PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 85, Number 3, July 1982

INEQUALITIES IN HILBERT MODULES OF MATRIX-VALUED FUNCTIONS

ADHEMAR BULTHEEL

ABSTRACT. The classical Cauchy-Schwarz inequality and extremality properties of reproducing kernels are generalized for a module of matrix-valued functions on which a matrix-valued inner product is defined. Reference to an application in the field of linear prediction of multivariate stochastic processes is made.

$\mathbf{F}$ denotes the field of real numbers $\mathbf{R}$ or the field of complex numbers $\mathbf{C}$. $K = \mathbf{F}^{n\times n}$ is the ring of $n \times n$ matrices with their entries in $\mathbf{F}$. We assume

(1) M is a (left) K-module of K-valued functions on a set Ω, and

(2) $\langle\cdot,\cdot\rangle$ is a K-valued nonnegative hermitian sesquilinear function on $M \times M$, i.e. for $F, G, H \in M$ and $A, B \in K$,

(P1) $\langle F, F\rangle \geq 0$,

(P2) $\langle F, G\rangle = \langle G, F\rangle^*$,

(P3) $\langle AF + BG, H\rangle = A\langle F, H\rangle + B\langle G, H\rangle$

(* denotes the adjoint of a matrix).

M with $\langle\cdot,\cdot\rangle$ is called a Hilbert module if $\langle F, F\rangle = 0$ iff $F = 0$. Note that $\langle\langle\cdot,\cdot\rangle\rangle = \operatorname{trace}\langle\cdot,\cdot\rangle$ is an $\mathbf{F}$-valued nonnegative hermitian sesquilinear functional when M is taken over the subring of matrices of the form aI, $a \in \mathbf{F}$. Thus the classical Cauchy-Schwarz inequality holds, i.e.

$$|\langle\langle F, G\rangle\rangle| \leq \langle\langle F, F\rangle\rangle^{1/2}\langle\langle G, G\rangle\rangle^{1/2} \qquad \forall F, G \in M.$$

Note further that $\langle F, F\rangle = 0 \Leftrightarrow \langle\langle F, F\rangle\rangle = 0$. Define the equivalence relation $\sim$ on M by $F \sim G$ iff $\langle F - G, F - G\rangle = 0$, and define for each $F \in M$ the equivalence class $[F]_e = \{G | G \sim F\}$. Then $[F]_e = F + [0]_e = \{F + G | G \sim 0\}$.

The following result is the K-valued generalization of the Cauchy-Schwarz inequality.

THEOREM 1. *For M, $\langle\cdot,\cdot\rangle$ as in* (1), (2) *we have*

$$\langle F, G\rangle\langle G, G\rangle^+\langle G, F\rangle \leq \langle F, F\rangle \qquad \forall F, G \in M.$$

(*If A is hermitian, A^+ denotes any hermitian $\{2\}$-inverse of A, i.e.* [**2**] *A^+ is hermitian and solves $A^+AA^+ = A^+$. In particular we could take for A^+ the unique Moore-Penrose generalized inverse.*)

PROOF. Properties (P1)–(P3) imply that

$$0 \leq \langle AF + BG, AF + BG\rangle = A\langle F, F\rangle A^* + B\langle G, G\rangle B^* - A\langle F, G\rangle B^* - B\langle G, F\rangle A^*.$$

Replace A by I and B by $\langle F, G\rangle\langle G, G\rangle^+$; then the result follows. □

Received by the editors July 21, 1981 and, in revised form, October 12, 1981.

1980 *Mathematics Subject Classification*. Primary 46C10, 47D15.

Key words and phrases. Cauchy-Schwarz, reproducing kernel, matrix optimization.

© 1982 American Mathematical Society
0002-9939/81/0000-0289/$02.25

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

As in the classical case of semidefinite inner product spaces, one can show with the aid of Theorem 1 that

$$[0]_e = M \cap M^\perp = M \cap \{F \in M | \langle F, G\rangle = 0 \, \forall G \in M\}$$

[3], i.e. the neutral and isotropic elements of M coincide. We have for $F_1 \sim F_2$ and $G_1 \sim G_2$ that $\langle F_1, G_1\rangle = \langle F_2, G_2\rangle$. $\langle [F]_e, [G]_e\rangle_e = \langle F, G\rangle$ is thus well defined and $M_e = \{[F]_e | F \in M\}$ is a Hilbert module under $\langle\cdot,\cdot\rangle_e$.

Let M, $\langle\cdot,\cdot\rangle$ be as in (1), (2). A $\langle\cdot,\cdot\rangle$-*compatible reproducing kernel* is a function $k(\cdot,\cdot) : \Omega \times \Omega \to K$ such that

(K1) $k(\cdot,\omega) \in M \, \forall\omega \in \Omega$, and

(K2) $\forall F \in M : \exists G \in [F]_e$ such that $\langle F, k(\cdot,\omega)\rangle = G(\omega) \, \forall \omega \in \Omega$.

Note when $[0]_e = \{0\}$, i.e. $\langle F, F\rangle = 0 \Leftrightarrow F = 0$; then $k(\cdot,\cdot)$ is the usual Aronszajn type (K-valued) reproducing kernel [1].

We now show that $k(\cdot,\cdot)$ solves some optimization problems as in the classical case [8, Theorem III.3, p. 44].

THEOREM 2. *Let M, $\langle\cdot,\cdot\rangle$, $k(\cdot,\cdot)$ be as in* (1), (2), (3). *Let $\alpha \in \Omega$ be fixed and let $k(\alpha,\alpha)$ be invertible. Then* (*with respect to the* $\leq$*-ordering of matrices*)

(a) $\min\{\langle F, F\rangle | F(\alpha) = I, \, F \in M\} = k(\alpha,\alpha)^{-1}$, *where the minimum occurs only for the elements $F \in [k(\alpha,\alpha)^{-1}k(\cdot,\alpha)]_e$ that satisfy $F(\alpha) = I$* (*so the minimizing value is unique when* $[0]_e = \{0\}$);

(b) *when* $[0]_e = \{0\}$, $\max\{F(\alpha)^*F(\alpha) | \langle F, F\rangle = I, \, F \in M\} = k(\alpha,\alpha)$, *where the maximum occurs only for F of the form $F(\cdot) = Ak(\cdot,\alpha)$ where A is any matrix such that $A^*A = k(\alpha,\alpha)^{-1}$* (*for example we may take* $A = (k(\alpha,\alpha)^{-1})^{1/2}$).

PROOF. (a) For any $F \in M$ with $F(\alpha) = I$, Theorem 1 gives

$$\langle F, F\rangle \geq \langle F, k(\cdot,\alpha)\rangle\langle k(\cdot,\alpha), k(\cdot,\alpha)\rangle^{-1}\langle k(\cdot,\alpha), F\rangle = k(\alpha,\alpha)^{-1}.$$

This lower bound for $\langle F, F\rangle$ is obtained by $F = S \triangleq k(\alpha,\alpha)^{-1}k(\cdot,\alpha)$ as one can easily verify. This proves that $k(\alpha,\alpha)^{-1}k(\cdot,\alpha)$ is a solution. Any other solution G has to satisfy $G(\alpha) = I$ and $\langle G, G\rangle = k(\alpha,\alpha)^{-1}$. The expansion of $\langle S - G, S - G\rangle$ shows that it equals zero, thus $G \in [S]_e$ and $G(\alpha) = I$.

(b) If we take in Theorem 1 $k(\cdot,\alpha)$ for F and if G satisfies $\langle G, G\rangle = I$, we have

$$k(\alpha,\alpha) = \langle k(\cdot,\alpha), k(\cdot,\alpha)\rangle \geq \langle k(\cdot,\alpha), G\rangle\langle G, G\rangle^+\langle G, k(\cdot,\alpha)\rangle = G(\alpha)^*G(\alpha).$$

The upper bound $k(\alpha,\alpha)$ for $G(\alpha)^*G(\alpha)$ is obtained for $G(\cdot) = Ak(\cdot,\alpha)$ with A as above. Clearly $\langle G, G\rangle = I$ and $G(\alpha)^*G(\alpha) = k(\alpha,\alpha)$. If F is any other solution of this problem, then it must also satisfy $F(\alpha)^*F(\alpha) = k(\alpha,\alpha)$ and $\langle F, F\rangle = I$. We claim that then F has the form $F(\cdot) = (F(\alpha)^*)^{-1}k(\cdot,\alpha)$ and thus is of the form required by the theorem. That F indeed has this form can be seen by checking that

$$\langle F(\cdot) - (F(\alpha)^*)^{-1}k(\cdot,\alpha), F(\cdot) - (F(\alpha)^*)^{-1}k(\cdot,\alpha)\rangle = 0$$

and thus $F(\cdot) = (F(\alpha)^*)^{-1}k(\cdot,\alpha)$ because $[0]_e = \{0\}$. □

THEOREM 3. *Let M, $\langle\cdot,\cdot\rangle$ be as in* (1), (2) *and let M_e be the Hilbert module as defined before. Let $\{\Phi_k\}_1^N$ be such that $\{[\Phi_k]_e\}_1^N$ span M_e over K and for $1 \leq j$, $k \leq N$, $\langle \Phi_j, \Phi_k\rangle = \delta_{jk}I$. Then $k(\omega_1,\omega_2) = \sum_{i=1}^N \Phi_i(\omega_2)\Phi_i(\omega_1)$ is a $\langle\cdot,\cdot\rangle$-compatible reproducing kernel for M.*

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

PROOF. Write $F \in M$ as $F = \sum_{k=1}^{N} A_k \Phi_k + \Phi$ with $\Phi \in [0]_e$. Then verify that $\langle F(\cdot), k(\cdot, \omega)\rangle = \sum_{k=1}^{N} A_k \Phi_k \in [F]_e$ and the theorem is proved. □

Application of Theorem 2(a). Consider the K-module M of matrix-valued functions analytic inside the unit disk of the complex plane and square integrable over the unit circle, i.e. for which we can define $\forall F, G \in M$,

$$\langle F, G\rangle = \int_{[-\pi,\pi)} F(e^{i\theta}) H(d\theta) G(e^{i\theta})^*$$

in the sense of [**9**]. $H(\cdot)$ is an $n \times n$, nontrivial, nonnegative, hermitian matrix-valued measure on the Borel sets of the real interval $[-\pi, \pi)$. Of course we have a Hilbert module M_e of equivalence classes as introduced before. $F \sim G$ iff $F = G$ a.e. with respect to $H(\cdot)$ as in [**9**]. Consider now the class of functions M_N spanned by $\{[B_k]_e\}_0^N$ over K such that $(M_N)_e$ is a submodule of M_e. Let $k_N(\cdot,\cdot)$ be a $\langle\cdot,\cdot\rangle$-compatible reproducing kernel for M_N. In M_N we minimize $\langle F, F\rangle$ with F constrained by the condition $F(\alpha) = 0$ for some $\alpha, |\alpha| < 1$. By Theorem 2(a), this optimum is $k_N(\alpha,\alpha)^{-1}$, and this is obtained by any function F of the class $[k_N(\alpha,\alpha)^{-1} k_N(\cdot,\alpha)]_e$ for which $F(\alpha) = I$.

We consider two important examples.

(a) Take $B_k = z^k I$, $k = 0, 1, \ldots, N$. M_N consists of matrix polynomials of degree at most N. The solution of the previously described optimization problem is given by Szegö's orthogonal polynomials [**5**] and has applications in AR filtering [**10**].

(b) Take $B_0 = I$, $B_k = ((z - \alpha_k)/(1 - \overline{\alpha}_k z)) B_{k-1}$, $|\alpha_k| < 1$, $k = 1, 2, \ldots, N$. M_N consists of rational matrix functions. The optimization problem is now related to the ARMA filtering problem [**7**]. Recursive construction of the reproducing kernels for increasing N is possible by the Nevanlinna-Pick algorithm [**6**, **4**]. It is a nice generalization of the Schur-Szegö recursion in the polynomial case.

One last remark: The solution given is possible only if $k_N(\alpha,\alpha)$ is invertible. In both previous examples B_0 is constant, and thus $k_0(\alpha,\alpha) = \Phi_0\Phi_0^*$ where $\Phi_0 \in K$ is a normalized version of B_0, e.g. $\Phi_0 = (\int_{[-\pi,\pi)} H(d\theta))^{-1/2} > 0$ if H is nontrivial. If $\{\Phi_k\}_0^N$ is an orthonormal generating set dervied from $\{B_k\}_0^N$ with the properties of Theorem 3, then

$$k_j(\alpha,\alpha) = \sum_{i=1}^{j} \Phi_i(\alpha)\Phi_i(\alpha)^* \geq k_0(\alpha,\alpha) > 0,$$

and thus $k_j(\alpha,\alpha)$ is nonsingular for $j > 0$.

REFERENCES

1. N. Aronzajn, *Theory of reproducing kernels*, Trans. Amer. Math. Soc. **68** (1950), 337–404.
2. A. Ben-Israel and T. N. E. Greville, *Generalized inverses, theory and applications*, Wiley, New York, 1974.
3. J. Bognar, *Indefinite inner product spaces*, Springer-Verlag, Berlin, 1974.
4. A. Bultheel, *Orthogonal matrix functions related to the multivariable Nevanlinna-Pick problem*, Bull. Soc. Math. Belg. Ser. B **31** (1980), 149–170.
5. P. Delsarte, Y. Genin and Y. Kamp, *Orthogonal polynomial matrices on the unit circle*, IEEE Trans. Circuits and Systems **CAS-25** (1978), 1051–1059.
6. ——, *The Nevanlinna-Pick problem for matrix valued functions*, SIAM J. Appl. Math. **36** (1978), 47–61.

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use

7. P. Dewilde and H. Dym, *Lossless chain scattering matrices and optimum linear prediction: the vector case*, Circuit Theory **9** (1981), 135–175.
8. H. Meschkowski, *Hilbertsche Räume mit Kernfunktionen*, Springer-Verlag, Berlin, 1962.
9. M. Rosenberg, *The square integrability of matrix valued functions with respect to a non-negative hermitian measure*, Duke J. Math. **31** (1964), 291–298.
10. N. Wiener and P. Masani, *The prediction theory of multivariable stochastic processes.* I: *Regularity conditions*, Acta Math. **98** (1957), pp. 111–150; II: *The linear predictor*, Acta Math. **99** (1958), 93–139.

Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200A, B-3030 Heverlee, Belgium

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use