Mathematical Methods in Data Science and Signal Processing

Homework Assignment 3

January 2, 2023

General instructions: Upload your solution text and code to Moodle via the dedicated submission box.

1. **Sparsity.** Prove the following proposition (see Sparsity and the parsimony principle presentation, slide 7). If x is s-sparse and $\text{spark}(\Phi) > 2s$, then x is the unique solution to ℓ_0 optimization problem for $y = \Phi x$.

2. **MRA.** Consider the following 1-D multi-reference alignment model. Suppose we acquire n observations drawn from the model

$$y_i = R_{\ell_i} x + \varepsilon_i, \quad i = 1, \ldots, n, \tag{1}$$

where

- $x \in \mathbb{R}^{15}$, a signal of length $L = 15$ with real entries, is a fixed signal we wish to estimate. In what follows, you can generate the signal by drawing random i.i.d. normal entries with zero mean and variance $\frac{1}{L}$, so that the expected norm is one.
- R_ℓ is a circular shift by ℓ entries. Namely, $(R_\ell x)[j] = x[(j-\ell) \bmod L]$ for $j = 0, \ldots, L-1$.
- The circular shifts are drawn from a generic (non-uniform) and unknown distribution ρ. In what follows, you can generate ρ by drawing L entries i.i.d. from a uniform distribution over $[0, 1]$, and normalizing their sum to 1, so that ρ will be a probability function.
- $\varepsilon_i \sim N(0, \sigma^2)$ are i.i.d.

We wish to estimate the signal x from $y_1, \ldots, y_n$, when ρ and the circular shifts $\ell_1, \ldots, \ell_n$ are unknown. What is the orbit of solutions? How would you define the estimation error?

Conduct the following experiments and display your results:

a) Generate $n = 10^4$ MRA observations following (1). Estimate the signal by expectation-maximization and by the method of moments. Calculate the estimation error obtained with each method. Run this experiment for several choices of σ between 0.1 and 10. For every fixed σ, repeat it 10 times. For all σ's and all repeats, you should use the same signal x and the same distribution ρ on the circular shifts. Plot the average estimation error (averaged over the repeats) as a function of σ.

b) Generate n MRA observations following (1) for n's ranging from $n = 10^2$ to $n = 10^5$ and a fixed $\sigma = 1$. Estimate the signal by expectation-maximization and by the method of moments and calculate the estimation error obtained by each method. Repeat the experiment 10 times for each fixed value of n. As before, use the same x and ρ for all n's and all repeats. Plot the average error (averaged over the repeats) as a function of n.

NOTE: The following paper might be a good starting point: https://ieeexplore.ieee.org/abstract/document/8590822/