An Application of Sylvester's Rank Inequality

Sidney H. Kung (sidneykung@yahoo.com), Cupertino, CA 95014

Concerning the diagonalization of matrices, the following results are well known: [D1] an $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors, and [D2] an $n \times n$ matrix A is diagonalizable if and only if for each eigenvalue λ_j $(j = 1, 2, \ldots, m)$, $\operatorname{rank}(\lambda_j I - A) = n - r_j$, where r_j is the multiplicity of λ_j.

To see whether or not a given matrix A meets these conditions, we normally: (i) find the characteristic polynomial $p(\lambda) = \det(\lambda I - A)$, (ii) find the distinct roots $\lambda_1, \lambda_2, \ldots, \lambda_m$ of $p(\lambda) = 0$, (iii) solve the homogeneous system of equations $(\lambda_i I - A)\vec{x} = \vec{0}$ corresponding to each λ_i. $i = 1, 2, \ldots, m$.

A less commonly used condition for diagonalization of a matrix in an undergraduate linear algebra course is the following:

Theorem 1. *Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be the distinct eigenvalues of an $n \times n$ matrix A. If*

$$\prod_{i=1}^{m} (\lambda_i I - A) = O, \tag{1}$$

where O is the $n \times n$ zero matrix, then A is diagonalizable.

As an example, we consider the matrix

$$A = \begin{bmatrix} 3 & 2 & -1 \\ -2 & -2 & 2 \\ 3 & 6 & -1 \end{bmatrix}.$$

It has two distinct eigenvalues $\lambda_1 = -4$ and $\lambda_2 = 2$. By direct computation,

$$(\lambda_1 I - A)(\lambda_2 I - A) = \begin{bmatrix} -7 & -2 & 1 \\ 2 & -2 & -2 \\ -3 & -6 & -3 \end{bmatrix} \begin{bmatrix} -1 & -2 & 1 \\ 2 & 4 & -2 \\ -3 & -6 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus, A is diagonalizable by Theorem 1.

doi:10.4169/college.math.j.42.2.148

Typically Theorem 1 is proved using the Jordan canonical form of matrix A (see, for example, [**2**]). Here we give a proof using an extended form of Sylvester's rank inequality. If A_1 and A_2 are two $n \times n$ matrices, then Sylvester's rank inequality (see [**1**]) states

$$\operatorname{rank}(A_1) + \operatorname{rank}(A_2) \leq n + \operatorname{rank}(A_1 A_2).$$

This inequality can be extended to m matrices $A_1, A_2, \ldots, A_m$:

$$\begin{aligned}\sum_{i=1}^{m} \operatorname{rank}(A_i) &\leq n + \operatorname{rank}(A_1 A_2) + \cdots + \operatorname{rank}(A_{m-1}) + \operatorname{rank}(A_m) \\ &\vdots \\ &\leq (m-2)n + \cdots + \operatorname{rank}(A_1 A_2 \cdots A_{m-1}) + \operatorname{rank}(A_m) \qquad (2) \\ &\leq (m-1)n + \operatorname{rank}(A_1 A_2 \cdots A_{m-1} A_m).\end{aligned}$$

Suppose that $\lambda_1, \lambda_2, \ldots, \lambda_m$ are distinct eigenvalues of matrix A and that condition (1) holds. Then, by applying (2) we have

$$\sum_{i=1}^{m} \operatorname{rank}(\lambda_i I - A) \leq (m-1)n.$$

Thus the number, l, of linearly independent eigenvectors of A is

$$\begin{aligned} l &= \sum_{i=1}^{m} [n - \operatorname{rank}(\lambda_i I - A)] \\ &= mn - \sum_{i=1}^{m} [\operatorname{rank}(\lambda_i I - A)] \\ &\geq mn - (m-1)n = n. \end{aligned}$$

However, l must be less or equal to n. Hence, $l = n$. By [D1], A is diagonalizable.

Equation (1) also provides a method for finding the eigenvectors associated with each eigenvalue. First note that (1) can be rewritten as

$$(\lambda_j I - A) \prod_{i=1, \neq j}^{m} (\lambda_i I - A) = O.$$

This shows that the column vectors of $\prod_{i=1, i\neq j}^{m} (\lambda_i I - A)$ are solutions of the system $(\lambda_j I - A)\vec{x} = \vec{0}$. According to [D2], $\operatorname{rank}(\lambda_j I - A) = n - r_j$, so the dimension of the solution space of $(\lambda_j I - A)\vec{x} = \vec{0}$ is r_j. So, $\operatorname{rank}\left(\prod_{i=1, i\neq j}^{m} (\lambda_i I - A)\right) \leq r_j$. However, by the extended Sylvester's rank inequality, we have

$$\begin{aligned}\operatorname{rank}\left(\prod_{i=1, i\neq j}^{m} (\lambda_i I - A)\right) &\geq \sum_{i=1, i\neq j}^{m} r(\lambda_i I - A) - (m-2)n \\ &= (m-1)n - \sum_{i=1, i\neq j}^{m} r_i - (m-2)n = r_j.\end{aligned}$$

Hence, rank $\left(\prod_{i=1,i\neq j}^{m}(\lambda_i I - A)\right)$ must equal to r_j. This implies that a Maximal Linearly Independent Set (MLIS) of the column vectors of matrix

$$\prod_{i=1,i\neq j}^{m}(\lambda_i I - A) \tag{3}$$

contains r_j vectors and comprises a basis for the eigenspace corresponding to eigenvalue λ_j.

For the previous example, we can reduce matrix $\lambda_1 I - A$ to echelon form and find its MLIS (see [**3**]):

$$\left\{\begin{bmatrix}-7\\2\\-3\end{bmatrix}, \begin{bmatrix}-2\\-2\\-6\end{bmatrix}\right\}.$$

This is a basis for the eigenspace of A corresponding to $\lambda_2 = 2$. Similarly, an MLIS for $\lambda_2 I - A$ is

$$\left\{\begin{bmatrix}-1\\2\\-3\end{bmatrix}\right\}.$$

This is a basis for the eigenspace of A associated with $\lambda_1 = -4$. Now let

$$P = \begin{bmatrix}-1 & -7 & -2\\ 2 & 2 & -2\\ -3 & -3 & -6\end{bmatrix}.$$

We obtain

$$P^{-1}AP = \begin{bmatrix}-4 & & \\ & 2 & \\ & & 2\end{bmatrix}.$$

Finally we remark that condition (1) is also necessary. A proof can be found in [**2**, p. 145].

Acknowledgments. The author would like to thank the referee and the editors for their suggestions that helped to improve the quality of this article.

Summary. Using two well known criteria for the diagonalizability of a square matrix plus an extended form of Sylvester's Rank Inequality, the author presents a new condition for the diagonalization of a real matrix from which one can obtain the eigenvectors by simply multiplying some associated matrices without solving a linear system of simultaneous equations.

References

1. F. E. Hohn, *Elementary Matrix Algebra*, 2nd ed., Macmillan, New York, 1964.
2. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1990.
3. D. C. Lay, *Linear Algebra and Its Applications*, 3rd ed., Addison Wesley, 2006.