# Stability of reduced order models in subspace identification


Ivan Markovsky[1], Jan C. Willems[1], Paolo Rapisarda[2] and Bart De Moor[1]

1 – Departement Elektrotechniek, Katholieke Universiteit Leuven, 2 – Departement Mathematics, Universiteit Maastricht


## Abstract


Given an observed input/output time-series $(u_t, y_t)$, $t = 0, 1, \ldots, t_f - 1$, $u_t \in \mathbb{R}^m$, $y_t \in \mathbb{R}^p$, find an approximate time-series model

$$\begin{aligned} x_{t+1} &= \hat{A}x_t + \hat{B}u_t, \qquad x_t \in \mathbb{R}^r, \\ y_t &= \hat{C}x_t + \hat{D}u_t. \end{aligned}$$

**Problem:**

Establish stability preservation and an error bound.

**Approach:**

Subspace identification, balanced truncated state sequence, least squares.

Equivalently, this is the problem of balanced reduction of the Most Powerful Unfalsified Model directly from the given data.


## Overview of deterministic subspace identification

Let the observed time-series $(u_t, y_t)$ be generated by the LTI system

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t, \qquad x_t \in \mathbb{R}^n, \\ y_t &= Cx_t + Du_t. \end{aligned} \tag{1}$$

For a given natural number $i > n$, define

$$X_\mathrm{p} := \begin{bmatrix} x_0 & x_1 & \cdots & x_t & \cdots & x_{j-1} \end{bmatrix} \quad \rightarrow \quad \text{sequence of initial conditions}$$

$$U := \begin{bmatrix} u_0 & u_1 & \cdots & u_t & \cdots & u_{j-1} \\ u_1 & u_2 & \cdots & u_{t+1} & \cdots & u_j \\ \vdots & \vdots & & \vdots & & \vdots \\ u_{2i-1} & u_{2i} & \cdots & u_{t+2i-1} & \cdots & u_{j+2i-2} \end{bmatrix} \quad \rightarrow \quad \text{every column is a sequence of inputs}$$

$$Y := \begin{bmatrix} y_0 & y_1 & \cdots & y_t & \cdots & y_{j-1} \\ y_1 & y_2 & \cdots & y_{t+1} & \cdots & y_j \\ \vdots & \vdots & & \vdots & & \vdots \\ y_{2i-1} & y_{2i} & \cdots & y_{t+2i-1} & \cdots & y_{j+2i-2} \end{bmatrix} \quad \rightarrow \quad \text{every column is the corresponding sequence of outputs}$$

Let $W := \left[\begin{smallmatrix} U \\ Y \end{smallmatrix}\right]$. $W(:, t)$ gives an I/O response with initial condition $x_t$. Split the data in $U$ and $Y$ into two equal length parts

$$U = \begin{bmatrix} U_\mathrm{p} \\ U_\mathrm{f} \end{bmatrix}, \qquad Y = \begin{bmatrix} Y_\mathrm{p} \\ Y_\mathrm{f} \end{bmatrix}.$$

Assume that $\operatorname{row\,span}(U_\mathrm{f}) \cap \operatorname{row\,span}(W_\mathrm{p}) = \{0\}$. Sufficient conditions is the persistency of excitation and $\operatorname{row\,span}(U_\mathrm{f}) \cap \operatorname{row\,span}(X_\mathrm{p}) = \{0\}$.

$$\mathbb{R}^j = \underbrace{\operatorname{row\,span}(W_\mathrm{p})}_{\substack{\text{response due to} \\ \text{past inputs and ini. cond.}}} \oplus \underbrace{\operatorname{row\,span}(U_\mathrm{f})}_{\substack{\text{response due} \\ \text{to future inputs}}} \oplus \underbrace{\operatorname{row\,span}\left(\begin{bmatrix} W_\mathrm{p} \\ U_\mathrm{f} \end{bmatrix}\right)^\perp}_{\text{noise}}$$

Define $\bar{Y}_\mathrm{f} := Y_\mathrm{f}/_{U_\mathrm{f}}W_\mathrm{p}$ (row-wise) as the component of $Y_\mathrm{f}$ lying in $\operatorname{row\,span}(W_\mathrm{p})$.

$$\bar{Y}_\mathrm{f} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{i-1} \end{bmatrix} \begin{bmatrix} x_i & x_{i+1} & \cdots & x_{i+j-1} \end{bmatrix} =: \Gamma X_\mathrm{f}$$

## Subspace identification algorithm

$\bar{Y}_\mathrm{f}$ contains free responses. Any rank revealing factorization

$$\bar{Y}_\mathrm{f} = LR^\top, \qquad L, R \in \mathbb{R}^{mi \times n}, \qquad L, R \text{ — full column rank},$$

defines a valid $\Gamma = L$ and a corresponding state sequence $X_\mathrm{f} = R^\top$.
The rank revealing factorization is not unique and this corresponds to the non uniqueness of $\Gamma$ and $X_\mathrm{f}$ due to the choice of the state space basis. We will use the SVD of $\bar{Y}_\mathrm{f}$ to compute the factorization.
With $X_\mathrm{f}$ computed, one can estimate $A$, $B$, $C$, and $D$ from a LS problem.

### Algorithm 1 (Deterministic subspace identification)

1. Input: the Hankel matrix $W$.
2. Compute the oblique projection $\bar{Y}_\mathrm{f} := Y_\mathrm{f}/_{U_\mathrm{f}}W_\mathrm{p}$.
3. Compute the SVD of the oblique projection $\bar{Y}_\mathrm{f} = U_1 S_1 V_1^\top$.
4. Define $\Gamma := U_1\sqrt{S_1}$ and $X_\mathrm{f} = \sqrt{S_1}V_1^\top$.
5. Find the parameters from the LS problem

$$\begin{bmatrix} \sigma X_\mathrm{f} \\ \sigma^* Y_\mathrm{f} \end{bmatrix} = \begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix} \begin{bmatrix} \sigma^* X_\mathrm{f} \\ \sigma^* U_\mathrm{f} \end{bmatrix}, \tag{2}$$

where

$$\sigma X_\mathrm{f} := \begin{bmatrix} x_{i+1} & x_{i+2} & \cdots & x_{i+j-1} \end{bmatrix} \quad \text{and} \quad \sigma^* X_\mathrm{f} := \begin{bmatrix} x_i & x_{i+1} & \cdots & x_{i+j-2} \end{bmatrix}.$$

6. Output: the estimates $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{D}$.

## Balanced model subspace identification algorithm

Let $V$ be a $j \times j$ to-be-determined matrix, and consider the SVD of $\bar{Y}_\mathrm{f}V$

$$\bar{Y}_\mathrm{f}V = U_1\Sigma_1 V_1^\top.$$

Define $\Gamma := U_1\sqrt{\Sigma_1}$ and $X_\mathrm{f} := \Gamma^+\bar{Y}_\mathrm{f}$, where $\Gamma^+$ is a right inverse of $\Gamma$.
$\bar{Y}_\mathrm{f} = \Gamma X_\mathrm{f}$, so that, $\Gamma$ is an extended observability matrix and $X_\mathrm{f}$ is a corresponding state sequence in certain basis.

$$\sqrt{\Sigma_1}V_1^\top = X_\mathrm{f}V \implies X_\mathrm{f}VV^\top X_\mathrm{f}^\top = \Sigma_1 \tag{3}$$

For $i \to \infty$ and $A$ Hurwitz, $X_\mathrm{f} = \Delta U_\mathrm{p} + A^i X_\mathrm{p} \approx \Delta U_\mathrm{p}$, where $\Delta := \begin{bmatrix} A^{i-1}B & A^{i-2}B & \cdots & AB & B \end{bmatrix}$ so from (3), we have

$$\Delta U_\mathrm{p}VV^\top U_\mathrm{p}^\top \Delta^\top \approx \Sigma_1.$$

Now, if we choose $V$ such that $U_\mathrm{p}VV^\top U_\mathrm{p}^\top = I$, then the controllability gramian becomes $\Delta\Delta^\top \approx \Sigma_1$ and is approximately equal to the observability gramian $\Gamma^\top\Gamma = \Sigma_1$, so the state-space basis fixed with the choice of $\Gamma$ and $X_\mathrm{f}$ is approximately balanced. A particular matrix $V$ is

$$V = U_\mathrm{p}^\top(U_\mathrm{p}U_\mathrm{p}^\top)^{-1}[I \;\; 0].$$

### Algorithm 2 (Identification of a balanced model)

1. Input: the Hankel matrix $W$.
2. Compute $V = U_\mathrm{p}^\top(U_\mathrm{p}U_\mathrm{p}^\top)^{-1}[I \;\; 0]$.
3. Compute the oblique projection $\bar{Y}_\mathrm{f} := Y_\mathrm{f}/_{U_\mathrm{f}}W_\mathrm{p}$.
4. Compute the SVD of $\bar{Y}_\mathrm{f}V = U_1 S_1 V_1^\top$.
5. Define $\Gamma := U_1\sqrt{S_1}$ and $X_\mathrm{f} = \Gamma^+\bar{Y}_\mathrm{f}$.
6. Find the parameters from the LS problem (2).
7. Output: the estimates $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{D}$.

## Reduced order model estimation by truncation and LS

In subspace identification, one finds first the state sequence $X$, and then $(A, B, C, D)$. Algorithm 2 finds even an approximate finite-time balanced system. But solving for $(A, B, C, D)$ requires solving equations of order equal to the order of the unreduced system. Reducing $X$ first, leads to equations of order equal to that of the reduced system!
For a given natural number $r < n$, partition the balanced state sequence $X_\mathrm{f}$ as follows

$$X_\mathrm{f} =: \begin{bmatrix} X_\mathrm{r} \\ X_\mathrm{trunc} \end{bmatrix}, \qquad \text{where} \qquad \operatorname{row\,dim}(X_\mathrm{r}) = r,$$

and find the reduced order model parameters $\hat{A}_\mathrm{r}$, $\hat{B}_\mathrm{r}$, $\hat{C}_\mathrm{r}$, $\hat{D}_\mathrm{r}$ by the least squares problem

$$\begin{bmatrix} \sigma X_\mathrm{r} \\ \sigma^* Y_\mathrm{f} \end{bmatrix} = \begin{bmatrix} \hat{A}_\mathrm{r} & \hat{B}_\mathrm{r} \\ \hat{C}_\mathrm{r} & \hat{D}_\mathrm{r} \end{bmatrix} \begin{bmatrix} \sigma^* X_\mathrm{r} \\ \sigma^* U_\mathrm{f} \end{bmatrix}. \tag{4}$$

While for the classical approach there is a proof of stability of the reduced order model and a bound for the $\mathcal{H}_\infty$ norm of the error system, for the alternative approach there are no such results yet. We state the following conjecture:

**General stability conjecture**

Let $(A, B, C, D)$ be a minimal, asymptotically stable, and balanced $n$-th order system. Given an infinite input/output sequence $(u_t, y_t)$, $t = 0, 1, \ldots$, where $u$ is persistently exciting of order at least $2n$, construct a balanced state sequence $X_\mathrm{f}$ via Algorithm 2. Solve the least squares problem (4) for the truncated state sequence $X_\mathrm{r}$. Then $\hat{A}_\mathrm{r}$ is Schur.

## Preliminary result: autonomous model

Theorem (Autonomous model). Consider the asymptotically stable model

$$x_{t+1} = Ax_t, \tag{5}$$

and let $x \in \ell_2(\mathbb{Z}, \mathbb{R}^n)$ be a trajectory of (5). Denote by $X$ the infinite matrix

$$X = \begin{bmatrix} x_0 & x_1 & \cdots & x_t & \cdots \end{bmatrix}.$$

For a given natural number $r < n$, partition $X$ as follows

$$X =: \begin{bmatrix} X_\mathrm{r} \\ X_\mathrm{trunc} \end{bmatrix}, \qquad \text{where} \qquad \operatorname{row\,dim}(X_\mathrm{r}) = r.$$

The solution

$$\hat{A}_\mathrm{r} = \sigma X_\mathrm{r} X_\mathrm{r}^\top (X_\mathrm{r}X_\mathrm{r}^\top)^{-1}$$

of the least squares problem

$$\min_{\hat{A}_\mathrm{r}} ||\sigma X_\mathrm{r} - \hat{A}_\mathrm{r}X_\mathrm{r}||_{\ell_2}^2$$

is Schur.

Proof: Since for any $X_\mathrm{r}$, $(\sigma X_\mathrm{r})(\sigma X_\mathrm{r})^\top \leq X_\mathrm{r}X_\mathrm{r}^\top$, we have

$$\begin{aligned} 0 &\leq \begin{bmatrix} X_\mathrm{r} \\ \sigma X_\mathrm{r} \end{bmatrix} \begin{bmatrix} X_\mathrm{r}^\top & (\sigma X_\mathrm{r})^\top \end{bmatrix} \\ &= \begin{bmatrix} X_\mathrm{r}X_\mathrm{r}^\top & X_\mathrm{r}(\sigma X_\mathrm{r})^\top \\ (\sigma X_\mathrm{r})X_\mathrm{r}^\top & (\sigma X_\mathrm{r})(\sigma X_\mathrm{r})^\top \end{bmatrix} \\ &\leq \begin{bmatrix} X_\mathrm{r}X_\mathrm{r}^\top & X_\mathrm{r}(\sigma X_\mathrm{r})^\top \\ (\sigma X_\mathrm{r})X_\mathrm{r}^\top & X_\mathrm{r}X_\mathrm{r}^\top \end{bmatrix}. \end{aligned}$$

The Schur complement of the (2,2) block in the right most matrix is

$$X_\mathrm{r}X_\mathrm{r}^\top - \underbrace{(\sigma X_\mathrm{r})X_\mathrm{r}^\top(X_\mathrm{r}X_\mathrm{r}^\top)^{-1}}_{\hat{A}_\mathrm{r}} X_\mathrm{r}(\sigma X_\mathrm{r})^\top \geq 0$$

Using the identity $X_\mathrm{r}(\sigma X_\mathrm{r})^\top = (X_\mathrm{r}X_\mathrm{r}^\top)\hat{A}_\mathrm{r}^\top$, we obtain a Stein inequality

$$X_\mathrm{r}X_\mathrm{r}^\top \geq \hat{A}_\mathrm{r}(X_\mathrm{r}X_\mathrm{r}^\top)\hat{A}_\mathrm{r}^\top,$$

which proves that $\hat{A}_\mathrm{r}^\top$ is Schur, so that $\hat{A}_\mathrm{r}$ is Schur. □

## Preliminary result: impulse response identification

Corollary (Impulse response case). Consider the model (1), with $(A, B, C)$ minimal and $A$ Schur. Let $u$ be the Kronecker delta function $\delta_t$ and $x_0 = 0$. Denote by $X$ the infinite matrix

$$X = \begin{bmatrix} x_0 & x_1 & \cdots & x_t & \cdots \end{bmatrix}.$$

For a given natural number $r < n$, partition $X$ as follows

$$X =: \begin{bmatrix} X_\mathrm{r} \\ X_\mathrm{trunc} \end{bmatrix}, \qquad \text{where} \qquad \operatorname{row\,dim}(X_\mathrm{r}) = r.$$

The solution

$$\hat{A}_\mathrm{r} = \sigma X_\mathrm{r} X_\mathrm{r}^\top (X_\mathrm{r}X_\mathrm{r}^\top - X_\mathrm{r}U^\top(UU^\top)^{-1}UX_\mathrm{r}^\top)^{-1} \tag{6}$$

of the least squares problem (4) is Schur.

Proof: The formula in the right-hand-side of (6) is obtained by explicitly computing the (1,1) block of the LS solution of (4)

$$\begin{bmatrix} \hat{A}_\mathrm{r} & \hat{B}_\mathrm{r} \\ \hat{C}_\mathrm{r} & \hat{D}_\mathrm{r} \end{bmatrix} = \begin{bmatrix} \sigma X_\mathrm{r} \\ Y_\mathrm{f} \end{bmatrix} \begin{bmatrix} X_\mathrm{r} \\ U_\mathrm{f} \end{bmatrix}^\top \left( \begin{bmatrix} X_\mathrm{r} \\ U_\mathrm{f} \end{bmatrix} \begin{bmatrix} X_\mathrm{r} \\ U_\mathrm{f} \end{bmatrix}^\top \right)^{-1}.$$

For $u = \delta$ and $x_0 = 0$, $X_\mathrm{r}U^\top(UU^\top)^{-1}UX_\mathrm{r}^\top = 0$, so that the estimate

$$\hat{A}_\mathrm{r} = \sigma X_\mathrm{r} X_\mathrm{r}^\top (X_\mathrm{r}X_\mathrm{r}^\top)^{-1}$$

has the same form as in the autonomous case. Under the assumptions of the theorem $u$ is persistently exciting and $x \in \ell_2(\mathbb{Z}, \mathbb{R}^n)$, so that the result of the Theorem applies and $\hat{A}_\mathrm{r}$ is Schur. □