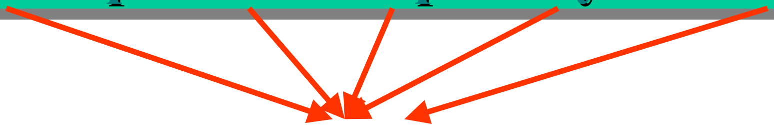
Subspace State-Space System IDentification

Author: Juan Carlos Gómez

Subspace State-Space System IDentification



4SID Methods

- **Properties**
  - They combine tools of **System Theory**, **Numerical Linear Algebra** and **Geometry** (projections).
  - They have their origin in **Realization Theory** as developed in the 60/70s (Ho & Kalman, 1966).
  - They provide reliable state-space models of **multivariable** LTI systems **directly** from input-output data.
  - They don’t require iterative optimization procedures → no problems with local minima, convergence and initialization.

- They don't require a particular (canonical) state-space realization → numerical conditioning improves.
- They require a modest computational load in comparison to traditional identification methods like PEM.
- The algorithms can be (they have been) efficiently implemented in software like **Matlab**.
- Main computational tools are QR and SVD.
- All subspace methods compute at some stage the **subspace** spanned by the columns of the extended observability matrix.
- The various algorithms (*e.g.,* N4SID, MOESP, CVA) differ in the way the extended observability matrix is estimated and also in the way it is used to compute the system matrices.

The system model

$$x_{k+1} = Ax_k + Bu_k + Ke_k$$
$$y_k = Cx_k + Du_k + e_k$$

State-space model in innovation form

The identification problem

To estimate the system matrices *(A, B, C, D)* and *K* , and the model order *n*, from an *(N+α-1)*-point data set of input and output measurements

$$\{u_k, y_k\}_{k=1}^{N+\alpha-1}$$

❑ Realization-based 4SID Methods

For a LTI system, a **minimal** state-space realization *(A, B, C, D)* completely defines the input-output properties of the system through

$$y_k = \sum_{\ell=0}^{\infty} h_\ell u_{k-\ell}$$ **convolution sum**

where the impulse response coefficients h_ℓ are related to the system matrices by

$$h_\ell = \begin{cases} D & , \quad \ell = 0 \\ CA^{\ell-1}B & , \quad \ell > 0 \end{cases}$$

$$H_{ij} = \begin{bmatrix} h_1 & h_2 & \cdots & h_j \\ h_2 & h_3 & \cdots & h_{j+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_i & h_{i+1} & \cdots & h_{i+j-1} \end{bmatrix}$$

Impulse Response Hankel Matrix

$$H_{ij} = \Gamma_i \mathbf{C}_j$$

Extended Observability Matrix ***(i>n)***

Extended Controlability Matrix ***(j>n)***

An estimate of the extended observability matrix can be computed by a **full rank** factorization of the impulse response Hankel matrix. This factorization is provided by the SVD of matrix H_{ij}.

$$H_{ij} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \approx U_1 \Sigma_1 V_1^T = \underbrace{\left(U_1 \Sigma_1^{1/2} \right)}_{\hat{\Gamma}_i} \underbrace{\left(\Sigma_1^{1/2} V_1^T \right)}_{\hat{\mathbf{C}}_j}$$

rank reduction

In the absence of noise, H_{ij} will be a rank n matrix, and Σ_1 will contain the n non-zero singular values $\rightarrow$ **model order is computed.** In the presence of noise, H_{ij} will have full rank and a rank reduction stage will be required for the model order determination.

Problems: it is necessary to measure or to estimate (for example, via correlation analysis) the impulse response of the system $\rightarrow$ **not good**

❑ Direct 4SID Methods

$$\mathbf{Y}_\alpha = \Gamma_\alpha \mathbf{X} + H_\alpha \mathbf{U}_\alpha + \mathbf{N}_\alpha$$ **fundamental equation** (1)

$$\mathbf{Y}_\alpha = \begin{bmatrix} y_1 & y_2 & \cdots & y_N \\ y_2 & y_3 & \cdots & y_{N+1} \\ \vdots & \vdots & \vdots & \vdots \\ y_{\alpha-1} & y_\alpha & \cdots & y_{N+\alpha-1} \end{bmatrix}$$

Output block Hankel matrix

(In a similar way are defined the Input block Hankel matrix $\mathbf{U}_\alpha$ and the Noise block Hankel matrix $\mathbf{N}_\alpha$.)

$$\Gamma_\alpha = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\alpha-1} \end{bmatrix}$$

Extended $(\alpha > n)$ **Observability Matrix**

$$\mathbf{X} = [x_1, x_2, \cdots, x_N]$$

State Sequence Matrix

$$H_\alpha = \begin{bmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ CA^{\alpha-2}B & CA^{\alpha-3}B & CA^{\alpha-4}B & \cdots & D \end{bmatrix}$$

Lower triangular block Toeplitz matrix of impulse responses (unknown).

❑ The main idea of Direct 4SID methods

In the absence of noise ($\boldsymbol{N}_\alpha = 0$), eq. (1) becomes

$$\mathbf{Y}_\alpha = \Gamma_\alpha \mathbf{X} + H_\alpha \mathbf{U}_\alpha \qquad (2)$$

and the part of the output which does not emanate from the state can be removed by multiplying (from the right) both sides of eq. (2) by the **orthogonal projection onto the null space of $\mathbf{U}_\alpha$**, *i.e.* by

$$\Pi^{\perp}_{\mathbf{U}^T_\alpha} \stackrel{\Delta}{=} I - \mathbf{U}^T_\alpha \left(\mathbf{U}_\alpha \mathbf{U}^T_\alpha\right)^{-1} \mathbf{U}_\alpha \stackrel{\Delta}{=} \mathbf{U}^{\perp}_\alpha$$

orthogonal projection

such that $\mathbf{U}_\alpha \mathbf{U}^{\perp}_\alpha = I$

This yields

$$\mathbf{Y}_\alpha \mathbf{U}^{\perp}_\alpha = \Gamma_\alpha \mathbf{X} \mathbf{U}^{\perp}_\alpha \qquad (3)$$

Note that the matrix on the left depends exclusively on the input-output data. Then, a **full rank factorization** of this matrix will provide an estimate $\hat{\Gamma}_\alpha$ of the extended observability matrix. Estimates of the corresponding system matrices can be obtained by resorting to the **shift invariance property** of the extended observability matrix, and by solving a system of linear equations in the least squares sense.

The factorization is provided by the SVD of the matrix on the left side

$$\mathbf{Y}_\alpha \mathbf{U}_\alpha^\perp = [U_1 \quad U_2]\begin{bmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix}V_1^T\\ V_2^T\end{bmatrix} \approx U_1\Sigma_1 V_1^T = \underbrace{\left(U_1\Sigma_1^{1/2}\right)}_{\hat{\Gamma}_\alpha}\left(\Sigma_1^{1/2}V_1^T\right) \quad (4)$$

rank reduction
(model order estimation)

(In the absence of noise $\Sigma_2 = 0$)

❑ Weighting Matrices

Row and column weighting matrices can be introduced in (4) before performing the SVD of the matrix in the left hand side. Any choice of positive-definite weighting matrices W_r and W_c will result in consistent estimates of the extended observability matrix.

$$W_r \mathbf{Y}_\alpha \mathbf{U}_\alpha^\perp W_c = [U_1 \quad U_2]\begin{bmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{bmatrix}\begin{bmatrix}V_1^T\\ V_2^T\end{bmatrix} \approx U_1\Sigma_1 V_1^T = \underbrace{\left(U_1\Sigma_1^{1/2}\right)}_{\hat{\Gamma}_\alpha}\left(\Sigma_1^{1/2}V_1^T\right)$$

change of coordinates in state-space

Existing algorithms employ the following choices for matrices W_r and W_c,

- **MOESP** (Verhaegen, 1994): $W_r = I, \quad W_c = \left(\frac{1}{N}\Phi\Pi_{U_\alpha^T}^\perp\Phi^T\right)^{-1}\Phi\Pi_{U_\alpha^T}^\perp$
- **CVA** (Larimore, 1990): $W_r = \left(\frac{1}{N}\mathbf{Y}_\alpha\Pi_{U_\alpha^T}^\perp\mathbf{Y}_\alpha^T\right)^{-1/2}, \quad W_c = \left(\frac{1}{N}\Phi\Pi_{U_\alpha^T}^\perp\Phi^T\right)^{-1/2}$
- **N4SID** (Van Overschee and de Moor, 1994):

$$W_r = I, \quad W_c = \left(\frac{1}{N}\Phi\Pi_{U_\alpha^T}^\perp\Phi^T\right)^{-1}\Phi$$

❑ Computation of the system matrices

Given an estimate $\hat{\Gamma}_\alpha$ of the extended observability matrix, estimates of the system matrices can be computed as:

- $\hat{C}$: first row block of $\hat{\Gamma}_\alpha$
- $\hat{A}$: solving in the least squares sense

$$\overline{\overline{\Gamma_\alpha}} = \underline{\underline{\Gamma_\alpha}}\hat{A}$$

shift-invariance property

- $\hat{B}$ and $\hat{D}$: solving a system of linear equations

❑ Presence of noise

In the presence of noise

$$\mathbf{Y}_\alpha = \Gamma_\alpha \mathbf{X} + H_\alpha \mathbf{U}_\alpha + \mathbf{N}_\alpha$$

and

$$\mathbf{Y}_\alpha \mathbf{U}_\alpha^\perp = \Gamma_\alpha \mathbf{X} + \mathbf{N}_\alpha \mathbf{U}_\alpha^\perp$$

noise term needs to be removed

The noise term can be removed by **correlating it away** with a suitable matrix. This can be interpreted as an **oblique projection.**