MAXIMUM LIKELIHOOD ESTIMATION WITH SIDE INFORMATION OF 1-D LAYERED MEDIA FROM NOISY IMPULSE REFLECTION RESPONSES

Andrew E. Yagle and Rajashri R. Joshi

Dept. of EECS, The University of Michigan, Ann Arbor MI 48109-2122 aey@eecs.umich.edu

ABSTRACT

We consider the problem of computing the maximum likelihood estimates of the reflection coefficients of a discrete 1-D layered medium from noisy observations of its impulse reflection response. We have side information in that a known subset of the reflection coefficients are known to be zero; this knowledge could come from either a priori knowledge of a homogeneous subregion inside the scattering medium, or from a thresholding operation in which noisy reconstructed reflection coefficients with absolute values below a threshold are known to be zero. Our procedure is simple, noniterative, and requires only solutions of systems of linear equations. Numerical examples are provided which demonstrate not only the operation of the algorithm, but also that the side information improves the reconstruction of unconstrained reflection coefficients as well as constrained ones, due to the nonlinearity of the problem.

1. INTRODUCTION

The one-dimensional inverse scattering problem for a discrete layered medium probed with an impulsive plane wave at normal incidence arises in many areas. For example, the radar reflection from a stratified lossless dielectric medium (such as an airplane skin) can be used to reconstruct the permittivity of the medium in each layer if the magnetic permeability is assumed to be constant. Another example is the well-known problem of reconstructing a lossless layered acoustic medium from its impulse reflection response.

The goal of this paper is to present a simple algorithm that computes the constrained maximum-likelihood estimate of the reflection coefficients of a discrete layered medium, from measurements of the impulse reflection response to which a Gaussian noise process with known mean and covariance has been added. The side information consists of a subset of the reflection coefficients whose values are assumed to be known. Our treatment of this problem in the exact case (all multiple reflections are included; no approximations are used other than low noise level) seems to be new.

It is important to note that due to the nonlinearity of the exact (all multiple reflections included) inverse scattering problem, we cannot solve this problem by simply setting the noisy reconstructed reflection coefficients to their known values. We could do this in the Born approximation (linearization of the inverse scattering problem), but the nonlinearity implies that alteration of (interface) reflection coefficients will produce a complicated alteration of the impulse reflection response. Furthermore, two almost-identical impulse reflection responses can arise from two very-different sequences of (interface) reflection coefficients. We show using examples that it is necessary to alter some of the non-constrained reflection coefficients, as well as the constrained ones, to obtain the maximum-likelihood estimate. We also show that the side information improves reconstruction of *unconstrained* reflection coefficients as well as constrained ones, due to the nonlinearity of the inverse scattering problem.

2. PROBLEM FORMULATION

Wave propagation in layered media is described by

$$\begin{bmatrix} d_n(i) \\ u_n(i) \end{bmatrix} = \frac{1}{t_n} \begin{bmatrix} 1 & -r_n \\ -r_n & 1 \end{bmatrix} \begin{bmatrix} d_{n-1}(i-0.5) \\ u_{n-1}(i+0.5) \end{bmatrix} \quad (2.1)$$

where the reflection coefficient $r_n = \frac{Z_n - Z_{n+1}}{Z_n + Z_{n+1}}$, the transmission coefficient $t_n = \sqrt{1-r_n^2}$ and $d_n(i)$ and $u_n(i)$ are the downgoing and upgoing waves at time i just below the nth interface. The physical interpretation of $d_n(i)$ and $u_n(i)$ as energy-normalized waves being scattered into each other should be apparent. For details about (2.1) and its derivation see [1]-[3].

The medium has a free (perfectly reflecting) surface; this is reasonable since in many applications there is a huge impedance mismatch between the medium and its exterior (e.g. between a solid and air). The medium is probed by an impulse δ_n, resulting in the medium's impulse reflection response k_n:

$$d_0(n) = \delta_n + k_n; \quad u_0(n) = k_n. \quad (2.2)$$

We assume we know that $r_{n_i} = 0$ for some known depths $n_1, n_2 \dots n_M$. Such information could come either from previous knowledge of a homogeneous slab inside the layered medium, or from thresholding small noisy estimates of r_n (if the noisy reconstructed $|r_n| < \eta$ for some threshold η set $r_n = 0$). Such thresholding strategies are common in signal processing.

We also assume we are given noisy observations

$$y_n = k_n + v_n, 1 \leq n \leq N = medium\,thickness \quad (2.3)$$

where v_n is a Gaussian noise process with known mean m_n and covariance $\lambda_{i,j} = E[(v_i - m_i)(v_j - m_j)]$. Since we can immediately convert the non-zero-mean problem into a zero-mean problem by replacing y_n with $y_n - m_n$ we assume without loss of generality in the sequel that $m_n = 0$. We also assume that the signal-to-noise level is high, i.e. $k_n^2 >> \lambda_{n,n} = \sigma_n^2$.

The goal is to reconstruct the maximum-likelihood estimates $\{\hat{r}_i, 1 \leq i \leq N\}$ of the r_n from the noisy observations $\{y_i, 1 \leq i \leq N\}$ of the impulse reflection response $\{k_i, 1 \leq i \leq N\}$, subject to the constraint of the side information $r_{n_i} = 0$ for some known depths $n_1, n_2 \dots n_M$.

3. PROBLEM SOLUTION

3.1. Log-likelihood function

Define vectors $\vec{y} = [y_1, y_2 \dots y_N]^T$, $\vec{v} = [v_1, v_2 \dots v_N]^T$, $\vec{k} = [k_1, k_2 \dots k_N]^T$, and $\vec{r} = [r_1, r_2 \dots r_N]^T$. Also define symmetric Toeplitz matrices Y, V, K whose first columns are $\vec{y}, \vec{v}, \vec{k}$, respectively. Also define the covariance matrix Λ whose $(i,j)th$ element is $\lambda_{i,j}$. Let $\hat{r}_n, \hat{k}_n, \hat{v}_n$ be maximum likelihood estimates of r_n, k_n, v_n, and $p_{\vec{x}}(\vec{x})$ be the probability density function for $\vec{x}$.

The log-likelihood function is $\log p_{\vec{y}|\vec{r}}(\vec{y}|\vec{r})$

$$= \log \frac{1}{(2\pi)^{N/2}\sqrt{|\Lambda|}} exp[-\frac{1}{2}(\vec{y} - \vec{k}(\vec{r}))^T \Lambda^{-1}(\vec{y} - \vec{k}(\vec{r}))]$$

where $\vec{k}(\vec{r})$ is the nonlinear functional relation between the r_n and the k_n specified by (2.1) and (2.2). That is, $\vec{k}$ is the reflection response resulting from reflection coefficients $\vec{r}$.

The maximum likelihood estimate $\hat{\vec{r}}$ of $\vec{r}$ is $\hat{\vec{r}} =$

$$\underset{\vec{r}}{argmax} \log p_{\vec{y}|\vec{r}}(\vec{y}|\vec{r}) = \underset{\vec{r}}{argmin} (\vec{y} - \vec{k}(\vec{r}))^T \Lambda^{-1}(\vec{y} - \vec{k}(\vec{r})).$$

Hence we must determine the $\vec{r}$ that minimizes the least-squares norm with respect to covariance matrix Λ between the resulting $\vec{k}(\vec{r})$ and the given data $\vec{y}$. That is, we must find the smallest perturbation of the given $\vec{y}$ consistent with the nonlinear constraints $r_{n_i} = 0$. It is not at all obvious how to do this at this point.

3.2. Use of Toeplitz systems of equations

The key to solving easily this nonlinearly-constrained least-squares minimization problem is to note that we can solve the 1-D discrete inverse scattering problem defined by (2.1) and (2.2) by solving the Toeplitz system of equations [2]-[5]

$$\begin{bmatrix} 1 & k_1 & k_2 & \dots & k_n \\ k_1 & 1 & k_1 & \dots & k_{n-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ k_n & k_{n-1} & k_{n-2} & \dots & 1 \end{bmatrix} \begin{bmatrix} a_n(0) \\ a_n(1) \\ \vdots \\ a_n(n) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (3.3)$$

for $n = 1, 2 \dots N$ and noting that $r_n = a_n(n) \prod_{i=1}^{n-1} t_i$ [1]-[3]. These nested Toeplitz systems of equations can easily be solved using the Levinson algorithm.

Write (3.3) as $K\vec{a} = \vec{t}$ where $\vec{a} = [a_n(0) \dots a_n(n)]^T$ and $\vec{t} = [1, 0 \dots 0]^T$. Also write (2.3) as $Y = K + V$. Here Y, K, V are submatrices of the Toeplitz matrices defined above (use n=current layer instead of N=total number of layers). When we attempt to reconstruct the medium from the noisy observations y_n, we obtain $Y\tilde{\vec{a}} = \vec{t}$, where tildes denote quantities associated with the noisy, unconstrained data.

Now we can proceed. We have

$$K\vec{a} = \vec{t} \rightarrow \vec{a} = K^{-1}\vec{t} = (Y - V)^{-1}\vec{t} \approx (Y^{-1} + Y^{-1}VY^{-1})\vec{t} \quad (3.4)$$

where we have used the approximation

$$(Y - V)^{-1} = (I - Y^{-1}V)^{-1}Y^{-1}$$

$$= (I + Y^{-1}V + (Y^{-1}V)^2 + \dots)Y^{-1} \approx Y^{-1} + Y^{-1}VY^{-1}. \quad (3.5)$$

We can expand $(I - Y^{-1}V)^{-1}$ in a power series in $Y^{-1}V$ provided $Y^{-1}V < I \rightarrow V < Y$ where $A < B$ means $A - B$ is positive definite. We can truncate the power series to one term provided $V << Y$, which is true provided the signal-to-noise ratio is large.

This series truncation is the only assumption required in our work. As long as $V < Y$ (which will surely be the case in practice) we can, if necessary, iterate our procedure. Since each iteration produces a consistent $\{k_n, r_n\}$, we can repeat the procedure as many times as desired, with the accuracy improving since the maximum singular value of $V^{-1}Y$ is getting smaller. In practice we have found one or two iterations to be sufficient to satisfy the constraints. We now have

$$\vec{a} = (Y^{-1} + Y^{-1}VY^{-1})\vec{t} = \tilde{\vec{a}} + Y^{-1}V\tilde{\vec{a}}. \quad (3.6)$$

Let J be an exchange matrix with ones on the main antidiagonal and zeros elsewhere. Recall $Y = JYJ$ for symmetric Toeplitz matrix. Premultiplying $[0, 0 \dots 1] =$

$\vec{t}^T J$ and $\tilde{\vec{a}}^T = \vec{t}^T Y^{-1}$, $r_n = a_n(n)\prod_{i=1}^{n-1} t_i$ gives

$$0 = r_n / \prod_{i=1}^{n-1} t_i = \tilde{r}_n / \prod_{i=1}^{n-1} \tilde{t}_i + (J\vec{\tilde{a}})^T V \vec{\tilde{a}} \quad n = n_1 \dots n_M \tag{3.7}$$

since the side information is that $r_n = 0$ for these n.

3.3. Linear system of equations

Now rewrite (3.7) as a linear system of equations

$$W\vec{v} = -\underline{\tilde{\vec{r}}} \tag{3.8}$$

where $\underline{\tilde{\vec{r}}}$ is a subvector of $\tilde{\vec{r}}$ consisting of the $n_1, n_2 \dots n_M$ elements of $\tilde{\vec{r}}$ and W is composed of known elements of $\tilde{\vec{a}}$. From (2.3) and the fact that maximum likelihood estimation commutes with nonlinear operations [6], it is clear that estimating r_n from y_n is equivalent to estimating k_n from y_n, which in turn is equivalent to estimating $v_n = y_n - k_n$ from y_n.

The linear system of equations (3.8) is clearly underdetermined. But it does express the side information in the form of a linear system of equations in $\vec{v}$. We know we wish to minimize (3.2) subject to this constraint. This shows immediately that we want the least-squares solution to (3.8), specifically the solution to (3.8) minimizing the norm defined in (3.2). This is

$$\hat{\vec{v}} = -\Lambda W^T (W\Lambda W^T)^{-1} \underline{\tilde{\vec{r}}} \tag{3.9}$$

which can easily be computed since the size of the system of equations is the number $M << N$ of constrained values of r_n.

In practice the additive Gaussian noise process will often be white. In this case $\Lambda = DIAG[\sigma^2]$ and (3.8) simplifies to the pseudoinverse

$$\hat{\vec{v}} = -W^T (WW^T)^{-1} \underline{\tilde{\vec{r}}}. \tag{3.10}$$

Note that if there is no noise in the data then the reconstructed $\tilde{r}_n$ will all be zero, so that $\hat{\vec{v}}$ will properly be a zero vector.

3.4. Summary of overall procedure

1. Run the Levinson algorithm to solve the nested systems of Toeplitz equations (3.3) using the noisy observations y_n instead of k_n. This results in $\{\tilde{a}_n(i), i = 1 \dots n, n = 1 \dots N\}$ and $\tilde{r}_n, n = 1 \dots n$; $\tilde{r}_n = a_n(n)\prod_{i=1}^{n-1} \tilde{t}_i$ is computed recursively.

2. Compute the pseudoinverse in (3.9) or (3.10), where W is determined from the quantities computed using the Levinson algorithm by rearranging (3.7). The solution is the maximum likelihood estimate of the additive noise v_n.

3. Compute $\hat{k}_n = y_n - \hat{v}_n$ where $\hat{v}_n$ is from the solution to (3.9) or (3.10). $\hat{k}_n$ is the maximum likelihood estimate of k_n.

4. Run the Levinson algorithm on $\hat{k}_n$. This computes the maximum likelihood estimates $\hat{r}_n$ of r_n subject to the constraint $r_{n_i} = 0$.

4. NUMERICAL EXAMPLES

The example is a continuous 1-D inverse scattering problem with $r(x)$ a realization of a $1/f$ fractal process except for $0.75 < x < 1.5$ where $r(x)$ was set to zero. The continuous problem was solved by discretizing the depth interval $0 < x < 3$ to 128 discrete points and running the discrete algorithm described above. The "true" $r(x)$ and corresponding "true" impulse reflection response $k(t)$ are shown in Figs. 1b and 1a, respectively; note $r(x) = 0$ for $0.75 < x < 1.5$. Zero-mean white Gaussian noise $v(t)$ was added to $k(t)$ to produce the noisy observation $y(t)$ in Fig. 1c. Our given information consists of $y(t)$ and $r(x) = 0$ for $0.75 < x < 1.5$.

The $r(x)$ reconstructed directly from $y(t)$ without using the side information is shown in Fig. 1d. Note that although $y(t)$ is not that noisy an observation of $k(t)$ (compare Figs. 1a and 1c), the reconstructed $r(x)$ shown in Fig. 1d is quite different from the true $r(x)$. Again, it is hard to see from Fig. 1d that $r(x)$ should be zero for $0.75 < x < 1.5$, even though we know this.

Our procedure was run twice on the noisy data $y(t)$. The results of the first run are shown in Figs. 1e and 1f. Note that in Fig. 1f the constraint is almost satisfied. Also note that the reconstruction of the *unconstrained* $r(x)$ outside the range $0.75 < x < 1.5$ is improved, especially for $0.5 < x < 0.75$ and $1.5 < x < 2.5$. This shows that the reconstruction of the *unconstrained* $r(x)$, as well as the constrained $r(x)$, is improved by our procedure. This is due to the nonlinearity of the inverse scattering problem: $r(x)$ affects not only $k(x)$ (as it does in the Born approximation) but all $k(t), t > x$ due to the multiple scattering, so different ranges of x and t affect each other.

The results of the second run are shown in Figs. 1g and 1h. There is virtually no change from Figs. 1e and 1f, except that the constraint is satisfied even more closely. This again suggests that a single run of the algorithm suffices.

5. ACKNOWLEDGMENT

This work was supported in part by the Office of Naval Research under grant #N00014-94-0959.

6. REFERENCES

[1] K.P. Bube and R. Burridge, "The one-dimensional inverse problem of reflection seismology," *SIAM Review* 25, 497-559, 1983.

[2] A.M. Bruckstein and T. Kailath, "Inverse scattering for discrete transmission-line models," *SIAM Review* 29, 359-389, 1987.

[3] A.M. Bruckstein, B.C. Levy and T. Kailath, "Differential methods in inverse scattering," *SIAM J. Applied Math.* 45, 312-335, 1985.