SASP-DAAP Homework #2

April 2022

The topic of this homework is acoustic source localization using microphone arrays. In the assignment folder you will find a `mat` file containing an array vector synthetically generated using 64 microphones. The microphones sample the acoustic field produced by two acoustic sources placed at different locations in space whose angular positions with respect to the array have to be determined. Your task is to apply both a Delay-And-Sum (DAS) beamformer and the MUSIC method to estimate the Direction Of Arrival (DOA) of both sources. You are required to complete the provided MATLAB script. You are then invited to comment the obtained results and answer a few questions.

1 Acoustic sources

In an anechoic room, we placed two acoustic sources at different locations in space, both lying on the same plane. Each source emits a single sine wave at a sampling frequency of $F_s = 8$ kHz. Namely, let $s_1(n)$ and $s_2(n)$ be the source signals defined as

$$s_i(n) = a_i \cdot \sin(2\pi f_{c_i} n/Fs), \qquad i = 1, 2 \quad n = 0, ..., N \tag{1}$$

where $\omega_{c_i} = 2\pi f_{c_i}$ is the angular frequency of the i-th source signal expressed in radians per second. Notice that $\omega_{c_1} \neq \omega_{c_2}$ and $a_1 \neq a_2$. Your task is to estimate the DOA for both source 1 and source 2 using a Uniform Linear Array.

2 Acoustic sources and the microphone array

The Uniform Linear Array (ULA) consists of M identical omnidirectional sensors uniformly spaced on a line coinciding with the x-axis of our coordinate reference system. The distance d between two adjacent microphones in the array was chosen so to avoid spatial aliasing for every possible angle of arrival and every signal frequency from 0 Hz up to the Nyquist frequency. The value of d is left to be determined as part of the homework.

The ULA lies on the same plane as the acoustic sources, therefore the spatial analysis can be limited to the 2-dimensional case. The frequency responses of the microphones in the array are assumed to be $H_1(\omega) = H_2(\omega) = ... = H_M(\omega) = 1$.

2.1 Model assumptions

As mentioned above, the acoustic space is considered to be anechoic, i.e., we can disregard the effect of reverberation and we can model the mixing process as instantaneous. Furthermore, the sources are placed in the far-field with respect to the ULA. Therefore, we can assume the acoustic wavefronts impinging on the array to be planar. Finally, the acoustic propagation is assumed to be homogeneous, with no dispersion.

As consequence of these assumptions, the channel response characterizing the propagation of each source signal from the i-th source to the m-th microphone amounts to a pure delay.

The speed of sound is $c = 340$ m/s.

2.2 DOA

In the 2D case, the DOA is defined as the angle between the y-axis and the line orthogonal to the acoustic wavefronts. With respect to our coordinate reference system, the DOA is $0°$ if the wavefronts are parallel to the array, it is positive if the source is placed in the second quadrant and it is negative if the source is located in the first quadrant. We consider only DOAs in the range $[-90°, 90°]$.

3 Array model

We define the array vector as

$$\mathbf{y}(\omega) = \mathbf{A}\mathbf{s}(\omega) + \mathbf{e}(\omega) \tag{2}$$

where $\mathbf{s}(\omega) = [S_1(\omega),\, S_2(\omega)]^T$ is the source vector containing the spectra of the source signals $s_1(n)$ and $s_2(n)$, $\mathbf{A}$ is a $M \times 2$ propagation matrix and $\mathbf{e}(\omega)$ is a vector of spatially white sensor noise with i.i.d. components and variance $\sigma_e^2 = 0.6$.

The file `array_data_64_mics.mat` contains the array vector obtained according to (2) using an array of $M = 64$ microphones with intra-microphone spacing d.

4 DOA estimation

Since the two source signals are sine waves, we can exploit the source localization techniques introduced for narrow-band signals. In particular, in this homework, we will focus on the DAS beamformers and MUSIC.

4.1 Frequency estimation

The two acoustic sources emit sine waves at two different frequencies. Thanks to this prior knowledge, we can limit our narrow-band analysis only to those (angular) frequencies. However, notice that ω_{c_1} and ω_{c_2} are unknown and you are required to estimate their value from the mixture signals (i.e., the provided array vector).

4.2 The Delay-and-sum beamformer

The idea behind the DAS beamformer is to perform spatial filtering so to minimize the power arriving from every direction while leaving the signal coming from a target angle unaltered. In general, this can be though as a band-pass filtering in the spatial domain. By varying the target angle, we are able to obtain a bank of spatial filters whose main lobes are steered towards every direction we are interested in examining.

By computing the power of the signal filtered with the spatial filters associated to all the candidate angles of arrival, we are able to estimate the directions associated to the highest impinging acoustic power. The source DOAs can be thus estimated as the direction associated with the largest impinging power at a certain frequency.

Analytically, this corresponds to finding the angle of the most prominent peak of the DAS pseudo-spectrum, which is defined as

$$p_{\text{DAS}}(\theta) = \frac{\mathbf{a}(\theta)^H \hat{\mathbf{R}} \mathbf{a}(\theta)}{M^2}, \quad \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \tag{3}$$

where $\hat{\mathbf{R}} = \frac{1}{N-1}\mathbf{y}(\omega)\mathbf{y}(\omega)^H$ is a sample estimate of the covariance matrix of the array data, $\mathbf{a}(\theta)$ is the steering vector defined in (4) and $\omega_{s_i} = \omega_{c_i} d \sin\theta / c$ is the normalized spatial frequency. Notice that, since we are analysing the pseudo-spectrum at the two source signal frequencies, the carrier frequency ω_{c_i} should be selected accordingly. Namely, $i = 1$ when we consider the first source and $i = 2$ when we consider the second one.

$$\mathbf{a}(\theta) = \left[1, e^{-j\omega_{s_i}}, ..., e^{-j(M-1)\omega_{s_i}}\right]^T \tag{4}$$

4.3 DOA candidates

Recalling the Delay-And-Sum beamformer theory, we have to evaluate the pseudo-spectrum for different angles θ, which are to be considered as potential candidates for our DOA estimate. In this homework, the angular resolution of such an analysis is to be set to 1° ranging from -90° to $+90^\circ$, for a total of 181 candidate angles.

4.4 The spatial response

After having estimated the DOA $\bar{\theta}_i$ for the i-th source, you are required to plot the spatial response (also known as *beam pattern*) for the angle of your DOA estimate. The spatial response is defined as $\mathbf{h}^H(\bar{\theta}_i)\mathbf{a}(\theta)$, where $\mathbf{h}(\bar{\theta}_i) = \mathbf{a}(\bar{\theta}_i)/M$ is the beamformer filter for the DOA angle $\bar{\theta}_i$ and $\mathbf{a}(\theta)$ is the steering vector as a function of the candidate angles selected with the angular resolution described in Section 4.3.

4.5 Spatial filtering

Once you have computed the beamformer filter $\mathbf{h}(\bar{\theta}_i)$ for both sources, we are ready to apply spatial filtering to the array data $\mathbf{y}(\omega)$. This spatial band-pass filter will reduce the power of every signal coming from directions different from the DOA. Therefore, if the DOA estimate is good enough, we may consider the output of the spatial filter $\hat{S}_i(\omega)$ as an estimate of the spectrum of the i-th source signal $s_i(n)$.

$$\hat{S}_i(\omega) = \mathbf{h}(\bar{\theta}_i)^H \mathbf{y}(\omega), \quad i = 1, 2 \tag{5}$$

The output of the filter is transformed back in the time domain and reproduced using `soundsc`. Notice that the noise terms in $\mathbf{e}(\omega)$ have a rather high variance; the sound won't be pleasing: turn down the volume of your loudspeakers! Notice that you might have to cast the time-domain signals to `real` before calling `soundsc`.

5 Parametric methods

In the second part of the MATLAB script, you are asked to implement the "MUltiple SIgnal Classification" (MUSIC) algorithm.

Key aspect of this source localization method is the parametrization of the covariance matrix $\mathbf{R}$ by means of the eigenvalue decomposition, i.e.,

$$\mathbf{R} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^H \tag{6}$$

MATLAB provides the funcion `eig` that returns the eigenvalues and eigenvectors of a matrix passed as an argument. Remember that, in order to easily implement MUSIC, we would like the eigenvalues to be sorted in descending order, i.e. $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_M$. The columns of $\mathbf{Q}$ (i.e., the corresponding eigenvectors) should be therefore rearranged accordingly.

In the following, let $N_{\text{src}} = 2$ denote the number of sources.

5.1 MUSIC algorithm

The MUSIC algorithm relies on the properties of the matrix $\mathbf{V}$ whose columns are the eigenvectors of the matrix $\mathbf{R}$ associated to the smallest $M - N_{\text{src}} - 1$ eigenvalues. The column space of $\mathbf{V}$ is known as *noise subspace*. Naturally, since the covariance matrix $\hat{\mathbf{R}}$ is an estimate of $\mathbf{R}$, the eigenvectors that span the noise subspace are also an estimate of the true ones. Namely, we can only compute $\hat{\mathbf{V}}$ which is an estimate of $\mathbf{V}$.

The estimate of the DOAs $\bar{\theta}_1$ and $\bar{\theta}_2$ are obtained by determining the angle associated to the most prominent peak in the MUSIC pseudo-spectrum defined in (7).

$$p_{\text{MUSIC}}(\theta) = \frac{1}{\mathbf{a}^H(\theta)\hat{\mathbf{V}}\hat{\mathbf{V}}^H\mathbf{a}(\theta)}, \quad \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \tag{7}$$

You are required to compute and plot the MUSIC pseudo-spectrum for both source signal frequencies ω_{c_1} and ω_{c_2}. Indeed, notice that the steering vector $\mathbf{a}(\theta)$ depends on the narrow-band frequency considered.

Questions

1. Discuss the differences between the Delay-and-sum approach to DOA estimation and the parametric methods such as MUSIC. What are the differences in terms of performance and pseudo-spectra? Are there any theoretical or practical reasons to prefer one approach over the other?

2. The microphone spacing d was selected to avoid spatial aliasing for every angle of arrival and for every frequency up to the Nyquist frequency. This is a rather conservative approach, could we have done differently?

3. Under the same simplifying assumptions made in this homework (e.g., no reverberation, ideal propagation, planar geometry, ideal sensors...), describe the procedure you would use to synthesize the array vector $\mathbf{y}(\omega)$ captured by an ULA with $M = 3$ microphones in the case of two sinusoidal source signals $s_1(n)$ and $s_2(n)$ at the same frequency ω_c with DOA $\bar{\theta}_1$ and $\bar{\theta}_2$, respectively. Please provide a few lines of MATLAB code.

Notes

- Please notice that you are required to provide your results in degrees while many MATLAB functions work with radians. Use the function `deg2rad` to convert degrees to radians, and the function `rad2deg` to convert radians to degrees.
- The instances of `real` in the MATLAB script are meant to remove the residual imaginary component left due to rounding errors. Other than that, you should not encounter any numerical problem in the plotting functions.
- the `pause` command stops the execution of the MATLAB script and waits for user input. Press any key to resume.
- Groups of two people are allowed; each group should submit a single zip files.
- Submissions should have the following format, e.g., `HW2_{name1}_{name2}.zip`
- Please be concise in writing the report (maximum two pages).