1-norm minimization and minimum-rank structured sparsity for symmetric and ah-symmetric generalized inverses: rank one and two

Luze Xu, Marcia Fampa, Jon Lee

Abstract  Generalized inverses are important in statistics and other areas of applied matrix algebra. A generalized inverse of a real matrix $A$ is a matrix $H$ that satisfies the Moore-Penrose (M-P) property $AHA = A$. If $H$ also satisfies the M-P property $HAH = H$, then it is called reflexive. Reflexivity of a generalized inverse is equivalent to minimum rank, a highly desirable property. We consider aspects of symmetry related to the calculation of various sparse reflexive generalized inverses of $A$. As is common, we use (vector) 1-norm minimization for both inducing sparsity and for keeping the magnitude of entries under control.

When $A$ is symmetric, a symmetric $H$ is highly desirable, but generally such a restriction on $H$ will not lead to a 1-norm minimizing reflexive generalized inverse. We investigate a block construction method to produce a symmetric reflexive generalized inverse that is structured and has guaranteed sparsity. Letting the rank of $A$ be $r$, we establish that the 1-norm minimizing generalized inverse of this type is a 1-norm minimizing symmetric generalized inverse when (i) $r = 1$ and when (ii) $r = 2$ and $A$ is nonnegative.

Another aspect of symmetry that we consider relates to another M-P property: $H$ is ah-symmetric if $AH$ is symmetric. The ah-symmetry property is sufficient for a generalized inverse to be used to solve the least-squares problem $\min\{\|Ax - b\|_2 : x \in \mathbb{R}^n\}$ using $H$, via $x := Hb$. We investigate a column block construction method to produce an ah-symmetric reflexive generalized inverse that is structured and has guaranteed sparsity. We establish that the 1-norm minimizing ah-symmetric generalized inverse of this type is...
a 1-norm minimizing ah-symmetric generalized inverse when (i) \( r = 1 \) and when (ii) \( r = 2 \) and \( A \) satisfies a technical condition.

1 Introduction

Generalized inverses are essential tools in statistics and other applications of matrix algebra. Of central importance is the celebrated Moore-Penrose (M-P) pseudoinverse, which can be used, for example, to calculate the least-squares solution of an over-determined system of linear equations. With respect to the system \( Ax = b \), the least-squares solution \( \min\{\|Ax - b\|_2 : x \in \mathbb{R}^n\} \) is given by \( x := A^+b \), where \( A^+ \) is the M-P pseudoinverse. Our motivating use case is that we have a very large (rank deficient) matrix \( A \) and multiple right-hand sides \( b \). And so we can see the value of having a sparse generalized inverse.

In what follows, we write \( \|H\|_1 \) to mean \( \|\text{vec}(H)\|_1 \) and \( \|H\|_{\max} \) to mean \( \|\text{vec}(H)\|_{\max} \) (in both cases, these are not the usual induced/operator matrix norms). We use \( I \) for an identity matrix and \( J \) for an all-ones matrix. Matrix dot product is indicated by \( \langle X, Y \rangle := \sum_{ij} x_{ij} y_{ij} = \text{trace}(X^\top Y) \). We use \( A[S,T] \) for the submatrix of \( A \) with row indices \( S \) and column indices \( T \); additionally, we use \( A[S,:] \) (resp., \( A[:,T] \)) for the submatrix of \( A \) formed by the rows \( S \) (resp., columns \( T \)). Finally, if \( A \) is symmetric, we let \( A[S] := A[S,S] \), the principal submatrix of \( A \) with row/column indices \( S \).

When a real matrix \( A \in \mathbb{R}^{m \times n} \) is not square or is square but not invertible, we consider “pseudoinverses” of \( A \) (see \cite{15}). The most well-known pseudoinverse is the M-P pseudoinverse independently discovered by A. Bjerhammar, E.H. Moore and R. Penrose (see \cite{1, 6, 14}). If \( A = U\Sigma V^\top \) is the real singular-value decomposition of \( A \) (see \cite{12}, for example), where \( U \in \mathbb{R}^{m \times m} \), \( V \in \mathbb{R}^{n \times n} \) are orthogonal matrices and \( \Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_p) \in \mathbb{R}^{m \times n} \) (\( p = \min\{m, n\} \)) with singular values \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0 \), then the M-P pseudoinverse of \( A \) can be defined as \( A^+ := V \Sigma^+ U^\top \), where \( \Sigma^+ := \text{diag}(\sigma_1^+, \sigma_2^+, \ldots, \sigma_p^+) \in \mathbb{R}^{n \times m} \), \( \sigma_i^+ := 1/\sigma_i \) for all \( \sigma_i \neq 0 \), and \( \sigma_i^+ := 0 \) for all \( \sigma_i = 0 \). The starting point for our investigation is the following celebrated result.

**Theorem 1 (see \cite{14})**. For \( A \in \mathbb{R}^{m \times n} \), the M-P pseudoinverse \( A^+ \) is the unique \( H \in \mathbb{R}^{n \times m} \) satisfying:

\[
\begin{align*}
AH &= A \\
HA &= H \\
(AH)^\top &= AH \\
(HA)^\top &= HA
\end{align*}
\]
Following [16], a **generalized inverse** is any $H$ satisfying $P_1$. Because we are interested in sparse $H$, $P_1$ is important to enforce, otherwise the completely sparse zero-matrix (which carries no information from $A$) always satisfies the other three M-P properties. A generalized inverse is **reflexive** if it satisfies $P_2$. Two very useful facts are: (i) if $H$ is a generalized inverse of $A$, then $\text{rank}(H) \geq \text{rank}(A)$, and (ii) a generalized inverse $H$ of $A$ is reflexive if and only if $\text{rank}(H) = \text{rank}(A)$ (see [16, Theorem 3.14]). Therefore, enforcing $P_2$ for a generalized inverse implies that the generalized inverse has minimum rank. A low-rank $H$ can be viewed as being more interpretable/explainable model (say in the context of the least-squares problem), so we naturally prefer reflexive generalized inverses (which have the least rank possible among generalized inverses). As we have said, we are interested in sparse generalized inverses. But structured sparsity of $H$ is even more valuable, as it can be viewed, in a different way, as being a more interpretable/explainable model.

Following [19], if $H$ satisfies $P_3$, we say that $H$ is **ah-symmetric**. That is, ah-symmetric means that $AH$ is symmetric. It is very important to know that not all of the M-P properties are required for a generalized inverse to exactly solve key problems. For example, if $H$ is any ah-symmetric generalized inverse, then $\hat{x} := Hb$ solves the least-squares problem $\min \{\|Ax - b\|_2 : x \in \mathbb{R}^n\}$ (see [2] or [9], for example).

**Previous work:** [5, 3, 4] used sparse-optimization techniques to give tractable right and left sparse pseudoinverses. [9] introduced the idea of sparse generalized inverses obtained by 1-norm minimization over convex relaxations of the M-P properties (also see [10]). [7] introduced “block solutions” giving reflexive generalized inverses with structured sparsity. They solved the 1-norm minimization problem for the rank-1 case and the rank-2 nonnegative case, and they gave an efficient approximation algorithm for the general-rank case. For symmetric reflexive generalized inverses, [19] gave a 1-norm approximation algorithm for the general-rank case, based on a local search over symmetric block solutions. Additionally, [19] introduced “column block solutions” giving reflexive ah-symmetric generalized inverses with structured sparsity. Furthermore, they gave an efficient 1-norm approximation algorithm for the general-rank case, based on a local search over column block solutions. [8] makes a detailed computational study of all these approximation algorithms (also see [19]).

The approximation algorithms aimed at minimizing $\|H\|_1$ from [7] and [19] (over $H$ satisfying $P_1+P_2$, or $P_1+P_2+H=H^T$, or $P_1+P_2+P_3$) carry out local searches (using appropriate local-search neighborhoods, over, respectively, block solutions, symmetric block solutions, or column block solutions — all reflexive and structured), but not with $\|H\|_1$ as a local-minimization criterion. In fact, in [19] it was demonstrated that such a criterion can fail to provide a decent approximation ratio for $r := \text{rank}(A) \geq 2$. However, the global minimizer with respect to the criterion $\|H\|_1$ might provide a better approximation ratio than the local search. If $r$ is very small, we could calculate a globally-minimum $\|H\|_1$ (over block solutions, symmetric block solutions, or
column block solutions, respectively) and hope that such a solution could be a 1-norm minimizing solution (satisfying $P_1 + P_2$, or $P_1 + P_2 + H = H^\top$, or $P_1 + P_2 + P_3$, respectively). In fact, [7] demonstrated this for the case of $P_1 + P_2$, when $r = 1$ and when $r = 2$ for nonnegative $A$.

**Our main contributions:** In §2, (i) we solve the 1-norm minimization problem for $P_1 + P_2 + H = H^\top$ (i.e., symmetric reflexive generalized inverses), in the rank-1 case and the rank-2 nonnegative case. In §3, (ii) we solve the 1-norm minimization problem for $P_1 + P_2 + P_3$ (i.e., ah-symmetric reflexive generalized inverses), in the rank-1 case and the rank-2 case under an efficiently-checkable technical condition, and (iii) we demonstrate that this technical condition is essentially necessary. Already, the rank-2 cases need side conditions and have rather complicated proofs. So some new insights will be needed to characterize optimality beyond the cases that we handle. Nevertheless, we believe that our results can be stepping stones to characterizing optimal solutions or getting better approximation ratios in further low-rank cases.

In what follows, a key tool that we employ is linear-optimization duality. Even though we are interested in reflexive generalized inverses, it is useful to consider $\min\{\|H\|_1 : P_1\} = \min\{\|H\|_1 : AHA = A\}$, which we re-cast as a linear-optimization problem (P) and its dual (D):

\[
\begin{align*}
\text{minimize} & \quad \langle J, H^+ \rangle + \langle J, H^- \rangle \\
\text{subject to} & \quad A(H^+ - H^-)A = A, \\
& \quad H^+, H^- \geq 0; \\
\end{align*}
\]

\[
\begin{align*}
\text{maximize} & \quad \langle A, W \rangle \\
\text{subject to} & \quad -J \leq A^\top W A^\top \leq J.
\end{align*}
\]

More compactly, we can recast (D) as: $\max\{\langle A, W \rangle : \|A^\top W A^\top\|_{\text{max}} \leq 1\}$.

## 2 Symmetric

Our starting point is the construction of symmetric block solutions via the following result.

**Theorem 2 ([19]).** For a symmetric matrix $A \in \mathbb{R}^{n \times n}$, let $r := \text{rank}(A)$. Let $\hat{A} := A[S]$ be any $r \times r$ nonsingular principal submatrix of $A$. Let $H \in \mathbb{R}^{n \times n}$ be equal to zero, except its submatrix with row/column indices $S$ is equal to $\hat{A}^{-1}$. Then $H$ is a symmetric reflexive generalized inverse of $A$. 


2.1 Rank 1

Next, we demonstrate that when \( \text{rank}(A) = 1 \), construction of a 1-norm minimizing symmetric reflexive generalized inverse can be based on the symmetric block construction over the diagonal elements of \( A \).

**Theorem 3.** Let \( A \) be an arbitrary rank-1 symmetric matrix, which is, without loss of generality, of the form \( A := uu^\top \), where \( 0 \neq u \in \mathbb{R}^n \). If \( i^* := \arg \max_i \{|u_i|\} = \arg \max_i \{|a_{ii}|\} \), then \( H := \frac{1}{u_{i^*}^2} e_{i^*} e_{i^*}^\top \), where \( e_{i^*} \in \mathbb{R}^n \) is a standard unit vector, is a symmetric reflexive generalized inverse of \( A \) with minimum 1-norm.

**Proof.** We consider \((P)\) and \((D)\). A feasible solution for \((P)\) is \( H^+ = \frac{1}{u_{i^*}^2} e_{i^*} e_{i^*}^\top \), \( H^- = 0 \). A feasible solution for \((D)\) is \( W = \frac{1}{u_{i^*}^2} e_{i^*} e_{i^*}^\top \), because \( \|A^\top WA\|_{\text{max}} = \frac{1}{u_{i^*}^2} \|A\|_{\text{max}} = 1 \). And the objective value of the dual solution is \( \langle A, W \rangle = \frac{1}{u_{i^*}^2} \cdot \frac{1}{u_{i^*}^2} = 1 \), which is the objective value of the primal solution. Therefore, by the weak-duality theorem of linear optimization, we have that \( H := H^+ - H^- \) is a generalized inverse of \( A \) with minimum 1-norm. By our construction, \( H \) is symmetric and reflexive. Therefore, \( H \) is a symmetric reflexive generalized inverse with minimum 1-norm. \( \square \)

Another way to view the rank-1 case is by using the Kronecker product to transform the constraint \( AHA = A \) into \( [A^\top \otimes A] \text{vec}(H) = \text{vec}(A) \). Note that \( \text{vec}(A) = \text{vec}(uu^\top) = u \otimes u \), and \( A^\top \otimes A = uu^\top \otimes uu^\top = [u \otimes u][u^\top \otimes u^\top] \). So the constraint becomes \( [u \otimes u] ([u^\top \otimes u^\top] \text{vec}(H)) = \frac{1}{u_{i^*}^2} e_{i^*} e_{i^*}^\top \). Thus the 1-norm minimization may be re-cast as \( \min \{|\text{vec}(H)|_1 : [u \otimes u]^\top \text{vec}(H) = 1 \} \), or \( \min \{|\text{vec}(H)|_1 : u^\top Hu = 1 \} \), or \( \min \{|H|_{\text{max}} : H = 1 \} \). By using the inequality \( x^\top y \leq \|x\|_\infty \|y\|_1 \), or \( \langle X, Y \rangle \leq \|X\|_{\text{max}} \|Y\|_1 \), we have \( \|H\|_1 \geq 1/\|uu^\top\|_{\text{max}} \) and the equality holds when \( H = \frac{1}{u_{i^*}^2} e_{i^*} e_{i^*}^\top \).

2.2 Rank 2

Generally, when \( \text{rank}(A) = 2 \), we cannot construct a 1-norm minimizing symmetric reflexive generalized inverse based on the symmetric block construction. For example, with

\[
A := \begin{bmatrix}
5 & 4 & 2 \\
4 & 5 & -2 \\
2 & -2 & 8
\end{bmatrix},
\]

we have a symmetric reflexive generalized inverse \( H := \frac{1}{81} A \) (because \( A^2 = 9A \)), with \( \|H\|_1 = \frac{2}{81} \). While the three symmetric reflexive generalized in-
verses based on the symmetric block construction have 1-norm equal to \( \frac{17}{36}, \frac{17}{36}, 2 \), all greater than \( \frac{34}{81} \).

Next, we demonstrate that under the natural but restrictive condition that \( A \) is non-negative, when \( \text{rank}(A) = 2 \), construction of a 1-norm minimizing symmetric reflexive generalized inverse can be based on the symmetric block construction over the \( 2 \times 2 \) principal submatrix of \( A \).

**Theorem 4.** Let \( A \) be an arbitrary rank-2 non-negative symmetric matrix. For any \( i_1, i_2 \in \{1, \ldots, n\} \), with \( i_1 < i_2 \), let \( \tilde{A} := A[i_1, i_2] \). If \( i_1, i_2 \) are chosen to minimize the 1-norm of \( \tilde{A}^{-1} \) among all nonsingular \( 2 \times 2 \) principal submatrices, then the \( n \times n \) matrix \( H \) constructed by Theorem 2 over \( \tilde{A} \), is a symmetric reflexive generalized inverse of \( A \) with minimum 1-norm.

**Proof.** Without loss of generality that \( \tilde{A} \) is in the north-west corner of \( A \). So we take \( A \) to have the form 
\[
\begin{bmatrix}
\tilde{A} & B \\
B^T & D
\end{bmatrix}
\]
Let \( M = 2I - J \) if \( \det(\tilde{A}) > 0 \) and \( M = J - 2I \) if \( \det(\tilde{A}) < 0 \), now we let
\[
W := \begin{bmatrix}
\tilde{W} & 0 \\
0 & 0
\end{bmatrix} := \tilde{A}^{-T} M \tilde{A}^{-T} \begin{bmatrix}
0 \\
0
\end{bmatrix}.
\]
The dual objective value
\[
\langle A, W \rangle = \text{trace}(A^T W) = \text{trace}(\tilde{A}^T \tilde{W}) = \text{trace}(M \tilde{A}^{-T}) = \langle M, \tilde{A}^{-1} \rangle = \|\tilde{A}^{-1}\|_1,
\]
i.e., \( \langle A, W \rangle = \|H\|_1 \). Also,
\[
A^T W A = \begin{bmatrix}
M & M \tilde{A}^{-T} B \\
B^T \tilde{A}^{-T} M & B^T \tilde{A}^{-T} M \tilde{A}^{-T} B
\end{bmatrix}.
\]
Clearly \( \|M\|_{\text{max}} \leq 1 \). Next, we consider \( \tilde{\gamma} := M \tilde{A}^{-T} \gamma = M \tilde{A}^{-1} \gamma \), where \( \gamma \) is an arbitrary column of \( B \). As \( \text{rank}(A) = 2 \) and \( \tilde{A} \) is nonsingular, we assume that \( \gamma = \tilde{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \). We may as well assume that \( x_1, x_2 \) are not both zero, because otherwise \( \tilde{\gamma} = 0 \) satisfying \( \|\tilde{\gamma}\|_{\text{max}} \leq 1 \). We have
\[
\tilde{\gamma} = M \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Rightarrow \|\tilde{\gamma}\|_{\text{max}} = |x_1 - x_2|.
\]
Consider \( \tilde{A} \) is chosen to minimize the 1-norm of \( \tilde{A}^{-1} \) among all nonsingular \( 2 \times 2 \) principal submatrices, we have
\[
\|\tilde{A}^{-1}\|_1 \leq \left\| \begin{bmatrix}
a_{11} & a_{11} x_1 + a_{12} x_2 \\
a_{11} x_1 + a_{12} x_2 & a_{11} x_1^2 + 2a_{12} x_1 x_2 + a_{22} x_2^2
\end{bmatrix}^{-1} \right\|_1
\] (1)
\begin{equation}
\|\tilde{A}^{-1}\|_1 \leq \left\| \begin{array}{cc}
a_{22} & a_{12}x_1 + a_{22}x_2 \\
a_{12}x_1 + a_{22}x_2 & a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2
\end{array} \right\|^{-1} \tag{2}
\end{equation}

Case 1. \(x_1 = 0\). If \(a_{12} = 0\), then \(a_{11}, a_{22} > 0\) because \(\tilde{A}\) is nonsingular. Using (1), we have
\[
\frac{a_{11} + 2a_{12} + a_{22}}{|\det(\tilde{A})|} \leq \frac{a_{11}(1 + x_1)^2 + 2a_{12}x_2(1 + x_1) + a_{22}x_2^2}{x_2^2|\det(\tilde{A})|}
\]
Simplifying, we obtain \(a_{11}(x_2^2 - 1) \leq 0 \Rightarrow |x_2| \leq 1\). If \(a_{12} > 0\), then \(x_2 \geq 0\) because \(a_{11}x_1 + a_{12}x_2 = a_{12}x_2 \geq 0\). Because \(x_1, x_2\) are not both zero, we have \(x_2 > 0\). Still using (1), we have \((a_{11}x_2 + a_{11} + 2a_{12}x_2)(x_2 - 1) \leq 0\) which implies \(x_2 \leq 1\).

Case 2. \(x_2 = 0\). Similarly by using (2), we have \(|x_1| \leq 1\) or \(0 < x_1 \leq 1\).

Case 3. \(x_2 \geq x_1, x_1, x_2 \neq 0\). Using (1), we have
\[
\frac{a_{11} + 2a_{12} + a_{22}}{|\det(\tilde{A})|} \leq \frac{a_{11}(1 + x_1)^2 + 2a_{12}x_2(1 + x_1) + a_{22}x_2^2}{x_2^2|\det(\tilde{A})|}
\]
Simplifying, we obtain \((a_{11}(x_2^2 + 1 + x_1) + 2a_{12}x_2)(x_2 - 1 - x_1) \leq 0\) because \(a_{11}(x_2 + 1 + x_1) + 2a_{12}x_2 = 2(a_{11}x_1 + a_{12}x_2) + a_{11}(1 + x_2 - x_1) > 0\), (it is zero only when \(a_{11} = 0\) and \(a_{11}x_1 + a_{12}x_2 = 0\), which implies \(a_{11} = a_{12} = 0\), a contradiction.) we obtain \(0 \leq x_2 - x_1 \leq 1\).

Case 4. \(x_2 < x_1, x_1, x_2 \neq 0\). Using (2), we have
\[
\frac{a_{11} + 2a_{12} + a_{22}}{|\det(\tilde{A})|} \leq \frac{a_{11}x_1^2 + 2a_{12}x_1(1 + x_2) + a_{22}(1 + x_2)^2}{x_1^2|\det(\tilde{A})|}
\]
Simplifying, we obtain \((a_{22}(x_1 + 1 + x_2) + 2a_{12}x_1)(x_1 - 1 - x_2) \leq 0\) because \(a_{22}(x_1 + 1 + x_2) + 2a_{12}x_1 = 2(a_{12}x_1 + a_{22}x_2) + a_{22}(1 - x_2 + x_1) > 0\), (it is zero only when \(a_{22} = 0\) and \(a_{12}x_1 + a_{22}x_2 = 0\), which implies \(a_{22} = a_{12} = 0\), a contradiction.) we obtain \(0 < x_1 - x_2 \leq 1\).

From the above, we show that \(|x_1 - x_2| \leq 1\), thus \(\|M\tilde{A}^{-\top}B\|_{\text{max}} = \|B^\top\tilde{A}^{-\top}M\|_{\text{max}} \leq 1\). Finally,
\[
\|B^\top\tilde{A}^{-\top}M\tilde{A}^{-\top}B\|_{\text{max}} = \frac{1}{2}\|B^\top\tilde{A}^{-\top}(4I - 2J)\tilde{A}^{-\top}B\|_{\text{max}}
\]
\[
= \frac{1}{2}\|B^\top\tilde{A}^{-\top}M^2\tilde{A}^{-\top}B\|_{\text{max}}
\]
\[
\leq \|B^\top\tilde{A}^{-\top}M\|_{\text{max}}\|M\tilde{A}^{-\top}B\|_{\text{max}} \leq 1
\]

Therefore, \(W\) is dual feasible. By the weak duality, we know that \(H\) is a symmetric generalized inverse of \(A\) with minimum 1-norm, which is also reflexive. \(\Box\)
3 ah-symmetric

Our starting point is the construction of column block solutions via the following result.

**Theorem 5 ([19]).** For $A \in \mathbb{R}^{m \times n}$, let $r := \text{rank}(A)$. For any $T$, an ordered subset of $r$ elements from $\{1, \ldots, n\}$, let $\hat{A} := A[:, T]$ be the $m \times r$ submatrix of $A$ formed by columns $T$. If $\text{rank}(\hat{A}) = r$, let $\hat{H} := (\hat{A}^\top \hat{A})^{-1} \hat{A}^\top$. The $n \times m$ matrix $H$ with all rows equal to zero, except rows $T$, which are given by $\hat{H}$, is an ah-symmetric reflexive generalized inverse of $A$.

Similarly as before, we note that it is useful to consider $\min \{\|H\|_1 : P1 + P3\} = \min \{\|H\|_1 : AHA = A, (AH)^\top = AH\}$, which we re-cast as a linear-optimization problem ($P_{ah}$) and its dual ($D_{ah}$):

**$P_{ah}$**

\[
\begin{align*}
\text{minimize} & \quad \langle J, H^+ \rangle + \langle J, H^- \rangle \\
\text{subject to} & \quad A(H^+ - H^-)A = A, \\
& \quad (H^+ - H^-)^\top A^\top = A(H^+ - H^-), \\
& \quad H^+, H^- \geq 0.
\end{align*}
\]

**$D_{ah}$**

\[
\begin{align*}
\text{maximize} & \quad \langle A, W \rangle \\
\text{subject to} & \quad -J \leq A^\top WA^\top + A^\top (V^\top - V) \leq J
\end{align*}
\]

We can see ($D_{ah}$) as: $\max \{\langle A, W \rangle : \|A^\top WA^\top + A^\top U\|_{\max} \leq 1, U^\top = -U\}$.

3.1 Rank 1

Next, we demonstrate that when $\text{rank}(A) = 1$, construction of a 1-norm minimizing ah-symmetric reflexive generalized inverse can be based on the column block construction.

**Theorem 6.** Let $A$ be an arbitrary $m \times n$, rank-$1$ matrix. For any $j \in \{1, \ldots, n\}$, let $\hat{a}$ be column $j$ of $A$. If $j$ is chosen to minimize the 1-norm of $\hat{a}^+$ among all columns except the zero columns, then the $n \times m$ matrix $H$ constructed by Theorem 5 over $\hat{a}$, is an ah-symmetric reflexive generalized inverse of $A$ with minimum 1-norm.

**Proof.** We prove a stronger result — that our constructed $H$ is a 1-norm minimizing ah-symmetric generalized inverse. By our construction, $H$ is reflexive, thus $H$ is an ah-symmetric reflexive generalized inverse with minimum 1-norm. To establish the minimum 1-norm of $H$, we consider the linear-optimization problems ($P_{ah}$) and ($D_{ah}$). As verified in Theorem 5, $H$ is a feasible solution for ($P_{ah}$), and its objective value is $\|H\|_1 = \|\hat{a}^+\|_1$ (it also satisfies the nonlinear equations ($P2$)).
The objective function of \((D_{ab})\) only depends on the variable \(W\). Feasibility of a \(W\) is equivalent to the existence of a skew-symmetric matrix \(U\) so that
\[
\|A^T W A^T + A^T U\|_{\text{max}} \leq 1 .
\]
(3)

Next, we are going to construct a dual feasible solution \(W\) with objective value \(\langle A, W \rangle = \|H\|_1\); then by the weak duality for linear optimization, we establish that \(H\) is optimal to \((P_{ab})\).

Let \(z := \text{sign}(\hat{a}^+)\). Suppose that \(\hat{a}_i\) is a nonzero element in \(\hat{a}\) with index \(i\). Let \(W\) be an \(m \times n\) matrix with all elements equal to zero, except the one in row \(i\) and column \(j\), which is given by \(\hat{w}\). Let \(U\) be an \(m \times m\) skew-symmetric matrix, with only row \(i\) and column \(i\) different from zero.

If \(\hat{w}\) and \(U\) are chosen to be
\[
\hat{w} := \frac{1}{\hat{a}_i} z(\hat{a}^+)^\top, \quad u_{ki} := \frac{1}{\hat{a}_i} (\hat{a}_k z(\hat{a}^+)^\top - z_k), \quad \forall \ k \neq i,
\]
then they satisfy
\[
\hat{a}_i \hat{w} \hat{a}^\top + \hat{a}^\top U = z .
\]
(4)

This is because for \(k \neq i\),
\[
\hat{a}_i \hat{w} \hat{a}_k + \hat{a}_i u_{ik} = z(\hat{a}^+)^\top \hat{a}_k + z_k - \hat{a}_k z(\hat{a}^+)^\top = z_k,
\]
and
\[
\hat{a}_i \hat{w} \hat{a}_k + \sum_{k \neq i} \hat{a}_k u_{ki} = z(\hat{a}^+)^\top (\hat{a}_i + \sum_{k \neq i} \hat{a}_k^2) - \sum_{k \neq i} \hat{a}_k z_k
\]
\[
= \frac{1}{\hat{a}_i} \left( z(\hat{a}^+)^\top (\hat{a}^\top \hat{a}) - z \hat{a} \right) + z_i = z_i ,
\]
and
\[
\text{trace}(A^T W) = \hat{a}_i \hat{w} = z(\hat{a}^+)^\top = \|\hat{a}^+\|_1 = \|H\|_1 .
\]

The dual constraint (3) can be written as
\[
\|\hat{a}^T W A^T + \hat{a}^T U\|_{\text{max}} \leq 1 ,
\]
(5)

and
\[
\|\hat{B}^T W A^T + \hat{B}^T U\|_{\text{max}} \leq 1 .
\]
(6)

From (4), we have that (5) is satisfied. To verify (6), let \(\hat{b} \in \mathbb{R}^m\) be an arbitrary column of \(\hat{B}\). As \(A\) has rank 1,
\[
\hat{b} = \alpha \hat{a} ,
\]
and
\[
\hat{b}^T W A^T + \hat{b}^T U = \alpha \hat{a}^T (W A^T + U) .
\]

Considering (4), we have that \(\|\hat{a}^T (W A^T + U)\|_{\text{max}} = 1\), and therefore,
\[
\|\hat{b}^T W A^T + \hat{b}^T U\|_{\text{max}} = |\alpha| .
\]
We have 
\[ \hat{a}^+ = (\hat{a}^\top \hat{a})^{-1} \hat{a}^\top = \frac{1}{\hat{a}^\top \hat{a}} \hat{a}^\top \].

We also have 
\[ \hat{b}^+ = (b^\top \hat{b})^{-1} \hat{b}^\top = \frac{1}{b^\top \hat{b}} \hat{b}^\top = \frac{1}{\alpha(\hat{a}^\top \hat{a})} \hat{a}^\top \].

From optimality of $H$, we have 
\[ \|\hat{a}^+\|_1 \leq \|\hat{b}^+\|_1 \]. Therefore
\[ \frac{1}{|\hat{a}^\top \hat{a}|} \|\hat{a}\|_1 \leq \frac{1}{|\alpha| |\hat{a}^\top \hat{a}|} \|\hat{a}\|_1 \].

So, $|\alpha| \leq 1$, which implies that (6) is satisfied. \hfill \Box

Before moving on to the rank-2 case, we generalize the choice of $\hat{w}, U$ satisfying (4) to the general rank-$r$ case.

**Theorem 7 ([19]).** Let $T$ be an ordered subset of $r$ elements from $\{1, \ldots, n\}$ and $\hat{A} := A[\cdot, T]$ be the $m \times r$ submatrix of an $m \times n$ matrix $A$ formed by columns $T$, and $\text{rank}(\hat{A}) = r$. There exists an $m \times n$ matrix $W$ and a skew-symmetric $m \times m$ matrix $U$ such that
\[ \hat{A}^\top W A^\top + \hat{A}^\top U = Z, \]
where $Z := \text{sign}(\hat{A}^+)$. Furthermore, $\langle A, W \rangle = \|\hat{A}^+\|_1$.

### 3.2 Rank 2

Generally, when $\text{rank}(A) = 2$, we cannot construct a 1-norm minimizing ah-symmetric reflexive generalized inverse based on the column block construction. Even under the condition that $A$ is non-negative, we have the following example:
\[ A = \begin{bmatrix} 1 & 3 & 8 \\ 2 & 2 & 8 \\ 3 & 1 & 8 \end{bmatrix}. \]

Note that $\text{rank}(A) = 2$ because $a_3 = 2a_1 + 2a_2$. We have an ah-symmetric reflexive generalize inverse with 1-norm $\frac{9}{8}$,
\[ H := \begin{bmatrix} -\frac{1}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & 0 & -\frac{1}{4} \\ \frac{1}{24} & \frac{1}{24} & \frac{7}{24} \end{bmatrix}. \]

However, the three ah-symmetric reflexive generalized inverses based on our column block construction have 1-norm $\frac{41}{27}, \frac{41}{27}, \frac{7}{9}$, respectively.
Next, we demonstrate that under an efficiently-checkable technical condition, when \( \text{rank}(A) = 2 \), construction of a 1-norm minimizing ah-symmetric reflexive generalized inverse can be based on the column block construction.

**Theorem 8.** Let \( A \) be an arbitrary \( m \times n \) rank-2 matrix. For any \( j_1, j_2 \in \{1, \ldots, n\} \), with \( j_1 < j_2 \), let \( \hat{A} := [\hat{a}_{j_1}, \hat{a}_{j_2}] \) be the \( m \times 2 \) submatrix of \( A \) formed by columns \( j_1 \) and \( j_2 \). Suppose that \( j_1, j_2 \) are chosen to minimize the 1-norm of \( \hat{H} := \hat{A}^+ \) among all \( m \times 2 \) rank-2 submatrices of \( A \). Every column \( \hat{b} \) of \( A \), can be uniquely written in the basis \( \hat{a}_{j_1}, \hat{a}_{j_2} \), say \( \hat{b} = \alpha \hat{a}_{j_1} + \beta \hat{a}_{j_2} \). Suppose that for each such column \( \hat{b} \) of \( A \), one of the following conditions holds on the associated \( \alpha, \beta \):

(i) \( |\alpha| + |\beta| \leq 1 \);
(ii) \( \hat{H}_{1j} \hat{H}_{2j} \leq 0 \) for \( j = 1, \ldots, m \), and \( \alpha \beta \geq 0 \);
(iii) \( \hat{H}_{1j} \hat{H}_{2j} \geq 0 \) for \( j = 1, \ldots, m \), and \( \alpha \beta \leq 0 \).

Then the \( n \times m \) matrix \( H \) constructed by Theorem 5 based on \( \hat{A} \), is an ah-symmetric reflexive generalized inverse of \( A \) with minimum 1-norm.

**Proof.** We prove a stronger result that our constructed \( H \) is a 1-norm minimizing ah-symmetric generalized inverse. By our construction, \( H \) is reflexive, thus \( H \) is an ah-symmetric reflexive generalized inverse with minimum 1-norm. To establish the minimum 1-norm of our constructed \( H \), we consider the dual pair of linear-optimization problems \( (P_{ah}) \) and \( (D_{ah}) \). As verified in Theorem 5, \( H \) is a feasible solution for \( (P_{ah}) \), and its objective value is \( \|H\|_1 = \|\hat{A}^+\|_1 \) (it also satisfies the nonlinear equations \( (P2) \)).

The objective-function of \( (D_{ah}) \) only depends on the variable \( W \). Feasibility of \( W \) is equivalent to the existence of a skew-symmetric matrix \( U \) satisfying

\[
\|A^TWA^T + A^TU\|_{\text{max}} \leq 1 .
\]

(7)

Next, we are going to construct a dual feasible solution \( W \) with objective value \( \langle A, W \rangle = \|H\|_1 \), then by the weak duality for linear optimization, we prove that \( H \) is optimal to \( (P_{ah}) \).

By Theorem 7, we can choose \( W \) and a skew-symmetric matrix \( U \) such that

\[
\hat{A}^TWA^T + \hat{A}^TU = Z ,
\]

(8)

and then

\[
\langle A, W \rangle = \|\hat{A}^+\|_1 = \|H\|_1 .
\]

The dual constraint \( (7) \) can be written as

\[
\|\hat{A}^TWA^T + \hat{A}^TU\|_{\text{max}} \leq 1 ,
\]

(9)

and

\[
\|\hat{B}^TWA^T + \hat{B}^TU\|_{\text{max}} \leq 1 .
\]

(10)
From (8), we have that (9) is satisfied. To verify (10), without loss of generality, let \((j_1, j_2) = (1, 2)\), and let \(\hat{b} \in \mathbb{R}^m\) be an arbitrary column of \(\hat{B}\) with \(\alpha\) and \(\beta\) such that \(\hat{b} = \alpha \hat{a}_1 + \beta \hat{a}_2\); thus
\[
\begin{align*}
\hat{b}^\top W A^\top + \hat{b}^\top U &= [\alpha \beta](\hat{A}^\top W A^\top + \hat{A}^\top U) = [\alpha \beta]Z = [\alpha \beta]\text{sign}(\hat{H}).
\end{align*}
\]
- For case (i), we have \(\|\hat{b}^\top W A^\top + \hat{b}^\top T\|_{\max} \leq (|\alpha| + |\beta|)\|Z\|_{\max} \leq 1\).
- For case (ii), because \(\hat{H}_{1j} \hat{H}_{2j} < 0\) for \(j = 1, \ldots, m\), we have \(\hat{b}^\top W A^\top + \hat{b}^\top U = (\alpha - \beta)\text{sign}(\hat{H}_{1j})\), and thus \(\|\hat{b}^\top W A^\top + \hat{b}^\top U\|_{\max} = |\alpha - \beta|\). Also we have \(\alpha \beta \geq 0\), so \(\|\hat{b}^\top W A^\top + \hat{b}^\top U\|_{\max} = |\alpha| - |\beta|\).
- For case (iii), because \(\hat{H}_{1j} \hat{H}_{2j} \geq 0\) for \(j = 1, \ldots, m\), we have \(\hat{b}^\top W A^\top + \hat{b}^\top U = (\alpha + \beta)\text{sign}(\hat{H}_{1j})\), and thus \(\|\hat{b}^\top W A^\top + \hat{b}^\top U\|_{\max} = |\alpha + \beta|\). Also we have \(\alpha \beta \leq 0\); so \(\|\hat{b}^\top W A^\top + \hat{b}^\top U\|_{\max} = |\alpha| - |\beta|\).

So to prove the dual feasibility, we only need to show that \(|\alpha| - |\beta| \leq 1\).

Let \(\hat{A}_{b/1} := [\hat{b} \hat{a}_2]\) and \(\hat{\delta}_{ij} := \hat{a}_i^\top \hat{a}_j\), for \(i, j = 1, 2\). We have
\[
\hat{A}^+ = (A^\top A)^{-1}A^\top = \begin{bmatrix}
\hat{a}_1^\top & \hat{a}_2^\top
\end{bmatrix}^{-1} \begin{bmatrix}
\hat{a}_1 & \hat{a}_2
\end{bmatrix}
\]
\[
= \frac{1}{\theta} \begin{bmatrix}
\delta_{12} - \delta_{12} & \hat{\delta}_{12} \\
-\delta_{12} & \delta_{12}
\end{bmatrix}
= \frac{1}{\theta} \begin{bmatrix}
\delta_{22} \hat{a}_1^\top - \delta_{12} a_2^\top & -\delta_{12} a_2^\top + \delta_{22} \hat{a}_1^\top \\
-\delta_{12} a_2^\top + \delta_{22} \hat{a}_1^\top & \delta_{22} \hat{a}_1^\top - \delta_{12} a_2^\top
\end{bmatrix}
\]
where \(\theta = \delta_{11} \delta_{22} - \delta_{12}^2\). We also have
\[
\hat{A}^+_{b/1} = (\hat{A}_{b/1}^\top \hat{A}_{b/1})^{-1} \hat{A}_{b/1}^\top
\]
\[
= \begin{bmatrix}
\hat{\alpha} \hat{a}_1^\top + \beta \hat{a}_2^\top \\
\hat{\alpha} \hat{a}_1^\top + \beta \hat{a}_2^\top
\end{bmatrix}
\]
\[
= \frac{1}{\theta} \begin{bmatrix}
\delta_{12} & -\delta_{12} \\
-\delta_{12} & \delta_{12}
\end{bmatrix}
= \frac{1}{\theta} \begin{bmatrix}
\delta_{22} \hat{a}_1^\top - \delta_{12} a_2^\top & -(\alpha \delta_{12} + \beta \delta_{22}) (\hat{a}_1^\top + \beta \hat{a}_2^\top)
\end{bmatrix}
\]
\[
= \frac{1}{\theta} \begin{bmatrix}
\delta_{22} \hat{a}_1^\top - \delta_{12} a_2^\top & -(\alpha \delta_{12} + \beta \delta_{22}) (\hat{a}_1^\top + \beta \hat{a}_2^\top)
\end{bmatrix}
\]
where
\[ \hat{\theta} = \delta_{22}(\alpha^2\delta_{11} + 2\alpha\beta\delta_{12} + \beta^2\delta_{22}) - (\alpha\delta_{12} + \beta\delta_{22})^2 \\
= \alpha^2\delta_{11}\delta_{22} + 2\alpha\beta\delta_{12}\delta_{22} + \beta^2\delta_{22}^2 - (\alpha^2\delta_{12}^2 + 2\alpha\beta\delta_{12}\delta_{22} + \beta^2\delta_{22}^2) \\
= \alpha^2(\delta_{11}\delta_{22} - \delta_{12}^2) = \alpha^2\theta . \]

From optimality of \( \hat{\theta} \), we have \( \|\hat{A}^+\|_1 \leq \|\hat{A}_{b/1}^+\|_1 \). Therefore

\[
\frac{1}{|\theta|} (\|\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T\|_1 + \| - \delta_{12}\hat{a}_1^T + \delta_{11}\hat{a}_2^T\|_1) \\
\leq \frac{|\alpha|}{|\theta|} (\|\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T\|_1 + \| - \alpha(\delta_{12}\hat{a}_1^T - \delta_{11}\hat{a}_2^T) - \beta(\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T)\|_1) \\
= \frac{1}{|\alpha\theta|} (\|\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T\|_1 + \| \alpha(-\delta_{12}\hat{a}_1^T + \delta_{11}\hat{a}_2^T) + \beta(-\delta_{22}\hat{a}_1^T + \delta_{12}\hat{a}_2^T)\|_1) \\
\leq \frac{1}{|\alpha\theta|} (\|\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T\|_1 + \| \alpha(-\delta_{12}\hat{a}_1^T + \delta_{11}\hat{a}_2^T)\|_1 + \| \beta(\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T)\|_1) \\
= \frac{1}{|\alpha\theta|} (\|\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T\|_1 + |\alpha|\| - \delta_{12}\hat{a}_1^T + \delta_{11}\hat{a}_2^T\|_1 + |\beta|\| - \delta_{22}\hat{a}_1^T + \delta_{12}\hat{a}_2^T\|_1) .
\]

So,

\[
|\alpha| (\|\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T\|_1 + \| - \delta_{12}\hat{a}_1^T + \delta_{11}\hat{a}_2^T\|_1) \\
\leq \|\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T\|_1 + |\alpha| (\|\delta_{12}\hat{a}_1^T - \delta_{11}\hat{a}_2^T\|_1 + |\beta|\|\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T\|_1) ,
\]

and

\[ |\alpha| - |\beta| \leq 1 . \tag{11} \]

Now, considering that \( \|\hat{A}^+\|_1 \leq \|\hat{A}_{b/2}^+\|_1 \), where \( \hat{A}_{b/2} := [\hat{a}_1 \hat{b}] \), we obtain

\[
\frac{1}{|\theta|} (\|\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T\|_1 + \| - \delta_{12}\hat{a}_1^T + \delta_{11}\hat{a}_2^T\|_1) \\
\leq \frac{|\beta|}{|\theta|} (\| - \delta_{12}\hat{a}_1^T + \delta_{11}\hat{a}_2^T\|_1 + \| \alpha(\delta_{12}\hat{a}_1^T - \delta_{11}\hat{a}_2^T) + \beta(\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T)\|_1) \\
\leq \frac{1}{|\beta\theta|} (\| - \delta_{12}\hat{a}_1^T + \delta_{11}\hat{a}_2^T\|_1 + |\alpha|\|\delta_{12}\hat{a}_1^T - \delta_{11}\hat{a}_2^T\|_1 + |\beta|\|\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T\|_1) .
\]

So,

\[
|\beta| (\|\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T\|_1 + \| - \delta_{12}\hat{a}_1^T + \delta_{11}\hat{a}_2^T\|_1) \\
\leq \| - \delta_{12}\hat{a}_1^T + \delta_{11}\hat{a}_2^T\|_1 + |\alpha|\|\delta_{12}\hat{a}_1^T - \delta_{11}\hat{a}_2^T\|_1 + |\beta|\|\delta_{22}\hat{a}_1^T - \delta_{12}\hat{a}_2^T\|_1 ,
\]

and

\[ |\beta| - |\alpha| \leq 1 . \tag{12} \]

From (11) and (12), we have \( |\alpha| - |\beta| \leq 1 . \) \( \square \)
Remark 1. The following example shows that we can allow different cases for each column $\hat{b}$. Let
\[
A := \begin{bmatrix}
2 & 3 & 1 & 5 \\
2 & 3 & 1 & 5 \\
2 & 5 & 2 & 7 \\
\end{bmatrix}.
\]
Then $\{i_1, i_2\} = \{1, 2\}$ minimizes the 1-norm of $\hat{H}$ with $\|\hat{H}\| = 3$. We have
\[
\hat{H} = \begin{bmatrix}
\frac{5}{2} & -\frac{5}{8} & -\frac{3}{4} \\
-\frac{1}{4} & -\frac{3}{8} & \frac{1}{2} \\
\end{bmatrix}
\]
satisfying $\hat{H}_{ij} \hat{H}_{kj} \leq 0$ for $j = 1, 2, 3$, and $\hat{b}_1 = [1, 1, 2] \hat{a} = -\frac{1}{4} \hat{a}_1 + \frac{1}{2} \hat{a}_2$ satisfies only case (i), and $\hat{b}_2 = [5, 5, 7] \hat{a} = \hat{a}_1 + \hat{a}_2$ satisfies only case (ii).

The technical sufficient condition in Theorem 8, while efficiently checkable, may seem rather complicated. But perhaps surprisingly, if $H$ having minimum 1-norm of $\hat{H}$ is an optimal solution to $\hat{(P_{ah})}$ (i.e., a 1-norm minimizing ah-symmetric generalized inverse of $A$, following our column block construction), then the condition is also necessary. So, for rank-2, there is no possibility of further generalizing the condition, in the context of proving the optimality of our chosen column block construction.

**Theorem 9.** Let $A$ be an arbitrary $m \times n$, rank-2 matrix. For any $j_1, j_2 \in \{1, \ldots, n\}$, with $j_1 < j_2$, let $\hat{A} := [\hat{a}_{j_1}, \hat{a}_{j_2}]$ be the $m \times 2$ submatrix of $A$ formed by columns $j_1$ and $j_2$. Suppose that $j_1, j_2$ are chosen to minimize the 1-norm of $\hat{H} := \hat{A}^+$ among all $m \times 2$ rank-2 submatrices of $A$. Suppose that the $n \times m$ matrix $H$ constructed by Theorem 5 based on $\hat{A}$, is an ah-symmetric generalized inverse of $A$ with minimum 1-norm. Every column $\hat{b}$ of $A$, can be uniquely written in the basis $\hat{a}_{j_1}, \hat{a}_{j_2}$, say $\hat{b} = \alpha \hat{a}_{j_1} + \beta \hat{a}_{j_2}$. Then for each such column $\hat{b}$ of $A$, one of the following conditions holds on the associated $\alpha, \beta$:

(i) $|\alpha| + |\beta| \leq 1$;
(ii) $\hat{H}_{ij_1} \hat{H}_{kj_2} \leq 0$ for $j = 1, \ldots, m$, and $\alpha \beta \geq 0$;
(iii) $\hat{H}_{ij_1} \hat{H}_{kj_2} \geq 0$ for $j = 1, \ldots, m$, and $\alpha \beta \leq 0$.

**Proof.** We consider the dual pair of linear-optimization problems $\hat{(P_{ah})}$ and $\hat{(D_{ah})}$. Because $H$ is an optimal solution to $\hat{(P_{ah})}$, by the complementary slackness, we have
\[
\langle J - A^T W A^T - A^T U, H^+ \rangle = 0,
\langle J + A^T W A^T + A^T U, H^- \rangle = 0,
\]
where $H^+ = \max\{H, 0\}$, $H^- = -\min\{H, 0\}$, $W, U$ is an optimal solution to $\hat{(D_{ah})}$ with $U^T = -U$. Along with $H^+, H^- \geq 0$ and dual feasibility, we have
\[(J - A^TWA^T - A^TU)_{ij}H_{ij}^+ = 0,\]
\[(J + A^TWA^T + A^TU)_{ij}H_{ij}^- = 0.\]

Thus,
\[(A^TWA^T + A^TU)_{ij} = \begin{cases} 1, & H_{ij} > 0, \\ -1, & H_{ij} < 0, \\ [-1, 1], & H_{ij} = 0. \end{cases}\]

Without loss of generality, assume that \((j_1, j_2) = (1, 2)\), and \(H = [\hat{H}; 0]\) with \(\hat{H} \in \mathbb{R}^{2 \times m}\). Let
\[(A^TWA^T + A^TU)[\{1, 2\}, :] = \hat{A}^TWA^T + \hat{A}^TU := Z.\]

For every column \(\hat{b}\) of \(A\), \(\hat{b} = \alpha\hat{a}_1 + \beta\hat{a}_2\) because \(\text{rank}(A) = 2\). Hence
\[\hat{b}^TWA^T + \hat{b}^TU = [\alpha \beta](\hat{A}^TWA^T + \hat{A}^TU) = [\alpha \beta]Z,\]

and we have \(\|\hat{b}^TWA^T + \hat{b}^TU\|_{\text{max}} \leq 1\).

- If for any \(j \in \{1, \ldots, m\}\), one of \(\hat{H}_{1j}, \hat{H}_{2j}\) is zero, then \(\hat{H}_{1j}\hat{H}_{2j} = 0\) for \(j = 1, \ldots, m\), thus for any \(\alpha, \beta\), either \(\alpha\beta \geq 0\) or \(\alpha\beta \leq 0\) holds.
- If \(\hat{H}_{1j}\hat{H}_{2j} \leq 0\) for \(j = 1, \ldots, m\), and \(\hat{H}_{1k}\hat{H}_{2k} < 0\) for some \(k\), then \(Z_{1k} = -Z_{2k} \neq 0\), and \(|\alpha - \beta| = |\alpha Z_{1k} + \beta Z_{2k}| \leq 1\), thus \(|\alpha| + |\beta| \leq 1\) or \(\alpha\beta \geq 0\).
- If \(\hat{H}_{1j}\hat{H}_{2j} \geq 0\) for \(j = 1, \ldots, m\), and \(\hat{H}_{1k}\hat{H}_{2k} > 0\) for some \(k\), then \(Z_{1k} = Z_{2k} \neq 0\), and \(|\alpha + \beta| = |\alpha Z_{1k} + \beta Z_{2k}| \leq 1\), thus \(|\alpha| + |\beta| \leq 1\) or \(\alpha\beta \leq 0\).

- Otherwise, we have both \(|\alpha - \beta| \leq 1\) and \(|\alpha + \beta| \leq 1\), which implies \(|\alpha| + |\beta| \leq 1\).

Hence \(\alpha, \beta\) must satisfy one of (i), (ii) and (iii).

\[\square\]

### 4 Conclusion and open questions

When \(A\) is symmetric, a 1-norm minimizing symmetric block solution is a 1-norm minimizing symmetric generalized inverse when (i) \(r = 1\) and when (ii) \(r = 2\) and \(A\) is nonnegative. A 1-norm minimizing column block solution is a 1-norm minimizing ah-symmetric generalized inverse when (i) \(r = 1\) and when (ii) \(r = 2\) and \(A\) satisfies a technical condition. It would be interesting to investigate the approximation ratio achieved by a 1-norm minimizing symmetric block solution and column block solution for general \(r\). Also characterizing optimality conditions beyond the cases that we studied is a nice challenge.

**Acknowledgements** M. Fampa was supported in part by CNPq grant 303898/2016-0. J. Lee was supported in part by ONR grant N00014-17-1-2296.
References

1. A. Bjerhammar, *Application of calculus of matrices to method of least squares with special reference to geodetic calculations*, Trans. Roy. Inst. Tech. Stockholm, (1951), p. 86 pp.

2. S. L. Campbell and C. D. Meyer, *Generalized inverses of linear transformations*, SIAM, 2009.

3. I. Dokmanić and R. Gribonval, *Beyond Moore-Penrose Part I: generalized inverses that minimize matrix norms*. http://arxiv.org/abs/1706.08349, 2017.

4. I. Dokmanić and R. Gribonval, *Beyond Moore-Penrose Part II: the sparse pseudoinverse*. https://hal.inria.fr/hal-01547283/file/pseudo-part2.pdf, 2017.

5. I. Dokmanić, M. Kolundžija, and M. Vetterli, *Beyond Moore-Penrose: sparse pseudoinverse*, in ICASSP 2013, pp. 6526–6530, 2013.

6. A. Dresden, *The fourteenth western meeting of the American Mathematical Society*, Bull. Amer. Math. Soc., 26 (1920), pp. 385–396.

7. M. Fampa and J. Lee, *On sparse reflexive generalized inverses*, Operations Research Letters, 46 (2018), pp. 605–610.

8. M. Fampa, J. Lee, G. Ponte, and L. Xu, *Experimental analysis of local search for sparse reflexive generalized inverses*. https://arxiv.org/abs/2001.03732, 2020.

9. V.K. Fuentes, M. Fampa, and J. Lee, *Sparse pseudoinverses via LP and SDP relaxations of Moore-Penrose*, in CLAIO 2016, 2016, pp. 343–350.

10. V.K. Fuentes, M. Fampa, and J. Lee, *Diving for sparse partially-reflexive generalized inverses*, in Optimization of Complex Systems: Theory, Models, Algorithms and Applications, H. A. Le Thi, H. M. Le, and T. Pham Dinh, eds., Springer, 2020, pp. 89–98.

11. T. Gkountouvas, V. Karakasis, K. Kourtis, G. Goumas, and N. Koziris, *Improving the performance of the symmetric sparse matrix-vector multiplication in multi-core*, in 2013 IEEE 27th Int. Symp. on Parallel and Distributed Proc., 2013, pp. 273–283.

12. G. Golub and C. Van Loan, *Matrix Computations (3rd Ed.)*, Johns Hopkins University Press, Baltimore, MD, USA, 1996.

13. B. K. Natarajan, *Sparse approximate solutions to linear systems*, SIAM journal on computing, 24 (1995), pp. 227–234.

14. R. Penrose, *A generalised inverse for matrices*, Proc. Cambridge Philos. Soc., 51 (1955), pp. 406–413.

15. C. Rao and S. Mitra, *Generalized Inverse Of Matrices and Its Applications*, Probability and Statistics Series, Wiley, 1971.

16. C. Rohde, *Contributions to the theory, computation and application of generalized inverses*, PhD thesis, University of North Carolina, Raleigh, N.C., May 1964. https://www.stat.ncsu.edu/information/library/mimeo.archive/ISMS_1964_392.pdf.

17. K. Schacke, *On the kronecker product*, Master’s thesis, University of Waterloo, (2004).

18. D. Williamson and D. Shmoys, *The Design of Approximation Algorithms*, Cambridge University Press, New York, NY, USA, 1st ed., 2011.

19. L. Xu, M. Fampa, J. Lee, and G. Ponte, *Approximate l1-norm minimization and minimum-rank structured sparsity for various generalized inverses via local search*. https://arxiv.org/abs/1903.05744, 2019.