The Annals of Statistics
1984, Vol. 12, No. 4, 1448–1466

ALL ADMISSIBLE LINEAR ESTIMATORS OF THE MEAN OF A GAUSSIAN DISTRIBUTION ON A HILBERT SPACE[1]

BY AVI MANDELBAUM

Cornell University

We consider linear estimators for the mean θ of a Gaussian distribution $N(\theta, C)$ on a Hilbert space, when the covariance operator C is known. It was argued in a previous work that the natural class of linear estimators is the class of measurable linear transformations. Using the simplest quadratic loss we prove that the linear estimator L is admissible if and only if the operator $C^{-1/2}LC^{1/2}$ is Hilbert-Schmidt, self-adjoint, its eigenvalues are all between 0 and 1 and two are equal to 1 at the most. As an application of the general theory, we investigate some linear estimators for the drift function of a Brownian motion.

1. Introduction.

1.1. Let $X = (X_1, \cdots, X_k)'$ be a Gaussian random vector with unknown mean $\theta = (\theta_1, \cdots, \theta_k)'$ and a known invertible covariance matrix C. Consider the problem of estimating θ under the loss

$$L(\theta, a) = (\theta - a)'C^{-1}(\theta - a). \tag{1.1}$$

Let L be a $k \times k$ real matrix. It follows from the results of Cohen [4] that

1.1.A. The linear estimator LX is admissible if and only if the matrix $C^{-1/2}LC^{1/2}$ is symmetric, its eigenvalues are all between 0 and 1 with equality at 1 for at most two of them.

Now consider the problem of estimating the mean θ of a Gaussian random vector X with values in an infinite dimensional separable Hilbert space. For example, X could be a Brownian motion with drift function θ and the loss (1.1) takes the form $\int [\theta'(t) - a'(t)]^2 dt$. (This special case actually inspired our study and is treated in Section 6). In a previous paper [11] it was argued that the natural class of linear estimators for θ is the class of measurable linear transformations on the Hilbert space. Our main result (Theorem 1 in Subsection 2.3) is the infinite dimensional analogue of 1.1.A. The necessity part is proved along the lines of [4]. The sufficiency is proved by applying Blyth's method to an equivalent problem in which one estimates infinite number of means $(\theta_1, \theta_2, \cdots)$ from independent normal observations $(X_1, X_2, \cdots)$.

1.2. The main difference between the finite dimensional problem and our infinite dimensional model is that in the latter we consider a parameter space

Received April 1984.
[1] Research supported in part by NSF Grants MCS 82-00031 and MCS 82-02286.
AMS 1980 *subject classifications*. Primary 62C15, 62C07, 62H12; secondary 60G15, 62H99.
Key words and phrases. Linear estimators, admissibility, Gaussian measures on a Hilbert space, measurable linear transformations, Brownian motion, Orenstein-Uhlenbeck process.

which is much smaller than the sample space. Another difference is that admissibility depends on the quadratic loss being used. For example, under the loss function analogue to (1.1), the usual estimator $\delta(X) = X$ is trivially inadmissible since its risk function is identically infinite. Admissibility with respect to the analogue loss of

$$L(\theta, a) = (\theta - a)'(\theta - a) \tag{1.2}$$

is harder to settle. We do not know, for example, if the estimators $\delta(X) = \alpha X$, $0 < \alpha < 1$ are admissible in the infinite dimensional model (see Subsections 7.2, 7.3 for further discussion).

1.3. In Section 2 we state our problem in decision theoretic terms and we formulate the main theorem. The necessity part of the main theorem is proved in Section 3 and the sufficiency in Section 5. In Section 4 the general model is reduced to an equivalent discrete model. The sufficiency proof of the main theorem actually follows from a more general admissibility result in the discrete set up (Theorem 3) which has little to do with the specific Gaussian setting. The general results are applied in Section 6 to investigate some linear estimators for the drift function of a Brownian motion. Finally, we discuss some generalizations and open problems in Section 7.

1.4. The structure of measurable linear transformations with respect to Gaussian measure was given in [11], as well as the facts about Gaussian measures on a Hilbert space which we use without proof. Additional helpful references are [8] and [12]. In the Hilbert space set up, admissibility of Bayes estimators under bounded subconvex loss function was considered in Le Cam [9]. The technique of reducing the general model to the one with a countable infinite number of observations has been widely used (see Grenander [7], for example). In a decision theoretic framework, the technique was used in Berger and Wolpert [1] who considered James-Stein estimators for the mean function of a Gaussian process under a quadratic weighted loss.

2. The main results.

2.1. The standard elements of an estimation problem are: a sample space $\mathscr{X}$, an action space $\mathscr{A}$, a parameter space Θ, a parametrized set of possible distributions $\{P_\theta, \theta \in \Theta\}$ and a loss function $L(\theta, a)$, $\theta \in \Theta$, $a \in \mathscr{A}$. Let H be a separable Hilbert space with inner product $(\cdot\,,\,\cdot)$ and norm $|\cdot|$. Denote by $N(\theta, C)$ the Gaussian distribution on H with mean θ and covariance operator C. The operator $C{:}H \to H$ must be linear compact operator which is positive semi-definite, self adjoint and trace class. We study the problem of estimating θ when C is known. Put in decision theoretic terms, both $\mathscr{X}$ and $\mathscr{A}$ equal H. The parameter space Θ consists of all θ for which the probability measure $P_\theta = N(\theta, C)$ is equivalent to $P_0 = N(0, C)$ (i.e., P_0 and P_θ have the same null sets). It is well known that

$$\Theta = C^{1/2}(H),$$

where $C^{1/2}$ is the positive square root of the operator C. Assume that C is injective.

Then the subspace Θ is itself a Hilbert space with the norm

$$\|\theta\| = |C^{-1/2}\theta|, \quad \theta \in \Theta. \tag{2.1}$$

The loss function we consider (in analogue to (1.1)) is

$$\begin{aligned} L(\theta, a) &= \|\theta - a\|^2, \quad \theta, a \in \Theta, \\ &= \infty \qquad\qquad \theta \in \Theta, \quad a \in H\backslash\Theta. \end{aligned} \tag{2.2}$$

A non-randomized *estimator* $\delta: H \to H$ is a mapping which is measurable with respect to the Borel σ-algebra of H. An estimator is evaluated by its risk function

$$R(\theta, \delta) = \int_H L(\theta, \delta(x))P_\theta(dx).$$

It is possible to express $R(\theta, \delta)$ in terms of P_0 by

$$R(\theta, \delta) = \int_H L(\theta, \delta(x + \theta))P_0(dx).$$

Estimators are identified if they are equal almost surely with respect to $P_0(P_0 - \text{a.s.})$. An estimator δ' is *as good as* δ if

$$R(\theta, \delta') \le R(\theta, \delta), \quad \theta \in \Theta. \tag{2.3}$$

An estimator δ' is *better* than δ if it is as good as δ and if (2.3) holds with strict inequality for at least one $\theta \in \Theta$. An estimator is *inadmissible* if there exists a better estimator, and it is *admissible* otherwise.

2.2. A *linear estimator* L is a nonrandomized estimator which is linear on a measurable subspace D_L with $P_0(D_L) = 1$. Using the terminology of [11], a linear estimator is a *measurable linear transformation* on H with respect to $P_0(P_0 - \text{mlt})$, and D_L is the *domain* of L. It is proved in [11] that L is a P_0-mlt if and only if the operator

$$T = LC^{1/2}: H \to H$$

is a linear operator which is Hilbert-Schmidt. Moreover, the Hilbert-Schmidt norm of T, $\|T\|_{\text{HS}}$, is given by

$$\|T\|_{\text{HS}}^2 = \int_H |Lx|^2 P_0(dx), \tag{2.4}$$

and the formal relations

$$(Lx, h) = (TC^{-1/2}x, h) = (C^{-1/2}x, T^*h), \quad h \in H$$

do have a rigorous interpretation (which is used in (3.9)).

2.3. The class of admissible linear estimators is described in the following

THEOREM 1. *A linear estimator L is admissible under the loss* (2.2) *if and only* *if the operator*

$$A = C^{-1/2}LC^{1/2}:H \to H \tag{2.5}$$

is Hilbert-Schmidt, self-adjoint and its eigenvalues are all between 0 *and* 1 *with two eigenvalues at the most equal to* 1.

REMARKS. (i) A Hilbert-Schmidt operator which is not self-adjoint need not have eigenvalues. Hence, in proving necessity, one needs to establish the self-adjointness of A before referring to eigenvalues.

(ii) When starting with a linear estimator for which A is not self-adjoint, the proof of necessity for Theorem 1 exhibits a better estimator for which A is self-adjoint.

(iii) If L is admissible, the operator $T = LC^{1/2}$ is trace-class and not only Hilbert-Schmidt. This follows from the facts that a product of two Hilbert-Schmidt operators is a trace class operator and that $C^{1/2}$ is Hilbert-Schmidt.

(iv) We shall prove that $L(\Theta) \in \Theta$ when L is admissible. Let

$$L_\Theta:\Theta \to \Theta$$

denote the restriction of L to Θ and consider Θ as a Hilbert space with the norm (2.1). Then, Theorem 1 can be reformulated as

THEOREM 1′. *A linear estimator is admissible if and only if the linear operator* L_Θ *has the same properties as the linear operator* A *in Theorem* 1.

We prove Theorem 1 in Sections 3 and 5.

3. Proof of necessity in Theorem 1.

3.1. Let L be an admissible linear estimator with domain D_L. First we prove that $A = C^{-1/2}LC^{1/2}$ is Hilbert-Schmidt. Then we show that A is self-adjoint. This guarantees that A has a spectral representation in terms of its eigenvalues from which we conclude that $0 \le A \le I$. Finally, we use a James-Stein estimator to improve upon L if three or more eigenvalues equal 1.

LEMMA 3.1. *If* $\delta:H \to H$ *is admissible under the loss* (2.2) *then*

$$\delta(x) \in \Theta \quad P_0 - a.s. \tag{3.1}$$

PROOF. There exist estimators with finite risk (for example, $\delta(x) \equiv 0$). Hence, $R(\theta, \delta) < \infty$ for some $\theta \in \Theta$ implying that (3.1) holds P_θ-a.s. Since P_θ and P_0 are equivalent, Lemma 3.1 follows.

We now prove that A is Hilbert-Schmidt. Let $\{e_i\}$ be an orthonormal basis of eigenvectors of C, then

$$\Theta = \{x \in H, \textstyle\sum_i (x, C^{-1/2}e_i)^2 < \infty\}. \tag{3.2}$$

Applying Lemma 3.1 to L, the P_0 measure of the measurable subspace

$D_L \cap L^{-1}(\Theta)$ equals 1. A measurable subspace with positive measure must contain Θ (see [12], page 142, for example), thus $L(\Theta) \subseteq \Theta$. Equivalently,

$$T(H) \subseteq \Theta,$$

which shows that $A = C^{-1/2}T$ is defined on all H. Let

$$Z_i = (Lx, C^{-1/2}e_i), \quad i = 1, 2, \cdots.$$

It follows from Lemma 3.1 and (3.2) that $\sum_i Z_i^2 < \infty$ P_0-a.s. Furthermore Lemma 1 in [11] shows that the variables Z_i are jointly normal with mean 0 and

$$EZ_i^2 = |T^*C^{-1/2}e_i|^2.$$

LEMMA 3.2. *Let $Z_1, Z_2, \cdots$ be jointly normal with mean 0. If $\sum_i Z_i^2 < \infty$ then $\sum_i EZ_i^2 < \infty$.*

PROOF. The vector $Z = (Z_1, Z_2, \cdots)$ is a centered Gaussian vector with values in the Hilbert space of square summable sequences. The sum $\sum_i EZ_i^2$ is the trace of the covariance of Z, hence it is finite.

We now have

$$\sum_i (T^*C^{-1/2}e_i, T^*C^{-1/2}e_i) < \infty.$$

That is, $T^*C^{-1/2}: H \to H$ is Hilbert-Schmidt and $A = (T^*C^{-1/2})^*$ must be as well.

3.2. To show that A is self-adjoint and $0 \le A \le I$, we construct (in Subsection 3.3) estimators that have these properties and are as good as L. Then we use the following lemma due to Farrell (see [3]).

LEMMA 3.3. *Let δ be an admissible estimator. If δ' is as good as δ, then $\delta' = \delta$ P_0-a.s.*

PROOF. As in Lemma 3.1, the risk of δ' is not identically infinite and $\delta'(x) \in \Theta$ P_0-a.s. If $a'' = \frac{1}{2}(a + a')$, $a' \neq a$, then

$$\|\theta - a''\|^2 < \tfrac{1}{2}\|\theta - a\|^2 + \tfrac{1}{2}\|\theta - a'\|^2.$$

Hence the estimator $\delta''(x) = \frac{1}{2}[\delta(x) + \delta'(x)]$ is better than δ unless $\delta' = \delta$ P_0-a.s.

Let $\tilde{L} = AC^{-1/2}$ be the P_0-mlt associated with the Hilbert-Schmidt operator A. Then $\tilde{L} = C^{-1/2}L$ and

$$\begin{aligned} R(\theta, L) &= \int_H |\tilde{L}(x + \theta) - C^{-1/2}\theta|^2 P_0(dx) \\ &= \int_H |\tilde{L}(x)|^2 P_0(dx) - 2\int (\tilde{L}x, (A - I)C^{-1/2}\theta) P_0(dx) \\ &\quad + |(A - I)C^{-1/2}\theta|^2. \end{aligned}$$

Applying (2.4) to $\tilde{L}$, we obtain an expression for the risk function of a linear estimator L:

$$R(\theta, L) = \|A\|_{HS}^2 + |(A - I)C^{-1/2}\theta|^2, \quad \theta \in \Theta. \tag{3.3}$$

3.3. The absolute value of a bounded linear operator $S:H \to H$ is defined as the self-adjoint operator

$$|S| = (S^*S)^{1/2},$$

(there should be no confusion with $|h|$, $h \in H$, since the absolute value is defined for operators only).

The following lemma is proved in Subsection 3.5:

LEMMA 3.4. *Let A be Hilbert-Schmidt. Then the operator*

$$\varphi(A) = I - |I - A|$$

is Hilbert-Schmidt, self-adjoint, $\varphi(A) \le I$ and

$$\|\varphi(A)\|_{HS}^2 \le \|A\|_{HS}^2, \tag{3.4}$$

$$|(\varphi(A) - I)h|^2 = |(A - I)h|^2, \quad h \in H. \tag{3.5}$$

Define a P_0-mlt L' by

$$L' = C^{1/2}\varphi(A)C^{-1/2}.$$

Using the expression (3.3), we get from Lemma 3.4 that L' is as good as L. By Lemma 3.3, $A = \varphi(A)$, proving that A is self-adjoint and $A \le I$. We prove the following in Subsection 3.6:

LEMMA 3.5. *Let A be self-adjoint and Hilbert-Schmidt. Then*

$$\psi(A) = |A|$$

is Hilbert-Schmidt, $\psi(A) \ge 0$

$$\|\psi(A)\|_{HS}^2 = \|A\|_{HS}^2, \tag{3.6}$$

$$|(\psi(A) - I)h|^2 \le |(A - I)h|^2, \quad h \in H. \tag{3.7}$$

Define a P_0-mlt L'' by

$$L'' = C^{1/2}\psi(A)C^{-1/2}.$$

By Lemma 3.5, L'' is as good as L, hence $A = \psi(A)$ which shows that $A \ge 0$.

3.4. We have proved that the operator A is Hilbert-Schmidt, self-adjoint and $0 \le A \le I$. A Hilbert-Schmidt operator is compact. By the spectral theorem for compact self-adjoint operators (see Gohberg and Goldberg [6], page 113, for example), there exists an orthonormal system $\{a_i\}$ of eigenvectors of A and corresponding real eigenvalues $[\alpha_i\}$ such that

$$Ah = \textstyle\sum_i \alpha_i(h, a_i)a_i, \quad h \in H. \tag{3.8}$$

The eigenvalues $\{\alpha_i\}$ are square summable ($\sum_i \alpha_i^2 = \| A \|_{\text{HS}}^2$) and $0 \le \alpha_i \le 1$, $i = 1, 2, \cdots$. Applying Lemma 1 of [11] to the P_0-mlt $\tilde{L} = AC^{-1/2}$, we obtain that P_0-a.s.

$$\tilde{L}x = \textstyle\sum_i (\tilde{L}x, a_i)a_i = \sum_i (C^{-1/2}x, A^*a_i) = \sum_i \alpha_i(C^{-1/2}x, a_i)a_i \tag{3.9}$$

where

$$\{(C^{-1/2}x, h), h \in H\} \tag{3.10}$$

is the white noise over H constructed in Subsection 2.6 of [11]. The random variables

$$Z_i = (C^{-1/2}x, a_i), \quad i = 1, 2, \cdots$$

are iid $N(0, 1)$ under the measure P_0. The P_0-mlt $L = C^{1/2}\tilde{L}$ has the representation

$$Lx = \textstyle\sum_i \alpha_i Z_i C^{1/2} a_i,$$

and

$$R(\theta, L) = \textstyle\sum_i \alpha_i^2 + \sum_i (1 - \alpha_i)^2(C^{-1/2}\theta, a_i).$$

Now suppose that three or more of the eigenvalues equal 1. Without loss of generality, let $\alpha_1 = \alpha_2 = \alpha_3 = 1$. Define

$$\delta_i(x) = (1 - (1/\textstyle\sum_{k=1}^3 Z_k^2))Z_i, \quad i = 1, 2, 3,$$

$$\delta_i(x) = \alpha_i Z_i, \qquad i = 4, 5, \cdots.$$

The James-Stein estimator

$$\delta(x) = \textstyle\sum_i \delta_i(x) C^{1/2} a_i \tag{3.11}$$

is better than L (see Stein [15], for example), contradicting the admissibility of L. This establishes the necessity part of Theorem 1.

REMARK. A representation of the form (3.11) is described in detail in Theorem 2, Subsection 4.3 (see also Remark (i) following it). Explicit calculations of the risk function in terms of this representation are given in the proof of Lemma 4.1 (see also [1]).

3.5. PROOF OF LEMMA 3.4. The equality (3.5) is a consequence of the relation

$$(\varphi(A) - I)^*(\varphi(A) - I) = (A - I)^*(A - I).$$

To prove that $\varphi(A)$ is Hilbert-Schmidt, we note that the operator $I + |I - A|$ has a bounded inverse (using the self-adjointness of $|I - A|$) and

$$\varphi(A) = [A^* + A - A^*A] \cdot [I + |I - A|]^{-1}.$$

Since $A^* + A - A^*A$ is Hilbert-Schmidt, $\varphi(A)$ is as well.

We now prove (3.4). It suffices to show that

$$((I - A)e_i, e_i) \le (|I - A| e_i, e_i)$$

for some orthonormal basis $\{e_i\}$ in H. By modifying the proof of Theorem 1.4 in

[8], one can write $I - A$ as UT where U is an isometry on the range of T, T is of the form

$$Th = \sum_i t_i(h, e_i)e_i, \quad h \in H,$$

e_i is an orthonormal basis in H and $t_i \geq 0$. Since T is self-adjoint and positive semi-definite, $T = (T^*U^*UT)^{1/2}$. Now (3.12) reduces to

$$(UTe_i, e_i) \leq (Te_i, e_i)$$

which follows from the relations $Te_i = t_ie_i$, $t_i \geq 0$, $(Ue_i, e_i) \leq 1$.

REMARK. Cohen's proof does not extend directly to the infinite dimensional situation because the right-hand side of the inequality (2.7) in [4] is infinite. However, it is possible to show that the mapping $\varphi(\cdot)$ is continuous in the Hilbert-Schmidt norm and to get (3.4) from [4] by a limiting argument.

3.6. PROOF OF LEMMA 3.5. If A is given by (3.8) then

$$\psi(A)h = \sum_i |\alpha_i| (h, a_i)a_i, \quad h \in H.$$

The Hilbert-Schmidt norm of A is given by

$$\| A \|_{\mathrm{HS}}^2 = \sum_i \alpha_i^2 = \| \psi(A) \|_{\mathrm{HS}}^2,$$

which proves (3.6). Now

$$(Ah, h) = \sum_i \alpha_i(h, a_i)^2 \leq \sum_i |\alpha_i| (h, a_i)^2 = (\psi(A)h, h), \quad h \in H,$$

which proves (3.7).

4. The discrete model.

4.1. Let $X_1, X_2, \cdots$ be independent normal random variables where X_i has mean θ_i and variance 1. Assume that

$$\sum_i \theta_i^2 < \infty. \tag{4.1}$$

Consider the problem of estimating $\theta^d = (\theta_1, \theta_2, \cdots)$ using the observation $X^d = (X_1, X_2, \cdots)$. The loss $L_d(\theta^d, a^d)$ in taking an action $a^d = (a_1, a_2, \cdots)$ is given by

$$L_d(\theta^d, a^d) = \sum_{i=1}^{\infty} (\theta_i - a_i)^2, \quad a^d \in \ell^2$$
$$= \infty, \qquad \text{otherwise},$$

where ℓ^2 is the space of the sequences which satisfy (4.1). An estimator δ^d is described by a sequence

$$\delta^d(x^d) = (\delta_1(x^d), \delta_2(x^d), \cdots), \quad x^d = (x_1, x_2, \cdots) \in \mathbb{R}^\infty$$

where $\delta_i(\cdot)$ are real-valued measurable functions with respect to the Kolmogorov

σ-algebra in $\mathbb{R}^\infty$. The risk of an estimator in the discrete problem will be denoted by $R_d(\theta^d, \delta^d)$. An estimator δ^d that is admissible is called D-admissible.

4.2. One can regard the discrete model as a special case of the model described in Subsection 2.1. Indeed, H can be taken as the space of sequences $h = (h_1, h_2, \cdots)$ such that

$$|h|^2 = \sum_i \lambda_i h_i^2 < \infty,$$

where $\{\lambda_i > 0\}$ is a fixed sequence of weights with $\sum_i \lambda_i < \infty$. On H consider the covariance operator

$$Ch = (\lambda_1 h_1, \lambda_2 h_2, \cdots).$$

Then $\Theta = C^{1/2}(H) = \ell^2$ and $\|\theta\|^2 = \sum_{i=1}^\infty \theta_i^2$. For $\theta \in \ell^2$, let P_θ stand for the Gaussian distribution on H with mean θ and covariance C. Then P_θ is the distribution of the sequence $X_1, X_2, \cdots$ described at the beginning of Subsection 4.1, viewed as an H-valued Gaussian vector. The loss (1.1) is now $\sum_i (\theta_i - a_i)^2$ (while (1.2) is $\sum_i \lambda_i(\theta_i - a_i)^2$).

4.3. We now show that the discrete model is statistically isomorphic to the model described in Subsection 2.1. Define a measurable mapping $\mathscr{D}: H \to \mathbb{R}^\infty$ by

$$\mathscr{D}x = ((C^{-1/2}x, e_1), (C^{-1/2}x, e_2), \cdots) \tag{4.2}$$

where $\{e_i\}$ is some orthonormal basis in H, and $\{(C^{-1/2}x, e_i)\}$ are as in (3.10).

Theorem 2. *An estimator δ^d is D-admissible if and only if the estimator $\delta: H \to H$ given by*

$$\delta(x) = \sum_i \delta_i(\mathscr{D}x) C^{1/2} e_i \tag{4.3}$$

is admissible in the model described in Subsection 2.1.

Remarks. (i) The basis $\{e_i\}$ is arbitrary. Hence, there are many discrete models that are isomorphic to our original model. A special basis was used in constructing the James-Stein estimator (3.11). This is further discussed in Subsection 4.5.

(ii) We only prove that if δ^d is D-admissible then δ is admissible. We never use the converse, the proof of which is very similar.

4.4. The proof of Theorem 2 is based on the following two lemmas.

Lemma 4.1. *If δ^d and δ are related by* (4.3) *and $\delta^d \in \ell^2$ a.s., then*

$$R(\theta, \delta) = R_d(D\theta, \delta^d) \tag{4.4}$$

where $D: \Theta \to \ell^2$ is the isometry onto ℓ^2 given by

$$D\theta = ((C^{-1/2}\theta, e_1), (C^{-1/2}\theta, e_2), \cdots).$$

Proof. The Hilbert space structure on ℓ^2 is the usual one. We have

$$\|\theta\|^2 = |C^{-1/2}\theta|^2 = \sum_i (C^{-1/2}\theta, e_i)^2,$$

hence D is an isometry, which is clearly onto. To prove (4.4), note that $\delta(x) \in \Theta$ P_0-a.s. and

$$\begin{aligned} L(\theta, \delta(x)) &= \|\delta(x) - \theta\|^2 = \sum_i [(C^{-1/2}\delta(x), e_i) - (C^{-1/2}\theta, e_i)]^2 \\ &= L_d(D\theta, \delta^d(\mathscr{D}x)) \end{aligned}$$

which yields

$$\begin{aligned} R(\theta, \delta) &= \int_H L_d(D\theta, \delta^d(\mathscr{D}x))P_\theta(dx) \\ &= \int_{\mathbb{R}^\infty} L_d(D\theta, \delta^d(x^d))dP_\theta \circ \mathscr{D}^{-1}. \end{aligned} \tag{4.5}$$

We now show that

$$P_\theta \circ \mathscr{D}^{-1} = \prod_i N(\theta_i, 1) \tag{4.6}$$

where $\theta^d = (\theta_1, \theta_2, \cdots) = D\theta$. Then (4.4) will follow from (4.5) immediately.

Let $(\lambda_1, \lambda_2, \cdots)$ be a sequence with finitely many λ_i nonzero, then

$$\begin{aligned} \int_{\mathbb{R}^\infty} \exp(\textstyle\sum_i \lambda_i x_i)\, dP_\theta \circ \mathscr{D}^{-1} &= \int_H \exp(\textstyle\sum_i \lambda_i(C^{-1/2}x, e_i))P_\theta(dx) \\ &= \textstyle\prod_i \exp(\lambda_i\theta_i + (1/2)\lambda_i^2), \end{aligned} \tag{4.7}$$

where the first equality follows from a change of variables and the last equality holds since under P_θ, the random variables $(C^{-1/2}x, e_i)$ are independent normal with mean θ_i and variance 1. The relation (4.7) identifies $P_\theta \circ \mathscr{D}^{-1}$ as the product measure in (4.6), and Lemma 4.1 is established.

Lemma 4.2. *Let $\delta: H \to H$ be an estimator such that $\delta \in \Theta$ P_0 − a.s. Then there exists an estimator δ^d which is related to δ by* (4.3).

Proof. Since $C^{-1/2}\delta$ is well defined P_0-a.s. we have

$$C^{-1/2}\delta(x) = \sum_i (C^{-1/2}\delta(x), e_i)e_i.$$

Hence

$$\delta(x) = \sum_i e_i(x)C^{1/2}e_i,$$

where $e_i(x) = (C^{-1/2}\delta(x), e_i): H \to \mathbb{R}^1$ are measurable functions. The Borel σ-algebra in H is equal to the σ-algebra generated by the mapping $\mathscr{D}$ defined in (4.2). This guarantees the existence of measurable functions $\delta_i: \mathbb{R}^\infty \to \mathbb{R}^1$ such that $e_i(x) = \delta_i(\mathscr{D}x)$, thus (4.3) holds.

Proof of Theorem 2. Assume that δ^d is admissible. As in the proof of Lemma 3.1, there exists θ^d such that $R_d(\theta^d, \delta^d) < \infty$ and $\delta^d \in \ell^2$ a.s. By Lemma

4.1, $R(\theta, \delta) < \infty$ where $\theta^d = D\theta$. If δ is inadmissible, there exists a $\tilde{\delta}$ which is better and also $\tilde{\delta} \in \Theta$ P_0-a.s. Let $\tilde{\delta}^d$ be the estimator associated with $\tilde{\delta}$ via Lemma 4.2. By Lemma 4.1, $\tilde{\delta}^d$ is better than δ^d, contradicting the admissibility of δ^d.

4.5. We now describe the simplest discrete analogue of an admissible linear estimator L. The operator $A = C^{-1/2}LC^{1/2}$ has the spectral decomposition (3.8). Without loss of generality we assume that $\{a_i\}$ is an orthonormal basis of H and this basis will be used to construct the discrete estimator L^d. As in (3.9) and as in the proof of Lemma 4.2:

$$e_i(x) = (C^{-1/2}x, A^*a_i) = \alpha_i(C^{-1/2}x, a_i),$$

which identifies δ^d as

$$\delta^d(x^d) = (\alpha_1 x_1, \alpha_2 x_2, \cdots). \tag{4.8}$$

From Theorems 1 and 2 we obtain

COROLLARY 4.3. *If δ^d is D-admissible, then*

$$\textstyle\sum_i \alpha_i^2 < \infty, \ 0 \le \alpha_i \le 1, \ i = 1, 2, \cdots$$

and two α_i at the most are equal to 1.

The next section is devoted to establishing the converse to Corollary 4.3.

5. Proof of sufficiency in Theorem 1.

5.1. Let $A: H \to H$ be a linear operator which has the properties described in Theorem 1. We prove that the P_0-mlt $L = C^{1/2}AC^{-1/2}$ is an admissible linear estimator. By Theorem 2 and the discussion in Subsection 4.4 it is sufficient to prove that

5.1.A. The estimator

$$\delta^d(x) = (x_1, \cdots, x_k, \alpha_{k+1}x_{k+1}, \alpha_{k+2}x_{k+2}, \cdots) \tag{5.1}$$

is admissible if $k = 0, 1, 2, 0 \le \alpha_i < 1$ for $i = k + 1, \cdots$ and $\sum_{i=k+1}^{\infty} \alpha_i^2 < \infty$. Consider, for example, the case $k = 0, 0 < \alpha_i < 1, i = 1, 2, \cdots$. In analogy to the finite dimensional situation, one may try to argue that δ^d is admissible, since it is the unique Bayes estimator with respect to the prior $G = \prod_{i=1}^{\infty} G_i$, where

$$G_i = N(0, \alpha_i/(1 - \alpha_i)). \tag{5.2}$$

However, our parameter space is ℓ^2 and G is a probability measure on ℓ^2 if and only if $\sum_i^{\infty} \alpha_i < \infty$ (see Subsection 7.2 for further discussion). Thus, the argument fails when $\sum_i \alpha_i = \infty$ and $\sum_i \alpha_i^2 < \infty$, but δ^d is still admissible. We resolve this difficulty with the aid of Blyth's method and we actually obtain a more general admissibility result for the discrete model (Theorem 3 in Subsection 5.4).

5.2. In the current and the following sections we describe some known results for the k-variate estimation problem discussed at the beginning of subsection 1.1 (see, for example, Brown [2] and [3]). Assume that

$$C = I.$$

Denote by $x^k = (x_1, \cdots, x_k)$ vectors in $\mathbb{R}^k$. Let $F(d\theta^k)$ be a finite measure on $\mathbb{R}^k$ such that

$$f(x^k) = \int_{\mathbb{R}^k} \exp(\textstyle\sum_{i=1}^k (x_i\theta_i - \tfrac{1}{2}\theta_i^2))F(d\theta^k) < \infty, \quad x^k \in \mathbb{R}^k. \tag{5.3}$$

We call the function $f(x^k)$ the *marginal density* associated with F. The unique Bayes estimator with respect to F, $\delta_F(x^k)$, has the form

$$\delta_F(x^k) = \nabla \log f(x^k). \tag{5.4}$$

5.3. Denote by $R(\theta^k, \delta)$ the risk function of the estimator

$$\delta(x^k) = (\delta_1(x^k), \cdots, \delta_k(x^k))$$

and put

$$R(F, \delta) = \int_{\mathbb{R}^k} R(\theta^k, \delta)F(d\theta^k).$$

Let $\delta(x^k)$ be an estimator with finite risk at the point 0 ($R(0, \delta) < \infty$); then δ is admissible if and only if there exists a sequence of finite measures $F^n(d\theta^k)$, $n = 1, 2, \cdots$ such that

5.2.A. The Bayes risk $R(F^n, \delta_{F^n}) < \infty$, $n = 1, 2, \cdots$.

5.2.B. The measure of the point set $\{0\}$, $F^n(\{0\}) \geq 1$, $n = 1, 2, \cdots$.

5.2.C. The sequence

$$\Delta_n = R(F^n, \delta) - R(F^n, \delta_{F^n}) = \int_{\mathbb{R}^k} \| \delta - \delta_{F^n} \|^2 f^n dP_0 \tag{5.5}$$

converges to 0, as $n \to \infty$. In (5.5), $\| \cdot \|^2$ is the sum of squares, $f^n(x^k)$ is the marginal density associated with F^n and $P_0(dx^k)$ is the k-variate normal distribution with mean 0 and covariance I. The second equality in (5.5) is essentially (1.3.2) of [2].

Generally, the method of proving admissibility by constructing a sequence of measures that satisfy 5.2.A–C is referred to as Blyth's method.

5.4. We now return to the infinte dimensional problem. All the results of Section 5.3 extend to this situation (see [3]). Till the end of the current section, only the discrete model is considered. To simplify notations, the sub(super)-script "d" will be omitted. Consistent with the notation in (5.5), denote by $P_0(dx)$ the product measure on $\mathbb{R}^\infty$ under which the coordinates are iid standard normal.

THEOREM 3. *Let $\delta^k(x^k) = (\delta_1(x^k), \cdots, \delta_k(x^k))$ be admissible in the k-variate model from Subsection 5.2. Let δ_i be Bayes with respect to G_i in the univariate model, $i = k+1, k+2, \cdots$. The estimator δ of the form*

$$\delta(x) = (\delta^k(x^k), \delta_{k+1}(x_{k+1}), \delta_{k+2}(x_{k+2}), \cdots)$$

is D-admissible if $R(0, \delta) < \infty$ and

$$\prod_{i=k+1}^{\infty} g_i(x_i) > 0 \ P_0\text{-a.s.}, \tag{5.6}$$

where g_i is the marginal density associated with G_i.

PROOF. Blyth's method is sufficient for admissibility. We construct a sequence of finite measures G^n on ℓ^2 such that 5.2.A–C hold. Let F^n be measures on $\mathbb{R}^k$ that satisfy 5.2.A–C applied to the k-variate admissible estimator $\delta^k(x^k)$. Put

$$H^n(d\theta^k) = F^n(d\theta^k)/\sqrt{\Delta_n}$$

where Δ_n is defined in 5.2.C. The prior G^n on ℓ^2 is defined as

$$G^n(d\theta) = 1_{\{0\}}(d\theta) + H^n(d\theta^k) \prod_{i=k+1}^{m_n} G_i(d\theta_i) \prod_{i=m_n+1}^{\infty} 1_{\{0\}}(d\theta_i) \tag{5.7}$$

where m_n is a divergent sequence to be specified later and $I_{\{0\}}(\cdot)$ denotes a measure on the appropriate probability space which assigns unit mass to 0. The Bayes estimate with respect to G^n, δ_{G^n}, can be calculated using (5.4). It is equal to

$$\delta_{G^n}(x) = (1 - 1/g^n)(\delta_1^n(x^k), \cdots, \delta_k^n(x^k), \delta_{k+1}(x_{k+1}), \cdots, \delta_{m_n}(x_{m_n}), 0, 0, \cdots)$$

where

$$g^n(x) = 1 + h^n(x^k) \prod_{i=k+1}^{m_n} g_i(x_i),$$

h^n is the marginal density associated with H^n and

$$(\delta_1^n(x^k), \cdots, \delta_k^n(x^k)) = \nabla \log h^n(x^k).$$

As in (5.5), we have

$$R(G^n, \delta) - R(G^n, \delta_{G^n}) = A_n + B_n + C_n,$$

where

$$A_n = \int_{\mathbb{R}^\infty} \sum_{i=1}^{k} \left[\delta_i(x^k) - \left(1 - \frac{1}{g^n}\right)\delta_i^n(x^k)\right]^2 g^n \, dP_0,$$

$$B_n = \int_{\mathbb{R}^\infty} \sum_{i=k+1}^{m_n} \left[\delta_i(x_i) - \left(1 - \frac{1}{g^n}\right)\delta_i(x_i)\right]^2 g^n \, dP_0,$$

$$C_n = \int_{\mathbb{R}^\infty} \sum_{i=m_n+1}^{\infty} [\delta_i(x_i) - 0]^2 g^n \, dP_0.$$

We proceed to show that for $n \to \infty$, A_n and B_n converge to 0 as long as $m_n \to \infty$ and that C_n converges to 0 if m_n diverges fast enough. The key fact is that as

$n \to \infty$ and as $m_n \to \infty$,

(5.8) $$g^n(x) \to \infty \quad P_0 - \text{a.s.}$$

Indeed, (5.8) follows from (5.6) and the relations

$$h^n(x^k) = f^n(x^k)/\sqrt{\Delta_n} \ge F^n(\{0\})/\sqrt{\Delta_n} \ge 1/\sqrt{\Delta_n} \to \infty, \quad n \to \infty.$$

Using the inequality $(a + b)^2 \le 2a^2 + 2b^2$ and that $g^n \ge 1$, we have

$$\begin{aligned} A_n &= \int_{\mathbb{R}^\infty} \textstyle\sum_{i=1}^k \left\{\left(1 - \frac{1}{g^n}\right)[\delta_i(x^k) - \delta_i^n(x^k)] + \frac{\delta_i(x^k)}{g^n}\right\}^2 g^n \, dP_0 \\ &\le 2 \int_{\mathbb{R}^\infty} \textstyle\sum_{i=1}^k \left(1 - \frac{1}{g^n}\right)^2 [\delta_i(x^k) - \delta_i^n(x^k)]^2 g^n \, dP_0 \\ &\quad + 2 \int \frac{1}{g^n} \textstyle\sum_{i=1}^k \delta_i(x^k)^2 \, dP_0 \le 2D_n + 2E_n \end{aligned}$$

where

$$D_n = \int_{\mathbb{R}^\infty} \textstyle\sum_{i=1}^k [\delta_i(x^k) - \delta_i^n(x^k)]^2 g^n \, dP_0,$$

$$E_n = \int_{\mathbb{R}^\infty} \frac{1}{g^n} [\textstyle\sum_{i=1}^k \delta_i(x^k)^2 + \sum_{i=k+1}^\infty \delta_i(x_i)^2] \, dP_0.$$

The measure $[\sum_{i=1}^k \delta_i(x^k)^2 + \sum_{i=k+1}^\infty \delta_i(x_i)^2] \, dP_0$ has total mass $R(0, \delta) < \infty$. Since $1/g^n \le 1$, from (5.8) and the dominated convergence theorem we conclude that $\lim E_n = 0$. By the independence, under P_0, of x^k and $\{x_i, i \ge k + 1\}$, by the equality $\int_{\mathbb{R}^\infty} \prod_{i=k+1}^{m_n} g_i(x_i) \, dP_0 = 1$ and the inequality $f^n(x^k) \ge 1$, as $n \to \infty$ we have

$$D_n = \int_{\mathbb{R}^k} \textstyle\sum_{i=1}^k [\delta_i(x^k) - \delta_i^n(x^k)]^2 (1 + h^n) \, dP_0 \le \Delta_n + \sqrt{\Delta_n} \to 0.$$

Hence $\lim A_n = 0$ as long as $n \to \infty$, $m_n \to \infty$. Now

$$B_n = \int_{\mathbb{R}^\infty} \frac{1}{g^n} \textstyle\sum_{i=k+1}^{m_n} \delta_i(x_i)^2 \, dP_0 \le E_n \to 0, \quad n \to \infty.$$

Finally,

(5.9) $$\begin{aligned} C_n &= \int_{\mathbb{R}^\infty} [\textstyle\sum_{i=m_n+1}^\infty \delta_i(x_i)^2] g^n(x^{m_n}) \, dP_0 \\ &= [1 + H_n(\mathbb{R}^k)] \cdot \left[\textstyle\sum_{i=m_n+1}^\infty \int_{\mathbb{R}^\infty} \delta_i(x_i)^2 \, dP_0\right]. \end{aligned}$$

The infinite sum in (5.9) is a tail of a convergent series since $R(0, \delta) < \infty$. It is possible, therefore, to choose $m_n \nearrow \infty$ fast enough so that $\lim C_n = 0$. Theorem 3 is now established.

5.5. We now complete the sufficiency part of Theorem 1 by proving 5.1.A. The estimator $\delta^k(x^k) = (x_1, \cdots, x_k)$ is admissible if $k = 1, 2$ (Stein [14]). The univariate estimator $\alpha_i x$, $0 \le \alpha_i < 1$ is Bayes with respect to the prior $N(0, \alpha_i/(1-\alpha_i))$. Hence δ is of the form considered in Theorem 3. Since the risk of δ is equal to

$$R(\theta, \delta) = k + \textstyle\sum_{i=k+1}^{\infty} [\alpha_i^2 + (1-\alpha_i)^2\theta_i^2], \quad \theta \in \ell^2,$$

and it is finite, we may conclude the admissibility of δ if we check (5.6). In our case

$$g_i(x_i) = \exp(\tfrac{1}{2}\alpha_i x_i^2 + \tfrac{1}{2}\log(1-\alpha_i))$$

and (5.6) holds if $\sum_{i=k+1}^{\infty} \alpha_i^2 < \infty$. Indeed, the random series $\sum_{i=k+1}^{\infty} \alpha_i(x_i^2 - 1)$ converges P_0-a.s. and the series $\sum_{i=k+1}^{\infty} [\alpha_i + \log(1-\alpha_i)]$ converges by the inequalities

$$0 \ge \alpha + \log(1-\alpha) \ge -(\alpha/(1-\alpha))^2, \quad 0 \le \alpha < 1,$$

and the fact that $\sum_i (\alpha_i/(1-\alpha_i))^2 < \infty$ if and only if $\sum_i \alpha_i^2 < \infty$. This concludes the proof of sufficiently in Theorem 1.

6. Linear estimators for the mean of the Wiener measure.

6.1. Let H be the space of functions on [0, 1] which are square integrable with respect to the Lebesgue measure. The Wiener measure P_θ on H is the Gaussian measure with mean $\theta \in H$ and covariance C given by

$$Cx(s) = \int_0^1 (s \wedge t)x(t)\,dt, \quad 0 \le s \le 1, \quad x \in H.$$

The parameter space $\Theta = C^{1/2}(H)$ consists of absolutely continuous functions θ with $\theta(0) = 0$ and derivative $\theta'(\cdot)$ which is square integrable. The inner product $\langle \cdot, \cdot \rangle$ in Θ induced by (2.1) is

$$\langle \theta, \eta \rangle = \int_0^1 \theta'(t)\eta'(t)\,dt, \quad \theta, \eta \in \Theta,$$

and the loss (2.2) is

$$\begin{aligned} L(\theta, a) &= \int_0^1 [\theta'(t) - a'(t)]^2\,dt, && \theta, a \in \Theta \\ &= \infty && \theta \in \Theta, \quad a \in H\backslash\Theta. \end{aligned} \tag{6.1}$$

The loss (6.1) is natural for engineering applications where the signal of interest is $\theta'(t)$ (see also comment 3 at the end of [1]). In Section 4 of [11] it was shown that the P_0-mlt are the Wiener integrals

$$Lx(s) = \int_0^1 \ell(s, t)\,dx(t), \quad x \in H, \tag{6.2}$$

where $\ell(\cdot, \cdot)$ is square integrable on $[0, 1] \times [0, 1]$. The restriction of L to Θ,

L_Θ, is

$$L_\Theta\theta(s) = \int_0^1 \ell(s, t)\theta'(t)\, dt, \quad \theta \in \Theta. \tag{6.3}$$

6.2. The admissible linear estimators are now described using Theorem 1′:

PROPOSITION 6.1. *The linear estimator* (6.2) *is admissible under the loss* (6.1) *if and only if the range of* L_Θ *in* (6.3) *is contained in* Θ *and with respect to the inner product* $\langle \cdot , \cdot \rangle$, L_Θ *is self-adjoint, Hilbert-Schmidt,* $0 \le L_\Theta \le I$, *and the eigenspace corresponding to the eigenvalue* 1 *is of dimension two at the most.*

Estimators of special importance are *nonanticipative* estimators. An estimator $\delta(x)$ is nonanticipative if $\delta(x)(t)$ depends only on $\{x(s),\ s \le t\}$. The linear estimator L in (6.2) is nonanticipative if and only if the kernel $\ell(s, t)$ is 0 for $t > s$. In that case L_Θ cannot be self-adjoint unless ℓ is 0. We have

COROLLARY 6.2. *There are no admissible linear estimators which are nonanticipative, except for the estimator* 0.

REMARKS. (i) Both Corollary 6.2 and the following Corollary 6.3 illustrate results which are basis free and which follow easily from the general theory.

(ii) Corollary (6.2) is, of course, not surprising, since nonanticipative estimators ignore so much information. It is of interest to investigate the problem of admissibility for nonanticipative estimators within the class of nonanticipative estimators.

6.3. The simplest class of linear estimators are *linear interpolators*. If the path x is observed, the linear interpolator L with knots

$$0 = t_0 < t_1 < \cdots < t_n \le 1$$

interpolates linearly the points

$$(0, 0), \quad (t_1, x(t_1)), \ \cdots, (t_n, x(t_n)), \quad (1, x(t_n)).$$

One can check that

$$\langle L_\Theta\theta, \eta \rangle = \sum_i \frac{[\theta(t_i) - \theta(t_{i-1})][\eta(t_i) - \eta(t_{i-1})]}{t_i - t_{i-1}};$$

hence L_Θ is self-adjoint and $\langle L_\Theta\theta, \theta \rangle \ge 0$. From the inequality

$$\left[\int_{t_{i-1}}^{t_i} h'(t)\, dt\right]^2 \le (t_i - t_{i-1}) \int_{t_{i-1}}^{t_i} h'(t)^2\, dt$$

it follows that $\langle L_\Theta\theta, \theta \rangle \le \langle \theta, \theta \rangle$. We get $0 \le L_\Theta \le I$. The eigenspace corresponding to eigenvalue 1 consists of all piecewise linear functions with knots at $\{0, t_1, \cdots, t_n\}$ and it is of dimension n. From Proposition 6.1 we obtain the following analogue of Stein's phenomenon:

COROLLARY 6.3. *A linear interpolator is admissible under the loss* (6.1) *if the number of knots, n, is* 1 *or* 2. *It is inadmissible if* $\eta \ge 3$.

7. Extensions and open problems.

7.1. We have investigated homogeneous linear estimators. A *nonhomogeneous linear estimator* δ is a *measurable affine transformation* with respect to P_0, i.e., δ is of the form

$$\delta(x) = Lx + h \tag{7.1}$$

where $h \in H$ and L is P_0-mlt (see Subsection 2.9 in [11]). In the univariate model, the estimator $\alpha x + h$ is admissible if and only if $0 \le \alpha < 1$ or $\alpha = 1$, $h = 0$. It follows that

THEOREM 4. *The nonhomogeneous linear estimator* (7.1) *is admissible under the loss* (2.2) *if and only if L is admissible, $h \in \Theta$ and $C^{-1/2}h$ is orthogonal to the eigenspace corresponding to the eigenvalue* 1 *of the operator A in* (2.5).

Only minor modifications in the proof of Theorem 1 are needed in order to establish Theorem 4.

7.2. Cohen [4] proved that admissible linear estimators are generalized Bayes. This is not so in the infinite dimensional model because of the restricted parameter space. As discussed in Subsection 5.1, the estimator

$$\delta(x) = (\alpha_1 x_1, \alpha_2 x_2, \cdots), \quad 0 < \alpha_i < 1, \quad \textstyle\sum_1 \alpha_i = \infty, \quad \sum_i \alpha_i^2 < \infty \tag{7.2}$$

is admissible. However, there exists no countably additive measure on $\Theta = \ell^2$ with respect to which δ is generalized Bayes. We believe that an admissible linear estimator is still generalized Bayes but with respect to a cylinder measure on Θ, i.e., a finitely additive measure Θ which is σ-additive on finite dimensional subspaces (see, for example, Kuo [8], page 92). In the finite dimensional case, a measure F is a generalized prior if (5.3) holds, and the generalized Bayes estimator with respect to F, δ_F, is defined as in (5.4). Something similar should hold in the infinite dimensional case for a Gaussian cylinder measure F (and maybe non-Gaussian as well).

7.3. In the finite dimensional case, an estimator which is admissible under one quadratic loss function is admissible under all. This is not true in infinite dimensions where an estimator may be admissible under some loss function and have infinite risk under another. We illustrate this point in the discrete setup. Consider the estimator δ^d in (4.8) and assume a *weighted* loss function

$$L(\theta, a) = \textstyle\sum_i \lambda_i(\theta_i - a_i)^2, \quad \sum_i \theta_i^2 < \infty,$$

where $\lambda_i > 0$, $i = 1, 2, \cdots$. In the model of Section 5, $\lambda_i = 1$, $i = 1, 2, \cdots$ and the estimator

$$\delta(x) = \alpha x, \quad x \in \mathbb{R}^\infty \ (0 < \alpha < 1),$$

is trivially inadmissible having infinite risk. It has finite risk when $\sum_i \lambda_i < \infty$. If

$$\lambda_i = q^i, \quad 0 < q < 1,$$

then αx is admissible when q is small enough, or equivalently, λ_i converges to 0 fast enough. We do not know, for example, what holds when

$$\lambda_i = 1/i^2, \ i = 1, 2, \cdots.$$

More generally, we do not know if the estimator $\delta(x) = \alpha x$, $x \in H$, $0 < \alpha < 1$, is admissible when the loss (2.2) is replaced by

$$L(\theta, a) = |\theta - a|^2,$$

where $|\cdot|$ is the norm in H.

7.4. In the finite dimensional case, a remarkable relation between estimators with a bounded risk function and diffusion processes was established by Brown [2]. This relation still holds for many estimators with unbounded risk (Srinivasan [13]), in particular for linear estimators. The diffusions corresponding to admissible linear estimators are the Ornstein-Uhlenbeck processes. Following formally the steps in [2] makes us believe that the processes corresponding to the linear estimators in infinite dimensions are the Ornstein-Uhlenbeck processes introduced by Malliavin [10] (see also Gaveau [5]), but we have no rigorous justification.

Acknowledgments. I am very grateful to L. Brown, E. Dynkin, R. Farrell, L. Gross, J. Hwang and I. Johnstone for their help and encouragement. Comments by the referee and the Associate Editor helped to improve the original version.

REFERENCES

[1] BERGER, J. and WOLPERT, R. (1982). Estimating the mean function of a Gaussian process and the Stein effect. *Proc. 3rd Purdue Symposium Statistical Decision Theory and Related Topics III*, Vol. 2. Edited by S. Gupta and J. Berger.

[2] BROWN, L. D. (1971). Admissible estimators, recurrent diffusions and insoluble boundary value problems. *Ann. Math. Statist.* **42** 855–903.

[3] BROWN, L. D. (1976). Lecture notes on statistical decision theory. Unpublished. Cornell University, Ithaca, New York.

[4] COHEN, A. (1966). All admissible linear estimates of the mean vector. *Ann. Math. Statist.* **37** 458–463.

[5] GAVEAU, B. (1981). Comptes rendus. *Acad. Sc. Paris*, t.293, serie I, 469–472.

[6] GOHBERG, I. and GOLDBERG, S. (1981). *Basic Operator Theory.* Birkhäusen, Boston.

[7] GRENANDER, U. (1981). *Abstract Inference.* Wiley, New York.

[8] KUO, H. (1975). Gaussian measures in Banach spaces. *Lecture Notes in Mathematics* **403**. Springer-Verlag, Berlin.

[9] LE CAM, L. (1983). Unpublished lecture notes. Berkeley University.

[10] MALLIAVIN, P. (1976). Stochastic calculus of variations and hypoelliptic operators. *Proc. Symp. Kyoto in Stochastic Differential Equations* 195–265. Kinokuniya and Wiley, New York.

[11] MANDELBAUM, A. (1983). Linear estimators and measurable linear transformations on a Hilbert space. *Z. Wahrsch. verw. Gebiete* **65** 385–397.

[12] ROZANOV, J. A. (1971). Infinite-dimensional Gaussian distribution. *Trans. Proc. Steklov Instit. Math.* **108.** American Mathematical Society.

[13] SRINIVASAN, C. (1978). Admissible generalized Bayes estimates and exterior boundary value problems. Purdue University, Dept. of Statistics. Mimeograph series no. 78-2.

[14] STEIN, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1** 197–206.

[15] STEIN, C. (1981). Estimation of the parameters of a multivariate normal distribution—I. Estimation of the means. *Ann. Math. Statist.* **9** 1135–1151.

GRADUATE SCHOOL OF BUSINESS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305