Direct estimation of the dynamical model from batch data from multiple experiments

October 31, 2019

Orthogonalisation and forward search are applied to all datasets to identify significant model terms prior to parameter estimation as described in [1].

Model structure for an individual dataset:

$$\underbrace{\boldsymbol{y}^k}_{\text{T}\times 1} = \underbrace{\Phi^k}_{\text{T}\times N}\underbrace{\theta^k}_{N\times 1}, \tag{1}$$

where T is the length of time-series, $k = 1,\ldots,K$ is the dataset index, N is identified number of significant terms. Model structure for K datasets:

$$\underbrace{\bar{\mathbf{Y}}}_{\text{T}K\times 1} = \underbrace{\bar{\Phi}}_{\text{T}K\times NK}\underbrace{\bar{\Theta}}_{NK\times 1}, \tag{2}$$

where the block matrices have the following structure:

$$\underbrace{\bar{\mathbf{Y}}}_{\text{T}K\times 1} = \begin{bmatrix} \underbrace{\boldsymbol{y}^1}_{\text{T}\times 1} \\ \underbrace{\boldsymbol{y}^2}_{\text{T}\times 1} \\ \vdots \\ \underbrace{\boldsymbol{y}^K}_{\text{T}\times 1} \end{bmatrix} ; \quad \underbrace{\bar{\Phi}}_{\text{T}K\times NK} = \begin{bmatrix} \underbrace{\Phi^1}_{\text{T}\times N} & \cdots & \cdots & \mathbb{O} \\ \mathbb{O} & \underbrace{\Phi^2}_{\text{T}\times N} & \cdots & \mathbb{O} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbb{O} & \cdots & \cdots & \underbrace{\Phi^K}_{\text{T}\times N} \end{bmatrix} ; \quad \underbrace{\bar{\Theta}}_{NK\times 1} = \begin{bmatrix} \underbrace{\theta^1}_{N\times 1} \\ \underbrace{\theta^2}_{N\times 1} \\ \vdots \\ \underbrace{\theta^K}_{N\times 1} \end{bmatrix}. \tag{3}$$

The relationship of the design parameters known from the experiments and the internal parameters of NARMAX model is defined by the following linear function:

$$\underbrace{\Theta}_{N\times K} = \underbrace{B}_{N\times L}\underbrace{A}_{L\times K}, \tag{4}$$

where A is the matrix where each row is a function of the vector of design parameters. The example structure is

$$A = \begin{bmatrix} \mathbb{I}_{K\times 1} & L_{K\times 1} & D_{K\times 1} & LD_{K\times 1} & L^2_{K\times 1} & D^2_{K\times 1} \end{bmatrix}^\top, \tag{5}$$

and where B denotes the matrix of unknown coefficients of a hypersurface of order L that maps a point in external parameter space, $\xi^k = (L_k, D_k)$, onto the point in the space of internal parameters,

θ^k. In can be seen that $\bar{\Theta} = \text{vec}(\Theta)$, then

$$\underbrace{\bar{\Theta}}_{NK \times 1} = \text{vec}\left(\underbrace{B}_{N \times L}\underbrace{A}_{L \times K}\right). \tag{6}$$

This vectorisation can be obtained using Kronecker product:

$$\text{vec}\left(\underbrace{B}_{N \times L}\underbrace{A}_{L \times K}\right) = (\underbrace{A^\top}_{K \times L} \otimes \underbrace{\mathbb{I}}_{N \times N})\underbrace{\text{vec}(B)}_{NL \times 1}. \tag{7}$$

Denoting the result of Kronecker product as $\underbrace{\mathbf{Kr}}_{NK \times NL} \triangleq (\underbrace{A^\top}_{K \times L} \otimes \underbrace{\mathbb{I}}_{N \times N})$ and vectorised coefficient matrix as $\underbrace{\bar{\mathbf{B}}}_{NL \times 1} \triangleq \text{vec}(B)$ yields the following:

$$\underbrace{\bar{\Theta}}_{NK \times 1} = \underbrace{\mathbf{Kr}}_{KN \times NL}\underbrace{\bar{\mathbf{B}}}_{NL \times 1}. \tag{8}$$

Substituting the above expression into (2) renders an expression that directly links the design parameters and the timeseries data

$$\underbrace{\bar{\mathbf{Y}}}_{TK \times 1} = \underbrace{\bar{\Phi}}_{TK \times NK}\underbrace{\mathbf{Kr}}_{KN \times NL}\underbrace{\bar{\mathbf{B}}}_{NL \times 1}, \tag{9}$$

where $\bar{\mathbf{B}}$ is the unknown vector and all other factors are known form the experiments or defined prior to structure identification.

The representation (9) allows estimating the coefficients in $\bar{\mathbf{B}}$ directly from the timeseries data bypassing the intermediate estimation of the internal coefficients in NARMAX model. It must be noted, however, that its not increasing the dimension of the matrix that guarantees identifiability of the system but its rank. The following condition must be satisfied:

$$\text{rk}(\bar{\Phi}\mathbf{Kr}) \geq NL. \tag{10}$$

The rank of the linear system (9) satisfies the following:

$$\text{rk}(\bar{\Phi}\mathbf{Kr}) \leq \min\left(\text{rk}(\bar{\Phi}), \text{rk}(\mathbf{Kr})\right), \tag{11}$$

where the rank of Kronecker product can be found as

$$\text{rk}(\mathbf{Kr}) = \text{rk}(\underbrace{A^\top}_{K \times L})\,\text{rk}(\underbrace{\mathbb{I}}_{N \times N}), \tag{12}$$

where the first factor is $\text{rk}(A^\top) = \min\{K, L\}$ (note that all rows and columns in the matrix A are known to be independent "by design"), and where second factor equals N. Given the above two expressions, it is evident that regardless of the length of the available time series, the identifiability of the system depends on the number of experiments, K, and the selected order of the hypersurface fitted to the points in the design parameter space, L. Moreover, if insufficient number of datapoints are available ($K < L$), the condition (10) can not be satisfied rendering the system (9) unidentifiable.