ESTIMATING THE PERIOD OF A PULSE TRAIN FROM A SET OF SPARSE, NOISY MEASUREMENTS

I. Vaughan L. Clarkson* Stephen D. Howard* Iven. M. Y. Mareels†

* Electronic Warfare Division, DSTO Salisbury, P.O. Box 1500, Salisbury 5108, Australia
† Department of Engineering, FEIT, The Australian National University, ACT 0200, Australia

ABSTRACT

The problem of estimation of the period of a pulse train from a set of sparse, noisy measurements of the times-of-arrival, and the association of pulse indices with the measurements, is considered for a pair of statistical models of the measurement process. We find that the estimation and association problem can be formulated as a simultaneous Diophantine approximation problem. We propose an algorithm for obtaining estimates and associations based on the LLL algorithm [5]. We also show a relationship with the maximisation of a certain trigonometric sum, which can be regarded as the periodogram of the measurements. We present some numerical results which indicate that the algorithm is able to make correct associations of pulse indices, and therefore accurate estimates of the period, with very high probability even for very sparse, short and noisy records. This is demonstrated in examples in which 99.9% of pulses are missing and as few as 9 pulses are recorded with average errors of up to 1% of the period on each measurement and yet the experimental frequency of correct association was more than 99%.

1. INTRODUCTION

Periodic pulse trains are a common feature of many physical systems. In this paper, we consider a situation in which a single periodic pulse train is observed and the *times-of-arrival* (TOA's) of pulses are measured. However, the pulses which are observed are not consecutive. It is assumed that some (and perhaps many) of the pulses were not observed. Additionally, we assume that TOA's are not measured perfectly, but are subject to random errors. The problem is to estimate the period of the pulse train from the data recorded and associate each of the measured TOA's with a pulse number or index, relative to the first observed, which takes account of the intervening missing pulses.

The motivating problem in this instance is that of passive radar surveillance. Typically, a radar emits a train of pulses in a periodic sequence. In the simplest and most frequently encountered case, the pulse train is purely periodic (the sequence has a length of one). Receivers for passive radar surveillance can often make use of the period, known as the *pulse repetition interval* (PRI), to identify the emitter and its mode of operation. However, for many conceivable reasons, it may not be possible to measure the TOA of each consecutive pulse. Pulses will be missing from the record. For certain receiver types, for example a scanning superheterodyne receiver, the record of pulses from a given pulse train may be extremely sparse. Additionally, measurement of the TOA will be subject to a variety of errors, such as thresholding effects caused by thermal noise or variability in received power or simply poor time resolution.

Very little has been published regarding the problem of estimating the period of a pulse train from sparse, noisy measurements. Indeed, the paper of Casey & Sadler [1] is the only widely published work of which the authors are aware. In that paper, a number of generalised Euclidean algorithms are proposed to recover the period. They demonstrated that, even for a very sparse record in which 98% of pulses are missing, the period can be reliably estimated. However, we have identified a number of areas in which their results can be improved. Firstly, we formulate statistical models for the measurement process. We make use of a fairly standard generalised Euclidean algorithm, derived from the LLL algorithm of Lenstra, Lenstra & Lovász [5] for lattice reduction, and explain its relationship to a method of maximum likelihood for estimation and association. Furthermore, the algorithm we propose is capable of reliably estimating the period and associating pulse indices for extremely sparse, noisy and short records.

2. SIGNAL MODEL

We will consider two signal models: a *simple* model and an *extended* model.

Consider a purely periodic pulse train with pulse repetition interval T and phase θ. Under this model, pulses are emitted at the times $iT + \theta$, $i \in \mathbb{Z}$. We assume that our record consists of n observations of pulses which are corrupted by noise and which are not necessarily consecutive. Thus, our observations are of random variables, which we shall denote $Z_1, Z_2, \ldots, Z_n$, such that

$$Z_i = s_i T + \theta + X_i$$

where the $s_i \in \mathbb{Z}$ are the indices of the observed pulses and the X_i are independent, identically distributed (i.i.d.) normal random variables representing the observation errors (noise) with zero mean and variance σ^2. We assume that the only *admissible* pulse indices are those which satisfy

$$0 = s_1 < s_2 < \cdots < s_n. \tag{1}$$

We will call this model the *simple* model.

We will find that, in general, no maximum likelihood estimate of the parameters exists in the simple model. We

The authors wish to acknowledge the funding of the activities of the Cooperative Research Centre for Robust and Adaptive Systems by the Australian Commonwealth Government under the Cooperative Research Centres Program.

require more prior information to obtain such an estimate. For this reason, we propose an *extended* model, in which we will also assume that the indices of the observed pulses are random variables. We assume that our TOA observations $Z_1, Z_2, \ldots, Z_n$ now have the form

$$Z_i = S_i T + \theta + X_i$$

where $S_1 = 0$ (a degenerate random variable),

$$S_{i+1} - S_i = Y_i + 1$$

for $i > 1$ and the Y_i are i.i.d. random variables from a geometric distribution with parameter λ. Furthermore, the X_i and the Y_i are assumed to be mutually independent. The parameter λ can be interpreted as the probability of a given pulse being observed.

3. PARAMETER ESTIMATION AND ASSOCIATION

Consider the method of maximum likelihood for estimation of the parameters T, θ and $\mathbf{s}$ in the simple model. The joint probability density function (p.d.f.) of the observations is

$$f\big(\mathbf{z};\ T,\theta,\sigma^2,\mathbf{s}\big) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\!\left(-\frac{\|\mathbf{z}-T\mathbf{s}-\theta\mathbf{1}_n\|^2}{2\sigma^2}\right) \tag{2}$$

where $\mathbf{z}$ is a column vector representing the possible values of the Z_i and $\mathbf{s}$ is a column vector representing the indices whose elements are assumed to be admissible according to (1). $\|\cdot\|$ denotes the Euclidean norm and $\mathbf{1}_n$ is the column vector of dimension n whose elements are all one.

If the vector of indices, $\mathbf{s}$, is known *a priori* then the problem is simply one of linear regression to estimate T and θ. If we set

$$\mathbf{Q} = \mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T \tag{3}$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix and observe that $\mathbf{Q}$ is a symmetric projection matrix then it can be shown that the estimates $\hat{\theta}$ and $\hat{T}$ can be expressed as

$$\hat{\theta} = \frac{1}{n}(\mathbf{z}-T\mathbf{s})^T\mathbf{1}_n \qquad \text{and} \qquad \hat{T} = \frac{\mathbf{z}^T\mathbf{Q}\mathbf{s}}{\mathbf{s}^T\mathbf{Q}\mathbf{s}}. \tag{4}$$

With our likelihood function being the joint p.d.f. in (2) and with $\mathbf{z}$ set to the observed values of the Z_i, we find that the likelihood function is maximised with respect to T and θ at $\hat{T}$ and $\hat{\theta}$ respectively. Thus, taking the logarithm of the likelihood function and discarding constants, we find that the likelihood function is maximised with respect to $\mathbf{s}$, T and θ when

$$F(\mathbf{s}) = \mathbf{z}^T\mathbf{Q}\mathbf{z} - \frac{\left(\mathbf{z}^T\mathbf{Q}\mathbf{s}\right)^2}{\mathbf{s}^T\mathbf{Q}\mathbf{s}} \tag{5}$$

is minimised. Clearly, $F(\mathbf{s}) \geqslant 0$. We also note from the form of (5) that maximisation of the likelihood function is equivalent to minimisation of $\sin^2\phi$ over all admissible $\mathbf{s}$, where ϕ is the angle between the vectors $\mathbf{Qz}$ and $\mathbf{Qs}$.

We will show in Section 4 that a unique maximum likelihood estimate does not exist in general, so we now consider what can be done using the extended model. In this model, we find that the joint p.d.f. is

$$g\big(\mathbf{z},\mathbf{s};\ T,\theta,\sigma^2,\lambda\big) = \lambda^{n-1}(1-\lambda)^{s_n-n+1} f\big(\mathbf{z};\ T,\theta,\sigma^2,\mathbf{s}\big) \tag{6}$$

when $\mathbf{s} \in \mathbb{Z}^n$ (now a column vector representing the possible values of the S_i) is admissible according to (1) and the joint p.d.f. is 0 otherwise. If we were interested only in finding maximum likelihood estimates for T and θ then we should maximise the likelihood function obtained by marginalising the p.d.f. $g(\cdot)$ over all admissible index vectors. However, suppose we want to simultaneously associate the observations with a set of pulse indices $\hat{\mathbf{s}}$. Then, in order to make our observations maximally likely, we should simply maximise $g(\cdot)$ over $\mathbf{s}$, T and θ. We call this *joint maximum likelihood estimation and association* (JMLEA).

For any postulated association $\mathbf{s}$, the maximum likelihood estimates for T and θ remain as they were in (4). So, by taking the logarithm of g and discarding constants, we find that the JMLEA is obtained with respect to $\mathbf{s}$, T and θ when the function

$$G(\mathbf{s}) = F(\mathbf{s}) + \kappa s_n \tag{7}$$

is minimised where

$$\kappa = -2\sigma^2\log(1-\lambda) > 0. \tag{8}$$

Clearly, $\kappa \approx 2\sigma^2\lambda$ when λ is small.

We will show that a JMLEA exists for the extended model. In Section 6 we demonstrate through numerical simulations that its statistical performance is very satisfactory, even under quite severe conditions.

4. FORMULATION AS A SIMULTANEOUS DIOPHANTINE APPROXIMATION PROBLEM

First, we define a point lattice. If $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N\}$ is a set linearly independent vectors in a vector space $\mathcal{E}$ then the set of points

$$\Omega = \sum_{i=1}^{N} \mathbf{b}_i\mathbb{Z}$$

is a *point lattice* (or simply lattice) of degree N and $\mathcal{B}$ is a basis of the lattice. We note that the basis of a lattice is not unique. The span of Ω is the subspace of $\mathcal{E}$ spanned by any basis of Ω.

Now, consider the simple model in which we hope to find maximum likelihood estimates for T and θ. Let $\mathbf{q}_i$ denote the i^{th} column of $\mathbf{Q}$ as defined in (3). Let $\{\mathbf{q}_2, \mathbf{q}_3, \ldots, \mathbf{q}_n\}$ be a basis of a lattice Ω in $\mathbb{R}^n$. Thus, any point $\mathbf{v}$ in the lattice Ω can be expressed $\mathbf{v} = \mathbf{Qs}$, where $\mathbf{s} \in \mathbb{Z}^n$ with $s_1 = 0$. We define a semi-norm and a norm,,

$$\|\mathbf{x}\|_1 = \|\mathbf{Mx}\| \qquad \text{and} \qquad \|\mathbf{x}\|_2 = \|\mathbf{x}\|, \tag{9}$$

where $\mathbf{M} = \mathbf{I}_n - \boldsymbol{\zeta}\boldsymbol{\zeta}^T$, $\boldsymbol{\zeta} = \mathbf{Qz}/\|\mathbf{Qz}\|$ and $\mathbf{z}$ is the vector of TOA measurements. We observe that $\|\cdot\|_1$ is the length of its argument when projected onto the plane orthogonal to $\mathbf{Qz}$.

It is not hard to verify that we can now rewrite $F(\cdot)$ from (5) as

$$F(\mathbf{s}) = \|\mathbf{Qz}\|_2^2 \frac{\|\mathbf{Qs}\|_1^2}{\|\mathbf{Qs}\|_2^2}.$$

Since $\mathbf{Qs} \in \Omega$ for any admissible $\mathbf{s}$, minimisation of $F(\cdot)$ is equivalent to minimisation of

$$F^*(\mathbf{v}) = \frac{\|\mathbf{v}\|_1}{\|\mathbf{v}\|_2}$$

over all admissible $\mathbf{v} \in \Omega$. The problem, posed in this way, is a *simultaneous Diophantine approximation* problem.

From the theory of simultaneous Diophantine approximation, it is then possible to show [2] that either there exists some non-zero lattice point $\mathbf{v}$ such that $F^*(\mathbf{v}) = 0$ or there exists a non-terminating sequence of lattice points $\mathbf{v}_j$ such that

$$\lim_{j\to\infty} F^*(\mathbf{v}_j) = 0 \qquad \text{and} \qquad \lim_{j\to\infty} \|\mathbf{v}_j\| = \infty. \tag{10}$$

Since $F^*(\mathbf{v}) = 0$ implies $F^*(k\mathbf{s}) = 0$ for all $k \in \mathbb{Z}$, we can now see that either there is no non-zero lattice points which minimises $F^*(\cdot)$ or an infinitude. Furthermore, it is possible to show that the probability of an infinitude is zero. We have neglected to consider the necessity that the lattice points be admissible to our model but it can be shown that this is assured if the observations are *time-ordered*, which is to say they satisfy $z_1 < z_2 < \cdots < z_n$.

Therefore, consider the extended model in which we impose a "cost" on the number of pulses missing from the record. Here, we set out to jointly estimate T and θ and associate the observation with a set of pulse indices $\mathbf{s}$ so as to maximise the likelihood of our observation (the JMLEA). As we described in Section 3, this involves the minimisation of $G(\mathbf{s})$ in (7) over all admissible $\mathbf{s}$.

We can show that a JMLEA must exist and there are at most a finite number of them. This is because the number of admissible index vectors $\mathbf{s}$ with $G(\mathbf{s}) \leqslant \nu$ for any $\nu \in \mathbb{R}$ must be finite.

Furthermore, under appropriate conditions, such as σ being sufficiently small with respect to T and n (see [2]), it can be shown that the JMLEA must be one of the *angular best approximations* from Ω with respect to $\|\cdot\|_1$ and $\|\cdot\|_2$. By this we mean that if $\mathbf{v}^* \in \Omega$ corresponds to a JMLEA then, for all other lattice point $\mathbf{v} \in \Omega$ with $\|\mathbf{v}\|_2 > 0$,

$$\begin{aligned} &\|\mathbf{v}\|_2 \leqslant \|\mathbf{v}^*\|_2 \;\Rightarrow\; F^*(\mathbf{v}) \geqslant F^*(\mathbf{v}^*) \\ \text{and} \qquad &F^*(\mathbf{v}) \leqslant F^*(\mathbf{v}^*) \;\Rightarrow\; \|\mathbf{v}\|_2 \geqslant \|\mathbf{v}^*\|_2 \end{aligned}$$

There is, with probability 1, a non-terminating sequence of best approximations which can be numbered $\mathbf{v}_j$ such that (10) is satisfied.

Now, for a related problem in best simultaneous Diophantine approximation [4] we know that the growth of successive best approximations is exponential. For this reason, an algorithm for find best approximations should converge rapidly to the JMLEA. Unfortunately, algorithms which find each and every best approximation are probably NP-hard [3].

We use an adaptation of the lattice reduction algorithm of Lenstra, Lenstra and Lovász [5], known as the LLL algorithm, for simultaneous Diophantine approximation. Details of the exact modifications of this algorithm for the period estimation problem are given in [2], but we can briefly describe the algorithm by saying that we iteratively apply the LLL algorithm to reduce the lattice Ω with respect to a norm which is a weighted sum of $\|\cdot\|_1$ and $\|\cdot\|_2$. On each iteration, we give successively more weight to $\|\cdot\|_1$, which tends to produce lattice points which are smaller and smaller with respect to $\|\cdot\|_1$ but larger and larger with respect to $\|\cdot\|_2$. It can be shown that the lattice points obtained in this way are good approximations. At each iteration, the function $G(\cdot)$ is evaluated for the index vectors associated with each new lattice point generated, and a running minimum is maintained. The algorithm terminates when each new lattice point corresponds to an index vector for which the right-hand term alone of $G(\cdot)$ in (7) is larger than the value of $G(\cdot)$ at the running minimum. The running minimum is then used to generate the estimates and the pulse index association. Usually, we will have found the JMLEA, but this cannot be guaranteed because the LLL reduction technique is known to miss some best approximations.

5. A RELATED TRIGONOMETRIC SUM

In this section, we point out the relationship between the simultaneous Diophantine approximation problem discussed in Section 4 and the maximisation (in magnitude) of the trigonometric sum

$$A(\omega) = \sum_{i=1}^{n} e^{-j2\pi z_i \omega}. \tag{11}$$

It is natural to wonder if a relationship of some kind exists, because the magnitude of (11) can be thought of as the periodogram of the function consisting of a train of impulses (Dirac delta functions) at the measured TOA's. A good candidate for the PRI of the observed pulse train should occur at the inverse of a frequency which maximises $|A(\omega)|$.

Now, given a vector of pulse indices, $\mathbf{s}$, we define the function

$$\omega(\mathbf{s}) = \frac{2\pi}{\hat{T}(\mathbf{s})} = 2\pi \frac{\mathbf{s}^T\mathbf{Qs}}{\mathbf{z}^T\mathbf{Qz}}.$$

Setting $\epsilon_i = \omega(\mathbf{s})(z_i - \overline{z}) - 2\pi(s_i - \overline{s})$, where $\overline{z}$ and $\overline{s}$ denote the arithmetic means of the z_i and s_i respectively, we note that

$$\sum_{i=1}^{n} \epsilon_i = 0 \qquad \text{and} \qquad \sum_{i=1}^{n} \epsilon_i^2 = \big(2\pi \|\mathbf{Qs}\|_1\big)^2.$$

We can employ these identities to find that

$$\begin{aligned} |A(\omega(\mathbf{s}))| &= \Big[\Big(\sum_{i=1}^{n} \cos\epsilon_i\Big)^2 + \Big(\sum_{i=1}^{n} \sin\epsilon_i\Big)^2\Big]^{1/2} \\ &= n - 2\pi^2 \|\mathbf{Qs}\|_1^2 + O\big(\|\mathbf{Qs}\|_1^4\big). \end{aligned}$$

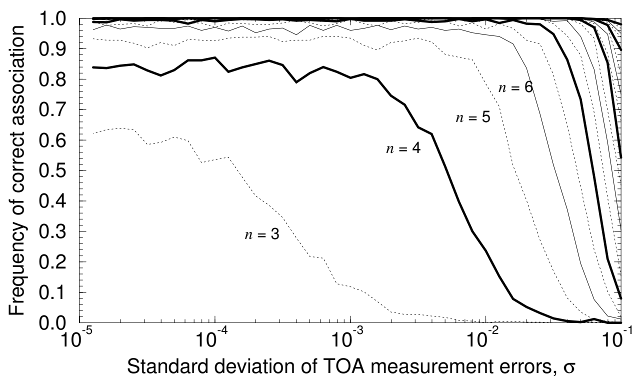

Figure 1: The experimental frequency of correct association of the pulse indices as a function of the standard deviation of the measurement errors, σ, for various numbers of pulses, n, with $T = 1$ and $\lambda = 10^{-3}$.

Thus we can see that there is indeed a link between the maximisation in magnitude of the trigonometric sum $A(\omega)$ (the periodogram) with the norm $\|\cdot\|_1$ used in formulating the simultaneous Diophantine problem. It also provides us with a way of visualising the behaviour our estimation and association algorithm in the frequency domain.

6. NUMERICAL RESULTS

We now assess the performance of our estimation and association algorithm with some numerical tests. The LiDIA library [6] was used for its implementations of variants of the LLL algorithm.

The first test performed was of the ability of the algorithm to find the correct association with the true pulse indices for various noise levels and various numbers of observed pulses under the condition $\lambda = 0.001$ and $T = 1$. Indeed, these settings for λ and T are used throughout this section. Figure 1 shows the results obtained. Each data point plotted is the experimental frequency of successful association from 500 trials. We can see that a success rate better than 99% is achieved even for a small number of observed pulses ($n > 8$) up to and exceeding a noise level of 1% of the PRI. Of course, given that the correct association is made with the indices, the variance of $\hat{T}$ is that which is obtained by linear regression. That is, $\operatorname{var}\hat{T} = \sigma^2 / \|\mathbf{Q}\hat{\mathbf{s}}\|^2$.

We can also see that, for any fixed n, the probability of correct association appears to approach an upper bound depending on n as σ is decreased. The upper bound represents the probability that the observed indices are not coprime. The limiting probability is related to the Riemann zeta function (see [1] for a discussion of its relationship with the PRI estimation problem). The probability that n numbers chosen at random are coprime asymptotically approaches 1 very quickly as n increases, and we witness this in Figure 1.

We now consider an interpretation of the behaviour of the lattice points produced by our algorithm in the frequency domain. Figure 2 shows the square of the magnitude of $A(f)$, $f = 2\pi\omega$, as defined in (11) for a particular set of observations with $n = 7$, $T = 1$, $\lambda = 10^{-3}$ and $\sigma = 10^{-2}$. The circles represent peaks corresponding to $\omega(\mathbf{s})$ for integer vectors $\mathbf{s}$ associated with lattice points produced by reduction steps in our algorithm. Our algorithm, then, considers a proportionally large number of low frequencies before "accelerating" away into the higher frequencies. In fact, the rate of increase in frequencies appears to be exponential, as expected. If a logarithmic scale had been used along the frequency axis, we would see that the frequency samples considered by the algorithm are roughly equally distributed over the range. For the set of observations used to generate Figure 2, the algorithm was able to correctly associate indices. Observe that the frequency corresponding to these indices (at $f = 0.999997$) is the largest peak in the periodogram for the range plotted.

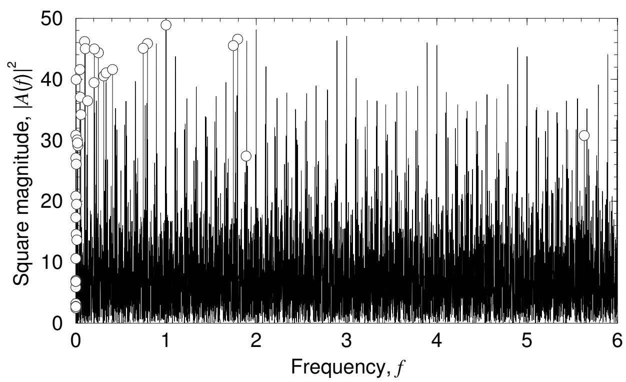

Figure 2: Interpretation of the outputs of the algorithm as a maximisation of the periodogram.

Finally, note that the number of different frequency points considered by the algorithm is 41. In contrast, the number of samples used to plot the full periodogram in Figure 2 was 6000.

7. REFERENCES

[1] S. D. Casey and B. M. Sadler, *A modified Euclidean algorithm for isolating periodicities from a sparse set of noisy measurements*, in Proc. ICASSP 95, vol. 3, 1995, pp. 1764–1767.

[2] I. V. L. Clarkson, *Approximation of Linear Forms by Lattice Points with Applications to Signal Processing*, PhD thesis, The Australian National University, 1996. In preparation.

[3] J. C. Lagarias, *The computational complexity of simultaneous Diophantine approximation problems*, in 23rd Annual Symposium on the Foundations of Computer Science, IEEE Computer Society, 1982, pp. 32–39.

[4] ——, *Best Diophantine approximations to a set of linear forms*, J. Austral. Math. Soc. (Series A), 34 (1983), pp. 114–122.

[5] A. K. Lenstra, H. W. Lenstra, and L. Lovász, *Factoring polynomials with rational coefficients*, Math. Ann., 261 (1982), pp. 515–534.

[6] LiDIA Group, *LiDIA — A Library for Computational Number Theory*, Universität des Saarlandes, 1995.