Robust least mean squares estimation of graph signals

Aalto University

Jari Miettinen, Sergiy Vorobyov and Esa Ollila

Department of Acoustics and Signal Processing, Aalto University, Finland

Motivation

Scope: Recovering graph signal from few sampled nodes.

Problem: The adjacency matrix $\mathbf{A}$ may not be accurately chosen.

Objective: Be more robust to errors in graph specification.

Data model

- The graph signal is sampled at M time points according to probabilistic sampling strategy. $\mathcal{S}[n]$ denotes the set of sampled nodes at time instant n and $\mathbf{D}_{\mathcal{S}[n]} = \mathrm{diag}(d_1[n], \dots, d_N[n])$, where $d_i[n] = 1$ if $i \in \mathcal{S}[n]$, and 0 otherwise. The model for observations:

$$\mathbf{y}[n] = \mathbf{D}_{\mathcal{S}[n]}(\mathbf{x} + \boldsymbol{\nu}[n]), \; n = 1, \dots, M,$$

  $\mathbf{x}$ is the graph signal of interest and $\boldsymbol{\nu}[n]$ is the measurement noise.
- We assume that $\mathbf{x}$ is bandlimited graph signal $\mathbf{x} = \mathbf{V}_{\mathcal{F}}\mathbf{s}$, where $\mathbf{V}_{\mathcal{F}}$ is $N \times |\mathcal{F}|$ matrix containing the eigenvectors corresponding to $|\mathcal{F}|$ smallest eigenvalues of the graph Laplacian matrix $\mathbf{L}$.

Graph LMS

- The optimization problem to be solved:

$$\min_{\mathbf{s}} \mathbb{E}\|\mathbf{D}_{\mathcal{S}[n]}(\mathbf{y}[n] - \mathbf{V}_{\mathcal{F}}\mathbf{s})\|^2.$$

- The update step for estimating $\mathbf{x}$:

$$\hat{\mathbf{x}}[n+1] = \hat{\mathbf{x}}[n] + \mu \mathbf{B}_{\mathcal{F}}\mathbf{D}_{\mathcal{S}[n]}(\mathbf{y}[n] - \hat{\mathbf{x}}[n]) \qquad (1)$$

  μ is step-size and $\mathbf{B}_{\mathcal{F}} = \mathbf{V}_{\mathcal{F}}\mathbf{V}_{\mathcal{F}}^H$.

Robust LMS estimation of graph signals

- Our method relies on that the LMS converges the faster, the better the matrix $\mathbf{B}_{\mathcal{F}}$ is related to the data.
- We use two criteria for the speed of convergence:

(a) Initial value $\hat{\mathbf{x}}_0[0] = \mathbf{0}$ and the aim is to maximize the energy

$$\|\hat{\mathbf{x}}_0[M]\|$$

of the estimate given by (1).

(b) Two random initial values $\hat{\mathbf{x}}_1[0]$ and $\hat{\mathbf{x}}_2[0]$ and the aim is to maximize the correlation

$$\mathrm{cor}(\hat{\mathbf{x}}_1[M], \hat{\mathbf{x}}_2[M]).$$

- We propose greedy algorithm to maximize above criteria over $\mathbf{A}$.
- To draw new search directions from the previous estimate of $\mathbf{A}$ we use the following model

$$\mathbf{A}_{\epsilon_1,\epsilon_2} = \mathbf{A} - \boldsymbol{\Delta}_{\epsilon_1} \odot \mathbf{A} + \boldsymbol{\Delta}_{\epsilon_2} \odot (\mathbf{1}_{N\times N} - \mathbf{A}) \qquad (2)$$

  $\boldsymbol{\Delta}_\epsilon$ is a random $N \times N$ matrix satisfying $\mathbb{P}([\boldsymbol{\Delta}_\epsilon]_{ij} = 1) = \epsilon$ and $\mathbb{P}([\boldsymbol{\Delta}_\epsilon]_{ij} = 0) = 1 - \epsilon$, and ϵ_1 and ϵ_2 are probabilities of removing and adding an edge from/to $\mathbf{A}$.

Algorithm

Input: $\mathbf{y}[1], \dots, \mathbf{y}[M]$, $\mathbf{D}_{\mathcal{S}[1]}, \dots, \mathbf{D}_{\mathcal{S}[M]}$, adjacency matrix estimate $\hat{\mathbf{A}}$, step size $\tilde{\epsilon}_1$, K

Initialisation : (a) $\hat{\mathbf{x}}_0[0] = \mathbf{0}$ (b) random initial values $\hat{\mathbf{x}}_1[0]$ and $\hat{\mathbf{x}}_2[0]$

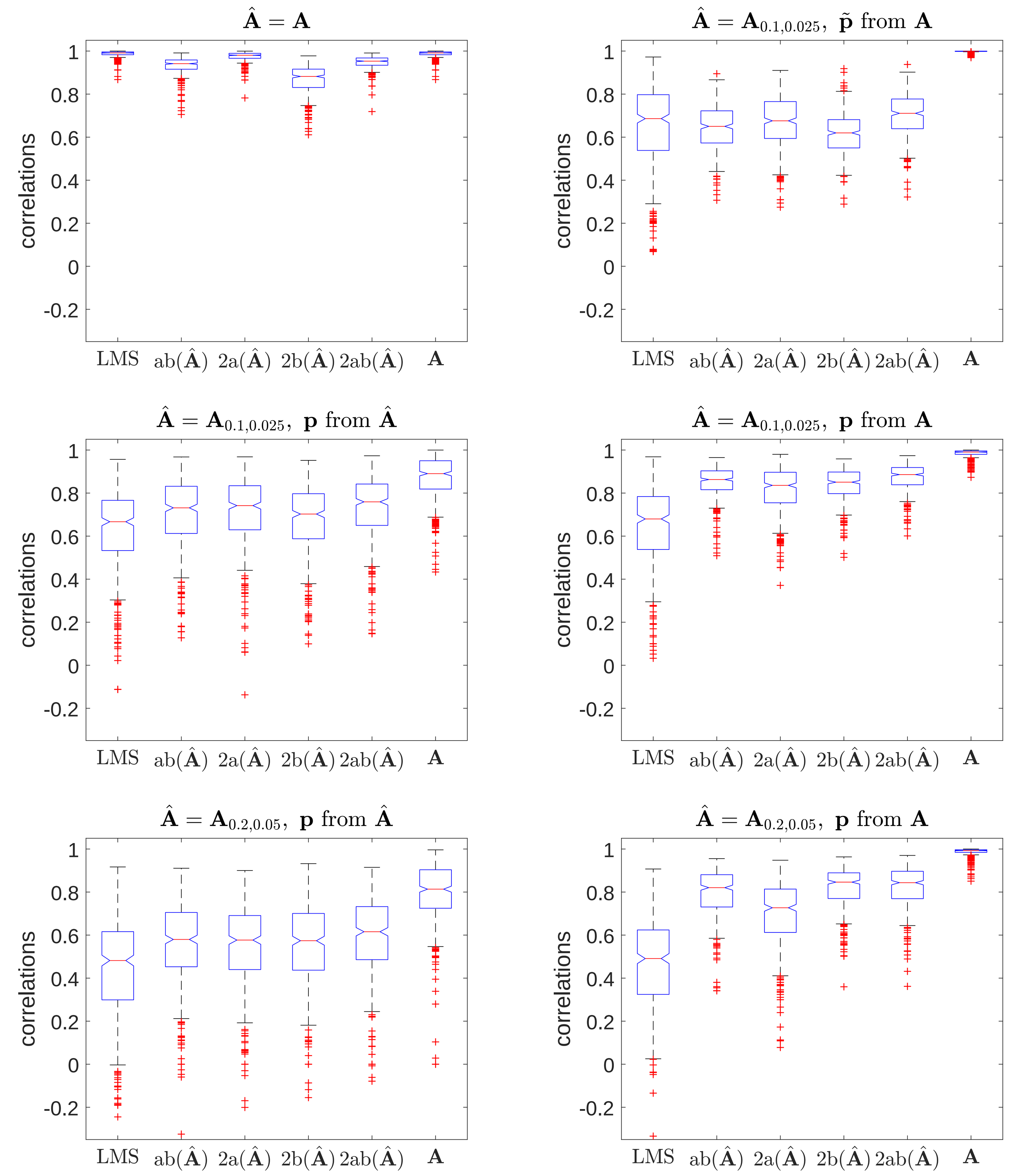
Repeat the following steps until $\hat{\mathbf{A}} = \hat{\mathbf{A}}_0$

1. $\hat{\mathbf{A}}_0 \leftarrow \hat{\mathbf{A}}$
2. Set $\tilde{\epsilon}_2^i = \frac{(i+1)\tilde{\epsilon}_1 w}{4(1-w)}$ for $i = 1, \dots, 5$, where $w = \frac{\|\hat{\mathbf{A}}\|_0}{N(N-1)}$
3. Draw a set of adjacency matrices $\hat{\mathbf{A}}_1, \dots, \hat{\mathbf{A}}_{5K}$ from the graph error model (2) around $\hat{\mathbf{A}}_0$ using $\tilde{\epsilon}_1$ and each of $\tilde{\epsilon}_2^i$ K times.
4. (a) Run (1) for $\mathbf{B}_{\mathcal{F}}^0, \mathbf{B}_{\mathcal{F}}^1, \dots, \mathbf{B}_{\mathcal{F}}^{5K}$ with initial value $\hat{\mathbf{x}}_0[0] = \mathbf{0}$.
   (b) Run (1) for $\mathbf{B}_{\mathcal{F}}^0, \mathbf{B}_{\mathcal{F}}^1, \dots, \mathbf{B}_{\mathcal{F}}^{5K}$ with initial values $\hat{\mathbf{x}}_1[0]$ and $\hat{\mathbf{x}}_2[0]$.
5. Find the adjacency matrix $\hat{\mathbf{A}}_{min}$ which yields largest value of
   (a) $\|\hat{\mathbf{x}}_0[M]\|$
   (b) $\mathrm{cor}(\hat{\mathbf{x}}_1[M], \hat{\mathbf{x}}_2[M])$
6. $\hat{\mathbf{A}} \leftarrow \hat{\mathbf{A}}_{min}$

Output: $\hat{\mathbf{x}}[M]$ given by LMS with $\hat{\mathbf{A}}$.

Simulation setup

- $N = 100$, $M = 50$, $|\mathcal{F}| = 5$.
- Noise covariance matrix is diagonal and the elements drawn uniformly on $[0.1, 0.2]$.
- $\mathbf{A}$ is Erdös–Rényi with probability parameter 0.1.
- The average sampling probability is $\mathbf{1}_N^\top \mathbf{p} = 0.125$ and about 80% of the probabilities are zeros, $\tilde{p} = \mathbf{1}_{\mathbf{p}>0}$.
- In Algorithm $\tilde{\epsilon}_1 = 0.01$ and $K = 4$.



Conclusion

- The effect of graph topology misspecification to graph signal recovery can be reduced by using the data to improve the interpolator matrix.
- Reconsidering only the edges related to sampled nodes is sufficient.
- The method would be more efficient if the sampled nodes were chosen based on the correct graph topology.

[1] P. Di Lorenzo, S. Barbarossa, P. Banelli, and S. Sardellitti. Adaptive least mean squares estimation of graph signals. *IEEE Trans. Signal Inf. Process. Netw.*, 2(4):555–568, 2016.

[2] P. Di Lorenzo, P. Banelli, E. Isufi, S. Barbarossa, and G. Leus. Adaptive graph signal processing: Algorithms and optimal sampling strategies. *IEEE Trans. Signal Process.*, 66(13):3584–3598, 2018.

[3] J. Miettinen, S. Vorobyov, and E. Ollila. Graph error effect in graph signal processing. In *ICASSP Calgary, Canada*, pages 4164–4168, 2018.