Multi-Kernel Regression Imputation on Manifolds via Bi-linear Modeling

Gaurav Shetty¹, Konstantinos Slavakis², Loris Cannelli³, Ukash Nakarmi⁴, Gesualdo Scutari⁴, Leslie Ying¹,⁵

UB: ¹Dept. of Electrical Eng., ²Dept. of Biomedical Eng., ³Dalle Molle Institute for Artificial Intelligence, ⁴Stanford University: Dept. of Electrical Eng. & Radiology; ⁵Purdue University: Dept. of Industrial Eng.

Introduction

What is missing data problem?
- Missing Data - Gaps in the observed/acquired data or data loss due to data delivery.
- Data acquisition and delivery processes are affected by sensor faults, connection errors in sensor networks, physical acquisition constraints, security attacks, environmental factors, and so on ...

Applications:
- Imaging Applications: MRI, CT, Remote Sensing, Seismic Imaging.
- Internet of Things (IoT): Data fusion in healthcare monitoring, Internet of Vehicles (IoV).
- Social Media: Recommender Systems, Sentiment Analysis, Social Network Analysis.

Impact:
- Imaging Applications: Incomplete data scanning leads to distorted, artifact-induced, low-resolution medical or geological imaging, not acceptable. Expensive and time consuming data acquisition schemes.
- Internet of Things: Can adversely affect the protocol implementation which plays a crucial role in remote, energy dependent sensor networks.
- Social Media: Loss of correlation information between features; Can lead to learning of incorrect models for data analysis which can further lead to incorrect biases and interpretations.

Main Idea

Consider \( y_t = \text{Measured time series data at time } t \).
- Non-linearly map all \( y_t \) to feature space where the similarity between the data points is exploited via kernel function \( k(\cdot, \cdot) \).
- Based on the assumption feature maps lie on the smooth low-dimensional manifold learn the compressed latent representations for the same.
- Latent Representations - Use the concept of tangent space to locally and affinely combine the neighboring points to describe each point on the manifold.
- Compression imposes low-rank structure to the data model, highly desirable in reconstruction problems.
- Reconstruct the high dimensional data point from latent representations.

Mathematical Formulation: The Dynamic-MRI Case

MR scanner time series data: \( Y \).
- dMRI Image series: \( X = F^{-1}(Y) \) related by inverse 2D Fourier Transform.
- Partially observed \( \tilde{S}(Y) \Rightarrow \text{Distorted, aliased, artifact induced } X \).
- \( F_k(\cdot) \Rightarrow 1\text{D Fourier Transform along temporal axis} \).

Inverse Problem Formulation

\[
\begin{align*}
\min_{X,Z,B} & \quad \frac{1}{2}\|X - DKB\|_2^2 + \lambda_{1}\|B\|_1 + \frac{1}{2}\|Z - FS(X)\|_2^2 + \lambda_{2}\|Z\|_1 \quad & \text{Data Fit Term} \\
s.t. & \quad \|Dn_{\theta}\| \leq C_{D}, \forall i \in \{1, \ldots, M\} \quad & \text{Bounding Constraint} \\
& \quad 1_{N_{\xi}}B_{i} = 1_{N_{\xi}}, \forall i \in \{1, \ldots, M\} \quad & \text{Affine Constraints} \\
& \quad \tilde{S}(Y) = FS(X) \quad & \text{MRI Data Relation} \\
\end{align*}
\]

Numerical Results

Validation

- Validation Numeric Metrics: NRMSE (voxel reconstruction error), HFEN (edge reconstruction error), M1 & M2 (sharpness measure) and SSIM (structural similarity).
- Quantitatively, the proposed kernel schemes present the best numbers when compared to the state-of-the-art methods.
- Qualitatively, the proposed scheme produces image reconstructions which are high-resolution, distortion-free, aliasing-free, artifact free and are very similar to the gold standard in regards to contrast and image structure.
- The proposed scheme consistently outperforms the other schemes for increasing number of missing values in the scanner data (acceleration rate).

Related Work

- This research was supported in part by the NSF grants 1714056 and 1718796.