ℓ_1 最適化を用いた循環構造を有するシステムのモデリング

石川 健太郎*・丸田 一郎**・足立 修一*

Modeling of Systems with Circulation Using ℓ_1 -norm Optimization

Kentaro ISHIKAWA*, Ichiro MARUTA** and Shuichi ADACHI*

In this paper, a new identification method for systems with a circulation structure is proposed. The output of the system is strongly correlated with that of the previous circulation period under the influence of the substances circulating in the system. Therefore, the model for such system contains higher order terms, which makes it difficult to identify the system when its circulation period is unknown. To deal with the systems with unknown circulation period, the proposed method utilizes the feature extraction method with ℓ_1 -norm optimization technique, which can estimate sparse parameter vector appears in the models with the circulation structure. A numerical simulation is given to show the effectiveness of the proposed method.

Key Words: ℓ_1 -norm optimization, sparsity, system identification, circulation

1. はじめに

本稿では内部に循環構造を有するシステム（物質が周期的に循環するような機構を有するシステム）のモデリングを考える。このようなシステムは化学プラント、生体システム、海洋循環など多岐にわたっており、これらの出力はいずれも 1 周期前の出力の影響を受ける。したがって、これを記述する ARX (Auto-Regressive with eXogenous) モデルなどの式誤差モデルは循環周期に対応した高次の回帰項（循環項）を含む。循環周期が既知であれば循環項を考慮したモデルを用いることで同定を行なうことができるが、一般に正確な循環周期は未知であるため、通常の間定手法では不要な回帰項を含む高次のモデルが推定されるといった問題が生じる。

そこで本稿では、循環構造を有するシステムの回帰項がスパース性を有する、すなわち大半の回帰項が 0 であることに着目し、圧縮センシングにおいて用いられる特徴量抽出手法である ℓ_1 最適化^{1), 2)}を応用することで、適切なモデルを得る手法を提案する。

2. ℓ_1 最適化によるスパース推定の概要

ある未知ベクトル $\mathbf{x} \in \mathbb{R}^n$ が連立方程式 $\mathbf{y} = \mathbf{A}\mathbf{x}$ を満たす場合を考える。ただし、ベクトル $\mathbf{y} \in \mathbb{R}^m$ と行列 $\mathbf{A} \in \mathbb{R}^{m \times n}$ は既知とする。また以下では $\mathbf{x}$ の ℓ_0 , ℓ_1 , ℓ_2 ノルムを、それぞれ $\mathbf{x}$ の非ゼロ要素の個数、 $\mathbf{x}$ の要素の絶対値和、 $\mathbf{x}$ の要素の二乗和平方根と定義し、それぞれ $\|\mathbf{x}\|_0$, $\|\mathbf{x}\|_1$, $\|\mathbf{x}\|_2$ と記す。

このとき、 $\mathbf{x}$ がスパース性を有する、すなわち $\|\mathbf{x}\|_0 \ll n$ であることが既知の場合には、つぎの ℓ_0 ノルム最小化問題を解くこと、すなわち ℓ_0 最適化によって $\mathbf{x}$ の正しい推定値を

Fig. 1 System with circulation

得ることができる。

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x} \quad (1)$$

しかし ℓ_0 最適化問題は一般に NP 困難である。計算論的な困難を回避するために、 ℓ_0 ノルムを ℓ_1 ノルムで置き換えた次式の ℓ_1 ノルム最小化問題を解くこと、つまり ℓ_1 最適化によって近似的に非ゼロ要素の個数を最小化することができる。

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x} \quad (2)$$

ℓ_1 最適化問題の解は $\mathbf{A}$ に関する適切な条件のもとで ℓ_0 最適化問題の解と一致することが示されており、スパース性を有するベクトルの推定手法として有用である²⁾。

3. 問題設定

同定対象とする「循環構造を有するシステム」について説明する。具体例として、Fig. 1 のように、円周状の流路を流れる流体に含まれる物質の濃度を、位置が固定されたセンサにより測定する場合を考える。このとき流出入がなく、かつ物質の拡散が無視できる場合には、ある時刻 k における濃度 $y(k)$ は 1 周前の濃度 $y(k-M)$ に一致する。ただし M は流体が 1 周するのに要する時間（循環周期）であり、簡単のためセンサのサンプリング周期は 1 秒とした。実際には流路の途中で物質の流入があり（以下ではまとめて入力 u として扱う）、また物質の拡散のために濃度 $y(k)$ はダイナミクスをもつので、このようなシステムは、たとえば

$$\begin{aligned} y(k) = & -a_1 y(k-1) - a_2 y(k-2) - a_M y(k-M) \\ & + b_1 u(k-1) + b_2 u(k-2) + w(k) \end{aligned} \quad (3)$$

のように循環構造を考慮した ARX モデルで記述することが

* 慶應義塾大学理工学部 横浜市港北区日吉 3-14-1

** 京都大学大学院情報学研究科 宇治市五ヶ庄

* Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama

** Graduate School of Informatics, Kyoto University, Gokasho, Uji

(Received December 4, 2012)

できる^(注1)．ここで $w(k) \in \mathbb{R}$ は正規性白色雑音である．なお、以下では $a_1 y(k-1)$, $a_2 y(k-2)$ のような循環構造によらない回帰項を「ダイナミクス項」, $a_M y(k-M)$ のような循環構造に対応した回帰項を「循環項」, また u に関する項を「入力項」と呼ぶこととする．

このとき循環周期 M が既知ならば, (3) 式は線形回帰モデルであるため最小二乗法を適用することで θ_0 を推定できる³⁾．しかし一般に循環周期や入力の次数は未知であるため, 通常の同定手法では回帰項, 入力項の最高次の次数に対して十分大きな値 m, n を設定してパラメータベクトルを

$$\theta = \begin{bmatrix} a_1 & a_2 & \cdots & a_m & b_1 & b_2 & \cdots & b_n \end{bmatrix}^T \quad (4)$$

とおき, $y - \Phi\theta$ に最小二乗法を適用して θ を推定する．ただし出力ベクトル y とデータ行列 Φ は時刻 $k = 1, 2, \dots, N$ までの入出力データから以下のように定義される^(注2)．

$$y = \begin{bmatrix} y(m+1) & y(m+2) & \cdots & y(N) \end{bmatrix}^T \quad (5)$$

$$\Phi = \begin{bmatrix} -y(m) & -y(m-1) & \cdots & -y(1) \\ -y(m+1) & -y(m) & \cdots & -y(2) \\ \vdots & \vdots & & \vdots \\ -y(N-1) & -y(N-2) & \cdots & -y(N-m) \\ u(m) & u(m-1) & \cdots & u(m-n+1) \\ u(m+1) & u(m) & \cdots & u(m-n+2) \\ \vdots & \vdots & & \vdots \\ u(N-1) & u(N-2) & \cdots & u(N-n) \end{bmatrix} \quad (6)$$

しかしこの場合, 得られた推定値 $\hat{\theta}$ はスパース性をもたない場合が多く, 不必要に複雑なモデルが推定されてしまう．

本稿の目的は, (3) 式のような循環構造を有するシステム (循環周期は未知) に関する入出力データが与えられたときに, 不要な項を含まない適切なモデルを得ることである．ただし議論を簡単にするために, 事前の解析によりダイナミクス項の次数 p (たとえば (3) 式では $p=2$) は既知とする．

4. ℓ_1 最適化を用いた循環構造を有するシステムの同定

前章で述べたような循環構造を有するシステムを考える．このとき, パラメータベクトル θ の真値はほとんどの要素が 0 のスパースなベクトルである．このような θ の推定手法として, LASSO (Least Absolute Shrinkage and Selection Operator) が知られており²⁾, これは最適化問題,

$$\min_{\theta} \|y - \Phi\theta\|_2 \quad \text{subject to} \quad \|\theta\|_1 \leq \delta \quad (7)$$

(δ は正定数) の解として θ を推定する手法である．

循環構造を有するシステムでは, θ がスパース性を有するのみならず, ダイナミクス項 (次数 p) の係数 $a_1, a_2, \dots, a_p$

が非ゼロの値をもつという特徴を有する．本稿ではこれらの特徴を利用した最適化問題,

$$\min_{\theta} (w_{yh} \|\theta_{yh}\|_1 + \|\theta_u\|_1) \quad \text{subject to} \quad \|y - \Phi\theta\|_2 \leq \varepsilon \quad (8)$$

に基づく推定を提案する．ただし,

$$\theta = \begin{bmatrix} \theta_{yl}^T & \theta_{yh}^T & \theta_u^T \end{bmatrix}^T \quad (9)$$

$$\theta_{yl} = \begin{bmatrix} a_1 & a_2 & \cdots & a_p \end{bmatrix}^T \quad (10)$$

$$\theta_{yh} = \begin{bmatrix} a_{p+1} & a_{p+2} & \cdots & a_m \end{bmatrix}^T \quad (11)$$

$$\theta_u = \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix}^T \quad (12)$$

とおいた．ここで $w_{yh} \in \mathbb{R}$ はユーザが与える非負の重みパラメータであり, これが大きいほど相対的に θ_{yh} の, また小さいほど相対的に θ_u のスパース性が増す．また ε は $y - \Phi\theta$ の出力誤差の上界を定める正定数であり, これもユーザが与えるパラメータである．

このように, ダイナミクス項の係数 θ_{yl} はすべて非ゼロの値をもつことから重みを 0 としてスパース化を行わず, ダイナミクス項を除いた出力項の係数 θ_{yh} と入力項の係数 θ_u に対しては重みを調節することで, 単に $\|\theta\|_1$ の最小化を行なうよりもスパースな解が求まることが期待される．(8) 式を「重み付き ℓ_1 最適化」と呼ぶこととする．

なお, (8) 式は凸計画問題として定式化できるため, 内点法などの既存のアルゴリズムを適用することで数値解を求めることが可能である²⁾．本稿では, 凸計画問題を解く MATLAB パッケージである CVX を利用した^{4), 5)}．

(8) 式を解いて推定値 $\hat{\theta}$ を求めたのち, その中で絶対値の大きい要素を非ゼロ要素とみなし, 対応する Φ の列だけを並べた部分行列を Φ_s とする．そして $y - \Phi_s \theta_s$ に最小二乗法を適用することで推定値 $\hat{\theta}_s$ を得る．このとき, 非ゼロ要素の推定が適切に行なわれていれば $\hat{\theta}_s$ と真値 θ_0 は一致する．

なお, 重み付き ℓ_1 最適化はパラメータベクトルの非ゼロ要素の特定, すなわちモデルに含める回帰項と入力項の決定に焦点を当てており, 係数値自体は最小二乗法により求めていることに注意する．以上の手順を **Algorithm 1** にまとめた．

Algorithm 1 : Weighted ℓ_1 -norm optimization

- 1: Solve the weighted ℓ_1 -norm optimization problem (8) and let the solution be $\hat{\theta}$.
 - 2: Set Φ_s to be the submatrix of Φ containing columns of Φ that correspond to the main peaks of $\hat{\theta}$.
 - 3: Calculate the least squares estimates $\hat{\theta}_s$, that is, $\hat{\theta}_s := \Phi_s^\dagger y$, where $\Phi_s^\dagger \triangleq (\Phi_s^T \Phi_s)^{-1} \Phi_s^T$ is a pseudo-inverse matrix of Φ_s .
-

5. 数値シミュレーション

循環構造を有するシステムの例として, ARX モデル

$$y(k) = 1.5y(k-1) - 0.7y(k-2) + 0.05y(k-20) \\ + 0.075y(k-21) - 0.035y(k-22) + 0.5u(k-1)$$

(注1) ここでは簡単のため 2 次の ARX モデルを仮定した．

(注2) 本稿では $m \geq n$ を仮定したが, $m < n$ の場合も y, Φ は同様に定義される (詳細な定義は省略)．

$$+ 0.25u(k-2) - 0.2u(k-3) + w(k) \quad (13)$$

で記述される対象システムを考える. このときダイナミクス項の次数 $p = 2$ となり, パラメータベクトルの真値 θ_0 は

$$\theta_0 = \begin{bmatrix} a_1 & a_2 & a_{20} & a_{21} & a_{22} & b_1 & b_2 & b_3 \end{bmatrix}^T \\ = \begin{bmatrix} -1.5 & 0.7 & -0.05 & -0.075 & 0.035 & 0.5 & 0.25 & -0.2 \end{bmatrix}^T \quad (14)$$

となる. (13) 式のモデルから得られる入出力データを用いて, 提案法の有効性を検証する. なおサンプリング周期は 1 秒, データ数 N は 1012, 入力 u は M 系列信号, 雑音 $w(k)$ は平均値 0, 分散 1×10^{-4} の正規性白色雑音とした. 入出力データの波形を Fig. 2 に示す. 図では最初の 200 秒を示した.

この入出力データを用いて, u を u の ℓ_2 ノルム, y を y の ℓ_2 ノルムで規格化したものに重み付き ℓ_1 最適化を適用した. なお設定パラメータは $m = 40$, $n = 20$, $w_{yh} = 0.5$, $\varepsilon = 0.009$ とした.

ここで, m , n は実問題において回帰項および入力項の最高次の次数, 22 および 3 が未知であることを考慮し, 十分大きな値として設定した. また w_{yh} は θ_{yh} , θ_u の両方がスパース化されるように, 若干の試行錯誤により決定した. また ε は解が存在する程度に小さい値, 具体的には雑音分散の平方根 (雑音の標準偏差) に近い値として設定した.

得られた結果を Fig. 3 に, またその拡大図を Fig. 4 に, それぞれ示す. また比較のため, 次式の ℓ_2 最適化問題を適用することで得られた推定値も両図に示した.

$$\min_{\theta} (w_{yh} \|\theta_{yh}\|_2 + \|\theta_u\|_2) \text{ subject to } \|\Phi\theta - y\|_2 \leq \varepsilon \quad (15)$$

ただし, 各パラメータの定義とその値は重み付き ℓ_1 最適化と同様とした.

Fig. 4 (b) をみると (15) 式の最適化では不要な項が推定されている. これはシステムのスパース性を考慮せずに, 単に高次のモデルとみなして解を計算しているためと考えられる. 一方, 提案法では θ_0 に対応する次数に大きなピークが現れており, 推定が正しく行なわれていることが確認できた.

このとき, 提案法により得られた, Φ_s に対応するパラメータベクトル θ_s は

$$\theta_s = \begin{bmatrix} a_1 & a_2 & a_{20} & a_{21} & a_{22} & b_1 & b_2 & b_3 \end{bmatrix}^T \quad (16)$$

となり θ_0 の要素と一致した. つづいて $y - \Phi_s \theta_s$ に最小二乗法を適用すると

$$\hat{\theta}_s = \begin{bmatrix} -1.5 & 0.7 & -0.05 & -0.075 & 0.035 & 0.5 & 0.25 & -0.2 \end{bmatrix}^T \quad (17)$$

となり, 正しく θ_0 を推定することができた.

以上の結果より, 循環構造を有するシステムに関する入出力データからの適切なモデリングが可能であると確認できた.

6. おわりに

本稿では, 循環構造を有するシステムの回帰項のスパース性に着目し, ℓ_1 最適化を用いたモデリング手法を提案した.

Fig. 2 Input and output signal of (13)

Fig. 3 Estimation result of parameter vector θ_0

(a) $a_{15}, \dots, a_{25}$.

(b) $b_1, \dots, b_{11}$.

Fig. 4 Enlarged view of Fig. 3

提案法では, ダイナミクス項の次数 p が既知であれば, 循環周期が未知であっても同定を行なえるという特長をもつ. この特長を数値シミュレーションにより確認した. 今後の課題としては, 決定に若干の試行錯誤を必要とする設定パラメータである w_{yh} の明確な設定基準を定めることがあげられる.

参考文献

- 1) E.J. Candès and M.B. Wakin: An introduction to compressive sampling, *IEEE Signal Process. Mag.*, **25-2**, 21/30 (2008)
- 2) 田中利幸: 圧縮センシングの数理, *IEICE 基礎・境界ソサイエティ Fundamentals Review*, **4-1**, 39/47 (2010)
- 3) 足立修一: システム同定の基礎, 東京電機大学出版局 (2009)
- 4) CVX Research, Inc.: CVX: Matlab software for disciplined convex programming, <http://cvxr.com/cvx> (2011)
- 5) M. Grant and S. Boyd: Graph implementations for nonsmooth convex programs, *Recent Advances in Learning and Control*, 95/110, Springer (2008).