Manuscript received Sep. 2, 2007; revised Dec. 15, 2007

A NARRATIVE APPROACH FOR SPEECH SIGNAL BASED MMSE ESTIMATION USING QUANTUM PARAMETERS

S.SASIKUMAR[1], S.KARTHIKEYAN[1], M.SUGANTHI[2], M.MADHESWARAN[3]
Department of Electronics and Communication Engineering
[1] P.S.N.A. College of Engineering and Technology, Dindigul-624622. INDIA.
[2]Thiagarajar College of Engineering, Madurai, INDIA.
[3]Muthayammal Engineering College, Rasipuram, INDIA.
Email: sonapoppy@yahoo.co.in, skarthik_82@rediffmail.com

Abstract -In this paper, the performance of different estimators in estimating the speech signal through Quantum parameters can be analyzed. The main objective is to estimate the speech signal by a set of linear and Non-linear estimators that are proposed to be efficient in performance. The Minimax mean square error estimator is designed to minimize the worst-case MSE. In an estimation context, the objective typically is to minimize the size of the estimation error, rather than that of the data error as a cause, in many practical scenarios the least-squares estimator is known to result in a large MSE. A comparative analysis between MMSE estimator with other linear and nonlinear estimators can be performed .The analysis proved that the MMSE estimator can outperform both from linear and nonlinear estimator.

Keywords: Minimax mean square error estimator, James-Stein Estimator, Maximum A Posteriori Estimation,, Quantum Constraints,Signal to noise ratio,Weighted least square estimation.

1. Introduction

The development in the field of signal processing is tremendous and quantum signal processing in particular has motivated the rigorous growth and research in the past few decades[1,2 and 3].The estimation using digital signal processing concepts have been the research area in the recent past. The quantum mechanical concepts have been shown more interest in the signal analysis due to its inherent properties [4, 5 and 6]. The introduction of Quantum mechanical concepts, which rely on estimation almost entirely on some of signal processing algorithms, are implemented with various techniques.

In many DSP applications we don't have complete or perfect knowledge of the signals we wish to process. We are faced with many unknowns and uncertainties like noisy measurements and unknown signal parameters [3]. We consider the class of estimation problems represented by the linear model

$$y = Hx + w \tag{1}$$

where x denotes a deterministic vector of unknown parameter, H denotes a known $n \times m$ matrix, and w denotes a zero-mean random vector with covariance C_w. It is well known that among all possible unbiased linear estimators, the LS estimator minimizes the variance [4]. However, this does not imply that the resulting variance or mean-squared error (MSE) is small, where the MSE of an estimator is the sum of the variance and the squared norm of the bias. Various modifications of the LS estimator for the case in which the data model is assumed to hold perfectly have been proposed [5]. Later [6], Stein showed that the LS estimator is inadmissible, i.e., for certain parameter values, other estimators exist with lower MSE. An explicit (nonlinear) estimator with this property, which is referred to as the James–Stein estimator, was later proposed and analyzed [7]. This work appears to have been the starting point for the study of alternatives to LS estimators. Among the more prominent alternatives is the shrunken estimator [8].Stochastic Gaussian Maximum Likelihood (ML) method [9] that deals with sub-Gaussian signals. Because of such uncertainties, the minimum mean-squared error and maximum posteriori estimators [10] cannot be used in many cases. The minimum mean-squared error linear estimator does not require this priori density. We present several analytical and numerical results demonstrating the superiority of minimax estimators over least-squares (LS) estimation. The results are related to and compared with other LS-dominating estimators, such as the James-Stein estimator [11]. A minimax mean-squared error (MSE) estimator is developed for estimating an unknown deterministic parameter vector in a linear model, subject to noise covariance uncertainties. The estimator is designed to minimize the worst-case MSE across all norm bounded parameter vectors and all noise covariance matrices, in a given region of uncertainty. The minimax estimator is shown to have the same form as the estimator that minimizes the worst-case MSE over all norm-bounded vectors for a least-favorable choice of the noise covariance matrix. [12]. In [13], minimum-mean-square-error (MMSE) channel estimation algorithm for OFDM systems is proposed. The algorithm adopts two-dimensional Hadamard transform (TDHT) instead of the conventional Fourier transform, and more noise interference can be filtered with the proposed scheme. In [14] an extended version of the Order Statistic Least Mean Square (OSLMS) algorithm involving ambiguous sorting is developed for speech signals. In [15], the problem of minimum mean-squared error (MMSE) estimation under convex constraints such as constraints on the estimated vector and constraints on the structure of the estimator there exist a simple closed form expression for the constrained MMSE estimator.

To improve the performance of the LS estimator at low to moderate SNR, we propose a modification of the LS estimate, in remainder of this paper is organized as follows: section 2 focuses on least square estimation, section 3 focuses weighted least square estimation and section 4 emphasizes on minimax mean square estimation, linear and nonlinear estimator like James–Stein estimator, shrunken, MAP estimator. In which we choose the estimator of x to minimize the total error variance in the observations y, subject to a constraint on the covariance of the error in the estimate of x.

The resulting estimator of x is derived as the minimax mean square estimation. This implies that for white Gaussian noise, there is no linear or nonlinear estimator with a smaller variance, or MSE, and the same bias as the MMSE estimator. In our method, the LS estimator resulting variance or mean-squared error (MSE) is small, where the MSE of an estimator is the sum of the variance and the squared norm of the bias. In comparison to the previous methods, the algorithm is more computationally efficient and highly parallelizable, which makes the algorithm more attractive for real time applications.

2. Least Square Estimation

Least squares estimation, also known as ordinary least squares analysis is a method for linear regression that determines the values of unknown quantities in a statistical model by minimizing the sum of the residuals (difference between the predicted and observed values) squared. A related method is the least mean squares (LMS) method. It occurs when the number of measured data is 1 and the gradient descent method is used to minimize the squared residual. LMS is known to minimize the expectation of the squared residual, with the smallest number of operations per iteration. However, it requires a large number of iterations to converge. Furthermore, many other types of optimization problems can be expressed in a least squares form, by either minimizing energy or maximizing entropy.

In least squares analysis, the objective is to minimize the following function

$$S = \sum_{i=1}^{n} (y_i - f(\vec{x}_i, \vec{a}))^2 . \tag{2}$$

subject to the parameter vector. The above minimization explains the origin of the name least squares. In regression analysis, one replaces the relation

$$f(x_i) = y_i \quad \text{by} \quad f(x_i) = y_i + \in_i \tag{3}$$

where the noise term ε is a random variable with mean zero. Note that we are assuming that the *x* values are exact, and all the errors are in the *y* values. Then the linear model which when using pure matrix notation becomes

$$Y = X\delta + \in \tag{4}$$

where ε is normally distributed with expected value 0 (i.e., a column vector of 0s) and variance $\sigma^2 I_n$, where I_n is the $n \times n$ identity matrix.

The least-squares estimator for δ is

$$\hat{\delta} = \left(X^T X\right)^{-1} X^T Y \tag{5}$$

and the sum of squares of residuals is

$$Y^T \left(I_n - X(X^T X)^{-1} X^T\right) Y. \tag{6}$$

3. Weighted Least Square Estimation

Weighted least square is a method of regression, similar to least squares in that it uses the same minimization of the sum of the residuals

$$S = \sum_{i=1}^{n} \left(y_i - f(x_i)\right)^2 \tag{7}$$

However, instead of weighting all points equally, they are weighted such that points with a greater weight contribute more to the fit:

$$S = \sum_{i=1}^{n} \omega_i \left(y_i - f(x_i)\right)^2 . \tag{8}$$

In a linear regression context,

$$f\left(x_i\right) = x_x^T \beta \tag{9}$$

then minimizing the weighted least squares

$$S = \sum_{i=1}^{n} \omega_i \left(y_i - x_i^{*T} \beta\right)_1^2 \tag{10}$$

is the same as minimizing the ordinary least squares

$$S = \sum_{i=1}^{n} \left(y_i^* - X_i^{*T} \beta\right)^2 . \tag{11}$$

In matrix notation, the weighted least squares estimator of β is

$$\hat{\beta} = \left(X^{*T} X^*\right)^{-1} X^{*T} Y^* = \left(X^T W X\right)^{-1} X^T W Y, \tag{12}$$

Say we are trying to solve for an over determined system which we can denote as

$$X_a = y . \tag{13}$$

We wish to solve for **a**. The least squares solution to this problem will be

$$a = \left(X^T X\right)^{-1} X^T Y \tag{14}$$

similar to least squares in that it uses the same minimization of the sum of the residuals instead of weighting all points equally, they are weighted such that points with a greater weight contribute more to the fit.

The WLS estimator has the additional property that it minimizes the variance from all *unbiased* estimators. Note, however, that a smaller MSE may be achieved by allowing for a bias. Hence we present our work, which shows MMSE is suitable for this criterion.

4. Minimax MSE Estimation

To develop an estimator that minimizes an objective directly related to the MSE, it is suggested to seek the linear estimator that minimizes the worst-case MSE over all possible values of x, assuming that C_w is known. The minimax estimator for the case in which $x^* T x \le U^2$, where T is an arbitrary positive definite weighting matrix; for $T = I$ the estimator reduces to,

$$\hat{x}_{MX}(C_w) = \frac{U^2}{U^2 + \gamma_0} (H^* C_w^{-1} H)^{-1} H^* C_w^{-1} y \tag{15}$$

where $\gamma_0 = Tr\{H^* C_w^{-1} H)^{-1}$ is the variance of the WLS estimator.

Let denote the unknown deterministic parameter vector in the model $y = Hx + w$, where H is a known $n \times m$ matrix, and W is a zero-mean random vector with covariance C_w. Let Q be an invertible matrix that jointly diagonalizes C_w and HH^*, so that $H = V \tilde{\sum} V^*$ for some unitary matrix V. Then the solution to problem for any convex set U such that $\Delta \succ 0$, is

$$\hat{x}_{MX}(C_w) = \frac{U^2}{U^2 + \sum_{i=1}^{m} \delta_i / \sigma_i^2} V \Sigma^{-1} Z^* Q^{-1} y \tag{16}$$

If C_w is any positive definite covariance matrix then the estimator $\hat{x}_{MX}(C_w)$ can be expressed as a shrinkage of the WLS estimator

$$\hat{x}_{MX}(C_w) = \alpha (H^* C_w^{-1} H)^{-1} H^* C_w^{-1} y \tag{17}$$

with $\Delta = diag(\delta_1, \delta_2, ...\delta_m)$ and Θ is an arbitrary diagonal matrix, and shrinkage factor α

$$\alpha = \frac{U^2}{U^2 + \sum_{i=1}^{m} \delta_i / \sigma_i^2} \tag{18}$$

In particular, choosing $C_w = \hat{C}_w$ where

$$\hat{C}_w = Q \begin{bmatrix} \Delta & 0 \\ 0 & \Theta \end{bmatrix} Q^* \tag{19}$$

where $\hat{x}_{MX}(C_w)$ is the minimax MSE estimator for fixed C_w.

Next, we note that for any $C_w \in \wp$ where $\wp$ is any positive definite covariance matrix of the form

$$\wp = \{C_w = Q \tilde{\Delta} Q^* \,\big|\, \tilde{\Delta} \in U\} \tag{20}$$

$$\sum_{i=1}^{m} \frac{\delta_i}{\sigma_i^2} = Tr\{H C_w^{-1} H)^{-1} \tag{21}$$

which is the variance of the WLS estimator with weight C^{-1}_w.

This discussion leads to find the worst-case covariance matrix $C_w \in \wp$ which maximizes the variance of the WLS estimator and designing a minimax estimator assuming known covariance matrix with $C_w = \hat{C}_w$ .As a final comment, if Q is unitary.

5. James-Stein Estimator

The James-Stein estimator is a nonlinear estimator which can be shown to dominate, or outperform, the "ordinary" (least squares) technique. The James-Stein estimator is given by

$$\hat{\theta}_{JS} = \left(1 - \frac{(m-2)\sigma^2}{\|Y\|^2}\right)Y. \qquad (22)$$

James and Stein demonstrated that the estimator presented above can be used when the variance σ^2 is unknown, by replacing it with the standard estimator of the variance.

6. Maximum A Posteriori Estimation

In statistics, the method of maximum posteriori (MAP, or posterior mode) estimation can be used to obtain a point estimate of an unobserved quantity on the basis of empirical data. It is closely related to Fisher's method of maximum likelihood (ML), but employs an augmented optimization objective which incorporates a prior distribution over the quantity one wants to estimate. MAP estimation can therefore be seen as a regularization of ML estimation.

Assume that we want to estimate an unobserved population parameter θ on the basis of observations *x*. Let *f* be the sampling distribution of *x*, so that *f*(*x* | θ) is the probability of *x* when the underlying population parameter is θ. Then the function

$$\theta \mapsto f(x|\theta) \qquad (23)$$

is known as the likelihood function and the estimate

$$\hat{\theta}_{ML}(x) = \arg\max_{\theta} f(x|\theta) \qquad (24)$$

as the maximum likelihood estimate of θ.

The method of maximum a posteriori estimation then estimates θ as the mode of the posterior distribution of this random variable:

$$\hat{\theta}_{MAP}(x) = \arg\max_{\theta} \frac{f(x|\theta)g(\theta)}{\int_{\Theta} f(x|\theta')g(\theta')d\theta'} = \arg\max_{\theta} f(x|\theta)g(\theta) \qquad (25)$$

The denominator of the posterior distribution does not depend on θ and therefore plays no role in the optimization. Observe that the MAP estimate of θ coincides with the ML estimate when the prior *g* is uniform. MAP estimates can be computed in several ways:

1. Analytically, when the mode(s) of the posterior distribution can be given in closed form. This is the case when conjugate priors are used.
2. By the numerical optimization such as the conjugate gradient method or Newton's method. This usually requires first or second derivatives, which have to be evaluated analytically or numerically.
3. From a modification of an expectation-maximization algorithm. This does not require derivatives of the posterior density.

7. Results and Discussion

The analog input is obtained through the channel 1 at a sample rate of 8000 and duration of 1.25 seconds and number of samples obtained from the speech signal is about 10000. The signal is obtained as a column vector. This column vector is converted into a square matrix. Now Hilbert transform is performed on this matrix so that the numerical values of the signal can be obtained. FFT is performed on the signal so that the spectral values of the signal can be obtained. As the concept of QSP is to be satisfied, now the spectral matrix is being converted into orthogonal matrix using Gram-Schmidt orthogonalization procedure. In the orthogonal matrix, white noise with zero mean and unit standard deviation added to the signal.

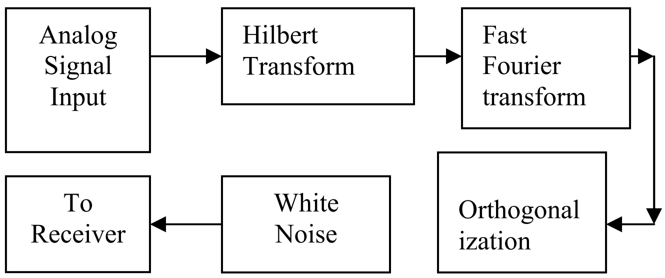

Fig 1. Block diagram for converting speech signal to orthogonal vectors

Input is a continuous speech signal given through microphone. This signal is plotted with its amplitude with respect to time in figure 2. After getting the original speech signal and the noise corrupted speech signal, we estimate the original signal using the estimation techniques.

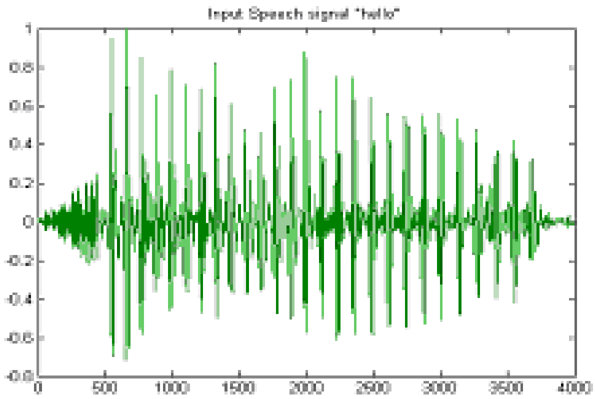

Fig 2. plot for the signal input.

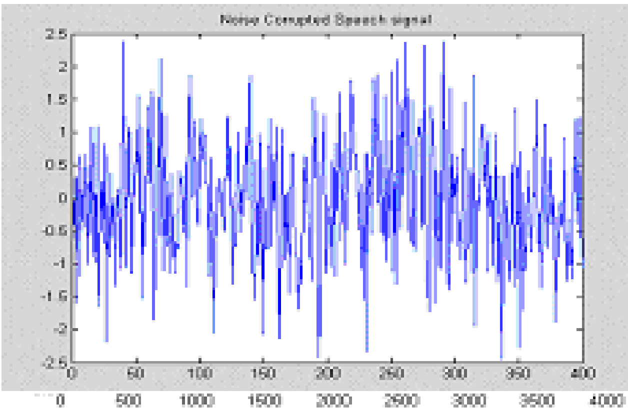

Fig 3. Noise corrupted Input Signal

Fig 4 shows the results of linear estimation of the speech signal which is compared with the result of the minimax mse estimator. Fig 4 and Fig 5 shows the comparison of the mean square error obtained from minimum mean square estimation and the other nonlinear estimators.

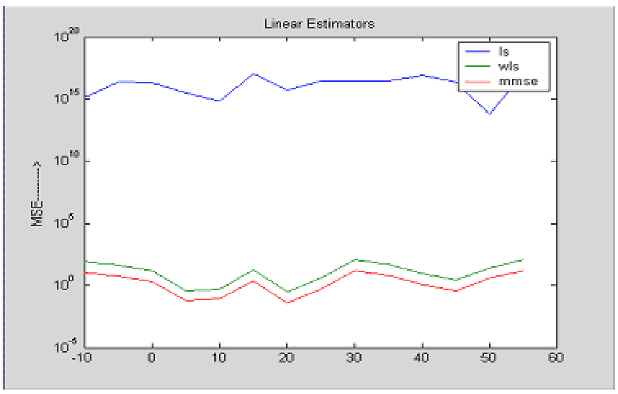

Fig 4. Linear Estimation and MMSE

James-Stein	0.00085	0.0002	3.31	8.809	7.9247
Shrunken	0.16284	0.1053	0.066	0.066	0.0157
MAP	1.2875	1.2374	1.614	1.378	1.3302
MMSE	0.0127	0.0986	0.613	0.001	6.9815

Table 1: MSE of Linear Estimation and MMSE

The plot of mean square error obtained for estimation of speech signal through least square, weighted least square ant the minimax mean square estimator. From the fig 4 we infer that minimax estimator is more efficient than the other two estimators considered.

In the other case we compared the performance of the MMSE with that of the most efficient nonlinear estimators James-stein, MAP and the shrunken estimators. Fig 5 shows the plot of the MSE of the three nonlinear estimators and the MMSE estimator. The plot of the MMSE estimation, from both the estimation schemes is inferred that the proposed MMSE estimation for quantum signal processing is still to be developed to achieve and perform better than the other estimation algorithms proposed in digital signal processing framework.

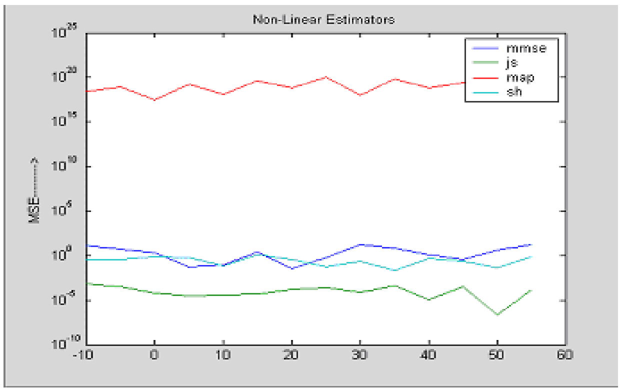

Fig 5. Non-Linear Estimation and MMSE

LS	9.2917	1.6502	1.1317	1.1317	3.4648
WLS	0.1608	0.6386	4.5294	0.01744	50.687
MMSE	0.01271	0.0982	0.6133	0.00153	6.9815

Table 2: MSE of Non-Linear Estimation and MMSE

Conclusion

In this paper, a comparative work can be performed with different estimation schemes in estimating the original speech signal from the noise corrupted signal. Considering the mean square error that result at the end of the estimation process as criteria to analyze the performance of the estimators. Mean square error is calculated from the difference between the true value and the estimated value of each estimator. The comparison of the performance of the minimax estimator with that of the other estimators that are proposed to be efficient in their performance with their own constraints. Furthermore, we explicitly construct estimators with multichannel through a set of linear transformations. As an application of our approach, we also develop Wiener type filters under certain restrictions, which allow for practical implementations. The present report shows that the proposed methods can also be implemented for MMSE multi-user detection for CDMA systems.

References

[1]. S. M. Kay, Fundamentals of Statistical Signal Processing Estimation Theory, Prentice-Hall, 1993.

[2].D.C.Griffiths, Introduction to Quantum mechanics, Prentice-Hall, Inc., 1995

[3].Y.C.Eldar and A.V.Oppenheim, “Covariance shaping Least Square Estimation”, IEEE transactions, Signal Processing, Sep 2001

[4].Y.C. Eldar, Quantum signal processing, Ph.D. thesis, MIT, Cambridge, MA, 2001.

[5].Y. C. Eldar and A. V. Oppenheim, "Quantum Signal Processing," Signal Processing Mag., Nov. 2002 vol. 19, pp. 12-32.

[6]. Y.C. Eldar, A.V. Oppenheim, D. Egnor, Orthogonal and projected orthogonal matched filter detection, IEEE Trans. Signal Process. (2001)

[7]. A. E. Hoerl and R.W. Kennard, “Ridge regression Biased estimation for nonorthogonal problems” Feb. 1970 vol12, pp. 55–67,

[8]. Y. C. Eldar and A. V. Oppenheim, “MMSE whitening and subspace whitening,” IEEE Trans. Inform. Theory, July 2003vol. 49, pp. 1846–1851.

[9]. Don Johnson “Signal Parameter Estimation. The Connexions Project and licensed under the Creative Commons Attribution License Connexions module”, Version 1.4, Aug 18, 2003

[10]. *Onkar Dabeer, Aditya Karnik,”*Consistent signal parameter estimation with 1-bit dithered sampling”, Tata Institute of Fundamental Research University of Waterloo Colaba, India Waterloo, N2L 3G1, Canada

[11]. Z. Ben-Haim and Y. C. Eldar, “Minimax estimators dominating the least squares estimator,” in Proc. Int. Conf. Acoustics. Speech, Signal Processing (ICASSP 2005), vol. 4, Mar. 18–23, 2005, pp. iv53–iv56.

[12] .Yonina C. Eldar, “Minimax MSE Estimation of Deterministic Parameters with Noise Covariance Uncertainties” IEEE transactions on signal processing, vol. 54, Pp-138-145.

[13]. Qihong ge, Liuguo yin, and Huazhong, yang “MMSE Channel Estimation Scheme Based on Two-Dimensional Hadamard Transform for OFDM Systems” Proceedings of the 6th WSEAS International Conference on Multimedia Systems & Signal Processing, China, 2006 (pp198-203)

[14]. Hirobumi Tanaka, Yuichiroh Ohhashi and Tetsuya Shimamura “Adaptive Non-Linear Prediction with Order Statistics for Speech in Impulsive Noise” Proceedings of the 2007 WSEAS Int. Conference on Circuits, Systems, Signal and Telecommunications, Australia, January 17-19,

[15]. T. Michaeli and Y. C. Eldar, "Minimum MSE Estimation with Convex Constraints," IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2007).