BULL. AUSTRAL. MATH. SOC. 62H99, 45H05
VOL. 11 (1974), 373-380.

Mercer's Theorem and Fredholm resolvents

C.S. Withers

Multivariate versions of Mercer's Theorem and the usual expansions of the resolvent and Fredholm determinant are shown to hold for an $n \times n$ symmetric kernel $N(x, y)$ with arbitrary domain in R^p under weakened continuity conditions. Further, the resolvent and determinant of $N(x, y) - a(x)b(y)$ are given in terms of those of $N(x, y)$.

1. Introduction

Our main result, given in §3, deals with eigenfunction expansions of the $n \times n$ matrix kernel $N(x, y)$ (Mercer's Theorem) and its iterates $N_j(x, y)$ and resolvent $N(x, y, \lambda)$. As well as allowing n to be arbitrary and the domain to be unbounded, we weaken the usual continuity assumptions of this important theorem, which has found applications in optimum detection theory (for example, Deutsch [2], p. 244) and statistics. These expansion formula are basic to the study of the distribution of the random variable $\int |X(t)|^2 dt$ where $X : R^p \to R^n$ is a Gaussian process with covariance N . Such random variables arise in connection with the asymptotic distribution and power of certain statistical tests; see Withers [8].

In §4 we give a simple but useful result: formulae for the resolvent and Fredholm determinant of $N(x, y) - a(x)b(y)$ in terms of those of $N(x, y)$, where $a(x)$, $b(y)$ are $n \times q$ and $q \times n$ functions. This

Received 9 July 1974.

section was motivated by the study of the random variable $\int |Y(t)|^2 dt$ where

$$Y(t) = X(t) + f(t) \int g(x)X(s)ds ,$$

and f, g are matrix functions, and X is as above; see Withers [9] for statistical applications where such variables arise naturally. The basic results of Fredholm integral equation theory for a matrix kernel $N(x, y)$ defined on a domain $\Omega \times \Omega$ where Ω is an arbitrary domain in p-dimensional euclidean space R^p appear to have been stated only for the case $n = p = 1$ and Ω a bounded interval. However the technique of deducing these results for general n is well known and some authors have realised such extensions are possible (*cf.* Riesz and Nagy [6], p. 145), although others have suggested that Ω must be bounded (*cf.* Pogorzelski [5], p. 95). We therefore begin with a summary of these results for general n and arbitrary domain in R^p .

2. Some basic results

We give here generalisations of some basic results. These are easily deduced by the method of Carleman [1] (given for $n = p = 1$, $\Omega = [a, b]$, N real), the technique of reducing to $n = 1$ (for example, Pogorzelski [5], p. 181) and the standard proofs.

Throughout this paper we shall use A^* to denote the conjugate transpose of a complex matrix $A = (A_{ij})$, and $\|\cdot\|$ to denote the norm defined by $\|A\|^2 = \sum |A_{ij}|^2$. All integrals will be with respect to Lebesgue measure over Ω , an arbitrary subset of R^p .

Given $\Omega \subset R^p$ consider a complex measurable $n \times n$ function $N(x, y)$ on $\Omega \times \Omega$ such that

$$0 < \iint \|N(x, y)\|^2 dxdy < \infty . \tag{1}$$

For f a complex measurable $n \times q$ function on Ω such that $\int \|f\|^2 < \infty$, let $Nf(x) = \int N(x, s)f(s)ds$ and $f(y)^*N = \int f(s)^*N(s, y)ds$.

https://doi.org/10.1017/S0004972700044002 Published online by Cambridge University Press

Let $N_j = N^{j-1}N$, $j \geq 1$, where $N^0 = I$, the identity operator. Then $N(x, y, \lambda)$, the *resolvent* of $N(x, y)$ exists and for y in Ω , $h(x) = N(x, y, \lambda)$ is the unique solution of

$$h = N(\cdot, y) + \lambda Nh ,$$

and for x in Ω , $g(y) = N(x, y, \lambda)$ is the unique solution of

$$g = N(x, \cdot) + \lambda gN .$$

When

(2) $$\sum \int |N_{ii}(x, x)| dx < \infty ,$$

then the *Fredholm determinant* $D(\lambda)$ exists and is given by

$$\frac{d}{d\lambda} \log D(\lambda) = -\int \operatorname{trace} N(x, x, \lambda) dx , \quad D(0) = 1 .$$

When

(3) $$N^*(y, x) = N(x, y) \quad \text{for } x, y \text{ in } \Omega ,$$

then there exist real numbers $\{\lambda_1, \lambda_2, \ldots\}$ (*eigenvalues*) and complex n-vectors on Ω , $\{\phi_1, \phi_2, \ldots\}$ (*eigenvectors*) satisfying

$$\lambda_i N\phi_i = \phi_i , \quad \int \phi_i^* \phi_j = \begin{cases} 1 , & i = j \\ 0 , & i \neq j \end{cases} , \quad 0 < |\lambda_1| \leq |\lambda_2| \leq \ldots ,$$

such that if

(4) $$\lambda N\phi = \phi , \quad \int |\phi|^2 < \infty , \ \phi \ n \times 1 ,$$

then $\lambda = \lambda_k$ for some k , and ϕ is a linear combination of those ϕ_r such that $\lambda_r = \lambda$.

When (3) and

(5) $$\sup_y \int \|N(x, y)\|^2 dx < \infty ,$$

then for x in Ω , and almost all y in Ω ,

$$N(x, y, \lambda) = N(x, y) + \lambda \sum_1^\infty \left\{\frac{\phi_i(x)\phi_i(y)^*}{\lambda_i(\lambda_i-\lambda)}\right\} , \quad \lambda \text{ not an eigenvalue,}$$

and

(6) $$N_j(x, y) = \sum_1^\infty \lambda_i^{-j}\phi_i(x)\phi_i(y)^* , \quad j \geq 2 ,$$

and the convergence of these series is (element-wise) absolute and uniform in Ω^2

3. Mercer's Theorem

Mercer's Theorem concerns the expansion of N in terms of its eigenfunctions and eigenvalues:

$$N(x, y) = \sum \phi_i(x)\phi_i(y)^*/\lambda_i .$$

Statements of the theorem in the literature all make unnecessary continuity and other assumptions. For example sometimes (5) is assumed (for example, Pogorzelski [5], p. 150). Our aim here is to impose as few conditions on N as seems possible. It is worth noting that a useful weakening of our continuity assumptions (9)-(11) may be made by excluding from Ω the set of points P at which they do not hold, provided P has Lebesgue measure zero.

Our version of Mercer's Theorem is as follows.

THEOREM 1. *Suppose* N *satisfies* (1), (3), *and the following*

(7) $$\int \phi^*N\phi = \iint \phi^*(x)N(x, y)\phi(y)dxdy \geq 0$$

for all complex $n \times 1$ *functions* ϕ *such that* $\int |\phi|^2 < \infty$,

(8) $$\sup_{x\in\Omega} \text{trace } N(x, x) < \infty ,$$

(9) $N(x, y)$ *is continuous at* $y = x \in \Omega$,

(10) $N_2(x, x)_{ii}$ *is continuous in* Ω, $1 \leq i \leq n$,

(11) $N_2(x, y)_{ii}$ *is continuous at* $y = x$ *in* Ω, $1 \leq i \leq n$,

https://doi.org/10.1017/S0004972700044002 Published online by Cambridge University Press

then for $\{\lambda_i, \phi_i\}$ *above,* $\{\phi_i\}$ *are continuous and* $\sum_1^\infty \lambda_i^{-1}\phi_i(x)\phi_i(y)^*$ *converges (elementwise) absolutely and uniformly in* Ω^2 *and equals* $N(x, y)$ *almost everywhere in* Ω^2 .

NOTES. (i) Since a uniformly convergent series of continuous functions converges to a continuous function, equality holds at continuity points of N , such as $x = y$.

(ii) Since (4) implies $\phi = \lambda^m N^m \phi$, (10) and (11) may be replaced by

$(11)^1$ for some $m \geq 1$, $N_{2m}(x, y)_{ii}$ and $N_{2m}(x, x)_{ii}$ are continuous at $y = x$ for x in Ω , $1 \leq i \leq n$; (*cf.* Hobson [3] who gives for continuity of $\{\phi_i\}$);

$(11)^2$ for some $m \geq 1$, $N_m(x, y)$ is continuous in x for almost all y in Ω and for $j \geq m$, N_j is bounded.

Proof of Theorem 1. We may without loss take $n = 1$. $\{\phi_i\}$ are continuous because (4) implies

$$|\phi(x)-\phi(y)|^2 \leq |\lambda|^2 I(x, y) \int |\phi|^2 ,$$

where $I(x, y) = \int |N(x, s)-N(y, s)|^2 ds \to 0$ by (3), (10), (11).

By (9) and the standard method (for example, [5], p. 151), $N(x, x)$ and $N(x, x) - \sum_1^q \lambda_i^{-1}|\phi_i(x)|^2$ are real and non-negative for $q \geq 1$. Hence by (8) for $\varepsilon > 0$, there exists M such that

$$\sum_M^\infty \lambda_i^{-1}|\phi_i(x)|^2 < \varepsilon \text{ in } \Omega ,$$

so that for $n_2 \geq n_1 \geq M$,

$$\sum_{n_1}^{n_2} \lambda_i^{-1}|\phi_i(x)\phi_i(y)^*| < \varepsilon \text{ in } \Omega^2 .$$

https://doi.org/10.1017/S0004972700044002 Published online by Cambridge University Press

Hence $\sum \lambda_i^{-1}\phi_i(x)\phi_i(y)^*$ converges absolutely and uniformly in Ω^2, so that by [5], pp. 130, 131, the sum equals $N(x, y)$ almost everywhere.

The usual expansions now follow:

COROLLARY. *Under the conditions of Theorem* 1,

(12) *for* $j \geq 1$, $N_j(x, y) = \sum \lambda_i^{-j}\phi_i(x)\phi_i(y)^*$ *almost everywhere in* Ω^2,

(13) $N(x, y, \lambda) = \sum (\lambda_i - \lambda)^{-1}\phi_i(x)\phi_i(y)^*$ *almost everywhere in* Ω^2, *if* λ *is not an eigenvalue,*

(14) *for* $j \geq 1$, $\int \text{trace } N_j(x, x)dx = \sum \lambda_i^{-j}$, *(possibly infinite for* $j = 1$ *), and the (elementwise) convergence in* (12) *and* (13) *is absolute and uniform.*

If also (2) *holds, that is,* $\int \text{trace } N(x, x)dx < \infty$, *then*

$$D(\lambda) = \prod_1^\infty (1 - \lambda/\lambda_i) . \tag{15}$$

4. The kernel $N(x, y) - a(x)b(y)$

Carleman [1] gave expansions in λ for $D(x, y, \lambda)$ and $D(\lambda)$ for $N = G + H$ and $N = G \cdot H$ in terms of multiple integrals of determinants, akin to Fredholm's series. Here we give more convenient formulae for $K(x, y, \lambda)$ and $D_K(\lambda)$, the resolvent and Fredholm determinant for the particular case

$$K(x, y) = N(x, y) - a(x)b(y) ,$$

where we assume a, b are $n \times q$ and $q \times n$ functions on Ω such that

$$\int \|a\|^2 < \infty , \quad \int \|b\|^2 < \infty ,$$

when $D_N(\lambda)$, $N(x, y, \lambda)$ are known, and where we set $D_N(\lambda) = D(\lambda)$ to avoid confusion with $D_K(\lambda)$.

THEOREM 2. *Let* N *satisfy* (1) *and* (2). *Let* T *denote the operator*

https://doi.org/10.1017/S0004972700044002 Published online by Cambridge University Press

$(I-\lambda N)^{-1}$ *so that*

$$Ta(x) = a(x) + \lambda \int N(x, y, \lambda)a(y)dy$$

and

$$b(x)T = b(x) + \lambda \int b(y)N(y, x, \lambda)dy .$$

Let $B(\lambda) = 1_q + \lambda \int bTa$, *where* $1_q = \text{diag}(1, \ldots, 1)$. *Then*

(16) $$K(x, y, \lambda) = N(x, y, \lambda) - Ta(x)B(\lambda)^{-1}b(y)T ,$$

for x, y *in* Ω *and* $D_K(\lambda) \neq 0$. *Also*

(17) $$D_K(\lambda) = D_N(\lambda) \cdot \det B(\lambda) .$$

Further, if $\det B(\lambda) = 0$, *then eigenfunctions of* K *with eigenvalue* λ *all have the form* Tac *where* $c \neq 0$ *is a* q*-vector such that*

$$B(\lambda)c = 0 .$$

NOTE. Michlin [4] has given a special case of (16) without proof.

Proof. Suppose $D_N(\lambda) \neq 0$ and $D_K(\lambda) = 0$.

Then $f = \lambda Kf$ has a non-trivial solution where $Kf(x) = \int K(x, y)f(y)dy$. Hence $c = \int bf \neq 0$ and $f = -\lambda Tac$. Hence $\det B(\lambda) = 0$.

Suppose $D_N(\lambda) \cdot \det B(\lambda) \neq 0$. Then $D_K(\lambda) \neq 0$ and for h such that $\int |h|^2 < \infty$, $f = h + \lambda Kf$ has solution

$$f = (I-\lambda K)^{-1}h = T(h-\lambda ac) .$$

Hence $c = Rh$ where $R = B(\lambda)^{-1}bT$, so that $(I-\lambda K)^{-1} = T(I-\lambda aR)$, which proves (16).

If $D_K(\lambda)D_N(\lambda) \neq 0$,

$$\frac{d}{d\lambda} \log\left(D_K(\lambda)/D_N(\lambda)\right) = - \int \text{trace}\left(K(x, x, \lambda)-N(x, x, \lambda)\right)dx$$

$$= \text{trace } B(\lambda)^{-1}C \text{ , by (14),}$$

https://doi.org/10.1017/S0004972700044002 Published online by Cambridge University Press

where $C = \int bT^2a$. For $|\lambda|$ small,

$$\frac{d}{d\lambda}\lambda T = I + 2\lambda N + 3\lambda^2N^2 + \ldots = T^2 ,$$

so that $C = d/d\lambda B(\lambda)$ for all λ by analytic continuation. (17) follows.

References

[1] T. Carleman, "Zur Theorie der linearen Integralgleichungen", *Math. Z.* 9 (1921), 196-217.

[2] Ralph Deutsch, *Estimation theory* (Prentice-Hall, Englewood Cliffs, New Jersey, 1965).

[3] E.W. Hobson, "On the representation of the symmetrical nucleus of a linear integral equation", *Proc. London Math. Soc.* (2) 14 (1915), 5-30.

[4] S. Michlin, "On the convergence of Fredholm series", *C.R. (Doklady) Acad. Sci. URSS (NS)* 42 (1944), 373-376.

[5] W. Pogorzelski, *Integral equations and their applications.* Vol. 1 (International Series of Monographs in Pure and Applied Mathematics, 88. Pergamon, Oxford, New York, Frankfurt; PWN-Polish Scientific Publishers, Warsaw, 1966).

[6] Frigyes Riesz and Béla Sz.-Nagy, *Functional analysis* (Frederick Ungar, New York, 1955).

[7] F. Smithies, *Integral equations* (Cambridge Tracts in Mathematics and Mathematical Physics, 49. Cambridge University Press, Cambridge, 1958).

[8] C.S. Withers, "The characteristic function of the L_2-norm of a Gaussian process", submitted.

[9] C.S. Withers, "On the asymptotic power of statistics which are L_2-norms", submitted.

Applied Mathematics Division,
Department of Scientific and Industrial Research,
Wellington, New Zealand.