D9.4

BOUNDS ON RESTORATION QUALITY USING a priori INFORMATION

H. Joel Trussell and Poorvi L. Vora

North Carolina State University
Department of Electrical and Computer Engineering, Box 7911,
Raleigh, NC 27695

ABSTRACT

A theoretical bound on the improvement possible due to knowledge of signal points, for the case of a gaussian signal in gaussian noise, has been calculated. Wiener and Projections onto Convex Sets restorations have been computed both with and without knowledge of the known points. The estimation errors are compared for different noise levels and blurs, and different numbers of known points.

1. INTRODUCTION

Classical signal restoration techniques (Wiener, Inverse) use little or no a priori information about the original signal. The newer techniques, among them Projection onto Convex Sets (POCS) and Fuzzy Sets, attempt to incorporate more a priori information into the restoration process. The results obtained using these techniques have been promising, though no particular technique is necessarily the most efficient possible with respect to utilisation of prior information.

Given some a priori knowledge about the signal, there will be a bound on the maximum possible improvement of the restoration due to utilisation of this knowledge. We have attempted to study the effect of complete knowledge of a portion of the signal vector.The theoretical improvement when the signal and the noise both have Gaussian pdfs is computed.

2. THE THEORETICAL BOUND ON THE IMPROVEMENT.

Consider the linear model:

$$g = Hf + n \tag{1}$$

where g and n, representing observations and observation noise respectively, are vectors of length M; f is the original signal of length N; and H, of dimension M x N, is a matrix representing "blurring" of the signal.

Let the signal have a Gaussian pdf with mean $\bar{f}$ and covariance K; let the noise be zero mean and let its pdf be Gaussian, with noise covariance matrix, K_n.

The problem may be stated thus:

Using the mean square error criterion, determine the improvement achieveable in the estimate if a portion of the signal vector is known. This problem is one of estimating the parameters of a multiple parameter signal, while knowing exactly some of them.

The Cramer-Rao bound will give an expression for the minimum mean square error possible using an unbiased estimate. For the case of a Gaussian a posteriori density, the bound is achieved by the MAP estimate, though this estimate may be biased [1, pg. 84].

For the case without knowledge of parameters:

$$f = [H^t Q_n H + Q_f]^{-1} [H^t Q_n g - Q_f \bar{f}] \tag{2}$$

where Q_n and Q_f are the inverses of the covariance matrices K_n and K respectively.

This is also the mmse(Wiener) estimate.

As the MAP estimate achieves the C-R bound, its mean square error (and hence the minimum mean square error) is given by the C-R bound [1, pg.80].

$$J^o = H^t Q_n H + Q_f \tag{3}$$

J^o is the Fisher information matrix. The total estimation error, e^2, is given by the trace of $J^{o\,-1}$, and under the proper assumptions may be written as:

$$e^2 = trace\ [H^t Q_n H + Q_f]^{-1} \tag{4}$$

$$e^2 = \sum_{i=1}^{N} \frac{1}{\frac{1}{\kappa_i^2} + \frac{\lambda_i}{\sigma_{ni}^2}} = \sum_{i=1}^{N} \frac{\kappa_i^2 \sigma_{ni}^2}{\kappa_i^2 \lambda_i + \sigma_{ni}^2} \tag{5}$$

where κ_i^2 is an eigenvalue of the matrix K, λ_i are eigenvalues of HH^t (and $H^t H$), and $\sigma_{n_i}^2$ are eigenvalues of Q_n. We have assumed $H^t Q_n H$,Q_f and $H^t H$ are simultaneously diagonalisable, and so are Q_n and $H^t H$. This is not a trivial assumption, and is obviously not true of most pairs of matrices. Note that if they commute, they are simultaneously diagonalisable [2, pg.207]. Note also, that if Q_f is diagonal, then they are simultaneously diagonalisable.

2.1 Estimating the signal vector when part of it is known.

Consider the effect of knowledge of part of the signal vector on the above error expression. Let:

$$f = [\, f_1^t \mid f_2^t \,]^t$$

where f_1 is the unknown part, and f_2 the known part, of f. Denote the lengths of f_1 and f_2 by n_1 and n_2 respectively, and note that $N = n_1 + n_2$.
Let:

$$H = [\, H_1 \mid H_2 \,]$$

where H_1 is M x n_1, and H_2 is M x n_2.
Then, from (1):

$$g = H_1 f_1 + H_2 f_2 + n \tag{6}$$

i.e.,

$$g_1 = H_1 f_1 + n = g - H_2 f_2 \tag{7}$$

This restoration problem is now one of constrained optimisation,its solution is obtained by solving:

$$\frac{\partial}{\partial f_1}[\ln p_{g_1 \mid f_1}(g_1 \mid f_1) + \ln p_{f_1 \mid f_2}(f_1 \mid f_2)] = 0 \tag{8}$$

Equation (8) represents the MAP estimate of f_1 given g_1 and f_2. It is different from the MAP estimate of f_1 given just g_1 as knowledge of f_2 will, in general, change the pdf of f_1, unless the two are independent. Consider the partitioning of K:

$$K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}$$

K_{11} is n_1 x n_1 and represents the covariance matrix of f_1. K_{22} is n_2 x n_2 and represents the covariance matrix of f_2. K_{12} is n_1 x n_2 and K_{21} is n_2 x n_1. As K is symmetric, $K_{21} = K_{12}^t$.

Now, $p_{f_1 \mid f_2}(f_1 \mid f_2)$ is Gaussian as f_1, f_2, and f are Gaussian. It is completely characterised by the mean vector and the covariance matrix, and,

$$p_{f_1 \mid f_2}(f_1 \mid f_2) = N(\overline{f_1} + K_{12}K_{22}^{-1}(f_2 - \overline{f_2}),$$

$$K_{11} - K_{12}K_{22}^{-1}K_{21}) \tag{9}$$

$$K_1 = K_{11} - K_{12}K_{22}^{-1}K_{21}$$

As the model is linear; the noise n, and the signal f_1, both Gaussian; the a posteriori density is also Gaussian. Hence the MAP estimate is the optimum mmse estimate as we observed earlier. Its error, e_1^2, may be obtained from (8), and on comparing (1) and (7),

$$e_1^2 = trace\ [H_1^t Q_n H_1 + Q_1]^{-1}$$

where $Q_1 = K_1^{-1}$.
Rewriting (5):

$$e_1^2 = \sum_{i=1}^{n_1} \frac{1}{\frac{1}{\delta_i^2} + \frac{\gamma_i}{\sigma_{ni}^2}} = \sum_{i=1}^{n_1} \frac{\delta_i \sigma_{ni}^2}{\delta_i \gamma_i + \sigma_{ni}^2} \tag{10}$$

where γ_i are eigenvalues of $H_1^t H_1$ and $H_1 H_1^t$, and δ_i are eigenvalues of K_1. Here we have assumed that K_1 and $H_1^t H_1$ are simultaneously diagonalisable, and that Q_n and $H_1 H_1^t$ are simultaneously diagonalisable, too. Further,

$$\delta_i = \tau_i^2 - \frac{\eta_i}{\rho_i^2} \tag{11}$$

Here we have assumed that:
i) $K_{12}K_{22}^{-1}K_{12}^t$, K_1 and $K_{12}K_{12}^t$
ii) K_{22}^{-1} and $K_{12}^t K_{12}$
are simultaneously diagonalisable. This is quite a rigid requirement, and may not always be met. τ_i^2 are eigenvalues of K_{11}; η_i are eigenvalues of $K_{12}^t K_{12}$; and ρ_i^2 are eigenvalues of K_{22}. From equations (10) and (11),

$$e_1^2 = \sum_{i=1}^{n_1} \frac{1}{\frac{1}{\tau_i^2 - \frac{\eta_i}{\rho_i^2}} + \frac{\gamma_i}{\sigma_{ni}^2}} \tag{12}$$

If the correlation between f_1 and f_2 is zero, we would still expect an improvement due to the fact that we now have M equations for n_1 parameters, which will help us reduce the uncertainty in estimating f_1. This should be reflected in λ_i and γ_i, and will not be noticed in the high noise case, as all the observations will be too noisy to make any difference-- the high noise case represents the influence of only correlation on error reduction.

Thus, clearly, the error, e_1^2 can only decrease with correlation. As e^2 is independent of the correlation, the non-trivial improvement can only increase with correlation. Thus, analysis of the error for the uncorrelated case will provide a lower bound on the best possible improvement. It will also provide an insight into the isolated case of error reduction due to a larger number of observations. There exist a number of problems in image and signal restoration where the known values are uncorrelated with the unknown values. Hence, this case merits further consideration.

2.2 The Difference in the Error Covariance Matrices for the Uncorrelated Case.

Consider a finite length point spread function (psf), which is shift-invariant and monotonic non-increasing. Such a psf is represented by a Toeplitz matrix H, where the diagonals do not increase away from the main diagonal. The aim is to obtain a measure of the improvement in the estimation of each unknown parameter given each is uncorrelated with the known parameters.

Following the same method that produced equation

(5), we obtain the Fischer information matrix for the constrained case:

$$J = H_1^t Q_n H_1 + Q_{11} \qquad (13)$$

where H_1 is defined in (7) and Q_{11} is the upper $n_1 \times n_1$ sub-matrix of Q_n.

Assume that both f and n represent i.i.d. variables, with variances σ_f^2 and σ_n^2. This is reasonable as the purpose here is to study the effect of the psf, and to observe the error difference and its relation to the psf and the number of known points; these assumptions are not expected to affect the conclusions.

The error covariance matrix may be segmented:

$$F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} = J^{-1}$$

F_{11} is $n_1 - (p-1) \times n_1 - (p-1)$. F_{22} is $p-1 \times p-1$. F_{12} is $n_1 - (p-1) \times p-1$ and F_{21} is $p-1 \times n_1 - (p-1)$.
To relate this error to that of the original problem we can write:

$$F^o = F + \begin{bmatrix} F_{12}XF_{12}^t & F_{12}XF_{22} \\ F_{22}XF_{12}^t & F_{22}XF_{22} \end{bmatrix} \qquad (14)$$

where

$$X = [R^{-1} - F_{22}]^{-1}$$

where R is a submatrix of $BC^{-1}B^t$,

$$B = H_1^t Q_n H_2$$

$$C = H_2^t Q_n H_2 + Q_{22}$$

and the difference is given by:

$$\Delta F = \begin{bmatrix} F_{12}XF_{12}^t & F_{12}XF_{22} \\ F_{22}XF_{12}^t & F_{22}XF_{22} \end{bmatrix} \qquad (15)$$

Each diagonal element of ΔF will represent the difference in the error in estimating the corresponding point without and with the knowledge. Because of the finite length psf, the observations corresponding to some points (call them the closer points) will be directly influenced by the known points. The error reduction at these points is expected to be considerable. Error reduction for the other (farther) points will be due only to error reduction at the closer points and will depend on the cross covariance terms. Hence, the error difference matrix for the closer points will depend only on the error covariance among the closer points (F_{22}) , as is indicated by the form of the lower right submatrix of ΔF. Also, the error difference matrix for the farther points depends on the cross covariance terms (F_{12}), as is indicated by the form of the upper left submatrix of ΔF.

It can be shown that ΔF is positive semi-definite. This implies that the error difference matrix is positive semi-definite, and that,as expected, there will not be an increase in error with knowledge.

Implementing the constrained Wiener would be normally done by assuming circularity of the matrices involved, but the assumption that both a matrix and a submatrix are circulant, is unreasonable. Hence, computationally tedious matrix multiplications and inversions need to be performed.

3. PROJECTIONS ONTO CONVEX SETS.

The technique of Projections onto Convex Sets (POCS) has been used successfully to obtain an iterative solution to the restoration problem when the signal is known, *a priori* , to satisfy constraints, each of which may be expressed as a closed convex set [4]. The set of signals with known points is convex, and hence POCS may be used to obtain a solution to this problem.

Assume the original signal is known, a priori, to belong to the intersection of m closed convex sets, each representing a constraint, i.e.

$$f \in C_0 = \bigcap_{i=1}^{m} C_i$$

C_i closed, convex. i = 1,2...m

$$f_{k+1} = P_{C_m} P_{C_{m-1}} \ldots P_{C_1} f_k$$

converges weakly to some $f_0 \in C_0$ if C_0 non-empty in an infinite dimensional space, where P_{C_i} is the projection onto C_i. The convergence is strong in a finite dimensional space [5]. The iterations should be terminated when the estimate is in the intersection of the sets.

The set of signals with known points is convex, and implementation of this constraint is extremely simple using POCS. Hence, though the POCS, being non-linear, is expected to have a larger error than the Wiener for the gaussian signal in gaussian noise, it would be worthwhile to compare the two cases.

4. EXPERIMENTAL VERIFICATION.

The theoretical bound for different signal types with different blurs and different noise levels was computed for a 32 point signal. Wiener and POCS restorations were performed on sample signals and the results compared with the bounds. Three kinds of signals: gaussian, binary (or pseudorandom)[8] and structured, all zero mean and of variance 1.0, were used. Each was degraded with gaussian blur of standard deviation 1 and 3, and noise levels of 20 and 40 dB. The error values are averages over 3 noise sequences per signal, 1 signal in the binary case, and 3 in the gaussian case. The sets of known signal mean, known signal variance, known noise mean and variance, and known points were used for the POCS restoration. These sets, except the known points set, are probabilistic, and as densities are gaussian, the three-sigma bounds are used to define the set [3]. The known signal mean , known signal variance, and known noise mean sets are not expected to affect the restoration greatly. In addition, for the pseudorandom signal, the set of known bounds was also used.

As this knowledge cannot be used by the Wiener, the POCS restorations for this signal are expected to be about as good as, if not better than, the Wiener restorations. the implementation of the Wiener filter uses matrices in order to exactly use the a priori knowledge. Circularity of the original blur matrix was assumed, but as it cannot also be assumed for the partitioned matrix, matrix inversions were performed. Restorations were performed with and without knowledge of 2 and 6 points, the known points being at the end of the signal. Mean square point estimation errors were compared.

Results are shown in tables 1-3.

Table 1
Mean square error.
Gaussian signal, SNR 40dB, std. deviation of blur 1.

Known points	POCS	Wiener	Theoretical
0	0.194	0.058	0.042
2	0.161	0.046	0.035

Known points	POCS	Wiener	Theoretical
0	0.169	0.051	0.042
6	0.155	0.041	0.033

Table 2
Mean square error.
Pseudorandom signal, SNR 20dB, std. deviation of blur 3.

Known points	POCS	Wiener	Theoretical
0	0.83	0.82	0.78
2	0.81	0.76	0.77

Known points	POCS	Wiener	Theoretical
0	0.78	0.75	0.78
6	0.75	0.70	0.75

Table 3
Fractional improvement (error difference as a fraction of original error).
Gaussian signal, SNR 40dB, 2 known points.

Std deviation of Blur	POCS	Wiener	Theoretical
1	0.17	0.2	0.17
3	0.02	0.04	0.03

The theoretical error difference is not a monotonic decreasing function of distance from known points, though it follows a distinct pattern. The error difference of the Wiener follows a similar pattern, while the POCS error behaves more randomly. Plots of all three error patterns are shown in figures 1-3.

5. CONCLUSIONS.

An improvement was observed in both the POCS and Wiener restorations with an increase in the number of known points. The mean square error of the POCS restoration is far from that of the Wiener for the Gaussian signal. POCS performs much better for the pseudorandom signal, but even so, it does not outperform the Wiener. Here it is worth noting that incorrect spectral information can degrade the Wiener to quite an extent. The results obtained are optimal for the Wiener. The fractional improvement (the error difference as a fraction of the original error) decreases as the blur worsens.

For low noise and low blurring, there is not much improvement in the POCS restoration with knowledge.

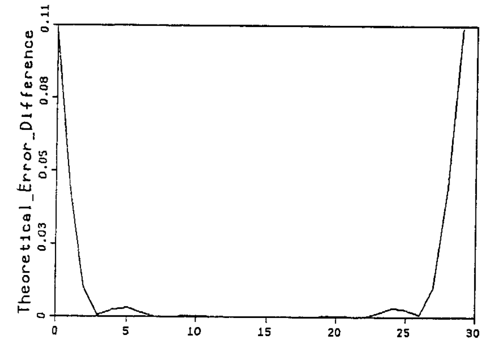

Fig._1_Point_Number

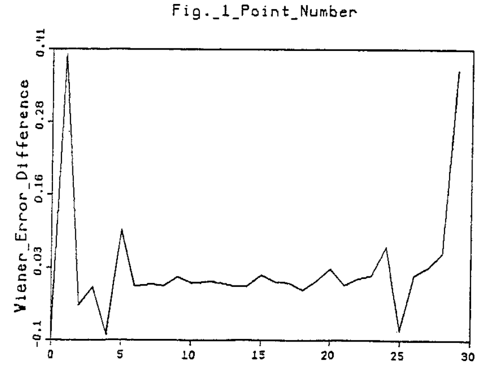

Fig._2._Point_Number

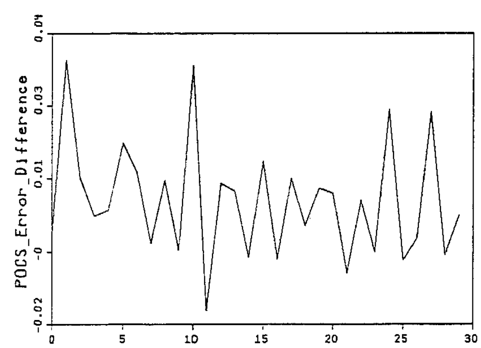

Fig_3_Point_Number

REFERENCES.

1. Van Trees, H. L., Detection, Estimation, and Modulation Theory: Wiley, 1968.
2. Hoffman, K. and Kunze, R., Linear Algebra, NJ: Prentice-Hall, 1971.
3. Trussell, H. J., and Civanlar, M. R. "The Feasible Solution in Signal Restoration," IEEE Trans. on ASSP Vol. ASSP-32, No. 2, pp. 201-212, April 1984.
4. Youla D. C., and Webb, H."Image restoration by the method of convex projections: Part 1-theory,"IEEE Trans. Med. Imag., Vol. MI-1,NO 2, pp. 81-94, Oct. 1982.
5. MacWilliams F. J. and Sloane N. J. A., "Pseudo-random Sequences and Arrays", IEEE Proceedings, Vol. 64, no. 12, Dec. 1976.