**Mathematical Methods**

**For**

**Data Science and Signal Processing**

**2022-2023**

**Assignment 3**

Or Trabelsi

ID1: 301613501

**Question 1**

Proposition

If is s-sparse and , then is the unique solution to optimization problem for .

Proof

First, let us recall the definition of the spark of a matrix.

The spark of a matrix is the smallest number of columns that are linearly dependent.

This means that if we have a matrix with , then no subset of columns of is linearly dependent.

Suppose is s-sparse and . We want to show that is the unique solution to the optimization problem for .

Since is s-sparse and , is a feasible solution to the optimization problem.

Suppose for the sake of contradiction that there exists another feasible solution such that and is s-sparse.

Then , which implies that is a linear combination of the columns of .

However, since , any set of columns of is linearly independent,   
so must have at most non-zero components.

Therefore, , and is the unique solution to the optimization problem, which completes the proof. ∎

**Question 2**

Expectation Maximization Algorithm

EM is an iterative algorithm that aims to find the marginalized maximum likelihood estimator and is used ubiquitously in many statistical models.

For the MRA model

and under the assumption that the translations are drawn from the uniform distribution, this algorithm takes a simple form and consists of two steps at each iteration.

**The E-step**

Given a current estimation , the first step (called the ***E-step***) computes a set of weights which can be understood as the translation distribution of each measurement , if was the underlying signal.

These weights are computed by

Where, is the normalization factor s.t. .

**The M-step**

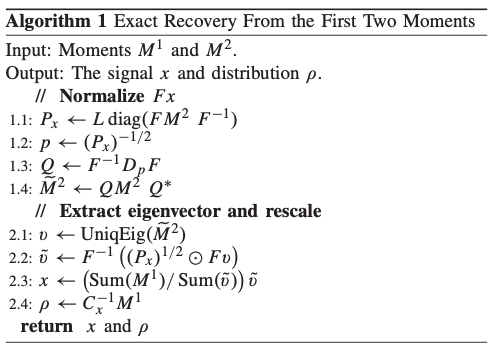
Then, the signal estimation is updated by marginalizing over the distributions and averaging (called the **M-step**).

Method of Moments Algorithm

This method is a spectral algorithm to estimate the signal, up to cyclic translation, from the first and second moments of the data which are given by

|  |  |
| --- | --- |
|  |  |
|  |  |

Based on those two moments, we can recover the signal by the following algorithm which is taken from the reference paper.



The Estimation Error

Our goal is to determine the sample complexity of , which we define to be the minimal number of measurements, as a function of the , required such that there is a sequence of estimators of with mean square error (MSE) converging to 0 as diverges.

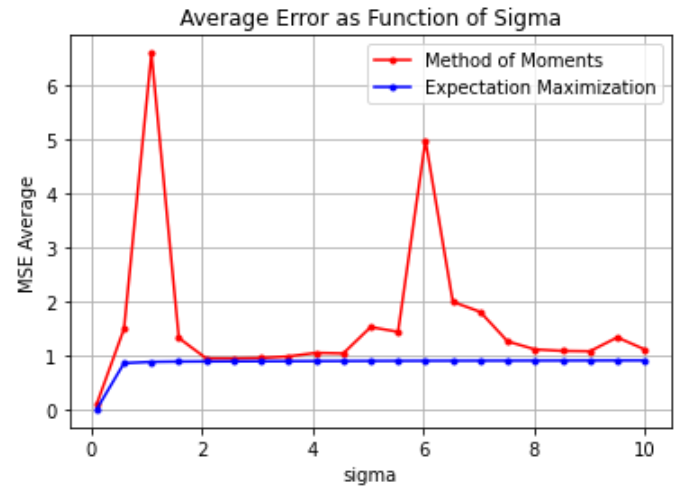
Since we are required to plot the average estimation error over the repeats, we also took the average of all the experiments’ , such that the final form is given by the following.

**Results**

Section A

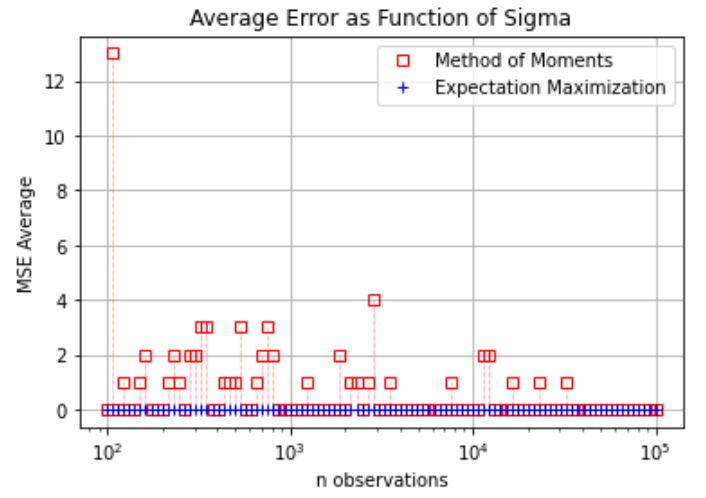
For random shifted observations with gaussian noise as mentioned above.

We estimated the signal by EM algorithm and the method of moments, and got the following results for the average estimation error over 10 experiments for several values between and . The averaged MSE error along all the experiments shows that for (approx.) the error is no longer increasing which means that for large we get very good estimation by the expectation maximization algorithm, while the method of moments required larger values for .   
Since we can also say that the results met the assumption \ claim in the paper that those 2 estimators will be relevant for small values.



Section B

Here we generated observations for a different range from to and estimated by both of the methods as that mentioned in , but now for fixed and repeat for each observations. As seen in the previous section, we can see that for this value of the EM error is equal to zero and the method of moments method is decreasing as the number of observations increase.



It was helpful to plot the data also for , which make it stands-out that also the EM algorithm is consistent s.t. for large number of observations the estimator will converge to the ground truth value.

|  |  |
| --- | --- |
| **Logarithmic scaling on Y-Axis** | **Without scaling on Y-Axis** |
|  |  |

**Reference**

This work is based on the following paper ([link](https://ieeexplore.ieee.org/abstract/document/8590822))

E. Abbe, T. Bendory, W. Leeb, J. M. Pereira, N. Sharon and A. Singer, "Multireference Alignments Easier With an Aperiodic Translation Distribution, in IEEE Transactions on Information Theory, vol. 65, no. 6, pp. 3565-3584, June 2019, doi: 10.1109/TIT.2018.2889674.