1 Overview

Main ideas:

1. Terminology: (i, j)-th entry, diagonal entries, main diagonal, diagonal matrix, zero matrix
2. Sums and scalar multiplication: equality of matrices, sums of matrices, scalar multiple of a matrix, properties of addition and scalar multiplication
3. Matrix multiplication: corresponds to composition of linear transformations, definition of matrix product, row-column rule for computing matrix product, properties of matrix multiplication
4. Powers of a matrix, transpose of a matrix

Examples in text:

1. sum of two matrices
2. scalar multiple of two matrices
3. matrix product using definition
4. size of matrix products
5. matrix product using row-column rule
6. non-commutativity of matrix product
7. transpose of a matrix

2 Discussion and Worked Examples

2.1 Matrix Multiplication

Suppose we have two linear transformations S and T, with corresponding standard matrices A and B. We would like to find a matrix corresponding to the composition $S \circ T$. Of course, the codomain of T must equal the domain of S in order for this to make sense.

If $S : \mathbb{R}^n \to \mathbb{R}^m$ and $T : \mathbb{R}^p \to \mathbb{R}^n$, then the composition $S \circ T$ maps $\mathbb{R}^p \to \mathbb{R}^m$.

$$\begin{tikzcd} \mathbb{R}^p \arrow[r, "T"] \arrow[rr, bend right, "S\circ T"'] & \mathbb{R}^n \arrow[r, "S"] & \mathbb{R}^m \end{tikzcd}$$

In this case, A is $(m \times n)$ and B is $(n \times p)$. We wish to find a $(m \times p)$ matrix, which we will call AB corresponding to $S \circ T$.

Let x be a vector in $\mathbb{R}^p$. Recall that the matrix-vector product Bx is a linear combination of the columns of B, with the entries of x as weights. If $b_1, \dots, b_p$ are the columns of B, this is

$$Bx \;=\; x_1\, b_1 \;+\; \dots \;+\; x_p\, b_p$$

Applying A next and using linearity,

$$A(Bx) \;=\; A\big(x_1\, b_1 \;+\; \dots \;+\; x_p\, b_p\big) \;=\; x_1\, Ab_1 \;+\; \dots \;+\; x_p\, Ab_p$$

This is a linear combination of the vectors $Ab_1, \dots, Ab_p$ with the entries of the vector x as weights. We can represent this as a matrix-vector product, where the columns of the matrix are the vectors $Ab_1, \dots, Ab_p$, and the vector is x:

$$\Big[Ab_1 | \dots | Ab_p\Big]\, x$$

Thus we have found a matrix which represents the successive application of B and A. We define the product AB of A and B to be this matrix, so that the AB corresponds to the successive actions of B and A:

$$AB \overset{\text{definition}}{=} \Big[Ab_1 | \dots | Ab_p \Big] \qquad \text{where } b_1, \dots, b_p \text{ are the columns of } B$$

Notice that each column of AB is a linear combination of the columns of A using weights from the corresponding column of B.

Example For the matrices A and B below, compute the matrix products AB and BA if they are defined:

$$A = \begin{bmatrix} 2 & 0 & -1 \\ 4 & -5 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

Since A is (2×3) and B is (2×2), the product AB is undefined, but

$$BA = \begin{bmatrix} 10 & -10 & 3 \\ 0 & -5 & 4 \end{bmatrix}$$

The **row-column rule** for computing matrix products (like the row-column rule for computing matrix-vector products) uses the dot product of the rows of the matrix on the left with the columns of the matrix on the right.

Properties of Matrix Multiplication Matrix multiplication satisfies many desirable properties, such as associativity, right and left distributivity with respect to matrix addition, and commutativity with scalar multiplication. The identity matrix $\mathbb{I}_n$ is a multiplicative identity. However, matrix multiplication does not satisfy all the properties that we are used to relying on. In particular,

1. Matrix multiplication is not commutative in general.
2. Cancellation laws do not hold for matrix multiplication, in general.
3. If a product of two matrices is zero, we cannot conclude that one of the two matrices is zero.

Activity Commutativity/non-commutativity of matrix multiplication and matrix powers.

2.2 Transpose of a Matrix

The transpose of a matrix A is the matrix whose columns are the rows of A and whose rows are the columns of A. The transpose of A is denoted A^T.

Example For the matrices A and B below, compute the transposes A^T, B^T, and compute matrix products $A^T B^T$ and $B^T A^T$ if they are defined:

$$A = \begin{bmatrix} 2 & 0 & -1 \\ 4 & -5 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

Computing the transforms of A and B is straightforward: just switch columns and rows.

$$A^T = \begin{bmatrix} 2 & 4 \\ 0 & -5 \\ -1 & 2 \end{bmatrix} \qquad B^T = \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$$

Notice that, since A is (2×3), A^T is (3×2) and since B is (2×2), B^T is also (2×2). Thus the matrix product $A^T B^T$ is defined and

$$A^T B^T = \begin{bmatrix} 10 & 0 \\ -10 & -5 \\ 3 & 4 \end{bmatrix}$$

but the matrix $B^T A^T$ is not defined. Comparing with the previous example, it is clear that

$$A^T B^T = (BA)^T$$

Note This is true in general: the transpose of a product of matrices is the product of the transposes of the matrices *in the reverse order.*