Application of Structured Total Least Squares for System Identification

Ivan Markovsky[1], Jan C. Willems[1], Sabine Van Huffel[1], Bart De Moor[1], and Rik Pintelon[2]

1 — Katholieke Universiteit Leuven, ESAT/SCD   2 — Vrije Universiteit Brussel, Department ELEC

The structured total least squares problem

STLS problem: $\min_{\Delta p, X} \|\Delta p\|^2 \quad \text{s.t.} \quad \mathcal{S}(p - \Delta p) \begin{bmatrix} X \\ -I_d \end{bmatrix} = 0$

$\mathcal{S}(p) = S_0 + \sum_{i=1}^{n_p} S_i p_i$ — structure specification, *e.g.*, $\mathcal{S}$ block-Hankel

Equivalent optimization problem:

$$\min_X \underbrace{r(X)\Gamma^{-1}(X)r(X)}_{f_0(X)}, \quad \text{where} \quad \Gamma(X) := G(X)G^\top(X) \qquad (*)$$

Theorem: Under the assumption that

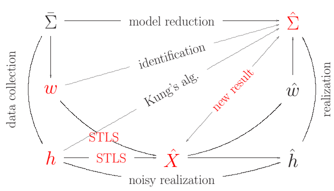
$$C = \begin{bmatrix} C^{(1)} \cdots C^{(q)} \end{bmatrix}, \quad \text{where } C^{(i)} \text{ is } \begin{cases} \text{block-Hankel/Toeplitz,} \\ \text{unstructured, or} \\ \text{exact} \end{cases}$$

Γ is block-Toeplitz and block-banded

the structure of Γ allows computation of $f_0(X)$ and $f_0'(X)$ in $O(m)$ flops
$\Longrightarrow$ fast algorithms for $(*)$

Application in MIMO system identification

Approximate modeling problems



$\hat\Sigma$ — approximating model w — observed general response
$\bar\Sigma$ — "true" model h — observed impulse response

Kernel subproblem: find a block-Hankel rank deficient matrix $\mathcal{H}(\hat w)$ approximating a given matrix $\mathcal{H}(w)$ $\rightsquigarrow$ STLS

$\mathcal{L}_{m,l}$ — the set of all LTI systems with m inputs and lag at most l
(m and l specify the complexity of the model class $\mathcal{L}_{m,l}$)

Identification problem: $\min_{\mathcal{B} \in \mathcal{L}_{m,l}} \left(\min_{\hat w} \|w - \hat w\|_{\ell_2}^2 \quad \text{s.t.} \quad \hat w \in \mathcal{B} \right)$

STLS problem: $\min_X \left(\min_{\hat w} \|w - \hat w\|_{\ell_2}^2 \quad \text{s.t.} \quad \mathcal{S}(\hat w) \left[\begin{smallmatrix} X \\ -I \end{smallmatrix} \right] = 0 \right)$

Theorem: Assume that $\mathcal{B} \in \mathcal{L}_{m,l}$ admits a kernel represent.

$$\mathcal{B} = \ker \Big(\sum_{i=0}^{l} R_i \sigma^i \Big), \quad \text{with } R_l =: \begin{bmatrix} Q_l & P_l \end{bmatrix}, \quad P_l \in \mathbb{R}^{p \times p} \text{ full rank}$$

and let

$$X^\top := -P_l^{-1} \begin{bmatrix} R_0 \cdots R_{l-1} \; Q_l \end{bmatrix}.$$

Then

$$w \in \mathcal{B}|_{[1,T]} \quad \Longleftrightarrow \quad \mathcal{H}_{l+1}^\top(w) \begin{bmatrix} X \\ -I \end{bmatrix} = 0.$$

Results on data sets from DAISY

DAISY (DAta base for Identification of SYstems)

Compared methods:

`stls` — the proposed method based on STLS
`pem` — the prediction error method of the Ident. Toolbox
`subid` — robust combined subspace algorithm

Considered examples.

#	Data set name	parameters T	m	p	l
1	Data of a simulation of the western basin of Lake Erie	57	5	2	1
2	Data of Ethane-ethylene destillation column	90	5	3	1
3	Data from an industrial dryer (Cambridge Control Ltd)	867	3	3	1
4	Wing flutter data	1024	1	1	5
5	Heat flow density through a two layer wall	1680	2	1	2
6	Simulation data of a pH neutralization process	2001	2	1	6
7	Data of a CD-player arm	2048	2	2	1
8	Data from a test setup of an industrial winding process	2500	5	2	2
9	Liquid-saturated steam heat exchanger	4000	1	1	2
10	Data from an industrial evaporator	6305	3	3	1
11	Continuous stirred tank reactor	7500	1	2	1
12	Model of a steam generator at Abbott Power Plant	9600	4	4	1

Misfit $M(w, \hat{\mathcal{B}})$ scaled by $M(w, \hat{\mathcal{B}}_{\texttt{stls}})$.

# Data set name	scaled misfit `gtls`		`pem`		`subid`	
1 Lake Erie	1	1	22.0	9.6	1.5	1.9
2 Destillation	1	1	17.5	14.4	3.1	3.7
3 Industrial dryer	1	1	1.2	1.1	1.2	1.1
4 Wing flutter	1	1.4	2.9	1	1.7	1.5
5 Heat flow	1	1	10.2	10.7	1.9	2.5
6 pH process	1	2.2	2.8	1	1.2	1.4
7 CD-player arm	1	1	1.4	1.4	1.1	1.2
8 Winding process	1	1	2.8	2.6	1.6	1.5
9 Exchanger	1	1	8.1	6.9	1.9	1.6
10 Evaporator	1	1	1.7	1.7	1.6	1.5
11 Tank reactor	1	1	51.5	39.0	2.3	1.6
12 Generator	1	1	3.3	3.1	2.4	2.6

100/100 — identification/validation
85/15 — identification/validation

Execution time scaled by $M(w, \hat{\mathcal{B}}_{\texttt{subid}})$.

# Data set name	scaled exec. time `gtls`		`pem`		`subid`	
1 Lake Erie	2.3	2.4	6.4	9.6	1	1
2 Destillation	5.7	4.4	19.7	15.8	1	1
3 Industrial dryer	22.5	19.8	20.8	19.7	1	1
4 Wing flutter	2.4	2.3	23.4	12.8	1	1
5 Heat flow	4.5	3.5	36.6	31.4	1	1
6 pH process	5.3	4.7	22.7	36.4	1	1
7 CD-player arm	6.2	13.7	38.2	34.5	1	1
8 Winding process	48.1	41.8	64.0	46.7	1	1
9 Exchanger	5.4	5.1	23.5	37.8	1	1
10 Evaporator	93.0	87.0	133	111	1	1
11 Tank reactor	32.3	29.0	124	118	1	1
12 Generator	288	244	205	207	1	1