The Annals of Probability
1982, Vol. 10, No. 2, 504–507

MINIMAX LINEAR SMOOTHING FOR CAPACITIES[1]

BY H. VINCENT POOR

University of Illinois at Urbana-Champaign

Minimax linear smoothers are considered for the problem of estimating a homogeneous signal field in an additive orthogonal noise field. A minimax game with the quadratic-mean estimation error as an objective function is used to formulate this problem. Uncertainty in signal and noise field spectra is modeled using general nonparametric classes of measures proposed by Huber and Strassen for the problem of minimax hypothesis testing. These classes, which are described in terms of Choquet alternating capacities of order 2, include the conventional models for spectral uncertainty and admit a general solution to the minimax linear smoothing problem.

1. Introduction. Suppose we oberve the random field $\{Y_z;\, z \in \mathbb{R}^n\}$ given for each $z \in \mathbb{R}^n$ by $Y_z = (S_z + N_z)$ where $\{S_z;\, z \in \mathbb{R}^n\}$ and $\{N_z;\, z \in \mathbb{R}^n\}$ are orthogonal random fields, each of which is second order, homogeneous, and quadratic-mean continuous. Suppose further that h is a complex-valued Borel-measurable function on $\mathbb{R}^n$, and that $\hat{S}_z$ denotes the linear estimate of S_z based on $\{Y_z;\, z \in \mathbb{R}^n\}$ which has transfer function h. Then the quadratic-mean estimation error associated with $\hat{S}_z$ is given by

$$(1)\qquad E\{|S_z - \hat{S}_z|^2\} = (2\pi)^{-n}\left[\int_{\mathbb{R}^n} |1 - h|^2\, dm_S + \int_{\mathbb{R}^n} |h|^2\, dm_N\right] \triangleq e(h; m_S, m_N)$$

where m_S and m_N are the spectral measures on $(\mathbb{R}^n, \mathscr{B}^n)$ associated (via Bochner's theorem [10, page 245]) with $\{S_z;\, z \in \mathbb{R}^n\}$ and $\{N_z;\, z \in \mathbb{R}^n\}$, respectively. For fixed m_S and m_N, the minimum possible value of $e(h; m_S, m_N)$ is achieved by the estimate with transfer function $\hat{h} = dm_S/d(m_S + m_N)$ and this minimum value is given by $(2\pi)^{-n} \int_{\mathbb{R}^n} \hat{h}\, dm_N$. (Note that $e(h; m_S, m_N) = (2\pi)^{-n} \int_{\mathbb{R}^n} \hat{h}\, dm_N + (2\pi)^{-n} \int_{\mathbb{R}^n} |\hat{h} - h|^2\, d(m_S + m_N)$.)

If, on the other hand, m_S and m_N are known only to be in classes $\mathscr{M}_S$ and $\mathscr{M}_N$, respectively, of spectral measures on $(\mathbb{R}^n, \mathscr{B}^n)$, then a reasonable design strategy is to find a linear estimate whose transfer function minimizes $\sup_{\mathscr{M}_S \times \mathscr{M}_N} e(h;\, m_S,\, m_N)$. Such an estimate will be a minimax linear smoother for $\mathscr{M}_S$ and $\mathscr{M}_N$. Certain aspects of this problem have been considered by Kassam and Lim [5] and by the author [6]. In this paper we consider the minimax linear smoothing problem for the situation in which the measure classes $\mathscr{M}_S$ and $\mathscr{M}_N$ are of the type generated by 2-alternating capacities as considered by Huber and Strassen [4] in the context of minimax hypothesis testing. Examples of this type of class include mixtures, Prohorov and Kolmogorov (variational) neighborhoods, and other previously considered models for spectral uncertainty. Here we apply the results of Huber and Strassen to find the structure of minimax linear smoothers for general models of this type.

2. The minimax smoother for capacity classes. In the following, Ω denotes a fixed subset of $\mathbb{R}^n$, $\mathscr{A}$ denotes the Borel σ-algebra on Ω, and $\mathscr{M}$ denotes the class of all finite measures on $(\Omega, \mathscr{A})$. Recall that a finite set function v on $\mathscr{A}$ is a 2-alternating capacity (see Choquet [1]) on $(\Omega, \mathscr{A})$ if it is increasing, continuous from below, continuous from above for closed sets, and if it satisfies $v(\phi) = 0$ and $v(A \cup B) + (A \cap B) \le v(A) + v(B)$ for all

Received February 1980; revised March 1981.

[1] This research was supported by the National Science Foundation under Grant ECS-79-16453 and by the Office of Naval Research under Contract N00014-81-K-0014.

AMS 1980 *subject classifications.* Primary 62M20; secondary 60G35, 93E11.

Key words and phrases. Minimax smoothing, Choquet capacities, random fields.

$A, B \in \mathscr{A}$. For a 2-alternating capacity v on $(\Omega, \mathscr{A})$ define the set $\mathscr{M}_v$ by

(2) $$\mathscr{M}_v = \{m \in \mathscr{M} \mid m(A) \le v(A) \text{ for all } A \in \mathscr{A}, \text{ and } m(\Omega) = v(\Omega)\}.$$

A number of properties of classes of the form of (2) have been developed by Huber and Strassen [4]. Note, for example, that $\mathscr{M}_v$ is weakly compact and that, if v is a measure, then $\mathscr{M}_v = \{v\}$.

For any pair (v_0, v_1) of 2-alternating capacities on $(\Omega, \mathscr{A})$ there exists a Radon-Nikodym derivative dv_1/dv_0, introduced in [4], which has the defining property that, for each $t \in [0, \infty]$,

(3) $$r_t(\{dv_1/dv_0 > t\}) = \inf_{A \in \mathscr{A}} r_t(A)$$

where $r_t(A) \triangleq (1 + t)^{-1}[tv_0(A) + v_1(A^c)]$. This derivative (which is a family of functions having the defining property (3)) is the basis for the minimax tests between capacity classes of the form of (2) as considered in [4]. Further properties and a generalization of this derivative have been considered by Rieder [7]. In this context we state the following result which is Theorem 4.1 of [4].

LEMMA 2.1 (Huber-Strassen). *Suppose v_S and v_N are 2-alternating capacities and π_0 is a version dv_S/dv_N. Then there exist measures $q_S \in \mathscr{M}_{v_S}$ and $q_N \in \mathscr{M}_{v_N}$ such that $\pi_0 \in dq_S/dq_N$ and such that*

$$q_S(\{\pi_0 < t\}) = v_S(\{\pi_0 < t\})$$

and

$$q_N(\{\pi_0 > t\}) = v_N(\{\pi_0 > t\})$$

for all $t \in [0, \infty]$.

Let $\mathscr{H}$ denote the class of all complex-valued $\mathscr{A}$-measurable functions on Ω. Lemma 2.1 leads to the following theorem:

THEOREM 2.2. *Suppose v_S and v_N are 2-alternating capacities on $(\Omega, \mathscr{A})$. Let π_0 be a version of dv_S/dv_N and choose (q_S, q_N) as in Lemma 2.1. Define $h_0 = \pi_0(1 + \pi_0)^{-1}$. Then $[h_0, (q_S, q_N)]$ is a saddle-point solution to the game*

$$\min_{h \in \mathscr{H}} \sup_{(m_S, m_N) \in \mathscr{M}_{v_S} \times \mathscr{M}_{v_N}} e(h; m_S, m_N)$$

where e is defined in (1), and thus h_0 is a minimax linear smoother for $\mathscr{M}_{v_S}$ and $\mathscr{M}_{v_N}$.

PROOF. Noting that $h_0 \in dq_S/d(q_S + q_N)$, we have directly that

$$e(h_0; q_S, q_N) \le e(h; q_S, q_N)$$

for all $h \in \mathscr{H}$. Thus, it is sufficient to show

(4) $$e(h_0; m_S, m_N) \le e(h_0; q_S, q_N)$$

for all $(m_S, m_N) \in \mathscr{M}_{v_S} \times \mathscr{M}_{v_N}$. Lemma 2.1 asserts that π_0 is stochastically smallest over $\mathscr{M}_{v_S}$ under q_S and is stochastically largest over $\mathscr{M}_{v_N}$ under q_N. Thus, since $|1 - h_0|^2 = (1 + \pi_0)^{-2}$ is decreasing in π_0 and $|h_0|^2 = \pi_0^2(1 + \pi_0)^{-2}$ is increasing in π_0, we have

$$\int_\Omega |1 - h_0|^2 \, dm_S \le \int_\Omega |1 - h_0|^2 \, dq_S$$

and

$$\int_\Omega |h_0|^2 \, dm_N \le \int_\Omega |h_0|^2 \, dq_N$$

for all $(m_S, m_N) \in \mathscr{M}_{v_S} \times \mathscr{M}_{v_N}$. Equation (4) and hence Theorem 2.2 follow.

Concerning the pair of measures singled out by Lemma 2.1, we may also state the following property.

THEOREM 2.3. *The pair* $(q_S, q_N) \in \mathcal{M}_{v_S} \times \mathcal{M}_{v_N}$ *satisfies the conclusion of Lemma* 2.1 *if and only if it maximizes*

$$\min_{h \in \mathcal{H}} e(h; m_S, m_N) = (2\pi)^{-n} \int_\Omega [dm_S/d(m_S + m_N)]\, dm_N$$

over all $(m_S, m_N) \in \mathcal{M}_{v_S} \times \mathcal{M}_{v_N}$.

PROOF. Define $f = dm_N/d(m_S + m_N)$. Then

$$\min_{h \in \mathcal{H}} e(h; m_S, m_N) = (2\pi)^{-n} \int_\Omega f\, dm_S = (2\pi)^{-n} \int_\Omega (f - f^2)\, d(m_S + m_N).$$

Since $C(x) = (x - x^2)$ is concave and twice continuously differentiable on [0, 1], Theorem 2.3 follows from Theorem 6.1 of [4].

Thus, in view of Theorem 2.3, the pair (q_S, q_N) singled out by Lemma 2.1 can be thought of as a least-favorable pair of spectral measures for minimax linear smoothing.

3. Discussion. Theorem 2.2 gives the general solution to the minimax linear smoothing problem for signal and noise uncertainty classes of the form of (2). Several useful examples of classes of this type are given by Huber and Strassen in [4], and other useful examples are given by Rieder [7], Strassen [8], and Vastola and Poor [9]. Some of the most commonly used examples of classes of the form $\mathcal{M}_v$ can be written as ε-neighborhoods of some nominal measure μ. Examples of capacity classes that have this structure are contaminated mixtures, variational neighborhoods, and Prohorov neighborhoods (see [4]). For this type of class, an uncertainty model will consist of a nominal pair (μ_S, μ_N) of signal and noise spectral measures with respective degrees ε_S and ε_N of uncertainty placed on the nominal measures. The derivative between capacities generating classes of this type is often of the form (see Huber [2, 3] and Rieder [7])

(5) $$\pi_0(\omega) = \max\{c', \min\{c'', \lambda(\omega)\}\}, \qquad w \in \Omega,$$

where λ is the Radon-Nikodym derivative between the nominal pair of measures (i.e., $\lambda \in d\mu_S/d\mu_N$) and c' and c'' are nonnegative constants with $c' \le c''$. If π_0 of (5) is a version of dv_S/dv_N, then Theorem 2.2 implies that a minimax linear smoother for $\mathcal{M}_{v_S}$ and $\mathcal{M}_{v_N}$ is given by

(6) $$h_0(\omega) = \max\{k', \min\{k'', h'(\omega)\}\}, \qquad \omega \in \Omega$$

where $k' = c'/(1 + c')$, $k'' = c''/(1 + c'')$ and $h' = \lambda/(1 + \lambda)$. Note that h' is the optimum smoother for the nominal model, and thus the minimax linear smoother for this case desensitizes the nominal smoother (to a degree depending on ε_S and ε_N) in those spectral regions where either μ_S or μ_N is dominant (i.e., where h' is near 1 or is near 0).

In the situations for which (5) is valid, (6) gives the transfer function of the minimax linear smoother. Suppose, for example, that $n = 1$, $\Omega = [-b, b]$ for some $b < \infty$, $c' < c''$, and h' is symmetric about $\omega = 0$ and is strictly decreasing on $[0, b]$. Then the minimax linear estimate of S_z determined by h_0 is given explicitly by

$$\hat{S}_z = \int_{-\infty}^{\infty} \bar{h}_0(z - t) Y_t\, dt$$

where $\bar{h}_0 \triangleq \mathscr{F}^{-1}\{h_0\}$ is given by

$$\bar{h}_0(t) = \bar{h}'(t) + k' \frac{[\sin(bt) - \sin(a't)]}{\pi t} + k'' \frac{\sin(a''t)}{\pi t}$$

$$- \int_{-\infty}^{\infty} \bar{h}'(t-\tau)[\sin(b\tau) - \sin(a'\tau) + \sin(a''\tau)](\pi\tau)^{-1}\, d\tau$$

with $\bar{h} = \mathscr{F}^{-1}\{h'\}$ and with a' [resp., a''] the positive solution to $h'(a') = k'$ [resp., $h'(a'') = k''$].

As a final comment we note that, although we assumed initially that the observation field was a continuous-parameter field, Theorems 2.2 and 2.3 are also directly applicable to the case in which the observation field is a discrete-parameter field (i.e., in which the time set is $\mathbb{Z}^n$) since this latter situations corresponds to the particular case of the analysis of Section 2 in which $\Omega = [-\pi, \pi]^n$.

4. Acknowledgment. The author would like to thank K. S. Vastola for several stimulating discussions concerning the techniques used in this paper. The helpful comments and suggestions of the anonymous referee also are gratefully acknowledged.

REFERENCES

[1] CHOQUET, G. (1953/54). Theory of capacities. *Ann. Inst. Fourier* **5** 131–292.
[2] HUBER, P. J. (1965). A robust version of the probability ratio test. *Ann. Math. Statist.* **36** 1753–1758.
[3] HUBER, P. J. (1968). Robust confidence limits. *Z. Wahrsch. verw. Gebiete* **10** 269–278.
[4] HUBER, P. J. and STRASSEN, V. (1973). Minimax tests and the Neyman-Pearson lemma for capacities. *Ann. Statist.* **1** 251–263.
[5] KASSAM, S. A. and LIM, T. L. (1977). Robust Wiener filters. *J. Franklin Inst.* **304** 171–185.
[6] POOR, H. V. (1980). On robust Wiener filtering. *IEEE Trans. Automat. Control* **AC-25** 521–526.
[7] RIEDER, H. (1977). Least favorable pairs for special capacities. *Ann. Statist.* **5** 909–921.
[8] STRASSEN, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist.* **36** 423–439.
[9] VASTOLA, K. S. and POOR, H. V. (1980). On generalized band models in robust detection and filtering. *Proc. 1980 Conf. on Inform. Sciences and Systems*, Princeton University, March 1980, 1–5.
[10] WONG, E. (1971). *Stochastic Processes in Information and Dynamical Systems.* McGraw-Hill, New York.

COORDINATED SCIENCE LABORATORY
UNIVERSITY OF ILLINOIS
1101 W. SPRINGFIELD AVENUE
URBANA, IL 61801