

線形システムの同定

1 はじめに

システムの解析，予測，制御，シミュレーション及び異常診断のため，システムの動特性を知る必要がある．システムの入出力信号の観測値に基づいて，ある評価規範のもとでそのシステムをもっともよく記述する数学モデルを決定することをシステム同定という．このような統計学に基づいた同定法は大量かつ複雑なデータ処理を必要とするが，デジタルコンピュータの発達はこのようなデータ処理を可能とした．したがって，この 20 年間に，システムの解析，予測，制御などの基礎となるシステム同定理論の進歩はめざましく，成熟しつつある．現在では，最小二乗法で代表されるシステムのパラメータ推定手法は実システムに使えるようになり，工学の基礎技術の一つに数えられるようになった．

本演習は，2 次の離散時間線形システムのパラメータ推定問題をシミュレーションによって検討する．

2 最小二乗法による線形システムのパラメータ推定

2.1 最小二乗法の導出

Fig. 1 に示すような線形システムの入出力関係を表す数式モデルの記述について考える．

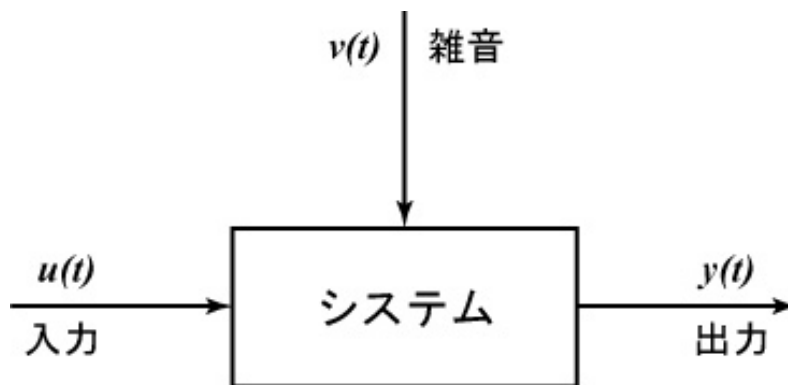


Figure 1: 線形システムの同定のためのモデル.

線形システムの次数 n を既知と仮定すると，その入出力関係は次式のような差分方程式で表される．

$$x_t = - \sum_{i=1}^n a_i x_{t-i} + \sum_{i=1}^n b_i u_{t-i} \quad (1)$$

¹この資料は，学部 3 年以上のレベルの学生が，講義を直接受けなくても，読んでわかるように配慮して作成したものです．最後まで落ち着いて読んでから始めてください．

ただし， u_t と x_t はそれぞれシステムの入出力信号の離散時刻 t におけるサンプル値であり， a_i と b_i はシステムモデルのパラメータである。

ここで，出力信号 x_t には，平均値ゼロの離散値確率雑音 v_t が加わると仮定する．

$$y_t = x_t + v_t \quad (2)$$

一般に，観測雑音 v_t はシステムの入力信号 u_t とは，無相関である．すなわち，

$$E[u_t v_t] = 0 \quad (3)$$

ただし， $E[\cdot]$ は期待値を意味する．

我々の目的は，システムの入出力の観測値 u_t と y_t からシステムのパラメータを推定することである．

ここで，(2) 式を (1) 式に代入すると，

$$y_t = -\sum_{i=1}^n a_i y_{t-i} + \sum_{i=1}^n b_i u_{t-i} + r_t \quad (4)$$

が得られる．ただし， $r_t = \sum_{i=0}^n a_i v_t$ ($a_0 = 1$) は観測雑音による式誤差である．

上式はまた次のようなベクトル形式で表現できる．

$$\begin{aligned} y_t &= \mathbf{z}_t^T \boldsymbol{\theta} + r_t \\ \mathbf{z}_t^T &= [-y_{t-1}, \dots, -y_{t-n}, u_{t-1}, \dots, u_{t-n}] \\ \boldsymbol{\theta}^T &= [a_1, \dots, a_n, b_1, \dots, b_n] \end{aligned} \quad (5)$$

システムのパラメータを推定するため，次のような二乗誤差評価を導入する．

$$J = \sum_{t=1}^N \rho^{N-t} r_t^2 = \sum_{t=1}^N \rho^{N-t} (y_t - \mathbf{z}_t^T \boldsymbol{\theta})^2 \quad (6)$$

ただし， $0 < \rho \leq 1$ は忘却係数である． N を現在時刻とみれば， ρ^{N-t} は過去値になるほど，二乗誤差に対する重みを小さくして推定値への寄与を小さくする役割を果たしており，一種の忘却機能を持っている．この忘却の効果のため，ゆっくり変動する時変パラメータの推定にも適用できる．

J を最小にする必要条件是 $\partial J / \partial \boldsymbol{\theta} = 0$ である．すなわち，

$$\frac{\partial J}{\partial \boldsymbol{\theta}} = -2 \sum_{t=1}^N \rho^{N-t} \mathbf{z}_t y_t + 2 \sum_{t=1}^N \rho^{N-t} \mathbf{z}_t \mathbf{z}_t^T \boldsymbol{\theta} = 0 \quad (7)$$

より， $\boldsymbol{\theta}$ の最小二乗推定値が得られる．

$$\hat{\boldsymbol{\theta}} = \left(\sum_{t=1}^N \rho^{N-t} \mathbf{z}_t \mathbf{z}_t^T \right)^{-1} \sum_{t=1}^N \rho^{N-t} \mathbf{z}_t y_t \quad (8)$$

上式が成り立つためには，

$$\left(\sum_{t=1}^N \rho^{N-t} \mathbf{z}_t \mathbf{z}_t^T \right)^{-1}$$

が存在しなければならない．一般に，入力信号が十分多くの周波数成分を含んでいれば，上記の逆行列が存在する．(8)式が J を最小化しているのを示すのは容易である．

$$J = \left[\boldsymbol{\theta} - \left(\sum_{t=1}^N \rho^{N-t} \mathbf{z}_t \mathbf{z}_t^T \right)^{-1} \sum_{t=1}^N \rho^{N-t} \mathbf{z}_t y_t \right] \left(\sum_{t=1}^N \rho^{N-t} \mathbf{z}_t \mathbf{z}_t^T \right) \left[\boldsymbol{\theta} - \left(\sum_{t=1}^N \rho^{N-t} \mathbf{z}_t \mathbf{z}_t^T \right)^{-1} \sum_{t=1}^N \rho^{N-t} \mathbf{z}_t y_t \right] + \sum_{t=1}^N \rho^{N-t} y_t^2 - \left(\sum_{t=1}^N \rho^{N-t} \mathbf{z}_t y_t \right)^T \left(\sum_{t=1}^N \rho^{N-t} \mathbf{z}_t \mathbf{z}_t^T \right)^{-1} \sum_{t=1}^N \rho^{N-t} \mathbf{z}_t y_t \quad (9)$$

と書けるので， $\hat{\boldsymbol{\theta}}$ は確かに上式を最小にしている．

2.2 推定値の漸近的性質

データ数が多くなるにつれて，確率極限

$$\text{plim}_{N \rightarrow \infty} \hat{\boldsymbol{\theta}} = \boldsymbol{\theta} \quad (10)$$

が成立するとき， $\hat{\boldsymbol{\theta}}$ は一致推定量であるという．

ここでは，忘却係数 $\rho = 1$ ，すなわち過去のデータを忘却しない場合について考察する．(8)式の推定値を(5)式に代入して，確率極限をとると，

$$\begin{aligned} \text{plim}_{N \rightarrow \infty} \hat{\boldsymbol{\theta}} &= \left(\text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{z}_t \mathbf{z}_t^T \right)^{-1} \left(\text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{z}_t \{ \mathbf{z}_t^T \boldsymbol{\theta} + r_t \} \right) \\ &= \boldsymbol{\theta} + \left(\text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{z}_t \mathbf{z}_t^T \right)^{-1} \text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{z}_t r_t \end{aligned} \quad (11)$$

が得られる．したがって， $\hat{\boldsymbol{\theta}}$ が一致推定量であるためには，ベクトル \mathbf{z}_t と式誤差 r_t は無相関でなければならない，すなわち，

$$\text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{z}_t r_t = 0 \quad (12)$$

という条件が成り立たなければならない．一般に上式が成立することが期待できないので，観測雑音のレベルが高い場合，最小二乗推定は一致推定量でなく，偏り(バイアス)を持つ．

2.3 補助変数法によるパラメータの一致推定

ベクトル \mathbf{z}_t と同サイズの補助変数ベクトル \mathbf{m}_t を導入し，次のように補助変数法による推定値を得る．

$$\hat{\boldsymbol{\theta}}_{IV} = \left(\sum_{t=1}^N \rho^{N-t} \mathbf{m}_t \mathbf{z}_t^T \right)^{-1} \sum_{t=1}^N \rho^{N-t} \mathbf{m}_t y_t \quad (13)$$

そこで，忘却係数 $\rho = 1$ ，すなわち過去のデータを忘却しない場合における補助変数推定値の漸近的性質について考察する．上式の推定値を(5)式に代入して，確率極限をとると，

$$\begin{aligned} \text{plim}_{N \rightarrow \infty} \hat{\boldsymbol{\theta}} &= \left(\text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{m}_t \mathbf{z}_t^T \right)^{-1} \left(\text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{m}_t \{ \mathbf{z}_t^T \boldsymbol{\theta} + r_t \} \right) \\ &= \boldsymbol{\theta} + \left(\text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{m}_t \mathbf{z}_t^T \right)^{-1} \text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N \mathbf{m}_t r_t \end{aligned} \quad (14)$$

が得られる。

以上より明らかに，補助変数ベクトル m_t が次のような条件を満足すれば，パラメータ推定値は一致推定量である。

(1) 逆行列

$$\left(\text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N m_t z_t^T \right)^{-1}$$

が存在する。

(2) 補助変数ベクトル m_t と式誤差 r_t は無相関である。すなわち，

$$\text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N m_t r_t = 0 \quad (15)$$

が成り立つ。

補助変数ベクトルの幾つかの構成：

(1) 遅延された入力を利用する。

$$m_t = [u_{t-d}, \dots, u_{t-d-n}, u_{t-1}, \dots, u_{t-n}] \quad (16)$$

ただし， $d > n$ 。

(2) 遅延された出力を利用する。

$$m_t = [y_{t-1-d}, \dots, y_{t-n-d}, u_{t-1}, \dots, u_{t-n}] \quad (17)$$

ただし， d の選び方は，明確ではない。

(3) 推定された出力を利用する。

$$m_t = [-\hat{x}_{t-1}, \dots, -\hat{x}_{t-n}, u_{t-1}, \dots, u_{t-n}] \quad (18)$$

ただし，

$$\hat{x}_t = -\sum_{i=1}^n \hat{a}_i \hat{x}_{t-i} + \sum_{i=1}^n \hat{b}_i u_{t-i} \quad (19)$$

ここで， \hat{a}_i 及び \hat{b}_i は何らかの方法で得たパラメータの推定値である（一致推定量でなくてもいい）。

本演習は(3)番目の構成法を利用する。詳細は後述する。

3 逐次推定アルゴリズム

前節で導出した最小二乗法および補助変数法は，データを蓄えておき，データ行列の逆行列を計算して推定値を得るバッチ処理である。このような方法は繰り返し計算を必要としないことから one-shot 法またはオフライン法とも呼ばれている。しかし，システムの実時間での監視や，異常の検出と診断には不向きである。

いま， N 個のデータから得られる推定値を $\hat{\theta}_N$ とする。そこで， $\hat{\theta}_N$ を過去の推定値及び時刻 N でのデータより求めることを考える。新しくデータが得られる度に直前の推定値を修正していく方式（逐次計算式）が実現されれば，オンライン推定や実時間推定への適用が可能である。

具体的導出過程は紙面の制限のため，参考文献に委ねたいが，最小二乗法と補助変数法に関する逐次推定アルゴリズムは以下ようになる。

$$\begin{aligned}
\hat{\theta}_N &= \hat{\theta}_{N-1} + \mathbf{L}_N \varepsilon_N \\
\varepsilon_N &= y_N - \phi_N^T \hat{\theta}_{N-1} \\
\mathbf{L}_N &= \frac{\mathbf{P}_{N-1} \psi_N}{\rho_N + \phi_N^T \mathbf{P}_{N-1} \psi_N} \\
\mathbf{P}_N &= \frac{1}{\rho_N} \left[\mathbf{P}_{N-1} - \frac{\mathbf{P}_{N-1} \psi_N \phi_N^T \mathbf{P}_{N-1}}{\rho_N + \phi_N^T \mathbf{P}_{N-1} \psi_N} \right]
\end{aligned} \tag{20}$$

ただし， \mathbf{P}_N は $2n \times 2n$ の行列で， \mathbf{L}_N は $2n \times 1$ のベクトルである．
上式において，逐次最小二乗法は次のように与えられる．

$$\phi_N = \mathbf{z}_N, \quad \psi_N = \mathbf{z}_N \tag{21}$$

また，逐次補助変数法は次のように与えられる．

$$\phi_N = \mathbf{z}_N, \quad \psi_N = \mathbf{m}_N \tag{22}$$

ここで，下記の点に留意しておくべきである．

- (1) 推定値 $\hat{\theta}_N$ は，1 時刻前の推定値 $\hat{\theta}_{N-1}$ に，式誤差の推定値 $y_N - \phi_N^T \hat{\theta}_{N-1}$ に比例した修正項を加えることによって生成される．したがって，バッチ処理に比べて記憶容量が少なくすむ．また，逐次計算では逆行列の計算を必要としない．
- (2) 上述の逐次アルゴリズムを実行するためには初期値の設定が必要である．通常，次のように選べばよい．

$$\begin{aligned}
\hat{\theta}_0 &= \mathbf{O} \\
\hat{\mathbf{P}}_0 &= \alpha \mathbf{I}
\end{aligned} \tag{23}$$

ただし， $\alpha = 10^3 \sim 10^5$ ．

- (3) 推定された出力を使って補助変数を構成する場合，推定されたシステムの安定性をチェックしなければならない．ここで，補助変数ベクトルは次のように構成されたとする．

$$\mathbf{m}_N = [-\hat{x}_{N-1}, \dots, -\hat{x}_{N-n}, u_{N-1}, \dots, u_{N-n}] \tag{24}$$

ただし，

$$\hat{x}_N = - \sum_{i=1}^n \bar{a}_{i,N-1} \hat{x}_{t-i} + \sum_{i=1}^n \hat{b}_{i,N-1} u_{t-i} \tag{25}$$

ここで， $\hat{b}_{i,N-1}$ は (20) 式の逐次推定アルゴリズムから得られた時刻 $N-1$ における推定値である． $\bar{a}_{i,N-1}$ は，安定性を考慮して，(20) 式の逐次推定アルゴリズムから得られた時刻 $N-1$ における推定値 $\hat{a}_{i,N-1}$ を修正したものである．時刻 N において，推定出力 \hat{x}_N を計算するとき，上記のシステムの極が安定でなければ，値が発散してしまう．したがって，システムの極，すなわち，

$$z^n + \bar{a}_{1,N-1} z^{n-1} + \dots + \bar{a}_{n,N-1} = 0 \tag{26}$$

の解が単位円の内部に入れば（複素数の半径が 1 より小さい），システムが安定である．

時刻 N において，推定出力 \hat{x}_N を計算するとき， $\hat{a}_{i,N-1}$ ($i = 1, 2, \dots, n$) で構成されたシステムの極が不安定であった場合，対策として， $\hat{a}_{i,N-1}$ ($i = 1, 2, \dots, n$) を安定な領域に射影することによって修正する．すなわち， $\hat{a}_{i,N-1}$ ($i = 1, 2, \dots, n$) が安定な極を構成していれば， $\bar{a}_{i,N-1} = \hat{a}_{i,N-1}$ ($i = 1, 2, \dots, n$) とする．そうでなければ， $\hat{a}_{i,N-1}$ ($i = 1, 2, \dots, n$) を安定な領域へ修正する．アルゴリズムは下記のとおりである．

Step a: If $\hat{a}_{i,N-1}(i = 1, 2, \dots, n)$ stable then
 $\bar{a}_{i,N-1} = \hat{a}_{i,N-1}(i = 1, 2, \dots, n)$
 Stop
 Else
 $\bar{a}_{i,N-1} = \bar{a}_{i,N-2} + 0.5 * (\hat{a}_{i,N-1} - \bar{a}_{i,N-2})$
 Endif
 Step b: If $\bar{a}_{i,N-1}(i = 1, 2, \dots, n)$ stable then
 Stop
 Else
 $\bar{a}_{i,N-1} = \bar{a}_{i,N-2} + 0.5 * (\bar{a}_{i,N-1} - \bar{a}_{i,N-2})$
 Endif
 Step c: Go to Step b.

注意：上述の修正は，推定出力の計算だけのために行われることに注意されたい．(20) 式の逐次推定アルゴリズムに対しては，推定されたシステムの安定性の如何に関わらず，このような修正をしない．

このように，システムの出力とパラメータを交互に推定することは Bootstrap (靴のひもかけ) 法という．

(4) 忘却係数 ρ_N も (20) 式の推定アルゴリズムの性能を大きく左右する．一般に，忘却係数が小さいほど，忘却機能が強く働くので，推定速度が速く，パラメータの変動にも対応できるが，システムに関する情報を速く忘却するので，観測雑音に弱く，推定値が振動的になる．通常，時不変システムのパラメータ推定において， $\rho_N = 1$ としても，差し支えないが，忘却係数を時間の経過とともに 1 に近づくように設定した方が，パラメータの推定値が速く真値へ収束する．

$$\rho_N = (1 - 0.01)\rho_{N-1} + 0.01, \quad \rho_0 = 0.95 \quad (27)$$

すなわち，最初の内に忘却係数を小さくすることによって推定速度を上げるが，推定値がある程度その真値に近づいたら，雑音の影響を軽減するために，忘却係数を 1 とすることによって推定速度を落とす．

一方，時変システムのパラメータ推定において， $\rho_N = 0.97 \sim 0.999$ と，1 よりやや小さい定数にすればよい．忘却係数が小さいほど，時変パラメータへの追従速度が速くなるが，雑音の影響を受けやすくなる．いずれにしても，雑音のレベルが高い場合，時変システムのパラメータ推定は困難である．

4 シミュレーション実験

次のような 2 次のシステムを考える．

$$\begin{aligned} x_t &= -a_1x_{t-1} - a_2x_{t-2} + b_1u_{t-1} + b_2u_{t-2} \\ y_t &= x_t + v_t \end{aligned} \quad (28)$$

ただし，

$$a_1 = -1.5, \quad a_2 = 0.7, \quad b_1 = 1.0, \quad b_2 = 0.5 \quad (29)$$

ここで，入力信号を $[0, 1]$ の範囲で一様分布する乱数とする．観測雑音はゼロ平均 (確認すること) の乱数で，入力信号と無相関 (確認すること) である．

雑音のレベルは一般に雑音信号比 (Noise Signal Ratio, NSR) で評価される．ここで，NSR は雑音 v_t の標準偏差と真の出力信号 x_t の標準偏差との比である．

上記の要領でシステムのシミュレータを作成すること。ただし、信号の初期値をすべてゼロとする。特に、雑音に課する仮定条件を確認すること。同じ実験をもう一回確認する場合、乱数の種数を前回のと同じでなければ、前回の結果が再現できないことに注意すること。

4.1 逐次最小二乗法によるパラメータ推定

同じ入力信号に対して、出力の雑音信号比 NSR をそれぞれ 0%、10%、20% とし、入出力信号の観測値 $\{u_t, y_t\} (t = 1, 2, \dots, N)$ を用いて、パラメータを逐次最小二乗法で推定せよ。ただし、 N は十分大きな整数 (2000 ~ 3000) とする。時刻を横軸として、各推定パラメータの挙動を、パラメータの真値とあわせて図示せよ。観測雑音によるパラメータ推定値への影響を考察せよ。

4.2 逐次最小二乗法による時変系のパラメータ推定

次のような 2 次の時変システムを考える。

$$\begin{aligned} x_t &= -a_{1,t}x_{t-1} - a_{2,t}x_{t-2} + b_{1,t}u_{t-1} + b_{2,t}u_{t-2} \\ y_t &= x_t + v_t \end{aligned} \tag{30}$$

ただし、

$$\begin{aligned} a_{1,t} &= -1.0 + 0.5 \sin(2\pi t/2000) \\ a_{2,t} &= 0.7 \\ b_{1,t} &= 1.0 \\ b_{2,t} &= 0.5 + 0.4 \cos(2\pi t/2000) \end{aligned} \tag{31}$$

同じ入力信号に対して、出力の雑音信号比 NSR をそれぞれ 0%、5%、10% とした場合、逐次最小二乗法で上記のシステムの時変パラメータを推定せよ。時刻を横軸として、各推定パラメータの挙動を、パラメータの真値とあわせて図示せよ。忘却係数及び観測雑音によるパラメータ推定値への影響を考察せよ。

4.3 逐次補助変数法によるパラメータ推定

逐次最小二乗法によるパラメータ推定の実験と同じ実験条件で、パラメータを逐次補助変数法で推定せよ。時刻を横軸として、各推定パラメータの挙動を、パラメータの真値とあわせて図示せよ。補助変数法と最小二乗法による結果を比較し、検討せよ。

参考文献

相良, 秋月, 中溝, 片山 (1981): システム同定 (計測自動制御学会).
 中溝高好 (1988): 信号解析とシステム同定 (コロナ社).
 相良, 和田, 中野 (1992): デジタル制御の基礎 (コロナ社).
 足立修一 (1993): ユーザーのためのシステム同定理論 (計測自動制御学会).
 片山徹 (1994): システム同定入門 (朝倉書店).