A new measure for distance to uncontrollability

Ivan Markovsky

University of Southampton

Outline

Linear time-invariant (LTI) systems

Let $\mathscr{B}$ be LTI system of order $\mathtt{n}$ with $\mathtt{m}$ inputs and $\mathtt{p}$ outputs. Define

- in discrete-time: $(\sigma x)(t) := x(t+1)$ — shift operator
- in continuous-time: $\sigma x := \mathrm{d}x/\mathrm{d}t$ — derivative operator

Two common representations of LTI systems are

- input/state/output representation

$$\mathscr{B}_{\mathrm{i/s/o}}(A,B,C,D) := \{\, \mathrm{col}(u,y) \mid \exists x,\ \sigma x = Ax + Bu,\ y = Cx + Du \,\}$$

- input/output representation

$$\mathscr{B}_{\mathrm{i/o}}(P,Q) := \{\, \mathrm{col}(u,y) \mid P(\sigma)y = Q(\sigma)u \,\}$$

where $P \in \mathbb{R}^{\mathtt{p}\times\mathtt{p}}[\xi]$, $\det(P) \neq 0$ and $Q \in \mathbb{R}^{\mathtt{p}\times\mathtt{m}}[\xi]$.

Controllability

Definition $\mathscr{B}$ is controllable if for all $w_1, w_2 \in \mathscr{B}$, $\exists\, w \in \mathscr{B}$, $\tau > 0$, such that $w_1(t) = w(t)$, for all $t < 0$ and $w_2(t) = w(t)$, for all $t \geq \tau$.

Think of w_1 as a given past traj. and w_2 as a desired future traj.

$\mathscr{B}$ controllable $\implies$ any given traj. can be steered to any desired trajectory

important condition for pole-placement, LQ, H_∞, ... control, *e.g.*,

controllability $\iff$ solvability of the state feedback pole-placement problem

Controllability test in terms of I/S/O representation

For numerically checking controllability of $\mathscr{B}$, we need to relate this property to the parameters of $\mathscr{B}$ in a particular representation.

Consider an I/S/O representation $\mathscr{B} = \mathscr{B}_{\mathrm{i/s/o}}(A, B, C, D)$.

A well known result is that

$\mathscr{B}_{\mathrm{i/s/o}}$ is controllable iff $\mathscr{C} := \begin{bmatrix} A & BA & \cdots & BA^{n-1} \end{bmatrix}$ is full rank

$\Longrightarrow$ checking controllability is a rank test problem for a structured matrix, which is a nonlinear transformation of A, B

Controllability test in terms of I/O representation

Consider an I/O representation $\mathscr{B} = \mathscr{B}_{\mathrm{i/o}}(P, Q)$.

A well known result is that

$\mathscr{B}_{\mathrm{i/o}}(P, Q)$ is controllable if and only if P and Q are coprime.

$\Longrightarrow$ checking controllability is a coprimness test problem for a pair of polynomial matrices.

Distance to rank deficiency

In numerical linear algebra, yes/no questions ($\mathscr{B}$ contr./uncontr.) are replaced by quantitative measures (distance of $\mathscr{B}$ to uncontr.)

Checking whether $\mathscr{C}$ is full rank is a yes/no question.

A corresponding quantitative measure is distance of $\mathscr{C}$ to rank deficiency: smallest $\|\Delta\mathscr{C}\|$, such that $\widehat{\mathscr{C}} := \mathscr{C} + \Delta\mathscr{C}$ is rank def.

However for $\|\Delta\mathscr{C}\|$ to be a meaningful measure for distance to uncontr., $\widehat{\mathscr{C}}$ has to be a controllability matrix for some system $\widehat{\mathscr{B}}$.

$\Longrightarrow$ $\Delta\mathscr{C}$ should have the same structured as $\mathscr{C}$.

Unstructured/structured low rank approximation

Consider a set of structured matrices $\mathbb{M}$ and define

$$d_r(A) := \min_{\Delta A \in \mathbb{M}} \|\Delta A\| \quad \text{subject to} \quad A + \Delta A \text{ has rank } r.$$

With $\mathbb{M} = \mathbb{R}^{m \times n}$, $d_r(A)$ is unstructured distance to rank-r matrices.

In special cases, unstructured $d_r(A)$ can be computed from the SVD of A

$$A = U \operatorname{diag}(\sigma_1, \ldots, \sigma_{\min(m,n)}) V^\top$$

- spectral norms: $d_r(A) = \sigma_{r+1}$
- Frobenius norm: $d_r(A) = \sqrt{\sigma_{r+1}^2 + \cdots + \sigma_{\min(m,n)}^2}$.

In general, $d_r(A)$ is difficult to compute.

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

$$d(A,B) := \min_{\widehat{A},\widehat{B}} \left\| \begin{bmatrix} A & B \end{bmatrix} - \begin{bmatrix} \widehat{A} & \widehat{B} \end{bmatrix} \right\|_{\mathrm{F}}$$

$$\text{subject to} \quad \mathscr{C}(\widehat{A},\widehat{B}) \text{ is rank deficient}$$

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

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

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

More general definition

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

where

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

Note: $d(A,B)$ is formally a special case of $d(\mathscr{B})$ with

$$\operatorname{dist}(\mathscr{B},\widehat{\mathscr{B}}) = \left\| \begin{bmatrix} A & B \end{bmatrix} - \begin{bmatrix} \widehat{A} & \widehat{B} \end{bmatrix} \right\|_{\mathrm{F}} \qquad \text{(Paige)}$$

however given $\mathscr{B}$ and $\widehat{\mathscr{B}}$, A,B and $\widehat{A},\widehat{B}$ are not uniquely defined

$\Longrightarrow$ (Paige) is not well defined

Special case: I/O representation of a SISO system

Consider a SISO system $\mathscr{B}$ with an input/output representation

$$p(\sigma)y = q(\sigma)u$$

With p monic, p,q are unique and

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

becomes a property of the pair of systems $(\mathscr{B},\widehat{\mathscr{B}})$.

The problem of computing $d(\mathscr{B}_{\text{i/o}}(p,q))$ becomes

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

Representations of $\mathscr{B}_{\text{i/o}}(\widehat{p},\widehat{q}) \in \overline{\mathscr{L}_{\text{ctrb}}}$

$\mathscr{B}_{\text{i/o}}(\widehat{p},\widehat{q}) \in \overline{\mathscr{L}_{\text{ctrb}}} \iff \widehat{p},\widehat{q}$ have common divisor c, $\deg(c) \geq 1$

$\iff \exists\, u,v$ such that $\widehat{p}(\xi) = u(\xi)c(\xi)$, $\widehat{q}(\xi) = v(\xi)c(\xi)$

$\iff \exists\, u,v$ such that $\widehat{p}(\xi)v(\xi) = \widehat{q}(\xi)u(\xi)$

$\iff \exists\, u,v$ such that $S(\widehat{p},\widehat{q}) \begin{bmatrix} v \\ -u \end{bmatrix} = 0$

$\iff S(\widehat{p},\widehat{q})$ is rank deficient

Equivalent problem to $(*)$

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

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

problem $(*)$ is a Sylvester structured low-rank approximation

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

Another equivalent problem to $(*)$

$$a(\xi)b(\xi) \iff \underbrace{\begin{bmatrix} a_0 & & & \\ a_1 & a_0 & & \\ \vdots & a_1 & \ddots & \\ a_{n_a} & \vdots & \ddots & a_0 \\ & a_{n_a} & & a_1 \\ & & \ddots & \vdots \\ & & & a_{n_a} \end{bmatrix}}_{T(a)} b \iff T(b)a$$

From $\quad \mathscr{B}_{\text{i/o}}(\widehat{p},\widehat{q}) \in \overline{\mathscr{L}_{\text{ctrb}}} \iff \exists\, u,v,$ s.t. $\begin{matrix} \widehat{p}(\xi) = u(\xi)c(\xi) \\ \widehat{q}(\xi) = v(\xi)c(\xi) \end{matrix}$,

we have

$$\min_{\widehat{p},\widehat{q},u,v,c} \left\| \begin{bmatrix} p \\ q \end{bmatrix} - \begin{bmatrix} \widehat{p} \\ \widehat{q} \end{bmatrix} \right\|_2 \quad \text{subject to} \quad \begin{bmatrix} \widehat{p} \\ \widehat{q} \end{bmatrix} = \begin{bmatrix} T(u) \\ T(v) \end{bmatrix} c \qquad (***)$$

Local optimization based algorithm

Theorem $\quad d\big(\mathscr{B}_{\text{i/o}}(p,q)\big)$ is equal to

$$\min_{c\in\mathbb{R}} \operatorname{trace}\left(\begin{bmatrix} p & q \end{bmatrix}^\top \left(I - T\left(\left[\begin{smallmatrix} c \\ 1 \end{smallmatrix}\right]\right) \left(T^\top\left(\left[\begin{smallmatrix} c \\ 1 \end{smallmatrix}\right]\right) T\left(\left[\begin{smallmatrix} c \\ 1 \end{smallmatrix}\right]\right) \right)^{-1} T^\top\left(\left[\begin{smallmatrix} c \\ 1 \end{smallmatrix}\right]\right) \right) \begin{bmatrix} p & q \end{bmatrix} \right).$$

Notes:

- $\widehat{p}$, $\widehat{q}$, u, v and the constraint are eliminated from $(***)$
- nonconvex nonlinear least squares problem
- solved numerically using local optimization methods
- the optimization variable is a scalar
- cost function evaluations: solve a structured LS problem
- exploiting structure, comput. complexity per iteration $O(n)$

Proof

Rewrite the constraint of $(***)$ as follows

$$\begin{bmatrix} \widehat{p} \\ \widehat{q} \end{bmatrix} = \begin{bmatrix} T(u) \\ T(v) \end{bmatrix} c \iff \begin{bmatrix} \widehat{p} & \widehat{q} \end{bmatrix} = T(c) \begin{bmatrix} u & v \end{bmatrix}$$

so that

$$\left\| \begin{bmatrix} p \\ q \end{bmatrix} - \begin{bmatrix} \widehat{p} \\ \widehat{q} \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} p & q \end{bmatrix} - T(c) \begin{bmatrix} u & v \end{bmatrix} \right\|_{\text{F}}$$

$(***)$ becomes an ordinary least-squares problem in u,v

$\rightsquigarrow$ closed form expression in c

Suboptimal initial approximations

can be computed from unstructured low rank approx. (SVD) of

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

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

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

Using the Sylvester matrix

$$\mathscr{B}_{\mathrm{i/o}}(p,q) \in \overline{\mathscr{L}_{\mathrm{ctrb}}} \quad \iff \quad \exists\, u,v \text{ such that } S(p,q)\begin{bmatrix} v \\ -u \end{bmatrix} = 0$$

1. Compute the SVD of $S(p,q) \rightsquigarrow$ approximate u and v.
2. Solve the LS problem $\begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} T(u) \\ T(v) \end{bmatrix} c.$
3. Define

$$\widehat{p}(\xi) = u(\xi)c_{\mathrm{ls}}(\xi) \qquad \text{and} \qquad \widehat{q}(\xi) = v(\xi)c_{\mathrm{ls}}(\xi).$$

Then

$$d\big(\mathscr{B}_{\mathrm{i/o}}(p,q)\big) \leq \|\operatorname{col}(p,q) - \operatorname{col}(\widehat{p},\widehat{q})\|_2.$$

Simulation example

Given $\mathscr{B}_{\mathrm{i/o}}(p,q)$, where

$$p(\xi) = \quad 0.058 + 0.684\xi + 2.745\xi^2 + 4.751\xi^3 + 3.622\xi^4 + 1.000\xi^5$$

$$q(\xi) = -0.134 - 1.408\xi - 5.149\xi^2 - 8.381\xi^3 - 6.092\xi^4 - 1.604\xi^5$$

we compute initial approximation $\mathscr{B}_{\mathrm{i/o}}(\widehat{p},\widehat{q})$ by one of the suboptimal methods and a locally optimal solution $\mathscr{B}_{\mathrm{i/o}}(\widehat{p}^*,\widehat{q}^*)$

Method	$d(\mathscr{B}_{\mathrm{i/o}}(\widehat{p},\widehat{q}))$	$d(\mathscr{B}_{\mathrm{i/o}}(\widehat{p}^*,\widehat{q}^*))$	# iter.	# cost fun. eval.
Sylvester	0.0109	0.0021	4	20
Bezout	0.0041	0.0007	3	12
Hankel	0.0026	0.0007	3	12
BMR	0.0051	0.0007	5	46

(BMR — Balanced model reduction)

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: $d(\mathscr{B}) := \min_{\widehat{\mathscr{B}} \in \overline{\mathscr{L}_{\mathrm{ctrb}}}} \operatorname{dist}(\mathscr{B},\widehat{\mathscr{B}})$
- 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, Hankel matrices, and balanced model reduction