2015

Some Properties of the Exchange Operator with respect to Structured Matrices defined by Indefinite Scalar Product Spaces

Hanz Martin C. Cheng  
Ateneo de Manila University, hanz_cheng@ateneo.com

Roden Jason David  
Ateneo de Manila University, rjdavid@ateneo.edu

Follow this and additional works at: https://repository.uwyo.edu/ela

Part of the Algebra Commons

Recommended Citation
DOI: https://doi.org/10.13001/1081-3810.2899

This Article is brought to you for free and open access by Wyoming Scholars Repository. It has been accepted for inclusion in Electronic Journal of Linear Algebra by an authorized editor of Wyoming Scholars Repository. For more information, please contact scholcom@uwyo.edu.
SOME PROPERTIES OF THE EXCHANGE OPERATOR WITH RESPECT TO STRUCTURED MATRICES DEFINED BY INDEFINITE SCALAR PRODUCT SPACES∗

HANZ MARTIN C. CHENG† AND RODEN JASON DAVID†

Abstract. The properties of the exchange operator on some types of matrices are explored in this paper. In particular, the properties of \( \text{exc}(A, p, q) \), where \( A \) is a given structured matrix of size \((p + q) \times (p + q)\) and \( \text{exc} : M \times N \times N \rightarrow M \) is the exchange operator are studied. This paper is a generalization of one of the results in [N.J. Higham. J-orthogonal matrices: Properties and generation. SIAM Review, 45:504–519, 2003].

Key words. Scalar product spaces, Structured matrices, Exchange operator.

AMS subject classifications. 15B10, 15B57, 46C20.

1. Introduction. In Mackey, Mackey and Tisseur [1] a scalar product \( \langle x, y \rangle_M \) defined by a non-singular matrix \( M \) is given by

\[
\langle x, y \rangle_M = \begin{cases} 
    x^T M y & \text{for bilinear forms}, \\
    x^* M y & \text{for sesquilinear forms}.
\end{cases}
\]

The adjoint of a matrix \( A \) with respect to the scalar product, denoted by \( A^*_M \) is defined by the property \( \langle Ax, y \rangle_M = \langle x, A^*_M y \rangle_M \). It can be shown that the matrix adjoint is given explicitly by

\[
A^*_M = \begin{cases} 
    M^{-1} A^T & \text{for bilinear forms}, \\
    M^{-1} A^* & \text{for sesquilinear forms}.
\end{cases}
\]

For the above cases, * is used to denote the conjugate transpose. An automorphism group \( G \), Lie algebra \( L \) and Jordan algebra \( J \) is associated with the scalar product. They are namely defined by:

\[
G : = \{ G \in \mathbb{K}^{n \times n} : \langle Gx, Gy \rangle_M = \langle x, y \rangle_M \ \forall x, y \in \mathbb{K}^n \} = \{ G \in \mathbb{K}^{n \times n} : G^* M = G^{-1} \},
\]

Received by the editors on February 8, 2015. Accepted for publication on May 8, 2015. Handling Editor: Michael Tsatsomeros.

†Department of Mathematics, Ateneo de Manila University, Quezon City, Philippines (hcheng@ateneo.edu, rjdavid@ateneo.edu).

304
Some Properties of the Exchange Operator With Respect to Structured Matrices

\[ \mathbb{L} : = \{ L \in \mathbb{K}^{n \times n} : \langle Lx, y \rangle_M = -\langle x, Ly \rangle_M \forall x, y \in \mathbb{K}^n \} = \{ L \in \mathbb{K}^{n \times n} : L^* M = -L \}, \]

\[ \mathbb{J} : = \{ J \in \mathbb{K}^{n \times n} : \langle Jx, y \rangle_M = \langle x, Jy \rangle_M \forall x, y \in \mathbb{K}^n \} = \{ J \in \mathbb{K}^{n \times n} : J^* M = J \}, \]

where \( \mathbb{K} = \mathbb{R} \) or \( \mathbb{C} \).

For this work, the automorphism group, Jordan algebra, and Lie algebra with respect to a scalar product defined by the nonsingular matrix \( M \) will be denoted as \( \mathbb{G}_M, \mathbb{J}_M, \) and \( \mathbb{L}_M \), respectively. Let \( n \) be a positive integer and let \( p, q \) be non-negative integers such that \( p + q = n \). Let

\[ \Sigma_{p,q} = \begin{bmatrix} I_p & O \\ O & -I_q \end{bmatrix}. \]

A matrix \( Q \in \mathbb{R}^{n \times n} \) is said to be \( J \)-orthogonal if and only if \( Q = \Sigma_{p,q} Q^T \Sigma_{p,q} \). A matrix \( A \in \mathbb{R}^{n \times n} \) can be partitioned in the form

\[ A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \]

where \( A_{11} \in \mathbb{R}^{p \times p}, A_{12} \in \mathbb{R}^{p \times q}, A_{21} \in \mathbb{R}^{q \times p}, \) and \( A_{22} \in \mathbb{R}^{q \times q} \). If \( A_{11} \) is nonsingular, then we define the exchange operator as

\[ \text{exc}(A, p, q) = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12} \\ A_{21}A_{11}^{-1} & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix}. \]

In some contexts, \( \text{exc}(A, p, q) \) is also known as the principal pivot transform of \( A \) relative to \( A_{11} \). Let \( B = \text{exc}(A, p, q) \). In Tsatsomeros [3] the exchange operator \( \text{exc}(A, p, q) \) is related to \( A \) such that for every \( x = (x_1^T, x_2^T)^T \) and \( y = (y_1^T, y_2^T)^T \) in \( \mathbb{C}^n \) partitioned conformally to \( A \),

\[ A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \]

if and only if

\[ B \begin{bmatrix} y_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ y_2 \end{bmatrix}. \]

2. Results.

**Theorem 2.1** (Higham [4]). Let \( A \in \mathbb{R}^{n \times n} \). If \( A \) is \( J \)-orthogonal then \( \text{exc}(A, p, q) \) is orthogonal. If \( A \) is orthogonal and \( A_{11} \) is nonsingular, then \( \text{exc}(A, p, q) \) is \( J \)-orthogonal.
First, we note that every $J$-orthogonal matrix is an element of the automorphism group $\mathcal{G}_{\Sigma_{p,q}}$ and every orthogonal matrix is an element of the automorphism group $\mathcal{G}_I$. Thus, Higham’s theorem is equivalent to stating that: if $A$ is an element of the automorphism group $\mathcal{G}_{\Sigma_{p,q}}$, then $\text{exc}(A, p, q) \in \mathcal{G}_I$. Also, if $A \in \mathcal{G}_I$, then it follows that $\text{exc}(A, p, q) \in \mathcal{G}_{\Sigma_{p,q}}$. We generalize the result above onto the automorphism group $\mathcal{G}_M$ and extend it to the Jordan algebra $\mathcal{J}_M$ and the Lie algebra $\mathcal{L}_M$, where $M = \begin{bmatrix} N_1 & O \\ O & N_2 \end{bmatrix}$ is a matrix such that $N_1 \in \mathbb{R}^{p \times p}$ and $N_2 \in \mathbb{R}^{q \times q}$ are both non-singular. Let $\hat{M} = \begin{bmatrix} N_1 & O \\ O & -N_2 \end{bmatrix}$.

**Theorem 2.2.** If $A \in \mathcal{G}_M$, then $\text{exc}(A, p, q) \in \mathcal{G}_{\hat{M}}$. Also, if $A \in \mathcal{G}_{\hat{M}}$, then $\text{exc}(A, p, q) \in \mathcal{G}_M$.

**Theorem 2.3.** If $A \in \mathcal{J}_M$, then $\text{exc}(A, p, q) \in \mathcal{J}_{\hat{M}}$. Also, if $A \in \mathcal{J}_{\hat{M}}$, then $\text{exc}(A, p, q) \in \mathcal{J}_M$.

**Theorem 2.4.** The following statements:

a. $A \in \mathcal{L}_M \implies \text{exc}(A, p, q) \in \mathcal{L}_{\hat{M}}$,

b. $A \in \mathcal{L}_{\hat{M}} \implies \text{exc}(A, p, q) \in \mathcal{L}_M$

are true if and only if $A$ is of the form $\begin{bmatrix} A_{11} & O_{p \times q} \\ O_{q \times p} & A_{22} \end{bmatrix}$.

We only prove Theorem 2.2 since the proof of Theorem 2.3 and Theorem 2.4 can be done similarly.

**Proof of Theorem 2.2.** $A \in \mathcal{G}_M \rightarrow A^{\star M} = A^{-1} or A^{\star M} A = I$. Taking the product, we get

(2.1) $N_1^{-1}A_{11}^*N_1A_{11} + N_1^{-1}A_{21}^*N_2A_{21} = I$,

(2.2) $N_2^{-1}A_{12}^*N_1A_{11} + N_2^{-1}A_{22}^*N_2A_{21} = O$,

(2.3) $N_1^{-1}A_{11}^*N_1A_{12} + N_1^{-1}A_{21}^*N_2A_{22} = O$,

(2.4) $N_2^{-1}A_{12}^*N_1A_{12} + N_2^{-1}A_{22}^*N_2A_{22} = I$.

We need to show that the product $\text{exc}(A, p, q)^{\star M} \text{exc}(A, p, q) = I$.

The $A_{11}$ block is given by

$$
N_1^{-1}A_{11}^*N_1A_{11} - N_1^{-1}A_{11}^*A_{21}^*N_2A_{21}A_{11}^{-1} = N_1^{-1}A_{11}^*N_1A_{11}^{-1} - N_1^{-1}A_{11}^*(N_1A_{11}^{-1} - A_{11}^*N_1) \text{ by } (2.1)
$$

$$
= I.
$$
The $A_{12}$ block is given by

$$\begin{align*}
-N_1^{-1}A_{11}^*N_1A_{11}^{-1}A_{12} + N_1^{-1}A_{11}^*A_{21}N_2(A_{22} - A_{21}A_{11}^{-1}A_{12}) & = -N_1^{-1}A_{11}^*N_1A_{11}^{-1}A_{12} - N_1^{-1}A_{11}^*(-A_{11}^*N_1A_{12}) \\
& \quad + N_1^{-1}A_{11}^*A_{21}N_2A_{21}A_{11}^{-1}A_{12} \text{ by } (2.3) \\
& = -N_1^{-1}A_{11}^*N_1A_{11}^{-1}A_{12} + A_{12} \\
& \quad + N_1^{-1}A_{11}^*(N_1 - A_{11}^*N_1A_{11})A_{11}^{-1}A_{12} \\
& = O.
\end{align*}$$

The $A_{21}$ block is given by

$$\begin{align*}
N_2^{-1}A_{12}^*A_{11}^{-1}A_{11} + N_2^{-1}A_{12}^*N_2A_{21}A_{11}^{-1} - N_2^{-1}A_{12}^*A_{11}^{-1}A_{21}N_2A_{21}A_{11}^{-1} & = N_2^{-1}A_{12}^*A_{11}^{-1} + (-N_2^{-1}A_{12}^*N_1A_{11})A_{11}^{-1} \\
& \quad - N_2^{-1}A_{12}^*A_{11}^{-1}A_{21}N_2A_{21}A_{11}^{-1} \text{ by } (2.2) \\
& = N_2^{-1}A_{12}^*A_{11}^{-1}N_1A_{11}^{-1} - N_2^{-1}A_{12}^*N_1 \\
& \quad - N_2^{-1}A_{12}^*A_{11}^{-1}(N_1 - A_{11}^*N_1A_{11})A_{11}^{-1} \text{ by } (2.1) \\
& = O.
\end{align*}$$

Finally, the $A_{22}$ block is given by

$$\begin{align*}
-N_2^{-1}A_{12}^*A_{11}^{-1}N_1A_{11}^{-1}A_{12} + N_2^{-1}A_{22}N_2A_{21}A_{11}^{-1} - N_2^{-1}A_{12}^*A_{11}^{-1}A_{21}N_2A_{21}A_{11}^{-1}A_{12} & = -N_2^{-1}A_{12}^*A_{11}^{-1}N_1A_{11}^{-1}A_{12} + N_2^{-1}A_{12}^*A_{11}^{-1}A_{21}N_2A_{21}A_{11}^{-1}A_{12} \\
& \quad - N_2^{-1}A_{12}^*A_{11}^{-1}A_{21}N_2A_{21}A_{11}^{-1} - N_2^{-1}A_{12}^*N_2A_{21}A_{11}^{-1}A_{12} \\
& \quad + N_2^{-1}A_{12}^*A_{11}^{-1}(N_1 - A_{11}^*N_1A_{11})A_{11}^{-1}A_{12} \text{ by } (2.1) \\
& = N_2^{-1}A_{12}^*A_{11}^{-1}N_1A_{11}^{-1} + N_2^{-1}A_{12}^*A_{11}^{-1}(-A_{11}^*N_1A_{11}) \\
& \quad - N_2^{-1}(-A_{12}^*N_1A_{11})A_{11}^{-1}A_{12} - N_2^{-1}A_{12}^*N_1A_{12} \\
& \quad \text{ by } (2.2) \text{ and } (2.3) \\
& = I \text{ by } (2.4).
\end{align*}$$

Thus, $\text{exc}(A,p,q) \in \mathbb{G}_{M'}$. The converse can be proven in a similar manner. \(\square\)

Let

$$K = \begin{bmatrix} O & N_1 \\ N_2 & O \end{bmatrix}, \quad \hat{K} = \begin{bmatrix} O & N_1 \\ -N_2 & O \end{bmatrix},$$

where $N_1 \in \mathbb{R}^{p \times p}$ and $N_2 \in \mathbb{R}^{p \times q}$ are both non-singular.

**Theorem 2.5.** If $A \in \mathbb{G}_K$, then $\text{exc}(A,p,q) \in \mathbb{J}_K$.

**Theorem 2.6.** If $A \in \mathbb{J}_K$, then $\text{exc}(A,p,q) \in \mathbb{G}_K$. 

We only prove Theorem 2.2, since the proof of Theorem 2.1 can be done similarly.

**Proof of Theorem 2.2.** $A \in G_K \implies A^\yectos = A^{-1}$ or $A^\yectos A = I$. Taking the product, we get

(2.5) \[ N_2^{-1}A_{22}N_2A_{11} + N_2^{-1}A_{12}^{-1}N_1A_{21} = I, \]
(2.6) \[ N_1^{-1}A_{22}N_2A_{11} + N_1^{-1}A_{11}^{-1}N_1A_{21} = 0, \]
(2.7) \[ N_2^{-1}A_{22}N_2A_{12} + N_2^{-1}A_{12}^{-1}N_1A_{22} = 0, \]
(2.8) \[ N_1^{-1}A_{21}^{-1}N_1A_{21} + N_1^{-1}A_{11}^{-1}N_1A_{22} = I. \]

We need to show that $K^{-1}exc(A, p, q)^*K = exc(A, p, q)$. Consider

\[
\begin{bmatrix}
O & -N_2^{-1} \\
N_1^{-1} & O
\end{bmatrix}
\begin{bmatrix}
A_{11}^{-1} & O \\
-A_{12}^{-1}A_{11}^{-1} & A_{22}^{-1} - A_{12}^{-1}A_{11}^{-1}A_{21}^{-1}
\end{bmatrix}
\begin{bmatrix}
O & N_1 \\
-N_2 & O
\end{bmatrix}
\]

Using equations (2.5) - (2.8), we have

\[
N_2^{-1}A_{22}N_2 - N_2^{-1}A_{12}^{-1}A_{11}^{-1}A_{21}^{-1}N_2
= A_{11}^{-1} - N_2^{-1}A_{12}^{-1}N_1A_{21}A_{11}^{-1} - N_2^{-1}A_{12}^{-1}A_{11}^{-1}A_{21}^{-1}N_2
\]

Also,

\[
N_2^{-1}A_{12}^{-1}A_{11}^{-1}N_1 = N_2^{-1}A_{12}^{-1}(A_{11}^{-1}A_{21}^{-1}N_2A_{12} + N_1A_{22})
= N_2^{-1}A_{12}^{-1}(A_{11}^{-1}A_{21}^{-1}N_2A_{12} + N_2^{-1}A_{12}^{-1}A_{12}^{-1}N_1A_{22})
= N_2^{-1}A_{12}^{-1}(A_{11}^{-1}A_{21}^{-1}N_2A_{12} - N_2^{-1}A_{12}^{-1}N_1A_{22})
= N_2^{-1}A_{12}^{-1}(-N_1A_{21}A_{11}^{-1}A_{12} - (A_{11}^{-1} - N_2^{-1}A_{12}^{-1}A_{12}^{-1})A_{12})
\]

and

\[
-N_1^{-1}A_{11}^{-1}A_{21}^{-1}N_2 = -N_1^{-1}(-N_1A_{21}A_{11}^{-1})
= A_{21}A_{11}^{-1}.
\]

Finally,

\[
N_1^{-1}A_{11}^{-1}N_1 = N_1^{-1}(N_1A_{22} + A_{11}^{-1}A_{21}^{-1}N_2A_{12})
= A_{22} + N_1^{-1}(A_{11}^{-1}A_{21}^{-1}N_2A_{12})
= A_{22} + N_1^{-1}(-N_1A_{21}A_{11}^{-1})A_{12}
= A_{22} - A_{21}A_{11}^{-1}A_{12}.
\]
Thus,
\[
\begin{bmatrix}
N_2^{-1}A_{22}N_2 - N_2^{-1}A_{12}^{-1}A_{11}^{-1}A_{21}N_2 & N_2^{-1}A_{12}^{-1}A_{11}^{-1}N_1 \\
-N_1^{-1}A_{11}^{-1}A_{21}^{-1}N_2 & N_1^{-1}A_{11}^{-1}N_1
\end{bmatrix}
= \begin{bmatrix}
A_{11}^{-1} & -A_{12}^{-1}A_{12} \\
A_{21}A_{11}^{-1} & A_{22} - A_{21}A_{11}^{-1}A_{12}
\end{bmatrix}
\]
or equivalently, \( \hat{K}^{-1}\text{exc}(A, p, q) = \text{exc}(A, p, q) \).

3. Application (The Schur complement method). Matrices that can be directly factored are limited in size due to large memory requirements. Thus, the difficulty of solving large-scale linear systems of equations using standard techniques arises. Preconditioned iterative solvers require less memory, but often suffer from slow convergence. One of the ideas that have been developed to address these problems is called the Schur complement method. Several parallel hybrid solvers have been developed based on the above mentioned idea [2].

In the Schur complement method, instead of solving the system
\[
\begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
= \begin{bmatrix}
y_1 \\
y_2
\end{bmatrix}
\]
directly, we consider
\[
\begin{bmatrix}
A_{11}^{-1} & -A_{12}^{-1}A_{12} \\
A_{21}A_{11}^{-1} & A_{22} - A_{21}A_{11}^{-1}A_{12}
\end{bmatrix}
\begin{bmatrix}
y_1 \\
x_2
\end{bmatrix}
= \begin{bmatrix}
x_1 \\
y_2
\end{bmatrix}
\]
and solve the system of equations
\[
A_{11}^{-1}y_1 - A_{11}^{-1}A_{12}x_2 = x_1, \\
A_{21}A_{11}^{-1}y_1 + (A_{22} - A_{21}A_{11}^{-1}A_{12})x_2 = y_2.
\]
In solving the above system, we need to compute \( (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \).

Suppose that \( p \) and \( q \) are integers such that \( q \leq p \) and
\[
M = \begin{bmatrix}
N_1 & O \\
O & N_2
\end{bmatrix}
\]
is a matrix such that \( N_1 \in \mathbb{R}^{p \times p} \) and \( N_2 \in \mathbb{R}^{q \times q} \) are both non-singular. Let
\[
\tilde{M} = \begin{bmatrix}
N_1 & O \\
O & -N_2
\end{bmatrix}.
\]
Suppose \( A \in G_M \). Since \( \text{exc}(A, p, q) \in G_{\tilde{M}} \) by Theorem 2.2 we know that \( (A_{22} - A_{21}A_{11}^{-1}A_{12})N_2^{-*}(-A_{11}^{-1}A_{12}) = A_{21}A_{11}^{-1}N_1^{-*}A_{11}^{-1} \). Suppose further that \( \text{rank}(A_{21}) = \)
Then we may compute
\[
(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}
\]
explicitly.

To compute \((A_{21}A_{11}^{-1}A_{12}^*)^{-1}\), we perform singular value decomposition on both \(A_{11}\) and \(A_{21}\). Write \(A_{11} = U\Sigma V^*\) and \(A_{21} = \hat{U}\hat{\Sigma}\hat{V}^*\), where \(U, V, \hat{U}\), and \(\hat{V}\) are unitary and

\[
\Sigma = \begin{bmatrix}
\sigma_1 & 0 & \ldots & 0 \\
0 & \sigma_2 & \ldots & 0 \\
0 & 0 & \ddots & 0 \\
0 & 0 & \ldots & \sigma_p
\end{bmatrix}, \quad \hat{\Sigma} = \begin{bmatrix}
\hat{\sigma}_1 & 0 & \ldots & 0 \\
0 & \hat{\sigma}_2 & \ldots & 0 \\
0 & 0 & \ddots & 0 \\
0 & 0 & \ldots & \hat{\sigma}_q
\end{bmatrix}
\]
such that \(\{\sigma_1, \sigma_2, \ldots, \sigma_p\}\) and \(\{\hat{\sigma}_1, \hat{\sigma}_2, \ldots, \hat{\sigma}_q\}\) are the singular values of \(A_{11}\) and \(A_{21}\), respectively. Then,

\[
A_{21}A_{11}^{-1}A_{12}^* = \hat{U}\hat{\Sigma}V^*\hat{\Sigma}^{-1}U^*U\Sigma^{-1}V^*V\hat{\Sigma}^*\hat{V}^*
\]

Finally,

\[
(A_{21}A_{11}^{-1}A_{12}^*)^{-1} = \hat{U}(\hat{\Sigma}^*)^*\hat{V}^*V\Sigma V^*\hat{V}\hat{\Sigma}^*\hat{U}^*
\]

where

\[
\hat{\Sigma}^+ = \begin{bmatrix}
\frac{1}{\hat{\sigma}_1} & 0 & \ldots & 0 \\
0 & \frac{1}{\hat{\sigma}_2} & \ldots & 0 \\
0 & 0 & \ddots & 0 \\
0 & 0 & \ldots & \frac{1}{\hat{\sigma}_q}
\end{bmatrix}
\]

Thus, given \(A \in \mathbb{C}_M\), the solution of \((3.11)\) is given explicitly by

\[
x_2 = (\Sigma V^*\hat{V}\hat{\Sigma}^*\hat{U}^*)^*(\Sigma V^*\hat{V}\hat{\Sigma}^*\hat{U}^*)(y_2 - A_{21}A_{11}^{-1}y_1),
\]

\[
x_1 = A_{11}^{-1}y_1 - A_{11}^{-1}A_{12}x_2.
\]
REFERENCES


