# INEQUALITIES IN HILBERT MODULES OF MATRIX VALUED FUNCTIONS.

by A. BULTHEEL

Report TW 52, June 1980

Applied Mathematics and Programming Division
Katholieke Universiteit Leuven (Belgium)

# 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.

Subject classification AMS(MOS) 46C10, 47D15 key words and phrases : Cauchy-Schwarz , reproducing kernel, matrix optimization.

[+] Afdeling Toegepaste Wiskunde en Programmatie,
K.U.Leuven, Celestijnenlaan 200A, B-3030 Belgium

Suppose $K=\mathbb{C}^{nxn}$ is the set of nxn complex constant matrices and M a (left and right K-) module of K-valued functions of a complex variable. We impose an Hilbert structure by introducing a (right) inner product which is a bilinear form $\Phi : MxM \to K$ such that for all $f, g, h \in M$ and $a, b \in K$ :

(P1) $\Phi(h,fa+gb) = \Phi(h,f)a + \Phi(h,g)b$

(P2) $\Phi(f,g) = \Phi(g,f)^{\star}$

(P3) $\Phi(f,f) \geqslant 0$ and $\mathrm{tr}\ \Phi(f,f) = 0 \Leftrightarrow f = 0$

The symbols $^{\star}$ and tr stand for adjoint and trace respectively, while $>(\geqslant)$ are used to indicate the partial ordering of hermitian matrices w.r.t. positive (semi) definiteness. We shall denote $\Phi(f,g)$ by $\langle f,g\rangle$ from now on.

An example of such an inner product which is of $L_2$ type can be found in [5]. The latter is an integral over the unit circle in $\mathbb{C}$ for an hermitian valued measure. This type of inner product is of some importance in the study of multivariable stationary stochastic processes [6].

Hilbert modules have been studied e.g. in [3].

We first give a generalization of the classical Cauchy-Schwarz inequality.

Theorem 1

$$\langle f,f\rangle \geqslant \langle f,g\rangle \langle g,g\rangle^{\dagger} \langle g,f\rangle$$

where $A^{\dagger}$ is the hermitian {2}-inverse of A [1] i.e. $A^{\dagger}$ is hermitian and such that $A^{\dagger}AA^{\dagger}=A^{\dagger}$.

Proof

Properties (P1)-(P3) for the inner product $\langle .,. \rangle$ imply that

$0 \leq \langle fa+gb, fa+gb \rangle = a^{\star}\langle f,f\rangle a + b^{\star}\langle g,g\rangle b - a^{\star}\langle f,g\rangle b - b^{\star}\langle g,f\rangle a.$

Replace a by I and b by $\langle g,g\rangle^{\dagger}\langle g,f\rangle$, then the result follows.

We now require that $\langle .,. \rangle$ is defined such that with the inner product $\ll .,. \gg = tr\langle .,. \rangle$ and the related norm, M becomes an ordinary Hilbert space. Sometimes, one must be very careful in the definition of $\langle .,. \rangle$ to meet this property [5]. If it is satisfied, we have a topology on M and we can speak of an open and closed submodule of M.

The following projection theorem is classical and taken from [5] :

Theorem 2

If E is a closed submodule of M, then $M = E \oplus E^{\perp}$ with

$E^{\perp} = \{f \in M \mid \langle f,g \rangle = 0 \ \forall g \in E\}$, the orthogonal complement of E. □

Before we proceed, let us introduce a terminology and a notation for matrix optimization problems.

If W is a set of arbitrary pxq matrices, and $\psi : W \mapsto K$ a map with hermitian values, then we say that the problem min $\psi(f)$ has a solution F in $V \subset W$ and a minimum $m \in K$ if

$$m = \psi(F) \leq \psi(f) \qquad \forall f \in V$$

and we denote this as

$$\delta[\min_{f \in V \subset W} \psi(f)] = \{F,m\}.$$

We call two optimization problems equivalent if they have the same set of solutions. If the solution of the optimization problem is unique, we write $\delta!$ instead of $\delta$. One can show that the following is true.

Lemma 1

Define

$$P1 : \min_{f \in V} \psi(f) \quad \text{and} \quad P2 : \min_{f \in V} tr\, \psi(f)$$

then if P1 has at least one solution, then P1 and P2 are equivalent. □

Of course, min can be replaced by max and because of the topology on M, also by inf or sup.

We are now ready to reformulate the projection theorem as the solution to an optimization problem. Indeed : if $f \in M$ and $E \subset M$ a submodule of M, then f can be written in a unique way as $f = f'+f''$ with $f' \in E$ and $f'' \in E^{\perp}$. We call $f' \stackrel{def}{=} \langle f|E \rangle$ the (right) projection of f onto E. This projection solves the following optimization problem.

Theorem 3

If $E \subset M$ is closed, $f \in M$ and

$$P : \inf_{p \in E} \langle f-p, f-p \rangle$$

then

$$\delta! [P] = \{\langle f|E\rangle, \langle r, r\rangle\}$$

with

$$r = f - \langle f|E\rangle \in E^{\perp}.$$

The proof is as in the scalar case and is trivial, except perhaps for the uniqueness. The latter follows from lemma 1 and the Hilbert space properties of M with the trace-inner product $\ll .,. \gg$. □

Now suppose that M has a (right) reproducing kernel [4] i.e. there exists some K valued function of two complex variables : k(x,y) such that

(K1) $k(x,y) \in M$ for constant $x$

(K2) $\langle k(x,.), f(.) \rangle = f(x)$ for all $f \in M$.

From the requirements (P1)-(P3) and (K1)-(K2), it follows trivially that the following properties are true.

Theorem 4

If $k(x,y)$ is a reproducing kernel for M, then

(1) if $k(x,y)$ exists, it is unique;

(2) $k(x,x) = \langle k(x,.), k(x,.) \rangle \geqslant 0$;

(3) $k(x,y) = k(y,x)^{\star}$;

(4) If E is a closed submodule of M and $k_E(x,y)$ its reproducing kernel, then for $f \in M$ :

$\langle f | E \rangle = \langle k_E(x,y), f \rangle$. □

We now show that $k(x,y)$ solves some optimization problems, as in the classical case [4].

Define for given $\alpha \in \mathbb{C}$ the following subsets of M :

$V_1(\alpha) = \{f \in M \mid f(\alpha) = I\}$ and $V_2 = \{f \in M \mid \langle f,f \rangle = I\}$

and the related optimization problems :

$P_1(\alpha)$ : $\min_{f \in V_1(\alpha)} \langle f,f \rangle$ and $P_2(\alpha)$ : $\max_{f \in V_2} f(\alpha)\ f(\alpha)^{\star}$,

then we have

Theorem 5

If $k(\alpha,\alpha)$ is invertible, then, with the above notations :

(1) $\delta! [P_1(\alpha)] = \{k(\alpha,y)\ k(\alpha,\alpha)^{-1},\ k(\alpha,\alpha)^{-1}\}$

(this is the projection of 0 onto $V_1(\alpha)$ )

(2) $\delta [P_2(\alpha)] = \{k(\alpha,y)(k(\alpha,\alpha)^{-1})^{L/2},\ k(\alpha,\alpha)\}$.

Here $A^{L/2}$ denotes any left square root of the positive semi definite matrix A, i.e. any square matrix B, such that $BB^{\star}=A$. With this degree of freedom, the above solution of $P_2(\alpha)$ exhausts the whole solution set.

Proof

(1) The Cauchy-Schwarz inequality gives for $h \in M$ :

$$\langle h,h\rangle \geqslant \langle h,k(\alpha,.)\rangle \ \langle k(\alpha,.),k(\alpha,.)\rangle^{-1} \ \langle k(\alpha,.),h\rangle$$

or

$$\langle h,h\rangle \geqslant h(\alpha)^{\star} \ k(\alpha,\alpha)^{-1} \ h(\alpha).$$

If $f(z) = h(z) \ h(\alpha)^{-1}$ and h varies over M with $h(\alpha)$ non singular, then $f(z)$ will vary over all $V_1(\alpha)$ matrices and therefore :

$$\langle f,f\rangle = (h(\alpha)^{\star})^{-1} \ \langle h,h\rangle \ h(\alpha)^{-1} \geqslant k(\alpha,\alpha)^{-1}$$

for all $f \in V_1(\alpha)$.

Now, because

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

we may conclude that $\langle f,f\rangle \geqslant \langle g,g\rangle$ for all $f \in V_1(\alpha)$ if $g = k(\alpha,.)k(\alpha,\alpha)^{-1}$.

This proves the existence of a solution to $P_1(\alpha)$ and uniqueness follows from theorem 3.

(2) If we take in the Cauchy-Schwarz inequality of theorem 1 :

$f = k(\alpha,.)$ and $g = h$, with $\det h(\alpha) \neq 0$ and $\det \langle h,h\rangle \neq 0$, then it follows that

$$h(\alpha) \ \langle h,h\rangle^{-1} \ h(\alpha)^{\star} \leqslant k(\alpha,\alpha).$$

If h varies over M with the given constraints, then $f = h.(\langle h,h\rangle^{-1})^{L/2}$

will vary over $V_2$ and thus

$$f(\alpha)\ f(\alpha)^* = h(\alpha)\ \langle h,h\rangle^{-1}\ h(\alpha)^* \leqslant k(\alpha,\alpha)$$

for all $f \in V_2$.

Now also $g = k(\alpha,.)\ k(\alpha,\alpha)^{-1} \in V_2$ and $g(\alpha)\ g(\alpha)^* = k(\alpha,\alpha)$, so that

$$f(\alpha)\ f(\alpha)^* \leqslant g(\alpha)\ g(\alpha)^* \quad \text{for all } f \in V_2 \text{ if } g = k(\alpha,.)\ k(\alpha,\alpha)^{-1}.$$

Which proves the theorem. □

The technique of matrix optimization has been used to give a least squares optimality interpretation for the multivariable Nevanlinna-Pick algorithm. This has applications in ARMA filtering. The extremality properties of the reproducing kernel play an essential role. The underlying inner product is one in the sense of Rosenberg [5]. For more details, consult [2].

Of course, we gave only a treatment involving "right" quantities, but it is a trivial task to reformulate the given definitions and properties in their left version.

References

1. A. Ben-Israel, T.N.E. Greville, Generalized inverses, theory and applications, J. Wiley & Sons, N.York, 1974.
2. A. Bultheel, Orthogonal matrix functions related to the multivariable Nevanlinna-Pick problem, Report TW49, may 1980, K.U.Leuven, Appl. Math. & Program. Division.
3. R. Delanghe, F. Brackx, "Hypercomplex function theory and Hilbert modules with reproducing kernel", Proc. London Math. Soc., V.37, 1978, pp. 545-576.
4. H. Meschkowski, Hilbertsche Räume mit Kernfunktionen, Springer Verlag, Berlin, 1962.

5. M. Rosenberg, "The square integrability of matrix valued functions with respect to a non-negative hermitian measure", Duke Math. Journ., V.31, 1964, pp. 291-298.

6. N. Wiener, P. Masani, "The prediction theory of multivariable stochastic processes", I regularity conditions, Acta Mathematica, V.98, 1957, pp.111-150.

II The linear predictor, Acta Mathematica, V.99, 1958, pp. 93-139.