Iterative constrained penalized likelihood estimation of parameters for CDMA

E. Khan, D. Slock

Department of Mobile Communication
Institute Eurecom
06904, Sophia-Antipolis, France

Abstract

We describe an iterative method for Maximum Likelihood (ML) parameter estimation corrupted by additive white Gaussian noise. In the objective function we subtract/ add Kullback-Leibler (KL) distance function or euclidean distance function to keep the old parameter set close to the new ones and can be considered as penalty term. The above augmented cost function can be maximized/ minimized over the constraint that the detected data vector lie on the sphere. We simplify this constraint function by using first order Taylor expansion at the old parameter value. The useful behavior of the proposed algorithm is verified by numerical experiments.

1 Introduction

Code Division Multiple Access (CDMA) is one of the most common multiple access techniques for wireless communication systems involving non orthogonal signalling. In CDMA system all resources are in principle available to all users simultaneously. The users are distinguished from each other by user specific signature sequences, modulating the transmitted data sysmbols using direct sequence spread spectrum techniques. In the past many iterative techniques have been considered. Talwar et al [1] proposed iterative least square with enumeration (ILSE), which solves the problem by estimating the channel by short training sequence or from previous estimate and find data sequence over all possible data in the finite alphabet (FA). They also proposed iterative least square with projection (ILSP), which also initially estimates the channel with the same method as for ILSE and treats the problem as continous optimization problem and projects the result onto the closest discrete alphabet. In [3], a constrained Maximum Likelihood problem was considered with the data vector to lie with in hypercube and called it as Box constrained ML. Similarly, they also proposed problem of maximizing likelihood function over sphere, i.e. confine the solution vector to lie within the sphere and project the solution vector on the sphere. Infact, in the sphere constrained problem the solution vector lies on the sphere and not in the interior of the constraining sphere (as is done in [2]). The other problem with their method is that small error in the solution vector can cause large error when projected on to sphere (provided the solution vector is well inside the sphere). In this paper we constrain the solution vector to lie almost (very close) on the sphere and jointly estimate the complex channel coefficients and data vector. The rest of the paper is organized as follows: The signal model for our problem is described in section 2. In section 3, we develop sphere constrained approximate penalized likelihood function. In section 4, we analyze the performance of the proposed method and simulations are presented.

2 Signal model

In this section, discrete-time base band uplink signal model for CDMA communication system is described. We consider asynchronous CDMA with single path channels. The signal is corrupted by the presence of the additive white Gaussian noise (AWGN) with zero mean and variance $\frac{N_o}{2} = \sigma^2$. The number of users in the system are assumed to be K. The processing gain, $N = T_d/T_c$, where T_d is symbol duration and T_c is the chip duration. The users transmit binary information symbol stream $d_k(n) \in \{-1, 1\}$, $n = 0, 1, ..., L-1$ is symbol interval index and L is the length of the data block. $s_k(n) = (s_k(nN+1).......s_k(n+1)N)^T$ with $s_k(i) \in (-1/\sqrt{N}, 1/\sqrt{N})$ is the spreading code of the user k to modulate n^{th} bit. In mobile radio channel, each transmission path encounters temporal and spatial fading [5]. Furthermore, each user is transmitting at a specific power level. In our single path K-user system this corresponds to each user being received with a random, time-dependent amplitude and phase, or equivalently, an arbitrary user k is affected by a random, time dependent complex channel coefficients, $c_k(i)$. The received base band signal can be written as

$$r = \sum_{i=0}^{L-1} \sum_{k=1}^{K} c_k(i) x_k(i) + n \quad (1)$$

$$r = \sum_{i=0}^{L-1} \sum_{k=1}^{K} c_k(i) d_k(i) \begin{pmatrix} 0_{iN+\tau_k} \\ s_k(i) \\ 0_{(L-i)N-\tau_k-1} \end{pmatrix} + n \quad (2)$$

The convenient matrix notation is given by

$$r = SCd + n \quad (3)$$

where data symbol vector is given by $d = (d_1(0), d_2(0), ..., d_K(L-1))^T = (d_1, d_2, ..., d_{LK})^T$ and C is $LK \times LK$ diagonal matrix containing the physical channel parameters. The complex channel coefficients $c_k(i)$ contains all the fading and attenuation effects of the radio channel. S is the matrix of transmitted waveforms with the column j expressed as

$$s_j = \begin{pmatrix} 0_{iN+\tau_k} \\ s_k(i) \\ 0_{(L-i)N-\tau_k-1} \end{pmatrix} \quad (4)$$

A minimal set of sufficient statistics of dimension LK is obtained through correlation, matched to the received signal. This also ensures the maximization of the SNR, i.e,

$$y = S^T r = S^T SCd + S^T n = RCd + z \quad (5)$$

where R is the correlation matrix and z is zero mean Gaussian vector with covariance $\sigma^2 R$.

3 Sphere constrained approximate ML

Given the set of data $y \in R^{LK}$, our goal is to find parameters that miximize the $logP(y|\theta)$ or minimize the negative of it. In iterative parameter estimation, given old set of parameters θ_i we need to find new set of parameters θ_{i+1} that improves the likelihood at each iteration. In our approach, we want the detected data vector to lie close to the sphere, therefore we also require that the new parameter vector to stay "close" to the old set of parameters. In order to achieve it, we encorporate a distance function , which can also be thought as penalty function. We now search for new set of parameters θ_{i+1} that minimizes the distance function summed with the negative loglikelihood function subject to spherical constraint. We will call this function as "augmented loglikelihood". More formally, the update is found by setting $\theta_{i+1} = arg_{min_\theta} L(\theta)$ where

$$L(\theta) = -logP(y|\theta) + d(\theta, \theta_i) + \lambda(d^T d - LK) \quad (6)$$

Lagrange multiplier λ is used to enforce the spherical constraint on symbols. The distance function $d(\theta, \theta_i)$ in our case is KL divergence but other distance function can also be used. The KL divergence is given by

$$d(\theta, \theta_i) = \int_y P(y|\theta) log \frac{P(y|\theta)}{P(y|\theta_i)} dy \quad (7)$$

We approximate the sphere constraint by the first order Taylor expansion around d_i (old parameter set), i.e.

$$d^T d - LK = (d^T d - LK)_{d_i} + (d - d_i)^T \nabla_d (d^T d - LK)|_{d=d_i} \quad (8)$$

where d_i is the value of parameter at iteration i. Substituting equation (7) and equation (8) in equation (6) we get

$$\begin{aligned} &L(\theta) = -logP(y|\theta) \\ &+ d(\theta, \theta_i) + \lambda((d - d_i)^T \nabla_d (d^T d - LK)|_{d=d_i}) \end{aligned} \quad (9)$$

The first order approximation is valid because distance function (penalty function) will force the new parameters to remain close to the old ones at each iteration and hence the estimated vector d will always be close to the surface of the sphere. The KL divergence after bit of algebra can be written in the following form

$$\begin{aligned} &d(\theta, \theta_i) = \tfrac{LK}{2} + \tfrac{1}{2} trace(I) \\ &+ \tfrac{1}{2\sigma^2}(m_\theta - m_{\theta_i})^T R^{-1} (m_\theta - m_{\theta_i}) \end{aligned} \quad (10)$$

where I is idendity matrix and m_θ is mean of the distribution. The above expression is convex function. Plugging in values from the received signal and omitting constant terms give

$$d(\theta, \theta_i) = \frac{1}{2\sigma^2}(RCd - R(Cd)_i)^T (Cd - (Cd)_i) \quad (11)$$

and

$$\begin{aligned} &logP(y|\theta) = \\ &\tfrac{LK}{2} log(2\pi) - \tfrac{1}{2\sigma^2}(y - RCd)^T R^{-1} (y - RCd) \end{aligned} \quad (12)$$

which after permuting C and d gives

$$\begin{aligned} &logP(y|\theta) = \\ &\tfrac{LK}{2} log(2\pi) - \tfrac{1}{2\sigma^2}(y - RDc)^T R^{-1} (y - RDc) \end{aligned} \quad (13)$$

where D is diagonal matrix with diagonal entries given by $(d_1(0), d_2(0) ...... d_k(L-1))$ and $c = diag(C)$ is vector composed of diagonal elements of matrix C. The loglikelihood equation can be further simplified as (after omitting constants)

$$\begin{aligned} &-logP(y|\theta) \\ &= \tfrac{1}{2\sigma^2}(y^T R^{-1} y - y^T Dc - c^T D^T y + c^T D^T RDc) \end{aligned} \quad (14)$$

Taking gradient with respect to c of the above function gives

$$-\nabla_c log P(y|\theta) = \frac{1}{\sigma^2}(-D^T y + D^T RDc) \quad (15)$$

The distance function after permuting C and d is written

$$d(\theta, \theta_i) = \frac{1}{2\sigma^2}(RDc - R(Dc)_i)^T(Dc - (Dc)_i) \quad (16)$$

the subscript i denotes the value of the parameters at i^{th} iteration. Rearranging and taking gradient with respect to c gives

$$\nabla_c d(\theta, \theta_i) = \frac{1}{\sigma^2}(D^T RDc - D^T R(Dc)_i) \quad (17)$$

Putting the above two gradient in the augmented log-likelihood equation and equating the resulting equation to zero gives

$$c = \frac{1}{2}(D^T RD)^{-1}(D^T y + D^T R(Dc)_i) \quad (18)$$

Similarly we take the gradient of the augmented log-likelihood function with respect to d and equating it to zero gives

$$d = (\frac{2}{\sigma^2}C^T RC)^{-1}(\frac{1}{\sigma^2}C^T y + \frac{1}{\sigma^2}C^T R(Cd)_i - 2\lambda d_i) \quad (19)$$

This expression is function of λ, i.e. Lagrange multiplier, which is given by

$$\lambda = \frac{h \pm \sqrt{h^2 - 4gj}}{2g} \quad (20)$$

where

$$h = 4e^T U d_i \quad (21)$$

$$j = e^T U e - LK \quad (22)$$

$$g = 4d_i^T U d_i \quad (23)$$

and $U = X^T X$, $e = a + v$ where X is

$$X = (\frac{2}{\sigma^2}C^T RC)^{-1} \quad (24)$$

,

$$a = \frac{1}{\sigma^2}C^T y \quad (25)$$

and

$$v = \frac{1}{\sigma^2}C^T R(Cd)_i \quad (26)$$

We also calculated the formulas for C and d when euclidean distance function is used instead of KL distance function. The euclidean distance between two parameters set is defined by

$$d(\theta, \theta_i) = \frac{1}{2}||\theta - \theta_i||^2 \quad (27)$$

The euclidean distance function after bit of simplification is written as

$$d(\theta, \theta_i) = \frac{1}{2}(c^T c + d^T d - 2c^T c_i - 2d^T d_i + c_i^T c_i + d_i^T d_i) \quad (28)$$

where the subscript i denotes the value of the parameter at the ith iteration. In our case the parameter set is given by $\theta = (c, d)$, where c is vector composed of diagonal elements of C. With the same procedure as is done for KL distance case, i.e., taking gradient of the distance function with respect to c and d and also taking gradient of the loglikelihood function with respect to c and d and plugging the results into augmented loglikelihood fuction and imposing the spherical constraint.The update equations for c and d are given by

$$c = (\frac{1}{\sigma^2}D^T RD + I)^{-1}(\frac{1}{\sigma^2}D^T y + c_i) \quad (29)$$

where I is identity matrix. Similarly for d, we have

$$d = (\frac{1}{\sigma^2}C^T RC + I)^{-1}(\frac{1}{\sigma^2}C^T y + d_i - 2\lambda d_i) \quad (30)$$

where λ is given by

$$\lambda = \frac{m \pm \sqrt{m^2 - 4ln}}{2l} \quad (31)$$

where $l = 4d_i^T U d_i$, $m = 4d_i^T UW$, $n = W^T UW - LK$ and $U = X^T X$. The expression for X abd W are as follows

$$X = (\frac{1}{\sigma^2}C^T RC + I)^{-1} \quad (32)$$

and

$$W = \frac{1}{\sigma^2}C^T y + d_i \quad (33)$$

The algorithm works as follows, 1) We start with the initial estimate of C_i and d_i, 2) We calculate C (the updated value), the updated value of C is used to calculate λ.These values are inturn plugged into update expression for d to get updated d. These two steps are continued until $||vec(C_{i+1} - C_i)|| < \delta$, where δ is small number. Noting that in the update equations for C and d (in case of KL distance), there are matrix inversions, i.e, we have to invert a matrix at each iteration which is computationally expensive. In the following lines we will derive low complexity algorithm by eliminating matrix inversion. This is done by polynomial expansion of the signature correlation matrix, R, i.e.

$$R^{-1} = (I+Q)^{-1} = \sum_{i=0}^{\infty} (-Q)^i \qquad (34)$$

where Q is equal to matrix R with diagonal elements put to zero and $Q^0 = I$, where I is identity matrix. If the elements of Q are small compared to one i.e. low cross-correlation. The matrix R^{-1} can be approximated by first order expansion (neglecting higher order terms), i.e.,

$$R^{-1} = I - Q \qquad (35)$$

In this way matrix inversion is replaced by adding two simple matrices.

4 Simulations

In this section we investigate the amplitude error and BER performance based on the simulations. The codes were selected at random and we considered two different scenarios. A lightly loaded case with six number of users as well as highly loaded case with, $K = 24$. In both the cases the processing gain was kept 32. We plot figure for the amplitude estimation error versus different values of SNR. As is clear from the figure (4), the estimation error decreases as the value of the SNR increases for both cases. However, estimation error of highly loaded case is more than lightly loaded case. We also simulated for BER for lightly loaded case. It is clear from the figure (1) and figure (3) that our receivers (with KL distance function and euclidean distance function) performs better than MMSE and the receiver proposed in [2] (they have the same performance). The MMSE receiver on average constrain the data vector to lie within sphere [2]. In [2] the authors considered the constraint that the data vector lie within sphere. On the other hand as we approximate the spherical constraint with the first order Taylor expansion but at the same time we do not want previous esitmated data vector to be far from the new estimate because of the distance function that we incorporated in the augmented likelihood function. Therfore we are always close to the sphere. Hence, we can consider our constraint to be shell region between two concentric hyperspheres, which is more constraining than sphere constraint proposed in [2]. We also plot the BER for approximate proposed receiver in figure (2). The approximation is done to reduce the complexity of the receiver.It is hoped that with the increase in the number of users better results are expected owing to the fact that the first order approximation of the sphere will almost lie on the surface of the sphere, i.e. we will be almost on the surface of the sphere. In the figure (2), we also compared the low complexity version of the algorithm with the MMSE. As is clear from the figure that it performs better than MMSE and the performance is almost identical with that of exact proposed receiver. In all the simulations for the BER the estimated values of the amplitudes were used. While in the case of MMSE true amplitudes were used in the simulations.

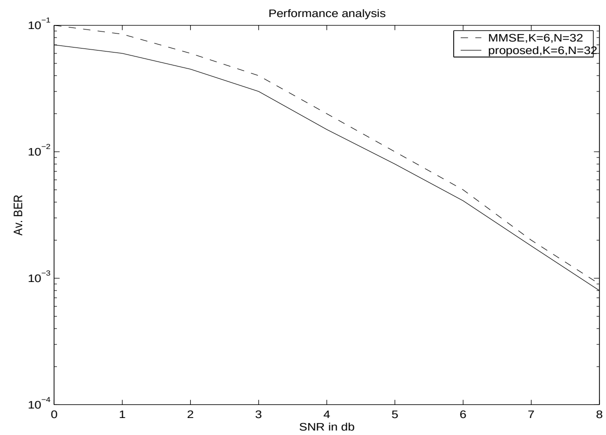

Figure 1: Average BER for MMSE and proposed receiver with KL didtance function

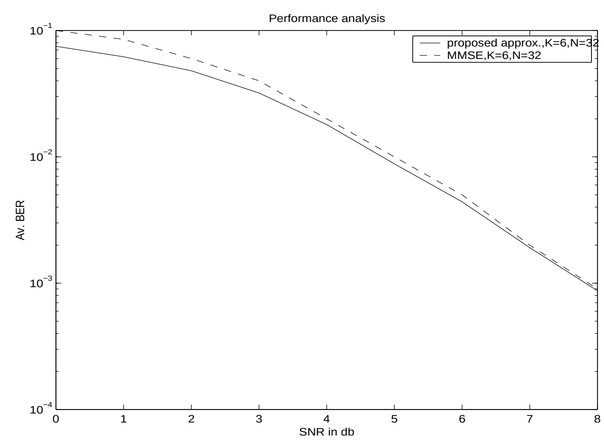

Figure 2: Average BER for MMSE and Approx. proposed receiver

References

[1] S .Talwar, M . Viberg, and A .Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array-Part-1: algorithms," IEEE trans. signal Processing, 44,pp. 1184-1197, 1996

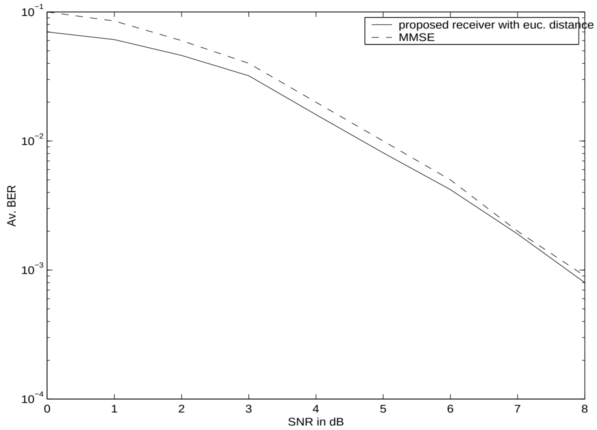

Figure 3: Average BER for MMSE and proposed receiver with euclidean distance function

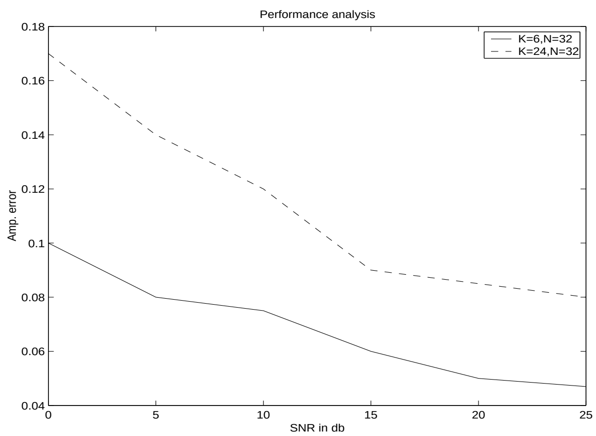

Figure 4: Amplitude error norm for proposed receiver

[2] P. H. Tan, L. K. Rasmussen and T. J. Lim, "Sphere-constrained maximum-likelihood detection in CDMA", in the proceedings of the IEEE Veh. Tech. Cof. (Tokyo, Japan), pp. 517-521, May 2000.

[3] P. H. Tan, L. K. Rasmussen and T. J. Lim, "Box-constrained maximum-likelihood detection in CDMA", in the proceedings of the The Zurich Seminar on Broadband Commun. (Zurich, Switzerland), pp. 55-62, Feb. 2000.

[4] P. H. Tan, L. K. Rasmussen and T. J. Lim, "Constrained maximum-likelihood detection in CDMA", IEEE Trans. Commun.,vol. 49, pp. 142-153, Jan. 2001.

[5] W.C. Jakes, "Microwave mobile communications", IEEE Press Classic Reissue, 1974.

[6] S .Talwar, and A .Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array-Part-11: performance analysis," IEEE trans. signal Processing, March, 1997.