Journal Watch
IEEE TSP, Jan(Volume 68), 2020

Unnikrishnan N

SPC Lab, Department of ECE
Indian Institute of Science

January 25, 2020

Variations on the Convolutional Sparse Coding Model by Ives Rey-Otero, Jeremias Sulam, and Michael Elad

Contributions

- Proposed alternate formulations of CSC in terms of mixed norms $l_2 - l_{1,\infty}$ and $l_{2,\infty} - l_1$ using local sparsity.
- Recovery algorithms using ADMM and PPXA are proposed.

System Model

$$\underset{\Gamma}{\text{minimize}} \frac{1}{2}\|X - D\Gamma\|_2^2 + \lambda\|\Gamma\|_1$$

where $D \in \mathbb{R}^{nm \times N}$, $n \ll N$ If D is a concatenation of m banded Circulant matrices, where each such matrix has a band of width $n \ll N$. As such, by simple permutation of its columns, such a dictionary consists of all shifted versions of a local dictionary D_L of size $n \times m$. Then the problem is known as CSC.

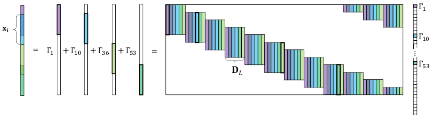


Figure: CSC model

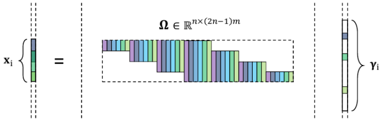


Figure: Stripe dictionary

$l_2 - l_{1,\infty}$ formulation

$$\mathbf{x}_i = \mathbf{R}_i\mathbf{X} = \mathbf{R}_i\mathbf{D}\boldsymbol{\Gamma} = \mathbf{R}_i\mathbf{D}\mathbf{S}_i^T\mathbf{S}_i\boldsymbol{\Gamma}$$

where x_i is the patch of length n extracted from i^{th} location using patch extractor operator R_i, S_i^T operator preserves the non zeros columns of R_iD and $\Omega = R_iDS_i^T$ is independant of i. Then problem can be reformulated as

$$\underset{\Gamma,\{\gamma_i\}}{\text{minimize}} \quad \frac{1}{2}\|Y - D\Gamma\|_2^2 + \lambda \max_i \|\gamma_i\|_1$$

Subject to $\forall i, \gamma_i = S_i\Gamma, i \in [N]$. Augmented lagrangian version

$$\frac{1}{2}\|Y - D\Gamma\|_2^2 + \lambda \max_i \|\gamma_i\|_1 + \frac{\rho}{2}\sum_i \|\gamma_i - S_i\Gamma + u_i\|_2^2$$

Tensor Completion From Regular Sub-Nyquist Samples by Charilaos I. Kanatsoulis, Xiao Fu, Nicholas D. Sidiropoulos, and Mehmet Akakaya

Contributions

- Study the task of sampling and reconstruction of signals that are tensors or tensor sampling.
- Generic as well as deterministic theoretical conditions (unlike CS and LRMC) are derived.
- Regular, equispaced and highly structured sampling strategies can be adopted - which has a much broader spectrum of applications in practice.

Problem Formulation

$$\mathbf{y} = sample(\underline{\mathbf{X}})$$

here $Sample(\cdot) : \mathbb{F}^{I \times J \times K} \to \mathbb{F}^{L}$, $L \ll IJK$. Applicable to both real and complex fields and higher order tensors. Goal is to study under what conditions and sampling strategies, identifying $\underline{\mathbf{X}}$ from $\mathbf{y}$ is possible.

Canonical Polyadic Decomposition (CPD)

$$\underline{\mathbf{X}} = \sum_{f=1}^{F} \mathbf{a}_f \odot \mathbf{b}_f \odot \mathbf{c}_f$$

where $\odot$ is outerproduct operator, $\mathbf{a} \in \mathbb{F}^{I}$, $\mathbf{b} \in \mathbb{F}^{J}$, $\mathbf{c} \in \mathbb{F}^{K}$. Here F is the minimum number of outerproducts needed to reconstruct $\underline{\mathbf{X}}$. CPD elementwise representation is

$$\underline{\mathbf{X}}(i,j,k) = \sum_{f=1}^{F} \mathbf{A}(i,f)\mathbf{B}(j,f)\mathbf{C}(k,f)$$

where $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2 \dots \mathbf{a}_F] \in \mathbb{F}^{I \times F}$, $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2 \dots \mathbf{b}_F] \in \mathbb{F}^{J \times F}$, $\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2 \dots \mathbf{c}_F] \in \mathbb{F}^{K \times F}$

Theorem 1

Uniqueness of $\underline{\mathbf{X}} = [ABC]$ decomposition is true when $I \geq J \geq K$ and $F \leq 2^{\lfloor \log_2 J \rfloor + \lfloor \log_2 K \rfloor - 2}$ almost surely.

Theorem 2

The decomposition $\underline{\mathbf{X}} = [ABC]$ is essentially unique with CP rank F if $k_A + k_B + k_C \geq 2F + 2$. where k_A is kruskal rank of $\mathbf{A}$

Main Idea

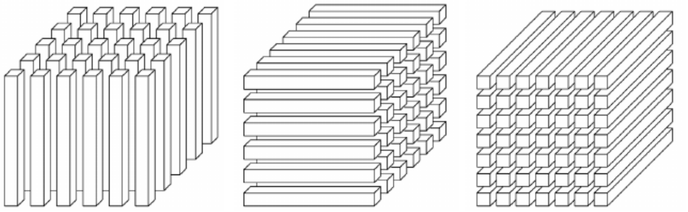
Uniqueness of CPD and relation between subsampled tensor and original tensor.

$$\underline{X}\left(\mathcal{S}_r, \mathcal{S}_c, \mathcal{S}_f\right) = \left[\boldsymbol{A}\left(\mathcal{S}_r, :\right), \boldsymbol{B}\left(\mathcal{S}_c, :\right), \boldsymbol{C}\left(\mathcal{S}_f, :\right)\right]$$

where $\mathcal{S}_r \subseteq \{1, \ldots, I\}$ rows, $\mathcal{S}_r \subseteq \{1, \ldots, J\}$ columns, $\mathcal{S}_r \subseteq \{1, \ldots, K\}$ fibers. One key observation is that the above sub-tensor can be decomposed to a sum of rank one terms of number equal to the rank of the original tensor.

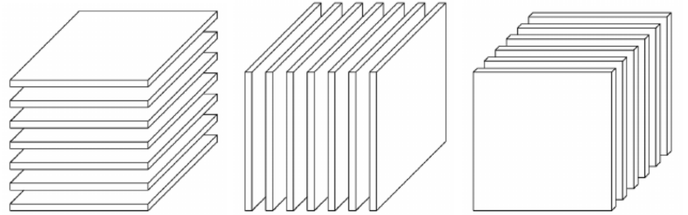
Results

- Introduced slab sampling, fiber sampling and entry sampling.
- Similar to matrix completion, the sample complexity for tensor signal reconstruction is mainly affected by the tensor rank and the tensor size, instead of signal bandwidth or sparsity. Unlike CS and LRMC, the proposed approach does not require incoherent sampling.
- Designing accelerated acquisition schemes for functional magnetic resonance imaging (fMRI) utilizing the proposed tensor sampling principles.



(a) Mode-1 (column) fibers: $\mathbf{x}_{:jk}$ (b) Mode-2 (row) fibers: $\mathbf{x}_{i:k}$ (c) Mode-3 (tube) fibers: $\mathbf{x}_{ij:}$

Fig. 2.1 *Fibers of a 3rd-order tensor.*



(a) Horizontal slices: $\mathbf{X}_{i::}$ (b) Lateral slices: $\mathbf{X}_{:j:}$ (c) Frontal slices: $\mathbf{X}_{::k}$ (or $\mathbf{X}_k$)

Figure: Tensor sampling

Sparse Bayesian Learning With Dynamic Filtering for Inference of Time-Varying Sparse Signals by Matthew R. OShaughnessy, Mark A. Davenport, and Christopher J. Rozell

Contributions

- Estimating time varying sparse signals using SBL.
- Key insight is that the estimate of $\mathbf{x}^t$ can be improved in a robust manner by injecting information from the estimate of the previous time step and a dynamics model into the hyperparameters of the SBL probability model.

System Model

$$\mathbf{y}^t = \mathbf{\Phi}\mathbf{x}^t + \mathbf{e}^t$$
$$\mathbf{x}^{t+1} = f_t(\mathbf{x}^t) + \mathbf{n}^t$$

where $\mathbf{y}^t, \mathbf{e}^t \in \mathbb{R}^M$, $\mathbf{x}^t \in \mathbb{R}^N$, $\mathbf{\Phi} \in \mathbb{R}^{M\times N}$ and $f_t(\cdot) : \mathbb{R}^N \to \mathbb{R}^N$. Here $M \ll N$ and $||\mathbf{x}^t||_0 \leq K < M$

Solution under SBL framework

System is corrupted under i.i.d gaussian noise $\mathbf{e}^t \sim \mathcal{N}(\mathbf{0}, \lambda\mathbf{I})$

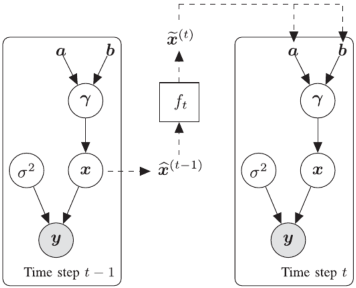
$$p(y|x,\lambda) = \mathcal{N}(\mathbf{\Phi}\mathbf{x}, \lambda\mathbf{I})$$
$$p(x_i|\gamma_i) = \mathcal{N}_{x_i}(0, \gamma_i)$$

$\mathbf{\Gamma} = diag(\gamma_i)$ is parameterised with $\{a_i, b_i\}_{i=1}^N$

SBL-DF

$$\widehat{\boldsymbol{x}}_{\mathrm{SBL}} = \mathbb{E}_{\boldsymbol{x}}[p(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\gamma},\lambda)] = \boldsymbol{\mu}$$
$$p(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\gamma},\lambda) = \mathcal{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$$

where $\boldsymbol{\Sigma} = \left(\boldsymbol{\Gamma}^{-1} + \lambda^{-1}\boldsymbol{\Phi}^T\boldsymbol{\Phi}\right)^{-1}$, $\boldsymbol{\mu} = \lambda^{-1}\boldsymbol{\Sigma}\boldsymbol{\Phi}^T\boldsymbol{y}$ and $\Gamma = diag\{\gamma_i\}$



$$\gamma_{\mathrm{dyn}} = \underset{\gamma}{\arg\min}\,\mathbb{E}\left[\|\widehat{\boldsymbol{x}} - \widetilde{\boldsymbol{x}}\|_2^2\right]$$

For simplification assume $\boldsymbol{\Phi}^T\boldsymbol{\Phi}$ is diagonal then $a_i = \xi$ and $b_i = \xi\widetilde{x}_i^2$
Here EM with pruning of μ and $\boldsymbol{\Sigma}$ followed by the $b_i = \xi\widetilde{x}_i^2$, where the parameter ξ represents how much weight the dynamics-based prediction is assigned in the evidence maximization procedure.

On the Convergence of a Bayesian Algorithm for Joint Dictionary Learning and Sparse Recovery by Geethu Joseph and Chandra R. Murthy

Contributions

- A novel algorithm for learning the sparsifying dictionary along with the sparse representation.
- Convergence guarantees of the dictionary update step using AM and ALS optimization procedures are derived. Stability of limit points irrespective of initializations are discussed.
- Convergence guarantees of entire algorithm (Dictionary update and sparse representation) is discussed.

Problem Statement

Recover both sparse vectors $\{\mathbf{x}_k\}_{k=1}^{K}$ and dictionary $\mathbf{A}$ from measurements $\{\mathbf{y}_k\}_{k=1}^{K}$

$$\mathbf{y}_k = \mathbf{A}\mathbf{x}_k + \mathbf{w}_k$$

where $\mathbf{y}^K = \{\mathbf{y}_k \in \mathbb{R}^m\}_{k=1}^{K}$, $\mathbf{x}^K = \{\mathbf{x}_k \in \mathbb{R}^N\}_{k=1}^{K}$, unknown $\mathbf{A} \in \mathbb{R}^{m\times N}$ with unit norm columns.
$\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$, $\mathbf{x}_k \sim \mathcal{N}(\mathbf{0}, diag(\gamma_k))$

$$\widehat{\mathbf{x}}_{k-SBL} = \mathbb{E}[\mathbf{x}_k|\mathbf{y}_k, \gamma_k, \widehat{\mathbf{A}}]$$

Final cost function is arrived by minimizing the negative log likelihood $-\log p(\mathbf{y}^K; \mathbf{\Lambda})$,
$\mathbf{\Lambda} = \{\boldsymbol{A}, \gamma_k; k = [K]\}$

$$T(\mathbf{\Lambda}) = \sum_{k=1}^{K} \log[|\sigma^2\mathbf{I} + \mathbf{A}\mathbf{\Lambda_k}\mathbf{A}^T|] + \mathbf{y}_k^T(\sigma^2\mathbf{I} + \mathbf{A}\mathbf{\Lambda_k}\mathbf{A}^T)^{-1}\mathbf{y}_k$$

- Sparse recovery using EM.
- Dictionary update is via Alternating Minimization or Armijo Line Search.

Other Interesting Papers

- A Block Sparsity Based Estimator for mmWave Massive MIMO Channels With Beam Squint.
- An Asymptotically Efficient Weighted Least Squares Estimator for Co-Array-Based DoA Estimation.
- A SpatialTemporal Subspace-Based Compressive Channel Estimation Technique in Unknown Interference MIMO Channels.
- Energy and Area-Efficient Recursive-Conjugate-Gradient-Based MMSE Detector for Massive MIMO Systems.
- LDA via L1-PCA of Whitened Data.
- Multi-Class Random Matrix Filtering for Adaptive Learning
- Nonlinear Filtering With Variable Bandwidth Exponential Kernels