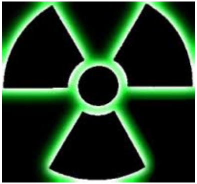
NucID: A direct approach to SYSID using Nuclear Norms.

Mohammad Naghsh, Ruben Cubo, Kristiaan Pelckmans

September 21, 2015



Overview

Today:

- NucID.
- Analysis.
- Algorithm.
- Extensions.
- Numerical results + discussion.

NucID

Why:

- Short time-series.
  - Incorporate structure!
  - Initial state estimation.
  - No averaging-out.
  - Many different short series (bio).
- Messy data.
  - Low SNRs!
  - Missing/irregularly sampled.
  - No averaging-out.
- Clear objective: to analysis.
- Convex optimization:
  - No local minima.
  - KKT conditions.
  - Algorithms.
- Direct extensions

NucID (Ct'd)

History of regularisation:

- Ill-posed problems $\|\cdot\|_2^2$.
- Regularisation vs. priors.
- Smoothing splines $\|\cdot\|_S^2$.
- Kernels $\|\cdot\|_K^2$.
- Compressed sensing $\|\cdot\|_1$.
- Nuclear norm $\|\cdot\|_*$.

Why simple models - because few even simpler ones to falsify!

NucID (Ct'd)

History of Nuclear norm analysis: Analysis: under which conditions (sampling) is $\| \cdot \|_* = \text{rank}(\cdot)$

- Fazel, 2001, 2002.
- Candes & Recht, 2009.
- Recht, Fazel & Parilo, 2010.
- Gross, 2011.

But specific for SI

- A. Hansson, Z. Liu, L. Vandeberghe, Verhaegen (2010 - ...)
- M. Fazel & al. (2010)
- I. Markovsky (no, it doesn't, 2012)
- L. Dai & K (no, not in general, 2014)
- Survey presentation of C. Rojas at ECC 2013.

NucID (Ct'd)

- FIR approximation:

$$y(t) = \sum_{\tau=0}^{\infty} h(\tau)u(t-\tau) \approx \sum_{\tau=0}^{d} h(\tau)u(t-\tau)$$

- In matrix form:

$$\text{Hankel}_d(h) = \begin{bmatrix} h_1 & h_2 & \dots & h_d \\ h_2 & \ddots & & \\ \vdots & & & \\ h_d & & & h_{2d-1} \end{bmatrix},$$

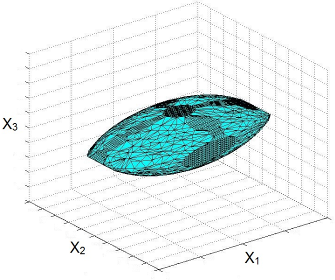
$$U_t(d) = \begin{bmatrix} u_t & \frac{u_{t-1}}{2} & \frac{u_{t-2}}{3} & \dots & \frac{u_{t-d+1}}{d} \\ \frac{u_{t-1}}{2} & \frac{u_{t-2}}{3} & & & \\ \frac{u_{t-2}}{3} & & \ddots & & \\ \vdots & & & & \\ \frac{u_{t-d+1}}{d} & & & & \frac{u_{t-2d}}{2d-1} \end{bmatrix},$$

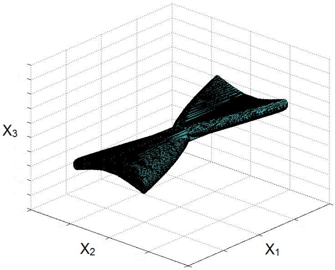
Then $\boxed{y_t \approx \sum_{\tau=0}^{d} h_\tau u_{t-\tau} = \text{tr}(\text{Hankel}_d(h) U_t(d))}$

NucID (Ct'd)

$$X = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix}$$

$\{X \text{ is Hankel}_2 : \|X\|_* \leq 1\}$ and $\{X \text{ is Hankel}_2 : \text{rank}(X) \leq 1, |x_1| \leq 1\}$





NucID (Ct'd)

NucID estimator:

- Observed dynamics as constraints.
- Minimum order model fitting data.
- So

$$\min_{h} \|\mathrm{Hankel}_d(h)\|_* \text{ s.t. } \mathrm{tr}\left(\mathrm{Hankel}_d(h) U_t(d)\right) = y(t).$$

or

$$\min_{H \text{ is } \mathrm{Hankel}_d} \|H\|_* \text{ s.t. } \mathrm{tr}\left(H U_t(d)\right) = y(t).$$

- If noise/approximation:

$$\min_{H \text{ is } \mathrm{Hankel}_d} \|H\|_* \text{ s.t. } \|\mathrm{tr}\left(H U_t(d)\right) - y(t)\|_2 \leq \epsilon$$

NucID (Ct'd)

Kalman-Ho realisation:

- State-space model:

$$\begin{cases} x_{t+1} = Ax_t + Bu_t \\ y_t = Cx_t. \end{cases}$$

$\forall t$, where $A \in \mathbb{R}^{m \times m}$ and B, C of appropriate size.

- Markov parameters

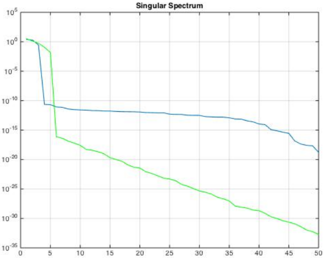
$$h(\tau) = CA^{\tau-1}B, \ \tau > 0.$$

- then

$$H = \text{Hankel}_d(h) = \mathcal{O}_{A,C}\mathcal{C}_{A,B}$$

- Projection + shift.
- MIMO.

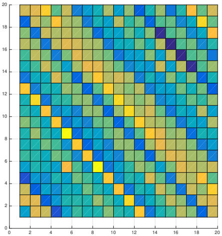
NucID (Ct'd)



Order estimation:

- Rank(H) $=$ McMillan order.
- Implementation dependent vs. thresholds.
- No-noise.
- Suggestive.
- Testing?

Analysis.



Nuclear norm:

- or $\|H\|_*$ can be written as

$$\|H\|_* = \sup_{\|G\|_2 \leq 1} \operatorname{tr}(GH)$$

(If H Hankel, then also G almost Hankel).

- $\{\lambda_i^\tau\}_\tau$ excited in at least $m(1+\epsilon)$ matrices $U_t(d)$.
- Conditions on u to give unique+correct minimal $\|H\|_*$ solution?

Algorithm.

ADMM [Hansson et al., 2012]:

- General problem:

$$\min_{h,H} \|H\|_* + \gamma \|Ah - b\|_2^2 \quad \text{s.t.} \quad H = \text{Hankel}(h)$$

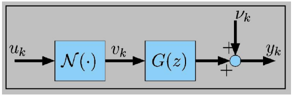
- Augmented Lagrangian:

$$\mathcal{L}(h, H, Z) = \|H\|_* + \gamma \|Ah - b\|_2^2 + \text{tr}\left(Z(H - \text{Hankel}(h))\right) + \frac{\rho}{2}\|H - \text{Hankel}_d(h)\|_2^2$$

- ADMM: Initialise H_0, h_0, Z_0.
  For $k = 1, 2, \ldots$
  1. $h_k = \arg\min_h \mathcal{L}(h, H_{k-1}, Z_{k-1})$
  2. $H_k = \arg\min_H \mathcal{L}(h_k, H, Z_{k-1})$
  3. Update $Z_k = Z_{k-1} + \rho\left(\text{Hankel}_d(h) - H\right)$.
- Implementation.

Extensions

- MIMO.
- Blind identification (SYSID 2015).
- Monotone Wiener systems.
- Hammerstein systems



$$y(t) = \sum_{\tau=1}^{\infty} h(\tau) f\left(u(t-\tau)\right)$$

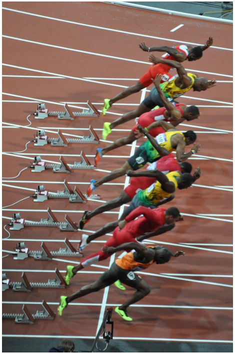
or

$$y(t) \approx \sum_{\tau=1}^{d} h(\tau) \sum_{j=1}^{m} b_j \phi_j\left(u(t-\tau)\right)$$

$$= \sum_{\tau=1}^{d} \sum_{j=1}^{m} c_{j,\tau} \phi_j\left(u(t-\tau)\right).$$

where $[c_{j,\tau} = b_j h(\tau)]_{j,\tau}$ is rank one.

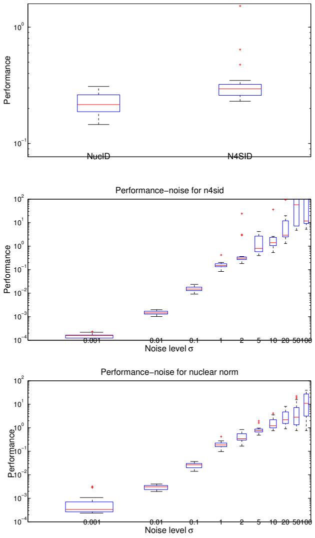
Extensions (Ct'd).



Estimating the cross-product of $\{h_\tau\}$ with initial samples (u_{-2}, u_{-1}, u_0) is rank-one:

$-$	$-$	$-$	$h_0 u_{-2}$	
$-$	$-$	$h_1 u_{-2}$	$h_0 u_{-1}$	
$-$	$h_2 u_{-2}$	$h_1 u_{-1}$	$h_0 u_0$	
$h_3 u_{-2}$	$h_2 u_{-1}$	$h_1 u_0$	$h_0 u_1$	$\rightarrow y(1)$
$h_3 u_{-1}$	$h_2 u_0$	$h_1 u_1$	$h_0 u_2$	$\rightarrow y(2)$
$h_3 u_0$	$h_2 u_1$	$h_1 u_2$	$h_0 u_3$	$\rightarrow y(3)$
$h_3 u_1$	$h_2 u_2$	$h_1 u_3$	$h_0 u_4$	$\rightarrow y(4)$
$h_3 u_2$	$h_2 u_3$	$h_1 u_4$	$h_0 u_5$	$\rightarrow y(5)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

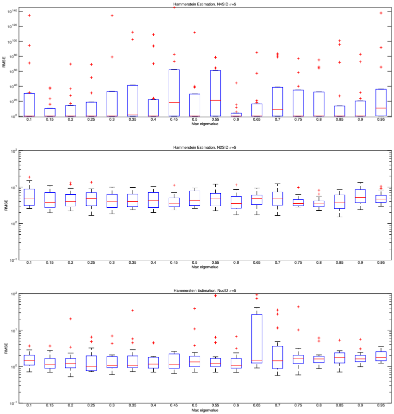
Results + discussion.



LTI case:

- SISO, no noise
  $n = 100, d = 100, n = 25$.
- SISO, with noise
  $n = 200, d = 100, n = 5, SNR = 1$.
- MIMO.

Results + discussion (Ct'd).



Hammerstein case:

- SISO, no noise $T = 500, d = 40, m = 10, n = 5.$
- SISO, with noise $T = 500, d = 40, m = 10, n = 5, SNR = 1.$

Conclusions

Take home:

- NucID.
- Analysis.
- Algorithm.
- Extensions.
- Numerical results + discussion.