ディジタル制御 第 13 回

システム同定

はじめに

- システム同定とは, 制御対象の数学モデルを定めることをいう. 制御理論それ自体とは趣が異なるが, モデルベースドコントロールでは不可欠な工程である.
- 今回の講義では, 参考文献にしたがい, システム同定の導入部分を紹介する.

- システム同定のためには確率および確率過程の知識が不可欠であるが, この講義ではこれらは既知と仮定する. また, 以下の議論では因果的な信号およびシステムのみを検討の対象とし, これについては一々断らない.

- 伝達関数を取り扱うときには, 前回に引き続き, 独立変数を z^{-1} とする (このようにした方が因果的なフィルタリングとの相性が良いため). また, 伝達関数の分母あるいは分子の係数をまとめたものをパラメータベクトルと解釈し, $\boldsymbol{\theta}$ と書くことがある. $G(z^{-1}, \boldsymbol{\theta})$ などの標記はこの意味である.

- $y(n) = G(z^{-1}, \boldsymbol{\theta})u(n)$ という式は, 以下の式が省略されたものと解釈する.

$$
\begin{aligned}
&y(n) \\
&= \mathcal{Z}^{-1}\left[G(z^{-1}, \boldsymbol{\theta})\mathcal{Z}\left[(u(n))_{n\in\mathbb{N}}\right]\right]\Big|_{\text{時刻 } n \text{ で評価}}
\end{aligned}
$$

- 信号 $(u(n))_{n\in\mathbb{N}}$ の **自己相関関数** および信号 $(u(n))_{n\in\mathbb{N}}$ と信号 $(y(n))_{n\in\mathbb{N}}$ の**相互相関関数**を それぞれ以下によって定義する:

$$\phi_{uu}(n,\tau) = \mathrm{E}[u(n+\tau)u(n)],$$
$$\phi_{yu}(n,\tau) = \mathrm{E}[y(n+\tau)u(n)].$$

ただし, $\mathrm{E}[\cdot]$ は集合平均である.

- 以下で自己および相互相関関数が現れるときには, 対象となる信号とシステムは**定常的**, すなわち相関関数が n に依存しないと仮定する. この仮定の下で, 第一引数を略し,

$$\phi_{uu}(\tau)$$

$$\phi_{yu}(\tau)$$

などと書く.

- 信号についてはエルゴード性を仮定する．これは, 信号に関する集合平均を時間平均で置き換えられるという意味である.

- これらの仮定の下で，例えば，次式が成り立つ．

$$
\begin{aligned}
\phi_{uu}(\tau) &= \mathrm{E}[u(n+\tau)u(n)] \\
&= \lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} u(k+\tau)u(k)
\end{aligned}
$$

モデル

- 今まで, 制御対象や補償器などの装置の 伝達関数表現や 状態空間表現に基づき 制御系の解析と設計について議論してきた.

- 伝達関数表現も，状態空間表現も，物理的実在としての装置そのものではなく，それを 抽象化 および 単純化 して数式で (近似的に) 表現したものである.

- 物理的実在物などを, 特定の目的で, 特定の観点から, 指定した形式で表現したものを, モデルと呼ぶ.

- たとえば, 制御における伝達関数モデルでは
  ...
  - ▷ 目的: 制御
  - ▷ 観点: 入出力関係
  - ▷ 形式: s, z(あるいは z^{-1}) の有理式

- 制御には, モデルに基づく モデルベースト制御 (Model Based Control; MBC) とモデルに基づかない モデルフリー制御 (Model Free Control; MFC) がある. 適切に構築された MBC に基づく制御系は MFC のそれより高性能であることが見込まれる. 一方で MBC は MFC より 解析や設計に手間がかかる.

- モデルを構築することをモデリングと呼ぶ. モデリングの手法は, 以下の 3 種類に大別される.
  - ▷ 物理モデリング
  - ▷ ブラックボックスモデリング
  - ▷ グレーボックスモデリング

- 物理モデリング (第一原理モデリング, ホワイトボックスモデリングとも呼ばれる) は, 物理法則などの 対象の挙動を支配する法則から直接モデルを導出する方法である. 精密なモデルが得やすい一方で, モデルが複雑すぎて制御には使いにくいこともある. この手法を取るためには, 対象の挙動を支配する法則に関する十分な知識が必要.

- ブラックボックスモデリング は, 対象の入出力関係のデータに基づき, 統計的手法によって モデルを導出する方法．複雑な対象では, その挙動を支配する法則に関する知見が不十分なことも多く, そのような場合にこの種の手法が必要となる．モデルは多くの場合単純で制御に使いやすい．実験によるデータ取得が可能な対象が既に存在することが前提となり, 装置自体を設計中の場合には使えない．

- グレーボックスモデリング は 物理モデリングとブラックボックスモデリングを組み合わせたもので, 対象の挙動を支配する法則が部分的にわかっているとき, 既知の部分はそのまま利用し, 未知の部分は実験によって定める手法である. 運動方程式によって マニピュレータの状態方程式の形を決めてから入出力関係に基づいて 状態方程式のパラメータを定めることはこれに該当する.

- ブラックボックスモデリングとグレーボックスモデリングの境界は曖昧.

- 物理モデリングは 個別の物理法則の問題なので, この講義ではこれ以上議論しない.

- 以下では, おもにブラックボックスモデリングについて述べる. ただし, その手法の多くは グレーボックスモデリングにも適用可能である.

モデルの分類

- 第 1 回で述べた動的システムの分類は そのままモデルの分類になっている.

- 別の観点からの分類になるが，その挙動が有限のパラメータによって特徴付けられるモデルをパラメトリックモデルと呼び，それ以外をノンパラメトリックモデルと呼ぶ.

- 伝達関数, 状態空間表現はパラメトリックモデルである.
- 周波数特性などの波形自体をモデルと解釈すると, ノンパラメトリックモデルになる.
- 人の思考を数式を使わずに表現する思考モデル, システムの特性をブロック線図やフローチャートなどの図で表現したグラフモデルもモデルの一種.

- 制御におけるモデルに要求されることは…
  - ▷ 対象の挙動を適切に表現できる
  - ▷ 制御系の解析と設計に使いやすい
  - ▷ 不確かさの表現ができる

第一の条件を満たすモデルは複雑になりがちだが, 第二の条件を満たすためには モデルは簡単な方がよい. バランスが重要.

システム同定

- システム同定 とは, ブラックボックスモデリング あるいはグレーボックスモデリングが選択された状況の下で, 対象とする動的システムの入出力関係の測定値 (多くの場合外乱が含まれる) を用い, 特定の観点から 対象と同一と見做すことができる数学モデルを作成することをいう.

- システム同定の理論は, 1 入力 1 出力線形システムについてはおおむね完成しており (この講義ではこの部分のみを取り扱う), 多入力多出力線形システムについても相当の研究の蓄積がある. 一方, 非線形システムの同定は研究段階.

- システム同定は時間領域でおこなわれることも周波数領域でおこなわれることもあるが, この講義では, 周波数特性の同定を除き, 前者のみを取り扱う.

システム同定の手順

- システム同定の手順は以下の通りである．これらの詳細は後で述べる．

  1. 準備: ステップ応答や周波数特性を測定し, 対象の大まかな特徴を把握する.
  2. 実験設計: ハードウェア, 入力, サンプリング周期等の選定.

3. システム同定実験: 対象の入出力関係のデータを収集.
4. データの前処理: データの切り出し, 異常値の除去, フィルタリング等.
5. 構造同定: モデルの形を決める.
6. モデルの同定: データに同定法を適用してモデルを決定する.
7. モデルの妥当性の評価.

実験設計

- このステップでは, ハードウェア, 入力, サンプリング周期等の選定をおこなう. ハードウェアの選定については特に述べることはないので, 残りの 2 点について述べる.

- **入力の選定**では, 対象からなるべく多くの情報を取得する入力を選定する．制御における入力とは役割が異なるので注意.

- 例として, $(u(n))_{n\in\mathbb{N}}$ を入力, $(y(n))_{n\in\mathbb{N}}$ を出力とし, 有限長のインパルス応答 $(h(n))_{n=0,\ldots,L}$ を持つシステム:

$$y(n) = \sum_{k=0}^{L} h(k)u(n-k)$$

を考える.

- システム同定では, 入出力の 測定値から

$$(h(n))_{n=0,\dots,L}$$

が一意的に 決定されるように入力を定める必要がある.

- $n \geq L$ に対し,

$$\boldsymbol{\theta} = (h(0), \ldots, h(L))^T$$
$$\boldsymbol{u}(n, n-L) = (u(n), \ldots, u(n-L))$$

とおく.

- $y(L), \ldots, y(L+K)$ をまとめると，以下のようになる．

$$\begin{pmatrix} y(L) \\ y(L+1) \\ \vdots \\ y(L+K) \end{pmatrix} = \begin{pmatrix} \boldsymbol{u}(L,0) \\ \boldsymbol{u}(L+1,1) \\ \vdots \\ \boldsymbol{u}(L+K,K) \end{pmatrix} \boldsymbol{\theta}$$

- 右辺左側の行列を $\mathbf{\Phi}(K)$ とする．同定が可能であるためには, $\mathbf{\Phi}(K)$ がフルランクであることが必要かつ十分である．

- $k \geq l$ に対し, $\boldsymbol{u}(k,l) = (u(k), \ldots, u(l))$ とし,

$$\boldsymbol{\Phi}(M,n) = \begin{pmatrix} \boldsymbol{u}(M-1,0) \\ \vdots \\ \boldsymbol{u}(M+n-1,n) \end{pmatrix}$$

とする.

- 同定の可否に関連し, PE 条件と呼ばれる条件を述べる.

$$\lim_{n\to\infty} \frac{1}{n}\mathbf{\Phi}^T(M,n)\mathbf{\Phi}(M,n)$$

が正定となることを次数 M の **PE** 条件と言う. M が同定すべきパラメータ数以上であることが同定が可能であるための必要条件である.

- 実用上は，同定のための入力には，擬似乱数を用いた2値信号が用いられることが多いが，正弦波などが用いられることもある．

- **サンプリング周期の選定**の仕方には厳密な規則はない．次ページに目安を列挙する．十分短かいサンプリング周期でデータを収集し，事後のデータ処理でサンプリング周期を変更する手法が取られることもある．

目安 1 準備の時点で測定した周波数特性を用い, そのバンド幅 (ゲインが定常ゲインの $1/\sqrt{2}$ となる角周波数) の 10 倍程度とする.

目安 2 対象のステップ応答が定常応答の 95% に到達する時刻を T_{95}, サンプリング周期を T_S としたとき,

$$\frac{1}{15}T_{95} \leq T_s < \frac{1}{4}T_{95}$$

のように T_S を選ぶ.

システム同定実験

- システム同定実験は, 個別の対象に 事前に選定した入力を適用してデータを収集する実験である. ここがもっとも大変な工程であり, 手間と費用がかかる上, 安全に関する配慮も必要であるが, 一般論として述べることはない.

データの前処理

- モデルの同定はデータに基づいて行なわれるが, モデルの有用性を高めるために, データに何らかの前処理を施すことが一般的. ただし, 恣意的なデータの処理によって問題が発生するリスクもあるため, 注意を要する工程でもある. 科学倫理的な観点から前処理 (の一部) が許容されない状況もあり得るので注意.

- 前処理の際には, まずデータに関する時間領域および周波数領域のグラフを描画し, 全体的な傾向を把握した上で, 必要があればデータに手を加える (すべてを必ず実施するというわけではない).

- 時間領域における典型的な処理は, 異常値の除去, 欠損データの補間, データの切り出し, スケーリングである. これらの意味は…

- ▷ **異常値 (外れ値, アウトライアー**とも呼ぶ) は, データの全体的な分布から極端に外れたデータを指す実用上の概念. 少数の異常値によってモデルのパラメータが大きく影響を受けることがあるので, 除去が望ましいが, 恣意的なデータの加工とならないよう注意する必要がある.

- ▷ 欠損データとは, データが欠落している部分のことを言う. これは, 通信障害など様々な原因で発生する. 典型的な対処は補間である.

- ▷ データの切り出しは, 時系列データの一部を抽出して利用することをいう. 異常値や欠損データない箇所を切り出すこともあれば, 外乱の影響が少ないと思われる箇所など条件が良い箇所を切り出すこともある.

▷ スケーリングとは, 入力と出力のどちらか一方の値を全体的に拡大ないし縮小することをいう. 入力と出力は一般に異なる物理量で, それらが取る値の範囲が大きく異なることがあるが, これは同定に悪影響を与える. そこで, 入力と出力が同程度の範囲に分布するように片方を拡大あるいは縮小するのがこの処理である.

- 周波数領域における典型的な処理は, 低周波および高周波外乱の除去と, サンプリングレートの変換. 高周波外乱の除去が必要であることは自明であろうが, 低周波側でも, センサの物理的特性などが原因となる定数あるいは緩やかに変動する外乱が含まれることがあり, 外乱の除去が必要となる. 低域通過フィルタと高域通過フィルタの組み合わせなどによって実現できる.

事前に高いサンプリング周波数でデータを取ってから 制御の目的に合わせてサンプリング周波数を変換することもある．このための処理をデシメーションと呼ぶ．デシメーションはディジタル信号処理の範疇なので，この講義ではこれ以上述べない．

構造同定

- このステップでは, まずノンパラメトリックモデルを用いるかパラメトリックモデルを用いるかを選択する．前者を選択した場合はここで終了．後者を選択した場合は, 更にモデルの構造と次数を選択する．

- パラメトリックモデルには, 非線形あるいは多入力多出力システムまで含めると, 多くのバリエーションがあるが, この講義は, 1入力1出力離散時間線形時不変システムで用いられる代表的なモデルの紹介に留める.

- 1入力といっても，実は外乱も入力の一種なので(ただし値は未知)，外乱と同定用の入力の対象への影響を分けて考える必要がある．

- $(\nu(n))_{n\in\mathbb{N}}$ を期待値零の白色雑音とし, システムに印加される外乱 $(w(n))_{n\in\mathbb{N}}$ は $(\nu(n))_{n\in\mathbb{N}}$ を以下のフィルタに通したものとする:
  $A(z^{-1}) = 1 + a_1 z^{-1} + \cdots + a_N z^{-N}$,
  $C(z^{-1}) = 1 + c_1 z^{-1} + \cdots + c_L z^{-L}$ として,

$$w(n) = \frac{C(z^{-1})}{A(z^{-1})}\nu(n).$$

- システム同定の分野では, 大抵の場合, $A(z^{-1})$ と $C(z^{-1})$ の定数項を 1 と仮定する (上式の赤字の部分).

- 上述のモデルは, 更に以下の 3 種類に分類される.
  - **AR モデル** $C(z^{-1}) = 1$ であるもの.
  - **MA モデル** $A(z^{-1}) = 1$ であるもの.
  - **ARMA モデル** それ以外

- AR は Auto Regressive, MA は Moving Average の略である. MA モデルの形のフィルタを, インパルス応答が有限時間で零に収束するので, Finite Impulse Response (FIR) フィルタと呼ぶことが多い. また, AR モデルあるいは ARMA モデルの形のフィルタを, インパルス応答が無限時間持続するという意味で, Infinite Impulse Response (IIR) フィルタと呼ぶことが多い.

- システム同定で用いる典型的な同定モデルは, 上述の AR モデルや ARMA モデルに同定用の外生入力を印加したもので, **ARX モデル**, **ARMAX モデル**と呼ばれる. 文字 X は外生入力 (eXogenous input) を表している. より一般的な **Box-Jenkins モデル**と呼ばれるものが用いられることもある (後述). また, 外乱入力 $(\nu(n))_{n\in\mathbb{N}}$ をフィルタに通さない**出力誤差モデル**と **FIR モデル**もある (後述).

- 外生入力 $u(n)$ と外乱 $\nu(n)$ が印加された同定モデルの (出力を $y(n)$ とする) の一般形は以下の通り.

$$y(n) = G(z^{-1}, \boldsymbol{\theta})u(n) + H(z^{-1}, \boldsymbol{\theta})\nu(n).$$

- 以下では, 伝達関数の形に基づいてこれらを分類するのだが, まずそのための準備をする.

- 伝達関数 $G(z^{-1}, \boldsymbol{\theta})$ は

$$G(z^{-1}, \boldsymbol{\theta}) = \frac{B(z^{-1})}{A(z^{-1})}$$

のように書かれているものとする.

- 伝達関数 $H(z^{-1}, \boldsymbol{\theta})$ は

$$H(z^{-1}, \boldsymbol{\theta}) = \frac{C(z^{-1})}{D(z^{-1})}$$

のように書かれているものとする.

- これから述べる分類は, $C(z^{-1})$ と $D(z^{-1})$ の取り方に応じたものである.

- パラメータ $\boldsymbol{\theta}$ は, $A(z^{-1})$, $B(z^{-1})$, $C(z^{-1})$ および $D(z^{-1})$ の係数である.

- システム同定の分野では, 通常は, $G(z^{-1}, \boldsymbol{\theta})$ に直達項がない場合のみを取り扱う．このため, 多くの場合, $B(z^{-1})$ は,

$$B(z^{-1}) = b_1 z^{-1} + \cdots b_M z^{-M}$$

のように取られる．z^0 の項がないことが直達項がないことを意味している．

- 伝達関数の分母多項式の最高次の係数は1であるものとする.

- 先に述べた一般形は, $H(z^{-1}, \boldsymbol{\theta})$ の分子および分母多項式の形に応じて, 以下の 5 種類に分類される.

ARX モデル

$$G(z^{-1}, \boldsymbol{\theta}) = \frac{B(z^{-1})}{A(z^{-1})}, \quad H(z^{-1}, \boldsymbol{\theta}) = \frac{1}{A(z^{-1})}$$

ARMAX モデル

$$G(z^{-1}, \boldsymbol{\theta}) = \frac{B(z^{-1})}{A(z^{-1})}, \quad H(z^{-1}, \boldsymbol{\theta}) = \frac{C(z^{-1})}{A(z^{-1})}$$

Box-Jenkins モデル

$$G(z^{-1}, \boldsymbol{\theta}) = \frac{B(z^{-1})}{A(z^{-1})}, \quad H(z^{-1}, \boldsymbol{\theta}) = \frac{C(z^{-1})}{D(z^{-1})}$$

出力誤差モデル

$$G(z^{-1}, \boldsymbol{\theta}) = \frac{B(z^{-1})}{A(z^{-1})}, \quad H(z^{-1}, \boldsymbol{\theta}) = 1$$

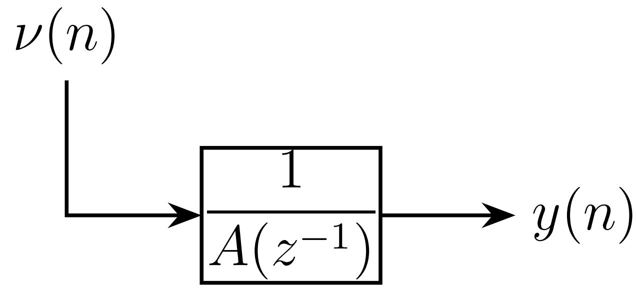
FIR モデル

$$G(z^{-1}, \boldsymbol{\theta}) = B(z^{-1}), \quad H(z^{-1}, \boldsymbol{\theta}) = 1$$

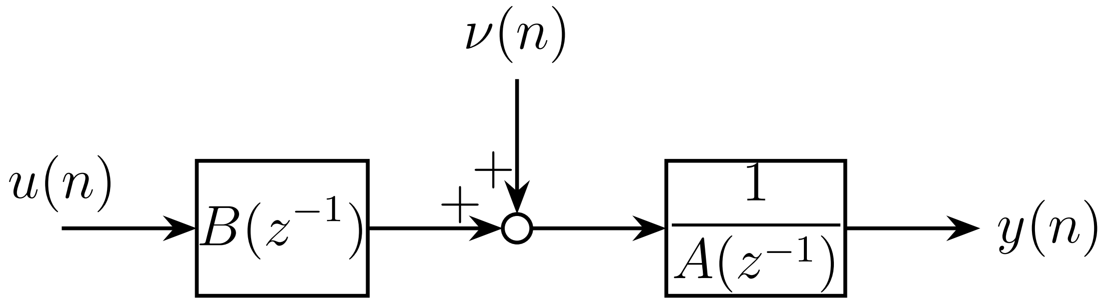
- 出力誤差モデルと FIR モデルは それぞれ Box-Jenkins モデルと ARX モデルの特別な場合.

- AR モデルと ARX モデルの比較

AR モデル

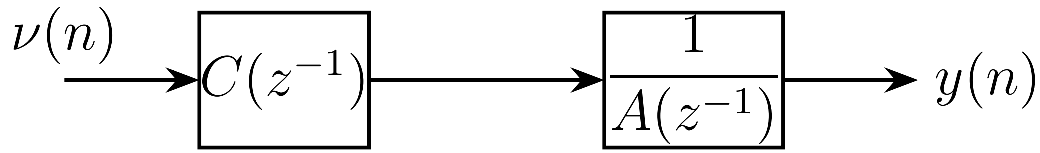


ARX モデル

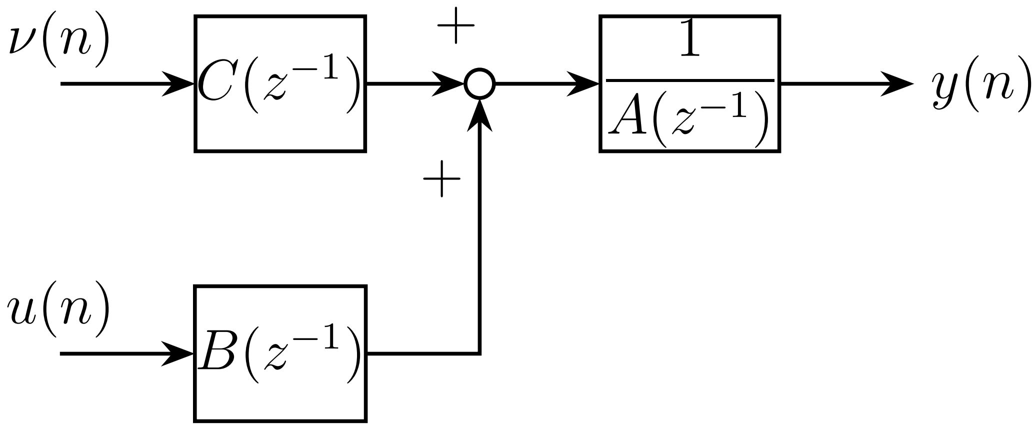


- ARMA モデルと ARMAX モデルの比較

ARMA モデル



ARMAX モデル



- ARX モデルを差分方程式で書き直すと,

$$
\begin{aligned}
& y(n) + a_1 y(n-1) + \cdots + a_N y(n-N) \\
& = b_1 u(n-1) + \cdots + b_M u(n-M) + \nu(n)
\end{aligned}
$$

となる. これを一般化して, 差分方程式に 外乱が直接加算されているモデルを, 式誤差モデルと呼ぶことがある.

- 入力誤差モデルと呼ばれるものが用いられることもある (詳細略).

- 上記以外にも様々なモデルがあり, また状態空間モデルが用いられることもあるが, この講義では説明を省略する.

- 実は, 構造同定とモデルの同定は 以下のようなループになっていると考えられる.

- 上記のループのように作業を進める場合, 単純なモデルから始めて, 必要に応じてモデルを複雑にするのが良いとされる．十分良いモデルが得られたらループを抜ける.

- モデルの次数の決定には, **Akaike's Information Theoretic Criterion (AIC)** などの基準が用いられるが, この講義では名前を挙げるのみに留める.

モデルの同定

- 以下では, ノンパラメトリックモデルおよびパラメトリックモデルの同定について述べる. 前者では, 外乱を含んだデータから 対象のインパルス応答あるいは周波数特性を推定する. 後者では, 選択されたモデルのパラメータを, 実験データに適合するよう定める.

ノンパラメトリックモデルの同定

§インパルス応答の同定

- 以下のシステムのインパルス応答の同定について述べる．エルゴード性と定常性が仮定されていることに注意.

$$y(n) = \sum_{k=0}^{\infty} g(k)u(n-k) + \nu(n).$$

- 同定用の入力 $(u(n))_{n\in\mathbb{N}}$ としては次の性質を持つものを選択する.

$$\phi_{uu}(\tau) = \begin{cases} \sigma_u^2 & \tau = 0, \\ 0, & \text{それ以外} \end{cases}$$

- $(u(n))_{n\in\mathbb{N}}$ と $(\nu(n))_{n\in\mathbb{N}}$ は無相関と仮定する.

以上の仮定のもとで…

$$
\begin{aligned}
&\phi_{yu}(\tau) \\
&\quad = \mathrm{E}\left[\left(\sum_{k=0}^{\infty} g(k)u(n+\tau-k) + \nu(n+\tau)\right) u(n)\right] \\
&\quad = \sum_{k=0}^{\infty} g(k)\phi_{uu}(\tau-k) = g(\tau)\sigma_u^2
\end{aligned}
$$

- したがって,

$$g(\tau) = \frac{\phi_{yu}(\tau)}{\sigma_u^2}$$

ただし, 相互相関関数は有限和で近似する.

$$\phi_{yu}(\tau) \simeq \frac{1}{K}\sum_{k=0}^{K-1} y(k+\tau)u(k)$$

- σ_u^2 については既知の筈であるが…

- 上記と同じ形の近似を用いて,

$$\sigma_u^2 \simeq \frac{1}{K}\sum_{k=0}^{K-1} u^2(k)$$

としてもよい．この場合,

$$g(\tau) \simeq \frac{\sum_{k=0}^{K-1} y(k+\tau)u(k)}{\sum_{k=0}^{K-1} u^2(k)}$$

となる.

§ 周波数特性の同定

- 周波数特性を同定するもっとも素朴な手段は, 入力を正弦波とし, その周波数を変えながら出力を測定すること. 簡単だが外乱には弱い.

- 相関関数を使うと, より雑音の影響を受けにくい同定ができる. 以下の式を再度用いる.

$$\phi_{yu}(\tau) = \sum_{k=0}^{\infty} g(k)\phi_{uu}(\tau - k)$$

- $\phi_{uu}(\tau)$ と $\phi_{yu}(\tau)$ の離散フーリエ変換を $\Phi_{uu}(e^{j\omega})$ と $\Phi_{yu}(e^{j\omega})$ とすると,

$$\Phi_{yu}(e^{j\omega}) = G(e^{j\omega})\Phi_{uu}(e^{j\omega}).$$

- $\phi_{uu}(\tau)$ と $\phi_{yu}(\tau)$ を有限の時間平均で近似したものを $\hat{\phi}_{uu}(\tau)$ と $\hat{\phi}_{yu}(\tau)$ とし，これらの離散フーリエ変換を $\hat{\Phi}_{uu}(e^{j\omega})$ と $\hat{\Phi}_{yu}(e^{j\omega})$ とすると, 以下の分母が零でなければ,

$$G(e^{j\omega}) \simeq \frac{\hat{\Phi}_{yu}(e^{j\omega})}{\hat{\Phi}_{uu}(e^{j\omega})}.$$

パラメトリックモデルの同定

- パラメトリックモデルでは, 過去の入出力データと $y(n) = G(z^{-1}, \boldsymbol{\theta})u(n) + H(z^{-1}, \boldsymbol{\theta})\nu(n)$ に基づき時刻 n における出力を推定し, それを実際の出力と比較することで, パラメータを同定する手法が標準的で, 推定誤差の分散を小さくする予測誤差法と推定誤差を過去のデータと無相関にする相関法に大別される.

- この講義では, 相関法については説明を省略し, 予測誤差法のみについて説明する.

§ 予測誤差法

- 先に述べたモデルの分類において, $H(z^{-1}, \boldsymbol{\theta})$ および $H^{-1}(z^{-1}, \boldsymbol{\theta})$ は安定で,

$$H(z^{-1}, \boldsymbol{\theta})\Big|_{z=\infty} = 1,$$

$(\nu(n))_{n\in\mathbb{N}}$ は期待値零, 分散 σ^2 の白色雑音とする.

- **出力の推定**とは，過去の入力 $(u(k))_{k<n}$ と出力 $(y(k))_{k<n}$ から現在の出力 $y(n)$ を推定するという意味で，**1 段先予測**と呼ばれる．以下でこれを導出する．

- ここで言う予測は, セルフチューニングコントロールの分野で用いられる ρ ステップ予測器とは違う言葉使いなので注意. システム同定の分野ではこのような呼び方が一般的.

- 当面はパラメータ $\boldsymbol{\theta}$ を既知とし, この条件の下で時刻 n における出力 $y(n)$ を過去のデータから推定する問題を考える.

- $k < n$ であれば,

$$\nu(k) = H^{-1}(z^{-1}, \boldsymbol{\theta}) \Big(y(k) - G(z^{-1}, \boldsymbol{\theta}) u(k) \Big)$$

  なので, $\nu(k)$ は過去のデータから計算できる.

- $\nu(n)$ は未知である.

- $H(z^{-1}, \boldsymbol{\theta})\Big|_{z=\infty} = 1$ と仮定したから,

$$H(z^{-1}, \boldsymbol{\theta}) = 1 + h_1 z^{-1} + h_2 z^{-2} + \cdots$$

のように展開され, 最初の項は 1 である.

- よって,

$$(H(z^{-1}, \boldsymbol{\theta}) - 1)\nu(n)$$

は過去の外乱 $\nu(k)$ $(k < n)$ の線形結合であり, 既知である.

- 未知の $\nu(n)$ を零とおいた推定値 (これが1段先予測された値) を $\hat{y}(n|n-1,\boldsymbol{\theta})$ とすると,

$$
\begin{aligned}
\hat{y}(n|n-1,\boldsymbol{\theta}) &= G(z^{-1},\boldsymbol{\theta})u(n) \\
&\quad + \left(H(z^{-1},\boldsymbol{\theta}) - 1\right)\nu(n) \\
&= G(z^{-1},\boldsymbol{\theta})u(n) \\
&\quad + \left(H(z^{-1},\boldsymbol{\theta}) - 1\right) \\
&\quad \cdot H^{-1}(z^{-1},\boldsymbol{\theta})\Big(y(n) - G(z^{-1},\boldsymbol{\theta})u(n)\Big)
\end{aligned}
$$

- 以上を整理すると, 次の 1 段先予測が得られる:

$$\hat{y}(n|n-1,\boldsymbol{\theta}) = \left(1 - H^{-1}(z^{-1},\boldsymbol{\theta})\right) y(n) + H^{-1}(z^{-1},\boldsymbol{\theta})G(z^{-1},\boldsymbol{\theta})u(n).$$

- $\varepsilon(n|\boldsymbol{\theta}) = y(n) - \hat{y}(n|n-1,\boldsymbol{\theta})$ と定義して先の式を代入すると, 以下のようになる.

$$\begin{aligned}&\varepsilon(n|\boldsymbol{\theta})\\ &\quad = H^{-1}(z^{-1},\boldsymbol{\theta})\left(y(n) - G(z^{-1},\boldsymbol{\theta})u(n)\right).\end{aligned} \tag{$\bigstar$}$$

- 予測誤差法では, 損失関数 $l(\varepsilon(n|\boldsymbol{\theta}))$ と呼ばれる関数を定義し, ある $K > 0$ に対し, 以下の $V_K(\boldsymbol{\theta})$ を最小とするパラメータ $\boldsymbol{\theta}$ を求める.

$$V_K(\boldsymbol{\theta}) = \frac{1}{K} \sum_{n=0}^{K-1} l(\varepsilon(n|\boldsymbol{\theta}))$$

- 損失関数は $\varepsilon(n|\boldsymbol{\theta})$ の影響を評価するための関数で, 理論的には何でもよいが, 実用上は,

$$l(\varepsilon(n|\boldsymbol{\theta})) = \varepsilon^2(n|\boldsymbol{\theta}) \tag{$\sharp$}$$

  がよく用いられる.

- K は PE 条件を参考にして定める.

- 上述の

$$V_K(\boldsymbol{\theta})$$

を損失関数と呼ぶ流儀もあるので注意.

- 以下では，損失関数として，つねに ($\sharp$) を用い，パラメータ推定問題を，

$$V_K(\boldsymbol{\theta}) = \frac{1}{K}\sum_{n=0}^{K-1}\varepsilon^2(n|\boldsymbol{\theta}) \qquad (\flat)$$

に関する最小化問題に帰着させる．

- **ARX モデルのパラメータ推定:** 予測誤差法によって ARX モデルのパラメータを推定する方法について述べる.

- ARX モデルでは,

$$
\begin{aligned}
G(z^{-1}, \boldsymbol{\theta}) &= \frac{B(z^{-1})}{A(z^{-1})} \\
H(z^{-1}, \boldsymbol{\theta}) &= \frac{1}{A(z^{-1})}
\end{aligned}
$$

であった.

- 先に述べた式

$$
\begin{aligned}
&\varepsilon(n|\boldsymbol{\theta}) \\
&\quad = H^{-1}(z^{-1}, \boldsymbol{\theta}) \left(y(n) - G(z^{-1}, \boldsymbol{\theta}) u(n) \right).
\end{aligned}
$$

にこれを代入すると,

$$
\varepsilon(n|\boldsymbol{\theta}) = A(z^{-1}) y(n) - B(z^{-1}) u(n) \qquad (\heartsuit)
$$

となる.

- 以下のように定義すると…

$$\boldsymbol{\theta} = (a_1, \ldots, a_N, b_1, \ldots, b_M)^T$$
$$\boldsymbol{\xi}(n) = (-y(n-1), \ldots, -y(n-N), \\ u(n-1), \ldots, u(n-M))$$

- 次ページの式が得られる．

$$
\begin{pmatrix} \varepsilon(n|\boldsymbol{\theta}) \\ \vdots \\ \varepsilon(n+K-1|\boldsymbol{\theta}) \end{pmatrix}
= \begin{pmatrix} y(n) \\ \vdots \\ y(n+K-1) \end{pmatrix} - \begin{pmatrix} \boldsymbol{\xi}(n) \\ \vdots \\ \boldsymbol{\xi}(n+K=1) \end{pmatrix} \boldsymbol{\theta}.
$$

- 前ページの式右辺第 1 項のベクトルを $\boldsymbol{\eta}$, 右辺第 2 項左側の行列を $\boldsymbol{\Xi}$ とする.

- 最小化すべき関数は

$$V_K(\boldsymbol{\theta}) = \frac{1}{K}\sum_{n=0}^{K-1}(\varepsilon(n|\boldsymbol{\theta}))^2$$

であったが, 先の記法を用いると,

$$V_K(\boldsymbol{\theta}) = \frac{1}{K}(\boldsymbol{\eta} - \boldsymbol{\Xi}\boldsymbol{\theta})^T(\boldsymbol{\eta} - \boldsymbol{\Xi}\boldsymbol{\theta})$$

となる.

- これを最小化するためには,

$$\frac{\partial V_K}{\partial \boldsymbol{\theta}} = \mathbf{0}$$

  を $\boldsymbol{\theta}$ について解けばよい.

- $\dfrac{\partial V_K}{\partial \boldsymbol{\theta}} = -\dfrac{2}{K}(\boldsymbol{\eta} - \boldsymbol{\Xi}\boldsymbol{\theta})^T \Xi$ だから, これを零とする $\boldsymbol{\theta}$ は $\Xi^T \Xi \theta = \Xi^T \eta$ の解, すなわち

$$\boldsymbol{\theta} = \left(\Xi^T \Xi\right)^{-1} \Xi^T \boldsymbol{\eta}$$

である. 上式右辺の行列が可逆であることが同定が可能であるための十分条件である.

- この方程式 (あるいは類似した形のもの) を正規方程式と呼ぶ.

- **FIR モデルのパラメータ推定:**　予測誤差法によって FIR モデルのパラメータを推定する方法について述べる.

- FIR モデルは ARX モデルの特別な場合なので, ARX モデルに対するパラメータ推定の方法がそのまま使える.

- ARX モデルのパラメータ $\boldsymbol{\theta}$ から $A(z^{-1})$ の定数項以外の係数を除き, $\boldsymbol{\xi}(n)$ から過去の出力の系列の部分を除いた上で,

$$\boldsymbol{\theta} = \left(\Xi^T \Xi\right)^{-1} \Xi^T \boldsymbol{\eta}$$

とすればよい.

- **ARMAX モデルのパラメータ推定:**　予測誤差法によって ARMAX モデルのパラメータを推定する方法について述べる.
- 再び以下の式を用いる.

$$
\begin{aligned}
&\varepsilon(n|\boldsymbol{\theta}) \\
&\quad = H^{-1}(z^{-1}, \boldsymbol{\theta}) \left(y(n) - G(z^{-1}, \boldsymbol{\theta}) u(n) \right).
\end{aligned}
$$

- $(\bigstar)$ に

$$G(z^{-1}, \boldsymbol{\theta}) = \frac{B(z^{-1})}{A(z^{-1})}$$

$$H^{-1}(z^{-1}, \boldsymbol{\theta}) = \frac{C(z^{-1})}{A(z^{-1})}$$

を代入すると…

- ARMAX モデルでは,

$$\varepsilon(n|\boldsymbol{\theta}) = \frac{A(z^{-1})}{C(z^{-1})}y(n) - \frac{B(z^{-1})}{C(z^{-1})}u(n) \qquad (\diamondsuit)$$

となる.

- 前ページの式を用い,

$$V_K(\boldsymbol{\theta}) = \frac{1}{K}\sum_{n=0}^{K-1}\varepsilon^2(n|\boldsymbol{\theta})$$

に関する最小化問題を解くことにより, パラメータ推定値が得られるのだが, ARX モデルと異なり, これは非線形計画問題になっている.

- 上述の問題は一般に解析的には解けないので，非線形計画問題の数値解法を用いて近似解を得る必要がある．

- この講義では非線形計画法には立ち入らない．詳細については [片山 (1994)] を参照せよ．

- **出力誤差モデルのパラメータ推定:**　予測誤差法によって出力誤差モデルのパラメータを推定する方法について述べる.

- 出力誤差モデルでは,

$$G(z^{-1}, \boldsymbol{\theta}) = \frac{B(z^{-1})}{A(z^{-1})}$$

$$H(z^{-1}, \boldsymbol{\theta}) = 1$$

であった.

- 先に述べた式

$$
\begin{aligned}
&\varepsilon(n|\boldsymbol{\theta}) \\
&\quad = H^{-1}(z^{-1}, \boldsymbol{\theta}) \left(y(n) - G(z^{-1}, \boldsymbol{\theta}) u(n) \right).
\end{aligned}
$$

にこれを代入すると…

- 出力誤差モデルでは,

$$\varepsilon(n|\boldsymbol{\theta}) = y(n) - \frac{B(z^{-1})}{A(z^{-1})}u(n) \qquad (\clubsuit)$$

となる.

- 前ページの式を用い,

$$V_K(\boldsymbol{\theta}) = \frac{1}{K}\sum_{n=0}^{K-1}\varepsilon^2(n|\boldsymbol{\theta})$$

に関する最小化問題を解くことにより, パラメータ推定値が得られるのだが, ARMAX モデルと同様に, これは非線形計画問題になっている.

- したがって, この形でパラメータを推定する場合には, 非線形計画法のアルゴリズムによって数値的な近似解を求める必要がある.

- 一方, 別の解法として, 外乱を無視すれば, 以下に述べるように, ARX モデルと同様に 最小二乗法によってパラメータ推定値を得ることができる. これについても述べておく.

- $y(n) = \dfrac{B(z^{-1})}{A(z^{-1})} u(n) + \nu(n)$ において, 機械的に $\nu(n) = 0$ とおくと,

$$A(z^{-1}) y(n) = B(z^{-1}) u(n)$$

となる.

- 以下のように定義すると…

$$
\begin{aligned}
\boldsymbol{\theta} &= (a_1, \ldots, a_N, b_1, \ldots, b_M)^T, \\
\boldsymbol{\xi}(n) &= (-y(n-1), \ldots, -y(n-N), \\
&\qquad u(n-1), \ldots, u(n-M))
\end{aligned}
$$

- 以下の式が得られる.

$$\begin{pmatrix} y(n) \\ \vdots \\ y(n+K-1) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\xi}(n) \\ \vdots \\ \boldsymbol{\xi}(n+K-1) \end{pmatrix} \boldsymbol{\theta}.$$

- 上式右辺のベクトルを $\boldsymbol{\eta}$, 右辺右側の行列を $\boldsymbol{\Xi}$ とする.

- 前ページの式は,

$$\boldsymbol{\eta} = \boldsymbol{\Xi}\boldsymbol{\theta}$$

と書き直せる.

- よって,

$$V_K(\boldsymbol{\theta}) = (\boldsymbol{\eta} - \boldsymbol{\Xi}\boldsymbol{\theta})^T(\boldsymbol{\eta} - \boldsymbol{\Xi}\boldsymbol{\theta})$$

とし, $V_K(\boldsymbol{\theta})$ を最小化することで, パラメータ推定値が得られる.

- これは AR モデルのパラメータ推定と同じ形であり, よって解も同じで.

$$\boldsymbol{\theta} = \left(\Xi^T \Xi\right)^{-1} \Xi^T \boldsymbol{\eta}$$

となる.

- 上述の方法は簡単ではあるが, AR モデルと異なり, 得られる推定値は統計的には良い性質を持たない (詳細は略す).

モデルの妥当性の評価

- モデルの妥当性の評価では，まず極と零の位置をチェックする．極と零点が近い位置にある場合には，モデルの次数が高すぎるという懸念がある．

- 次に, 以下の式で定義される量 (残差と呼ばれる) を用いてモデルの妥当性を調べる.

$$H(z^{-1}, \boldsymbol{\theta})^{-1}(y(n) - G(z^{-1}, \boldsymbol{\theta})u(n))$$

同定がうまくいっている場合には, 残差が白色かつ入力と独立になっている筈である. そうでない場合には再検討が必要.

- モデルが妥当であると判断されれば, システム同定に係る一連の手順は終了．そうでない場合には, 構造同定のステップに戻って作業をやり直す.

- システム同定実験のやり直しということも考えられなくはないが, この工程は手間と費用がかかるため, 一般には気軽にできるものではない.

参考文献

- 足立, システム同定の基礎, 東京電機大学出版局, 2009
- 片山, システム同定入門, 朝倉書店, 1994
- 片山, システム同定, 朝倉書店, 2004
- L. Ljung, System Identification, 2/e, Prentice-Hall, 1999.

- 電子情報通信学会知識ベース

  http://www.ieice-hbkb.org/portal/

- Selecting a Model Structure in the System Identification Process

  http://www.ni.com/product-documentation/4028/en/