10. Model Parametrization

- Model classification
- General model structure
- Uniqueness properties

Model Classification

- SISO/MIMO models
- Linear/Nonlinear models
- Parametric/Nonparametric models
- Time invariant/Time varying models
- Time domain/Frequency domain models
- Lumped/Distributed parameter models
- Deterministic/Stochastic models

General model structure

$$\mathcal{M}(\theta): \quad y(t) = G(q^{-1};\theta)u(t) + H(q^{-1};\theta)e(t)$$
$$\mathbf{E}\, e(t)e(s)^* = \Lambda(\theta)\delta_{t,s}$$

- $y(t)$ is ny-dimensional output
- $u(t)$ is nu-dimensional input
- $e(t)$ is an i.i.d. random variable with zero mean (white noise)
- q^{-1} is backward shift operator
- H, G, Λ are functions of the parameter vector θ
- This model is a genearal linear model in u and e

Feasible set of parameters

θ take the values such that

- H^{-1} and $H^{-1}G$ are asymptotically stable
- $G(0;\theta) = 0$ and $H(0;\theta) = I$
- $\Lambda(\theta) \succeq 0$

General SISO model structure

$$A(q^{-1})y(t) = \frac{B(q^{-1})}{F(q^{-1})}u(t) + \frac{C(q^{-1})}{D(q^{-1})}e(t), \quad \mathbf{E}\, e(t)e(t)^* = \lambda^2$$

where

$$A(q^{-1}) = 1 + a_1 q^{-1} + \ldots a_p q^{-p}$$

$$B(q^{-1}) = b_1 q^{-1} + b_2 q^{-2} + \ldots + b_n q^{-n}$$

$$C(q^{-1}) = 1 + c_1 q^{-1} + \ldots + c_m q^{-m}$$

$$D(q^{-1}) = 1 + d_1 q^{-1} + \ldots + d_s q^{-s}$$

$$F(q^{-1}) = 1 + f_1 q^{-1} + \ldots + f_r q^{-r}$$

Special cases

Output error structure

$$y(t) = \frac{B(q^{-1})}{F(q^{-1})}u(t) + e(t)$$

In this case $H(q^{-1};\theta) = 1$

The output error is the difference between the measurable output $y(t)$ and the model output $B(q^{-1})/F(q^{-1})u(t)$

If $A(q^{-1}) = 1$

$$y(t) = \frac{B(q^{-1})}{F(q^{-1})}u(t) + \frac{C(q^{-1})}{D(q^{-1})}e(t)$$

- G and H have no common paramater
- possible to estimate G consistently even if the choice of H is not appropriate

ARMAX models

An autoregressive moving average model with an exogenous input:

$$A(q^{-1})y(t) = B(q^{-1})u(t) + C(q^{-1})e(t)$$

where

$$\begin{aligned} A(q^{-1}) &= I + a_1 q^{-1} + \ldots a_p q^{-p} \\ B(q^{-1}) &= b_1 q^{-1} + b_2 q^{-2} + \ldots + b_n q^{-n} \\ C(q^{-1}) &= I + c_1 q^{-1} + \ldots + c_m q^{-m} \end{aligned}$$

with $\mathbf{E}\, e(t)e(t)^* = \lambda^2 I$

The parameter vector is

$$\theta = (a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_n, c_1, c_2, \ldots, c_m)$$

(the noise covariance could be a parameter to be estimated too)

Special cases of ARMAX models

- Autoregressive (AR) models

$$A(q^{-1})y(t) = e(t)$$

- Moving average (MA) models

$$y(t) = C(q^{-1})e(t)$$

- Finite impulse response (FIR) models

$$y(t) = B(q^{-1})u(t) + e(t)$$

- Autoregressive with exogenous input (ARX) models

$$A(q^{-1})y(t) = B(q^{-1})u(t) + e(t)$$

State-space models

A linear stochastic model:

$$x(t+1) = A(\theta)x(t) + B(\theta)u(t) + \nu(t)$$
$$y(t) = C(\theta)x(t) + \eta(t)$$

$\nu(t)$ and $\eta(t)$ are white noise sequences with zero means and

$$\mathbf{E} \begin{bmatrix} \nu(t) \\ \eta(t) \end{bmatrix} \begin{bmatrix} \nu(s) \\ \eta(s) \end{bmatrix}^* = \begin{bmatrix} R_1(\theta)\delta_{t,s} & R_{12}(\theta)\delta_{t,s} \\ R_{12}^*(\theta)\delta_{t,s} & R_2(\theta)\delta_{t,s} \end{bmatrix}$$

- $\nu(t)$ is the *process noise*
- $\eta(t)$ is the *measurement noise*
- needs to transform to the so-called *innovation form* to compare with the standard model

Choosing a class of model structures

Important factors:

- **Flexibility:** the model structure should describe most of the different system dynamics expected in the application
- **Parsimony:** the model should contain the smallest number of free parameters required to explain the data adequately
- **Algorithm complexity:** the form of model structure can considerably influence the computational cost
- **Properties of the criterion function:** for example, the asymptotic properties of prediction-error method depends crucially on the criterion function and the model structure

Uniqueness properties

Within a model structure, we are concerned with the problem of adequately and uniquely describing a given system

Define $\mathcal{D}$ the set of θ for which $(\hat{G}, \hat{H}, \hat{\Lambda})$ gives a *perfect description* of the true system

Three possibilities of this set can occur:

- the set $\mathcal{D}$ is empty or underparametrization
- the set $\mathcal{D}$ contains one point
- the set $\mathcal{D}$ consists of several points or overparametrization

Uniqueness properties for a scalar ARMA model

Let the true ARMA model is given by

$$A(q^{-1})y(t) = C(q^{-1})e(t), \quad \mathbf{E}\, e(t)^2 = \lambda^2$$

$\mathcal{D}$ is the set of $\hat{A}, \hat{B}, \hat{C}, \hat{\lambda}$ for which

$$\frac{C(q^{-1})}{A(q^{-1})} = \frac{\hat{C}(q^{-1})}{\hat{A}(q^{-1})}, \quad \hat{\lambda^2} = \lambda^2$$

In order for these equalities to have a solution, we must have

$$\text{deg}(\hat{A}) \geq \text{deg}(A), \quad \text{deg}(\hat{C}) \geq \text{deg}(C)$$

or,

$$n^* \triangleq \min\left\{\text{deg}(\hat{A}) - \text{deg}(A), \text{deg}(\hat{C}) - \text{deg}(C)\right\} \geq 0$$

- A and C have no common factor
- $\frac{C(q^{-1})}{A(q^{-1})}$ and $\frac{\hat{C}(q^{-1})}{\hat{A}(q^{-1})}$ must have the same poles and zeros

These implies

$$\hat{A}(q^{-1}) = A(q^{-1})D(q^{-1}), \quad \hat{C}(q^{-1}) = C(q^{-1})D(q^{-1})$$

where $D(q^{-1})$ has arbitrary coefficients

$$\mathrm{deg}(D) = \min\{\mathrm{deg}(\hat{A}) - \mathrm{deg}(A), \mathrm{deg}(\hat{C}) - \mathrm{deg}(C)\} = n^*$$

- $n^* > 0$: infinitely many solutions of $\hat{C}, \hat{A}, \hat{\lambda}$ (by varying D)
- $n^* = 0$: this gives $D(q^{-1}) = 1$, or at least one of $\hat{A}$ and $\hat{C}$ has the same degree as the true polynomial

Nonuniqueness of general state-space models

Consider the multivariable model

$$\begin{aligned} x(t+1) &= A(\theta)x(t) + B(\theta)u(t) + \nu(t) \\ y(t) &= C(\theta)x(t) + \eta(t) \end{aligned}$$

where $\nu(t)$ and $\eta(t)$ are mutually independent white noise with zero means and covariance R_1, R_2 resp.

Also consider a second model

$$\begin{aligned} z(t+1) &= \bar{A}(\theta)z(t) + \bar{B}(\theta)u(t) + \bar{\nu}(t) \\ y(t) &= \bar{C}(\theta)z(t) + \eta(t) \end{aligned}$$

where $\mathbf{E}\,\bar{\nu}(t)\bar{\nu}(s)^* = \bar{R}_1\delta_{t,s}$ and

$$\bar{A} = QAQ^{-1}, \quad \bar{B} = QB, \quad \bar{C} = CQ^{-1}, \quad \bar{R}_1 = QR_1Q^*$$

for some nonsingular matrix Q

The two models are equivalent:

- they have the same transfer function from u to y

$$G(q^{-1}) = \bar{C}(qI - A)^{-1}\bar{B} = CQ^{-1}(qI - QAQ^{-1})^{-1}QB = C(qI - A)^{-1}B$$

- the outputs y from the two models have the same second-order properties, $i.e.$, the spectral densities are the same

$$\begin{aligned} S_y(\omega) &= \bar{C}(e^{i\omega} - \bar{A})^{-1}\bar{R}_1(e^{i\omega} - \bar{A})^{-*}\bar{C}^* + R_2 \\ &= CQ^{-1}(e^{i\omega} - \bar{A})^{-1}QR_1Q^*(e^{i\omega} - \bar{A})^{-*}Q^{-*}C^* + R_2 \\ &= C[Q^{-1}(e^{i\omega} - \bar{A})Q]^{-1}R_1[Q^*(e^{i\omega} - \bar{A})^*Q^{-*}]^{-1}C^* + R_2 \\ &= C(e^{i\omega} - A)^{-1}R_1(e^{i\omega} - A)^{-*}C^* + R_2 \end{aligned}$$

The model is not unique since Q can be chosen arbitrarily

References

Chapter 6 in
T. Söderström and P. Stoica, *System Identification*, Prentice Hall, 1989