Near optimal finite time identification of arbitrary linear dynamical systems

Tuhin Sarkar & Alexander Rakhlin

Massachusetts Institute of Technology

June 12, 2019
Plan

1. Problem Definition
2. Analysis and Techniques
3. Results
4. Main Results
5. Poster Details
Linear Time Invariant (LTI) Systems

LTI systems appear in autoregressive processes, control and RL systems. Formally,

$$X_{t+1} = AX_t + \eta_{t+1}$$  (1)

- $X_t, \eta_t \in \mathbb{R}^n$. $X_t$ is state vector, $\eta_t$ is noise vector.
- $A$ is state transition matrix : characterizes the LTI system.
- Assume $\{\eta_t\}_{t=1}^{\infty}$ is isotropic and subGaussian.
Linear Time Invariant (LTI) Systems

LTI systems appear in autoregressive processes, control and RL systems. Formally,

\[ X_{t+1} = AX_t + \eta_{t+1} \quad (1) \]

- \( X_t, \eta_t \in \mathbb{R}^n \). \( X_t \) is state vector, \( \eta_t \) is noise vector.
- \( A \) is state transition matrix : characterizes the LTI system.
- Assume \( \{\eta_t\}_{t=1}^{\infty} \) is isotropic and subGaussian.
Linear Time Invariant (LTI) Systems

LTI systems appear in autoregressive processes, control and RL systems. Formally,

\[ X_{t+1} = AX_t + \eta_{t+1} \]  \hspace{1cm} (1)

- \( X_t, \eta_t \in \mathbb{R}^n \). \( X_t \) is state vector, \( \eta_t \) is noise vector.
- \( A \) is state transition matrix: characterizes the LTI system.
- Assume \( \{\eta_t\}_{t=1}^{\infty} \) is isotropic and subGaussian.
Learning $A$ from data

**Goal**: Learn $A$ from $\{X_t\}_{t=1}^T$

$$\hat{A} = \inf_{A_o} \sum_{t=1}^{T} ||X_{t+1} - A_o X_t||_2^2$$

Estimation error

$$E = A - \hat{A} = \left( \sum_{t=1}^{T} \eta_{t+1} X_t^\top \right) \left( \sum_{t=1}^{T} X_t X_t^\top \right)^+$$  \hspace{1cm} (2)

Error analysis hard: $\{X_t\}_{t=1}^T$ are not independent.
Faradonbeh et. al. (2017). Finite time identification in unstable linear systems.


Past works fail to capture correct behavior for all $A$. 
Main Technique

The analysis proceeds in two steps:

- Show invertibility of sample covariance matrix:
  \[ \sum_{t=1}^{T} X_t X_t^\top \approx f(T)I. \]

- Show the following for self–normalized martingale term:
  \[
  \left( \sum_{t=1}^{T} \eta_{t+1} X_t^\top \right) \left( \sum_{t=1}^{T} X_t X_t^\top \right)^{-1/2} = O(1)
  \]
The analysis proceeds in two steps:

- Show invertibility of sample covariance matrix:
  \[ \sum_{t=1}^{T} X_t X_t^\top \approx f(T)I. \]

- Show the following for self-normalized martingale term:

\[
\left( \sum_{t=1}^{T} \eta_{t+1} X_t^\top \right) \left( \sum_{t=1}^{T} X_t X_t^\top \right)^{-1/2} = O(1)
\]
Let $\rho_i(A)$ be the absolute value of $i^{th}$ eigenvalue of $A$ with $\rho_i(A) \geq \rho_{i+1}(A)$. Then

- $\rho_i \in S_0 \implies \rho_i(A) \leq 1 - C/T$
- $\rho_i \in S_1 \implies \rho_i(A) \in [1 - C/T, 1 + C/T]$
- $\rho_i \in S_2 \implies \rho_i(A) \geq 1 + C/T$

**Theorem**

- $\rho_i(A) \in S_0 \implies \sum_{t=1}^{T} X_t X_t^\top = \Theta(T)$
- $\rho_i(A) \in S_1 \implies \sum_{t=1}^{T} X_t X_t^\top = \Omega(T^2)$
- $\rho_i(A) \in S_2 \implies \sum_{t=1}^{T} X_t X_t^\top = \Omega(e^{aT})$ (under necessary and sufficient “regularity” conditions only)
Let $\rho_i(A)$ be the absolute value of $i^{th}$ eigenvalue of $A$ with $\rho_i(A) \geq \rho_{i+1}(A)$. Then

- $\rho_i \in S_0 \implies \rho_i(A) \leq 1 - C/T$
- $\rho_i \in S_1 \implies \rho_i(A) \in [1 - C/T, 1 + C/T]$
- $\rho_i \in S_2 \implies \rho_i(A) \geq 1 + C/T$

**Theorem**

- $\rho_i(A) \in S_0 \implies \sum_{t=1}^{T} X_t X_t^\top = \Theta(T)$
- $\rho_i(A) \in S_1 \implies \sum_{t=1}^{T} X_t X_t^\top = \Omega(T^2)$
- $\rho_i(A) \in S_2 \implies \sum_{t=1}^{T} X_t X_t^\top = \Omega(e^{aT})$ (under necessary and sufficient “regularity” conditions only)
Let $\rho_i(A)$ be the absolute value of $i^{th}$ eigenvalue of $A$ with $\rho_i(A) \geq \rho_{i+1}(A)$. Then

- $\rho_i \in S_0 \implies \rho_i(A) \leq 1 - C/T$
- $\rho_i \in S_1 \implies \rho_i(A) \in [1 - C/T, 1 + C/T]$
- $\rho_i \in S_2 \implies \rho_i(A) \geq 1 + C/T$

**Theorem**

- $\rho_i(A) \in S_0 \implies \sum_{t=1}^{T} X_t X_t^\top = \Theta(T)$
- $\rho_i(A) \in S_1 \implies \sum_{t=1}^{T} X_t X_t^\top = \Omega(T^2)$
- $\rho_i(A) \in S_2 \implies \sum_{t=1}^{T} X_t X_t^\top = \Omega(e^{aT})$ (under necessary and sufficient “regularity” conditions only)
Self Normalized Martingale

Theorem (Abbasi-Yadkori et. al. 2011)

Let $V$ be a deterministic matrix with $V \succ 0$. For any $0 < \delta < 1$ and \{\eta_t, X_t\}^T_{t=1}$ defined as before, we have with probability $1 - \delta$

$$
\| (\bar{Y}_{T-1})^{-1/2} \sum_{t=0}^{T-1} X_t \eta'_{t+1} \|_2 
\leq R \sqrt{8n \log \left( \frac{5 \det(\bar{Y}_{T-1})^{1/2} \det(V)^{-1/2}}{\delta^{1/n}} \right)}
$$

(3)

where $\bar{Y}^{-1}_\tau = (Y_\tau + V)^{-1}$ and $R^2$ is the subGaussian parameter of $\eta_t$. 
Combining the previous results (and a few more matrix manipulations) we show

**Theorem**

- \( \rho_i(A) \in S_0 \cup S_1 \cup S_2 \implies \|E\|_2 = O(T^{-1/2}) \)
- \( \rho_i(A) \in S_1 \cup S_2 \implies \|E\|_2 = O(T^{-1}) \)
- \( \rho_i(A) \in S_2 \implies \|E\|_2 = O(e^{-aT}) \) (under necessary and sufficient “regularity” conditions only)
Combining the previous results (and a few more matrix manipulations) we show

**Theorem**

- $\rho_i(A) \in S_0 \cup S_1 \cup S_2 \implies \|E\|_2 = O(T^{-1/2})$
- $\rho_i(A) \in S_1 \cup S_2 \implies \|E\|_2 = O(T^{-1})$
- $\rho_i(A) \in S_2 \implies \|E\|_2 = O(e^{-aT})$ (under necessary and sufficient “regularity” conditions only)
Combining the previous results (and a few more matrix manipulations) we show

**Theorem**

- $\rho_i(A) \in S_0 \cup S_1 \cup S_2 \implies \|E\|_2 = O(T^{-1/2})$
- $\rho_i(A) \in S_1 \cup S_2 \implies \|E\|_2 = O(T^{-1})$
- $\rho_i(A) \in S_2 \implies \|E\|_2 = O(e^{-aT})$ (under necessary and sufficient “regularity” conditions only)
Main Result 2

Regularity condition: All eigenvalues greater than one should have geometric multiplicity one.

**Theorem**

*If the regularity conditions are violated then OLS is inconsistent.*

OLS cannot learn $A = \rho I$ where $\rho \geq 1.5$. $E$ has a non-trivial probability distribution.
Please come to our poster at Pacific Ballroom #193 at 6.30 pm today.

Thank you!