Least-squares globally optimal misfit modelling for SISO systems

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

1 Introduction

Previous work of our research group showed how the least-squares optimal realisation of an autonomous linear time-invariant (LTI) system can be obtained by solving a multiparameter eigenvalue problem (MEVP), the eigen-tuples of which characterise the globally optimal solution(s) [3]. We extend this methodology to the single-input single-output (SISO) setting, resulting in an approach for least-squares globally optimal misfit modelling for SISO systems. Thereby we address a longstanding open problem of least-squares optimality in misfit modelling.

2 Misfit modelling

We parametrise LTI models by means of a rational transfer function $b(z)/a(z)$ in the time-delay operator z^{-1}, which maps the input to output. This model description is a special case (misfit-only) of the more general misfit versus latency framework [4]. The degree n of the numerator and the denominator of the transfer function determines the complexity of the model. Given a sequence of input-output data $\boldsymbol{w} = (\boldsymbol{u}, \boldsymbol{y}) \in \mathbb{R}^N \times \mathbb{R}^N$ and a specified model complexity n, we want to identify the model for which the 2-norm of the corresponding misfit is minimal. The misfit $\tilde{\boldsymbol{w}} = (\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{y}}) \in \mathbb{R}^N \times \mathbb{R}^N$ is the 'modification' that has to be made to the given data such that the resulting sequence $\hat{\boldsymbol{w}} = (\hat{\boldsymbol{u}}, \hat{\boldsymbol{y}}) \in \mathbb{R}^N \times \mathbb{R}^N$ satisfies the model equations of the identified model exactly. A schematic overview is given in Figure 1. Misfit modelling has close links to Willems' behavioural approach [6, 5] on the one hand and (structured) total least-squares on the other [2, 1].

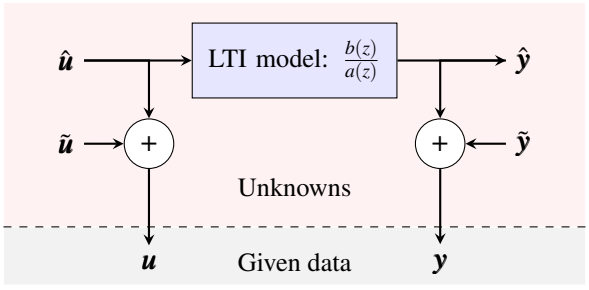

Figure 1: Schematic overview of the identification problem in the misfit setting. The given input-output data $(\boldsymbol{u}, \boldsymbol{y})$ is decomposed in a regular part $(\hat{\boldsymbol{u}}, \hat{\boldsymbol{y}})$ that satisfies the LTI model exactly, and a misfit part $(\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{y}})$, the 2-norm of which should be minimized. In spirit, the misfit idea is close to the dynamic errors-in-variables approach, but we do not presume any a priori statistical assumption.

3 Problem formulation

Several approaches to solve the misfit modelling optimisation problem have been proposed in the literature, both for the parametrisation used here [2, 1], and for a behavioural (isometric) state-space parametrisation [5]. However, these iterative methods remain heuristic, in the sense that convergence to a globally optimal model is not guaranteed.

The misfit modelling problem can be written as:

$$\begin{aligned} \min_{\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{y}}, a_i, b_i} \quad & \sigma^2 = ||\tilde{\boldsymbol{u}}||_2^2 + ||\tilde{\boldsymbol{y}}||_2^2 = \sum_{k=0}^{N-1} (u_k - \hat{u}_k)^2 + \sum_{k=0}^{N-1} (y_k - \hat{y}_k)^2, \\ \text{s.t.} \quad & \hat{y}_{k+n} + a_1 \hat{y}_{k+n-1} + \ldots + a_n \hat{y}_k = b_0 \hat{u}_{k+n} + \\ & \qquad b_1 \hat{u}_{k+n-1} + \ldots + b_n \hat{u}_k \quad \forall k \in \{0, \ldots, N-n-1\}, \end{aligned}$$

with $a_i, b_i \in \mathbb{R}, i = 1, \ldots, n$ the coefficients of $b(z)/a(z)$. Obviously, this optimisation problem is nonlinear, yet, upon closer inspection, its nature is multivariate polynomial: the objective function is a sum of squares, and all unknowns (the $\hat{u}_k$'s, $\hat{y}_k$'s and the transfer function coefficients) also appear as multivariate polynomials in the constraints. Upon invoking Lagrange multipliers, the first order necessary conditions for optimality are also multivariate polynomials, the common roots of which characterise the globally optimal solutions of the identification problem. Rooting such problems is equivalent to solving an MEVP.

We will elaborate on these lines of thought in more detail in our presentation and we will reveal the special properties of the optimal model and associated data sequences, some of which have already been briefly mentioned in [1].

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 grants SB/1SA1319N, SB/1S93918, and SB/1S1319N), 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] Bart De Moor. Dynamic total linear least squares. *IFAC Symposium on System Identification*, 3(10):159 – 164, 1994.

[2] Bart De Moor. Total least squares for affinely structured matrices and the noisy realization problem. *IEEE Transactions on Signal Processing*, 42(11):3104 – 3113, 1994.

[3] 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.

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

[5] Berend Roorda. *Global Total Least Squares - A method for the construction of open approximate models from vector time series*. PhD thesis, Erasmus Universiteit Rotterdam, 1995.

[6] Jan C. Willems. From time series to linear system, Part I: Finite dimensional linear time invariant systems; II: Exact modelling; III: Approximate modelling. *Automatica*, 22/23(5,6,1):561 – 580, 675 – 694, 87 – 115, 1986, 1987.