FASTMAP: A FAST, APPROXIMATE MAXIMUM *A POSTERIORI* PROBABILITY PARAMETER ESTIMATOR WITH APPLICATION TO ROBUST MATCHED-FIELD PROCESSING

Brian F. Harrison

Naval Undersea Warfare Center
Code 2121
Newport, RI 02841 USA

Richard J. Vaccaro and Donald W. Tufts

Department of Electrical Engineering
University of Rhode Island
Kingston, RI 02881 USA

ABSTRACT

In many estimation problems, the set of unknown parameters can be divided into a subset of desired parameters and a subset of nuisance parameters. Using a maximum *a posteriori* (MAP) approach to parameter estimation, these nuisance parameters are integrated out in the estimation process. This can result in an extremely computationally-intensive estimator. This paper proposes a method by which computationally-intensive integrations over the nuisance parameters required in Bayesian estimation may be avoided under certain conditions. The propsed method is an approximate MAP estimator which is much more computationally efficient than direct, or even Monte Carlo, integration of the joint posteriori distribution of the desired and nuisance parameters. As an example of its efficiency, we apply the fast algorithm to matched-field source localization in an uncertain environment.

1. INTRODUCTION

In many estimation problems, the set of unknown parameters can be divided into two subsets. The parameters of major interest comprise one subset and the remaining parameters comprise the second subset. The remaining parameters only serve to complicate the problem and are referred to as nuisance parameters. In the maximum *a posteriori* (MAP) approach to parameter estimation, these nuisance parameters are treated as random variables with assumed prior probability density functions and integrated out in the process of estimating the desired parameters [1], [2]. Depending on the number of nuisance parameters, practical implementation of the MAP estimator can be extremely computationally-intensive. In the next section, we will show that under certain conditions a computationally-efficient approximation to the MAP estimator can be obtained. Efficiency is achieved by approximately performing the integration off-line prior to the processing of data observations.

We will derive the MAP estimator following the development of Richardson and Nolte [3]. Consider the following data model

$$\mathbf{y} = s\mathbf{a}(\mathbf{\Theta}, \mathbf{\Psi}) + \mathbf{n}, \tag{1}$$

where $\mathbf{y}$ is a $N \times 1$ vector of observed data, s is a Gaussian distributed complex random variable, $\mathbf{a}(\mathbf{\Theta}, \mathbf{\Psi})$ is a $N \times 1$ vector parameterized by the vectors $\mathbf{\Theta}$ and $\mathbf{\Psi}$, and $\mathbf{n}$ is a $N \times 1$ vector of Gaussian random variables. This model is applicable to array processing and time-series analysis problems. We will assume that the parameters of interest are contained in $\mathbf{\Theta}$, while $\mathbf{\Psi}$ contains the nuisance parameters. Each of the nuisance parameters is assumed to be a random variable with a known uniform probability distribution. We can then express the *a posteriori* probability density function (pdf) of $\mathbf{\Theta}$ as

$$p(\mathbf{\Theta}|\mathbf{y}) = \frac{p(\mathbf{\Theta})}{p(\mathbf{y})} p(\mathbf{y}|\mathbf{\Theta}). \tag{2}$$

Assuming that s, $\mathbf{\Psi}$, and $\mathbf{\Theta}$ are all independent of each other, the *a posteriori* pdf of $\mathbf{\Theta}$ is given by

$$p(\mathbf{\Theta}|\mathbf{y}) = C(\mathbf{y})p(\mathbf{\Theta}) \int_{\mathbf{\Psi}} \frac{\exp\left\{\frac{H(\mathbf{\Theta},\mathbf{\Psi})}{G(\mathbf{\Theta},\mathbf{\Psi})}\right\}}{G(\mathbf{\Theta},\mathbf{\Psi})} p(\mathbf{\Psi})\, d\mathbf{\Psi}, \tag{3}$$

where $C(\mathbf{y})$ is a normalization constant, $p(\mathbf{\Theta})$ is assumed uniform over the parameter space of interest, and

$$\begin{aligned} H(\mathbf{\Theta},\mathbf{\Psi}) &= \left|\tfrac{\sigma_A}{\sqrt{2}}\mathbf{a}^H(\mathbf{\Theta},\mathbf{\Psi})\mathbf{R}_\mathbf{n}^{-1}\mathbf{y}\right|^2, \\ G(\mathbf{\Theta},\mathbf{\Psi}) &= \sigma_A^2\mathbf{a}^H(\mathbf{\Theta},\mathbf{\Psi})\mathbf{R}_\mathbf{n}^{-1}\mathbf{a}(\mathbf{\Theta},\mathbf{\Psi}) + 1, \end{aligned} \tag{4}$$

where σ_A^2 is the signal amplitude variance and $\mathbf{R}_\mathbf{n} = \sigma_n^2 E\{\mathbf{n}\mathbf{n}^H\}$ where σ_n^2 is the noise variance. The maximum *a posteriori* (MAP) estimate of $\mathbf{\Theta}$ is obtained by maximizing (3) over $\mathbf{\Theta}$, i.e.,

$$\hat{\mathbf{\Theta}} = \arg\max_{\mathbf{\Theta}} \int_{\mathbf{\Psi}} \frac{\exp\left\{\frac{H(\mathbf{\Theta},\mathbf{\Psi})}{G(\mathbf{\Theta},\mathbf{\Psi})}\right\}}{G(\mathbf{\Theta},\mathbf{\Psi})} p(\mathbf{\Psi})\, d\mathbf{\Psi}. \tag{5}$$

If the vectors $\mathbf{a}(\mathbf{\Theta},\mathbf{\Psi})$ are normalized to have unit norm and $\mathbf{R}_\mathbf{n}$ equals the identity matrix, equation (5) can be written as

$$\hat{\mathbf{\Theta}} = \arg\max_{\mathbf{\Theta}} \int_{\mathbf{\Psi}} \exp\{F(\mathbf{\Theta},\mathbf{\Psi})\}\, p(\mathbf{\Psi})\, d\mathbf{\Psi}, \tag{6}$$

where

$$F(\mathbf{\Theta},\mathbf{\Psi}) = \frac{\frac{\sigma_A^2}{2}\left|\mathbf{a}^H(\mathbf{\Theta},\mathbf{\Psi})\mathbf{y}\right|^2}{\sigma_A^2 + 1}. \tag{7}$$

For numerical implementation of this estimator, we assume that $p(\mathbf{\Psi})$ is a uniformly distributed pdf and approximate the integral by a summation

$$\hat{\mathbf{\Theta}} = \arg\max_{\mathbf{\Theta}} \sum_{j=1}^{M} \exp\{F(\mathbf{\Theta},\mathbf{\Psi}_j)\}, \tag{8}$$

where the $\mathbf{\Psi}_j$ are vectors of the nuisance parameters sampled from their probability distributions. A Monte Carlo approach to computing (8) was proposed in [5]. Thus M inner products between the vectors $\mathbf{a}(\mathbf{\Theta},\mathbf{\Psi}_j)$, $j = 1,\ldots,M$ and $\mathbf{y}$ are computed for each point in parameter space $\mathbf{\Theta}$. We remark that usually $M \gg N$.

2. COMPUTATIONALLY-EFFICIENT MAXIMUM *A POSTERIORI* PARAMETER ESTIMATION

The exponential in (8) can be approximated over a finite interval by a linear approximation of the form

$$\exp\{F(\mathbf{\Theta},\mathbf{\Psi}_j)\} \approx a\{F(\mathbf{\Theta},\mathbf{\Psi}_j)\} + b. \tag{9}$$

for some constants $a > 0$ and b. Substituting this approximation into (8) gives

$$\hat{\mathbf{\Theta}} = \arg\max_{\mathbf{\Theta}}\ Mb + \frac{a\sigma_A^2}{2\sigma_A^2 + 2} \sum_{j=1}^{M} \left|\mathbf{a}^H(\mathbf{\Theta},\mathbf{\Psi}_j)\mathbf{y}\right|^2, \tag{10}$$

or equivalently

$$\hat{\mathbf{\Theta}} = \arg\max_{\mathbf{\Theta}} \sum_{j=1}^{M} \left|\mathbf{a}^H(\mathbf{\Theta},\mathbf{\Psi}_j)\mathbf{y}\right|^2. \tag{11}$$

Notice that if the perturbations of the $\mathbf{a}(\mathbf{\Theta},\mathbf{\Psi}_j)$ vectors for each $\mathbf{\Theta}$ over the realizations $\mathbf{\Psi}_j$ are small, then the variability of the values of $\left|\mathbf{a}^H(\mathbf{\Theta},\mathbf{\Psi}_j)\mathbf{y}\right|^2$ will also be small. Clearly, the magnitude of the perturbations depends on the sensitivity of $\mathbf{a}(\mathbf{\Theta},\mathbf{\Psi}_j)$ to the realizations of $\mathbf{\Psi}_j$ and would be dependent on the particular problem of interest. When the perturbations are small, (11) will be a close approximation to (8). Equation (11) can be rewritten as

$$\hat{\mathbf{\Theta}} = \arg\max_{\mathbf{\Theta}} \mathbf{y}^H\mathbf{R}_\mathbf{a}(\mathbf{\Theta})\mathbf{y}, \tag{12}$$

where

$$\mathbf{R}_\mathbf{a}(\mathbf{\Theta}) = \sum_{j=1}^{M} \mathbf{a}(\mathbf{\Theta},\mathbf{\Psi}_j)\mathbf{a}^H(\mathbf{\Theta},\mathbf{\Psi}_j). \tag{13}$$

It has been our experience that the matrix $\mathbf{R}_\mathbf{a}(\mathbf{\Theta})$ is usually well approximated by a matrix of lower rank. This can be verified by computing the eigen decomposition $\mathbf{R}_\mathbf{a}(\mathbf{\Theta}) = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H$. Small eigenvalues in Σ can be set to zero [6]. A similar approach is used in [5] for computing a low-rank approximation to an environmental constraint matrix used by the MV-EPC processor. After setting the small eigenvalues in $\mathbf{\Sigma}$ to zero, we define $\mathbf{\Sigma}_1$ to be a diagonal matrix containing the non-zero eigenvalues

and $\mathbf{U}_1$ as a matrix of the corresponding eigenvectors. A low-rank approximation to $\mathbf{R_a}(\mathbf{\Theta})$ is given by $\tilde{\mathbf{R}}_\mathbf{a}(\mathbf{\Theta}) = \mathbf{U}_1\mathbf{\Sigma}_1\mathbf{U}_1^H$. Substituting this approximation into (12) and rewriting the expression as the norm-squared of a matrix-vector product gives

$$\hat{\mathbf{\Theta}} = \arg\max_{\mathbf{\Theta}} \|\mathbf{B_a}(\mathbf{\Theta})\mathbf{y}\|^2, \qquad (14)$$

where $\mathbf{B_a}(\mathbf{\Theta}) = \mathbf{\Sigma}_1^{\frac{1}{2}}\mathbf{U}_1^H$.

The advantage to using the approximate MAP estimator of (14) over the MAP estimator of (8) is that the computation of $\mathbf{B_a}(\mathbf{\Theta})$ can be done off-line before the processing of data. The number of on-line inner product computations over each point in the desired parameter space $\mathbf{\Theta}$ is reduced and equal to q, the rank of $\mathbf{B_a}(\mathbf{\Theta})$, where $q \leq N$. In contrast, (8) must compute M inner products over each point $\mathbf{\Theta}$ on-line, where $M \gg N \geq q$. Thus, (14) is a computationally-efficient approximation to the MAP estimator and will be called the *fast, approximate* MAP estimator or FASTMAP. We conjecture that $\mathbf{R_a}(\mathbf{\Theta})$ rank deficient implies that the linear approximation in (9) is accurate. This is a topic of current investigation. FASTMAP has been applied effectively in the array processing technique known as matched-field processing [7] for robustness to environmental uncertainty. A number of examples of this application, including experimental as well as simulated data, are given in [8].

3. A MATCHED-FIELD PROCESSING APPLICATION

In this section we will demonstrate the utility of the FASTMAP estimator using the array processing technique known as matched-field processing (MFP) . MFP is a model-based source localization method which is a function of the ocean environmental parameters. Precise knowledge of the environmental parameters are required. When the environmental parameters are uncertain, a MAP approach to MFP can be used for robustness to the uncertainty. Using the model of (1) for this case, $\mathbf{\Theta}$ would contain the source location parameters of range and depth and the uncertain environmental parameters would be contained in $\mathbf{\Psi}$. The vector $\mathbf{a}(\mathbf{\Theta}, \mathbf{\Psi})$ is called a *replica vector* in the matched-field literature.

PARAMETER	GENLMIS-S	GENLMIS
water depth	102.5±1 m	102.5±2.5 m
surface sound speed	1500±2.5 m/s	1500±2.5 m/s
bottom sound speed	1480±2.5 m/s	1480±2.5 m/s
uppermost layer sound speed	1600±10 m/s	1600±50 m/s
lowermost layer sound speed	1750±10 m/s	1750±100 m/s
bottom attenuation	.35±0.1 dB/λ	.35±0.25 dB/λ
bottom density	1.75±0.1	1.75±0.25

Table 1: Uncertainty intervals of environmental parameters.

We will use the Naval Research Laboratory benchmark environment to test our algorithm. This is a simulated shallow-water environment, containing uncertainties, designed as a standardized testbed for MFP algorithms. It contains seven uncertain environmental paramaters. The details of this environment can be found in [5],[8],[9]. We will use the two cases of environmental uncertainty, GENLMIS-S and GENLMIS defined in Table 1.

We have shown previously that computationally-efficient smoothing of the range/depth ambiguity surface can provide robustness to the gridding of the surface [10]. This smoothing has been applied to the FASTMAP estimator in the formulation of a two-step MFP algorithm called MU-RSWP [8]. It is this algorithm, of which FASTMAP is the basis, that we will demonstrate here. The performance of MU-RSWP was compared to that of the conventional Bartlett processor over 100 Monte Carlo trials for both cases. The Bartlett processor assumed the nominal values of the environmental parameters. In each trial, a unique randomly selected source position and environmental realization was chosen. For the GENLMIS-S case, Bartlett achieved a correct localization percent-

age of 67%, while MU-RSWP achieved 100%. For the larger uncertainty GENLMIS case, Bartlett achieved 34%, while MU-RSWP achieved 80%. The results for MU-RSWP are also consistant with the theoretical correct localization percentages given in [5] for the MAP estimator.

A computational comparison of MU-RSWP and the MAP estimator has also been performed using the GENLMIS case [8]. The basis for this comparison was the number of inner products between replica vectors and the data vector required to process a single snapshot of array data. It was shown that MU-RSWP provided a substantial 75:1 computational savings over the MAP estimator.

4. CONCLUSION

The FASTMAP estimator, a computationally efficient approximation to the maximum *a posteriori* (MAP) estimator, was presented. It was shown that the FASTMAP estimator is much more computationally efficient than direct, or even Monte Carlo, integration of the joint posteriori distribution. A matched-field processing application of the FASTMAP estimator was shown to illustrate its effectiveness and efficiency.

5. REFERENCES

[1] H.L. Van Trees, *Detection, Estimation, and Modulation Theory*, Part I, J. Wiley, 1968.

[2] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, 1993.

[3] A.M. Richardson and L.W. Nolte, "*A posteriori* probability source localization in an uncertain sound speed, deep ocean environment," *J. Acoust. Soc. Amer.*, vol. 89, 2280-2284, 1991.

[4] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 1984.

[5] J.A. Shorey, L.W. Nolte, and J.L. Krolik, "Computationally efficient monte carlo estimation algorithms for matched field processing in uncertain ocean environments," *J. Comput. Acous.*, vol. 2, pp. 285-314, 1994.

[6] B. Noble and J.W. Daniel, *Applied Linear Algebra*, Prentice-Hall, 1988.

[7] A.B. Baggeroer, W.A. Kuperman, and P.N. Mikhalevsky, "An overview of matched field methods in ocean acoustics," *IEEE Journal of Oceanic Engr.*, vol. 18, pp. 401-424, 1993.

[8] B.F. Harrison, *Robust Matched-Field Processing: A Subspace Approach*, PhD dissertation, The University of Rhode Island, 1996.

[9] M.B. Porter and A. Tolstoy, "The matched field processing benchmark problems," *J. Comput. Acous.*, vol. 2, pp. 161-185, 1994.

[10] B.F. Harrison, R.J. Vaccaro, and D.W. Tufts, "Robust source localization in an acoustic waveguide," *J. Acoust. Soc. Amer.,* vol. 100, pp. 384-391, 1996.