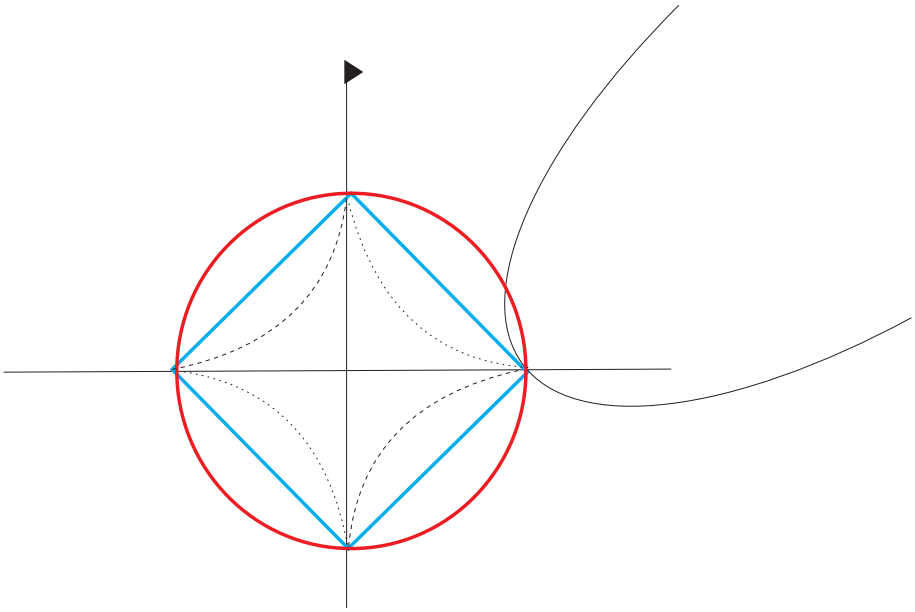
CHALMERS



Consistency analysis and regularization parameter selection for direction-of-arrival estimation Based on compressed sensing techniques

Ashkan Panahi

Signal Processing Group
Department of Signals and Systems
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden

EXE037/2010

Abstract

In this report we address the well known problem of estimating the Direction Of Arrival (DOAs) from sensor array data, and discuss it as a linear regression problem by assuming a narrow band, far field model of the sources. Recently, the Least Absolute Shrinkage and Selection Operator (LASSO) is proposed for such a problem as an approximation to the computationally hard, but precise solution of Maximum Likelihood (ML). We explain the implementation using convex programming tools, and show the superiority of this method by discussing the theoretical bounds of consistency and comparing the results to that of conventional estimators. Furthermore, we give two different solutions for selecting the regularization parameter in the LASSO method. First, by viewing the problem as a model order selection and using Minimum Description Length (MDL) principle and second, by taking the Bayesian nature of regularization into account and deriving an ML estimator for the regularization parameter. Finally, we discuss about the theoretical consistency of such a method for the DOA estimation in some extreme cases. We particularly show that for the noiseless case the method may not be consistent for very close sources. This shows the fundamental resolution of LASSO for the DOA estimation.

Keywords: DOA, Localization, Linear regression, LASSO, Regularization, Consistency analysis, Compressed sensing, Sparse Analysis, Convex programming etc.

CONTENTS

1 INTRODUCTION

Estimating the direction of signal waves arriving from a set of relatively far sources is a well-studied problem in the literature. This problem might be applied to a variety of applications such as radar and sonar (Trees 1968), wireless communications (Poor and G. W. Wornell 1998), and seismology (Robinson and Treitel 1980). These signals are measured by an array of sensors with a relatively small dimensions. According to the nature of the signals, there exists a variety of methods to treat such a problem. From the frequency spectrum point of view, the signal might be classified as a narrow-band or wide-band signal. However, there is always the possibility of dividing a wide frequency band into many narrow sub-bands, and treat the resulting multiple narrow-band signals either jointly or separated. Thus, in this work we focus on the narrow-band far field model of the received data.

Forming a linear regression estimation problem, the DOA in the desired scenario has been studied by a variety of techniques such as Maximum likelihood (Schweppe 1968), subspace methods (Viberg and Ottersten 1991), and beamforming techniques (Van Veen and Buckley 1988, Viberg and Krim 1997). In this work we investigate the newly introduced method of LASSO (Tibshirani 1996) as a solution to the DOA estimation problem which has been proposed in (Malioutov *et al.* 2005). The LASSO method is also know as *Basis Pursuit*, first introduced in (Chen and Donoho 1994). We view the estimation problem as a shrinkage-operator-approximation to the ML estimation, by reformulating the DOA problem in a sparse framework. Later we will discuss about another insight into the method as a Bayesian estimation.

Because the shrinkage operator is a regularization in essence, the choice of the regularization parameter is one of the fundamental obstacles to implement this method. To overcome this restriction, we view the problem as a model order selection for which many methods have been proposed, such as Generalized Likelihood Ratio Test (GLRT)(Stoica *et al.* 2004), Generalized Information Criterion (GIC)(Akaike 1973), Minimum Description Length (MDL)(Rissanen 2006), etc. We propose using MDL to treat the problem as a separated detection-estimation problem, and illustrate its performance by simulation. We also propose an ML estimation of the parameter which is more beneficial in small sample size.

Finally we examine the behavior of the method by a mathematical analysis in two extreme cases of noiseless data and large number of sensors respectively. We show that the method is not consistent for any choice of directions and we will give a consistency conditions which is in a close relationship to the concept of resolution. By this method we prove the superiority of LASSO as compared to the conventional beamforming methods but also that it is inferior to, e.g. subspace methods at high enough SNR.

2 NOTATION

Abbreviations

AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
DOA	Direction Of Arrival
FFT	Fast Fourier Transform
GIC	Generalized Information Criterion
LLN	Law of Large Numbers
LTI	Linear Time Invariant
MAP	Maximum a Posteriori Probability
MDL	Minimum Description Length
ML	Maximum Likelihood
MUSIC	Multiple Signal Classification
NML	Normalized Maximum Likelihood
PDF	Probability Density Function
SC	Stochastic Complexity
SML	Stochastic Maximum Likelihood
SNR	Signal to Noise Ratio
SOC	Second Order Cone
USC	Universal Stochastic Complexity

Notations

$\mathbf{a}(\theta)$	steering vector at direction θ
$\mathbf{A}$	matrix of steering vectors
$\mathbf{A}^g$	matrix of steering vectors sampled at the grid
m	number of sensors
n	number of sources
N	number of gridpoints
$\mathbf{n}$	noise vector
s	source waveform
$\mathbf{s}^g$	sparse generalized source vector
x	received data value
θ_i	direction of source i
θ_i^g	direction of grid point i
θ	vector of true directions
Θ	the set of all possible θ vectors
ϕ	electrical angle
λ	regularization parameter
σ^2	noise variance
$\hat{}$	estimation of the variable

3 MODELING

In this chapter we discuss the models for the data generated by an array of sensors in different physical scenarios. We also review some popular estimators and their theoretical performances.

The data model is a mathematical representation of the physical process of generating the data. This model may contain stochastic (e.g. noise) as well as deterministic parameters. The deterministic parameters might be known or unknown beforehand. On the other hand, the estimator is an algorithm aimed to estimate an unknown parameter in the model using the known parameters including the received data. Such a problem sometimes is referred as an *inverse* problem. Normally, an exact estimation is impossible due to the fact that generally, the total number of the unknown parameters (stochastic and deterministic) is greater than the number of equations provided by the model. However, if the number of equations is at least equal to the number of unknown deterministic parameters, it is possible to define and compute a stochastic grade of accuracy (e.g. error variance). In this case we can talk about the *best* estimator in a certain stochastic sense. Such a model might be referred as a *well-behaved* model. For the real application, an inexact nature of the data model increases the inaccuracy of the estimation which is known as the *model noise*.

3.1 Data Model

Our model is based on the far-field representation of the narrow band signal. A narrow band signal can be expressed as $u(t) = \mathrm{Re}(\tilde{u}(t)e^{j\omega t})$ where $\tilde{u}(t)$ is a baseband complex signal known as complex envelope, and ω is the carrier frequency. Figure 3.1 shows the structure of sources and receivers in a linear, and constant medium. For the case of multiple sources the received data can be modeled as the superposition of the received data for each individual source. Assuming a far field source, the wave about the receiver array can be approximated by a plane wave traveling from the source to the origin. Denoting the complex envelope field at the origin by $s(t)$, the field at a location $\vec{x}$ is related to that at the origin by $\tilde{u}(\vec{x}, t) = \tilde{u}(0, t)e^{-j\vec{k}.\vec{x}} = s(t)e^{-j\vec{k}.\vec{x}}$, where $\vec{k}$ is the wave vector, and $\vec{k}.\vec{x}$ is the scalar product. Also note that due to the linear nature of the medium, the field at the origin is the output of a Linear Time Invariant (LTI) system with a narrow band source input. This response is an attenuation and time shifting of the original signal, which does not distort it.

Next, suppose a set of m sensors at locations $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m$. The sensors are assumed to be ideal except for an additive noise term. We assume a special configuration of the sensors, all of which are oriented along a unique axis. We set this direction as the x coordinate. The sensor locations can then be written as $\vec{x}_i = d_i\vec{u}_x$ where $\vec{u}_x$, is the unit vector of x axis, and d_i is the separation between the i^{th} sensor and the

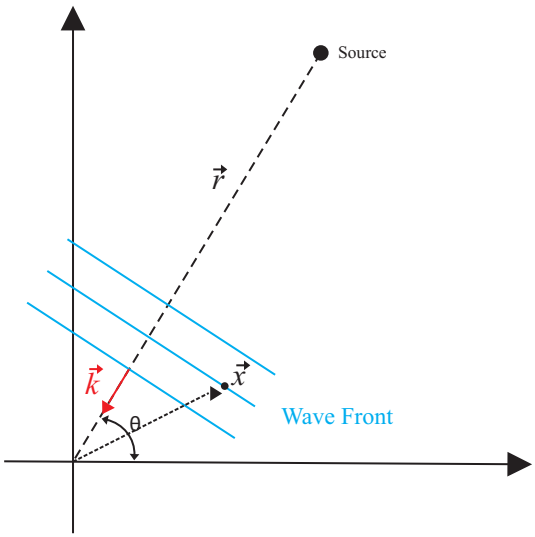


Figure 3.1. The plane wave of a far field source arround the origin

origin. Such an array may be referred to as a *linear array* . The received data at this sensor array can be written as,

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_m(t) \end{bmatrix} = \begin{bmatrix} e^{-j\vec{k}.\vec{x}_1} \\ e^{-j\vec{k}.\vec{x}_2} \\ \vdots \\ e^{-j\vec{k}.\vec{x}_m} \end{bmatrix} s(t) = \begin{bmatrix} e^{j\frac{2\pi}{\lambda}d_1\cos\theta} \\ e^{j\frac{2\pi}{\lambda}d_2\cos\theta} \\ \vdots \\ e^{j\frac{2\pi}{\lambda}d_m\cos\theta} \end{bmatrix} s(t) = \mathbf{a}(\theta)s(t) \tag{3.1}$$

where λ is the wavelength. The vector $\mathbf{a}(\theta)$ is known as the *steering vector*. For the case of n sources on directions $\theta_1, \ldots, \theta_n$, the received data will be

$$\begin{aligned} \mathbf{x}(t) &= \sum_{i=1}^{n} \mathbf{a}(\theta_i)s_i(t) \\ &= [\mathbf{a}(\theta_1)\ \mathbf{a}(\theta_2)\ldots\ \mathbf{a}(\theta_n)] \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_n(t) \end{bmatrix} = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) \end{aligned} \tag{3.2}$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$. We are also interested in a special, but important case of the Uniform Linear Array, in which $d_k = (k-1)\Delta$ for an arbitrary positive separation factor Δ, in which case, for simplicity, we introduce $\phi = \frac{2\pi}{\lambda}\Delta\cos\theta$. A common choice for the separation factor is $\Delta = \frac{\lambda}{2}$, in which case we deal with a half-wavelength array with $\phi = \pi\cos\theta$. Finally, due to the model and the physical receiver noise the overall model of the received data can be written as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \tag{3.3}$$

where the noise vector $\mathbf{n}(t)$ is a stochastic, unknown complex variable. We assume that it is a zero-mean, circularly symmetric, Gaussian vector with the covariance matrix equal to $\sigma^2\mathbf{I}$.

The number of the unknown deterministic parameters (i.e. source signals and directions) in such a model is $3n$ (the directions and the real and imaginary parts of the source waveforms[1]). Thus, according to the discussion in the introduction of this chapter, this model is well-behaved model if $n \leq \frac{2m}{3}$. This is a general assumption we make through the rest of this work. However, if the source signals were treated as stochastic parameters the criterion would change to $n \leq m$. In this case there exists a stochastic prior knowledge about the sources.

3.2 Conventional Estimation Techniques

One of the most straightforward and accurate estimation techniques for unknown deterministic parameters in (3.3) the is Maximum Likelihood (ML) principle. The ML method uses the likelihood function, $f(\mathbf{x};\mathbf{s},\boldsymbol{\theta})$ due to the stochastic model of the noise and maximizes it over the all possible deterministic parameters. Assuming a temporally uncorrelated received data, the likelihood function for the model (3.3) is given by

$$f(\mathbf{x};\mathbf{s},\boldsymbol{\theta}) = \frac{1}{\sqrt{\pi(\sigma^2)^m}^T} e^{-\frac{\sum_{t=1}^{T} |\mathbf{x(t)}-\mathbf{A}(\boldsymbol{\theta})\mathbf{s(t)}|^2}{\sigma^2}} \tag{3.4}$$

and the ML estimator is given by

$$(\{\hat{\mathbf{s}}(t)\}_{t=1}^{T}, \hat{\boldsymbol{\theta}}) = \arg \min_{\{\mathbf{s}(t)\}_{t=1}^{T}, \boldsymbol{\theta}} \sum_{t=1}^{T} |\mathbf{x(t)} - \mathbf{A}(\boldsymbol{\theta})\mathbf{s(t)}|^2 \tag{3.5}$$

Denoting the trace operator by $\mathbf{tr}(.)$, this can be further simplified as

$$\begin{aligned} \hat{\mathbf{s}}(t) &= \mathbf{A}^{\dagger}(\hat{\boldsymbol{\theta}})\mathbf{x}(t), \\ \hat{\boldsymbol{\theta}} &= \arg\min_{\Theta} \mathbf{tr}(\Pi^{\perp}_{\mathbf{A}(\boldsymbol{\theta})}\hat{\mathbf{R}}) \end{aligned} \tag{3.6}$$

where $\hat{\mathbf{R}} = \frac{\sum_{t=1}^{T} \mathbf{x}(t)\mathbf{x}^H(t)}{T}$ is the sampled correlation matrix. Further, $\Pi^{\perp}_{\mathbf{A}(\theta)}$ and $\mathbf{A}^{\dagger}(\boldsymbol{\theta})$ are the perpendicular projection and pseudo inverse matrices of $\mathbf{A} = \mathbf{A}(\boldsymbol{\theta})$ respectively, defined as

$$\begin{aligned} \mathbf{A}^{\dagger}(\boldsymbol{\theta}) &= (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H \\ \Pi^{\perp}_{\mathbf{A}(\boldsymbol{\theta})} &= \mathbf{I} - \mathbf{A}\mathbf{A}^{\dagger} \end{aligned} \tag{3.7}$$

Although ML gives a high-performance estimation of directions and sources, it is computationally inappropriate because the cost function, $\mathbf{tr}(\Pi^{\perp}_{\mathbf{A}(\boldsymbol{\theta})}\hat{\mathbf{R}})$ in (3.6) possesses many local minima on its domain. Accordingly, many suboptimal methods

[1]Note that here we treated the noise variance, σ as a known parameter. However, in many practical problems and even later in this work σ is unknown and the number of unknown parameters will be $3n+1$.

have been proposed to solve the DOA estimation problem. Note that if a good estimation of directions is available, the sources are evaluated by the first equation in (3.6), which can be interpreted as a zero forcing signal separation method. In this work we discuss the Beamforming and Subspace methods briefly as popular methods of avoiding ML complexity. It is worth noting that there exist some variations to ML which increase the applicability of it such as Nonlinear Least Squares (NLLS) implementation methods which may be implemented by Space-Alternating Generalized Expectation-Maximization Algorithm (SAGE) (Fessler and Hero 1994) and Stochastic ML (SML)(Ottersten *et al.* 1992).However the computational problem is still dominant for these methods.

3.2.1 Beamforming Techniques

The idea of beamforming is inspired from the time domain signal processing according to the similarity between time and uniform sampling in a ULA. Consider the inner product of different steering vectors of the array

$$|\mathbf{a}^H(\theta_1)\mathbf{a}(\theta_2)| = \frac{|e^{jm\Delta\phi} - 1|}{|e^{j\Delta\phi} - 1|} = \frac{|\sin\frac{m\Delta\phi}{2}|}{|\sin\frac{\Delta\phi}{2}|} = e(\Delta\phi) \tag{3.8}$$

where $\Delta\phi = \phi_2 - \phi_1$. The function $e(\Delta\phi)$ can be seen to decrease when the directions get far from each other. We can therefore find the energy component received from a certain direction, θ in space by computing the inner product of the data to the steering vector, $\mathbf{a}(\theta)$ and its average energy. For many snapshots ($T \to \infty$), we have

$$P(\theta) = \frac{1}{T}\sum_{t=1}^{T} |\mathbf{a}^H(\theta)\mathbf{x}(t)|^2 = \mathbf{a}^H(\theta)\mathbf{R}_x\mathbf{a}(\theta) = \mathbf{a}^H(\theta)\mathbf{A}(\theta)\mathbf{R}_s\mathbf{A}^H(\theta)\mathbf{a}(\theta) + m\sigma^2 \tag{3.9}$$

where $\mathbf{R}_x$ and $\mathbf{R}_s$ are the covariance matrices of $\mathbf{x}$ and $\mathbf{s}$ respectively. Thus, considering approximately perpendicular steering vectors we claim that the peak points of the spatial spectrum, $P(\theta)$ are at least close to the true directions of the sources. However, the quality of method is limited because of firstly, the bias introduced by superposition of the shifted patterns in Figure 3.2.1 and secondly, the resolution limitation due to the beam width of the antenna pattern. The latter may cause two peaks to merge if the true directions are too close to each other.

3.2.2 The MUSIC Method

The multiple signal classification (MUSIC) method is based on the fact that the ML criterion in (3.5) can be viewed as the problem of finding the best n-dimensional subspace describing the data in an m-dimensional space. The received data snapshots form a cloud of an ellipsoid around the subspace of noiseless received data which is the range space of $\mathbf{A}(\boldsymbol{\theta})$. For high SNRs the minor axis of the ellipse, known as the noise subspace, is related only to the components of noise while the major axis,

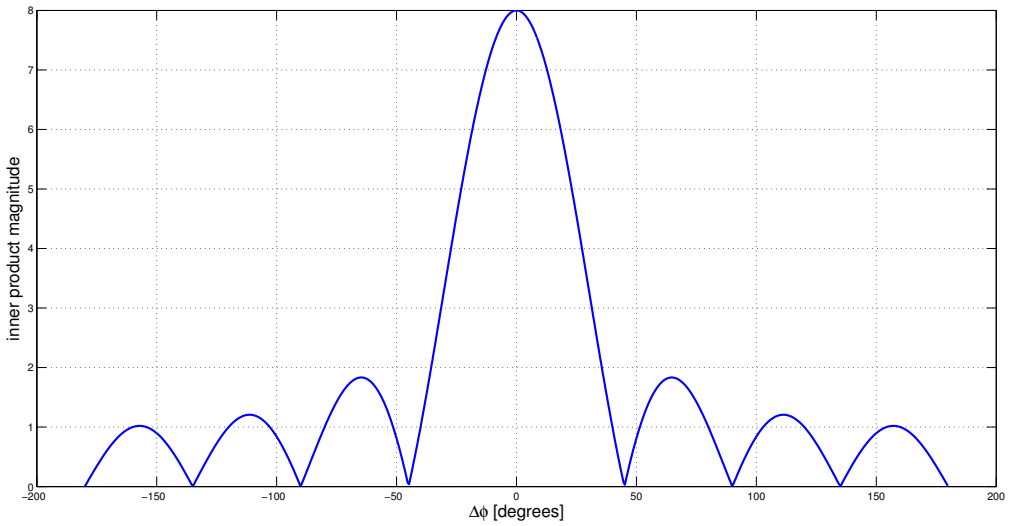


Figure 3.2. Beam pattern for 8 sensors of a ULA array versus the electrical angle separation

known as signal subspace, is related to both signal and noise components. It can be deduced that the signal subspace is approximately formed by a linear combination of the directions in which the data has a maximum-variance component, also known as principal components. This directions can be found by the following maximization.

$$\hat{\mathbf{w}} = \arg \max_{\|\mathbf{w}\|_2=1} \sum_t |\mathbf{w}^H \mathbf{x}(t)|^2 = \arg \max_{\|\mathbf{w}\|_2=1} \mathbf{w}^H \mathbf{R}_x \mathbf{w} \tag{3.10}$$

where $\mathbf{R}_x$ is the sampled covariance matrix. It is easy to show that (e.g. in (G. H. Golub 1996)) the solution for this maximization given by the eigenvector related to the maximum eigenvalue of $\mathbf{R}_x$. We conclude that the signal subspace is approximately the span of the eigenvectors related to the dominant values in the spectrum of $\mathbf{R}_x$. This space can be found using the eigenvalue decomposition of it which can be implemented by Singular Value Decomposition (SVD). As the result, $\mathbf{R}_x$ can be written as

$$\mathbf{R} = \mathbf{U}^{\mathbf{H}} \mathbf{\Sigma} \mathbf{U} \tag{3.11}$$

where $\mathbf{U}$ is the orthogonal matrix of eigenvectors and $\mathbf{\Sigma}$ is a diagonal matrix whose diagonal elements are the singular values, $\sigma_1^2 \geq \sigma_2^2 \geq \ldots \geq \sigma_m^2$. Suppose $\sigma_1, \ldots, \sigma_n$ are the dominant singular values and $\mathbf{U} = [\mathbf{U}_s \ \mathbf{U}_n]$ where $\mathbf{U}_s$ is an $m \times n$ matrix. Then the MUSIC spectrum is written as

$$P_{\mathrm{MU}}(\theta) = \frac{1}{\mathbf{a}^H(\theta) \mathbf{U}_n \mathbf{U}_n^H \mathbf{a}(\theta)} \tag{3.12}$$

The directions of the signals then can be estimated by locating the peaks in the spectrum.

4 SPARSE MODELING

In the previous chapter we reviewed some popular methods for DOA estimation. In this chapter we show that by reformulating the model in (3.3) it is possible to get a different approximation of ML estimation. This new approximation is based on the fact that the solution of an optimization problem containing the ℓ_1 norm is most likely to be sparse because of the singularity of the absolute value function at zero. On the other hand, due to the convex nature of the ℓ_1 norm it is possible to perform the ℓ_1 optimization problems in a computationally efficient way using convex optimization techniques. The main benefit of such a treatment is that the Nonlinear Least Squares (NLLS) problem of the maximum likelihood is transformed to a linear one by introducing a large dictionary matrix of all possible basis vectors. In the last part of this chapter we show the result of applying this new method to the simulated data which shows that the sparse solution given by ℓ_1 optimization is a "good" estimation of the actual data. However, the theoretical discussion is postponed to the next chapter.

4.1 Discretization and Grids

The problem of finding directions from a continuous interval of the real numbers can be solved by choosing a finite, discrete subset of index points in the interval and finding the closest index point to the real value. The precision of this method is directly related to how close the index points are. We give a formal definition of such a tessellation.

Definition 1. A grid $G = \{\theta_i^{(g)}\}_{i=0}^{N}$ is an increasing sequence of $N+1$ directions with $\theta_0^{(g)} = 0$ and $\theta_N^{(g)} = \pi$. We also denote the maximum distance between consecutive grid points (i.e $\theta_{i+1}^{(g)} - \theta_i^{(g)}$) by $d(G)$ as a measure of fineness of the grid.

Obviously, for every direction $\theta \in [0\ 2\pi]$ there exists a grid point, $\theta_i^{(g)}$ so that $|\theta_i^{(g)} - \theta| \leq \frac{d(G)}{2}$ .Thus for a half-wavelength ULA sensor array,

$$\Delta\phi = \pi \left|\cos\theta_i^g - \cos\theta\right| \leq \left|\pi \sin\frac{d(G_N)}{2}\right| \left|\sin\left(\theta + \frac{d(G_N)}{2}\right)\right| \leq \pi\frac{d(G_N)}{2} \tag{4.1}$$

According to the definition (3.8) we obtain

$$|\mathbf{a}^H(\theta)\mathbf{a}(\theta_i^g)| \geq e(\frac{d(G_N)}{2}) \tag{4.2}$$

This means that in (3.3), by projecting each basis vector ($\mathbf{a}(\theta_k)$), can be written as

$$\mathbf{a}(\theta_k) = \frac{1}{m}\mathbf{a}(\theta_{i_k}^g)\left(\mathbf{a}^H(\theta_k)\mathbf{a}(\theta_{i_k}^g)\right) + \mathbf{b}_k \tag{4.3}$$

for some $\theta^g_{i_k} \in G$ where $\mathbf{b}_k$ is the orthogonal projection of $\mathbf{a}(\theta_k)$ on $\mathbf{a}(\theta^g_{i_k})$. We alos have $\|\mathbf{b}_k\| \leq \sqrt{1 - e^2(\frac{d(G_N)}{2})}$. Then, (3.3) can be written as

$$\mathbf{x}(t) = \sum_{k=1}^{n} \mathbf{a}(\theta^g_{i_k}) \underbrace{\left(\mathbf{a}^H(\theta_k)\mathbf{a}(\theta^g_{i_k})\right) s_k(t)}_{s'_k(t)} + \underbrace{\sum_{k=1}^{n} \mathbf{b}_k s_k(t) + \mathbf{n}(t)}_{\mathbf{n}'(t)} \tag{4.4}$$

The discretized estimation is actually the estimation of the set of indexes $I = \{i_1, i_2, \ldots, i_n\}$ corresponding to the closest basis vectors in the grid G to the actual basis vectors. We call these indexes and their corresponding basis vectors the *active indexes* and *active basis* respectively. When $N \to \infty$, if $d(G_N) \to 0$ it can be seen that $\mathbf{b}_k \to 0$, thus $\mathbf{n}'(t) \to \mathbf{n}(t)$, and $\theta^g_{i_k} \to \theta_k$. From now on we assume that the real directions are on the grid and use the model in (4.4). For simplicity, we denote $\theta^g_{i_k}$ and $\mathbf{a}(\theta^g_{i_k})$ by θ_k and $\mathbf{a}_k$ respectively. Later we will discuss the overall performance of the estimation by looking at the true directions which may not be on the grid. Finally, note that for every grid, G_N we have $d(G_N) \geq \frac{\pi}{N}$ because we have

$$d(G_N) = \max_i \theta^{(g)}_{i+1} - \theta^{(g)}_i \geq \frac{\sum_{i=0}^{N-1} \theta^{(g)}_{i+1} - \theta^{(g)}_i}{N} = \frac{\pi}{N} \tag{4.5}$$

4.2 Sparse Representation

As described above we are going to estimate the set of n indexes $I = \{i_1, i_2, \ldots, i_n\}$ by assuming (4.4) as a model. The corresponding directions will be the estimates of the true directions in (3.3). Now, (4.4) can be written as

$$\begin{aligned} \mathbf{x}(t) &= \sum_{i \in I} \mathbf{a}(\theta^g_i) s'_k(t) + \mathbf{n}'(t) \\ &= \sum_{i \in I} \mathbf{a}(\theta^g_i) s'_k(t) + \sum_{i \notin I} \mathbf{a}(\theta^g_i) 0 + \mathbf{n}'(t) \end{aligned} \tag{4.6}$$

Defining

$$s^g_i(t) = \begin{cases} s'_k(t) & i = i_k \in I \\ 0 & i \notin I \end{cases} \tag{4.7}$$

(4.6) can be simplified as

$$\mathbf{x}(t) = \sum_{i=0}^{N} \mathbf{a}(\theta^g_i) s^g_i(t) + \mathbf{n}'(t) \tag{4.8}$$

Roughly speaking, although the received vector, $\mathbf{x}(t)$ is very close to the span of n basis vectors, it can also be interpreted as a linear combination of all N basis vectors. Note that because $N > m$ the set of all basis vectors[1] are not linearly

[1]Such an overcomplete basis is usually called a frame, but here we stick to the more informal nomenclature.

independent which indicates that the linear decomposition is not unique. However, in the desired decomposition most of the components, s_i^g are equal to zero. More precisely, the number of nonzero elements in the vector $\mathbf{s}^g = [s_1^g \; s_2^g \; \dots s_N^g]^T$, known as the ℓ_0 norm and denoted by $\|\mathbf{s}^g\|_0$ is equal to n. Such a vector with many zero entries may be referred to as a *sparse* vector. The nonzero elements of $\mathbf{s}^g$ are known as *active sources* or *active indexes*. In summary the sparse representation of the model (3.3) can be written as

$$\mathbf{x}(t) = \mathbf{A}^g \mathbf{s}^g(t) + \mathbf{n}'(t) \qquad , \qquad \|\mathbf{s}^g(t)\|_0 = n \tag{4.9}$$

where $\mathbf{A}^g$ is the matrix of all basis vectors in the grid. Under which constraints is (4.9) a well behaved model? To answer this question we first introduce the following definition.

Definition 2. A matrix $\mathbf{A}$ is said to be *k-ambiguous* iff

- The submatrix of any k columns of $\mathbf{A}$ is full-rank.
- There exists a submatrix of $k+1$ columns of $\mathbf{A}$ which is rank deficient.

Note that a model is well-behaved if for each realization of the stochastic parameters the unknown deterministic parameters has a unique solution in terms of the known and stochastic ones. However, the whole unknown parameters may not have a unique solution in terms of the known ones. The following theorem gives a criterion for the uniqueness of $\mathbf{s}^g$ in (4.9).

Theorem 1. *Suppose $\mathbf{A}^g$ is a k-ambiguous matrix. In (4.9), $\mathbf{s}^g$ has a unique solution in terms of $\mathbf{x}$ and $\mathbf{n}'$ if and only if $n \leq \frac{k}{2}$.*

Proof. First assume that $n \leq \frac{k}{2}$ and there exists two different vectors, $\mathbf{s}_1^g$, and $\mathbf{s}_2^g$ satisfying the model. Then we have

$$\mathbf{x}(t) - \mathbf{n}'(t) = \mathbf{A}^g \mathbf{s}_1^g = \mathbf{A}^g \mathbf{s}_2^g \tag{4.10}$$

which implies that $\mathbf{A}^g(\mathbf{s}_1^g - \mathbf{s}_2^g) = 0$. On the other hand

$$\|\mathbf{s}_1^g\|_0 = \|\mathbf{s}_2^g\|_0 \leq \frac{k}{2} \rightarrow \|\mathbf{s}_1^g - \mathbf{s}_2^g\|_0 \leq k \tag{4.11}$$

This means that the vector $\mathbf{A}^g(\mathbf{s}_1^g - \mathbf{s}_2^g) = 0$ is a linear combination of at most k columns of $\mathbf{A}^g$. Because $\mathbf{A}^g$ is k-ambiguous, these columns are linearly independent so $\mathbf{s}_1^g - \mathbf{s}_2^g = 0$.

Next, suppose $n > \frac{k}{2}$. Then, there exists a set of $k+1$ linearly dependent columns in $\mathbf{A}^g$. In other words there exists a vector $\mathbf{d}^g$ with $\|\mathbf{d}^g\|_0 = k+1$ and $\mathbf{A}^g\mathbf{d}^g = 0$. Now because $n > \frac{k}{2}$ we can easily construct two vectors $\mathbf{s}_1^g$, and $\mathbf{s}_2^g$ so that $\|\mathbf{s}_1^g\|_0 = \|\mathbf{s}_2^g\|_0 = n$ and $\mathbf{d}^g = \mathbf{s}_1^g - \mathbf{s}_2^g$. As can be seen, both of these vectors are equivalent in the model (4.9) which proves the converse part. □

Remark. *For the case of m sensors in a ULA, $\mathbf{A}^g$ is m-ambiguous because each m columns of $\mathbf{A}^g$ form a Vandermonde matrix (Horn and Johnson 1991) which is always full rank. In this case the criterion is $n \leq \frac{m}{2}$ as stated in Chapter 3.*

4.3 Maximum Likelihood in Sparse Framework

It is now natural to express the ML estimator introduced in (3.5) in terms of the new vectors of sparse representation. Note that (3.5) combines the data generated in different snapshots by adding up the noise magnitudes. However, the desired directions, $\boldsymbol{\theta}$ are temporally constant. On the other hand, the ℓ_0 norm of $\mathbf{s}^g(t)$ does not give any information about the places of the sources. We are not only concerned about the number of active sources at each snapshot, but also the places of them. In other words we require an inactive index to be zero at all times and if a vector is active at some time it should be treated as an active source for all times. Accordingly, we call an index, a *generally active* if it is active at least in one snapshot. Introducing

$$\gamma_i = \sqrt{\sum_{t=1}^{T} \frac{1}{T} |s_i^g(t)|^2} \tag{4.12}$$

it can be seen that an index, i is generally active if and only if $\gamma_i \neq 0$. Thus, the number of generally active sources is equal to $\|\boldsymbol{\gamma}\|_0$ where $\boldsymbol{\gamma} = [\gamma_1 \ \gamma_2 \ \dots \ \gamma_N]^T$. Using this definition we can express the ML estimator as follows

$$\hat{\mathbf{s}}^g = \arg\min_{\{\mathbf{s}^g(t)\}_{t=1}^T} \tfrac{1}{T} \textstyle\sum_{t=1}^{T} \|\mathbf{x}(t) - \mathbf{A}^g \mathbf{s}^g(t)\|_2^2$$

$$\text{subject to} \tag{4.13}$$

$$\|\boldsymbol{\gamma}\|_0 = n$$

To solve this optimization one may attempt to perform an exhaustive search over all possible sets of generally active basis, and for each set compute the least square solution using the first equation of (3.6). The estimated active indexes are those that minimize (4.13) among all possible choices. However, this method is computationally costly, especially as the number of grids and sensors increase. Note that by reformulating ML as (3.6), the problem of local minima is not solved. Each combination of active sources in the exhaustive search basically gives a local minimum in ML.

4.4 Solution Based on ℓ_1 Regularization

The optimization introduced in (4.13) can be viewed as a regularization, which is basically adding more constraints in order to make the solution unique. Note that the sparsity condition in (4.9) is essential because without it the uniqueness of the model is lost. This also can be seen in the ML equation (4.13), where without the constraint $\|\boldsymbol{\gamma}\|_0 = n$ the least square criterion can be driven to zero because the range space of $\mathbf{A}^g$ is the whole of $\mathbb{R}^m$. Thus, it an be said that (4.13) is a linear least square method regularized by the constraint $\|\boldsymbol{\gamma}\|_0 = n$. In this chapter we first give a brief introduction to the regularization technique.

4.4.1 Regularization

A regularization is an additional constraint, usually in the form of a cost function which can be expressed in different equivalent forms. However, there always exists a so called *regularization parameter*, which controls the relative importance of the additional cost. Suppose a function $F(\theta)$ is to be optimized under an additional constraint of $G(\theta) = \mu$, where μ is the regularization parameter. Common methods of expressing the regularization are

$$\min_{G(\theta)=\mu} F(\theta) \tag{4.14}$$

and

$$\min_{\theta} F(\theta) + \lambda G(\theta) \tag{4.15}$$

The term $F(\theta) + \lambda G(\theta) = L(\theta, \lambda)$ is known as the Lagrangian form of the optimization. There is a close relationship between the solutions of (4.14) and (4.15), which is stated as follows.

Theorem 2. *Suppose for some value of λ the solution of (4.15) is $\theta = \theta_0$. Then the solution of (4.14) with $\mu = G(\theta_0)$ is $\theta = \theta_0$ as well.*

Proof. Suppose the solution to (4.14) with $\mu = G(\theta_0)$ is $\theta = \theta_1$ then $G(\theta_1) = \mu = G(\theta_0)$ and

$$F(\theta_1) \leq F(\theta_0) \tag{4.16}$$

On the other hand

$$F(\theta_0) + \lambda G(\theta_0) \leq F(\theta_1) + \lambda G(\theta_1) \rightarrow F(\theta_0) \leq F(\theta_1) \tag{4.17}$$

combining the two above results, we see that $F(\theta_0) = F(\theta_1)$. This result shows that $\theta = \theta_0$ is a global minimum point for (4.14) as well. Further note that if we know that the global minimum is unique then θ_0 is the unique one. Otherwise, there might be other global minima, θ_1 with $F(\theta_0) = F(\theta_1)$. □

To illustrate the result of Theorem 2 we apply it to the sparse ML optimization. In this case the linear expression (4.15) can be written as

$$\min_{\{\mathbf{s}^g(t)\}_{t=1}^T} \sum_{t=1}^{T} \|\mathbf{x}(t) - \mathbf{A}^g \mathbf{s}^g(t)\|_2^2 + \lambda \|\boldsymbol{\gamma}\|_0 \tag{4.18}$$

By changing λ we are able to change the solution of the above equation. If for some λ the number of generally active indexes, $\|\boldsymbol{\gamma}\|_0$ is equal to n, then it is also the solution to the sparse ML, problem, (4.13). However, there might not exist any λ giving $\|\boldsymbol{\gamma}\|_0 = n$. The next theorem equips us with a tool to search efficiently for a suitable regularization parameter, λ.

Theorem 3. *In (4.15), $G(\theta)$ is a decreasing function of λ meaning that if θ_1 and θ_2 are the solutions for λ_1 and λ_2 respectively with $\lambda_1 \leq \lambda_2$ then $G(\theta_2) \leq G(\theta_1)$.*

Proof. From the minimality of θ_1 and θ_2 we can conclude

$$F(\theta_1) + \lambda_1 G(\theta_1) \leq F(\theta_2) + \lambda_1 G(\theta_2) \tag{4.19}$$

and

$$F(\theta_2) + \lambda_2 G(\theta_2) \leq F(\theta_1) + \lambda_2 G(\theta_1) \tag{4.20}$$

Manipulating these two equations we find the following result.

$$\lambda_1(G(\theta_1) - G(\theta_2)) \leq F(\theta_2) - F(\theta_1) \leq \lambda_2(G(\theta_1) - G(\theta_2)) \tag{4.21}$$

Because $\lambda_1 \leq \lambda_2$ we conclude that $G(\theta_1) - G(\theta_2) \geq 0$ which proves the theorem. □

4.4.2 Regularization Using the ℓ_1 Norm

The ML estimator introduced in (4.13) and (4.18) is still computationally inappropriate. Hence, an approximation to these optimization problems should be made. The relation between the ℓ_1 and ℓ_0 norm has been known and studied for a long time. The ℓ_1 norm of a vector is the sum of the absolute values of its components

$$\|\mathbf{s}\|_1 = \sum_{i=1}^{n} |s_i| \tag{4.22}$$

The ℓ_1 norm has been used for some basic linear regression problems. Basically, what to do is to substitute the ℓ_0 norm with the ℓ_1 norm, and it is expected to get a sparse solution due to the singularity of the absolute function at zero. The behavior and consistency of this method will be discussed in the next chapter in details. The ℓ_1 regularization may be written either as

$$\begin{gathered} \hat{\mathbf{s}}^g = \arg \min_{\{\mathbf{s}^g(t)\}_{t=1}^T} \sum_{t=1}^T \|\mathbf{x}(t) - \mathbf{A}^g \mathbf{s}^g(t)\|_2^2 \\ \text{subject to} \quad \|\boldsymbol{\gamma}\|_1 \leq \mu \end{gathered} \tag{4.23}$$

or

$$\hat{\mathbf{s}}^g = \min_{\{\mathbf{s}^g(t)\}_{t=1}^T} \sum_{t=1}^{T} \|\mathbf{x}(t) - \mathbf{A}^g \mathbf{s}^g(t)\|_2^2 + \lambda \|\boldsymbol{\gamma}\|_1 \tag{4.24}$$

The main advantage with this approximation is that both (4.23) and (4.24) belong to the class of convex optimization problems, which are generally solvable in a quite fast way. More accurately, both can be reformulated to the Second Order Cone (SOC) programming framework, which provides the opportunity for implementation using SOC toolboxes such as the SEDUMI toolbox (Sturm 1998).

4.4.3 The overall algorithm

Because the ℓ_1 regularization is an approximation of the ℓ_0 one, once the estimation problem is solved by (4.23) or (4.24), the generally active indexes can be obtained from this solution. Note that the $\boldsymbol{\gamma}$ vector is the estimated RMS value received from each direction, and can be treated as the spatial spectrum of the data. So, the directions can be found by finding the peaks of this spectrum and thresholding. The regularization parameter should be chosen so that the number of the generally active sources is equal to n. Although Theorem 3 states that for these optimization problems, $\|\boldsymbol{\gamma}\|_1$ is a decreasing function of the regularization parameter, λ we expect the $\|\boldsymbol{\gamma}\|_0$ to be decreasing as well. This fact enables us to perform the search for the appropriate λ faster using a bisection method. In summary the method can be explained as follows

1. Choose an initial regularization parameter and a threshold, τ.
2. Use (4.23) or (4.24) to get an estimate of $\mathbf{s}^g(t)$.
3. Evaluate the $\boldsymbol{\gamma}$ vector using (4.12) and locate its peaks.
4. Choose the peaks in γ which are greater than τ. If the number of peaks is not equal to n, modify λ due to the monotonicity property and go to Step 2.
5. Given the estimated directions, the source waveforms are estimated using the first equation of (3.6).
6. It is also possible to modify $\boldsymbol{\gamma}$ using the newly estimated sources, and recompute peaks and go to step 4 until the estimated active indexes are not changed anymore.
7. End.

4.5 ℓ_1-regularization as a MAP estimator

Many regularized conditions can be viewed as Maximum A Posteriori (MAP) criteria, by assuming the additional cost function as a result of a prior knowledge about the data. This context is commonly referred to as a Bayesian estimation. Following the work of (Tibshirani 1996), in this section we give a stochastic view of the ℓ_1-regularization for later use. First we introduce the complex Laplace distribution

Definition 3. A complex random variable S is said to be distributed according to a complex Laplacian distribution if

$$f_S(s) = \frac{\alpha}{2\pi} e^{-\alpha|s|} \tag{4.25}$$

For a vector, $\mathbf{S}$, of k independent Laplacian random variables the joint p.d.f will be

$$f_{\mathbf{S}}(\mathbf{s}) = \left(\frac{\alpha}{2\pi}\right)^k e^{-\alpha\|\mathbf{s}\|_1} \tag{4.26}$$

Suppose in the model (4.9) the sources are Laplacian and the noise is Gaussian. Then the a posterior probability of $\boldsymbol{S}$ given $\boldsymbol{X}$ can be written as

$$f_{\mathbf{S}|\mathbf{X}}(\mathbf{s}|\mathbf{x}) = \frac{f_{\mathbf{X}|\mathbf{S}}(\mathbf{x}|\mathbf{s})f_{\mathbf{S}}(\mathbf{s})}{f_{\mathbf{X}}(\mathbf{x})} \propto e^{-\frac{\|\mathbf{x}-\mathbf{A}^g\mathbf{s}\|^2}{2\sigma^2}-\alpha\|\mathbf{s}\|_1} \tag{4.27}$$

The MAP estimate is defined as the maximizing argument of the a posteriori distribution. In this case the MAP estimator can therefore be written as

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \|\mathbf{x} - \mathbf{A}^g\mathbf{s}\|^2 + 2\alpha\sigma^2\|\mathbf{s}\|_1 \tag{4.28}$$

The relation to (4.15) is obvious by putting $\lambda = 2\alpha\sigma^2$. Note that the real sparse source, $\mathbf{s}^g$ is in general not a Laplacian random vector, and this shows the main difference between ℓ_1 and ℓ_0 optimization. How can the solution to a MAP estimator be different from the prior distribution? The answer is that although the MAP estimator is designed to combine the prior information by the ML estimator to get a typical solution by this prior knowledge, its performance is guaranteed only when the known variables are typical by the given model. As can be seen the observed vector, $\mathbf{x}$ in our model is actually made by a sparse vector, $\mathbf{s}^g$ which would be unlikely to occur if $\mathbf{s}^g$ were a Laplace vector. In other words ℓ_1 regularization utilizes MAP in a data space region where the performance of MAP is not guaranteed. However it still works well. The intuition we made in this chapter will be used later on.

5 IMPLEMENTATION ISSUES

In this chapter we will review some practical issues related to the ℓ_1 regularization method for linear regression problems. In the previous chapter we introduced some logical steps to devise this method and gave a complete procedure of implementing such an algorithm. In this section we focus on the method of implementation within the SOC framework, and then we discuss the choice of regularization parameter when the number of sources is an unknown deterministic parameter.

5.1 SOC Adaptation

Now, we are going to explain how to represent the ℓ_1 regularization in a SOC framework. We first need to define a Second Order Cone (SOC)(Boyd and Vandenberghe 2004):

Definition 4. An m-dimensional second order cone, $\mathscr{C}$ is a generalization of the ordinary 3-dimensional cone as follows

$$\mathscr{C} = \left\{ (x_1, x_2, \ldots, x_m)^T \quad | \quad \sum_{i=1}^{m-1} |x_i|^2 \leq \rho^2 |x_m|^2 \right\} \tag{5.1}$$

Definition 5. Given a block vector $\mathbf{x} = \left[\mathbf{x}_1^T \ \mathbf{x}_2^T \ \ldots \ \mathbf{x}_k^T\right]^T$, where each of the vectors $\mathbf{x}_i$ is either a free vector or it is confined to a second order cone. An SOC program is an optimization over such an $\mathbf{x}$ of the form

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad &\mathbf{A}\mathbf{x} = \mathbf{b} \qquad \text{and} \qquad \mathbf{x}_i \in \mathscr{C}_i \end{aligned} \tag{5.2}$$

Such an optimization can be solved efficiently using computer toolboxes such as SeDuMi, which we will use for generating the simulation in Chapter 7. The question is how to represent the optimization problem in (4.23) and (4.24) within the SOC framefork. To do so we write down the sparse model as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}^g(t) + \mathbf{n}(t) = [\mathbf{A} \quad \mathbf{I}] \begin{bmatrix} \mathbf{s}^g(t) \\ \mathbf{n}(t) \end{bmatrix} \tag{5.3}$$

which forms the linear constraint in (5.2). Now, for each direction index $i = 1, \ldots, N$ we introduce the following cones

$$\mathscr{C}_i = \left\{ \begin{bmatrix} s_i^g(1) \\ s_i^g(2) \\ \ldots \\ s_i^g(T) \\ \gamma_i \end{bmatrix} : \sum_{t=1}^{T} |s_i^g(t)|^2 \leq \gamma_i^2 \right\} \tag{5.4}$$

and also

$$\mathscr{C}_N = \left\{ (\mathbf{n}(1), \mathbf{n}(2), \ldots, \mathbf{n}(T), \nu) : \sum_{t=1}^{T} \|n(t)\|_2^2 \leq \nu^2 \right\} \tag{5.5}$$

For (4.14) an additional linear equation of the following form should be added to the linear constraints

$$\sum_{i=1}^{N} \gamma_i = \mu \tag{5.6}$$

and the minimization (4.14) is easily written as

$$\begin{array}{c} \hat{\mathbf{s}} = \arg \min\limits_{\mathbf{s},\mathbf{n},\{\gamma_i\}_{i=1}^n,\nu} \nu \\ \text{subject to} \quad \mathbf{x}(t) = \mathbf{A}\mathbf{s}^g(t) + \mathbf{n}(t) \quad , \quad \sum_{i=1}^N \gamma_i = \mu. \end{array} \tag{5.7}$$

The minimization (4.15) can also be written as

$$\begin{array}{c} \hat{\mathbf{s}} = \arg \min\limits_{\mathbf{s},\mathbf{n},\{\gamma_i\}_{i=1}^n,\nu} \nu + \lambda \sum_{i=1}^N \gamma_i \\ \text{subject to} \quad \mathbf{x}(t) = \mathbf{A}\mathbf{s}^g(t) + \mathbf{n} \end{array} \tag{5.8}$$

5.2 SVD Based Dimension Reduction

Although ℓ_1 regularization reduces the complexity of estimation dramatically, its complexity grows fast as the number of snapshots grows, because the number of constraints grows fast. For a large number of snapshots, it is practically impossible to implement such a method. In this case it is proposed in (Malioutov 2003) to construct a small set of snapshots by linearly combining different snapshots and use these transformed snapshots as an input to the ℓ_1 regularization procedure. As can be seen from the model of the data, any linear combination of the sources results in a data with the similar model but a different noise level:

$$\sum \alpha(t)\mathbf{x}(t) = \mathbf{A} \sum \alpha(t)\mathbf{s}(t) + \sum \alpha(t)\mathbf{n}(t) \tag{5.9}$$

for any set of coefficients $\alpha(t)$. By a wise choice of these coefficients it is possible to increase the SNR, which will compensate the decrease in the number of snapshots. In other terms, we are to remove the noise subspace while keeping the signal subspace constant. This might not be possible, especially when the number of sources is unknown. However, it is possible to form a candidate signal subspace wider than the real one by assuming an upper bound for the number of the sources. This upper bound is dented by K. As proposed in (Malioutov 2003), we can use the SVD of the matrix $\mathbf{X} = [\boldsymbol{x}(1)\ \boldsymbol{x}(2)\ \ldots\ \boldsymbol{x}(T)]$ to find the candidate signal subspace. Note

that in the noise-free case, $\mathbf{X}$ is in the range space of $\mathbf{A}(\theta)$, $\mathscr{R}_A$. Let us compute the SVD of the data matrix $\mathbf{X}$:

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H \tag{5.10}$$

For low noise or a large number of snapshots we expect the dominant sources to be in relationship with the signal subspace, while the smaller singular values form the noise subspace. Introducing a matrix $\mathbf{D} = \begin{bmatrix} \mathbf{I}_{K\times K} \\ 0 \end{bmatrix}$, the signal subspace can be written as the range space of the following matrix.

$$\mathbf{X}' = \mathbf{U}\mathbf{\Sigma}\mathbf{D} = \mathbf{X}\mathbf{V}\mathbf{D} \tag{5.11}$$

We use each of the K columns of $\mathbf{X}'$ as one snapshot of data.

5.3 Choosing the regularization parameter

The relation between the regularization parameter and the number of sources is an unsolved problem, which might be regarded as a point for further research. As discussed in Chapter 4, the regularization parameter can be found by recursive use of the convex regularized optimization and updating the parameter, λ by a suitable method such as bisectioning search. Generally, such a method is computationally fast, because of both the fast implementation method of convex programming in each recursion, and fast convergence of the bisection method. Normally, there exists an interval of suitable parameter, λ in which the ℓ_1 regularization gives exactly the same n active basis. In such an interval, changing λ may only cause a change in the estimated waveforms, $\mathbf{s}(t)$. However, when the number of sources is unknown, it should be estimated first. From this point of view the estimation of the regularization parameter with an unknown number of sources is fundamentally a model order selection problem.

There is a variety of methods for model order selection such as the General Information Criterion (GIC)(Stoica *et al.* 2004), Bayesian Information Criterion (BIC)(Akaike 1973) , Minimum Description Length (MDL)(Wax and Ziskind 1989), etc. All of these methods result in an additional optimization, which is basically a least-square criterion regularized by some function of the estimated number of the sources. In this section we explain the MDL method in details, and then apply it to the ℓ_1 estimation framework. Finally we will discuss the possibility of modifying the MDL criterion by taking the MAP point of view of the ℓ_1 regularization into account.

5.3.1 Minimum Description Length

MDL is a method for choosing a model that best describs the received data among a set of possible models. There exists a variety of intuitions behind this method all end-

ing up with the same formulation. The initial idea, introduced by (Rissanen 2006), is to find a minimum-length expression, with which the original data is reconstructible. A general formulation of such a problem is to find the minimum-length input code to a universal Turing machine (Boolos and Jeffrey 1989) which gives the desired data as the output. This minimum length is known as the *Kolmogorov complexity* (Cover and Thomas 2006). However, it is shown that such a general problem is unsolvable (Li and Vitanyi 1997). Thus, Rissanen introduced a more confined question of the minimal addressing for a piece of data using a set of predefined models (Rissanen 1986). This minimal address is well-known as the *Universal Complexity* of the data associated with the given model. The goal is then to find the model with minimum universal complexity among a set of models. To give a formal representation we introduce the following definitions.

Definition 6. A model $\mathscr{M} = \{f(\mathbf{x};\theta)|\theta \in \Theta\}$ is an indexed set of distributions. The index set Θ can vary from a finite set to $\mathbb{R}^n$. In our case of discretized DOAs, it can be the set of all n distinct indexes.

Now we define a class of encoding techniques for the received data and give the length of the representations due to this encoding. For a continuous set of data, the encoder first quantizes the message to N levels. Although for our encoders the length of the encoded data, $L^{(N)}(\mathbf{x})$ tends to infinity as N grows, but its difference to the constant-length coding length, $L_{\mathrm{CL}}(\mathbf{x}) = \log N$ tends to a finite number, which may be called the *differential length* of $\mathbf{x}$. The differential length shows how many bits in length we gain by a special encoding scheme.

Our scheme is actually the prefix codes introduced and discussed in (Cover and Thomas 2006). For such a code we can find the differential length as follows.

Definition 7. Related to each p.d.f $f(\mathbf{x};\theta)$ of the model $\mathscr{M}$, the *Stochastic Complexity* $L_\theta(\mathbf{x})$ is the differential length of the optimal prefix code according to $f(\mathbf{x};\theta)$, and is given by

$$L_\theta(\mathbf{x}) = -\log f(\mathbf{x};\theta) \tag{5.12}$$

The definition of stochastic complexity brings some new interpretations to the previously discussed stochastic methods. For instance, the ML estimate can be interpreted as finding the minimum-length representation of data over all possible likelihoods. Note that for a model $\mathscr{M}$, ML gives the best index, θ matching the data. To sum up, from a compression point of view, ML estimation implies that a model is matched to the data if its representation length is as small as possible, which is a scientific statement of Occam's Razor[1].

5.3.2 Universal Coding and Universal Modeling

As discussed in the previous section, there is a close relationship between the code length of a data set and the best parameter selection for it. It is also possible to

[1]Occam's razor is a principle that the simplest explanation is usually the true one

talk about how much a data is related to a model when the parameter is unknown. From the duality, it is identical to ask how much this data can be compressed in the model in the same situation. It turns out that by choosing an appropriate distribution $f_{\mathbf{X}}(\mathbf{x})$, we should construct a "best" universal code regardless of the parameters, and find the stochastic complexity of the data, $\bar{L}(\mathbf{x})$ due to this code. We are not going to probe the details carefully here. However, in (Shtarkov 1987) it is shown that the best universal code in the sense of two-stage-addressing is given by the Normalized Maximum Likelihood (NML) as follows

$$f_{\mathrm{NML}}(\mathbf{x}) = \frac{f(\mathbf{x};\hat{\theta}(\mathbf{x}))}{\int f(\mathbf{x};\hat{\theta}(\mathbf{x}))\mathrm{d}^m\mathbf{x}} \tag{5.13}$$

where $\hat{\theta}(\mathbf{x})$ is the ML estimate of the parameter $\boldsymbol{\theta}$ based on the data, $\mathbf{x}$. The NML is a natural universal model for the data, since it assigns the highest possible likelihood to each data vector $\mathbf{x}$. Because the resulting function does not have the identity integral it should be normalized. For the optimal universal model, the stochastic complexity of $\mathbf{x}$ is called the Universal Stochastic Complexity (USC), and is given by

$$\mathrm{USC}(\mathbf{x}) = -\log f_{\mathrm{NML}}(\mathbf{x}) = -\log f(\mathbf{x};\hat{\theta}(\mathbf{x})) + \log \int f(\mathbf{x};\hat{\theta}(\mathbf{x}))\mathrm{d}^m\mathbf{x} \tag{5.14}$$

From the duality of the description length and model fitting, one can propose the USC as a measure of the fit of the model to data. If there exists a set of models for a given data, the one with the smallest USC should be chosen. This is exactly the MDL model order selection principle.

5.3.3 MDL for DOA Estimation

Now we apply the MDL criterion derived in (5.14) to our problem of interest. First we give a theorem first developed by Risannen to approximate the second term in (5.14).

Theorem 4. *For the case of an ν-dimensional continuous parameter vector and an μ-dimensional observation vector we asymptotically have*

$$\log \int f(\mathbf{x};\hat{\theta}(\mathbf{x}))\, d^{\mu}\mathbf{x} \approx \frac{\nu}{2}\log\frac{\mu}{2\pi} + \int \sqrt{|I(\theta)|}d\theta \tag{5.15}$$

where $|I(\theta)|$ is the determinant of the Fisher information matrix, I (Cover and Thomas 2006) evaluated at θ.

Proof. The proof is given in (Rissanen 1996). □

In spite of the above approximation, computing the Fisher information and the integral is complex. Instead, the whole stochastic complexity can be approximated by the following expression for a large number, T of m-dimensional observed vectors as shown in (Rissanen 1996).

$$\begin{aligned} \mathrm{USC}(\mathbf{x}) &= \log f(\mathbf{x};\hat{\theta}(\mathbf{x})) + \frac{1}{2}\nu \log \mu \\ &= -\frac{\sum_{t=1}^{T} \|\mathbf{x}(t) - \mathbf{A}^g \mathbf{s}(t)^g\|_2^2}{\sigma^2} + mT \log (2\pi\sigma^2) + \frac{1}{2} nT \log mT \end{aligned} \tag{5.16}$$

Closely related to Bayesian Information Criterion (BIC), This approximation is widely used, and gives very good result when the number of snapshots is high enough for the asymptotic code length assumption to hold. In this equation the noise variance, σ^2 is treated as a known variable. However, in many problems this is not true and the MDL criterion should be recomputed using the ML estimator of σ^2. The new MDL is easily computed by substituting $\hat{\sigma^2} = \frac{\sum_{t=1}^{T} \|\mathbf{x}(t) - \mathbf{A}^g \mathbf{s}(t)^g\|_2^2}{mT}$ into (5.16). Neglecting the constant terms, the MDL criterion can be expressed as

$$\mathrm{MDL}(\mathbf{x}, n) = mT \log \sum_{t=1}^{T} \|\mathbf{x}(t) - \mathbf{A}^g \mathbf{s}^g(t)\|_2^2 + \frac{1}{2} nT \log mT \tag{5.17}$$

5.4 A New ML Estimator of λ

In the previous chapter we gave a method for choosing a proper regularization parameter using its relationship to the number of sources. The MDL technique used for this purpose is based on the model fitting to the data, which requires a large number of received data. However, the MAP condition introduced in chapter 4 can help to get an alternative estimate of the regularization parameter. As discussed before, the ℓ_1 regularization method can be expressed in terms of a maximum a posteriori probability with a Laplacian prior on the data. We also mentioned that because of the mismatch between the sparse model of the data and the Laplacian assumption, our MAP estimator is utilized in a data region which is not designed for. To overcome this problem we introduce a truncated prior as follows:

$$f_{\mathbf{S}^g}(\mathbf{s}^g) = \begin{cases} K e^{-\mu \|\mathbf{s}^g\|_1} & \|\mathbf{s}^g\|_0 = n \\ 0 & \|\mathbf{s}^g\|_0 \neq n \end{cases} \tag{5.18}$$

where K is a proper normalization scalar. When $\mathbf{s}^g$ is a complex n-dimensional vector, $K = (\frac{\lambda^2}{2\pi})^n$. The negative logarithm of the MAP criterion for $T = 1$ can be written as

$$-\log \left[f_{\mathbf{X}|\mathbf{S}}(\mathbf{x}|\mathbf{s}) f_{\mathbf{S}}(\mathbf{s}) \right] = \frac{\|\mathbf{x} - \mathbf{A}^g \mathbf{s}^g\|_2^2}{\sigma^2} + \mu \|\mathbf{s}^g\|_1 - n \log \frac{\mu^2}{2\pi} + m \log \pi\sigma^2 \tag{5.19}$$

This expression can also be written in terms of an alternative parameter $\lambda = \mu\sigma^2$ as

$$\begin{aligned} -\log\left[f_{\mathbf{X}|\mathbf{S}}(\mathbf{x}|\mathbf{s})f_{\mathbf{S}}(\mathbf{s})\right] &= \frac{\|\mathbf{x}-\mathbf{A}^g\mathbf{s}^g\|_2^2+\lambda\|\mathbf{s}^g\|_1}{\sigma^2} - n\log\frac{\lambda^2}{2\pi\sigma^4} + m\log\pi\sigma^2 \\ &= \frac{\|\mathbf{x}-\mathbf{A}^g\mathbf{s}^g\|_2^2+\lambda\|\mathbf{s}^g\|_1}{\sigma^2} + (m+2n)\log\sigma^2 - n\log\frac{\lambda^2}{2\pi} + m\log\pi \quad (5.20) \end{aligned}$$

Let us denote

$$\mathscr{K}(\lambda) = \min_{\mathbf{s}^g} \|\mathbf{x}-\mathbf{A}^g\mathbf{s}^g\|_2^2+\lambda\|\mathbf{s}^g\|_1 \quad (5.21)$$

which can be computed by convex optimization methods (Boyd and Vandenberghe 2004). The Maximum a posteriori probability is obtained by minimizing (5.20) over the unknown parameters, λ, σ^2

$$\hat{\lambda}, \hat{\sigma}^2 = \arg\min_{\lambda,\sigma^2} \frac{\mathscr{K}(\lambda)}{\sigma^2} + (m+2n)\log\sigma^2 - n\log\frac{\lambda^2}{2\pi} \quad (5.22)$$

Solving with respect to σ^2 we get

$$\hat{\sigma}^2 = \frac{\mathscr{K}(\lambda)}{m+2n} \quad (5.23)$$

and substituting into (5.22) we have

$$\hat{\lambda} = \arg\min_{\lambda} (m+2n)\log\frac{\mathscr{K}(\lambda)}{m+2n} + n(2-\log\frac{\lambda^2}{2\pi}) \quad (5.24)$$

Note that since n is implicitly a function of λ, all terms in (5.24) are needed. The minimization in (5.24) is our prosed new method to estimate λ, and at the same time the number of sources. If is important to note that in practice a threshold must be used to select n as the number of significant values in $\mathbf{s}^g$.

6 CONSISTENCY ANALYSIS

We introduced the ℓ_1 regularization technique as an approximation to the exact ℓ_0 one. Many questions may arise concerning the performance of this new method. In this section we are going to discuss the general behavior of the ℓ_1 regularization by analyzing its solution for some asymptotic cases. We also give a detailed discussion on the consistency of this method in two asymptotic cases, namely large number of sensors and high SNR respectively.

6.1 Mathematical Representation of the Solution

In this section we focus on the solution of (4.13). One may try to minimize this using standard differentiation. However, because the absolution function is non-differentiable at zero, this method fails. Although in (Malioutov 2003), a closed form differential solution is given, it is beneficial from neither a practical nor an analytical point of view. Note that the ℓ_1 regularization is valuable because of its singularity at zero, since most of the optimum parameters will be zero. Alternatively, assuming nonzero values for some special indexes and fixing others to zero we can solve the optimization by differentiating. Then, for each choice of active indexes a minimum is obtained and the global one is chosen later. Mathematically, one can propose for each choice, $I = \{i_1, i_2, \dots, i_n\}$ of n indexes and the corresponding $\boldsymbol{\theta} = \{\theta^g_{i_1}, \theta^g_{i_2}, \dots, \theta^g_{i_n}\}$, to solve

$$\{\hat{\mathbf{s}}(t, \boldsymbol{\theta})\}_{t=1}^{T} = \arg\min_{\mathbf{s}(t)} \frac{1}{T} \sum_{t=1}^{T} \|\mathbf{x}(t) - \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t)\|_2^2 + \lambda \sum_{i=1}^{n} \gamma_i \tag{6.1}$$

where

$$\gamma_i = \sqrt{\frac{1}{T} \sum_{t=1}^{T} |s_i(t)|^2} \tag{6.2}$$

with nonzero components of $\mathbf{s}(t)$. This might be interpreted as a marginal minimum for (4.13). Then the minimum cost function for this choice of active indexes can be computed as

$$V_{\text{LASSO}}(\boldsymbol{\theta}) = \frac{1}{T} \sum_{t=1}^{T} \|\mathbf{x}(t) - \mathbf{A}(\boldsymbol{\theta})\hat{\mathbf{s}}(t, \boldsymbol{\theta})\|_2^2 + \lambda \sum_{i=1}^{n} \sqrt{\frac{1}{T} \sum_{t=1}^{T} |\hat{s}_i(t, \boldsymbol{\theta})|^2} \tag{6.3}$$

Then, the active basis vectors of the solution to (4.13) are given by the minimum point of $V_{\text{LASSO}}(\boldsymbol{\theta})$. Practically, this method is equivalent to the exhaustive search introduced in Chapter 3 which is not favorable. On the other hand, we will get a good analytical overview of the method. Also, note that due to the convex nature of (4.13) there exists a unique local minimum which is also the global one. This shows

that for a certain choice of active basis different from that of the global minimum, either local minima over the waveform, $\mathbf{s}$ do not exist (which means that some indexes should be neglected) or this minimum value can be reduced more by adding more indexes. We summarize the result of the minimization for each active index set in the following theorem.

Theorem 5. *Consider the optimization problem (6.1). For each generally active index $I = \{i_1, i_2, \ldots, i_n\}$, and the corresponding active directions $\theta_k = \theta^g_{i_k}$ the solution for $\hat{\mathbf{s}}(t, \boldsymbol{\theta})$ is given by*

$$\hat{\mathbf{s}}(t,\boldsymbol{\theta}) = \left(\frac{\lambda}{2}\,\mathbf{I} + \boldsymbol{\Gamma}\mathbf{A}^H(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})\right)^{-1} \boldsymbol{\Gamma}\mathbf{A}^H(\boldsymbol{\theta})\mathbf{x}(t) \tag{6.4}$$

$$= \boldsymbol{\Gamma}\mathbf{A}(\boldsymbol{\theta})\left(\frac{\lambda}{2}\mathbf{I} + \mathbf{A}(\boldsymbol{\theta})\boldsymbol{\Gamma}\mathbf{A}^H(\boldsymbol{\theta})\right)^{-1} \mathbf{x}(t) \tag{6.5}$$

where $\boldsymbol{\Gamma}$ is a diagonal matrix with $\Gamma_{ii} = \gamma_i \geq 0$ which are given in (6.2). Denoting $\frac{\lambda}{2}\mathbf{I} + \mathbf{A}(\boldsymbol{\theta})\boldsymbol{\Gamma}\mathbf{A}^H(\boldsymbol{\theta}) = \boldsymbol{\Lambda}$ and to compute $\boldsymbol{\Gamma}$ we also have

$$\mathbf{a}^H(\theta_i)\boldsymbol{\Lambda}^{-1}\mathbf{R}_x\boldsymbol{\Lambda}^{-1}\mathbf{a}^H(\theta_i) = 1 \qquad i = 1, 2, \ldots, n \tag{6.6}$$

where $\mathbf{R}_x = \frac{1}{T}\sum_{t=1}^{T}\mathbf{x}(t)\mathbf{x}^H(t)$.

Proof. Equating the derivative with respect to the real and imaginary parts of $\mathbf{s}(t)$ to zero and using the definition (6.2) we have

$$2\mathbf{A}^H(\boldsymbol{\theta})(\mathbf{x}(t) - \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t)) = \lambda\boldsymbol{\Gamma}^{-1}\mathbf{s}(t) \qquad t = 1, 2, \ldots, T \tag{6.7}$$

which implies that

$$\begin{aligned}\mathbf{s}(t) &= \left(\frac{\lambda}{2}\boldsymbol{\Gamma}^{-1} + \boldsymbol{A}^H(\boldsymbol{\theta})\boldsymbol{A}(\boldsymbol{\theta})\right)^{-1}\mathbf{A}^H(\boldsymbol{\theta})\mathbf{x}(t) \\ &= \left(\frac{\lambda}{2}\mathbf{I} + \boldsymbol{\Gamma}\boldsymbol{A}^H(\boldsymbol{\theta})\boldsymbol{A}(\boldsymbol{\theta})\right)^{-1}\boldsymbol{\Gamma}\mathbf{A}^H(\boldsymbol{\theta})\mathbf{x}(t)\end{aligned} \tag{6.8}$$

from the identity

$$(\mathbf{I} + \mathbf{PQ})^{-1}\,\mathbf{P} = \mathbf{P}\,(\mathbf{I} + \mathbf{QP})^{-1} \tag{6.9}$$

(6.8) can also be written as

$$\mathbf{s}(t) = \boldsymbol{\Gamma}\mathbf{A}^H(\boldsymbol{\theta})\left(\frac{\lambda}{2}\mathbf{I} + \mathbf{A}(\boldsymbol{\theta})\boldsymbol{\Gamma}\mathbf{A}^H(\boldsymbol{\theta})\right)^{-1}\mathbf{x}(t) = \boldsymbol{\Gamma}\mathbf{A}(\boldsymbol{\theta})\boldsymbol{\Lambda}^{-1}\mathbf{x}(t) \tag{6.10}$$

Now, from the definition of γ_i we have

$$\gamma_i^2 = \frac{1}{T}\sum_{t=1}^{T}|s_i(t)|^2 = \frac{1}{T}\gamma_i^2\sum_{t=1}^{T}|\mathbf{a}^H(\theta_i)\boldsymbol{\Lambda}^{-1}\mathbf{x}(t)|^2 \tag{6.11}$$

which also implies that

$$\frac{1}{T}\sum_{t=1}^{T}\mathbf{a}^H(\theta_i)\mathbf{\Lambda}^{-1}\mathbf{x}(t)\mathbf{x}^H(t)\mathbf{\Lambda}^{-1}\mathbf{a}(\theta_i)=1 \tag{6.12}$$

This proves the theorem by taking the constant terms with respect to t out from the summation. □

From Theorem 5, we can reach the solution of the exact maximum likelihood method by letting λ tend to zero in (6.4). We can infer that the regularization parameter, λ plays a contradictory role in the model. On one hand, it reduces the computational complexity by introducing a convex measure of the sparsity. On the other hand, it is expected that by introducing λ, the solution departs from the optimal ML solution. To discuss this behavior clearly, we try to simplify the solution in (6.4) for the special case of the orthogonal steering vectors.

6.1.1 The Case of Orthogonal Basis Vectors

Suppose all the basis vectors in the array manifold are orthogonal. This might be obtained by letting m go to infinity. Note that for such a situation to be possible, the number of steering vectors at the grid, N, should be less than the number of sensors, m. In this case, (4.9) is not under determined anymore and the exact ML estimate can be found using the fact that

$$\Pi^{\perp}_{\mathbf{A}(\theta)}=\mathbf{I}-\frac{1}{m}\sum_{i=1}^{n}\mathbf{a}(\theta_i)\mathbf{a}^H(\theta_i) \tag{6.13}$$

where we have used that $\|\mathbf{a}(\theta)\|_2^2 = m$. After some manipulations, the second equation of (3.5) can be written as

$$\hat{\boldsymbol{\theta}}=\arg\max_{\boldsymbol{\theta}}\sum_{i=1}^{n}\mathbf{a}^H(\theta_i)\mathbf{R}_{\mathbf{x}}\mathbf{a}(\theta_i) \tag{6.14}$$

which is solved by finding the n highest maxima (peaks) of the function $\mathbf{a}^H(\theta)\mathbf{R}_{\mathbf{x}}\mathbf{a}(\theta)$. This shows that for a fixed grid, the conventional beamforming method is asymptotically identical to ML as the number of sensors goes to infinity.

The general solution to the ℓ_1 regularization approach, given in (6.4), can also be simplified in the case of orthogonal basis vectors as follows.

Theorem 6. *If all steering vectors are orthogonal with* $\|\mathbf{a}(\theta)\|_2=\sqrt{m}$ *the active basis of the solution to (4.15) is given by the set*

$$\boldsymbol{\theta}_{th}(\lambda)=\left\{\theta_i^g\,|\sqrt{\mathbf{a}^H(\theta_i^g)\mathbf{R}_x\mathbf{a}(\theta_i^g)}\geq\frac{\lambda}{2}\ ,i=1,2,\ldots,N\right\} \tag{6.15}$$

Proof. We remind that estimating the directions is performed by first computing the marginal minima over waveforms, $\hat{\mathbf{s}}(t,\boldsymbol{\theta})$ for each possible active direction combination $\boldsymbol{\theta}$, then computing the cost function for each marginal minimum, and finally exhaustive search over all possible active direction combinations for the minimum value of the cost function, $V_{\text{LASSO}(\boldsymbol{\theta})}$ at the marginal minimum point. For some active directions there might not be any minimum, $\hat{\mathbf{s}}(t,\boldsymbol{\theta})$. Note that because for the directions related to the global minimum there exists a marginal one, we limit our exhaustive search to the active sets for which the marginal minimum exists. Such a direction set, $\boldsymbol{\theta}$, might be referred to as a *candidate* active set. In the case of orthogonal basis vectors, for every choice of $\boldsymbol{\theta}$ we have

$$\mathbf{A}^H(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta}) = m\,\mathbf{I} \tag{6.16}$$

where $\mathbf{I}$ is the identity matrix of a proper size. Then (6.4) can also be written as

$$\hat{\mathbf{s}}(t,\boldsymbol{\theta}) = \left(\frac{\lambda}{2}\mathbf{I} + m\boldsymbol{\Gamma}\right)^{-1}\boldsymbol{\Gamma}\mathbf{A}^H(\hat{\boldsymbol{\theta}})\mathbf{x}(t) \tag{6.17}$$

which further can be simplified as

$$\hat{s}_i(t,\boldsymbol{\theta}) = \frac{\gamma_i}{\frac{\lambda}{2}+m\gamma_i}\mathbf{a}^H(\theta_i)\mathbf{x}(t) = \eta_i\mathbf{a}^H(\theta_i)\mathbf{x}(t) \tag{6.18}$$

From the definition of γ_i we can after some manipulations write

$$\left(\frac{\lambda}{2}+m\gamma_i\right)^2 = \frac{1}{T}\sum_{t=1}^{T}|\mathbf{a}^H(\theta_i)\mathbf{x}(t)|^2 = \mathbf{a}^H(\theta_i)\mathbf{R}_x\mathbf{a}(\theta_i) \tag{6.19}$$

Note that we are limited to $\gamma_i > 0$ which also means that a certain choice of active basis, $\boldsymbol{\theta}$ with the corresponding indexes, I is a candidate set if and only if

$$\forall i \in I \quad \sqrt{\mathbf{a}^H(\theta_i^g)\mathbf{R}_x\mathbf{a}(\theta_i^g)} \geq \frac{\lambda}{2} \tag{6.20}$$

Or alternatively, $\boldsymbol{\theta} \subset \boldsymbol{\theta}_t(\lambda)$. Besides, for a candidate set, $\boldsymbol{\theta}$, one can compute the cost function at the local minimum, $\hat{\boldsymbol{s}}(t,\boldsymbol{\theta})$ in (4.15) as

$$\begin{aligned}
V_{\text{LASSO}}(\boldsymbol{\theta}) &= \frac{1}{T}\sum_{t=1}^{T}\|\mathbf{x}(t)-\mathbf{A}(\boldsymbol{\theta})\hat{\boldsymbol{s}}(t,\boldsymbol{\theta})\|_2^2 + \lambda\sum_{i=1}^{n}\gamma_i \\
&= \frac{1}{T}\sum_{t=1}^{T}\|\mathbf{x}(t)-\sum_{i=1}^{n}\hat{s}_i(t,\boldsymbol{\theta})\mathbf{a}(\theta_i)\|_2^2 + \lambda\sum_{i=1}^{n}\gamma_i \\
&= \sum_{t=1}^{T}\left\|\left[\underbrace{\mathbf{I}-\sum_{i=1}^{n}\eta_i\mathbf{a}(\theta_i)\mathbf{a}^H(\theta_i)}_{\mathbf{P}}\right]\mathbf{x}(t)\right\|_2^2 + \lambda\sum_{i=1}^{n}\gamma_i \\
&= \text{Tr}\left[\mathbf{P}\mathbf{R}_x\mathbf{P}^H\right] + \lambda\sum_{i=1}^{n}\gamma_i = \text{Tr}\left[\mathbf{P}^H\mathbf{P}\mathbf{R}_x\right] + \lambda\sum_{i=1}^{n}\gamma_i
\end{aligned} \tag{6.21}$$

Because the steering vectors are orthogonal we get

$$\mathbf{P}^H\mathbf{P} = \mathbf{I} + \sum_{i=1}^{N} (m\eta_i^2 - 2\eta_i)\mathbf{a}(\theta_i)\mathbf{a}^H(\theta_i) \tag{6.22}$$

Also note that from (6.19) we have

$$m\eta_i^2 - 2\eta_i = m\left(\frac{1}{m} - \eta_i\right)^2 - \frac{1}{m} = \frac{\frac{\lambda^2}{4m}}{(\frac{\lambda}{2} + m\gamma_i)^2} - \frac{1}{m} = \frac{\frac{\lambda^2}{4m}}{\mathbf{a}^H(\hat{\theta}_i)\mathbf{R}_x\mathbf{a}(\hat{\theta}_i)} - \frac{1}{m} \tag{6.23}$$

Combining this results with the previous form of the cost function we get

$$\begin{aligned} V_{\text{LASSO}}(\boldsymbol{\theta}) &= \text{Tr}[\mathbf{R}_x] + \frac{\lambda^2 n}{4m} - \frac{1}{m}\sum_{i=1}^{n} \mathbf{a}^H(\theta_i)\mathbf{R}_x\mathbf{a}(\theta_i) \\ &+ \frac{\lambda}{m}\sum_{i=1}^{n} \sqrt{\mathbf{a}^H(\theta_i)\mathbf{R}_x\mathbf{a}(\theta_i)} - \frac{\lambda^2 n}{2m} = \text{Tr}[\mathbf{R}] \\ &- \frac{1}{m}\sum_{i=1}^{n} \left(\sqrt{\mathbf{a}^H(\theta_i)\mathbf{R}\mathbf{a}(\theta_i)} - \frac{\lambda}{2}\right)^2 \end{aligned} \tag{6.24}$$

Let us denote the active direction set containing the global minimum by $\hat{\boldsymbol{\theta}}$. It is actually the DOA estimates at the grid. Earlier, we showed that the DOA estimate, $\hat{\boldsymbol{\theta}}$ is a subset of $\boldsymbol{\theta}_{\text{th}}(\lambda)$ because it contains a local minimum. Note that in (6.24) the terms related to different directions θ_i, are decoupled, so that adding more directions from $\boldsymbol{\theta}_{\text{th}}(\lambda)$ may decrease $V(\boldsymbol{\theta})$, because each element of $\boldsymbol{\theta}_{\text{th}}(\lambda)$ satisfies the criterion (6.20). Thus,

$$\forall \boldsymbol{\theta} \subset \boldsymbol{\theta}_t(\lambda) \quad V(\boldsymbol{\theta}) \geq V(\boldsymbol{\theta}_{\text{th}}(\lambda)) \tag{6.25}$$

which shows that $\boldsymbol{\theta}_t(\lambda)$ is the set of active basis related to the global minimum. □

The above theorem shows that the behavior of the ℓ_1 regularization method when the steering vectors are orthogonal is to form the spatial power spectrum, $P(\theta) = \mathbf{a}^H(\theta)\mathbf{R}\mathbf{a}(\theta)$ and threshold it by $\frac{\lambda^2}{4}$. By changing threshold it is possible to control the number of active directions. The higher the λ is the fewer sources are chosen. In order to get the true sources $\boldsymbol{\theta}_0$, within the active directions set, λ should be chosen so that

$$\max_{\theta \notin \theta_0} P(\theta) < \frac{\lambda}{2} \leq \min_{\theta \in \theta_0} P(\theta) \tag{6.26}$$

It is also worth noting that the solutions for the signal values in (6.18) are not equal to the true ones due to the scaling factor η_i. The scaling factor vanishes only when λ tends to zero. However, due to the criterion (6.26) this is generally impossible. Hence, we generally conclude that the ℓ_1 regularization method is not a good estimator of the waveforms, while it may detect the directions precisely. That is why in Chapter 4 we propose the LS estimate for the signal values after estimating the directions. Here, we introduce some examples of such a situation which may help to get more insight into the overall method.

Example 1 In the absence of noise, the model of the received data can be written as $\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta}_0)\mathbf{s}$. Then, we have

$$P(\theta) = \sqrt{\mathbf{a}^H(\theta)\mathbf{A}(\boldsymbol{\theta}_0)\mathbf{R}_s\mathbf{A}^H(\boldsymbol{\theta}_0)\mathbf{a}(\theta)} \tag{6.27}$$

where $\mathbf{R}_s = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbf{s}(t)\mathbf{s}^H(t)$, and from the orthogonality (6.27) can be simplified as

$$P(\theta) = \begin{cases} \sqrt{\frac{1}{T}\sum_{t=1}^{T} |s_i(t)|^2} & \theta = \theta_i \in \theta_0 \\ 0 & \theta \notin \theta_0 \end{cases} \tag{6.28}$$

Then the criterion for correct estimation in (6.26) becomes

$$0 < \lambda < 2 \min P(\theta_i) \tag{6.29}$$

It is better to chose a smaller value of λ so that the scaling factor η_i in (6.18) gets close to 1. However, it can not be zero.

Example 2 If the number of snapshots is large we can use the benefit of the Law of Large Numbers (LLN). Note that when $T \to \infty$ the matrix $\mathbf{R}_x$ converges to the true correlation matrix

$$\mathbf{R}_x = \mathbf{A}(\boldsymbol{\theta}_0)\mathbf{R}_s\mathbf{A}^H(\boldsymbol{\theta}_0) + \mathbf{R}_n \tag{6.30}$$

where $\mathbf{R}_s = \frac{1}{T}\sum_{t=1}^{T} \mathbf{n}(t)\mathbf{n}^H(t)$. Denoting

$$P_n(\theta) = \sqrt{\mathbf{a}^H(\theta)\mathbf{R}_n\mathbf{a}(\theta)} \tag{6.31}$$

we can write the criterion as

$$\max_{\theta\notin\boldsymbol{\theta}_0} P_n(\theta) < \frac{\lambda}{2} \leq \min_{\theta\in\boldsymbol{\theta}_0} \sqrt{P(\theta)^2 + P_n(\theta)^2} \tag{6.32}$$

Here, $P_n(\theta)$ is the spacial power spectrum of the noise, and it usually assumed to be uniform over all angles. In such a situation there always exists a proper λ for a consistent estimation.

6.1.2 Noiseless Solution

Another important limiting case is when the noise in the model is zero. It is generally expected for a "good" estimator in a well-behaved model to give the true estimate when the stochastic parameters are known. The noise-free case is such a situation in which the stochastic noise parameters are known to be zero. When the basis vectors are nonorthogonal, analyzing the method using the solution in (6.4) is complicated. Instead, for the noise free model another important method is proposed based on the uncertainty principle (Donoho and Huo 2001). Here we explain the main result for a general sparse linear regression problem, which can also be found in (Elad and Bruckstein 2002).

First, note that in this case the ℓ_1 regularization algorithm in (4.9) can be written as follows, due to the fact that the noise is known to be zero

$$\begin{aligned} \hat{\mathbf{s}}^g(t) &= \arg\min_{\mathbf{s}^g(t)} \|\gamma\|_1 \\ \text{subject to} \quad & \mathbf{A}^g\mathbf{s}^g(t) = \mathbf{x}(t) \end{aligned} \tag{6.33}$$

First, we develop a condition for the one snapshot case ($T = 1$). The generalization to the multiple snapshots case is straightforward. Suppose the received data $\mathbf{x}(t)$ is the linear combination of n distinct bases in $\mathbf{A}^g$ with indexes, $I = \{i_1, i_2, \ldots, i_n\}$. The following theorem of (Elad and Bruckstein 2002) gives a condition for the true estimation. Before introducing the theorem we define a new notation as follows

$$\|\mathbf{s}\|_I = \sum_{i \in I} |s_i| \tag{6.34}$$

Note that

$$\|\mathbf{s}\|_1 = \|\mathbf{s}\|_I + \|\mathbf{s}\|_{I^c} \tag{6.35}$$

where I^c is the complement set of I.

Theorem 7. *(Donoho and Huo 2001) For a given set of active basis, I of the true data, the solution to (6.33) is correct regardless of the signal values if and only if*

$$\forall \mathbf{x} \in \mathscr{N}_{A^g}, \quad \|\mathbf{x}\|_I \leq \|\mathbf{x}\|_{I^c} \tag{6.36}$$

where $\mathscr{N}_{A^g}$ is the null space of $\mathbf{A}^g$.

Proof. First suppose for an index set, I the condition is satisfied. Suppose there exists a signal vector, $\mathbf{s}^g$ with active vectors at I for which the solution is some different vector, $\mathbf{s}^{g\prime}$. Then we have $\mathbf{A}^g\mathbf{s}^g = \mathbf{A}^g\mathbf{s}^{g\prime}$ and $\|\mathbf{s}^{g\prime}\|_1 < \|\mathbf{s}^g\|_1$. Introducing $\mathbf{x} = \mathbf{s}^{g\prime} - \mathbf{s}^g$ we note that $\mathbf{x} \in \mathscr{N}_{A^g}$ and

$$\|\mathbf{x} + \mathbf{s}^g\|_1 < \|\mathbf{s}^g\|_1 \tag{6.37}$$

This can be also written as

$$\|\mathbf{s}^g\|_I > \|\mathbf{x} + \mathbf{s}^g\|_I + \|\mathbf{x}\|_{I^c} > \|\mathbf{s}^g\|_I - \|\mathbf{x}\|_I + \|\mathbf{x}\|_{I^c} \tag{6.38}$$

The last inequality is the result of the triangle inequality. This means that $\|\mathbf{x}^g\|_{I^c} < \|\mathbf{x}^g\|_I$ which is in contradiction with the assumption.

Second, suppose there exists $\mathbf{x} \in \mathscr{N}_{A^g}$ so that $\|\mathbf{x}^g\|_{I^c} < \|\mathbf{x}^g\|_I$. Let us denote the active basis for this vector by J. Assume a vector $\mathbf{s}^g$ with I as active indexes, and the additional property that its elements at the indexes in $I \cap J$ are the negative of the corresponding elements in $\mathbf{x}$ so that $\|\mathbf{x} + \mathbf{s}^g\|_{I\cap J} = 0$ and also $\|\mathbf{s}^g\|_{I\cap J} = \|\mathbf{x}\|_I$. Then

$$\begin{aligned} \|\mathbf{x} + \mathbf{s}^g\|_1 &= \|\mathbf{x} + \mathbf{s}^g\|_{I-J} + \|\mathbf{x} + \mathbf{s}^g\|_{J-I} \\ &= \|\mathbf{s}^g\|_{I-J} + \|\mathbf{x}\|_{I^c} < \|\mathbf{x}\|_I + \|\mathbf{s}^g\|_{I-J} \\ &= \|\mathbf{s}\|_{I\cap J} + \|\mathbf{s}^g\|_{I-J} = \|\mathbf{s}^g\|_1 \end{aligned} \tag{6.39}$$

This shows that there exists a signal vector for which the estimate is different from the true one both in directions and values. □

The criterion (6.36) of the above theorem can also be restated as the maximum of $\frac{\|\mathbf{x}\|_I}{\|\mathbf{x}\|_1}$ to be less than $\frac{1}{2}$. Finding all suitable active bases for which the theorem is satisfied has not been done yet. However, it is expected that the criterion is a trade off between the number of active basis vectors and the minimum angle of them. The less the number of sources is, the closer they can be.

To find the maximum of $\frac{\|\mathbf{x}\|_I}{\|\mathbf{x}\|_1}$ in the null space we can assume without lose of generality that $\|\mathbf{x}\|_1 = 1$. Now the problem is reformulated as finding the maximum of $\|\mathbf{x}\|_I$ in the intersection, $\mathscr{S}$ of the Null space hyper plane and the convex polygon of $\|\mathbf{x}\|_1 = 1$ in an N-dimensional space.Note that $\mathscr{S}$ is a convex polygon and $\|\mathbf{x}\|_I$ is a piecewise linear function. Hence, its maximum occurs on one of the vertexes of $\mathscr{S}$. But each vertex of $\mathscr{S}$ is a vector $\mathbf{x}$ in the null space with at most $m+1$ active indexes. We conclude that to search for the maximum value we only need to search the null space over the vectors with at most $m+1$ nonzero elements. Suppose $\mathbf{A}^g$ is m-ambiguous which means that each vertex point contains exactly $m+1$ nonzero elements. Corresponding to each set, $J = \{j_1, j_2, \ldots, j_{m+1}\}$ of $m+1$ active indexes the vertex can be found as follows

$$[\mathbf{a}_{j_1}\ \mathbf{a}_{j_2}\ \ldots \mathbf{a}_{j_m}] \begin{bmatrix} x_{j_1} \\ x_{j_2} \\ \cdots \\ x_{j_m} \end{bmatrix} = -x_{j_{m+1}} \mathbf{a}_{j_{m+1}} \tag{6.40}$$

This time we can assume without lose of generality that $-x_{j_{m+1}} = 1$ and $j_{m+1} \notin I$. Thus, the above equation will admit a unique solution and the criterion can be written as

$$T(J, I) = \frac{\sum_{j_k \in I} |x_{j_k}|}{\sum_{k=1}^{m} |x_{j_k}| + 1} \tag{6.41}$$

where x_{j_k} is given by (6.40). One important conclusion is that consistency for $\sigma^2 = 0$ cannot be guaranteed for all scenarios. In Section 7.3 we get an empirical graph of the consistency region by ℓ_1 regularization.

6.2 Consistency Conditions

In this chapter we are going to analyze the conditions under which the ℓ_1 regularization is guaranteed to have a solution sufficiently close to the actual DOA parameters.

Orthogonality and the case of large number of sensors

We earlier stated that the distinct basis vectors become orthogonal as the number of sensors grows. On the other hand, in the previous section we showed that the ℓ_1 regularization method gives the true directions for the case of orthogonal basis vectors. Now, we discuss the consistency for a large number of sensors by giving a

more accurate definition for this case. The main question in this chapter is weather or not the case of large number of sensors and orthogonality are equivalent in a ULA.

Assume an infinite sequence of ULA sensor arrays, all with the same separation distance, d between sensors. Suppose the number of sensors for the m^{th} array is m. Suppose a grid G with N points is used for all arrays. We denote the steering vectors for these arrays as

$$\mathbf{a}_m(\theta) = \begin{bmatrix} 1 \\ e^{j\phi} \\ e^{j2\phi} \\ \vdots \\ e^{j(m-1)\phi} \end{bmatrix} \qquad m = 1, 2, \ldots, \infty \tag{6.42}$$

where we assumed that the reference point is on the first sensor and $\phi = 2\pi\frac{d}{\lambda}\cos\theta$. We also define $\phi_i = 2\pi\frac{d}{\lambda}\cos\theta_i^g$ which are the electrical angles for the points on the grid. We assume that the sources are on the grid points. The other case should be treated differently. This set of problems are solved by the ℓ_1 regularization method with a sequence of regularization parameters, $\{\lambda_m\}_{m=1}^{\infty}$. Note that for any two distinct indexes i, and j it is easy to see that

$$\lim_{m\to\infty} \frac{1}{m} |\mathbf{a}_m^H(\theta_i^g)\mathbf{a}_m(\theta_j^g)| = 0 \tag{6.43}$$

For any choice of active basis for the m^{th} ULA we can write

$$\mathbf{A}_m^H\mathbf{A}_m = m\mathbf{I} + m\mathbf{E}_m(\theta) \tag{6.44}$$

Where for simplicity the argument for $\mathbf{A}$ is neglected. Due to (6.43), for every $\epsilon > 0$ there exists a sufficiently large M for which for every choice of the active indexes, $\hat{\theta}$

$$m > M \Rightarrow \forall i, j \ \epsilon_{i,j} = (\mathbf{E}_m)_{i,j} \leq \epsilon \tag{6.45}$$

This also can be written as

$$\lim_{m\to\infty} \sup_{\boldsymbol{\theta}} \|\mathbf{E}_m(\boldsymbol{\theta})\| = 0 \tag{6.46}$$

For each direction, θ we also introduce

$$y_m(\theta, t) = \frac{1}{m}\mathbf{a}_m^H(\theta)\mathbf{x}(t) \tag{6.47}$$

We denote the true and the estimated DOAs by $\boldsymbol{\theta}_0$ and $\hat{\boldsymbol{\theta}}$ respectively. Now, according to the model of the signal in (3.3) it can be inferred that

$$\lim_{m\to\infty} y_m(\theta_i^g, t) = \lim_{m\to\infty} \frac{1}{m}\mathbf{a}_m^H(\theta_i^g)\left[\sum_{j=1}^{n} \mathbf{a}_m(\theta_{0j})s_j(t) + \mathbf{n}_m(t)\right] = \begin{cases} s_j(t) & \theta_i^g = \theta_{0j} \in \boldsymbol{\theta}_0 \\ 0 & \text{Otherwise} \end{cases} \tag{6.48}$$

Note that the term related to the noise goes to zero almost surely as explained in (Stout 1974). Then, the solution (6.4) can be written as

$$\hat{\mathbf{s}}(\boldsymbol{\theta}, t) = \left(\frac{\lambda_m}{2m}\boldsymbol{\Gamma}^{-1} + \mathbf{I} + \mathbf{E}_m\right)^{-1} \mathbf{y}_m(\boldsymbol{\theta}, t) \tag{6.49}$$

where $\mathbf{y}_m(\boldsymbol{\theta}, t) = [y(\theta_1, t), y(\theta_2, t), \ldots, y(\theta_n, t)]^T$. To state the theorem we need a lemma to be proved before hand.

Lemma 1. *For every $\epsilon > 0$ there exists an M so that for every direction set $\boldsymbol{\theta}$ and $m > M$ we almost surely have*

$$\|\hat{\mathbf{s}}(\boldsymbol{\theta}, t)\|_2 \leq \sqrt{\sum_{j|\theta_j \in \boldsymbol{\theta} \cap \boldsymbol{\theta}_0} |s_j(t)|^2} + \epsilon \leq C \tag{6.50}$$

where C is a constant not relating to any of the variables.

Proof. Equation (6.49) can be written as

$$\left(\frac{\lambda_m}{2m}\boldsymbol{\Gamma}^{-1} + \mathbf{I} + \mathbf{E}_m\right) \hat{\mathbf{s}}(t) = \mathbf{y}_m(t) \tag{6.51}$$

Then we can write

$$\begin{aligned} \|\mathbf{y}_m(t)\|_2 &\geq \|\hat{\mathbf{s}}(t)\|_2 \sigma_{\min}\left(\frac{\lambda_k}{2m_k}\boldsymbol{\Gamma}^{-1} + \mathbf{I} + \mathbf{E}_m\right) \\ &\geq \|\hat{\mathbf{s}}(t)\|_2 \sigma_{\min}\left(\mathbf{I} + \mathbf{E}_m\right) \end{aligned} \tag{6.52}$$

where $\sigma_{\min}(.)$ is the minimum singular value of its argument. The last inequality holds because $\frac{\lambda_m}{2m}\boldsymbol{\Gamma}^{-1}$ is positive definite. Now note that for each $\mathbf{x}$,

$$\mathbf{x}^H(\mathbf{I} + \mathbf{E}_m)\mathbf{x} = \|\mathbf{x}\|_2^2 + \mathbf{x}^H(\mathbf{E}_m)\mathbf{x} \geq (1 - \epsilon)\|\mathbf{x}\|_2^2 \tag{6.53}$$

Where the last inequality is due to (6.45). this shows that the minimum eigenvalue is greater than $1 - \epsilon$ and so according to (6.48) we have

$$\|\hat{\mathbf{s}}(t)\|_2 \leq \frac{\sqrt{\sum_{j|\theta_j \in \boldsymbol{\theta}} |s_j(t)|^2} + \epsilon}{1 - \epsilon} \leq \sqrt{\sum_{j|\theta_j \in \boldsymbol{\theta}} |s_j(t)|^2} + \epsilon' \leq C \tag{6.54}$$

where ϵ' tends to zero as ϵ does and C can be for example $\sqrt{\sum_j |s_j(t)|^2} + 1$. □

In the above lemma we showed that for a sufficiently large number of sensors, m. The estimated source vector is always bounded regardless of the estimated position of sources. Now the following theorem shows the consistency.

Theorem 8. *Consider a sequence of arrays satisfying 6.43. Assume that the point sources are exactly on the grid points, then there exists a corresponding sequence of regularization parameters $\{\lambda_m\}_{m=1}^{\infty}$ for which the solutions to the ℓ_1 regularization problem converges to the exact signals and directions.*

Proof. Writing $\mathbf{E}_k = [\mathbf{e}_1\ \mathbf{e}_2\ \ldots \mathbf{e}_n]^H$, from (6.51) we have

$$\forall r,\ \left(\frac{\lambda_m}{2m\gamma_r}+1\right)\hat{s}_r(\boldsymbol{\theta},t) = y_m(\theta_r,t) - \mathbf{e}_r^H\hat{\mathbf{s}}(\boldsymbol{\theta},t) \tag{6.55}$$

It turns out that

$$\left(\frac{\lambda_m}{2m\gamma_r}+1\right)^2\sum_{t=1}^{T}|\hat{s}_r(\boldsymbol{\theta},t)|^2 = \sum_{t=1}^{T}|y_m(\theta_r,t) - \mathbf{e}_r^H\hat{\mathbf{s}}(\boldsymbol{\theta},t)|^2 \tag{6.56}$$

According to the definition of γ_r we can write

$$\lim_{m\to\infty}\left(\frac{\lambda_m}{2m}+\gamma_r\right) = \lim_{m\to\infty}\sqrt{\frac{1}{T}\sum_{t=1}^{T}|y_m(\theta_r,t) - \mathbf{e}_r^H\hat{\mathbf{s}}(\boldsymbol{\theta},t)|^2} \tag{6.57}$$

Now, due to Lemma 1, for sufficiently large k, $\hat{\mathbf{s}}(\boldsymbol{\theta},t)$ is bounded and according to (6.45) the term $\mathbf{e}_r^H\hat{\mathbf{s}}(\boldsymbol{\theta},t)$ tends to zero at infinity. Also from (6.48) we have

$$\lim_{m\to\infty}\left(\frac{\lambda_m}{2m}+\gamma_r\right) = \begin{cases} \sqrt{\frac{1}{T}\sum_{t=1}^{T}|s_i(t)|^2} & \theta_r = \theta_{0i} \\ 0 & \text{Otherwise} \end{cases} \tag{6.58}$$

Note that as a part of the marginal solution, $\boldsymbol{\Gamma}$ is a function of the choice of the active sources, $\boldsymbol{\theta}$. We here showed that the $\boldsymbol{\Gamma}$s of all marginal solutions uniformly converge to their limits. Suppose λ_m is chosen so that $\lim_{m\to\infty}\frac{\lambda_m}{m} = 0$ Then we get the limit of the $\gamma_r(\boldsymbol{\theta})$ as the left hand side of (6.58). Note that because $\hat{\mathbf{s}}(\boldsymbol{\theta},t)$ is a continuous function of λ, $\boldsymbol{\Gamma}$, and $\mathbf{y}_k$, from (6.49) it can be derived that

$$\lim_{m\to\infty}\hat{\mathbf{s}}(\boldsymbol{\theta},t) = \lim_{m\to\infty}\mathbf{y}_m(\boldsymbol{\theta},t) \tag{6.59}$$

On the other hand the cost function of the ℓ_1 regularization methods can be written as

$$\begin{aligned} V_{\text{LASSO}}(\boldsymbol{\theta}) &= \frac{1}{T}\sum_{t=1}^{T}\|\mathbf{x}(t) - \mathbf{A}\hat{\mathbf{s}}(\boldsymbol{\theta},t)\|_2^2 + \lambda_m\sum_{i=1}^{n}\gamma_i(\boldsymbol{\theta}) \\ &= \frac{1}{T}\sum_{t=1}^{T}\|\mathbf{x}(t)\|_2^2 + m\left[\frac{1}{T}\sum_{t=1}^{T}\left(\hat{\mathbf{s}}^H(\boldsymbol{\theta},t)(\mathbf{I}+\mathbf{E}_k)\hat{\mathbf{s}}(\boldsymbol{\theta},t)\right.\right. \\ &- \left.\left.2\text{Re}(\mathbf{y}^H(\boldsymbol{\theta},t)\hat{\mathbf{s}}(\boldsymbol{\theta},t))\right) + \frac{\lambda_m}{m}\sum_{i=1}^{n}\gamma_i\right] \end{aligned} \tag{6.60}$$

Because the maximum does not change by shifting and scaling, according to (6.59) and (6.45), and the fact that $\lim_{m\to\infty}\frac{\lambda_m}{m} = 0$ we get

$$\lim_{m\to\infty}V_{\text{LASSO}}(\boldsymbol{\theta}) \propto -\lim_{m\to\infty}\sum_{t=1}^{T}\mathbf{y}_m^H(\boldsymbol{\theta},t)\mathbf{y}_m(\boldsymbol{\theta},t) \tag{6.61}$$

From (6.45) it is obvious that there exists a sufficiently large M such that for every $m > M$ the global maximum occurs at the true directions ($\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0$). Also, from (6.45) and (6.59) it can be seen that for the true directions $\hat{\mathbf{s}}(t)$ tends to the true values. □

There are some key points in the proof of the above theorem. Firstly, the large m approaches the situation of orthogonality, where it is possible to explain the behavior of the method by beamforming and thresholding. Secondly, according to the term $\mathbf{A}^H\mathbf{x}$ in (6.4), the noise term vanishes as m grows, which allows the threshold λ to be chosen small so that the scaling factor η_i in (6.18) tends to 1.

What if one of the source directions is not on the grid points? If the grid is constant this means that the received signal term $\mathbf{a}(\theta_i)s_i$ will become orthogonal to all basis vectors and will be disappeared from $\mathbf{A}^H\mathbf{x}$. In the other words, the sensitivity range of each basis vector to its adjacent directions decreases by growing m, so that a source with a certain distance from the grid points eventually becomes invisible by the basis vectors. To cope with this problem one may propose to increase the density of the grid gradually by m so that the point can be seen at least by one basis vector, but this implicitly means that the consecutive basis vectors gradually get closer and will not be orthogonal anymore even if m tends to infinity!

To sum up, we can judge that generally, the large number of sensors case is not equivalent to an orthogonal basis set. Instead, it achieves the true values of the parameters by a remarkable increase in the effective Signal to Noise Ratio (SNR) through a growing diversity of the received data. From this point of view it can be inferred that the general consistency for a large number of sensors is related to the noise-free case as well as the orthogonal case.

6.3 High SNR Consistency

We bring this chapter to conclusion by giving a discussion about the consistency of the regularization method when the SNR is high enough. In Section 6.1.2 we learned the condition under which a set of active indexes can be recovered by the method. Here we show that with the same condition it is possible to recover the directions on a grid for sufficiently small noise by thresholding the estimated power spectrum, γ. We adopt the equivalent form of ℓ_1 regularization similar to (6.33) and we confine ourselves to the case of one snapshot. Consider the following formulation

$$\begin{aligned} \hat{\mathbf{s}}_\delta(\mathbf{x}) &= \arg\min_{\mathbf{s}} \|\mathbf{s}\|_1 \\ \text{subject to} \quad & \|\mathbf{A}\mathbf{s} - \mathbf{x}\| \leq \delta \end{aligned} \tag{6.62}$$

We show the consistency for such an optimization but before that we need some observations.

Lemma 2. *Given a matrix* $\mathbf{A}$ *and a vector* $\mathbf{y} \in \mathscr{R}_A$ *there always exists a vector* $\mathbf{x}$

such that $\mathbf{Ax} = \mathbf{y}$*, and*

$$\|\mathbf{x}\|_2 \leq \frac{\|\mathbf{y}\|_2}{\sigma_{min}(\mathbf{A})} \tag{6.63}$$

where $\sigma_{min}(\mathbf{A})$ *denotes the minimum nonzero singular value of* $\mathbf{A}$*.*

Proof. Using Singular Value Decomposition (SVD), we can write $\mathbf{A} = \mathbf{USV}^H$, where $\mathbf{U}$ and $\mathbf{V}$ are two orthogonal matrices and $\mathbf{S}$ can be written as

$$\mathbf{S} = \begin{pmatrix} \mathbf{\Sigma} & 0 \\ 0 & 0 \end{pmatrix} \tag{6.64}$$

where $\mathbf{\Sigma}$ is the diagonal matrix of nonzero singular values. Then, $\mathbf{Ax} = \mathbf{y}$ implies

$$\mathbf{USV}^H\mathbf{x} = \mathbf{y} \tag{6.65}$$

defining $\mathbf{x}' = \mathbf{V}^H\mathbf{x}$ and $\mathbf{y}' = \mathbf{U}^H\mathbf{y}$ we get

$$\begin{pmatrix} \mathbf{\Sigma} & 0 \\ 0 & 0 \end{pmatrix} \mathbf{x}' = \begin{pmatrix} \mathbf{\Sigma} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x}'_t \\ \mathbf{x}'_b \end{pmatrix} = \begin{pmatrix} \mathbf{y}'_t \\ \mathbf{y}'_b \end{pmatrix} \tag{6.66}$$

where we divide $\mathbf{x}'$ and $\mathbf{y}'$ into two blocks of suitable dimensions. It can be seen in (6.66) that $\mathbf{y}'_b = \mathbf{0}$ and $\Sigma\mathbf{x}'_t = \mathbf{y}'_t$. To get the smallest norm we can choose $\mathbf{x}'_b = \mathbf{0}$ and $\mathbf{x}'_t = \Sigma^{-1}\mathbf{y}'_t$. Then, we have

$$\|\mathbf{x}'\|_2 = \|\mathbf{x}'_t\|_2 \leq \frac{\|\mathbf{y}'_t\|_2}{\sigma_m(A)} = \frac{\|\mathbf{y}'\|_2}{\sigma_m(A)} \tag{6.67}$$

Because $\mathbf{U}$ and $\mathbf{V}$ are orthogonal, $\|\mathbf{x}'\|_2 = \|\mathbf{x}\|_2$ and $\|\mathbf{y}'\|_2 = \|\mathbf{y}\|_2$, which implies (6.63). □

Note that for a ULA with a uniform grid we can find the nonzero singular values by computing the eigenvalues of $\mathbf{A}^g\mathbf{A}^{gH}$. We have

$$\frac{1}{N}(\mathbf{A}^{gH}\mathbf{A}^g)_{p,q} = \frac{1}{N}\sum_{i=1}^{N} e^{j(p-q)\phi_i} \tag{6.68}$$

which for a large N, is a Riemann sum for the integral $\int_0^\pi e^{j(p-q)\phi}d\phi = \pi\delta_{p,q}$. This shows that for a large value of N the minimum singular value of $\mathbf{A}^g$ is about $\sqrt{N\pi}$. Now we can state the consistency condition as follows

Theorem 9. *Assume that the active indexes for the true sources at the grid points satisfy the condition of Theorem 7 with strict inequality. Then, for every* $\epsilon > 0$ *there exist a pair of* $\mu, \delta > 0$ *such that if* $\|\mathbf{n}\| \leq \mu$ *we get* $\|\hat{\mathbf{s}}_\delta(\mathbf{x}) - \mathbf{s}\|_1 \leq \epsilon$*, where* $\hat{\mathbf{s}}_\delta(\mathbf{x})$ *is defined in (6.62)*

Proof. If we choose $\mu \leq \delta$, then we are sure that

$$\|\mathbf{x} - \mathbf{A}^g\mathbf{s}\|_2 = \|\mathbf{n}\|_2 \leq \mu \leq \delta \tag{6.69}$$

which shows that the actual source vector, $\mathbf{s}$ is in the feasible region of (6.62). Thus,

$$\|\mathbf{s}_\delta(\mathbf{x})\|_1 \leq \|\mathbf{s}\|_1 \tag{6.70}$$

Now we show that the difference, $\mathbf{d} = \mathbf{s} - \mathbf{s}_\delta(\mathbf{x})$, between the true value and the estimated one can be decomposed into two vectors ($\mathbf{d} = \boldsymbol{\Delta} + \boldsymbol{\nu}$) for which $\boldsymbol{\Delta} \in \mathscr{N}_A$ and $\|\boldsymbol{\nu}\|_2$ is small. We have

$$\|\mathbf{x} - \mathbf{A}^g\mathbf{s}_\delta(\mathbf{x})\|_2 = \|\mathbf{A}^g\mathbf{s} + \mathbf{n} - \mathbf{A}^g\mathbf{s}_\delta(\mathbf{x})\|_2 \leq \delta \tag{6.71}$$

which implies that

$$\|\mathbf{A}^g\mathbf{s} - \mathbf{A}^g\mathbf{s}_\delta(\mathbf{x})\|_2 \leq \delta + \mu \tag{6.72}$$

Due to Lemma 2 there exists a vector ν so that $\mathbf{A}\nu = \mathbf{A}\mathbf{d}$ and $\|\nu\|_2 \leq \frac{\|\mathbf{A}^g\mathbf{d}\|_2}{\sigma_m(\mathbf{A}^g)}$ and from (6.72)

$$\|\nu\|_2 \leq \frac{\|\mathbf{A}^g\mathbf{d}\|_2}{\sigma_m(\mathbf{A}^g)} \leq \frac{\delta + \mu}{\sigma_m(\mathbf{A}^g)} = \delta' \tag{6.73}$$

Also note that $\boldsymbol{\Delta} = \mathbf{d} - \nu \in \mathscr{N}_A$.

Now we can write (6.70) as

$$\|\mathbf{s} - \mathbf{d}\|_1 \leq \|\mathbf{s}\|_1 \tag{6.74}$$

Using the notation introduced in (6.34), (6.74) becomes

$$\|\mathbf{s} - \mathbf{d}\|_I + \|\mathbf{d}\|_{I^c} \leq \|\mathbf{s}\|_I \tag{6.75}$$

and by the triangle inequality we get,

$$\|\mathbf{s}\|_I - \|\mathbf{d}\|_I + \|\mathbf{d}\|_{I^c} \leq \|\mathbf{s} - \mathbf{d}\|_I + \|\mathbf{d}\|_{I^c} \leq \|\mathbf{s}\|_I \tag{6.76}$$

which leads to

$$\|\mathbf{d}\|_{I^c} \leq \|\mathbf{d}\|_I \tag{6.77}$$

Substituting the decomposition for $\mathbf{d}$, we get

$$\|\boldsymbol{\Delta} + \nu\|_{I^c} \leq \|\boldsymbol{\Delta} + \nu\|_I \tag{6.78}$$

Again, using the triangle inequality and simplifying, we have

$$\|\boldsymbol{\Delta}\|_{I^c} \leq \|\boldsymbol{\Delta}\|_I + \|\nu\|_1 \tag{6.79}$$

Now according to the fact that $\|\nu\|_1 \leq \sqrt{m}\|\nu\|_2$ and due to (6.73) we have

$$\|\boldsymbol{\Delta}\|_1 \leq 2\|\boldsymbol{\Delta}\|_I + \sqrt{m}\delta' \tag{6.80}$$

If Theorem 7 holds true with strict inequality, then because of the discrete nature of the grid there exists a sufficiently small but constant bound κ such that we have $\frac{\|\Delta\|_I}{\|\Delta\|_1} \leq \frac{1}{2} - \kappa$. Combining this result with (6.80) we have

$$\|\Delta\|_1 \leq \frac{\sqrt{m}\delta'}{2\kappa} \tag{6.81}$$

and finally, since $\mathbf{s} - \hat{\mathbf{s}}_\delta((x)) = \boldsymbol{\Delta} + \boldsymbol{\nu}$, we obtain

$$\|\mathbf{s}_\delta(\mathbf{x}) - \mathbf{s}\|_1 \leq \|\Delta\|_1 + \|\nu\|_1 \leq \left(\frac{1}{\kappa} + 1\right)\sqrt{m}\delta' \tag{6.82}$$

And the result follows. □

7 COMPUTER SIMULATIONS

In this chapter we discuss the results for the previously introduced LASSO-based DOA estimation using simulated data. All the simulations are performed in the MATLAB software using the SeDuMi toolbox. For the case of multiple snapshots we used the dimension reduction technique introduced in section 5.

7.1 Comparison to the Classical Methods

In this section we present the result of comparing LASSO to the conventional well-known methods of MUSIC and beamforming. LASSO is not expected to provide a better performance as compared to the ML criterion. However, as we stated, the superiority of the new method is clear when we compare the computational effort for ML and LASSO. Figure 7.1 shows a typical spectrogram for the LASSO method. The regularization parameter is chosen as the smallest value giving true number of sources which is also correspondent to the smaller noise term, $\|\mathbf{x} - \mathbf{As}\|$. As can be

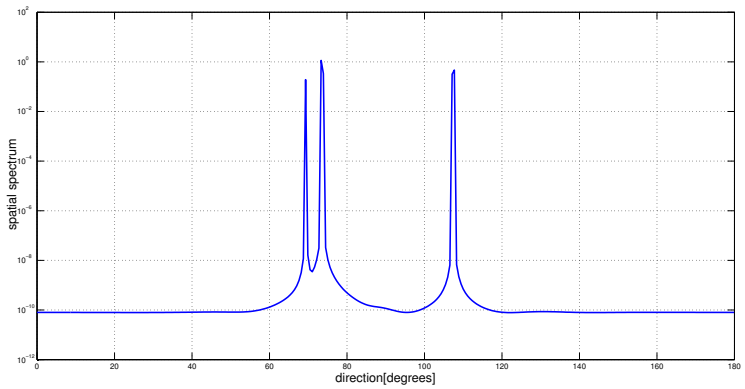


Figure 7.1. The spatial spectrum of LASSO for one snapshot, and SNR=10dB.

seen, if a source point is not exactly on the grid there always exists a set of nonzero values around the true direction. This problem can be overcome by first choosing the peak points of the spectrum and then thresholding.

Resolution

The resolution is the minimum angle separation between different sources which can be recovered by the method. We investigate the resolution of the LASSO method by introducing three sources at 70°, 80°, and 108° degrees and running the simulation with both one and multiple snapshots for a ULA of 8 sensors and half-wavelength separation. Figure 7.2 shows the resulting spatial spectra for LASSO and beam-

forming

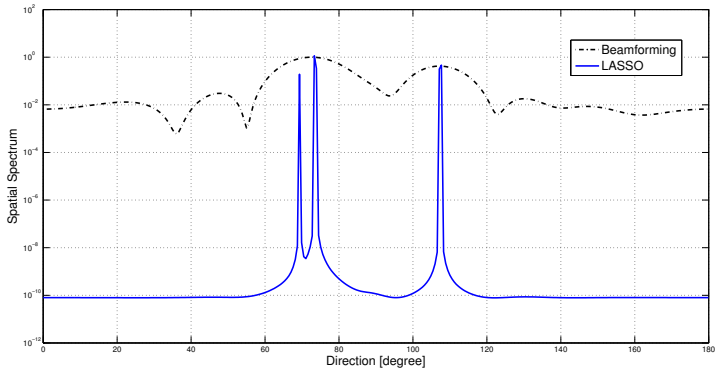
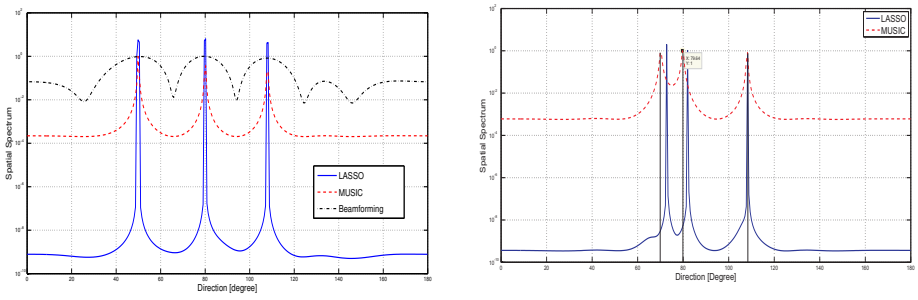


Figure 7.2. The spatial spectra of LASSO and beamforming. True sources at $\theta = \{70°, 80°, 180°\}$, 8-element standard ULA,T=1 snapshot.

Figure 7.2 shows that the LASSO technique can resolve the sources with 10° separation, while two peaks will merge in the beamforming spectrum. Figure 7.3 shows the effect of applying LASSO to multiple snapshot data. The vertical lines show the true DOAs.



(a) (b)

Figure 7.3. Spatial spectra for LASSO comparing to a) Beamforming and MUSIC with $\boldsymbol{\theta} = \{50°, 80°, 108°\}$. b) MUSIC by closer sources. Scenario as in 7.2 but $T = 100$ and SNR=7dB.

Investigating the multiple snapshot case with $T = 100$, Figure 7.3(a) shows that the LASSO method works well when the sources are sufficiently separated. However, as the angle between sources gets smaller, a bias appears in the estimated angles (Figure 7.3(b)). As we show later in this section this bias is a result of non-consistency in the method.

Effect of noise

In this section we show the effect of decreasing the SNR by simulating the scenario of Figure 7.2 but increasing the noise power. Note that it is known that for both beamforming and MUSIC methods, the effect of noise will decrease with a large number of snapshots, so we choose a relatively small number of snapshots $T = 10$. The result is shown in Figure 7.4. In this figure note that although the LASSO

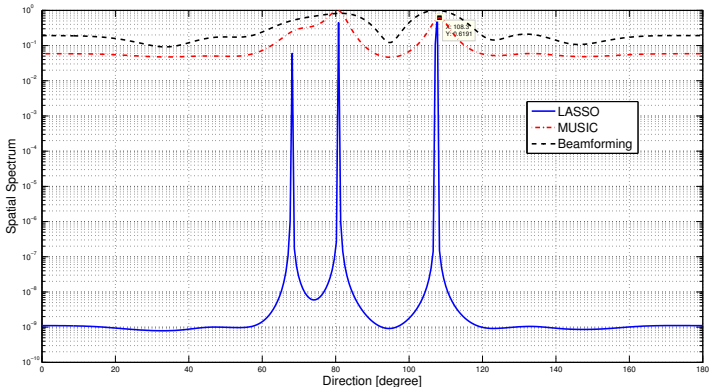


Figure 7.4. Spatial spectra of the LASSO ,beamforming, and MUSIC methods. SNR=0dB.

method can resolve 3 sources unlike the other methods, there is still an error due to the low SNR.

7.2 Regularization Parameter Selection

In this section we show the result of applying the regularization parameter selection methods previously introduced.

Multiple snapshot case

As proposed in Chapter 5, for a large number of snapshots the MDL criterion can be used to find a proper value of the regularization parameter, λ. For this purpose, we set up a simulation of 3 sources with $T = 100$, and performed LASSO for different values of λ. For each such a result we compute the USC given in (5.16) and plot it versus λ. The result is shown in Figure 7.5. Note that the region between vertical lines with triangle mark is the region where the true number of sources is obtained i.e. $n = 3$.

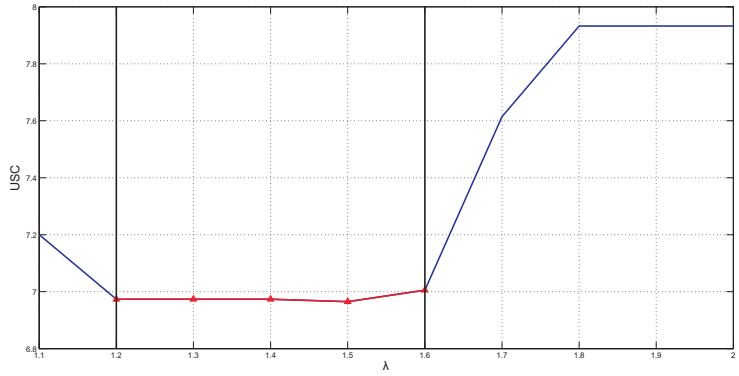


Figure 7.5. The universal stochastic complexity for 3 sensors and SNR=3dB.

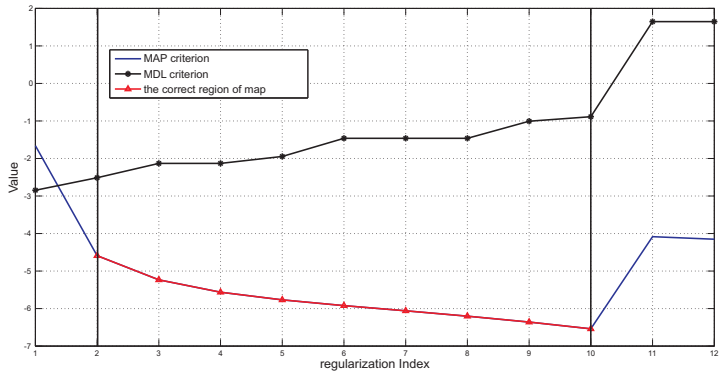


Figure 7.6. The MAP and MDL criteria for one snapshot and SNR=10dB.

One snapshot case

For the one snapshot case we apply the same procedure as the MDL simulation but this time compute (5.24) for each optimization for different λ. The result is given in (7.6). As before the region between the vertical lines shows the true detection area. As it can be seen unlike the MDL criterion, MAP has a the tendency to chose the largest possible λ corresponding to the most sparse solution. It is also worth noting that the MAP estimation for λ is sensitive to the distribution of the source. For a pdf with tails stronger than Laplacian distribution the method may loose its precision. Here we used a Gaussian variable to generate the result.

7.3 A Theoretical Resolution Bound

Finally we try to find the consistency region of all active sets, I satisfying theorem 7. To do so we solve the optimization problem

$$\min_{\mathbf{x}\in\mathcal{N}_A, \|x\|_I=1} \|\mathbf{x}\|_1 \tag{7.1}$$

which using the regularization equivalence conditions in chapter 4 can be written as

$$\min_{\mathbf{x}\in\mathcal{N}_A} \|\mathbf{x}\|_1 + \lambda\|\mathbf{x}\|_I \tag{7.2}$$

with a small value of λ. This problem is convex and can be done using convex optimization techniques. The result is shown in figure where we find the closest separation between two sources at different angles resolvable by the method i.e. the resolution. We compared this result to the beam width of the beamformer for different number of sensors, which is approximately the resolution of the beamforming technique.

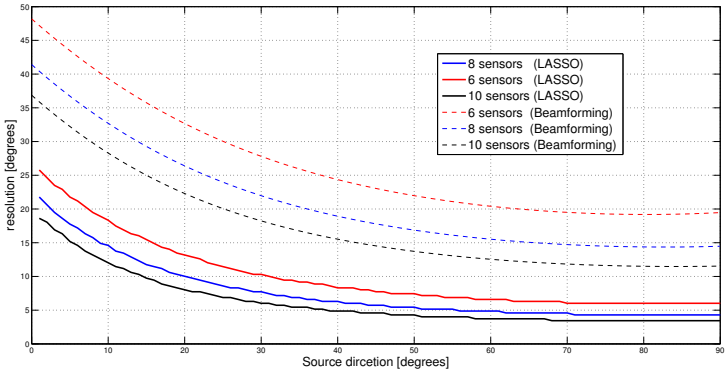


Figure 7.7. The fundumental resolution of LASSO compared to Beamforming.

It can be seen that for $T = 1$ and very high SNR the LASSO method has a much better resolution than the beamforming method. Also, note that the resolution for both method depends on the DOA, but for beamforming it neither depends on the SNR nor the number of snapshots.

8 CONCLUSION

In this report we discussed the DOA estimation solution by viewing the unknown parameters as a sparse vector and using the Least Absolute (ℓ_1) norm as a Shrinkage and Selection Operator, which justifies the terminology LASSO (**??**). The main idea is to replace the abstract functional form by the mechanism of an operator choosing from a sufficiently large look-up table. As described in Section 4, this can be shown mathematically as a product of a very large look-up matrix, $\mathbf{A}^g$ to a sparse selection vector $\mathbf{s}^g$. Forming a sparse regularization problem, the solution can be found by the LASSO method using convex programming. Next, we addressed the question of selecting the regularization parameter λ, in the LASSO method, and showed that it can be solved by a two-stage procedure of detecting the number of sources and estimating their directions. We introduced MDL as a strong criterion for detecting the number of sources based on the estimation of directions and signals. We also viewed LASSO as a Bayesian estimator with a truncated Laplacian prior, and devised an ML criterion for selecting λ. From a mathematical analysis we found that:

- For the asymptotic case of orthogonal bases, the method is merely a combination of beamforming and thresholding by λ. Note that in this case the beamforming gives a perfect solution if the noise is almost spatially white (its spatial spectrum is approximately constant).

- For the case of a large number of sensors, if the source point is on the grid the method is consistent through the asymptotic orthogonality of the bases both to each other and to the noise vector. We proved uniform convergence for such a case by increasing the number of sensors.

- For the high SNR case we showed that the method is not consistent. We introduced a previously devised criterion for consistency and provided a numerical (but not analytical) method of implementing it using convex programming. This criterion is used to find the fundamental resolution of the LASSO method for DOA estimation.

- We showed that if the point is not on the grid the case of large number of sensors resembles the noiseless case. However, the noiseless case with an arbitrary source is not discussed.

From a simulation study, we found that:

- The LASSO method for DOA estimation is a high resolution method compared to the beamforming technique. It also does not depend to the source correlation compared to MUSIC, and is less sensitive to the noise level.

- For a large number of snapshots, the SVD-based dimension reduction reduces the computation time for LASSO dramatically.
- The MDL criterion is suitable for a large number of received data. For less number of snapshots, the MAP criterion might be used but it is sensitive to the true distribution of the sources.
- There is a fundumental resolution limit for this method which can be computed.

8.1 Suggestions For Future Work

The LASSO-based technique for DOA estimation has a great potential as demonstrated herein as well as in (Malioutov 2003). However, much work remains before its properties are fully understood. The following lists a few suggestions:

- According to the fact that we are only interested in special solutions to the LASSO problem is it possible to perform the search for the regularization parameter by a faster method?
- What is the behavior of the method when the DOA is not on the grid?
- What is the theoretical performance of the method with stochastic waveforms?

Finally, a deeper investigation about the bias introduced by the dimension reduction, the Bayesian interpretation of the LASSO as well as the theoretical discussion for an arbitrary source direction, and an analytical performance analysis is proposed by the author as interesting topic for future research.

Bibliography

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *Proc. 2nd Int. Symp. Information Theory*, pp 267–281.

Boolos, G. and Jeffrey, R. (1989). *Computability and Logic*. 3 edn. Cambridge University Press.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge university press.

Chen, S. and Donoho, D.L. (1994). On basis pursuit.. Technical report. Department of Statistics, Stanford University.

Cover, Thomas M. and Thomas, Joy A. (2006). *Elements of information theory*. 2 edn. Wiley-Interscience.

Donoho, D.L. and Huo, X. (2001). Uncertainty principles and ideal atomic decomposition. *Information Theory, IEEE Transactions on*, **47**(7), 2845 –2862.

Elad, M. and Bruckstein, A.M. (2002). A generalized uncertainty principle and sparse representation in pairs of bases. *Information Theory, IEEE Transactions on*, **48**(9), 2558 – 2567.

Fessler, J.A. and Hero, A.O. (1994). Space-alternating generalized expectation-maximization algorithm. *Signal Processing, IEEE Transactions on*, **42**(10), 2664 –2677.

G. H. Golub, C. F. Van Loan (1996). *Matrix computations*. 3rd edn. The John Hopkins university press.

Horn, Roger A. and Johnson, Charles R. (1991). *Topics in matrix analysis*. Cambridge University Press.

Li, M. and Vitanyi, P. (1997). *An Introduction to Kolmogorov Complexity and Its Applications*. 2nd edn. Springer-Verlag.

Malioutov, D., Cetin, M. and Willsky, A.S. (2005). A sparse signal reconstruction perspective for source localization with sensor arrays. *Signal Processing, IEEE Transactions on*, **53**(8), 3010 – 3022.

Malioutov, Dimitry M. (2003). A sparse signal reconstruction perspective for source localization with sensor arrays. Master's thesis. Massachusetts Institute of Technology Dept. of Electrical Eng.

Ottersten, B., Viberg, M. and Kailath, T. (1992). Analysis of subspace fitting and ml techniques for parameter estimation from sensor array data. *Signal Processing, IEEE Transactions on*, **40**(3), 590 –600.

Poor, H. V. and G. W. Wornell, Eds. (1998). *Wireless Communications: Signal Processing Perspectives*. Prentice Hall.

Rissanen, J. (1986). Stochastic complexity and modeling. *The Annals of Statistics*, **14**, 1080–1100.

Rissanen, J. (2006). Information and complexity in statistical modeling. *Inf. theory Workshop, ITW Punta del Este.*, pp 351 –351.

Rissanen, J.J. (1996). Fisher information and stochastic complexity. *Information Theory, IEEE Transactions on*, **42**(1), 40 –47.

Robinson, E. A. and Treitel, S. (1980). *Geophysical Signal Analysis*. Prentice Hall.

Schweppe, F. (1968). Sensor-array data processing for multiple-signal sources. *Information Theory, IEEE Transactions on*, **14**(2), 294 – 305.

Shtarkov, Y. M. (1987). Universal sequential coding of single messages. *Problems of Information Transmission*, **23**(3), 3–17.

Stoica, P., Selen, Y. and Li, Jian (2004). On information criteria and the generalized likelihood ratio test of model order selection. *Signal Processing Letters, IEEE*, **11**(10), 794 – 797.

Stout, W. F. (1974). *Almost Sure Convergence*. Academic press, New York.

Sturm, Jos F. (1998). Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones.

Tibshirani, Robert (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B, (Methodological)*, **58**(1), 267–288.

Trees, H. L. Van (1968). *Detection, Estimation, and Modulation Theory, Part III: Radar-Sonar Signal Processing and Gaussian Signals in Noise*. John Wiley and Sons.

Van Veen, B.D. and Buckley, K.M. (1988). Beamforming: a versatile approach to spatial filtering. *ASSP Magazine, IEEE*, **5**(2), 4 –24.

Viberg, M. and Krim, H. (1997). Two decades of statistical array processing. In: *Signals, Systems Computers, 1997. Conference Record of the Thirty-First Asilomar Conference on*. vol. 1. pp 775 –777 vol.1.

Viberg, M. and Ottersten, B. (1991). Sensor array processing based on subspace fitting. *Signal Processing, IEEE Transactions on*, **39**(5), 1110 –1121.

Wax, M. and Ziskind, I. (1989). Detection of the number of coherent signals by the mdl principle. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, **37**(8), 1190 –1196.