

EE 771 : Recent Topics in Analytical Signal Processing

Assignment 4

Arka Sadhu - 140070011

April 26, 2018

1 Q1

Given

$$g(x) = \sum_{k=-M}^{k=M} a[k] \exp(j2\pi kx)$$

We can write $y = [g(t_1), g(t_2) \dots g(t_n)]^T$ as $y = Ax$ where $x = [a[-M], \dots, a[M]]^T$. Here matrix A is a $n \times 2M + 1$ matrix and its i^{th} row and j^{th} column is given by $A_{i,j} = \exp(j2\pi(-M+j-1)t_i)$.

It is visible that A is full rank matrix for any $n \leq 2M + 1$ since the matrix A can be shown to be equivalent to a vandermonde matrix with some minor manipulations. Suppose that x is s -sparse vector, and we take $2s$ measurements, this gives us A of dimensions $2s \times 2M + 1$. If we assume that $2s \leq 2M + 1$ then we have A to be full rank with rank $2s$.

Assume that we have two different solutions: x_1 and x_2 both of which are s -sparse. This gives us $y = Ax_1$ and $y = Ax_2$ and therefore $A(x_1 - x_2) = 0$. Clearly, $x_1 - x_2$ can be at most $2s$ sparse vector. But rank of A is more than $2s$, and therefore we must have $x_1 - x_2 = 0$ as the only solution. That is $x_1 = x_2$ which implies A results in unique x if $2s$ measurements are done.

2 Q2

Given

$$g(x) = \sum_{k=0}^M a_k x^k$$

and $x \in [0, 1]$

In this case again we will have $y = Ax$ where $x = [a[0] \dots a[M]]^T$ and A will be of dimensions $n \times M + 1$. Its i^{th} row and j^{th} columns are given as: $A_{i,j} = t_i^{j-1}$.

3 Q3

The main difference between L1-minimization and OMP is that the former is independent of the sparsity. This is evident from the equation as well where the only constraint is because of the regularization which again is not a hard constraint.

In practice what is found is (reference: Simon Book Pg 73), OMP is much faster for small values of s . In such cases the L1-minimization may take more time as it is not able to utilize the hard constraints on the sparsity. On the other hand, if the sparsity is mildly larger, L1-minimization may beat the OMP algorithm in terms of computation time.

4 Q4

We are given the definition of coherence as:

$$\mu_1(s) = \max_{i \in [N]} \max \left\{ \sum_{k \in S} |\langle a_i, a_k \rangle|, S \subset [N], \text{card}(S) = s, i \notin S \right\}$$

Need to show that for any s-sparse vector x :

$$(1 - \mu_1(s - 1))\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \mu_1(s - 1))\|x\|_2^2$$

This is equivalent to showing for $\|x\|_2^2 = 1$:

$$(1 - \mu_1(s - 1)) \leq \|Ax\|_2^2 \leq (1 + \mu_1(s - 1))$$

So for the remainder of the proof assume $\|x\|_2^2 = 1$.

We define

$$\mu_{1,i} = \max\left\{\sum_{k \in S} |\langle a_i, a_k \rangle|, S \subset [N], \text{card}(S) = s, i \notin S\right\}$$

Thus

$$\mu_1 = \max_{i \in [N]} \mu_{1,i}$$

We first note:

$$\|Ax\|_2^2 = \langle A_s x_s, A_s x_s \rangle = x_s^T A_s^T A_s x_s = x_s^T \hat{A}_s x_s$$

Here $\hat{A}_s = A_s^T A_s$. Also if A_s and x_s are not real, then transpose is replaced by conjugate transpose everywhere. Denote the result by F . We get:

$$F = x_s^T \hat{A}_s x_s = \sum_{i \in S} \sum_{j \in S} x_i \hat{A}_{i,j} x_j$$

We further note that $\hat{A}_{i,j} = |\langle a_i, a_j \rangle|$ which implies $\hat{A}_{i,i} = 1$. This allows us to rewrite F as:

$$F = \sum_{i \in S} x_i^2 + \sum_{i \in S} \sum_{j \in S-i} x_i |\langle a_i, a_j \rangle| x_j$$

$$F = 1 + \sum_{i \in S} \sum_{j \in S-i} x_i |\langle a_i, a_j \rangle| x_j$$

We observe that

$$\sum_{j \in S-i} |\langle a_i, a_j \rangle| x_j \leq \mu_{1,i}(s - 1)$$

as the equality is obtained only if we have the x_j corresponding to the maximum value of $|\langle a_i, a_j \rangle|$ as 1, else LHS is always less than RHS. This gives us:

$$F \leq 1 + \sum_{i \in S} \mu_{1,i}(s - 1)x_i$$

Again, by similar observation as before we get:

$$\sum_{i \in S} \mu_{1,i}(s - 1)x_i \leq \mu_1(s - 1)$$

Thus we get

$$F \leq 1 + \mu_1(s - 1)$$

For lower bound instead of having the corresponding coefficient x_j and x_i as 1 we have them set to -1. This results instead

$$F \geq 1 - \mu_1(s - 1)$$

Hence we have proved:

$$1 - \mu_1(s - 1) \leq \|Ax\|_2^2 \leq 1 + \mu_1(s - 1)$$

5 Q5

Coherence of matrix A is defined as:

$$\mu(A) = \max_{1 \leq i \neq j \leq N} |\langle a_i, a_j \rangle|$$

We know that a DFT matrix is orthogonal matrix. Therefore for any $i \neq j$ we have $|\langle a_i, a_j \rangle| = 0$. Therefore the coherence is also 0.

6 Q6

We have (t_k, y_k) measurements for $k = 1, 2, \dots, m$. Also it is known that the function $f(t)$ is sparse in the dictionary $\phi_1(t), \dots, \phi_N(t)$. Thus it is prudent to consider:

$$f(t) = \sum_{k=1}^N a_k \phi_k(t)$$

We also have $y = f(t)$. Combining the two we get:

$$y = \Phi x$$
$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$
$$x = \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix}$$

Φ is of dimensions $m \times N$. $\Phi_{i,j} = \phi_j(t_i)$