A GENERALIZED MAXIMUM LIKELIHOOD ESTIMATION ALGORITHM FOR PASSIVE DOPPLER-BEARING TRACKING

Xiao–Jiao Tao, Cai–Rong Zou and Zhen–Ya He *
DSP Division, Department of Radio Engineering
Southeast University
Nanjing 210018, P.R. China
Tel: +86-25-6631700 Ext. 2481
Fax: +86-25-7712719

ABSTRACT

Estimation of the trajectory of a target from a passive sonar's bearings and frequency measurements in the presence of multivariate normally distributed noise, with unknown inhomogeneous general covariance, is modelled as a nonlinear multiresponse parameter estimation problem. It is shown that Maximum Likelihood Estimation in this case is identical to optimizing a determinant criterion which has a concise form and contains no elements of unknown covariance matrix. An effective Guass-Newton type algorithm, using only the first-order derivatives of the model function, is presented to implement such estimation. The simulation shows that the proposed approach is superior to the traditional estimation methods especially under the condition of strong inhomogeneity of noise covariance and high correlation between different types of measurement noises.

1. INTRODUCTION

Passive Doppler-bearing tracking (DBT) is the determination of the trajectory of a target solely from noise-corrupted bearings and Doppler-shifted frequency measurements of signals originating from the target. DBT have two advantages over conventional bearings-only tracking (BOT): no requirement of a maneuver of the ownship for the observability of target motion parameter, and the more precise estimate obtained [2]. Traditionally, it always assumes that: the noises are independent and normally distributed while the variances of noises are time-shifted invariant [1]-[5]. Thus, Maximum Likelihood Estimation (MLE) for BOT is just the nonlinear least square estimation [1], and MLE for DBT is of a weighted least square form [5]. However, this assumption may not be realistic. It could be reasonable to assume that the variance of the same type measurement noise is constant within a relative short period but not for all the time, as both the target's state relative to the ownship and the environmental conditions may have changed remarkably. Furthermore, assuming independent noises for different type measurements at the same instant could not be justified, as all of these measurements come out of the same noise-contaminated signal and "asymptotically uncorrelated" property does not hold for the situations where only small number of samples are available.

In this paper, we present a generalized MLE algorithm to solve the DBT problem under the more general assumption that the noises are multivariate-normally distributed and have the unknown inhomogeneous covariance matrix with any forms. Although the combination of linearization, poor condition and small signal-to-noise ratio is one of reasons for using nonrecursive estimation method with favorable numerical properties [5], the most important one for using MLE in this paper is its efficiency to deal with the arbitrary forms of unknown inhomogeneous noise covariance matrix. Such advantage can be shown in the two steps of the algorithm: First, the generalized MLE is simply identical to minimizing a concise criterion, so-called determinant criterion, which does not contain any terms of unknown noise covariance matrix. Secondly, a novel Guass-Newton type algorithm is presented to implement such iterative optimizing procedure, where it needs only the first-order derivatives of the model function to construct the gradient and an approximate Hessian through a proposed simple transformation. Thus, the proposed approach gives more precise ouput and needs less computations than conventional batch process method.

2. THEORY

The measurement equation can be written in a general form, referred to as the so-called multiresponse model,

$$Y = F(X) + Z \tag{1}$$

* This work was supported by the Climbing Programme – National Key Project for Fundamental Research in China under Grant NSC 92097.

$$Y = \begin{bmatrix} \beta_0 & f_0^0 & f_0^1 & \cdots & f_0^{M-1} \\ \beta_1 & f_1^0 & f_1^1 & \cdots & f_1^{M-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_{N-1} & f_{N-1}^0 & f_{N-1}^1 & \cdots & f_{N-1}^{M-1} \end{bmatrix}$$

where $X = [R, \beta, v, \theta]^T$ denotes the relative state vector of target to ownship (range R, bearing β, speed v and course θ). Y is $N \times M$ observation matrix whose (i,j)-th element denotes the ith type of measurements at the jth instant. β_i and f_i^j denotes bearing angle and frequency measurement respectively. F is the $N \times M$ model function matrix depending on the unknown parameter X, and Z is the $N \times M$ noise matrix.

In this tracking problem the purpose is to estimate the state vector X from a set of measurements Y. Such estimate $\hat{X}$ are given by the values of X which optimize some criterion. The criterion will depend on the assumptions about the noise Z. For example, for BOT we make the stringent assumption that Z are normally distributed and independent with the same variance σ^2, then least squares is appropriate and it leads to the trace criterion, i.e. minimizing $tr(Z^T Z)$. This paper extends the noise assumption to the more general case, as listed below.

Assumption 1: The observation period N is uniformly decomposed into a finite number of segments n, during which the noise Z_k $(k = 0, 1, \cdots, K)$ of measurements are multivariate normally distributed with the unknown covariance, i.e. $Z_k \sim N(0, \Omega_k)$.

Assumption 2: At the same instant, noises of different types of measurements are correlated, $\{\Omega_k\}_{ij} \neq 0 \quad \forall i \neq j \quad (k = 0, \cdots, K)$

Assumption 3: In the k-th segment, the covariance of noises is constant but it differs from that of another segment k' $(k' \neq k)$, i.e. inhomogeneous covariance as $\Omega_k \neq \Omega_{k'}$ $(k' \neq k)$.

Assumption 4: Noises of measurements at the different instants are independent, i.e. $E[Z_k^T Z_{k'}] = 0 \quad \forall k \neq k' \quad (k, k' = 0, \cdots, K)$

Obviously, the joint probability density function for the N observations, conditional on all the unknown parameters, is,

$$p(Y|X,\Omega) \propto \prod_{k=0}^{K} |\Omega_k|^{-n/2} exp\left[-\frac{tr(Z_k \Omega_k^{-1} Z_k^T)}{2}\right] \quad (2)$$

where the vertical bars $|\cdot|$ denotes a determinant. Then the MLE is equivalent to maximize the loglikelihood function,

$$L(X, \Omega_0^{-1}, \cdots, \Omega_K^{-1}) = constant + \sum_{k=0}^{K} \left[\frac{N_k}{2} ln|\Omega_k^{-1}| - \frac{tr(Z_k \Omega_k^{-1} Z_k^T)}{2}\right] \quad (3)$$

Write the last term as $tr(Z_k^T Z_k \Omega_k^{-1})$ and differentiate the entire expression with respect to the (i,j)th element σ_k^{ij} of Ω_k^{-1}. Using the result [6]

$$\frac{\partial ln|\Omega_k^{-1}|}{\partial \sigma_k^{ij}} = \{\Omega_k\}_{ij} \quad (4)$$

allows to write,

$$\frac{\partial L(X, \Omega_0^{-1}, \cdots, \Omega_K^{-1})}{\partial \sigma_k^{ij}} = \frac{n}{2}\{\Omega_k\}_{ij} - \frac{1}{2}\{Z_k^T Z_k\}_{ij} \quad (5)$$

Then, setting this derivative to zero provides the conditional estimate,

$$\{\hat{\Omega}_k\} = \frac{\{Z_k^T Z_k\}}{n} \quad (6)$$

which, when substituted into (3), gives the conditional loglikelihood function,

$$L(X, \hat{\Omega}_0^{-1}, \cdots, \hat{\Omega}_K^{-1}) = constant - \sum_{k=0}^{K} \frac{n}{2} ln|Z_k^T Z_k| \quad (7)$$

the MLE is then obtained by minimizing the determinant criterion,

$$\Phi(X) = \prod_{k=0}^{K} |Z_k^T Z_k| \quad (8)$$

3. NUMERICAL ALGORITHM

Newton-type method are well-established iterative techniques designed to solve general minimization problems as (8). Here, to overcome the difficulty of non-positive definite Hessian, a modified Newton method based on the modified Cholesky decomposition of Hessian [9] is utilized. A typical Newton-type algorithm consists of the basic step of calculating the gradient vector $g^{(k)}$ and Hessian matrix $H^{(k)}$ from first two order derivatives of model function in each iteration, and consequently this step needs much more computations. This paper gives an efficient evaluation of $g^{(k)}$ and $H^{(k)}$ using only first-order derivatives of the model through a simple transformation.

First, the determinant criterion is written as,

$$\Phi(X) = \prod_{k=0}^{K} \phi_k \quad (9)$$

where

$$\phi_k = |Z_k^T Z_k|$$

Take a QR decomposition of Z_k $(k = 0, \cdots, K)$,

$$Z_k = Q_k R_k \tag{10}$$

then following the results [7],

$$\frac{\partial \phi_k}{\partial x_i} = 2\phi_k \sum_{m=1}^{M} g_{k,i,mm} \tag{11}$$

$$\begin{aligned} \frac{\partial^2 \phi_k}{\partial x_i \partial x_j} &= 4\phi_k \sum_{m=1}^{M} g_{k,i,mm} \sum_{m=1}^{M} g_{k,j,mm} \\ &+ 2\phi_k \left[- \sum_{m=1}^{M} \sum_{n=1}^{M} g_{k,i,mn} g_{k,j,nm} \right. \\ &+ \left. \sum_{m=M+1}^{N_k} \sum_{n=1}^{M} g_{k,i,mn} g_{k,j,nm} \right] \end{aligned} \tag{12}$$

where $g_{k,i,mn}$ denotes the (m,n)th element of the $N \times M$ matrix,

$$G_{k,i} = Q_k^T \frac{\partial Z_k}{\partial x_i} R_k^{-1} \tag{13}$$

allows us to write,

$$\begin{aligned} g_i &= \frac{\partial \Phi(X)}{\partial x_i} \\ &= 2\Phi \sum_{k_1=0}^{K} \sum_{m=1}^{M} g_{k_1,i,mm} \end{aligned} \tag{14}$$

$$\begin{aligned} H_{ij} &= \frac{\partial^2 \Phi}{\partial x_i \partial x_j} \\ &= \Phi \sum_{k_1=0}^{K} \left[4 \sum_{m=1}^{M} g_{k_1,i,mm} \left(\sum_{k2=0}^{K} \sum_{m=1}^{M} g_{k_2,j,mm} \right) \right. \\ &+2 \left(- \sum_{m=1}^{M} \sum_{n=1}^{M} g_{k_1,i,mn} g_{k_1,j,nm} \right. \\ &+ \left. \left. \sum_{m=M+1}^{N_k} \sum_{n=1}^{M} g_{k_1,i,mn} g_{k_1,j,nm} \right) \right] \end{aligned} \tag{15}$$

Equation (15) permits very efficient evaluation of H, because once QR decomposition of Z_k $(k = 0, \cdots, K)$ is done and the matrices $G_{k,i}$ are formed, it is only necessary to collect a few inner products. Although Q_k^T occurs as a factor in (13), the matrix Q_k is not explicitly formed; instead, $Q_k^T \frac{\partial Z_k}{\partial x_i}$ can be formed by applying Householder transformations to $\frac{\partial Z_k}{\partial x_i}$ when taking QR decomposition of Z_k.

4. SIMULATION RESULTS

The performance of above approach is investigated using synthesized data from three different scenarios and it is compared with the traditional MLE method, i.e. Weighted Least Squares (WLS) method [5], which ignores the correlation of different types of noises and their inhomogeneity. Target travels in a northern direction at a constant speed of 6.0 kn from the initial location that bearing and range equal to 118.1^o and 17.0 nautical miles, respectively. The measurement period $T = 2$ minutes and at each instant one bearing and two frequencies (central frequencies are 300Hz, 600Hz respectively) is generated. The speed of sound in water are assumed to be constant $c = 3000kn$. 1) The first scenario is a one-leg problem, where the own ship move in a southern direction at a constant speed of 10kn and there is a total of 30 measurements at the end of the track. This problem is characterized by high noise level ($\sigma_\beta = 5.0^o$, $\sigma_f = 0.3Hz$), but the different types of measurement noises are independent. The simulation result is shown in Table I. The bias and deviation (variance) of MLE are slightly different from that of WLS. 2) The second scenario is same as first scenario, except that the noises of different measurements are correlated. The noise covariance matrix is,

$$\Omega = \begin{bmatrix} \sigma_\beta^2 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_f^2 & \sigma_2 \\ \sigma_1 & \sigma_2 & \sigma_f^2 \end{bmatrix}$$

where $\sigma_1 = \rho\sigma_\beta\sigma_f$, $\sigma_2 = \rho\sigma_f\sigma_f$, $\sigma_\beta = 5.0^o$, $\sigma_f = 0.3Hz$ and $\rho = 0.9$. Table II shows that both the bias and deviation of WLS remarkably exceed that of MLE. It is clear that ignorance of the noise correlation will degrade the performance of estimates. 3) In the third scenario, the ownship changes its course after 1 hour from south to east while keeping its speed fixed at 10kn. At the end of this two-leg scenario, there is a total of 60 measurements available, 30 in each leg. The covariance of noises is inhomogeneous, while in the first leg it is same as the second scenario and then changes its value in the second leg as $\sigma_\beta = 6.0^o$, $\sigma_f = 0.6Hz$ and $\rho = 0.5$. Observe in Table III that the deviations of WLS are approximately two times larger than those of the proposed MLE. Thus, the proposed MLE which considers the inhomogeneity of noise covariance improves the performance of estimation. Although two frequency bias of MLE exceed that of WLS, such bias become not very important as they are very small compared with the exact values of frequency and the deviations.

5. CONCLUSION

We have presented and solved the problem of TMA under more general noise assumption that the noises of different type measurements are correlated and covariance of noises is time-variant. The proposed MLE is efficient as it is simplified as an optimization problem to minimize a determinant criterion which contains no terms of unknown noise covariance matrix. To obtain such estimates, we introduced a Guass-Newton method which needs only the first order derivatives of model function to construct the gradient and Hessian, thus it needs less computations.

TABLE I
Simulation Results for the One-Leg Scenario

	Bias		Deviation	
	MLE	WLS	MLE	WLS
R (NMI)	-0.1623	-0.1580	1.9519	1.9249
β (°)	0.2142	0.2127	2.1447	2.0987
v (kn)	0.4651	0.4377	1.2405	1.1978
θ (°	2.2452	2.0649	23.4669	22.9276
f_1 (Hz)	0.0221	0.0197	0.2799	0.2714
f_2 (Hz)	0.0539	0.0493	0.5493	0.5313

TABLE II
Simulation Results for the One-Leg Scenario

	Bias		Deviation	
	MLE	WLS	MLE	WLS
R (NMI)	-0.0501	-0.0589	1.1911	2.0230
β (°)	0.0121	0.1953	1.6974	2.6358
v (kn)	0.3226	1.1178	0.7725	1.8218
θ (°)	0.1465	0.5428	20.1596	31.8039
f_1 (Hz)	-0.0004	-0.0049	0.2336	0.4103
f_2 (Hz)	0.0028	0.0138	0.4552	0.8167

TABLE III
Simulation Results for the Two-Leg Scenario

	Bias		Deviation	
	MLE	WLS	MLE	WLS
R (NMI)	-0.0112	-0.0262	0.4562	1.0781
β (°)	0.0811	0.0837	1.1461	2.0990
v (kn)	0.0030	0.08766	0.3241	0.5476
θ (°)	0.0772	0.3823	3.7597	11.5160
f_1 (Hz)	0.0011	-0.0002	0.0553	0.1244
f_2 (Hz)	0.0026	-0.0003	0.0843	0.2327

References

[1] Nardone, S. C., Lindgren, A. G., and Gong, K. F., "Fundamental properties and performance of conventional bearings-only target motion analysis", *IEEE Trans. on Automatic control*, Vol. 29, 775-787, Sept. 1984.

[2] Passerieux, J. M., Pillon, D., Blanc-Benon, P., and Jauffret, C., "Target motion analysis with bearings and frequencies measurements via instrumental variable estimator", In *Proc ICASSP'89*, Glasgow, Scotland, May 1989.

[3] Jauffret, C., and Bar-Shalom, Y., "Track formation with bearing and frequency measurements in clutter", *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 26, 999-1009, Nov. 1990.

[4] Chan, Y. T., and Rudnicki, S. W., "Bearings-only and doppler-bearing tracking using instrumental variables", *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 28, 1076-1082, Oct. 1992.

[5] De Vlieger, J. H., and Gmelig Meyling, R. H. J., "Maximum likelihood estimation for long-range target tracking using passive sonar measurements", *IEEE Trans. on Signal Processing*, Vol. 5, 1216-1224, May. 1992.

[6] Bard, Y. *Nonlinear Parameter Estimation.* New York: Academic Press, 1974.

[7] Bates, D. M., and Watts, D. G. *Nonlinear regression analysis and its applications.* New York: Wiley, 1988.

[8] Box, G. E. P., and Tiao, G. C. *Bayesian inference in statistical analysis.* Reading, Mass: Addison-Wesley, 1973.

[9] Gill, P. E., Murray, W., and Wright, M. H. *Practical optimization.* London: Academic Press, 1981.