TECHNICAL REPORT NO. 420

August 1975
A CANONICAL FORM FOR THE PROBLEM
OF ESTIMATING SMOOTH SURFACES

by

Grace Wahba

Department of Statistics
University of Wisconsin-Madison

This work was supported by the U.S. Air Force Office of Scientific Research under Grant AF-AFOSR-2363-B
A CANONICAL FORM FOR THE PROBLEM
OF ESTIMATING SMOOTH SURFACES

by
Grace Wahba
Department of Statistics
University of Wisconsin-Madison

July, 1975


This work was supported by the U. S. Air Force Office of Scientific
Research under Grant AF-AFOSR-2363-B
ABSTRACT

We show how the problem of estimating a smooth surface on a rectangle in Euclidean p-space, which is measured discretely and with normally distributed errors, reduces to the problem of estimating the mean of a multivariate normal vector.

Two empirical Bayes type estimators are noted, and it is observed that cross-validation is useful in certain cases.
1. The Problem of Estimating a Smooth Surface

Our model is

\[ y(t) = f(t) + e(t), \quad t \in T \]

where \( T \) is a rectangle in Euclidean p-space\(^1\) and

i) \( e(t) \sim \mathcal{N}(0, \sigma^2) \), i.i.d., \( t \in T \).

\( \sigma^2 \) may be known or unknown. \( f(t) \) is either a smooth function in a given reproducing kernel Hilbert space \( \mathcal{H}_Q \) with reproducing kernel \( Q(s,t) \) or a stochastic process with \( \mathbb{E}f(t) = 0, \mathbb{E}f(s)f(t) = bQ(s,t), \) \( b \) unknown. It is instructive to compare the two situations.

\( Q(s,t) \) is given by

\[ Q(s,t) = \sum_{\nu=1}^{\infty} \lambda_\nu \phi_\nu(s) \phi_\nu(t) \]

where \( \{\phi_\nu\}_{\nu=1}^{\infty} \) is an orthonormal set of continuous functions on \( L_2[T] \), \( \lambda_\nu > 0 \) and

\[ \lambda_\nu = o(\nu^{-2m}) \]

for some fixed \( m \geq 2 \).\(^2\) \( f \in \mathcal{H}_Q \) iff \( f \in \text{span}\{\phi_\nu\} \) and

\[ \sum_{\nu=1}^{\infty} \frac{f_\nu^2}{\lambda_\nu} < \infty, \]

where the generalized Fourier coefficients \( f_\nu \) are given by

\(^1\) \( T \) can be much more general, specifically any compact metric space on which can be defined an infinite sequence of continuous \( L_2 \)-orthonormal functions.

\(^2\) Our analysis can be carried out for other decay rates of \( \lambda_\nu \), e.g., \( \lambda_\nu = o(e^{-\alpha \nu}) \).
\[ f_v = \int_{\mathcal{T}} \phi_v(s) f(s) \, ds. \quad (1) \]

We consider the two (distinct) cases

\begin{itemize}
  \item[ii)] \( f \in \mathcal{H}_Q \) and \( \sum_{v=1}^{\infty} \frac{f_v^2}{\lambda_v^2} < \infty \)
  \item[ii')] \( f(t) = \sum_{v=1}^{\infty} f_v \phi_v(t), \quad f_v \sim \mathcal{N}(0, b\lambda_v) \) independent.
\end{itemize}

The smoothing problem is to recover an estimate \( \hat{f}(t) \) of \( f(t), t \in \mathcal{T}, \) given observations \( y(t), t \in \mathcal{T}_n, \) where \( \mathcal{T}_n \) is an \( n \)-point subset of \( \mathcal{T}. \) The loss when \( \hat{f} \) is used is \( \int_{\mathcal{T}} (f(t) - \hat{f}(t))^2 \, dt. \) In this note we demonstrate how this problem can (large \( n \)) be reduced to the problem of estimating the mean of a multivariate normal, thus the extensive literature on this latter problem (see Efron and Morris [5] and Hudson [7] and the bibliographies there) can be brought to bear on the problem. We suggest a simple estimate for the \( \sigma^2 \) known case which looks reasonable for both ii) and ii'). When \( \sigma^2 \) is unknown, we note that an estimator derived from cross validation as in Wahba and Wold [12] is good for ii). An idea of Anderson and Bloomfield [1] [2] applies to ii').
2. The Estimates

We define a one-parameter family of estimates, \( f_n, \lambda, \lambda \geq 0 \) for \( f \) as follows:

\[
f_{n, \lambda}(t) = \sum_{v=1}^{n} \frac{\hat{f}_v}{1 + \lambda/\lambda_v} \phi_v(t), \quad t \in T
\]

where

\[
\hat{f}_v = \lambda_v (\phi_v(t_1), \ldots, \phi_v(t_n)) Q_n^{-1} \begin{pmatrix}
y(t_1) \\
\vdots \\
y(t_n)
\end{pmatrix}
\]

and \( Q_n \) is the \( n \times n \) matrix with \( i, j \)th entry \( Q(t_i, t_j) \). The \( \{\hat{f}_v\} \) should be viewed as the sample generalized Fourier coefficients and the formula for \( \hat{f}_v \) as a quadrature formula for the integral

\[
\frac{1}{\lambda} \int_0^1 \phi_v(s) g(s) \, ds
\]

given \( g(t_1), g(t_2), \ldots, g(t_n) \).

We have

\[
\hat{f}_v = f_{vn} = \lambda_v (\phi_v(t_1), \ldots, \phi_v(t_n)) Q_n^{-1} \begin{pmatrix}
f(t_1) \\
\vdots \\
f(t_n)
\end{pmatrix}
\]

\( v = 1, 2, \ldots, n \).

It can be shown that

\[
f_{vn} = \frac{1}{\lambda} \int_0^1 \phi_v(t) (P_T f)(t) \, dt, \quad v = 1, 2, \ldots
\]

where \( P_T f \) is the orthogonal projection in \( \mathcal{H}_Q \) of \( f \) onto the
subspace \( V_n = \text{span}(Q_{t_1}(\cdot)) \), where \( Q_{t_1}(\cdot) = Q(t_1, \cdot) \). (For calculations of this type see [9], [11] and references cited there.) Sometimes

\[
\hat{f}_v \approx \frac{1}{n} \sum_{v=1}^{n} \phi_v(t_1) y(t_1).
\]

Furthermore, by Parseval's theorem,

\[
\sum_{v=1}^{\infty} (f_v - \hat{f}_v)^2 = \int_T \left[f(t) - (P_n f)(t)\right]^2 dt. \tag{1}
\]

Convergence properties of \( f = P_n f \) when \( T = [0,1] \) may be found in [9], [10], [11], the quantity \((1)\) is \( O(n^{-2m-1}) \) when ii) holds if the maximum distance between two neighboring points is \( O(1/n) \). Under the model iii'), \( (1+\lambda/n)^{-1} \).

\( \hat{f}_v \) can be viewed as a good approximation to the posterior mean of \( f_v \)

and \( \hat{f}_n, \lambda(t) \) as a good approximation to the posterior mean of \( f(t) \)

when \( \lambda = \sigma^2/nb \).

Letting \( \Gamma \) be the nxn matrix with \( v \)th entry \( \phi_v(t_1) \), and \( D \) be

the nxn diagonal matrix with \( vv \)th entry \( \lambda_v \), we have that the covariance matrix \( \Sigma \) of \( (\hat{f}_1, \ldots, \hat{f}_n) \) is

\[
\Sigma = \sigma^2 D \Gamma D' (\Gamma D \Gamma')^{-1} B^{-1} \Gamma D
\]

where the \( i,j \)th entry of \( B \) is \( \sum_{v=n+1}^{\infty} \lambda_v \phi_v(t_i) \phi_v(t_j) \).

If \( \frac{1}{n} \sum_{v=1}^{n} \phi_v(t_1) \) is uniformly bounded, then \( \text{Trace } B = O(n^{-(2m-2)}) \). Then, \( 2/ \) to a good approximation,

\( 2/ \) This is the only place \( m \geq 2 \) is used. Elsewhere we only use \( m \geq 1 \).
\[ \sum \sigma^2 D_f (D_f')^{-2} \Gamma D = \sigma^2 (\Gamma \Gamma')^{-1} \]

The loss when \( f_{n, \lambda}(t) \) is used is given by

\[ \int_T (f(t) - f_{n, \lambda}(t))^2 \, dt , \]

and the expected loss, \( R(\lambda) \) is given by

\[
R(\lambda) = \sum_{v=n+1}^{\infty} f_v^2 + \sum_{v=1}^{n} \left( f_v - \frac{\lambda f_{\hat{V}_v}}{\hat{\lambda}_v + \lambda} \right)^2
\]

\[
= \sum_{v=n+1}^{\infty} f_v^2 + \sum_{v=1}^{n} \left( f_v - \frac{\lambda f_{\hat{V}_v}}{\hat{\lambda}_v + \lambda} \right)^2 + \sum_{v=1}^{n} \var_\lambda \frac{\lambda^2 \var_\lambda f_{\hat{V}_v} f_v}{(\hat{\lambda}_v + \lambda)^2} + 2 \sum_{v=1}^{n} \var_\lambda f_{\hat{V}_v} f_v
\]

\[
= \left\{ \sum_{v=n+1}^{\infty} f_v^2 + \sum_{v=1}^{n} \frac{\lambda^2 (f_v - f_{\hat{V}_v})^2}{(\hat{\lambda}_v + \lambda)^2} + 2 \sum_{v=1}^{n} \frac{\lambda f_{\hat{V}_v} f_v}{(\hat{\lambda}_v + \lambda)^2} \right\}
\]

\[
+ \left\{ \lambda^2 \sum_{v=1}^{n} \frac{f_v^2}{(\lambda + \lambda_v)^2} + \sum_{v=1}^{n} \frac{\lambda^2 \var_\lambda f_{\hat{V}_v}}{(\lambda + \lambda_v)^2} \right\}
\]

The first term in brackets is bounded in absolute value by

\[
\sum_{v=n+1}^{\infty} f_v^2 + \sum_{v=n+1}^{\infty} f_{\hat{V}_v}^2 + \int_T (f(t) - \hat{f}_n(t))^2 \, dt + 2\lambda \left( \sum_{v=1}^{n} \frac{f_v^2}{\lambda_v} \right) \left( \sum_{v=1}^{n} \frac{\lambda^2 (f_v - f_{\hat{V}_v})^2}{\lambda^2} \right)^{1/2}
\]

and we shall suppose that it is negligible compared to the second term in brackets as \( n \to \infty \). This is true in all the examples we know of whenever the points in \( T_n \) become dense in \( T \).
Suppose further, that the \( \{t_i\} \) are regularly enough spaced so that

\[
\frac{1}{n} \sum_{i=1}^{n} \phi_{\mu}(t_i) \phi_{\mu}(t_i) \sim \int \phi_{\mu}(t) \phi_{\mu}(t) \, dt
\]

\[
= 1, \quad \mu = \nu
\]

\[
= 0, \quad \mu \neq \nu.
\]

Regularity conditions on the distribution of the \( t_i \)'s would be required for this. Then

\[
\eta' \eta \sim nI, \quad \text{var} \; \hat{f}_{\nu} \sim \frac{\sigma^2}{n}.
\]

Thus whenever (1) is very small, and (2) holds approximately, we have reduced the problem to the "canonical" form

\[
\hat{f}_{\nu} \sim \mathcal{N}(f_{\nu}, \sigma^2/n), \quad \text{independent}
\]

with either

\[
\text{ii)} \quad \sum_{\nu=1}^{\infty} \frac{f^2_{\nu}}{\lambda^2_{\nu}} < \infty
\]

or

\[
\text{ii') } \quad f_{\nu} \sim \mathcal{N}(0, b \lambda_{\nu}).
\]

In either case ii) or ii'), we estimate \( f_{\nu} \) by \( \hat{f}_{\nu}(1 + \lambda/\lambda_{\nu})^{-1} \), with expected loss
\[ R(\lambda) = \mathbb{E} \left( \sum_{v=1}^{n} \left( f_v - \frac{\hat{f}_v}{(1 + \lambda/\lambda_v)} \right)^2 \right) \approx \lambda^2 \sum_{v=1}^{n} \left( \frac{\hat{f}_v^2}{(\lambda_v + \lambda)^2} + \frac{\sigma^2}{n} \sum_{v=1}^{n} \frac{\lambda_v^2}{(\lambda_v + \lambda)^2} \right)^2 
\]  
(An argument resulting in an expression similar to (3) can be found in Cogburn and Davis [3].)

If \( \sigma^2 \) is known, and \( \sum_{v=1}^{n} (f_v - f_{vn})^2 \) negligible, then an unbiased estimate of \( R(\lambda) \) of (3) is \( \hat{R}(\lambda) \) given by

\[ \hat{R}(\lambda) = \lambda^2 \frac{\sum_{v=1}^{n} \hat{f}_v^2}{(\lambda_v + \lambda)^2} + \frac{\sigma^2}{n} \sum_{v=1}^{n} \frac{\lambda_v^2 - \lambda^2}{(\lambda_v + \lambda)^2} \]

and it is reasonable to suppose that the minimizer of \( \hat{R}(\lambda) \) would provide a good choice of \( \lambda \) for either model ii) or ii'). If (2) does not hold, then \( \text{var} \hat{f}_v = \sigma^2/n \) must be replaced by \( \text{var} \hat{f}_v = \sigma^2 \gamma_{vv} \) where \( \gamma_{vv} \) is the \( vv \)th entry of \( (\Gamma' \Gamma)^{-1} \), and \( \hat{R}(\lambda) \) becomes

\[ \hat{R}(\lambda) = \lambda^2 \frac{\sum_{v=1}^{n} \hat{f}_v^2}{(\lambda_v + \lambda)^2} + \frac{\sigma^2}{n} \sum_{v=1}^{n} \frac{\lambda_v^2 - \lambda^2 \gamma_{vv}}{(\lambda_v + \lambda)^2} \]

Suppose ii') holds along with (2) and \( \sigma^2 \) and \( b \) are unknown. Then a maximum likelihood estimate for \( \lambda = \sigma^2/nb \) can be obtained using in the likelihood function the distribution

\[ \hat{f}_v \sim \mathcal{N}(0, b(\lambda_v + \sigma^2/\lambda b)) = \mathcal{N}(0, b(\lambda_v + \lambda)), \text{ independent}. \]

The estimate for \( \lambda \) is the minimizer of

\[ \frac{\lambda \sum_{v=1}^{n} \left( \frac{\hat{f}_v^2}{(\lambda_v + \lambda)^2} \right)}{\left[ \frac{n}{(\lambda_v + \lambda)^{1/n}} \right] \prod_{v=1}^{n} \left( \lambda_v + \lambda \right)} \]
This idea is to be found in Andersen and Bloomfield [1] [2].

Suppose (1) and (2) holds. If \( \lambda_v = h^{-1}(v) v^{-2m} \) where \( a \le h \le b \), then

\[
\sum_{v=1}^{n} \frac{\lambda_v^2}{(\lambda_v^2 + \lambda)^2} = \sum_{v=1}^{n} \frac{1}{(1 + \lambda h(v) v^{2m})} \approx \frac{c}{\lambda^{1/2m}} \int_{0}^{\infty} \frac{dx}{(1 + x^{2m})^2}
\]

where \( b^{-1/2m} \le c \le a^{-1/2m} \). It is then not hard to show (see [14] for details) that the minimizer \( \lambda^* \) of \( R(\lambda) \) of (3) satisfies

\[
\lambda^* = \left[ \frac{\sigma^2}{4m k_m} \left( \frac{1}{n} \sum_{v=1}^{n} \frac{1}{\lambda_v^2} \right) \right]^{2m/(4m+1)} (1 + o(1)),
\]

with \( o(1) \rightarrow 0 \) as \( n \rightarrow \infty \), and so

\[
R(\lambda^*) = o(n^{-4m/(4m+1)}).
\]

Let

\[
\nu(\lambda) = \frac{\lambda^2 \sum_{v=1}^{n} \frac{x^2}{(\lambda_v + \lambda)^2}}{\left( \frac{1}{n} \sum_{v=1}^{n} \frac{1}{(\lambda_v + \lambda)} \right)^2}.
\]

It is shown in [12][13] that

\[
\nu(\lambda) \sim \sum_{k=1}^{n} \left( f_{n,k}^{(k)}(t_k) - y(t_k) \right)^2 \omega_{kk}(\lambda)
\]

where \( f_{n,k}^{(k)} \) is \( f_{n-1,k} \) where the kth data point \( y(t_k) \) has been omitted, and
\[ \omega_{kk}(\lambda) = m_{kk}(\lambda)/n \sum_{j=1}^{n} m_{jj}(\lambda) \]

where \( m_{jj}(\lambda) \) is the \( jj \)th entry of \( (Q_n + n\lambda I)^{-1} \). The minimizer of \( V(\lambda) \) may thus be viewed as a cross-validation estimate of \( \lambda \). It is shown in [12][13] that, if \( \tilde{\lambda} \) is the minimizer of \( EV(\lambda) \), then \( \tilde{\lambda} = \lambda^*(1 + o(1)) \), where \( o(1) \to 0 \) as \( n \to \infty \).


For computational purposes, when \( T \) is the unit cube in Euclidean \( p \)-space, it may be convenient to let \( \mathcal{H}_Q = \mathcal{H}_R \times \mathcal{H}_R \times \cdots \mathcal{H}_R \) where \( \mathcal{H}_R \) is a reproducing kernel Hilbert space of functions on \([0,1]\). When \( p = 2 \) and \( s = (s_1, s_2) \), \( t = (t_1, t_2) \), then

\[ Q(s, t) = Q(s_1, s_2; t_1, t_2) = R(s_1, t_1) R(s_2, t_2). \]

where \( R(s, t) \) is the reproducing kernel for \( \mathcal{H}_R \). If \( R(s, t) = \sum_{\nu=1}^{\infty} \lambda_{\nu} \varphi_{\nu}(s) \varphi_{\nu}(t) \), then the eigenfunctions and eigenvalues of \( Q \) are given by

\[ \lambda_{\mu \nu} = \lambda_{\mu} \lambda_{\nu}, \quad \mu, \nu = 1, 2, \ldots \]

\[ \varphi_{\mu \nu}(s, t) = \varphi_{\mu}(s) \varphi_{\nu}(t). \]

See Cogburn and Davis [3], Golomb [6] for handy reproducing kernels for spaces of periodic functions on \([0,1]\). If \( R(s, \cdot) \) is a spline function (see, e.g. [4][8]) then \( Q(s, \cdot) \) will be a tensor product spline.
Under model ii'), the exact posterior mean $f_{n,\lambda}(\sim_t)$, say, of $f(\sim_t)$ when $\lambda = \sigma^2/nb$ is given by

$$f_{n,\lambda}(\sim_t) = (Q_{\sim_{t_1}}(\sim_t), \ldots, Q_{\sim_{t_n}}(\sim_t)) (Q_n + n\lambda I)^{-1} \begin{pmatrix} y(\sim_{t_1}) \\ \vdots \\ y(\sim_{t_n}) \end{pmatrix}$$

of which $f_{n,\lambda}$ is a good approximation. $f_{n,\lambda}$ is in the subspace $V_n$. If the $\sim_{t_i}$ are irregularly spaced and $n$ is very large the following procedure ($p = 2$), which reduces to the model to a regression model, may be computationally simpler without much loss in accuracy. Let

$$V_{kk} = \text{span}(Q_{\sim}(\cdot), \sim \in T_{kk}) \text{ where } T_{kk} = \begin{pmatrix} \frac{1}{k} \\ \frac{1}{k} \\ \vdots \\ \frac{1}{k} \end{pmatrix},$$

$i, j = 0, 1, \ldots, k, (k+1)^2 = q < n$. If, e.g., $R(s, \cdot)$ is a cubic spline, then $V_{kk}$ is a space of bi-cubic (tensor product) splines. Choose any convenient basis, say $(\omega_{\sim}(t))_{\sim=1}^q$ for $V_{kk}$. Then

$$f(\sim) = \sum_{\sim=1}^q \omega_{\sim}(t) \beta_{\sim} + (f - P_{T_{kk}} f)(\sim),$$

for some $\{\beta_{\sim}\}$ where $P_{T_{kk}} f$ is the projection of $f$ onto $V_{kk}$. If $R(s, t)$ "behaves like" a Green's function for a $2m_0$th order linear differential operator, \(\frac{\partial}{\partial t}\) (which happens for $R(s, \cdot)$ a polynomial spline of degree $2m_0 - 1$) then it can be shown for model ii) that

$$|f(\sim) - P_{V_{kk}} f(\sim)| \leq O(k^{-(2m_0 - \frac{1}{2})}).$$

\[\text{Then the eigenvalues for } R(s, t) \text{ are } O(\nu^{-2m_0}).\]
(See [10] for some of the details.) Thus a good approximation to the original model ii) is the regression model,

\[ y(t) = \sum_{v=1}^{q} \omega_v(t) \beta_v + \epsilon(t), \quad t \in T \]  

(4)

If \( X \) is the \( n \times q \) matrix with \( v \)th entry \( \omega_v(t) \) and (4) is a good approximation to the original model, then the prior \( \text{ii}' \) is approximately equivalent to a zero mean Gaussian prior on the \( \{\beta_v\} \) with covariance \( \Sigma_{\beta \beta} \) approximately satisfying \( bQ_b \approx bX\Sigma_{\beta \beta}X' \). Under model \( \text{ii}' \), then, the posterior mean of \( \beta = (\beta_1', \ldots, \beta_q') \) is \( \beta_\lambda \) given by

\[ \beta_\lambda = (X'X + \lambda \Sigma_{\beta \beta}^{-1})^{-1}X'y \]

where \( \lambda = \frac{\sigma^2}{b} \). \( \Sigma_{\beta \beta}^{-1} \) can be approximated e.g. by

\[ \Sigma_{\beta \beta}^{-1} \approx X' \Gamma \begin{pmatrix} \lambda_1 & 0 \\ \vdots & \ddots \\ 0 & \lambda_q \end{pmatrix} \Gamma'X \]

where \( \Gamma \) is the \( q \times n \) dimensional matrix whose rows are the first \( q \) eigenvectors of \( Q_n \) and \( \lambda_1', \ldots, \lambda_q \) are the first (largest) \( q \) eigenvalues of \( Q_n \).

Acknowledgment

We wish to thank the Department of Statistics, Stanford University, for their generous hospitality while this was being written, and Professor Ingram Olkin for stimulating the author's interest in the problem of smoothing surfaces.
REFERENCES


13. Wahba, Grace (1975), The approximate solution of linear operator equations when the data are noisy, University of Wisconsin, Madison, Mathematics Research Center TSR No. 1567.
<table>
<thead>
<tr>
<th>REPORT DOCUMENTATION PAGE</th>
<th>READ INSTRUCTIONS BEFORE COMPLETING FORM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. REPORT NUMBER</td>
<td>Technical Report #420</td>
</tr>
<tr>
<td>2. GOVT ACCESSION NO.</td>
<td></td>
</tr>
<tr>
<td>4. TITLE (and Subtitle)</td>
<td>A CANONICAL FORM FOR THE PROBLEM OF ESTIMATING SMOOTH SURFACES</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>6. PERFORMING ORG. REPORT NUMBER</td>
<td></td>
</tr>
<tr>
<td>7. AUTHOR(s)</td>
<td>Grace Wahba</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>9. PERFORMING ORGANIZATION NAME AND ADDRESS</td>
<td>Department of Statistics University of Wisconsin 1234 W. Dayton Street</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>13. NUMBER OF PAGES</td>
<td>14</td>
</tr>
<tr>
<td>14. MONITORING AGENCY NAME &amp; ADDRESS (if different from Controlling Office)</td>
<td></td>
</tr>
<tr>
<td>15. SECURITY CLASS. (of this report)</td>
<td></td>
</tr>
<tr>
<td>15a. DECLASSIFICATION/DOWNGRADE SCHEDULE</td>
<td></td>
</tr>
<tr>
<td>16. DISTRIBUTION STATEMENT (of this report)</td>
<td>This document has been approved for public release and sale; its distribution is unlimited.</td>
</tr>
<tr>
<td>17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)</td>
<td></td>
</tr>
<tr>
<td>18. SUPPLEMENTARY NOTES</td>
<td></td>
</tr>
<tr>
<td>19. KEY WORDS (Continue on reverse side if necessary and identify by block number)</td>
<td>estimating surfaces splines cross-validation</td>
</tr>
<tr>
<td>20. ABSTRACT (Continue on reverse side if necessary and identify by block number)</td>
<td>We show how the problem of estimating a smooth surface on a rectangle in Euclidean p-space, which is measured discretely and with normally distributed errors, reduces to the problem of estimating the mean of a multivariate normal vector. Two empirical Bayes type estimators are noted, and it is observed that cross-validation is useful in certain cases.</td>
</tr>
</tbody>
</table>