz] から 4k [Hz] といわれています。主な、声や音の周波数を挙げておくと、

²20 [Hz] から 20k [Hz] というのは、非常に耳の良い人の値で、一般的な人は 20 [Hz] から 15k [Hz] 程度です。ちなみに、他の動物では、

イヌ:	50k [Hz],
ネコ:	60k [Hz],
コウモリ:	120k [Hz],
イルカ:	150k [Hz]

と言われています。

³固定電話で電話越しに声を聞くと少し違った声に聞こえるのは、伝送可能な周波数帯域が 300~3,400 [Hz] となっているためです。電話機が登場した当時は、話の内容が伝わればよいという考えで、音質までは余り考慮されていませんでした。

- 20.0 [Hz]: 音には聴こえない周波数,
- 27.5 [Hz]: ピアノの1番低い鍵盤周波数,
- 440 [Hz]: 時報の「ぴ」の周波数,
- 880 [Hz]: 時報の「ぽーん」の周波数,
- 1,000 [Hz]: 男性の声の主成分の周波数,
- 2,000 [Hz]: 女性の声の主成分の周波数,
- 4,096 [Hz]: ピアノの1番高い鍵盤の周波数

となっています。

以上のことを考慮しながら、音のサンプリング、離散フーリエ変換、離散フーリエ積分を行なっ ていくことにしましょう。まず、音(声)のサンプリングには、Windowsに付属している図 5.1 の ソフトウェア「サウンドレコーダー」を使用します。「スタート」メニューから、「プログラム」→ 「アクセサリ」→「エンターテイメント」→「サウンドレコーダー」の順に選択し、サウンドレ コーダーを起動してください(コンピュータのマイク端子にマイクを差し込んでおくのを忘れな いでください)。使い方は簡単なので詳細な説明は省きますが、2つほど注意点を挙げておきます。

∢{€Sound -	サウンド レ	コーダー		_ 🗆 🗙
ファイル(<u>F</u>)	編集(<u>E</u>)	エフェクタ(<u>S</u>)	ヘルプ(円)	
位置: 0.00 秒				長ざ 0.00 秒

図 5.1: サウンドレコーダー

1つは、マイクから録音できない場合です。マイクから録音できるようにするには、以下の設定 作業を行なってください。サウンドレコーダーのメニューから、「編集」→「オーディオのプロパ ティ」の順に選択し、図 5.2 の「オーディオのプロパティ」ウインドウを開きます。

オーディオのプロ	//ティ
オーディオデ	パイス
- 音の再生 で ()	 優先するデバイス(<u>P</u>): Crystal WDM Audio
-録音	
	優先するデバイス(<u>R</u>):
18	Crystal WDM Audio
	「長鹿の」 詳細の
MIDI 音調	業の再生
1	優先するデバイス(E):
	Microsoft GS Wavetable SW Synth
	音量(1) バージョン情報(8)
□ 優先デ	バイスのみ使用する(山)
	OK キャンセル 適用(益)

図 5.2: オーディオのプロパティ

さらに、「録音」の項目の「音量」をクリックし、図 5.3 の「録音コントロール」ウインドウを開きます。「マイク」の「選択」にチェックが入っているか確認し、入っていなければ「マイク」の 「選択」をチェックし、入力装置をマイクに設定します。

🗿 録音コントロール			_ 🗆 🗙
オプション(P) ヘルフ	<u>и(H)</u>		
モノラル ミキサー	CD プレーヤー	マイク	ライン入力
バランス:	バランス:	バランス:	バランス:
▶──	♦ – /– ♦		
音量:	音量:	音量:	音量:
□ 選択(<u>S</u>)	□ 選択(5) (☑ 選択(S)	□ 選択(<u>S</u>)
Crystal WDM Audio			

図 5.3: 録音コントロール

もう1つは、録音した音をファイルに保存する場合です。「ファイル」メニューから「名前を付け て保存」を選択すると、図 5.4 の「名前を付けて保存」ウインドウが開きます。

名前を付けて保存				<u>1</u>	2 N
保存する場所(1):	デスクトップ	•		* 💷 •	
マイ ドキュメント マイ コンピュータ マイ コンピュータ					
ファイル名(N):			•	保存(S)	
ファイルの種類(工):	サウンド ファイル (*.wav)		-	キャンセル	
形式: PC	M 22.050 kHz, 8 ビット, モノラル	変更(<u>0</u>)	\supset		

図 5.4: 名前を付けて保存

さらに、「形式」の項目の「変更」ボタンをクリックし、図 5.5 の「サウンドの選択」ウインドウ を開きます。

サウンドの選択		? ×
サウンド名(N): [無題]	✓ 名前を付けて保存(S)	削除(円)
形式(<u>F</u>): 属性(<u>A</u>):	PCM 8000 kHz, 8 ビット、モノラル 7 KB/秒	▼
	OK キャンセル	

図 5.5: サウンドの選択

このウインドウからサンプリング周波数を選ぶことができます。「形式」には Mathematica で読 み込み可能な「PCM」⁴を選択してください。また、「属性」には「8.000kHz, 8bit, モノラル」⁵を 選択してください (この選択を本テキストのデフォルトとします)。設定が終わったら、「OK」ボ タンをクリックし、図 5.4 の「名前を付けて保存」ウインドウに戻り、「保存する場所」と「ファ イル名」を適切に設定し、「保存」ボタンをクリックしてファイルに保存してください (ファイル の保存形式は「Microsoft wave 形式」で、拡張子は「.wav」です)。

サウンドレコーダーで離散化された波の作成を終えたら、Mathematica を使って離散フーリ エ変換や離散フーリエ積分を行なってみましょう。まず、離散化された波(保存したファイル) を Mathematica に読み込みましょう。ファイルを読み込むには、SetDirectory[]命令を使って ファイルのあるディレクトリに変更し、Import[]命令を使って離散的な数値列を配列として読 み込みます(詳細については、ホームページにリンクされている「テキスト関連プログラム11 (Mathematica ノートブック)」を参照のこと)。

離散化された波をグラフで見るには、ListPlot[]命令を使います。例えば、

ListPlot[x, PlotRange -> {Automatic, {-1, 1}}];

を実行すると図 5.6 のグラフが描かれます。ただし、変数 x には時間領域の離散化された波が配列形式で記憶されているものとします。



図 5.6: 離散化された波の確認

⁴アナログ信号の波形を時系列に区切り (標本化といいます)、その区切った波形の高さを 2 値 (2 進数) に変換して (量子化して)、その信号を表現する変調方式です。

⁵サンプリング周波数には、8k, 11k, 12k, 16k, 22k, 24k, 32k, 44k [Hz] のいずれかが選択できます (この節の最初 の表を参照のこと)。量子化ビット数には、8 ビットと 16 ビットのいずれかが選択できます (8 ビットでは 256 = 2^8 段 階で、16 ビットでは 65536 = 2^{16} 段階で、量子化されます)。モノラルとステレオの違いは、モノラルは 1 チャンネ ル、ステレオは 2 チャンネルとなります (時系列に対して、扱える独立した波の個数をチャンネル (ch) といいます)。

離散化された波を音として聞くには、ListPlay[]命令を使います。例えば、

ListPlay[x, SampleRate -> 8000];

を実行すると図 5.7 のグラフが描かれ、音(離散化された波)が再生されます。



図 5.7: 離散化された波の音の再生

離散フーリエ変換を行なうには、Fourier[]命令を使います。例えば、

lx = Chop[Fourier[x, FourierParameters -> {-1, -1}]];

を実行します。なお、Chop[]命令は、ゼロに近い近似実数を厳密な整数0に置き換えます。周波数領域における離散フーリエ変換(離散周波数スペクトル密度)のグラフを描くと図5.8および図5.9のようになります。グラフを描くには、

ListPlot[Re[lx], PlotRange -> {Automatic, {Min[Re[lx]], Max[Re[lx]]}}];

および

ListPlot[Im[lx], PlotRange -> {Automatic, {Min[Im[lx]], Max[Im[lx]]}}];

を実行します。



周波数領域において、波の特徴を調べるには各周波数スペクトル (各周波数の振幅)の大きさを調 べることが重要となります。したがって、各周波数ごとに絶対値を取ってグラフを描くと図 5.10 のようになります (1000 番目あたりの周波数で一番大きい周波数スペクトルとなっています)。な お、グラフを描くには、

ListPlot[Abs[lx], PlotRange -> {Automatic, {0, Max[Abs[lx]]}}]





図 5.10: 離散フーリエ変換の絶対値 |X_k|

*1秒間に 8000 回サンプリングしていることから、離散化された波の要素数を 8000 で割ると、観 測区間 (録音時間) T_0 [秒] が求まります。すなわち、 $\Delta f = 1/T_0$ [Hz] が求められます。

離散フーリエ積分を行なうには、InverseFourier[]命令を使います。例えば、

x = Chop[InverseFourier[lx, FourierParameters -> {-1, -1}]];

を実行します (グラフなどの描き方は、離散フーリエ変換の場合と同様に行なうことができます)。 なお、(離散フーリエ積分によって時間領域に戻した)離散化された波をファイルに保存するには、

Export["output.wav", Sound[SampledSoundList[x, 8000]], "WAV"];

を実行します。

具体的な計算例については、ホームページにリンクされている 「テキスト関連プログラム11 (Mathematica ノートブック)」を ご覧ください。