Globally optimal misfit identification of multidimensional difference equations

Lukas Vanpoucke and Bart De Moor

KU Leuven - Department of Electrical Engineering (ESAT-STADIUS)

Abstract

It has been shown before that the globally optimal least-squares misfit identification of single output, autonomous difference equations with constant coefficients can be formulated as a polynomial optimization problem, due to the polynomial nature of these models. The stationary points of such optimization problems comprise the solution set of a system of multivariate polynomials. Moreover, for this identification problem, the resulting system of equations can be written as a particular class of polynomial systems: a so-called multiparameter eigenvalue problem (MEP). In the case of a finite solution set, such polynomial systems can be solved using the linear algebra-based block Macaulay method. This poster extends this methodology to the misfit identification of autonomous m-dimensional (mD) difference equations. A parametrization is first proposed based on a generalization of the Cayley-Hamilton theorem. Additionally, we outline the MEP formulation for the globally optimal identification problem, for which a numerical example is provided.

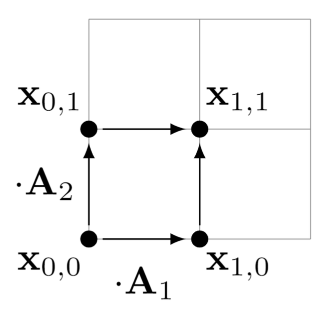
Multidimensional autonomous systems

- **Multidimensional difference equations:** Let $y_{k,l} : \mathbb{Z}^2 \mapsto \mathbb{R}$ be a signal of a single-output, real, autonomous, linear, multishift invariant system. $y_{k,l}$ can then be characterized as the kernel of a polynomial matrix $\mathcal{R}(\sigma_1, \sigma_2)$ in the shift operators σ_1, σ_2 [7, 5].

$$\sigma_1 \cdot y_{k,l} = y_{k+1,l} \qquad \sigma_2 \cdot y_{k,l} = y_{k,l+1} \qquad \mu_i \in R[\sigma_1, \sigma_2]$$

$$0 - \mathcal{R}(\sigma_1, \sigma_2) \cdot y_{k,l} \to \begin{cases} \mu_1(\sigma_1, \sigma_2) \cdot y_{k,l} = 0 \\ \vdots \\ \mu_q(\sigma_1, \sigma_2) \cdot y_{k,l} = 0 \end{cases}$$

- **State-space model:** An n-th order linear, autonomous, multishift invariant system can also be described as a state-space model of the form below [1]:



$$\boldsymbol{x}_{k+1,l} = \boldsymbol{A}_1 \cdot \boldsymbol{x}_{k,l} \qquad \boldsymbol{x}_{k,l+1} = \boldsymbol{A}_2 \cdot \boldsymbol{x}_{k,l} \qquad y_{k,l} = \boldsymbol{C} \cdot \boldsymbol{x}_{k,l}$$

where $\boldsymbol{A}_1, \boldsymbol{A}_2 \in \mathbb{R}^n$ and $\boldsymbol{A}_1\boldsymbol{A}_2 = \boldsymbol{A}_2\boldsymbol{A}_1$.

Since any family of commuting matrices is simultaneously triangularizable [3] and the state space model is invariant under a simultaneous similarity transform, we can assume $\boldsymbol{A}_1, \boldsymbol{A}_2$ to be upper triangular, an eigenvalue-revealing format, without loss of generality.

- **Generalized Cayley-Hamilton** [4]: Let $\chi_{\boldsymbol{A}}$ denote the characteristic polynomial of matrix $\boldsymbol{A}$. Substituting in $c_1 \cdot x + c_2 \cdot y$, the expression can be rewritten as a polynomial in c_1, c_2, with polynomial coefficients $\mu_p(x, y)$.

$$\boldsymbol{A} = c_1\boldsymbol{A}_1 + c_2\boldsymbol{A}_2$$
$$\chi_{\boldsymbol{A}}(c_1 \cdot x + c_2 \cdot y) = \sum_{\alpha+\beta=n} \mu_p(x, y) \cdot c_1^\alpha c_2^\beta$$

Let (λ_1, λ_2) be a pair of eigenvalues, that is, entries at the same diagonal index, of $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, respectively. Clearly, $c_1 \cdot \lambda_1 + c_2 \cdot \lambda_2$ is an eigenvalue of $\boldsymbol{A}$, implying the following relation, irrespective of the values of c_1, c_2.

$$\chi_{\boldsymbol{A}}(c_1 \cdot \lambda_1 + c_2 \cdot \lambda_2) = \sum_{\alpha+\beta=n} \mu_p(\lambda_1, \lambda_2) \cdot c_1^\alpha c_2^\beta = 0$$

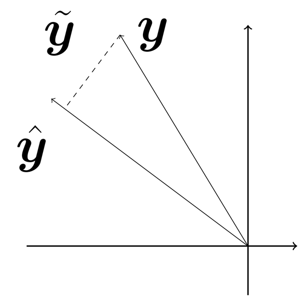
As this polynomial in c_1, c_2 is identically zero $\forall c_1, c_2 \in \mathbb{C}$, it holds that $\mu_p(\lambda_1, \lambda_2) = 0,\ \forall p \in \{1, ..., \binom{2+n-1}{2-1}\}$. Since the solutions of these polynomials describe the system modes, they are valid difference equations, under light assumptions.

Misfit identification

- Using the parameterization above, we apply the least squares misfit identification framework to identify models from the given data [2]:

1. Split the given output sequence $\mathbf{y}$ into an *exact* data sequence $\hat{\boldsymbol{y}}$ and a *misfit* data sequence $\tilde{\boldsymbol{y}}$:
$$\boldsymbol{y}_{k,l} = \hat{\boldsymbol{y}}_{k,l} + \tilde{\boldsymbol{y}}_{k,l}.$$
2. Constrain the *exact* data sequence to adhere to the predefined model parameterization.
3. Optimize to find the model parameters describing the exact data sequence closest to the given data, in a least squares sense.

$$\min_{\hat{\mathbf{y}}} ||\tilde{\boldsymbol{y}}||_2^2$$
$$\text{s.t.} \begin{cases} \boldsymbol{y} = \tilde{\boldsymbol{y}} + \hat{\boldsymbol{y}} \\ \mu_p(\sigma_1, \sigma_2) \cdot \hat{y}_{k,l} = 0,\ \forall k, l, p \end{cases}$$



The polynomial nature of both the constraints and the objective function leads to a **polynomial optimization problem**, the stationary points of which can be formulated as a system of multivariate polynomials.

Multiparameter eigenvalue problem

- **Multiparameter eigenvalue problem:** The second set of constraints can be expressed in terms of a Macaulay Matrix, illustrated below for the simple case of $\mu(\sigma_1, \sigma_2) = \lambda - \sigma_1$.

$$\boldsymbol{M}(\boldsymbol{\lambda}) \cdot \hat{\boldsymbol{y}} = \begin{array}{c} \\ \mu \\ \sigma_1 \cdot \mu \\ \sigma_2 \cdot \mu \end{array} \begin{bmatrix} (0,0) & (1,0) & (0,1) & (2,0) & (1,1) & (0,2) \\ \lambda & -1 & & & & \\ & \lambda & & -1 & & \\ & & \lambda & & -1 & \end{bmatrix} \begin{bmatrix} \hat{y}_{0,0} \\ \hat{y}_{1,0} \\ \hat{y}_{0,1} \\ \hat{y}_{2,0} \\ \hat{y}_{1,1} \\ \hat{y}_{0,2} \end{bmatrix} = \mathbf{0}$$

Let the superscript $^{\lambda_i}$ denote the partial derivative operation w.r.t. λ_i. Using auxiliary variables $\boldsymbol{f}$, the stationary points are then the solutions of the following system of polynomials.

$$\begin{bmatrix} \boldsymbol{M}^{\lambda_i}\boldsymbol{M}^\intercal + \boldsymbol{M}\boldsymbol{M}^{\lambda_i\intercal} & \boldsymbol{M}\boldsymbol{M}^\intercal & \boldsymbol{M}^{\lambda_i}\boldsymbol{y} \\ \boldsymbol{y}^\intercal\boldsymbol{M}^{\lambda_i\intercal} & \boldsymbol{y}^\intercal\boldsymbol{M}^\intercal & 0 \\ \boldsymbol{M}\boldsymbol{M}^\intercal & 0 & \boldsymbol{M}\boldsymbol{y} \end{bmatrix} \begin{bmatrix} \boldsymbol{f} \\ \boldsymbol{f}^{\lambda_i} \\ -1 \end{bmatrix} = \mathbf{0} \quad \forall i \in \{1, ..., 2n\}$$

This system is an MEP, as illustrated below for a first-order model, where the polynomial matrix is split up in terms of the monomials involved.

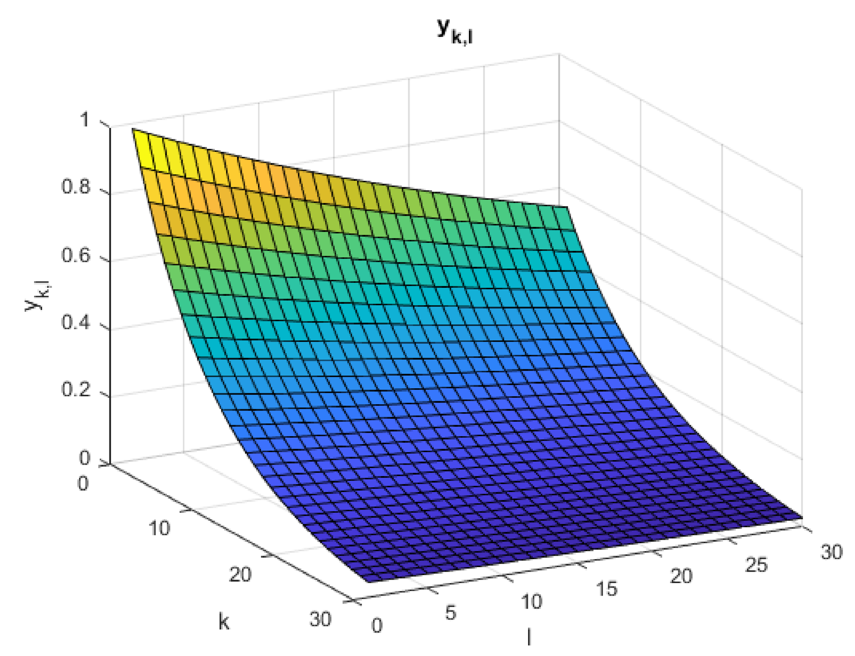
$$\left(\sum_{\alpha+\beta\leq 2} \boldsymbol{M}_{\alpha,\beta} \cdot \lambda_1^\alpha \lambda_2^\beta \right) \cdot \boldsymbol{v} = \mathbf{0} \quad \text{with} \quad ||\boldsymbol{v}|| = 1$$

Numerical example

We start from three data points from the signal $y_{k,l} = 0.98^k \cdot 0.9^l$ and fit a first-order model. Solving the MEP as a system of polynomials using a symbolic solver from this noiseless data yields the correct modes as global optimum.

$$\lambda_1 = 0.98 \qquad \lambda_2 = 0.9$$

The system of equations has an affine, positive dimensional solution set of maximizing stationary points, such that the MEP-specific block-Macaulay method [6] cannot solve it.



Applications and further research

Further research:

- Different solvers, e.g. homotopy continuation.
- Reformulate the problem to eliminate the infinite number of affine solutions.

Applications:

- Benchmarking existing heuristic methods, as it is a computationally expensive approach.
- Studying the properties of the globally optimal solutions can be studied, potentially leading to faster algorithms.

References

[1] E. Fornasini, P. Rocha, and S. Zampieri. "State Space Realization of 2-D Finite-Dimensional Behaviours". In: *SIAM Journal on Control and Optimization* 31.6 (1993), pp. 1502–1517.

[2] Philippe Lemmerling and Bart De Moor. "Misfit versus latency". In: *Automatica* 37.12 (2001), pp. 2057–2067.

[3] N. H. McCoy. "On the characteristic roots of matric polynomials". In: *Bulletin of the American Mathematical Society* 42.8 (1936), pp. 592–600.

[4] "On the Generalized Cayley-Hamilton Theorem". In: *IEEE Transactions on Automatic Control* 31 (2 1986), pp. 156–157.

[5] P Rocha and J C Willems. "State for 2-D systems". In: *Linear Algebra and Its Applications* 122-124 (1989), pp. 1003–1038.

[6] Christof Vermeersch and Bart De Moor. "Two complementary block Macaulay matrix algorithms to solve multiparameter eigenvalue problems". In: *Linear Algebra and its Applications* 654 (2022), pp. 177–209.

[7] J.C. Willems. "From time series to linear system - Part I. Finite dimensional linear time invariant systems". In: *Automatica* 22 (5 1986), pp. 561–580.

This work was supported in part by the KU Leuven: Research Fund (C16/15/059, C3/19/053, C24/18/022, C3/20/117, C3I-21-00316, iBOF/23/064), Industrial Research Fund (13-0260, IOFm/16/004, IOFm/20/002), and LRD bilateral industrial projects; in part by Flemish Government agencies: FWO (S005319, I013218N, T001919N), EWI, and VLAIO (HBC.2019.2204, HBC.2021.0076); and in part by the European Commission (ERC Adv. Grant under grant 885682).