Subspace Identification of Structured State-Space Models with Unknown Inputs (含未知输入的结构化状态空间模型辨识)

Chengpu Yu (俞成浦)

Key Laboratory of Intelligent Control and Decision of Complex Systems

Beijing Institute of Technology, China

June 24, 2019

Outline

Outline

Background

Structured state-space model: a spatio-temporal model with structured interconnection in the space domain and dynamic response in the time domain.

Applications of the structured state-space model identification:

1. Provide network models for general implementation of networked systems.
2. Estimate parameters involved in physical models.
3. Approximate continuous physical objects by finite-element networks.

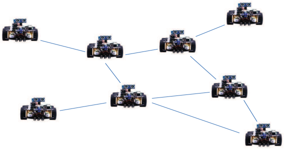


Figure: Multi-agent network

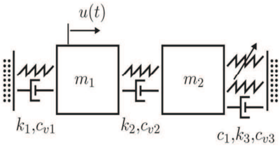


Figure: Mass-spring model

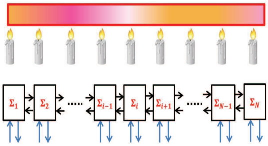


Figure: Finite-element network

Background

Traditional SysID methods require (noisy) input-output measurements; however, with unknown inputs, the identification problem (identifiability analysis + identification method development) becomes challenging.

Examples of structured systems with unknown inputs are: local network system with unknown interconnections, compartmental temperature field with people movement, etc.

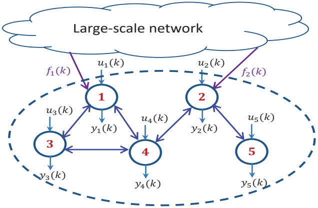


Figure: Local network

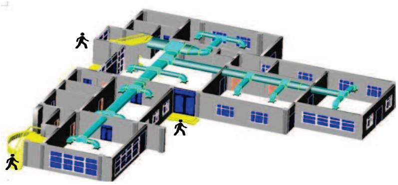


Figure: HVAC system

Outline

Problem formulation

The concerned state-space system model:

$$x(k+1) = A(\theta)x(k) + B(\theta)u(k) + Hf(k)$$
$$y(k) = C(\theta)x(k) + w(k)$$

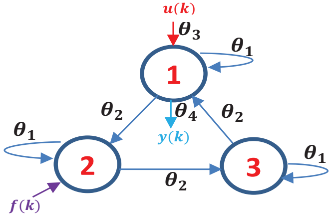
$u(k)$-measurable input, $f(k)$-unknown input, $y(k)$-measurable output. The system matrices $A(\theta)$, $B(\theta)$ and $C(\theta)$ are affinely parameterized w.r.t. $\theta \in \mathbb{R}^l$:

$$A(\theta) = A_0 + A_1\theta_1 + \cdots + A_l\theta_l,\ B(\theta) = B_0 + B_1\theta_1 + \cdots + B_l\theta_l,\ C(\theta) = C_0 + C_1\theta_1 + \cdots + C_l\theta_l.$$

Example 1

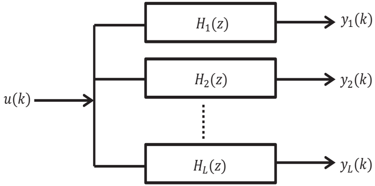
$$A(\theta) = \begin{bmatrix} \theta_1 & 0 & \theta_2 \\ \theta_2 & \theta_1 & 0 \\ 0 & \theta_2 & \theta_1 \end{bmatrix}, B(\theta) = \begin{bmatrix} \theta_3 \\ 0 \\ 0 \end{bmatrix}$$

$$H = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, C(\theta) = \begin{bmatrix} \theta_4 \\ 0 \\ 0 \end{bmatrix}^T.$$



- The unknown input $f(k)$ is assumed to be deterministic. No prior knowledge of $f(k)$ is available except the persistent excitation condition.
- In the literature, only the blind SIMO system identification framework can handle this identification problem.

$$
\begin{aligned}
y_1(k) &= H_1(z)u(k) \\
y_2(k) &= H_2(z)u(k) \\
&\vdots \\
y_L(k) &= H_L(z)u(k)
\end{aligned}
$$



Since the SIMO model is a decoupled network model, it cannot cope with the identification of general structured state-space models.

Outline

Identifiability analysis 1 - Full state observation

Given the full state observation, an equivalent state-space model can be written as

$$x(k+1) = \underbrace{(A + HQ_1)}_{\hat{A}} x(k) + \underbrace{(B + HQ_2)}_{\hat{B}} u(k) + H \underbrace{[f_k - Q_1 x(k) - Q_2 u(k)]}_{\hat{f}(k)}$$

$$\hat{y}(k) = Cx(k)$$

where $Q_1 \in \mathbb{R}^{r\times n}$ and $Q_2 \in \mathbb{R}^{r\times m}$ are ambiguity matrices, and $\hat{y}(k) = y(k) - w(k)$ denotes the noise-free output.

Without any structural constraints, the system matrices cannot be identified even if the states can be fully observed. For example, when (A, H) is controllable, the system poles can be arbitrarily assigned.

Identifiability analysis 1 - Full state observation

Theorem 1. Let $\mathcal{P}_H^{\perp} = I - H(H^T H)^{-1} H^T$. Suppose that the state sequence $x(k)$ can be fully observed, the original identification problem boils down to estimating the parameter vector θ from the following modified state-space model

$$\mathcal{P}_H^{\perp} x(k+1) = \mathcal{P}_H^{\perp} A(\theta) x(k) + \mathcal{P}_H^{\perp} B(\theta) u(k)$$
$$\hat{y}(k) = C(\theta) x(k).$$

Example 1 (continue)

$$\begin{bmatrix} x_1(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} \theta_1 & 0 & \theta_2 \\ 0 & \theta_2 & \theta_1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} \theta_3 \\ 0 \end{bmatrix} u(k)$$
$$\hat{y}(k) = \theta_4 x_1(k).$$

The state-equation of the second agent is removed due to the unknown input signal; however, θ can be estimated because of its duplicates in other state equations. When $H = [1\ 1\ 1]^T$, the system parameters are still identifiable.

Identifiability analysis 2 - No direct state observation

The past and future data equations are given as

$$\hat{Y}_p = OX_p + T_u U_p + T_f F_p$$
$$\hat{Y}_f = OX_f + T_u U_f + T_f F_f$$

where

$$O = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{s-1} \end{bmatrix}, \quad T_u = \begin{bmatrix} 0 & & \\ CB & \ddots & \\ \vdots & \ddots & 0 \\ CA^{s-2}B & \cdots & CB \end{bmatrix}, \quad T_f = \begin{bmatrix} 0 & & \\ CH & \ddots & \\ \vdots & \ddots & 0 \\ CA^{s-2}H & \cdots & CH \end{bmatrix},$$

$$X_p = [x(1)\ x(2)\ \cdots\ x(h)], \quad X_f = [x(s)\ x(s+1)\ \cdots\ x(s+h-1)],$$

$$Y_p = \begin{bmatrix} y(1) & y(2) & \cdots & y(h) \\ y(2) & y(3) & \cdots & y(h+1) \\ \vdots & \vdots & \iddots & \vdots \\ y(s) & y(s+1) & \cdots & y(s+h-1) \end{bmatrix}, \quad Y_f = \begin{bmatrix} y(s) & \cdots & y(s+h-1) \\ y(s+1) & \cdots & y(s+h) \\ \vdots & \iddots & \vdots \\ y(2s-1) & \cdots & y(2s+h-2) \end{bmatrix}.$$

Lemma 1. Assume that CH has full column rank and the state-space model described by the matrix of tuples $(A, H, C, 0)$ is strongly observable , i.e., $\mathrm{rank}[O \quad T_f] = n + \mathrm{rank}[T_f]$. Then, the row subspace of the state space sequence can be estimated as

$$\mathrm{Row}[X_f] = \mathrm{Row}\begin{bmatrix} U_p \\ \hat{Y}_p \end{bmatrix} \cap \mathrm{Row}\begin{bmatrix} U_f \\ \hat{Y}_f \end{bmatrix}.$$

The above lemma indicates that the system state can be estimated up to a similarity transformation, i.e.,

$$x(k) = Q\hat{x}(k)$$

where $\hat{x}(k)$ is the state estimate and Q is the similarity transformation matrix.

Given the state estimate $\hat{x}(k)$, the state-space model can be written as

$$
\begin{aligned}
Q\hat{x}(k+1) &= AQ\hat{x}(k) + Bu(k) + Hf(k) \\
\hat{y}(k) &= CQ\hat{x}(k).
\end{aligned}
$$

By Theorem 1, the identifiability of θ in the original state-space model is equivalent to the identifiability of (θ, Q) of the following state-space model

$$
\begin{aligned}
\mathcal{P}_H^{\perp} Q\hat{x}(k+1) &= \mathcal{P}_H^{\perp} A(\theta)Q\hat{x}(k) + \mathcal{P}_H^{\perp} B(\theta)u(k) \\
\hat{y}(k) &= C(\theta)Q\hat{x}(k).
\end{aligned}
$$

Theorem 2. Let the low-rank factorization of $\mathcal{P}_H^{\perp}$ be given as $\mathcal{P}_H^{\perp} = U_H V_H^T$ where $U_H \in \mathbb{R}^{n\times(n-r)}$ and $V_H \in \mathbb{R}^{n\times(n-r)}$ have full column rank. Then, the above state-space model is identifiable if and only if, for any nonsingular matrices $Q^* \in \mathbb{R}^{n\times n}$ and $\Pi \in \mathbb{R}^{(n-r)\times(n-r)}$, the following equations

$$
\begin{aligned}
&\Pi V_H^T Q^* = V_H^T Q, \quad \Pi V_H^T A(\theta^*)Q^* = V_H^T A(\theta)Q \\
&\Pi V_H^T B(\theta^*) = V_H^T B(\theta), \quad C(\theta^*)Q^* = C(\theta)Q
\end{aligned}
$$

yield that $\theta = \theta^*$, $Q = Q^*$ and $\Pi = I$.

Example 2 The identifiability of the state-space model with the following structured system matrices is considered:

$$A(\theta) = \begin{bmatrix} \theta_1 & 0 & \theta_2 \\ \theta_2 & \theta_1 & 0 \\ 0 & \theta_2 & \theta_1 \end{bmatrix}, \; B(\theta) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \; H = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \; C(\theta) = \begin{bmatrix} \theta_3 & 1 & 0 \\ 0 & 0 & \theta_3 \end{bmatrix}.$$

Choosing the matrix V_H as $V_H^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, we have the following equation group

$$\begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} \begin{bmatrix} q_1^* \\ q_3^* \end{bmatrix} = \begin{bmatrix} q_1 \\ q_3 \end{bmatrix} \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix},$$

$$\begin{bmatrix} \theta_1^* & 0 & \theta_2^* \\ 0 & \theta_2^* & \theta_1^* \end{bmatrix} \begin{bmatrix} q_1^* \\ q_2^* \\ q_3^* \end{bmatrix} = \begin{bmatrix} \theta_1 & 0 & \theta_2 \\ 0 & \theta_2 & \theta_1 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix},$$

$$\begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} \theta_3^* & 1 & 0 \\ 0 & 0 & \theta_3^* \end{bmatrix} \begin{bmatrix} q_1^* \\ q_2^* \\ q_3^* \end{bmatrix} = \begin{bmatrix} \theta_3 & 1 & 0 \\ 0 & 0 & \theta_3 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}.$$

For any non-singular matrices Π and Q^*, it can be verified that $Q = Q^*$ and $\theta = \theta^*$.

Outline

Subspace identification method

Following identifiability analysis, the subspace identification method will be developed by two steps:

1. State estimation using the subspace intersection;
2. Parameter estimation by difference-of-convex optimization.

State estimation

By Lemma 1, the system state can be estimated as follows

$$\mathrm{Row}[X_f] = \mathrm{Row}\begin{bmatrix} U_p \\ \hat{Y}_p \end{bmatrix} \cap \mathrm{Row}\begin{bmatrix} U_f \\ \hat{Y}_f \end{bmatrix}.$$

The subspace intersection can be computed by two steps:

1. Take the SVD as follows

$$\begin{bmatrix} U_p \\ Y_p - W_p \\ U_f \\ Y_f - W_f \end{bmatrix} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$$

2. Partition U_2 into two sub-matrices of the same size $U_2 = \begin{bmatrix} U_{21}^T & U_{22}^T \end{bmatrix}^T$, the intersection can be computed as

$$\hat{X}_f = U_{21}^T \begin{bmatrix} U_p \\ Y_p - W_p \end{bmatrix}.$$

Lemma 2 If the matrix $[O \quad T_f]$ is a tall matrix, then the variance of the measurement noise can be estimated as

$$\sigma_w^2 = \lambda_{min}\left[\lim_{h\to\infty}\frac{1}{h}Y_pY_p^T - Y_pU_p^T(U_pU_p^T)^{-1}U_pY_p^T\right] \tag{1}$$

where $\lambda_{min}(\cdot)$ represents the least eigenvalue.

Then, an unbiased estimate of the subspace intersection can be computed as follows.

$$\lim_{h\to\infty}\frac{1}{h}\begin{bmatrix} U_p \\ Y_p \\ U_f \\ Y_f \end{bmatrix}\begin{bmatrix} U_p \\ Y_p \\ U_f \\ Y_f \end{bmatrix}^T - \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sigma_w^2 I & 0 & \begin{bmatrix} 0 & 0 \\ \sigma_w^2 I & 0 \end{bmatrix} \\ 0 & 0 & 0 & 0 \\ 0 & \begin{bmatrix} 0 & \sigma_w^2 I \\ 0 & 0 \end{bmatrix} & 0 & \sigma_w^2 I \end{bmatrix}$$

$$= \begin{bmatrix} U_1 & U_2 \end{bmatrix}\begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix}\begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix}.$$

Partition U_2 into $U_2 = \begin{bmatrix} U_{21}^T & U_{22}^T \end{bmatrix}^T$, we have that

$$\hat{X}_f = U_{21}^T\begin{bmatrix} U_p \\ Y_p \end{bmatrix} = QX_f + \underbrace{U_{21}^T\begin{bmatrix} 0 \\ W_p \end{bmatrix}}_{\text{estimation error}}.$$

Parameter estimation

Define the following row sequences

$$\bar{x}_h = \hat{x}(s+2 : s+h+1), \;\; x_h = \hat{x}(s+1 : s+h), \;\; u_h = u(s+1 : s+h), \;\; y_h = y(s+1 : s+h).$$

Substituting the estimated state into the original state-space model yields

$$\begin{aligned} \mathcal{P}_H^{\perp} Q\bar{x}_h &= \mathcal{P}_H^{\perp} A(\theta) Q x_h + \mathcal{P}_H^{\perp} B(\theta) u_h + \eta_1 \\ y_h &= C(\theta) Q x_h + w_h + \eta_2, \end{aligned}$$

where η_1 and η_2 are the state estimation errors that are asymptotically uncorrelated with the inputs and states.

Step 1. The parameters can be estimated by solving the following optimization problem

$$\min_{Q,\theta} \quad \left\| \mathcal{P}_H^{\perp} Q\bar{x}_h - \mathcal{P}_H^{\perp} A(\theta) Q x_h - \mathcal{P}_H^{\perp} B(\theta) u_h \right\|_F^2 + \| y_h - C(\theta) Q x_h \|_F^2 .$$

Step 2. The bilinear estimation problem is equivalent to a rank-constrained optimization problem

$$\min_{Q_i,\mathcal{A},\mathcal{C},\theta,\Gamma} \left\| \mathcal{P}_H^{\perp} Q\bar{x}_h - \mathcal{P}_H^{\perp}\mathcal{A}x_h - \mathcal{P}_H^{\perp}B(\theta)u_h \right\|_F^2 + \|y_h - \mathcal{C}x_h\|_F^2$$

$$s.t. \quad \mathcal{A} = A_0 Q + \sum_{i=1}^{l} A_i Q_i, \quad \mathcal{C} = C_0 Q + \sum_{i=1}^{l} C_i Q_i$$

$$\Gamma = \begin{bmatrix} 1 & \theta_1 & \cdots & \theta_l \\ vec(Q) & vec(Q_1) & \cdots & vec(Q_l) \end{bmatrix}$$

$$\mathrm{rank}[\Gamma] = 1 \;\; (\|\Gamma\|_* - \|\Gamma\|_2 = 0)$$

Step 3. By treating the nonnegative constraint $\|\Gamma\|_* - \|\Gamma\|_2$ as a penalty, we can obtain that

$$\min_{Q_i,\mathcal{A},\mathcal{C},\theta,\Gamma} \left\| \mathcal{P}_H^{\perp} Q\bar{x}_h - \mathcal{P}_H^{\perp}\mathcal{A}x_h - \mathcal{P}_H^{\perp}B(\theta)u_h \right\|_F^2 + \|y_h - \mathcal{C}x_h\|_F^2 + \lambda\left(\|\Gamma\|_* - \|\Gamma\|_2\right)$$

$$s.t. \quad \mathcal{A} = A_0 Q + \sum_{i=1}^{l} A_i Q_i, \quad \mathcal{C} = C_0 Q + \sum_{i=1}^{l} C_i Q_i$$

$$\Gamma = \begin{bmatrix} 1 & \theta_1 & \cdots & \theta_l \\ vec(Q) & vec(Q_1) & \cdots & vec(Q_l) \end{bmatrix}$$

This optimization problem is solved by the sequential convex programming approach.

Outline

Blind identification of an SIMO model

Consider the following SIMO model $y(k) = \sum_{i=1}^{L} h_i u(k-i) + w(k)$ or

$$x(k+1) = Ax(k) + Hf(k), \quad y(k) = Cx(k) + w(k)$$

where

$$A = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix}, \quad H = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} h_1 & h_2 & \cdots & h_L \end{bmatrix}.$$

The data equation has the following extended observability matrix and convolution matrix

$$O = \begin{bmatrix} h_1 & \cdots & h_L \\ 0 & \ddots & \vdots \\ \vdots & \ddots & h_1 \\ 0 & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}, T_f = \begin{bmatrix} 0 & & & & \\ h_L & 0 & & & \\ \vdots & \ddots & \ddots & & \\ h_1 & \cdots & h_L & 0 & \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & h_1 & \cdots & h_L \end{bmatrix}.$$

It can be seen that $[O \quad T_f]$ is exactly a convolution matrix of the FIR model.

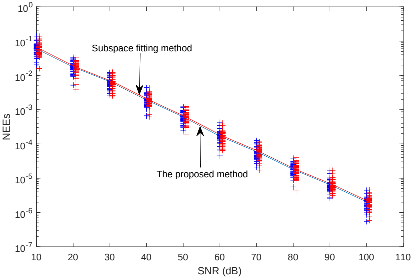


Figure: Comparison between the subspace fitting method and the proposed method.

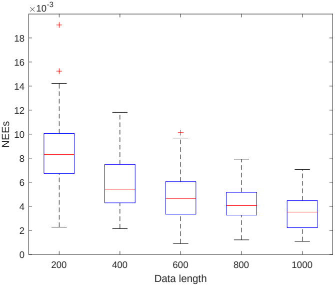


Figure: Performance of the proposed method against the data length.

Structured state-space model

Consider Example 2 with the system matrices

$$A(\theta) = \begin{bmatrix} \theta_1 & 0 & \theta_2 \\ \theta_2 & \theta_1 & 0 \\ 0 & \theta_2 & \theta_1 \end{bmatrix},\ B(\theta) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\ H = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},\ C(\theta) = \begin{bmatrix} \theta_3 & 1 & 0 \\ 0 & 0 & \theta_3 \end{bmatrix}$$

and the parameter vector $\theta = \begin{bmatrix} 0.5 & 0.3 & 1 \end{bmatrix}^T$.

It has been verified that this structured state-space model is identifiable in the presence of the unknown input signal.

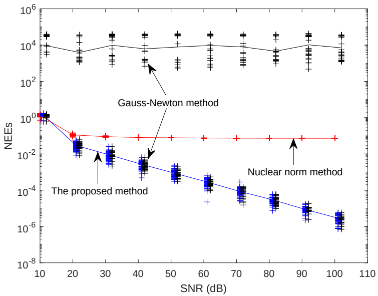


Figure: Performance comparison among the Gauss-Newton method, nuclear-norm method and the proposed method.

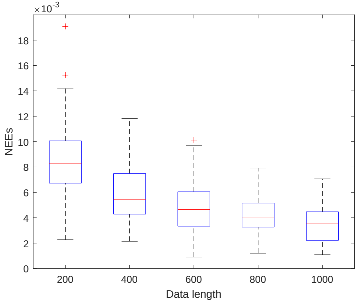


Figure: The recovered unknown input signal at SNR=30dB.

Conclusions

Conclusions

1. Identifiability conditions for the structured state-space models
2. Subspace-based state estimation with noise compensation
3. Difference-of-convex programming approach for the bilinear estimation

Future work

Simultaneous system structure reconstruction and the system parameter estimation using (partially) measurable input-output data.

THANK YOU!