Chapter 5

Positive Semidefinite (and Definite) $m$-Symmetric Matrices using Schur Complement in Minkowski Space $\mathbb{M}$

5.1 Introduction

The idea of Schur complement goes back to Sylvester in 1851. The term Schur complement was introduced by E.Haynsworth [41]. Carlson et al. [15] defined the generalized Schur complement and many authors studied their applications in statistics, matrix theory, electrical network theory, discrete-time regulator problem, sophisticated techniques and some other fields. For interesting results concerning with Schur complements one may refer [2], [41], [15]. Ando [2] gave an idea about the generalized Schur complements in 1979. Banachiewicz expressed the inverse of a partitioned matrix in terms of Schur complement.
When the partitioned matrix $A$, given by $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, where $A_{11} \in C^{m \times n}$ and $A_{22} \in C^{p \times q}$ are nonsingular, then $S$ is nonsingular and

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} + A_{12}^{-1}A_{12}S^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}S^{-1} \\ -S^{-1}A_{21}A_{11}^{-1} & S^{-1} \end{bmatrix},$$

where $S$ denotes the generalized Schur complement of $A$ which is defined by

$$S = A_{22} - A_{21}A_{11}^{-1}A_{12}, \quad \text{where} \quad A_{11}^{-1} \in A_{11}\{1\}.$$

Suppose that the square matrix $M$ written as $2 \times 2$ block matrix

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix},$$

(5.1)

where $A$ is a $s \times s$ matrix and $D$ is a $t \times t$ matrix, with $n = s + t$, where $A$ and $D$ are square matrices, but $B$ and $C$ are not square unless $n = m$. Entries are generally assumed to be complex. If $A$ is square and nonsingular, then $M$ can be decomposed as

$$M = \begin{bmatrix} I_m & 0 \\ CA^{-1} & I_t \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{bmatrix} \begin{bmatrix} I_m & A^{-1}B \\ 0 & I_t \end{bmatrix}.$$ (5.2)

This decomposition is often called Aitken block-diagonalization formula in the literature, see Puntanen and Styan [72]. Moreover, if both $M$ and $A$ are nonsingular, then the Schur complement $S = D - CA^{-1}B$ is nonsingular too, and the inverse of $M$ can be written in the following form

$$M^{-1} = \begin{bmatrix} I_m & -A^{-1}B \\ 0 & I_t \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ 0 & S^{-1} \end{bmatrix} \begin{bmatrix} I_m & 0 \\ -CA^{-1} & I_t \end{bmatrix} = \begin{bmatrix} A^{-1} + A^{-1}BS^{-1}CA^{-1} & -A^{-1}BS^{-1} \\ -S^{-1}CA^{-1} & S^{-1} \end{bmatrix}. $$ (5.3)

This well-known formula is called the Banachiewicz inversion formula for the inverse of a nonsingular matrix in the literature, see Puntanen and
Styan [72], and can be found in most linear algebra books. The two formulas in (5.2) and (5.3) and their consequences are widely used in manipulating partitioned matrices and their operations. When both $A$ and $M$ in (5.1) are singular, the two formulas in (5.2) and (5.3) can be extended to generalized inverses of matrices. Here we have made a similar study by using Minkowski inverses. Section 5.3 and 5.4 gives a brief introduction of the partitioned matrix and the block diagonalization of the Schur complement in Minkowski space. Section 5.5 and 5.6 gives a complete characterization of the partitioned inverses and Minkowski inverses of a positive semidefinite (and Definite) m-symmetric matrices using Schur complement in Minkowski space $\mathcal{M}$.

### 5.2 Schur Complement

If $A$ is nonsingular, the Schur complement of $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ with respect to $A$ is defined as

$$M/A = D - CA^{-1}B.$$  \hfill (5.4)

If $D$ is nonsingular, the Schur complement of $M$ with respect to $D$ is defined as

$$M/D = A - BD^{-1}C.$$  \hfill (5.5)

Matrices (5.4) and (5.5) are called the Schur complement of $A$ in $M$ and the Schur complement of $D$ in $M$ respectively.
5.3 Partitioned Matrix in Minkowski Space

Let \( M \sim \) be an \( n \times n \) block matrix written as
\[
M \sim = \begin{bmatrix}
A^- & -C^-G_1 \\
-G_1B^- & D^*
\end{bmatrix},
\]
where \( A^- \) is a \( s \times s \) matrix and \( D^* \) is a \( t \times t \) matrix with \( n = s + t \). (so \(-C^-G_1 \) is a \( s \times t \) matrix and \(-G_1B^- \) is a \( t \times s \) matrix.) We can try to solve the linear system
\[
\begin{bmatrix}
A^- & -C^-G_1 \\
-G_1B^- & D^*
\end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} p \\ q \end{bmatrix},
\]
That is
\[
A^-x - C^-G_1y = p, \quad (5.6)
\]
\[
-G_1B^-x + D^*y = q. \quad (5.7)
\]
Assuming that \( D^* \) is invertible, then we first solve for \( y \) getting
\[
y = (D^*)^{-1}(q + G_1B^-x)
\]
substituting \( y \) in equation (5.6) we get,
\[
A^-x - C^-G_1(D^*)^{-1}(q + G_1B^-x) = p
\]
\[
A^-x - C^-G_1(D^*)^{-1}q - C^-G_1(D^*)^{-1}G_1B^-x = p
\]
\[
(A^- - C^-G_1(D^*)^{-1}G_1B^-)x - C^-G_1(D^*)^{-1}q = p
\]
\[
(A^- - C^-G_1(D^*)^{-1}G_1B^-)x = p + C^-G_1(D^*)^{-1}q
\]
\[
x = (A^- - C^-G_1(D^*)^{-1}G_1B^-)^{-1}(p + C^-G_1(D^*)^{-1}q)
\]
\[
x = (A^- - C^-G_1(D^*)^{-1}B^-)^{-1}(p + C^-G_1(D^*)^{-1}q)
\]
\[
x = (A^- - C^-D^-)^{-1}B^-)^{-1}(p + C^-G_1(D^*)^{-1}q)
\]
Now substitute \( x \) value in (5.6) we get

\[
A^{-1}(A^{-1} - C^{-1}(D^{-1}B^{-1})^{-1}(p + C^{-1}G_1(D^{-1})^{-1}q) - C^{-1}G_1(D^{-1})^{-1}(q + G_1B^{-1}x) = p
\]

\[
y = (D^{-1})^{-1}(q + G_1B^{-1}(A^{-1} - C^{-1}(D^{-1}B^{-1})^{-1}(p + C^{-1}G_1(D^{-1})^{-1}q))
\]

The matrix equation \( A^{-1} - C^{-1}(D^{-1}B^{-1})^{-1} \) is the Schur complement of \( D^{-1} \) in \( M^{-1} \). If \( A^{-1} \) is non-singular, the Schur complement of \( M^{-1} \) with respect to \( A^{-1} \) is defined as

\[
M^{-1}/A^{-1} = D^{-1} - G_1B^{-1}(A^{-1})^{-1}C^{-1}G_1
\]  \( (5.8) \)

If \( D^{-1} \) is non-singular, then the Schur complement of \( M^{-1} \) with respect to \( D^{-1} \) is defined as

\[
M^{-1}/D^{-1} = A^{-1} - C^{-1}(D^{-1}B^{-1})^{-1}
\]  \( (5.9) \)

Matrices (5.8) and (5.9) are called the Schur complement of \( A^{-1} \) in \( M^{-1} \) and the Schur complement of \( D^{-1} \) in \( M^{-1} \).

The equations written as

\[
x = (A^{-1} - C^{-1}(D^{-1}B^{-1})^{-1}(p + C^{-1}G_1(D^{-1})^{-1}q
\]

\[
y = (D^{-1})^{-1}(q + G_1B^{-1}(A^{-1} - C^{-1}(D^{-1}B^{-1})^{-1}(p + C^{-1}G_1(D^{-1})^{-1}q))
\]

yield a formula for the inverse of \( M^{-1} \) in terms of the Schur complement of \( D^{-1} \) in \( M^{-1} \), namely

\[
\begin{bmatrix}
A^{-1} & -C^{-1}G_1 \\
-G_1B^{-1} & D^{-1}
\end{bmatrix}^{-1} = 
\begin{bmatrix}
(A^{-1} - C^{-1}(D^{-1}B^{-1})^{-1})^{-1} & (A^{-1} - C^{-1}(D^{-1}B^{-1})^{-1}C^{-1}G_1(D^{-1})^{-1})^{-1} \\
(D^{-1})^{-1}G_1B(A^{-1} - C^{-1}(D^{-1}B^{-1})^{-1}C^{-1}G_1(D^{-1})^{-1}) & (D^{-1})^{-1} + (D^{-1})^{-1}G_1B(A^{-1} - C^{-1}(D^{-1}B^{-1})^{-1}C^{-1}G_1(D^{-1})^{-1})^{-1}
\end{bmatrix}
\]

That is

\[
\begin{bmatrix}
A^{-1} & -C^{-1}G_1 \\
-G_1B^{-1} & D^{-1}
\end{bmatrix}^{-1} = \frac{1}{(D^{-1})^{-1}G_1B} \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix}
(A^{-1} - C^{-1}(D^{-1}B^{-1})^{-1})^{-1} & 0 \\
0 & (D^{-1})^{-1}
\end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}
\]

63
It follows immediately that
\[
\begin{bmatrix}
A^- & -C^-G_1 \\
-G_1B^- & D^-
\end{bmatrix} =
\begin{bmatrix}
I & -C^-G_1(D)^{-1} \\
0 & I
\end{bmatrix}
\begin{bmatrix}
A^- -C^-((D)^{-1}B^- & 0 \\
0 & D^-
\end{bmatrix}
=\begin{bmatrix}
I & 0 \\
-(D)^{-1}G_1B^- & I
\end{bmatrix}.
\]

### 5.4 Block Diagonalization of a Schur Complement in Minkowski Space \( \mathcal{M} \)

The following block diagonalization clearly display the Schur complement role in Minkowski space. If \( A^- \) is non-singular, then
\[
\begin{bmatrix}
I & 0 \\
G_1B^-(A^-)^{-1} & I
\end{bmatrix}
\begin{bmatrix}
A^- & -C^-G_1 \\
-G_1B^- & D^-
\end{bmatrix}
\begin{bmatrix}
I & (A^-)^{-1}C^-G_1 \\
0 & I
\end{bmatrix} =
\begin{bmatrix}
A^- & 0 \\
0 & M^-/A^-
\end{bmatrix}.
\]

That is
\[
M^- =\begin{bmatrix}
A^- & -C^-G_1 \\
-G_1B^- & D^-
\end{bmatrix} =
\begin{bmatrix}
I & 0 \\
G_1B^-(A^-)^{-1} & I
\end{bmatrix}
\begin{bmatrix}
A^- & 0 \\
0 & M^-/A^-
\end{bmatrix} \begin{bmatrix}
I & (A^-)^{-1}C^-G_1 \\
0 & I
\end{bmatrix},
\]

where as if \( D^* \) is non-singular,
\[
\begin{bmatrix}
I & -C^-G_1(D^*)^{-1} \\
0 & I
\end{bmatrix}
\begin{bmatrix}
M^-/D^* & 0 \\
0 & D^*
\end{bmatrix}
\begin{bmatrix}
I & 0 \\
(D^*)^{-1}G_1B^- & I
\end{bmatrix} =
\begin{bmatrix}
M^-/D^* & 0 \\
0 & D^*
\end{bmatrix}.
\]

That is
\[
M^- =\begin{bmatrix}
A^- & -C^-G_1 \\
-G_1B^- & D^-
\end{bmatrix} =
\begin{bmatrix}
I & 0 \\
0 & I
\end{bmatrix}
\begin{bmatrix}
M^-/D^* & 0 \\
0 & D^*
\end{bmatrix} \begin{bmatrix}
I & 0 \\
-(D^*)^{-1}G_1B^- & I
\end{bmatrix}.
\]

All of these can be verified by matrix multiplication.

### 5.5 Partitioned Inversion

If \( A^- \) is non-singular, then
\[
(M^-)^{-1} = \begin{bmatrix}
A^- & -C^-G_1 \\
-G_1B^- & D^-
\end{bmatrix}^{-1},
\]
\[
\begin{bmatrix}
I & (A^\sim)^{-1}C^{-1}G_1 \\
0 & I
\end{bmatrix}
\begin{bmatrix}
(A^\sim)^{-1} & 0 \\
0 & (M^\sim/A^\sim)^{-1}
\end{bmatrix}
\begin{bmatrix}
I & 0 \\
(G_1B^\sim)(A^\sim)^{-1} & I
\end{bmatrix},
\]

\[
= \begin{bmatrix}
(A^\sim)^{-1} & (A^\sim)^{-1}C^{-1}G_1(M^\sim/A^\sim)^{-1} \\
0 & (M^\sim/A^\sim)^{-1}
\end{bmatrix}
\begin{bmatrix}
I & 0 \\
(G_1B^\sim)(A^\sim)^{-1} & I
\end{bmatrix},
\]

\[
= \begin{bmatrix}
(A^\sim)^{-1} + (A^\sim)^{-1}C^{-1}G_1(M^\sim/A^\sim)^{-1}(G_1B^\sim)(A^\sim)^{-1} & (A^\sim)^{-1}C^{-1}G_1(M^\sim/A^\sim)^{-1} \\
(M^\sim/A^\sim)^{-1}G_1B^\sim(A^\sim)^{-1} & (M^\sim/A^\sim)^{-1}
\end{bmatrix}.
\]

Suppose that \(D^\sim\) is non-singular, then

\[
(M^\sim)^{-1} = \begin{bmatrix}
A^\sim & -C^{-1}G_1 \\
-G_1B^\sim & D^\sim
\end{bmatrix}^{-1}
\]

\[
= \begin{bmatrix}
(M^\sim/D^\sim)^{-1} & (M^\sim/D^\sim)^{-1}C^{-1}G_1(A^\sim)^{-1} \\
(D^\sim)^{-1}G_1B^\sim(M^\sim/D^\sim)^{-1} & (D^\sim)^{-1}(G_1B^\sim)(M^\sim/D^\sim)^{-1}C^{-1}G_1(D^\sim)^{-1}
\end{bmatrix}.
\]

If \(A^\sim\) and \(D^\sim\) are both Schur complements then \(M^\sim/A^\sim\) and \(M^\sim/D^\sim\) are all invertible. By comparing the above two expressions for \((M^\sim)^{-1}\), we get the (non-obvious) formula,

\[
(A^\sim - C^{-1}(D^\sim)^{-1}B^\sim)^{-1} = (A^\sim)^{-1} - (A^\sim)^{-1}G_1B^\sim(D^\sim - G_1B^\sim(A^\sim)^{-1}C^{-1}G_1)^{-1}(-G_1B^\sim)(A^\sim)^{-1}.
\]

Using this formula, we obtain another expression for the inverse of \(M^\sim\) involving the Schur complements of \(A^\sim\) and \(D^\sim\),

\[
\begin{bmatrix}
A^\sim & -C^{-1}G_1 \\
-G_1B^\sim & D^\sim
\end{bmatrix}^{-1} = 
\begin{bmatrix}
(A^\sim - C^{-1}(D^\sim)^{-1}B^\sim)^{-1} & (A^\sim)^{-1}C^{-1}G_1(D^\sim - G_1B^\sim(A^\sim)^{-1}C^{-1}G_1) \\
-(D^\sim - G_1B^\sim(A^\sim)^{-1}C^{-1}G_1)^{-1}(-G_1B^\sim)(A^\sim)^{-1} & (D^\sim - G_1B^\sim(A^\sim)^{-1}C^{-1}G_1)^{-1}
\end{bmatrix}.
\]
If we set \( D^* = I \) and \(-C^* G_1\) to \( C^* G_1\) we get

\[
(A^* - C^* B^*)^{-1} = (A^*)^{-1} + (A^*)^{-1} C^* G_1 (I + G_1 B^* (A^*)^{-1} C^* G_1)^{-1} (-G_1 B^*) (A^*)^{-1},
\]

a formula known as the Matrix Inversion formula.

**Proposition 5.5.1** For any m-symmetric matrix \( M^* \) of the form

\[
M^* = \begin{bmatrix}
A^* & -C^* G_1 \\
-G_1 C & -G_1 B^*
\end{bmatrix},
\]

If \(-G_1 B^*\) is invertible then the following properties hold.

(i) \( M^* > 0 \) if and only if \((-G_1 B^*) > 0 \) and \( A^* + C^* G_1 (G_1 B^*)^{-1} (G_1 C) > 0\),

(ii) If \((-G_1 B^*) > 0\), then \( M^* \geq 0 \) if and only if \( A^* + C^* G_1 (G_1 B^*)^{-1} (G_1 C) \geq 0\).

**Proof**: We know that for any m-symmetric matrix \( T \) and any invertible matrix \( N \), the matrix \( T \) is positive definite \((T > 0)\) if and only if \( NTN^* \) is positive definite. That is \( NTN^* > 0\). But a block diagonal matrix is positive definite if and only if each diagonal block is positive definite. Hence (i) satisfied. Similarly, we can show that for any m-symmetric matrix \( T \) and any invertible matrix \( N \), we have \( T \geq 0 \) if and only if \( NTN^* \geq 0\).

**Proposition 5.5.2** For any m-symmetric matrix \( M^* \) of the form

\[
M^* = \begin{bmatrix}
A^* & -C^* G_1 \\
-G_1 C & -G_1 B^*
\end{bmatrix},
\]

If \( A^* \) is invertible then the following properties hold.

(i) \( M^* > 0 \) if and only if \( A^* > 0 \) and \(-G_1 B^* - G_1 C (A^*)^{-1} C^* G_1 > 0\),

(ii) If \( A^* > 0 \), then \( M^* \geq 0 \) if and only if \(-G_1 B^* - G_1 C (A^*)^{-1} C^* G_1 \geq 0\).
Proof: When \(-G_1B^-\) is singular (or \(A^-\) is singular), it is still possible to characterize when a symmetric matrix \(M^-\), as above is positive semidefinite but this requires using a version of the Schur complement involving the Minkowski inverse of \(-G_1B^-\), namely \(A^- + C^-G_1(G_1B^-)^mG_1C\) (or the Schur complement, \(-G_1B^- + G_1C(A^-)^mC^-G_1\) of \(A^-\)).

5.6 Minkowski Inverses

Every square \(n \times n\) matrix \(M^-\) has a singular value decomposition (SVD). We can write \(M^- = U\Sigma V^\sim\), where \(U\) and \(V\) are orthogonal matrices and \(\Sigma\) is a diagonal matrix of the form \(\Sigma = \text{diag}(\sigma_1, \sigma_2, ..., \sigma_r, 0, ..., 0)\) and \(r\) is the rank of \(M^-\). The \((\sigma_i)^\sim\)'s are called the singular values of \(M^-\) and they are the positive square roots of the non zero eigenvalues of \(M^-M\) and \(MM^-\). Also \(U\) and \(V\) are not unique. Furthermore, the columns of \(V\) are eigenvectors of \(MM^-\) and the columns of \(U\) are eigenvectors of \(M^-M\). \(M^- = U\Sigma V^\sim\) is a singular value decomposition of \(M^-\).

If \(\text{rk}(M^-M) = \text{rk}(M^-) = \text{rk}(MM^-)\), then \((M^-)^m\) exists. Then we define the Minkowski inverse \((M^-)^m\) of \(M^-\) by \((M^-)^m = V\Sigma^\sim U^-\), where \(\Sigma^\sim = \text{diag}((\sigma_1)^\sim, (\sigma_2)^\sim, ..., (\sigma_r)^\sim, 0, ..., 0)\). Clearly when \(M^-\) has rank \(r = n\). \(M^-\) is invertible, then \((M^-)^{-1} = (M^-)^m\). That is, \((M^-)^m\) is a Minkowski inverse of \(M^-\). By Theorem 2.3.5 [63], \((M^-)^m\) is uniquely defined in terms of \(M^-\). (The same \((M^-)^m\) is obtained for all possible SVD of \(M^-\)). It is easy to check that \(M^- (M^-)^m M^- = M^-\), \((M^-)^m M^- (M^-)^m = (M^-)^m\) and both \(M^- (M^-)^m\) and \((M^-)^m M^-\) are symmetric matrices.
In fact

\[ M^\sim (M^\sim)^m = U\Sigma V^\sim \Sigma^\sim U^\sim = U\Sigma \Sigma^\sim U^\sim = U \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} U^\sim \]

and

\[ (M^\sim)^m M^\sim = V \Sigma^\sim U^\sim U \Sigma V^\sim = U \Sigma^\sim \Sigma V^\sim = V \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} V^\sim. \]

We immediately get \((M^\sim (M^\sim)^m)^2 = \Sigma (M^\sim)^m \), \((M^\sim)^m M^\sim)^2 = (M^\sim)^m M^\sim\). So both \(M^\sim (M^\sim)^m\) and \((M^\sim)^m M^\sim\) are orthogonal projections (since they both are symmetric). We claim that \((M^\sim)^m\) is the orthogonal projection onto \(\ker(M^\sim)^\perp\), the orthogonal complement of \(\ker(M^\sim)\). Obviously, \(\text{range}(M^\sim) (M^\sim)^m \subseteq \text{range}(M^\sim)\) and for any \(y = M^\sim x \in \text{range}(M^\sim)\), as \(M^\sim (M^\sim)^m M^\sim = M^\sim\), we have \(M^\sim (M^\sim)^m y = M^\sim (M^\sim)^m M^\sim x = M^\sim x = y\). Therefore \(M^\sim (M^\sim)^m y = y\). So the image of \(M^\sim (M^\sim)^m\) is indeed the range of \(M^\sim\). It is also clear that \(\ker(M^\sim) \subseteq \ker((M^\sim)^m M^\sim)\) and since \(M^\sim (M^\sim)^m M^\sim = M^\sim\). We also have \(\ker((M^\sim)^m M^\sim) \subseteq \ker(M^\sim)\) and so \(\ker((M^\sim)^m M^\sim) = \ker(M^\sim)\). Since \((M^\sim)^m M^\sim\) is Hermitian, \(\text{range}((M^\sim)^m M^\sim) = \ker((M^\sim)^m M^\sim)^\perp = \ker(M^\sim)^\perp\) as claimed. It will also be useful to see that \(\text{range}(M^\sim) = \text{range}(M^\sim (M^\sim)^m)\) consists of all vector \(y \in C^n\) such that \(U^\sim y = \begin{bmatrix} Z^- \\ 0 \end{bmatrix}\) with \(Z \in C^r\). Indeed if

\( \begin{align*}
y &= M^\sim x, \text{ then } U^\sim y &= U^\sim M^\sim x = U^\sim U \Sigma V^\sim x = \Sigma V^\sim x \\
 &= \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} V^\sim x = \begin{bmatrix} Z^- \\ 0 \end{bmatrix},
\end{align*} \)

where \(\Sigma_r\) is the \(r \times r\) diagonal matrix \(\text{diag}(\sigma_1, \ldots, \sigma_r)\).

Conversely, if \(U^\sim y = \begin{bmatrix} Z^- \\ 0 \end{bmatrix}\) then \(y = U \begin{bmatrix} Z^- \\ 0 \end{bmatrix}\) and

\[ M^\sim (M^\sim)^m y = U \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} U^\sim y = U \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} U^\sim U \begin{bmatrix} Z^- \\ 0 \end{bmatrix}. \]
\[ M^\ast (M^\ast)^m y = y \] which shows that \( y \) belongs to the range of \( M^\ast \).

Similarly, we claim that \( \text{range}((M^\ast)^m M^\ast) = \ker(M^\ast)^\perp \) consists of all vector \( y \in \mathbb{C}^n \) such that \( V^\ast y = \begin{bmatrix} Z^- \\ 0 \end{bmatrix} \), with \( Z \in \mathbb{C}^r \).

If \( y = (M^\ast)^\ast M^\ast u \), then \( y = (M^\ast)^\ast M^\ast u = V \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} V^\ast u = V \begin{bmatrix} Z^- \\ 0 \end{bmatrix} \), for some \( Z \in \mathbb{C}^r \).

Conversely, if \( V^\ast y = \begin{bmatrix} Z^- \\ 0 \end{bmatrix} \) then \( y = V \begin{bmatrix} Z^- \\ 0 \end{bmatrix} \) and so, \( (M^\ast)^m M^\ast V \begin{bmatrix} Z^- \\ 0 \end{bmatrix} = V \begin{bmatrix} I_r & 0 \\ 0 & 0_{n-r} \end{bmatrix} V^\ast V \begin{bmatrix} Z^- \\ 0 \end{bmatrix} = V \begin{bmatrix} Z^- \\ 0 \end{bmatrix} = y \), which shows that \( y \in \text{range}((M^\ast)^m M^\ast) \). If \( M^\ast \) is a symmetric matrix, then in general there is no SVD, \( U \Sigma V^\ast \) of \( M^\ast \) with \( U = V \). However, if \( M^\ast \geq 0 \), then the eigenvalues of \( M^\ast \) are nonnegative and so the nonzero eigenvalues of \( M^\ast \) are equal to the singular values of \( M^\ast \) and singular value decompositions of \( M^\ast \) of the form \( M^\ast = U \Sigma V^\ast \). In this case \( U \) and \( V \) are unitary matrices, \( M^\ast M \) and \( MM^\ast \) are Hermitian orthogonal projections. If \( M^\ast \) is a normal matrix which means that \( M^\ast M = MM^\ast \), then there is an intimate relationship between singular value decompositions of \( M^\ast \) and block diagonalization of \( M^\ast \). If \( M^\ast \) is a normal matrix, then it can be block diagonalized with respect to an orthogonal matrix \( U \), as \( M^\ast = U \Lambda U^\ast \), where \( \Lambda \) is the block diagonal matrix, where \( \Lambda = \text{diag}(B_1, B_2, \ldots, B_n) \), consisting either of 2 \( \times \) 2 blocks of the form

\[
B_j = \begin{bmatrix} \lambda_j & -\mu_j \\ \mu_j & \lambda_j \end{bmatrix}
\]

with \( \mu_j \neq 0 \) or one dimensional blocks \( B_k = \lambda_k \). Assume that \( B_1, B_2, \ldots, B_p \) are 2 \( \times \) 2 blocks and that \( \lambda_{2p+1}, \ldots, \lambda_n \) are the scalar entries. We know that the numbers \( \lambda_j \pm i\mu_j \), and the \( \lambda_{2p+k} \) are the eigenvalues of \( A \).

Let \( \rho_{2j-1} = \rho_{2j} = \sqrt{(\lambda_j)^2 + (\mu_j)^2} \) for \( j = 1, 2, \ldots, p \).
\[ \rho_{2p+j} = \lambda_j \text{ for } j = 1, 2, ..., n - 2p \] and assume that the blocks are ordered so that \( \rho_1 \geq \rho_2 \geq ... \geq \rho_n \). Then it is easy to see that
\[
UU^\sim = U^\sim U = U \Lambda U^\sim U^\sim U^\sim = U \Lambda \Lambda^\sim U^\sim, \text{ with } \Lambda \Lambda^\sim = \text{diag}(\rho_1^2, ..., \rho_n^2).
\]
So, the singular values \( \sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n \) of \( A \) which are the nonnegative square roots of the eigenvalues of \( AA^\sim \), are such that \( \sigma_j = \rho_j, \ 1 \leq j \leq n \).

For any \( A \in \mathbb{C}^{n \times n} \) in \( m \), If \( A \) is \( m \)-normal if and only if \( A^\sim GA = AGA^\sim \) (By Theorem 2.1.3. [2]). We can define the diagonal matrices
\[
\Sigma = \text{diag}(\sigma_1, ..., \sigma_r, 0, ..., 0) \text{ where } r = rk(A), \sigma_1 \geq ... \geq \sigma_r \geq 0, \text{ and } \\
B = \text{diag}(\sigma_1^{-1} B_1, ..., \sigma_{2p}^{-1} B_p, 1, ..., 1) \text{ so that } B \text{ is an orthogonal matrix and } \\
\Lambda = B \Sigma = (B_1, ..., B_p, \lambda_{2p+1}, ..., \lambda_r, 0, ..., 0). \text{ But then, we can write } \\
A = U \Lambda U^\sim = UB \Sigma U^\sim \text{ and we if let } V = UB, \text{ as } U \text{ is orthogonal and } \\
B \text{ is also orthogonal, } V \text{ is also orthogonal and } A = V \Sigma U^\sim \text{ is an SVD for } A. \text{ Now, we get } \\
A^m = U \Sigma^m V^\sim = U \Sigma^m B^\sim U^\sim. \text{ However, since } B \text{ is an orthogonal matrix, } (B)^\sim = (B)^m \text{ and a simple calculation shows that } \\
\Sigma^m B^\sim = \Sigma^m B^{-1} = \Lambda^m, \text{ which yields the formula } A^m = U \Lambda^m U^\sim. \text{ Also observe that if we write } \\
\Lambda_r = (B_1, ..., B_p, \lambda_{2p+1}, ..., \lambda_r), \text{ then } \Lambda_r \text{ is invertible and } \\
\Lambda^m = \begin{bmatrix} (\Lambda_r)^{-1} & 0 \\ 0 & 0 \end{bmatrix}.
\]
Therefore, the Minkowski inverse of a normal matrix can be computed directly from any block diagonalization of \( A \), as claimed. Next we will use the Minkowski inverses to generalize the result to the symmetric matrices
\[
M^\sim = \begin{bmatrix} A^\sim & -C^\sim G_1 \\ -G_1 C & -G_1 B^\sim \end{bmatrix}
\]
where \( -G_1 B^\sim \text{ (or } A^\sim \) is singular.

**Proposition 5.6.1** If \( P \) is an invertible symmetric matrix, then the function
\[
f(x) = \frac{1}{2} x^\sim P x + x^\sim b \text{ has a minimum value if and only if } P \geq 0, \text{ in which}
\]
Case this optimal value is obtained for a unique value of \( x \), namely
\[
x^* = -P^{-1}b,
\]
and with \( f(P^{-1}b) = -\frac{1}{2}b^\sim P^{-1}b \).

Proof: Observe that
\[
\frac{1}{2}(x + P^{-1}b)^\sim P(x + P^{-1}b) = \frac{1}{2}x^\sim Px + x^\sim b + \frac{1}{2}b^\sim P^{-1}b.
\]
Thus, \( f(x) = \frac{1}{2}x^\sim Px + x^\sim b = \frac{1}{2}(x + P^{-1}b)^\sim P(x + P^{-1}b) - \frac{1}{2}b^\sim P^{-1}b \).

If \( P \) has some negative eigenvalue, say \(-\lambda \) (with \( \lambda \geq 0 \)), if we pick any eigenvector \( u \) of \( P \) associated with \( \lambda \), then for any \( \alpha \in \mathbb{R} \) with \( \alpha \neq 0 \), if we let \( x = \alpha u - P^{-1}b \), then as \( Pu = -\lambda u \), we get
\[
f(x) = \frac{1}{2}(x + P^{-1}b)^\sim P(x + P^{-1}b) - \frac{1}{2}b^\sim P^{-1}b,
\]
\[
= \frac{1}{2}\alpha u^\sim Pu - \frac{1}{2}b^\sim P^{-1}b,
\]
\[
= -\frac{1}{2}\alpha^2 \lambda u^2 - \frac{1}{2}b^\sim P^{-1}b,
\]
and as \( \alpha \) can be made as large as we want and \( \lambda \geq 0 \), we see that \( f \) has no minimum. Consequently, in order for \( f \) to have a minimum, we must have \( P \geq 0 \). In this case, as \((x + P^{-1}b)^\sim P(x + P^{-1}b) \geq 0 \), it is clear that the minimum value of \( f \) is achieved when \( x + P^{-1}b = 0 \). (since the equation \( f(x) \) is of symmetric form consisting of eigen vectors. The minimum value can be determined by \( f'(x) = 0 /x > 0. \) ) That is \( x = -P^{-1}b \).

**Proposition 5.6.2** If \( P \) is a symmetric matrix, then the function
\[
f(x) = \frac{1}{2}x^\sim Px + x^\sim b
\]
has a minimum value if and only if \( P \geq 0 \) and \((I - PP^m)b = 0 \), in which case this minimum value is \( P^* = -\frac{1}{2}b^\sim P^m b \).

Furthermore, if \( P = U^\sim \Sigma U \) is an SVD of \( P \), then the optimal value is achieved by all \( x \in C^n \) of the form \( x = -P^m b + U^\sim \begin{bmatrix} 0 \\ Z \end{bmatrix} \), for any \( Z \in C^{n-r} \) where \( r \) is the rank of \( P \).

Proof: The case where \( P \) is invertible is taken care of by proposition (5.6.1). so, we may assume that \( P \) is singular. If \( P \) has rank \( r < n \), then
we can diagonalize $P$ as $P = U^\sim \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} U$, where $U$ is an orthogonal matrix and where $\Sigma_r$ is an $r \times r$ diagonal invertible matrix. Then, we have

$$f(x) = \frac{1}{2} x^\sim U^\sim \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} U x + x^\sim U^\sim U b$$

$$= \frac{1}{2} (U x)^\sim \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} U x + (U x)^\sim U b.$$  

If we write $U x = \begin{bmatrix} y \\ z \end{bmatrix}$ and $U b = \begin{bmatrix} c \\ d \end{bmatrix}$ with $y, c \in \mathbb{C}^r$ and $z, d \in \mathbb{C}^{n-r}$, we get,

$$f(x) = \frac{1}{2} (U x)^\sim \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} U x + (U x)^\sim U b$$

$$= \frac{1}{2} (y^\sim, z^\sim) \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} + \begin{bmatrix} y^\sim, z^\sim \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix},$$

$$= \frac{1}{2} y^\sim \Sigma_r y + y^\sim c + z^\sim d.$$  

For $y = 0$, we get $f(x) = z^\sim d$, so if $d \neq 0$, the function $f$ has no minimum. Therefore, if $f$ has a minimum, then $d = 0$. However, $d = 0$ means that $U b = \begin{bmatrix} c \\ 0 \end{bmatrix}$ and we know that $b$ is in the range of $P$ (here $U$ is $U^\sim$) which is equivalent to $(I - P P^\sim) b = 0$. If $d = 0$, then

$$f(x) = \frac{1}{2} y^\sim \Sigma_r y + y^\sim c$$

and as $\Sigma_r$ is invertible, By proposition 5.6.1, the function $f$ has a minimum if and only if $\Sigma_r \geq 0$, which is equivalent to $P \geq 0$. Therefore, we proved that if $f$ has a minimum, then $(I - P P^\sim)b = 0$ and $P \geq 0$. Conversely, if $(I - P P^\sim)b = 0$ and $P \geq 0$. To prove that $f$ does have a minimum. When the above conditions hold, the minimum is
achieved if \( y = -\Sigma^{-1} c, z = 0 \) and \( d = 0 \).

i.e, For \( x^* \) given by \( U x^* = \begin{bmatrix} -\Sigma^{-1} c \\ 0 \end{bmatrix} \) and \( U b = \begin{bmatrix} c \\ 0 \end{bmatrix} \).

From which we deduce that \( x^* = -U\begin{bmatrix} \Sigma^{-1} c \\ 0 \end{bmatrix} = -U\begin{bmatrix} \Sigma^{-1} c & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} c \\ 0 \end{bmatrix} = -(U)\begin{bmatrix} \Sigma^{-1} c \\ 0 \end{bmatrix} U b = -P^m b \)

and the minimum value of \( f \) is \( f(x^*) = -\frac{1}{2} b^\sim P^m b \).

For any \( x \in C^n \) of the form \( x = -P^m b + U^\sim \begin{bmatrix} 0 \\ z \end{bmatrix} \) for any \( z \in C^{n-r} \).

Our previous calculations shows that \( f(x) = -\frac{1}{2} b^\sim P^m b \). When a symmetric matrix \( \begin{bmatrix} A^- & -C^\sim G_1 \\ -G_1 C & -G_1 B^- \end{bmatrix} \) is positive semidefinite. Thus we want to know when the function

\[
f(x, y) = (x^\sim, y^\sim) \begin{bmatrix} A^- & -C^\sim G_1 \\ -G_1 C & -G_1 B^- \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}
\]

\[
f(x, y) = x^\sim A^- x - 2x^\sim C^\sim G_1 y - y^\sim G_1 B^- y
\]

has a minimum with respect to both \( x \) and \( y \). Holding \( y \) constant, proposition 5.6.2 implies that \( f(x, y) \) has a minimum if and only if \( A \geq 0 \) and \( (I - AA^-)By = 0 \) and then, the minimum value is

\[
f(x^*, y) = -y^\sim B^- A^m By + y^\sim cy = y^\sim (c - B^- A^m B)y.
\]

Since we want \( f(x, y) \) to be uniformly bounded from below for all \( x, y \) we must have \( (I - AA^m)B = 0 \). Now, \( f(x^*, y) \) has a minimum if and only if \( A \geq 0, \ (I - AA^m)B = 0, \ C - B^- A^m B \geq 0 \). Therefore we established that \( f(x, y) \) has a minimum over all \( x, y \) if and only if \( A \geq 0, \ (I - AA^m)B = 0, \ C - B^- A^m B \geq 0 \). A similar reasoning applies if we first minimize with respect to \( y \) and then with respect to \( x \), but this time, the Schur complement \( A^- - C^\sim G_1 (G_1 B^-)^m (C^\sim G_1)^- \), of \( -G_1 B^- \) is involved. Putting all these
facts together we get our main result.

**Theorem 5.6.1** Given any symmetric matrix \( M^\sim = \begin{bmatrix} A^\sim & -C^\sim G_1 \\ -G_1 C & -G_1 B^\sim \end{bmatrix} \), the following conditions are equivalent.

(i) \( M^\sim \succeq 0 \) (\( M^\sim \) is positive semidefinite.)

(ii) \( A^\sim \succeq 0, \ (I - A^\sim (A^\sim)^m) (-C^\sim G_1) = 0, \)
\[ -G_1 B^\sim - G_1 C (A^\sim)^m C^\sim G_1 \succeq 0. \]

(iii) \( -G_1 B^\sim \succeq 0, \ (I - G_1 B^\sim B G_1) (-G_1 C)^* = 0. \)

**Proof:** If \( M^\sim \succeq 0 \), then by Proposition 5.5.1, it is clear that the above conditions are equivalent.