Journal Watch
IEEE TSP Nov.1, 1, 2019

Unnikrishnan N

SPC Lab, Department of ECE
Indian Institute of Science

November 19, 2019

Low-complexity Graph Sampling With Noise and Signal Reconstruction via Neumann Series by Fen Wang, Gene Cheung, and Yongchao Wang

Contributions

- Sampling matrix is obtained by solving modified A-optimality criteria based on Neumann series
- Select nodes greedily without large matrix inversion using matrix inversion lemma.
- Dynamic subset sampling introduced –extended sampling case.

Setting

- $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, W\}$ a graph with N nodes indexed by $\mathcal{V} = \{1, 2, ...N\}$. $\mathcal{E}$ specifies the set of connected node pairs(i, j) and W is adjacency matrix.
- Connected, undirected graphs with no self-loops are assumed.
- Laplacian matrix $L = D - W$ and its eigen decomposition $L = V\Sigma V^T$ with eigen values in increasing order. Where D is diagonal degree matrix.
- The GFT of a graph signal $x \in \mathbf{R}^N$ is defined as $\tilde{x} = V^T x$.
- A K-bandlimited graph signal can be expressed as $x = V_K \tilde{x}_K$ where V_K is first K columns of V matrix and $\tilde{x}_K$ is first K elements of $\tilde{x}$ vector. K-Bandlimited low pass filter is defined as $T = V_K V_K^T$

Definition To sample M elements from x to produce $x_S = Cx \in \mathbb{R}^M$ with a sample set $S \subseteq \mathcal{V}$ and $|S| = M$, define a $M \times N$ 0-1 binary sampling matrix C as

$$\mathbf{C}_{ij} = \begin{cases} 1, & j = \mathcal{S}(i) \\ 0, & \text{otherwise} \end{cases}$$

A sampled K-bandlimited graph signal can now be written as $x_S = CV_K\tilde{x}_K$ with $\text{rank}(CV_K) = K$. Least square solution of a noisy observation

$$y_S = x_S + n_S$$
$$\hat{x}_{LS} = V_K(CV_K)^{\dagger} y_S$$
$$R_{\hat{x}_{LS}} = V_K[(CV_K)^T CV_K]^{-1} V_K$$

Problem formulation

A-optimality criteria

$$\mathbf{C}^* = \underset{\mathbf{C}\in\mathbb{F}^{M\times N}}{\arg\min}\ \text{Tr}\left(\left[(\mathbf{CV}_K)^{\top}\mathbf{CV}_K\right]^{-1}\right)$$

Augmented A-optimality criteria

$$\mathbf{C}^* = \underset{\mathbf{C}\in\mathbb{F}^{\mathrm{M}\times\mathrm{N}}}{\arg\min}\ \text{Tr}\left(\left[(\mathbf{CV}_K)^{\top}\mathbf{CV}_K + \mu\mathbf{I}\right]^{-1}\right)$$

Equivalent Augmented A-optimality

$$\mathcal{S}^* = \underset{S:|S|=M}{\arg\min}\ \text{Tr}\left(\mathbf{T}_{\mathcal{S}} + \mu\mathbf{I}\right)^{-1}$$

where $T_S = CV_K(CV_K)^T$. This T_S is a low pass filter that can be approximated with fast Graph fourier transform (truncated Jacobi algorithm).

$$\min_{\hat{\mathbf{\Lambda}},\mathbf{S}_1,\ldots,\mathbf{s}_J} \left\|\mathbf{L} - \mathbf{S}_1 \ldots \mathbf{S}_J\hat{\mathbf{\Lambda}}\mathbf{S}_J^{\top} \ldots \mathbf{S}_1^{\top}\right\|_F^2$$

where S_j are givens rotation matrices and $\hat{\mathbf{\Lambda}}$ is a diagonal matrix

Modified A-optimality problem

$$\mathcal{S}^* = \underset{\mathcal{S}:|\mathcal{S}|=M}{\arg\min}\ \mathrm{Tr}\left(\mathbf{T}_{\mathcal{S}}^{\mathrm{FGFT}} + \mu\mathbf{I}\right)^{-1}$$

where $\mathbf{T}_{\mathcal{S}}^{\mathrm{FGFT}}$ is computed using fast Graph fourier transform.

Other Points

- GFS Graph Signal Sampling Algorithm using matrix inversion lemma discussed.
- Static subset sampling problem is extended to dynamic case
- Reconstruction MSE Evaluation of graph signal for various algorithms were discussed

A Low-Complexity Nonlinear Least Mean Squares Filter Based on a Decomposable Volterra Model by Felipe Chaud Pinheiro and Cassio Guimaraes Lopes

Goal

- Decomposable volterra model is used for modelling non-linearity.
- Steepest descend and LMS formulations were done.
- Necessary condition for stability is derived and thus step size also.

Volterra Model

In a nonlinear system with input signal $u(i)$ and output signal $y(i)$, a Volterra series representation is given by

$$y(i) = y_0 + y_1(i) + y_2(i) + y_3(i) + \cdots$$

where $y0$ is a constant and

$$y_k(i) = \sum_{0 \le i_1, \ldots, i_k < M} \mathcal{H}_k(i_1, \ldots, i_k) u(i - i_1) \cdots u(i - i_k)$$

where $\mathcal{H}_k$ is volterra kernal and its order of complexity for computation is $O(M^K)$

Decomposable Volterra Model(DVM)

$$u_i^{\otimes k} = \underbrace{u_i \otimes \cdots \otimes u_i}_{k \text{ times}} \quad (1 \times M^k)$$

$$u_i \triangleq \begin{bmatrix} u(i) & u(i-1) & \ldots & u(i-1+M) \end{bmatrix} \quad (1 \times M)$$

$$y_k(i) = u_i^{\otimes k} h_k$$

$$h_K = w_1 \otimes \cdots \otimes w_K \quad (M^K \times 1)$$

where $w_1, \ldots, w_K$ are $M \times 1$ vectors.
The effect of the decomposability hypothesis is to make the output y(i) computable via the product of the output of K linear FIR filters(with overall complexity $O(KM)$), as follows:

$$\begin{aligned} y(i) = u_i^{\otimes K} h_K &= (\underbrace{u_i \otimes \cdots \otimes u_i}_{K \text{ times}})(w_1 \otimes \cdots \otimes w_K) \\ &= (u_i w_1) \otimes \cdots \otimes (u_i w_K) \\ &= (u_i w_1) \cdots (u_i w_K) \end{aligned}$$

Problem formulation

$$\min_w \mathbb{E} \left| \boldsymbol{d} - \boldsymbol{u}^{\otimes K} w \right|^2 \quad \text{s.t.} \quad w \text{ being decomposable}$$

$$J(w_1, \ldots, w_K) = R_d - w^* R^*_{u^K d} - R_{u^K d} w + w^* R_{u^K} w$$

$$\frac{\partial J}{\partial w_s} = \frac{\partial J}{\partial w} \cdot \frac{\partial w}{\partial w_s}$$

$$\frac{\partial J}{\partial w} = -R_{u^K d} + w^* R_{u^K} \quad (1 \times M^K)$$

For each $1 \leq s < K$, the factor $W^s = \frac{\partial w}{\partial w_s}$ is an $M^k \times M$ Jacobian matrix.

Steepest Descend Case

$$w_{s,i} = w_{s,i-1} - \mu \left[\frac{\partial J}{\partial w_s} \Big|_{i-1} \right]^*$$
$$= w_{s,i-1} + \mu W_{i-1}^{(s)*} [R_{du^K} - R_{u^K} w_{i-1}]$$

For each $1 \leq s < K$

LMS

$$\widetilde{R}_{u^K,i} = u_i^{\otimes K *} u_i^{\otimes K}, \quad \widetilde{R}_{du^K,i} = u_i^{\otimes K *} d(i)$$
$$w_{s,i} = w_{s,i-1} + \mu W_{i-1}^{(s)*} [u_i^{\otimes K *} d(i) - u_i^{\otimes K *} u_i^{\otimes K} w_{i-1}]$$
$$= w_{s,i-1} + \mu \left(u_i^{\otimes K} W_{i-1}^{(s)} \right)^* [d(i) - u_i^{\otimes K} w_{i-1}]$$

Other points

- Initializations of w discussed.
- Convergence and stability discussed.
- Detailed steady state analysis done.

Non-Negative Orthogonal Greedy Algorithms by Thanh Thi Nguyen, Jrme Idier, Charles Soussen, and El-Hadi Djermoune

Contributions

- Define a class of non-negative orthogonal greedy algorithms with their structural properties.
- Propose a fast and exact implementation based on active set resolution of non-negative LS.

Problem formulation

$$\min_{\boldsymbol{x}\geq\boldsymbol{0}} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|_2^2 \quad \text{s.t.} \quad \|\boldsymbol{x}\|_0 \leq K. \qquad (\ell_0+)$$

Consider the NNLS problem related to support S

$$\min_{\boldsymbol{x}\geq\boldsymbol{0}} \|\boldsymbol{y}-\boldsymbol{H}\boldsymbol{x}\|^2 \quad \text{s.t.} \quad \operatorname{supp}(\boldsymbol{x}) \subset S.$$

$\hat{x}_S^+$ is a solution if and only if the KKT conditions are satisfied:

$$\begin{cases} H_C^{\mathrm{t}}(\boldsymbol{y}-\boldsymbol{H}\hat{\boldsymbol{x}}_S^+) & = \boldsymbol{0} \\ H_{S\setminus C}^{\mathrm{t}}(\boldsymbol{y}-\boldsymbol{H}\hat{\boldsymbol{x}}_S^+) & \leq \boldsymbol{0} \end{cases}$$

where $C := \operatorname{supp}\left(\hat{\boldsymbol{x}}_S^+\right) \subset S$ call $\boldsymbol{r}_S^+ = \boldsymbol{y}-\boldsymbol{H}\hat{\boldsymbol{x}}_S^+$

General Structure of NNOGA

Algorithm 1: General Structure of a Non-negative Orthogonal Greedy Algorithm to Solve (ℓ_0+).

input : $\boldsymbol{y}, H, K$
output: $\boldsymbol{x}$

$\boldsymbol{x} \leftarrow \mathbf{0}$; $S \leftarrow \emptyset$; $\boldsymbol{r}_S \leftarrow \boldsymbol{y}$;
while $|S| < K$ *and* $\max_{i \in \bar{S}} \boldsymbol{h}_i^{\mathrm{t}} \boldsymbol{r}_S > 0$ **do**
  Select an index $\ell \in \bar{S}$ by a selection rule $\mathcal{S}(\boldsymbol{y}, H, S)$;
  $S \leftarrow S \cup \{\ell\}$;
  $\boldsymbol{x} \leftarrow \hat{\boldsymbol{x}}_S^+$;
  $S \leftarrow \mathrm{supp}(\boldsymbol{x})$;
  $\boldsymbol{r}_S = \boldsymbol{y} - H\boldsymbol{x}$;
end

Greedy algorithms can be interpreted as descent algorithms dedicated to the minimization of the residual norm using supports of growing size. Contrary to the unconstrained case, the selection of only some atoms in $\bar{S}$ may produce a decrease of the residual at a given iteration of an NNOG algorithm.

Definition 1 For a given support S, define the set of indices corresponding to descending atoms as follows:

$$\mathcal{D}_S = \left\{ i \in \{1, \ldots, n\}, \|\boldsymbol{r}^+_{S \cup \{i\}}\| < \|\boldsymbol{r}^+_S\| \right\}$$

where $\mathcal{D}_S \subset \bar{S}$

Definition 2 A descent selection rule is a function S(y, H, S) that takes its values in $\mathcal{D}_S$.

One possible descend selection rule is

$$\mathcal{S}_1(S) \in \arg\max_{i \notin S} \boldsymbol{h}_i^{\mathrm{t}} \boldsymbol{r}_S$$

Fast Computations

- Descend selection rule.
- Active set NNLS
- Warm start (other than zero vector start with ULS solution with positive entries)
- Descend rule can be parallisable.

Comparison with other non negative techniques such as NN-SP, NN-CoSAMP are done.

Other Interesting Papers

- Power Leakage Elimination for Wideband mmWave Massive MIMO-OFDM Systems: An Energy-Focusing WindowApproach
- Double Bayesian Smoothing as Message Passing .
- Efficient Equalization of Time-Varying Channels in MIMO OFDM Systems.
- Sparse Recovery and Dictionary Learning From Nonlinear Compressive Measurements