TEL AVIV UNIVERSITY

Mathematical Methods

For

Data Science and Signal Processing

2022-2023

Assignment 3

Or Trabelsi

ID1: 301613501

Question 1

Proposition

If x is s-sparse and $spark(\Phi) > 2s$, then x is the unique solution to l_0 optimization problem for $y = \Phi x$.

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 Φ with $spark(\Phi) > 2s$, then no subset of s columns of Φ is linearly dependent.

Suppose x is s-sparse and $spark(\Phi) > 2s$. We want to show that x is the unique solution to the l_0 optimization problem for $y = \Phi x$.

Since x is s-sparse and $y = \Phi x$, x is a feasible solution to the optimization problem.

Suppose for the sake of contradiction that there exists another feasible solution x' such that $y = \Phi x'$ and x' is s-sparse.

Then $\Phi(x - x') = 0$, which implies that $x - x'$ is a linear combination of the columns of Φ.

However, since $spark(\Phi) > 2s$, any set of $s + 1$ columns of Φ is linearly independent, so $x - x'$ must have at most s non-zero components.

Therefore, $x = x'$, and x is the unique solution to the l_0 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

$$y_i = R_{l_i} x + \epsilon_i, \qquad i = 1, \dots, n$$

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 x_{k-1}, the first step (called the ***E-step***) computes a set of weights which can be understood as the translation distribution of each measurement y_j, if x_{k-1} was the underlying signal.

These weights are computed by

$$w_k^{l,j} = C_k^j e^{-\frac{1}{2\sigma^2}\left\|R_l x_{k-1}+y_j\right\|_2^2}$$

Where, C_k^j is the normalization factor s.t. $\sum_l w_k^{l,j} = 1$.

The M-step

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

$$x_k = \frac{1}{N}\sum_{j=1}^{N}\sum_{l=0}^{L-1} w_k^{l,j} R_l^{-1} y_j$$

<u>Method of Moments Algorithm</u>

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

$$\widehat{M}^1 = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad y_i - \mathit{the\ i'th\ observation\ of\ the\ n\ observations}$$

$$\widehat{M}^2 = \frac{1}{n}\sum_{i=1}^{n} y_i y_i^T - \sigma^2 I \qquad I - L \times L\ \mathit{identity\ matrix}$$

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

Algorithm 1 Exact Recovery From the First Two Moments

Input: Moments M^1 and M^2.
Output: The signal x and distribution ρ.
// **Normalize** Fx
1.1: $P_x \leftarrow L\,\mathrm{diag}(FM^2\ F^{-1})$
1.2: $p \leftarrow (P_x)^{-1/2}$
1.3: $Q \leftarrow F^{-1} D_p F$
1.4: $\widetilde{M}^2 \leftarrow QM^2\ Q^*$
// **Extract eigenvector and rescale**
2.1: $v \leftarrow \mathrm{UniqEig}(\widetilde{M}^2)$
2.2: $\tilde{v} \leftarrow F^{-1}\left((P_x)^{1/2} \odot Fv\right)$
2.3: $x \leftarrow \left(\mathrm{Sum}(M^1)/\,\mathrm{Sum}(\tilde{v})\right)\tilde{v}$
2.4: $\rho \leftarrow C_x^{-1} M^1$
return x and ρ

The Estimation Error

Our goal is to determine the sample complexity of y_i, which we define to be the minimal number of measurements, as a function of the σ, required such that there is a sequence of estimators $\{\hat{x}_N\}$ of x with mean square error (MSE) converging to 0 as N diverges.

$$MSE_r = \frac{1}{\|x\|_2^2} \mathbb{E}\left\{\min_{s\in\mathbb{Z}_L} \|x - \hat{x}\|_2^2\right\}, \qquad for\ r = 1, \dots, R$$

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

$$MSE_{avg} = \frac{1}{\mathrm{R}} \sum_{r=1}^{\mathrm{R}} MSE_r$$

Results

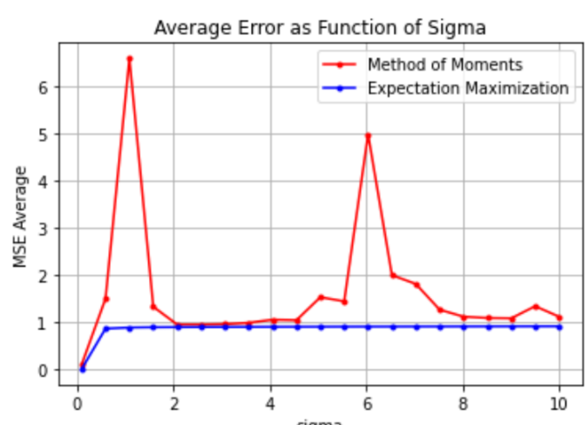
Section A

For 10^4 random shifted observations with gaussian noise as mentioned above.

$$y_i = R_{l_i} x + \epsilon_i, \qquad i = 1, \dots, n$$

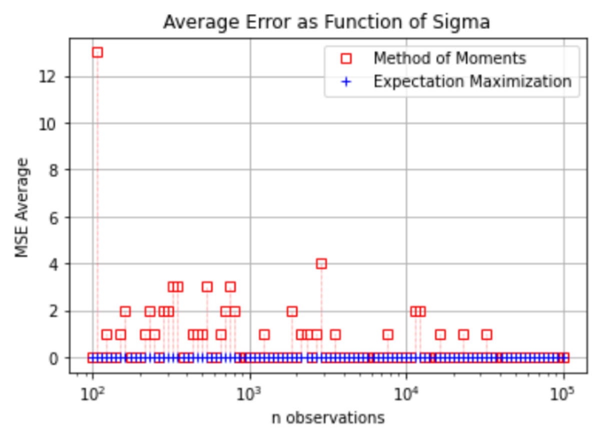
We estimated the signal x by EM algorithm and the method of moments, and got the following results for the average estimation error over 10 experiments for several σ values between 0.1 and 10. The averaged MSE error along all the experiments shows that for $\sigma \geq 0.8$ (approx.) the error is no longer increasing which means that for large σ we get very good estimation by the expectation maximization algorithm, while the method of moments required larger values for σ.

Since $\sigma \propto \frac{1}{SNR}$ we can also say that the results met the assumption \ claim in the paper that those 2 estimators will be relevant for small SNR values.

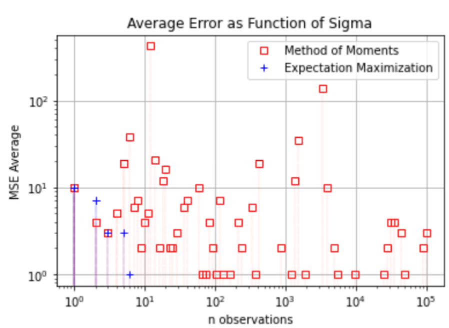
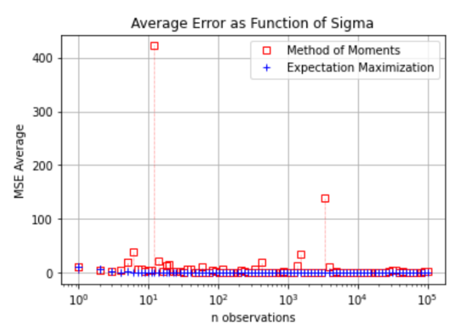


<u>Section B</u>

Here we generated observations for a different range from 10^2 to 10^5 and estimated x by both of the methods as that mentioned in A, but now for fixed $\sigma = 1$ and repeat $10\ times$ for each n observations. As seen in the previous section, we can see that for this value of σ 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 $n < 10^2$, 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.

<table>
<tr><th>Logarithmic scaling on Y-Axis</th><th>Without scaling on Y-Axis</th></tr>
<tr><td>
</td><td>
</td></tr>
</table>

Reference

This work is based on the following paper (link)

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.