

LEAST SQUARES ESTIMATION OF POLYNOMIAL PHASE SIGNALS VIA STOCHASTIC TREE-SEARCH

Dawei Huang, Simon Sando, Lian Wen

CiSSaIM, Queensland University of Technology, Brisbane, Australia, 4001.

ABSTRACT

Estimating the parameters for a constant amplitude, polynomial-phase signal with additive Gaussian noise is considered. The difficulty in this problem is that there are many unobserved integers when a linear regression model is used for wrapped phases [1]. Analysing the least squares target function based on the regression model, we use the differencing approach [3] to simplify it. Thus a tree-search algorithm can be used to find the solution of the least squares problem. To reduce the computational complexity, statistical inference methods are applied. Then an attractive recursive algorithm is derived. Simulation results show that this algorithm works at a lower SNR than that for existing methods.

1. INTRODUCTION

Estimating the coefficients of a polynomial-phase signal in the presence of noise has arisen from many applications in signal processing [1-6]. Assume that we have observations

$$z_t = Ae^{i\sum_{j=0}^p \theta_j t^j} + n_t, \quad t = 0, 1, 2, \dots, T-1,$$

where A is a constant, $\Theta = [\theta_0, \dots, \theta_p]$ is the unknown parameter vector to be estimated, and $\{n_t\}$ is a complex white normal sequence with mean zero and variance $2\sigma_n^2$. For estimating Θ uniquely, we have to place some constraints. Since $e^{i\theta}$ is a periodic function in 2π , and since for any integers k_j here exist rationals α_j 's such that

$$\frac{2k_m\pi}{m!}t^m + 2\pi \sum_{j=0}^{m-1} \alpha_j t^j = 0 \pmod{2\pi}. \quad (1)$$

For example, $\pi t + \pi t^2 = t(t+1)\pi = 0 \pmod{2\pi}$. Then $[\theta_p, \dots, \theta_m + \frac{2m\pi}{m!}, \theta_{m-1} + \alpha_{m-1}, \dots, \theta_0 + \alpha_0]$ corresponds to the same observations as Θ_p . Thus, we ask

$$-\frac{\pi}{j!} \leq \theta_j < \frac{\pi}{j!}, \quad j = 0, 1, \dots \quad (2)$$

This problem has received considerable attention. A direct approach is applying the linear regression technique to unwrapped signals [1, 2]. However, the existing unwrapping methods may not work in the presence of noise. To avoid unwrapping, differencing in phase via the multiplication of suitable consecutive observations has been suggested [2-5]. Unfortunately, multiplication introduces a new noise term to the estimation problem (see (4) in [6]). This term can only be ignored when the SNR is high. The theoretical analysis as well as the simulation in [6] show that this approach may be efficient (attain the Cramer-Rao lower bound) at SNR higher than 20db only.

The difficulty of this problem is that there are many unobserved variables. Based on the observations, we can only obtain the wrapped phase of z_t :

$$y_t = \begin{cases} \arctan \frac{Im(z_t)}{Re(z_t)}, & Re(z_t) \geq 0, \\ \arctan \frac{Im(z_t)}{Re(z_t)} + \pi, & Re(z_t) < 0, Im(z_t) > 0, \\ \arctan \frac{Im(z_t)}{Re(z_t)} - \pi, & Re(z_t) < 0, Im(z_t) < 0 \end{cases} \quad (3)$$

which is distributed on $[-\pi, \pi]$. The unwrapped phase of z_t is related to y_t through an unobserved integer process x_t :

$$y_t + 2x_t\pi = \angle z_t = \sum_{j=0}^p \theta_j t^j + u_t, \quad t = 0, 1, \dots, T-1, \quad (4)$$

where u_t is the perturbation. According to [1], $\{u_t\}$ may be assumed to be a white normal sequence with mean zero and variance $\frac{\sigma_n^2}{\lambda^2}$ when the SNR is not too low. Thus, if we know $\{x_t, t = 0, 1, \dots, T-1\}$, standard linear regression analysis can be used for estimating the parameters θ_j 's and the variances of the estimators attain the Cramer-Rao lower bound [1].

In this paper, we combine both regression and differencing techniques with the least squares criterion to estimate $\{x_t, t = 0, 1, \dots, T-1\}$. To determine the $\{x_t\}$ that maximises the likelihood function among all possible integer sequences, we develop a stochastic tree-search method for this problem. Using Levinson recursion, we show that the levels in the tree needed to

decide an x_t can be fixed and do not depend on T . This avoids the increasing of computational complexity as T increases. Furthermore, we decide the branches (nodes) and leaves (ends) of the tree by statistical inference. This reduces the size of the tree, and hence the computational complexity dramatically. Finally, we develop a very simple recursive algorithm to calculate the value of the target function at each node and leaf.

In Section 2, we derive a target function in terms of only $\{x_t, t = 0, 1, 2, \dots, T-1\}$ from the least squares criterion. To analyze this target function, a differencing technique is applied. This target function can then be rewritten as a sum of residual squares, where the k^{th} residual depends on $\{x_t, t = 0, 1, 2, \dots, k\}$ only. In Section 3, we develop a finite level searching tree based on statistical inference. Section 4 summarizes the algorithm and shows some simulation results.

2. TARGET FUNCTION

To solve the regression problem in (4), let $X_T = [x_0, x_1, \dots, x_{T-1}]'$ and $Y_T = [y_0, y_1, \dots, y_{T-1}]'$. We need to consider the least squares problem

$$\begin{aligned} & \min_{x_t's} \min_{\theta_j's} \sum_{t=0}^{T-1} \left(y_t + 2x_t\pi - \sum_{j=0}^p \theta_j t^j \right)^2 \\ & = \min_{X_T} \min_{\Theta} \|Y_T + 2\pi X_T - H_T \Theta\|^2, \end{aligned} \quad (5)$$

$$\text{where } H_t = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t-1 & \dots & (t-1)^p \end{bmatrix}.$$

Under the normality of $\{u_t\}$ [1], this approach is equivalent to the maximum likelihood approach. The double minimization in (5) can be performed in two steps. First, for a given X_T , the Θ yielding the minima is given by

$$\hat{\Theta} = (H_T' H_T)^{-1} H_T' (Y_T + 2\pi X_T), \quad (6)$$

so we only need to choose X_T to minimize

$$= \left\| \left[I - H_T (H_T' H_T)^{-1} H_T' \right] (Y_T + 2\pi X_T) \right\|^2.$$

Notice that $P_T = I - H_T (H_T' H_T)^{-1} H_T'$ is the projection operator onto the linear space S_T that is orthogonal to the linear space spanned by the column vectors in H_T . To investigate the matrix P_T , we need to find a basis for S_T . The differencing idea in [4, 5, 6] can be

used for this purpose, but with a higher order. Let

$$d_k = \left[\underbrace{0, \dots, 0}_k, \underbrace{1, -\binom{p+1}{1}, \binom{p+1}{2}, \dots, (-1)^{p+1}}_{p+2}, \underbrace{0, \dots, 0}_{T-k-p-2} \right]'$$

where $k = 0, 1, 2, \dots, T-p-2$. When $p = 2$, we have

$$d_k = \left[\underbrace{0, \dots, 0}_k, \underbrace{1, -3, 3, -1, 0, \dots, 0}_{T-k-4} \right]', \quad k = 0, 1, \dots, T-4.$$

Since they consist of the coefficients of the $(p+1)^{th}$ differencing operator $\nabla^{p+1} = (1-B)^{p+1}$, where B is the backshift operator, one can confirm that

$$d_k' H_T = 0, \quad k = 0, 1, \dots, T-p-2,$$

and so $\{d_0, d_1, \dots, d_{T-p-2}\}$ is a basis of the $(T-p-1)$ -dimensional Euclidean subspace S_T .

Based on $\{d_0, \dots, d_k\}$, we can find a standard orthogonal basis of S_T :

$$e_k = \frac{d_k - \sum_{j=1}^{k \wedge (p+1)} (e_{k-j}' d_k) e_{k-j}}{\left\| d_k - \sum_{j=1}^{k \wedge (p+1)} (e_{k-j}' d_k) e_{k-j} \right\|}, \quad (7)$$

where $k = 0, 1, \dots, T-p-2$. and $a \wedge b \equiv \min(a, b)$. We can then show that

$$P_T = \sum_{k=0}^{T-p-2} e_k e_k',$$

and hence we choose X_T to minimize

$$\left\| \sum_{k=0}^{T-p-2} e_k e_k' (Y_T + 2\pi X_T) \right\|^2 = \sum_{k=0}^{T-p-2} [e_k' (Y_T + 2\pi X_T)]^2. \quad (8)$$

3. STOCHASTIC TREE-SEARCH

Now we consider how to find the minima of (8) recursively. Let $\sigma_k = \frac{\sigma_T}{A} \left\| d_k - \sum_{j=1}^{k \wedge (p+1)} (e_{k-j}' d_k) e_{k-j} \right\|$ and

$$\begin{aligned} w_k &= e_k' (Y_T + 2\pi X_T) \\ &= \frac{d_k' (Y_T + 2\pi X_T) - \sum_{j=1}^{k \wedge (p+1)} (e_{k-j}' d_k) w_{k-j}}{\sigma_k}. \end{aligned} \quad (9)$$

From (7) it is clear that w_k only depends on Y_{k+p+2} and X_{k+p+2} . Also, suppose that X_{k+p+2} satisfies (4),

and let $U_T = [u_0, \dots, u_{T-1}]'$. Since d_k is orthogonal to H_T , we have

$$w_k = \frac{d_k'(H_T\Theta + U_T) - \sum_{j=1}^{k \wedge (p+1)} (e'_{k-j} d_k) w_{k-j}}{\sigma_k} \\ = \frac{d_k' U_T - \sum_{j=1}^{k \wedge (p+1)} (e'_{k-j} d_k) w_{k-j}}{\sigma_k},$$

and so $\{w_k\}$ is a white standard normal sequence. If we know $X_{k+p+1} = [x_0, \dots, x_{k+p}]'$, with probability 0.999, x_{k+p+1} should be one of the integers to satisfy $|w_k| < 3.2905$. Thus, let

$$c_k = \left(\sum_{j=0}^p (-1)^j \binom{p+1}{j} (y_{k+j} + 2\pi x_{k+j}) \right. \\ \left. + (-1)^{p+1} y_{k+p+1} - \sum_{j=1}^{k \wedge (p+1)} (e'_{k-j} d_k) w_{k-j} \right) / \sigma_k.$$

According to (9), we have

$$\frac{(c_k - 3.2905) \sigma_k}{2\pi} \leq (-1)^p x_{k+p+1} \leq \frac{(c_k + 3.2905) \sigma_k}{2\pi}. \quad (10)$$

If there are n_k integers satisfying (9), then corresponding to a given vector X_{t+p+1} is a tree branch

$$x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_{k+p} \rightarrow \begin{cases} x_{k+p+1}^{(1)} \\ \vdots \\ x_{k+p+1}^{(n_k)} \end{cases}.$$

In practice, we do not know X_{k+p+1} , and so we have to do the tree-search step by step.

Firstly, we consider the initial values $\{x_0, x_1, \dots, x_p\}$. From (2) and the consideration in the end of section 4, we take $x_k = 0$, $k = 0, 1, \dots, p$. Then, we choose $x_{p+1}^{(j_1)}$ according to (10). For each $X_{p+2}^{(j_1)} = [0, \dots, 0, x_{p+1}^{(j_1)}]'$, we choose $x_{p+2}^{(j_1, j_2)}$, so on. Thus, we establish a searching tree and calculate the residual $w_{p+t}^{(j_1, \dots, j_t)}$ at each leaf $x_{p+t}^{(j_1, \dots, j_t)}$.

Secondly, we consider how to end the tree-search at a certain depth of the tree. If we continue the above procedure for all samples, the computational complexity may be too high, with the number of leaves in the tree possibly increasing exponentially. Fortunately, the contribution of $x_{p+t}^{(j_1, \dots, j_t)}$ to the residual $w_{p+t+N}^{(j_1, \dots, j_t)}$ decreases to zero quickly as N increases.

Let $D_t = [d_0, d_1, \dots, d_t]$, $\Gamma_t = D_t' D_t$, and $\Phi_t = [\phi_{t,t}, \phi_{t,t-1}, \dots, \phi_{t,1}, 1]'$. Φ_t satisfies $\Gamma_t \Phi_t = [0, \dots, 0, s_t^2]'$,

where $s_t^2 = \left\| d_k - \sum_{j=1}^{k \wedge (p+1)} (e'_{k-j} d_k) e_{k-j} \right\|^2$, and thus

$$D_t' \left(\frac{1}{s_t} D_t \Phi_t \right) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ s_t \end{bmatrix}, \quad \left(\frac{1}{s_t} D_t \Phi_t \right)' \left(\frac{1}{s_t} D_t \Phi_t \right) = 1.$$

ie. $\frac{1}{s_t} D_t \Phi_t$ is a linear combination of d_k , $k = 0, 1, \dots, t$; which is orthogonal to all d_k , $k = 0, 1, \dots, t-1$ and has norm one. Therefore,

$$e_t = \frac{1}{s_t} D_t \Phi_t. \quad (11)$$

The elements in e_t are the prediction coefficients of the $MA(p+1)$ series $u_t = \sum_{j=0}^{t \wedge (p+1)} (-1)^j \binom{p+1}{j} \varepsilon_{t-j}$, where $\varepsilon_t = u_t - Proj_{\{u_k, 0 \leq k < t\}} u_t$. Since this $MA(p+1)$ is a purely non-deterministic series, we conclude that

$$\phi_{t+N, k} \rightarrow 0 \text{ for } k = t+N, \dots, N, \text{ and} \\ s_{t+N}^2 \rightarrow 1 \text{ as } N \rightarrow \infty. \quad (12)$$

Also, since $w_{t+N}^{(j_1, \dots, j_t, j_{t+1}, \dots, j_{t+N})} = e'_{t+N} (Y_T + 2\pi X_T)$ and (11,12), we can choose $x_{p+t}^{(j_1, \dots, j_t)}$ by minimizing

$$L \left(X_{p+t+N}^{(j_1, \dots, j_t)} \right) = \sum_{k=t+1}^{t+N} w_k^2 \left(X_{k+p}^{(j_1, \dots, j_k)} \right), \quad (13)$$

ie, we only need a fixed depth tree-search, say N , for deciding each x_{p+t} .

Thirdly, to further reduce the computational complexity, we may end certain branches before achieving the depth N . Indeed, if $\{x_t\}$ satisfies (4), $\sum_{k=m+1}^{m+t} w_k^2$ should be a χ_t^2 random variable, so if the value is not significant at a prespecified level, say 99.99%, we can delete $X_{p+t}^{(j_1, \dots, j_t)}$ from the candidate list and not proceed along this branch of the tree.

4. ALGORITHM

Let $\{s_k\}$ be defined as in section 3, and

$$r_s = (-1)^s \sum_{j=s}^{p+1-s} \binom{p+1}{j} \binom{p+1}{j-2}. \quad (14)$$

Put $s_0^2 = r_0$ and $a_{k,t} = 0$ for all $k < 1$ or $t > k \wedge (p+1)$. We calculate

$$a_{k,t} = \frac{r_t' - \sum_{j=1}^{k \wedge (p+1-t)} a_{k-t, j} a_{k, t+j}}{s_{k-t}} \\ s_k^2 = r_0 - \sum_{j=1}^{k \wedge (p+1-t)} a_{k, j}^2. \quad (15)$$

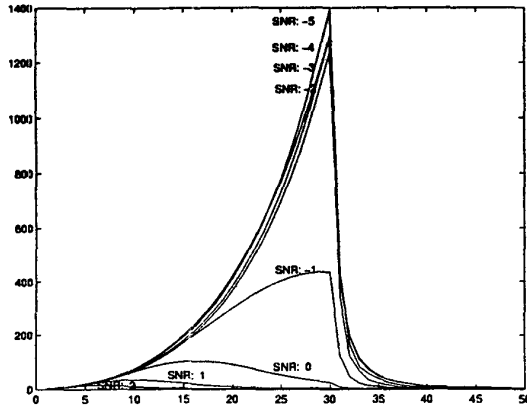


Figure 1: Computational overhead. Average number of leaves per search vs time window. The window length is 30, with an overlap of 29.

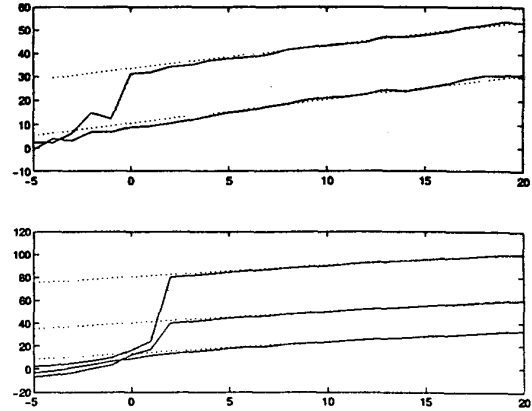


Figure 2: Performance of Estimators. Inverse of MSE vs SNR in dB for (a): linear phase with $T=24$ and 100 repetitions. (b): quadratic phase with $T=100$ and 1000 repetitions. Search depths are 20 and 40 respectively.

where $t = k \wedge (p+1), \dots, 1$. We can then show that

$$c_k = \frac{A}{\sigma_n s_k} \left(\sum_{j=0}^p (-1)^j \binom{p+1}{j} (y_{k+j} + 2\pi x_{k+j}) + (-1)^{p+1} y_{k+p+1} - \sum_{j=1}^{k \wedge (p+1)} (e'_{k-j} d_k) w_{k-j} \right).$$

Starting from the initial values $x_0 = \dots = x_p = 0$, we choose $x_{p+1}^{(j_1)}, x_{p+2}^{(j_1, j_2)}, \dots, x_{p+N}^{(j_1, \dots, j_N)}$ satisfying (10). The corresponding residuals can be obtained by

$$w_k^{(j_1, \dots, j_k)} = c_k^{(j_1, \dots, j_k)} + \frac{(-1)^{p+1} A x_{k+p+1}^{(j_1, \dots, j_k)}}{\sigma_n s_k} \quad (16)$$

Let X_{p+N}^* be the minima of (13), where \hat{x}_{p+1} is the $(p+1)^{th}$ component of X_{p+N}^* . Starting from $x_1 = \dots = x_p = 0$ and $x_{p+1} = \hat{x}_{p+1}$, we repeat the above procedure for estimating \hat{x}_{p+2} and so on.

Once we obtain the estimator \hat{X}_T , Θ can be obtained by (6). However, we must satisfy the constraints in (2). If the estimators are out of the range, we correct them by subtracting $\frac{2k_j \pi}{j!}$ and $2\alpha_k \pi$ respectively.

Simulations have been done and part results are shown. The computational complexity is shown in figure 1, while the mean square error (MSE) is compared to the Cramer-Rao bounds in figure 2. Compared with the results in [3, 6], we can see significant improvements in the threshold for attaining the Cramer-Rao bound.

5. ACKNOWLEDGMENTS

The authors would like to thank Dr. A. M. Zoubir, who motivated this research. This research was supported by a Large ARC grant.

6. REFERENCES

- [1] Tretter S. A. "Estimating the Frequency of a Noisy Sinusoid by Linear Regression", *IEEE Transactions on Information Theory*, 31:832-835, 1985.
- [2] Djuric P.M. and Kay S.M. "Parameter Estimation of Chirp Signals", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38:2118-2126, 1990.
- [3] Kay S.M. "A Fast and Accurate Single Frequency Estimator", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37:1987-1990, 1989.
- [4] Peleg S. and Porat B. "Estimation and Classification of Polynomial-Phase Signals", *IEEE Transactions on Information Theory*, 37:422-428, 1991.
- [5] Peleg S. and Friedlander B. "The Discrete Polynomial-Phase Transform", *IEEE Transactions on Signal Processing*, 43:1901-1914, 1995.
- [6] Clarkson, V. Kootsookos P.J. and Quinn B.G. "Analysis of the Variance Threshold of Kay's Weighted Linear Predictor Frequency Estimator", *IEEE Transactions Signal Processing*, 42:2370-2379, 1994.