REDUCING THE RANK OF \((A - \lambda B)^1\)

GERALD L. THOMPSON AND ROMAN L. WEIL

Abstract. The rank of the \(n \times n\) matrix \((A - \lambda I)\) is \(n - J(\lambda)\) when \(\lambda\) is an eigenvalue occurring in \(J(\lambda) \geq 0\) Jordan blocks of the Jordan normal form of \(A\). In our principal theorem we derive an analogous expression for the rank of \((A - \lambda B)\) for general, \(m \times n\), matrices. When \(J(\lambda) > 0\), \(\lambda\) is a rank-reducing number of \((A - \lambda I)\). We show how the rank-reducing properties of eigenvalues can be extended to \(m \times n\) matrix expressions \((A - \lambda B)\). In particular we give a constructive way of deriving a polynomial \(P(\lambda, A, B)\) whose roots are the only rank-reducing numbers of \((A - \lambda B)\). We name this polynomial the characteristic polynomial of \(A\) relative to \(B\) and justify that name.

The rank of the \(n \times n\) matrix \((A - \lambda I)\) is \(n - J(\lambda)\) when \(\lambda\) is an eigenvalue appearing on the main diagonal of \(J(\lambda) \geq 0\) distinct Jordan blocks in the Jordan normal form of \(A\). See, e.g., Wilkinson [2, p. 11]. Thus an eigenvalue, \(\lambda\), of \(A\) is a rank-reducing number of \((A - \lambda I)\).

In this paper we show how the rank-reducing properties of eigenvalues can be extended to \(m \times n\) matrix expressions \((A - \lambda B)\). In particular we give a constructive way of deriving a polynomial \(P(\lambda, A, B)\) whose roots are the only rank-reducing numbers of \((A - \lambda B)\). We have named this polynomial the characteristic polynomial of \(A\) relative to \(B\). In our principal theorem we derive a simple expression for the rank of \((A - \lambda B)\). Finally we relate our results to the normal form derived by Gantmacher.

Consider the rank-reducing scalars \(\lambda\) of \((A - \lambda B)\). If the determinant \(\vert B \vert \neq 0\), then the problem of finding rank-reducing numbers is equivalent to an ordinary eigenvalue problem. For if \(\vert B \vert \neq 0\), the rank of \((A - \lambda B)\) is the same as that of \((B^{-1}A - \lambda I)\) and the rank-reducing \(\lambda\) of \((A - \lambda B)\) are the eigenvalues of \(B^{-1}A\). But this analysis breaks down if \(B\) is singular or, more generally, if \(A\) and \(B\) are rectangular. In this paper we shall consider rectangular \(A\) and \(B\), \(m \times n\). Let \(\text{rank}(A) = p\) and \(\text{rank}(B) = r\).

Lemma 1. For all \(m \times n\) \(A\) and \(B\), there exist nonsingular \(S_1\) and \(T_1\)
such that $S_1(A - \lambda B) T_1$ can be partitioned into

$$
\begin{bmatrix}
r & t & q \\
r & E_{11} & E_{12} & 0 & 0 \\
s & E_{21} & 0 & 0 & 0 \\
q & 0 & 0 & I_q & 0 \\
0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
- \lambda
\begin{bmatrix}
I_r & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\end{bmatrix}
\tag{1}
$$

where $q \leq \min(m - r, n - r, \phi)$, $E_{12}$ is in column echelon form and rank $(E_{12}) = t$, and $E_{21}$ is in row echelon form and rank $(E_{21}) = s$. Any of the numbers $s$, $t$, $q$, may be zero and $r + s + q \leq m$, $r + t + q \leq n$.

**Proof.** In Theorem 2.2 of [1] we proved a similar statement. An outline of the proof is as follows. In the expression $A - \lambda B$ use elementary row and column operations to put $B$ in diagonal form, ignoring their effects on $A$. Thus $A - \lambda B$ becomes

$$
\begin{bmatrix}
C_{11} & C_{12} \\
C_{21} & C_{22} \\
\end{bmatrix}
- \lambda
\begin{bmatrix}
I_r & 0 \\
0 & 0 \\
\end{bmatrix}
\tag{2}
$$

Now do the same kind of operation on $C_{22}$, except putting the identity matrix last so that $C_{22}$ becomes

$$
\begin{bmatrix}
t & q \\
s & 0 & 0 \\
q & 0 & I_q \\
\end{bmatrix}
\tag{3}
$$

By further row and column operations (in particular transferring all zero rows to the bottom and all zero columns to the right) change $C_{12}$ to $(E_{12} 0)$ where $E_{12}$ is in column echelon form and change $C_{21}$ to $[E_{21}]$ where $E_{21}$ is in row echelon form. Note that each successive level of transformation does not destroy the results of the preceding transformations.

**Lemma 2.** If $s + t > 0$ in (1), then there are nonsingular $S_2$ and $T_2$ such that $S_2(A - \lambda B) T_2$ can be partitioned into

$$
\begin{bmatrix}
r - s & s & t & q \\
r - t & C_{11} & 0 & 0 & 0 & 0 \\
t & 0 & 0 & I_t & 0 & 0 \\
s & 0 & I_s & 0 & 0 & 0 \\
q & 0 & 0 & 0 & I_q & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
- \lambda
\begin{bmatrix}
D_{11} & D_{12} & 0 & 0 & 0 \\
D_{21} & D_{22} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\tag{4}
$$
Proof. In Theorem 2.4 of [1] we prove a similar statement. An outline of the proof is as follows. Permute and partition $E_{12}$ and $E_{21}$ in (1) so that the first $2 \times 2$ block of (1) becomes

$$
\begin{bmatrix}
  r-s & s & t \\
  r-t & F_{11} & F_{12} \\
  t & F_{21} & F_{22} \\
  s & F_{31} & I_s
\end{bmatrix}
- \lambda
\begin{bmatrix}
  P_{11} & P_{12} & 0 \\
  P_{21} & P_{22} & 0 \\
  0 & 0 & 0
\end{bmatrix}
$$

where $P$ is a permutation matrix and $I_t$ and $I_s$ are constructed from the leading ones in the echelon forms $E_{12}$ and $E_{21}$. Further row and column operations lead to (2) where the $(r-t) \times (r-s)$ matrices $C_{11}$ and $D_{11}$ are defined by

$$
C_{11} = F_{11} - F_{12}F_{31} - F_{13}F_{21} + F_{12}F_{22}F_{31},
$$

$$
D_{11} = P_{11} - P_{12}F_{31} - F_{13}P_{21} + F_{12}P_{22}F_{31}.
$$

Lemma 3. (a) If $s+t=0$ in (1), then any number $\lambda$ reduces the ranks of both $(A-\lambda B)$ and $(E_{11}=\lambda I)$ by the number of Jordan blocks containing that $\lambda$ in the Jordan normal form of $E_{11}$.

(b) If $s+t>0$, then $\lambda$ is a rank-reducing number for $(A-\lambda B)$ if and only if it is a rank-reducing number for $(C_{11}-\lambda D_{11})$ in (2).

Proof. (a) If $s+t=0$, then the second row and second column of (1) are empty and the result is an obvious consequence of the classical results on the Jordan normal form; e.g., see [2].

(b) Suppose $s+t>0$. Define the row vector

$$
w = (w_1, w_2, w_3, w_4, w_5)
$$

partitioned the same way the rows of (2) are. Multiplying $S_2(A-\lambda B)T_2=0$ by $w$, we obtain from expression (2) the equations

$$
\begin{align*}
w_1(C_{11}-\lambda D_{11}) - w_2\lambda D_{21} &= 0, \\
-w_3\lambda D_{12} - w_2\lambda D_{22} + w_3I_s &= 0, \\
w_2I_t &= 0, \\
-w_4\lambda D_{12} &= 0.
\end{align*}
$$

Here $w_5$ is arbitrary and can be ignored. From (6), $w_4=0$ and from (5), $w_2=0$. The solution of (4) determines $w_2(=w_2\lambda D_{12})$ and imposes no constraint on $w_1$. Hence equations (3) to (6) reduce to simply $w_1(C_{11}-\lambda D_{11})=0$, regardless of the value of $\lambda$. From this it follows that the row rank (and hence the rank) of $(A-\lambda B)$ and $(C_{11}-\lambda D_{11})$ satisfy
\[ \text{rank}(A - \lambda B) = \text{rank}(C_{11} - \lambda D_{11}) + q + s + t, \]

from which the lemma follows.

**Lemma 4.** Assume \((A - \lambda B)\) has at least one rank-reducing number. Then either \(s + t = 0\) in (1) or there is a strictly smaller square matrix \((A^{(k)} - \lambda B^{(k)})\) such that \(s_k + t_k = 0\), \(B^{(k)}\) is an identity matrix, and \(\lambda\) is a rank-reducing number for \((A - \lambda B)\) if and only if it is a rank-reducing number for \((A^{(k)} - \lambda B^{(k)})\).

**Proof.** If \(s + t > 0\), define \((A^{(1)} - \lambda B^{(1)}) = (A - \lambda B)\) and define \((A^{(2)} - \lambda B^{(2)}) = (C_{11} - \lambda D_{11})\), where \(C_{11}\) and \(D_{11}\) are defined in (2) of Lemma 2. Note that \((A^{(2)} - \lambda B^{(2)})\) is of dimension strictly smaller than that of \((A^{(1)} - \lambda B^{(1)})\); by Lemma 3 these two matrices have the same rank-reducing numbers. Now put \((A^{(2)} - \lambda B^{(2)})\) into its normal form (1) and let its block dimensions be \(r_2, s_2, t_2\), and \(q_2\). If \(s_2 + t_2 = 0\), the lemma is proved; and if \(s_2 + t_2 > 0\) we can reapply the same procedure and define a strictly smaller matrix \((A^{(3)} - \lambda B^{(3)})\) that has the same rank-reducing numbers, and so on. At the \(j\)th step construct \((A^{(j)} - \lambda B^{(j)})\) whose size is \((r_j + s_j) \times (r_j + t_j)\) where \(r_j + s_j = r_{j-1} - t_{j-1}\) and \(r_j + t_j = r_{j-1} - s_{j-1}\). Since dimensions are reduced at each step the construction process must obviously stop after a finite number of steps with a matrix \((A^{(k)} - \lambda B^{(k)})\) such that \(s_k + t_k = 0\). Because of the assumption that there is a rank-reducing number, it follows that \(r_k > 0\) and \(B^{(k)}\) is an identity matrix. By the construction and Lemma 3, the rank-reducing numbers of \((A - \lambda B)\) are the same as those of \((A^{(k)} - \lambda B^{(k)})\).

**Remark 1.** If \((A - \lambda B)\) has no rank-reducing numbers, the procedure outlined in the proof of Lemma 4 can be carried out, but will terminate at the \((k + 1)\)st stage whenever \(r_k = s_k\) or \(r_k = t_k\) or both.

**Remark 2.** When \(s_k + t_k = 0\), \((A^{(k)} - \lambda B^{(k)})\) is square and \(B^{(k)}\) is an identity matrix. Hence if \((A - \lambda B)\) has at least one rank-reducing number, the procedure outlined in the proof of Lemma 4 will lead to an ordinary eigenvalue problem \((A^{(k)} - \lambda B^{(k)})\) where \(B^{(k)}\) is an identity matrix. Therefore all rank-reducing numbers of \((A^{(k)} - \lambda B^{(k)})\), and those of \((A - \lambda B)\) can be found by using ordinary eigenvalue techniques.

**Definition 1.** For \(m \times n\) matrices \(A\) and \(B\) we define the characteristic polynomial \(P(\lambda, A, B)\) of \(A\) relative to \(B\) as follows:

(a) If there are no rank-reducing numbers of \((A - \lambda B)\), then \(P(\lambda, A, B) = 1\).

(b) If there is at least one rank-reducing number of \((A - \lambda B)\), let \(P(\lambda, A, B)\) be the characteristic polynomial of \(E_{11}^{(k)}\) which is given
by the normal form of \( A^{(k)} - \lambda B^{(k)} \) where \( k \) is the smallest such that \( s_k + t_k = 0 \). (Note that \( E_{11}^{(k)} \) is the same as \( E_{11} \) in (1).)

Justification for the name characteristic polynomial is given below.

**Definition 2.** For any complex number \( \lambda \) define

(a) \( J(\lambda) = 0 \) if \( \lambda \) is not a root of \( P(\lambda, A, B) = 0 \),

(b) \( J(\lambda) = \) the number of Jordan blocks containing \( \lambda \) in the Jordan normal form of \( E_{11}^{(k)} \) where \( k \) is the smallest such that \( s_k + t_k = 0 \).

**Theorem.** For any complex number \( \lambda \)

\[
\text{rank}(A - \lambda B) = r + q - J(\lambda)
\]

where \( r \) and \( q \) are defined in (1).

**Proof.** If \( \lambda \) is not a rank-reducing number, the rank of the submatrix

\[
\begin{bmatrix}
E_{11} - \lambda I_r & E_{12} \\
E_{21} & 0
\end{bmatrix}
\]

in (1) is \( r \), so \( \text{rank}(A - \lambda B) = r + q \). If \( \lambda \) is a rank-reducing number then repeated applications of Lemmas 3 and 4 show that \( \lambda \) reduces the rank of the derived \( (E_{11}^{(k)} - \lambda I_{r}) \) matrix by \( J(\lambda) \) and hence reduces the rank of \( (A - \lambda B) \) by \( J(\lambda) \). This completes the proof.

Obviously the Cayley-Hamilton theorem, which states that a square matrix satisfies its own characteristic polynomial equation, cannot be extended to rectangular matrices since powers \( (> 1) \) of a rectangular matrix are not defined. However, the less emphasized property of a square matrix \( A \) that the roots of \( P(\lambda, A, I) = 0 \) are identical with the rank-reducing numbers of \( (A - \lambda I) \) can be extended to rectangular matrices as we have shown in this paper. Thus the corresponding fact that the roots of \( P(\lambda, A, B) = 0 \) are identical with the rank-reducing numbers of \( (A - \lambda B) \) provides the justification for calling \( P(\lambda, A, B) \) the characteristic polynomial of \( A \) with respect to \( B \).

The matrix \( (A - \lambda B) \) has been called a matrix pencil. Gantmacher [1, Chapter 12] develops a theory of matrix pencils and applies his results to quadratic forms and differential equations. Gantmacher transforms the \( m \times n \) matrix \( (A - \lambda B) \) to a rational, Jordan-like normal form as follows. Let the nonsingular \( m \times m \) matrix \( S \) and the nonsingular \( n \times n \) matrix \( T \) represent a series of elementary row and column operations, respectively. Then there exist \( S \) and \( T \) so that \( S(A - \lambda B)T \) is (rectangular) block diagonal with as many as \( (p + q + r + 2) \) blocks:
\[ \{ h[0], L_{s+1}, \ldots, L_{s+p}, L^T_{\eta h+1}, \ldots, L^T_{\eta h+p}, N_{\mu 1}, \ldots, N_{\mu r}, (J-\alpha I) \} \]

All other entries in \( S(A-\lambda B)T \) are zero. Neither the first nor the last block need exist for a given system nor need \( p, q, r \) be positive. The matrix \( S(B-\alpha A)T \) is the "direct sum" of the diagonal blocks. The block

\[ L_{h[0]} \]

has \( h \) rows and \( g \) columns and all its elements are zero. The block \( L_{s+1} \) has \( s+1 \) rows and \( s+1+1 \) columns with structure

\[ L_s = \begin{bmatrix} \lambda & 1 & 0 \\ \lambda & 1 & \cdot \\ \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot \\ \cdot & \cdot & \lambda & 1 \end{bmatrix} \]

and rank \( \epsilon \). The block \( L^T_{\eta h+j} \) has \( \eta h+j+1 \) rows and \( \eta h+j \) columns with structure

\[ L^T_\eta = \begin{bmatrix} \lambda & 0 \\ 1 & \lambda \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 0 & 1 & \lambda \end{bmatrix} \]

and rank \( \eta \). The square block \( N_{\mu k} \) is \( \mu_k \) by \( \mu_k \) with structure

\[ N_\mu = \begin{bmatrix} 1 & \lambda & 0 \\ \mu & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & \cdot & \cdot \end{bmatrix} \]

and rank \( \mu \). The final block \( J-\alpha I \) is an ordinary square eigensystem in Jordan normal form.

The Jordan normal form of the matrix \( (A^{(k)}-\lambda B^{(k)}) \) derived by our techniques when \( s_k+1\mu_k = 0 \) is identical to the matrix \( (J-\lambda I) \) in the Gantmacher normal form. The ranks of the matrices \( L, L^T, \) and...
N of the Gantmacher normal form are independent of \( \lambda \), so these matrices do not have rank-reducing numbers. Therefore, the only rank-reducing numbers of \((A - \lambda B)\) are those of \((J - \lambda I)\). Of course the rank-reducing numbers of \((J - \lambda I)\) are the eigenvalues of \(J\).

References


Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

University of Chicago, Chicago, Illinois 60637