INDEPENDENT COMPONENT ANALYSIS

Machine Learning

IIIy CSE IDP and IDDMP

Refer : Stephen Marsland

Motivation

- **Blind Source Separation**:
  Many real-world signals arise from multiple underlying sources that mix together before observation.
- **Limitations of PCA/FA:**
  PCA and Factor Analysis produce uncorrelated components but not necessarily statistically independent ones.
- **Goal:**
  Recover original independent source signals from observed mixtures (e.g., separate voices at a crowded event).

What is Independent Component Analysis ?

- A technique to find a **linear transformation** of the data such that the resulting components are **statistically independent**.
- Goes beyond uncorrelatedness (covariance) by enforcing **independence** (joint probability factorizes).
- **Key assumption:** Observations are linear mixtures of underlying independent sources.
- Solves the cocktail party problem under the condition that the number of sensors matches the number of sources.

Blind Source Separation & Cocktail Party Problem

- **Concept**: Observed data are mixtures of independent physical processes.
- **Statistical Independence**: For components b_i and b_j, independence implies $E[b_i , b_j] = E[b_i] E[b_j]$ (not just zero covariance).
- **Cocktail Party Analogy**:
  - Multiple overlapping sounds (voices, music, noise) are recorded by several microphones.
  - ICA recovers individual source signals provided the number of microphones ≥ number of sources.
- **Applications**:
  - Audio signal processing (speech separation)
  - Biomedical signal analysis (EEG/MEG)
  - Financial data modeling

Mixing Model Formulation

- **Sources and Observations**:
  - Two underlying source signals over tim (s_1^t, s_2^t)
  - Two microphone recording: (x_1^t, x_2^t)
- **Linear Mixing Equations**:

$$\begin{aligned} x_1 &= as_1 + bs_2, \\ x_2 &= cs_1 + ds_2, \end{aligned}$$

- **Matrix Form:**

$\mathbf{x} = A\,\mathbf{s}$, where $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

- **Demixing:**

  Ideal recovery would use $\mathbf{s} = A^{-1}\mathbf{x}$,but since is unknown, we estimate a demixing matrix (approximate A^{-1}).

Key Observations about Mixtures

- **Dependence**: The observed mixtures (x) are correlated, even though the source signals (s) are independent
- **Central Limit Effect**: Mixtures tend to appear more Gaussian (normal) than the original sources, irrespective of source distributions (Central Limit Theorem).
- **Complexity**: Mixed signals have more complex distributions than each individual source.

Identifying Source Components

- **Principle 1**: True source factors should be independent; finding independent components suggests original sources.

- **Principle 2**: Non-Gaussianity indicates source signals (mixtures are more Gaussian by Central Limit Theorem).

- **Measurement of Independence**:
  - Mutual Information: Quantifies dependence between variables

Entropy Fundamentals

- **Why Entropy?**
  - Entropy measures the randomness or uncertainty in a variable.
  - Gaussian variables have the highest entropy among all variables with the same variance.
  - Therefore, lower entropy implies deviation from Gaussianity, which helps identify independent sources.

$$H(y) = -\int g(y) \log g(y) \, dy$$

- **Entropy Definition :**

where $g(y)$ is the probability density function of y.

Negentropy & Approximation

- **Negentropy**: Measures non-Gaussianity by comparing entropy of the signal with a Gaussian variable of the same v $J(y) = H(z) - H(y)$

  where H(·) is entropy and z is a Gaussian variable with the same variance as y.

- The higher the negentropy, the more non-Gaussian the signal is, suggesting it might be a source

- **Approximation for Practical Use:** $J(y) \approx \big(E[G(y)] - E[G(z)]\big)^2$

  with nonlinearity:

  - $G(u) = \frac{1}{a}\log\cosh(a\,u)$ for 1 ≤ a ≤ 2
  - $G'(u) = \tanh(a\,u)$

  Where **a** is a tuning parameter (1 ≤ a ≤ 2) and **u** is the argument to G

FastICA Algorithm

- **Overview**:
  - FastICA is an efficient fixed-point algorithm for estimating the demixing matrix W.
- **Key Steps**:
  - **Centering & Whitening**: Preprocess data to zero mean and unit covariance.
  - **Iteration**: Update weight vector w using $w \leftarrow E[\, x\, G(w^T x)\,] - E[G'(w^T x)]\, w$
  - **Normalization**: Enforce **|| w || = 1** and decorrelate different w's.
- **Implementations:**
  - Available in Python libraries (e.g., scikit-learn, MDP package).
  - Widely used in signal processing and data analysis applications.

Summary

- **ICA Objective**: Decompose observed signals into statistically independent components.
- **Core Principles**:
  - Independence over mere uncorrelatedness.
  - Exploit non-Gaussianity (negentropy) to identify sources.
- **Workflow**:
  - Preprocess (center, whiten).
  - Optimize demixing matrix W (e.g., FastICA).
  - Recover source estimates $s = W \cdot x$.
- **Strengths & Applications**:
  - Effective in blind source separation (audio, biomedical signals, finance).
  - Scalable, fast algorithms (e.g., FastICA).