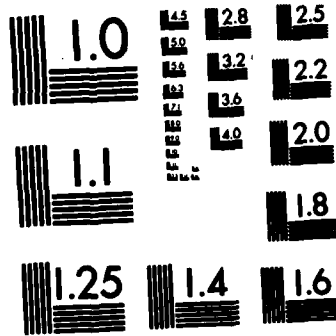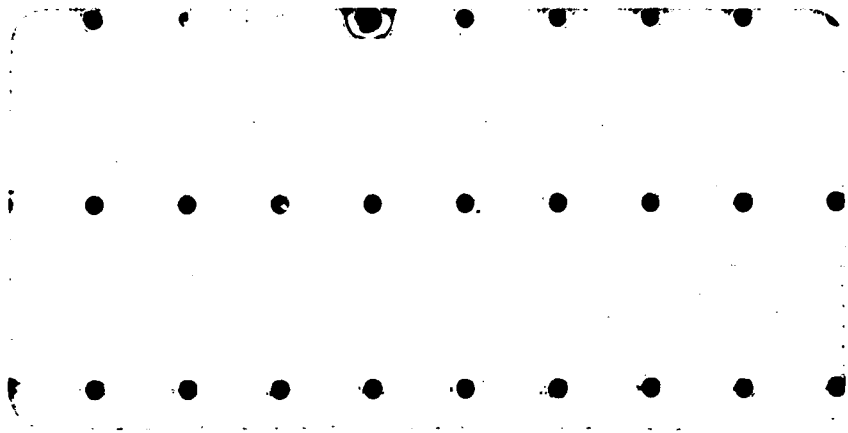MICROCOPY RESOLUTION TEST CHART
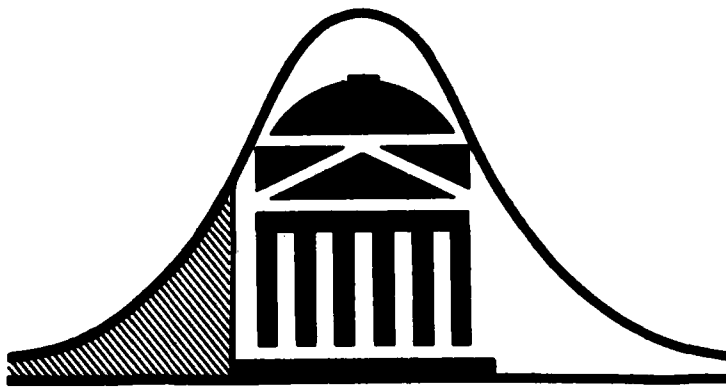
NATIONAL BUREAU OF STANDARDS-1963-A

SOUTHERN METHODIST UNIVERSITY

DTIC
ELECTE
MAR 1 1983
S D
D

DEPARTMENT OF STATISTICS

DALLAS, TEXAS 75275

83 02 023 171

ON THE RELATIONSHIP BETWEEN FUNCTIONS WITH
THE SAME OPTIMAL KNOTS IN SPLINE WITH
PIECEWISE POLYNOMIAL APPROXIMATION

by

R. L. Eubank

DEPARTMENT OF STATISTICS
Southern Methodist University
Dallas, Texas 75275

ON THE RELATIONSHIP BETWEEN FUNCTIONS WITH
THE SAME OPTIMAL KNOTS IN SPLINE AND
PIECEWISE POLYNOMIAL APPROXIMATION

by

R. L. Eubank
Department of Statistics, Southern Methodist University, Dallas, Texas

Running head. Functions with the same optimal knots.

Abstract. The relationship between functions with the same optimal knots
for $L_2[0,1]$ approximation by k-th order splines or piecewise polynomials
is investigated. It is shown that, if two functions have positive con-
tinuous k-th derivatives, they will have the same optimal knots if and
only if they differ by a polynomial of order k. An application to design
selection for continuous time regression is considered and extensions to
L_p approximation are also provided.

1. __Introduction__   Consider the $L_2[0,1]$ approximation of two functions by splines or piecewise polynomials with free knots. In this note we investigate the relationship between functions, f and g, having the same optimal knots.

Let P_N^k and S_N^k denote, respectively, the set of all k-th order piecewise polynomials and k-th order splines with N-1 distinct knots in the interior of [0,1]. For simplicity, a generic notation, f_N, will be used to indicate the best $L_2[0,1]$ approximation to f from either P_N^k or the closure of S_N^k, $\bar{S}_N^k$; its meaning will be clear from context. The properties of f_N for P_N^k have been studied by Chow [5] and Burchard and Hale [4]. Similar results for $\bar{S}_N^k$ can be found in Barrow, et al. [2] and Barrow and Smith [3]. All these papers, however, have focused on the behavior of f_N and its corresponding __optimal__ __knots__ for f in various function classes. For instance, Chow [5] has established conditions on f that insure unique optimal knots for each N whereas Barrow and Smith [3] have provided an asymptotic characterization of the behavior of the optimal knots in terms of f's k-th derivative. In constrast, this paper focuses on what is, in a certain sense, an inverse of these problems. Rather than asking how f's properties characterize the behavior of it's optimal knots we ask, instead, to what extent the optimal knots characterize f. The answer is, essentially, that two functions with the same optimal knots for all N must differ by at most a k-th order polynomial. More specifically we show the following:

__Theorem 1.__ __Let__ $f,g \in L_2[0,1] \cap C^k[0,1]$ __with__ $f^{(k)}$ __and__ $g^{(k)}$ __both__ __positive__ __on__ [0,1]. __Then__, f __and__ g __have__ __the__ __same__ __optimal__ __knots__ __for__ $L_2[0,1]$ __approximation__ __from__ P_N^k (__or__ __from__ $\bar{S}_N^k$) __for__ __all__ N __if__ __and__ __only__ __if__ __there__ __exists__ __constants__ $\alpha_0, \ldots, \alpha_k$, ($\alpha_k \neq 0$) __such__ __that__

$$f(t) = \sum_{i=0}^{k-1} \alpha_i t^i + \alpha_k g(t) \ , \ t \in [0,1]. \tag{1.1}$$

The proof of Theorem 1 is given in the next section. Although this

theorem is of interest in its own right it also has statistical applications, one of which is briefly discussed in Section 3. Another application is to the statistical problems of grouping, spacing and stratification considered by Adatia and Chan [1] who, when k=1, have obtained a special case of this result. It follows from Eubank [6] that Theorem 1 can be used to obtain their Theorem 5 under alternative conditions. Section 4 contains some concluding remarks and suggestions for future research. An extension of Theorem 1 to $L_p[0,1]$ approximation is also provided.

2. Proof of Theorem

Let us first prove Theorem 1 for approximation from $\bar{S}_N^k$ as the case of P_N^k will follow similarly. Throughout this and the subsequent section we use $||\cdot||$ to denote the usual $L_2[0,1]$ norm and D_N for the set of 'all possible' knot choices, i.e.,

$$D_N = \{(t_0,t_1,\ldots,t_N): 0=t_0 < t_1 < \cdots < t_N =1\}. \qquad (2.1)$$

For $T \in D_N$ let S_T^k denote the $L_2[0,1]$ orthogonal projector for the set of all k-th order splines having knots at T. Then, a best approximation f_N is characterized by

$$||f-f_N|| = \inf_{T \in D_N} ||f-S_T^k f||. \qquad (2.2)$$

Note that although $S_T^k f$ has noncoincident knots, the knots for f_N need not be distinct.

Now let $f(t) = \sum_{i=0}^{k-1} \alpha_i t^i + \alpha_k g(t)$ for some constants $\alpha_0,\ldots,\alpha_k$ with $\alpha_k \neq 0$. Then $||f-S_T^k f|| = |\alpha_k| \, ||g-S_T^k g||$ for all $T \in D_N$. Consequently, f and g have the same optimal knots for all N and the direct implication has been verified.

To see the necessity of condition (1.1) let $\{f_N\}$ be a sequence of best approximations to f and $0 = t_0^N < t_1^N \leq \cdots \leq t_{N-1}^N < t_N^N = 1$ the optimal knots corresponding to f_N. By assumption these will also be optimal knots for g. Define, as in Barrow and Smith [3], a piecewise linear function $\underline{T}^N \in [0,1]$ satisfying $\underline{T}^N(i/N) = t_i^N$, $i=0,\ldots,N$, and let $\underline{s}_N$ denote its right continuous inverse. It then follows from Theorem 3 of [3] that, since $f^{(k)} > 0$, as $N \to \infty$

$$\underline{s}_N(\tau) \to \int_0^\tau \{f^{(k)}(t)\}^{2/2k+1} dt / \int_0^1 \{f^{(k)}(x)\}^{2/2k+1} dx, \text{ uniformly on } [0,1]. \quad (2.3)$$

But, since the $\{t_i^N\}_{i=0}^N$ are also optimal knots for g,

$$\underline{s}_N(\tau) \to \int_0^\tau \{g^{(k)}(t)\}^{2/2k+1} dt / \int_0^1 \{g^{(k)}(x)\}^{2/2k+1} dx. \quad (2.4)$$

The result now follows by equating (2.3) and (2.4) and differentiating.

To prove Theorem 1 for P_N^k we need only note that an analog of Theorem 3 in Barrow and Smith [3] can be shown to hold, in this case, by using their proof in conjunction with Theorem 1.1 of [4]. The required result can then be obtained as in the previous arguments.

3. An application to regression design for time series. In this section an application of Theorem 1 to the problem of optimal design for continuous time regression will be presented. The situation is as follows. A stochastic process, Y, is assumed to have the form

$$Y(t) = \beta f(t) + X(t), \quad t \in [0,1], \quad (3.1)$$

where f is a known regression function, β is an unknown parameter and X is a zero mean process with covariance kernel

$$R(s,t) = (k-1)!^{-2} \int_0^1 (s-u)_+^{k-1} (t-u)_+^{k-1} du, \quad (3.2)$$

with $x_+^{k-1} = x^{k-1}$ if $x \geq 0$ and zero otherwise. The objective is to obtain an estimate of the unknown parameter β. However, suppose it is possible only to sample the Y process, as well as it k-1 quadratic mean derivatives, at some set of noncoincident time points $T \in D_N$. Thus an estimator must be based solely on the observations $Y_T^k = \{Y^{(i)}(t); t \in T, i=0,\ldots,k-1\}$. For this purpose the best linear unbiased estimator of β can be obtained through the use of generalized least squares (c.f. Sacks and Ylvisaker [10].) This estimator will be denoted by $\hat{\beta}_{k,T}$. Since the choice of T is usually at the discretion of the experimenter, the problem arises as to how T should be selected. Let $V(\hat{\beta}_{k,T})$ denote the variance of $\hat{\beta}_{k,T}$; then, an optimal

estimator can be obtained by choosing $T^* \varepsilon D_N$ to satisfy

$$V(\hat{\beta}_{k,T^*}) = \inf_{T \varepsilon D_N} V(\hat{\beta}_{k,T}). \tag{3.3}$$

Under the regularity condition that f admits the representation

$$f(t) = (k-1)!^{-1} \int_0^1 f^{(k)}(u)(t-u)_+^{k-1} du \tag{3.4}$$

for some $f^{(k)} \varepsilon L_2[0,1]$ problem (3.3) is amenable to analysis and has been studied by Sacks and Ylvisaker [10], Wahba [11], Eubank, Smith and Smith [7] and others. We now investigate, with the aid of Theorem 1, the relationship between regression models having identical optimal designs for all N.

It follows from Eubank, Smith and Smith [7] that for R of the form (3.2), $T \varepsilon D_N$ and f satisfying (3.4) we may write $V(\hat{\beta}_{k,T}) = ||P_T^k f^{(k)}||^{-2}$ where P_T^k is the $L_2[0,1]$ orthogonal projector for the set of all piecewise polynomials of order k with knots at T. As $||P_T^k f^{(k)}||^2 = ||f^{(k)}||^2 - ||f^{(k)} - P_T^k f^{(k)}||^2$ problem (3.3) can be restated as: find $(f^{(k)})_N$ such that

$$||f^{(k)} - (f^{(k)})_N|| = \inf_{T \varepsilon D_N} || f^{(k)} - P_T^k f^{(k)}|| \tag{3.5}$$

with T^* then provided by the knots for $(f^{(k)})_N$. In view of Theorem 1 this has the following immediate consequence.

<u>Corollary</u>. <u>Let</u> f_1 <u>and</u> f_2 <u>be</u> <u>two</u> <u>regression</u> <u>functions</u> <u>satisfying</u> (3.4) <u>with</u> $f_i^{(k)} \varepsilon C^k[0,1]$ <u>and</u> $f_i^{(2k)}$ <u>positive</u> <u>on</u> [0,1] <u>for</u> i = 1,2. <u>Also let</u> $\hat{\beta}_{k,T}^1$ <u>and</u> $\hat{\beta}_{k,T}^2$ <u>denote</u> <u>the</u> <u>BLUE's</u> <u>of</u> β <u>corresponding</u> <u>to</u> <u>when</u> f, <u>in</u> (3.1), <u>is</u> f_1 <u>and</u> f_2 <u>respectively</u>. <u>Then</u>, $\hat{\beta}_{k,T}^1$ <u>and</u> $\hat{\beta}_{k,T}^2$ <u>have</u> <u>the</u> <u>same</u> <u>optimal</u> <u>designs</u> <u>for</u> <u>all</u> N <u>if and only if</u>

$$f_1(t) = \sum_{i=0}^{k-1} \alpha_i t^{k+i} + \alpha_k f_2(t) , \quad t \varepsilon[0,1].$$

Thus, not only does the regression function dictate the properties of the optimal designs but, conversely, the optimal designs also characterize the regression function up to a 2k-th order polynomial. The corollary can be extended to a wider class of processes through use of results in [8].

4. **Concluding remarks**. Theorem 1 has a straightforward extension to the case of best L_p [0,1] approximation, $1 \le p \le \infty$, from P_N^k. Using the notation and results in Theorem 1.1 of [4] and the techniques of Section 2 it is possible to show the following:

Theorem 2. Assume that either $f,g \in W^{k,1}(0,1)$ or else $f,g \in W^{k,1,loc}(0,1) \cap L_p[0,1]$ and $|f^{(k)}|, |g^{(k)}|$ are monotone a.e. with $|f^{(k)}|^\sigma$ and $|g^{(k)}|^\sigma$ integrable for $\sigma = (k+p^{-1})^{-1}$. For $p = \infty$ also assume that $f,g \in C[0,1]$. If, in addition, $(f^{(k)}g^{(k)}) \ge 0$ (or $-(f^{(k)}g^{(k)}) \ge 0$, as appropriate) a.e., then f and g have the same optimal knots for all N if and only if (1.1) holds.

Note that when p=2 Theorem 2 implies Theorem 1 for P_N^k. An analog of Theorem 2 for $\bar{S}_N^k$ and $1 \le p < \infty$ can be obtained using results in [9].

More general versions of Theorems 1 and 2 are desirable. In particular, relaxation of the conditions imposed on the sign change behavior of $f^{(k)}$ and $g^{(k)}$ would be of interest. Unfortunately, the asymptotic behavior of the optimal knots serves only to distinguish between functions whose kth derivatives are not proportional in absolute value. Thus, relaxation of these restrictions will require the use of information provided by the optimal knots for finite N. To illustrate the ideas involved consider the function $f(x)=x$, $x \in [0,1]$ and define g(x) to be f(x) on [0,.5] and 1-f(x) on [.5,1]. Then, for k=1 and p=2, say, we have $|f'| = |g'|$ so that f and g are indistinguishable in terms of the asymptotic behavior of their optimal knots. However, the optimal knots can still be seen to separate these two functions by comparing the optimal knots for f when N=2, viz 1/3 and 2/3, to the breakpoints 1/4 and 3/4 for g.

References

[1] A. Adatia and L. K. Chan. Relations between stratified, grouped
     and selected order statistics samples. Scand. Actuar. J. (1981),
     193-202.

[2] D. L. Barrow, C. K. Chui, P. W. Smith, and J. D. Ward. Unicity of
     best mean approximation by second order splines with variable
     knots. Math. Comp. 32 (1978), 1131-1143.

[3] D. L. Barrow and P. W. Smith. Asymptotic properties of best $L_2[0,1]$
     approximation by splines with variable knots. Quart.
     Appl. Math. 36 (1978), 293-304.

[4] H. G. Burchard and D. F. Hale. Piecewise polynomial approximation
     on optimal meshes. J. Approx. Theory 14 (1975), 123-147.

[5] J. Chow. Uniqueness of best $L_2[0,1]$ approximation by piecewise
     polynomials with variable breakpoints. Math. Comp. (1982),
     to appear.

[6] R. L. Eubank. Optimal grouping, spacing, stratification and
     piecewise constant approximation. Tech. Rep. No.164, Department
     of Statistics, Southern Methodist Univ., Dallas, Texas (1982).

[7] R. L. Eubank, P. L. Smith and P. W. Smith. Uniqueness and eventual
     uniqueness of optimal designs in some time series models.
     Ann. Statist. 9 (1981), 486-493.

[8] R. L. Eubank, P. L. Smith and P. W. Smith. A note on optimal and
     asymptotically optimal designs for certain time series models.
     Ann. Statist. 10 (1982), to appear.

[9] D. D. Pence and P. W. Smith. Asymptotic properties of best $L_p[0,1]$
     approximation by splines. SIAM J. Math. Anal. 13 (1982), 409-420.

[10] J. Sacks and D. Ylvisaker. Designs for regression problems with
     correlated errors III. Ann. Math. Statist. 41 (1970), 2057-2074.

[11] G. Wahba. On the regression design problem of Sacks and Ylvisaker.
     Ann. Math. Statist. 42 (1971), 1035-1053.

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
167	*AD25002*	

4. TITLE (and Subtitle)	5. TYPE OF REPORT & PERIOD COVERED
ON THE RELATIONSHIP BETWEEN FUNCTIONS WITH THE SAME OPTIMAL KNOTS IN SPLINE AND PIECEWISE POLYNOMIAL APPROXIMATION	Technical Report
	6. PERFORMING ORG. REPORT NUMBER
	167

7. AUTHOR(s)	8. CONTRACT OR GRANT NUMBER(s)
RANDALL L. EUBANK	N00014-82-k-0207

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Southern Methodist University Dallas, Texas 75275	NR-042-479

11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE
	October 1982
Office of Naval Research Arlington, VA 22217	13. NUMBER OF PAGES
	6

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 purposes 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)

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

The relationship between functions with the same optimal knots for $L_2[0,1]$ approximation by k-th order splines or piecewise polynomials is investigated. It is shown that, if two functions have positive continuous k-th derivatives, they will have the same optimal knots if and only if they differ by a polynomial of order k. An application to design selection for continuous time regression is considered and extensions to L_p approximation are also provided.

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73
S/N 0102-014-6601 |

END

FILMED

3-83

DTIC