Data-driven simulation

using the nuclear norm heuristic

Philippe Dreesen[1] Ivan Markovsky[1] Konstantin Usevich[2]

[1]Vrije Universiteit Brussel (VUB) Dept. ELEC
[2]Université de Lorraine, CNRS, CRAN

Contents

LTI systems and Hankel matrices

The block-Hankel matrix encodes LTI

Linear time-invariant system

- input $\boldsymbol{u}_k \in \mathbb{R}^m$
- output $\boldsymbol{y}_k \in \mathbb{R}^p$
- lag ℓ (minimal)

$$\begin{aligned} a_0 \boldsymbol{y}_k + a_1 \boldsymbol{y}_{k+1} + \cdots + a_\ell \boldsymbol{y}_{k+\ell} \\ = b_0 \boldsymbol{u}_k + \cdots + b_\ell \boldsymbol{u}_{k+\ell}, \text{for } k \leq 1 \end{aligned}$$

The Hankel matrix $\boldsymbol{H}_L$ of a sequence $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)$

$$H_L(x) = \begin{bmatrix} x_1 & x_2 & \cdots & x_{N-L+1} \\ x_2 & x_3 & \cdots & x_{N-L+2} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \cdots & x_N \end{bmatrix}$$

The block-Hankel matrix encodes LTI

Linearity and time-invariance lead to (block-)Hankel structure

$$H_L(u,y) = \left[\begin{array}{cccc} u_1 & u_2 & \cdots & u_{N-L+1} \\ u_2 & u_3 & \cdots & u_{N-L+2} \\ \vdots & \vdots & \ddots & \vdots \\ u_L & u_{L+1} & \cdots & y_N \\ \hline y_1 & y_2 & \cdots & y_{N-L+1} \\ y_2 & y_3 & \cdots & y_{N-L+2} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \cdots & x_N \end{array}\right] = \left[\begin{array}{c} H_L(u) \\ \hline H_L(y) \end{array}\right]$$

For i/o data from an LTI system with lag ℓ

$$\left[\begin{array}{ccc|ccc} b_0 & \cdots & b_\ell & -a_0 & \cdots & -a_\ell \end{array}\right] \left[\begin{array}{c} H_{\ell+1}(u) \\ \hline H_{\ell+1}(y) \end{array}\right] = 0$$

and

$$\operatorname{rank} H_L(u,y) = mL + p\ell$$

Low-rank Mosaic Hankel matrix completion

Given is an i/o dataset $(\boldsymbol{u}^d, \boldsymbol{y}^d)$ of a system

Find the output $\boldsymbol{y}^s$ corresponding to a simulation input $\boldsymbol{u}^s$

Data-driven simulation as low-rank completion problem

Given i/o data $(\boldsymbol{u}^d, \boldsymbol{y}^d)$ and inputs $\boldsymbol{u}^s$

Find $\boldsymbol{y}^s$ such that

$$\operatorname{rank}\left[\begin{array}{c|c} H(\boldsymbol{u}^d) & H(\boldsymbol{u}^s) \\ \hline H(\boldsymbol{y}^d) & H(\boldsymbol{y}^s) \end{array}\right] = \operatorname{rank}\left[\begin{array}{c} H(\boldsymbol{u}^d) \\ \hline H(\boldsymbol{y}^d) \end{array}\right]$$

Notice that similar data-driven variations are possible, such as output tracking and Kalman filtering with missing data [7]

Convex relaxations

Nuclear norm for matrix completion

Matrix rank minimization

$$\begin{aligned} \underset{X}{\text{minimize}} \quad & \text{rank } X \\ \text{subject to} \quad & X_{ij} = D_{ij}, \text{ for } (i,j) \in \Omega \end{aligned}$$

This is an NP-hard problem [1, 6]...

Nuclear norm minimization

$$\begin{aligned} \underset{X}{\text{minimize}} \quad & \|X\|_\star \\ \text{subject to} \quad & X_{ij} = D_{ij}, \text{ for } (i,j) \in \Omega \end{aligned}$$

with the *nuclear norm* $\|X\|_\star = \sum \sigma_i(X)$

This is a convex problem, and perfect recovery is possible [1] (assuming random known/missing element locations)

Nuclear norm for matrix completion

Nuclear norm minimization

$$\begin{aligned} \underset{X}{\text{minimize}} \quad & \|X\|_\star \\ \text{subject to} \quad & X_{ij} = D_{ij}, \text{ for } (i,j) \in \Omega \end{aligned}$$

with the *nuclear norm* $\|X\|_\star = \sum \sigma_i(X)$

- Introduced around 2002 by Fazel and Boyd [4, 3, 8]
- Nuclear norm approximation used for system identification [5]
- Related to compressed sensing:

  sparsity in matrix singular values spectrum

Mosaic-Hankel matrix completion

Nuclear norm Mosaic-Hankel completion

Data-driven simulation as nuclear norm problem

Given i/o data (u^d, y^d) and simulation input u^s

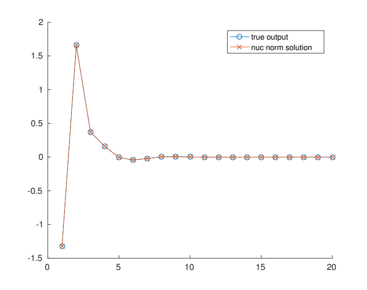
Find y^s from

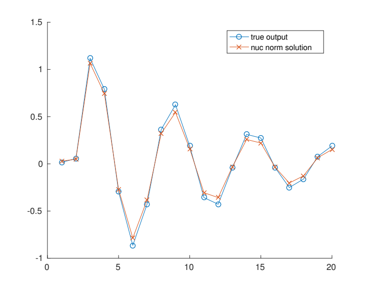
$$\underset{y^s}{\text{minimize}} \left\| \left[\begin{array}{c|c} H(u^d) & H(u^s) \\ \hline H(y^d) & H(y^s) \end{array} \right] \right\|_\star$$

Data-driven impulse response simulation works well

Given a given data sequence $(\boldsymbol{u}^d, \boldsymbol{y}^d)$ of length $N_d = 80$ from a random system (`drss`) of order $n = 4$

Find the response of the system $\boldsymbol{y}^s$ to an impulse input $\boldsymbol{u}^s$ $(N_s = 20)$

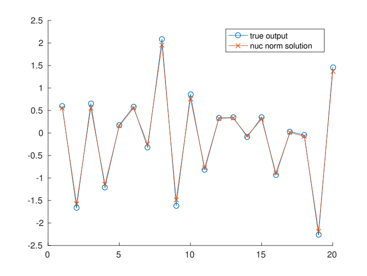


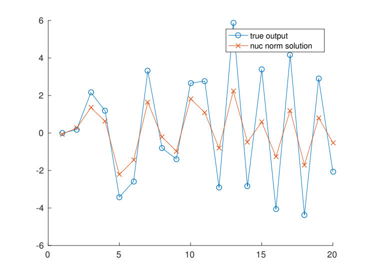


Data-driven random input simulation often fails

Given a given data sequence $(\mathbf{u}^d, \mathbf{y}^d)$ of length $N_d = 80$ from a random system (`drss`) of order $n = 4$

Find the response of the system $\mathbf{y}^s$ to a random input $\mathbf{u}^s$ $(N_s = 20)$





So far...

- Data-driven signal processing bypasses modeling step
- Impulse simulation works but random simulation often fails
- Minimal nuclear norm does not always impose low rank

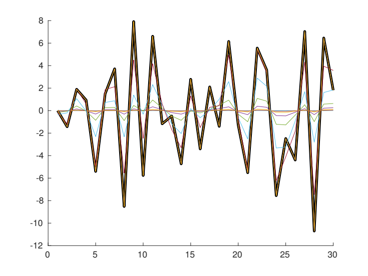
Scaling the available data changes the recovery error

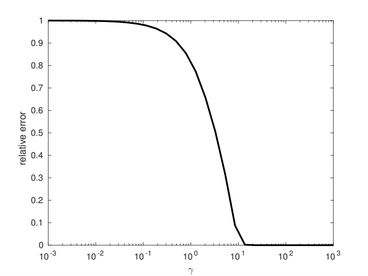
Data-driven simulation using NN (with data scaling by γ)

Given i/o data $(\boldsymbol{u}^d, \boldsymbol{y}^d)$ and simulation input $\boldsymbol{u}^s$

Find $\boldsymbol{y}^s$ from

$$\underset{\boldsymbol{y}^s}{\text{minimize}} \left\| \begin{bmatrix} \gamma H(\boldsymbol{u}^d) & H(\boldsymbol{u}^s) \\ \hline \gamma H(\boldsymbol{y}^d) & H(\boldsymbol{y}^s) \end{bmatrix} \right\|_\star$$





Guarantees for perfect recovery

Hankel matrix completion (Usevich-Comon [9])

Consider sum-of-exponentials $x_k = \sum_{j=1}^{n} c_j \lambda_j^k$ ("impulse response")

Given $x_1, \dots, x_L$
Find $x_{L+1}, \dots, x_{L+M}$

$$H = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_L \\ x_2 & x_3 & \iddots & x_L & x_{L+1} \\ x_3 & \iddots & \iddots & \iddots & \vdots \\ \vdots & x_L & \iddots & \iddots & x_{L+M-1} \\ x_L & x_{L+1} & \cdots & x_{L+M-1} & x_{L+M} \end{bmatrix}$$

Do rank and nuclear norm minimization have the same solution?
Yes, provided stability conditions on poles $\lambda_j < 1$ [2, 9]

Can we develop a similar result for data-driven simulation?

Affine matrix structure

The Hankel matrix $\mathcal{H}(y)$ has an affine matrix structure

$$\mathcal{H}(y) = \left[\begin{array}{c|c} \gamma H(u^d) & H(u^s) \\ \hline \gamma H(y^d) & H(y) \end{array}\right] = S_0 + \sum_{k=1}^{N_s} y_k S_k$$

Specifically, we have

$$\begin{aligned} S_0 &= \left[\begin{array}{c|c} H_L(u_d) & H_L(u_s') \\ H_L(y_d) & 0 \end{array}\right] \\ S_1 &= \left[\begin{array}{c|c} 0 & 0 \\ 0 & H_L(0_{L-1} \wedge (1, 0, \ldots, 0)) \end{array}\right] \\ &\vdots \\ S_{N_s} &= \left[\begin{array}{c|c} 0 & 0 \\ 0 & H_L(0_{L-1} \wedge (0, \ldots, 0, 1)) \end{array}\right] \end{aligned}$$

(prepending with zeros for initial state)

Guarantees for successful recovery

Lemma (Gillard and Usevich, 2018)

Let $\mathbf{y}^\star \in \mathbb{R}^{N_s}$ *and the compact SVD of* $\mathcal{H}(\mathbf{y}^\star)$ *be given by*

$$\mathcal{H}(\mathbf{y}^\star) = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top.$$

Further, let $\mathbf{P} = \mathbf{I} - \mathbf{U}\mathbf{U}^\top$, $\mathbf{Q} = \mathbf{I} - \mathbf{V}\mathbf{V}^\top$, *and* $\mathbf{B} = \mathbf{U}\mathbf{V}^\top$.

Then the following statements hold

- $\mathbf{y}^\star$ *is a global optimizer if* $\exists \mathbf{M}$ *with* $\|\mathbf{M}\|_2 \leq 2$, *and*

$$\langle \mathbf{P}\mathbf{M}\mathbf{Q} + \mathbf{B}, \mathbf{S}_k \rangle = 0, \quad \forall k = 1, \ldots, N_s.$$

- *If* $\|\mathbf{M}\|_2 < 1$, *then* $\mathbf{y}^\star$ *is the unique minimizer*

Existence of the matrix $\mathbf{M}$ is a *certificate* for optimality

Conditions for optimality

Verifying if minimal rank solution $y^\star$ minimizes the nuclear norm?

Corollary (Optimality certificate)

Let M_{min} be obtained from

$$\begin{array}{ll} \underset{M}{\text{minimize}} & \|M\|_2 \\ \text{subject to} & \underbrace{S^\top (Q(\gamma) \otimes P(\gamma))}_{\mathcal{A}(\gamma)} \operatorname{vec}(M) = \underbrace{S^\top \operatorname{vec}(B)}_{b(\gamma)} \end{array}$$

Then $y^\star$ minimizes the nuclear norm iff $\|M_{min}\|_2 \leq 1$.
If, in addition, $\|M_{min}\|_2 < 1$, then it is the unique solution.

Here $P(\gamma)$ and $Q(\gamma)$ are as in lemma, with vectorized constraints $\langle PMQ + B, S_k \rangle = 0$ with $S = \begin{bmatrix} \operatorname{vec}(S_1) & \cdots & \operatorname{vec}(S_{N_s}) \end{bmatrix}$.

Conditions for optimality

It is difficult to minimize the two-norm, but we can solve the least squares problem (we have $\|M\|_2 \leq \|M\|_F$), see [9, 8]

Corollary (Least-squares optimality certificate)

Let $M^\star$ be the minimizer of

$$\begin{aligned} \underset{M}{\text{minimize}} \quad & \|M\|_F = \sqrt{\text{vec}(M)^\top \text{vec}(M)} \\ \text{subject to} \quad & \underbrace{S^\top (Q(\gamma) \otimes P(\gamma))}_{\mathcal{A}(\gamma)} \text{vec}(M) = \underbrace{S^\top \text{vec}(B)}_{b(\gamma)} \end{aligned} \tag{1}$$

If $\|M^\star\|_2 < 1$, then $y^\star$ is the solution of the nuclear norm minimization.

Conclusion

Summary

- Data-driven simulation bypasses explicitly building model
- Leverage low-rank properties via Hankel matrix
- Convex relaxation via nuclear norm minimization
- Scale available data set [R1]
- Guarantees for perfect recovery [R2]

[R1] P. D. and I. Markovsky. "Data-driven simulation using the nuclear norm heuristic". Proc International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2019), Brighton, UK, 2019.

[R2] P. D. and I. Markovsky and K. Usevich. "Data-driven simulation using nuclear norm matrix completion: guarantees for successful recovery", submitted to Workshop on Low-Rank Models and Applications (LRMA19), Mons, Belgium, 2019.

Thank you!

Data-driven simulation

using the nuclear norm heuristic

Philippe Dreesen[1] Ivan Markovsky[1] Konstantin Usevich[2]

[1]Vrije Universiteit Brussel (VUB) Dept. ELEC
[2]Université de Lorraine, CNRS, CRAN

References i

E. J. Candès and B. Recht.
Exact matrix completion via convex optimization.
Found. Comput. Math., 9(6):717–772, April 2009.

L. Dai and K. Pelckmans.
On the nuclear norm heuristic for a Hankel matrix completion problem.
Automatica, 51:268–272, 2015.

M. Fazel.
Matrix rank minimization with applications.
PhD thesis, Stanford University, 2002.

M. Fazel, H. Hindi, and S. Boyd.
A rank minimization heuristic with application to minimum order system approximation.
volume 6 of *Proceedings American Control Conference*, pages 4734–4739, 2001.

M. Fazel, T. K. Pong, D. Sun, and P. Tseng.
Hankel matrix rank minimization with applications to system identification and realization.
SIAM J. Matrix Anal. Appl, 34(3):946–977, 2013.

N. Gillis and F. Glineur.
Low-rank matrix approximation with weights or missing data is NP-hard.
SIAM J. Matrix Anal. Appl., 32(4):1149–1165, 2011.

I. Markovsky.
A missing data approach to data-driven filtering and control.
IEEE Trans. Automat. Contr., 62(4):1972–1978, April 2017.

B. Recht, M. Fazel, and P. A. Parrilo.
Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization.
SIAM Rev., 52(3):471–501, 2010.

K. Usevich and P. Comon.
Hankel low-rank matrix completion: Performance of the nuclear norm relaxation.
IEEE J. Sel. Top. Signal Process., 10(4):637–646, June 2016.