COMPUTATIONAL ASPECTS OF CONSTRAINED ESTIMATION. (U)

MAR 82  L M CABRAL

UNCLASSIFIED  NUSC/TR-6068
Computational Aspects of Constrained Estimation

L. M. Cabral
Combat Control Systems Department
PREFACE

This investigation was conducted under NUSC IR/IED Project No. A75930, "Applications of Regularization Techniques to Problems of ASW," principal investigator—L.M. Cabral (Code 3513).

The technical reviewer for this report was E.J. Hilliard, Jr. (Code 351).

REVIEWED AND APPROVED: 1 MARCH 1982

W. A. VonWinkle
Associate Technical Director for Technology
Hybrid estimation strategies can exploit both deterministic and statistical a priori and (more generally) exogenous information for the solution of improperly stated problems. Computationally, problems remain improperly stated due to numerical round-off or algorithmic non-computability. A compact quadratic optimization procedure is presented that resolves this problem, as well as decreases computation time.
# TABLE OF CONTENTS

- **INTRODUCTION** .......................................................... 1
- **GENERAL LINEAR PROBLEM** ........................................... 3
- **GENERALIZED INVERSE SOLUTION** ................................... 7
- **GENERALIZED CONSTRAINED ESTIMATION** .......................... 9
  - Regularization ............................................................ 10
  - Quadratic Optimization Methods ................................... 12
- **RECURSIVE ESTIMATES** ................................................ 17
- **SUMMARY** ................................................................... 19
- **REFERENCES** .................................................................. 20
COMPUTATIONAL ASPECTS OF CONSTRAINED ESTIMATION

INTRODUCTION

The representation of physical systems by logical-mathematical models and the subsequent realization of these models as computational programs dominates engineering and scientific activities. Generally, a model characterizes a cause-effect relationship as a mapping of a set $C$ into a set $E$ ($\text{Im}: C \rightarrow E$). Problems concerned with such mappings are variously classified depending upon the information available and the information sought. The direct (normal or analysis) problem seeks to generate the set of effects from a set of causes, while those problems concerned with reversing the cause-effect relationship are termed inverse or indirect problems. Synthesis (identification) problems require the determination of the laws or mappings that govern cause-effect relationships.$^2$ A fundamental concern is the awareness of conditions for which solutions to such problems are revealed.

A problem is said to be "well-posed" in the sense of Hadamard$^*$ if a unique solution exists and arbitrary small changes in model parameters lead to correspondingly small variations in the solution.$^{3,4}$ This latter condition of stability is significant considering that it is quite standard to assure the existence and uniqueness of solutions and yet accommodate the possibility of results having little physical relevance. The viewpoint is maintained that physical systems that lead to problem statements having potentially unstable solutions are not unusual and that methods for the appraisal and resolution of the problem exist and add little to the computational burden of the problem.

$^*$A further characterization of Hadamard's conditions is given by Rutman$^7$ where for $A$, a linear map in Banach space, $b \in \text{Im}A$ (existence) and $\ker A = 0$ (uniqueness) imply an algebraic well-posedness, and $\text{Im}A = \overline{\text{Im}A}$ (continuous dependence of the solution) implies a topological well-posedness.
Inverse problems comprise the majority of those problems that may be considered ill-posed, viz., unique solutions that are unstable. Usually, the analysis of direct problems of quadrature and problems of synthesis are easily reformulated as inverse problems. Examination of the causes of ill-posedness suggests the following classifications: problems that are shown to be ill-posed as stated (e.g., by arguments of the Riemann-Lebesque lemma); problems having parameters contaminated by noise; and problems for which computational solutions are sought and thus vitiated by errors of truncation. Most often, ill-posedness results from combinations of these factors, which in some cases provide illusions of reliable analysis, but in all cases compound initial difficulties.

Instances of ill-posed problems are frequently encountered in applied situations, numerous examples of which can be found. Often, these applications require solving systems of simultaneous linear and non-linear equations, ordinary and partial differential equations, and integral equations. Of contemporary interest are industrial applications that include inverse scattering problems, ocean acoustic tomography, seismology, inverse wave propagation, image restoration (typically in problems requiring restoration beyond the diffraction limit), mathematical optimization, data query, and optimal control. A common subproblem of continuing mathematical interest is the solution of the general linear problem, which is seen to be a finite-dimensional representation of the Fredholm integral equation of the first kind. In the case of non-linear problems, linearization techniques or other innovations are often successful in reducing the problem to a linear equivalent. A main consideration is the investigation of computational methods appropriate for the solution of linear ill-posed problems and the application of methods of constrained estimation.
Problems for the solution of Fredholm integral equations of the first kind, that is,

\[ \int_{\eta_1}^{\eta_2} K(t,s)f(s)ds = g(t)(\xi_1 \leq t \leq \xi_2) \]  

(1)

occur frequently whenever input data \( f(s) \) are to be determined from measurements \( g(t) \) obtained from the output of a device or instrument; the kernel function \( K(t,s) \) categorizes the measurement process. As a practical example, consider the problem of estimating the acoustic field \( N(\phi) \), from beam measurements \( M(\gamma) \) of a line array consisting of \( k \) equally spaced elements. This problem results in the expression

\[ M(\gamma) = \int_{0}^{2\pi} \frac{\sin^2 \left\{ k\pi \xi / \tau [\cos(\phi - \psi) - \cos(\gamma)] \right\} N(\phi)d\phi, }{k\sin^2 \left\{ \pi \xi / \tau [\cos(\phi - \psi) - \cos(\gamma)] \right\} } \]

where \( \phi \) is the azimuth measurement, \( \gamma \) corresponds to the beam steering direction, \( \psi \) is the array heading, \( \tau \) is the wavelength of the input, and \( \xi \) is the array element spacing. For continuous kernels, this problem is ill-posed.

Algebraization of (1) proceeds by application of an appropriate rule of quadrature \( w_j \), which results in the approximation

\[ \sum_{j=0}^{m} w_j K(t_i,s_j)f(s_j) = g(t_i) \quad (i = 1,2,\ldots,n). \]

A typical selection of abscissae \( s_j \) is given by \( s_j = a + h(j - 1) \), where the displacement \( h = (\eta_2 - \eta_1)/(m - 1) \). For a Simpson's rule of quadrature,
\[ w_j = \begin{cases} 
\frac{h}{3} & \text{(for } j = 1 \text{ or } m) \\
\left[3 + (-1)^j\right]\frac{h}{3} & \text{(otherwise)} 
\end{cases} \]

and \( m \) is odd. There then remains the selection of the quantization parameters \( n \) and \( m \), which are chosen to be small enough to reduce the degree of numerical truncation, but large enough to assure adequate resolution and representation of the physical process. In Hunt \cite{Hunt1984} this dilemma, encountered in the naive solution of equation (1), is demonstrated and further characterizes the ill-posedness of such problems.

A compact representation of a finite-dimensional linear system is provided by a matrix expression, which for (1) results in

\[ Ax = b. \] (2)

Conditions for the existence of solutions to (1) follow by Picard's criteria.\cite{Picard1896} Alternatively, equation (2) admits a solution if and only if the coefficient matrix \([A]\) and the augmented matrix \([A:b]\) are of equal rank; i.e., the vector \( b \) is a linear combination of the columns of \( A \), in which case the linear system is consistent. If, in addition, the matrix \( A \) is square and of full rank, then the solution to (2) is unique and can be generated by application of Cramer's rule (or the inversion of \( A \)). For an overdetermined linear system, unique "solutions" are provided by transforming equation (2) into a system of normal equations

\[ A'Ax = A'b, \] (3)

which is seen to be mathematically equivalent to the linear least-squares problem.

But Cramer's rule is applicable only if the determinant \( \det(A) \) of the square coefficient matrix \( A \), in (2), is non-zero. In situations where the determinant is zero, the matrix is said to be non-regular or singular. In linear systems having small determinants, \( A^{-1} \) contains elements of large amplitude that amplify small variations in the measurement vector \( b \).\cite{Hunt1984}
This deterioration in the conditioning of the linear system (ill-conditioning) is seen as the numerical manifestation of ill-posedness and is manifested by increasing interdependence among the rows of the coefficient matrix or smoothness of the kernel in equation (1). Symmetrization only exacerbates this condition since, for A (a square matrix), $\det(A'A) = \det(A)^2$. Adequate appraisal of conditioning is then seen as prerequisite to the generation of solutions to linear systems.

A useful method of gaging the sensitivity of a linear system is provided by the relative condition number $K$ where

$$K = \frac{\|A\| \|A^{-1}\|}{\|A^{-1}\|}.$$  \hspace{1cm} (4)

Typically, the norm of the real square matrix $A$ corresponds to the maximum of the sums of the moduli in the rows of $A$. The matrix $A$ represents a linear mapping, or transformation, of an arbitrary vector $x$ into the vector $Ax$. The norm of $A$ is then a measure of the distortion under the linear transformation. By a simple perturbation analysis of equation (2),

$$\frac{\|\delta x\|}{\|x\|} \leq K \frac{\|\delta b\|}{\|b\|},$$

by which for large $K$ it is seen that small relative variations in $A$ and $b$ are magnified in $x$, a restatement of the condition of ill-posedness. However, determining equation (4) is complicated by the uncertainty in the estimate of $\|A^{-1}\|$ when $K$ is large.

Alternatively, the condition number can be determined by the square root of the ratio of the largest eigenvalue of $AA'$ to its smallest eigenvalue. It is then apparent that ill-conditioning is associated with eigenvalues close to zero. A dominant eigenvalue $\lambda_m$ of $D = AA'$, where $D$ is similar to a diagonal matrix, can be determined from the sequence

$$w_m = Dw_{m-1}/\lambda_m,$$

where $w_0$ is an arbitrary vector and $\lambda_m$ is the maximum element of $Dw_{m-1}$.  
Similarly, the sequence

\[ u_n = (D - \lambda_n I)u_{n-1}/\lambda_n \]

provides an estimate of the minimal eigenvalue \( \lambda_n \). For non-diagonal type matrices, deflation procedures may be appropriate for determining the minimal eigenvalue.\(^3\)
Often, the occurrence of ill-posed problems in practice motivates the application of multiple precision computations generating costly failures and a skepticism in the promise of computing machinery. Alternative remedies (e.g., preconditioning via scaling), the use of orthogonal (or nearly orthogonal) base functions, large number arithmetic, interval arithmetic, direct search methods, and forward/backward error analysis may provide some insight into the characteristics of the solution, but usually at excessive cost. Even problems of moderate dimensionality can rapidly exhaust available memory capacities as machine truncation errors are compounded. Invariably, information is lost as real numbers (viz., irrational numbers) are mapped onto the sieve-like range space that persists in the computational environment. Similar effects result as measurements become contaminated by noise. Attempts to ameliorate the effects of lost information are frustrated unless additional information is furnished.

Although existence of a solution to equation (3) is assured for an arbitrary coefficient matrix, uniqueness is provided if and only if the coefficient matrix is of full rank. Since the set of all least-squares solutions forms a closed convex set, a unique element can be selected that solves equation (3). An appropriate selection is given by the least-squares solution of equation (2), which is of minimum norm variously denoted by the generalized inverse or pseudo-inverse solution.

For an arbitrary \( m \times n \) matrix \( A \), the generalized inverse is defined by the unique \( n \times m \) matrix \( A^+ \), which satisfies Penrose's lemmas:

\[
\begin{align*}
A^+A^+ &= A^+ \\
AA^+ &= A \\
&&(AA^+)' &= AA^+ \\
&&(A^+A)' &= A^+A 
\end{align*}
\]
Computationally advantageous is the representation of $A^+$ in terms of the singular-value decomposition of $A$, that is,

$$A = UL^{1/2}V'. \quad (5)$$

The matrices $U$ and $V'$ result from columns formed by the eigenvectors of $AA'$ and $A'A$ and $L$ is an $m \times n$ matrix composed of a $k \times k$ ($k = \text{rank}(A)$) diagonal matrix of the corresponding eigenvalues with the remainder zero-filled.\(^{12}\) The generalized inverse is then given by

$$A^+ = V'L^{-1/2}U'. \quad (6)$$

By Lagrange minimization, the least-squares solution of (2) having minimum norm results in the expression

$$A^+u = (A'A + \mu^2I)^{-1}A' \cdot \quad (7)$$

Substituting (5) into (7), and noting that $U$ and $V$ are unitary, results in

$$A^+u = V'L^{1/2}[L + \mu^2I]^{-1}U', \quad \text{which is equivalent to equation (6) as } \mu \text{ approaches zero.}$$

Equation (7) corresponds to the ridge inverse (compare the Levenberg-Marquardt procedure\(^{52}\) and damped least-squares\(^{53}\)) as applied in ridge regression,\(^{54}\) or the constrained estimate of the solution of (2) and is seen to be equivalent to an approximate generalized inverse.\(^{55}\) The ridge estimate may be seen to be a type of weighted average between the input data and supplemental information.\(^{56}\) This idea can be further extended to include varied qualities of information that conform to the anticipated nature of the solution.
GENERALIZED CONSTRAINED ESTIMATION

The application of supplemental information to the solution of ill-posed problems is quite common. For example, Wiener theory requires statistics concerning the signal and noise processes. Such quantitative inputs (a priori information), when available, may not be sufficient for the solution of ill-posed problems; additional information will often be required.

Generally, all relevant factors characterizing the nature of the solution may be described as exogenous information. Because of the variety of information that may influence the problem, generalized solution methods are rare. An early approach by Kreisel seeks smooth solutions to equation (1) utilizing exponential weighting functions that effectively neutralize the severe oscillations prevalent in ill-posed solutions. Moreover, smooth solutions are readily justified as representative of most physical processes.

Similar constraints motivated the development by Philips and Twomey for equation (1) in the form

\[ \min \left\{ \|Ax - b\| + \alpha \phi^2(x) \right\} \tag{8} \]

where \( \phi^2(x) \) corresponds to (side) conditions imposed on the least-squares solution, and \( \alpha > 0 \) is the degree to which such conditions should influence the solution. For smoothness constraints, \( \phi^2(x) = \| \phi \|_2^2 \), where \( \phi \) corresponds to a second-difference approximation to the solution and is of the form

\[
\phi = \begin{bmatrix}
1 & -2 & 1 \\
1 & -2 & 1 \\
1 & -2 & 1 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}.
\tag{9}
\]
A general solution to (8) is given by

$$x = (A'A + \alpha \Omega)^{-1}(A'b + \alpha c) ,$$

where c corresponds to a known (absolute) bias and \(\Omega = \phi'\phi\). The potential of using alternative constraints was also suggested by Twomey (e.g., \((x, \Omega x)\) is equivalent to the variance of the solution where

$$\Omega = \begin{cases} 
(n-1)/n, & i = j \\
-1/n, & i \neq j \quad (1, j = 1, 2, \ldots, n)
\end{cases}$$

and where n is the number of measurements. Constraints in terms of entropy functions can be found in Smith.\(^6\)

REGULARIZATION

The approach indicated in (8) was also independently investigated by Tikhonov}\(^6\) and resulted in his formalizing the method of regularization.* By regularization is meant an adjustment of the initial problem that admits proper solutions. Justification of the method is given in various sources\(^7, 10, 64-66\) where a general form of the smoothing functional in (8) is given by

$$\phi^2(x) = \int_0^1 \sum_{j=0}^p a_j(\varepsilon) \left[ \frac{d^j x(\varepsilon)}{d\varepsilon^j} \right]^2 \, d\varepsilon < \infty .$$

Thus, for \(a_0 = 0, a_1 = 0, a_2 = 1,\) and \(p = 2,\) there results the smoothing functional corresponding to (9) and in a similar manner the estimation of equation (7) is defined for \(a_0 = p_0 = 0.\)

*Also called the Tikhonov-Miller method.\(^62, 65\)
The method of regularization is given a number of statistical interpretations in which the Tikhonov method is supplemented by a Wiener technique or a Bayesian strategy. An estimate for the problem of equation (9) can be given by

\[
x_\alpha = (A'WA + \alpha H)^{-1}A'Wb,
\]

where \( W \) is a weighting coefficient matrix generated from known statistics concerning the measurements \( b \). In Edenhofer and Varah H corresponds to the assumed known covariance matrix of the data \( x \), and \( W \) corresponds to the covariance matrix of the measurements (a particular form of the weighting matrix).

A problem that remains is the selection of the regularization parameter in (8). This formulation is seen to correspond to the Lagrange minimization for the problem:

\[
\min_{x_\alpha} \left\{ \phi(x_\alpha); \|Ax_\alpha - b\|^2 = \|e\|^2 \right\},
\]

where the vector \( e \) represents the measurement error. A suitable \( \alpha \) is then chosen so that the equality in (11) is satisfied. Iterative procedures for determining \( \alpha \) then require an estimate of \( e \). Varah suggests an interactive graphics approach for the solution of the basic problem of equation (1). Graphical display of a problem may be successful because it motivates the analyst to employ perceptual abilities that the alphanumeric display precludes. This approach can be extended to provide a subjective method for approximating \( \alpha \), as well as applying additional regularization criteria.
Direct methods for estimating $\alpha$ are often desirable; but without adequate noise estimates they are less accurate. Labianca\textsuperscript{12} provides an empirical approach where $\alpha$ is selected as a fraction of the maximum singular-value of $A'A$, the fraction 0.0136 being a suitable choice.

**QUADRATIC OPTIMIZATION METHODS**

For many physical problems a non-negative solution (e.g., probability distributions, signal spectra) is required. An $n$-step method for selecting a regularized solution given non-negativity constraints is offered by Turchin.\textsuperscript{75} This basic approach is further exploited by Rutman\textsuperscript{90} to include additional qualities of constraints in a selective manner if required.

Solution of the general linear problem (2) with non-negativity constraints is equivalent to the problem

$$\min_x \{ Z(x) = \frac{1}{2} (x,Dx) - (q,x) : x \geq 0 \} \tag{13}$$

where $D = A'A$ and $q = A'b$. A regularized solution requires that $D = A'WA + \alpha H$ and $q = A'WB$. For the situation where the solution $x^0$ to the unconstrained problem is negative, a proximal non-negative solution $x^m$ to $x^0$ exists** and is determined by the following algorithm:

\begin{itemize}
  \item [\textsuperscript{*}Equivalently,] $Z(x) = \frac{1}{2} (Dx - 2q,x)$.
  \item [\textsuperscript{**}A solution at the origin is always possible.]
\end{itemize}
Algorithm I

Step 0: Given D and $x^0$, then $q = Dx^0$, $S = \{i : x^0_i > 0\}$ and $x^0_i = \max[0, x^0_i]$, $\forall i$.

Step 1: Set $x^r = T^{-1}z$, where $T = [d_{ij}]$ and $z = [q_i]$, $i, j \in S$.

Step 2: If $x^r_i > 0$, $\forall i$, then go to Step 4.

Step 3: Set $S = S\{j\}: r_j = \min[x^P_i/(x^P_i - x^r_i)]$, $x^P_i = x^P_i + (x^r_i - x^P_i)r_j$, and go to Step 1.

Step 4: If $e^i(Tx^r - z) > 0$, $\forall i$, then go to Step 6.

Step 5: Set $S = SU\{j\}: r_j \min[e^i(Tx^r - z)]$, and go to Step 1.

Step 6: Set $x^m = x^r$.

The vector $e^i$ with ith element unity and all others zero is called the unit column matrix.

Additional constraints can be accommodated for equation (13) by introducing the transformation $Rx = x$, where $R$ is non-singular. Equation (13) then becomes

$$\min_{\tilde{x}} \{Z(\tilde{x}) = 1/2(\tilde{x}, \tilde{\delta}x) - (\tilde{q}, \tilde{x}) : \tilde{x} \geq 0\},$$

which is essentially reduced to the constraints of the type of equation (13). The utility of the transformation is demonstrated by examining several transformation constraint matrix pairs $R$ and $R^{-1}$.
Monotonicity constraints often occur in physical problems (e.g., polynomials with positive coefficients and cumulative probability distributions). A monotonically increasing function \( x(c) \) has the property that \( dx(c)/dc \geq 0 \). Choosing

\[
R^{-1} = \begin{bmatrix}
-1 & 1 \\
-1 & 1 \\
& \ddots & \ddots \\
& & -1 & 1 \\
& & & \ddots & 1
\end{bmatrix}
\]

corresponds to imposing the constraint that the first differences of the restored function \( x \) be non-negative. The matrix

\[
R = \begin{bmatrix}
-1 & -1 & \ldots & -1 & 1 \\
-1 & -1 & \ddots & -1 & 1 \\
& \ddots & \ddots & \ddots & \ddots \\
& & & -1 & 1 \\
& & & & 1
\end{bmatrix}
\]

performs the incremental summations of the restored differences.

A similar development leads to unimodality constraints given by the matrices

\[
R^{-1} = \begin{bmatrix}
-1 & 1 \\
& \ddots & \ddots \\
& & -1 & 1 \\
& & & \ldots & 1 \\
& & & & -1 \\
& & & & & 1
\end{bmatrix}
\]

and
where \( n \) is odd and the row consisting of the unitary element corresponds to the mode position.

Convexity constraints are imposed by requiring the second derivative of the restored function to be non-negative. This leads to the matrices

\[
R_{-1} = \begin{bmatrix}
-1 & -1 & \cdots & 1 \\
-1 & \cdots & \cdots & 1 \\
\vdots & \ddots & \ddots & \vdots \\
1 & \cdots & \cdots & -1 \\
1 & \cdots & -1 & -1 \\
\end{bmatrix}
\]

and

\[
R = [r_{ij}] = [r_{ji}] = [-i(n - j + 1)/(n + 1)], \quad i \geq j.
\]

Concavity constraints are developed in a similar manner.

Non-negativity/non-positivity constraints are simply developed by matrices of the type \( R = \text{diag}\{\ldots1,\ldots-1,\ldots\} \). Further selectivity is imposed by limiting the set of active constraints to the intersection of the set of selected coordinates and the set of positive constraints \( S \) in Algorithm I.

After determining \( x^m \), a new value of the regularization parameter \( \alpha^* \) is recomputed to account for the adjustment in the estimate \( x^0 \) due to non-negativity and additional constraints. In Turchin\textsuperscript{61} the method for recomputing the regularization parameter is given by
\[ a^* = a(n^*/n)^3, \]

where \( n^* \) is the number of elements in the estimate \( x^m \) that differ from zero. This new value of the regularization parameter is then used to compute an adjusted estimate \( x^m \) using, for example, equation (12).

It is sometimes observed that, although a finite algorithm is designed to terminate with a solution, in practice convergence to a solution does not occur. Such is the case with Algorithm I. The property is exhibited that a problem may be mathematically well-posed and remain numerically ill-conditioned. Algorithm I will require inversion of matrices of the order up to \( n-1 \) for which truncation errors are prevalent. Because this pathology is unavoidable, it is always desirable to minimize its effects.

The following iterative algorithm is intended to solve the problem of equation (13) while avoiding matrix inversion:

**Algorithm II**

Step 0: Given \( D \) and \( x^0 \), then \( q = Dx^0 \). Set \( \varepsilon = 0 \) and

\[ x_i^0 = 0, \quad \forall i = 1,2,\ldots,n. \]

Step 1: Set \( k = \{ j: q_j/d_{jj} = \max_i[q_i/d_{ii}], \ j \neq \varepsilon \} \).

Step 2: If \( q_k/d_{kk} < 0 \), then go to Step 4.

Step 3: Set \( \varepsilon = k; \) set \( x_i^0 = x_i^0 + q_k/d_{kk} \);

\[ q_i = q_i - d_{ik}q_k/d_{kk}, \quad \forall i; \quad \text{and go to Step 1.} \]

Step 4: Set \( x^m = x^0 \).
Since the iterative solution begins at the origin, a gradient search (Step 1) selects a coordinate providing a feasible solution. A corresponding translation of the coordinate system (Step 3) adjusts the feasible solution to coincide with the origin. All such displacements that do not result in a non-feasible solution are then accumulated to form the optimum solution.

RECURSIVE ESTIMATES

In many applications, computational considerations of memory size or processing time require alternative formalizations of the processing algorithms in order to realize practical solutions to problems. Recursive estimates often provide such efficiencies with little increase in algorithmic complexity. For classical least-squares and the generalized inverse, the recursive procedure will, with minor modification, provide for the deletion of particular data points without recomputation of all data. In control settings, the Kalman filter—a recursive counterpart to the Wiener filter—is often used. Recursive interpretations for the methods of regularization are similarly desirable.

On-line techniques of regularization appropriate for identification and input signal recovery have been previously observed and found to require a redefinition of the method of regularization. A simple approach corresponding to the recursive approximate generalized inverse is developed and is applicable to problems of the type found in Radhakrishna.

An alternative representation for the solution to the problem of (8) is given by the least-squares solution of

$$[A' \cdot \Phi']' [x; 0]' = [b; 0]'$$,
assuming no bias. An additional measurement \( b_{n+1} \) requires a corresponding addition \( A_{(n+1)} = [a_{n+1,1} \ldots a_{n+1,n}] \) to the coefficient matrix. The recursive estimate \( x_{(n+1)} \) is then given by

\[
X_{(n+1)} = x(n) + P_{(n+1)} A'_{(n+1)}(b_{(n+1)} - A_{(n+1)} x(n))
\]

\[
P_{(n+1)} = P(n) - P(n) A'_{(n+1)}
\]

\[
P(n) = (Q'(n) Q(n))^{-1},
\]

where \( Q'(n) = [A'^t \sqrt{\phi}] \).
SUMMARY

Methods of regularization are shown to be extensions of deterministic least-squares estimation. Conventional approaches are limited by the availability of additional information that must be introduced into the problem to produce a solution. Similar limitations hold for methods of regularization; however, these methods have the advantage of utilizing qualitative inputs and relaxing quantitative information requirements. More notably, these methods provide a resource for resolving problems that are mathematically ill-posed and consequently inappropriate for solution by conventional methods.

Other methods comparable with methods of regularization (e.g., the augmented Galerkin method and the singular value decomposition with truncation or damping) are either less effective or prove to be computationally burdensome. Algorithm II and the transformation methods in Rutman provide an indication of the utility of the Philips-Tikhonov-Turchin approach available at minimal expense. Although industrial application of regularization has been generally limited to image restoration, additional developments will broaden the application to include control settings and the solution of nonlinear problems.
REFERENCES


REFERENCES (Cont'd)


REFERENCES (Cont'd)


REFERENCES (Cont'd)


REFERENCES (Cont'd)


REFERENCES (Cont'd)


REFERENCES (Cont'd)


REFERENCES (Cont'd)


REFERENCES (Cont'd)


REFERENCES (Cont'd)


<table>
<thead>
<tr>
<th>Addressee</th>
<th>No. of Copies</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNR (ONR-100)</td>
<td>1</td>
</tr>
<tr>
<td>CNM (MAT-O8T, MAT-O8T1)</td>
<td>2</td>
</tr>
<tr>
<td>NAVSEA (SEA-O03, SEA-63R)</td>
<td>2</td>
</tr>
<tr>
<td>NSWC, Dahlgren</td>
<td>1</td>
</tr>
<tr>
<td>NSWC, White Oak</td>
<td>1</td>
</tr>
<tr>
<td>NRL, Washington, DC</td>
<td>1</td>
</tr>
<tr>
<td>NORDA, Bay St. Louis</td>
<td>1</td>
</tr>
<tr>
<td>DTNSRDC, Bethesda</td>
<td>1</td>
</tr>
<tr>
<td>NADC, Warminster</td>
<td>1</td>
</tr>
<tr>
<td>NOSC, San Diego</td>
<td>1</td>
</tr>
<tr>
<td>NPS, Monterey</td>
<td>1</td>
</tr>
<tr>
<td>APL, Univ. of Washington</td>
<td>1</td>
</tr>
<tr>
<td>ARL, Penn State</td>
<td>1</td>
</tr>
<tr>
<td>DARPA</td>
<td>1</td>
</tr>
<tr>
<td>DTIC, Alexandria</td>
<td>2</td>
</tr>
<tr>
<td>Stanford Univ., Stanford, CA (Attn: T. Kailath)</td>
<td>1</td>
</tr>
<tr>
<td>TASC, McLean, VA (Attn: S. Marple)</td>
<td>1</td>
</tr>
<tr>
<td>R. Streit, Menlo Park, CA</td>
<td>1</td>
</tr>
</tbody>
</table>