Solving Homogeneous Systems
with Sub-matrices

Massoud Malek

Mathematics, California State University
East Bay, Hayward, CA 94542, USA

Abstract

We show that linearly independent solutions of $\mathcal{M}X = \theta$, where $\mathcal{M}$ is an $m \times n$ matrix, may be found by the largest non-singular sub-matrix of $\mathcal{M}$. With this method, we may also obtain eigenvectors and generalized eigenvectors corresponding to an eigenvalue $\lambda$. Finally, we shall explain how to construct a generalized modal matrix, to obtain a Jordan canonical form of a square matrix without solving a system except for finding the ranks.

Mathematics Subject Classification: 15B51

Keywords: Eigenvalues and eigenvectors; eigenspace; defective eigenvalue; defective matrix; generalized eigenvectors; generalized modal matrix; Jordan canonical form

1 Introduction

In all that follows the $n \times n$ identity matrix is denoted by $I_n$. A permutation matrix $\mathcal{P}$ is obtain from the identity matrix, by permuting some of its rows or columns. The zero column vector is denoted by $\theta$ and the $m \times n$ zero matrix is denoted by $\mathbb{Z}_{m \times n}$.

Let $\lambda$ be an eigenvalue of the $n \times n$ real or complex matrix $\mathcal{A}$. An eigenvector of $\mathcal{A}$ corresponding to the eigenvalue $\lambda$ is a non-trivial solution of
\((A - \lambda I_n)u = \theta\). The set of all such vectors is called the eigenspace corresponding to the eigenvalue \(\lambda\). The algebraic multiplicity of an eigenvalue is its multiplicity in the characteristic polynomial and its geometric multiplicity is the dimension of its eigenspace.

An eigenvalue is said to be defective, if its geometric multiplicity is less than its algebraic multiplicity. Matrices with some defective eigenvalues are called defective. These matrices are not diagonalizable.

If \(\lambda\) is defective, then for an integer \(k > 1\), any nonzero vector \(u(\lambda, k)\) satisfying:

\[
A^k_{\lambda} u(\lambda, k) = \theta \quad \text{with} \quad A^{k-1}_{\lambda} u(\lambda, k) \neq \theta
\]

is called a generalized eigenvector of order \(k\) corresponding to the eigenvalue \(\lambda\). Clearly the generalized eigenvector of order one is just an eigenvector.

In this paper, we show that linearly independent solutions of the homogeneous linear system \(MX = \theta\), where \(M\) is an \(m \times n\) matrix, may be obtain by using the largest non-singular sub-matrix of \(M\). The same method may be used to find all the eigenvectors and generalized eigenvectors of a square matrix corresponding to an eigenvalue.

If the rank of the \(m \times n\) matrix \(M\) is \(n\), then the dimension of the nullity of \(M\) is zero; this clearly implies that \(\theta\) is the only solution of \(MX = \theta\).

Given the \(m \times n\) matrix \(M\) of rank \(h < n\); we partition the matrix \(M\) or \(N = P M Q\) (permutation-equivalent to \(M\)), into

\[
M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \quad \text{or} \quad N = P M Q = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix},
\]

where at least one of the \(M_{ij}\) or \(N_{ij}\) blocks is a non-singular \(h \times h\) matrix.

2 Results

Let \(U\) be an \(n \times (n - h)\) matrix, where its columns represent \((n - h)\) linearly independent solutions of \(MX = \theta\). Clearly \(U\) is not unique.

Our main result explains how to find \(U\), using the inverse of an \(h \times h\) block of \(M\).

**Theorem 1.** Suppose the rank of the \(m \times n\) matrix \(M\) is \(h < n\).

1. If the block \(M_{11}\) is a non-singular \(h \times h\) matrix; then \(U = \begin{bmatrix} -M_{11}^{-1} M_{12} \\ I_{n-h} \end{bmatrix} \).

2. If the block \(M_{12}\) is a non-singular \(h \times h\) matrix; then \(U = \begin{bmatrix} I_{n-h} \\ -M_{12}^{-1} M_{11} \end{bmatrix} \).
3. If the block $\mathcal{M}_{21}$ is a non-singular $h \times h$ matrix; then

$$ U = \begin{bmatrix} \mathcal{M}_{11}^{-1} & \mathcal{M}_{12} \\ I_{n-h} & \end{bmatrix} . $$

4. If the block $\mathcal{M}_{22}$ is a non-singular $h \times h$ matrix; then

$$ U = \begin{bmatrix} I_{n-h} \\ \mathcal{M}_{22}^{-1} & \mathcal{M}_{12} \\ \end{bmatrix} . $$

**Proof.** Clearly, the rank of any $n \times (n-h)$ matrix containing $I_{n-h}$ is $n-h$.

(1) We have

$$ \begin{bmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} \\ -I_{n-h} & \end{bmatrix} = \begin{bmatrix} -\mathcal{M}_{11}^{-1} & \mathcal{M}_{12} \\ I_{n-h} & \end{bmatrix} = \mathcal{M}_{11} \mathcal{M}_{11}^{-1} \mathcal{M}_{12} + \mathcal{M}_{12} = Z_{m \times n-h} . $$

Since each row of $[ \mathcal{M}_{21} \mathcal{M}_{22} ]$ is a linear combination of the rows of $[ \mathcal{M}_{11} \mathcal{M}_{12} ]$, we have:

$$ \mathcal{M} \begin{bmatrix} -\mathcal{M}_{11}^{-1} \mathcal{M}_{12} \\ I_{n-h} \end{bmatrix} = \begin{bmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} \\ \mathcal{M}_{21} & \mathcal{M}_{22} \end{bmatrix} \begin{bmatrix} -\mathcal{M}_{11}^{-1} \mathcal{M}_{12} \\ I_{n-h} \end{bmatrix} = Z_{m \times n-h} . $$

Thus

$$ U = \begin{bmatrix} -\mathcal{M}_{11}^{-1} \mathcal{M}_{12} \\ I_{n-h} \end{bmatrix} . $$

The proofs of (2) through (4) are similar to the proof of the first part. \( \square \)

The matrix

$$ \mathcal{M} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} $$

of rank 2 cannot be partitioned into

$$ \mathcal{M} = \begin{bmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} \\ \mathcal{M}_{21} & \mathcal{M}_{22} \end{bmatrix} , $$

where one of the $\mathcal{M}_{ij}$ blocks is a $2 \times 2$ non-singular sub-matrix. But by interchanging the second and the third column, we may obtain a matrix with at least one

In our next corollary, we address this case.

**Corollary 1.** Let $\mathcal{P}$ and $\mathcal{Q}$ be two permutation matrices which make the block $\mathcal{N}_{11}$ of the matrix

$$ \mathcal{N} = \mathcal{P} \mathcal{M} \mathcal{Q} = \begin{bmatrix} \mathcal{N}_{11} & \mathcal{N}_{12} \\ \mathcal{N}_{21} & \mathcal{N}_{22} \end{bmatrix} , $$

an $h \times h$ invertible block. Then $U = \mathcal{Q} V$, where

$$ V = \begin{bmatrix} -\mathcal{N}_{11}^{-1} \mathcal{N}_{12} \\ I_{n-h} \end{bmatrix} . $$
**Proof.** By using the proof of Theorem 1 part (1) on the matrix $N$, we obtain

$$NV = Z_{m \times n - h}.$$  

The remainder of the proof then follows from:

$$MU = M[QV] = P^t N^t [QV] = P^t [NV] = P^t Z_{m \times n - h} = Z_{m \times n - h}.$$  

**Remark 1.** If the matrix $N$ in Corollary 1 is obtained from $M$ by using only some row operations (i.e., $Q = I_n$); then $U = V$.

**Eigenvectors.** Let $\lambda$ be an eigenvalue of the $n \times n$ matrix $A$ with geometric multiplicity $m$, and let $U_\lambda(A)$ be an $n \times m$ matrix, where its columns represent $m$ linearly independent eigenvectors of $A$ corresponding to the eigenvalue $\lambda$.

Since eigenvectors of $A$ corresponding to the eigenvalue $\lambda$ are solutions of the homogeneous linear system $(A - \lambda I_n)u = \theta$; by replacing the matrix $M$ with $A_\lambda = A - \lambda I_n$ and $h$ with $n - m$ in Theorem 1 or Corollary 1, one may obtain $U_\lambda(A)$.

**Remark 2.** With the conventional method, the matrix in order to obtain $m$ linearly independent eigenvectors of a matrix corresponding to an eigenvalue; one must first solve a homogenous system of linear equations which has infinitely many solutions. Then one must verify if they are linearly independent.

Contrary to the conventional method, the matrix $U_\lambda(A)$ gives all the linearly independent eigenvectors associated to the eigenvector $\lambda$. Also, the higher the geometric multiplicity of $\lambda$ is, the task of finding linearly independent eigenvectors associated to $\lambda$ becomes easier.

**Generalized Eigenvectors.** The fact that generalized eigenvectors are solutions of some homogeneous linear systems, Theorem 1 or Corollary 1 may be used to find them. Here is how:

**Step 1.** Obtain the matrix $A^k_\lambda = (A - \lambda I_n)^k$.

**Step 2.** Find $r_k$, the rank of the matrix $A^k_\lambda$.

**Step 3.** If $r_k = 0$ (i.e., the matrix $A^k_\lambda$ is the zero matrix), then any vector $v$ satisfying $A^{k-1}_\lambda v \neq \theta$ is a generalized eigenvector of $A$ of order $k$. If not, replace the matrix $M$ with $A^k_\lambda$ and $h$ with $n - r_k$ in Theorem 1 or Corollary 1; any column $v$ of $U_\lambda(A^k_\lambda)$ satisfying

$$A^k_\lambda v \neq \theta$$

represents a generalized eigenvector of order $k$ of $A$ corresponding to the eigenvalue $\lambda$. 


Jordan Canonical Form. Any $n \times n$ defective matrix $A$ can be put in Jordan canonical form by a similarity transformation, i.e.,

$$M^{-1}AM = J = \begin{bmatrix} J_1 & & J_2 & & & \cdots \ & & & \ & J_m & & & & \end{bmatrix},$$

where $J_i = \begin{bmatrix} \lambda_i & 1 & & & & \ & \lambda_i & \ddots & & & \ & & \ddots & 1 \end{bmatrix}$ is of size $n_i \times n_i$ with $\sum_{i=1}^{m} n_i = n$. The blocks $J_1, J_2, \ldots, J_m$ are called Jordan blocks. The matrix $M$ is called a generalized modal matrix for $A$ and is obtained from eigenvectors and generalized eigenvectors of the matrix $A$. The order of the largest Jordan block of $A$ corresponding to an eigenvalue $\lambda$ is called the index of $\lambda$. It is the smallest value of $k \in \mathbb{N}$ such that

$$\text{rank} \left( A - \lambda I_n \right)^k = \text{rank} \left( A - \lambda I_n \right)^{k+1}.$$

Let $k$ be the smallest positive integer such that $A_\lambda^k u = \theta$. Then the sequence

$$A_\lambda^{k-1} u, A_\lambda^{k-2} u, \ldots, A_\lambda^2 u, A_\lambda u, u$$

is called a Jordan chain of linearly independent generalized eigenvectors of length $k$.

To find a Jordan chain of length $k$ corresponding to a defective eigenvalue $\lambda$, one must solve the equation $A_\lambda v = u_\lambda$. If there are more eigenvectors associated with the defective eigenvalue $\lambda$, then it is not always clear which eigenvector must be chosen to solve a non-homogeneous linear system in order to produce the generalized eigenvector.

Instead of starting with an eigenvector which may or may not produce a Jordan chain of length $k$, we start with the matrix $A_\lambda^k$, which has a smaller non-singular block than $A_\lambda^{k-1}$. The matrix $U_\lambda(A, k)$ has a column $j$ such that for $i = 1, 2, \ldots, k-1$, the $j^{th}$ columns of $A_\lambda^i U_\lambda(A, k)$ produce a Jordan chain of length $k$.

We shall explain this procedure in more detail with two examples.

**Example 1.** Consider the matrix $A = \begin{bmatrix} -7 & -4 & 6 & 9 \\ -11 & 0 & 6 & 9 \\ -11 & -4 & 10 & 9 \\ -11 & -4 & 6 & 13 \end{bmatrix}$ with the characteristic polynomials $K_A(\lambda) = (\lambda - 4)^4$. Thus $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 4$. 

Solving homogeneous systems with sub-matrices
Let \( \mathcal{A}_4 = \mathcal{A} - 4I_4 \), then

\[
\mathcal{A}_4 = \begin{bmatrix}
-11 & -4 & 6 & 9 \\
-11 & -4 & 6 & 9 \\
-11 & -4 & 6 & 9 \\
-11 & -4 & 6 & 9
\end{bmatrix}
\quad \text{and} \quad
\mathcal{A}_4^2 = \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}.
\]

Since the rank of \( \mathcal{A}_4 \) is one, then the geometric multiplicity of \( \lambda = 4 \) in \( \mathcal{A} \) is three; so according to Theorem 1, we need a \( 1 \times 1 \) non-singular block of \( \mathcal{A}_4 \). Hence

\[
U_{\mathcal{A}}(4, 1) = \frac{1}{11} \begin{bmatrix}
-4 & 6 & 9 \\
11 & 0 & 0 \\
0 & 11 & 0 \\
0 & 0 & 11
\end{bmatrix}
\implies u_1 = \begin{bmatrix}
-4 \\
11 \\
0 \\
0
\end{bmatrix}, \quad u_2 = \begin{bmatrix}
6 \\
0 \\
11 \\
0
\end{bmatrix}, \quad \text{and} \quad u_3 = \begin{bmatrix}
9 \\
0 \\
0 \\
11
\end{bmatrix}.
\]

Now, to obtain a generalized modal matrix of \( \mathcal{A} \), we need a generalized eigenvector associated with one of the above eigenvectors. With the conventional method, we have to solve one of the following matrix equations:

\[
(\mathcal{A} - 4I_4) v = u_i \quad \text{for} \quad i = 1, 2, 3,
\]

Notice that all the rows of the matrix \( \mathcal{A}_4 = (\mathcal{A} - 4I_4) \) are the same but the entries of \( u_i \)'s are different. This means that we must find another eigenvector in order to produce a generalized modal matrix.

Since \( \mathcal{A}_4^2 \) is a zero matrix, any non-zero vector which is not an eigenvector of \( \mathcal{A} \) becomes a generalized eigenvector of order one. So we may choose for example the vector \( e_1 = [1 \ 0 \ 0 \ 0]^t \) as our generalized eigenvector. Then from \( e_1 \), we get a new eigenvector

\[
v = \mathcal{A}_4 e_1 = [-11 \ -11 \ -11 \ -11]^t.
\]

By using \( v \), \( e_1 \), \( u_2 \), and \( u_3 \), we construct the generalized modal matrix \( \mathcal{M} = [v \ e_1 \ u_2 \ u_3] \). Then we have

\[
\mathcal{M} = \begin{bmatrix}
-11 & 1 & 6 & 9 \\
-11 & 0 & 0 & 0 \\
-11 & 0 & 11 & 0 \\
-11 & 0 & 0 & 11
\end{bmatrix}
\quad \text{with} \quad
\mathcal{M}^{-1} \mathcal{A} \mathcal{M} = J(\mathcal{A}) = \begin{bmatrix}
4 & 1 & 0 & 0 \\
0 & 4 & 0 & 0 \\
0 & 0 & 4 & 0 \\
0 & 0 & 0 & 4
\end{bmatrix}.
\]

We conclude our paper with an example of a defective matrix with different eigenvalues.
Example 2. Consider the matrix \( A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 4 & 3 & 2 & 0 & 0 \\ 5 & 4 & 3 & 2 & 0 \\ 6 & 5 & 4 & 3 & 2 \end{bmatrix} \) with characteristic polynomial \( K_A(\lambda) = (\lambda - 1)^2(\lambda - 2)^3 \).

Define \( A_1 = A - I_5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 3 & 1 & 0 & 0 \\ 5 & 4 & 3 & 1 & 0 \\ 6 & 5 & 4 & 3 & 1 \end{bmatrix} \) and \( A_2 = A - 2I_5 = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 3 & -1 & 0 & 0 & 0 \\ 4 & 3 & 0 & 0 & 0 \\ 5 & 4 & 3 & 0 & 0 \\ 6 & 5 & 4 & 3 & 0 \end{bmatrix} \).

The geometric multiplicities of both \( \lambda_1 = 1 \) and \( \lambda_2 = 2 \) are one. So we need

\[
A_1^2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 13 & 3 & 1 & 0 & 0 \\ 29 & 13 & 6 & 1 & 0 \\ 52 & 29 & 17 & 6 & 1 \end{bmatrix}, \quad \text{and} \quad A_2^3 = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 9 & -1 & 0 & 0 & 0 \\ -14 & 3 & 0 & 0 & 0 \\ -4 & -5 & 0 & 0 & 0 \\ 53 & 8 & 0 & 0 & 0 \end{bmatrix}.
\]

By using Theorem 1 for \( A_1^2 \) and \( A_2^3 \), we obtain

\[
V = U_A(1, 2) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -13 & -3 \\ 49 & 5 \\ -125 & -8 \end{bmatrix} \quad \text{and} \quad W = U_A(2, 3) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.
\]

We have

\[
A_1 V = \begin{bmatrix} 0 & 0 \\ 3 & 0 \\ -9 & 0 \\ 15 & 0 \\ -24 & 0 \end{bmatrix}, \quad A_2 W = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 3 & 0 & 0 \\ 4 & 3 & 0 \end{bmatrix}, \quad \text{and} \quad A_2^2 W = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.
\]
To construct a generalized modal matrix of $\mathcal{A}$ from $V$ and $W$, we must use the first columns of $\mathcal{A}_1 V, V, \mathcal{A}_2^2 W, \mathcal{A}_2 W,$ and $W$, in that order.

$$\mathcal{M} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \\ -9 & -13 & 0 & 0 & 1 \\ 15 & 49 & 0 & 3 & 0 \\ -24 & -125 & 9 & 4 & 0 \end{bmatrix} \quad \text{with} \quad \mathcal{M}^{-1} \mathcal{A} \mathcal{M} = J(\mathcal{A}) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$  

Notice that no linear system was solved and there was no need to find any eigenvector of the matrix $\mathcal{A}$ directly.

References


Received: April 29, 2018; Published: June 28, 2018