Question 1

Given Matrices

Matrix **A** has two rows and three columns, so its dimensions are 2×3. Matrix **B** has four rows and two columns, so its dimensions are 4×2.

Calculating the Dimensions of the Required Matrices

Dimensions of the Matrix BA For this multiplication to be defined, the number of columns in matrix **B** (which is 2) must match the number of rows in matrix **A** (which is also 2). The resulting matrix will have dimensions 4×3. **Answer:** 4×3

Dimensions of the Transposed Matrix B^T The transpose of a matrix swaps its rows and columns. Since matrix **B** has dimensions 4×2, its transposed form will have dimensions 2×4. **Answer:** 2×4

Dimensions of the Matrix $B^T A$ The dimensions of matrix B^T are 2×4. The dimensions of matrix A are 2×3. The number of columns in B^T (which is 2) matches the number of rows in A (which is also 2), so the multiplication is valid. The resulting matrix will have dimensions 2×3. **Answer:** 2×3

Dimensions of the Matrix $A^T B$ The dimensions of matrix A^T are 3×2. The dimensions of matrix B are 4×2. For the multiplication to be defined, the number of columns in A^T (which is 2) must match the number of rows in B (which is 4), but they do not match. Therefore, this multiplication is **not defined**. **Answer:** Not defined

—

Question 1.1

First, we define the two matrices:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix}$$

1. BA

Since the number of columns in B matches the number of rows in A, the multiplication is valid, and the result will be a 4×3 matrix:

$$BA = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

Result:

$$BA = \begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} & b_{11}a_{13} + b_{12}a_{23} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} & b_{21}a_{13} + b_{22}a_{23} \\ b_{31}a_{11} + b_{32}a_{21} & b_{31}a_{12} + b_{32}a_{22} & b_{31}a_{13} + b_{32}a_{23} \\ b_{41}a_{11} + b_{42}a_{21} & b_{41}a_{12} + b_{42}a_{22} & b_{41}a_{13} + b_{42}a_{23} \end{bmatrix}$$

2. B^T

The transpose of matrix B is:

$$B^T = \begin{bmatrix} b_{11} & b_{21} & b_{31} & b_{41} \\ b_{12} & b_{22} & b_{32} & b_{42} \end{bmatrix}$$

3. $B^T A$

Since the number of columns in B^T matches the number of rows in A, this multiplication is valid, resulting in a 2×3 matrix:

$$B^T A = \begin{bmatrix} b_{11} & b_{21} & b_{31} & b_{41} \\ b_{12} & b_{22} & b_{32} & b_{42} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

Result:

$$B^T A = \begin{bmatrix} b_{11}a_{11} + b_{21}a_{21} & b_{11}a_{12} + b_{21}a_{22} & b_{11}a_{13} + b_{21}a_{23} \\ b_{12}a_{11} + b_{22}a_{21} & b_{12}a_{12} + b_{22}a_{22} & b_{12}a_{13} + b_{22}a_{23} \end{bmatrix}$$

4. $A^T B$

The dimensions of A^T are 3×2, while the dimensions of B are 4×2. Since the number of columns in A^T (which is 3) does not match the number of rows in B (which is 4), the multiplication is **not defined**.

Question 1.2

(a)

Given:

$$x^{(i)} = \begin{bmatrix} x_1^{(i)} & x_2^{(i)} \end{bmatrix} \in R^{1 \times 2}$$

and

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \in R^{2\times 1}$$

Computing $x^{(i)}\theta$:

$$x^{(i)}\theta = \begin{bmatrix} x_1^{(i)} & x_2^{(i)} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$$

Since the number of columns in $x^{(i)}$ (2) matches the number of rows in θ (2), the multiplication is valid.

$$x^{(i)}\theta = x_1^{(i)}\theta_1 + x_2^{(i)}\theta_2 \in R$$

The result is a scalar.
Computing $X\theta$:
If we stack all samples into a matrix X:

$$X = \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(n)} \end{bmatrix} \in R^{n\times 2}$$

then:

$$X\theta = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} \\ x_1^{(2)} & x_2^{(2)} \\ \vdots & \vdots \\ x_1^{(n)} & x_2^{(n)} \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$$

Since the number of columns in X (2) matches the number of rows in θ (2), the multiplication is valid.

$$X\theta = \begin{bmatrix} x_1^{(1)}\theta_1 + x_2^{(1)}\theta_2 \\ x_1^{(2)}\theta_1 + x_2^{(2)}\theta_2 \\ \vdots \\ x_1^{(n)}\theta_1 + x_2^{(n)}\theta_2 \end{bmatrix} \in R^{n\times 1}$$

The result is a column vector with dimensions $n \times 1$.
—

(b)

We know that:

$$z^T z = \sum_{i=1}^{n} (z^{(i)})^2$$

which represents the sum of squares of a set of numbers, written in vector-matrix form.

Now consider the following function:

$$J = \sum_{i=1}^{n} \left(x^{(i)}\theta - y^{(i)} \right)^2$$

First, we rewrite J in terms of a vector:

$$J = \sum_{i=1}^{n} \left(x^{(i)}\theta - y^{(i)} \right)^2$$

Next, compute the gradient $\nabla_\theta J$, i.e., differentiate J with respect to the vector θ.