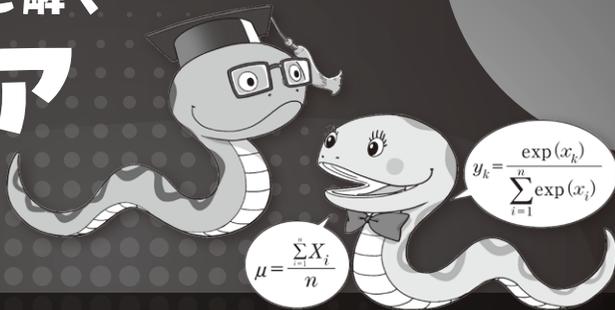


数式とコードと図でひも解く

エンジニア
数学第2回 音の信号処理(2) …
ダイナミック・レンジを小さくするコンプレッサ

川村 新

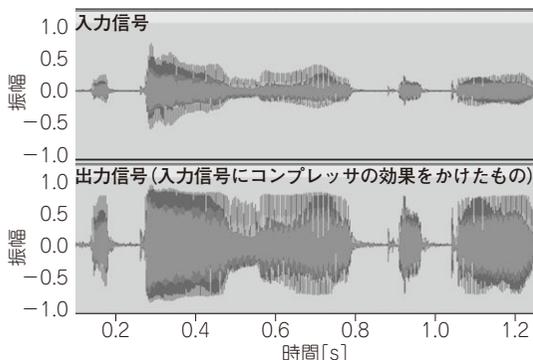


図1 実行結果の波形

本連載では、よく使用される技術や、専門外だとあまりふれない技術の数学を、対応するプログラムとともに解説していきます。(編集部)

● コンプレッサとは?…声も爆発音もほどよい音量にしたい

TVや動画を見ていて、声小さく聞こえにくい場面があると、ボリュームを上げて音量を大きくします。ところが、この状態で視聴を続けていると、物音や爆発音などの音量が大きくなりすぎて、慌ててボリュームを絞って音量を下げる場合があります。

このように、音声のダイナミック・レンジ(振幅の最大値と最小値の幅)が大きいと、ちょうど良い音量を保持することが難しくなります。そこで、ポップス音楽などでは、小さい振幅の音声部分だけを大きくして聞こえやすくしています。この技術は、音声のダイナミック・レンジを小さく圧縮する処理となるので、コンプレッサ(圧縮器)と呼ばれています(図1)。

コンプレッサは、音声だけでなく画像を見やすくするためにも使われる、とても応用範囲の広い技術です。

● 原理の数式

時刻 n における入力信号を $x(n)$ 、処理後の出力信号 $y(n)$ とします。ただし、これらはデジタル信号

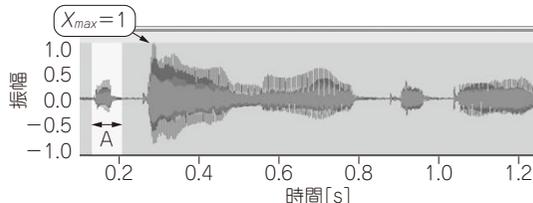


図2 単純な振幅の増幅

であり n は整数です。ここで、信号の絶対値の最大値を1とします。最初に単純なボリューム調整を考えます。この場合、入力信号を K (≥ 0)倍して出力すればよいので、

$$y(n) = Kx(n) \dots \dots \dots (1)$$

と書けます。ここで最大値が1なので、 K の範囲は、

$$0 \leq K \leq \frac{1}{x_{max}} \dots \dots \dots (2)$$

に制限されます。ここで、 x_{max} は、 $x(n)$ の絶対値の最大値を表します。

例として、図2のような信号を考え、範囲Aの音量を大きくしたいとします。しかし $x_{max} = 1$ より、不等式(2)から K の最大値は1となります。よって、式(1)で範囲Aの振幅を大きくすることはできません。

そこで別の方法を考えます。まず、 $x(n)$ 、 $y(n)$ について、両者の符号が同一とします。スルー出力の場合、入力と出力の絶対値は同じになるので、絶対値の入出力関係は、次の式となります。

$$|y(n)| = |x(n)| \dots \dots \dots (3)$$

ここで絶対値の最大値は1です。図3に式(3)の関係を示します。スルーなので振幅の増減はありません。

次に、しきい値 T_h ($0 < T_h < 1$)を導入し、 T_h 以上の入力に対して、傾き a ($0 < a < 1$)の直線に対応させてみます。このときの入出力関係を図4に示します。これにより、出力変化の幅、つまりダイナミック・レンジが圧縮されます。別の言い方をすれば、小振幅の音声と大振幅の音声の差が縮まります。

傾き a の直線は、点 (T_h, T_h) を通るので、 $(y - T_h)$