Misfit identification of autonomous multidimensional systems

Lukas Vanpoucke, Bart De Moor

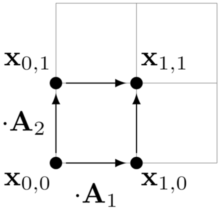
lukas.vanpoucke@esat.kuleuven.be, bart.demoor@esat.kuleuven.be

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Table of Contents

Autonomous multidimensional systems

- We aim to identify multidimensional (mD) systems of the following form [1]:

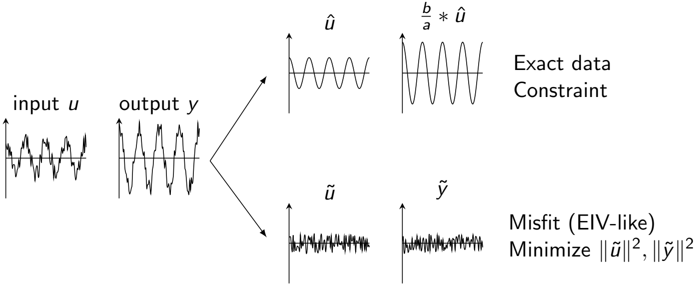


$$\begin{aligned} \boldsymbol{x}_{k+1,l} &= \boldsymbol{A}_1 \cdot \boldsymbol{x}_{k,l} \\ \boldsymbol{x}_{k,l+1} &= \boldsymbol{A}_2 \cdot \boldsymbol{x}_{k,l} \\ y_{k,l} &= \boldsymbol{C} \cdot \boldsymbol{x}_{k,l} \end{aligned}$$

where $\boldsymbol{A}_1 \boldsymbol{A}_2 = \boldsymbol{A}_2 \boldsymbol{A}_1$.

- Any finite 2D autonomous system can be realized in the above state space form [4].

Misfit identification



$$\min_{a,b,\hat{u}} \alpha \|\tilde{y}\|_2^2 + \beta \|\tilde{u}\|_2^2$$

$$\text{s.t.} \begin{cases} y_k = \tilde{y}_k + \hat{y}_k \\ u_k = \tilde{u}_k + \hat{u}_k \\ a(\sigma)\hat{y}_k = b(\sigma)\hat{u}_k \end{cases}$$

- *Exact data* follows the predefined model class exactly.
- First order optimality conditions (FONC) $\rightarrow$ system of multivariate polynomials.

Multiparameter eigenvalue problems

- The polynomial systems obtained have a special structure, e.g [2]:

$$(\boldsymbol{A}_0 + \nu \cdot \boldsymbol{A}_1 + \mu \cdot \boldsymbol{A}_2)\boldsymbol{x} = \boldsymbol{0}$$

- Solve using the block-Macaulay method [6]:
  1. Determine null space of the Macaulay matrix.
  2. Use shift invariance to formulate eigenvalue problem.

$$\begin{array}{c} \\ \cdot 1 \\ \cdot \nu \\ \cdot \mu \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & \nu & \mu & \nu^2 & \nu\mu & \mu^2 \end{array} \\ \begin{bmatrix} \boldsymbol{A}_0 & \boldsymbol{A}_1 & \boldsymbol{A}_2 & & & \\ & \boldsymbol{A}_0 & & \boldsymbol{A}_1 & \boldsymbol{A}_2 & \\ & & \boldsymbol{A}_0 & & \boldsymbol{A}_1 & \boldsymbol{A}_2 \end{bmatrix} \end{array} \begin{bmatrix} 1 \\ \nu \\ \mu \\ \nu^2 \\ \nu\mu \\ \mu^2 \end{bmatrix} = \boldsymbol{0}$$

Table of Contents

Assumption

- Assume $\boldsymbol{A}_1, \boldsymbol{A}_2$ to be simultaneously diagonizable $\rightarrow$ assume diagonal $\boldsymbol{A}_1, \boldsymbol{A}_2$.
  - Generic case of commuting matrices.

$$\boldsymbol{A}_1 = \boldsymbol{V}\boldsymbol{D}_1\boldsymbol{V}^{-1}$$
$$\boldsymbol{A}_2 = \boldsymbol{V}\boldsymbol{D}_2\boldsymbol{V}^{-1}$$

- Allows the following parameterization of the output in terms of the eigenvalues $\lambda_j^{(i)}$:

$$y_{k,l} = \sum_{j=1}^{n} c_j \xi_j \left(\lambda_j^{(1)}\right)^k \left(\lambda_j^{(2)}\right)^l .$$

Finding the global optimum

- Define the following matrices of Vandermonde vectors and their derivatives.

$$\mathbf{\Lambda} = \begin{bmatrix} 1 & \dots & 1 \\ \lambda_1^{(1)} & \dots & \lambda_n^{(1)} \\ \lambda_1^{(2)} & \dots & \lambda_n^{(2)} \\ \left(\lambda_1^{(1)}\right)^2 & \dots & \left(\lambda_n^{(1)}\right)^2 \\ \lambda_1^{(1)}\lambda_1^{(2)} & \dots & \lambda_n^{(1)}\lambda_n^{(2)} \\ \vdots & \vdots & \vdots \end{bmatrix}, \quad \mathbf{\Lambda}^{(\lambda)} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 1 & 0 \\ 0 & 1 & \dots & 0 & 1 \\ 2\lambda_1^{(1)} & 0 & \dots & 2\lambda_n^{(1)} & 0 \\ \lambda_1^{(2)} & \lambda_1^{(1)} & \dots & \lambda_n^{(2)} & \lambda_n^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

- The FONC of the misfit modeling framework then yields the following MEVP:
  - Compute all solutions to the FONC and select **global optimum**.

$$\begin{bmatrix} \mathbf{\Lambda}^\mathsf{T}\mathbf{\Lambda} & \mathbf{\Lambda}^\mathsf{T}\boldsymbol{y} \\ {\mathbf{\Lambda}^{(\lambda)}}^\mathsf{T}\mathbf{\Lambda} & {\mathbf{\Lambda}^{(\lambda)}}^\mathsf{T}\boldsymbol{y} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ -1 \end{bmatrix} = \mathbf{0}$$

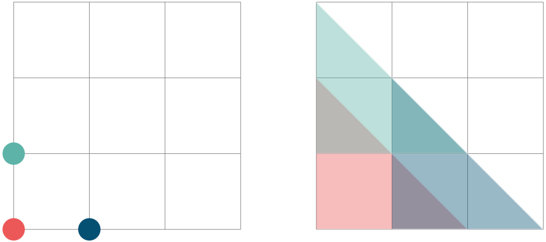
Characterization of the optimum

- The following orthogonality property holds for the misfit sequence:

$$\tilde{\boldsymbol{y}} \perp \text{Range}(\begin{bmatrix} \boldsymbol{\Lambda} & \boldsymbol{\Lambda}^{(\lambda)} \end{bmatrix}).$$

- Indicates **structure** in the optimal misfit sequence.
- $\exists \boldsymbol{M}$, with $\boldsymbol{M}$ a Macaulay matrix s.t. $\boldsymbol{M} \begin{bmatrix} \boldsymbol{\Lambda} & \boldsymbol{\Lambda}^{(\lambda)} \end{bmatrix} = \mathbf{0}$ [2].
  - Interpretation: $\tilde{\boldsymbol{y}}$ results from an mD FIR filter bank.

$$\tilde{\boldsymbol{y}} = \boldsymbol{M}^\mathsf{T} \boldsymbol{g} \quad \longrightarrow$$



$\boldsymbol{g}$ $\qquad$ $\boldsymbol{M}^\mathsf{T}\boldsymbol{g}$

Characterization of the optimum

- Reminiscent of a finite dimensional version of Walsh's theorem [5, 3].
- Walsh's theorem (1D, infinite dimensional signals):
  - Optimal $\tilde{\boldsymbol{y}}$ results from double FIR filtering.
  - FIR filter kills modes corresponding to its zeros.
  - $\rightarrow$ Characterization of FONC in terms of interpolatory conditions in the z-domain.
- Open problem: exact structure of $\boldsymbol{M}$.
  - Need a parametrization in terms of the difference equations.

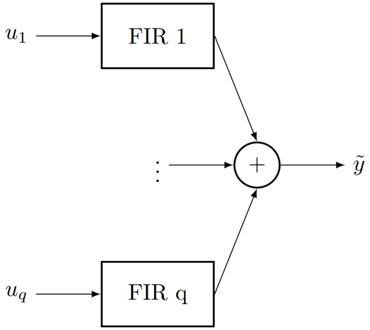


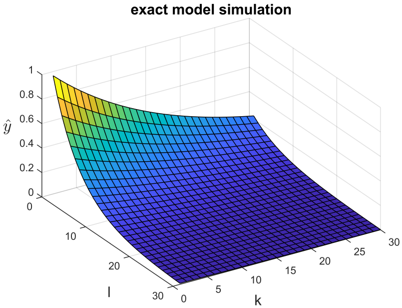
Table of Contents

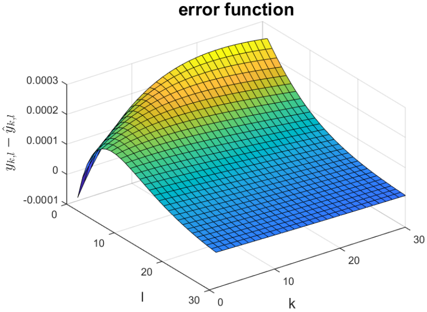
Example: misfit modeling

- Proof of concept example:

$$A_1 = 0,95 \quad A_2 = 0.85 \quad C = 1 \quad x_{0,0} = 1$$

- Misfits sequence constructed using the orthogonality properties
  - Signal to noise ratio: $\|\hat{y}\|/\|\tilde{y}\| = 1.9$
- The constructed system is recovered up to numerical errors





PDE example

- Data: second order solution to the heat equation.
- We identify a first order model.
- Time behaviour is captured nicely, second order spatial behaviour less so.

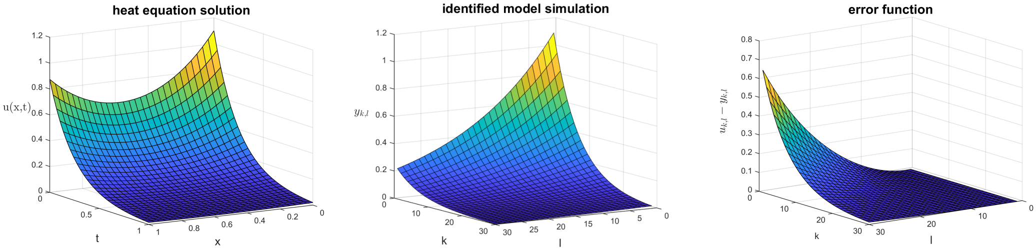


Table of Contents

Conclusion and applications

- Application: benchmarking heuristic methods.
- Only a few data points can be used.
- Parameterization of the difference equations is better suited.
  - Degree of polynomial system is independent on the number of data points.
  - Identification of systems with multiplicities in their modes.
- Characterized the optimal misfit sequence from a system theoretic point of view.

Questions?

Derivations

- Objective function:

$$\min_{\boldsymbol{\lambda}_i, \boldsymbol{x}_{0,0}} \quad \sigma^2 = \|\tilde{\boldsymbol{y}}\|_2^2 = \|\boldsymbol{y} - \boldsymbol{\Lambda}\boldsymbol{x}_{0,0}\|_2^2$$

$$= \boldsymbol{y}^*\boldsymbol{y} + \boldsymbol{x}_{0,0}^*\boldsymbol{\Lambda}^*\boldsymbol{\Lambda}\boldsymbol{x}_{0,0} - \boldsymbol{y}^*\boldsymbol{\Lambda}\boldsymbol{x}_{0,0} - \boldsymbol{x}_{0,0}^*\boldsymbol{\Lambda}^*\boldsymbol{y}$$

- FONC:

$$0 = \frac{\delta\sigma^2}{\delta\lambda_{i,k}} = \boldsymbol{x}_{0,0}^*\boldsymbol{\Lambda}^*\frac{\delta\boldsymbol{\Lambda}}{\delta\lambda_{i,k}}\boldsymbol{x}_{0,0} - \boldsymbol{y}^*\frac{\delta\boldsymbol{\Lambda}}{\delta\lambda_{i,k}}\boldsymbol{x}_{0,0} \qquad i = 1 \dots n, \quad k = 1 \dots m$$

$$0 = \frac{\delta\sigma^2}{\delta\xi_i} = \boldsymbol{x}_{0,0}^*\boldsymbol{\Lambda}^*\boldsymbol{\Lambda}\frac{d\boldsymbol{x}_{0,0}}{d\xi_i} - \boldsymbol{y}^*\boldsymbol{\Lambda}\frac{d\boldsymbol{x}_{0,0}}{d\xi_i} \qquad i = 1 \dots n$$

- MEVP formulation: $\boldsymbol{x}_{0,0} \neq \boldsymbol{0}$ (maximizing solution) $\rightarrow$ "divide" by $\boldsymbol{x}_{0,0}$ in first set of equations.

Derivations

- Orthogonality of misfits (similarly for $\boldsymbol{\Lambda}$):

$$\frac{\delta \boldsymbol{\Lambda}}{\delta \lambda_{i,k}} = \begin{bmatrix} \mathbf{0} & \frac{\delta \boldsymbol{\lambda}_i^{\boldsymbol{\alpha}_0}}{\delta \lambda_{i,k}} & \mathbf{0} \\ \mathbf{0} & \frac{\delta \boldsymbol{\lambda}_i^{\boldsymbol{\alpha}_1}}{\delta \lambda_{i,k}} & \mathbf{0} \\ \vdots & \underbrace{\vdots}_{row\ i} & \vdots \end{bmatrix} = [0 \ \boldsymbol{v}_{i,k} \ 0]$$

- Rewrite first set of equations, using $\hat{\boldsymbol{y}} = \boldsymbol{\Lambda} \boldsymbol{x}_{0,0}$ and $\boldsymbol{y} = \hat{\boldsymbol{y}} + \tilde{\boldsymbol{y}}$:

$$0 = -(\boldsymbol{y}^* - \hat{\boldsymbol{y}}^*) \boldsymbol{v}_{i,k} \xi_i \quad i = 1, ..., n \ \ k = 1, ..., m$$

In matrix form, for all k, i:

$$\tilde{\boldsymbol{y}}^* [\boldsymbol{v}_{1,1} \ \boldsymbol{v}_{1,2} \ ... \ \boldsymbol{v}_{2,1}, ...] = \tilde{\boldsymbol{y}}^* \boldsymbol{\Lambda}^{(\lambda)}$$

Bibliography I

[1] Joseph A Ball and Victor Vinnikov. "Overdetermined multidimensional systems: state space and frequency domain methods". In: *Mathematical systems theory in biology, communications, computation, and finance*. 2003, pp. 63–119.

[2] Katrien De Cock and Bart De Moor. "Multiparameter Eigenvalue Problems and Shift-invariance". In: *IFAC-PapersOnLine, IFAC-PapersOnLine, Part of Special Issue: 24th International Symposium on Mathematical Theory of Networks and Systems (MTNS2020)* 54.9 (2021), pp. 159–165.

[3] Bart De Moor. "Least squares optimal realisation of autonomous LTI systems is an eigenvalue problem". In: *Communications in Information and Systems* 20.2 (2020), pp. 163–207.

Bibliography II

[4] 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.

[5] Phillip Regalia. "Adaptive IIR filtering in Signal Processing and Control". In: CRC Press, 1994, p. 129.

[6] Christof Vermeersch and Bart De Moor. "Two complementary block Macaulay matrix algorithms to solve multiparameter eigenvalue problems". In: (2022). KU Leuven, Leuven, Belgium, Tech. Rep.