Nonparametric Independent Process Analysis

Zoltán Szabó
Faculty of Informatics, Eötvös Loránd University,
Pázmány Péter sétány 1/C, Budapest, H-1117, Hungary
email: zzoli@cs.elte.hu
web: http://nipg.inf.elte.hu

Barnabás Póczos
School of Computer Science, Carnegie Mellon University,
5000 Forbes Ave, 15213, Pittsburgh, PA, USA
email: bapoczos@cs.cmu.edu
web: http://www.autonlab.org

1. Introduction

• Linear dynamical systems (LDS): \( x_t = f(x_{t-1},...x_{t-L},u_t) + \varepsilon_t \)
  – Limitations: linear dynamics, Gaussian driving noise.

• Non-Gaussian driving noises:
  – ICA = separation of mixed non-Gaussian, one-dimensional sources: \( x_t \sim ASA \).
  – Limitations: Unknown, nonparametric dynamics is hardly touched: stationary + ergodic sources, constrained mixing.

• Our contributions:
  – ISA with nonparametric, asymptotically stationary dynamics.
  – Unknown and possibly different dimensional components.
  – Simple separation based solution: kernel regression + ISA.

2. Problem

• Task: estimate linearly mixed (\( f \)) multidimensional sources (\( s_t \)) with independent driving noises (\( e_t \)) of unknown functional autoregressive dynamics (\( f \)) with independent driving noises (\( e_t \)):
  \[
  s_t = f(s_{t-1},...,s_{t-L},e_t) + e_t, \\
  x_t = A_s s_t.
  \]
  \( (1), (2) \)

• Assumptions: \( A_s \): full column rank; \( e_t \): independent \( e_t \).

• Goal (IAR-IPA): estimate \( A_s \) from observations \( x_t \):
  – Special cases:
    – if \( f \) were known, linear: autoregressive IPA (AR-IPA).
    – If order \( L = 0 \): traditional ISA.
    – ISA with one-dimensional independent subspaces (\( d = 1 \)): ICA.

3. Method

• In the ISA special case: ISA separation principle
  – ICA = ICA up to permutation – conjecture of Cardoso (’98).

• We derive a similar reduction scheme for the IAR-IPA problem:
  – IAR-IPA = IAR identification + ISA.

• According to (1)-(2) \( x_t \) is IAR with innovation \( e_t \):
  \[
  x_t = A(s_{t-1},...,s_{t-L}) + A_0 + e_t = g(s_{t-1},...,s_{t-L}) + e_t.
  \]
  Idea: (3) = nonparametric regression problem.

4. Illustration

• Dataset: d-geom (\( d_1 = 2, d_2 = 3, d_3 = 4 \), ikeda (\( M = 2, d_n = 2 \)): see Fig. 3.

• Performance (Amari-index): ISA ambiguities \( \rightarrow \) measure the block-permutation property of \( G = W_{G,A} \).

• Experiences:
  – d-geom (different dimensional sources; 12D):
    – Amenable for sample size \( T \geq 100,000 \), see Fig. 4.
    – ikeda (Fig. 5).
  – IAR-IPA: can not find the proper subspaces.
  – LDS: EM + Kalman smoother \( \rightarrow \) Amari-index = 0.48 \( \approx \) poor.
  – IAR-IPA: precise estimation for sample number \( T \geq 100,000 \), Amari-index \( \approx 0.855 \).

Figure 3: Datasets. Left: d-geom. Right: ikeda.

Figure 4: Illustration on the d-geom dataset. Left: Amari-index. Right: Hinton-diagram of \( G \).

Figure 5: Illustration on the ikeda dataset. (a): Amari-index. (b): Observation, \( x_t \). (c): Hinton-diagram of \( G \). (d): Estimated subspaces \( \hat{A}_s \), IAR-IPA.