Abstract

Let $R = P \text{diag}(\gamma_0 I_{m_0}, \gamma_1 I_{m_1}, \ldots, \gamma_{k-1} I_{m_{k-1}})P^{-1} \in \mathbb{C}^{m \times m}$ and $S_{\sigma} = Q \text{diag}(\gamma_0(0) I_{n_0}, \gamma_1(1) I_{n_1}, \ldots, \gamma_{k-1}(k-1) I_{n_{k-1}})Q^{-1} \in \mathbb{C}^{n \times n}$, where $m_0 + m_1 + \cdots + m_{k-1} = m$, $n_0 + n_1 + \cdots + n_{k-1} = n$, $\gamma_0, \gamma_1, \ldots, \gamma_{k-1}$ are distinct complex numbers, and $\sigma: \mathbb{Z}_k \rightarrow \mathbb{Z}_k = \{0, 1, \ldots, k-1\}$. We say that $A \in \mathbb{C}^{m \times n}$ is $(R, S_{\sigma})$-commutative if $RA = AS_{\sigma}$. We characterize the class of $(R, S_{\sigma})$-commutative matrices and extend results obtained previously for the case where $\gamma_\ell = e^{2\pi i \ell/k}$ and $\sigma(\ell) = \alpha \ell + \mu \pmod{k}$, $0 \leq \ell \leq k-1$, with $\alpha, \mu \in \mathbb{Z}_k$. Our results are independent of $\gamma_0, \gamma_1, \ldots, \gamma_{k-1}$, so long as they are distinct; i.e., if $RA = AS_{\sigma}$ for some choice of $\gamma_0, \gamma_1, \ldots, \gamma_{k-1}$ (all distinct), then $RA = AS_{\sigma}$ for arbitrary of $\gamma_0, \gamma_1, \ldots, \gamma_{k-1}$.

MSC: 15A09; 15A15; 15A18; 15A99

Keywords: Commute; Eigenvalue problem; Least Squares problem; Moore–Penrose Inverse; $(R, S_{\sigma})$-commutative; Singular value decomposition

*e-mail: wtrench@trinity.edu*
1 Introduction

A matrix $A = [a_{rs}]_{r,s=0}^{n-1} \in \mathbb{C}^{n \times n}$ is said to be centrosymmetric if

$$a_{n-r-1,n-s-1} = a_{rs}, \quad 0 \leq r, s \leq n - 1,$$

or centro-skewsymmetric if

$$a_{n-r-1,n-s-1} = -a_{rs}, \quad 0 \leq r, s \leq n - 1.$$

The study of such matrices is facilitated by the observation that $A$ is centrosymmetric (centro-skewsymmetric) if and only if $JA = AJ$ ($JA = -AJ$), where $J$ is the flip matrix, with ones on the secondary diagonal and zeroes elsewhere. Several authors [2, 3, 4, 5, 8, 10, 13, 25] used this observation to show that centrosymmetric and centro-skewsymmetric matrices can be written as $A = PCP^{-1}$, where $P$ diagonalizes $J$ and $C$ has a useful block structure. We will discuss this further in Example 3.

Following this idea, other authors [6, 11, 12, 14, 24] considered matrices satisfying $RA = AR$ or $RA = -AR$, where $R$ is a nontrivial involution; i.e., $R = R^{-1} \neq \pm I$. We continued this line of investigation in [15, 16, 17, 19], and extended it in [18, 20], defining $A \in \mathbb{C}^{m \times n}$ to be $(R, S)$-symmetric ($(R, S)$-skew symmetric) if $RA = AS$ ($RA = -AS$), where $R \in \mathbb{C}^{m \times m}$ and $S \in \mathbb{C}^{n \times n}$ are nontrivial involutions. We showed that a matrix $A$ with either of these properties can be written as $A = PCQ^{-1}$, where $P$ and $Q$ diagonalize $R$ and $S$ respectively and $C$ has a useful block form.

Chen [7] and Fasino [9] studied matrices $A \in \mathbb{C}^{n \times n}$ such that $RAR^* = \zeta^\mu A$, where $R$ is a unitary matrix that satisfies $R^k = I$ for some $k \leq n$ and $\zeta = e^{2\pi i/k}$. In [21] we studied matrices $A \in \mathbb{C}^{m \times n}$ such that $RA = \zeta^\mu AS$, where

$$R = P \text{ diag } \left( I_{m_0}, \zeta I_{m_1}, \ldots, \zeta^{k-1} I_{m_{k-1}} \right) P^{-1},$$  \hspace{1cm} (1)

$$S = Q \text{ diag } \left( I_{n_0}, \zeta I_{n_1}, \ldots, \zeta^{k-1} I_{n_{k-1}} \right) Q^{-1},$$  \hspace{1cm} (2)

$$m_0 + m_1 + \cdots + m_{k-1} = m, \quad n_0 + n_1 + \cdots + n_{k-1} = n, \hspace{1cm} (3)$$

and $\alpha, \mu \in \mathbb{Z}_k = \{0, 1, \ldots, k - 1\}$.

Finally, motivated by a problem concerning unilevel block circulants [22], in [23] we considered matrices $A \in \mathbb{C}^{m \times n}$ such that $RA = \zeta^\mu AS^\alpha$, with $\alpha, \mu \in \mathbb{Z}_k$. We called such matrices $(R, S, \alpha, \mu)$-symmetric, and showed that $A$ has this property if and only if

$$A = \sum_{\ell=0}^{k-1} P_{\alpha \ell + \mu (\text{mod } k)} F_\ell \hat{Q}_\ell \quad \text{with} \quad F_\ell \in \mathbb{C}^{\alpha \ell + \mu (\text{mod } k) \times n_\ell}, \quad 0 \leq \ell \leq k - 1, \hspace{1cm} (4)$$

which has useful computational and theoretical applications. ($P_0, \ldots, P_{k-1}$ and $\hat{Q}_0, \ldots, \hat{Q}_{k-1}$ are defined in Section 2, specifically, (7)-(10).) The class of $(R, S, \alpha, \mu)$-symmetric matrices includes, for example, centrosymmetric, skew-centrosymmetric,
We can now write

\[ R = P \text{ diag}(\gamma_0 I_{m_0}, \gamma_1 I_{m_1}, \ldots, \gamma_{k-1} I_{m_{k-1}}) P^{-1} \in \mathbb{C}^{m \times m} \quad (5) \]

and

\[ S = Q \text{ diag}(\gamma_0 I_{n_0}, \gamma_1 I_{n_1}, \ldots, \gamma_{k-1} I_{n_{k-1}}) Q^{-1} \in \mathbb{C}^{n \times n}, \quad (6) \]

where \( \gamma_0, \gamma_1, \ldots, \gamma_{k-1} \) are distinct complex numbers, except when there is an explicit statement to the contrary. We define

\[ R_\sigma = P \text{ diag}(\gamma_{\sigma(0)} I_{m_0}, \gamma_{\sigma(1)} I_{m_1}, \ldots, \gamma_{\sigma(k-1)} I_{m_{k-1}}) P^{-1} \]

and

\[ S_\sigma = Q \text{ diag}(\gamma_{\sigma(0)} I_{n_0}, \gamma_{\sigma(1)} I_{n_1}, \ldots, \gamma_{\sigma(k-1)} I_{n_{k-1}}) Q^{-1}, \]

where \( \sigma : \mathbb{Z}_k \rightarrow \mathbb{Z}_{k} \).

We can partition

\[ P = \begin{bmatrix} P_0 & P_1 & \cdots & P_{k-1} \end{bmatrix}, \quad Q = \begin{bmatrix} Q_0 & Q_1 & \cdots & Q_{k-1} \end{bmatrix} \quad (7) \]

\[ P^{-1} = \begin{bmatrix} \hat{P}_0 \\ \hat{P}_1 \\ \vdots \\ \hat{P}_{k-1} \end{bmatrix} \quad \text{and} \quad Q^{-1} = \begin{bmatrix} \hat{Q}_0 \\ \hat{Q}_1 \\ \vdots \\ \hat{Q}_{k-1} \end{bmatrix}, \quad (8) \]

where

\[ P_r \in \mathbb{C}^{m \times m}, \quad \hat{P}_r \in \mathbb{C}^{m \times m}, \quad \hat{P}_r P_s = \delta_{rs} I_{m_r}, \quad 0 \leq r, s \leq k - 1, \quad (9) \]

\[ Q_r \in \mathbb{C}^{n \times n}, \quad \hat{Q}_r \in \mathbb{C}^{n \times n}, \quad \text{and} \quad \hat{Q}_r Q_s = \delta_{rs} I_{n_r}, \quad 0 \leq r, s \leq k - 1. \quad (10) \]

We can now write

\[ R = \sum_{\ell=0}^{k-1} \gamma_\ell \hat{P}_\ell, \quad R_\sigma = \sum_{\ell=0}^{k-1} \gamma_{\sigma(\ell)} \hat{P}_\ell, \quad (11) \]

\[ S = \sum_{\ell=0}^{k-1} \gamma_\ell \hat{Q}_\ell, \quad \text{and} \quad S_\sigma = \sum_{\ell=0}^{k-1} \gamma_{\sigma(\ell)} \hat{Q}_\ell. \quad (12) \]
Definition 1 In general, if $U \in \mathbb{C}^{m \times m}$, $V \in \mathbb{C}^{n \times n}$, and $A \in \mathbb{C}^{m \times n}$, we say that $A$ is $(U, V)$-commutative if $UA = AV$. In particular, we say that $A \in \mathbb{C}^{m \times n}$ is $(R, S_\sigma)$-commutative if $RA = AS_\sigma$. If $\sigma$ is the identity (i.e., $RA = AS$), we say that $A$ is $(R, S)$-commutative. If $A, R \in \mathbb{C}^{n \times n}$ and $RA = AR$, we say – as usual – that $A$ commutes with $R$.

3 Necessary and sufficient conditions for $(R, S_\sigma)$-commutativity

Theorem 1 $A \in \mathbb{C}^{m \times n}$ is $(R, S_\sigma)$-commutative if and only if

$$A = P \left( [C_{rs}]_{r,s=0}^{k-1} \right) Q^{-1}, \quad \text{where} \quad C_{rs} \in \mathbb{C}^{m \times n},$$

and

$$C_{rs} = 0 \quad \text{if} \quad r \neq \sigma(s), \quad 0 \leq r, s \leq k - 1.$$  \hfill (14)

Proof. Any $A \in \mathbb{C}^{m \times n}$ can be written as in (13) with $C = P^{-1}AQ$ partitioned as indicated. If

$$D = \text{diag} \left( \gamma_0 I_{m_0}, \gamma_1 I_{m_1}, \ldots, \gamma_{k-1} I_{m_{k-1}} \right)$$

and

$$D_\sigma = \text{diag} \left( \gamma_{\sigma(0)} I_{n_0}, \gamma_{\sigma(1)} I_{n_1}, \ldots, \gamma_{\sigma(k-1)} I_{n_{k-1}} \right),$$

then

$$RA = (PDP^{-1})(PCQ^{-1}) = PDCQ^{-1} = P \left( [\gamma_r C_{rs}]_{r,s=0}^{k-1} \right) Q^{-1}$$

and

$$AS_\sigma = (PCQ^{-1})(QD_\sigma Q^{-1}) = PCD_\sigma Q^{-1} = P \left( [\gamma_{\sigma(s)} C_{rs}]_{r,s=0}^{k-1} \right) Q^{-1}.$$  \hfill (15)

Therefore $RA = AS_\sigma$ if and only if $[\gamma_r - \gamma_{\sigma(s)}] C_{rs} = 0$, $0 \leq r, s \leq k - 1$, which is equivalent to (14), since $\gamma_0, \gamma_1, \ldots, \gamma_{k-1}$ are distinct. \hfill $\square$

The following theorem is a convenient reformulation of Theorem 1.

Theorem 2 $A \in \mathbb{C}^{m \times n}$ is $(R, S_\sigma)$-commutative if and only if

$$A = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_\ell \hat{Q}_\ell \quad \text{with} \quad F_\ell \in \mathbb{C}^{m_{\sigma(\ell)} \times n_\ell}, \quad 0 \leq \ell \leq k - 1,$$  \hfill (15)

in which case

$$F_\ell = \hat{P}_{\sigma(\ell)} AQ_\ell, \quad 0 \leq \ell \leq k - 1,$$  \hfill (16)

and

$$RA = AS_\sigma = \sum_{\ell=0}^{k-1} \gamma_{\sigma(\ell)} P_{\sigma(\ell)} F_\ell \hat{Q}_\ell$$  \hfill (17)

for arbitrary $\gamma_0, \gamma_1, \ldots, \gamma_{k-1}$.
(R, S₀)-commutative matrices

PROOF. From (13), an arbitrary \( A \in \mathbb{C}^{m \times n} \) can be written as

\[
A = \sum_{s=0}^{k-1} \sum_{r=0}^{k-1} P_r C_r s \hat{Q}_l.
\]  

From Theorem 1, \( A \) is \((R, S₀)\)-commutative if and only if \( C_r s = 0 \) if \( r \neq \sigma(s) \), in which case (18) reduces to (15) with \( F_l = C_{\sigma(l), l} \in \mathbb{C}^{m_{\sigma(l)} \times n_l} \). From (10) and (15), \( AQ_l = P_{\sigma(l)} F_l \), \( 0 \leq \ell \leq k-1 \), so (9) with \( r = \sigma(\ell) \) implies (16). Eqns. (9)–(12) and (15) imply (17).

Example 1 If \( \sigma \) is the permutation

\[
\sigma = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 4 & 0 & 2 & 5 \end{pmatrix} = (0, 1, 3)(2, 4)(5),
\]

then (15) becomes

\[
A = P_1 F_0 \hat{Q}_0 + P_3 F_1 \hat{Q}_1 + P_4 F_2 \hat{Q}_2 + P_0 F_3 \hat{Q}_3 + P_2 F_4 \hat{Q}_4 + P_5 F_5 \hat{Q}_5,
\]

with

\[
F_0 \in \mathbb{C}^{m_1 \times n_0}, \quad F_1 \in \mathbb{C}^{m_3 \times n_1}, \quad F_2 \in \mathbb{C}^{m_4 \times n_2},
\]
\[
F_3 \in \mathbb{C}^{m_0 \times n_3}, \quad F_4 \in \mathbb{C}^{m_2 \times n_4}, \quad F_5 \in \mathbb{C}^{m_5 \times n_5},
\]

and

\[
RA = AS_\sigma = \gamma_1 P_1 F_0 \hat{Q}_0 + \gamma_2 P_3 F_1 \hat{Q}_1 + \gamma_4 P_4 F_2 \hat{Q}_2 + \gamma_0 P_0 F_3 \hat{Q}_3 + \gamma_2 P_2 F_4 \hat{Q}_4 + \gamma_5 P_5 F_5 \hat{Q}_5
\]

for arbitrary \( \gamma_0, \ldots, \gamma_5 \).

Example 2 If

\[
\sigma = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 0 & 1 & 2 & 0 \end{pmatrix}
\]

(which is not a permutation), then (15) becomes

\[
A = P_2 F_0 \hat{Q}_0 + P_1 F_1 \hat{Q}_1 + P_0 F_2 \hat{Q}_2 + P_1 F_3 \hat{Q}_3 + P_2 F_4 \hat{Q}_4 + P_0 F_5 \hat{Q}_5,
\]

with

\[
F_0 \in \mathbb{C}^{m_2 \times n_0}, \quad F_1 \in \mathbb{C}^{m_1 \times n_1}, \quad F_2 \in \mathbb{C}^{m_0 \times n_2},
\]
\[
F_3 \in \mathbb{C}^{m_1 \times n_3}, \quad F_4 \in \mathbb{C}^{m_2 \times m_4}, \quad F_5 \in \mathbb{C}^{m_0 \times n_5},
\]

and

\[
RA = AS_\sigma = \gamma_1 P_2 F_0 \hat{Q}_0 + \gamma_1 P_1 F_1 \hat{Q}_1 + \gamma_0 P_0 F_2 \hat{Q}_2 + \gamma_1 P_1 F_3 \hat{Q}_3 + \gamma_2 P_2 F_4 \hat{Q}_4 + \gamma_0 P_0 F_5 \hat{Q}_5
\]

for arbitrary \( \gamma_0, \ldots, \gamma_5 \).
Example 3  All results obtained by assuming that $R$ and $S$ are involutions (and therefore have eigenvalues 1 and $-1$) can just as well be obtained by assuming only that $R$ and $S$ have the same two distinct eigenvalues, with possibly different multiplicities. The original idea in this area of research has its origins in the observation that $A$ is centrosymmetric (skew-centrosymmetric) if and only if $AJ = JA$ ($AJ = -JA$). Since $J^2 = I$, these conditions can just as well be written as $JAJ = A$ ($JAJ = -A$); however, this and the invertibility of $J$ are irrelevant. To illustrate this, suppose $n = 2r$, in which case

$$J = \begin{bmatrix} P_0 & P_1 \\ \end{bmatrix} \begin{bmatrix} I_r & 0 \\ 0 & -I_r \\ \end{bmatrix} \begin{bmatrix} P_0^T \\ P_1^T \\ \end{bmatrix}$$

(i.e., $\widehat{P}_0 = P_0^T$ and $\widehat{P}_1 = P_1^T$), where

$$P_0 = \frac{1}{\sqrt{2}} \begin{bmatrix} I_r \\ J_r \\ \end{bmatrix} \quad \text{and} \quad P_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} I_r \\ -J_r \\ \end{bmatrix}.$$ 

Starting from this, it can be shown $AJ = JA$ (or, equivalently, $A$ is centrosymmetric) if and only if

$$A = \begin{bmatrix} P_0 & P_1 \\ \end{bmatrix} \begin{bmatrix} B_0 & 0 \\ 0 & B_1 \\ \end{bmatrix} \begin{bmatrix} P_0^T \\ P_1^T \\ \end{bmatrix} = P_0 B_0 P_0^T + P_1 B_1 P_1^T$$

(19)

with $B_0, B_1 \in \mathbb{C}^{r \times r}$. However, Theorem 2 implies that $A$ has the form (19) if $RA = AR$ for some $R$ of the form

$$R = \begin{bmatrix} P_0 & P_1 \\ \end{bmatrix} \begin{bmatrix} \gamma_0 I_r & 0 \\ 0 & \gamma_1 I_r \\ \end{bmatrix} \begin{bmatrix} P_0^T \\ P_1^T \\ \end{bmatrix}$$

with $\gamma_0 \neq \gamma_1$, in which case

$$RA = AR = \gamma_0 P_0 B_0 P_0^T + \gamma_1 P_1 B_1 P_1^T.$$ 

for arbitrary $\gamma_0$ and $\gamma_1$.

According to the classical theorem, $AJ = -JA$ (or, equivalently, $A$ is skew-centrosymmetric) if and only if

$$A = \begin{bmatrix} P_0 & P_1 \\ \end{bmatrix} \begin{bmatrix} 0 & C_1 \\ C_0 & 0 \\ \end{bmatrix} \begin{bmatrix} P_0^T \\ P_1^T \\ \end{bmatrix} = P_1 C_0 \widehat{P}_0 + P_0 C_1 \widehat{P}_1$$

(20)

with $C_0, C_1 \in \mathbb{C}^{r \times r}$. Now let $\sigma(0) = 1$ and $\sigma(1) = 0$, so

$$R_\sigma = \begin{bmatrix} P_0 & P_1 \\ \end{bmatrix} \begin{bmatrix} \gamma_1 I_r & 0 \\ 0 & \gamma_0 I_r \\ \end{bmatrix} \begin{bmatrix} P_0^T \\ P_1^T \\ \end{bmatrix}.$$ 

Theorem 2 implies that $A$ has the form (20) if and only if $RA = AR_\sigma$ for some $\gamma_0$ and $\gamma_1$ with $\gamma_0 \neq \gamma_1$, in which case

$$RA = AR_\sigma = \gamma_1 P_1 C_0 \widehat{P}_0 + \gamma_0 P_0 C_1 \widehat{P}_1$$

for all $\gamma_0$ and $\gamma_1$. 
(R, S_0)-commutative matrices

Example 4 Let \( R = [\delta_{r,s-1}(\text{mod } k)]_{r,s=0}^{k-1} \), which is the 1-circulant with first row

\[
\begin{bmatrix}
0 & 1 & 0 & \cdots & 0
\end{bmatrix}.
\]

By the Ablow-Brenner theorem [1], \( C \in \mathbb{C}^{k \times k} \) is an \( \alpha \)-circulant \( C = [c_{j-\alpha r(\text{mod } k)}]_{r,s=0}^{k-1} \) if and only if \( RC = CR^\alpha \). Since

\[
R = P \ \text{diag}(1, \zeta, \zeta^2, \ldots, \zeta^{k-1}) \ P^* \]

where

\[
P = \begin{bmatrix}
p_0 & p_1 & \cdots & p_{k-1}
\end{bmatrix}
\]

with \( p_\ell = \frac{1}{\sqrt{k}} \begin{bmatrix}
1 \\
\zeta^\ell \\
\zeta^{2\ell} \\
\vdots \\
\zeta^{(k-1)\ell}
\end{bmatrix}, \quad 0 \leq \ell \leq k-1,
\]

and

\[
R^\alpha = P \ \text{diag}(1, \zeta^\alpha, \zeta^{2\alpha}, \ldots, \zeta^{(k-1)\alpha}) \ P^*.
\]

the Ablow-Brenner theorem can be interpreted to mean that \( C \) is \( (R, R_0) \)-commutative with \( \sigma(\ell) = \alpha \ell \) (mod \( k \)), \( 0 \leq \ell \leq k-1 \). Therefore Theorem 2 implies that

\[
C = \sum_{\ell=0}^{k-1} p_{\alpha \ell(\text{mod } k)} \ f_{\ell} \ f_{\ell}^* ,
\]

where \( f_0, f_1, \ldots, f_{k-1} \) are scalars. As a matter of fact, if

\[
R = P \ \text{diag}(\gamma_0, \gamma_1, \ldots, \gamma_{k-1}) \ P^* \]

with arbitrary \( \gamma_0, \gamma_1, \ldots, \gamma_{k-1} \), then

\[
RC = CR_\sigma = \sum_{\ell=0}^{k-1} \gamma_{\alpha \ell(\text{mod } k)} \ p_{\alpha \ell(\text{mod } k)} \ f_{\ell} \ f_{\ell}^*.
\]

Example 5 Let \( R \) and \( S \) be as in (1) and (2) and let \( \sigma(\ell) = \alpha \ell + \mu \) (mod \( k \)), so

\[
S_\sigma = Q \ \text{diag} \left( \zeta^\mu I_{m_0}, \zeta^{\alpha+\mu} I_{m_1}, \ldots, \zeta^{(k-1)\alpha+\mu} I_{m_{k-1}} \right) \ Q^{-1}.
\]

Then the \( (R, S, \alpha, \mu) \)-symmetric matrix \( A \) in (4) is \( (R, S_0) \)-commutative. More generally, if \( R \) and \( S \) are as in (5) and (6) and \( \sigma(\ell) = \alpha \ell + \mu \) (mod \( k \)), then

\[
RA = AS_\sigma = \sum_{\ell=0}^{k-1} \gamma_{\alpha \ell + \mu(\text{mod } k)} \ p_{\alpha \ell + \mu(\text{mod } k)} \ F_{\ell} \ \widehat{Q}_{\ell}
\]

for arbitrary \( \gamma_0, \gamma_1, \ldots, \gamma_{k-1} \).
Renaming the variables in Theorem 2 yields the following theorem.

**Theorem 3** If \( \rho : \mathbb{Z}_k \rightarrow \mathbb{Z}_k \), then \( B \in \mathbb{C}^{n \times m} \) is \((S, R)\)-commutative if and only if

\[
B = \sum_{\ell=0}^{k-1} Q_{\rho(\ell)} G_{\ell} \hat{P}_{\ell} \quad \text{with} \quad G_{\ell} \in \mathbb{C}^{n \times m}, \quad 0 \leq \ell \leq k-1, \quad (21)
\]

in which case

\[
G_{\ell} = \hat{Q}_{\rho(\ell)} B P_{\ell}, \quad 0 \leq \ell \leq k-1,
\]

and

\[
SB = BR_{\rho} = \sum_{\ell=0}^{k-1} \gamma_{\rho(\ell)} Q_{\rho(\ell)} G_{\ell} \hat{P}_{\ell}
\]

for arbitrary \( \gamma_0, \gamma_1, \ldots, \gamma_{k-1} \).

### 4 General Results

**Remark 1** If \( \sigma \) or \( \rho \) is a permutation of \( \mathbb{Z}_k \), we can replace \( \ell \) by \( \sigma(\ell) \) or \( \ell \) by \( \rho(\ell) \) in a summation \( \sum_{\ell=0}^{k-1} \), as in the proof of the following theorem, where “\( \circ \)” denotes composition; i.e., \( \sigma \circ \rho(\ell) = \sigma(\rho(\ell)) \) and \( \rho \circ \sigma(\ell) = \rho(\sigma(\ell)) \). Also,

\[
\hat{P}_{\sigma(\ell)} P_{\sigma(\ell)} = \delta_{rs} I_{n_{\sigma(\ell)}} \quad \text{and} \quad \hat{Q}_{\rho(\ell)} Q_{\rho(\ell)} = \delta_{rs} I_{n_{\rho(\ell)}}, \quad 0 \leq r, s \leq k-1, \quad (22)
\]

if and only if \( \sigma \) and \( \rho \) are permutations. We will use this frequently without specifically invoking it.

**Theorem 4** Suppose \( A \in \mathbb{C}^{m \times n} \) is \((R, S_{\sigma})\)-commutative and \( B \in \mathbb{C}^{n \times m} \) is \((S, R_{\rho})\)-commutative. Then: (a) \( AB \) is \((R, R_{\sigma\circ\rho})\)-commutative if \( \rho \) is a permutation and (b) \( BA \) is \((S, S_{\rho\circ\sigma})\)-commutative if \( \sigma \) is a permutation.

**PROOF.** From Theorems 2 and 3, our assumptions imply that \( A \) is as in (15) and \( B \) is as in (21). If \( \rho \) is a permutation then replacing \( \ell \) by \( \rho(\ell) \) in (15) yields

\[
A = \sum_{\ell=0}^{k-1} P_{\sigma(\rho(\ell))} F_{\rho(\ell)} \hat{Q}_{\rho(\ell)}.
\]

From this, (21), and (22),

\[
AB = \sum_{\ell=0}^{k-1} P_{\sigma(\rho(\ell))} F_{\rho(\ell)} G_{\ell} \hat{P}_{\ell},
\]

so (9) and (11) imply that

\[
R(AB) = (AB) R_{\sigma\circ\rho} = \sum_{\ell=0}^{k-1} \gamma_{\sigma(\rho(\ell))} P_{\sigma(\rho(\ell))} F_{\rho(\ell)} G_{\ell} \hat{P}_{\ell},
\]
(\(R, S_\alpha\))-commutative matrices

which proves (a).

If \(\sigma\) is a permutation, replacing \(\ell\) by \(\sigma(\ell)\) in (21) yields

\[
B = \sum_{\ell=0}^{k-1} Q_{\rho(\sigma(\ell))} G_{\sigma(\ell)} \hat{F}_{\alpha(\ell)}.
\]

From this, (15), and (22),

\[
BA = \sum_{\ell=0}^{k-1} Q_{\rho(\sigma(\ell))} G_{\sigma(\ell)} F_\ell \hat{Q}_\ell.
\]

so (10) and (12) imply that

\[
S(BA) = (AB) S_{\rho\sigma} = \sum_{\ell=0}^{k-1} Q_{\rho(\sigma(\ell))} G_{\sigma(\ell)} F_\ell \hat{Q}_\ell.
\]

which proves (b). \(\Box\)

**Corollary 1** If \(\sigma\) is a permutation, \(A \in \mathbb{C}^{m \times n}\) is \((R, S_\alpha)\)-commutative, and \(B \in \mathbb{C}^{n \times m}\) is \((S, R_{\alpha-1})\)-commutative, then \(AB\) commutes with \(R\) and \(BA\) commutes with \(S\).

**Theorem 5** Suppose \(j > 1\) and \(A_j \in \mathbb{C}^{m \times m}\) is \((R, R_{\alpha_j})\)-commutative, where \(\sigma_j\) is a permutation if \(j > 1\). Then \(A_1 A_2 \cdots A_j\) is \((R, R_{\alpha_1 \alpha_2 \cdots \alpha_j})\)-commutative; specifically, if

\[
A_j = \sum_{\ell=0}^{k-1} P_{\sigma_j(\ell)} F^{(j)}_\ell \hat{P}_\ell,
\]

then

\[
A_1 A_2 = \sum_{\ell=0}^{k-1} P_{\sigma_1 \sigma_2(\ell)} F^{(1)}_{\sigma_2(\ell)} F^{(2)}_\ell \hat{P}_\ell,
\]

\[
A_1 A_2 A_3 = \sum_{\ell=0}^{k-1} P_{\sigma_1 \sigma_2 \sigma_3(\ell)} F^{(1)}_{\sigma_2 \sigma_3(\ell)} F^{(2)}_{\sigma_3(\ell)} F^{(3)}_\ell \hat{P}_\ell,
\]

and, in general,

\[
A_1 A_2 \cdots A_j = \sum_{\ell=0}^{k-1} P_{\sigma_1 \sigma_2 \cdots \sigma_j(\ell)} F^{(1)}_{\sigma_2 \cdots \sigma_j(\ell)} \cdots F^{(j-1)}_{\sigma_j(\ell)} F^{(j)}_\ell \hat{P}_\ell.
\]

**Proof.** To minimize complicated notation, suppose

\[
B_j = \sum_{\ell=0}^{k-1} P_{\sigma_1 \sigma_2 \cdots \sigma_j(\ell)} G^{(j)}_\ell \hat{P}_\ell
\]
10

William F. Trench

for some \( j \geq 1 \). Since \( \sigma_{j+1} \) is a permutation, we can replace \( \ell \) by \( \sigma_{j+1}(\ell) \) to obtain

\[
B_j = \sum_{\ell=0}^{k-1} P_{\sigma_1 \sigma_2 \cdots \sigma_{j+1}(\ell)} G_{\sigma_{j+1}(\ell)} \hat{P}_{\sigma_{j+1}(\ell)}.
\]

Therefore, from (23) with \( j \) replaced by \( j + 1 \),

\[
B_j A_{j+1} = \left( \sum_{\ell=0}^{k-1} P_{\sigma_1 \sigma_2 \cdots \sigma_{j+1}(\ell)} G_{\sigma_{j+1}(\ell)} \hat{P}_{\sigma_{j+1}(\ell)} \right) \left( \sum_{\ell=0}^{k-1} P_{\sigma_{j+1}(\ell)} F_{(j+1),\ell} \hat{P}_{\ell} \right) = \sum_{\ell=0}^{k-1} P_{\sigma_1 \sigma_2 \cdots \sigma_{j+1}(\ell)} G_{(j+1),\ell} \hat{P}_{\ell} \text{ with } G_{(j+1),\ell} = G_{\sigma_{j+1}(\ell)} F_{(j+1),\ell}.
\]

This provides the basis for a straightforward induction proof of the assertion. \( \square \)

**Corollary 2** If \( \sigma \) is a permutation, \( A \in \mathbb{C}^{m \times m} \) is \((R, R_{\sigma})\)-commutative, and \( j \) is a positive integer, then \( A^j \) is \((R, R_{\sigma})\)-commutative, explicitly,

\[
A^j = \sum_{\ell=0}^{k-1} P_{\sigma^j(\ell)} F_{\sigma^{j-1}(\ell)} \cdots F_{\sigma(\ell)} F_{\ell} \hat{P}_{\ell}
\]

and

\[
RA = AR = \sum_{\ell=0}^{k-1} \gamma_{\sigma^j(\ell)} P_{\sigma^j(\ell)} F_{\sigma^{j-1}(\ell)} \cdots F_{\sigma(\ell)} F_{\ell} \hat{P}_{\ell}
\]

for arbitrary \( \gamma_0, \gamma_1, \ldots, \gamma_{k-1} \).

5 **Generalized Inverses and Singular Value Decompositions**

If \( A \) is an arbitrary complex matrix then \( A^- \) is a reflexive inverse of \( A \) if \( AA^- A = A \) and \( A^- AA^- = A^- \). The Moore-Penrose inverse \( A^\dagger \) of \( A \) is the unique matrix that satisfies the Penrose conditions

\[
(AA^\dagger)^* = AA^\dagger, \quad (A^\dagger A)^* = AA^\dagger, \quad AA^\dagger A = A, \quad \text{and} \quad A^\dagger AA^\dagger = A^\dagger.
\]

**Theorem 6** Suppose \( \sigma \) is a permutation and \( A \in \mathbb{C}^{m \times n} \) is \((R, S_{\sigma})\)-commutative, so

\[
A = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\ell} \hat{Q}_{\ell},
\]

by Theorem 2. Let \( F_0^-, F_1^- \cdots, F_{k-1}^- \) be reflexive inverses of \( F_0, F_1, \ldots, F_{k-1} \), and define

\[
B = \sum_{\ell=0}^{k-1} Q_{\ell} F_{\ell} \hat{P}_{\sigma(\ell)}.
\]
Then $B$ is a reflexive inverse of $A$. Moreover, if $P$ and $Q$ are unitary, then

$$A^\dagger = \sum_{\ell=0}^{k-1} Q_{\ell} F_{\ell}^\dagger P_{\sigma(\ell)}^*. \tag{27}$$

**Proof.** From (9), (10), (22), (25), and (26),

$$AB = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\ell} F_{\ell}^\dagger \hat{P}_{\sigma(\ell)} \quad BA = \sum_{\ell=0}^{k-1} Q_{\ell} F_{\ell}^\dagger F_{\ell} \hat{Q}_{\ell}. \tag{28}$$

$$ABA = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\ell} F_{\ell}^\dagger F_{\ell} \hat{Q}_{\ell} = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\ell} \hat{Q}_{\ell} = A. \tag{29}$$

and

$$BAB = \sum_{\ell=0}^{k-1} Q_{\ell} F_{\ell}^\dagger F_{\ell} F_{\ell}^\dagger \hat{P}_{\sigma(\ell)} = \sum_{\ell=0}^{k-1} Q_{\ell} F_{\ell}^\dagger \hat{P}_{\sigma(\ell)} = B. \tag{30}$$

The last two equations show that $B$ is a reflexive inverse of $A$. If $P$ and $Q$ are unitary and we redefine

$$B = \sum_{\ell=0}^{k-1} Q_{\ell} F_{\ell}^\dagger P_{\sigma(\ell)}^*,$$

then (28)–(30) become

$$AB = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\ell} F_{\ell}^\dagger P_{\sigma(\ell)}^*, \quad BA = \sum_{\ell=0}^{k-1} Q_{\ell} F_{\ell}^\dagger F_{\ell} Q_{\ell}^*, \tag{31}$$

$$ABA = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\ell} F_{\ell}^\dagger F_{\ell} Q_{\ell}^* = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\ell} Q_{\ell}^* = A,$$

and

$$BAB = \sum_{\ell=0}^{k-1} Q_{\ell} F_{\ell}^\dagger F_{\ell} F_{\ell}^\dagger P_{\sigma(\ell)}^* = \sum_{\ell=0}^{k-1} Q_{\ell} F_{\ell}^\dagger P_{\sigma(\ell)}^* = B.$$

Moreover, from (31)

$$(AB)^* = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} (F_{\ell} F_{\ell}^\dagger)^* P_{\sigma(\ell)}^* = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\ell} F_{\ell}^\dagger P_{\sigma(\ell)}^* = AB$$

and

$$(BA)^* = \sum_{\ell=0}^{k-1} Q_{\ell} (F_{\ell}^\dagger F_{\ell})^* Q_{\ell}^* = \sum_{\ell=0}^{k-1} Q_{\ell} F_{\ell}^\dagger F_{\ell} Q_{\ell}^* = BA.$$

Therefore $B = A^\dagger$, which implies (27).
Corollary 3 If \( \sigma \) is a permutation, \( P \) and \( Q \) are unitary, and \( A \in \mathbb{C}^{m \times n} \) is \( (R, S_\sigma) \)-commutative, then \( A^\dagger \) is \( (S, R_{\sigma^{-1}}) \)-commutative.

PROOF. From (9)–(11), (22), and (27),

\[
SA^\dagger = A^\dagger R_{\sigma^{-1}} = \sum_{\ell=0}^{k-1} y_\ell Q_\ell F_\ell^\dagger P_{\sigma(\ell)}^*
\]

for arbitrary \( y_0, y_1, \ldots, y_{k-1} \). \( \square \)

Remark 2 It is well known – and straightforward to verify – that if \( G \in \mathbb{C}^{p \times q} \) and \( \text{rank} \ G = q \), then \( G^\dagger = (G^*G)^{-1}G^* \). Hence, (27) implies the following corollary.

Corollary 4 In addition to the assumptions of Theorem 6, suppose that \( \text{rank} \ F_\ell = n_\ell \), \( 0 \leq \ell \leq k - 1 \) (or, equivalently, \( \text{rank} \ A = n \)). Then

\[
A^\dagger = \sum_{\ell=0}^{k-1} Q_\ell (F_\ell^* F_\ell)^{-1} F_\ell^* P_{\sigma(\ell)}^*.
\]

Theorem 7 Suppose \( \sigma \) is a permutation, \( P \) and \( Q \) are unitary, and \( A \) is \( (R, S_\sigma) \)-commutative and therefore of the form

\[
A = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_\ell Q_\ell^*.
\]

by Theorem 2. Let

\[
F_\ell = \Omega_\ell \Gamma_\ell \Phi_\ell^*, \quad 0 \leq \ell \leq k - 1,
\]

with

\[
\Omega_\ell \in \mathbb{C}^{m_{\sigma(\ell)} \times m_{\sigma(\ell)}}, \quad \Gamma_\ell \in \mathbb{C}^{m_{\sigma(\ell)} \times n_\ell}, \quad \text{and} \quad \Phi_\ell \in \mathbb{C}^{n_\ell \times n_\ell}, \quad 0 \leq \ell \leq k - 1,
\]

be singular value decompositions of \( F_\ell, 0 \leq \ell \leq k - 1 \). Let

\[
\Omega = \begin{bmatrix}
P_{\sigma(0)} \Omega_0 & P_{\sigma(1)} \Omega_1 & \cdots & P_{\sigma(k-1)} \Omega_{k-1}
\end{bmatrix}
\]

and

\[
\Phi = \begin{bmatrix}
Q_0 \Phi_0 & Q_1 \Phi_1 & \cdots & Q_{k-1} \Phi_{k-1}
\end{bmatrix}
\]

Then

\[
A = \Omega \text{ diag}(\Gamma_0, \Gamma_1, \ldots, \Gamma_{k-1}) \Phi^*
\]

is a singular value decomposition of \( A \), except that the singular values are not necessarily arranged in decreasing order. Thus, for \( 0 \leq \ell \leq k - 1 \), each singular value of \( F_\ell \) is a singular value of \( A \) with an associated left singular vector in the column space of \( P_{\sigma(\ell)} \) and a right singular vector in the column space of \( Q_\ell \).

We invoke the first equality in (22) repeatedly in the proof of the following theorem.
Theorem 8 Suppose \( \sigma \) is a permutation, \( P \) is unitary, and \( A \in \mathbb{C}^{m \times m} \) is \((R, R_\sigma)\)-commutative, so
\[
A = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_\ell P_\ell^*.
\]
by Theorem 2. Then:
(i) \( A \) is Hermitian if and only if \( F_{\sigma(\ell)}^* F_{\sigma(\ell)} = F_\ell P_\ell^* \), \( 0 \leq \ell \leq k - 1 \).
(ii) \( A \) is normal if and only if \( F_{\sigma(\ell)}^* F_{\sigma(\ell)} = F_\ell F_\ell^* \), \( 0 \leq \ell \leq k - 1 \).
(iii) \( A \) is EP (i.e., \( AA^\dagger = A^\dagger A \)) if and only if \( F_{\sigma(\ell)}^* F_{\sigma(\ell)} = F_\ell F_\ell^* \), \( 0 \leq \ell \leq k - 1 \).

PROOF. Since \( R \) is unitary, Theorems 2 and 6 imply that
\[
A = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_\ell P_\ell^*. \quad A^* = \sum_{\ell=0}^{k-1} P_\ell F_\ell^* F_{\sigma(\ell)}^* \quad \text{and} \quad A^\dagger = \sum_{\ell=0}^{k-1} P_\ell F_\ell^* P_{\sigma(\ell)}^*.
\]
Replacing \( \ell \) by \( \sigma(\ell) \) in the second sum in (32) yields
\[
A^* = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\sigma(\ell)}^* F_{\sigma(\ell)}^*.
\]
and comparing this with the first sum in (32) yields (i).

From (32),
\[
AA^* = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_\ell F_{\sigma(\ell)}^* F_{\sigma(\ell)}^* P_{\sigma(\ell)}^*
\]
and
\[
A^* A = \sum_{\ell=0}^{k-1} P_\ell F_\ell^* F_\ell^* P_\ell^* = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\sigma(\ell)}^* F_{\sigma(\ell)}^* P_{\sigma(\ell)}^*.
\]
Comparing the second sum here with (33) yields (ii).

From (33),
\[
AA^\dagger = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_\ell F_\ell^* P_{\sigma(\ell)}^*
\]
and
\[
A^\dagger A = \sum_{\ell=0}^{k-1} P_\ell F_\ell^* F_\ell^* P_\ell^* = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\sigma(\ell)}^* F_{\sigma(\ell)}^* P_{\sigma(\ell)}^*.
\]
Comparing the second sum here with (34) yields (iii). \( \square \)

6 Solving \( AZ = w \) and the least-squares problem

Throughout this section \( \sigma \) is a permutation and \( A \in \mathbb{C}^{m \times n} \) is \((R, S_\sigma)\)-commutative, and can therefore be written as in (15).
If \( z \in \mathbb{C}^n \) and \( w \in \mathbb{C}^m \), we write
\[
z = Qu = \sum_{\ell=0}^{k-1} Q_{\ell} u_{\ell} \quad \text{and} \quad w = Pv = \sum_{\ell=0}^{k-1} P_{\ell} v_{\ell},
\] (35)
with \( u_{\ell} \in \mathbb{C}^n \) and \( v_{\ell} \in \mathbb{C}^m \), \( 0 \leq \ell \leq k - 1 \).

**Theorem 9** If (35) holds then

(a) \( A'z = w \) if and only if (b) \( F_{\ell} u_{\ell} = v_{\sigma(\ell)} \), \( 0 \leq \ell \leq k - 1 \). (36)

**Proof.** From (10), (15), and (35),
\[
A'z - w = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\ell} u_{\ell} - \sum_{\ell=0}^{k-1} P_{\ell} v_{\ell} = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\ell} u_{\ell} - \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} v_{\sigma(\ell)}
\]
\[
= \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} (F_{\ell} u_{\ell} - v_{\sigma(\ell)}),
\] (37)

so (36)(b) implies (36)(a). From (22) and (37),
\[
F_{\ell} u_{\ell} - v_{\sigma(\ell)} = \hat{P}_{\sigma(\ell)} (A'z - w), \quad 0 \leq \ell \leq k - 1,
\]
so (36)(a) implies (36)(b). \( \square \)

Since \( F_{\ell} \in \mathbb{C}^{m_{\sigma(\ell)} \times n_{\ell}} \), \( 0 \leq \ell \leq k - 1 \), (36) implies the following theorem.

**Theorem 10** \( A \) is invertible if and only if \( m_{\sigma(\ell)} = n_{\ell} \) and \( F_{\ell} \) is invertible, \( 0 \leq \ell \leq k - 1 \) (which, from (3), implies that \( m = n \)). In this case,
\[
A^{-1} = \sum_{\ell=0}^{k-1} Q_{\ell} F_{\ell}^{-1} \hat{P}_{\sigma(\ell)}
\] (38)
and the solution of \( A'z = w \) is
\[
z = \sum_{\ell=0}^{k-1} Q_{\ell} F_{\ell}^{-1} v_{\sigma(\ell)}.\]

Moreover, \( A^{-1} \) is \( (S, R_{\sigma-1}) \)-commutative; specifically,
\[
SA^{-1} = A^{-1} R_{\sigma-1} = \sum_{\ell=0}^{k-1} \gamma_{\ell} Q_{\ell} F_{\ell}^{-1} \hat{P}_{\sigma(\ell)}
\]
for arbitrary \( \gamma_0, \gamma_1, \ldots, \gamma_{k-1} \).
If \( m = n \) and \( R = S \) (so \( A \) is \((R, R_0)\)-commutative), then (38) becomes

\[
A^{-1} = \sum_{\ell=0}^{k-1} P_{\ell} F_{\ell}^{-1} \hat{P}_{\sigma(\ell)}.
\]

In this case,

\[
A^{-j} = \sum_{\ell=0}^{k-1} P_{\ell} F_{\ell}^{-1} F_{\sigma(\ell)}^{-1} \cdots F_{\sigma^{j-1}(\ell)}^{-1} \hat{P}_{\sigma^j(\ell)},
\]

which can be verified by simply multiplying the right hand side by \( A^j \) as written in (24).

Before turning to the least squares problem for \( A \), we review some elementary facts about the least squares problem for a matrix \( G \in \mathbb{C}^{p \times q} \) and a given \( u \in \mathbb{C}^p \); i.e., find \( v \in \mathbb{C}^q \) such that

\[
\| Gv - u \| = \min_{\xi \in \mathbb{C}^q} \| G\xi - u \|,
\]

where \( \| \cdot \| \) is the 2-norm. An arbitrary \( v \in \mathbb{C}^{p \times q} \) can be written as

\[
v = G^\dagger u + G(v - G^\dagger u),
\]

so

\[
\| Gv - u \|^2 = \| (GG^\dagger - I_p)u \|^2 + \| G(v - G^\dagger u) \|^2.
\]

since

\[
(G(v - G^\dagger u))^* (GG^\dagger - I_p)u = [GG^\dagger G(v - G^\dagger u)]^* (GG^\dagger - I_p)u
\]

\[
= [G(v - G^\dagger u)]^* GG^\dagger (GG^\dagger - I_p)u
\]

and

\[
G^\dagger (GG^\dagger - I_p) = G^\dagger GG^\dagger - G^\dagger = 0.
\]

Hence,

\[
\min_{\xi \in \mathbb{C}^q} \| G\xi - u \| = \| (GG^\dagger - I)u \|,
\]

and this minimum is attained with a given \( v \) if and only if \( v = G^\dagger u + h \) where \( Gh = 0 \). In this case, \( \| v \|^2 = \| G^\dagger u \|^2 + \| h \|^2 \) since

\[
h^* G^\dagger u = h^* G^\dagger GG^\dagger u = (G^\dagger Gh)^* G^\dagger u = 0,
\]

so \( v_0 = G^\dagger u \) is the unique solution of (37) with minimal norm, and is therefore called the optimal solution. From Remark 2, \( v_0 = (G^* G)^{-1} G^* u \) if \( \text{rank}(G) = q \). If \( P \) is unitary then (37) implies that

\[
\| Az - w \|^2 = \sum_{\ell=0}^{k-1} \| F_{\ell} u_\ell - v_{\sigma(\ell)} \|^2,
\]
so the least squares problem for $A$ and a given $w$ reduces to $k$ independent least squares problems for $F_{\ell} \in \mathbb{C}^{m_{\sigma(\ell)} \times n_{\ell}}$ and a given $v_{\sigma(\ell)} \in \mathbb{C}^{m_{\sigma(\ell)}}$, $0 \leq \ell \leq k - 1$. Therefore,

$$\|Az - w\| = \min_{\xi \in \mathbb{C}^n} \|A\xi - w\|$$

if and only if

$$z = \sum_{\ell=0}^{k-1} Q_{\ell} (F_{\ell}^\dagger v_{\sigma(\ell)} + h_{\ell}),$$

where $F_{\ell} h_{\ell} = 0$, $0 \leq \ell \leq k - 1$. If $Q$ is also unitary, then

$$\|z\|^2 = \sum_{\ell=0}^{k-1} \|F_{\ell}^\dagger v_{\sigma(\ell)} + h_{\ell}\|^2 = \sum_{\ell=0}^{k-1} \|F_{\ell}^\dagger v_{\sigma(\ell)}\|^2 + \sum_{\ell=0}^{k-1} \|h_{\ell}\|^2,$$

so the unique optimal (least norm) solution of the least squares problem is

$$z = \sum_{\ell=0}^{k-1} Q_{\ell} F_{\ell}^\dagger v_{\sigma(\ell)},$$

which can be written as

$$z = \sum_{\ell=0}^{k-1} Q_{\ell} (F_{\ell}^* F_{\ell})^{-1} F_{\ell}^* v_{\sigma(\ell)} \quad \text{if rank}(F_{\ell}) = n_{\ell}, \quad 0 \leq \ell \leq k - 1,$$

or, equivalently, if rank($A$) = $n$.

### 7 The eigenvalue problem

Throughout this section $A \in \mathbb{C}^{m \times m}$ is $(R, R_{\sigma})$-commutative, and can therefore be written as

$$A = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\ell} \hat{P}_{\ell} \quad \text{where} \quad F_{\ell} \in \mathbb{C}^{m_{\sigma(\ell)} \times n_{\ell}}, \quad 0 \leq \ell \leq k - 1,$$

(39)

and $\sigma$ is a permutation.

An arbitrary $z \in \mathbb{C}^m$ can be written as

$$z = \sum_{\ell=0}^{k-1} P_{\ell} u_{\ell} \quad \text{with} \quad u_{\ell} \in \mathbb{C}^{m_{\ell}}, \quad 0 \leq \ell \leq k - 1.$$

Therefore (9) and (39) imply that

$$Az - \lambda z = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} F_{\ell} u_{\ell} - \lambda \sum_{\ell=0}^{k-1} P_{\ell} u_{\ell} = \sum_{\ell=0}^{k-1} P_{\sigma(\ell)} (F_{\ell} u_{\ell} - \lambda u_{\sigma(\ell)});$$

(40)
hence, \(Az = \lambda z\) if and only if

\[F_\ell u_\ell = \lambda u_{\sigma(\ell)}, \quad 0 \leq \ell \leq k - 1.\]

We first consider the case where \(\sigma\) is the identity. The next three theorems are essentially restatements of results from [21], recast so as to be consistent with viewpoint that we have taken in this paper.

Let \(C_\ell\) denote the column space of \(P_\ell\) and let \(C = \bigcup_{\ell=0}^{k-1} C_\ell\).

**Theorem 11** If \(A\) commutes with \(R\) then \(\lambda\) is an eigenvalue of \(A\) if and only if \(\lambda\) is an eigenvalue of one or more of the matrices \(F_0, F_1, \ldots, F_{k-1}\). Assuming this to be true, let

\[S_A(\lambda) = \{\ell \in \{0, 1, \ldots, k-1\} \mid \lambda\text{ is an eigenvalue of } F_\ell\}.\]

If \(\ell \in S_A(\lambda)\) and \(\{u^{(1)}_\ell, u^{(2)}_\ell, \ldots, u^{(d_\ell)}_\ell\}\) is a basis for the set \(\{u_\ell \in \mathbb{C}^{m_\ell \times m_\ell} \mid F_\ell u_\ell = \lambda u_\ell\}\),

then \(P_\ell u^{(1)}_\ell, P_\ell u^{(2)}_\ell, \ldots, P_\ell u^{(d_\ell)}_\ell\) are linearly independent \(\lambda\)-eigenvectors of \(A\). Moreover,

\[\bigcup_{\ell \in S_A(\lambda)} \{P_\ell u^{(1)}_\ell, P_\ell u^{(2)}_\ell, \ldots, P_\ell u^{(d_\ell)}_\ell\}\]

is a basis for the \(\lambda\)-eigenspace of \(A\). Finally, \(A\) is diagonalizable if and only if \(F_0, F_1, \ldots, F_{k-1}\) are all diagonalizable. In this case, \(A\) has \(m_\ell\) linearly independent eigenvectors in \(C_\ell\), \(0 \leq \ell \leq k - 1\).

It seems useful to consider the case where \(A\) is diagonalizable more explicitly.

**Theorem 12** Suppose a diagonalizable matrix \(A\) commutes with \(R\) and and \(F_\ell = \Omega_\ell D_\ell \Omega^{-1}_\ell\) is a spectral decomposition of \(F_\ell\), \(0 \leq \ell \leq k - 1\). Let

\[\Omega = \begin{bmatrix} P_0 \Omega_0 & P_1 \Omega_1 & \cdots & P_{k-1} \Omega_{k-1} \end{bmatrix}\]

Then

\[A = \Omega \left( \bigoplus_{\ell=0}^{k-1} D_\ell \right) \Omega^{-1}\]

with

\[\Omega^{-1} = \begin{bmatrix} \Omega_0^{-1} \hat{P}_0 \\ \Omega_1^{-1} \hat{P}_1 \\ \vdots \\ \Omega_{k-1}^{-1} \hat{P}_{k-1} \end{bmatrix}\]

is a spectral decomposition of \(A\).

**Remark 3** It is well known that commuting diagonalizable matrices are simultaneously diagonalizable. Theorem 12 makes this explicit, since since \(\Omega R \Omega^{-1}\) and \(\Omega A \Omega^{-1}\) are both diagonal.
The original version of the following theorem, which dealt with centrosymmetric matrices, is due to Andrew [2, Theorem 6]. The proof is practically identical to Andrew’s original proof.

**Theorem 13**

(i) If $A$ commutes with $R$ and $\lambda$ is an eigenvalue of $A$, then the $\lambda$-eigenspace of $S$ has a basis in $C$.

(ii) If $A$ has $n$ linearly independent eigenvectors in $C$, then $A$ commutes with $R$.

**Proof.** (i) See Theorem 11. (ii) If $z \in C$ then $Rz = \gamma_t z$ for some $\ell \in \mathbb{Z}_k$. If $Az = \lambda z$ and $Rz = \gamma_t z$, then

$$RAz = \lambda Rz = \lambda \gamma_t z \quad \text{and} \quad ARz = \gamma_t Az = \gamma_t \lambda z;$$

hence, $RAz = ARz$. Now suppose that $A$ has $n$ linearly independent eigenvectors

$$\{z_1, z_2, \ldots, z_n\} \text{ in } C.$$

Then we can write an arbitrary $z \in C^n$ as $z = \sum_{i=1}^n a_i z_i$. Since $RAz_i = ARz_i, 1 \leq i \leq n$, it follows that $RAz = ARz$. Therefore $AR = RA$.

For the remainder of this section we assume that $A$ is $(R, R_\sigma)$-commutative and $\sigma$ is a permutation other than the identity.

The following theorem shows that finding the null space of $A$ reduces to finding the null spaces of $F_0, F_1, \ldots, F_{k-1}$.

**Theorem 14** If $A$ is $(R, R_\sigma)$-commutative and $\sigma$ is a permutation then $Az = 0$ if and only if $z = \sum_{i=0}^{k-1} F_\ell u_\ell$, where

$$F_\ell u_\ell = 0, \quad 0 \leq \ell \leq k - 1; \quad (41)$$

hence, the null space if $A$ is independent of $\sigma$ (so long as $\sigma$ is a permutation).

**Proof.** Clearly, (41) implies that $Az = 0$ without any assumption on $\sigma$. For the converse, note from (22) and (40) that if $\sigma$ is a permutation then $P_{\sigma(\ell)} Az = F_\ell u_\ell$, $0 \leq \ell \leq k - 1$, so $Az = 0$ implies (41).

Henceforth we assume that $\lambda \neq 0$. In this case, suppose that $\sigma$ has $p$ orbits $\sigma_0, \ldots, \sigma_{p-1}$. If $p = 1$, then $\sigma$ is a $k$-cycle and $Z_k = \{\sigma^j(0) | 0 \leq j \leq k - 1\}$. In any case, if $\ell_r \in \sigma_r, 0 \leq r \leq p - 1$, then $Z_k = \sigma_0 \cup \cdots \cup \sigma_{p-1}$, where

$$\sigma_r = \{\sigma^j(\ell_r) | 0 \leq j \leq k_r - 1\}, \quad 0 \leq r \leq p - 1,$$

and $k_0 + \cdots + k_{p-1} = k$. It is important to note that

$$\sigma^{k_r}(\ell_r) = \ell_r, \quad 0 \leq r \leq p - 1; \quad (42)$$

and $k_0, k_1, \ldots, k_{p-1}$ are respectively the smallest positive integers for which these equalities hold. In Example 1, $p = 3$, $\sigma_0 = \{0, 1, 3\}, \sigma_1 = \{2, 4\}, \sigma_2 = \{5\}$, so $k_0 = 3, k_1 = 2, k_3 = 1$. $Z_k = \sigma_0 \cup \sigma_1 \cup \sigma_2$, and we may choose $\ell_0 = 0, \ell_1 = 2$, and $\ell_2 = 5$. 
To solve the eigenvalue problem, we rearrange the terms in $z = \sum_{\ell=0}^{k-1} P_{\ell} u_{\ell}$ as

$$z = \sum_{r=0}^{p-1} z_r \quad \text{with} \quad z_r = \sum_{j=0}^{k-1} P_{\sigma^j(\ell_r)} u_{\sigma^j(\ell_r)}, \quad 0 \leq r \leq p-1, \quad (43)$$

and rearrange the terms in (39) as

$$A = \sum_{r=0}^{p-1} A_r \quad \text{with} \quad A_r = \sum_{j=0}^{k-1} P_{\sigma^{j+1}(\ell_r)} F_{\sigma^j(\ell_r)} \bar{P}_{\sigma^j(\ell_r)}, \quad 0 \leq r \leq p-1. \quad (44)$$

Since (9) implies that $A_r A_s = 0$ if $r \neq s$, we can replace (44) by

$$A = A_0 \oplus A_1 \oplus \cdots \oplus A_{p-1};$$

hence, $Az = \lambda z$ if and only if

$$A_r z_r = \lambda z_r, \quad 0 \leq r \leq p-1.$$

Therefore, the eigenvalue problem for $A$ reduces to $p$ independent eigenvalue problems for $A_0, A_1, \ldots, A_{p-1}$.

From (43) and (44), $A_r z_r = \lambda z_r$ if and only if

$$\sum_{j=0}^{k-1} P_{\sigma^{j+1}(\ell_r)} F_{\sigma^j(\ell_r)} u_{\sigma^j(\ell_r)} = \lambda \sum_{j=0}^{k-1} P_{\sigma^j(\ell_r)} u_{\sigma^j(\ell_r)} = \lambda \sum_{j=0}^{k-1} P_{\sigma^{j+1}(\ell_r)} u_{\sigma^{j+1}(\ell_r)},$$

which is equivalent to

$$F_{\sigma^j(\ell_r)} u_{\sigma^j(\ell_r)} = \lambda u_{\sigma^{j+1}(\ell_r)}, \quad 0 \leq j \leq k_r - 1. \quad (45)$$

If $k_r = 1$ then $\sigma(\ell_r) = \ell_r$ and (44) becomes $F_{\ell_r} u_{\ell_r} = \lambda u_{\ell_r}$; hence, if $(\lambda, u_{\ell_r})$ is an eigenpair of $F_{\ell_r}$ then $z_r = P_{\ell_r} u_{\ell_r}$ is an $\lambda$-eigenvector of $A$.

If $k_r > 1$ then (42) and (44) imply that

$$G_r u_{\ell_r} = \lambda^k u_{\ell_r}, \quad \text{where} \quad G_r = F_{\sigma^{k_r-1}(\ell_r)} \cdots F_{\sigma(\ell_r)} F_{\ell_r} \in \mathbb{C}^{m_{\ell_r} \times m_{\ell_r}}.$$

Therefore, if $v$ is a nonzero eigenvalue of $G_r$ and $\zeta = e^{2\pi i / k_r}$, then $v^{1/k}, v^{1/k} \zeta, \ldots, v^{1/k} \zeta^{k_r-1}$ are distinct eigenvalues of $A_r$ (and therefore of $A$). If $\lambda$ is any one of these eigenvalues, then the corresponding eigenvector $z_r$ of $A_r$ (and therefore of $A$) is given by (43), where $u_{\sigma^j(\ell_r)}, 1 \leq j \leq k_r-1$, can be computed recursively from (44) as

$$u_{\sigma^j(\ell_r)} = \frac{1}{\lambda} F_{\sigma^{j-1}(\ell_r)} u_{\sigma^{j-1}(\ell_r)}, \quad 1 \leq j \leq k_r - 1.$$

References


(R, S_{o})-commutative matrices


