Selective signal interference nullification and subspace reconstruction

Problem B: Spatial estimation and separation

Team: Bella y Sensual
Coach: Ignacio Monardes
Authors: Martín Andrighetti
Ignacio Muñoz
Vicente Opazo
Benjamín Rubio

Selective signal interference nullification and subspace reconstruction

1 Abstract

This article proposes a solution to problem B: Spatial estimation and separation, of the IMC Huawei challenges. The problem revolves around reconstructing a target subspace amidst interfering signals given access to a specific non-linear system in form of a oracle the solution can use. Leveraging mathematical formulations the solution we propose strategically combines outcomes from multiple oracle calls and gives a specific formulation on how to recover a basis for the solution with high probability. Practical considerations, such as tolerance for error and sample size determination, are also addressed. The proposed method achieves optimal results on the contest platform.

Keywords: Signal processing, Interference nullification, Subspace reconstruction, Mathematical modeling, Sampling.

2 Introduction

In wireless communication systems, spatial separation and estimation play pivotal roles in scenarios such as interference suppression and channel estimation. The ability to effectively separate target signals from interference signals is crucial for ensuring reliable communication in real world applications. This problem becomes more important every day with the exponential increase of wireless signals in use around the globe, a pattern that will continue for the foreseeable future.

In this article, we address a solution Problem B: Spatial estimation and separation, which revolves around spatial estimation and separation in the presence of multiple signals with varying spatial characteristics for a particular model of non-linear signal systems. We were given access to a platform to test solutions to the problem in some real world inspired datasets.

Specifically, the challenge describes a non-linear system with hidden parameters, such that its output combines the output of target and interference signals. Contestants are required to propose an algorithm that can access this system via oracle calls, the algorithm has to be able to recover a representation for the subspace generated by a certain set of parameters in the system. Additionally, additional parameters are present in the system that represent interference signals, the proposed algorithm has to be able to nullify the effect of the interference when computing the target subspace.

In this article we first do a full mathematical characterization of the problem and then describe the solution our team implemented during the contest. The solution exploits the particular characterization of the non-linear system in such a way that we are able to nullify the effect of the interference signals simply by considering certain combinations of input and output coordinates of the oracle. Finally, we address multiple practical considerations of the proposed solution. The proposed solution achieved a full score on the contest platform for problem B.

3 Glossary

- Given $n \in \mathbb{N}$, we define the set $[n] = \{1, \ldots, n\}$.
- Given $n \in \mathbb{N}$ We denote the i-th canonical vector by $\mathbf{e}_i^n$, that is $\mathbf{e}_i^n \in \mathbb{C}^n$ and

$$(\mathbf{e}_i^n)_j = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}$$

- Given complex vectors $\{\mathbf{v}_i\}_{i \in I}$ we define

$$\operatorname{span}\{\mathbf{v}_i : i \in I\} = \left\{ \mathbf{x} : \exists \{\alpha_i\}_{i \in I} \in \mathbb{C}^{|I|} \text{ such that } \sum_{i \in I} \alpha_i \mathbf{v}_i = \mathbf{x} \right\}.$$

- Given a matrix $\mathbf{X} \in \mathbb{C}^{m \times n}$ we define its Frobenius norm as

$$\|\mathbf{X}\|_F = \sqrt{\sum_{(i,j) \in [m] \times [n]} |\mathbf{X}_{i,j}|^2}.$$

- Given a matrix $\mathbf{X} \in \mathbb{C}^{m \times n}$ and naturals $i \leq j \in [m]$, $k \leq \ell \in [n]$ we define the sub-matrix $\mathbf{X}[i : j, k : \ell] \in \mathbb{C}^{(j-i+1) \times (\ell-k+1)}$ by $(\mathbf{X}[i : j, k : \ell])_{a,b} = \mathbf{X}_{a-i+1, b-k+1}$. Further, for simplicity we use $\mathbf{X}[:, k : \ell] = \mathbf{X}[1 : m, k : \ell]$, $\mathbf{X}[i : j, :] = \mathbf{X}[i : j, 1 : n]$, $\mathbf{X}[i, k : \ell] = \mathbf{X}[i : i, k : \ell]$ and $\mathbf{X}[i : j, k] = \mathbf{X}[i : j, k : k]$.

4 Problem Description

The challenge requires to do parameter estimation in a particular signal system model. We start by doing a full mathematical formulation of the problem that will be used in the rest of the article.

The entire problem can be formally reduced to the following. There exists unknown parameters $N \in \mathbb{N}$, $\mathbf{H}_1^{(t)} \in \mathbb{C}^{1\times 32}, \mathbf{H}_2^{(t)} \in \mathbb{C}^{1\times 32}, \mathbf{H}_3^{(t)} \in \mathbb{C}^{32\times 1}$ for all $t \in [N]$; $\mathbf{S}^{(1)}, \mathbf{S}^{(2)} \in \mathbb{C}^{32\times 300}$ and $c_j^{(t)} \in \mathbb{C}$ for all $j \in [300]$ and $t \in [N]$. Each of the values of $t \in [N]$ represent different signals in the system. We are given access to a function $\mathbf{Y}$ that receives 2 matrices $\mathbf{W}_1 \in \mathbb{C}^{32\times N_{stream,1}}, \mathbf{W}_2 \in \mathbb{C}^{32\times N_{stream,2}}$ with $\|\mathbf{W}_1\|_F = \|\mathbf{W}_2\|_F = 1$ and $N_{stream,k} \in [32]$ is selected by the user, and its output is defined by the following formula

$$\mathbf{Y}(\mathbf{W}_1, \mathbf{W}_2) = \sum_{t\in[N]} \mathbf{H}_3^{(t)} \times f_t\left(\mathbf{H}_1^{(t)} \times \mathbf{X}_1(\mathbf{W}_1),\ \mathbf{H}_2^{(t)} \times \mathbf{X}_2(\mathbf{W}_2)\right)$$

Where

$$\mathbf{X}_k(\mathbf{W}_k) = \sqrt{\frac{32}{N_{stream,k}}}\mathbf{W}_k^* \times \mathbf{S}^{(k)}[1 : N_{stream,k}, :]$$

And for all $t \in [N]$ the function $f_t : \mathbb{C}^{1\times 300} \times \mathbb{C}^{1\times 300} \to \mathbb{C}^{1\times 300}$, is defined by

$$f_t(\mathbf{a}, \mathbf{b}) = \left[c_1^{(t)}\mathbf{a}_1^2(\mathbf{b}_1)^* \ \dots \ c_{300}^{(t)}\mathbf{a}_{300}^2(\mathbf{b}_{300})^*\right]$$

We are assured that vectors $\mathbf{H}_1^{(t)}, \mathbf{H}_2^{(t)}, \mathbf{H}_3^{(t)}$ satisfy the following property, for each $t \in [N]$ either all coordinates of $\mathbf{H}_1^{(t)}, \mathbf{H}_2^{(t)}, \mathbf{H}_3^{(t)}$ are nonzero (abs value order at least 10^{-5}) or there exists $j \in [32]$ such that all their coordinates are zero (abs value order 10^{-9}) except their j-th one (abs value order 1). For the first scenario we say that t is a "target" signal and in the other t is an "interference" signal.

We are tasked with retrieving $\mathbf{T} := \text{span}\left\{\mathbf{H}_1^{(t)} : t \text{ is "target"}\right\}$, more specifically, given that we are guaranteed that $\dim(\mathbf{T}) = |\{t : t \text{ is "target"}\}|$, we need to find $\dim(\mathbf{T})$ vectors in $\mathbb{C}^{1\times 32}$ such that they form a basis of $\mathbf{T}$.

Some additional properties of the parameters and the system were omitted as they do not affect and are not used by our proposed solution.

5 Assumptions

For the analysis of our solution we use the following assumptions:

1. In order to make the analysis easier we assume a stricter definition for "interference" signals, that is, there exists $j \in [32]$ such that all their coordinates are **exactly** zero except their j-th one which is **exactly** of absolute value 1.

2. The values in the unknown parameters in the system are sufficiently uncorrelated such that taking random coordinates of the matrices gives different values with high probability. We will use this to obtain the target span space in our solution.

At the end of the next section we explain how the errors incurred by this assumptions can be reduced in practice.

6 Solution

6.1 Theoretical Solution

In order to introduce the proposed solution note that using $N_{stream,1} = N_{stream,2} = 1$, the equations presented in the previous section give

$$\mathbf{Y}(\mathbf{e}_k^{32}, \mathbf{e}_\ell^{32})_{i,j} = 32\sqrt{32} \sum_{t\in[N]} c_j^{(t)} \mathbf{H}_{3,i}^{(t)} \left(\mathbf{S}_{1,j}^{(1)} \mathbf{H}_{1,k}^{(t)}\right)^2 \left(\mathbf{S}_{1,j}^{(2)} \mathbf{H}_{2,\ell}^{(t)}\right)^*$$

And with $k \neq k' \in [32]$

$$\mathbf{Y}\left(\frac{\mathbf{e}_k^{32} + \mathbf{e}_{k'}^{32}}{\sqrt{2}}, \mathbf{e}_\ell^{32}\right)_{i,j} = 16\sqrt{32} \sum_{t\in[N]} c_j^{(t)} \mathbf{H}_{3,i}^{(t)} \left(\mathbf{S}_{1,j}^{(1)} \left(\mathbf{H}_{1,k}^{(t)} + \mathbf{H}_{1,k'}^{(t)}\right)\right)^2 \left(\mathbf{S}_{1,j}^{(2)} \mathbf{H}_{2,\ell}^{(t)}\right)^*$$

Then lets define

$$\mathbf{y}_{i,j,k,k',\ell} := \begin{cases} 2\mathbf{Y}\left(\frac{\mathbf{e}_k^{32} + \mathbf{e}_{k'}^{32}}{\sqrt{2}}, \mathbf{e}_\ell^{32}\right)_{i,j} - \mathbf{Y}(\mathbf{e}_k^{32}, \mathbf{e}_\ell^{32})_{i,j} - \mathbf{Y}(\mathbf{e}_{k'}^{32}, \mathbf{e}_\ell^{32})_{i,j} & \text{if } k \neq k' \\ 2\mathbf{Y}(\mathbf{e}_k^{32}, \mathbf{e}_\ell^{32})_{i,j} & \text{if } k = k' \end{cases}$$

Then

$$\mathbf{y}_{i,j,k,k',\ell} = 64\sqrt{32} \sum_{t\in[N]} c_j^{(t)} \mathbf{H}_{3,i}^{(t)} \left(\mathbf{S}_{1,j}^{(1)}\right)^2 \mathbf{H}_{1,k}^{(t)} \mathbf{H}_{1,k'}^{(t)} \left(\mathbf{S}_{1,j}^{(2)} \mathbf{H}_{2,\ell}^{(t)}\right)^*$$

Therefore, taking $\alpha_{i,j,k,\ell}^{(t)} := 64\sqrt{32} c_j^{(t)} \mathbf{H}_{3,i}^{(t)} \left(\mathbf{S}_{1,j}^{(1)}\right)^2 \mathbf{H}_{1,k}^{(t)} \left(\mathbf{S}_{1,j}^{(2)} \mathbf{H}_{2,\ell}^{(t)}\right)^*$ we get

$$\hat{\mathbf{y}}_{i,j,k,\ell} := [\mathbf{y}_{i,j,k,1,\ell} \ \cdots \ \mathbf{y}_{i,j,k,32,\ell}] = \sum_{t\in[N]} \alpha_{i,j,k,\ell}^{(t)} \mathbf{H}_1^{(t)}$$

In other words, $\hat{\mathbf{y}}_{i,j,k,\ell} \in \operatorname{span}\left\{\mathbf{H}_1^{(t)} : t \in [N]\right\}$. Further, using assumption 1 in the previous section note that if we take $i \neq \ell$ we get $\alpha_{i,j,k,\ell}^{(t)} \neq 0$ if and only if t is a "target" signal because otherwise (i.e. on an "interference" signal) at most one of $\mathbf{H}_{3,i}^{(t)}$ and $\mathbf{H}_{2,\ell}^{(t)}$ has absolute value greater than zero. Therefore, taking $i \neq \ell$ we make sure to get $\hat{\mathbf{y}}_{i,j,k,\ell} \in \operatorname{span}\left\{\mathbf{H}_1^{(t)} : t \text{ is "target"}\right\} = \mathbf{T}$.

We know how to obtain a vector in $\mathbf{T}$. Now, using assumption 2 in the previous section, we see that taking enough samples, lets say M, and random indexes $\{i_n\}_{n\in[M]} \in$

$[32]^M, \{j_n\}_{n\in[M]} \in [300]^M, \{k_n\}_{n\in[M]} \in [32]^M, \{\ell_n\}_{n\in[M]} \in [32]^M$ such that $i_n \neq \ell_n$ for all $n \in [M]$, then with high probability we obtain

$$\operatorname{span}\{\hat{\mathbf{y}}_{i_n,j_n,k_n,\ell_n} : n \in [M]\} = \mathbf{T}$$

Now, lets assume we fix a particular set of M samples such that the corresponding $\hat{\mathbf{y}}$ vectors span $\mathbf{T}$. Lets define $\hat{\mathbf{Y}} := \begin{bmatrix}\hat{\mathbf{y}}^T_{i_1,j_1,k_1,\ell_1} & \cdots & \hat{\mathbf{y}}^T_{i_M,j_M,k_M,\ell_M}\end{bmatrix} \in \mathbb{C}^{32\times M}$, then we can recover a basis for $\mathbf{T}$ using an singular value decomposition (SVD) of the matrix. In other words, we obtain matrices $\mathbf{U} \in \mathbb{C}^{32\times 32}, \mathbf{\Sigma} \in \mathbb{C}^{32\times M}, \mathbf{V} \in \mathbb{C}^{M\times M}$ such that $\hat{\mathbf{Y}} = \mathbf{U\Sigma V}^*$, the columns $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices, and $\mathbf{S}$ is a rectangular diagonal matrix with $\{\mathbf{\Sigma}_{i,i}\}_{i\in[32]}$ equal to the singular values of $\hat{\mathbf{Y}}$ ordered in decreasing absolute value. Further, the columns of $\mathbf{U}$ form a unitary basis of the columns of $\hat{\mathbf{Y}}$ and they represent the left-singular vectors associated to the singular values in $\mathbf{\Sigma}$ (in the same order). Then a solution to the problem is given by

$$\left\{\mathbf{U}[:,i]^T : i \in [32] \wedge \mathbf{\Sigma}_{i,i} \neq 0\right\}.$$

Note that $\mathbf{\Sigma}$ will have exactly $\dim(\mathbf{T})$ nonzero entries and $\mathbf{U}$ has exactly $\dim(\mathbf{T})$ nonzero columns, therefore we can use the SVD decomposition to infer the number of "target" signals.

This gives a complete solution to the problem. In terms of complexity, the bulk of the computation lies in making M oracle calls to the system for each sampled set of indexes, on each call out computation is dominated by reading the output in constant time (as the size is always 32×300). From the results of the calls we reconstruct $\hat{\mathbf{Y}}$ and then we obtain its SVD decomposition to obtain the solution, which can be done in $O(M^2)$.

6.2 Practical Considerations

In our solution we assumed that taking indexes $i \neq \ell$ makes $\hat{\mathbf{y}}_{i,j,k,\ell} \in \mathbf{T}$, but that relies in assumption 1 which is not exactly correct in the problem setting. In reality we get that $\hat{\mathbf{y}}_{i,j,k,\ell}$ is close to $\mathbf{T}$, i.e. there exists $P_{\mathbf{T}}(\hat{\mathbf{y}}_{i,j,k,\ell}) \in \mathbf{T}$ (projection to $\mathbf{T}$) such that $|\hat{\mathbf{y}}_{i,j,k,\ell} - P_{\mathbf{T}}(\hat{\mathbf{y}}_{i,j,k,\ell})|$ is close to 0 (order of 10^{-9}). This difference makes it such that the basis produced does not exactly generate the same subspace as $\mathbf{T}$.

Similarly, even ignoring the error just described, given that the parameters of the system are unknown, we don't have a direct way of knowing how many samples are enough to obtain a full representation of $\mathbf{T}$.

Both of these problems can be minimized by increasing the number of samples as much as we can, intuitively, assuming that that the error induced by the first assumption is randomly distributed, with enough samples, the closer the first $\dim(\mathbf{T})$ vectors of the SVD decomposition represent $\mathbf{T}$. Note that because of these deviations, $\mathbf{\Sigma}$ will almost surely have exactly 32 nonzero singular values, and $\mathbf{U}$ will almost surely have complete rank. Therefore, in order to accurately determine how many columns of $\mathbf{U}$ we need to select, we can use a tolerance parameter $\varepsilon > 0$, and the solution we use in practice is given by

$$\left\{\mathbf{U}[:, i]^T : i \in [32] \wedge |\mathbf{\Sigma}_{i,i}| < \varepsilon\right\}.$$

6.3 Solution summary and results

The solution proposed can be summarized with the following steps:

1. Generate M samples of indexes $\{i_n\}_{n\in[M]} \in [32]^M, \{j_n\}_{n\in[M]} \in [300]^M, \{k_n\}_{n\in[M]} \in [32]^M, \{\ell_n\}_{n\in[M]} \in [32]^M$ such that $i_n \neq \ell_n$ for all $n \in [M]$.
2. Obtain vectors $\hat{\mathbf{y}}_{i_n,j_n,k_n,\ell_n}$ for all $n \in [M]$ and construct matrix $\hat{\mathbf{Y}}$. This involves around $2M \times 32$ oracle calls, but this can be reduced by considering sample indexes which reuse oracle calls in the first step.
3. Obtain an SVD decomposition $\hat{\mathbf{Y}} = \mathbf{U\Sigma V}^*$.
4. Given a parameter $\varepsilon > 0$ which can be tuned in testing, return the set of row vectors $\left\{\mathbf{U}[:, i]^T : i \in [32] \wedge |\mathbf{\Sigma}_{i,i}| < \varepsilon\right\}$. In the testing suite we more specifically return $\mathbf{L}_{est}$, a matrix with rows as the described solution row vectors.

This solution achieves a full score of 1000000 points in the challenge.