Review Material

Chapters 1-3 of Matrix Analyis Textbook

Textbook

Example 1.2.1

Problem: Solve the following system using Gaussian elimination with back substitution:

$$\begin{aligned} v - w &= 3, \\ -2u + 4v - w &= 1, \\ -2u + 5v - 4w &= -2. \end{aligned}$$

Solution: The associated augmented matrix is

$$\left(\begin{array}{ccc|c} 0 & 1 & -1 & 3 \\ -2 & 4 & -1 & 1 \\ -2 & 5 & -4 & -2 \end{array}\right).$$

Since the first pivotal position contains 0, interchange rows one and two before eliminating below the first pivot:

$$\left(\begin{array}{ccc|c} 0 & 1 & -1 & 3 \\ -2 & 4 & -1 & 1 \\ -2 & 5 & -4 & -2 \end{array}\right) \xrightarrow{\textit{Interchange } R_1 \textit{ and } R_2} \left(\begin{array}{ccc|c} -2 & 4 & -1 & 1 \\ 0 & 1 & -1 & 3 \\ -2 & 5 & -4 & -2 \end{array}\right) R_3 - R_1$$

$$\longrightarrow \left(\begin{array}{ccc|c} -2 & 4 & -1 & 1 \\ 0 & 1 & -1 & 3 \\ 0 & 1 & -3 & -3 \end{array}\right) R_3 - R_2 \longrightarrow \left(\begin{array}{ccc|c} -2 & 4 & -1 & 1 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & -2 & -6 \end{array}\right).$$

Back substitution yields

$$\begin{aligned} w &= \frac{-6}{-2} = 3, \\ v &= 3 + w = 3 + 3 = 6, \\ u &= \frac{1}{-2}(1 - 4v + w) = \frac{1}{-2}(1 - 24 + 3) = 10. \end{aligned}$$

Textbook

Example 1.3.1

Problem: Apply the Gauss–Jordan method to solve the following system:

$$\begin{aligned} 2x_1 &+ 2x_2 &+ 6x_3 &= 4, \\ 2x_1 &+ x_2 &+ 7x_3 &= 6, \\ -2x_1 &- 6x_2 &- 7x_3 &= -1. \end{aligned}$$

Solution: The sequence of operations is indicated in parentheses and the pivots are circled.

$$\left(\begin{array}{rrr|r} 2 & 2 & 6 & 4 \\ 2 & 1 & 7 & 6 \\ -2 & -6 & -7 & -1 \end{array}\right) R_1/2 \longrightarrow \left(\begin{array}{rrr|r} 1 & 1 & 3 & 2 \\ 2 & 1 & 7 & 6 \\ -2 & -6 & -7 & -1 \end{array}\right) \begin{array}{l} \\ R_2 - 2R_1 \\ R_3 + 2R_1 \end{array}$$

$$\longrightarrow \left(\begin{array}{rrr|r} 1 & 1 & 3 & 2 \\ 0 & -1 & 1 & 2 \\ 0 & -4 & -1 & 3 \end{array}\right) (-R_2) \longrightarrow \left(\begin{array}{rrr|r} 1 & 1 & 3 & 2 \\ 0 & 1 & -1 & -2 \\ 0 & -4 & -1 & 3 \end{array}\right) \begin{array}{l} R_1 - R_2 \\ \\ R_3 + 4R_2 \end{array}$$

$$\longrightarrow \left(\begin{array}{rrr|r} 1 & 0 & 4 & 4 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & -5 & -5 \end{array}\right) -R_3/5 \longrightarrow \left(\begin{array}{rrr|r} 1 & 0 & 4 & 4 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & 1 \end{array}\right) \begin{array}{l} R_1 - 4R_3 \\ R_2 + R_3 \\ \\ \end{array}$$

$$\longrightarrow \left(\begin{array}{rrr|r} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

Therefore, the solution is $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$.

Textbook

Example 2.1.1

Problem: Apply modified Gaussian elimination to the following matrix and circle the pivot positions:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 & 3 & 3 \\ 2 & 4 & 0 & 4 & 4 \\ 1 & 2 & 3 & 5 & 5 \\ 2 & 4 & 0 & 4 & 7 \end{pmatrix}.$$

Solution:

$$\begin{pmatrix} 1 & 2 & 1 & 3 & 3 \\ 2 & 4 & 0 & 4 & 4 \\ 1 & 2 & 3 & 5 & 5 \\ 2 & 4 & 0 & 4 & 7 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 3 \\ 0 & 0 & -2 & -2 & -2 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & -2 & -2 & 1 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 3 \\ 0 & 0 & -2 & -2 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 1 & 3 & 3 \\ 0 & 0 & -2 & -2 & -2 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Row Echelon Form

An $m \times n$ matrix $\mathbf{E}$ with rows $\mathbf{E}_{i*}$ and columns $\mathbf{E}_{*j}$ is said to be in *row echelon form* provided the following two conditions hold.

- If $\mathbf{E}_{i*}$ consists entirely of zeros, then all rows below $\mathbf{E}_{i*}$ are also entirely zero; i.e., all zero rows are at the bottom.

- If the first nonzero entry in $\mathbf{E}_{i*}$ lies in the j^{th} position, then all entries below the i^{th} position in columns $\mathbf{E}_{*1}, \mathbf{E}_{*2}, \ldots, \mathbf{E}_{*j}$ are zero.

These two conditions say that the nonzero entries in an echelon form must lie on or above a stair-step line that emanates from the upper-left-hand corner and slopes down and to the right. The pivots are the first nonzero entries in each row. A typical structure for a matrix in row echelon form is illustrated below with the pivots circled.

$$\begin{pmatrix} * & * & * & * & * & * & * & * \\ 0 & 0 & * & * & * & * & * & * \\ 0 & 0 & 0 & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Rank of a Matrix

Suppose $\mathbf{A}_{m\times n}$ is reduced by row operations to an echelon form $\mathbf{E}$. The ***rank*** of $\mathbf{A}$ is defined to be the number

$$\begin{aligned} rank\,(\mathbf{A}) &= \text{ number of pivots} \\ &= \text{ number of nonzero rows in } \mathbf{E} \\ &= \text{ number of basic columns in } \mathbf{A}, \end{aligned}$$

where the ***basic columns*** of $\mathbf{A}$ are defined to be those columns in $\mathbf{A}$ that contain the pivotal positions.

Problem: Determine the rank, and identify the basic columns in

Textbook

Example 2.1.2

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 2 & 4 & 2 & 2 \\ 3 & 6 & 3 & 4 \end{pmatrix}.$$

Solution: Reduce $\mathbf{A}$ to row echelon form as shown below:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 2 & 4 & 2 & 2 \\ 3 & 6 & 3 & 4 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \mathbf{E}.$$

Consequently, $rank(\mathbf{A}) = 2$. The pivotal positions lie in the first and fourth columns so that the basic columns of $\mathbf{A}$ are $\mathbf{A}_{*1}$ and $\mathbf{A}_{*4}$. That is,

$$\text{Basic Columns} = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} \right\}.$$

Pay particular attention to the fact that the basic columns are extracted from $\mathbf{A}$ and not from the row echelon form $\mathbf{E}$.

Reduced Row Echelon Form

A matrix $\mathbf{E}_{m\times n}$ is said to be in *reduced row echelon form* provided that the following three conditions hold.

- $\mathbf{E}$ is in row echelon form.
- The first nonzero entry in each row (i.e., each pivot) is 1.
- All entries above each pivot are 0.

A typical structure for a matrix in reduced row echelon form is illustrated below, where entries marked * can be either zero or nonzero numbers:

$$\begin{pmatrix} 1 & * & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 1 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 1 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$\mathbf{E_A}$ Notation

For a matrix $\mathbf{A}$, the symbol $\mathbf{E_A}$ will hereafter denote the unique reduced row echelon form derived from $\mathbf{A}$ by means of row operations.

Problem: Determine $\mathbf{E_A}$, deduce $rank\,(\mathbf{A})$, and identify the basic columns of

Textbook

Example 2.2.2

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 2 & 3 & 1 \\ 2 & 4 & 4 & 6 & 2 \\ 3 & 6 & 6 & 9 & 6 \\ 1 & 2 & 4 & 5 & 3 \end{pmatrix}.$$

Solution:

$$\begin{pmatrix} 1 & 2 & 2 & 3 & 1 \\ 2 & 4 & 4 & 6 & 2 \\ 3 & 6 & 6 & 9 & 6 \\ 1 & 2 & 4 & 5 & 3 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 2 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 2 & 2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 2 & 3 & 1 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 2 & 2 & 3 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 2 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Therefore, $rank\,(\mathbf{A}) = 3$, and $\{\mathbf{A}_{*1}, \mathbf{A}_{*3}, \mathbf{A}_{*5}\}$ are the three basic columns.

If such relationships exist, then the columns are said to be "linearly dependent." Otherwise the columns are "linearly independent."

Column Relationships in A and $\mathbf{E_A}$

- Each nonbasic column $\mathbf{E}_{*k}$ in $\mathbf{E_A}$ is a combination (a sum of multiples) of the basic columns in $\mathbf{E_A}$ to the left of $\mathbf{E}_{*k}$. That is,

$$\mathbf{E}_{*k} = \mu_1 \mathbf{E}_{*b_1} + \mu_2 \mathbf{E}_{*b_2} + \cdots + \mu_j \mathbf{E}_{*b_j}$$

$$= \mu_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \mu_j \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_j \\ \vdots \\ 0 \end{pmatrix},$$

  where the $\mathbf{E}_{*b_i}$'s are the basic columns to the left of $\mathbf{E}_{*k}$ and where the multipliers μ_i are the first j entries in $\mathbf{E}_{*k}$.

- The relationships that exist among the columns of $\mathbf{A}$ are exactly the same as the relationships that exist among the columns of $\mathbf{E_A}$. In particular, if $\mathbf{A}_{*k}$ is a nonbasic column in $\mathbf{A}$, then

$$\mathbf{A}_{*k} = \mu_1 \mathbf{A}_{*b_1} + \mu_2 \mathbf{A}_{*b_2} + \cdots + \mu_j \mathbf{A}_{*b_j}, \qquad (2.2.3)$$

  where the $\mathbf{A}_{*b_i}$'s are the basic columns to the left of $\mathbf{A}_{*k}$, and where the multipliers μ_i are as described above—the first j entries in $\mathbf{E}_{*k}$.

Textbook

Example 2.2.3

Problem: Write each nonbasic column as a combination of basic columns in

$$\mathbf{A} = \begin{pmatrix} 2 & -4 & -8 & 6 & 3 \\ 0 & 1 & 3 & 2 & 3 \\ 3 & -2 & 0 & 0 & 8 \end{pmatrix}.$$

Solution: Transform $\mathbf{A}$ to $\mathbf{E_A}$ as shown below.

$$\begin{pmatrix} 2 & -4 & -8 & 6 & 3 \\ 0 & 1 & 3 & 2 & 3 \\ 3 & -2 & 0 & 0 & 8 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -2 & -4 & 3 & \frac{3}{2} \\ 0 & 1 & 3 & 2 & 3 \\ 3 & -2 & 0 & 0 & 8 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -2 & -4 & 3 & \frac{3}{2} \\ 0 & 1 & 3 & 2 & 3 \\ 0 & 4 & 12 & -9 & \frac{7}{2} \end{pmatrix} \rightarrow$$

$$\begin{pmatrix} 1 & 0 & 2 & 7 & \frac{15}{2} \\ 0 & 1 & 3 & 2 & 3 \\ 0 & 0 & 0 & -17 & -\frac{17}{2} \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 2 & 7 & \frac{15}{2} \\ 0 & 1 & 3 & 2 & 3 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 2 & 0 & 4 \\ 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix}$$

The third and fifth columns are nonbasic. Looking at the columns in $\mathbf{E_A}$ reveals

$$\mathbf{E}_{*3} = 2\mathbf{E}_{*1} + 3\mathbf{E}_{*2} \quad \text{and} \quad \mathbf{E}_{*5} = 4\mathbf{E}_{*1} + 2\mathbf{E}_{*2} + \frac{1}{2}\mathbf{E}_{*4}.$$

The relationships that exist among the columns of $\mathbf{A}$ must be exactly the same as those in $\mathbf{E_A}$, so

$$\mathbf{A}_{*3} = 2\mathbf{A}_{*1} + 3\mathbf{A}_{*2} \quad \text{and} \quad \mathbf{A}_{*5} = 4\mathbf{A}_{*1} + 2\mathbf{A}_{*2} + \frac{1}{2}\mathbf{A}_{*4}.$$

You can easily check the validity of these equations by direct calculation.

Rather than depending on geometry to establish consistency, we use Gaussian elimination. If the associated augmented matrix $[\mathbf{A}|\mathbf{b}]$ is reduced by row operations to a matrix $[\mathbf{E}|\mathbf{c}]$ that is in row echelon form, then consistency—or lack of it—becomes evident. Suppose that somewhere in the process of reducing $[\mathbf{A}|\mathbf{b}]$ to $[\mathbf{E}|\mathbf{c}]$ a situation arises in which the only nonzero entry in a row appears on the right-hand side, as illustrated below:

$$\text{Row } i \longrightarrow \left(\begin{array}{cccccc|c} * & * & * & * & * & * & * \\ 0 & 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \alpha \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \end{array}\right) \longleftarrow \alpha \neq 0.$$

If this occurs in the i^{th} row, then the i^{th} equation of the associated system is

$$0x_1 + 0x_2 + \cdots + 0x_n = \alpha.$$

For $\alpha \neq 0$, this equation has no solution, and hence the original system must also be inconsistent (because row operations don't alter the solution set). The converse also holds. That is, if a system is inconsistent, then somewhere in the elimination process a row of the form

$$(\,0 \quad 0 \quad \cdots \quad 0 \quad | \quad \alpha\,), \quad \alpha \neq 0 \tag{2.3.1}$$

must appear. Otherwise, the back substitution process can be completed and a solution is produced. There is *no* inconsistency indicated when a row of the form $(0\;0\;\cdots\;0\,|\,0)$ is encountered. This simply says that $0 = 0$, and although this is no help in determining the value of any unknown, it is nevertheless a true statement, so it doesn't indicate inconsistency in the system.

Consistency

Each of the following is equivalent to saying that $[\mathbf{A}|\mathbf{b}]$ is consistent.

- In row reducing $[\mathbf{A}|\mathbf{b}]$, a row of the following form never appears:

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & | & \alpha \end{pmatrix}, \quad \text{where} \quad \alpha \neq 0. \tag{2.3.2}$$

- $\mathbf{b}$ is a nonbasic column in $[\mathbf{A}|\mathbf{b}]$. (2.3.3)
- $rank[\mathbf{A}|\mathbf{b}] = rank\,(\mathbf{A})$. (2.3.4)
- $\mathbf{b}$ is a combination of the basic columns in $\mathbf{A}$. (2.3.5)

Textbook

Example 2.3.1

Problem: Determine if the following system is consistent:

$$
\begin{aligned}
x_1 + x_2 + 2x_3 + 2x_4 + x_5 &= 1, \\
2x_1 + 2x_2 + 4x_3 + 4x_4 + 3x_5 &= 1, \\
2x_1 + 2x_2 + 4x_3 + 4x_4 + 2x_5 &= 2, \\
3x_1 + 5x_2 + 8x_3 + 6x_4 + 5x_5 &= 3.
\end{aligned}
$$

Solution: Apply Gaussian elimination to the augmented matrix $[\mathbf{A}|\mathbf{b}]$ as shown:

$$
\left(\begin{array}{ccccc|c}
1 & 1 & 2 & 2 & 1 & 1 \\
2 & 2 & 4 & 4 & 3 & 1 \\
2 & 2 & 4 & 4 & 2 & 2 \\
3 & 5 & 8 & 6 & 5 & 3
\end{array}\right)
\longrightarrow
\left(\begin{array}{ccccc|c}
1 & 1 & 2 & 2 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 2 & 2 & 0 & 2 & 0
\end{array}\right)
$$

$$
\longrightarrow
\left(\begin{array}{ccccc|c}
1 & 1 & 2 & 2 & 1 & 1 \\
0 & 2 & 2 & 0 & 2 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 \\
0 & 0 & 0 & 0 & 0 & 0
\end{array}\right).
$$

Because a row of the form $(0 \;\; 0 \;\; \cdots \;\; 0 \;\; | \;\; \alpha)$ with $\alpha \neq 0$ never emerges, the system is consistent. We might also observe that $\mathbf{b}$ is a nonbasic column in $[\mathbf{A}|\mathbf{b}]$ so that $rank[\mathbf{A}|\mathbf{b}] = rank(\mathbf{A})$. Finally, by completely reducing $\mathbf{A}$ to $\mathbf{E_A}$, it is possible to verify that $\mathbf{b}$ is indeed a combination of the basic columns $\{\mathbf{A}_{*1}, \mathbf{A}_{*2}, \mathbf{A}_{*5}\}$.

Addition of Matrices

If $\mathbf{A}$ and $\mathbf{B}$ are $m \times n$ matrices, the *sum* of $\mathbf{A}$ and $\mathbf{B}$ is defined to be the $m \times n$ matrix $\mathbf{A}+\mathbf{B}$ obtained by adding corresponding entries. That is,

$$[\mathbf{A}+\mathbf{B}]_{ij} = [\mathbf{A}]_{ij} + [\mathbf{B}]_{ij} \quad \text{for each } i \text{ and } j.$$

For example,

$$\begin{pmatrix} -2 & x & 3 \\ z+3 & 4 & -y \end{pmatrix} + \begin{pmatrix} 2 & 1-x & -2 \\ -3 & 4+x & 4+y \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ z & 8+x & 4 \end{pmatrix}.$$

Properties of Matrix Addition

For $m \times n$ matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, the following properties hold.

Closure property:	$\mathbf{A} + \mathbf{B}$ is again an $m \times n$ matrix.
Associative property:	$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$.
Commutative property:	$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$.
Additive identity:	The $m \times n$ matrix $\mathbf{0}$ consisting of all zeros has the property that $\mathbf{A} + \mathbf{0} = \mathbf{A}$.
Additive inverse:	The $m \times n$ matrix $(-\mathbf{A})$ has the property that $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$.

Scalar Multiplication

The product of a scalar α times a matrix $\mathbf{A}$, denoted by $\alpha\mathbf{A}$, is defined to be the matrix obtained by multiplying each entry of $\mathbf{A}$ by α. That is, $[\alpha\mathbf{A}]_{ij} = \alpha[\mathbf{A}]_{ij}$ for each i and j.

For example,

$$2\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 6 \\ 0 & 2 & 4 \\ 2 & 8 & 4 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 & 4 \\ 6 & 8 \\ 0 & 2 \end{pmatrix}.$$

Properties of Scalar Multiplication

For $m \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ and for scalars α and β, the following properties hold.

Closure property:	$\alpha\mathbf{A}$ is again an $m \times n$ matrix.
Associative property:	$(\alpha\beta)\mathbf{A} = \alpha(\beta\mathbf{A})$.
Distributive property:	$\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$. Scalar multiplication is distributed over matrix addition.
Distributive property:	$(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$. Scalar multiplication is distributed over scalar addition.
Identity property:	$1\mathbf{A} = \mathbf{A}$. The number 1 is an identity element under scalar multiplication.

Transpose

The *transpose* of $\mathbf{A}_{m\times n}$ is defined to be the $n \times m$ matrix $\mathbf{A}^T$ obtained by interchanging rows and columns in $\mathbf{A}$. More precisely, if $\mathbf{A} = [a_{ij}]$, then $[\mathbf{A}^T]_{ij} = a_{ji}$. For example,

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}.$$

It should be evident that for all matrices, $\left(\mathbf{A}^T\right)^T = \mathbf{A}$.

Hermitian Transpose

For $\mathbf{A} = [a_{ij}]$, the ***conjugate matrix*** is defined to be $\overline{\mathbf{A}} = [\overline{a}_{ij}]$, and the ***conjugate transpose*** of $\mathbf{A}$ is defined to be $\bar{\mathbf{A}}^T = \overline{\mathbf{A}^T}$. From now on, $\bar{\mathbf{A}}^T$ will be denoted by A^{H}, so $[\mathrm{A}^{\mathrm{H}}]_{ij} = \overline{a}_{ji}$. For example,

$$\begin{pmatrix} 1-4\mathrm{i} & \mathrm{i} & 2 \\ 3 & 2+\mathrm{i} & 0 \end{pmatrix}^{\mathrm{H}} = \begin{pmatrix} 1+4\mathrm{i} & 3 \\ -\mathrm{i} & 2-\mathrm{i} \\ 2 & 0 \end{pmatrix}.$$

$(\mathrm{A}^{\mathrm{H}})^{\mathrm{H}} = \mathbf{A}$ for all matrices, and $\mathrm{A}^{\mathrm{H}} = \mathbf{A}^T$ whenever $\mathbf{A}$ contains only real entries. Sometimes the matrix A^{H} is called the ***Hermitian*** of $\mathbf{A}$.

Properties of the Transpose/Hermitian Transpose

If $\mathbf{A}$ and $\mathbf{B}$ are two matrices of the same shape, and if α is a scalar, then each of the following statements is true.

$$(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T \quad \text{and} \quad (\mathbf{A}+\mathbf{B})^{\mathrm{H}} = \mathbf{A}^{\mathrm{H}} + \mathbf{B}^{\mathrm{H}}. \tag{3.2.1}$$

$$(\alpha\mathbf{A})^T = \alpha\mathbf{A}^T \quad \text{and} \quad (\alpha\mathbf{A})^{\mathrm{H}} = \overline{\alpha}\mathbf{A}^{\mathrm{H}}. \tag{3.2.2}$$

Symmetries

Let $\mathbf{A} = [a_{ij}]$ be a square matrix.

- $\mathbf{A}$ is said to be a *symmetric matrix* whenever $\mathbf{A} = \mathbf{A}^T$, i.e., whenever $a_{ij} = a_{ji}$.

- $\mathbf{A}$ is said to be a *skew-symmetric matrix* whenever $\mathbf{A} = -\mathbf{A}^T$, i.e., whenever $a_{ij} = -a_{ji}$.

- $\mathbf{A}$ is said to be a *hermitian matrix* whenever $\mathbf{A} = \mathbf{A}^{\mathrm{H}}$, i.e., whenever $a_{ij} = \overline{a}_{ji}$. This is the complex analog of symmetry.

- $\mathbf{A}$ is said to be a *skew-hermitian matrix* when $\mathbf{A} = -\mathbf{A}^{\mathrm{H}}$, i.e., whenever $a_{ij} = -\overline{a}_{ji}$. This is the complex analog of skew symmetry.

For example, consider

$$\mathbf{A} = \begin{pmatrix} 1 & 2+4\mathrm{i} & 1-3\mathrm{i} \\ 2-4\mathrm{i} & 3 & 8+6\mathrm{i} \\ 1+3\mathrm{i} & 8-6\mathrm{i} & 5 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 1 & 2+4\mathrm{i} & 1-3\mathrm{i} \\ 2+4\mathrm{i} & 3 & 8+6\mathrm{i} \\ 1-3\mathrm{i} & 8+6\mathrm{i} & 5 \end{pmatrix}.$$

Linear Functions

Suppose that $\mathcal{D}$ and $\mathcal{R}$ are sets that possess an addition operation as well as a scalar multiplication operation—i.e., a multiplication between scalars and set members. A function f that maps points in $\mathcal{D}$ to points in $\mathcal{R}$ is said to be a ***linear function*** whenever f satisfies the conditions that

$$f(x+y) = f(x) + f(y) \tag{3.3.1}$$

and

$$f(\alpha x) = \alpha f(x) \tag{3.3.2}$$

for every x and y in $\mathcal{D}$ and for all scalars α. These two conditions may be combined by saying that f is a linear function whenever

$$f(\alpha x + y) = \alpha f(x) + f(y) \tag{3.3.3}$$

for all scalars α and for all $x, y \in \mathcal{D}$.

One of the simplest linear functions is $f(x) = \alpha x$, whose graph in $\Re^2$ is a straight line through the origin. You should convince yourself that f is indeed a linear function according to the above definition. However, $f(x) = \alpha x + \beta$ does not qualify for the title "linear function"—it is a linear function that has been translated by a constant β. Translations of linear functions are referred to as ***affine functions.*** Virtually all information concerning affine functions can be derived from an understanding of linear functions, and consequently we will focus only on issues of linearity.

Linearity is encountered at every turn. For example, the familiar operations of differentiation and integration may be viewed as linear functions. Since

$$\frac{d(f+g)}{dx} = \frac{df}{dx} + \frac{dg}{dx} \quad \text{and} \quad \frac{d(\alpha f)}{dx} = \alpha\frac{df}{dx},$$

the differentiation operator $D_x(f) = df/dx$ is linear. Similarly,

$$\int (f+g)dx = \int f dx + \int g dx \quad \text{and} \quad \int \alpha f dx = \alpha \int f dx$$

means that the integration operator $I(f) = \int f dx$ is linear.

There are several important matrix functions that are linear. For example, the transposition function $f(\mathbf{X}_{m\times n}) = \mathbf{X}^T$ is linear because

$$(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T \quad \text{and} \quad (\alpha\mathbf{A})^T = \alpha\mathbf{A}^T$$

(recall (3.2.1) and (3.2.2)). Another matrix function that is linear is the *trace* function presented below.

Textbook

Example 3.3.1

The ***trace*** of an $n \times n$ matrix $\mathbf{A} = [a_{ij}]$ is defined to be the sum of the entries lying on the main diagonal of $\mathbf{A}$. That is,

$$trace\,(\mathbf{A}) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}.$$

Problem: Show that $f(\mathbf{X}_{n\times n}) = trace\,(\mathbf{X})$ is a linear function.

Solution: Let's be efficient by showing that (3.3.3) holds. Let $\mathbf{A} = [a_{ij}]$ and $\mathbf{B} = [b_{ij}]$, and write

$$\begin{aligned} f(\alpha\mathbf{A} + \mathbf{B}) &= trace\,(\alpha\mathbf{A} + \mathbf{B}) = \sum_{i=1}^{n} [\alpha\mathbf{A} + \mathbf{B}]_{ii} = \sum_{i=1}^{n} (\alpha a_{ii} + b_{ii}) \\ &= \sum_{i=1}^{n} \alpha a_{ii} + \sum_{i=1}^{n} b_{ii} = \alpha \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = \alpha\, trace\,(\mathbf{A}) + trace\,(\mathbf{B}) \\ &= \alpha f(\mathbf{A}) + f(\mathbf{B}). \end{aligned}$$

$$\mathbf{R} = \begin{pmatrix} r_1 & r_2 & \cdots & r_n \end{pmatrix} \quad \text{and} \quad \mathbf{C} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix},$$

the ***standard inner product*** of $\mathbf{R}$ with $\mathbf{C}$ is defined to be the scalar

$$\mathbf{RC} = r_1c_1 + r_2c_2 + \cdots + r_nc_n = \sum_{i=1}^{n} r_ic_i.$$

For example,

$$\begin{pmatrix} 2 & 4 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = (2)(1) + (4)(2) + (-2)(3) = 4.$$

Matrix Multiplication

- Matrices $\mathbf{A}$ and $\mathbf{B}$ are said to be *conformable* for multiplication in the order $\mathbf{AB}$ whenever $\mathbf{A}$ has exactly as many columns as $\mathbf{B}$ has rows—i.e., $\mathbf{A}$ is $m \times p$ and $\mathbf{B}$ is $p \times n$.

- For conformable matrices $\mathbf{A}_{m\times p} = [a_{ij}]$ and $\mathbf{B}_{p\times n} = [b_{ij}]$, the *matrix product* $\mathbf{AB}$ is defined to be the $m \times n$ matrix whose (i,j)-entry is the inner product of the i^{th} row of $\mathbf{A}$ with the j^{th} column in $\mathbf{B}$. That is,

$$[\mathbf{AB}]_{ij} = \mathbf{A}_{i*}\mathbf{B}_{*j} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}.$$

- In case $\mathbf{A}$ and $\mathbf{B}$ fail to be conformable—i.e., $\mathbf{A}$ is $m \times p$ and $\mathbf{B}$ is $q \times n$ with $p \neq q$—then no product $\mathbf{AB}$ is defined.

For example, if

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}_{2\times 3} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{pmatrix}_{3\times 4}$$

inside ones match

shape of the product

then the product $\mathbf{AB}$ exists and has shape 2×4. Consider a typical entry of this product, say, the (2,3)-entry. The definition says $[\mathbf{AB}]_{23}$ is obtained by forming the inner product of the second row of $\mathbf{A}$ with the third column of $\mathbf{B}$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ \boxed{a_{21} \quad a_{22} \quad a_{23}} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \boxed{\begin{matrix} b_{13} \\ b_{23} \\ b_{33} \end{matrix}} & b_{14} \\ b_{21} & b_{22} & & b_{24} \\ b_{31} & b_{32} & & b_{34} \end{pmatrix},$$

so

$$[\mathbf{AB}]_{23} = \mathbf{A}_{2*}\mathbf{B}_{*3} = a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} = \sum_{k=1}^{3} a_{2k}b_{k3}.$$

For example,

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & -4 \\ -3 & 0 & 5 \end{pmatrix}, \ \mathbf{B} = \begin{pmatrix} 1 & 3 & -3 & 2 \\ 2 & 5 & -1 & 8 \\ -1 & 2 & 0 & 2 \end{pmatrix} \Longrightarrow \mathbf{AB} = \begin{pmatrix} 8 & 3 & -7 & 4 \\ -8 & 1 & 9 & 4 \end{pmatrix}$$

Textbook

Example 3.5.1

The cancellation law (3.5.3) fails for matrix multiplication. If

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}, \quad \text{and} \quad \mathbf{C} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix},$$

then

$$\mathbf{AB} = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix} = \mathbf{AC} \quad \text{but} \quad \mathbf{B} \neq \mathbf{C}$$

in spite of the fact that $\mathbf{A} \neq \mathbf{0}$.

Rows and Columns of a Product

Suppose that $\mathbf{A} = [a_{ij}]$ is $m \times p$ and $\mathbf{B} = [b_{ij}]$ is $p \times n$.

- $[\mathbf{AB}]_{i*} = \mathbf{A}_{i*}\mathbf{B}$ $\big[($ i^{th} row of $\mathbf{AB}$ $)=($ i^{th} row of $\mathbf{A}$ $) \times \mathbf{B}\big]$. (3.5.4)

- $[\mathbf{AB}]_{*j} = \mathbf{AB}_{*j}$ $\big[($ j^{th} col of $\mathbf{AB}$ $)= \mathbf{A} \times ($ j^{th} col of $\mathbf{B}$ $)\big]$. (3.5.5)

- $[\mathbf{AB}]_{i*} = a_{i1}\mathbf{B}_{1*} + a_{i2}\mathbf{B}_{2*} + \cdots + a_{ip}\mathbf{B}_{p*} = \sum_{k=1}^{p} a_{ik}\mathbf{B}_{k*}$. (3.5.6)

- $[\mathbf{AB}]_{*j} = \mathbf{A}_{*1}b_{1j} + \mathbf{A}_{*2}b_{2j} + \cdots + \mathbf{A}_{*p}b_{pj} = \sum_{k=1}^{p} \mathbf{A}_{*k}b_{kj}$. (3.5.7)

These last two equations show that rows of $\mathbf{AB}$ are combinations of rows of $\mathbf{B}$, while columns of $\mathbf{AB}$ are combinations of columns of $\mathbf{A}$.

For example, if $\mathbf{A} = \begin{pmatrix} 1 & -2 & 0 \\ 3 & -4 & 5 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 3 & -5 & 1 \\ 2 & -7 & 2 \\ 1 & -2 & 0 \end{pmatrix}$, then the second row of $\mathbf{AB}$ is

$$[\mathbf{AB}]_{2*} = \mathbf{A}_{2*}\mathbf{B} = \begin{pmatrix} 3 & -4 & 5 \end{pmatrix} \begin{pmatrix} 3 & -5 & 1 \\ 2 & -7 & 2 \\ 1 & -2 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 3 & -5 \end{pmatrix},$$

and the second column of $\mathbf{AB}$ is

$$[\mathbf{AB}]_{*2} = \mathbf{AB}_{*2} = \begin{pmatrix} 1 & -2 & 0 \\ 3 & -4 & 5 \end{pmatrix} \begin{pmatrix} -5 \\ -7 \\ -2 \end{pmatrix} = \begin{pmatrix} 9 \\ 3 \end{pmatrix}.$$

This example makes the point that it is wasted effort to compute the entire product if only one row or column is called for. Although it's not necessary to compute the complete product, you may wish to verify that

$$\mathbf{AB} = \begin{pmatrix} 1 & -2 & 0 \\ 3 & -4 & 5 \end{pmatrix} \begin{pmatrix} 3 & -5 & 1 \\ 2 & -7 & 2 \\ 1 & -2 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 9 & -3 \\ 6 & 3 & -5 \end{pmatrix}.$$

Distributive and Associative Laws

For conformable matrices each of the following is true.

- $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$ (left-hand distributive law).
- $(\mathbf{D}+\mathbf{E})\mathbf{F} = \mathbf{DF}+\mathbf{EF}$ (right-hand distributive law).
- $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$ (associative law).

Textbook

Example 3.6.1

Linearity of Matrix Multiplication. Let $\mathbf{A}$ be an $m \times n$ matrix, and f be the function defined by matrix multiplication

$$f(\mathbf{X}_{n\times p}) = \mathbf{AX}.$$

The left-hand distributive property guarantees that f is a linear function because for all scalars α and for all $n \times p$ matrices $\mathbf{X}$ and $\mathbf{Y}$,

$$\begin{aligned} f(\alpha\mathbf{X}+\mathbf{Y}) &= \mathbf{A}(\alpha\mathbf{X}+\mathbf{Y}) = \mathbf{A}(\alpha\mathbf{X}) + \mathbf{AY} = \alpha\mathbf{AX} + \mathbf{AY} \\ &= \alpha f(\mathbf{X}) + f(\mathbf{Y}). \end{aligned}$$

Of course, the linearity of matrix multiplication is no surprise because it was the consideration of linear functions that motivated the definition of the matrix product at the outset.

Identity Matrix

The $n \times n$ matrix with 1's on the main diagonal and 0's elsewhere

$$\mathbf{I}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

is called the ***identity matrix*** of order n. For every $m \times n$ matrix $\mathbf{A}$,

$$\mathbf{A}\mathbf{I}_n = \mathbf{A} \quad \text{and} \quad \mathbf{I}_m\mathbf{A} = \mathbf{A}.$$

The subscript on $\mathbf{I}_n$ is neglected whenever the size is obvious from the context.

Analogous to scalar algebra, we define the 0^{th} power of a square matrix to be the identity matrix of corresponding size. That is, if $\mathbf{A}$ is $n \times n$, then

$$\mathbf{A}^0 = \mathbf{I}_n.$$

Positive powers of $\mathbf{A}$ are also defined in the natural way. That is,

$$\mathbf{A}^n = \underbrace{\mathbf{A}\mathbf{A}\cdots\mathbf{A}}_{n \text{ times}}.$$

The associative law guarantees that it makes no difference how matrices are grouped for powering. For example, $\mathbf{A}\mathbf{A}^2$ is the same as $\mathbf{A}^2\mathbf{A}$, so that

$$\mathbf{A}^3 = \mathbf{A}\mathbf{A}\mathbf{A} = \mathbf{A}\mathbf{A}^2 = \mathbf{A}^2\mathbf{A}.$$

Also, the usual laws of exponents hold. For nonnegative integers r and s,

$$\mathbf{A}^r\mathbf{A}^s = \mathbf{A}^{r+s} \quad \text{and} \quad (\mathbf{A}^r)^s = \mathbf{A}^{rs}.$$

We are not yet in a position to define negative or fractional powers, and due to the lack of conformability, powers of nonsquare matrices are never defined.

Textbook

Example 3.6.2

A Pitfall. For two $n \times n$ matrices, what is $(\mathbf{A}+\mathbf{B})^2$? **Be careful!** Because matrix multiplication is not commutative, the familiar formula from scalar algebra is not valid for matrices. The distributive properties must be used to write

$$\begin{aligned}(\mathbf{A}+\mathbf{B})^2 &= \underbrace{(\mathbf{A}+\mathbf{B})}(\mathbf{A}+\mathbf{B}) = \underbrace{(\mathbf{A}+\mathbf{B})}\mathbf{A} + \underbrace{(\mathbf{A}+\mathbf{B})}\mathbf{B} \\ &= \mathbf{A}^2 + \mathbf{BA} + \mathbf{AB} + \mathbf{B}^2,\end{aligned}$$

and this is as far as you can go. The familiar form $\mathbf{A}^2+2\mathbf{AB}+\mathbf{B}^2$ is obtained only in those rare cases where $\mathbf{AB} = \mathbf{BA}$. To evaluate $(\mathbf{A}+\mathbf{B})^k$, the distributive rules must be applied repeatedly, and the results are a bit more complicated—try it for $k = 3$.

Reverse Order Law for Transposition

For conformable matrices $\mathbf{A}$ and $\mathbf{B}$,

$$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T.$$

The case of conjugate transposition is similar. That is,

$$(\mathbf{AB})^{\mathrm{H}} = \mathbf{B}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}.$$

Textbook

Example 3.6.4

For every matrix $\mathbf{A}_{m\times n}$, the products $\mathbf{A}^T\mathbf{A}$ and $\mathbf{A}\mathbf{A}^T$ are symmetric matrices because

$$\left(\mathbf{A}^T\mathbf{A}\right)^T = \mathbf{A}^T\mathbf{A}^{T^T} = \mathbf{A}^T\mathbf{A} \quad \text{and} \quad \left(\mathbf{A}\mathbf{A}^T\right)^T = \mathbf{A}^{T^T}\mathbf{A}^T = \mathbf{A}\mathbf{A}^T.$$

Textbook

Example 3.6.5

Trace of a Product. Recall from Example 3.3.1 that the trace of a square matrix is the sum of its main diagonal entries. Although matrix multiplication is not commutative, the trace function is one of the few cases where the order of the matrices can be changed without affecting the results.

Problem: For matrices $\mathbf{A}_{m \times n}$ and $\mathbf{B}_{n \times m}$,

$$trace\,(\mathbf{AB}) = trace\,(\mathbf{BA}).$$

Note: This is true in spite of the fact that $\mathbf{AB}$ is $m \times m$ while $\mathbf{BA}$ is $n \times n$. Furthermore, this result can be extended to say that any product of conformable matrices can be permuted *cyclically* without altering the trace of the product. For example,

$$trace\,(\mathbf{ABC}) = trace\,(\mathbf{BCA}) = trace\,(\mathbf{CAB}).$$

However, a noncyclical permutation may not preserve the trace. For example,

$$trace\,(\mathbf{ABC}) \neq trace\,(\mathbf{BAC}).$$

Block Matrix Multiplication

Suppose that $\mathbf{A}$ and $\mathbf{B}$ are partitioned into submatrices—often referred to as *blocks*—as indicated below.

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1r} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{s1} & \mathbf{A}_{s2} & \cdots & \mathbf{A}_{sr} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \cdots & \mathbf{B}_{1t} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \cdots & \mathbf{B}_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B}_{r1} & \mathbf{B}_{r2} & \cdots & \mathbf{B}_{rt} \end{pmatrix}.$$

If the pairs $(\mathbf{A}_{ik}, \mathbf{B}_{kj})$ are conformable, then $\mathbf{A}$ and $\mathbf{B}$ are said to be *conformably partitioned*. For such matrices, the product $\mathbf{AB}$ is formed by combining the blocks exactly the same way as the scalars are combined in ordinary matrix multiplication. That is, the (i, j)-block in $\mathbf{AB}$ is

$$\mathbf{A}_{i1}\mathbf{B}_{1j} + \mathbf{A}_{i2}\mathbf{B}_{2j} + \cdots + \mathbf{A}_{ir}\mathbf{B}_{rj}.$$

Textbook

Example 3.6.6

Block multiplication is particularly useful when there are patterns in the matrices to be multiplied. Consider the partitioned matrices

$$\mathbf{A} = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right) = \begin{pmatrix} \mathbf{C} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix}, \quad \mathbf{B} = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 4 \end{array}\right) = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{C} & \mathbf{C} \end{pmatrix},$$

where

$$\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{C} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Using block multiplication, the product $\mathbf{AB}$ is easily computed to be

$$\mathbf{AB} = \begin{pmatrix} \mathbf{C} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{C} & \mathbf{C} \end{pmatrix} = \begin{pmatrix} 2\mathbf{C} & \mathbf{C} \\ \mathbf{I} & \mathbf{0} \end{pmatrix} = \left(\begin{array}{cc|cc} 2 & 4 & 1 & 2 \\ 6 & 8 & 3 & 4 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right).$$

Textbook

Example 3.6.7

Reducibility. Suppose that $\mathbf{T}_{n\times n}\mathbf{x} = \mathbf{b}$ represents a system of linear equations in which the coefficient matrix is *block triangular*. That is, $\mathbf{T}$ can be partitioned as

$$\mathbf{T} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{pmatrix}, \quad \text{where} \quad \mathbf{A} \text{ is } r \times r \text{ and } \mathbf{C} \text{ is } n - r \times n - r. \tag{3.6.3}$$

If $\mathbf{x}$ and $\mathbf{b}$ are similarly partitioned as $\mathbf{x} = \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix}$, then block multiplication shows that $\mathbf{Tx} = \mathbf{b}$ reduces to two smaller systems

$$\begin{aligned} \mathbf{A}\mathbf{x}_1 + \mathbf{B}\mathbf{x}_2 &= \mathbf{b}_1, \\ \mathbf{C}\mathbf{x}_2 &= \mathbf{b}_2, \end{aligned}$$

so if all systems are consistent, a block version of back substitution is possible—i.e., solve $\mathbf{C}\mathbf{x}_2 = \mathbf{b}_2$ for $\mathbf{x}_2$, and substituted this back into $\mathbf{A}\mathbf{x}_1 = \mathbf{b}_1 - \mathbf{B}\mathbf{x}_2$, which is then solved for $\mathbf{x}_1$.

Matrix Inversion

For a given square matrix $\mathbf{A}_{n\times n}$, the matrix $\mathbf{B}_{n\times n}$ that satisfies the conditions

$$\mathbf{AB} = \mathbf{I}_n \quad \text{and} \quad \mathbf{BA} = \mathbf{I}_n$$

is called the *inverse* of $\mathbf{A}$ and is denoted by $\mathbf{B} = \mathbf{A}^{-1}$. Not all square matrices are invertible—the zero matrix is a trivial example, but there are also many nonzero matrices that are not invertible. An invertible matrix is said to be *nonsingular*, and a square matrix with no inverse is called a *singular matrix*.

Existence of an Inverse

For an $n \times n$ matrix $\mathbf{A}$, the following statements are equivalent.

- $\mathbf{A}^{-1}$ exists ($\mathbf{A}$ is nonsingular). (3.7.5)
- $rank\,(\mathbf{A}) = n$. (3.7.6)
- $\mathbf{A} \xrightarrow{\text{Gauss–Jordan}} \mathbf{I}$. (3.7.7)
- $\mathbf{Ax} = \mathbf{0}$ implies that $\mathbf{x} = \mathbf{0}$. (3.7.8)

Properties of Matrix Inversion

For nonsingular matrices $\mathbf{A}$ and $\mathbf{B}$, the following properties hold.

- $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$. (3.7.13)
- The product $\mathbf{AB}$ is also nonsingular. (3.7.14)
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ (the reverse order law for inversion). (3.7.15)
- $\left(\mathbf{A}^{-1}\right)^{T} = \left(\mathbf{A}^{T}\right)^{-1}$ and $\left(\mathbf{A}^{-1}\right)^{*} = (\mathbf{A}^{*})^{-1}$. (3.7.16)

Products of Nonsingular Matrices Are Nonsingular

If $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ are each $n \times n$ nonsingular matrices, then the product $\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k$ is also nonsingular, and its inverse is given by the reverse order law. That is,

$$(\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k)^{-1} = \mathbf{A}_k^{-1}\cdots\mathbf{A}_2^{-1}\mathbf{A}_1^{-1}.$$

Sherman–Morrison Formula

- If $\mathbf{A}_{n\times n}$ is nonsingular and if $\mathbf{c}$ and $\mathbf{d}$ are $n \times 1$ columns such that $1 + \mathbf{d}^T\mathbf{A}^{-1}\mathbf{c} \neq 0$, then the sum $\mathbf{A} + \mathbf{c}\mathbf{d}^T$ is nonsingular, and

$$\left(\mathbf{A} + \mathbf{c}\mathbf{d}^T\right)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{c}\mathbf{d}^T\mathbf{A}^{-1}}{1 + \mathbf{d}^T\mathbf{A}^{-1}\mathbf{c}}. \tag{3.8.2}$$

- The *Sherman–Morrison–Woodbury formula* is a generalization. If $\mathbf{C}$ and $\mathbf{D}$ are $n \times k$ such that $(\mathbf{I} + \mathbf{D}^T\mathbf{A}^{-1}\mathbf{C})^{-1}$ exists, then

$$(\mathbf{A} + \mathbf{C}\mathbf{D}^T)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}(\mathbf{I} + \mathbf{D}^T\mathbf{A}^{-1}\mathbf{C})^{-1}\mathbf{D}^T\mathbf{A}^{-1}. \tag{3.8.3}$$

Textbook

Example 3.8.1

Problem: Start with $\mathbf{A}$ and $\mathbf{A}^{-1}$ given below. Update $\mathbf{A}$ by adding 1 to a_{21}, and then use the Sherman–Morrison formula to update $\mathbf{A}^{-1}$:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \quad \text{and} \quad \mathbf{A}^{-1} = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix}.$$

Solution: The updated matrix is

$$\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \mathbf{A} + \mathbf{e}_2 \mathbf{e}_1^T.$$

Applying the Sherman–Morrison formula yields the updated inverse

$$\mathbf{B}^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{e}_2\mathbf{e}_1^T\mathbf{A}^{-1}}{1 + \mathbf{e}_1^T\mathbf{A}^{-1}\mathbf{e}_2} = \mathbf{A}^{-1} - \frac{[\mathbf{A}^{-1}]_{*2}[\mathbf{A}^{-1}]_{1*}}{1 + [\mathbf{A}^{-1}]_{12}}$$

$$= \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix} - \frac{\begin{pmatrix} -2 \\ 1 \end{pmatrix} \begin{pmatrix} 3 & -2 \end{pmatrix}}{1 - 2} = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}.$$