Adaptive Least Mean Squares Estimation of Graph Signals

ASP Preliminary Report (Anoushka Vyas - 20171057)

***Abstract*— The aim of this project is to propose a least mean squares (LMS) strategy for adaptive estimation of signals defined over graphs. Assuming the graph signal to be band-limited, over a known bandwidth, the method enables reconstruction, with guaranteed performance in terms of mean-square error, and tracking from a limited number of observations over a subset of vertices.**
Furthermore, to cope with the case where the bandwidth is not known beforehand, we propose a method that performs a sparse online estimation of the signal support in the (graph) frequency domain, which enables online adaptation of the graph sampling strategy.

I. ASSUMPTIONS

The signal is initially assumed to be perfectly band-limited, i.e. its spectral content is different from zero only on a limited set of frequencies F and the non-zero frequency nodes are known.

II. LOCALIZATION OPERATORS

Given a subset of vertices $S \subseteq V$, we define a vertex-limiting operator as the diagonal matrix:

$$D_S = diag(1_S) \tag{1}$$

where 1_S is the set indicator vector, whose $i-th$ entry is equal to one, if $x_0 \in S$, or zero otherwise. Similarly, given a subset of frequency indices $F \subseteq V$, we introduce the filtering operator:

$$B_F = U\Sigma_F U^{\mathrm{H}} \tag{2}$$

where Σ_F is a diagonal matrix defined as:

$$\Sigma_F = diag(1_F) \tag{3}$$

III. MODEL

Let us consider a signal $x_0 \in C^N$ defined over the graph $G = (V, E)$. The signal is initially assumed to be perfectly band-limited, i.e. its spectral content is different from zero only on a limited set of frequencies F. Let us consider partial observations of signal x_0 , i.e. observations over only a subset of nodes. Denoting with S the sampling set (observation subset), the observed signal at time n can be expressed as:

$$y[n] = D(x_0 + v[n]) = DBx_0 + Dv[n] \tag{4}$$

where D is the vertex-limiting operator defined in (1) , which takes nonzero values only in the set S, and $v[n]$ is a zero-mean, additive noise with covariance matrix C_v . The second equality in (4) comes from the bandlimited assumption, i.e. $Bx_0 = x_0$, with B denoting the operator in (2) that projects onto the (known) frequency set F .

IV. OPTIMIZATION

Following an LMS approach , the optimal estimate for x_0 can be found as the vector that solves the following optimization problem:

$$\begin{aligned} &\min_{\boldsymbol{x}} \ \mathbf{E}\,\|\boldsymbol{y}[n] - \mathbf{D}\mathbf{B}\boldsymbol{x}\|^2 \\ &\text{s.t.} \quad \mathbf{B}\boldsymbol{x} = \boldsymbol{x}, \end{aligned}$$

V. SAMPLING STRATEGIES

The properties of the LMS algorithm strongly depend on the choice of the sampling set S, i.e. on the vertex limiting operator D. The sampling strategy must be carefully designed in order to:

- enable reconstruction of the signal;
- guarantee stability of the algorithm; and
- impose a desired mean-square error at convergence.

Both the number of samples and their location is fundamental for the performance of the algorithm. Three algorithms were used to find the samples to be taken for reconstruction algorithm:

- Minimum Mean square deviation
- Maximization of determinant
- Maximization of the minimum eigenvalue

A. Minimum Mean square deviation

The method iteratively selects the samples from the graph that lead to the largest reduction in terms of steady state MSD. The steady state MSD is given as:

$$\begin{aligned} \mathrm{MSD} &= \lim_{n\to\infty} \mathbf{E}\|\tilde{\boldsymbol{x}}[n]\|^2 = \lim_{n\to\infty} \mathbf{E}\|\tilde{\boldsymbol{s}}[n]\|^2 \\ &= \mu^2 \mathrm{vec}(\mathbf{G})^T (\mathbf{I} - \mathbf{Q})^{-1} \mathrm{vec}(\mathbf{I}). \end{aligned}$$

$$\begin{aligned} \mathbf{G} &= \mathbf{U}_{\mathcal{F}}^H \mathbf{D} \mathbf{C}_v \mathbf{D} \mathbf{U}_{\mathcal{F}} \\ \mathbf{Q} &= (\mathbf{I} - \mu\, \mathbf{U}_{\mathcal{F}}^H \mathbf{D} \mathbf{U}_{\mathcal{F}}) \otimes (\mathbf{I} - \mu\, \mathbf{U}_{\mathcal{F}}^H \mathbf{D} \mathbf{U}_{\mathcal{F}}). \end{aligned}$$

where $\widetilde{s[n]}$ is the GFT of $x[n] - x_0$ and U_F the matrix having as columns the eigenvectors of the Laplacian matrix associated to the frequency indices F.

Sampling strategy 1: Minimization of MSD

Input Data : M, the number of samples.
Output Data : $\mathcal{S}$, the sampling set.
Function : initialize $\mathcal{S} \equiv \emptyset$
while $|\mathcal{S}| < M$
$s = \arg\min_j \ \mathrm{vec}(\mathbf{G}(\mathbf{D}_{\mathcal{S}\cup\{j\}}))^T (\mathbf{I} - \mathbf{Q}(\mathbf{D}_{\mathcal{S}\cup\{j\}}))^{\dagger}\mathrm{vec}(\mathbf{I})$;
$\mathcal{S} \leftarrow \mathcal{S} \cup \{s\}$;
end

B. *Maximization of determinant*

The rationale underlying this strategy is to design a well suited basis for the graph signal that we want to estimate. This criterion coincides with the maximization of the the pseudo-determinant of the matrix $U_F{}^{\mathrm{H}}DU_F$ i.e. the product of all non zeros eigenvalues. The proof is as follows:

$$\begin{aligned}\mathrm{MSD} &= \mu^2 \mathrm{vec}(\mathbf{G})^T(\mathbf{I}-\mathbf{Q})^{-1}\mathrm{vec}(\mathbf{I})\\ &= \mu^2 \mathrm{vec}(\mathbf{G})^T \mathbf{V}(\mathbf{I}-\mathbf{\Lambda})^{-1}\mathbf{V}^H \mathrm{vec}(\mathbf{I})\\ &= \mu^2 \sum_{i=1}^{|\mathcal{F}|^2} \frac{p_i \cdot q_i}{1-\lambda_i(\mathbf{Q})}\end{aligned}$$

$$\boldsymbol{p} = \{p_i\} = \mathbf{V}^H \mathrm{vec}(\mathbf{G}),\ \boldsymbol{q} = \{q_i\} = \mathbf{V}^H \mathrm{vec}(\mathbf{I}).$$

$$\lambda_i(\mathbf{Q}) = \left(1-\mu\lambda_k(\mathbf{U}_{\mathcal{F}}^H\mathbf{D}\mathbf{U}_{\mathcal{F}})\right)\left(1-\mu\lambda_l(\mathbf{U}_{\mathcal{F}}^H\mathbf{D}\mathbf{U}_{\mathcal{F}})\right)$$

Sampling strategy 2: Maximization of $\left|\mathbf{U}_{\mathcal{F}}^H\mathbf{D}\mathbf{U}_{\mathcal{F}}\right|_+$

Input Data : M, the number of samples.
Output Data : $\mathcal{S}$, the sampling set.
Function : initialize $\mathcal{S} \equiv \emptyset$
while $|\mathcal{S}| < M$
$s = \arg\max_j \ \left|\mathbf{U}_{\mathcal{F}}^H\mathbf{D}_{\mathcal{S}\cup\{j\}}\mathbf{U}_{\mathcal{F}}\right|_+$;
$\mathcal{S} \leftarrow \mathcal{S} \cup \{s\}$;
end

For minimum MSD we want theese eigenvalues to be as far from 1 as possible. Thus, the eigenvalues of matrix $U_F{}^{\mathrm{H}}DU_F$ has to be as far away from zero as possible. Thus, it equivalent to maximizing the determinant.

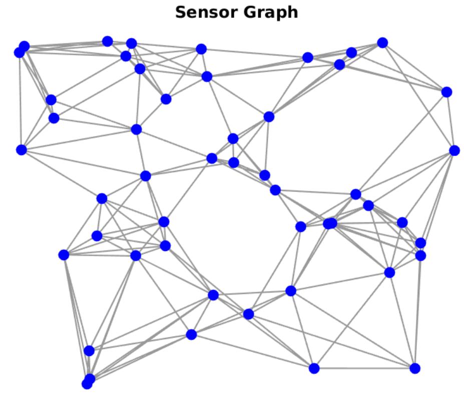
C. *Maximization of the minimum eigenvalue*

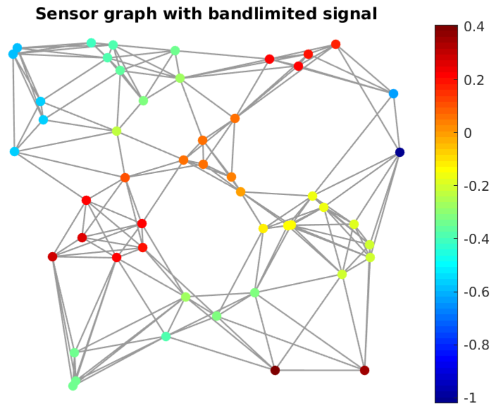
Sampling strategy 3: Maximization of $\lambda_{\min}^{+}\left(\mathbf{U}_{\mathcal{F}}^H\mathbf{D}\mathbf{U}_{\mathcal{F}}\right)$

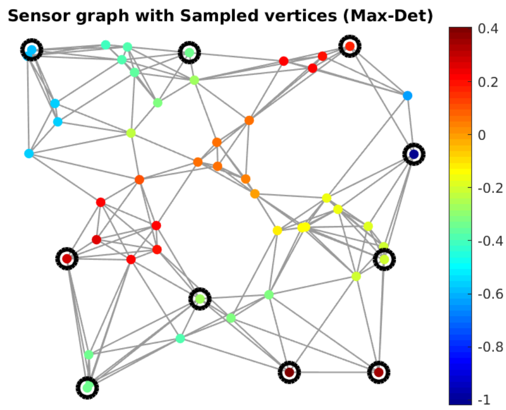
Input Data : M, the number of samples.
Output Data : $\mathcal{S}$, the sampling set.
Function : initialize $\mathcal{S} \equiv \emptyset$
while $|\mathcal{S}| < M$
$s = \arg\max_j \ \lambda_{\min}^{+}\left(\mathbf{U}_{\mathcal{F}}^H\mathbf{D}_{\mathcal{S}\cup\{j\}}\mathbf{U}_{\mathcal{F}}\right)$;
$\mathcal{S} \leftarrow \mathcal{S} \cup \{s\}$;
end

VI. RESULTS

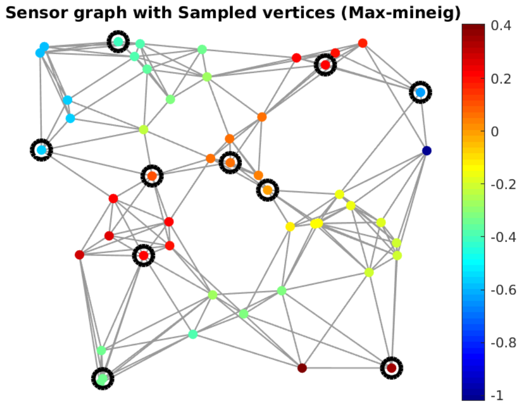
Number of nodes is 5, Bandwidth is 10 and learning rate is 0.5



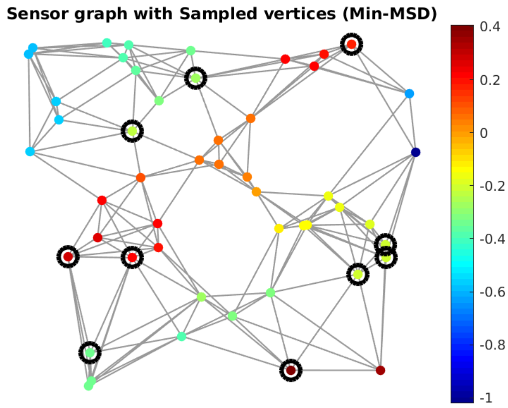





```
Sampled set of vertices for Max-Det Algorithm
    46
    36
    23
    50
    26
     3
     6
    43
     7
    48
```




```
Sampled set of vertices for Max-mineig Algorithm
    28
     9
    49
    12
    39
     2
    32
    18
    15
    46
```




```
Sampled set of vertices for Min-MSD Algorithm
    16
    47
    15
    36
    43
    24
    44
     6
     8
    48
```

VII. REMAINING TASKS

- Implementing the LMS algorithm after finding the desired D matrix.

Algorithm 1: LMS algorithm for graph signals

Start with $\boldsymbol{x}[0] \in \mathcal{B}_{\mathcal{F}}$ chosen at random. Given a sufficiently small step-size $\mu > 0$, for each time $n > 0$, repeat:

$$\boldsymbol{x}[n+1] = \boldsymbol{x}[n] + \mu\, \mathbf{BD}\left(\boldsymbol{y}[n] - \boldsymbol{x}[n]\right) \tag{12}$$

where the step size is:

$$0 < \mu < \frac{2}{\lambda_{\max}\left(\mathbf{U}_{\mathcal{F}}^{H}\mathbf{D}\mathbf{U}_{\mathcal{F}}\right)},$$

- Observation about convergence in mean square sense of the whole signal and the mean square sense node wise.
- Further tasks include situations where we consider the graph is fixed, and the spectral content of the signal can vary over time in an unknown manner and then trying to estimate the signal.