A Note on Smoothness Priors
and Nonlinear Regression

by

R. L. Eubank

Technical Report No. SMU-DS-TR-195
Department of Statistics ONR Contract

December 1985

Research sponsored by the Office of Naval Research
Contract N00014-85-K-0340

Reproduction in whole or in part is permitted for
any purpose of the United States Government

This document has been approved for public
release and sale; its distribution is unlimited.

Department of Statistics
Southern Methodist University
Dallas, Texas 75275

A Note on Smoothness Priors

and Nonlinear Regression

by

R. L. Eubank[1]

Southern Methodist University

ABSTRACT

A generalization of a nonparametric estimator proposed by Shiller (1984) is considered. The estimator is shown to be directly related to penalized least-squares estimation and spline smoothing. Some simplifications concerning the computation of Shiller's estimator and corresponding posterior covariances are indicated.

Running Head: Smoothness Priors

Key Words: Bayesian prediction, nonparametric regression, partial splines, penalized least squares, spline smoothing.

[1]R. L. Eubank is Associate Professor, Department of Statistics, Southern Methodist University, Dallas, Texas 75275. This research was supported under Office of Naval Research Contract N00014-82-K-0207. The author would like to thank the editors and referees for helpful comments and suggestions which improved the original version of this paper.

1. INTRODUCTION

In a recent article in this journal Shiller (1984) considered the problem of estimation from a partially nonlinear regression model. He derived a nonparametric estimation procedure using Bayesian methodology and smoothness priors. In this note we derive a generalization of Shiller's estimator. The results we obtain reveal additional properties for Shiller's estimation technique and establish the connection between his estimator and spline smoothing. We also indicate some simplifications regarding the computation of his estimator and the associated conditional variances and covariances which arise from his Bayesian model.

Let y_i and $(t_i, \underline{x}_i')$, $i = 1,\ldots,n$, denote observations on a scalar response variable, Y, and p + 1 independent variables, t, $X_1,\ldots, X_p$, respectively. Assume that $a \leq t_1 < \ldots < t_n \leq b$, for finite constants a and b, and that the observations follow the partially nonlinear model

$$\underline{y} = \underline{f} + X\underline{\gamma} + \underline{\varepsilon}, \tag{1.1}$$

where $\underline{y}' = (y_1,\ldots,y_n)$, $X' = [\underline{x}_1,\ldots,\underline{x}_n]$ is a known $p \times n$ matrix, $\underline{\gamma}$ is an unknown vector of coefficients, $\underline{f} = (f(t_1),\ldots,f(t_n))'$ is a vector of values for some unknown function, f, and $\underline{\varepsilon}' = (\varepsilon_1,\ldots,\varepsilon_n)$ is a vector of zero mean, normal, random variables which are uncorrelated and have common variance σ^2. The objective is to obtain an estimate of the unknown regression function

$$\mu(t,\underline{x}) = f(t) + \underline{x}'\underline{\gamma},\ t \in [a,b]. \tag{1.2}$$

An estimator for μ in (1.2) has been derived by Shiller (1984) under the assumption that f is smooth in the sense that its slope does not change too rapidly. When i) there are no repeated observations, ii) we are not estimating μ at points where there are no observations and iii) certain restrictions (detailed in the Theorem below) are imposed on X, his estimator of $\underline{f}$ and $\underline{\gamma}$ can be written as

$$\begin{bmatrix} \hat{\underline{f}} \\ \hat{\underline{\gamma}} \end{bmatrix} = \begin{bmatrix} (I+\lambda R'H^{-1}R) & X \\ X' & X'X \end{bmatrix}^{-1} \begin{bmatrix} \underline{y} \\ X'\underline{y} \end{bmatrix}, \qquad (1.3)$$

where R and H are matrices involving differences among the t_i that are defined in his paper, and λ is a scalar parameter that must be specified by the user (or determined from the data). Shiller shows that his estimator is a cubic spline and it is clear from (1.3) that it is also a penalized least-squares estimator. However, it is by no means obvious from (1.3) how (or even if) his estimator is related to smoothing splines.

In the next section we discuss a general class of estimators for μ derived from a penalized least-squares criterion. This criterion allows for the assumption of varying degrees of smoothness corresponding to prior beliefs about the nature of f. The resulting estimators of f are generalizations of smoothing spline estimators. Shiller's estimator (1.3) is shown to correspond to the cubic case.

In Section 3 we discuss Shiller's Bayesian model in more detail and derive simplified expressions for the posterior covariance kernel of the regression function. An alternative Bayesian model which ties the estimator to polynomial regression is mentioned and the penalized least-squares and Bayesian perspectives are contrasted.

Shiller (1984) also considers estimators of $\underline{f}$ and $\underline{\gamma}$ derived using what are termed discrete smoothness priors. We note in passing that penalized least-squares and minimax properties of these estimators follow from work by Rice (1982) and Engle, et al. (1983).

2. PENALIZED LEAST SQUARES

Assume that f in (1.2) is a nonrandom function that admits m-1 absolutely continuous derivatives and has a square integrable mth derivative on [a,b], i.e., assume that $f \in W_2^m[a,b]$, where

$$W_2^m [a,b] = \{f : f^{(j)} \text{ abs. cont., } j = 0,\ldots,m-1, \int_a^b f^{(m)}(t)^2 dt < \infty\} .$$

In this case one approach to the estimation of μ would be to minimize the penalized least-squares criterion

$$\sum_{i=1}^{n} (y_i - \underline{x}_i' \underline{\gamma} - f(t_i))^2 + \lambda \int_a^b f^{(m)}(t)^2 \, dt, \quad \lambda > 0 , \tag{2.1}$$

with respect to $\underline{\gamma}$ and $f \in W_2^m [a,b]$.

Criterion (2.1) combines a measure of fidelity to the data $(\sum_{i=1}^{n}(y_i - \underline{x}_i' \underline{\gamma} - f(t_i))^2)$ with a measure of the smoothness of f. The presence of the term, $\lambda \int_a^b f^{(m)}(t)^2 dt$, penalizes a function which is not smooth in the sense of having a large mth derivative. Since $\int_a^b f^{(m)}(t)^2 dt$ vanishes if and only if f is a polynomial of order m, criterion (2.1) can be viewed as penalizing potential estimators of μ which depart too much from an mth order polynomial model in t.

Reasons for allowing general values of m in (2.1), rather that just m = 2, can be tied to prior beliefs about the nature of f. If, for example, one believes f to be approximately linear then m = 2 will suffice. In contrast, the belief that f is nearly quadratic would prompt the use of m = 3. The added flexibility of general m should be of value in areas such as response surface analysis where some knowledge of the response function is often available. We should also mention that once m has been chosen a choice of λ reflects our faith in the value selected for m. Larger values of λ force stricter adherence to polynomial behavior than smaller values. In settings where values of m and λ cannot be selected a priori, data driven methods, such as generalized cross-validation (Craven and Wahba 1979, Wahba 1984a) can be used to effect a choice for one or both parameters.

Wahba (1984a, b, 1985) has coined the name (univariate) <u>partial splines</u> for the estimators which arise from (2.1). For comparison with Shiller's approach we require an explicit form for these estimators which is detailed in the following theorem. The proof is both short and instructive and therefore included for completeness.

<u>Theorem</u>. Let $NS^{2m-1}(t_1, \ldots, t_n)$ denote the set of natural splines of degree 2m-1 with knots at $t_1, \ldots, t_n$, i.e., $NS^{2m-1}(t_1, \ldots, t_n)$ is the set of all piecewise polynomials with breakpoints at $t_1, \ldots, t_n$ which are of degree 2m-1 in (t_1, t_n), degree m-1 outside this interval, and possess 2m-2 continuous derivatives. Given a basis, $B_1, \ldots, B_n$, for $NS^{2m-1}(t_1, \ldots, t_n)$ define $n \times n$ matrices B and G and the $n \times m$ matrix T by

$$B = \{B_j(t_i)\}_{i,j=1,n} \quad ,$$

$$G = \{ \int_a^b B_i^{(m)}(t)\, B_j^{(m)}(t)dt\}_{i,j=1,n} \, ,$$

and

$$T = \{t_i^{j-1}\}_{\substack{i=1,n \\ j=1,m}} \quad .$$

Then, if $n \geq m$ there is a unique minimizer for (2.1) if and only if $[X, T]$ has full column rank. The resulting estimator of μ, $\mu_\lambda(t,\underline{x}) = f_\lambda(t) + \underline{x}'\underline{\gamma}_\lambda$, is in $NS^{2m-1}(t_1,\ldots,t_n)$ as a function of t and is given by

$$f_\lambda(t) = (B_1(t),\ldots,B_n(t))\ (B'B + \lambda G)^{-1}B'(\underline{y} - X\underline{\gamma}_\lambda) \qquad (2.2a)$$

and

$$\underline{\gamma}_\lambda = [X'(I-H(\lambda))X]^{-1}X'(I-H(\lambda))\underline{y}\ , \qquad (2.2b)$$

where

$$H(\lambda) = B(B'B + \lambda G)^{-1}B'.$$

Proof. Since $n \geq m$, the unique minimizer of (2.1) for any fixed, but arbitrary, $\underline{\gamma}$ is

$$(B_1(t),\ldots,B_n(t))(B'B + \lambda G)^{-1}B'(\underline{y} - X\underline{\gamma})\ ,$$

(see e.g., Kimeldorf and Wahba 1971, Wahba 1978 or Lyche and Schumaker 1973). Inserting this expression into (2.1) and minimizing with respect to $\underline{\gamma}$ leads to the system of equations

$$[X'(I-H(\lambda))X]\underline{\gamma} = X'(I-H(\lambda))\underline{y}\ .$$

It can be shown (see, e.g., Wahba 1978) that $I-H(\lambda)$ is a positive semi-definite matrix of rank n−m which satisfies $[I-H(\lambda)]T = 0$. The assumptions of the theorem are therefore both necessary and sufficient for $\underline{\gamma}_\lambda$ to be uniquely defined. □

It follows from the preceding proof that f_λ is the smoothing spline fit to $\underline{y} - X\underline{\gamma}_\lambda$. When p=0 the univariate smoothing spline estimator of f results. Consequently, the estimator (2.2) provides an extension of the smoothing spline concept.

To establish the connection between estimator (2.2) and Shiller's estimator let us restrict attention, for the moment, to the case of $m = 2$. Under the so called canonical basis for the natural cubic splines (see, e.g., Utreras 1979) $B_j(t_i) = \delta_{ij}$ so that $B = I$. It follows from Demmler and Reinsch (1975), for example, that $G = R'H^{-1}R$ in (1.3) for this basis. Substitution into (2.2) then reveals that his estimator corresponds to our estimator under the canonical basis and is therefore a partial cubic spline estimator of μ. Shiller did not explicitly assume that [X,T] has full rank. The need for this condition is a consequence of Theorem 1.

To evaluate his estimator of f at points other than those where observations were taken, Shiller used dummy variables to represent these additional points along the function. This has the drawback that an equation system of dimension larger than necessary for estimation of $\underline{f}$ and $\underline{\gamma}$ alone must be solved and that no explicit expression is available for f_λ as a function of t (see, however, his Section 4).

Equation (2.2a) provides an explicit expression for f_λ in terms of natural spline basis functions. From this we conclude that to estimate μ we can first compute $\underline{\gamma}_\lambda$ and then obtain the smoothing spline fit to $\underline{y} - X\underline{\gamma}_\lambda$ as our estimator of f. The estimator $\mu_\lambda(t,\underline{x}) = f_\lambda(t) + \underline{x}'\underline{\gamma}_\lambda$ can then be easily evaluated for any values of t and $\underline{x}$ once these computations are concluded. The number of calculations required is the same as is needed for estimation of $\underline{f}$ and $\underline{\gamma}$ alone. This gives an alternative to the approach suggested by Shiller that can provide computational savings and allows for the use of other bases which may be more computationally convenient. It should be noted that when computing f_λ and $\underline{\gamma}_\lambda$ one would not actually compute the inverse matrices in (2.2) but would instead utilize efficient methods for solving the normal equations from which (2.2) derives.

By expressing f_λ in terms of basis functions it is possible to provide a mathematical explanation for results in Shiller's Sections 3 and 4. He showed that adding or deleting evaluation points for f to his estimation procedure where there were no observations did not effect other estimated values of f or $\underline{\gamma}_\lambda$. We now see that this follows from the fact that the coefficients of the basis functions under the canonical basis are $f_\lambda(t_1),\ldots,$ $f_\lambda(t_n)$. The fact that f_λ is the unique element of $NS^{2m-1}(t_1, \ldots t_n)$ which interpolates $f_\lambda(t_1),\ldots,f_\lambda(t_n)$ provides an explanation for the interpolation properties of the estimator discussed in his Section 4.

Returning to the case of general m, it is well known that a smoothing spline reduces to a polynomial regression estimator when the smoothing parameter becomes infinite. Using this fact we see that, as $\lambda \to \infty$, $f_\lambda(t) + \underline{x}'\underline{\gamma}_\lambda$ reduces to the usual least-squares estimator of $\mu(t,\underline{x})$ for the case when f is assumed to be a polynomial of order m. This illustrates a relationship between estimator (2.2) and polynomial regression and generalizes results by Shiller for $m = 2$ (see his Section 6).

There are several generalizations of the partial spline estimator discussed in this section. Of particular importance are partial thin plate splines which can be used to estimate f and $\underline{\gamma}$ when f is nonlinear in more than one variable. Details can be found in Wahba (1984a, b, 1985). We note that this approach may be more appropriate in many cases then the suggestion by Shiller (1984, p. 609) that an additively separable model be used when f is nonlinear in several variables.

A final point worth mention is that the partial spline estimator enjoys certain minimax properties. Using work by Speckman (1982) or Li (1982) the estimator can be shown to provide a type of worst case protection against departures from a polynomial model in t.

3. POSTERIOR COVARIANCES AND RELATIONSHIPS BETWEEN MODELS

In this final section we study Shiller's Bayesian model in more detail. In particular, an expression is obtained for the posterior covariance kernel of his estimator for general m. Some comments on an alternative model and the relationship between the Bayesian and penalized least-squares approach are also given.

For general m, Shiller's model can be described as follows. Let $\{Z(t) : t \in [a,b]\}$ be the (m-1)-fold integral of a Weiner process. More specifically, Z is a zero mean, normal process with covariance kernel

$$Q(s,t) = [(m-1)!]^{-2} \int_a^s (s-u)^{m-1} (t-u)^{m-1}\, du, \qquad s \le t\ . \tag{3.1}$$

Then, μ is assumed to have the same prior distribution as the stochastic process $\sigma_s Z(t + \eta) + \underline{x}'\underline{\gamma}$, where σ_s and η are positive constants and $\underline{\gamma}$ is uncorrelated with Z and has a zero mean, normal distribution with variance-covariance matrix $\sigma_s^2 \alpha I$. The covariance between $\mu(t, \underline{x})$ and $\mu(s,\underline{x}^*)$ under this model is $\sigma_s^2(Q(s,t) + \alpha\underline{x}'\underline{x}^* + \underline{t}'C(\eta)\underline{s})$, where $\underline{t} = (1, \ldots, t^{m-1})'$, $\underline{s} = (1, \ldots, s^{m-1})'$ and $C(\eta)$ is the $m \times m$ matrix with typical element

$\left.\frac{\partial^{i+j}Q(s,t)}{\partial s^i \partial t^j}\right|_{s=t=\eta}$, $i,j = 0, \ldots, m-1$. Letting α and η tend to infinity gives μ_λ with $\lambda = \sigma^2/\sigma_s^2$ as the posterior mean of μ.

An alternative to Shiller's model, which generalizes work by Wahba (1978, 1983), is to assume that μ has the same prior distribution as the process $\Sigma_{j=0}^{m-1} \theta_j t^j + \sigma_s Z(t) + \underline{x}'\underline{\gamma}$. The polynomial coefficients are uncorrelated with Z and $\underline{\gamma}$ and have a m-variate normal distribution with mean $\underline{0}$ and variance-covariance matrix $\sigma_s^2 \delta I$. When α, $\delta \to \infty$, μ_λ with $\lambda = \sigma^2/\sigma_s^2$ is found to be the posterior mean of μ for this model as well. However, the two formulations are not equivalent in general since, for example, the covariance between $\mu(t, \underline{x})$ and $\mu(s,\underline{x}^*)$ in this case is $\sigma_s^2(Q(s,t) + \alpha\underline{x}'\underline{x}^* + \delta\underline{t}'\underline{s})$. This latter model is also of value since it makes the connection between the estimator μ_λ and polynomial regression transparent. Blight and Ott (1975) and Wahba (1978) give discussions of the motivations for this type of polynomial regression model.

It is important to note that there is a distinct difference between both Bayesian models discussed in this section and the assumption that $f \in W_2^m[a,b]$ which was employed in Section 2. It can be shown (Wahba 1983) that the sample paths of f under the Bayesian formulation cannot lie in $W_2^m[a,b]$. Nonetheless, Wahba (1983) has shown, in the case of $p = 0$ (i.e., no linear term in the model), that "standard errors" derived using a Bayesian approach can be quite useful for interval estimation of an unknown regression function from $W_2^m[a,b]$. Thus, we now provide an expression for the covariance kernel of μ under our two Bayesian models.

Let T and X be as defined in Sections 1 and 2 and set $\lambda = \sigma^2/\sigma_s^2$,

$$U = [T, X] ,$$

$$\underline{u}_t = (\underline{t}', \underline{x}')', \qquad \underline{u}_s^* = (\underline{s}', \underline{x}^{*\prime}),$$

$$\underline{Q}_t = (Q(t,t_1), \ldots, Q(t,t_n))', \quad \underline{Q}_s = (Q(s,t_1), \ldots, Q(s, t_n))'$$

and

$$Q_n = \{Q(t_i,t_j)\}_{i,j=1,n} .$$

Keeping σ^2 fixed for the moment, it can then be shown that under either of the models discussed above

$$\begin{aligned} \mathrm{Cov}(\mu(t,\underline{x}), \mu(s,\underline{x}^*)|\underline{y},\sigma^2) = (\sigma^2/\lambda) \{&Q(s,t) + \underline{u}_t' M\underline{u}_s^* \\ &- \underline{u}_t' MU'Q_n^{-1}\underline{Q}_s - \underline{Q}_t'Q_n^{-1}UM\underline{u}_s^* - \underline{Q}_t'P\underline{Q}_s\} \qquad (3.2) \\ &+ \sigma^2(\underline{u}_t'MU'Q_n^{-1} + \underline{Q}_t'P)A(\lambda)(Q_n^{-1}UM\underline{u}_s^* + P\underline{Q}_s) , \end{aligned}$$

where $A(\lambda)$ is the $n \times n$ matrix which transforms $\underline{y}$ to the vector of fitted values $(\mu_\lambda(t_1,\underline{x}_1), \ldots, \mu_\lambda(t_n,\underline{x}_n))'$,

$$M = (U'Q_n^{-1}U)^{-1}$$

and

$$P = Q_n^{-1} - Q_n^{-1}UMU'Q_n^{-1} .$$

The proof of (3.2) is similar to that of Theorem 2 in Wahba (1983). Using the covariances for μ described above, one first applies Lemma 1 of Wahba (1983) and then uses matrix identities similar to her expressions (2.12) -

(2.15) to establish the result. We note that this effectively extends her Theorem 2 to the case of univariate partial splines.

If we follow Shiller (1984) and use the noninformative prior he suggested for $h = \sigma^{-2}$ the only change in (3.2) is that σ^2 is replaced by $\hat{\sigma}^2$, the inverse of the posterior mean of h. Using properties of natural spline interpolants, it can be shown that

$$\hat{\sigma}^2 = (n-p-m)^{-1} \{ \sum_{i=1}^{n} (y_i - \mu_\lambda(t_i,\underline{x}_i))^2 + \lambda \int_a^b f_\lambda^{(m)}(t)^2 dt\} . \qquad (3.3)$$

Thus $(n-p-m)\hat{\sigma}^2$ is just criterion (2.1) evaluated at μ_λ. This "estimator" of σ^2 should be compared to $\tilde{\sigma}^2 = \Sigma_{i=1}^n (y_i - \mu_\lambda(t_i,\underline{x}_i))^2/\mathrm{tr}(I - A(\lambda))$ which is suggested by Wahba (1984a, 1985).

When m = 2, (3.2) is the covariance kernel for Shiller's estimator. In contrast to his approach, the variances and covariances are given as explicit functions of s, t, $\underline{x}$ and $\underline{x}^*$. This should be particularly useful when interval estimates are to be constructed for functionals of μ such as integrals, derivatives or evaluation at points not in the original estimation procedure.

REFERENCES

Craven, P. and Wahba, G. (1979), "Smoothing Noisy Data With Spline Functions: Estimating the Correct Degree of Smoothing by the Method of Generalized Cross-Validation," Numerische Mathematik, 31, 377-403.

Blight, B. J. N. and Ott, L. (1975), "A Bayesian Approach to Polynomial Regression," Biometrika, 62, 79-88.

Demmler, A. and Reinsch, C. (1975), "Oscillation Matrices With Spline Smoothing," Numerische Mathematik, 24, 375-382.

Engle, R. F., Granger, C. W.J., Rice, J. and Weiss, A. (1984), "Nonparametric Estimates of the Relation Between Weather and Electricity Demand," manuscript.

Kimeldorf, G. and Wahba, G. (1971), "Some Results on Tchebycheffian Spline Functions," Journal of Mathematical Analysis and Applications, 33, 82-95.

Li, K. C. (1982), "Minimaxity of the Method of Regularization on Stochastic Processes," Annals of Statistics, 10, 937-942.

Lyche, T. and Schumaker, L. (1973), "Computation of Smoothing and Interpolating Natural Splines Via Local Bases," SIAM Journal of Numerical Analysis, 10, 1027-1038.

Rice, J. (1982), "An Approach to Peak Area Estimation," Journal of Research of the National Bureau of Standards, 87, 53-65.

Shiller, R. J. (1984), "Smoothness Priors and Nonlinear Regression," Journal of the American Statistical Association, 79, 609-615.

Speckman, P. (1982), "Minimax Estimation of Linear Functionals in a Hilbert Space," Manuscript.

Utreras, F. (1979), "Cross-Validation Techniques for Smoothing Spline Functions in One or Two Dimensions," in Smoothing Techniques for Curve Estimation, eds. T. Gasser and M. Rosenblatt, Lecture Notes in Mathematics No. 757, Springer, New York, 196-232.

Wahba, G. (1978), "Improper Priors, Spline Smoothing and the Problem of Guarding Against Model Errors in Regression," Journal of the Royal Statistical Society, Series B, 40, 364-372.

Wahba, G. (1983), "Bayesian "Confidence Intervals" for the Cross-Validated Smoothing Spline," Journal of the Royal Statistical Society, Series B, 45, 133-150.

Wahba, G. (1984a), "Cross-Validated Spline Methods for the Estimation of Multivariate Functions from Data on Functionals,'" in Statistics: An Appraisal, Proceedings 50th Anniversary Conference Iowa State Statistical Laboratory, eds. H. A. David and H. T. David, Iowa State University Press, 205-235.

Wahba, G. (1984b), "Partial Spline Models for the Semiparametric Estimation of Functions of Several Variables," in Statistical Analysis of Time Series, Tokyo: Institute of Statistical Mathematics, 319-329.

Wahba, G. (1985), Comments to Peter J. Huber, Projection Pursuit, to appear in the Annals of Statistics.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER SMU-DS-TR-195	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Note on Smoothness Priors and Nonlinear Regression		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) R. L. Eubank		8. CONTRACT OR GRANT NUMBER(s) ONR-N00014-85-K-0340
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Southern Methodist University Dallas, Texas 75275		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-479
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		12. REPORT DATE December 1985
		13. NUMBER OF PAGES 14
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted for any putpose of The United States Government.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Bayesian Prediction, Nonparametric Regression, Partial Splines, Penalized Least Squares, Spline Smoothing.

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

A generalization of a nonparametric estimator proposed by Shiller (1984) is considered. The estimator is shown to be directly related to penalized least-squares estimation and spline smoothing. Some simplifications concerning the computation of Shiller's estimator and corresponding posterior covariances are indicated.