Greedy identification of linear dynamical systems

Matthieu Blanke, Marc Lelarge

DI ENS, INRIA Paris

The problem

Linear dynamics in $\mathbb{R}^d$:

$$x_{t+1} = A_\star x_t + B_\star u_t + w_t, \quad 0 \leq t < T$$

with noise $w_t \sim \mathcal{N}(0, \sigma^2 I_d)$, control variables $u_t \in \mathbb{R}^m$ subject to the power constraint

$$\frac{1}{T}\sum_{t=0}^{T-1} \|u_t\|^2 \leq \gamma^2,$$

and the **unknown parameters** of the dynamics $\theta_\star = (A_\star \; B_\star) \in \mathbb{R}^{d \times (d+m)}$, which are estimated from the observed trajectory (x_t).

Goal Find the best inputs (u_t) to drive the system towards a **maximally informative** trajectory for the estimation of $\theta_\star$.

Motivation

In complex controlled systems (*e.g.* aircraft, robot, ...), the unknown parameters are estimated in an identification phase. The estimation must be **fast** and **sample-efficient**:

- collecting observations is **costly** (think of an aircraft test flight),
- the linear regime is a **short-term** approximation,
- the algorithm should work **on-line**.

Other prospectives include model-based reinforcement learning (LQR) and bandits pure exploration.

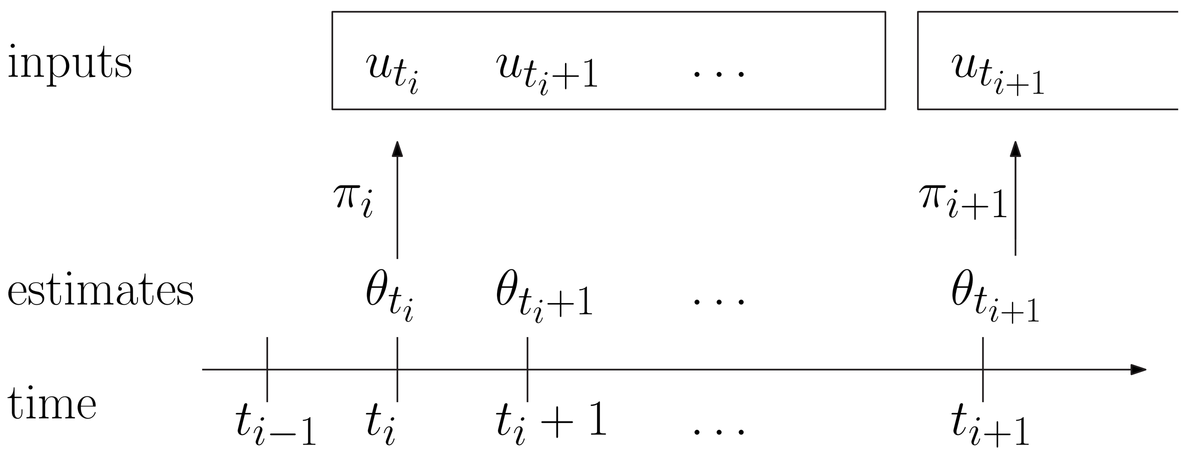
Test flight of a A330neo, source: Airbus.

Related work Theory of optimal design in the 1970s [1, 3], focused on single-input systems and on input design in the frequency domain. Growing interest from the machine learning community lately [5, 6] with theoretical (asymptotic) bounds. Identification algorithm TOPLE proposed in [6].

Identification and planning

Sequential identification: choose a **policy** π_i adaptively at times $t = \{t_i\}$ with the current estimate θ_i.

Planning: π_i is chosen to minimize a cost F.



Algorithm 1 Sequential identification

inputs $\hat{\theta}$, $\{t_i\}$, F, θ_0, π_0
output final estimate θ_T
for $0 \leq i \leq n-1$ **do**
  run the true system from $t_i + 1$ to t_{i+1} with inputs $u_t = \pi_i(x_{1:t}, u_{1:t-1})$
  $\theta_i = \hat{\theta}(x_{1:t_i}, u_{1:t_i-1})$ ▷ estimation
  π_i solves $\min_{\pi \in \Pi_\gamma} F(\pi; \theta_i, t_{i+1})$ ▷ planning
end for

Optimal design

By the **linear structure** of the problem, a natural estimator is the **least squares estimator**

$$\hat{\theta}(\tau)^\top = M_{T-1}^{-1} \sum_{t=0}^{T-1} z_t x_{t+1}^\top$$

with $z_t = \begin{pmatrix} x_t \\ u_t \end{pmatrix}$ and $M_t = \sum_{s=0}^{t} z_s z_s^\top$.

A natural cost functional is given by the theory of **optimal experiment design**:

$$F(\pi; \theta, t) = -\Phi\left(\mathbb{E}_\theta\left[M_t\right]\right)$$

for $\Phi(M) = -\mathrm{tr}(M^{-1})$ (A-optimality) or $\Phi(M) = \log\det M$ (D-optimality).

Greedy approach

Greedy planning: optimize the cost functional one-step-ahead, *i.e.* set $t_i = i$. Leads to

$$\max_{u \in \mathbb{R}^m} \quad \Phi\left(M_{t-1} + z(u) z(u)^\top\right)$$

such that $z(u) = \begin{pmatrix} x_t \\ u \end{pmatrix}$ and $\|u\|^2 = \gamma^2$.

Upsides Can be solved at minimal cost. The learning algorithm runs online: u_t is chosen on-the-fly with the current knowledge θ_t.

Downsides Greedy approximation, no theoretical guarantees.

Algorithm 2 Greedy sequential identification

output final estimate θ_T
for $0 \leq t \leq T-1$ **do**
  $u_t \in \operatorname{argmax}_{\|u\|^2 = \gamma^2} \Phi\left(M_{t-1} + \mathbb{E}_\theta[z_t z_t^\top]\right)$
  play u_t, observe z_{t+1}
  $M_t = M_t + z_{t+1} z_{t+1}^\top$
  $\theta_{t+1}^\top = M_{t+1}^{-1}\left(M_t \theta_t^\top + z_t y_t^\top\right)$
end for

Proposition For D-optimality and A-optimality, greedy planning reduces to

$$\min_{u \in \mathbb{R}^m} \quad u^\top Q u - 2 b^\top u$$

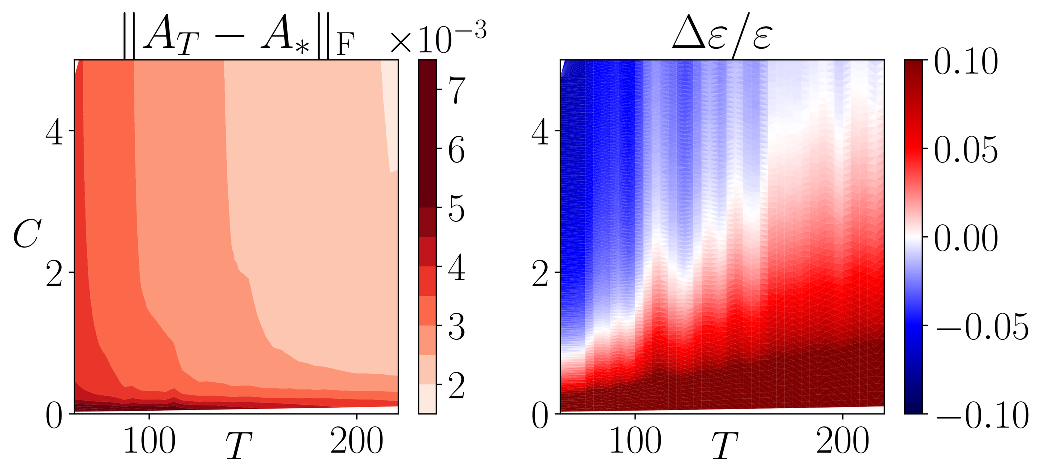
$$\text{such that} \quad \|u\|_2^2 = \gamma^2$$

with $Q \in \mathbb{R}^{m \times m}$ and $b \in \mathbb{R}^m$ simple functions of M_{t-1} and θ_t.

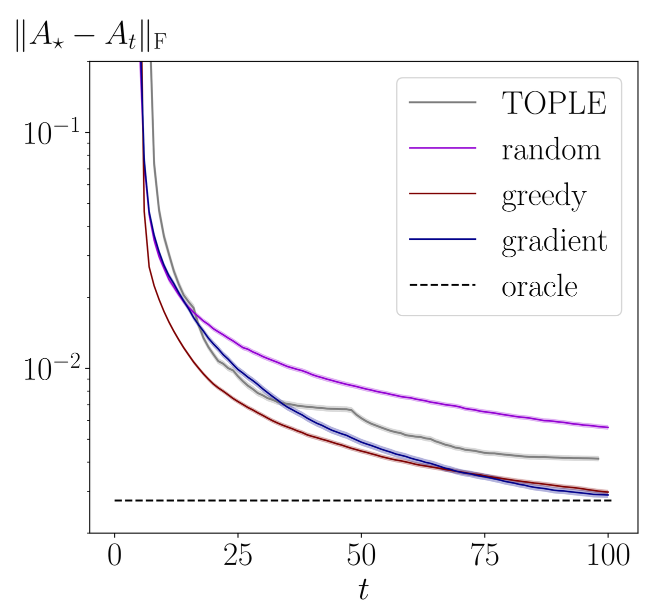
Characterization of the minimizers by the Lagrange multiplier theorem. Efficient numerical solution at the cost of an eigenvector decomposition and a scalar root-finding search.

Results

We compare our greedy algorithm to gradient-based approaches. Two resources: number of observations T and compute C. Larger C means more gradient steps. Average over 1000 random matrices.



Left Performance of gradient. **Right** Relative performance of gradient vs greedy. Positive means greedy is better.



Performance of various algorithms versus the number of observations averaged over 1000 random matrices.

Real-life system: lateral system of a a C-8 Buffalo aircraft. The lateral motion is controlled by the aileron and rudder angles [2]. The number of observations is limited: $T = 125$. The results are averaged over 1000 trials.

	Random	TOPLE	Gradient	Greedy	Oracle
Estimation error	1.1×10^{-1}	8.6×10^{-2}	8.3×10^{-2}	8.2×10^{-2}	8.0×10^{-2}
Computation time	1	55.7	25	1.13	100

Conclusion

- We devised a fast, sample-efficient algorithm for system identification.
- No theoretical guarantees but good performance in a practical framework with limited observations.
- Prospective: extension to LQR with unknown parameters ? to non-linear systems ?

References

[1] G.C. Goodwin and R.L. Payne. *Dynamic System Identification: Experiment Design and Data Analysis*. Developmental Psychology Series. Academic Press, 1977.

[2] NK Gupta, RK Mehra, and WE Hall Jr. Application of optimal input synthesis to aircraft parameter identification, 1976.

[3] R. Mehra. Optimal inputs for linear system identification. *IEEE Transactions on Automatic Control*, 19(3):192–200, 1974.

[4] Friedrich Pukelsheim. *Optimal design of experiments*. SIAM, 2006.

[5] Andrew Wagenmaker and Kevin Jamieson. Active learning for identification of linear dynamical systems, 2020.

[6] Andrew Wagenmaker, Max Simchowitz, and Kevin Jamieson. Task-optimal exploration in linear dynamical systems, 2021.