Chapter 2

More Properties of Matrices and Matrix Arithmetic

When dealing with matrices, we should already know that

multiplication is NOT commutative

$$\mathbf{AB} \neq \mathbf{BA}$$

Distributive and Associative Properties

However, matrix multiplication DOES satisfy the distributive and associative properties:

- $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$ (*distributive*)
- $(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA}+\mathbf{CA}$ (*distributive*)
- $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C} = \mathbf{ABC}$ (*associative*)

Example: Distributive and Associative Properties

Simplify:

$$\mathbf{A}(\lambda\mathbf{I} + \mathbf{A}^{-1}\mathbf{C})$$

$$\begin{aligned}\mathbf{A}(\lambda\mathbf{I} + \mathbf{A}^{-1}\mathbf{C}) &= \mathbf{A}\lambda\mathbf{I} + \mathbf{A}(\mathbf{A}^{-1}\mathbf{C}) \\ &= \lambda\mathbf{A} + \mathbf{C}\end{aligned}$$

Properties of the Matrix Transpose

There are also few properties that you should be familiar with regarding the matrix transpose operation:

- $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
- $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ *(reverse-order)*
- $(\mathbf{ABCD})^T = \mathbf{D}^T\mathbf{C}^T\mathbf{B}^T\mathbf{A}^T$ (*reverse-order, any number of factors*)
- $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$ (*Often, we write* $\mathbf{A}^{-T}$)
- $(\alpha\mathbf{A})^T = \alpha\mathbf{A}^T$ (*transpose of scalar is same scalar*)
- $(\mathbf{A}^T)^T = \mathbf{A}$

Properties of the Matrix Transpose

There are a few properties that you should be familiar with regarding the matrix transpose operation:

- $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
- $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$ (*reverse-order*)
- $(\mathbf{ABCD})^T = \mathbf{D}^T\mathbf{C}^T\mathbf{B}^T\mathbf{A}^T$ (*reverse-order, any number of factors*)
- $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$ (*Often, we write* $\mathbf{A}^{-T}$)
- $(\alpha\mathbf{A})^T = \alpha\mathbf{A}^T$ (*transpose of scalar is same scalar*)
- $(\mathbf{A}^T)^T = \mathbf{A}$

Example: Is this matrix symmetric?

1. $\mathbf{X}^T\mathbf{X}$
   - Is the transpose equal to the original?

$$(\mathbf{X}^T\mathbf{X})^T = \mathbf{X}^T(\mathbf{X}^T)^T = \mathbf{X}^T\mathbf{X} \qquad \text{yes!}$$

2. $\mathbf{I} + \mathbf{x}\mathbf{x}^T$
   - Is the transpose equal to the original?

$$(\mathbf{I} + \mathbf{x}\mathbf{x}^T)^T = \mathbf{I}^T + (\mathbf{x}\mathbf{x}^T)^T = \mathbf{I} + \mathbf{x}\mathbf{x}^T \qquad \text{yes!}$$

3. If $\mathbf{A}$ is symmetric ($\mathbf{A} = \mathbf{A}^T$) and $\mathbf{B}$ is symmetric ($\mathbf{B} = \mathbf{B}^T$), is the product $\mathbf{AB}$ symmetric?
   - Is the transpose equal to the original?

$$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T = \mathbf{BA} \qquad \text{no!}$$

Example: Simplifying Expressions

Simplify the expression $(\mathbf{A}^T\mathbf{B})^T\mathbf{A}^{-1}$

$$
\begin{aligned}
(\mathbf{A}^T\mathbf{B})^T\mathbf{A}^{-1} &= \mathbf{B}^T(\mathbf{A}^T)^T\mathbf{A}^{-1} && \textit{reverse-order law} \\
&= \mathbf{B}^T\mathbf{A}\mathbf{A}^{-1} && (\mathbf{A}^T)^T = \mathbf{A} \\
&= \mathbf{B}^T. && \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}
\end{aligned}
$$

Check Your Understanding

1. Simplify the following expressions:
   - $\mathbf{C}^{-1}[(\mathbf{A}+\mathbf{B})(\mathbf{C}^T)]^T$
   - $(\mathbf{X}^T\mathbf{X})^T(\mathbf{X}^T\mathbf{X})^{-1}$
2. Determine whether the following matrices are symmetric:
   - $\mathbf{X}\mathbf{X}^T$
   - $\mathbf{A}+\mathbf{A}^T$
   - **ABA** if both **A** and **B** are symmetric.

Check Your Understanding - Solution

1. Simplify the following expressions:
   - $\mathbf{C}^{-1}[(\mathbf{A}+\mathbf{B})(\mathbf{C}^T)]^T = \mathbf{A}^T + \mathbf{B}^T$
   - $(\mathbf{X}^T\mathbf{X})^T(\mathbf{X}^T\mathbf{X})^{-1} = \mathbf{I}$
2. Determine whether the following matrices are symmetric:
   - $\mathbf{X}\mathbf{X}^T$ yes!
   - $\mathbf{A}+\mathbf{A}^T$ yes!
   - $\mathbf{ABA}$ if both $\mathbf{A}$ and $\mathbf{B}$ are symmetric yes!.

Matrix Powers

We define/compute powers of a matrix in the same way we do for scalars:

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

$$\mathbf{A}^1 = \mathbf{A}$$

$$\mathbf{A}^2 = \mathbf{AA}$$

$$\mathbf{A}^3 = \mathbf{AAA}$$

When powering a product of matrices, exercise caution:

$$(\mathbf{AB})^2 = (\mathbf{AB})(\mathbf{AB}) = \mathbf{ABAB} \neq \mathbf{A}^2\mathbf{B}^2$$

Chapter 2

Arithmetic with Partitioned Matrices

Partitioned Matrices

We will often want to consider a matrix as a collection of either rows or columns (or perhaps "blocks") rather than individual elements. When we write $\mathbf{A} = (\mathbf{A}_1|\mathbf{A}_2|\dots|\mathbf{A}_n)$ we are viewing the matrix $\mathbf{A}$ as collection of column vectors, $\mathbf{A}_i$, in the following way:

$$\mathbf{A} = (\mathbf{A}_1|\mathbf{A}_2|\dots|\mathbf{A}_n) = \begin{array}{c} \begin{array}{cccc} \mathbf{A}_1 & \mathbf{A}_2 & \dots & \mathbf{A}_n \end{array} \\ \left(\begin{array}{c|c|c|c} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{n2} & \dots & a_{mn} \end{array}\right) \end{array}$$

Similarly, we can write $\mathbf{A}$ as a collection of row vectors:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 \\ \hline \mathbf{A}_2 \\ \hline \mathbf{A}_3 \\ \hline \vdots \\ \hline \mathbf{A}_m \end{pmatrix} = \begin{array}{c} \mathbf{A}_1 \\ \hline \mathbf{A}_2 \\ \hline \mathbf{A}_3 \\ \hline \vdots \\ \hline \mathbf{A}_m \end{array} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \hline a_{21} & a_{22} & & a_{2n} \\ \hline a_{31} & a_{32} & \dots & a_{3n} \\ \hline \vdots & \vdots & & \vdots \\ \hline a_{m1} & a_{n2} & \dots & a_{mn} \end{pmatrix}$$

We could even draw divisions in the matrix to partition it into *blocks*:

$$\mathbf{A} = \left(\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{array}\right) = \left(\begin{array}{c|c} \mathbf{B} & \mathbf{C} \\ \hline \mathbf{D} & \mathbf{F} \end{array}\right)$$

Why is this useful? As long as two matrices are partitioned *conformably*, we can actually multiply them as if the partitioned blocks are entries in a new matrix:

$$\left(\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{array}\right) \left(\begin{array}{c} v_1 \\ v_2 \\ \hline v_3 \\ v_4 \end{array}\right)$$

$$= \left(\begin{array}{c|c} \mathbf{B} & \mathbf{C} \\ \hline \mathbf{D} & \mathbf{F} \end{array}\right) \left(\begin{array}{c} \mathbf{u} \\ \hline \mathbf{w} \end{array}\right)$$

$$= \left(\begin{array}{c} \mathbf{Bu} + \mathbf{Cw} \\ \hline \mathbf{Du} + \mathbf{Fw} \end{array}\right)$$

If you work out all of the multiplication and addition from the previous example, you'll see we aren't really saving any time in the computation. But, we save time in the notation and often we can represent unique concepts by ordering the rows and columns into special blocks.

You'll recall from a previous worksheet that you've worked through several views of matrix multiplication. These can always be found in general from partitioned matrix arithmetic.

Partitioned Matrices: Arithmetic

Let's consider matrix multiplication a few ways, starting with our original formulation.

$$\mathbf{X} = (\mathbf{X}_1|\mathbf{X}_2|\mathbf{X}_3|\dots|\mathbf{X}_p) \quad \text{and} \quad \mathbf{X}^T = \begin{pmatrix} \mathbf{X}_1^T \\ \hline \mathbf{X}_2^T \\ \hline \mathbf{X}_3^T \\ \hline \vdots \\ \hline \mathbf{X}_p^T \end{pmatrix}$$

So, what can we say about the matrix product $\mathbf{X}^T\mathbf{X}$?

Partitioned Matrices: Arithmetic

$$\mathbf{X}^T\mathbf{X} = \begin{pmatrix} \mathbf{X}_1^T \\ \hline \mathbf{X}_2^T \\ \hline \mathbf{X}_3^T \\ \hline \vdots \\ \hline \mathbf{X}_p^T \end{pmatrix} (\mathbf{X}_1|\mathbf{X}_2|\mathbf{X}_3|\dots|\mathbf{X}_p)$$

Is this partitioning *conformable* for multiplication? Always want to multiply *Row* $\times$ *Column*. If we do that, we will be multiplying $\mathbf{X}_i^T\mathbf{X}_j$ at each step. Do these calculations make sense?

$$\underset{1\times n}{\mathbf{X}_i^T} \;\; \underset{n\times 1}{\mathbf{X}_j}$$

Yes! So the partitioning is conformable for multiplication.

Partitioned Matrices: Arithmetic

$$\mathbf{X}^T\mathbf{X} = \left(\begin{array}{c} \mathbf{X}_1^T \\ \hline \mathbf{X}_2^T \\ \hline \mathbf{X}_3^T \\ \hline \vdots \\ \hline \mathbf{X}_p^T \end{array}\right) (\mathbf{X}_1|\mathbf{X}_2|\mathbf{X}_3|\dots|\mathbf{X}_p)$$

$$= \begin{pmatrix} \mathbf{X}_1^T\mathbf{X}_1 & \mathbf{X}_1^T\mathbf{X}_2 & \mathbf{X}_1^T\mathbf{X}_3 & \dots & \mathbf{X}_1^T\mathbf{X}_p \\ \mathbf{X}_2^T\mathbf{X}_1 & \mathbf{X}_2^T\mathbf{X}_2 & \mathbf{X}_2^T\mathbf{X}_3 & \dots & \mathbf{X}_2^T\mathbf{X}_p \\ \mathbf{X}_3^T\mathbf{X}_1 & \mathbf{X}_3^T\mathbf{X}_3 & \mathbf{X}_3^T\mathbf{X}_3 & \dots & \mathbf{X}_3^T\mathbf{X}_p \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{X}_p^T\mathbf{X}_1 & \mathbf{X}_p^T\mathbf{X}_2 & \mathbf{X}_p^T\mathbf{X}_3 & \dots & \mathbf{X}_p^T\mathbf{X}_p \end{pmatrix}$$

Partitioned Matrices: Arithmetic

$$
= \begin{pmatrix}
\mathbf{x}_1^T\mathbf{x}_1 & \mathbf{x}_1^T\mathbf{x}_2 & \mathbf{x}_1^T\mathbf{x}_3 & \dots & \mathbf{x}_1^T\mathbf{x}_p \\
\mathbf{x}_2^T\mathbf{x}_1 & \mathbf{x}_2^T\mathbf{x}_2 & \mathbf{x}_2^T\mathbf{x}_3 & \dots & \mathbf{x}_2^T\mathbf{x}_p \\
\mathbf{x}_3^T\mathbf{x}_1 & \mathbf{x}_3^T\mathbf{x}_3 & \mathbf{x}_3^T\mathbf{x}_3 & \dots & \mathbf{x}_3^T\mathbf{x}_p \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\mathbf{x}_p^T\mathbf{x}_1 & \mathbf{x}_p^T\mathbf{x}_2 & \mathbf{x}_p^T\mathbf{x}_3 & \dots & \mathbf{x}_p^T\mathbf{x}_p
\end{pmatrix}
$$

Diagonal elements contain the sum of squares for each column (variable).

Another way that we will end up using partitioned matrices is to represent consider the opposite formulation where the matrix on the right is partitioned into columns and the matrix on the left is partitioned into rows.

$$\mathbf{U}_{m\times r}\mathbf{V}^T_{r\times n}$$

$$(\mathbf{U}_1|\mathbf{U}_2|\mathbf{U}_3|\dots|\mathbf{U}_r)\begin{pmatrix}\mathbf{V}_1^T\\ \hline \mathbf{V}_2^T\\ \hline \mathbf{V}_3^T\\ \hline \vdots\\ \hline \mathbf{V}_r^T\end{pmatrix}$$

Partitioned Matrices: Arithmetic

$$\mathbf{U}_{m\times r}\mathbf{V}^T_{r\times n}$$

$$(\mathbf{U}_1|\mathbf{U}_2|\mathbf{U}_3|\ldots|\mathbf{U}_r)\begin{pmatrix}\mathbf{V}_1^T \\ \hline \mathbf{V}_2^T \\ \hline \mathbf{V}_3^T \\ \hline \vdots \\ \hline \mathbf{V}_r^T\end{pmatrix}$$

$$= \mathbf{U}_1\mathbf{V}_1^T + \mathbf{U}_2\mathbf{V}_2^T + \mathbf{U}_3\mathbf{V}_3^T + \cdots + \mathbf{U}_r\mathbf{V}_r^T$$

Let's back up a minute...

$$\mathbf{U}_{m\times r}\mathbf{V}^T_{r\times n}$$

$$(\mathbf{U}_1|\mathbf{U}_2|\mathbf{U}_3|\dots|\mathbf{U}_r)\begin{pmatrix}\mathbf{V}_1^T\\ \hline \mathbf{V}_2^T\\ \hline \mathbf{V}_3^T\\ \hline \vdots\\ \hline \mathbf{V}_r^T\end{pmatrix}$$

Is this partitioning *conformable* for multiplication? Again, want to multiply *Row* $\times$ *Column*. If we do that, we will be multiplying $\mathbf{U}_i\mathbf{V}_j^T$ at each step. Do these calculations make sense?

$$\underset{m\times 1}{\mathbf{U}_i}\ \underset{1\times n}{\mathbf{V}_j^T}$$

Sure! Each element in the sum will be an $m \times n$ matrix!

Check Your Understanding

Compute the following matrix product, using block multiplication:

$$\left(\begin{array}{cc|c} 1 & -1 & 2 \\ 0 & 2 & 3 \\ \hline 0 & 0 & 1 \end{array}\right) \left(\begin{array}{cc|c} -2 & 1 & -1 \\ 0 & 1 & 2 \\ \hline 0 & 0 & 1 \end{array}\right)$$

Check Your Understanding - Solution

$$\left(\begin{array}{cc|c} 1 & -1 & 2 \\ 0 & 2 & 3 \\ \hline 0 & 0 & 1 \end{array}\right)\left(\begin{array}{cc|c} -2 & 1 & -1 \\ 0 & 1 & 2 \\ \hline 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{cc|c} -2 & 0 & -1 \\ 0 & 2 & 7 \\ \hline 0 & 0 & 1 \end{array}\right)$$

(1,1)-block: $\begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} -2 & 1 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \end{pmatrix}\begin{pmatrix} 0 & 0 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}$

(1,2)-block: $\begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \end{pmatrix}(1) = \begin{pmatrix} -1 \\ 7 \end{pmatrix}$

(2,1)-block: $\begin{pmatrix} 0 & 0 \end{pmatrix}\begin{pmatrix} -2 & 1 \\ 0 & 1 \end{pmatrix} + (1)\begin{pmatrix} 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}$

(2,2)-block: $\begin{pmatrix} 0 & 0 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \end{pmatrix} + (1)(1) = 1$