Section 2.5. Partitioned Matrices

Note. In this section we consider matrices which we partition into four submatrices. We consider their algebraic properties, inverses, determinants, and some eigenvalue properties. These ideas are also covered in Theory of Matrices; see my online notes on Section 3.1. Basic Definitions and Notations where a partitioned matrix is defined (see page 5), and Section 3.4 More on Partitioned Square Matrices: The Schur Complement where additional properties are given. .

Definition. Matrix $\mathcal{A} = \mathcal{A}(n \times p)$ can be *partitioned* into four matrices $\mathcal{A}_{11}$, $\mathcal{A}_{12}$, $\mathcal{A}_{21}$, and $\mathcal{A}_{22}$ as

$$\mathcal{A} = \begin{pmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{pmatrix}.$$

where $\mathcal{A}_{ij} = \mathcal{A}_{ij}(n_i \times p_j)$ where $i, j \in \{1, 2\}$, $n_1 + n_2 = n$, and $p_1 + p_2 = p$.

Theorem 2.5.A. Algebraic Properties of Partitioned Matrices.

Let $\mathcal{A}_{ij} = \mathcal{A}_{ij}(n_i \times p_j)$ and $\mathcal{B}_{ij} = \mathcal{B}_{ij}(n_i \times p_j)$ be partitioned as given above (with corresponding parts of the same sizes). Then:

$$\mathcal{A} + \mathcal{B} = \begin{pmatrix} \mathcal{A}_{11} + \mathcal{B}_{11} & \mathcal{A}_{12} + \mathcal{B}_{12} \\ \mathcal{A}_{21} + \mathcal{B}_{21} & \mathcal{A}_{22} + \mathcal{B}_{22} \end{pmatrix},$$

$$\mathcal{B}^\top = \begin{pmatrix} \mathcal{B}_{11}^\top & \mathcal{B}_{21}^\top \\ \mathcal{B}_{12}^\top & \mathcal{B}_{22}^\top \end{pmatrix}, \text{ and}$$

$$\mathcal{A}\mathcal{B}^\top = \begin{pmatrix} \mathcal{A}_{11}\mathcal{B}_{11}^\top + \mathcal{A}_{12}\mathcal{B}_{12}^\top & \mathcal{A}_{11}\mathcal{B}_{21}^\top + \mathcal{A}_{12}\mathcal{B}_{22}^\top \\ \mathcal{A}_{21}\mathcal{B}_{11}^\top + \mathcal{A}_{22}\mathcal{B}_{12}^\top & \mathcal{A}_{21}\mathcal{B}_{21}^\top + \mathcal{A}_{22}\mathcal{B}_{22}^\top \end{pmatrix}.$$

The first two properties are fairly clear. A proof of the third property follows from Theorem 3.2.2 of my online notes for Theory of Matrices (MATH 5090) on Section 3.2. Multiplication of Matrices and Multiplication of Vectors and Matrices.

Theorem 2.5.B. Let $\mathcal{A} = \mathcal{A}(p \times p)$ be nonsingular and suppose

$$\mathcal{A} = \begin{pmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{pmatrix},$$

where $\mathcal{A}_{11}$ and $\mathcal{A}_{22}$ are square. Then

$$\mathcal{A}^{-1} = \begin{pmatrix} \mathcal{A}^{11} & \mathcal{A}^{12} \\ \mathcal{A}^{21} & \mathcal{A}^{22} \end{pmatrix},$$

where the parts of $\mathcal{A}^{-1}$ in terms of the parts of $\mathcal{A}$ are:

$$\begin{aligned} \mathcal{A}^{11} &= (\mathcal{A}_{11} - \mathcal{A}_{12}\mathcal{A}_{22}^{-1}\mathcal{A}_{21})^{-1} = \mathcal{B}^{-1} \\ \mathcal{A}^{12} &= -\mathcal{B}^{-1}\mathcal{A}_{12}\mathcal{A}_{22}^{-1} \\ \mathcal{A}^{21} &= -\mathcal{A}_{22}^{-1}\mathcal{A}_{21}\mathcal{B}^{-1} \\ \mathcal{A}^{22} &= \mathcal{A}_{22}^{-1} + \mathcal{A}_{22}^{-1}\mathcal{A}_{21}\mathcal{B}^{-1}\mathcal{A}_{12}\mathcal{A}_{22}^{-1} \end{aligned}$$

Note. Härdle and Simar make a comment about an "alternative expression" for $\mathcal{A}^{-1}$ which results from interchanging $\mathcal{A}_{11}$ and $\mathcal{A}_{22}$. I think it is a reference to the Schur complement and its use in finding the inverse of a partitioned matrix; for details, see my online Theory of Matrices (MATH 5090) notes on Section 3.4. More on Partitioned Square Matrices: The Schur Complement.

Theorem 2.5.C. Suppose square matrix $\mathcal{A}$ has partition

$$\mathcal{A} = \begin{pmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{pmatrix}$$

where $\mathcal{A}_{11}$ and $\mathcal{A}_{22}$ are square. Let $\mathcal{B} = \mathcal{A}_{11} - \mathcal{A}_{12}\mathcal{A}_{22}^{-1}\mathcal{A}_{21}$, as in Theorem 2.5.B.

1. If $\mathcal{A}_{11}$ is nonsingular, then

$$|\mathcal{A}| = |\mathcal{A}_{11}|\,|\mathcal{A}_{22} - \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\mathcal{A}_{12}|.$$

2. If $\mathcal{A}_{22}$ is nonsingular, then

$$|\mathcal{A}| = |\mathcal{A}_{22}|\,|\mathcal{A}_{11} - \mathcal{A}_{12}\mathcal{A}_{22}^{-1}\mathcal{A}_{21}|.$$

The proof of claim 1 is given in Theory of Matrices (see Theorem 3.4.3 in Section 3.4. More on Partitioned Square Matrices: The Schur Complement). The proof of claim 2 is similar (apply a sequence of row interchanges and column interchanges to $\mathcal{A}$).

Note. As a special case, suppose $\mathcal{B} = \mathcal{B}((p+1) \times (p+1))$ is partitioned as

$$\mathcal{B} = \begin{pmatrix} 1 & b^\top \\ a & \mathcal{A} \end{pmatrix}$$

where a and b are $p \times 1$ vectors and $\mathcal{A}$ is non-singular. We then have by Theorem 2.5.C(2) that $|\mathcal{B}| = |\mathcal{A}|\,|(1) - (b^\top)\mathcal{A}^{-1}(a)|$. We can perform elementary row operations on $\mathbb{B}$ using the first row to eliminate all entries under the 1 to get

$$\mathcal{B} = \begin{pmatrix} 1 & b^\top \\ a & \mathcal{A} \end{pmatrix} \sim \begin{pmatrix} 1 & b^\top \\ 0 & \mathcal{A} - ab^\top \end{pmatrix},$$

from which we can calculate the determinate by expanding along the first column to get $|\mathcal{B}| = |A - ab^\top|$, So we have

$$|\mathcal{B}| = |\mathcal{A} - ab^\top| = |\mathcal{A}|\,|1 - b^\top \mathcal{A}^{-1} a|.$$

Note. Let $\mathcal{A} = \mathcal{A}(n \times p)$ and $\mathcal{B}(p \times n)$ be matrices where $n \geq p$. Consider $\begin{pmatrix} -\lambda\mathcal{I}_n & -\mathcal{A} \\ \mathcal{B} & \mathcal{I}_p \end{pmatrix}$. By Theorem 2.5.C we get two representations of the determinant of this matrix:

$$\begin{vmatrix} -\lambda\mathcal{I}_n & -\mathcal{A} \\ \mathcal{B} & \mathcal{I}_p \end{vmatrix} = |-\lambda\mathcal{I}_n|\,|(\mathcal{I}_p) - (\mathcal{B})(-\lambda\mathcal{I}_n)^{-1}(-\mathcal{A})| = (-\lambda)^n \left|\mathcal{I}_p + \mathcal{B}\left(\frac{1}{-\lambda}\right)\mathcal{I}_n\mathcal{A}\right|$$

$$= (-\lambda)^n \left|\mathcal{I}_p - \left(\frac{1}{\lambda}\right)\mathcal{B}\mathcal{A}\right| = (-\lambda)^{-n}\left|\frac{1}{-\lambda}(\mathcal{B}\mathcal{A} - \lambda\mathcal{I}_p)\right|$$

$$= (-\lambda)^n \left(\frac{1}{-\lambda}\right)^p |\mathcal{B}\mathcal{A} - \lambda\mathcal{I}_p| = (-\lambda)^{n-p}|\mathcal{B}\mathcal{A} - \lambda\mathcal{I}_p|$$

and

$$\begin{vmatrix} -\lambda\mathcal{I}_n & -\mathcal{A} \\ \mathcal{B} & \mathcal{I}_p \end{vmatrix} = |\mathcal{I}_p|\,|(-\lambda\mathcal{I}_n) - (-\mathcal{A})(\mathcal{I}_p)^{-1}(\mathcal{B})| = |-\lambda\mathcal{I}_n + \mathcal{A}\mathcal{B}| = |\mathcal{A}\mathcal{B} - \lambda\mathcal{I}_n|.$$

Combining we have

$$\begin{vmatrix} -\lambda\mathcal{I}_n & -\mathcal{A} \\ \mathcal{B} & \mathcal{I}_p \end{vmatrix} = (-\lambda)^{n-p}|\mathcal{B}\mathcal{A} - \lambda\mathcal{I}_p| = |\mathcal{A}\mathcal{B} - \lambda\mathcal{I}_n|.$$

Since $|\mathcal{A}\mathcal{B} - \lambda\mathcal{I}_n|$ is the characteristic polynomial for matrix $\mathcal{A}\mathcal{B}$, and $(-\lambda)^{n-p}|\mathcal{B}\mathcal{A} - \lambda\mathcal{I}_p|$ involves the characteristic polynomial of matrix $\mathcal{B}\mathcal{A}$ then we see that the n eigenvalues of $\mathbb{A}\mathbb{B}$ include the p eigenvalues of $\mathcal{B}\mathcal{A}$ along with the eigenvalue 0 with multiplicity $n - p$. This partially gives the following theorem.

Theorem 2.6. For $\mathcal{A} = \mathcal{A}(n \times p)$ and $\mathcal{B} = \mathcal{B}(p \times n)$, the nonzero eigenvalue of $\mathcal{AB}$ and $\mathcal{BA}$ are the same and have the same multiplicity. If x is an eigenvector of $\mathcal{AB}$ for an eigenvalue $\lambda \neq 0$, then $y = \mathcal{B}x$ is an eigenvector of $\mathcal{BA}$.

Note. Another proof of the first claim of Theorem 2.6 can be found in the Theory of Matrices (MATH 5090) class; see Theorem 3.8.2(7) in Section 3.8. Eigenanalysis; Canonical Factorizations, though the proof is to be given in Exercise 3.16. The claim about eigenvectors is straightforward:

$$\mathcal{BA}y = \mathcal{BA}(\mathcal{B}x) = \mathcal{B}(\mathcal{AB}x) = \mathcal{B}(\lambda x) = \lambda(\mathcal{B}x) = \lambda y.$$

Notice that $\lambda \neq 0$ is necessary because we cannot have 0 as an eigenvector (by definition).

Corollary 2.2. For $\mathcal{A}(n \times p)$, $\mathcal{B}(q \times n)$, $a(p \times 1)$, and $b(q \times 1)$ we have $\text{rank}(\mathcal{A}ab^\top\mathcal{B}) \leq 1$. The nonzero eigenvalue, if it exists, equals $b^\top\mathcal{BA}a$ (with eigenvector $\mathcal{A}a$).

Revised: 5/2/2021