# A NONITERATIVE MAXIMUM LIKELIHOOD PARAMETER ESTIMATOR OF SUPERIMPOSED CHIRP SIGNALS

*Supratim Saha, Steven Kay.*

Department of Electrical and Computer Engineering, University of Rhode Island

**ABSTRACT**

We address the problem of parameter estimation of superimposed chirp signals in noise. The approach used here is a computationally modest implementation of a maximum likelihood (ML) technique. The ML technique for estimating the complex amplitudes, chirping rates and frequencies reduces to a separable optimization problem where the chirping rates and frequencies are determined by maximizing a compressed likelihood function which is a function of only the chirping rates and frequencies. Since the compressed likelihood function is multidimensional, its maximization via grid search is impractical. We propose a non-iterative maximization of the compressed likelihood function using importance sampling. Simulation results are presented for a scenario involving closely spaced parameters for the individual signals.

## 1. INTRODUCTION

Chirp signals are encountered in many different engineering applications including radar, active sonar and passive sonar systems. The problem of parameter estimation of chirp signals has received a great deal of attention, [3]. These approaches have been proven to be effective in the sense that they achieve the Cramer Rao Lower Bound (CRLB). However most of these approaches are designed for a single chirp signal. Parameter estimation of superimposed chirp signals is a difficult signal processing problem. The need for determining the parameters of superimposed chirp signals arises in passive sensor array systems, where it has been shown in [6] that the problem of range and direction of arrival estimation for moderately far, broadside targets reduces to that of estimating the parameters of sums of chirp signals. Liang and Arun [5] have also addressed an iterative maximum likelihood (ML) approach to this problem. Rank reduction techniques were used to get good initial parameter estimates, which were then used in a maximum likelihood procedure to get the final estimates. Although the approach has been shown to achieve good results at high SNRs, there is no guarantee that the global optimum will be achieved.

Our aim in this paper is to develop a non iterative computationally modest implementation of a ML estimator for the chirp signal parameters. To develop the estimator, we first show that the data model involves estimation of linear and nonlinear parameters of a partial general linear model [1]. The complex amplitudes form the linear parameter vector and the chirp rates and frequencies form the nonlinear parameter vector. The parameter estimation gets decoupled, where the nonlinear parameter vector needs to be estimated first by maximizing a compressed likelihood function involving only the chirp rate and frequencies as unknown parameters. The complex amplitudes are obtained from the estimates of chirp rates and frequencies. In this paper we focus on estimation of chirp rates and frequencies only. The straightforward implementation of the maximization of the compressed likelihood function involves a grid search which is impractical and whose computational complexity increases with the number of signals. To carry out this maximization non-iteratively we use a global optimization theorem proposed in [2]. To efficiently implement the optimization, we use Monte Carlo Importance Sampling [4]. It is observed that the technique produces good estimates for the unknown parameters even in cases where the individual parameters are closely spaced. Furthermore, the computational burden is quite modest.

## 2. PROBLEM DEFINITION

A sequence $x\left[n\right], n = 0, \ldots, N-1$ is observed having the following parametric representation.

$$x\left[n\right] = \sum_{i=1}^{p} A_i \exp\left[j(2\pi f_i n + \pi m_i n^2)\right] + w\left[n\right] \qquad (1)$$

where the parameters, chirp rate $m_i$ ( $0 \le m_i \le 2$), frequency $f_i$ ($0 \le f_i \le 1$) and the complex amplitudes $A_i$ for $i = 1, \ldots p$ are unknown. The noise $w\left[n\right]$, $n = 0, \ldots, N-1$ is a segment of a zero mean complex Gaussian random process. The aim is to obtain maximum likelihood estimates of the chirp rate $m_i$ and frequency $f_i$, for $i = 1, \ldots, p$ from $x\left[n\right]$ for $n = 0, \ldots, N-1$.

The data described by (1) can be expressed in matrix form as

$$\mathbf{x} = \mathbf{H}(\boldsymbol{\alpha}, \boldsymbol{\beta})\boldsymbol{\theta} + \mathbf{w} \qquad (2)$$

where $\boldsymbol{\alpha} = [f_1 \dots f_p]^T$, and $\boldsymbol{\beta} = [m_1 \dots m_p]^T$. The matrix $\mathbf{H}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ can be expressed as

$$\mathbf{H}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = [\mathbf{e}(\alpha_1, \beta_1) \dots \mathbf{e}(\alpha_p, \beta_p)] \quad (3)$$

where the vector $\mathbf{e}(\alpha_i, \beta_i)$ is given by

$$\mathbf{e}(\alpha_i, \beta_i) = \begin{bmatrix} \exp(j2\pi f_i(0) + j\pi m_i(0)^2) \\ \exp(j2\pi f_i(1) + j\pi m_i(1)^2) \\ \dots \\ \exp(j2\pi f_i(N-1) + j\pi m_i(N-1)^2) \end{bmatrix} \quad (4)$$

Since the noise is assumed to be additive white Gaussian with variance $\sigma^2$, the probability density function of the data vector $\mathbf{x}$ in (2) parameterized by $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}$, is given by, $p(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta})$, which is equal to

$$\frac{1}{\pi^N \sigma^N} \exp\left[ -\frac{1}{\sigma^2} (\mathbf{x} - \mathbf{H}(\boldsymbol{\alpha}, \boldsymbol{\beta})\boldsymbol{\theta})^H (\mathbf{x} - \mathbf{H}(\boldsymbol{\alpha}, \boldsymbol{\beta})\boldsymbol{\theta}) \right] \quad (5)$$

Hence the likelihood function of the data $L(\mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta})$ is given by

$$L(\mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}) \propto p(\mathbf{x}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}) \quad (6)$$

The joint ML estimates of $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}$ is obtained by maximizing $L(\mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta})$. From (5) and (6) this joint maximization is equivalent to following step

$$\min_{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}} (\mathbf{x} - \mathbf{H}(\boldsymbol{\alpha}, \boldsymbol{\beta})\boldsymbol{\theta})^H (\mathbf{x} - \mathbf{H}(\boldsymbol{\alpha}, \boldsymbol{\beta})\boldsymbol{\theta})$$

The parameter vectors $\boldsymbol{\alpha}, \boldsymbol{\beta}$ which appear in the matrix $\mathbf{H}$ are nonlinearly related to $\mathbf{x}$ whereas the parameter vector $\boldsymbol{\theta}$ is linearly related to $\mathbf{x}$. It is known that for such kinds of joint parameter estimation problems as in (2), the parameter estimation procedure gets decoupled [1] where estimation of the unknown nonlinear parameters is done first and the estimated nonlinear parameters are inserted in the matrix $\mathbf{H}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ to obtain the linear parameter estimate. The estimates of the two nonlinear parameters are obtained as [1]

$$\left[ \hat{\boldsymbol{\alpha}}_{mle}, \hat{\boldsymbol{\beta}}_{mle} \right] = \max_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \left[ \mathbf{x}^H (\mathbf{H} \left[ \mathbf{H}^H \mathbf{H} \right]^{-1} \mathbf{H}^H) \mathbf{x} \right] \quad (7)$$

The function in the right hand side (RHS) of (7) is called the compressed likelihood function $L_c(\boldsymbol{\alpha}, \boldsymbol{\beta})$. It can be observed from (7) that obtaining $\left[ \hat{\boldsymbol{\alpha}}_{mle}, \hat{\boldsymbol{\beta}}_{mle} \right]$ will require a multidimensional grid search over the two parameter vectors. It is because of the lack of closed form solution that the proposed approaches for these kinds of problems have been iterative. Liang and Arun [5] have reported a 3 stage iterative ML approach for this problem. Pincus [2] showed that for a function of several variables having many local maxima, it is possible to have a closed form expression for the variables fetching the global maximum. Motivated by the result of [2], we develop a non-iterative estimator for $\left[ \hat{\boldsymbol{\alpha}}_{mle}, \hat{\boldsymbol{\beta}}_{mle} \right]$.

## 3. GLOBAL OPTIMIZATION THEOREM

The theorem proposed by Pincus [2] was used for obtaining the maximum/minimum of a multidimensional function. It is stated as follows. We apply this theorem to obtain the estimates of the vectors $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ which maximize the compressed likelihood function $L_c(\boldsymbol{\alpha}, \boldsymbol{\beta})$. Based on the theorem [2], the estimates of $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\beta}}$ are given by

$$[\hat{\boldsymbol{\alpha}}]_i = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} [\boldsymbol{\alpha}]_i \, \bar{L}_c'(\boldsymbol{\alpha}, \boldsymbol{\beta}) d\boldsymbol{\alpha} d\boldsymbol{\beta} \quad (8)$$

and

$$[\hat{\beta}]_i = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} [\beta]_i \, \bar{L}_c'(\boldsymbol{\alpha}, \boldsymbol{\beta}) d\boldsymbol{\alpha} d\boldsymbol{\beta} \quad (9)$$

where

$$\bar{L}_c'(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \lim_{\rho \to \infty} \frac{\exp(\rho L_c(\boldsymbol{\alpha}, \boldsymbol{\beta}))}{\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp(\rho L_c(\boldsymbol{\alpha}, \boldsymbol{\beta})) d\boldsymbol{\alpha} d\boldsymbol{\beta}} \quad (10)$$

and $[\hat{\boldsymbol{\alpha}}]_i$ and $[\hat{\boldsymbol{\beta}}]_i$ are the $i^{th}$ components of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. It can be observed from (8) and (9) that the theorem provides a closed form expression for obtaining the parameters that maximize the function but its evaluation requires computation of a multidimensional integral. However it can be noted that the integrations involved in (8) and (9) are closely related with integrations involved in probability theory to compute expected values of random variables having a joint probability density function (PDF). This is because the normalized function $\bar{L}_c'(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is positive and has all the properties of a joint PDF. However the parameter vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are not random. Thus the normalized function is termed a pseudo-PDF. Using this concept, the Monte Carlo techniques can be used to replace the multidimensional integrations in (8) and (9). The simplest Monte Carlo approach would require generation of random vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ distributed according to the joint PDF $\bar{L}_c'(\boldsymbol{\alpha}, \boldsymbol{\beta})$. However, $\bar{L}_c'(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is a highly nonlinear function of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. As a result, direct generation of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ is not easy and one needs to resort to *other* Monte Carlo techniques which generate samples according to some simpler PDF and use those samples to estimate the means. Importance Sampling belongs to this class of Monte Carlo techniques and has been proved to be a highly effective tool in evaluation of integrals in Bayesian theory [4]. We thus use importance sampling described in the next section to efficiently evaluate the estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

## 4. IMPORTANCE SAMPLING

The importance sampling approach is based on the observation that integrals of the type $\int h(\mathbf{x}) \bar{L}'(\mathbf{x}) \mathbf{dx}$ can be expressed as

$$\int h(\mathbf{x})\bar{L}'(\mathbf{x})\mathbf{dx} = \int h(\mathbf{x})\frac{\bar{L}'(\mathbf{x})}{\bar{g}(\mathbf{x})}\bar{g}(\mathbf{x})d\mathbf{x} \tag{11}$$

where $\bar{g}(\mathbf{x})$ is assumed to possess all the properties of a PDF. Then, the right-hand-side of (11) can be expressed as the expected value of $h(\mathbf{x})\frac{\bar{L}'(\mathbf{x})}{\bar{g}(\mathbf{x})}$ , with respect to the pseudo-PDF $\bar{g}(\mathbf{x})$. The function $\bar{g}(\mathbf{x})$ is called the normalized importance function. Unlike $\bar{L}'(\mathbf{x})$, which in general is a nonlinear function of $\mathbf{x}$, $\bar{g}(\mathbf{x})$ can be chosen to be some simple function of $\mathbf{x}$, so that realizations of $\mathbf{x}$ can be easily generated. Then, the value of the integral in (11) can be found by the Monte Carlo approximation

$$\frac{1}{K}\sum_{k=1}^{K} h(\mathbf{x}_k)\frac{\bar{L}'(\mathbf{x}_k)}{\bar{g}(\mathbf{x}_k)} \tag{12}$$

where $\mathbf{x}_k$ is the $k^{th}$ realization of the vector $\mathbf{x}$ distributed according to the pseudo-PDF $\bar{g}(\mathbf{x})$. The value of $K$ needed for a good approximation depends on the choice of $g$. Typically, $\bar{g}(\mathbf{x})$ should be chosen similar to $\bar{L}(\mathbf{x})$, as this reduces the variance of the estimate given by (12). However, another important point to keep in mind when choosing $\bar{g}(\mathbf{x})$ is that it should be simple enough so that $\mathbf{x} \sim \bar{g}(\mathbf{x})$ can be easily generated [4] .

The ideas expressed by (11) and (12) can be applied for the estimation of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, once the importance function for this problem is defined. In particular, if the normalized importance function is $\bar{g}(\boldsymbol{\alpha},\boldsymbol{\beta})$, then the estimates of the coordinates of the vector $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ computed using this importance function is expressed as

$$[\hat{\boldsymbol{\alpha}}]_i = \frac{1}{R}\sum_{k=1}^{R}[\boldsymbol{\alpha}_k]_i\frac{\bar{L}_c'(\boldsymbol{\alpha}_k,\boldsymbol{\beta}_k)}{\bar{g}(\boldsymbol{\alpha}_k,\boldsymbol{\beta}_k)} \tag{13}$$

and

$$\left[\hat{\boldsymbol{\beta}}\right]_i = \frac{1}{R}\sum_{k=1}^{R}[\boldsymbol{\beta}_k]_i\frac{\bar{L}_c'(\boldsymbol{\alpha}_k,\boldsymbol{\beta}_k)}{\bar{g}(\boldsymbol{\alpha}_k,\boldsymbol{\beta}_k)} \tag{14}$$

where $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$ are the $k^{th}$ realizations of the vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ distributed according to the importance function $\bar{g}(\boldsymbol{\alpha},\boldsymbol{\beta})$. The normalized importance function $\bar{g}(\boldsymbol{\alpha},\boldsymbol{\beta})$ needs to be chosen so that the samples $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$ can be easily generated. Furthermore, $\bar{g}(\boldsymbol{\alpha},\boldsymbol{\beta})$ should be a close approximation to $\bar{L}_c'(\boldsymbol{\alpha},\boldsymbol{\beta})$. Since

$$L_c(\boldsymbol{\alpha},\boldsymbol{\beta}) = \mathbf{x}^H\mathbf{H}(\boldsymbol{\alpha},\boldsymbol{\beta})\left[\mathbf{H}^H(\boldsymbol{\alpha},\boldsymbol{\beta})\mathbf{H}(\boldsymbol{\alpha},\boldsymbol{\beta})\right]^{-1}\mathbf{H}^H(\boldsymbol{\alpha},\boldsymbol{\beta})\mathbf{x}$$

it is obvious that the function $L_c'(\boldsymbol{\alpha},\boldsymbol{\beta})$ is not separable in $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. However, if we force the matrix $\mathbf{H}^H(\boldsymbol{\alpha},\boldsymbol{\beta})\mathbf{H}(\boldsymbol{\alpha},\boldsymbol{\beta})$ to be an identity matrix, then the function $L_c(\boldsymbol{\alpha},\boldsymbol{\beta})$ will become separable in $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. This is the main idea behind choosing the importance function. Thus the importance function is chosen by forcing the $p \times p$ matrix $\mathbf{H}^H(\boldsymbol{\alpha},\boldsymbol{\beta})\mathbf{H}(\boldsymbol{\alpha},\boldsymbol{\beta})$ to be $N\mathbf{I}_p$, where $\mathbf{I}_p$ is a $p \times p$ identity matrix. To make the importance function similar to the function $L_c'(\boldsymbol{\alpha},\boldsymbol{\beta})$, we choose the importance function as

$$g(\boldsymbol{\alpha},\boldsymbol{\beta}) = \exp(\rho\mathbf{x}^H\mathbf{H}(\boldsymbol{\alpha},\boldsymbol{\beta})\frac{\mathbf{I}}{N}\mathbf{H}^H(\boldsymbol{\alpha},\boldsymbol{\beta})\mathbf{x}) \tag{15}$$

and its normalized version as $\bar{g}(\boldsymbol{\alpha},\boldsymbol{\beta})$

$$\bar{g}(\boldsymbol{\alpha},\boldsymbol{\beta}) = \frac{g(\boldsymbol{\alpha},\boldsymbol{\beta})}{\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty}g(\boldsymbol{\alpha},\boldsymbol{\beta})d\boldsymbol{\alpha}d\boldsymbol{\beta}} \tag{16}$$

## 5. ESTIMATION OF PARAMETERS

As a result of this choice of $g(\boldsymbol{\alpha},\boldsymbol{\beta})$ in (15), the importance function now becomes separable in $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ and can be expressed as

$$\bar{g}(\boldsymbol{\alpha},\boldsymbol{\beta}) = \prod_{i=1}^{p}\bar{g}(\alpha_i,\beta_i) \tag{17}$$

This enables generation of $p$ independent samples of $(\alpha_i,\beta_i)$, distributed according to the joint PDF $\bar{g}(\alpha_i,\beta_i)$, with the condition that no two of $(\alpha_i,\beta_i)$ are the same. The $p$ parameter vectors $(\alpha_i,\beta_i)$, for $i = 1,\dots,p$ need to be distinct. This assumption is necessary for this problem as otherwise the matrix $\mathbf{H}^H(\boldsymbol{\alpha},\boldsymbol{\beta})\mathbf{H}(\boldsymbol{\alpha},\boldsymbol{\beta})$ will become singular. The variables $\alpha_i$ and $\beta_i$ are generated jointly using the following three steps

1. Evaluate the two-dimensional joint PDF $\bar{g}(\alpha,\beta)$ at $M \times M$ discrete set of points on a rectangular grid and obtain the marginal PDF $\bar{g}(\alpha)$ as

   $$\bar{g}(\alpha_l) = \sum_{i=1}^{M}\bar{g}(\alpha_l,\beta_i)\delta\beta_i$$

   for $l = 1,\dots M$. From the marginal PDF $\bar{g}(\alpha_l)$ obtain the cumulative distribution function $G(\alpha)$, as

   $$G(\alpha) = \int_0^{\alpha}\bar{g}(x)dx$$

   by approximating the integral as a sum.

2. From the marginal PDF $\bar{g}(\alpha_l)$ so obtained in step 1, obtain the conditional PDF $\bar{g}(\beta|\alpha)$ as

   $$\bar{g}(\beta_k|\alpha_l) = \frac{\bar{g}(\beta_k,\alpha_l)}{\bar{g}(\alpha_l)}$$

   Evaluate the cumulative distribution function of the conditional PDF $\bar{g}(\beta|\alpha)$ as

   $$G(\beta_k|\alpha_l) = \int_0^{\alpha}\bar{g}(x|\beta_k)dx$$

   for all $k, l = 1,\dots,M$.

3. Generate $u_1, u_2, \sim U[0,1]$ and obtain $\alpha = G^{-1}(u_1)$ and $\beta = G^{-1}(u_2|\alpha)$. Repeat this steps $p$ times to obtain a realization of the vector $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, each of which has dimension $p \times 1$.

4. Repeat step 3 $R$ times to obtain $R$ realizations of the vector $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

Now these realizations can be used in (13 ) and (14) to obtain the estimates which are essentially the linear means of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. However we do not use (13) and (14). Rather, we make further use of the limited ranges of $\alpha_i$ and $\beta_i$ in reducing the computations. Since $0 \leq \alpha_i \leq 1$ and $0 \leq \beta_i \leq 2$, they posses the properties of a circular random variable. Circular mean also alleviates the bias [7]. We thus compute the circular means and obtain the angle of the circular means to compute $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. The expressions for the estimates based on the circular mean definition are given by,

$$[\hat{\boldsymbol{\alpha}}]_i = \frac{1}{2\pi}\angle\frac{1}{R}\sum_{k=1}^{R}\exp(j2\pi\,[\boldsymbol{\alpha}_k]_i)\frac{L_c'(\boldsymbol{\alpha}_k,\boldsymbol{\beta}_k)}{g(\boldsymbol{\alpha}_k,\boldsymbol{\beta}_k)} \quad (18)$$

and

$$\left[\hat{\boldsymbol{\beta}}\right]_i = \frac{2}{2\pi}\angle\frac{1}{R}\sum_{k=1}^{R}\exp(j2\frac{\pi}{2}\,[\boldsymbol{\beta}_k]_i)\frac{L_c'(\boldsymbol{\alpha}_k,\boldsymbol{\beta}_k)}{g(\boldsymbol{\alpha}_k,\boldsymbol{\beta}_k)} \quad (19)$$

It should be noted that using (18) and (19), the normalizing constants are no longer required as a result of the $\angle$ operator, thus reducing the computational burden considerably.

## 6. SIMULATION RESULTS AND CONCLUSIONS

We present an example of estimation of parameters of two equipower closely spaced chirp signals for which $A_1 = 1, f_1 = 0.6, m_1 = 0.71$. and $A_2 = 1, f_2 = 0.63, m_2 = 0.75$. The data record length is 50 and SNR is chosen to be 10 dB. In Figure 1, we plot the coordinates of the estimates of $(f_1, m_1)$ and $(f_2, m_2)$. The x-axis refers to frequency whereas y-axis refers to chirp rate. The true coordinates is shown by circles. The estimate for 100 realizations is plotted. It can be observed that the technique is always able to resolve the signals and the estimates lie very near the actual signal parameters. The number of realizations $R$ needed to obtain estimates from (18) and (19) for the simulation was 3500 and the minimum value of $\rho$ was 0.4.

However more exhaustive simulations need to be done to asses the merit of the method, by finding the threshold SNR for a given data record length at which the CRLB is no longer achieved. This is currently under investigation.

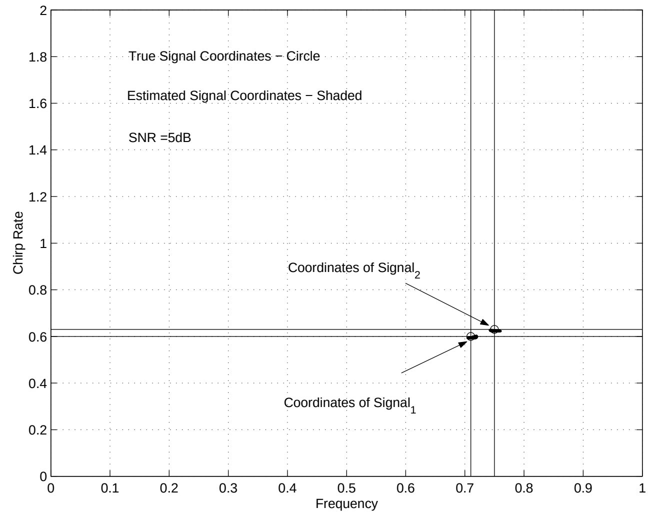


**Fig. 1**. Plot of True Parameters and 100 realizations of estimates

## 7. REFERENCES

[1] S. Kay, "Fundamentals of Statistical Signal Processing": Estimation Theory, Prentice Hall, Englewood Cliffs, NJ, 1993.

[2] M. Pincus, "A Closed Form Solution for Certain Programming Problems", *Oper. Research*, pp. 690-694, 1962.

[3] C.S. Ramalingam, "Nonstationary Signal Analysis", Chapter 2, Doctoral Thesis, University of Rhode Island, 1995.

[4] L. Stewart, "Bayesian Analysis using Monte Carlo Integration: a Powerful Methodology for Handling Some Difficult Problems", *American Statistician*, pp. 195–200, 1983.

[5] R.M. Liang and K.S. Arun, "Parameter Estimation for Superimposed Chirp Signals", Proc. of International Conference on Acoustics, Speech and Signal Processing, 273-276, 1992.

[6] K.S. Arun and R.M. Liang, "A Closed Form Solution for Instantaneous Amplitude and Frequency Estimation: Performance Bounds and applications to source localization *Proc. SPIE*, July 1991.

[7] B.C. Lowell, P.J. Kootsookos, R.C. Williamson, "The Circular Nature of Discrete-Time Frequency Estimates", *Pro. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, pp. 3369-3372, May 1991.