Connections between some results on the generalized linear least squares problem

Citation for published version (APA):
Connections between some results on the generalized linear least squares problem

by

A.D. Poley

T.H.-Report 81-WSK-01
May 1981
Abstract
This paper deals with the generalized least squares problem

"find \( f \) which solves \( \min \{ \| Kf \|_2 \mid \| M(\mathbf{L}f - \mathbf{h}) \|_2 \) is minimal \}."" 

It is shown that the solution can be written as \( f = L^+_M \mathbf{h} \), where \( L^+_M \) is a solution matrix such that, in case of nonuniqueness (i.e. \( N(ML) \cap N(K) \neq \{0\} \)), \( f \) has minimal Euclidean norm. This \( L^+_M \) is uniquely determined by Penrose-like conditions.

1980 AMS Mathematics Subject Classification. Primary 65F20; Secondary 62J05.
1. Introduction

In this paper we will consider the generalized least squares problem
(1.1.1) "find a vector \( f \) which minimizes \( \| M(Lf - h) \|_2 \)."

If \( ML \) does not have full column rank, then \( f \) is not uniquely determined by
(1.1.1) and we can prescribe additional conditions for \( f \); for instance, we can consider the problem
(1.1.2) "find \( f \) which solves \( \min \{ \| Kf \|_2 \mid \| M(Lf - h) \|_2 \text{ is minimal} \} \)."

In these problems \( K, M \) and \( L \) have to satisfy no other conditions than that
their dimensions fit together. However, it is immediately clear that, if \( N(ML) \cap N(K) \neq \{0\} \),
even problem (1.1.2) has no unique solution.

To start with, in §2 we consider a statistical problem similar to (1.1.1)
and in solving that we follow the methods used by C.C. Paige [1].

In §3 we use similar methods to solve the more general problem (1.1.1).

Afterwards we simplify the solution and obtain a good starting point for
attacking problem (1.1.2) in §4. First we solve this problem under
the condition \( N(ML) \cap N(K) = \{0\} \) (cf. Eldén [2]), next we consider the
properties of the solution found without assuming that condition.

In §5 we discuss what Penrose-like conditions the solution matrix corre­
sponding to problem (1.1.2) satisfies, and under which extra condition
a matrix, which satisfies these conditions, is unique. Also we compare our
results with those of Ben-Israel and Greville [3, Sec. 3.3].

Finally, in the last paragraph we consider the solution of problem (1.1.2)
by use of Lagrange multipliers.

We list a few notations to be used throughout this report:
* \( R(A) \) and \( N(A) \) stand for the range and nullspace of a matrix \( A \), respecti­

vely.
* $\| \cdot \|$ means the Euclidean vector norm.
* $\| A \|_F = (\tr(A^HA))^{\frac{1}{2}}$ is the Frobenius matrix norm.
* $A \in \mathbb{R}^n$ or $A \in \mathbb{C}^n$ means $A$ has full column rank or full row rank, respectively.
* A matrix $U$ is called left-unitary when $U^HU = I$.
* $S^\perp$ means the orthogonal complement of a subspace $S$.
* Superscripts to a matrix refer to the Penrose conditions that matrix satisfies. These conditions are

$$(1) \quad AXA = A,$$
$$(2) \quad XAX = X,$$
$$(3) \quad (AX)^H = AX,$$
$$(4) \quad (XA)^H = XA.$$  

For instance, if $X$ satisfies (1) and (3), we write $X = A^{(1,3)}$.

The pseudoinverse (or Moore-Penrose inverse) $A^{(1,2,3,4)}$ which satisfies all four conditions, will be denoted by $A^+$.

* The $M,K$-weighted pseudoinverse of $L$ is defined by

$$L^+_{MK} = (I - (K,E_0^+)K)(ML)^+M,$$

where $E_0 = I - (ML)^+ML$, cf. Eldén [2, §2].

2. Paige's method of solution

Paige [1] considers the following stochastic model. Let $W$ be a nonnegative definite Hermitian matrix and $w$ a stochastic with $\mathbb{E}(w) = 0$, $\mathbb{E}(ww^H) = \sigma^2 W$. Let $y = CX + w$, with $C$ a known matrix and $x$ a fixed but unknown parameter-vector. The problem now is to

(2.1.1) "estimate $x$ from a realisation of $y".$

Let $W = BB^H$ with $B \in \mathbb{R}^n$, then (since first and second moments of $w$ and $Bv$ are equal) Paige's model can be reduced to the following equivalent form.
Let v be a stochast with \( \mathcal{E}(v) = 0 \) and \( \mathcal{E}(vv^H) = c^2 I \). Let \( y = Cx + Bv \), with C a known matrix and x a parameter-vector. In stochastics it is shown (see Appendix) that problem (2.1.1) leads to

"find to a given \( y \in \mathbb{R}((|C|B)) \) vectors x and v that minimize \( \|v\|^2 \) under the condition \( y = Cx + Bv \)."

The essence of Paige's solution method is to decompose C as

\[
C = QR = (Q_1|Q_2) \begin{pmatrix} R_1 \\ 0 \end{pmatrix} = Q_1R_1,
\]

with Q unitary and \( R_1 \in \mathbb{L}_R \), and to define

\[
T = \begin{pmatrix} T_1 \\ T_2 \end{pmatrix} := \begin{pmatrix} Q_1^H \\ Q_2^H \end{pmatrix} B.
\]

Then the condition \( y = Cx + Bv \) is equivalent to

(i) \( Q_1^H y = T_1 v + R_1 x \),
(ii) \( Q_2^H y = T_2 v \).

Since \( R_1 \) has full row rank, there exists a solution x of (i) to every y and v. Hence (i) does not constrain v and the compatibility condition

\( y \in \mathbb{R}(C) + \mathbb{R}(B) \) is equivalent to \( Q_2^H y \in \mathbb{R}(T_2) \). As is well known, the minimum 2-norm vector v that satisfies the compatible system (ii) is

\[
\mathcal{V} := T_2^* Q_2^* y.
\]

So x has to satisfy

\[
(2.1.3) \quad R_1 x = (Q_1^H - T_1 T_2^* Q_2^H) y = Q_1^H (I - B Q_2 H^B Q_2^H) y.
\]

The solution of (2.1.3) with minimal 2-norm can be written as

\[
\mathcal{X} := R_1^+ Q_1^H (I - B Q_2^H B^+ Q_2^H) y = C^+ (I - B Q_2 H^B Q_2^H) y.
\]

The matrix \( C^+ (I - B Q_2 H^B Q_2^H) \) will be called the solution matrix to
problem (2.1.1).

Remark. The condition $B \in RR$ is necessary for the reduction of Paige's model to its equivalent form, however, it does not play any role in the solution method of problem (2.1.2).

3. Application of Paige's method to problem (1.1.1)

3.1. Now we will derive the solution of problem (1.1.1) by a method similar to that used in §2. By defining $v := -\overline{M}(Lf - h)$ we can formulate (1.1.1) as

\[(3.1.1) \text{"find } f \text{ and } v \text{ that solve } \min_{f,v} \{ \|v\| \mid Mlf + v = Mh \}.\]

So we can use the results of §2, with $B = I$, $C = M = QR = Q_1R_1$ with $Q$ unitary and $R_1 \in LR$. If we substitute this in (2.1.3), we find that $f$

must satisfy

\[(3.1.2) \quad R_1f = Q_1^H(I - Q_2Q_2^H)Mh = Q_1^HMh.\]

The solution with minimum 2-norm of (3.1.2) is

\[\hat{f} := R_1^+Q_1^HMh.\]

Since $R_1^+Q_1^H = (ML)^+$, we thus obtain the well-known minimum 2-norm solution of (1.1.1),

\[(3.1.3) \quad \hat{f} = (ML)^+Mh.\]

3.2. If we want to obtain the solution matrix of problem (1.1.1) in a form similar to that of problem (2.1.2) in §2 (e.g. for reasons of symmetry, see §5.3), we observe that (1.1.1) is also equivalent to

\[(3.2.1) \text{"find } f \text{ and } v \text{ that solve } \min_{f,v} \{ \|v\| \mid M^+MLf + M^+v = M^+Mh \}.\]

Using again the method of §2, now with $C = M^+ML = \tilde{Q}\tilde{R} = Q_1\tilde{R}_1$ with $\tilde{Q}$ unitary and $\tilde{R}_1 \in LR$, we find after substitution in (2.1.3) that $f$

must satisfy
The solution with minimal 2-norm of (3.2.2) (and so of (1.1.1)) is

$$\mathbf{f} := (\mathbf{M}'\mathbf{M}L)^+(I - \mathbf{M}'(\tilde{Q}_2^H\mathbf{M}^+)^+\tilde{Q}_2^H)\mathbf{M}^+\mathbf{M}h,$$

which should be equal to (3.1.3). Indeed, we have

$$\mathbf{M}'\mathbf{M}L(I - \mathbf{M}^+(\tilde{Q}_2^H\mathbf{M}^+)^+\tilde{Q}_2^H)\mathbf{M}^+ = (\mathbf{M}L)^+,$$

which can be proved by verifying that the left-hand side of (3.2.4) satisfies the four \((\mathbf{M}L)^+\)-Penrose conditions:

1. \[
\mathbf{M}'\mathbf{M}L(I - \mathbf{M}^+(\tilde{Q}_2^H\mathbf{M}^+)^+\tilde{Q}_2^H)\mathbf{M}^+ = \mathbf{M}'\mathbf{M}L(I - \mathbf{M}^+(\tilde{Q}_2^H\mathbf{M}^+)^+\tilde{Q}_2^H)\mathbf{M}^+ = \mathbf{M}'\mathbf{M}L \quad \text{(since } \tilde{Q}_2^H\mathbf{M}^+\mathbf{M}L = 0),
\]

2. \[
(\mathbf{M}'\mathbf{M}L)^+(I - \mathbf{M}^+(\tilde{Q}_2^H\mathbf{M}^+)^+\tilde{Q}_2^H)\mathbf{M}^+\mathbf{M}'\mathbf{M}L = (\mathbf{M}'\mathbf{M}L)^+(I - \mathbf{M}^+(\tilde{Q}_2^H\mathbf{M}^+)^+\tilde{Q}_2^H)\mathbf{M}^+.
\]

3. \[
\mathbf{M}'\mathbf{M}L(\mathbf{M}'\mathbf{M}L)^+(I - \mathbf{M}^+(\tilde{Q}_2^H\mathbf{M}^+)^+\tilde{Q}_2^H)\mathbf{M}^+ = \mathbf{M}(I - \mathbf{Q}_2\mathbf{Q}_2^H)(I - \mathbf{M}^+(\tilde{Q}_2^H\mathbf{M}^+)^+\tilde{Q}_2^H)\mathbf{M}^+ = \mathbf{M}(I - \mathbf{M}^+(\tilde{Q}_2^H\mathbf{M}^+)^+\tilde{Q}_2^H)\mathbf{M}^+ = \mathbf{M}(I - \mathbf{M}^+(\tilde{Q}_2^H\mathbf{M}^+)^+\tilde{Q}_2^H)\mathbf{M}^+ \quad \text{is Hermitian},
\]

4. \[
(\mathbf{M}'\mathbf{M}L)^+(I - \mathbf{M}^+(\tilde{Q}_2^H\mathbf{M}^+)^+\tilde{Q}_2^H)\mathbf{M}^+\mathbf{M}'\mathbf{M}L = (\mathbf{M}'\mathbf{M}L)^+(I - \mathbf{M}^+(\tilde{Q}_2^H\mathbf{M}^+)^+\tilde{Q}_2^H)\mathbf{M}^+\mathbf{M}'\mathbf{M}L \quad \text{is Hermitian}.
\]

Remark. In (3.2.1) we could, instead of \(\mathbf{M}^+\), take an arbitrary \((1)\)-inverse \(\mathbf{M}^{(1)}\). Then the condition in (3.2.1) is equivalent to

$$v = \mathbf{M}(h - Lf) + (I - \mathbf{M}^{(1)})z,$$

with \(z\) arbitrary. Under this condition, \(\|v\|\) is minimal for any fixed \(f\) (and varying \(z\)) at \(v = -\mathbf{M}(Lf - h)\) if and only if

$$\mathbf{M}'(I - \mathbf{M}^{(1)})\mathbf{M} = 0,$$

that is, if \(\mathbf{M}^{(1)}\) is also a 3-inverse of \(\mathbf{M}\).

Conclusion. Problem (1.1.1) is equivalent to
"find f and v that solve

\[(3.2.1') \]
\[\min\{\|v\| \mid M^{(1,3)}MLf - M^{(1,3)}v = M^{(1,3)}Mh\}\]

with \(M^{(1,3)}\) an arbitrary \((1,3)\)-inverse of \(M\).

With \(C = M^{(1,3)}ML = Q'R' = Q'R', \) \(Q'\) unitary, \(R' \in LR\), we find

\[(3.2.3')f = (M^{(1,3)}ML)^+(I - M^{(1,3)}(Q_2^HM^{(1,3)} + Q_2^HM^{(1,3)})M^{(1,3)}Mh).\]

Indeed, we have

\[(3.2.4')(M^{(1,3)}ML)^+(I - M^{(1,3)}(Q_2^HM^{(1,3)} + Q_2^HM^{(1,3)})M^{(1,3)}Mh)^+(ML)^+ ,\]

which can be proved in the same way as (3.2.4).

4. Application of Paige's method to problem (1.1.2)

4.1. Now we consider the problem (1.1.2) under the extra condition

\(N(ML) \cap N(K) = \{0\}\). We remark that the vector \(f\) which minimizes (1.1.2), is that solution of (3.1.1) which minimizes \(\|Kf\|\). So we can start from condition (3.1.2),

\[R_1f = Q_1^HMh\]

and reformulate the problem as

"find \(f\) that minimizes \(\|Kf\| \mid R_1f = Q_1^HMh\)."

Let \(R_1 = \tilde{R}P^H = (\tilde{R} \mid 0) \begin{pmatrix} P^H \\ 0 \end{pmatrix} = \tilde{R} \begin{pmatrix} P^H \\ P_2^H \end{pmatrix} \), with \(P\) unitary and \(\tilde{R}\) regular (this is possible since \(R_1 \in LR\)).

Now define \(g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} = \begin{pmatrix} P^H \\ P_2^H \end{pmatrix} f.\)

Then \(Kf = KP_1g_1 + KP_2g_2\) and \(R_1f = \tilde{R}_1g_1 = Q_1^HMh\), which implies \(g_1 = \tilde{R}_1^{-1}Q_1^HMh.\) Then our problem becomes
(4.1.1) "find \( g_2 \) that minimizes \( \| K P_1 R_1^{-1} Q_1 H M h + K P_2 g_2 \| \)."

It is clear from the definition of \( R_1 \) that \( R(ML) = R(P_2) \), hence the condition \( N(ML) \cap N(K) = \{0\} \) is equivalent to \( K P_2 \in RR \).

So problem (4.1.1) has the unique solution

\[
      g_2 = -(K P_2)^+ K P_1 R_1^{-1} Q_1 H M h.
\]

Thus we find under the condition \( N(ML) \cap N(K) = \{0\} \) the unique solution for problem (1.1.2) to be

\[
      f = P_1 g_1 + P_2 g_2 = L_{MK}^+ h,
\]

where

\[
(4.1.2) L_{MK}^+ = (I - P_2 (K P_2)^+ K) P_1 R_1^{-1} Q_1 H M = (I - P_2 (K P_2)^+ K) (M L)^+ M.
\]

Since \( ML = Q_1 R_1 = Q_1 R_1 P_1^H \), we have

\[
      E_0 = (I - (M L)^+ M L) = P_2 P_2^H,
\]

\[
      (KE_0)^+ = (K P_2 P_2^H)^+ = P_2 (K P_2)^+,
\]

which shows that (4.1.2) is another form of the \( M,K \)-weighted pseudo-inverse of \( L \) as defined by Eldén [2, §2].

4.2. The method of §4.1 can also be applied to problem (1.1.2) directly.

Let \( ML = \tilde{L} P_1^H = (\tilde{L}_1 \mid 0) \begin{pmatrix} P_1^H \\ P_2^H \end{pmatrix} \) with \( P \) unitary (so \( R(P_2) = N(ML) \))

and \( \tilde{L}_1 \in RR \) (\( P_1 \) is the same as in §4.1!).

Then define \( P_1^H f =: g = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \)

so \( K f = K P_1 g_1 + K P_2 g_2 \), \( M L f = \tilde{L}_1 g_1 \).
Since \( \tilde{L}_1 \in \mathbb{R}^n \), \( \| \tilde{L}_1 g_1 - Mh \| \) is minimal iff \( g_1 = \tilde{L}_1^* Mh \). Then our problem becomes

\[ (4.2.1) \text{ "find } g_2 \text{ that minimizes } \| KP_1 \tilde{L}_1^* Mh + KP_2 g_2 \|. \]

Again, \( N(ML) \cap N(K) = \{0\} \) implies \( KP_2 \in \mathbb{R}^n \) and problem (4.1.1) has the unique solution

\[ g_2 = - (KP_2)^+ KP_1 \tilde{L}_1^* Mh. \]

Since \( ML = \tilde{L}_1 P_1^H \), \( P_1 \tilde{L}_1^* = (ML)^+ \), we again find, under the condition \( N(ML) \cap N(K) = \{0\} \), that problem (1.1.2) has the unique solution

\[ f = L^+_MK h, \]

where \( L^+_MK \) is given by (4.1.2).

4.3. If \( N(ML) \cap N(K) \neq \{0\} \), then \( L^+_MK h \) (with \( L^+_MK \) defined by (4.1.2)) still solves problem (1.1.2), but it is not the unique solution, since the component of \( f \) in the intersection of the two null-spaces is arbitrary. We shall derive a special property that characterizes \( L^+_MK h \) among all solutions of problem (1.1.2). The first point where non-uniqueness occurs in the derivation of §4.1 is that the solutions of (4.1.1) are given by

\[ g_2 = -(KP_2)^+ KP_1 \tilde{R}_1^Q_1^H Mh + (I - (KP_2)^+ KP_2) z, \]

with \( z \) arbitrary. This leads to the result

\[ f = L^+_MK h + P_2 (I - (KP_2)^+ KP_2) z = L^+_MK h + \tilde{P}_2 z, \]

where \( \tilde{P}_2 := P_2 (I - (KP_2)^+ KP_2) \). Now we remark that

\[ (L^+_MK)^+ P_2 = M^H (ML)^+ H (I - K^H (KP_2)^+ P_2^H) P_2 (I - (KP_2)^+ KP_2). \]
since $R(P)$ = $N(ML)$. Therefore

which leads to the conclusion:

\[ f = \hat{h} \]

is that $f$ that satisfies (1.1.2) and has minimal 2-norm.

5. Generalized Penrose conditions and their solutions

5.1. It follows from §4 that problem (1.1.2) admits of a solution matrix $X$ in the following sense:

\[ (5.1.1) \quad \text{"for each } h, f := Xh \text{ is a solution to problem (1.1.2)."} \]

We shall now characterize $X$ directly. If (for all $h$) $f := Xh$ minimizes

\[ \| M(Lf - h) \| \]

then for all $h$ and all $\delta f$,

\[ \| M(Lf - h) \|^2 \leq \| M(LX - I)h + ML \delta f \|^2 , \]

which is true iff

\[ (5.1.2) \quad (ML)H(ML - I) = 0 . \]

If $X$ satisfies (5.1.2) then $f$ minimizes $\| M(Lf - h) \|$ iff $f = Xh + \delta f$

with $ML \delta f = 0$, hence, iff $f = Xh + (I - (ML)^+ML)z$, $z$ arbitrary.

Consequently, $f = Xh$ is (for all $h$) a solution to problem (1.1.2) iff for all $h$ and all $z$,

\[ \| KXh \|^2 \leq \| KXh + K(I - (ML)^+ML)z \|^2 \]

or, equivalently, iff

\[ (5.1.3) \quad (KX)H(K(I - (ML)^+ML) = 0 . \]
Hence, $X$ is a solution matrix to problem (1.1.2) iff $X$ satisfies (5.1.2) and (5.1.3).

Since in the original problem $L$ only occurs in the combination $ML$, we may assume without loss of generality that $R(L) \cap N(M) = \{0\}$ (or equivalently $N(L) = N(ML)$).

**Lemma.** The conditions (5.1.2) and (5.1.3) are equivalent to (5.1.2) and (5.1.3')

$$\begin{align*}
\text{(5.1.3')} &\quad (XX)^H K(I - XL) = 0 .
\end{align*}$$

**Proof.** i) If (5.1.3') holds, then

$$0 = (XX)^H K(I - XL)(I - (ML)^+ ML) = (XX)^H K(I - (ML)^+ ML),$$

since $N(L) = N(ML)$, i.e. (5.1.3) holds.

ii) If $X$ satisfies (5.1.2), then

$$0 = (ML)^H M(I - LX)L = (ML)^H ML(I - XL),$$

hence also

$$ML(I - XL) = 0 .$$

So if (5.1.3) holds, then

$$0 = (XX)^H K(I - (ML)^+ ML)(I - XL) = (XX)^H K(I - XL),$$

i.e. (5.1.3') holds.

Now we observe that

(5.1.2) is equivalent to \( \begin{cases} 
LXL = L , \\
M^H MLX is Hermitian ; 
\end{cases} \)

(5.1.3') is equivalent to \( \begin{cases} 
K(XLX - X) = 0 , \\
K^H KXL is Hermitian . 
\end{cases} \)

So $\hat{X}$ satisfies (5.1.1) iff $\hat{X}$ satisfies the four Penrose-like conditions:

\( \begin{align*}
(5.1.4) &\quad \begin{cases} 
(1) \quad LXL = L , \\
(2) \quad K(XLX - X) = 0 , \\
(3) \quad M^H MLX is Hermitian , \\
(4) \quad K^H KXL is Hermitian . 
\end{cases} 
\end{align*} \)
5.2. Let us now search for a general solution of the conditions (5.1.4).

First we remark that (5.1.4) is equivalent to (5.1.2) and (5.1.3).

In §3.1 we had $ML = QR = Q_1R_1$ and in §4.1 we had $R_1 = \tilde{R}_1P_1^H = \tilde{R}_1P_1^H$.

Consequently $N(ML) = N(P_1^H) = R(P_2)$ and

$$I - (ML)^+ML = P_2P_2^H.$$ 

If $X_0$ satisfies (5.1.2), then the general solution of (5.1.3) is

$$\tilde{X} = X_0 + P_2Z,$$

with $Z$ arbitrary. Then (5.1.3) becomes

$$P_2P_2^H\tilde{X}^H\tilde{X} = 0,$$

or

$$(KP_2)^H(K(X_0 + P_2z) = 0.$$ 

This implies that

$$Z = -(KP_2)^+KX_0 + (I - (KP_2)^+KP_2)Z',$$

with $Z'$ arbitrary. So the general solution of (5.1.2) and (5.1.3) is

$$\tilde{X} = (I - P_2(KP_2)^+K)(X_0 + P_2Z'),$$

with $Z'$ arbitrary.

We can take $X_0 = (ML)^+M$, which satisfies (5.1.2). Then for the corresponding

$$\tilde{X}_0 = (I - P_2(KP_2)^+K)(ML)^+M$$

we obtain

$$\tilde{X}_0\tilde{X}_0 = (I - P_2(KP_2)^+K)P_1P_1^H\tilde{X}_0 = \tilde{X}_0.$$
which in general is not true for other solutions \( \hat{X} \). Since the general solution \( \hat{X} \) can be written as

\[
\hat{X} = (I - P_2(KP_2)^+K)P_1P_1^HH(ML)^+M + P_2(I - (KP_2)^+KP_2)Z',
\]

and

\[
(P_2(I - (KP_2)^+KP_2))^H(I - P_2(KP_2)^+K)P_1 = 0,
\]

we see that the special \( \hat{X}_0 \) is the solution of (5.1.5) with the smallest F-norm. Regarding uniqueness, \( \hat{X} \) is independent of \( Z' \) and therefore unique iff

\[
P_2(I - (KP_2)^+KP_2) = 0 \quad \text{or equivalently} \quad N(KP_2^2) = N(P_2) = \{0\}.
\]

The latter is equivalent to \( R(P_2) \cap N(K) = \{0\} \), so to \( N(ML) \cap N(K) = \{0\} \).

**Conclusion.** If \( N(ML) = N(L) \) and \( N(L) \cap N(K) = \{0\} \), then \( \hat{X} = L_{MK}^+ \) as defined by (4.1.2) is the unique matrix satisfying the conditions (5.1.5).

**Remarks.** 1. Without the assumption \( R(L) \cap N(M) = \{0\} \), we can maintain §5.1 if we replace \( L \) by \( M^{(1)}ML \) in (5.1.3'), with \( M^{(1)} \) an arbitrary (1)-inverse of \( M \). Then

\[
(5.1.2) \text{ is equivalent to } \begin{cases} M(LXL - L) = 0, \\
M^{H}MLX \text{ is Hermitian}; 
\end{cases}
\]

and

\[
(5.1.3') \text{ is equivalent to } \begin{cases} K(XM^{(1)}MLX - X) = 0, \\
K^{H}KX^{(1)}ML \text{ is Hermitian}; 
\end{cases}
\]

and (5.1.5) becomes

\[
(1) M(LXL - L) = 0, \quad (3) M^{H}MLX \text{ is Hermitian},
\]

\[
(2) K(XM^{(1)}MLX - X) = 0, \quad (4) K^{H}KX^{(1)}ML \text{ is Hermitian}.
\]

2. Instead of looking for a solution matrix \( X \) for problem (1.1.2), we may also consider the problem

"find \( X \) that minimizes \( \|KX\|_F \mid \|M(I - LX)\|_F \text{ is minimal}"."
It is easily shown that $X$ solves this problem iff $X$ satisfies (5.1.2) and (5.1.3).

5.3. Ben-Israel and Greville [3, Sec. 3.3] find the same conditions (5.1.5) for the solution matrix of problem (1.1.2), although they restrict themselves to the case that $K^H K$ and $M^H M$ are regular, so $K$ and $M \in \mathbb{R}^r$, which implies $N(L) = N(ML)$ and $N(ML) \cap N(K) = \{0\}$. To construct the solution, they take the Cholesky factorisation of $K^H K$ and $M^H M$,

$$
K^H K = R_K^H R_K, \quad M^H M = R_M^H R_M,
$$

with $R_K$ and $R_M$ both regular matrices, and then find

$$
L^+_{MK} = R_K^{-1} (R_M L R_K^{-1})^+ R_M,
$$

see [3, Ex. 3.39]. By observing that

$$
K = (U_1 \mid U_2) \begin{pmatrix} R_K & 0 \\ 0 & 0 \end{pmatrix}, \quad M = (V_1 \mid V_2) \begin{pmatrix} R_M & 0 \\ 0 & 0 \end{pmatrix},
$$

with $U = (U_1 \mid U_2)$ and $V = (V_1 \mid V_2)$ both unitary, we can write

$$
L^+_{MK} = R_K^{-1} U_1^H (R_M L R_K^{-1})^+ V_1^H V_1 R_M
$$

$$
= R_K^{-1} U_1^H R_M^H R_K^{-1} U_1^H V_1^H V_1 R_M = K^+ (MLK^+)^+ M.
$$

It is easy to verify that under the condition $K$ and $M \in \mathbb{R}^r$,

$$
L^+_{IK} = K^+ (LK^+)^+ = (I - P_2(KP_2)^+ K)L^+
$$

with $P_2$ left-unitary and $R(P_2) = N(L)$, and

$$
L^+_{ML} = (ML)^+ M = L^+(I - M^+ (Q_2^H M^+ Q_2^H)) M.
$$
with $Q_2$ left-unitary and $R(Q_2) = N(L^H)$. Since under the same condition (even if only $N(ML) = N(L)$) $L^+_{MK} = L^+_{IK}L^+_{MI}$, we can give four alternative expressions for $L^+_{MK}$:

(5.3.1) $L^+_{MK} = K^+(MLK)^+M$ (cf. [3, Ex. 3.39])

(5.3.2) $= (I - P_2(KP_2)^+K)(ML)^+M$ (cf. [2])

(5.3.3) $= K^+(LK^+)(I - M^+(Q_2^HM^+)^+Q_2^H)$

(5.3.4) $= (I - P_2(KP_2)^+K)L^+(I - M^+(Q_2^HM^+)^+Q_2^H)$,

where $P_2$ and $Q_2$ are left-unitary matrices with $R(P_2) = N(L)$ and $R(Q_2) = N(L^H)$. We now ask whether the expressions (5.3.1) - (5.3.4) still satisfy our Penrose conditions under less stringent conditions than KERR and MERR. We already saw that this is true for (5.3.2) under the only condition $N(ML) = N(L)$. It is easily found that the Penrose conditions are satisfied by (5.3.1) and (5.3.3) if $K \in \text{RR}$, and by (5.3.4) if $M \in \text{RR}$.

We can settle the difficulty $M \notin \text{RR}$ (but still $K \in \text{RR}$), by not considering $Lf - h$, but $M^+(MLf - h)$ instead. So by analogy we obtain the solution matrix

$$L^+_{MK} = (M^+ML)^+_{IK}M^+ML(M^+ML)^+_{MI}M^+M = (M^+ML)^+_{IK}M^+MLL^+_{MI}. $$

Thus (5.3.4) becomes

(5.3.4')$L^+_{MK} = (I - P_2(KP_2)^+K)(M^+ML)^+(I - N^+(Q_2^HM^+)^+Q_2^H),$

where now $Q_2$ and $P_2$ are left-unitary matrices, with $R(P_2) = N(M^+ML)$ and $R(Q_2) = N((M^+ML)^H)$. With these $P_2$ and $Q_2$, expressions (5.3.1) - (5.3.3) remain the same.

If also $K \notin \text{RR}$, we have a greater problem in finding alternative expressions.
If we restrict ourselves to $f \in R(K^+) = (N(K))^\perp$, then there is no problem, since the restriction of $K$ to $R(K^+)$ is injective. Then we define $f := K^+Kz$, and solve the problem

"find $z$ that minimizes \{ norm of $Kz$ \mid norm of $MLK^+Kz - Mh$ is minimal \} ."

The solution of this problem is

$$z = (LK^+K)^+_{MK} h + Sy ,$$

with $R(S) = N(K)$ and $y$ arbitrary. So $f = K^+Kz = K^+(LK^+K)^+_{MK} h$ is unique. As before we have four expressions for the solution matrix:

\begin{align}
(5.3.5) & \quad K^+(LK^+K)^+_{MK} = K^+(MLK^+)^+_{MK} M \\
(5.3.6) & \quad = K^+(I - KP_2(KP_2)^+)K(MLK^+)^+_{MK} M \\
(5.3.7) & \quad = K^+(M^+MLK^+)^+(I - M^+(Q_2^H M^+)^+Q_2^H)M^+M \\
(5.3.8) & \quad = K^+(I - KP_2(KP_2)^+)K(M^+MLK^+)^+M^+(I - (Q_2^H M^+)^+Q_2^H M^+)M ,
\end{align}

where $P_2$ and $Q_2$ are left-unitary matrices, with $R(P_2) = N(M^+MLK^+)K$ and $R(Q_2) = N((M^+MLK^+)^H)$.

Without the restriction $f \in R(K^+)$, we were not able to find a generalization for the expressions (5.3.1) and (5.3.3).

6. Optimization theory

6.1. We now want to solve the problem (1.1.2) using some optimization theory. A general form for this problem is

"find $x$ that minimizes \{ $f(x) \mid g(x) = b$ \},"

where $f$ and $g$ are sufficiently smooth functions. The theory of Lagrange multipliers states that a solution $x$ of this problem corresponds with the $x$-coordinates of a saddle point of the Lagrange functional
We can find the saddle points from

\[ \nabla_x L = 0 : Df(x) + \nabla_x g(x) = 0 , \]
\[ \nabla_z L = 0 : g(x) = b , \]

and so we have to solve \( x \) from this system.

Now let us return to our problem. We saw in §3 that, using (3.1.2), problem (1.1.2) can be formulated as

(6.1.1) "find \( f \) that minimizes \( \|Kf\|^2 \) is minimal and \( MLf + v = Mh \)."

Let us first consider the "inner" problem

\[ \min \{ \|v\|^2 \mid MLf + v = Mh \} , \]

which is equivalent to

(6.1.2) \[ \min \{ \| (0 \mid I) \left( \begin{array}{c} f \\ v \end{array} \right) \|^2 \mid (ML \mid I) \left( \begin{array}{c} f \\ v \end{array} \right) = Mh \}. \]

We define \( T := (0 \mid I) \), \( g := \left( \begin{array}{c} f \\ v \end{array} \right) \), \( H := (ML \mid I) \), \( b := Mh \), then problem (6.1.2) becomes

(6.1.3) \[ \min \{ \| Tg \|^2 \mid Hg = b \}. \]

The Lagrange functional corresponding to (6.1.3) is

(6.1.4) \[ L(g, z) = \nabla_t^H T^H g - z^H(Hg - b) \]

and we find the saddle points from

\[ \nabla_g L = 0 : g = T^H H = 0; \quad (ML)^H z = 0, z = v ; \]
\[ \nabla_z L = 0 : Hg = b; \quad MLf + v = Mh . \]

So the solution \( f \) of the system
(6.1.5) \[
\begin{pmatrix}
I & ML \\
(ML)^H & 0
\end{pmatrix}
\begin{pmatrix}
z \\
f
\end{pmatrix} =
\begin{pmatrix}
Mh \\
0
\end{pmatrix}
\]
also solves problem (6.1.2) and \( v = z \).

Returning to our original problem (6.1.1), we define
\[
p := \begin{pmatrix} z \\ f \end{pmatrix},
c := \begin{pmatrix} I & ML \\
(ML)^H & 0
\end{pmatrix},
S = (0 | K),
k := \begin{pmatrix} Mh \\
0
\end{pmatrix},
\]
then (6.1.1) can be shortly written as
\[
(6.1.6) \min \|Sp\| \mid Cp = k.
\]
The Lagrange functional is now
\[
L(p,r) = \frac{1}{2} p^H S^H Sp + r^H (Cp - k),
\]
and we find the saddle points from
\[
\nabla_p L = 0 : p^H S^H S + r^H C = 0; \quad S^H Sp + C^H r = 0;
\]
\[
\nabla_r L = 0 : Cp = k.
\]

Then the solution \( p \) of the system
\[
\begin{pmatrix}
S^H S & C^H \\
C & 0
\end{pmatrix}
\begin{pmatrix}
p \\
v
\end{pmatrix} =
\begin{pmatrix}
0 \\
k
\end{pmatrix}
\]
also solves (6.1.6). After substitution of the expressions for \( S, C, p \) and \( k \), this system becomes
\[
\begin{pmatrix}
0 & 0 & I & ML \\
0 & K & (ML)^H & 0 \\
I & ML & 0 & 0 \\
(ML)^H & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
z \\
f \\
r_1 \\
r_2
\end{pmatrix} =
\begin{pmatrix}
0 \\
0 \\
Mh \\
0
\end{pmatrix}
\]
where the vector \((z^H, r_1^H, r_2^H)^H\) is a vector of Lagrange multipliers.

Now let \( U_2 \) be left-unitary and \( R(U_2^H) = N(C) \), then it is easy to verify (cf. Eldén [2, corr. 3.3]) that
In the same way as before, we find

\( (6.1.7) \) \[ \begin{pmatrix} \mathbf{S}^H & \mathbf{C}^H \\ \mathbf{C} & 0 \end{pmatrix} = \begin{pmatrix} \mathbf{U}_2((\mathbf{S}\mathbf{U}_2)^\dagger \mathbf{S}\mathbf{U}_2)^\dagger \mathbf{U}_2^H & \mathbf{C}_\mathbf{IS}^+ \\ \mathbf{C}_\mathbf{IS}^+ \mathbf{S}^H & -\mathbf{S}^H \mathbf{C}_\mathbf{IS}^+ \end{pmatrix}, \]

where \( \mathbf{C}_\mathbf{IS}^+ = (\mathbf{I} - \mathbf{U}_2(\mathbf{S}\mathbf{U}_2)^\dagger \mathbf{S})\mathbf{C}^+ \).

In the same way as before, we find

\( (6.1.8) \) \[ \mathbf{C}^+ = \begin{pmatrix} \mathbf{I} & \mathbf{M} \mathbf{L} \\ \mathbf{M} \mathbf{L}^H & 0 \end{pmatrix}^+ = \begin{pmatrix} \mathbf{I} - (\mathbf{M} \mathbf{L})(\mathbf{M} \mathbf{L})^+ & (\mathbf{M} \mathbf{L})^\dagger \\ (\mathbf{M} \mathbf{L})^\dagger & -\mathbf{M} \mathbf{L} \end{pmatrix}. \]

We know that \( \mathbf{U}_2\mathbf{U}_2^H = \mathbf{I} - \mathbf{C}^+ \mathbf{C} = \mathbf{I} - \mathbf{C} \mathbf{C}^+ \) (since \( \mathbf{C} \) is Hermitian), and so

\[ \mathbf{U}_2\mathbf{U}_2^H = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{I} - (\mathbf{M} \mathbf{L})^\dagger \mathbf{M} \mathbf{L} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{P}^\perp_2 \end{pmatrix}, \]

with \( \mathbf{P}^\perp_2 \) a left-unitary matrix such that \( \mathbf{R}^\perp(\mathbf{P}^\perp_2) = \mathbf{N}(\mathbf{M} \mathbf{L}) \).

So, we can take \( \mathbf{U}_2 = \begin{pmatrix} \mathbf{0} \\ \mathbf{P}^\perp_2 \end{pmatrix} \).

By combining the previous results we find for \( \begin{pmatrix} \mathbf{S}^H & \mathbf{C}^H \\ \mathbf{C} & 0 \end{pmatrix}^+ \) the formula displayed in fig. 1.

Consequently, the solution of problem (6.1.1) (and so of (1.1.2)) with minimum 2-norm is again found to be

\[ \mathbf{f} = (\mathbf{I} - \mathbf{P}^\perp_2(\mathbf{K} \mathbf{P}^\perp_2)^\dagger \mathbf{K})(\mathbf{M} \mathbf{L})^\dagger \mathbf{M} \mathbf{h}, \]

where \( \mathbf{P}^\perp_2 \) is left-unitary and \( \mathbf{R}^\perp(\mathbf{P}^\perp_2) = \mathbf{N}(\mathbf{M} \mathbf{L}) \).
\[
\begin{pmatrix}
S_{\text{H}_s} & C_{\text{H}}^+ \\
C & 0
\end{pmatrix} = \\
\begin{pmatrix}
0 & 0 & (C^+)^{11} & (C^+)^{12} \\
0 & P_2((\text{KP}_2)^+)^H & (I - P_2(\text{KP}_2)^+)^K(C^+)^{21} & (I - P_2(\text{KP}_2)^+)^K(C^+)^{22} \\
(C^+)^{11} & (C^+)^{12} & (I - P_2(\text{KP}_2)^+)^K(C^+)^{21} & -(C^+)^{12}K(I - KP_2(\text{KP}_2)^+K(C^+)^{21} \\
(C^+)^{21} & (C^+)^{22} & (I - P_2(\text{KP}_2)^+)^K(C^+)^{21} & -(C^+)^{22}K(I - KP_2(\text{KP}_2)^+K(C^+)^{22}
\end{pmatrix}
\]

with \( C^+ = \begin{pmatrix} I & \text{ML}^+ \\ \text{ML}^H & 0 \end{pmatrix} = \begin{pmatrix} (I - \text{ML}(\text{ML})^+) & (\text{ML})^{+H} \\ \text{ML}^+ & -(\text{ML})^{+H} \end{pmatrix} \)
6.2. The method of Lagrange multipliers can also be applied directly to problem (1.1.2),
"find \( f \) that minimizes \( \| Kf \| \), \( \| M(Lf - h) \| \) is minimal."
The "inner" problem is easy to solve, since
\[ \| M(Lf - h) \| \text{ is minimal iff } MLf = ML(ML)^+ Mh. \]
Hence, problem (1.1.2) becomes
(6.2.1) "find \( f \) that minimizes \( \| Kf \|^2 \text{ iff } MLf = ML(ML)^+ Mh.\)"

The corresponding Lagrange functional is
\[ L(f,r) = \frac{1}{2} f^H K^H Kf + r^H (MLf - ML(ML)^+ Mh). \]

Then the solution \( f \) of the system
\[
\begin{pmatrix}
K^H K & (ML)^H \\
ML & 0
\end{pmatrix}
\begin{pmatrix}
f \\
v
\end{pmatrix}
= \begin{pmatrix}
0 \\
ML(ML)^+ Mh
\end{pmatrix}
\]
satisfies problem (6.2.1).

Similar to (6.1.7) it is easy to verify that
\[
\begin{pmatrix}
K^H K & (ML)^H \\
ML & 0
\end{pmatrix}^+
= \begin{pmatrix}
-P_2((KP_2)^H K P_2)^+ P_2 & (ML)^+_{IK} \\
((ML)^+_{IK})^H & (K(ML)^+_{IK})^H K(ML)^+_{IK}
\end{pmatrix},
\]
where \( (ML)^+_{IK} = (I - P_2(KP_2)^+ K)(ML)^+ \) and \( P_2 \) is left-unitary with
\( R(P_2) = N(ML). \) So again we find for the solution of problem (1.1.2)
with minimum 2-norm,
\[
f = (ML)^+_{IK} ML(ML)^+ Mh = (I - P_2(KP_2)^+ K)(ML)^+ Mh.
\]

Appendix. Equivalence of problems (2.1.1) and (2.1.2)
1. We start from problem (2.1.1), i.e.
"estimate \( x \) from a realisation of \( y = Cx + \omega. \)"
Here \( C \) is a known matrix; \( \omega \) is a stochastic with \( \mathbb{E}(\omega) = 0, \mathbb{E}(\omega \omega^H) = \sigma^2 \omega, \)
where \( W \) is a nonnegative definite Hermitian matrix. As in §2 we set \( W = BB^H \) with \( B \in \mathbb{R}^R \), and \( w = Bv \), then \( v \) is a stochastic with \( C(v) = 0 \), \( C(vv^H) = \sigma^2 I \).

**Definition.** A linear function \( \varphi : x \rightarrow p^H x \) is called **estimable** if there is a linear function \( \psi : y \rightarrow q^H y \) such that the stochastic vector \( y = Cx + Bv \) satisfies \( C(q^Hy) = p^H x \).

Then \( \psi(y) \) is called **linear unbiased estimator** (LUE) of \( \varphi(x) \).

Since for arbitrary \( q \) we have \( C(q^Hy) = q^HCx \), \( \varphi \) is estimable iff \( p \in \text{R}(C^H) \), and \( \psi(y) = q^Hy \) is a LUE of \( \varphi(x) = p^H x \) iff \( C^HQ = p \). For any \( q \) satisfying the latter condition we have

\[
C((q^Hy - p^H x)^2) = C((q^HBv)^2) = \sigma^2 q^H BB^H q = \sigma^2 q^H Wq.
\]

**Definition.** \( \hat{\varphi} : y \rightarrow q^Hy \) is called **best linear unbiased estimator** (BLUE) of \( \varphi(x) = p^H x \) if

\[
q^H Wq = \min \{ q^H Wq \mid C^HQ = p \}.
\]

Using a Lagrange multiplier \( \tau \), we find that \( \hat{\varphi} \) is BLUE of \( \varphi \) iff \( q \) and \( \tau \) satisfy

\[
(A.1.1) \quad \begin{pmatrix} W & C \\ C^H & 0 \end{pmatrix} \begin{pmatrix} q \\ \tau \end{pmatrix} = \begin{pmatrix} 0 \\ p \end{pmatrix}.
\]

**Remark.** It is clear from (A.1.1) that \( Wq \in \text{R}(C) \) and the system is compatible iff \( p \in \text{R}(C^H) \).

**Definition.** An observation \( y \) is called **compatible** iff \( y \in \text{R}(B|C) \), so iff \( y \in \text{R}(W|C) \).
Lemma. Consider the matrix $A = \begin{pmatrix} W & C \\ C^H & 0 \end{pmatrix}$, where $W = BB^H$.

Then $\begin{pmatrix} r \\ x \end{pmatrix} \in N(A)$ iff $r \in N \begin{pmatrix} B^H \\ C^H \end{pmatrix} \land x \in N(C)$ ,

and $\begin{pmatrix} y \\ z \end{pmatrix} \in R(A)$ iff $y \in R((B|C)) \land z \in R(C^H)$.

Proof. $\begin{pmatrix} W & C \\ C^H & 0 \end{pmatrix} \begin{pmatrix} r \\ x \end{pmatrix} = 0$ is equivalent to $\begin{pmatrix} BB^H r + Cx = 0 \\ r^H C = 0 \end{pmatrix}$.

This implies

$$0 = r^H C x = -r^H B B^H r,$$

so $B^H r = 0$, and consequently $C x = 0$.

Therefore

$$\begin{pmatrix} r \\ x \end{pmatrix} \in N(A) \iff [r \in N \begin{pmatrix} B^H \\ C^H \end{pmatrix} \land [x \in N(C) = (R(C^H))],$$

which implies that

$$\begin{pmatrix} y \\ z \end{pmatrix} \in R(A) \iff y \in R((B|C)) \land z \in R(C^H).$$

The lemma implies that the system

$$(A.1.2) \begin{pmatrix} W & C \\ C^H & 0 \end{pmatrix} \begin{pmatrix} \hat{p} \\ \hat{x} \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix}$$

has a solution iff $y$ is compatible. If $y$ is compatible, $\hat{x}$ is to be determined from $(A.1.2)$ modulo $N(C)$.

Now let $y$ be compatible and let $\phi : x \rightarrow p^H x$ be estimable, then $\hat{\phi}$, $\hat{p}$ and $\hat{\mathcal{R}}$ are determined as above, and we have

$$\hat{\phi}(y) = \hat{\phi}^H y = \hat{\phi}^H (W \hat{p} + C \hat{x}) = -\hat{p}^H C^T \hat{x} + p^H \hat{x} = p^H \hat{x} = \varphi(\mathcal{R}).$$

So for a given compatible $y$ and with $\hat{x}$ as above, the BLUE for any estimable function $\phi : x \rightarrow p^H x$, is $p^H \hat{x}$. In particular, $C x$ is estimable and its BLUE
is $\mathcal{Y} := Cx$.

2. Now we start from problem (2.1.2), i.e.

"find to a given $y \in \mathbb{R}((C|B))$ vectors $x$ and $v$ that minimize $\|v\|
under the condition $y = Cx + Bv$".

By definition $y$ is compatible. Using a Lagrange multiplier $r$, we find that
$v, r$ and $x$ have to satisfy the system

$$
\begin{bmatrix}
-I & B^H & 0 \\
B & 0 & C \\
0 & C^H & 0
\end{bmatrix}
\begin{bmatrix}
v \\
r \\
x
\end{bmatrix} =
\begin{bmatrix}
0 \\
y \\
0
\end{bmatrix}.
$$

(A.2.1)

This system is equivalent to

$$
\begin{bmatrix}
W & C \\
C^H & 0
\end{bmatrix}
\begin{bmatrix}
r \\
x
\end{bmatrix} =
\begin{bmatrix}
y \\
0
\end{bmatrix},
v = B^H r.
$$

(A.2.2)

So the solutions $x$ and $v$ of problem (2.1.2) are $x = \mathcal{X}$ and $v = B^H \mathcal{P}$, with
$\mathcal{X}$ and $\mathcal{P}$ as determined in the first section of this appendix.

**Conclusion.** Since both problems ((2.1.1) and (2.1.2)) lead to the same
solution set for $x$, the equivalence of both is evident.

For a more professional look into these matters, see e.g. [4, Ch. 3].

**Acknowledgement.** The first draft of this paper has been written as a student's
paper under the guidance of Prof. G.W. Veltkamp. I am grateful to him for
leaving me some notes on the subject which were the start of this work, for
the stimulating discussions and finally, for his comments on the manuscript.

**References.**

