# An algorithm for approximate common divisor computation

Ivan Markovsky and Sabine Van Huffel

K.U.Leuven, ESAT-SISTA





## Outline







## Definition of distance to uncontrollability

Notation:

| | | |
|---|---|---|
| $\mathscr{L}$ | — | set of LTI systems |
| $\mathscr{L}_{\text{ctrb}}/\overline{\mathscr{L}_{\text{ctrb}}}$ | — | set of controllable/uncontrollable LTI systems |
| $\text{dist}(\mathscr{B},\widehat{\mathscr{B}})$ | — | measure for the distance from $\mathscr{B}$ to $\widehat{\mathscr{B}}$ |

Definition: Given $\mathscr{B} \in \mathscr{L}$ and $\text{dist}(\cdot,\cdot)$, define

$$d(\mathscr{B}) := \min_{\widehat{\mathscr{B}}} \text{dist}(\mathscr{B},\widehat{\mathscr{B}}) \quad \text{subject to} \quad \widehat{\mathscr{B}} \in \overline{\mathscr{L}_{\text{ctrb}}}$$

to be the distance of $\mathscr{B}$ to uncontrollability (w.r.t. to $\text{dist}(\cdot,\cdot)$)

Note: $d(\mathscr{B})$ is a representation free notion, however,

in order to compute it one has to use a representation of $\mathscr{B}$





## Input/output and input/state/output representations

Let $\sigma$ be the

- shift operator $(\sigma x)(t) = x(t+1)$ (in discrete-time) or
- derivative operator $\sigma x = dx/dt$ (in continuous-time)

I/O representation: $\forall \mathscr{B} \in \mathscr{L}$, $\exists$ polynomials $P,Q$, $\det(P) \neq 0$, $\text{degree}(P) > \text{degree}(Q)$, and a permutation matrix $\Pi$, such that

$$\mathscr{B} = \mathscr{B}(P,Q,\Pi) := \{\Pi \operatorname{col}(u,y) \mid P(\sigma)y = Q(\sigma)u\}$$

I/S/O representation: $\forall \mathscr{B} \in \mathscr{L}$, $\exists$ matrices $A,B,C,D$, and a permutation matrix $\Pi$, such that

$$\mathscr{B} = \mathscr{B}(A,B,C,D,\Pi) := \{\Pi \operatorname{col}(u,y) \mid \exists x, \quad \sigma x = Ax + Bu, \quad y = Cx + Du\}$$

## About the mappings $\mathscr{B} \mapsto (P,Q)$ and $\mathscr{B} \mapsto (A,B,C,D)$

Passing among representations are classical problems
e.g., $(P,Q) \mapsto (A,B,C,D)$ is a realization problem

The parameters $P,Q$ and $A,B,C,D$ are not unique, so that
$\mathscr{B} \mapsto (P,Q)$ and $\mathscr{B} \mapsto (A,B,C,D)$ are one-to-many mappings

Notable exception, for a SISO system, taking $p$ monic,

$(p,q)$ is in a one-to-one relation with $\mathscr{B}$

Next, we will use this special case for computing $d(\mathscr{B})$, w.r.t.

$$\operatorname{dist}(\mathscr{B},\widehat{\mathscr{B}}) := \sqrt{\|p-\widehat{p}\|_2^2 + \|q-\widehat{q}\|_2^2}$$





## $\widehat{\mathscr{B}} \in \mathscr{L}_{\text{ctrb}} \iff p,q$ coprime

Our problem becomes:

$$\min_{\widehat{p},\widehat{q}} \left\| \begin{bmatrix} p \\ q \end{bmatrix} - \begin{bmatrix} \widehat{p} \\ \widehat{q} \end{bmatrix} \right\|_2 \quad \text{subject to} \quad \mathscr{B}(\widehat{p},\widehat{q}) \in \overline{\mathscr{L}_{\text{ctrb}}}$$

or equivalently

$$\min_{\widehat{p},\widehat{q}} \left\| \begin{bmatrix} p \\ q \end{bmatrix} - \begin{bmatrix} \widehat{p} \\ \widehat{q} \end{bmatrix} \right\|_2 \quad \text{subject to} \quad \begin{array}{c} \widehat{p} \text{ and } \widehat{q} \text{ have} \\ \text{a common factor} \\ \text{of degree} \geq 1 \end{array}$$





## Equivalent problem: structured low-rank approx.

$\widehat{\mathscr{B}} \in \overline{\mathscr{L}_{\text{ctrb}}}$ is equivalent to rank deficiency of the Sylvester matrix

$$S(\widehat{p},\widehat{q}) := \begin{bmatrix} \widehat{p}_0 & & & & \widehat{q}_0 & & & \\ \widehat{p}_1 & \widehat{p}_0 & & & \widehat{q}_1 & \widehat{q}_0 & & \\ \vdots & \widehat{p}_1 & \ddots & & \vdots & \widehat{q}_1 & \ddots & \\ \widehat{p}_n & \vdots & \ddots & \widehat{p}_0 & \widehat{q}_n & \vdots & \ddots & \widehat{q}_0 \\ & \widehat{p}_n & & \widehat{p}_1 & & \widehat{q}_n & & \widehat{q}_1 \\ & & \ddots & \vdots & & & \ddots & \vdots \\ & & & \widehat{p}_n & & & & \widehat{q}_n \end{bmatrix}$$

our problem becomes a Sylvester structured low-rank approx.

$$\min_{\widehat{p},\widehat{q},w} \left\| \begin{bmatrix} p \\ q \end{bmatrix} - \begin{bmatrix} \widehat{p} \\ \widehat{q} \end{bmatrix} \right\|_2 \quad \text{subject to} \quad S(\widehat{p},\widehat{q}) \begin{bmatrix} w \\ 1 \end{bmatrix} = 0 \qquad (**)$$





## Paige's distance to uncontrollability

C. C. Paige defined in

*Properties of numerical algorithms related to computing controllability, IEEE-AC, vol. 26, 1981*

the following measure for distance of $\mathscr{B}$ to uncontrollability

$$d(A,B) := \text{minimize}_{\widehat{A},\widehat{B}} \left\| \begin{bmatrix} A & B \end{bmatrix} - \begin{bmatrix} \widehat{A} & \widehat{B} \end{bmatrix} \right\|_{\text{F}}$$
$$\text{subject to} \quad (\widehat{A},\widehat{B}) \text{ is uncontrollable}$$

many papers on computing $d(A,B)$ (98 citations in WoS)

However, $d(A,B)$ depends on the choice of the state space basis!

$\implies d(A,B)$ not a genuine property of the pair of systems $(\mathscr{B},\widehat{\mathscr{B}})$

## Equivalent optimization problem

**Theorem** The optimization problem

$$\min_{\substack{\widehat{p},\widehat{q},c\\ p_{\text{red}},q_{\text{red}}}} \left\| \begin{bmatrix} p \\ q \end{bmatrix} - \begin{bmatrix} \widehat{p} \\ \widehat{q} \end{bmatrix} \right\|_2 \quad \text{subject to} \quad \begin{aligned} \widehat{p} &= p_{\text{red}} c \\ \widehat{q} &= q_{\text{red}} c \\ \deg(c) &= k \end{aligned}$$

is equivalent to

$$\min_c \operatorname{trace}\left( \begin{bmatrix} p & q \end{bmatrix}^\top \left( I - T(c)\left(T^\top(c)T(c)\right)^{-1} T^\top(c) \right) \begin{bmatrix} p & q \end{bmatrix} \right)$$

where $T(c) \in \mathbb{R}^{(n+1)\times n}$ is a lower triangular banded Toeplitz matrix with first column equal to $\operatorname{col}(c,1,0,\ldots,0)$.





## Comments

- $\widehat{p}$, $\widehat{q}$, $p_{\text{red}}$, $q_{\text{red}}$, and the constraint are eliminated
- nonconvex nonlinear least squares problem
- can be solved numerically using local optimization methods
- cost function evaluations: solve a structured LS problem
- exploiting structure, comput. complexity $O(n)$ per iteration

$\rightsquigarrow$ an algorithm for approximate GCD computation

the algorithm needs an initial approximation for $c(\xi)$





## Suboptimal initial approximations

Computable from unstructured low-rank approx. (SVD) of

1. Sylvester matrix $S(p,q)$
2. Bezout matrix $B(p,q)$
3. Hankel matrix $H(h)$

$$B(p,q) := \begin{bmatrix} p_1 & \cdots & p_n \\ \vdots & \iddots & \\ p_n & & 0 \end{bmatrix} \begin{bmatrix} q_0 & \cdots & q_{n-1} \\ & \ddots & \vdots \\ 0 & & q_{n-1} \end{bmatrix} - \begin{bmatrix} q_1 & \cdots & q_n \\ \vdots & \iddots & \\ q_n & & 0 \end{bmatrix} \begin{bmatrix} p_0 & \cdots & p_{n-1} \\ & \ddots & \vdots \\ 0 & & p_{n-1} \end{bmatrix}$$

$$H(h) := \begin{bmatrix} h_1 & h_2 & \cdots & h_n \\ h_2 & h_3 & \iddots & h_{n+1} \\ \vdots & \iddots & \iddots & \vdots \\ h_n & h_{n+1} & \cdots & h_{2n} \end{bmatrix}, \qquad \frac{q(\xi)}{p(\xi)} = \sum_{t=0}^{\infty} h_t \xi^{-t-1}$$





## Using the Sylvester matrix

$$\begin{aligned} \mathscr{B}(p,q) \in \overline{\mathscr{L}_{\text{ctrb}}} &\iff p,q \text{ have common divisor } c,\ \deg(c) \geq 1 \\ &\iff \exists\, p_{\text{red}}, q_{\text{red}} \text{ such that } p(\xi)q_{\text{red}}(\xi) = q(\xi)p_{\text{red}}(\xi) \\ &\iff \exists\, p_{\text{red}}, q_{\text{red}} \text{ such that } S(p,q)\begin{bmatrix} q_{\text{red}} \\ -p_{\text{red}} \end{bmatrix} = 0 \\ &\iff S(p,q) \text{ is low-rank} \end{aligned}$$

After computing $p_{\text{red}}, q_{\text{red}}$ from the SVD, we solve the LS problem

$$\begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} T(p_{\text{red}}) \\ T(q_{\text{red}}) \end{bmatrix} c \qquad \left( \begin{array}{l} T(\cdot) \text{ is lower triangular} \\ \text{banded Toeplitz matrix} \end{array} \right)$$

and define $\widehat{p}(\xi) = p_{\text{red}}(\xi)c_{\text{ls}}(\xi)$ and $\widehat{q}(\xi) = q_{\text{red}}(\xi)c_{\text{ls}}(\xi)$. Then

$$d\left(\mathscr{B}(p,q)\right) \leq \|\operatorname{col}(p,q) - \operatorname{col}(\widehat{p},\widehat{q})\|_2$$

# Conclusions

- motivation: replace the statement "$\mathscr{B}$ contr./uncontr." with a quantitative one "distance of $\mathscr{B}$ to uncontrollability"
- the definition invariably considered in the literature is not representation invariant
- behavioral measure: $d(\mathscr{B}) := \min_{\widehat{\mathscr{B}} \in \overline{\mathscr{L}_{\text{ctrb}}}} \text{dist}(\mathscr{B}, \widehat{\mathscr{B}})$
- in the SISO case, $d(\mathscr{B})$ can be defined in terms of the normalized I/O representation $p(\sigma)y = q(\sigma)u$
- the computation of $d(\mathscr{B})$ leads to a nonlinear least squares problem, which cost function evaluation is $O(n)$
- SVD upper bounds, based on the Sylvester, Bezout, and Hankel matrices



Motivation: distance to uncontrollability An algorithm for approximate GCD computation

Thank you

SISTA