MAXIMUM ENTROPY SPECTRUM ESTIMATION FROM NOISY CORRELATION MEASUREMENTS

Kai-Bor Yu, Xinhua Zhuang, Robert M. Haralick

Department of Electrical Engineering
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

ABSTRACT

A computational method has been developed for solving the maximum entropy spectrum estimation with uncertainty in the correlation measurements. This approach depends on first solving an unconstrained optimization problem and then an iterative search for the zero of the constraint equation along a well-defined trajectory. This trajectory is governed by a system of differential equations and the unconstrained optimal solution as the initial condition. The direction of search can be determined easily by evaluating the constraint equation at each iteration. In solving the trajectory, the Gauss-Seidel iterative algorithm is used to solve the symmetric system of equations. The choice of the previous estimate as the initial estimate greatly accelerates the convergence rate of this scheme.

I. INTRODUCTION

The purpose of this paper is to present a computational approach for solving the maximum entropy method (MEM) spectrum estimation when there is uncertainty in the correlation measurements. The standard MEM is computationally efficient and known for its high resolution properties [1]. The underlying principle of the MEM can be put into the Bayesian framework as a method for statistical inference. A discussion of the rational of this method can be found in Jayne's paper [2].

The traditional algorithm requires the availability of accurate correlation function on a uniform grid. However, the correlation coefficients are usually estimated from samples taken from non-uniform grid, thus an algorithm that takes into the account of the non-uniformity and the noisy nature of the correlation coefficients is of practical interest. A discussion of the MEM spectral estimation for nonuniformly spaced correlation measurements can be found in [3], [4], [5], and for noisy correlation can be found in [6] and [7]. Lang and McClellan [6] have investigated the dual approach while Wornecke and D'Addario [7] have investigated the primal problem. Lim and Malik [5] have used an alternating projection type of algorithm which, though not proven to converge, has been proven to converge in practice.

This paper presents an algorithm for solving a similar problem discussed in [7]. The problem addressed in [7] is in the context of Fourier synthesis in radio astronomy while we are concerned here with the spectral density estimation from correlation coefficients. The problem posed is a difficult constrained optimization problem in n + 1 dimensional space. Our approach to avoid this difficulty is to convert the problem into a m-dimensional unconstrained optimization problem together with a 1-dimensional search problem for satisfying the constraint equation.

II. THE MEM SPECTRUM ESTIMATION PROBLEM

The entropy of a Gaussian process with power spectrum \( S(k) \) is

\[
H(S) = \int_K \log S(k) \, dk
\]

where \( \mu \) is a positive measure and the spectral support \( K \) is a compact subset of \( \mathbb{R}^D \). Thus \( K \) is the region of the frequency space in which power is assumed to be present. The MEM spectrum estimate is that spectrum which maximizes \( H(S) \) subject to the following correlation constraints,

\[
r(\delta_i) = \int_K S(k) e^{-j k \cdot \delta_i} \, dk + \sigma_i^2
\]

\[i = 1, 2, \ldots, m\]

where \( r(\delta_i)'s \) are the correlation measurements, and \( \sigma_i^2 \)'s are independent, zero-mean random variables with variances \( \sigma_i^2 \). The above system of constraints can be replaced by the following single constraint which is determined by the noise statistics

\[
\sum_{i=1}^{m} \sigma_i^{-2} |r(\delta_i) - \hat{r}(\delta_i)|^2 = m
\]
where
\[ \hat{r}(\xi) = \int_{K} S(k) e^{j\xi \cdot \hat{K}} \, dk. \]

By the Central Limit Theorem, the above equality constraint will be very nearly satisfied when \( m \) is large.

By using the discrete formulation, we have the following constrained optimization problem:

\[
\begin{align*}
\text{max} \quad & J(\lambda; S_1, \ldots, S_n) = \Delta A \sum_{k=1}^{n} \log S_k - \lambda \cdot q(S_1, \ldots, S_n) \\
\text{subject to the constraints} \quad & q(S_1, \ldots, S_n) = \sum_{i=1}^{m} \left| r(\xi_i) - \Delta A \sum_{k=1}^{n} S_k e^{j\xi_i \cdot \hat{K}} \right| - m \\
& = (\Delta A)^2 \frac{2}{\sigma_0^2} p - (\Delta A) \sigma_0 d + \sum_{i=1}^{m} \frac{2}{\sigma_0^2} - m.
\end{align*}
\]

where \( \xi \) is the spatial frequency, \( \Delta A \) is the volume element and \( \hat{K} \) is the spatial coordinate, \( P_{1K} \) and \( d_1 \) are given by

\[
P_{1K} = 2 \sum_{i=1}^{m} \frac{1}{\sigma_i^2} \cos (1 - k) \xi_i
\]

and

\[
d_1 = 2 \sum_{i=1}^{m} \frac{1}{\sigma_i^2} r(\xi_i) \cos 1 \xi_i
\]

\[
S = (S_1 \ldots S_n)^T, \quad d = (d_1 \ldots d_n)^T, \quad P = (P_{1K})_{mn}
\]

\(
P \) can be proved to be positive semidefinite. No explicit solution is known for the constrained optimization problems just posed, and numerical solution of such problems is very difficult. The conventional approach is to vary \( \lambda \) and \( \{S_i\} \) in some iterative way. This approach is very time consuming because a nonlinear optimization procedure needed to be carried out for each \( \lambda \), and moreover some heuristic arguments needed to be introduced for the iterations on \( \lambda \).

Let us prove the uniqueness of the solution for the constrained optimization problem just posed, and numerical solution of such problems is very difficult. The conventional approach is to vary \( \lambda \) and \( \{S_i\} \) in some iterative way. This approach is very time consuming because a nonlinear optimization procedure needed to be carried out for each \( \lambda \), and moreover some heuristic arguments needed to be introduced for the iterations on \( \lambda \).

**Lemma 1** (Uniqueness) For any fixed \( \lambda > 0 \), the function \( J(\lambda; S_1, \ldots, S_n) \) has at most one maximum point.

**Proof**

\[
\nabla^2 J = -\Delta A \text{diag} \left[ \frac{1}{S_1^2}, \ldots, \frac{1}{S_n^2} \right] - \lambda (\Delta A)^2 P < 0
\]

since \( \text{diag} \left[ \frac{1}{S_1^2}, \ldots, \frac{1}{S_n^2} \right] \) is positive definite and \( \frac{1}{S_i^2} \) is positive semi-definite.

**Lemma 2** For any \( \lambda > 0 \), \( S^0 = (S_1^0 \ldots S_n^0) \) maximizes \( J(\lambda; S_1, \ldots, S_n) \) if and only if

\[
\lambda \Delta A \left[ S^0 - \left( \frac{1}{S_1^0}, \ldots, \frac{1}{S_n^0} \right) \right] = \lambda \Delta A \left[ S_1^0 \ldots S_n^0 \right]
\]

Proof. Using \( \nabla^2 J < 0 \), \( J \) is a concave function, and (1) is derived by setting \( \nabla J = 0 \).

**Theorem 1.** For some \( \lambda > 0 \), if \( S_1^0 \ldots S_n^0 \) maximizes \( J(\lambda; S_1, \ldots, S_n) \), then \( (\lambda_0; S_1^0 \ldots S_n^0) \) uniquely determines a trajectory maximum point of \( J \) corresponding to various values of \( \lambda > 0 \) which are governed by the following system,

\[
\frac{dS}{d\lambda} = d - (\Delta A) P S
\]

where \( S = (S_1 \ldots S_n) \) and \( S^0 = (S_1^0 \ldots S_n^0) \) and \( P \) is given by

\[
P = \text{diag} \left[ \frac{1}{S_1^0}, \ldots, \frac{1}{S_n^0} \right] + \lambda (\Delta A) P
\]

Proof. From the necessary and sufficient condition (1), we take the derivative of \( S \) with respect to \( \lambda \) and after some manipulations we obtain the differential equations (2) and maximal point of the unconstrained problem as initial condition.
The above system of differential equation enables us to find the other solution of \( J(\lambda; S_1, \ldots, S_n) \) when \( \lambda \) varies. In fact, by using a particular solution of (1) as the initial condition, the whole trajectory of the optimal solution \( S_j \) for varying \( \lambda \) can be computed rather accurately by using Euler's method. We can derive an iteration algorithm for solving the trajectory by discretizing equation (2).

\[
E_k (\delta^{k+1} - \delta^k) = (\lambda^{k+1} - \lambda^k) (d - (DA) P S_k^k)
\]

where \( E_k \) is the evaluation of \( E \) at \( S = \delta^k \).

The constrained optimization can then be solved by varying \( \lambda \) using Newton's method until the constraint equation is satisfied. The direction of search in \( \lambda \) can be obtained by taking the derivative of \( q \) with respect to \( \lambda \).

**Theorem 2.**

Along the path determined by (2), \( dq/\lambda < 0 \).

**Proof.**

\[
\frac{dq}{\lambda} = \nabla q - \frac{dS}{\lambda}
\]

\[
= -(DA)[d - (DA) P S]_T E^{-1} (d - (DA) P S) S < 0
\]

Since along the optimal trajectory, \( S \) satisfies the necessary and sufficient condition for the unconstrained optimum, thus we have from (1)

\[
[d - (DA) P S] = \frac{1}{S_1} \quad \frac{1}{\ldots} \quad \frac{1}{S_n}
\]

and together with \( E \) is positive definite, we obtain the above conclusion.

From Theorem 1 and 2, we derive the following algorithm.

**Algorithm**

1. Initial conditions. \( \lambda_0, S^0 \) (from unconstrained optimization)
2. If \( q(\lambda; S_1, \ldots, S_k) \) is the optimal solution, \( \epsilon \) is a given small number.
3. Iteration on \( \lambda^{k+1} \). If \( q(\lambda^{k+1}; S_1, \ldots, S_k) > 0 \), then \( \lambda^{k+1} \lambda_k \), otherwise \( \lambda^{k+1} = \lambda_k \).
4. Iteration on \( S^{k+1} \). From equation (3), we have

\[
E_k S^{k+1} = b_k
\]

where

\[
b_k = \frac{1}{S_k} \quad \frac{1}{\ldots} \quad \frac{1}{S_n}
\]

+ \( (2\lambda^{k+1} - \lambda_k) P S_k^k \)

5. Go back to (2)

By using this scheme, hopefully we could find a maximum point \( \dot{S}^* \) along the optimal trajectory such that \( q(S^*) = 0 \). In the choice of initial condition, \( \lambda_0 \) should be chosen such that the unconstrained maximization could be easily solved and \( q(S^0, \ldots, S^n) \) is as close to zero as possible. Large values of \( \lambda \) cause the total squared residual of the solution to be small. Some heuristic arguments and prior information may be injected into the choice of \( \lambda_0 \). A discussion of some strategies in choosing \( \lambda_0 \) in the context of Fourier synthesis can be found in [6].

We would also like to point out that in carrying out Euler's method, the inverse \( E_k \) needed not be obtained. In fact this system of symmetric linear equations (3) can be solved by using Cholesky method. This method still requires a lot of computation time and storage space. Thus we propose to use Gauss-Seidel procedure [8] for solving (4). This is an even more efficient algorithm which makes use of the positive definite and symmetric nature of \( E_k \). This algorithm can converge very fast when the initial condition is closed to the desired solution.

Let us look at the equation (3) again

\[
E_k x = b_k
\]

where \( x \) is the desired solution vector \( S^{k+1} \). We set forth the following iterative scheme

\[
x_0 = S^k
\]

\[
x_i^{m+1} = \sum_{i=1}^{n} \left( \frac{1}{(F_k)_{ij}} - \sum_{i=1}^{n} (E_k)_{ij} x_j^{m+1} \right)
\]

\[
\sum_{i=1}^{n} (F_k)_{ij} x_j^{m+1}
\]

It is noted that the algorithm will converge to the desired solution \( S^{k+1} \) for any arbitrary initial condition. However a choice of \( S^k \) as the initial condition will greatly accelerate the convergence rate. This is due to the fact that \( S^{k+1} \) is close to \( S^k \) when \( |\lambda^{k+1} - \lambda_k| \) is appropriately small. Thus it is reasonable to expect only a few iterations are required for the convergence when \( S^{k+1} \) is sufficiently close to \( S^k \).

**IV. DISCUSSION**

We have presented a computational approach for solving the primal problem of the MEM spectrum estimation when there is uncertainty in the correlation measurements. The problem posed is a constrained optimization in \( n + 1 \)-dimensional space. Instead of solving this problem directly, our approach depends first on solving an unconstrained optimization problem in a \( n \)-dimensional space and then an iterative search.
for satisfying the constraint equation in a 1-dimensional space. The unconstrained optimization can be solved by using any conventional algorithms. Heuristic arguments and prior information may be injected in the choice of $\lambda_0$. In solving the constraint equation, the spectral estimates are updated by an iteration procedure along the optimal trajectory for varying $\lambda$. This optimal trajectory is governed by a system of differential equations and the unconstrained solution as the initial condition. The direction of this 1-dimensional search along the optimal trajectory can be obtained easily by choosing the sign of the constraint equation. From the property that $\frac{d\lambda}{d\lambda}$ is negative along the optimal trajectory, the search can be carried out along one direction of $\lambda$ until the estimates are closed to the solution. In solving the spectrum estimates iteratively along the optimal trajectory, Gauss-Seidel algorithm is used for solving this positive symmetric system of equations. This algorithm will converge to the desired solution for any arbitrary initial condition. However, the choice of spectrum coefficients at the previous iteration as the initial estimate will greatly speed up the convergence rate. This is due to the fact that the spectrum estimates updates are close to each other when the $\lambda_{k+1}$ is appropriately close to $\lambda$. Thus, it is reasonable to expect that only a few iterations are required for the convergence of the Gauss-Seidel algorithm at each iteration.

REFERENCES


