制御のためのシステム同定 7章 モデルの選定法と妥当性の検証

川勝孝也

2024年6月13日

予測誤差法

予測誤差

時刻nとして、予測誤差ϵで構成される損失関数$J(\boldsymbol{\theta})$を最小化する$\boldsymbol{\theta}$を求める

$$\epsilon(n, \boldsymbol{\theta}) = y(n) - \hat{y}(n|\boldsymbol{\theta}) \tag{1}$$

対数尤度

予測誤差が平均0で分散σ^2の正規分布に従う場合、対数尤度は式(2)で求まる

$$\log \mathcal{L}(\boldsymbol{\theta}) = -\frac{N}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{n=1}^{N} \epsilon^2(n, \boldsymbol{\theta}) \tag{2}$$

損失関数

予測誤差に正規分布を仮定した最尤推定は、式(3)の最小2乗法と等価になる

$$\arg\max_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}) = \arg\min_{\boldsymbol{\theta}} J(\boldsymbol{\theta}) = \arg\min_{\boldsymbol{\theta}} \sum_{n=1}^{N} \{y(n) - \hat{y}(n|\boldsymbol{\theta})\}^2 \tag{3}$$

周波数特性

線形モデル

z^{-1}を1周期の遅延、nを時刻として、出力$y(n)$と操作$u(n)$と雑音$w(n)$の関係

$$y(n) = G(z, \boldsymbol{\theta})u(n) + H(z, \boldsymbol{\theta})w(n) \tag{4}$$

周波数重み関数

式(4)で、パーセバルの定理より、損失関数に周波数重み関数$W(\omega)$が現れる

$$J(\boldsymbol{\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} W(\omega, \boldsymbol{\theta}) \left| \frac{Y(\omega)}{U(\omega)} - G(\omega, \boldsymbol{\theta}) \right|^2 d\omega \tag{5}$$

$$W(\omega, \boldsymbol{\theta}) = |H(e^{j\omega}, \boldsymbol{\theta})|^{-2} |U(\omega)|^2$$

式(4)が低域通過特性で、操作uが白色性なら、低周波の同定精度が低下する

式誤差モデル

立式

z^{-1}を1周期の遅延、nを時刻として、出力$y(n)$と操作$u(n)$と雑音$w(n)$の関係

$$A(z)y(n) = B(z)u(n) + w(n) = w(n) + \sum_{k=1}^{\infty} b(k)u(n)z^{-k} \tag{6}$$

求解

最適解$\hat{\boldsymbol{\theta}}$が求まる条件は、式(7)の行列$R$が正定値で、損失関数$J(\boldsymbol{\theta})$が凸関数

$$\begin{aligned}
J(\boldsymbol{\theta}) &= \sum_{n=1}^{N} \left\{ y(n) - \boldsymbol{\theta}^\top \boldsymbol{\varphi}(n) \right\}^2 = c(N) - 2\boldsymbol{\theta}^\top \boldsymbol{f}(N) + \boldsymbol{\theta}^\top R(N)\boldsymbol{\theta} \\
R(N) &= \sum_{n=1}^{N} \left[\begin{array}{c|c} \boldsymbol{\varphi}_y(n)\boldsymbol{\varphi}_y^\top(n) & \boldsymbol{\varphi}_y(n)\boldsymbol{\varphi}_u^\top(n) \\ \hline \boldsymbol{\varphi}_u(n)\boldsymbol{\varphi}_y^\top(n) & \boldsymbol{\varphi}_u(n)\boldsymbol{\varphi}_u^\top(n) \end{array} \right] \\
\hat{\boldsymbol{\theta}}(N) &= R^{-1}(N)\boldsymbol{f}(N)
\end{aligned} \tag{7}$$

持続的励起性

正定値の条件

操作uがk_y+k_u次の持続的励起性で、安定かつ可観測なら、行列Rは正定値

$$\boldsymbol{\varphi}_y(n)=\begin{bmatrix}-y(n-1)\\ \vdots \\ -y(n-k_y)\end{bmatrix},\quad \boldsymbol{\varphi}_u(n)=\begin{bmatrix}u(n-1)\\ \vdots \\ u(n-k_u)\end{bmatrix} \tag{8}$$

持続的励起性

自己相関ϕ_{uu}で、式(9)の行列R_kが正則の場合、操作uはk次の持続的励起性

$$R_k=\begin{bmatrix}\phi_{uu}(0) & \phi_{uu}(1) & \cdots & \phi_{uu}(k-1)\\ \phi_{uu}(1) & \phi_{uu}(0) & \cdots & \phi_{uu}(k-2)\\ \vdots & \vdots & \ddots & \vdots \\ \phi_{uu}(k-1) & \phi_{uu}(k-2) & \cdots & \phi_{uu}(0)\end{bmatrix} \tag{9}$$

正弦波は2次の持続的励起性で、次数を高めるには、最大周期列信号を使う

状態空間モデル

状態方程式

離散量nを時刻として、行列A, B, Cで記述される状態空間モデルを考えよう

$$\begin{aligned} \boldsymbol{x}(n) &= A\boldsymbol{x}(n) + B\boldsymbol{u}(n), \\ \boldsymbol{y}(n) &= C\boldsymbol{x}(n) \end{aligned} \tag{10}$$

ハンケル行列

インパルス応答$g(n)$でハンケル行列Hを構成し、可制御・可観測行列に分解

$$H = \begin{bmatrix} g(1) & g(2) & \cdots & g(k+1) \\ g(2) & g(3) & \cdots & g(k+2) \\ \vdots & \vdots & \ddots & \vdots \\ g(k+1) & g(k+2) & \cdots & g(2k+1) \end{bmatrix} = \begin{bmatrix} B \\ AB \\ \vdots \\ A^k B \end{bmatrix} \begin{bmatrix} C^\top \\ (CA)^\top \\ \vdots \\ (CA^k)^\top \end{bmatrix}^\top \tag{11}$$

可制御・可観測行列からB, Cを取得して、シフト不変性を利用してAを算出

構造同定

事前に得られた物理知識に基づき、モデルの構造を決める必要がある

手順

1. 同定を行うモデルの構造を選定
2. 無駄時間やモデルの次数を決定

選定

交差検証を実施して、情報量基準に注目して、最適なモデル構造を選ぶ

評価

極零相殺や残差の白色性に問題がある場合、別のモデル構造を検討する

圧縮

平衡実現により、ハンケル特異値に注目して、モデルの次数を削減する

式誤差 (ARX) モデル

立式

z^{-1}を1周期の遅延、nを時刻として、出力$y(n)$と操作$u(n)$と雑音$w(n)$の関係

$$A(z)y(n) = B(z)u(n) + w(n) \tag{12}$$

得失

- 雑音と操作の伝達関数の極が共通なので、入力側の外乱に適する
- 線形回帰なので最適化の計算量が低く、パラメータも削減できる
- 雑音に脆弱なので、高次だと過学習し、低次だと精度が悪化する

特性

周波数重み関数が高域通過特性を持つので、プレフィルタによる補正が必要

$$W(\omega, \boldsymbol{\theta}) = |A(\omega, \boldsymbol{\theta})|^2 |U(\omega)|^2 \tag{13}$$

式誤差 (ARMAX) モデル

立式

z^{-1}を1周期の遅延、nを時刻として、出力$y(n)$と操作$u(n)$と雑音$w(n)$の関係

$$A(z)y(n) = B(z)u(n) + C(z)w(n) \tag{14}$$

得失

- 雑音と操作の伝達関数の極が共通なので、入力側の外乱に適する
- 雑音専用の多項式を割り当てるので、モデルの柔軟性が改善する
- 雑音と操作が共通の極を持ち、出力誤差モデルよりも雑音に脆弱

特性

周波数重み関数が高域通過特性を持つので、プレフィルタによる補正が必要

$$W(\omega, \boldsymbol{\theta}) = |A(\omega, \boldsymbol{\theta})|^2 |C(\omega, \boldsymbol{\theta})|^{-2} |U(\omega)|^2 \tag{15}$$

出力誤差 (OE) モデル

立式

z^{-1}を1周期の遅延、nを時刻として、出力$y(n)$と操作$u(n)$と雑音$w(n)$の関係

$$y(n) = B(z)F^{-1}(z)u(n) + w(n) \tag{16}$$

得失

- 雑音と操作の伝達関数の極が独立なので、出力側の外乱に適する
- 雑音の分離に優れ、開ループ実験なら動特性を正確に推定できる
- 非線形最適化なので、計算量が重く、局所解に陥る可能性がある

特性

周波数重み関数が全域通過特性を持つので、プレフィルタによる補正が不要

$$W(\omega, \boldsymbol{\theta}) = |U(\omega)|^2 \tag{17}$$

出力誤差 (BJ) モデル

立式

z^{-1}を1周期の遅延、nを時刻として、出力$y(n)$と操作$u(n)$と雑音$w(n)$の関係

$$y(n) = B(z)F^{-1}(z)u(n) + C(z)D^{-1}(z)w(n) \tag{18}$$

得失

- 雑音と操作の伝達関数の極が独立なので、出力側の外乱に適する
- 雑音の分離に優れ、開ループ実験なら動特性を正確に推定できる
- 非線形最適化なので、計算量が重く、局所解に陥る可能性がある

特性

周波数重み関数が全域通過特性を持つ場合、プレフィルタによる補正が不要

$$W(\omega, \boldsymbol{\theta}) = |C(\omega, \boldsymbol{\theta})|^{-2}|D(\omega, \boldsymbol{\theta})|^{2}|U(\omega)|^{2} \tag{19}$$

情報量基準

赤池情報量基準

k次モデルをN個の標本で同定して、誤差が分散σ^2の正規分布に従う場合、

$$\mathrm{AIC} = -2\log\mathcal{L}(\theta) + 2k \approx N\log\sigma^2 + 2k \tag{20}$$

ベイズ情報量基準

k次モデルをN個の標本で同定して、誤差が分散σ^2の正規分布に従う場合、

$$\mathrm{BIC} = -2\log\mathcal{L}(\theta) + k\log N \approx N\log\sigma^2 + k\log N \tag{21}$$

最終予測誤差

k次モデルをN個の標本で同定して、誤差が分散σ^2の正規分布に従う場合、

$$\mathrm{FPE} = \frac{N+k}{N-k}\sigma^2 \approx \exp\frac{\mathrm{AIC}}{N} \tag{22}$$

極零相殺

極零相殺

同定したモデルの極と零点が重複する場合は、モデルの次数が過剰と考える

$$G(s) = \frac{N(s)}{D(s)} = \frac{s+2}{(s+2)(s+5)} = \frac{1}{s+5} \tag{23}$$

信頼区間

極と零点がぴったり重なるのは稀なので、信頼区間の重複の有無を評価する

$$G(s) = \frac{s+2.01}{(s+1.99)(s+5)} \approx \frac{1}{s+5} \tag{24}$$

残差解析

残差の白色性

残差$e(n)$の白色性も、同定したモデルの妥当性の評価で、重要な根拠となる

$$\begin{aligned} e(n) &= H^{-1}(z)\left\{y(n) - G(z)u(n)\right\} \\ y(n) &= G(z)u(n) + H(z)w(n) \end{aligned} \tag{25}$$

残差の独立性

- 出力誤差モデルでは、残差$e(n)$と操作$u(n)$の独立性を重視して評価する
- フィードバック系の場合は、残差$e(n)$と操作$u(n)$が相関を持つので注意

推定誤差

雑音誤差

観測雑音やシステム雑音など、確率的な外乱に起因する、同定結果の曖昧性

$$
\begin{aligned}
y(n) &= G_0(z)u(n) + w(n) = G(z,\boldsymbol{\theta})u(n) + \xi(n) \\
P(N) &= \mathbf{E}\left[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^\top \right] \approx \frac{\sigma^2}{N}\bar{R}^{-1} \\
\bar{R} &= \mathbf{E}\left[\left(\frac{d}{d\boldsymbol{\theta}}\hat{y}(n|\boldsymbol{\theta}) \right) \left(\frac{d}{d\boldsymbol{\theta}}\hat{y}(n|\boldsymbol{\theta}) \right)^\top \right]
\end{aligned} \tag{26}
$$

分散σ^2の雑音のもとで、推定値$\hat{\boldsymbol{\theta}}$の誤差の共分散$P(N)$は、標本数$N$と反比例

系統誤差

モデルの次数など、システムの構造的な差異に起因する、同定結果の曖昧性

$$
\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{2\pi}\int_{-\pi}^{\pi} W(\omega)\left|G(e^{j\omega},\boldsymbol{\theta}) - G_0(e^{j\omega})\right|^2 d\omega \tag{27}
$$

系統誤差により、伝達関数G, G_0が乖離して、周波数重み関数$W(\omega)$が現れる

対角正準形

状態方程式

連続量tを時刻として、行列A, B, Cで記述される状態空間モデルを考えよう

$$\begin{aligned}\dot{\boldsymbol{x}}(t) &= A\boldsymbol{x}(t) + B\boldsymbol{u}(t), \\ \boldsymbol{y}(t) &= C\boldsymbol{x}(t)\end{aligned} \tag{28}$$

対角正準形

行列Aの固有値を$\lambda_1, \ldots, \lambda_k$として、式(29)の正則行列$T$で、行列$A$を対角化

$$\begin{aligned}\boldsymbol{x}(t) &= T\hat{\boldsymbol{x}}(t) \\ \dot{\hat{\boldsymbol{x}}}(t) &= \hat{A}\hat{\boldsymbol{x}}(t) + \hat{B}\boldsymbol{u}(t) \ \ \text{where} \ \ \hat{A} = T^{-1}AT = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_k) \\ \boldsymbol{y}(t) &= \hat{C}\hat{\boldsymbol{x}}(t)\end{aligned} \tag{29}$$

制御・観測

行列$\hat{B}, \hat{C}$で固有値λ_kに対応する要素が0だと、状態$\hat{x}_k$の制御・観測が不可能

モデル圧縮

可制御性グラム行列

行列W_cが正則なら、そのシステムの内部の状態$\boldsymbol{x}(t)$を操作$\boldsymbol{u}(t)$で制御できる

$$W_c = \int_0^\infty (e^{At}B)(e^{At}B)^\top dt \Rightarrow AW_c + W_cA^\top + BB^\top = 0 \tag{30}$$

可観測性グラム行列

行列W_oが正則なら、そのシステムの内部の状態$\boldsymbol{x}(t)$を出力$\boldsymbol{y}(t)$で観測できる

$$W_o = \int_0^\infty (Ce^{At})^\top (Ce^{At}) dt \Rightarrow A^\top W_o + W_oA + C^\top C = 0 \tag{31}$$

平衡実現による圧縮

適当な行列Sで、式(32)の対角化を行うと、ハンケル特異値$\sigma_1, \ldots, \sigma_k$を得る

$$\hat{W}_c = SW_cS^\top = \hat{W}_o = S^{-\top}W_oS^{-1} = \mathrm{diag}\left(\sigma_1, \ldots, \sigma_k\right) \tag{32}$$

可制御性の証明

十分条件

式(33)の操作$\boldsymbol{u}(t)$により、時刻$t_0<0$から0にかけて、状態$\boldsymbol{x}_0$から$\boldsymbol{x}(0)$に到達

$$\begin{aligned} \boldsymbol{u}(t) &= (e^{-At}B)^\top W_c^{-1}(\boldsymbol{x}(0) - e^{-At_0}\boldsymbol{x}_0) \\ \boldsymbol{x}(t) &= e^{At}\left(\boldsymbol{x}(0) + \int_0^t e^{-A\tau}B\boldsymbol{u}(\tau)d\tau\right) \Rightarrow \lim_{t_0\to-\infty}\boldsymbol{x}(t_0) = \boldsymbol{x}_0 \end{aligned} \tag{33}$$

必要条件

行列W_cが非正則なら、式(34)で状態$\boldsymbol{x}_0$から$\boldsymbol{0}$に到達する操作$\boldsymbol{u}(t)$は存在せず

$$\begin{aligned} &\exists \boldsymbol{x}_0 \neq \boldsymbol{0},\ \ \boldsymbol{x}_0^\top W_c \boldsymbol{x}_0 = 0 \Rightarrow (e^{A\tau}B)^\top \boldsymbol{x}_0 = \boldsymbol{0} \\ &\qquad \lim_{t\to\infty}\boldsymbol{x}(t) = \lim_{t\to\infty} e^{At}\left(\boldsymbol{x}_0 + \int_0^t e^{-A\tau}B\boldsymbol{u}(\tau)d\tau\right) = \boldsymbol{0} \\ &\qquad \Rightarrow \boldsymbol{x}_0^\top\boldsymbol{x}_0 = -\lim_{t\to\infty}\int_0^t \boldsymbol{x}_0^\top e^{-A\tau}B\boldsymbol{u}(\tau)d\tau = 0\ \perp\ \boldsymbol{x}_0 \neq \boldsymbol{0} \end{aligned} \tag{34}$$

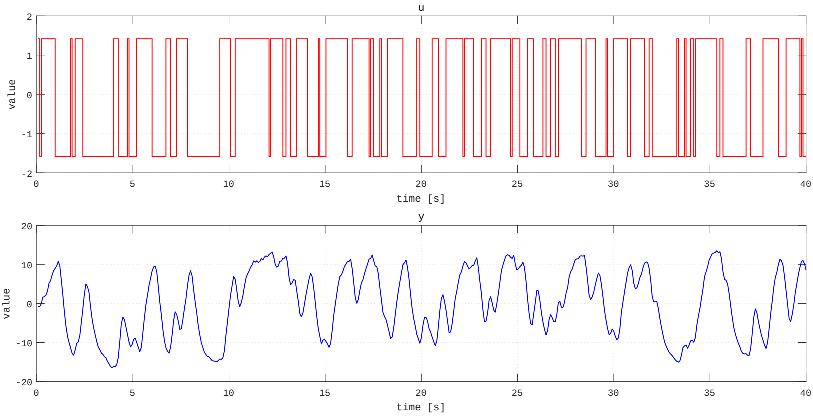
例題

ドライヤーの例題を通して、モデル構造の選定と妥当性の評価を行う

```
load dry2;
ze = dtrend(dry2(  1:500));
zv = dtrend(dry2(501:end));
```

実験

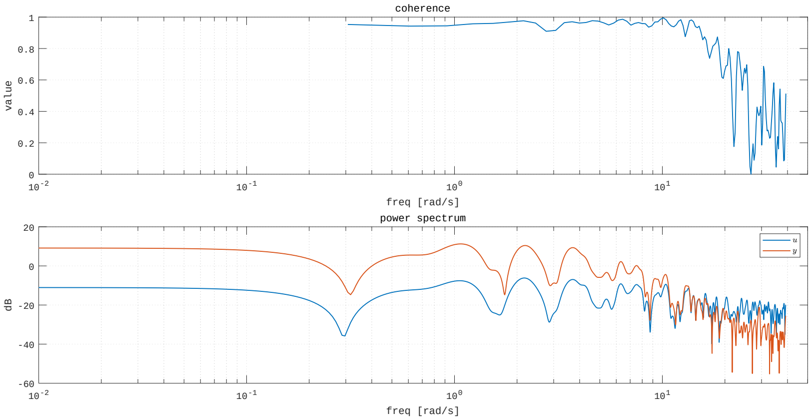
適当な最大周期列信号をシステムに入力して、実験データを取得する




```
idplot(ze);
```

周波数特性

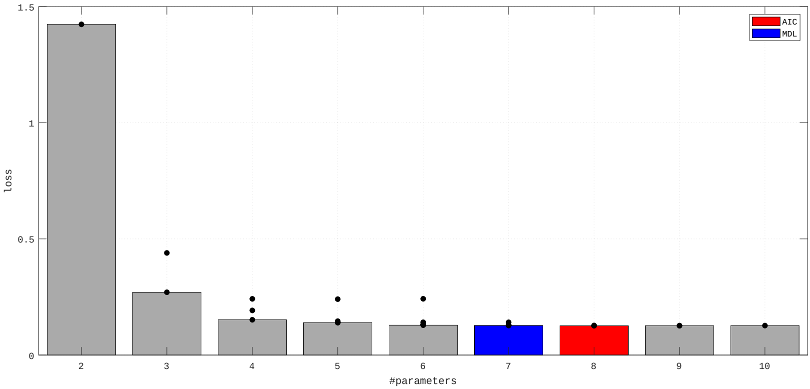
パワースペクトルとコヒーレンスを描画して、周波数特性を確認する




```
mscohere(ze.u, ze.y, [], [], [], 1/ze.Ts);
```

無駄時間と次数

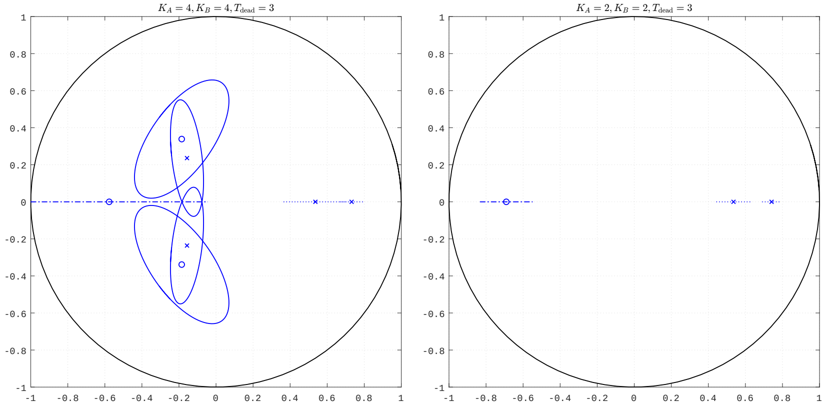
損失関数により、無駄時間を決定してから、モデルの次数を決定する




```
th = arx(ze, selstruc(arxstruc(ze, zv, struc(1:5, 1:5, 3)), 0));
```

極零相殺

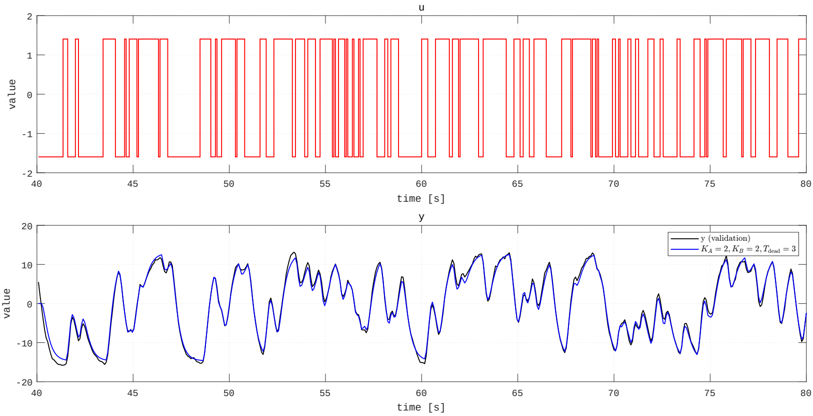
極と零点の信頼区間の重複が見られるので、モデルの次数を削減する




```
zpplot(th2zp(th), 3);
```

時間領域で評価

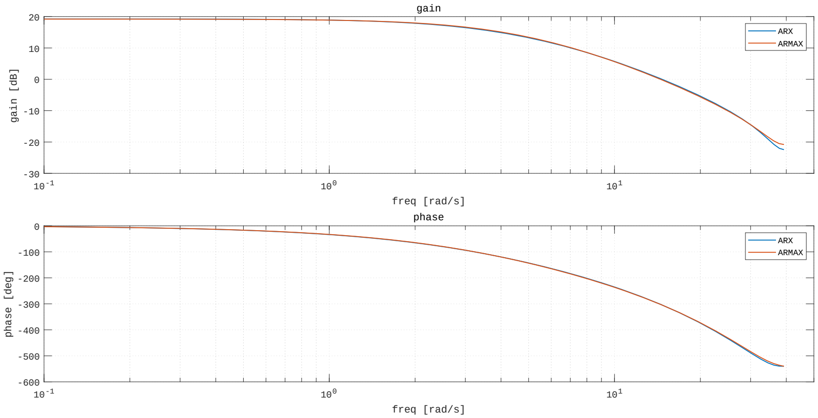
雑音なしの状態で時間発展を求め、実験データとの適合性を評価する




```
compare(zv, th);
```

周波数領域で評価

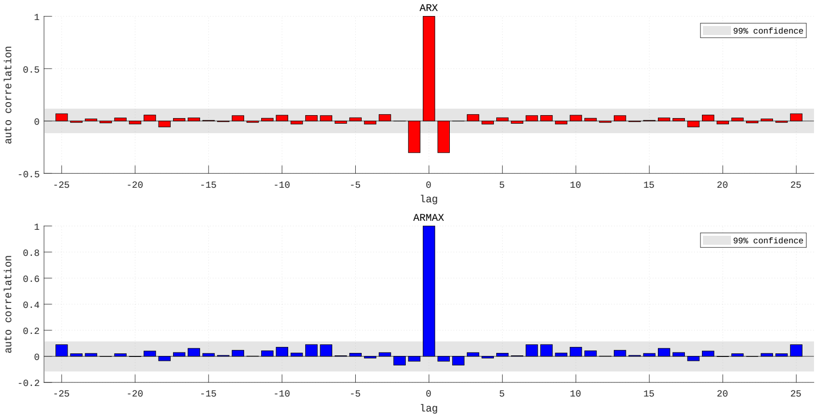
モデル構造を変えながら、ボード線図を描画して、信頼性を評価する




```
bodeplot(th);
```

残差解析

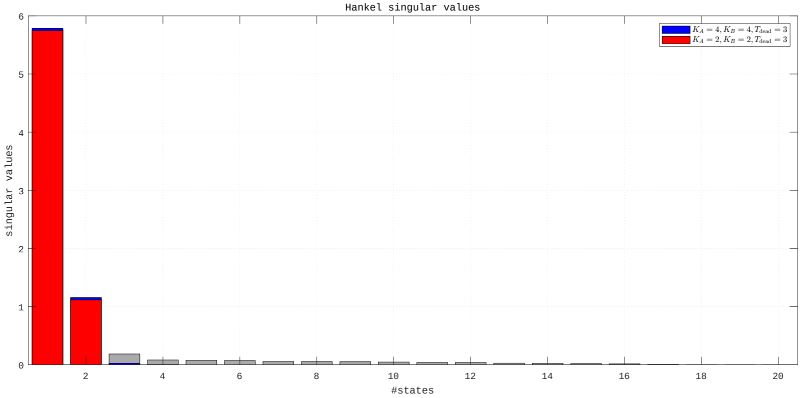
残差の自己相関を求め、残差の白色性を評価して、妥当性を棄却する




```
resid(zv, arx(ze, [2 2 3]));
```

ハンケル特異値

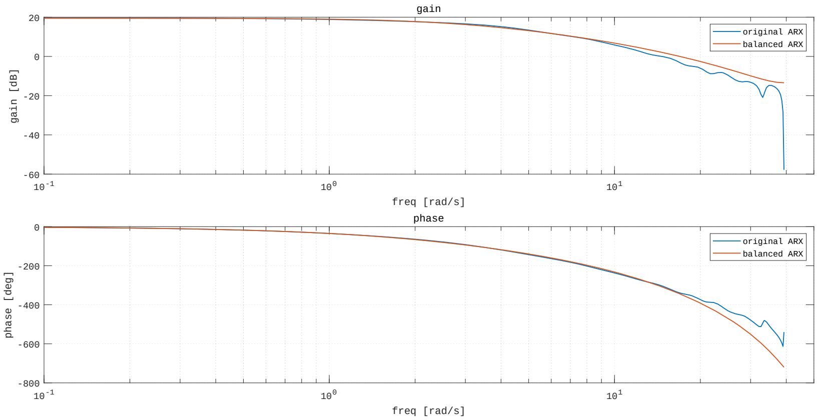
ハンケル特異値に注目すれば、雑音に適合した部分空間を特定できる




```
[thb, g] = balreal(arx(ze, [20 20 3]));
```

平衡実現

平衡実現により、可制御性・可観測性に注目して、次数を削減できる




```
bodeplot(xelim(balreal(arx(ze, [20 20 3])), 3:20));
```