Connection between dynamical total least squares as a multi-parameter eigenvalue problem and as a Riemannian SVD

42nd Benelux meeting on systems and control

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

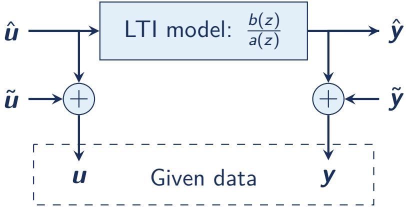
Sem Viroux
(sem.viroux@esat.kuleuven.be)

Bart De Moor, *Fellow IEEE, SIAM, IFAC*
(bart.demoor@esat.kuleuven.be)

March 21, 2023

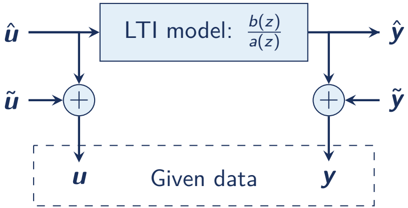
Dynamical total least squares (DTLS)

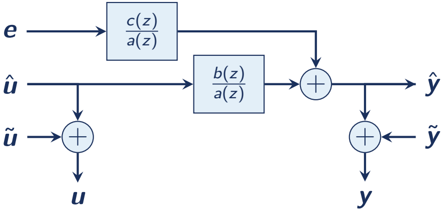
- Given input/output data: $\boldsymbol{u}, \boldsymbol{y} \in \mathbb{R}^N$
- Unknown model parameters: $\boldsymbol{a} \in \mathbb{R}^n, \boldsymbol{b} \in \mathbb{R}^{n+1}$
- Approximating data $\hat{\boldsymbol{u}}$, $\hat{\boldsymbol{y}}$
- Misfits $\tilde{\boldsymbol{u}}$, $\tilde{\boldsymbol{y}}$



Dynamical total least squares (DTLS)

- Given input/output data: $\boldsymbol{u}, \boldsymbol{y} \in \mathbb{R}^N$
- Unknown model parameters: $\boldsymbol{a} \in \mathbb{R}^n, \boldsymbol{b} \in \mathbb{R}^{n+1}$
- Approximating data $\hat{\boldsymbol{u}}$, $\hat{\boldsymbol{y}}$
- Misfits $\tilde{\boldsymbol{u}}$, $\tilde{\boldsymbol{y}}$



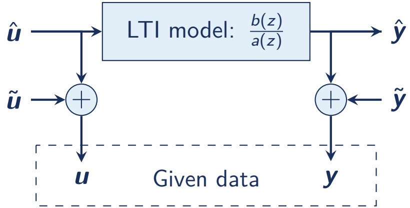


- Special case in misfit-versus-latency framework

Lemmerling et al. 2001 [6]

Dynamical total least squares (DTLS)

- Given input/output data: $\boldsymbol{u}, \boldsymbol{y} \in \mathbb{R}^N$
- Unknown model parameters: $\boldsymbol{a} \in \mathbb{R}^n, \boldsymbol{b} \in \mathbb{R}^{n+1}$
- Approximating data $\hat{\boldsymbol{u}}$, $\hat{\boldsymbol{y}}$
- Misfits $\tilde{\boldsymbol{u}}$, $\tilde{\boldsymbol{y}}$



DTLS

$$\min_{\boldsymbol{a},\boldsymbol{b}} \quad \sigma^2 = \|\tilde{\boldsymbol{u}}\|_2^2 + \|\tilde{\boldsymbol{y}}\|_2^2 = \|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2^2 + \|\boldsymbol{y} - \hat{\boldsymbol{y}}\|_2^2,$$

$$\text{s.t.} \quad \hat{y}_{k+n} + \sum_{i=1}^{n} a_i \hat{y}_{k+n-i} = \sum_{i=0}^{n} b_i \hat{u}_{k+n-i} \qquad \text{for } k = 0, \ldots, N-n-1.$$

DTLS as a multi-parameter eigenvalue problem (MEVP)

DTLS

$$\min_{\boldsymbol{a},\boldsymbol{b}} \quad \sigma^2 = \|\tilde{\boldsymbol{u}}\|_2^2 + \|\tilde{\boldsymbol{y}}\|_2^2 = \|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2^2 + \|\boldsymbol{y} - \hat{\boldsymbol{y}}\|_2^2,$$

$$\text{s.t.} \quad \hat{y}_{k+n} + \sum_{i=1}^{n} a_i \hat{y}_{k+n-i} = \sum_{i=0}^{n} b_i \hat{u}_{k+n-i} \qquad \text{for } k = 0, \ldots, N-n-1.$$

$$\boldsymbol{w} = \begin{bmatrix} y_{N-1} & u_{N-1} & y_{N-2} & u_{N-2} & \ldots & y_0 & u_0 \end{bmatrix}^{\mathsf{T}}$$

and $\tilde{\boldsymbol{w}}$ and $\hat{\boldsymbol{w}}$ similarly.

$$\sigma^2 = \tilde{\boldsymbol{w}}^{\mathsf{T}} \tilde{\boldsymbol{w}}$$

Using a similar derivation as in De Moor 2020 [5]

DTLS as a multi-parameter eigenvalue problem (MEVP)

$$\mathbf{T}_{ab} = \begin{bmatrix} 1 & -b_0 & a_1 & -b_1 & \dots & a_n & -b_n & & & & & & & \\ & & 1 & -b_0 & \dots & a_{n-1} & -b_{n-1} & a_n & -b_n & & & & & \\ & & & \ddots & \ddots & & & & \ddots & \ddots & & & & \\ & & & & 1 & -b_0 & a_1 & -b_1 & \dots & a_n & -b_n & & & \\ & & & & & & 1 & -b_0 & \dots & a_{n-1} & -b_{n-1} & a_n & -b_n \end{bmatrix}$$

such that $\mathbf{T}_{ab}\hat{\boldsymbol{w}} = \mathbf{0}$

Using a similar derivation as in De Moor 2020 [5]

DTLS as a multi-parameter eigenvalue problem (MEVP)

$$\mathbf{T}_{ab} = \begin{bmatrix} 1 & -b_0 & a_1 & -b_1 & \dots & a_n & -b_n & & & & & & & \\ & & 1 & -b_0 & \dots & a_{n-1} & -b_{n-1} & a_n & -b_n & & & & & \\ & & & \ddots & \ddots & & & \ddots & \ddots & & & & & \\ & & & & 1 & -b_0 & a_1 & -b_1 & \dots & a_n & -b_n & & & \\ & & & & & & 1 & -b_0 & \dots & a_{n-1} & -b_{n-1} & a_n & -b_n \end{bmatrix}$$

such that $\mathbf{T}_{ab}\hat{\boldsymbol{w}} = \mathbf{0}$

$$\begin{aligned} \mathbf{T}_{ab}\tilde{\boldsymbol{w}} &= \mathbf{T}_{ab}\left(\boldsymbol{w} - \hat{\boldsymbol{w}}\right) \\ &= \mathbf{T}_{ab}\boldsymbol{w} \end{aligned} \quad \Rightarrow \quad \begin{aligned} \tilde{\boldsymbol{w}} &= {\mathbf{T}_{ab}}^{\dagger}\mathbf{T}_{ab}\boldsymbol{w} \\ &= \mathbf{T}_{ab}^{\top}\left(\mathbf{T}_{ab}\mathbf{T}_{ab}^{\top}\right)^{-1}\mathbf{T}_{ab}\boldsymbol{w} \end{aligned}$$

Using a similar derivation as in De Moor 2020 [5]

DTLS as a multi-parameter eigenvalue problem (MEVP)

$$\mathbf{T}_{ab} = \begin{bmatrix} 1 & -b_0 & a_1 & -b_1 & \dots & a_n & -b_n & & & & & & \\ & & 1 & -b_0 & \dots & a_{n-1} & -b_{n-1} & a_n & -b_n & & & & \\ & & & \ddots & \ddots & & & \ddots & \ddots & & & & \\ & & & & 1 & -b_0 & a_1 & -b_1 & \dots & a_n & -b_n & & \\ & & & & & & 1 & -b_0 & \dots & a_{n-1} & -b_{n-1} & a_n & -b_n \end{bmatrix}$$

such that $\mathbf{T}_{ab}\hat{\boldsymbol{w}} = \mathbf{0}$

$$\begin{aligned} \mathbf{T}_{ab}\tilde{\boldsymbol{w}} &= \mathbf{T}_{ab}\left(\boldsymbol{w} - \hat{\boldsymbol{w}}\right) \\ &= \mathbf{T}_{ab}\boldsymbol{w} \end{aligned} \quad \Rightarrow \quad \begin{aligned} \tilde{\boldsymbol{w}} &= {\mathbf{T}_{ab}}^{\dagger}\mathbf{T}_{ab}\boldsymbol{w} \\ &= \mathbf{T}_{ab}^{\mathsf{T}}\left(\mathbf{T}_{ab}\mathbf{T}_{ab}^{\mathsf{T}}\right)^{-1}\mathbf{T}_{ab}\boldsymbol{w} \end{aligned}$$

$$\sigma^2 = \tilde{\boldsymbol{w}}^{\mathsf{T}}\tilde{\boldsymbol{w}} = \boldsymbol{w}^{\mathsf{T}}\mathbf{T}_{ab}^{\mathsf{T}}\left(\mathbf{T}_{ab}\mathbf{T}_{ab}^{\mathsf{T}}\right)^{-1}\mathbf{T}_{ab}\boldsymbol{w}$$

Using a similar derivation as in De Moor 2020 [5]

DTLS as a multi-parameter eigenvalue problem (MEVP)

$$\sigma^2 = \tilde{\boldsymbol{w}}^\top \tilde{\boldsymbol{w}} = \boldsymbol{w}^\top \mathbf{T}_{ab}^\top \left(\mathbf{T}_{ab}\mathbf{T}_{ab}^\top\right)^{-1} \mathbf{T}_{ab}\boldsymbol{w}$$

Auxilary variables:

$$\mathbf{D}_{ab} = \mathbf{T}_{ab}\mathbf{T}_{ab}^\top$$

$$\boldsymbol{f} = \mathbf{D}_{ab}^{-1}\mathbf{T}_{ab}\boldsymbol{w}$$

$$\Rightarrow$$

$$\begin{bmatrix} \mathbf{D}_{ab} & \mathbf{T}_{ab}\boldsymbol{w} \\ \boldsymbol{w}^\top \mathbf{T}_{ab}^\top & \sigma^2 \end{bmatrix} \begin{bmatrix} \boldsymbol{f} \\ -1 \end{bmatrix} = \mathbf{0}$$

Using a similar derivation as in De Moor 2020 [5]

Multi-parameter eigenvalue problem (MEVP)

$$\begin{bmatrix} \mathbf{D}_{ab} & \mathbf{T}_{ab}\boldsymbol{w} \\ \boldsymbol{w}^\top \mathbf{T}_{ab}^\top & \sigma^2 \end{bmatrix} \begin{bmatrix} \boldsymbol{f} \\ -1 \end{bmatrix} = \mathbf{0}$$

Using a similar derivation as in De Moor 2020 [5]
De Cock and De Moor 2021 [1]

Multi-parameter eigenvalue problem (MEVP)

$$\begin{bmatrix} \mathbf{D}_{ab} & \mathbf{T}_{ab}\boldsymbol{w} \\ \boldsymbol{w}^\top \mathbf{T}_{ab}^\top & \sigma^2 \end{bmatrix} \begin{bmatrix} \boldsymbol{f} \\ -1 \end{bmatrix} = \mathbf{0}$$

$$\frac{\partial}{\partial a_i} \Downarrow \frac{\partial}{\partial b_i}$$

MEVP

(e.g., if $n = 1$)

$$\underbrace{\left(\mathbf{A}_0 + \mathbf{A}_1 a_1 + \mathbf{A}_2 a_1^2 + \mathbf{B}_{00} + \mathbf{B}_{10} b_0 + \mathbf{B}_{01} b_1 + \mathbf{B}_{20} b_0^2 + \mathbf{B}_{11} b_0 b_1 + \mathbf{B}_{02} b_1^2\right)}_{\mathcal{M}_{ab}} \boldsymbol{v} = \mathbf{0}$$

- constant coefficient matrices $\mathbf{A}$ and $\mathbf{B}$
- multivariate matrix polynomial $\mathcal{M}_{\boldsymbol{ab}}$
- $\boldsymbol{v}$ contains $\boldsymbol{f}$ and derivatives thereof
- eigenvalues given by powers of $\boldsymbol{a}$, $\boldsymbol{b}$
- solutions via rank deficiency of $\mathcal{M}_{\boldsymbol{ab}}$

Using a similar derivation as in De Moor 2020 [5]
De Cock and De Moor 2021 [1]

Dynamic TLS (DTLS) & structured TLS (STLS)

DTLS

$$\min_{\boldsymbol{a},\boldsymbol{b}} \quad \sigma^2 = \|\tilde{\boldsymbol{u}}\|_2^2 + \|\tilde{\boldsymbol{y}}\|_2^2 = \|\boldsymbol{u} - \hat{\boldsymbol{u}}\|_2^2 + \|\boldsymbol{y} - \hat{\boldsymbol{y}}\|_2^2,$$

$$\text{s.t.} \quad \hat{y}_{k+n} + \sum_{i=1}^{n} a_i \hat{y}_{k+n-i} = \sum_{i=0}^{n} b_i \hat{u}_{k+n-i} \qquad \text{for } k = 0, \dots, N-n-1.$$

STLS

$$\min_{\boldsymbol{a},\boldsymbol{b}} \quad \sigma^2 = \left\|\tilde{\mathbf{H}}\right\|_F^2 = \left\|\mathbf{H} - \hat{\mathbf{H}}\right\|_F^2,$$

$$\text{s.t.} \quad \hat{\mathbf{H}} \begin{bmatrix} 1 \\ \boldsymbol{a} \\ -\boldsymbol{b} \end{bmatrix} = \mathbf{0}.$$

- $\mathbf{H} = [\mathbf{Y} \mid \mathbf{U}], \quad \hat{\mathbf{H}} = [\hat{\mathbf{Y}} \mid \hat{\mathbf{U}}], \quad \tilde{\mathbf{H}} = [\tilde{\mathbf{Y}} \mid \tilde{\mathbf{U}}]$
- $\mathbf{U}, \hat{\mathbf{U}}, \tilde{\mathbf{U}}$: Hankel matrices constructed from $\boldsymbol{u}, \hat{\boldsymbol{u}}, \tilde{\boldsymbol{u}}$
- $\mathbf{Y}, \hat{\mathbf{Y}}, \tilde{\mathbf{Y}}$: Hankel matrices constructed from $\boldsymbol{y}, \hat{\boldsymbol{y}}, \tilde{\boldsymbol{y}}$
- constraint imposes rank-deficiency of $\hat{\mathbf{H}}$

STLS & the Riemannian SVD (RiSVD)

STLS

$$\min_{\boldsymbol{a},\boldsymbol{b}} \quad \sigma^2 = \left\|\tilde{\mathbf{H}}\right\|_F^2 = \left\|\mathbf{H} - \hat{\mathbf{H}}\right\|_F^2,$$

$$\text{s.t.} \quad \hat{\mathbf{H}} \begin{bmatrix} 1 \\ \boldsymbol{a} \\ -\boldsymbol{b} \end{bmatrix} = \mathbf{0}.$$

$\Leftrightarrow$

RiSVD

$$\mathbf{H}\boldsymbol{v} = \mathbf{D}_{\boldsymbol{v}}\boldsymbol{u}\tau, \quad \boldsymbol{u}^\mathsf{T}\mathbf{D}_{\boldsymbol{v}}\boldsymbol{u} = 1,$$
$$\mathbf{H}^\mathsf{T}\boldsymbol{u} = \mathbf{D}_{\boldsymbol{u}}\boldsymbol{v}\tau, \quad \boldsymbol{v}^\mathsf{T}\mathbf{D}_{\boldsymbol{u}}\boldsymbol{v} = 1.$$

- Riemannian singular triplets $(\boldsymbol{u}, \tau, \boldsymbol{v})$
- RiSVD triplets correspond to critical points of STLS
- $\boldsymbol{v} = \begin{bmatrix}\boldsymbol{a}^\mathsf{T} & -\boldsymbol{b}^\mathsf{T}\end{bmatrix}^\mathsf{T}$
- $\mathbf{D}_{\boldsymbol{u}}$ and $\mathbf{D}_{\boldsymbol{v}}$ positive definite weighting matrices quadratic in the elements of $\boldsymbol{u}$ and $\boldsymbol{v}$
- $\tau^2 = \sigma^2$

Theorem 1 of De Moor 1994 [3]

Solving the RiSVD

- Heuristic solution methods:
  - Inverse iteration algorithm
  - Christiaan's flow (CF)
  - Dependent on initial guess
- Convergence (to global minimizer) not proven
  - "Works in most cases"
  - RiSVD triplets correspond to equilibria
  - CF: triplet with minimal τ^2 is stable

De Moor 1993 [2]
De Moor 1995 [4]

RiSVD

$$\mathbf{H}\boldsymbol{v} = \mathbf{D}_{\boldsymbol{v}}\boldsymbol{u}\tau, \quad \boldsymbol{u}^{\mathsf{T}}\mathbf{D}_{\boldsymbol{v}}\boldsymbol{u} = 1,$$
$$\mathbf{H}^{\mathsf{T}}\boldsymbol{u} = \mathbf{D}_{\boldsymbol{u}}\boldsymbol{v}\tau, \quad \boldsymbol{v}^{\mathsf{T}}\mathbf{D}_{\boldsymbol{u}}\boldsymbol{v} = 1.$$

- Riemannian singular triplets $(\boldsymbol{u}, \tau, \boldsymbol{v})$
- RiSVD triplets correspond to critical points of STLS
- $\boldsymbol{v} = \begin{bmatrix}\boldsymbol{a}^{\mathsf{T}} & -\boldsymbol{b}^{\mathsf{T}}\end{bmatrix}^{\mathsf{T}}$
- $\mathbf{D}_{\boldsymbol{u}}$ and $\mathbf{D}_{\boldsymbol{v}}$ positive definite weighting matrices quadratic in the elements of $\boldsymbol{u}$ and $\boldsymbol{v}$
- $\tau^2 = \sigma^2$

Research questions

- Can a proof be formulated for the heuristic algorithms in the RiSVD problem setting?
- Is there any connection between both formulations?

MEVP

$$\mathcal{M}_{\boldsymbol{ab}}\boldsymbol{v} = \mathbf{0}$$

RiSVD

$$\mathbf{H}\boldsymbol{v} = \mathbf{D}_{\boldsymbol{v}}\boldsymbol{u}\tau, \quad \boldsymbol{u}^{\mathsf{T}}\mathbf{D}_{\boldsymbol{v}}\boldsymbol{u} = 1,$$
$$\mathbf{H}^{\mathsf{T}}\boldsymbol{u} = \mathbf{D}_{\boldsymbol{u}}\boldsymbol{v}\tau, \quad \boldsymbol{v}^{\mathsf{T}}\mathbf{D}_{\boldsymbol{u}}\boldsymbol{v} = 1.$$

References I

Katrien De Cock and Bart De Moor.
Multiparameter eigenvalue problems and shift-invariance.
IFAC-PapersOnLine, 54(9):159–165, 2021.

Bart De Moor.
Structured total least squares and L_2 approximation problems.
Linear Algebra and its Applications, 188-189:163–205, 1993.

Bart De Moor.
Dynamic total linear least squares.
IFAC Proceedings Volumes, 27(8):1121–1126, 1994.

Bart De Moor.
Continuous-time algorithms for the Riemannian SVD.
In *Proceedings of 1995 34th IEEE Conference on Decision and Control*, pages 1084–1090, New Orleans, LA, USA, 12 1995. IEEE.

References II

Bart De Moor.
Least squares optimal realisation of autonomous LTI systems is an eigenvalue problem.
Communications in Information and Systems, 20(2):163–207, 2020.

Philippe Lemmerling, Bart De Moor, and Sabine Van Huffel.
The structured total least-squares approach for non-linearly structured matrices.
Numerical Linear Algebra with Applications, 9(4):321–332, 2002.