Acta Appl Math (2010) 111: 65–81
DOI 10.1007/s10440-009-9532-3

A Geometrical Approach to Indefinite Least Squares Problems

Juan Ignacio Giribet · Alejandra Maestripieri · Francisco Martínez Pería

Received: 11 February 2009 / Accepted: 10 June 2009 / Published online: 20 June 2009

Abstract Given Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, a (bounded) closed range operator $C : \mathcal{H} \to \mathcal{K}$ and a vector $y \in \mathcal{K}$, consider the following indefinite least squares problem: find $u \in \mathcal{H}$ such that $\langle B(Cu - y), Cu - y\rangle = \min_{x\in\mathcal{H}} \langle B(Cx - y), Cx - y\rangle$, where $B : \mathcal{K} \to \mathcal{K}$ is a bounded selfadjoint operator.

This work is devoted to give necessary and sufficient conditions for the existence of solutions of this abstract problem. Although the indefinite least squares problem has been thoroughly studied in finite dimensional spaces, the geometrical approach presented in this manuscript is quite different from the analytical techniques used before. As an application we provide some new sufficient conditions for the existence of solutions of an $\mathcal{H}^\infty$ estimation problem.

Keywords Least squares · Oblique projections · Selfadjoint operators · Weighted generalized inverses

Mathematics Subject Classification (2000) Primary 46C20 · 93E24 · Secondary 15A09 · 47B50

Partially supported by PIP 5272 (CONICET), UBACyT I023, ANPCyT 1728 PICT06.
Partially supported by PIP 5272 (CONICET), UBACyT I023.
Partially supported by PIP 5272 (CONICET), UNLP 11 X472.

J.I. Giribet · A. Maestripieri
Departamento de Matemática, FI-UBA, Buenos Aires, Argentina

J.I. Giribet
e-mail: jgiribet@fi.uba.ar

J.I. Giribet · A. Maestripieri · F. Martínez Pería
IAM-CONICET, Saavedra 15, 3rd floor, 1083 Buenos Aires, Argentina

A. Maestripieri
e-mail: amaestri@fi.uba.ar

F. Martínez Pería (✉)
Departamento de Matemática, FCE-UNLP, La Plata, Argentina
e-mail: francisco@mate.unlp.edu.ar

1 Introduction

Given Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$, a closed range operator $C \in L(\mathcal{H}, \mathcal{K})$, a selfadjoint operator $B \in L(\mathcal{K})$ and a vector $y \in \mathcal{K}$, we say that a vector $u \in \mathcal{H}$ is a B-least squares solution (B-LSS) of the equation $Cx = y$ if it satisfies

$$\langle B(Cu - y), Cu - y\rangle = \min_{x \in \mathcal{H}} \langle B(Cx - y), Cx - y\rangle. \tag{1}$$

Observe that, if B is a fundamental symmetry, i.e. $B = B^* = B^{-1}$, (1) is an optimization problem in Krein spaces that may not admit any solution, even in the finite dimensional case, see [15, 16].

In the last years there has been an increasing interest among engineers in solving practical problems which involve the minimization of certain linear functionals in indefinite metric spaces. In particular, the introduction of Krein spaces in $\mathcal{H}^\infty$ estimation and control techniques made possible to adapt traditional tools of control theory to $\mathcal{H}^\infty$ control problems (see [17] for a complete exposition on this subject). Moreover, some aspects of the indefinite metric spaces theory have provided an explanation to important issues in adaptive filter theory, see [14, 16].

A problem related to B-least squares problems is the *linear state estimation in $\mathcal{H}^\infty$ spaces*. Given the output signal $y = \{y_n\}_{n\in\mathbb{N}} \in \ell^2(\mathbb{C}^q)$ it is intended to estimate the *state of the system* $x = \{x_n\}_{n\in\mathbb{N}} \in \ell^2(\mathbb{C}^m)$ assuming that it satisfies the following dynamics:

$$\begin{aligned} x_{n+1} &= A_n x_n + D_n u_n, \\ y_n &= C_n x_n + v_n, \end{aligned} \tag{2}$$

where $\{A_n\}_{n\in\mathbb{N}} \subseteq L(\mathbb{C}^m, \mathbb{C}^m)$, $\{D_n\}_{n\in\mathbb{N}} \subseteq L(\mathbb{C}^p, \mathbb{C}^m)$ and $\{C_n\}_{n\in\mathbb{N}} \subseteq L(\mathbb{C}^m, \mathbb{C}^q)$ are known sequences of operators and the initial condition $x_0 \in \mathbb{C}^m$ and the signals $u = \{u_n\}_{n\in\mathbb{N}} \in \ell^2(\mathbb{C}^p)$ and $v = \{v_n\}_{n\in\mathbb{N}} \in \ell^2(\mathbb{C}^q)$ are unknown. A broad class of physical systems can be described by (2), which are commonly called a *state space* representation of the system.

The linear state estimation problem consists in finding $F \in L(\ell^2(\mathbb{C}^q), \ell^2(\mathbb{C}^m))$ such that $\hat{x} := Fy$, approximates (in some sense) the state of the system $x = \{x_n\}_{n\in\mathbb{N}}$. There are several different techniques to tackle the linear estimation problem, each one depending on the criterion of approximation selected, for instance the *suboptimal estimation in $\mathcal{H}^\infty$ problem* focuses on, given $\gamma > 0$ find, if it is possible, $\hat{x} \in \ell^2(\mathbb{C}^m)$ such that:

$$\sup_{x_0 \in \mathbb{C}^m, u \in \ell^2(\mathbb{C}^p), v \in \ell^2(\mathbb{C}^q)} \frac{\sum_n \|\hat{x}_n - x_n\|^2}{\|x_0\|^2 + \sum_n \|u_n\|^2 + \sum_n \|v_n\|^2} < \gamma. \tag{3}$$

The above equation can be interpreted as trying to estimate a bound γ for the ratio between the estimation error and the unknown parameters in the system, which are represented by the signals u, v and the initial condition x_0. The smaller γ is, the better the estimation problem can be solved. It is not difficult to see that there exists a solution of the suboptimal estimation in $\mathcal{H}^\infty$ problem if there exists a B-LSS for an appropriate equation $Cx = y$ in some Hilbert space (see Example 3.2).

This work is devoted to give necessary and sufficient conditions for the existence of B-LSS of the equation $Cx = y$. Although the indefinite least squares problem has been thoroughly studied in finite dimensional spaces [3, 15, 24], the geometrical approach presented in this manuscript is quite different from the analytical techniques used before.

Given a closed range operator $C \in L(\mathcal{H}, \mathcal{K})$ and a vector $y \in \mathcal{K}$, the classical least squares problem consists in finding the minimal norm $u \in \mathcal{H}$ such that

$$\|Cu - y\| = \min_{x \in \mathcal{H}} \|Cx - y\|. \tag{4}$$

Observe that $\min_{x \in \mathcal{H}} \|Cx - y\|$ is the distance between the vector y and the (closed) range of C (hereafter denoted $R(C)$) so (4) says that $\|y - Cu\|^2 = \|(I - P)y\|^2$, where P is the orthogonal projection onto $R(C)$. Moreover, (4) holds if and only if $Cu = Py$. Therefore, the (unique) solution of the least squares problem is given by $u = C^\dagger y$, where $C^\dagger$ is the Moore-Penrose inverse of C.

More generally, if $A \in L(\mathcal{K})$ is a (semidefinite) positive operator, a weighted least squares solution of the equation $Cx = y$ is a vector $u \in \mathcal{H}$ such that

$$\|Cu - y\|_A = \min_{x \in \mathcal{H}} \|Cx - y\|_A, \tag{5}$$

where $\| \, \|_A$ is the seminorm on $\mathcal{K}$ defined by $\|x\|_A = \|A^{1/2}x\| = \langle Ax, x\rangle^{1/2}$. This problem was studied in [6, 12] and some applications to statistical problems can be found in [4, 5, 21, 23]. Notice that (5) is equivalent to

$$\langle A(Cu - y), Cu - y\rangle = \min_{x \in \mathcal{H}} \langle A(Cx - y), Cx - y\rangle,$$

so it can be seen as a particular case of (1).

Our approach to the B-least squares problem is, essentially, the same presented above for the classical least squares problem. Using the geometrical properties of the B-selfadjoint projections (selfadjoint projections with respect to the sesquilinear form we are considering) we can find the B-LSS of $Cx = y$ provided that a B-selfadjoint projection onto $R(C)$ exists. However, the existence of these projections is not a trivial fact and it has been studied in several papers, see [7–9, 13].

It is also worthwhile remarking that this manuscript does not intend to provide algorithms to estimate the solutions of the B-least squares problem but to relate the existence of such solutions to some abstract geometrical conditions such as angles between subspaces, oblique projections and generalized inverses.

The paper is organized as follows. Section 2 contains some preliminary results, mainly on B-selfadjoint projections. In Sect. 3 we recall that a necessary condition for the existence of B-LSS is that $R(C)$ is a B-nonnegative subspace (i.e. $\langle x, x\rangle_B \geq 0$ for every $x \in R(C)$). Under this hypothesis, given $y \in \mathcal{K}$, the B-LSS of the equation $Cx = y$ coincide with the solutions of the normal equation associated to the problem:

$$C^* B(Cx - y) = 0.$$

Furthermore, if $\mathcal{K} = R(C) + R(BC)^\perp$ and $y \in \mathcal{K} \setminus R(C)$, we show that $u \in \mathcal{H}$ is a B-LSS of $Cx = y$ if and only if $Cu = Qy$ for some B-selfadjoint projection Q with $R(Q) = R(C)$. We also prove that, for a fixed vector $y \in \mathcal{K}$, the set of solutions of the normal equation is an affine manifold, parallel to the nullspace of BC.

Finally, a minimization problem among the B-LSS of the equation $Cx = y$ is presented. If $A \in L(\mathcal{H})$ is a selfadjoint operator, we look for those $w \in \mathcal{H}$ which are B-LSS of $Cx = y$ and satisfy

$$\langle w, w\rangle_A \leq \langle u, u\rangle_A, \quad \text{for every } B\text{-LSS } u \in \mathcal{H} \text{ of } Cx = y.$$

A vector $w \in \mathcal{H}$ satisfying the above conditions is called an AB-least squares solution (AB-LSS) of the equation $Cx = y$. It is shown that $w \in \mathcal{H}$ is an AB-LSS of $Cx = y$ if and only if $w = (I - Q)C^{\dagger}Py$, where P and Q are appropriate B-selfadjoint and A-selfadjoint projections, respectively. In this case, the operator $D = (I - Q)C^{\dagger}P \in L(\mathcal{K}, \mathcal{H})$ can be seen as a "weighted inverse" of C because it is a solution of

$$CXC = C, \qquad XCX = X, \qquad A(XC) = (XC)^*A, \qquad B(CX) = (CX)^*B.$$

The solutions of the above equations have been studied by X. Sheng and G. Chen [25] in finite dimensional spaces. They can also be seen as an extension of the weighted inverses considered, for positive weights, by Eldén [11] (in finite dimensional spaces) and by G. Corach et al. [6] (in infinite dimensional Hilbert spaces). See also the book by A. Ben-Israel and T.N.E. Greville [1] and the paper by X. Mary [20].

Section 4 is devoted to study the solutions of the normal equation without the restriction of $R(C)$ being B-definite. In this case, there no longer exist B-LSS of the equation $Cx = y$. However, the solutions of the normal equation $C^*B(Cx - y) = 0$ can be related, under certain decomposability condition on $R(C)$, to the solutions of a min-max problem.

2 Preliminaries

Along this work $\mathcal{H}$ and $\mathcal{K}$ denote complex (separable) Hilbert spaces, $L(\mathcal{H}, \mathcal{K})$ is the algebra of bounded linear operators from $\mathcal{H}$ into $\mathcal{K}$, $L(\mathcal{H}) = L(\mathcal{H}, \mathcal{H})$ and $CR(\mathcal{H}, \mathcal{K})$ is the set of (bounded linear) closed range operators from $\mathcal{H}$ into $\mathcal{K}$. If $T \in L(\mathcal{H}, \mathcal{K})$ then $T^* \in L(\mathcal{K}, \mathcal{H})$ denotes the adjoint operator of T, $R(T)$ stands for the range of T and $N(T)$ for its nullspace.

Consider the following subsets of $L(\mathcal{H})$: let $L(\mathcal{H})^+$ be the cone of (semidefinite) positive operators, $L(\mathcal{H})^s$ the (real) vector space of selfadjoint operators and denote by $\mathcal{Q}$ the set of (oblique) projections, i.e. $\mathcal{Q} = \{Q \in L(\mathcal{H}) \ : \ Q^2 = Q\}$.

If $\mathcal{S}$ and $\mathcal{T}$ are two closed subspaces of $\mathcal{H}$, denote by $\mathcal{S} \dotplus \mathcal{T}$ the direct sum of $\mathcal{S}$ and $\mathcal{T}$, $\mathcal{S} \oplus \mathcal{T}$ the (direct) orthogonal sum of them and $\mathcal{S} \ominus \mathcal{T} := \mathcal{S} \cap (\mathcal{S} \cap \mathcal{T})^{\perp}$. If $\mathcal{H} = \mathcal{S} \dotplus \mathcal{T}$, the oblique projection onto $\mathcal{S}$ along $\mathcal{T}$, $P_{\mathcal{S}||\mathcal{T}}$, is the projection with $R(P_{\mathcal{S}||\mathcal{T}}) = \mathcal{S}$ and $N(P_{\mathcal{S}||\mathcal{T}}) = \mathcal{T}$. In particular, $P_{\mathcal{S}} := P_{\mathcal{S}||\mathcal{S}^{\perp}}$ is the orthogonal projection onto $\mathcal{S}$. If $C \in CR(\mathcal{H}, \mathcal{K})$, $C^{\dagger}$ denotes the Moore-Penrose inverse of C.

Given $B \in L(\mathcal{H})^s$ consider the sesquilinear form in $\mathcal{H} \times \mathcal{H}$ defined by $\langle x, y\rangle_B := \langle Bx, y\rangle$, for $x, y \in \mathcal{H}$. If $\mathcal{S}$ is a closed subspace of $\mathcal{H}$, the *B-orthogonal companion* to $\mathcal{S}$ is given by

$$\mathcal{S}^{\perp_B} := \{x \in \mathcal{H} \ : \ \langle x, s\rangle_B = 0 \ \text{ for every } s \in \mathcal{S}\}.$$

It holds that $\mathcal{S}^{\perp_B} = B^{-1}(\mathcal{S}^{\perp}) = B(\mathcal{S})^{\perp}$. Given two closed subspaces $\mathcal{S}$ and $\mathcal{T}$ of $\mathcal{H}$, we say that $\mathcal{S}$ is *B-orthogonal* to $\mathcal{T}$ if $\mathcal{T} \subseteq \mathcal{S}^{\perp_B}$ and denote it by $\mathcal{S} \perp_B \mathcal{T}$.

A vector $x \in \mathcal{H}$ is *B-positive* if $\langle x, x\rangle_B > 0$. A subspace $\mathcal{S}$ of $\mathcal{H}$ is B-positive if every $x \in \mathcal{S}$, $x \neq 0$, is a B-positive vector. B-nonnegative, B-neutral, B-negative and B-nonpositive vectors (and subspaces) are defined analogously.

Also, the B-isotropic part of $\mathcal{S}$ is defined by $\mathcal{N} := \{x \in \mathcal{S} : \langle x, s\rangle_B = 0 \ \forall s \in \mathcal{S}\}$. Observe that $\mathcal{N} = \mathcal{S} \cap \mathcal{S}^{\perp_B}$ and there exist closed subspaces $\mathcal{S}$ of $\mathcal{H}$ such that $\mathcal{N} \neq \{0\}$.

Definition 2.1 Given $B \in L(\mathcal{H})^s$, a closed subspace $\mathcal{S}$ of $\mathcal{H}$ is said to be *B-decomposable* if it can be represented as the B-orthogonal direct sum of a B-neutral subspace $\mathcal{S}_0$, a B-positive subspace $\mathcal{S}_+$ and a B-negative subspace $\mathcal{S}_-$, i.e.

$$\mathcal{S} = \mathcal{S}_0 \dotplus \mathcal{S}_+ \dotplus \mathcal{S}_-.$$

It is important to notice that not every subspace $\mathcal{S}$ of $\mathcal{H}$ is B-decomposable, see [18, Example 1.33]. Observe that, if $\mathcal{S}$ is B-decomposable then $\mathcal{S}_0=\mathcal{N}$, see [18] for a complete exposition on this subject.

An operator $T\in L(\mathcal{H})$ is B-*selfadjoint* if $\langle Tx,y\rangle_B=\langle x,Ty\rangle_B$ for every $x,y\in\mathcal{H}$. It is easy to see that T satisfies this condition if and only if $BT=T^*B$.

Definition 2.2 Let $B\in L(\mathcal{H})^s$ and $\mathcal{S}$ be a closed subspace of $\mathcal{H}$. The pair $(B,\mathcal{S})$ is *compatible* if there exists a B-selfadjoint projection with range $\mathcal{S}$, i.e. if the set

$$\mathcal{P}(B,\mathcal{S}):=\{Q\in\mathcal{Q}\,:\,R(Q)=\mathcal{S},\ BQ=Q^*B\}$$

is not empty.

Notice that a projection Q is B-selfadjoint if and only if $N(Q)\subseteq R(Q)^{\perp_B}$, see [7, Lemma 3.2]. Then, $(B,\mathcal{S})$ is compatible if and only if

$$\mathcal{H}=\mathcal{S}+B(\mathcal{S})^{\perp}.$$

In [10], given two closed subspaces $\mathcal{S}$ and $\mathcal{T}$ of $\mathcal{H}$, the *minimal angle* between $\mathcal{S}$ and $\mathcal{T}$ is defined as the angle in $[0,\pi/2]$ whose cosine is

$$c_0(\mathcal{S},\mathcal{T})=\sup\{|\langle x,y\rangle|\,:\,x\in\mathcal{S},\|x\|\le 1,\ y\in\mathcal{T},\|y\|\le 1\}.$$

By [9, Theorem 2.15], the compatibility of the pair $(B,\mathcal{S})$ is equivalent to the following angle condition:

$$c_0(\mathcal{S}^{\perp},\overline{B(\mathcal{S})})<1.$$

Given a compatible pair $(B,\mathcal{S})$, the B-isotropic part $\mathcal{N}$ of $\mathcal{S}$ coincides with $\mathcal{S}\cap N(B)$ and the Hilbert space $\mathcal{H}$ can be decomposed as $\mathcal{H}=\mathcal{S}\dotplus(B(\mathcal{S})^{\perp}\ominus\mathcal{N})$, so the following oblique projection is well defined:

$$P_{B,\mathcal{S}}:=P_{\mathcal{S}||B(\mathcal{S})^{\perp}\ominus\mathcal{N}}. \tag{6}$$

Observe that $P_{B,\mathcal{S}}\in\mathcal{P}(B,\mathcal{S})$ because $R(P_{B,\mathcal{S}})=\mathcal{S}$ and $N(P_{B,\mathcal{S}})\subseteq B(\mathcal{S})^{\perp}$. In what follows, we state several results about the set $\mathcal{P}(B,\mathcal{S})$ which will be needed later.

Theorem 2.3 *Let $B\in L(\mathcal{H})^s$ and $\mathcal{S}$ be a closed subspace of $\mathcal{H}$ such that $(B,\mathcal{S})$ is compatible. Then $\mathcal{P}(B,\mathcal{S})$ is an affine manifold that can be parametrized as*

$$\mathcal{P}(B,\mathcal{S})=P_{B,\mathcal{S}}+L(\mathcal{S}^{\perp},\mathcal{N}),$$

where $L(\mathcal{S}^{\perp},\mathcal{N})$ is viewed as a subspace of $L(\mathcal{H})$. Moreover, $P_{B,\mathcal{S}}$ has minimal norm in $\mathcal{P}(B,\mathcal{S})$.

Proof See [7, Theorem 3.5]. □

Proposition 2.4 *Let $B\in L(\mathcal{H})^s$ and $\mathcal{S}$ be a closed subspace of $\mathcal{H}$ such that $(B,\mathcal{S})$ is compatible. If $Q\in\mathcal{P}(B,\mathcal{S})$ then*

$$Q=P_{B,\mathcal{S}\ominus\mathcal{N}}+P_{\mathcal{N}||\mathcal{S}\ominus\mathcal{N}+N(Q)}. \tag{7}$$

Proof It is a particular case of [9, Proposition 3.5]. □

Proposition 2.5 *Let $B \in L(\mathcal{H})^s$ and $\mathcal{S}$ be a closed subspace of $\mathcal{H}$ such that $(B, \mathcal{S})$ is compatible. If $x \in \mathcal{H}$ then $(I - P_{B,\mathcal{S}})x$ is the unique minimal norm element in the set*

$$\{(I - Q)x : Q \in \mathcal{P}(B, \mathcal{S})\}. \tag{8}$$

Proof It is analogous to the proof of [8, Theorem 3.2]. □

Proposition 2.6 *Let $B \in L(\mathcal{H})^s$ and $\mathcal{S}$ be a closed subspace of $\mathcal{H}$. Then, $(B, \mathcal{S})$ is compatible if and only if there exists a (unique) orthogonal decomposition $\mathcal{S} \ominus N(B) = \mathcal{S}_+ \oplus \mathcal{S}_-$, where $\mathcal{S}_+$ is a (closed) B-positive subspace, $\mathcal{S}_-$ is a (closed) B-negative subspace, $(B, \mathcal{S}_\pm)$ is compatible and $\mathcal{S}_+ \perp_B \mathcal{S}_-$.*

Proof See [19, Theorem 5.1 and Proposition 5.2]. □

3 Least Squares Problems

In this section, given $C \in CR(\mathcal{H}, \mathcal{K})$, $y \in \mathcal{K}$ and an operator $B \in L(\mathcal{K})^s$, we are interested in characterizing, if there is any, the B-LSS of $Cx = y$, i.e. those vectors $u \in \mathcal{H}$ such that

$$\langle B(Cu - y), Cu - y\rangle = \min_{x \in \mathcal{H}} \langle B(Cx - y), Cx - y\rangle. \tag{9}$$

This kind of problems has been previously studied both in finite and infinite dimensional spaces. The following are some problems that can be translated into B-least squares problems.

Example 3.1 [24] Given two invertible Hermitian matrices $\Pi \in \mathbb{C}^{m\times m}$ and $W \in \mathbb{C}^{n\times n}$, a column vector $y \in \mathbb{C}^n$, and an arbitrary matrix $T \in \mathbb{C}^{n\times m}$, A.H. Sayed et al. studied the following minimization problem: characterize those vectors $z_0 \in \mathbb{C}^m$ such that

$$z_0^* \Pi^{-1} z_0 + (y - Tz_0)^* W^{-1}(y - Tz_0) = \min_{z \in \mathbb{C}^m} \left[z^* \Pi^{-1} z + (y - Tz)^* W^{-1}(y - Tz) \right].$$

Observe that the above problem can be restated as: characterize those vectors $z_0 \in \mathbb{C}^m$ such that

$$\langle B(Cz_0 - w), Cz_0 - w\rangle = \min_{z \in \mathbb{C}^m} \langle B(Cz - w), Cz - w\rangle,$$

where $w = \begin{pmatrix} 0 \\ y \end{pmatrix} \in \mathbb{C}^p$, $B = \begin{pmatrix} \Pi^{-1} & 0 \\ 0 & W^{-1} \end{pmatrix} \in \mathbb{C}^{p\times p}$, $C = \begin{pmatrix} I \\ T \end{pmatrix} \in \mathbb{C}^{p\times m}$ and $p = m + n$.

Example 3.2 (Suboptimal linear estimation in $\mathcal{H}^\infty$, [22]) In what follows we show that there exists a solution of the suboptimal estimation in $\mathcal{H}^\infty$ problem (see the Introduction for the statement of this problem) if there exists a B-LSS for an appropriate equation $Cx = y$ in some Hilbert space.

Given $i \geq j \geq 0$, consider $\Phi_{i,j} \in L(\mathbb{C}^m, \mathbb{C}^m)$ and $\Gamma_{i,j} \in L(\mathbb{C}^p, \mathbb{C}^m)$ defined by

$$\Phi_{i,j} = \begin{cases} A_{i-1} \dots A_j & \text{if } i - 1 > j \\ A_j & \text{if } i - 1 = j \\ I & \text{if } i = j \end{cases} \quad \text{and} \quad \Gamma_{i,j} = \Phi_{i,i-j+1} D_{i-j},$$

respectively. Notice that, for every $n \geq 1$, x_n can be calculated as:

$$x_n = \Phi_{n,0} x_0 + \sum_{k=1}^{n} \Gamma_{n,n-k+1} u_{k-1}.$$

Moreover, if $\Phi \in L(\mathbb{C}^m, \ell^2(\mathbb{C}^m))$ and $\Gamma \in L(\ell^2(\mathbb{C}^p), \ell^2(\mathbb{C}^m))$ are defined by $\Phi z = \{\Phi_{n,0} z\}_{n\in\mathbb{N}}$, $z \in \mathbb{C}^m$, and $\Gamma w = \{\sum_{k=1}^{n} \Gamma_{n,n-k+1} w_{k-1}\}_{n\in\mathbb{N}}$, $w = \{w_n\}_{n\in\mathbb{N}} \in \ell^2(\mathbb{C}^p)$, respectively, then

$$x = \Phi x_0 + \Gamma u.$$

Also, $y = v + H_1 x_0 + H_2 u$ where $H_1 \in L(\mathbb{C}^m, \ell^2(\mathbb{C}^q))$ and $H_2 \in L(\ell^2(\mathbb{C}^p), \ell^2(\mathbb{C}^q))$ are given by $H_1 z = \{C_n \Phi_{n,0} z\}_{n\in\mathbb{N}}$, $z \in \mathbb{C}^m$, and $H_2 w = \{\sum_{k=1}^{n} C_n \Gamma_{n,n-k+1} w_{k-1}\}_{n\in\mathbb{N}}$, $\{w_n\}_{n\in\mathbb{N}} \in \ell^2(\mathbb{C}^p)$.

Let $\mathcal{H} = \mathbb{C}^m \oplus \ell^2(\mathbb{C}^p) \oplus \ell^2(\mathbb{C}^q)$ and $B \in L(\mathcal{H})^s$ the operator with block matrix representation given by

$$B = \begin{pmatrix} I + H_1^* H_1 - \frac{1}{\gamma}\Phi^*\Phi & H_1^* H_2 - \frac{1}{\gamma}\Phi^*\Gamma & -H_1^* + \frac{1}{\gamma}\Phi^* F \\ H_2^* H_1 - \frac{1}{\gamma}\Gamma^*\Phi & I + H_2^* H_2 - \frac{1}{\gamma}\Gamma^*\Gamma & -H_2^* + \frac{1}{\gamma}\Gamma^* F \\ -H_1 + \frac{1}{\gamma}F^*\Phi & -H_2 + \frac{1}{\gamma}F^*\Gamma & I - \frac{1}{\gamma}F^* F \end{pmatrix}.$$

Then, considering the sesquilinear form induced by B, it is easy to see that

$$\begin{aligned} J((x_0, u, y)) &= \langle (x_0, u, y), (x_0, u, y)\rangle_B \\ &= \|x_0\|^2_{\mathbb{C}^m} + \|u\|^2_{\ell^2(\mathbb{C}^p)} + \|y - H_1 x_0 - H_2 u\|^2_{\ell^2(\mathbb{C}^q)} - \frac{1}{\gamma}\|Fy - \Phi x_0 - \Gamma u\|^2_{\ell^2(\mathbb{C}^m)}. \end{aligned}$$

Thus, in (3) we look for those $\hat{x} = Fy$ such that $J((x_0, u, y)) > 0$. Therefore, we are interested in studying if $J(z) = \langle z, z\rangle_B$, $z \in \mathcal{H}$ attains a minimum and if it is positive. Observe that, given the output of the system $y = \{y_n\}_{n\in\mathbb{N}}$, the minimum of $J(z)$ has to be found among the vectors $z \in \mathcal{H}$ such that $P_{\mathcal{S}^\perp} z = (0, 0, y)$, where $\mathcal{S}^\perp = \{0\} \oplus \{0\} \oplus \ell^2(\mathbb{C}^q)$, i.e. a solution is a vector $z_0 \in \mathcal{H}$ such that $P_{\mathcal{S}^\perp} z_0 = (0, 0, y)$ and $\langle z_0, z_0\rangle_B = \min\{\langle z, z\rangle_B : P_{\mathcal{S}^\perp} z = (0, 0, y)\}$ or, rewriting the problem, some $w_0 \in \mathcal{H}$ such that

$$\langle \tilde{y} - P_{\mathcal{S}} w_0, \tilde{y} - P_{\mathcal{S}} w_0\rangle_B = \min_{w\in\mathcal{H}} \langle \tilde{y} - P_{\mathcal{S}} w, \tilde{y} - P_{\mathcal{S}} w\rangle_B,$$

where $\tilde{y} = (0, 0, y) \in \mathcal{H}$.

The next result establishes necessary and sufficient conditions for the existence of B-LSS of the equation $Cx = y$. Similar results have been presented in [6, Remark 4.3] for (semidefinite) positive weights and in [2, Theorem 8.4] for indefinite metric spaces.

Lemma 3.1 *Let $C \in CR(\mathcal{H}, \mathcal{K})$, $B \in L(\mathcal{K})^s$ and $y \in \mathcal{K}$. Then, $u \in \mathcal{H}$ is a B-LSS of the equation $Cx = y$ if and only if $R(C)$ is B-nonnegative and $y - Cu \in R(BC)^\perp$. Hence, if $u, v \in \mathcal{H}$ are two B-LSS of $Cx = y$ then $C(u - v) \in \mathcal{N} = R(C) \cap R(BC)^\perp$.*

Proof Let $u \in \mathcal{H}$ be a B-LSS of $Cx = y$. If $x \in \mathcal{H}$ and $\alpha \in \mathbb{R}$, then

$$\begin{aligned} \langle Cu - y, Cu - y\rangle_B &\leq \langle Cu + \alpha Cx - y, Cu + \alpha Cx - y\rangle_B \\ &= \langle Cu - y, Cu - y\rangle_B + 2\alpha \operatorname{Re}\langle Cu - y, Cx\rangle_B + \alpha^2 \langle Cx, Cx\rangle_B. \end{aligned}$$

Therefore, $2\alpha\,\mathrm{Re}\langle Cu-y, Cx\rangle_B + \alpha^2\langle Cx, Cx\rangle_B \geq 0$ for every $\alpha \in \mathbb{R}$, and a standard argument shows that $\mathrm{Re}\langle Cu-y, Cx\rangle_B = 0$. In the same way, considering $\beta = i\alpha$, $\alpha \in \mathbb{R}$, it follows that $\mathrm{Im}\langle Cu-y, Cx\rangle_B = 0$. Then, $\langle Cu-y, Cx\rangle_B = 0$ and $\langle Cx, Cx\rangle_B \geq 0$ for every $x \in \mathcal{H}$.

Conversely, suppose that $R(C)$ is B-nonnegative and there exists $u \in \mathcal{H}$ such that $y - Cu \in R(BC)^\perp$. Then, for every $x \in \mathcal{H}$,

$$\begin{aligned}\langle y-Cx, y-Cx\rangle_B &= \langle y-Cu, y-Cu\rangle_B + \langle C(u-x), C(u-x)\rangle_B \\ &\geq \langle y-Cu, y-Cu\rangle_B.\end{aligned}$$

Therefore, u is a B-LSS of $Cx = y$.

Observe that, if u and v are B-LSS of $Cx = y$ then $y - Cu, y - Cv \in R(BC)^\perp$. Thus, $C(u-v) = (y-Cv) - (y-Cu) \in R(C) \cap R(BC)^\perp = \mathcal{N}$. □

Let $C \in CR(\mathcal{H},\mathcal{K})$, $B \in L(\mathcal{K})^s$ and $y \in \mathcal{K}$. Given $u \in \mathcal{H}$, notice that $y - Cu \in R(BC)^\perp = N(C^*B)$ if and only if u is a solution of the *normal equation*

$$C^*B(Cx - y) = 0.$$

The following propositions characterize the solutions of the above equation.

Proposition 3.2 *Let $C \in CR(\mathcal{H},\mathcal{K})$ and $B \in L(\mathcal{K})^s$. Given $y \in \mathcal{K}$, consider the normal equation*

$$C^*B(Cx - y) = 0. \tag{10}$$

1. *If $u \in \mathcal{H}$ is a solution of (10), then the set of solutions of (10) coincides with $u + N(C^*BC)$.*
2. *Equation (10) admits a solution for every $y \in \mathcal{K}$ if and only if the pair $(B, R(C))$ is compatible, or equivalently, $c_0(R(C)^\perp, N(C^*B)^\perp) < 1$.*
3. *If $(B, R(C))$ is compatible and $y \in \mathcal{K} \setminus R(C)$, then $u \in \mathcal{H}$ is a solution of (10) if and only if there exists $Q \in \mathcal{P}(B, R(C))$ such that $Cu = Qy$. In this case, the set of solutions of (10) is $u + N(BC)$.*

Proof

1. It is trivial.
2. Notice that there is a solution of $C^*B(Cx-y) = 0$ for every $y \in \mathcal{K}$ if and only if $\mathcal{K} = R(C) + R(BC)^\perp$, and the last condition is equivalent to $c_0(R(C)^\perp, N(C^*B)^\perp) < 1$ (see the Preliminaries).
3. For $y \in \mathcal{K}$ and $u \in \mathcal{H}$ suppose that there exists $Q \in \mathcal{P}(B, R(C))$ such that $Cu = Qy$. Then, $y - Cu = (I - Q)y \in N(Q) \subseteq R(BC)^\perp$ (see Preliminaries), i.e. u is a solution of $C^*B(Cx - y) = 0$.

Conversely, for $y \in \mathcal{K} \setminus R(C)$, let u be a solution of $C^*B(Cx-y) = 0$. Then, $y = Cu + z$ with $z \in R(BC)^\perp$, $z \notin R(C)$. Since $z \in R(BC)^\perp \setminus R(C)$ and $\mathcal{K} = R(C) + R(BC)^\perp$, it is easy to see that there exists a closed subspace $\mathcal{S}$ of $R(BC)^\perp$ such that $z \in \mathcal{S}$ and $\mathcal{H} = R(C) \dot{+} \mathcal{S}$. Therefore, $Q = P_{R(C)||\mathcal{S}} \in \mathcal{P}(B, R(C))$ and

$$Qy = Q(Cu + z) = Cu.$$

Finally, notice that $N(C^*BC) = N(BC)$ because $(B, R(C))$ is compatible. Therefore, the set of solutions of $C^*B(Cx - y) = 0$ coincides with $u + N(BC)$. □

Given $C \in CR(\mathcal{H}, \mathcal{K})$ and $B \in L(\mathcal{K})^s$, suppose that $(B, R(C))$ is compatible. It follows from item *3.* in the above proposition that, given $y \in \mathcal{K}$, if $Cu = Qy$ for some $Q \in \mathcal{P}(B, R(C))$, then u is a solution of the normal equation. But the converse is no longer true when $y \in R(C)$. In this case, $Qy = y$ for every $Q \in \mathcal{P}(B, R(C))$ and the set of solutions of $Cx = y$ is $C^\dagger y + N(C)$ but the set of solutions of the normal equation $C^*B(Cx - y) = 0$ can be parametrized as $C^\dagger y + N(BC)$.

The next result characterizes the set of solutions of the normal equation by means of a family of inner inverses of C.

Proposition 3.3 *Let $C \in CR(\mathcal{H}, \mathcal{K})$ and $B \in L(\mathcal{K})^s$ such that the pair $(B, R(C))$ is compatible. Given $y \in \mathcal{K} \setminus R(C)$, $u \in \mathcal{H}$ is a solution of $C^*B(Cx - y) = 0$ if and only if there exists a solution $D \in L(\mathcal{K}, \mathcal{H})$ of*

$$CXC = C, \qquad BCX = (CX)^*B,$$

such that $Dy = u$.

Proof Given $y \in \mathcal{K} \setminus R(C)$, suppose that $u = Dy$, with $D \in L(\mathcal{K}, \mathcal{H})$ satisfying $CDC = C$ and $BCD = (CD)^*B$. It is easy to see that $Q = CD$ is a B-selfadjoint projection. Furthermore, $R(Q) \subseteq R(C) = R(CDC) \subseteq R(Q)$ i.e. $Q \in \mathcal{P}(B, R(C))$. Then $Cu = Qy$ with $Q \in \mathcal{P}(B, R(C))$ and, by Proposition 3.2, u is a solution of the normal equation.

Conversely, if $u \in \mathcal{H}$ is a solution of $C^*B(Cx - y) = 0$, there exists $Q \in \mathcal{P}(B, R(C))$ such that $Cu = Qy$. Then, $u = C^\dagger Qy + z$ where $z \in N(C)$. Consider an operator $T \in L(\mathcal{K}, \mathcal{H})$ with $R(T) \subseteq N(C)$ such that $Ty = z$, and define $D = C^\dagger Q + T$. Since $CD = C(C^\dagger Q + T) = Q$ it is easy to see that D is a solution of

$$CXC = C, \quad BCX = (CX)^*B,$$

and $Dy = C^\dagger Qy + Ty = u$. □

As it was stated in Lemma 3.1, a necessary condition for the existence of B-LSS of the equation $Cx = y$ is that $R(C)$ is B-nonnegative. Under this hypothesis, and applying the previous results in this section, the following description of the B-least squares problem is obtained.

Theorem 3.4 *Given $B \in L(\mathcal{K})^s$, let $C \in CR(\mathcal{H}, \mathcal{K})$ such that $R(C)$ is B-nonnegative. Then, the following conditions hold:*

1. *Given $y \in \mathcal{K}$, $u \in \mathcal{H}$ is a B-LSS of the equation $Cx = y$ if and only if u is a solution of the normal equation $C^*B(Cx - y) = 0$.*
2. *There exists a B-LSS of the equation $Cx = y$ for every $y \in \mathcal{K}$ if and only if the pair $(B, R(C))$ is compatible. In this case, if $y \in \mathcal{K} \setminus R(C)$, $u \in \mathcal{H}$ is a B-LSS of $Cx = y$ if and only if there exists $Q \in \mathcal{P}(B, R(C))$ such that $Cu = Qy$.*
3. *If $(B, R(C))$ is compatible, the set of B-LSS of the equation $Cx = y$ coincides with the affine manifold $u + N(BC)$, where u is any fixed B-LSS of $Cx = y$.*

In [15], given $B \in L(\mathcal{H})^s$, a (closed) B-nonnegative subspace $\mathcal{S}$ of $\mathcal{H}$ and $y \in \mathcal{H}$, B. Hassibi et al. studied the problem of finding vectors $u \in \mathcal{S}$ such that

$$\langle u - y, u - y\rangle_B = \min_{x\in\mathcal{H}} \langle P_{\mathcal{S}} x - y, P_{\mathcal{S}} x - y\rangle_B, \tag{11}$$

where $P_{\mathcal{S}}$ is the orthogonal projection onto $\mathcal{S}$. They were particularly interested in cases where there is a unique solution of the problem. By Lemma 3.1, $u \in \mathcal{S}$ satisfies (11) if and only if $y - u \in R(BP_{\mathcal{S}})^{\perp}$. It is easy to see that this condition holds if and only if

$$P_{\mathcal{S}} B P_{\mathcal{S}} u = P_{\mathcal{S}} B y. \tag{12}$$

When $\mathcal{H}$ is finite dimensional, if there exists a unique solution $u \in \mathcal{S}$ of (12) for some $y_0 \in \mathcal{H}$, then the operator $P_{\mathcal{S}} B P_{\mathcal{S}}|_{\mathcal{S}}$ is injective. Therefore, $P_{\mathcal{S}} B P_{\mathcal{S}}|_{\mathcal{S}}$ is invertible, and there exists a unique solution of (12) for every $y \in \mathcal{H}$.

If $\mathcal{H}$ is an infinite dimensional Hilbert space this may be not true, since $P_{\mathcal{S}} B P_{\mathcal{S}}|_{\mathcal{S}}$ may be injective but not invertible. In fact, if $\mathcal{H}$ is an infinite dimensional Hilbert space, there is a solution of (12) for every $y \in \mathcal{H}$ if and only if the equation

$$(P_{\mathcal{S}} B P_{\mathcal{S}})X = P_{\mathcal{S}} B$$

admits a solution in $L(\mathcal{H})$, or equivalently, the pair $(B, \mathcal{S})$ is compatible; see [7, Proposition 3.3].

3.1 Minimizing in the Set of B-LSS

In the following paragraphs we study a minimization problem in the set of B-LSS of $Cx = y$. Given $y \in \mathcal{K}$, $C \in CR(\mathcal{H}, \mathcal{K})$, $A \in L(\mathcal{H})^s$ and $B \in L(\mathcal{K})^s$ suppose that $(B, R(C))$ is compatible and $R(C)$ is B-nonnegative. Also, recall that if $(B, R(C))$ is compatible then $N(C^*BC) = N(BC)$.

Definition 3.5 An element $w \in \mathcal{H}$ is an AB-least squares solution (hereafter AB-LSS) of $Cx = y$ if w is a B-LSS of $Cx = y$ and

$$\langle w, w\rangle_A \le \langle u, u\rangle_A, \quad \text{for every } B\text{-LSS } u \text{ of } Cx = y.$$

Before characterizing the set of AB-LSS of $Cx = y$, assuming that $(B, R(C))$ is compatible, we define a particular B-LSS of $Cx = y$ based on $P_{B,R(C)}$, the minimal norm element of $\mathcal{P}(B, R(C))$.

Definition 3.6 Let $C \in CR(\mathcal{H}, \mathcal{K})$, $B \in L(\mathcal{K})^s$ and $y \in \mathcal{K}$. Suppose that the pair $(B, R(C))$ is compatible. Then,

$$u_y := C^{\dagger} P_{B,R(C)} y \tag{13}$$

is the *minimal* B-LSS of the equation $Cx = y$.

Observe that $Cu_y = P_{B,R(C)} y$, so that u_y is a B-LSS of $Cx = y$ (see Theorem 3.4). The following result characterizes it.

Proposition 3.7 *Given $y \in \mathcal{K}$, $u_y \in \mathcal{H}$ is the unique B-LSS of $Cx = y$ in $N(C)^\perp$ which satisfies*

$$\|y - Cu_y\| = \min\{\|y - Cu\| : u \text{ is a } B\text{-LSS of } Cx = y\}. \tag{14}$$

Proof If $y \in R(C)$ it is easy to see that u_y satisfies (14). Let $y \in \mathcal{K} \setminus R(C)$. If u_0 is a B-LSS of $Cx = y$ then there exists $Q \in \mathcal{P}(B, R(C))$ such that $Cu_0 = Qy$ and, by Proposition 2.5,

$$\|y - Cu_0\| = \|(I - Q)y\| \geq \|(I - P_{B,R(C)})y\| = \|y - Cu_y\|.$$

Then, u_0 satisfies $\|y - Cu_0\| = \min\{\|y - Cu\| : u$ is a B-LSS of $Cx = y\}$ if and only if $Qy = P_{B,R(C)}y$, i.e., $Cu_0 = Cu_y$, or equivalently, $u_0 \in u_y + N(C)$. □

Corollary 3.8 *Let $C \in CR(\mathcal{H}, \mathcal{K})$, $B \in L(\mathcal{K})^s$ and $y \in \mathcal{K}$. Suppose that $(B, R(C))$ is compatible and $R(C)$ is B-nonnegative. Then, the following conditions are equivalent*:

1. $\mathcal{N} = \{0\}$;
2. *$R(C)$ is B-positive*;
3. *u_y is the unique B-LSS in $N(C)^\perp$ of $Cx = y$.*

Proof Since $C^*BC \in L(\mathcal{H})^+$, $Cx \in \mathcal{N}$ if and only if $\langle Cx, Cx\rangle_B = 0$. Therefore, item *1* is equivalent to item *2*. On the other hand, item *1* is equivalent to item *3* by Theorem 3.4. □

The following proposition provides a parametrization of the set of AB-LSS of $Cx = y$ under suitable conditions.

Proposition 3.9 *If $A \in L(\mathcal{H})^s$ then, there exists an AB-LSS of the equation $Cx = y$ for every $y \in \mathcal{K}$ if and only if $(A, N(BC))$ is compatible and $N(BC)$ is A-nonnegative. Moreover, if $y \in \mathcal{K} \setminus R(BC)^\perp$, w is an AB-LSS of $Cx = y$ if and only if there exists $Q \in \mathcal{P}(A, N(BC))$ and $P \in \mathcal{P}(B, R(C))$ such that*

$$w = (I - Q)C^\dagger Py.$$

Proof Suppose that, for every $y \in \mathcal{K}$, there exists an AB-LSS $w_y \in \mathcal{H}$ of the equation $Cx = y$, i.e.

$$\langle w_y, w_y\rangle_A = \min_{u \in u_y + N(BC)} \langle u, u\rangle_A, \tag{15}$$

where $u_y = C^\dagger P_{B,R(C)}y$ is the minimal B-LSS of the equation $Cx = y$. Let $z_y \in N(BC)$ such that $w_y = u_y + z_y = u_y + Ez_y$, where $E = P_{N(BC)}$ is the orthogonal projection onto $N(BC)$. Then,

$$\begin{aligned}\langle u_y + Ez_y, u_y + Ez_y\rangle_A = \langle w_y, w_y\rangle_A &= \min_{z \in N(BC)} \langle u_y + z, u_y + z\rangle_A \\ &= \min_{x \in \mathcal{H}} \langle u_y + Ex, u_y + Ex\rangle_A.\end{aligned}$$

Therefore, z_y is an A-LSS of the equation $Ex = -u_y$. Hence, by Lemma 3.1, $R(E) = N(BC)$ is A-nonnegative.

The compatibility of $(B, R(C))$ implies that $\{u_y : y \in \mathcal{K}\} = N(C)^\perp$, therefore the equation $Ex = z$ admits an A-LSS for every $z \in N(C)^\perp$. Also, the equation $Ex = z$ admits an exact solution (which is also an A-LSS) for every $z \in N(C)$ because $N(C) \subseteq R(E)$. Thus, $Ex = z$ admits an A-LSS for every $z \in \mathcal{K}$ and, applying Theorem 3.4, it follows that $(A, N(BC))$ is compatible.

Observe that if $y \notin R(BC)^{\perp}$ then $u_y \notin R(E) = N(BC)$: In fact, $u_y \in N(BC)$ if and only if $BP_{B,R(C)}y = 0$ and $N(BP_{B,R(C)}) = R(BP_{B,R(C)})^{\perp} = R(BC)^{\perp}$, because $P_{B,R(C)}$ is B-selfadjoint. Therefore, by Theorem 3.4, if $y \notin R(BC)^{\perp}$ then there exists $Q \in \mathcal{P}(A, N(BC))$ such that $Ez_y = -Qu_y$. So,

$$w_y = u_y + z_y = u_y + Ez_y = (I - Q)u_y = (I - Q)C^{\dagger}P_{B,R(C)}y.$$

Conversely, suppose that $(A, N(BC))$ is compatible, $N(BC)$ is A-nonnegative and let $Q \in \mathcal{P}(A, N(BC))$. If $P \in \mathcal{P}(B, R(C))$, observe that $(I - Q)C^{\dagger}P = (I - Q)C^{\dagger}P_{B,R(C)}$. Indeed, if $P \in \mathcal{P}(B, R(C))$, there exists $Z \in L(R(C)^{\perp}, \mathcal{N})$ such that $P = P_{B,R(C)} + Z$ (see Theorem 2.3). Then, $(I - Q)C^{\dagger}P = (I - Q)C^{\dagger}P_{B,R(C)}$ because $(I - Q)C^{\dagger}Z = 0$.

Given $y \in \mathcal{K}$, consider $w = (I - Q)C^{\dagger}P_{B,R(C)}y$. Then, $w \in u_y + N(BC)$ and therefore, it is a B-LSS of $Cx = y$ (see Theorem 3.4).

On the other hand, given any B-LSS u of $Cx = y$, there exists $z \in N(BC)$ such that $u = u_y + z = u_y + Qz$ and

$$\begin{aligned}\langle u, u\rangle_A &= \langle (I - Q)u_y + Q(u_y + z), (I - Q)u_y + Q(u_y + z)\rangle_A \\ &= \langle w, w\rangle_A + 2\,\mathrm{Re}\langle (I - Q)u_y, Q(u_y + z)\rangle_A + \langle Q(u_y + z), Q(u_y + z)\rangle_A \\ &= \langle w, w\rangle_A + \langle Q(u_y + z), Q(u_y + z)\rangle_A \geq \langle w, w\rangle_A,\end{aligned}$$

because $R(Q) \perp_A N(Q)$ and $R(Q) = N(BC)$ is A-nonnegative. Thus, w is an AB-LSS of $Cx = y$. □

It follows from the proof of Proposition 3.9 that, given $y \in \mathcal{K} \setminus R(BC)^{\perp}$, the set of AB-LSS of the equation $Cx = y$ is

$$\{(I - Q)C^{\dagger}P_{B,R(C)}y : \; Q \in \mathcal{P}(A, N(BC))\}. \tag{16}$$

On the other hand, if $y \in R(BC)^{\perp}$ then $u_y \in N(BC)$. So, the problem of finding an AB-LSS of $Cx = y$ translates into finding a vector $u \in N(BC)$ such that $\langle u, u\rangle_A = 0$, see (15). Hence, if $y \in R(BC)^{\perp}$, the set of AB-LSS is $\mathcal{M} = N(A) \cap N(BC)$ and (16) describes a proper subset of the set of AB-LSS of $Cx = y$. However, it contains the minimal norm AB-LSS of $Cx = y$.

Proposition 3.10 *Let $A \in L(\mathcal{H})^s$ such that $(A, N(BC))$ is compatible and $N(BC)$ is an A-nonnegative subspace of $\mathcal{H}$. If $y \in \mathcal{K}$ then $v_y = (I - P_{A,N(BC)})C^{\dagger}P_{B,R(C)}y$ is the unique minimal norm element of the set of AB-LSS of $Cx = y$. Moreover, if $\mathcal{M} = \{0\}$ then v_y is the unique AB-LSS of $Cx = y$.*

Proof If $y \in \mathcal{K} \setminus R(BC)^{\perp}$ and v is an AB-LSS of $Cx = y$, there exists $Q \in \mathcal{P}(A, N(BC))$ such that $v = (I - Q)u_y$. By Proposition 2.5,

$$\|v\| = \|(I - Q)u_y\| \geq \|(I - P_{A,N(BC)})u_y\| = \|v_y\|,$$

and $\|v\| = \|v_y\|$ if and only if $v = v_y$. Therefore, $v_y = (I - P_{A,N(BC)})C^{\dagger}P_{B,R(C)}y$ is the unique minimal norm AB-LSS of $Cx = y$.

On the other hand, if $y \in R(BC)^{\perp}$ then $u_y \in N(BC)$. Therefore, v_y is the minimal norm AB-LSS of $Cx = y$ because $v_y = 0$. □

Proposition 3.3 characterize the B-LSS of $Cx = y$ in terms of a set of inner inverses of C. The end of this section is devoted to parametrize the set of AB-LSS of $Cx = y$ in terms of weighted generalized inverses.

Definition 3.11 Given $C \in CR(\mathcal{H}, \mathcal{K})$ and weights $A \in L(\mathcal{H})^s$ and $B \in L(\mathcal{K})^s$, $D \in L(\mathcal{K}, \mathcal{H})$ is a *weighted generalized inverse of* C if D is a solution of

$$CXC = C, \qquad XCX = X, \qquad A(XC) = (XC)^* A, \qquad B(CX) = (CX)^* B. \tag{17}$$

The above equations can be seen as an extension of the previous definitions given for positive weights by Eldén [11] (in finite dimensional spaces) and by G. Corach et al. [6] (in infinite dimensional Hilbert spaces).

The following theorem presents conditions for the existence of weighted generalized inverses (respect to selfadjoint weights A and B) of a closed range operator C and characterizes those inverses in terms of the Moore-Penrose inverse of C and the sets of A-selfadjoint and B-selfadjoint projections. The proof is omitted since it is analogous to the proofs given in [6, Sect. 3].

Theorem 3.12 *Given $C \in CR(\mathcal{H}, \mathcal{K})$, $A \in L(\mathcal{H})^s$ and $B \in L(\mathcal{K})^s$ there exists $D \in L(\mathcal{K}, \mathcal{H})$ such that D is a solution of* (17) *if and only if $(A, N(C))$ and $(B, R(C))$ are compatible pairs. In this case,*

$$GI(C, A, B) = \{(I - Q)C^{\dagger} P : \ Q \in \mathcal{P}(A, N(C)) \text{ and } P \in \mathcal{P}(B, R(C))\}$$

is the set of all bounded linear solutions of (17).

Finally, if $y \in \mathcal{K} \setminus R(BC)^{\perp}$, it is possible to parametrize the set of AB-LSS of the equation $Cx = y$ using a set of weighted generalized inverses.

Proposition 3.13 *Let $C \in CR(\mathcal{H}, \mathcal{K})$, $A \in L(\mathcal{H})^s$ and $B \in L(\mathcal{K})^s$ such that $(B, R(C))$ and $(A, N(BC))$ are compatible, $R(C)$ is B-nonnegative and $N(BC)$ is A-nonnegative. Given $y \in \mathcal{K} \setminus R(BC)^{\perp}$, the set of AB-LSS of the equation $Cx = y$ is given by*

$$\{Ty : T \in GI(P_{R(C)\ominus\mathcal{N}} C, A, B)\}.$$

Proof By Proposition 3.9 we know that given $y \in \mathcal{K} \setminus R(BC)^{\perp}$, the set of AB-LSS of $Cx = y$ is

$$\{(I - Q)C^{\dagger} P y : \ Q \in \mathcal{P}(A, N(BC)), P \in \mathcal{P}(B, R(C))\}.$$

Given $P \in \mathcal{P}(B, R(C))$, let $E = P_{\mathcal{N} || R(C)\ominus\mathcal{N} + N(P)}$. If $\tilde{C} = P_{R(C)\ominus\mathcal{N}} C$ then $R(\tilde{C}) = R(C) \ominus \mathcal{N}$ and, by Proposition 2.4, it follows that $P = P_{B, R(\tilde{C})} + E$. On the other hand, for any $Q \in \mathcal{P}(A, N(BC))$, $(I - Q)C^{\dagger} E = 0$, because $R(C^{\dagger} E) \subseteq N(BC)$. Then the set of AB-LSS of $Cx = y$ can be written as

$$\{(I - Q)C^{\dagger} P_{B, R(\tilde{C})} y : \ Q \in \mathcal{P}(A, N(BC))\}. \tag{18}$$

Next, we show that, if $Q \in \mathcal{P}(A, N(BC))$ then $(I-Q)C^\dagger = (I-Q)\tilde{C}^\dagger$. If $\tilde{C} = P_{R(C)\ominus\mathcal{N}}C$ then it is easy to see that $\tilde{C}^\dagger = P_{N(P_{R(C)\ominus\mathcal{N}}C)^\perp}C^\dagger = P_{N(BC)^\perp}C^\dagger$ because

$$N(P_{R(C)\ominus\mathcal{N}}C) = C^{-1}(\mathcal{N}) = N(BC).$$

Furthermore, if $Q \in \mathcal{P}(A, N(BC))$ then $(I-Q)P_{N(BC)} = 0$. So, $(I-Q)\tilde{C}^\dagger = (I-Q)P_{N(BC)^\perp}C^\dagger = (I-Q)C^\dagger$.

Observe that $R(\tilde{C})$ is a B-non-degenerated subspace of $\mathcal{K}$ (i.e. its B-isotropic part is trivial). Therefore, $\mathcal{P}(B, R(\tilde{C})) = \{P_{B,R(\tilde{C})}\}$ and it is easy to see that the set given in (18) coincides with $\{Ty : T \in GI(\tilde{C}, A, B)\}$. □

4 A Min-Max Solution of the Equation $Cx = y$

In Sect. 3 we established a relationship between the B-LSS of $Cx = y$ and the solutions of the normal equation $C^*B(Cx - y) = 0$. More precisely, supposing that $R(C)$ is a B-nonnegative subspace of $\mathcal{K}$, we proved that $u \in \mathcal{H}$ is a B-LSS of $Cx = y$ if and only if it is a solution of the normal equation $C^*B(Cx - y) = 0$. Notice that, if $R(C)$ is a B-nonpositive subspace of $\mathcal{K}$, vectors $u \in \mathcal{H}$ satisfying

$$\langle B(Cu - y), Cu - y\rangle = \max_{x\in\mathcal{H}} \langle B(Cx - y), Cx - y\rangle,$$

can be characterized following the same ideas as in the B-least squares problem. In fact, similar results to those given in Theorem 3.4 can be established mutatis mutandis.

The purpose of this section is to study the solutions of the normal equation without the restriction of $R(C)$ being B-definite. We will prove that the solutions of the normal equation $C^*B(Cx - y) = 0$ can be related to the solutions of a min-max problem.

Definition 4.1 Let $C \in CR(\mathcal{H}, \mathcal{K})$ and $B \in L(\mathcal{K})^s$ such that $R(C)$ is B-decomposable. Given $y \in \mathcal{K}$, a vector $u \in \mathcal{H}$ is a B-min-max solution (B-MMS) of the equation $Cx = y$ if

$$\langle y - Cu, y - Cu\rangle_B = \min_{s\in\mathcal{S}_+}\max_{t\in\mathcal{S}_-}\langle y - (s+t), y - (s+t)\rangle_B \tag{19}$$

$$= \max_{t\in\mathcal{S}_-}\min_{s\in\mathcal{S}_+}\langle y - (s+t), y - (s+t)\rangle_B \tag{20}$$

where $R(C) = \mathcal{N} \dotplus \mathcal{S}_+ \dotplus \mathcal{S}_-$ is a decomposition as in Definition 2.1.

Remark 4.2 Given a decomposition $R(C) = \mathcal{N} \dotplus \mathcal{S}_+ \dotplus \mathcal{S}_-$ as above, observe that $\mathcal{N} \dotplus \mathcal{S}_+$ is a closed subspace and

$$\max_{t\in\mathcal{S}_-}\min_{s\in\mathcal{S}_+}\langle y - (s+t), y - (s+t)\rangle_B = \max_{t\in\mathcal{S}_-}\min_{x\in\mathcal{S}_+ + \mathcal{N}}\langle y - (x+t), y - (x+t)\rangle_B. \tag{21}$$

Indeed, for a fixed $t \in \mathcal{S}_-$,

$$\min_{x\in\mathcal{S}_+ + \mathcal{N}}\langle y - (x+t), y - (x+t)\rangle_B \le \min_{s\in\mathcal{S}_+}\langle y - (s+t), y - (s+t)\rangle_B$$

because $\mathcal{S}_+ \subseteq \mathcal{S}_+ + \mathcal{N}$. On the other hand, if $w \in \mathcal{S}_+ + \mathcal{N}$ satisfies

$$\langle y - (w+t), y - (w+t)\rangle_B = \min_{x\in\mathcal{S}_+ + \mathcal{N}}\langle y - (x+t), y - (x+t)\rangle_B,$$

then, by Lemma 3.1, $y-(w+t)$ is B-orthogonal to $\mathcal{S}_+ + \mathcal{N}$, that is $\langle y-(w+t), x\rangle_B = 0$ for every $x \in \mathcal{S}_+ + \mathcal{N}$. Suppose that $w = s_0 + n_0$, with $s_0 \in \mathcal{S}_+$ and $n_0 \in \mathcal{N}$. Hence, $\langle (y-t)-w, n_0\rangle_B = 0$ and $\langle n_0, n_0\rangle_B = 0$ because $n_0 \in \mathcal{N}$. Therefore,

$$\begin{aligned}\langle y-(w+t), y-(w+t)\rangle_B &= \langle (y-t)-s_0, (y-t)-s_0\rangle_B \\ &\geq \min_{s\in\mathcal{S}_+} \langle y-(t+s), y-(t+s)\rangle_B.\end{aligned}$$

So, considering the maximum over the vectors $t \in \mathcal{S}_-$, (21) follows.

Also, notice that the decomposition $R(C) = \mathcal{N} \dotplus \mathcal{S}_+ \dotplus \mathcal{S}_-$ in Definition 2.1 is not necessarily unique. However, the following result shows that the B-MMS definition is independent of the selected decomposition. Furthermore, it characterizes the B-MMS of the equation $Cx = y$. Along the following paragraphs, $\mathcal{Z}$ denotes the set of B-neutral vectors in $\mathcal{K}$.

Theorem 4.3 *Let $C \in CR(\mathcal{H}, \mathcal{K})$ and $B \in L(\mathcal{K})^s$ such that $R(C)$ is B-decomposable. Given $y \in \mathcal{K}$, $u \in \mathcal{H}$ is a B-MMS of $Cx = y$ if and only if $u \in u_0 + C^{-1}(\mathcal{Z})$, where $u_0 \in \mathcal{H}$ is a solution of the normal equation $C^*B(Cx-y) = 0$.*

Proof Suppose that $u \in \mathcal{H}$ is a B-MMS of $Cx = y$, Then, by Remark 4.2,

$$\langle y-Cu, y-Cu\rangle_B = \max_{t\in\mathcal{S}_-} \min_{s\in\mathcal{S}_+} \langle y-(s+t), y-(s+t)\rangle_B,$$

where $R(C) = \mathcal{N} \dotplus \mathcal{S}_+ \dotplus \mathcal{S}_-$ is a decomposition as in Definition 2.1. Fixed $t \in \mathcal{S}_-$, by Lemma 3.1, it follows that $\min_{s\in\mathcal{S}_+} \langle (y-t)-s, (y-t)-s\rangle_B$ is attained at $s_0(t) \in \mathcal{S}_+$ if and only if $\langle y-t-s_0(t), x\rangle_B = 0$ for every $x \in \mathcal{S}_+$. Then, $\langle y-s_0(t), x\rangle_B = \langle y-t-s_0(t), x\rangle_B = 0$ for every $x \in \mathcal{S}_+$ because $\mathcal{S}_-$ is B-orthogonal to $\mathcal{S}_+$. Therefore, $s_0(t) = s_0$ is the unique vector in $\mathcal{S}_+$ which satisfies

$$\langle (y-t)-s_0, (y-t)-s_0\rangle_B = \min_{s\in\mathcal{S}_+} \langle (y-t)-s, (y-t)-s\rangle_B.$$

Hence, $\langle y-Cu, y-Cu\rangle_B = \max_{t\in\mathcal{S}_-} \langle (y-s_0)-t, (y-s_0)-t\rangle_B$.

Analogously, this maximum is attained at $t_0 \in \mathcal{S}_-$ if and only if $\langle y-s_0-t_0, t\rangle_B = 0$ for every $t \in \mathcal{S}_-$. Moreover, by Remark 4.2, $\langle y-s_0-t_0, x\rangle_B = 0$ for every $x \in \mathcal{S}_- + \mathcal{N}$.

Let $u_0 \in \mathcal{H}$ such that $Cu_0 = s_0 + t_0$. Since $R(C) = \mathcal{N} \dotplus \mathcal{S}_+ \dotplus \mathcal{S}_-$, it is easy to see that $\langle y-Cu_0, z\rangle_B = 0$ for every $z \in R(C)$. Hence, u_0 is a solution of the normal equation $C^*B(Cx-y) = 0$ and

$$\langle y-Cu, y-Cu\rangle_B = \langle y-Cu_0, y-Cu_0\rangle_B.$$

Since $\langle y-Cu_0, z\rangle_B = 0$ for every $z \in R(C)$, then

$$\begin{aligned}\langle y-Cu_0, y-Cu_0\rangle_B &= \langle y-Cu, y-Cu\rangle_B \\ &= \langle (y-Cu_0)+(Cu_0-Cu), (y-Cu_0)+(Cu_0-Cu)\rangle_B \\ &= \langle y-Cu_0, y-Cu_0\rangle_B + \langle Cu_0-Cu, Cu_0-Cu\rangle_B,\end{aligned}$$

and the above equation holds if and only if $\langle Cu_0-Cu, Cu_0-Cu\rangle_B = 0$. Therefore, $u \in u_0 + C^{-1}(\mathcal{Z})$.

Conversely, consider a solution $u_0 \in \mathcal{H}$ of the normal equation $C^*B(Cx - y) = 0$ and let $u \in u_0 + C^{-1}(\mathcal{Z})$. Then, $\langle Cu - y, Cu - y\rangle_B = \langle Cu_0 - y, Cu_0 - y\rangle_B$.

Suppose that $R(C) = \mathcal{N} \dot{+} \mathcal{S}_+ \dot{+} \mathcal{S}_-$ is a decomposition as in Definition 2.1, and let $x_0 \in \mathcal{S}_+ + \mathcal{N}$ and $t_0 \in \mathcal{S}_-$ such that $Cu_0 = x_0 + t_0$. Since $\langle y - Cu_0, x\rangle_B = 0$ for every $x \in R(C)$ and $\mathcal{S}_+ + \mathcal{N}$ is B-orthogonal to $\mathcal{S}_-$, it is easy to see that $\langle y - x_0 - t, x\rangle_B = \langle y - x - t_0, t\rangle_B = 0$ for every $x \in \mathcal{S}_+ + \mathcal{N}$ and $t \in \mathcal{S}_-$. Then, considering the equation $Px = y - t_0$ (where P is the orthogonal projection onto $\mathcal{S}_+ + \mathcal{N}$) it follows by Lemma 3.1 that

$$\begin{aligned}\langle y - Cu, y - Cu\rangle_B &= \langle y - x_0 - t_0, y - x_0 - t_0\rangle_B \\ &= \min_{x\in\mathcal{H}} \langle y - t_0 - Px, y - t_0 - Px\rangle_B \\ &= \min_{s\in\mathcal{S}_+ + \mathcal{N}} \langle y - t_0 - s, y - t_0 - s\rangle_B \\ &= \min_{s\in\mathcal{S}_+} \langle y - t_0 - s, y - t_0 - s\rangle_B \\ &= \min_{s\in\mathcal{S}_+ + \mathcal{N}} \max_{t\in\mathcal{S}_-} \langle s + t - y, s + t - y\rangle_B. \qquad \square\end{aligned}$$

Corollary 4.4 *Let $C \in CR(\mathcal{H}, \mathcal{K})$ and $B \in L(\mathcal{K})^s$ such that $R(C)$ is B-decomposable. Then, there exists a B-MMS of $Cx = y$ for every $y \in \mathcal{K}$ if and only if $(B, R(C))$ is compatible. In this case, u is a B-MMS of $Cx = y$ if and only if $Cu \in P_{B,R(C)}y + \mathcal{Z}$.*

Proof Suppose that $(B, R(C))$ is compatible. If u is a B-MMS of $Cx = y$ then $u = u_0 + z$, where u_0 is a solution of the normal equation $C^*B(Cx - y) = 0$ and $z \in C^{-1}(\mathcal{Z})$. By Proposition 3.2, $u_0 = C^\dagger P_{B,R(C)}y + n$ with $n \in N(BC) = C^{-1}(\mathcal{N})$. Hence, $Cu = Cu_0 + Cz = P_{B,R(C)}y + C(n + z) \in P_{B,R(C)}y + \mathcal{Z}$.

Conversely, if $Cu = P_{B,R(C)}y + z$ and $z \in \mathcal{Z}$ then $u = C^\dagger P_{B,R(C)}y + (C^\dagger z + P_{N(C)}u) \in u_y + C^{-1}(\mathcal{Z})$, where u_y is the minimal solution of the normal equation $C^*B(Cx - y) = 0$. Then, by Theorem 4.3, u is a B-MMS of the equation $Cx = y$. $\square$

Acknowledgements The authors would like to acknowledge the referee, whose comments helped us to improve this work.

References

1. Ben-Israel, A., Greville, T.N.E.: Generalized Inverse: Theory and Applications. Springer, Berlin (2003)
2. Bognár, J.: Indefinite Inner Product Spaces. Springer, Berlin (1974)
3. Chandrasekaran, S., Gu, M., Sayed, A.H.: A stable and efficient algorithm for the indefinite least squares problem. SIAM J. Matrix Anal. Appl. **20**(2), 354–362 (1998)
4. Chipman, J.S.: On least squares with insufficient observations. J. Am. Stat. Assoc. **59**, 1078–1111 (1964)
5. Chipman, J.S.: Specification problems in regression analysis. In: Proc. Sympos. Theory and Application of Generalized Inverses of Matrices (Lubbock, Texas, 1968), pp. 114–176. Texas Tech. Press, Lubbock (1968)
6. Corach, G., Maestripieri, A.: Weighted generalized inverses, oblique projections and least squares problems. Numer. Funct. Anal. Optim. **26**(6), 659–673 (2005)
7. Corach, G., Maestripieri, A., Stojanoff, D.: Oblique projections and Schur complements. Acta Sci. Math. (Szeged) **67**, 337–256 (2001)
8. Corach, G., Maestripieri, A., Stojanoff, D.: Oblique projections and abstract splines. J. Approx. Theory **117**, 189–206 (2002)
9. Corach, G., Maestripieri, A., Stojanoff, D.: A classification of projectors. Banach Cent. Publ. **67**, 145–160 (2004)
10. Dixmier, J.: Etudes sur les variétés et opérateurs de Julia, avec quelques applications. Bull. Soc. Math. France **77**, 11–101 (1949)

11. Eldén, L.: Perturbation theory for the least squares problem with linear equality constraints. SIAM J. Numer. Anal. **17**(3), 338–350 (1980)
12. Golub, G.H., Van Loan, C.F.: An analysis of the total least squares problem. SIAM J. Numer. Anal. **17**(6), 883–893 (1980)
13. Hassi, S., Nordström, K.: On projections in a space with an indefinite metric. Linear Algebra Appl. **208/209**, 401–417 (1994)
14. Hassibi, B., Sayed, A.H., Kailath, T.: $\mathcal{H}^\infty$ optimality criteria for LMS and backpropagation. Adv. Neural Inf. Process. Syst. **6**, 351–359 (1994)
15. Hassibi, B., Sayed, A.H., Kailath, T.: Linear estimation in Krein spaces—Part I: Theory. IEEE Trans. Autom. Control **41**(1), 18–33 (1996)
16. Hassibi, B., Sayed, A.H., Kailath, T.: Linear estimation in Krein spaces—Part II: Application. IEEE Trans. Autom. Control **41**(1), 33–49 (1996)
17. Hassibi, B., Sayed, A.H., Kailath, T.: Indefinite-Quadratic Estimation and Control. A Unified Approach to $\mathcal{H}^2$ and $\mathcal{H}^\infty$ Theories. Studies in Applied and Numerical Mathematics. SIAM, Philadelphia (1999)
18. Iokhvidov, I.S., Azizov, T.Ya.: Linear Operators in Spaces with an Indefinite Metric. Wiley, New York (1989)
19. Maestripieri, A., Martínez Pería, F.: Decomposition of selfadjoint projections in Krein spaces. Acta Sci. Math. **72**, 611–638 (2006)
20. Mary, X.: Moore-Penrose inverse in Krein spaces. Integral. Equ. Oper. Theory **60**(3), 419–433 (2008)
21. Morley, T.D.: A Gauss-Markov theorem for infinite-dimensional regression models with possibly singular covariance. SIAM J. Appl. Math. **37**(2), 257–260 (1979)
22. Nagpal, K.M., Khargonekar, P.P.: Filtering and smoothing in an $\mathcal{H}^\infty$ setting. IEEE Trans. Autom. Control **36**, 152–166 (1991)
23. Rao, C., Mitra, S.K.: Theory and application of constrained inverse of matrices. SIAM J. Appl. Math. **24**, 473–488 (1973)
24. Sayed, A.H., Hassibi, B., Kailath, T.: Fundamental Inertia Conditions for the Minimization of Quadratic Forms in Indefinite Metric Spaces. Oper. Theory: Adv. Appl. Birkhauser, Cambridge (1996)
25. Sheng, X., Chen, G.: The generalized weighted Moore-Penrose inverse. J. Appl. Math. Comput. **25**, 407–413 (2007)