Empirical Intrinsic Geometry for Nonlinear Signal Analysis

Ronen Talmon
Department of Electrical Engineering
Technion – IIT

The 4th Annual International TCE Conference
May 2014
Manifold Learning

- Suppose we have a set of 15 images of a toy:
Manifold Learning

• Sort according to the “intrinsic” variable:
Manifold Learning

• A challenging computer vision task
  – High dimensional
  – Small number of samples

• Manifold learning provides a data-driven unsupervised solution, that reveals the intrinsic/natural parameterization of the set
  – Without assuming a-priori knowledge
  – Viewing the samples merely as vectors of numbers
Manifold Learning

• Data points in a high-dimensional space
  – Images, songs, finance and biomedical data

• The data do not fill the high-dimensional space “uniformly”
  – The dimension is chosen by the user or the acquisition system
Manifold Learning

• Data points in a high-dimensional space
  – Images, songs, finance and biomedical data

• The data do not fill the high-dimensional space “uniformly”
  – The dimension is chosen by the user or the acquisition system
  – Often, the data lie on a low-dimensional manifold, conveying its intrinsic degrees of freedom
Manifold Learning

• **Input**: a set of high dimensional samples

• **Two parts:**
  – **Local Step**: define pairwise affinities between two individual samples
  – **Global Step**: EVD of a kernel

• **Output**: a low dimensional representation of the samples (embedding)
Intrinsic Geometry
Intrinsic Geometry

\[ p(i) \mapsto (x(i), y(i)) \]
Intrinsic Geometry

\[ p(i) \leftrightarrow (x(i), y(i)) \]
\[ p(i) \leftrightarrow (r(i), \theta(i)) \]
Intrinsic Geometry

We look for a canonical representation:
1. Independent of the original coordinate system

\[ p(i) \mapsto (x(i), y(i)) \]
\[ p(i) \mapsto (r(i), \theta(i)) \]
\[ p(i) \mapsto (\varphi_1(i), \varphi_2(i)) \]
Intrinsic Geometry

We look for a canonical representation:

1. Independent of the original coordinate system
2. That gives a parameterization of the intrinsic geometric structure of the data (underlying manifold)
Audio Signal Processing

• Any musical instrument is controlled by several intrinsic variables

• For example:
  – A flute is controlled by covering holes
  – A violin is controlled by the length of the strings
  – The speech system is controlled by the position of the tongue, lips, teeth, etc.
Audio Signal Processing

• View the signal as the output of a dynamical system

\[ u(t) \xrightarrow{\theta(t)} z(t) \]
Problem Setting

• Given measurements $z(t) \in \mathbb{R}^n$

• Suppose the measurements are locally stationary and depend upon hidden variables $\theta(t) \in \mathbb{R}^d$

• Our main goal is to empirically recover the hidden variables $\theta(t)$

• The primary focus is to build a pairwise distance

$$d(z(t), z(\tau)) \approx \|\theta(t) - \theta(\tau)\|^2$$
Single Channel Localization

[with I. Cohen and S. Gannot]
Single Channel Localization
Single Channel Localization

\[ u(t) \rightarrow \theta(t) \rightarrow z(t) \]

\[ \theta = \text{The direction of arrival (DOA) angle} \]
Single Channel Localization

\[ \text{DOA [Angles }^\circ\text{]} = \arccos(\psi) \]

Obtained Representation

[with I. Cohen and S. Gannot]
Biomedical Application

• Identifying preseizure states in epileptic patients from intracranial electroencephalography (icEEG)
  – No definitive ground truth

• Open questions:
  – Whether such states exist
  – Whether they can be detected in icEEG signals
Problem Setting

• In the EEG application:
  
  - $z(t)$ is the measured EEG signal
  
  - $\theta(t)$ is a representation of the brain activity or the physiological state of the subject
Biomedical Application

• Build intrinsic modeling/geometry
  – In a data-driven (unsupervised) manner, find a canonical coordinate system
  – Aiming to be independent of measurement and instrumental modalities
Biomedical Application

• The Goal:
  • Recover latent variables with a physical/physiological meaning
  • Through the description of the data in the new coordinate system
Intracranial EEG in Epilepsy Research

[with S. Mallat, H. Zaveri and R. Coifman]
Intracranial EEG in Epilepsy Research

[with S. Mallat, H. Zaveri and R. Coifman]
Consider a set of points on a 2-D torus in $\mathbb{R}^3$, which are samples of a Brownian motion on the torus.

We attach to each point a Riemannian metric that is driven by the dynamics (the Brownian motion), and therefore, it is invariant to the coordinate system.
Intrinsic Geometry

We attach to each point a Riemannian metric that is driven by the dynamics (the Brownian motion), and therefore, it is invariant to the coordinate system.
Intrinsic Geometry

Euclidean Distance between the 3D coordinates

Mahalanobis Distance between the Angles

Euclidean Distance between the Angles
Intrinsic Geometry
Intrinsic Geometry

Dynamics!
Contact 1

Contact 2

The common factors of uncorrelated/independent measurements are the intrinsic variables/state

Multiple views can replace our assumption that the intrinsic dynamics is driven by independent unit variance noise
Intracranial EEG in Epilepsy Research

Contact 1

Contact 2

Combined

[with R. Lederman]
Intracranial EEG in Epilepsy Research

[with R. Lederman]
Multimodal Signal Analysis

• **Sleep stage identification:**
  – Measurements taken in a specialized sleep laboratory:
    • EEG of brain waves
    • Electrooculography (EOG) of eye movements
    • Electromyography (EMG) of skeletal muscle activity
    • Respiratory Activity
Multimodal Signal Analysis

- **Sleep stage identification:**
  - Identify the 6 sleep stages: awake, REM, and 4 stages from light to deep sleep
Multimodal Signal Analysis

- *Sleep stage identification:*

Two (scalp) EEG contacts

[with H.-T. Wu]
Multimodal Signal Analysis

- Sleep stage identification:

Air Flow

Abdominal EMG

Chest EMG

[with H.-T. Wu]
• Joint work with:
  – Ronald Coifman (Yale)
  – Stephane Mallat (ENS-Paris)
  – Israel Cohen (Technion)
  – Sharon Gannot (Bar-Ilan)
  – Hitten Zaveri (Yale)
  – Hau-tieng Wu (Stanford)
  – Roy Lederman (Yale)

Thank you!