Connection between DTLS as an MEVP and as a Riemannian SVD

Sem Viroux and Bart De Moor, *Fellow IEEE, SIAM, IFAC*
Center for Dynamical Systems, Signal Processing, and Data Analytics (STADIUS),
Department of Electrical Engineering (ESAT), KU Leuven, 3001 Leuven, Belgium
sem.viroux@esat.kuleuven.be, bart.demoor@esat.kuleuven.be

1 Introduction

Previous work of our research group has shown how the least-squares optimal realisation of an autonomous linear time-invariant (LTI) system can be obtained by solving a multiparameter eigenvalue problem (MEVP) [3]. The globally optimal solution(s) of the non-convex optimisation problem can be found via the eigentuples of the MEVP. An alternative formulation of the problem, based on a structured total least-squares problem [1], is given by a so-called Riemannian singular value decomposition (SVD), which can be solved using a heuristic method [2]. We compare both approaches, investigate how they relate to each other, and try to extend the latter such that global optimality is guaranteed.

2 Dynamic total least squares (DTLS)

Given a sequence of single-input single-output (SISO) data $\boldsymbol{w} = \begin{bmatrix}\boldsymbol{u}^\mathsf{T} & \boldsymbol{y}^\mathsf{T}\end{bmatrix}^\mathsf{T} \in \mathbb{R}^{2N}$, we wish to identify an LTI model that has $\hat{\boldsymbol{w}} = \begin{bmatrix}\hat{\boldsymbol{u}}^\mathsf{T} & \hat{\boldsymbol{y}}^\mathsf{T}\end{bmatrix}^\mathsf{T} \in \mathbb{R}^{2N}$, an approximation of the observed data, in its behaviour (i.e., $\hat{\boldsymbol{w}}$ exactly satisfies the model equations). The model is chosen so that the 2-norm of the misfit $\tilde{\boldsymbol{w}} = \boldsymbol{w} - \hat{\boldsymbol{w}}$ is minimised.

The dynamic total least squares (DTLS) realisation problem is formulated as follows:

$$\begin{aligned} \min_{\boldsymbol{a},\boldsymbol{b}} \quad & \sigma^2 = \|\tilde{\boldsymbol{u}}\|_2^2 + \|\tilde{\boldsymbol{y}}\|_2^2 = \|\boldsymbol{u}-\hat{\boldsymbol{u}}\|_2^2 + \|\boldsymbol{y}-\hat{\boldsymbol{y}}\|_2^2 \,, \\ \text{s.t.} \quad & \sum_{i=0}^{n} a_i \hat{y}_{k+n-i} = \sum_{i=0}^{n} b_i \hat{u}_{k+n-i} \quad k = 0,\ldots,N-n-1 \,, \\ & a_0 = 1 \,, \end{aligned} \tag{1}$$

where $\boldsymbol{a},\boldsymbol{b} \in \mathbb{R}^{n+1}$ contain the coefficients of the n-th order difference equation governing the dynamics of the LTI model. This model class is a special case of the misfit-versus-latency framework [4]. Figure 1 schematically illustrates the relations given in (1).

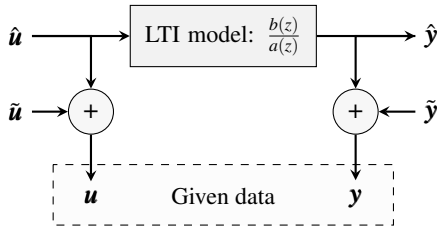

Figure 1: Schematic overview of the identification problem. The given input/output data $(\boldsymbol{u},\boldsymbol{y})$ is decomposed in a regular part $(\hat{\boldsymbol{u}},\hat{\boldsymbol{y}})$ that satisfies the LTI model, and a misfit part $(\tilde{\boldsymbol{u}},\tilde{\boldsymbol{y}})$, the 2-norm of which is to be minimised.

3 MEVP

Notice that the DTLS problem of (1) is of multivariate polynomial nature: the objective function is a sum of squares, and all the constraints are multivariate polynomials in the unknowns ($\hat{\boldsymbol{u}}$, $\hat{\boldsymbol{y}}$, $\boldsymbol{a}$, and $\boldsymbol{b}$). Upon invoking Lagrange multipliers, the first order necessary conditions for optimality can be composed. This results in a system of multivariate polynomial equations, which can be reformulated as an MEVP, the eigentuples of which characterise all the stationary points of the identification problem [3]. After evaluation of the eigentuples in the objective function, the globally optimal solution(s) can be identified.

4 STLS & Riemannian SVD

An alternative, but equivalent, approach to solve the DTLS problem is given by approximating a "double Hankel" matrix $\mathbf{A} = \begin{bmatrix}\mathbf{Y} & \mathbf{U}\end{bmatrix}$ constructed from the observed data by a rank deficient matrix $\mathbf{B}$, which has the same affine matrix structure as $\mathbf{A}$, such that the objective σ^2 in equation (1) is minimal. The minimising matrix $\mathbf{B}$ is related to the singular triplet $(\boldsymbol{u},\tau,\boldsymbol{v})$ with smallest singular value of the so-called Riemannian SVD [2]:

$$\begin{aligned} \mathbf{A}\boldsymbol{v} &= \mathbf{D}_{\boldsymbol{v}}\boldsymbol{u}\tau \,, \quad \boldsymbol{u}^\mathsf{T}\mathbf{D}_{\boldsymbol{v}}\boldsymbol{u} = 1 \,, \\ \mathbf{A}^\mathsf{T}\boldsymbol{u} &= \mathbf{D}_{\boldsymbol{u}}\boldsymbol{v}\tau \,, \quad \boldsymbol{v}^\mathsf{T}\mathbf{D}_{\boldsymbol{u}}\boldsymbol{v} = 1 \,, \end{aligned}$$

where τ is the (unknown) singular value and $\boldsymbol{u}$ and $\boldsymbol{v}$ are the corresponding singular vectors. The latter corresponds to the model parameters: $\boldsymbol{v} = \begin{bmatrix}\boldsymbol{a}^\mathsf{T} & \boldsymbol{b}^\mathsf{T}\end{bmatrix}^\mathsf{T}$. The positive definite matrices $\mathbf{D}_{\boldsymbol{v}}$ and $\mathbf{D}_{\boldsymbol{u}}$ are nonlinear weighting matrices, with elements quadratic in the components of resp. $\boldsymbol{v}$ and $\boldsymbol{u}$.

We derive the connection between both formulations and investigate whether we can we use root-finding algorithms to solve for the smallest singular triplet of the Riemannian SVD, leading to the globally optimal solution(s) of the DTLS problem.

Acknowledgements

This work was supported in part by the KU Leuven: Research Fund (project C16/15/059), and several Leuven Research and Development bilateral industrial projects, in part by Flemish Government agencies: FWO (EOS project G0F6718N (SeLMA), and PhD grant SB/1SA1319N), EWI (Flanders AI Research Program), and in part by the European Commission (EU Research Council under the European Union's Horizon 2020 research and innovation programme (ERC Adv. Grant) under grant 885682).

References

[1] B. De Moor. Structured total least squares and l2 approximation problems. *Linear algebra and its applications*, 188(C):163–205, 1993.

[2] B. De Moor. Dynamic total linear least squares. *IFAC Symposium on System Identification*, 3(10):159–164, 1994.

[3] B. De Moor. Least squares optimal realisation of autonomous LTI systems is an eigenvalue problem. *Comm. in Inf. and Systems*, 20(2):163–207, 2020.

[4] P. Lemmerling and B. De Moor. Misfit versus latency. *Automatica*, 37:2057–2067, 2001.