COMPACT COVARIANCE OPERATORS

CHARLES R. BAKER AND IAN W. McKEAGUE

ABSTRACT. Let $B$ be a real separable Banach space and $R : B^* \to B$ a covariance operator. All representations of $R$ in the form $\sum \epsilon_n \otimes \epsilon_n$, $(\epsilon_n, n > 1) \subset B$, are characterized. Necessary and sufficient conditions for $R$ to be compact are obtained, including a generalization of Mercer's theorem. An application to characteristic functions is given.

1. Introduction. The study of covariance operators is a major component in the theory of probability measures on Banach spaces [10], [9], [1]. The covariance operator of a strong second-order measure is always compact [2]; however, the covariance operator of a weak second-order measure need not be compact. In this paper we first characterize series representations of covariance operators, and then give a set of necessary and sufficient conditions for a covariance operator to be compact. The classical Mercer's theorem [7] can be obtained as an immediate corollary. These results are then applied to extend a result of Prohorov and Sazanov [6] on relative compactness of probability measures from Hilbert space to Banach space.

2. Definitions and notation. $B$ is a real separable Banach space with norm $\| \cdot \|$ and topological dual $B^*$. A linear operator $R : B^* \to B$ is a covariance operator if $R$ is symmetric and nonnegative: $\langle Ru, v \rangle = \langle u, Rv \rangle$ and $\langle Ru, u \rangle \geq 0$, for all $u, v$ in $B^*$. A probability measure $\mu$ on the Borel $\sigma$-field of $B$ is said to be weak second-order if $\int_B \langle x, u \rangle^2 \, d\mu(x) < \infty$, for all $u$ in $B^*$; $\mu$ is strong second-order if $\int_B ||x||^2 \, d\mu(x) < \infty$. Every weak second-order measure $\mu$ has a mean element $m$ in $B$ and a covariance operator $R : B^* \to B$ [9], [10], defined by

$$\langle m, v \rangle = \int_B \langle x, v \rangle \, d\mu(x), \quad \langle Ru, v \rangle_2 = \int_B \langle x-m, u \rangle \langle x-m, v \rangle \, d\mu(x),$$

for all $u, v$ in $B^*$. Strong second-order measures have compact covariances; the strong second-order property is not necessary in order that $\mu$ have compact covariance.

For a covariance operator $R : B^* \to B$ it is well known [8], [1], that there exists a separable Hilbert space $H \subset B$ such that the natural injection $j : H \to B$ is continuous and $R = jj^*$. $H$ is the RKHS of $R$ and is the completion of range$(R)$ with respect to the inner product $\langle \cdot, \cdot \rangle_H$ defined by $\langle Ru, Rv \rangle_H = \langle Ru, v \rangle$.

Received by the editors January 27, 1981 and, in revised form, March 10, 1981.

AMS (MOS) subject classifications (1970). Primary 60B05, 47B05, 60B10.

Key words and phrases. Covariance operator, compact operator, probability in Banach spaces.

1 This research was supported by the Office of Naval Research under Contract N00014-75-C-0491.
$I_H$ will denote the identity on $H$. For $u, v$ in $B$, $z$ in $B^*$ (resp. in $H$), $(u \otimes v)(z) = \langle v, z \rangle u$ (resp., $\langle v, z \rangle_H u$). If $T$ is any map $r(T) \equiv \text{range}(T)$. $\tau$ is the linear topology on $B^*$ determined by a neighborhood base at zero of the form $V_{cA}(0) = \{ f \in B^* : \sup_{x \in C} \langle f, x \rangle^2 < \epsilon \}$ for all $\epsilon > 0$ and all compact sets $C \subset B$ ($\tau$ is the topology of uniform convergence on compact sets). For a given covariance operator $R: B^* \to B$, $q_R$ is the real-valued quadratic functional on $B^*$ defined by $q_R f = \langle Rf, f \rangle$. The notation $R = \Sigma_n e_n \otimes e_n$ for $\{e_n, n > 1\} \subset B$ means that the sequence $(\Sigma_n e_n \otimes e_n)$ converges to $R$ in the strong operator topology: $\Sigma_n \langle e_n, f \rangle e_n \to Rf$ in the norm topology of $B$, for all $f$ in $B^*$. $I_H = \Sigma_n e_n \otimes e_n$ has a similar interpretation. If $\{u_n, n > 1\}$ is any orthonormal basis for $H$, then $R = \Sigma_j u_n \otimes j u_n$ [9]. $K_H$ will denote the unit ball in $H$.

If $\mu$ is a probability measure on the Borel $\sigma$-field of $B$, its characteristic functional $\hat{\mu}$ is defined as $\hat{\mu}(x) = \int_B e^{ix'y} \, d\mu(y)$, for $x$ in $B^*$.

3. Representation of covariance operators. In this section, $R$ is an arbitrary covariance operator.

**Theorem 1.** $R = \Sigma_n e_n \otimes e_n$ if and only if $e_n = jv_n$, $v_n \in H$ for $n > 1$, and $I_H = \Sigma_n v_n \otimes v_n$.

**Proof.** It suffices to show that the stated conditions are necessary for $R = \Sigma_n e_n \otimes e_n$. Suppose $R = \Sigma_n e_n \otimes e_n$, and fix $e_k$. Let $P_k = e_k \otimes e_k$. To show $e_k \in \text{range}(j)$, let (as in [3]) $D: r(j^*) \to B$ be defined by $Df = P_k f$. Then $\|Df\|^2 = \|P_k f\|^2 = \|e_k\|^2 \|e_k\|^2 < \|e_k\|^2 \|e_n\|^2 = \|e_k\|^2 \|f\|^2_H$. Thus $D$ can be extended to a continuous linear map from $r(j^*) = H$ into $B$. From its definition, $Dj^* = P_k$, so $P_k = jD^*$ and thus $e_k \in \text{range}(j)$.

To see that $I_H = \Sigma v_n \otimes v_n$, where $jv_n = e_n$, $n > 1$, define $Q_N = \Sigma^N v_n \otimes v_n$. $Q_N = Q_N^*$ and $Q_N > 0$, so $Q_N^{1/2}$ exists. $\|Q_N^{1/2} j^* f\|_H^2 = \Sigma_n \langle f, e_n \rangle^2 \|j^* f\|_H^2$, so that $\|Q_N^{1/2}\| < 1$ and $\|Q_N^{1/2} x\|_H \to \|x\|_H$ for all $x$ in $r(j^*)$. Thus, 

$$\left| \sum_{1}^{N} (v_n \otimes v_n) j^* f - j^* f \right|_H^2 = \|Q_N j^* f - j^* f\|_H^2,$$

which converges to zero as $N \to \infty$ for any fixed $f$ in $B^*$. Thus, $\Sigma v_n \otimes v_n = I_H$ on $r(j^*)$, and the result follows by $\overline{r(j^*)} = H$. □

**Remark.** Suppose $E$ is a locally convex topological vector space, $R: E' \to E$ is a covariance operator, and $R = jj^*$, where $j: H \to E$ is the injection and $H$ is the RKHS of $R$. $R$ will have such a representation, for example, if $E$ is quasi-complete [8]. In this case, it is easily shown that Theorem 1 holds without modification.

The representation $I_H = \Sigma v_n \otimes v_n$ does not require that $\{v_n, n > 1\}$ be a CONS in $H$; however, sufficient conditions for $\{v_n, n > 1\}$ to be a CONS in $H$ can be given.

**Proposition 1.** Suppose $I_H = \Sigma v_n \otimes v_n$; the following are equivalent:

1. $\|v_k\|_H = 1$.
2. $v_k \not\subset \text{sp}\{v_n, n \neq k\}$.
3. $v_k \perp \text{sp}\{v_n, n \neq k\}$.

If any of the above conditions holds for all $k > 1$, then $\{v_n, n > 1\}$ is a CONS in $H$.  

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
4. Compact covariance operators.

THEOREM 2. Suppose $R = \sum e_n \otimes e_n, \{e_n, n \geq 1\} \subset B$. Let $\{v_n, n \geq 1\} \subset H$ be such that $e_n = jv_n, n \geq 1$. The following are equivalent:

(1) $R$ is compact;
(2) $j$ is compact;
(3) $[K_R] \subset B$ is compact in $B$;
(4) the series $\sum v_n \otimes jv_n$ converges uniformly in $H$ on bounded subsets of $B^*$;
(5) $\sum e_n \otimes e_n$ converges to $R$ uniformly in $B$ on bounded subsets of $B^*$;
(6) $q_R$ is $w^*$-continuous on bounded subsets of $B^*$;
(7) $q_R$ is $\tau_c$-continuous.

Proof. (1) $\Rightarrow$ (2). Suppose $f_n \rightarrow f$ in the $w^*$ topology of $B^*$, where $||f_n|| < k$ for all $x$. Then $||j^*f_n - j^*f||_H^2 = \langle R(f_n - f), (f_n - f) \rangle < 2k||R(f_n - f)||_B$; since $R$ is compact, $j^*f_n \rightarrow j^*f$ in $H$ [4, p. 486] and thus $j$ is compact.

(2) $\Rightarrow$ (3). $j$ compact implies $[K_R]$ is relatively compact in $B$. Since $K_R$ is weakly compact in $H$ and $j$ is weakly continuous, $[K_R]$ is weakly compact in $B$, and thus closed.

(3) $\Rightarrow$ (2) by definition.

(2) $\Rightarrow$ (4). By Theorem 1, $\sum v_n \otimes v_n = I_H$. Set $Q_N = \sum v_n \otimes v_n$. If $A \subset B^*$ is bounded, then $j^*A$ is compact; by Dini's theorem $||Q_N^{1/2}x||_H \leq ||x||_H$ uniformly on $j^*[A]$. Hence $||Q_N - I||j^*x||_H^2 \leq ||j^*x||_H^2 - ||Q_N^{1/2}j^*x||_H^2 \rightarrow 0$ uniformly on $A$.

(4) $\Rightarrow$ (5), since $j$ is continuous.

(5) $\Rightarrow$ (1), since $R$ is the uniform limit of compact operators.

(2) $\Leftarrow$ (6) follows from the fact that $j$ is compact if and only if $j^*f_n \rightarrow 0$ in the norm topology of $H$ for all bounded generalized sequences $(f_n)$ in $B^*$ which are $w^*$ convergent to zero [4, p. 486], and $q_R(f_n) = ||j^*f_n||_H^2$.

(1) $\Rightarrow$ (7). Suppose $R$ is compact. Writing $C = j[K_R]$, $C$ is compact in $B$. $q_R(f) = \langle R(f), f \rangle = ||j^*f||_H^2 = \sup_{x \in C} \langle j^*f, x \rangle_H^2 = \sup_{x \in C} \langle f, x \rangle^2$. Thus $q_R$ is $\tau_c$-continuous at zero. $\tau_c$-continuity of $q_R$ follows from $q_R(f_n) = q_R(f_n - f) - q_R(f) + 2\langle Rf, f_n \rangle$.

(7) $\Rightarrow$ (1). Suppose $q_R$ is $\tau_c$-continuous. Using (6), $R$ is compact if $q_R$ is $w^*$ continuous at 0 on bounded subsets of $B^*$. But $B$ is separable so that the $w^*$ topology on bounded subsets of $B^*$ is metrizable and it suffices to consider sequences. Suppose $f_n \rightarrow w^* 0$ and $||f_n|| < k$. Let $L$ be an arbitrary compact subset of $B$. Since $(f_n)$ is bounded in $B^*$ the $f_n$ are equicontinuous and uniformly bounded as continuous functions on $L$. Thus, by the Arzela-Ascoli theorem [4, p. 266], $(f_n)$ is relatively compact as a subset of $C^*(L)$. Thus since $f_n \rightarrow w^* 0$, $f_n$ converges to 0 uniformly on $L$. Therefore $f_n \rightarrow 0$ and $q_R(f_n) \rightarrow 0$. This completes the proof of Theorem 2.

Remarks. (1) Suppose $r: [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ is continuous, symmetric and positive definite. For fixed $t \in [0, 1]$, let $r_t(x) = x_1$ for $x$ in $C[0, 1]$; $||r_t|| = 1$. A compact covariance operator $R: C^*[0, 1] \rightarrow C[0, 1]$ is defined by $[R\mu](t) = \int_0^1 r(t, s) \, d\mu(s)$ for any $\mu$ in $C^*[0, 1]$ (by Arzela-Ascoli theorem). Thus for $s, t \in [0, 1]$, $\langle R\pi_t, \pi_s \rangle = r(t, s)$. The integral operator in $L_2[0, 1]$ corresponding to the kernel $r$ has continuous orthonormal eigenvectors $\{y_n, n \geq 1\}$ and associated
nonzero eigenvalues \( \{ \lambda_n, n > 1 \} \); it is well known that \( \{ \sqrt{\lambda_n} y_n, n > 1 \} \) is a CONS in the RKHS \( H \) of \( R \). Thus, from Theorem 2, \( \sum_{n=1}^{N} \lambda_n y_n(t)y_n(s) \) converges uniformly to \( r(t, s) \) for all \( t, s \) in \([0, 1]\). This is the classical Mercer’s theorem [7, pp. 245–246].

(2) The fact that the unit ball of \( H \) is compact in \( B \) when \( R \) is compact was proved by Kuelbs [5] under the assumption that \( R \) is the covariance of a strong second-order measure.

5. Characteristic functionals. Let \( \Lambda \) denote a family of probability measures on \( B \) (separable Banach) and \( \hat{\Lambda} \) the corresponding family of characteristic functionals.

**Theorem 3.** Let \( B \) be a separable Banach space. Then the following are equivalent:

(a) There exists a topology \( \tau \) on \( B^* \) such that for each family \( \Lambda \) of probability measures on \( B \), \( \hat{\Lambda} \) is equicontinuous in this topology if and only if \( \Lambda \) is relatively compact in the topology of weak convergence.

(b) \( B \) is finite dimensional.

**Proof.** As in the Hilbert space case (see [6, Lemma 2]), \( \tau_c \) is the weakest topology on \( B^* \) such that relative compactness of \( \Lambda \Rightarrow \) equicontinuity of \( \hat{\Lambda} \). Suppose that (a) holds. Then \( \tau_c \subset \tau \) and \( \tau_c \)-equicontinuity of \( \hat{\Lambda} \) implies relative compactness of \( \Lambda \). Now let \( R: B^* \to B \) be any compact covariance operator. Let \( \{ e_n \} \) be a CONS in the RKHS of \( R \). Define \( \mu_k \) to be the zero mean Gaussian measure on \( B \) with covariance operator \( \sum_k e_n \otimes e_n \). Then \( \{ \mu_k \} \) is \( \tau_c \)-equicontinuous by Theorem 2 and \( \{ \mu_k \} \) is relatively compact. Therefore \( R \) is the covariance of a Gaussian probability measure on \( B \) and, by [9, Theorem 11], \( B \) is finite dimensional. \( \square \)

Theorem 3 extends a result of Prohorov and Sazonov [6] who proved it for Hilbert spaces.

**References**


Department of Statistics, University of North Carolina, Chapel Hill, North Carolina 27514