New algorithms for estimating spacecraft position using scanning techniques for Deep Space Network antennas

Lingji Chen[a], Nanaz Fathpour[b] and Raman K. Mehra[a]

[a]Scientific Systems Company Inc., 500 W. Cummings Park, Ste 3000, Woburn, MA 01801, USA;
[b]Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

ABSTRACT

As more and more nonlinear estimation techniques become available, our interest is in finding out what performance improvement, if any, they can provide for practical nonlinear problems that have been traditionally solved using linear methods. In this paper we examine the problem of estimating spacecraft position using conical scan (conscan) for NASA's Deep Space Network antennas. We show that for additive disturbances on antenna power measurement, the problem can be transformed into a linear one, and we present a general solution to this problem, with the least square solution reported in literature as a special case. We also show that for additive disturbances on antenna position, the problem is a truly nonlinear one, and we present two approximate solutions based on linearization and Unscented Transformation respectively, and one "exact" solution based on Markov Chain Monte Carlo (MCMC) method. Simulations show that, with the amount of data collected in practice, linear methods perform almost the same as MCMC methods. It is only when we artificially reduce the amount of collected data and increase the level of noise that nonlinear methods show significantly better accuracy than that achieved by linear methods, at the expense of more computation.

Keywords: conical scan (conscan), spacecraft position estimation, Deep Space Network antennas, linearization, Unscented Transformation, Markov Chain Monte Carlo (MCMC), Metropolis-Hastings (MH) algorithm, simulations

1. INTRODUCTION

Recent years have witnessed a tremendous growth of research and development in nonlinear estimation theory and techniques. In dealing with a general probability distribution, for example, a class of sampling algorithms known as Markov Chain Morte Carlo (MCMC) has been extensively used. The availability of efficient algorithms and ever-increasing computational power often indicates that such treatments are not prohibitively expensive. For some problems, MCMC provides the only viable solution. However, even for problems that have had successful solutions obtained with linear/Gaussian approximations, a natural question to ask is how much improvement in performance, if any, sophisticated nonlinear techniques can offer.

This paper is an exercise in examining one such problem, namely, the estimation of spacecraft position using scanning techniques for Deep Space Network antennas. As described in the references[1–3] , the NASA Deep Space Network antennas have spacecraft trajectory programmed into them to form the antenna command. To compensate for disturbances and determine the true position of the spacecraft, circular movements are added to the antenna command trajectory in a technique known as conical scanning (conscan). From the sinusoidal variations in the power of the signal received from the spacecraft by the antenna, the true spacecraft position can then be estimated. Since conscan is much faster than such disturbances as thermal deformations that cause the antenna not pointing precisely towards the spacecraft, the estimation problem can be mathematically described as follows. In Figure 1, the origin of the coordinate represents where the antenna would point to normally (when not performing conscan), and a_i is its position at sampling instant i during a conscan period. The spacecraft position s is unknown and assumed constant in the chosen coordinate system. Power measurement p_i is taken at each a_i, $i = 1, 2, \ldots, N$, and it is a nonlinear function of the norm of the pointing error $e_i \triangleq a_i - s$:

$$p_i = f(||e_i||) \tag{1}$$

Emails: chen@ssci.com, Nanaz.Fathpour@jpl.nasa.gov, rkm@ssci.com

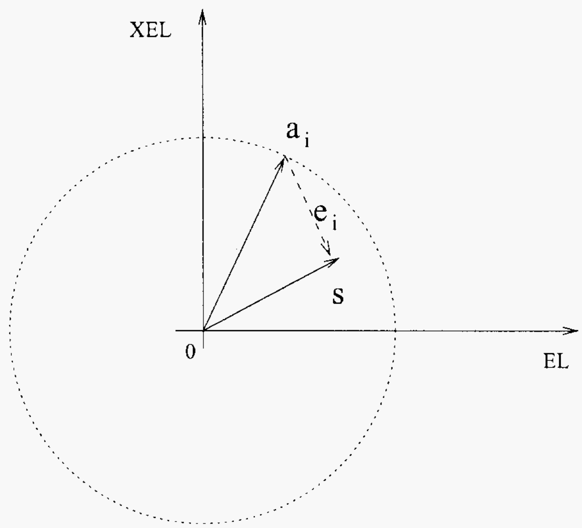

Figure 1. Illustration of the conscan problem set up

The nonlinear function $f(\cdot)$ can be approximated at several levels. In this paper, as in the reference[1], we consider only the following model:

$$p_i = p_0 \left(1 - \frac{\mu}{h^2} ||e_i||^2\right) \tag{2}$$

where p_0, μ and h are known constants.

Both the power measurement p_i itself and the antenna position a_i can be disturbed, and the problem is to determine the spacecraft position s with such measurements. In this paper only the batch mode processing is considered, *i.e.*, data collected from a full conscan period is used all at once to compute the true spacecraft position.

In the reference[1], least square solutions were obtained for the case of power disturbances, and the same solutions seemed to have been applied to the case of antenna position disturbances as well. In this paper we show that the effect of the two types of disturbances are different: The former leads to an intrinsically linear problem for the model (2), while the latter leads to an intrinsically nonlinear problem. We provide linear solutions to the problem, as well as nonlinear solutions based on MCMC. Simulations were performed using the parameters reported in the reference[1]. The results show that for the given sampling rate and noise level, linear methods are essentially as accurate as the nonlinear methods. It is only when we artificially decrease the sampling rate and increase the noise level that nonlinear methods yield significantly more accurate results (with increased computational cost). Although the results are not surprising, we hope that this exercise still can provide some reassurance and possibly insights to practitioners and theorists in the field.

2. POWER DISTURBANCES ONLY – LINEAR PROBLEM

In this section we assume that the antenna position a_i is not perturbed, and only the receiver power measurement p_i is corrupted by noise:

$$p_i = p_0 \left(1 - \frac{\mu}{h^2} ||e_i||^2\right) + v_i \tag{3}$$

where

$$e_i \triangleq a_i - s \tag{4}$$

and v_i is i.i.d. normal with mean 0 and variance σ^2. Since we consider batch processing, we define

$$p \triangleq [p_1, p_2, \ldots, p_n]^T$$

for an entire conscan period with n samples. We would also like to pose the problem in a more general Bayesian framework as follows. Assume that

$$s \text{ has a prior distribution that is normal with mean 0 and covariance } P. \tag{5}$$

We want to characterize the posterior distribution of s given noisy measurement p, and define suitable estimate of s based on the posterior distribution.

Now we introduce the following assumption: The samples a_i are symmetric on the circle in Figure 1, *i.e.*,

$$\sum_{i=1}^{n} a_i = 0 \tag{6}$$

The above holds true in the simulations in the reference[1] , but was not a requirement there.

From (3) and (4) it follows that

$$p_i = p_0 \left(1 - \frac{\mu}{h^2}(r^2 + s^T s)\right) + \frac{2p_0\mu}{h^2} a_i^T s + v_i \tag{7}$$

where r is the radius of the circle in Figure 1, and consequently $a_i^T a_i = r^2$, $i = 1, 2, \ldots, n$.

Define

$$\bar{p} \triangleq \frac{1}{n}\sum_{i=1}^{n} p_i = p_0 \left(1 - \frac{\mu}{h^2}(r^2 + s^2)\right) + \frac{1}{n}\sum_{i=1}^{n} v_i \tag{8}$$

where the assumption (6) is used.

Comment: The first term in the sum on the right hand side of (8) was called "mean power" in the reference[1] , and the above $\bar{p}$ was considered an approximation to the mean power. Here we introduce $\bar{p}$ simply as an algebraic mean of all p_i given by (7), $i = 1, 2, \ldots, n$, and this will lead to an exact solution to a linear problem, as shown in the following.

Define

$$z_i \triangleq p_i - \bar{p} = \frac{2p_0\mu}{h^2} a_i^T s + \xi_i \tag{9}$$

where

$$\xi_i \triangleq v_i - \frac{1}{n}\sum_{i=1}^{n} v_i$$

Then we have

$$z = As + \xi \tag{10}$$

where

$$z \triangleq [z_1, z_2, \ldots, z_n]^T$$

$$A \triangleq \frac{2p_0\mu}{h^2}[a_1, a_2, \ldots, a_n]^T$$

and

$$\xi \triangleq [\xi_1, \xi_2, \ldots, \xi_n]^T$$

Thus the problem has been reduced to a linear one, *i.e.*, estimating s from (10). This is an *exact* formulation for the measurement model (3), with a symmetry assumption (6) that can easily be satisfied in practice.

Before we present the solution to this problem, we first review some general results for easy reference[4] .

Review 1: If x and y are jointly Gaussian with mean

$$\begin{bmatrix} m_x \\ m_y \end{bmatrix}$$

and covariance

$$\begin{bmatrix} P_{xx} & P_{xy} \\ P_{xy}^T & P_{yy} \end{bmatrix}$$

then the conditional distribution of x conditioned on y is Gaussian with mean

$$m_{x|y} = m_x + P_{xy}P_{yy}^{-1}(y - m_y) \tag{11}$$

and covariance

$$P_{x|y} = P_{xx} - P_{xy}P_{yy}^{-1}P_{xy}^T \tag{12}$$

Review 2: If x is Gaussian with mean m_x and covariance P, and v is independent Gaussian noise with mean 0 and covariance R, then given a measurement $z = Hx + v$, the posterior distribution of x is Gaussian with mean

$$m_{x|z} = m_x + PH^T(HPH^T + R)^{-1}(z - Hm_x)$$

and covariance

$$P_{x|z} = P - PH^T(HPH^T + R)^{-1}HP$$

Solution: Let the prior distribution of s be given by (5) and the measurement vector z be given by (10). Then the posterior distribution of s is Gaussian with mean

$$m_{s|z} = PA^T(APA^T + R)^{-1}z = (P^{-1} + A^TR^{-1}A)^{-1}A^TR^{-1}z \tag{13}$$

and covariance

$$P_{s|z} = (P^{-1} + A^TR^{-1}A)^{-1}$$

where R is the covariance matrix of ξ.

Special case: A special case of the above is the following: If n is large enough, $R \approx \sigma^2 I$ where σ is the variance of power measurement noise v_i and I is the identity matrix. If we assume no prior knowledge of s, we can set $P = \infty$. Thus we have

$$m_{s|z} = (A^TA)^{-1}A^Tz \tag{14}$$

This is the least square solution given in the reference[1] .

3. POSITION DISTURBANCES – NONLINEAR PROBLEM

In the previous section the antenna position a_i is calculated from the radius of the circle in Figure 1, the velocity and the sampling time. In this section we consider the case when the true antenna position is perturbed from its nominal position, so that the pointing error is given by

$$e_i = a_i + v_i - s \tag{15}$$

where the random vector v_i is i.i.d. Gaussian with mean 0 and covariance P_v. For simplicity of presentation we omit the disturbance term in the power measurement. From (2) and (15) we have

$$p_i = p_0\left(1 - \frac{\mu}{h^2}(r^2 + s^Ts)\right) + \frac{p_0\mu}{h^2}\left(2a_i^Ts + 2v_i^Ts - 2a_i^Tv_i - v_i^Tv_i\right) \tag{16}$$

Define

$$\bar{v} \triangleq \frac{1}{n}\sum_{i=1}^{n} v_i, \quad \xi_i \triangleq 2a_i^Tv_i + v_i^Tv_i, \quad \bar{\xi} \triangleq \frac{1}{n}\sum_{i=1}^{n}\xi_i$$

Let

$$\bar{p} \triangleq \frac{1}{n}\sum_{i=1}^{n} p_i = p_0\left(1 - \frac{\mu}{h^2}(r^2 + s^Ts)\right) + \frac{p_0\mu}{h^2}\left(2\bar{v}^Ts - \bar{\xi}\right)$$

Define

$$z_i \triangleq p_i - \bar{p} = \frac{2p_0\mu}{h^2}a_i^Ts + \frac{2p_0\mu}{h^2}(v_i^T - \bar{v}^T)s + \frac{p_0\mu}{h^2}(\bar{\xi} - \xi_i) \tag{17}$$

This is a nonlinear non-Gaussian problem, since the noise v_i becomes multiplicative, and the noise term ξ_i is no longer Gaussian.

We propose in the following three solutions to this problem.

3.1. Approximate solution by linearization

Since the mean of s and v_i are both zero, the linearized equation for (17) is given by

$$z_i = \frac{2p_0\mu}{h^2} a_i^T s + \frac{2p_0\mu}{h^2} \left(\frac{1}{n} \sum_{i=1}^{n} a_i^T v_i - a_i^T v_i \right) \tag{18}$$

Let

$$\eta_i \triangleq \frac{2p_0\mu}{h^2} \left(\frac{1}{n} \sum_{i=1}^{n} a_i^T v_i - a_i^T v_i \right)$$

Then (18) will be in the same form as (9), and the solution can be obtained as in Section 2.

3.2. Approximate solution by Unscented Transformation

Define

$$z \triangleq [z_1, z_2, \ldots, z_n]^T$$

and

$$v \triangleq [v_1^T, v_2^2, \ldots, v_n^T]^T,$$

and we can state the problem as follows. Given a jointly Gaussian distribution of s (the prior) and v, we want to find the joint distribution of s and z and in turn the conditional distribution of s on z, in order to obtain an estimate of s.

In the previous section the joint distribution of s and z is obtained by linearizing the equation (17). Another way of achieving this is through the so called "Unscented Transformation:"

> **Review 3:** Unscented Transformation: If x has a Gaussian distribution, after a nonlinear transformation $y = f(x)$, y may no longer have a Gaussian distribution. If we want to approximate the distribution of y with a Gaussian one, we can proceed by linearizing the transformation $f(\cdot)$. Another way is to use some deterministically chosen points x_i (called *Sigma Points*) to represent the random variable x, and calculate their transformed values $y_i = f(x_i)$, $i = 1, 2, \ldots, 2N + 1$ where N is the dimension of x. Then these y_i can be fit with a Gaussian distribution.

We will not go into the details in this paper, but refer interested readers to[5] .

3.3. "Exact" solution by Markov Chain Monte Carlo (MCMC)

Any numeric solution is an approximate solution to some extent, but here we use the word "exact" to indicate that we are trying to obtain a general probability distribution rather than its Gaussian approximation. A general distribution can be represented by independent samples, which we will review first.

> **Review 4:** Monte Carlo integration: If we can draw samples x_i, $i = 1, 2, \ldots, N$, of a random variable x that has a probability density function (pdf) $f(x)$, then some statistics of x (for example, its mean) can be estimated using these samples:
>
> $$\int \phi(x) f(x) dx \approx \frac{1}{N} \sum_{i=1}^{N} \phi(x_i)$$
>
> It can be readily shown that such an estimate is unbiased and converges to the true answer as N goes to infinity.

For some distributions such as uniform or Gaussian, it is easy to draw samples from them, even when the random variable is in high dimensions. For a general pdf $f(\cdot)$, however, sampling it may be very difficult. Moreover, we often can calculate $f(\cdot)$ only up to a normalizing constant. Therefore we often resort to a *proposal distribution* $g(\cdot)$ which is easy to sample from. In the following reviews[6, 7] we first describe a simple technique called *rejection sampling* in order to get familiar with the concept of a proposal. Following this we will introduce the more sophisticated *Metropolis-Hastings (MH) Algorithm*, which is one of the most popular MCMC methods.

Review 5: Rejection Sampling: Let $M > 0$ be a known constant such that $f(x) \leq Mg(x)$ for all x. Take a sample x^* from the proposal distribution $g(\cdot)$, and independently take a sample u from a uniform distribution between 0 and 1. If $u < \frac{f(x^*)}{Mg(x^*)}$, then accept x^* as a sample of $f(\cdot)$; otherwise, reject it. Repeat the process until the desired number of samples have been obtained.

Intuitively, if we cover the x-$f(x)$ plane with uniformly distributed points, then those under the curve $Mg(x)$ will have their x-coordinates distributed according to $g(\cdot)$, and those that are picked by using $u < \frac{f(x^*)}{Mg(x^*)}$ in the above Review will fall under the curve of $f(x)$, which means that their x-coordinates wil be distributed according to (a normalized pdf) $f(\cdot)$.

For high dimensions the above method does not perform well, due to the low acceptance rate and other reasons. The MH Algorithm is more efficient.

Review 6: MH Algorithm: Choose a set of initial samples $\{x_1^{(0)}, x_2^{(0)}, \ldots, x_N^{(0)}\}$. Each sample will then go through a Markov Chain, *i.e.*, from step i to step $i+1$, $x_j^{(i)}$ will become $x_j^{(i+1)}$ according to some transitional probability, $j = 1, 2, \ldots, N$. When L number of steps have been completed, the set $\{x_1^{(L)}, x_2^{(L)}, \ldots, x_N^{(L)}\}$ will be taken as samples of the desired distribution.

The MH Algorithm defines the transition from $x^{(i)}$ to $x^{(i+1)}$ (where we have omitted the subscript j for simplicity) as follows: Let $f(\cdot)$ be the desired distribution (possibly unnormalized). Take a sample x^* from a proposal distribution $g(x^*|x^{(i)})$. Take a sample u from a uniform distribution between 0 and 1. If $u < \min\left\{1, \frac{f(x^*)g(x^{(i)}|x^*)}{f(x^{(i)})g(x^*|x^{(i)})}\right\}$ then $x^{(i+1)} = x^*$. Otherwise $x^{(i+1)} = x^{(i)}$.

In the conscan problem, we are interested in obtaining the posterior pdf $f(s|z)$ of the spacecraft position s, given the measurement vector $z \triangleq [z_1, \ldots, z_n]^T$ from (17), and the prior distribution of s from (5). By Bayes rule, this pdf can be calculated as follows:

$$f(s|z) = \frac{f(z|s)f(s)}{\text{normalizing constant}} \tag{19}$$

The prior term $f(s)$ can be calculated from a Gaussian pdf. Upon examining the measurement equation

$$z_i \triangleq p_i - \bar{p} = \frac{2p_0\mu}{h^2} a_i^T s + \frac{2p_0\mu}{h^2}(v_i^T - \bar{v}^T)s + \frac{p_0\mu}{h^2}(\bar{\xi} - \xi_i)$$

we find that the likelihood term

$$f(z|s) = \prod_{i=1}^{n} f(z_i|s)$$

can be calculated from a *non-central χ-square ditribution*, assuming a noise covariance matrix $P_v = \sigma_v^2 I$. Note that for this likelihood calculation, s is given, so we can also work with the original power measurement

$$p_j = p_0\left(1 - \frac{\mu}{h^2}(r^2 + s^T s)\right) + \frac{p_0\mu}{h^2}\left(2a_j^T s + 2v_j^T s - 2a_j^T v_j - v_j^T v_j\right), j = 1, 2, \ldots, n$$

Thus we can calculate the desired distribution up to a normalizing constant. To use the MH Algorithm, we need to specify a proposal distribution $g(\cdot|x)$. We adopt the following:

$$g(\cdot|x) \text{ is Gaussian with mean } x \text{ and covariance } P. \tag{20}$$

where P is the initial covariance of s defined in (5).

4. SIMULATIONS

We have conducted several groups of simulations for the nonlinear problem discussed in the previous section. Each simulation consists of the following steps:

1. Choose the noise parameter σ_v for the covariance matrix $P_v = \sigma_v^2 I$ of the position disturbance.
2. Choose the number of measurements n taken in a conscan period.
3. Repeat the following 15 times:
   (a) Randomly choose an "unknown" spacecraft position s according to the prior distribution. (The covariance P in (5) is chosen such that 99% of the time s lies within the conscan circle in Figure 1.)
   (b) Calculate the **L**east **S**quare estimate (similar to (14) but for Section 3.1), the general **Linear** estimate (similar to (13) but for Section 3.1), and the **MCMC** estimate (mean of the samples in Section 3.3). Calculate the Eucleadian distance from an estimate to the true position, and record it for each solution in each simulation.
4. Calculate the mean and standard deviation of the Eucleadian distances for each solution.

In the MCMC simulation, $N = 100$ particles were chosen to run $L = 500$ steps. The rest of the parameters that have not been specified are chosen as in the reference[1] . The simulation results have been tabulated for the cases $\sigma_v = 1.5$ and $\sigma_v = 0.3$ respectively, in Table 1.

points	3000		300		30		10	
	mean	std	mean	std	mean	std	mean	std
LS	0.06461	0.02943	0.2448	0.1268	0.7805	0.5281	1.3949	1.0323
Linear	0.06377	0.02980	0.2233	0.1185	0.4951	0.2976	0.5705	0.3139
MCMC	0.06417	0.02917	0.2086	0.1032	0.4404	0.2568	0.5173	0.2954

points	3000		300		30		10	
	mean	std	mean	std	mean	std	mean	std
LS	0.008553	0.004643	0.02734	0.01279	0.1196	0.05627	0.1946	0.08264
Linear	0.008647	0.004598	0.02736	0.01251	0.1184	0.05262	0.1760	0.08010
MCMC	0.008640	0.004590	0.02705	0.01283	0.1227	0.05813	0.1670	0.07329

Table 1. Estimation errors for the three algorithms on different numbers of measurement points, for high level of noise (top) and low level of noise (bottom)

We observe the following from the table:

- When the noise level is low, all methods perform more or less the same.
- When the noise level is high, but there is plenty of data ($n = 3000$ as in the reference[1]), all methods perform more or less the same. It is only when the conscan sampling rate is artificially dropped to have less data points ($n = 300$, $n = 30$ and $n = 10$) that nonlinear methods offer better accuracy.

5. CONCLUSIONS AND FUTURE WORK

As more and more nonlinear estimation techniques become available, our interest is in finding out what performance improvement, if any, they can provide for practical nonlinear problems that have been traditionally solved using linear methods. In this paper we have examined the problem of estimating spacecraft position using conical scan (conscan) for NASA's Deep Space Network antennas. We showed that for additive disturbances on antenna power measurement, the problem can be transformed into a linear one, and we presented a general solution to this problem, with the least square method reported in reference[1] as a special case. We also showed that for additive disturbances on antenna position, the problem is a truly nonlinear one, and we presented two approximate solutions based on linearization and Unscented

Transformation respectively, and one "exact" solution based on Markov Chain Monte Carlo (MCMC) method. Simulations showed that, with the amount of data collected in practice[1] , linear methods perform almost the same as MCMC methods. It is only when we artificially reduced the amount of collected data and increased the level of noise when nonlinear methods show significantly better accuracy than that achieved by linear methods, at the expense of more computation.

We hope that this exercise will add confidence to the practioners who are using linear methods to solve their problems, and also provide an example of what additional computation can achieve for problems with sparse data and high noise level.

6. ACKNOWLEDGEMENT

The authors would like to thank Dr. David Bayard and Dr. Wodek Gawronski who introduced us to the conscan problem and for their many useful discussions on the subject matter. This work was partially prepared at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

1. W. Gawronski and E. M. Craparo, "Three scanning techniques for deep space network antennas to estimate spacecraft position," *The Interplanetary Network Progress Report* **42-147**, pp. 1–17, 2001.
2. D. B. Eldred, "An improved conscan algorithm based on a kalman filter," *The Telecommunications and Data Acquisition Progress Report* **42-116**, pp. 223–231, 1993.
3. L. S. Alvarez, "Analysis of open-loop conical scan pointing error and variance estimators," *The Telecommunications and Data Acquisition Report* **06-32**, pp. 81–90, 1993.
4. P. S. Maybeck, *Stochastic models, estimation, and control*, vol. 141 of *Mathematics in Science and Engineering*, 1979.
5. S. Julier, J. Uhlmann, and H. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control* **45**(3), pp. 477 –482, 2000.
6. C. Andrieu, N. de Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Machine Learning* (50), 2003.
7. C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer Texts in Statistics, Springer-Verlag, 2004.