

視聴覚的情報検索システムの 時系列情報処理への適用

統計数理研究所 濱 田 義 保

(1981年11月 受付)

はじめに

大量の時系列データが与えられたとする。この時、データがどのような時間変化をしているかを解析し、そのデータを与えるシステムの将来の変動を予測することが、時系列解析の目的である。では、与えられたデータをそのまま解析に供して良いものであろうか。これに対する答は否である。なんとすれば、与えられたデータの中には、誤データ等、不要なものが含まれていることがあり、あらかじめ、必要なデータのみを分離しておかねばならない。また、データの動きを説明するための時系列モデルを求める時にも、前もって、データを定常時系列とみなすべきか、非定常時系列とみなすべきか分っていた方が、好都合である。さらに、線形モデル、非線形モデルのいずれをあてはめるべきか分っていた方が良い。

では、データの数値の列をながめて、上に述べたようなことを実行するのは可能であろうか。そのようなことは、至難の業であると考えの方が自然であろう。アナログ的にデータを図示し、その時間変化を見て、判断する方が、はるかに能率的である。そこで考え出されたのが、視聴覚的情報検索システムである。このシステムは、統計数理研究所における特別研究「大量時系列データの統計的情報処理に関する研究」の一環として、赤池弘次の発案によりアイ電子測器株式会社の協力を得て作製された。現在このシステムはなお試用改良の段階にあるが、その現状について報告するのが本稿の目的である。

視聴覚的情報検索システムの構成

システムの機器構成は図1のようになっている。この中で、視聴覚的情報検索において、特に重要な働きをする、モニタ装置と音声出力装置の仕様について説明する。

モニタ装置は、表示用の3現象オシロスコープと2048ワード×2チャンネルのデータを記憶する外部メモリ及び、D/Aコンバータから成っている。オシロスコープのチャンネル1、チャンネル2に、それぞれメモリの各チャンネル中に保存されたデータが同時に表示される。画面に表示されるデータ数は各チャンネル1024データであるが、ダイヤルの操作により表示開始点を前後に動かすことによって2048データすべての様子

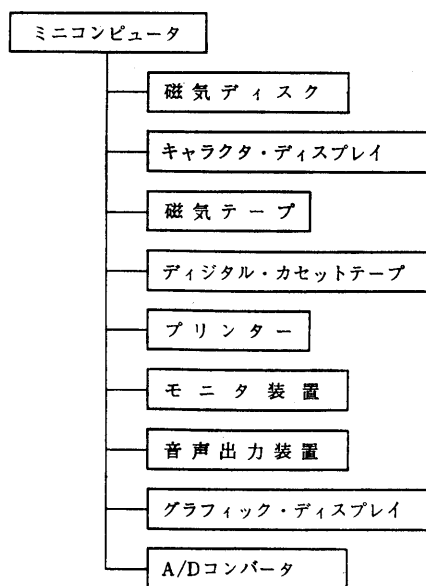


図1 視聴覚的情報検索システムの構成

を見ることができる。

オシロスコープのチャンネル3には、チャンネル1,2に表示されているデータの中から、必要な部分を選択するためのウィンドが表示される。ウィンドの先頭と末尾は、それぞれ独立に4データ単位に設定できる。表示された1024データの一部あるいは全部を解析の対象として用いる場合には、装置に付随したコレクト・キーを押す。この操作により、ウィンド中に含まれるデータ部分がミニコンピュータ部に送出される。表示データ全部が不要な時は、コレクト・キーの代りにスキップ・キーを押す。

コレクト・キーまたはスキップ・キーを押すと、それまで表示されていたデータに続く2048ワード×2チャンネルのデータがミニコンピュータからモニタ装置のメモリに読み込まれ表示される。そこで、これまで述べてきた操作を繰返せば、解析に必要なデータを容易に選択することができる。

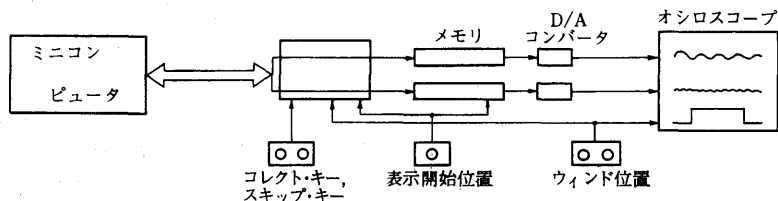


図2 モニタ装置の構成

音声出力装置は、データから算出された自己回帰モデルの白色雑音に対する応答を音声出力としてとり出す目的で作製された。ミニコンピュータによって計算された応答出力をDA変換した後、ローパスフィルターにより平滑化したものを音声として出力する。大量の定常時系列の各部分のスペクトル特性を自己回帰モデルのパラメータ値として記録し、そのスペクトルを図示すると同時に対応する音声を出力すれば、後の検索は大いに容易になるであろう。ただしこの音声出力装置は現在のところまだ試験的な段階にあり、更に改良が計画されている。

ソフトウェアシステムの開発

モニタ装置、音声出力装置を効果的に利用するために、次のような機能を持つプログラムを新たに開発した。1変数の時系列データの検索を行うためのプログラムについて説明する。データをチャンネル1にのみ表示し、チャンネル2はあけておく。コレクト・キーおよびスキップ・キーを用いることにより連続的にデータを見、誤データ等を検出する。ミニコンピュータのキーボードからの指示によって、誤データ等を含まない部分を解析に用いるデータとして選択する。このようにして選択されたデータは再びチャンネル1に表示される。表示されたデータの時間変化を見ることにより、定常時系列とみなすべきか非定常時系列とみなすべきかのおおよその判定ができるであろう。非定常時系列と判定した時は、スキップ・キーを押すとミニコンピュータで階差をとる計算をした後、階差をとった時系列をチャンネル2に表示できるので、階差をとったものが定常になるかどうか、判定すればよい。階差の階差をチャンネル2に表示することもできるので、定常時系列と判定できるまで逐次この操作を行えばよい。定常時系列と判定した時はコレクト・キーを押すと、チャンネル2に表示されているデータをミニコンピュータに送込み、解析に用いることができる。

データの定常性を調べるために便利のように、連続した4096データの前半をチャンネル1に、後半をチャンネル2に表示させるようにした。この時、両チャンネルの同じx座標のところに2048

ステップだけ、時間の異ったデータが表示されていることになるので、2つのチャンネルの時間変化の様子を比べることにより、容易に定常性の検討ができる。

2変数以上の時系列データを扱う時のためには、チャンネル1,2に適当に選んだ2組のデータを表示させるプログラムを用意し、2変数間の相関の概要をつかむことができるようにした。

以上では、入力データを処理するために開発したプログラムについてのみ述べてきた。計算結果の表示に、モニタ装置を利用するためのプログラムもいくつか開発した。チャンネル1に入力データ、チャンネル2に解析の結果得られた自己回帰モデルを使ったシミュレーションの結果を表示し、比較できるようにした。また、解析の結果得られるパワースペクトルをチャンネル1に表示するプログラムも用意した。この時、チャンネル2には、スペクトルのレベルを知るための基準線を表示する工夫を行った。

音声出力装置を利用して、入力データの定常性を調べるためのプログラムも開発した。これを用いれば、入力データをいくつかのブロックに分けて、各部分の時間変化を音声出力として聞くことができる。もし入力データが非定常であれば、音の性格が変化するので容易に認識することができる。

以上に述べたようなソフトウェアシステムを開発したことにより、視聴覚的情報検索システムの効率的利用が、初めて可能になった。

応 用 例

このシステムの応用例について報告する。データを処理するための解析プログラムとしては、赤池弘次等 [1] によって開発された時系列解析と制御のためのプログラムパッケージ TIMSAC-74 の中の NONST を会話型にし、視聴覚的情報検索システムを利用できるように、機能を拡張させたものを用いる。NONST の持っている元の機能について説明する。このプログラムは、尾崎統等 [2] の局所的に定常な時系列を解析するためのモデルのあてはめを行うためのものである。 N 個の観測値から成る時系列データが与えられたとする。これを長さ n の局所的なブロック $N_1, N_2, \dots, N_l (N_1 + \dots + N_l = N, N_1 = \dots = N_{l-1} = n, 1 \leq N_l \leq n)$ に分割して、各ブロックを定常時系列と考えて、自己回帰モデルをあてはめる。この時、1番目のブロックと2番目のブロックを別々の時系列モデルにあてはめるべきかの判断は、情報基準量 AIC [3] を用いて、次のように行う。1番目、2番目のブロックを別々の自己回帰モデルにあてはめた時の AIC を AIC_1 として、2つのブロックを合わせたものに自己回帰モデルをあてはめた時の AIC を AIC_2 とする。 AIC_1 と AIC_2 の大きさを比較して、 $AIC_1 < AIC_2$ ならば、各ブロックに別々の自己回帰モデルをあてはめる。逆に、 $AIC_2 < AIC_1$ ならば、2つのブロックに同じ自己回帰モデルをあてはめる。次に3番目のブロックの扱い方を説明する。 $AIC_1 < AIC_2$ の場合を考える。2番目のブロックと3番目のブロックを別々の自己回帰モデルにあてはめるべきかどうかを、前と同様に AIC を用いて行えばよい。 $AIC_2 < AIC_1$ の時は、1番目、2番目のブロックを合わせたブロックと3番目のブロックを別々の自己回帰モデルにあてはめるべきかどうかを、やはり AIC を用いて行えばよい。4番目以降のブロックについて考える時も同様である。

NONST の持っているもうひとつの機能は、それぞれのブロックにあてはめた自己回帰モデルの係数を用いて、パワースペクトルの計算を行うことである。

視聴覚的情報検索システムの利点を生かし、次のように NONST を利用している。入力データを磁気テープから読込む。このデータをモニタ装置のチャンネル1に2048データずつ表示させて、先に述べたモニタ装置の機能を利用して解析に必要なデータを作成する。再び、この作

られたデータをモニタ装置に表示させて、データの定常性を調べる。この時、逐次階差をとった結果がチャンネル2に表示されるので、目で見ても、データのレベルが定常になったと判断できるまで階差をとる。このようにしてできた階差を NONST の入力データとする。入力データは、目で見ると定常でも、スペクトル特性の詳細は実際には明らかではないので、局所的に自己回帰モデルをあてはめて行く NONST を解析に用いている。ブロックごとに計算して行く時に、得られる自己回帰モデルのパワースペクトルをモニタ装置に表示させているので、スペクトルの形の変化を見ると、データ全体が定常か、局所的に定常かの判定をすることができる。

定常性の判定に、得られた自己回帰モデルを使ったシミュレーションの結果を音声出力装置を通して、音声として聞けるようにしているのも、音の性質の変化も用いることができる。

このプログラムの使用例として、モニタ装置の表示を、ポラロイドカメラによって、写真にとった結果を示しておく。入力データとしては、あらかじめ決めておいたパラメータ値を持つ3次の自己回帰モデルのシミュレーションを行い、そこから得られた時系列を一回積分したものの（部分和列）を用いている。写真Aは、チャンネル1に入力データ、チャンネル2には、一回だけ階差をとった結果を表示したものである。チャンネル2は、一回階差をとると定常になることを示しているが、これは入力データが、定常時系列のシミュレーションで得られた結果を一回積分したものであるから当然のことである。チャンネル3に表示されているのが、ウィンドである。

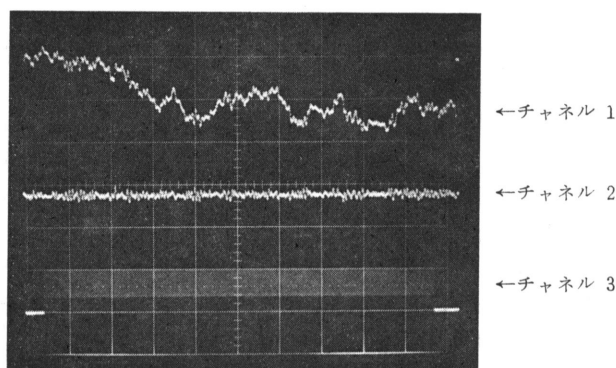


写真 A チャンネル 1; 入力データ
チャンネル 2; 一回階差
チャンネル 3; ウィンド

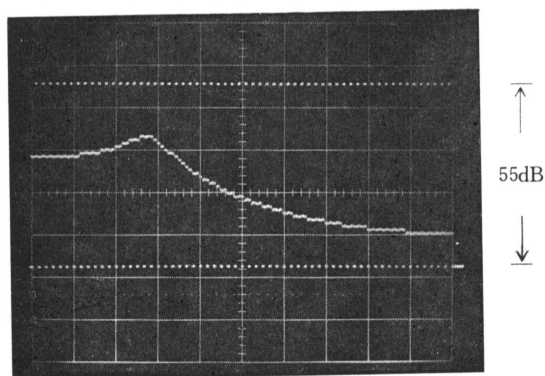


写真 B パワースペクトル

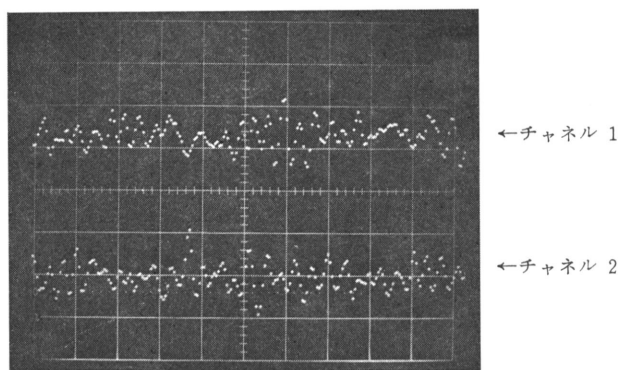


写真 C チャンネル 1; 解析したデータ
チャンネル 2; シミュレーション

写真Bは、写真Aのチャンネル2に表示されているデータを、ウィンドで指定される部分だけ取り出し、NONSTの入力データとして解析した結果得られるパワースペクトルである。水平な、平行線の間隔は、55dBである。写真Cのチャンネル1,2はそれぞれ、計算に用いたデータと解析の結果得られた自己回帰モデルのシミュレーションの結果の表示である。このように、シミュレーションの結果をモニタに表示することもできるし、さらに、データの時間変化の予測をし、その結果をモニタ装置に表示することもできる。

お わ り に

現状では、以上に述べたようなことを視聴覚的情報検索システムを用いて行うことができるまでにプログラムの整備が進められている。

トレンド成分の動きの追跡が主な問題となるようなデータでは、階差をとって定常化することにより、意味がない。この場合には、赤池弘次等[4]によって開発された季節変動調整のためのプログラムBAYSEAの機能の一部を用いて、トレンド成分を除去した後のデータをNONSTの入力とすればよい。このような処理もできるように、システムを発展させるべく目下努力中である。

データ処理に視聴覚的な方法を活用することのメリットは著しく大きい。本稿を読まれた方が、このようなシステムの利用に関心を持たれるようになることを願って本稿を終る。

本研究は第5研究部長赤池弘次の指導の下にすすめられたものである。ソフトウェア開発について一部、矢田誠司君の協力を得た。また、レフェリーの適切な御指摘に感謝します。

参 考 文 献

- [1] Akaike, H., Arahata, E. and Ozaki, T. (1975, 1976). TIMSAC-74, A Time Series Analysis and Control Program Package, *Computer Science Monographs* No. 5, No. 6, Institute of Statistical Mathematics, Tokyo.
- [2] Ozaki, T. and Tong, H. (1975). On the fitting of non-stationary autoregressive models in time series analysis, *Proceeding of the 8th Hawaii International Conference on System Science*, Western Periodical Company.
- [3] Akaike, H. (1974). A new look at the statistical model identification, *IEEE Trans. Automat. Contr.*, AC-19, 716-723.
- [4] Akaike, H. and Ishiguro, M. (1980). BAYSEA, A Bayesian Seasonal Adjustment Program, *Computer Science Monographs* No 13, Institute of Statistical Mathematics, Tokyo.

On an Audio-visual Information Processing
System for Time Series Analysis

Yoshiyasu Hamada

(The Institute of Statistical Mathematics)

In time series analysis the original data may contain wrong or unnecessary parts. These parts must be identified and removed before analysing the data. It seems to be impossible to remove these parts from the whole data by seeing a sequence of numerical values of them. At the Institute of Statistical Mathematics an audio-visual information processing system is developed based on the proposal of Dr. Akaike. In this paper the construction of this system is briefly explained. The software developed for this system is described and its use is demonstrated by examples.