11. Partitioned Matrices

Given an $(n \times n)$ matrix, we can break it up into blocks of 'submatrices'. For instance, we can write

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 2 & 6 \\ 2 & 8 & 2 & 1 \\ 3 & 1 & 2 & 4 \\ 4 & 2 & 1 & 3 \\ 3 & 1 & 1 & 7 \end{bmatrix} = \left[\begin{array}{c|ccc} 1 & 3 & 2 & 6 \\ 2 & 8 & 2 & 1 \\ \hline 3 & 1 & 2 & 4 \\ 4 & 2 & 1 & 3 \\ 3 & 1 & 1 & 7 \end{array}\right] = \begin{bmatrix} \mathbf{a}_{11} & \mathbf{A}_{12} \\ \mathbf{a}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

where $\mathbf{a}_{11} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{A}_{12} = \begin{bmatrix} 3 & 2 & 6 \\ 8 & 2 & 1 \end{bmatrix}$, $\mathbf{a}_{21} = \begin{bmatrix} 3 \\ 4 \\ 3 \end{bmatrix}$, $\mathbf{A}_{22} = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 1 & 3 \\ 1 & 1 & 7 \end{bmatrix}$.

There are of course many ways of partitioning the same matrix.

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 2 & 6 \\ 2 & 8 & 2 & 1 \\ 3 & 1 & 2 & 4 \\ 4 & 2 & 1 & 3 \\ 3 & 1 & 1 & 7 \end{bmatrix} = \left[\begin{array}{c|ccc} 1 & 3 & 2 & 6 \\ 2 & 8 & 2 & 1 \\ \hline 3 & 1 & 2 & 4 \\ 4 & 2 & 1 & 3 \\ 3 & 1 & 1 & 7 \end{array}\right] = \left[\begin{array}{cc|cc} 1 & 3 & 2 & 6 \\ 2 & 8 & 2 & 1 \\ 3 & 1 & 2 & 4 \\ \hline 4 & 2 & 1 & 3 \\ 3 & 1 & 1 & 7 \end{array}\right].$$

Addition of Partitioned Matrices

Consider two $(n \times m)$ matrices $\mathbf{A}$ and $\mathbf{B}$

$$\mathbf{A} = \begin{bmatrix} \underbrace{\mathbf{A}_{11}}_{n_1 \times m_1} & \underbrace{\mathbf{A}_{12}}_{n_1 \times m_2} \\ \underbrace{\mathbf{A}_{21}}_{n_2 \times m_1} & \underbrace{\mathbf{A}_{22}}_{n_2 \times m_2} \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} \underbrace{\mathbf{B}_{11}}_{n_1 \times m_1} & \underbrace{\mathbf{B}_{12}}_{n_1 \times m_2} \\ \underbrace{\mathbf{B}_{21}}_{n_2 \times m_1} & \underbrace{\mathbf{B}_{22}}_{n_2 \times m_2} \end{bmatrix}$$

where $n_1 + n_2 = n$ and $m_1 + m_2 = m$. We emphasize that **A** and **B** are of the same size and partitioned in a similar fashion. Then

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} \underbrace{\mathbf{A}_{11} + \mathbf{B}_{11}}_{n_1 \times m_1} & \underbrace{\mathbf{A}_{12} + \mathbf{B}_{12}}_{n_1 \times m_2} \\ \underbrace{\mathbf{A}_{21} + \mathbf{B}_{21}}_{n_2 \times m_1} & \underbrace{\mathbf{A}_{22} + \mathbf{B}_{22}}_{n_2 \times m_2} \end{bmatrix}. \qquad (11.1)$$

If **A** and **B** are of the same size but partitioned differently, we can of course still add $\mathbf{A} + \mathbf{B}$. However, the formula on the right-hand side will not be valid.

Multiplication of Partitioned Matrices

Consider two matrices $\mathbf{A}$ and $\mathbf{B}$ of dimensions $(n \times p)$ and $(p \times m)$ respectively. Suppose they are partitioned as follows:

$$\mathbf{A} = \begin{bmatrix} \underbrace{\mathbf{A}_{11}}_{n_1 \times p_1} & \underbrace{\mathbf{A}_{12}}_{n_1 \times p_2} \\ \underbrace{\mathbf{A}_{21}}_{n_2 \times p_1} & \underbrace{\mathbf{A}_{22}}_{n_2 \times p_2} \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} \underbrace{\mathbf{B}_{11}}_{p_1 \times m_1} & \underbrace{\mathbf{B}_{12}}_{p_1 \times m_2} \\ \underbrace{\mathbf{B}_{21}}_{p_2 \times m_1} & \underbrace{\mathbf{B}_{22}}_{p_2 \times m_2} \end{bmatrix}.$$

We emphasize that the column-wise partition of **A** must match the row-wise partition of **B**. Then

$$\mathbf{AB} = \begin{bmatrix} \underbrace{\mathbf{A}_{11}}_{n_1 \times p_1} & \underbrace{\mathbf{A}_{12}}_{n_1 \times p_2} \\ \underbrace{\mathbf{A}_{21}}_{n_2 \times p_1} & \underbrace{\mathbf{A}_{22}}_{n_2 \times p_2} \end{bmatrix} \begin{bmatrix} \underbrace{\mathbf{B}_{11}}_{p_1 \times m_1} & \underbrace{\mathbf{B}_{12}}_{p_1 \times m_2} \\ \underbrace{\mathbf{B}_{21}}_{p_2 \times m_1} & \underbrace{\mathbf{B}_{22}}_{p_2 \times m_2} \end{bmatrix} = \begin{bmatrix} \underbrace{\mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21}}_{n_1 \times m_1} & \underbrace{\mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22}}_{n_1 \times m_2} \\ \underbrace{\mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21}}_{n_2 \times m_1} & \underbrace{\mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22}}_{n_2 \times m_2} \end{bmatrix}. \tag{11.2}$$

The partitioned matrix multiplication follows in similar fashion to the usual matrix multiplication, with “rows diving into columns”. Care nonetheless must be taken to ensure that the submatrices are compatible for multiplication.

Transposition of Partitioned Matrices

$$\text{If } \mathbf{A} = \begin{bmatrix} \underbrace{\mathbf{A}_{11}}_{n_1 \times p_1} & \underbrace{\mathbf{A}_{12}}_{n_1 \times p_2} \\ \underbrace{\mathbf{A}_{21}}_{n_2 \times p_1} & \underbrace{\mathbf{A}_{22}}_{n_2 \times p_2} \end{bmatrix}, \text{ then } \mathbf{A}^{\mathrm{T}} = \begin{bmatrix} \underbrace{\mathbf{A}_{11}^{\mathrm{T}}}_{p_1 \times n_1} & \underbrace{\mathbf{A}_{21}^{\mathrm{T}}}_{p_1 \times n_2} \\ \underbrace{\mathbf{A}_{12}^{\mathrm{T}}}_{p_2 \times n_1} & \underbrace{\mathbf{A}_{22}^{\mathrm{T}}}_{p_2 \times n_2} \end{bmatrix}$$

<u>Example</u> Let X be a data matrix containing *n* observations of three variables X_1, X_2, and X_3. The data for each variable are stored in columns. Let x_{ij} represent the *i*th observation of variable X_j

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & x_{n3} \end{bmatrix}.$$

We can write this matrix to emphasize observations, or variables. If

$$\mathbf{X}_1 = \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \\ \vdots \\ x_{n1} \end{bmatrix}, \ \mathbf{X}_2 = \begin{bmatrix} x_{12} \\ x_{22} \\ x_{32} \\ \vdots \\ x_{n2} \end{bmatrix}, \text{ and } \mathbf{X}_3 = \begin{bmatrix} x_{13} \\ x_{23} \\ x_{33} \\ \vdots \\ x_{n3} \end{bmatrix},$$

i.e., $\mathbf{X}_j$ contains all *n* observations of variable X_j, then we can write **X** as the partitioned matrix

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \mathbf{X}_3 \end{bmatrix}.$$

If we let

$$\mathbf{x}_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ x_{i3} \end{bmatrix},\ i = 1, 2, \ldots, n\,.$$

That is, $\mathbf{x}_i$ is the (column) vector containing the *i*th observations of all three variables. Then we can write

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^{\mathrm{T}} \\ \mathbf{x}_2^{\mathrm{T}} \\ \mathbf{x}_3^{\mathrm{T}} \\ \vdots \\ \mathbf{x}_n^{\mathrm{T}} \end{bmatrix}$$

which emphasizes the observation number.

In statistics, the matrix $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ plays a very important role. Using our partition by variable, we have

$$\mathbf{X}^{\mathrm{T}}\mathbf{X} = \begin{bmatrix} \mathbf{X}_1^{\mathrm{T}} \\ \mathbf{X}_1^{\mathrm{T}} \\ \mathbf{X}_1^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \mathbf{X}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1^{\mathrm{T}}\mathbf{X}_1 & \mathbf{X}_1^{\mathrm{T}}\mathbf{X}_2 & \mathbf{X}_1^{\mathrm{T}}\mathbf{X}_3 \\ \mathbf{X}_2^{\mathrm{T}}\mathbf{X}_1 & \mathbf{X}_2^{\mathrm{T}}\mathbf{X}_2 & \mathbf{X}_2^{\mathrm{T}}\mathbf{X}_3 \\ \mathbf{X}_3^{\mathrm{T}}\mathbf{X}_1 & \mathbf{X}_3^{\mathrm{T}}\mathbf{X}_2 & \mathbf{X}_3^{\mathrm{T}}\mathbf{X}_3 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_{i1}^2 & \sum_{i=1}^n x_{i1}x_{i2} & \sum_{i=1}^n x_{i1}x_{i3} \\ \sum_{i=1}^n x_{i1}x_{i2} & \sum_{i=1}^n x_{i2}^2 & \sum_{i=1}^n x_{i2}x_{i3} \\ \sum_{i=1}^n x_{i1}x_{i3} & \sum_{i=1}^n x_{i2}x_{i3} & \sum_{i=1}^n x_{i3}^2 \end{bmatrix} \tag{11.3}$$

Using our partition by observations, we can write

$$\mathbf{X}^{\mathrm{T}}\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & \cdots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} \mathbf{x}_1^{\mathrm{T}} \\ \mathbf{x}_2^{\mathrm{T}} \\ \mathbf{x}_3^{\mathrm{T}} \\ \vdots \\ \mathbf{x}_n^{\mathrm{T}} \end{bmatrix} = \sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i^{\mathrm{T}} = \begin{bmatrix} \sum_{i=1}^n x_{i1}^2 & \sum_{i=1}^n x_{i1}x_{i2} & \sum_{i=1}^n x_{i1}x_{i3} \\ \sum_{i=1}^n x_{i1}x_{i2} & \sum_{i=1}^n x_{i2}^2 & \sum_{i=1}^n x_{i2}x_{i3} \\ \sum_{i=1}^n x_{i1}x_{i3} & \sum_{i=1}^n x_{i2}x_{i3} & \sum_{i=1}^n x_{i3}^2 \end{bmatrix}. \tag{11.4}$$

The expression $\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i^{\mathrm{T}}$ is a compact way of writing $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ that emphasizes the fact that each element is a sum of the squared observations, or the product of observations across variables.

Inverse of Partitioned Matrix

Suppose $\mathbf{A}$ is invertible, and suppose it is partitioned in the following manner:

$$\mathbf{A} = \begin{bmatrix} \underbrace{\mathbf{A}_{11}}_{n_1\times n_1} & \underbrace{\mathbf{A}_{12}}_{n_1\times n_2} \\ \underbrace{\mathbf{A}_{21}}_{n_2\times n_1} & \underbrace{\mathbf{A}_{22}}_{n_2\times n_2} \end{bmatrix}$$

We emphasize that the partition is such that the diagonal (top left to bottom right) blocks are square. The off-diagonal blocks need not be square. Then

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}_{11}^{-1} - \mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1} \\ -(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1} \end{bmatrix}. \tag{11.5}$$

We can apply Gauss-Jordan elimination to derive (11.5):

$$\left[\begin{array}{cc|cc}\mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{I} & \mathbf{0}\\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{0} & \mathbf{I}\end{array}\right]$$

$$\xrightarrow[{[2]=[2]-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}[1]}]{}$$

$$\left[\begin{array}{cc|cc}\mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} & -\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I}\end{array}\right]$$

$$\xrightarrow[{[1]=[1]-\mathbf{A}_{12}(\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}[2]}]{}$$

$$\left[\begin{array}{cc|cc}\mathbf{A}_{11} & \mathbf{0} & \mathbf{I}+\mathbf{A}_{12}(\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{12}(\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\\ \mathbf{0} & \mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} & -\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{I}\end{array}\right]$$

$$\xrightarrow[{\substack{[1]=\mathbf{A}_{11}^{-1}[1]\\ [2]=(\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}[2]}}]{}$$

$$\left[\begin{array}{cc|cc}\mathbf{I} & \mathbf{0} & \mathbf{A}_{11}^{-1}+\mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\\ \mathbf{0} & \mathbf{I} & -(\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & (\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\end{array}\right]$$

An alternative, equivalent, formula is

$$\mathbf{A}^{-1}=\begin{bmatrix}(\mathbf{A}_{11}-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} & -(\mathbf{A}_{11}-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}(\mathbf{A}_{11}-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} & \mathbf{A}_{22}^{-1}-\mathbf{A}_{22}^{-1}\mathbf{A}_{21}(\mathbf{A}_{11}-\mathbf{A}_{11}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\end{bmatrix} \tag{11.6}$$

which can be derived applying Gauss-Jordan elimination "from the bottom up", i.e., first step is $[1]=[1]-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}$, and so on.

Determinant of Partitioned Matrix It can be shown that the determinant of a block triangular matrix of the form

$$\mathbf{A}=\begin{bmatrix}\underbrace{\mathbf{A}_{11}}_{n_1\times n_1} & \underbrace{\mathbf{A}_{12}}_{n_1\times n_2}\\ \underbrace{\mathbf{0}}_{n_2\times n_1} & \underbrace{\mathbf{A}_{22}}_{n_2\times n_2}\end{bmatrix}$$

is $|\mathbf{A}|=|\mathbf{A}_{11}||\mathbf{A}_{22}|$. This fact, together with the fact that the elementary (block) row operation of adding a multiple of a (block) row to another (block) row does not change the determinant, can be used to derive the determinant of the partitioned matrix

$$\mathbf{A}=\begin{bmatrix}\underbrace{\mathbf{A}_{11}}_{n_1\times n_1} & \underbrace{\mathbf{A}_{12}}_{n_1\times n_2}\\ \underbrace{\mathbf{A}_{21}}_{n_2\times n_1} & \underbrace{\mathbf{A}_{22}}_{n_2\times n_2}\end{bmatrix}.$$

We have $\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix}\xrightarrow[{[2]=[2]-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}[1]}]{}\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{0} & \mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\end{bmatrix}$

which gives

$$|\mathbf{A}|=|\mathbf{A}_{11}||\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}|. \tag{11.7}$$

If $\mathbf{A}_{11}$ is singular, we can use

$$|\mathbf{A}|=|\mathbf{A}_{22}||\mathbf{A}_{11}-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}|. \tag{11.8}$$

Exercise

1. Derive (11.6).

2. Given $\mathbf{A} = \begin{bmatrix} \underbrace{\mathbf{A}_{11}}_{n_1 \times n_1} & \underbrace{\mathbf{A}_{12}}_{n_1 \times n_2} \\ \underbrace{\mathbf{A}_{21}}_{n_2 \times n_1} & \underbrace{\mathbf{A}_{22}}_{n_2 \times n_2} \end{bmatrix}$, show that

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}_{11}^{-1} - \mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1} \\ -(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} & -(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1} & \mathbf{A}_{22}^{-1} - \mathbf{A}_{22}^{-1}\mathbf{A}_{21}(\mathbf{A}_{11} - \mathbf{A}_{11}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \end{bmatrix}$$

by taking the products $\mathbf{AA}^{-1}$ and showing that the result is the identity matrix.

3. Derive (11.8).