

Genetic Algorithm を用いた連続時間信号の分析に関して

後藤 龍二^{*1} 佐藤 裕二^{*2}

本稿では、(1)時系列信号の分析問題、(2)誤差を含んだ観測データのみからの解析問題、(3)陰関数を含む複雑関数の問題などの課題を全て含んだ問題を対象とした GA の適用可能性を探るために、音などの連続時間信号の特徴分析への適用可能性を実験的に調査した結果に関して報告する。音などの時間的に連続な信号の特徴分析を行うためには、その信号に含まれる周波数成分およびその周波数成分ごとの振幅、位相を分析する必要がある。この分析のためには、上記(1),(2),(3)の問題を全て解決しなければならない。連続時間信号の分析の用途として、音声や未知音源の音紋解析、機械装置の作動音、体内音の分析による異常検出などが考えられる。これらの信号は一定時間間隔で周期的に変化する周期信号が多い。また長時間では非周期信号であるが、数サイクルの短時間においては周期信号であるものも多い。実験結果から、GA を用いて、連続時間周期信号の周波数成分、振幅、位相の分析および解析結果の評価が可能であることを確認することができた。

Analysis of Time Series Signal with Genetic Algorithms

Ryuji Goto^{*1} and Yuji Sato^{*2}

The objective of this research is to investigate the applicability of genetic algorithms(GA) to problems that encompass all of time-series signal analysis, analysis from noisy observational data, complex and implicit function solutions through computer simulation. To know the features of signals like sound, it is necessary to analyze frequency component, amplitude and phase of each component. This capability might be applied to classify of sonic source and voice, to analysis failure mode of mechanical system and physical diagnosis etc. Many of these time-series signals are periodic. Through the simulation, we could certify applicability of GA to analyze the features of time series signal, to appraise analytical accuracy.

1 . はじめに

我々は、(1)時系列信号の分析問題、(2)誤差を含んだ観測データのみからの解析問題、(3)陰関数を含む複雑関数の問題などの課題を全て含んだ問題解法への遺伝的アルゴリズムの適用可能性を探る研究を行っている。音などの連続時間信号の特徴を分析するためには、信号に含まれる周波数成分および周波数成分の振幅、位相を求める必要がある。機械装置の作動音や移動体の機関音などの多くは一定時間ごとに繰り返される周期信号であり、長時間では周期信号でない連続時間信号でも短時間では周期信号である場合が多い。従来から連続時間信号の分析にはフーリエ変換が用いられている。フーリエ変換が可能であるためには、基本周波数の解析や窓関数による畳み込み演算が必要である。窓関数は信号の種類により適合性を評価する必要がある。フーリエ変換の分析は、オープンループであり、結果の評価は別途行う必要がある。本稿では、周期的連続時間信号を等時間間隔でサンプリングすることにより、離散的時系列デジタル信号に変換し、GA により基本周波数成分および高調波成分の振幅、位相を分析することを提案し、その評価結果に関して報告する。以下第 2 章で解くべき問題の定式化を行い、第 3 章で遺伝的アルゴリズムの適用方法、第 4 章で評価シミュレーションおよび考察、第 5 章でまとめを行う。

2 . 信号解析の定式化

図 1 は連続時間周期信号 $x(t)$ の例を示す。 T は周期である。 $x(t)$ の基本周波数を f_p とすれば、 $f_p = 1/T$ であり、角周波数を w_0 とすれば $w_0 = 2\pi f_p$ である。 $x(t)$ は正余弦波関数で級数展開す

[著者所属] *1: 法政大学大学院 ITPC Hosei-Graduate School IT Professional Course

*2: 法政大学情報科学部 Faculty of Computer and Information Science, Hosei University

ることができるので、式(1)で表すことができる。
 式(1)は信号を基本周波数およびその高調波成分で表すことができることを示す。式(1)はオイラー式により式(2)に変換することができる。 a_0 は直流成分であり、 a_k, b_k は直交する振幅成分である。 $\sqrt{a_k^2 + b_k^2}$ が振幅である。 a_k は位相である。これらの式は、積分表現の式であり、プログラム処理には不適当であるのでプログラムで扱い易い離散表現で表わす必要がある。

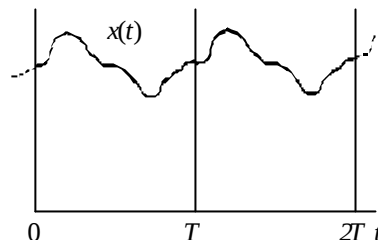


図1 連続時間周期信号

$$x(t) = \sum_{k=-\infty}^{\infty} C_k e^{jk\omega_0 t} \quad \text{ここで} \quad C_k = \frac{1}{T} \int_0^T x(t) e^{-jk\omega_0 t} dt \quad k = -\infty, \dots, -3, -2, -1, 0, 1, 2, 3, \dots, \infty \quad \dots\dots(1)$$

$$x(t) = a_0 + \sum_{k=1}^{\infty} (a_k \cos k\omega_0 t + b_k \sin k\omega_0 t), \quad a_0 = \frac{1}{T} \int_0^T x(t) dt, \quad a_k = \frac{2}{T} \int_0^T x(t) \cos k\omega_0 t dt, \quad b_k = \frac{2}{T} \int_0^T x(t) \sin k\omega_0 t dt$$

$$a_k = \tan^{-1}(b_k / a_k) \quad \dots\dots\dots(2)$$

図2において、連続時間周期信号 $x(t)$ を時間間隔 Δt で1周期に N 回サンプリングするものとする。サンプル数 N は $T / \Delta t$ となり、各サンプリング値は $x(n\Delta t)$ で表される。ここで n は0から始まる整数である。サンプリングの範囲は予想される周期 T を超える値とする。

式(2)は式(3)の離散表現で近似することができる。
 a_k, b_k は k, f_p で決まる値であり、 f_p が決まれば k 固有の値となる。式(3)において、 Δt はシステム条件によって予め決定される。その一例としては「サンプリング定理」で示されるナイキスト周波数（信号に含まれる最高高周波数の2倍）がある。信号に観測誤差が含まれるような場合は、更に高い周波数でのサンプリングが必要である。信号は一般には高い高調波になるほど振幅が小さくなる。また、特徴分析は予想される周波帯で行うため高周波ノイズ成分などの帯域外周波数成分を濾波する場合が多い。その場合、式(3)は有限項となる。

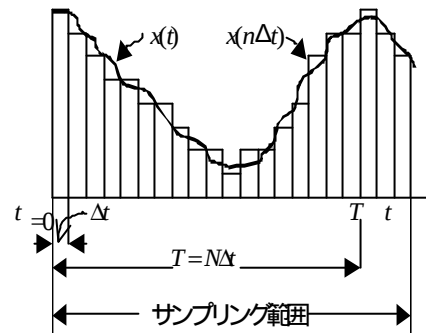


図2 信号のサンプリング

$$x(n\Delta t) = A_0 + A_1 \{ \sin(2p f_p n\Delta t) + a_1 \} + A_2 \{ \sin(4p f_p n\Delta t) + a_2 \} + A_3 \{ \sin(6p f_p n\Delta t) + a_3 \} \\ + A_4 \{ \sin(8p f_p n\Delta t) + a_4 \} + A_5 \{ \sin(10p f_p n\Delta t) + a_5 \} + \dots\dots\dots$$

$$\text{ここで} \quad A_0 = \Delta t f_p \sum_{n=0}^{1/\Delta t f_p} x(n\Delta t) \quad A_k = \sqrt{a_k^2 + b_k^2} \quad a_k = \tan \frac{b_k}{a_k} \quad k = 1, 2, 3, \dots, \infty \quad \dots\dots\dots(3)$$

3. 遺伝的アルゴリズムの適用方法

3.1 信号の構成要素

式(3)において未知数は $f_p, A_0, a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4, a_5, b_5, \dots, \infty$ である。このシミュレーションは、実験的であるので、求めるべき未知数を $f_p, A_0, a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4, a_5, b_5$ とする。

3.2 染色体の構造

前記した未知数 $f_p, A_0, a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4, a_5, b_5$ のうち A_0 は未知数であるが、式(3)で示す

とあり未知数 f_p の従属変数である。したがって、 A_0 は GA において染色体とせず f_p に従属して成長させる変数として扱うこととし、残りの $f_p, a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4, a_5, b_5$ を一つの個体を構成する染色体とする。染色体は 11 種類となる。構造を図 3 に示す。基本周波数の最大値は 600Hz としビット

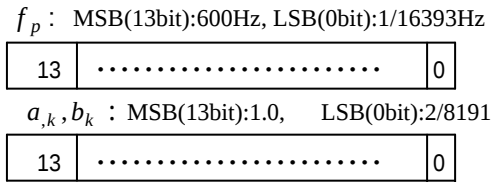


図 3 染色体の構造

操作の分解能を 600/16393Hz とした。各振幅成分である a_k, b_k の最大値を 2.0 とし、ビット操作の分解能を 1/16393 とした。この値は、信号の観測誤差が 0.2% であることを想定しているの、解析結果をそれに相当する精度とするためである。

3.3 適合度関数

サンプリングで観測した振幅値 $ox(n\Delta t)$ と、GA により推定した $f_p, a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4, a_5, b_5$ の推定値 $esf_p, esa_1, esb_1, esa_2, esb_2, esa_3, esb_3, esa_4, esb_4, esa_5, esb_5$ に基づいて求めた振幅値 $esx(n\Delta t)$ との差を求め、全てのサンプリング点の平均値の逆数を適合度と定義した。適合度関数を式(4)に示す。ここで M はサンプリング数、 C は適当な比例定数である。適合度 F は上記のとおり、観測した信号波形と解析し

$$F = C \frac{M}{\sum_{n=0}^{M-1} \{ox(n\Delta t) - esx(n\Delta t)\}^2} \quad \dots(4)$$

た波形の全サンプリング点の誤差を直接表すので、適合度により直ちに解析精度を知ることができる。

3.4 遺伝子操作

ある世代の個体群から次の世代の個体群を選択する方法として、適合度上位の個体を一定の割合で保留し、残りを淘汰する。淘汰した個体数をルーレット選択により補って、個体総数を保った。淘汰される個体数を M 個とすると、今の世代の個体から $M/2$ ペアの個体を重複を許してルーレット選択し、それぞれに対して、1 点交叉と突然変異を施して不足を補う M 個の個体とした。この遺伝的操作は、全ての染色体に対して独立に行った。さらに一定世代以降では、直前世代の最大適合度の個体がある時点における最適値であるので、それを構成する染色体を乱数で振って新しい染色体を作り、それらの染色体で組み立てた新個体で上記 M 個の一部を置き換え、収束の高速化を行った。

4 評価シミュレーション

4.1 評価方法

シミュレーションフローを図 4 に示す。観測値発生部において、理論値 $f_p, A_0, a_1, b_1, a_2, b_2, a_3, b_3, a_4, b_4, a_5, b_5$ を初期設定し $x(n\Delta t)$ の理論値を求める。

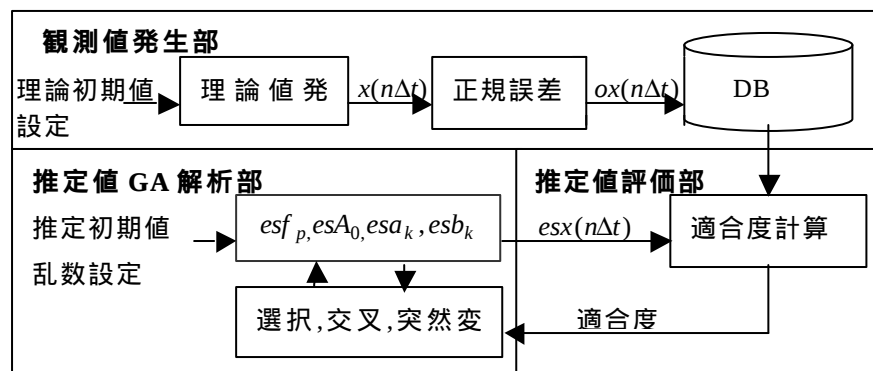


図 4 シミュレーションフロー

観測誤差としてランダムな正規分布誤差を付加し、観測値 $ox(n\Delta t)$ を求める。推定値 GA 解析部に初代の染色体をランダムに設定し、それらの値を初代の $esf_p, esa_1, esb_1, esa_2, esb_2, esa_3, esb_3, esa_4, esb_4, esa_5$

, esb_5 とする。これらに基づいて推定振幅 $esx(n\Delta t)$ を求め推定値評価部に送る。推定値評価部では $ox(n\Delta t)$ と $esx(n\Delta t)$ から適合度を計算する。適合度が予め設定した精度を満たさないときは推定 GA 解析部で、選択、交叉、突然変異により新しい世代の染色体に更新する。適合度が予め設定した精度を満たしたときまたは、前もって設定した世代に達したときシミュレーションは終了する。

4.2 シミュレーション条件

シミュレーションの条件は、つぎのとおりである。

- (1) 個体数=3600, エリート個体数=60, 最大世代数=60, 観測誤差=0.2%, 交叉率=0.8, 突然変異率=0.0005 サンプリグ周波数=ナイキスト周波数 5.0kHz の 0.95 倍, 1.0 倍, 1.3 倍, 1.6 倍, 2.0 倍
- (2) 理論値; $f_p = 500\text{Hz}$, $a_1=1.0, b_1=1.0, a_2=0.7, b_2=-0.7, a_3=-0.4, b_3=-0.4, a_4=-0.2, b_4=0.2, a_5=0.1, b_5=0.1$

4.3 シミュレーション結果

(1) 適合度の推移

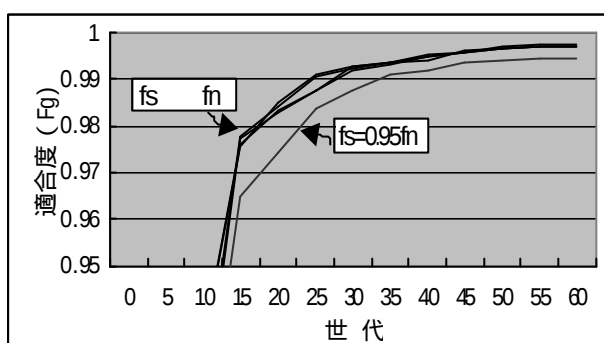


図5 適合度の世代推移

最大適合度の世代推移を図5に示す。この図の適合度 Fg は式(4)の F を $Fg=1-1/F$ で正規化した値である。適合度曲線はサンプリグ周波数 f_s がナイキスト周波数 f_n の 0.95, 1.0, 1.3, 1.6, 2.0 倍の場合を示す。ナイキスト周波数より大きいサンプリグ周波数の場合は、適合度は世代に従って収束しており、最終誤差は約 0.25% となっている。ナイキスト周波数より低い 0.95 倍の場合は、収束性が悪く、最終誤差は約 0.5% となっている。なお、この適合度は 10 回のシミュレーションの平均値である。

(2) 染色体の解析精度

ナイキスト周波数以上のサンプリグ周波数では最終解析結果の誤差はつぎのとおりであった。

基本周波数誤差：0.1Hz 振幅成分誤差：0.25% 位相角：0.3°

4.4 考 察

シミュレーション結果から、つぎの判断ができる。

- (1) 観測した信号の誤差 0.2% に対して解析結果の振幅成分誤差は 0.25% であり、GA 解析による誤差の増分は 0.1% 程度と考えられる。
- (2) サンプリグ周波数は、ナイキスト周波数以上でないと解析精度が低下する。このことは、一般のデジタル信号処理においてエイリアシング防止のために適用される「サンプリグ定理」を、GA による信号処理においても適用する必要があることを示している。

5 ま と め

(1) 時系列信号の分析問題, (2) 誤差を含んだ観測データのみからの解析問題, (3) 陰関数を含む複雑関数の問題などの課題を全て含んだ問題解法へ GA の適用可能性を探る一例として、音などの連続時間信号の特徴 (周波数成分および周波数成分の振幅、位相) 分析の可能性をシミュレーションにより探求した。その結果、GA による分析が可能であり、適合度により解析精度の評価で可能であることを確認できた。今後は、更に多様な条件によってシミュレーションを行う必要がある。特に解析可能高調波次数、観測誤差に対する耐性、処理時間の短縮の検討が重要である。処理時間に関しては、正余弦波の象限対象性を利用し、計算の重複を回避する方法が効果的と考えられる。分析範囲、分析精度、処理時間等の観点からフーリエ変換との比較検討を行い、改善していく必要がある。