DEPARTMENT OF STATISTICS
University of Wisconsin
1210 West Dayton St.
Madison, WI 53706

TECHNICAL REPORT NO. 852

March 1989

Regularization and Cross Validation Methods for Nonlinear, Implicit, Ill-Posed Inverse Problems [1]

by

Grace Wahba

To appear in the Proceedings of the
Seventh Mathematical Geophysics Seminar,
"Model Optimization in Exploration Geophysics"
Andreas Vogel, Ed.
Vieweg, Wiesbaden-Braunschweig

[1]Research supported in part by AFOSR under grant AFOSR-87-0171, and by NASA under contract NAG5-316.

Regularization and Cross Validation Methods for Nonlinear, Implicit, Ill-Posed Inverse Problems

Grace Wahba*
University of Wisconsin-Madison

Cross validated regularization methods for ill posed inverse problems are reviewed. These methods have been extended to the parameter estimation problem for partial differential equations. In the p. d. e. problem one observes discrete, possibly noisy data on the solution, the forcing function, and the boundary values, and wishes to estimate a (distributed) coefficient in the equation. Some directions for further extension of the method are suggested.

1 Introduction

In this primarily survey paper we review some recent developments in regularization and cross validation methods for certain non-linear, implicit, ill posed inverse problems. The nonlinear implicit ill posed inverse problems we are concerned with arise when one has a partial differential equation, either time dependent or steady state, which typically, models some flow. One observes discrete, noisy values of the solution, and possibly, some forcing function, and one wishes to estimate some distributed parameter coefficient

*Supported by the AFOSR under Grant 87-0171

in the equation. This parameter represents some physical quantity like permeability, transmittivity, or, in other contexts, velocity field and mixing coefficients.

The kind of regularization methods we are talking about are also called penalized output least squares. An early basic reference for this method is Kravaris and Seinfeld(1985). O'Sullivan(1986,1987,1988) has expanded on this approach and, among other contributions, in a key step has recently shown how to use an implicit differentiation to allow the use of cross validation methods to choose the usually crucial regularization or smoothing parameter(s). In this paper we will begin by reviewing work in a general, linear, explicit ill posed inverse problem, then go on to survey extensions to mildly nonlinear ill posed inverse problems. We will describe how these methods are extended to certain implicit problems in the estimation of the aforementioned distributed parameter coefficients, and then we will suggest a few directions for the extension of this approach.

2 The explicit linear ill posed inverse problem

There is an extensive literature on explicit linear ill posed inverse problems. We mention only a very few references relevant to the discussion here. An extensive bibliography concerning the approaches discussed here can be found in the forthcoming book Wahba(1989). A good place to enter the literature concerning methods discussed here is O'Sullivan(1986). See also Anderson, deHoog and Lukas(1980), Wahba(1977,1982,1985), and references cited there.

We suppose

$$y_i = \int_{\mathcal{T}} K(t_i, s) f(s) ds + \epsilon_i, \quad i = 1, 2, ..., n \tag{2.1}$$

where y_i is observed, $K(t,s)$ is known exactly, f is to be estimated, and the ϵ_i are independent errors with mean 0 and common, generally unknown variance. Here we will suppose that the ϵ_i's are Gaussian random variables but this assumption has been relaxed in various papers, see, for example Wahba(1985a) and references cited there. f is supposed to be in some Hilbert space $\mathcal{H}$ of real valued functions on $\mathcal{T}$. the natural assumption is that $\mathcal{H}$ is a reproducing kernel Hilbert space (a Hilbert space in which all the evaluation functionals are bounded), but we will not assume that the reader is familiar with these spaces. Of course the assumption that K is known exactly begs some very important practical questions, which we ignore here.

We will let $\{B_l\}_{l=1}^L$ be a set of L basis functions which will be used to approximate f,

$$f \sim \sum_{k=1}^{K} c_k B_k. \tag{2.2}$$

If $\mathcal{T}$ is an interval of the real line, then B-splines (see deBoor(1978)), which are hill-functions, are generally the basis functions of choice. If $\mathcal{T}$ is a compact set in Euclidean d space then tensor products of B-splines, thin plate basis functions (Wahba(1980)), or other basis functions may be used. The type and number of basis functions can be an important issue and is not discussed here. See, for example Nychka et al (1984) for further discussion of this issue. In the method of regularization, f is estimated as the minimizer in $\mathcal{H}$ of

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \int_{\mathcal{T}} K(t_i,s)f(s)ds)^2 + \lambda J(f) \tag{2.3}$$

where $J(f)$ is a suitable penalty functional, usually a seminorm or norm in $\mathcal{H}$. An explicit representation for the minimizer of 2.3 may be found in Kimeldorf and Wahba(1971) provided the reproducing kernel for $\mathcal{H}$ is known. In problems involving large data sets, and in nonlinear generalizations of 2.3 , the approximation 2.2 is substituted into 2.3 which is then treated as an optimization problem in c. If $\mathcal{T}$ is an interval of the real line,

$$J(f) = \int_{\mathcal{T}} (f''(t))^2 dt \tag{2.4}$$

is a popular "smoothness" penalty functional.

In Euclidean d-space both thin plate and tensor product spline penalty functionals are popular. In E^2, the thin plate penalty functional analogous to 2.4 is

$$J(f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (f_{x_1x_1}^2 + 2f_{x_1x_2}^2 + f_{x_2x_2}^2)dx_1dx_2. \tag{2.5}$$

Penalty functionals which are seminorms in a reproducing kernel Hilbert space have a Bayesian interpretation, see Kimeldorf and Wahba(1970), Wahba(1978), which may aid in choosing them in given practical situations.

Substituting 2.2 into 2.3 results in the problem: Find c to minimize

$$\frac{1}{n} \parallel y - Xc \parallel^2 + \lambda c'Jc \tag{2.6}$$

where X is the $n \times L$ matrix with ijth entry

$$\int_{\mathcal{T}} K(t_i,s)B_j(s)dt \tag{2.7}$$

and with some abuse of notation, J is now the $n \times n$ matrix such that

$$c'Jc = J(\sum_k c_k B_k). \tag{2.8}$$

Under the assumption that the intersection of the null spaces of X and J is empty, we have that the minimizer of 2.6, call it c_λ is

$$c_\lambda = (X'X + n\lambda J)^{-1} X'y. \tag{2.9}$$

As anyone who has ever tried to solve a real ill posed inverse problem knows, the result can be quite sensitive to λ. We use the method of generalized cross validation(GCV), (Craven and Wahba(1979), Golub, Heath and Wahba(1979)) to obtain a good estimate of λ from the data. This estimate is computed as follows: Let $A(\lambda)$be the $n \times n$ matrix relating the data vector y to the predicted data vector $\hat{y}$, where

$$\hat{y}_i = \int K(t_i, s) \sum_k c_{\lambda k} B_k(s) ds. \tag{2.10}$$

Then

$$A(\lambda) = X(X'X + \lambda J)^{-1} X'. \tag{2.11}$$

the GCV estimate of λ is obtained as the minimizer of the cross validation function

$$V(\lambda) = \frac{\frac{1}{n} \| (I - A(\lambda) y \|^2}{(\frac{1}{n} Trace(I - A(\lambda)))^2}. \tag{2.12}$$

This estimate has optimality properties for choosing the λ which minimizes the (true) predictive mean square error

$$\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - \int K(t_i, s) f(s) ds)^2. \tag{2.13}$$

See, for example Craven and Wahba(1979), Utreras(1981), Li(1986). Under somewhat general (but not completely general) circumstances it also has favorable properties with respect to the minimization of the true mean square error

$$\int (f(s)ds - f_\lambda(s)ds)^2 \tag{2.14}$$

where f_λ is the estimate of f, see Wahba and Wang(1987). A review of methods for niminizing 2.12, suitable when n is of the order of hundreds and larger, can be found in Gu et al(1988)

3 The nonlinear explicit ill posed inverse problem

Now replace 2.1 by a nonlinear equation,

$$y_i = N_i(f) + \epsilon_i \quad i = 1, 2, ..., n \tag{3.1}$$

where N_i is a nonlinear functional, for example

$$N_i(f) = \int_T K(t_i, s, f(s))ds, \tag{3.2}$$

say. Again approximating f by

$$f \sim \sum_{k=1}^{L} c_k B_k, \tag{3.3}$$

f is estimated as the minimizer of

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - N_i(c))^2 + \lambda c' J c, \tag{3.4}$$

where, with some abuse of notation we let

$$N_i(c) = N_i(\sum c_k B_k). \tag{3.5}$$

For fixed λ, 3.4 may be minimized numerically by a Newton iteration. Let $c^{(l)} = (c_1^{(1)}, c_2^{(l)}, ..., c_K^{(l)})$ be the lth iterate, and

$$N_i(c) \sim N_i(c^{(l)}) + \sum_{k=1}^{N} \frac{\partial N_i}{\partial c_k}(c_k - c_k^{(l)}). \tag{3.6}$$

Let $X^{(l)}$ be the $n \times N$ matrix with ikth entry $\left.\frac{\partial N_i}{\partial c_k}\right|_{c=c^{(l)}}$ and let

$$y^{(l)} = y - \begin{pmatrix} N_1(c^{(l)}) \\ \vdots \\ N_n(c^{(l)}) \end{pmatrix} + X^{(l)} c^{(l)}. \tag{3.7}$$

Then after the lth step $c^{(l)}$ is found to minimize

$$\frac{1}{n} \| y^{(l)} - X^{(l)} c \|^2 + n\lambda c' J c \tag{3.8}$$

and so

$$c^{(l+1)} = (X'^{(l)}X^{(l)} + n\lambda J)^{-1}X'^{(l)}y^{(l)}. \tag{3.9}$$

At convergence, say at the Lth step, we have

$$A^{(L)}(\lambda) = X'^{(L)}(X'^{(L)}X^{(L)} + n\lambda J)^{-1}X'^{(L)}. \tag{3.10}$$

The GCV estimate of λ for nonlinear problems is then the minimizer of

$$V^{(L)}(\lambda) = \frac{\frac{1}{n} \| (I - A^{(L)}(\lambda)y \|^2}{(\frac{1}{n}Trace(I - A^{(L)}(\lambda)))^2}. \tag{3.11}$$

See O'Sullivan and Wahba(1985). Of course, as in most nonlinear problems not everything is guaranteed, and reasonably good starting guesses may be required. Nevertheless good results have been obtained in some practical cases, see O'Sullivan and Wahba.

Linear inequality constraints on the c's obtained from physical considerations of the properties of α may be imposed by using a programming algorithm to minimize 3.4, or rather 3.8 subject to these constraints. The GCV function at convergence is obtained by setting the active constraints to equality constraints and finding the influence matrix $A(\lambda)$ for the solution of the equality constrained quadratic minimization problem. See Villalobos and Wahba(1987) for more details. The solution of an ill posed convolution equation with a positive solution showed dramatic improvement when positivity constraints were included, see Wahba(1982). If there are non physically meaningful minimizers to the unconstrained quadratic optimization problem, the imposition of meaningful inequality constraints can be important.

4 The nonlinear implicit, ill posed inverse problem (system identification)

The key new idea in this Section, which is at 4.14, allows the use of GCV in the system identification problem. The result given here has been adapted from O'Sullivan (1986a). Kravaris and Seinfeld (1985) have proposed the method of 4.4 below, which, adopting the nomenclature of the field, might be called the penalized output least squares method. Another important recent reference is O'Sullivan (1987), where convergence properties of the method are discussed.

The dynamic flow of fluid through a porous medium is modelled by a diffusion equation

$$\frac{\partial u(\mathbf{x},t)}{\partial t} - \frac{\partial}{\partial \mathbf{x}}\left\{\alpha(\mathbf{x})\frac{\partial}{\partial \mathbf{x}}u(\mathbf{x},t)\right\} = q(\mathbf{x},t),\ \mathbf{x}\in\Omega, t\in[t_{min},t_{max}] \tag{4.1}$$

subject to prescribed initial and boundary conditions, for example $u(\mathbf{x},0) = u_0(\mathbf{x})$ (initial condition) and $\frac{\partial u}{\partial w} = 0$ where w is the direction normal to the boundary. Here, if $\mathbf{x} = (x_1,\dots,x_d)$ then $\frac{\partial}{\partial \mathbf{x}} = \sum_{j=1}^{d}\frac{\partial}{\partial x_j}$. Here u is, say, pressure, q represents a forcing function (injection of fluid into the region), and α is the transmittivity or permeability of the medium. If u_0 and q are known exactly, then for fixed α in some appropriate class, u is determined (implicitly) as a function of α. Typically α must be non-negative to the physically meaningful, and sufficiently positive for there to be measurable flow.

The practical problem is, given measurements

$$y_{ij} = u(\mathbf{x}(i),t_j,\alpha) + \epsilon_{ij} \tag{4.2}$$

on u, and the initial boundary functions, and q, estimate α.

We remark that if $\frac{\partial}{0\mathbf{x}}u(\mathbf{x},t)$ is 0 for $\mathbf{x}$ in some region $\Omega_0 \subset \Omega$, all t, then there is no information in the experiment concerning $\alpha(\mathbf{x})$ for $x\in\Omega_0$. Although the algorithm below may provide an estimate for $\alpha(\mathbf{x})$ for $x\in\Omega_0$, in this case the information is coming from the prior, and not the experiment. This is an extremely important practical problem, see e.g. the references in O'Sullivan (1986a) and Kravaris and Seinfeld (1985).

The problem will be solved approximately in the span of a suitable set of N basis functions

$$\alpha(\mathbf{x}) = \sum_{k=1}^{N} c_k B_k(\mathbf{x}),$$

and since α must be non negative, we put a sufficiently large number of linear inequality constraints on $c = (c_1,\dots,c_N)$, that is

$$\sum_{k=1}^{N} c_k B_k(\mathbf{x}) \geq 0 \tag{4.3}$$

for $\mathbf{x}$ in some finite set, so that the estimate is positive. If stronger information than just positivity is known, then it should be used. We seek to find c subject to 4.3 to minimize

$$\sum_{ij}(y_{ij} - u(\mathbf{x}(i),t_j,c))^2 + \lambda c'\Sigma c, \tag{4.4}$$

where $c'\Sigma c = ||P_1\alpha||^2$. For the moment we suppose that u_0 and q are known exactly. Then

$$u(\mathbf{x}(i), t_j, \alpha) \simeq u(\mathbf{x}(i), t_j; c)$$

is a non linear functional of c, but only defined implicitly. If $u(\mathbf{x}(i), t_j; c)$ could be linearized about some reasonable starting guess

$$\alpha_0(\mathbf{x}) = \sum_{k=1}^{N} c_k^{(0)} B_k(\mathbf{x})$$

then the methods of Section 3 could be used to numerically find the minimizing c_λ and to choose λ by GCV.

Given a guess $c^{(l)}$ for c, we would like to be able to linearize about $c^{(l)}$,

$$u(\mathbf{x}(i), t_j; c) \simeq u(\mathbf{x}(i), t_j; c^{(l)}) + \sum_k X_{ijk}(c_k - c_k^{(l)}), \tag{4.5}$$

where

$$X_{ijk} = \frac{\partial u}{\partial c_k}(\mathbf{x}(i), t_j; c)\Big|_{c=c^{(l)}}. \tag{4.6}$$

If this could be done, then c and λ could be determined, at least in principle, via the constrained Gauss Newton iteration and the GCV procedure described in Section 3.

Let

$$L_c = \frac{\partial}{\partial t} - \frac{\partial}{\partial \mathbf{x}}\left\{\sum_{k=1}^{N} c_k B_k(\mathbf{x}) \frac{\partial}{\partial \mathbf{x}}\right\}, \tag{4.7}$$

let

$\mathcal{B} = u :$ u satisfies the given initial and boundary conditions

$\mathcal{B}_0 = u :$ u satisfies homogeneous initial and boundary conditions

and let

$$\delta_k = (0, \cdots, 0, \delta, 0, \cdots, 0),\ \delta \text{ in the kth position}$$

Let u_c be the solution to

$$L_c u_c = q, \quad u_c \in \mathcal{B}, \tag{4.8}$$

let $u_{c+\delta_k}$ be the solution to

$$L_{c+\delta_k} u_{c+\delta_k} = q, \quad u_{c+\delta_k} \in \mathcal{B} \tag{4.9}$$

and let

$$h_{c,k}(\delta) = \frac{u_{c+\delta_k} - u_c}{\delta}. \tag{4.10}$$

Observe that

$$L_{c+\delta_k} = L_c - \delta \frac{\partial}{\partial \mathbf{x}} B_k(x) \frac{\partial}{\partial \mathbf{x}}, \tag{4.11}$$

then substituting 4.7 into 4.10 gives

$$(L_c - \delta \frac{\partial}{\partial \mathbf{x}} B_k(\mathbf{x}) \frac{\partial}{\partial \mathbf{x}})(u_c + \delta h_{c,k}(\delta)) = q \tag{4.12}$$

$$u_c + \delta h_{c,k}(\delta) \in \mathcal{B}.$$

Subtracting 4.8 from 4.12 gives

$$L_c h_{c,k}(\delta) = \frac{\partial}{\partial \mathbf{x}} B_k(\mathbf{x}) \frac{\partial}{\partial \mathbf{x}} (u_c + \delta h_{c,k}(\delta)) \tag{4.13}$$

and, assuming that we can take limits as $\delta \to 0$, and letting $\lim_{\delta \to 0} h_{c,k}(\delta) = h_{c,k}$ gives that $h_{c,k}$ is the solution to the problem

$$L_c h = \frac{\partial}{\partial \mathbf{x}} B_k(\mathbf{x}) \frac{\partial}{\partial \mathbf{x}} u_c$$

$$h \in \mathcal{B}_0$$

Thus if everything is sufficiently "nice" $X_{ijk}^{(l)}$ can be obtained by solving

$$L_{c^{(l)}} h = \frac{\partial}{\partial \mathbf{x}} B_k(\mathbf{x}) \frac{\partial}{\partial \mathbf{x}} u_{c^{(l)}}, \quad h \in \mathcal{B}_0 \tag{4.14}$$

and evaluating the solution $h_{c,k}$ at $\mathbf{x}(i), t_j$. O'Sullivan (1988) has carried out this program on a one-dimensional example.

We emphasize that this is a non linear ill posed problem complicated by the fact that the degree of non linearity as well as the degree of ill-posedness can depend fairly strongly on the unknown solution. To see more clearly some of the issues involved, let us examine a problem sitting in Euclidean n-space which has many of the features of the system identification problem. Let $X_1, \ldots, X_N$ be $N \le n$ matrices each of dimension $n - M \times n$, let B be an $M \times n$ matrix of rank M, and let $u \in E^n$, $q \in E^{n-M}$ and $b \in E^M$ be related by

$$\left(\sum_{k=1}^{N} c_k X_k \right) u = q \tag{4.15}$$

$$Bu = b.$$

Think of c, q and b as stand-ins for α; the forcing function; and the initial/boundary conditions, respectively.

Suppose q and b are known exactly and it is known a priori that $c_k \geq \alpha_k > 0$, $k = 1, \ldots, N$, and that this condition on the c_k's ensures that the matrix $\begin{pmatrix} \sum_k c_k X_k \\ B \end{pmatrix}$ is invertible. Suppose that one observes

$$y_i = u_i + \epsilon_i, \quad i = 1, \ldots, n$$

where u_i is the ith component of u. Letting $\Psi_{ij}(c)$ be the ijth entry of $\begin{pmatrix} \sum_n c_k X_k \\ \cdots \\ B \end{pmatrix}^{-1}$, we may estimate c as the minimizer of

$$\sum_i^n \left(y_i - \sum_{j=1}^{n-M} \Psi_{ij}(c) q_j - \sum_{j=n-M+1}^{n} \Psi_{ij}(c) b_{j-(n-M)} \right)^2 + \lambda c' \Sigma c, \tag{4.16}$$

subject to $c_k \geq \alpha_k$. The ability to estimate the c's can be expected to be quite sensitive to the true values of c as well as q and b.

Returning to the original system identification problem, we now consider the case where the boundary conditions are not completely known. If (as in a one dimensional, steady state problem) there are only $M << n$ unknowns in the initial/boundary values, then the analogue of 4.16 could (in principle) be minimized with respect to c and $b = (b_1, \ldots, b_M)$.

More generally, suppose that the forcing function q and the boundary conditions $\frac{\partial u}{\partial w} = 0$ are known exactly, but the initial conditions $u(\mathbf{x}, 0) = u_0(\mathbf{x})$ are observed with error, that is

$$z_i = u_0(\mathbf{x}(i)) + \epsilon_i$$

Modelling $u_0(\mathbf{x})$ as

$$u_0(\mathbf{x}) \simeq \sum_{\nu=1}^{M} b_\nu \tilde{B}_\nu(\mathbf{x}) \tag{4.17}$$

where the $\tilde{B}_\nu$ are appropriate basis functions (not necessarily the same as before) and letting $b = (b_1, \ldots, b_M)$, we have

$$u \simeq u(\mathbf{x}, t; c, b)$$

and we want to choose b, and c subject to appropriate constraints, to minimize

$$\frac{1}{n}\left\{\sum_{ij}(y_{ij}-u(\mathbf{x}(i),t_j;c,b))^2+\sum_i(z_i-\sum_\nu b_\nu\tilde{B}_\nu(\mathbf{x}(i)))^2\right\}+\lambda_1 c'\Sigma c+\lambda_2 b'\tilde{\Sigma} b \tag{4.18}$$

where $b'\tilde{\Sigma}b$ is an appropriate penalty on u_0. The penalty functionals $c'\Sigma c$ and $b'\tilde{\Sigma}b$ may be quite different, since the first contains prior information about the permeability and the second about the field. (This expression assumes that all the measurements have the same variance, but weights can be used if this is not the case).

For fixed λ_1 and λ_2 this minimization can in principle be done as before, provided we have a means of calculating

$$\tilde{X}_{ij\nu}=\frac{\partial u}{\partial b_\nu}(\mathbf{x}(i),t_j;c;b). \tag{4.19}$$

The $\tilde{X}_{ij\nu}$ can be found by the same method used for the X_{ijk}. Let $u_{c,b}$ be the solution to the problem

$$L_c u_{c,b}=q,\ \ \frac{\partial u_{c,b}}{\partial w}=0,\ \ u_{c,b}(\mathbf{x},0)=\Sigma b_\nu\tilde{B}_\nu(\mathbf{x}). \tag{4.20}$$

Let $\delta_\nu=(0,\cdots,\delta,\cdots 0),\delta$ in the νth position, and let $u_{c,b+\delta_\nu}$ be the solution to

$$L_c u_{c,b+\delta_\nu}=q,\ \frac{\partial u_{c,b+\delta_\nu}}{\partial w}=0,u_{c,b+\delta_\nu}(\mathbf{x},0)=\delta\tilde{B}_\nu+\left(\sum_\mu b_\mu\tilde{B}_\mu(\mathbf{x})\right) \tag{4.21}$$

and let

$$\tilde{h}_{c,b,\nu}(\delta)=\frac{u_{c,b+\delta_\nu}-u_{c,b}}{\delta}. \tag{4.22}$$

Then, subtracting 4.21 from 4.20 as before, we see that $\lim_{\delta\to 0}\tilde{h}_{c,b,\nu}(\delta)=\tilde{h}_{c,b,\nu}$ is the solution to the problem

$$L_c u=0,\ \ \frac{\partial u}{\partial w}=0,u(\mathbf{x},0)=\tilde{B}_\nu(\mathbf{x}). \tag{4.23}$$

Then the $\tilde{X}_{ij,\nu}$ are obtained by evaluating $\tilde{h}_{c,b,\nu}$ at $\mathbf{x}(i),t_j$. $V(\lambda_1,\lambda_2)$ can at least in principle be minimized to estimate good values of λ_1 and λ_2 by GCV. Some numerical methods for minimizing V as a function of multiple smoothing parameters may be found in Gu and Wahba(1988).

References

Anderson, R., F. deHoog, and M. Lukas (1980) *The Application and Numerical Solution of Integral Equations,* Sijthoff and Noordhoff, Germantown.

Craven, P. and G. Wahba (1979) Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation, *Numer. Math., 31,* 377–403.

deBoor, C. (1978) *A Practical Guide to Splines,* Springer-Verlag, New York.

Golub, G., M. Heath, and G. Wahba (1979) Generalized cross validation as a method for choosing a good ridge parameter, *Technometrics, 21,* 215–224.

Gu, C., D. Bates, Z. Chen, and G. Wahba (1988) *The computation of GCV functions through Householder tridiagonalization with application to the fitting of interaction spline models,* Technical Report 823, Statistics Dept., University of Wisconsin, Madison, WI, to appear, SIAM J. Matrix Analysis.

Gu, C. and G. Wahba (1988) *Minimizing GCV/GML scores with multiple smoothing parameters via the Newton method,* Technical Report 847, Dept. of Statistics, University of Wisconsin, Madison, WI.

Kimeldorf, G. and G. Wahba (1970) Spline functions and stochastic processes, *Sankya Ser. A, 32, Part 2,* 173–180.

Kimeldorf, G. and G. Wahba (1971) Some results on Tchebycheffian spline functions, *J. Math. Anal. Applic., 33,* 82–95.

Kravaris, C. and J. Seinfeld (1985) Identification of parameters in distributed parameter systems by regularization, *SIAM J. Control Opt., 23,* 217–241.

Li, K. C. (1986) Asymptotic optimality of C sub L and generalized cross validation in ridge regression with application to spline smoothing, *Ann. Statist., 14,* 1101–1112.

Nychka, D., G. Wahba, S. Goldfarb, and T. Pugh (1984) Cross-validated spline methods for the estimation of three dimensional tumor size distributions from observations on two dimensional cross sections, *J. Am. Stat. Assoc., 79,* 832–846.

O'Sullivan, F. (1986) A statistical perspective on ill-posed inverse problems, *Statistical Science, 1,* 502–527.

O'Sullivan, F. (1987) *Constrained non-linear regularization with application to some system identification problems,* Technical Report 99, Dept. of Statistics, University of California, Berkeley, Berkeley, CA.

O'Sullivan, F. (1988) *Parameter estimation in parabolic and hyperbolic equations,* Technical Report 127, Dept. of Statistics, University of Washington, Seattle, WA.

O'Sullivan, F. and G. Wahba (1985) A cross validated Bayesian retrieval algorithm for non-linear remote sensing, *J. Comput. Physics, 59,* 441–455.

Utreras, F. (1981) Optimal smoothing of noisy data using spline functions, *SIAM J. Sci. Stat. Comput., 2,* 349–362.

Wahba, G. (1977) Practical approximate solutions to linear operator equations when the data are noisy, *SIAM J. Numer. Anal., 14,* 651–667.

Wahba, G. (1978) Improper priors, spline smoothing and the problem of guarding against model errors in regression, *J. Roy. Stat. Soc. Ser. B, 40,* 364–372.

Wahba, G. (1980) Spline bases, regularization, and generalized cross validation for solving approximation problems with large quantities of noisy data, in *Approximation Theory III,* edited by W. Cheney, pp. 905–912, Academic Press.

Wahba, G. (1982) Constrained regularization for ill posed linear operator equations, with applications in meteorology and medicine, in *Statistical Decision Theory and Related Topics, III, Vol.2,* edited by S. Gupta and J. Berger, pp. 383–418, Academic Press.

Wahba, G. (1985a) Comments to P. Huber, Projection pursuit, *Ann. Statist., 13,* 518–521.

Wahba, G. (1985b) A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem, *Ann. Statist., 13,* 1378–1402.

Wahba, G. and Y. Wang (1987) *When is the optimal regularization parameter insensitive to the choice of the loss function?,* Technical Report 809, Dept. of Statistics, University of Wisconsin, Madison, WI.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Unlimited
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report No. 852	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION University of Wisconsin-Madison 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Department of Statistics University of Wisconsin-Madison Madison, WI 53706	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING / SPONSORING ORGANIZATION AFOSR/PKZ 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code) Bldg. 410 Bolling AFB, Washington, D.C. 20332-6448	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
REGULARIZATION AND CROSS VALIDATION METHODS FOR NONLINEAR, IMPLICIT, ILL-POSED INVERSE PROBLEMS

12. PERSONAL AUTHOR(S)
Grace Wahba

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Scientific Interim	FROM ____ TO ____	March 1989	14

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Cross validated regularization methods for ill posed inverse problems are reviewed. These methods have been extended to the parameter estimation problem for partial differential equations. In the p. d. e. problem one observes discrete, possibly noisy data on the solution, the forcing function, and the boundary values, and wishes to estimate a (distributed) coefficient in the equation. Some directions for further extension of the method are suggested.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code) / 22c. OFFICE SYMBOL