

加速度脈波データの非線形時系列モデリング

Study on Nonlinear Time Series Models for Acceleration Plethysmogram

経営システム工学専攻 西海宏樹
NISHIGAI Hiroki

1 はじめに

現在、様々な分野において時系列解析が適用されている。しかし、時系列データの中には単純なシステムで生成されているものだけとは限らないので、各時系列データの特徴を考慮した時系列解析が必要となってくる。例えば時系列データが非線形モデルに従うと考えられる場合、線形モデリングでは不十分で、非線形モデリングが必要となってくる。

本研究では、非線形性を有していると考えられる指先の加速度脈波（指脈）データに対する時系列モデリングについて検討する。一般に指脈データは非線形性を有していると言われているが、どのような非線形モデルが適切かどうかについては明らかではない。また、指脈データの特徴として強い周期性を示すことがあげられる。しかし、周期は必ずしも一定ではなく、わずかに変化していることが多い。本研究では、非線形モデルとして階層型ニューラルネットワークを用い、指脈データに対して周期の変動を取り入れた非線形時系列モデリングについて考察する。

2 データ

時系列データとして、PC用の測定システム「Salus APG」により実際に測定した指脈データを用いた。ただし、計算時間の短縮のため間引きしたデータに対し解析を行う。4時点以上置いて間引くと時系列の特徴が変化するため、ここでは3時点置きに間引いた。図1に、1本の指脈データの例を示す。系列長は1327である。

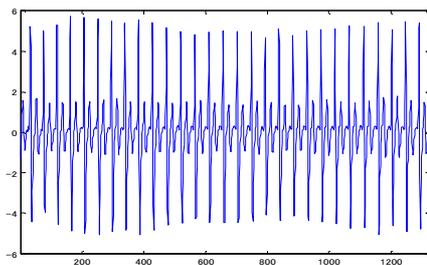


図1: 指脈の時系列データ

また、図2は、図1の指脈データの標本自己相関関数のグラフである。

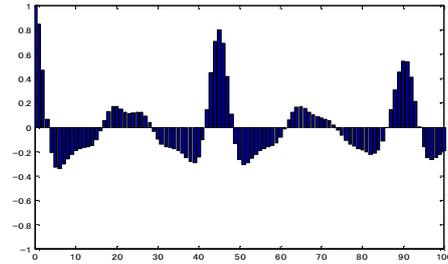


図2: 標本自己相関関数

図1、2からわかるように、指脈データは周期性が強い。そこで、周期性を除去するために、およそその周期だと思われるラグ d の差分演算を行う。ここで d は標本自己相関関数を参考にした。図3に差分をとった指脈データを示す。

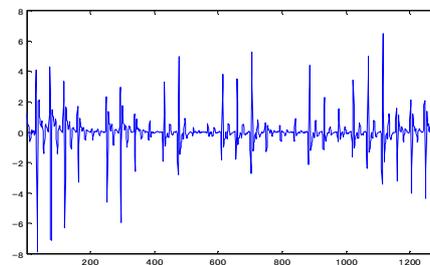


図3: 差分をとった指脈データ ($d = 45$)

図3からわかるように、周期が除去できている部分と、除去できていない部分が混在している。周期が完全に除去できないのは、周期が変化しているためだと考えられる。このような時系列の場合、モデリングのために周期の変化を取り入れることが有効だと考えられる。

3 モデル

対象となる指脈の時系列を $\{x_t\}$ とおく。本研究では、具体的に次のようなモデルを想定した。

(1) AR(90) モデル

90 次の線形自己回帰モデル

$$x_t = a_0 + a_1 x_{t-1} + a_2 x_{t-2} + \dots + a_{90} x_{t-90} + e_t \quad (1)$$

ここで、 $\{e_t\}$ はホワイトノイズである。

(2) NN モデル

非線形自己回帰型ニューラルネットワークモデル

$$x_t = f(x_{t-1}, x_{t-2}, x_{t-19}, x_{t-44}, x_{t-45}, x_{t-89}, x_{t-90}) + e_t \quad (2)$$

ニューラルネットワークの推定においては、変数が増加するとモデルの推定が困難になるので、サブセットのラグを用いることとした。

(3) 周期変動 NN モデル

周期変動を取り入れた NN モデル

周期の変化を取り入れるため、以下のようなモデルを導入する。

最初に確率過程 $\{S_n\}$ の存在を仮定する。 $\{S_n\}$ の取る値は $\{x_t\}$ の周期である可能性のある値で、観測はされない。 $T_n = \sum_{i=1}^n S_i$ として $\{x_t\}$ は $T_{n-1} < t \leq T_n$ のとき次式に従うものとする。(ただし $T_0 = 0$)

$$x_t = f(x_{t-1}, x_{t-2}, x_{t-19}, x_{t-l_4}, x_{t-(l_4+1)}, x_{t-l_6}, x_{t-(l_6+1)}) + e_t \quad (3)$$

$$\begin{cases} l_4 = [aS_n + bS_{n-1}] \\ l_6 = [aS_n + S_{n-1} + bS_{n-2}] \end{cases}$$

ここで a, b は

$$a + b = 1, a : b = (t - T_{n-1}) : (T_n - t)$$

を満たす数で、 $[x]$ は x の少数第一位を四捨五入した整数を意味する。

4 推定

本研究では系列長 1327 のデータのうち $150 \leq t \leq 1027$ の値をモデルの推定に用いた。

AR モデルは周期やパラメータの数を検討し、次数を 90 として最小二乗法を用いて推定した。

NN モデルの推定は、中間層のサイズが 15 でシ

グモイド関数を、出力層は線形関数とした。パラメータの推定方法としては、従来の BP (バックプロパゲーション) 法よりも効率の良いといわれている LM (Levenberg-Marquardt) 法 (MATLAB) を用いる。重みとバイアス項の初期値は、ランダムに 1000 回与え、推定されたモデルの中から平均二乗誤差 (MSE) が最も小さいものを選択した。

また、周期変動 NN モデルについては、まず、周期である可能性のある値をとる確率過程 $\{S_n\}$ を $\{x_t\}$ のピークの間隔を抽出することで推定する。入力ラグ以外は NN モデルと同様の推定方法を用いた。

5 推定結果

推定された $\{S_n\}$ の分布を図 4 に示す。また、推定された各モデルより得られた残差誤差を表 1 に示す。

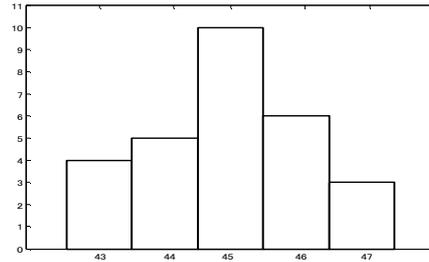


図 4: $\{S_n\}$ の分布

図 4 から、 $\{S_n\}$ は 45 を中心に 43 から 47 まで分布している。このことから周期が変化していることがわかる。

表 1: 残差誤差

	残差誤差
AR(90) モデル	0.00127
NN モデル	0.00148
周期変動 NN モデル	0.000769

NN モデルのあてはまりは AR(90) と比較し、わずかに悪い結果となっているが、どちらのモデルの方が指脈データのモデリングに有効かはわからない。周期変動 NN モデルはさらにあてはまりがよくなっており、周期変動 NN モデルは指脈データのモデリングに有効な可能性が高いといえる。

6 シミュレーション

第5節で推定された各モデルが指脈データに対してどの程度あてはまっているのか、以下のようなシミュレーションを行った。

N 時点までの時系列から $N+1$ の値を推定する一期先の推定値は、

$$\hat{x}_{N+1|N} = f(x_{N+1-l_1}, x_{N+1-l_2}, \dots, x_{N+1-l_p}) \quad (4)$$

によって与えられる (l_1, \dots, l_p はラグ)。ここでは、得られた推定値を次の観測値とみなして逐次的に代入を繰り返し、 k 期先の推定を行うこととする。 N 時点までの時系列による x_{N+k} の推定値を $\hat{x}_{N+k|N}$ とおくと ($k > 0$)、次式が得られる。

$$\hat{x}_{N+k|N} = f(\hat{x}_{N+k-l_1|N}, \hat{x}_{N+k-l_2|N}, \dots, \hat{x}_{N+k-l_p|N}) \quad (5)$$

ただし、 f は第5節で推定された線形自己回帰モデルもしくはニューラルネットワークであり、 $i \leq N$ に対しては $\hat{x}_{i|N} = x_i$ とする。ここでは $N = 150$ として $k = 4850$ までのシミュレーションを行った。

また、 $\{S_n\}$ は指脈データのピークの間隔から推定しているため、 $\{S_n\}$ を推定することの出来ない範囲については、標本自己相関関数よりおよそその周期であると考えられる 45 とした。

シミュレーション結果を以下に示す。

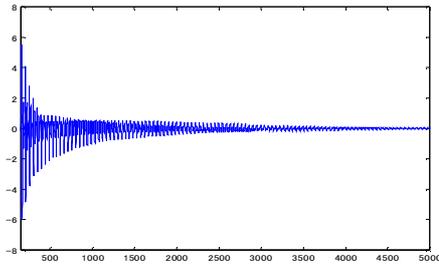


図 5: AR(90) モデルの結果

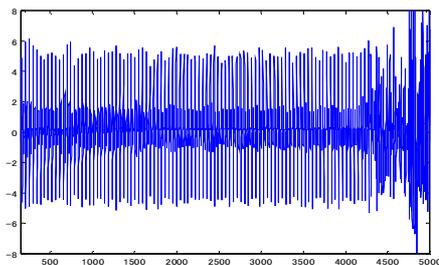


図 6: NN モデルの結果

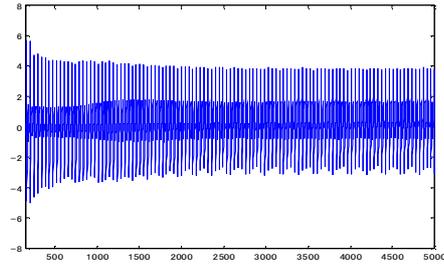


図 7: 周期変動 NN モデルの結果

図5より、AR(90) モデルでシミュレーションを行った結果、指脈データの振幅が減衰していくのがわかる。

図6より、NNモデルでシミュレーションを行った結果、指脈データの特徴をうまく捉えていることがわかる。ただ、シミュレーションを続けると指脈データの波形が崩れ、安定していないことがわかる。

図7より、周期変動 NN モデルでシミュレーションを行った結果、指脈データの振幅が減衰しているが AR(90) モデルほどではなく、安定した推定ができていることがわかる。

これらのことから、指脈データのモデリングには非線形モデルを用いることが有効であることがわかる。さらに、周期変動を取り入れたモデリングも必要となってくるのがわかる。

7 予測

周期変動 NN モデルの有効性を検討するために、指脈データの予測について考察した。

予測に用いる周期変動 NN モデルは、第5節・第6節と同様、系列長 1327 の指脈データのうち $150 \leq t \leq 1027$ の値をモデリングに用いた。残りの $1028 \leq t \leq 1327$ は予測評価のためのテストデータとして用いた。

まず一期先 (1028 時点) の予測値を求め、得られた予測値を次の観測値とみなして逐次的に代入を繰り返して 1327 時点まで予測を行う。ただし、 $\{S_n\}$ は予測せず、指脈データの周期と考えられる値、45 として固定する。加えて、正確な意味での予測ではなくなるが、予測評価のテストデータである $1028 \leq t \leq 1327$ から推定された $\{S_n\}$ を用いて予測も行った。 $\{S_n\}$ を推定することの出来ない範囲については、指脈データの周期と考

られる値、45 とした。

以上の予測結果を次に示す。

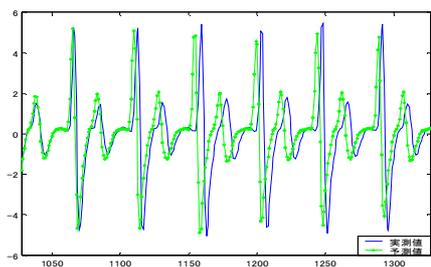


図 8: 予測結果 ($\{S_n\}$ 固定)

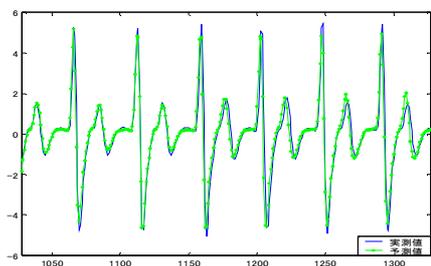


図 9: 予測結果 ($\{S_n\}$ 変動)

表 2: 周期変動 NN モデルでの予測誤差

	予測誤差
$\{S_n\}$ 固定	5.4345
$\{S_n\}$ 変動	0.39938

以上の結果から、 $\{S_n\}$ を固定して予測を行うよりも、 $\{S_n\}$ を実測値通り変動させた方が予測がうまくいっていることがわかる。このことから、周期変動 NN モデルを用いて $\{x_t\}$ をより正確に予測するためには、 $\{S_n\}$ を予測することが必要になってくると考えられる。

そこで、ここでは $\{S_n\}$ が単純マルコフ過程に従っているという仮説をたて、マルコフ性の検定を行う。ただし、1本の指脈データから推定できる $\{S_n\}$ だけでは十分なマルコフ性の検定を行うことができない。よって、ここでは対象指脈データと一緒に連続して測定した同一人物の指脈データ6本を合わせてマルコフ性の検定を行った。 χ^2 検定の結果、単純マルコフ過程であるという仮説は棄却された。従って、より複雑なモデルをあてはめる必要がある。しかし、単純マルコフ過程を想定することは予測の観点からは有効である可能性がある。

8 考察

指脈データに対して、ニューラルネットワークによる非線形モデリングが有効であることがわかった。特に線形自己回帰モデルによるモデリングと比べて、ニューラルネットワークによるモデリングの方が中・長期的にみてうまく当てはまっている。理由として、指脈データは非線形性を有していて、非線形モデルであるニューラルネットワークそのものの中に指脈データの特徴を組み込むことができたためだと考えられる。しかし、周期が変化する指脈データのように複雑なデータに対しては、基本的な NN モデルだけではデータの特徴を完全に組み込むことはできない。その場合、周期変動 NN モデルのような工夫が必要となる。周期変動 NN モデルにおいて $\{S_n\}$ に一定値を用いたものと、変動する実測値を用いたもので比較したとき、実測値を用いた場合の方が $\{x_t\}$ の予測がうまくいっていることから、周期変動 NN モデルそのものの中に指脈データの特徴が組み込むことができたということがわかる。このことから、指脈データのモデリングに対してはニューラルネットワークの中でも、入力ラグの一部を $\{S_n\}$ に依存させ周期変動を取り入れた NN モデルが有効であることがわかった。

本研究では、指脈に関する医学的な知識を前提にしないで、ニューラルネットワークの非線形モデルとしての有効性について検討してきた。指脈データを医学的な面から捉えることについては今後の課題である。また、指脈データに対するニューラルネットワークによる非線形モデリングが有効なことから、指脈データ以外の時系列データへの応用も考えられる。

参考文献

- [1] 若林繁生, 渡邊則生.(2005)「ニューラルネットワークを用いた指尖脈波の時系列モデリング」, 統計関連学会連合大会 講演報告集,24.
- [2] 西海宏樹, 渡邊則生.(2006)「指尖加速度脈波データの時系列モデリング」, 統計関連学会連合大会 講演報告集,163.
- [3] 渡邊則生著「ソフトコンピューティングと時系列解析」シーエーピー出版 2003