01: Einstein Summation Convention and the Symmetry of the Dot Product

Chapter Goal: To introduce a new, efficient notation for summations (**Einstein Summation Convention**) and use it to gain a new perspective on matrix multiplication and the Dot Product.

---

1. Core Idea: A New Way to Write Matrix Multiplication

- **Problem:** Writing the Σ (sigma) symbol for summations in matrix multiplication is long and repetitive.
- **Einstein Summation Convention:** A notational "shortcut".
- **The Rule:** If an index (like `j`) appears repeated in a single term, it is implicitly assumed that there is a summation (Σ) across all possible values of that index.
- **Example:**
  - Standard Notation: $(AB)_{ik} = \sum_j A_{ij}B_{jk}$
  - Einstein Notation: $(AB)_{ik} = A_{ij}B_{jk}$
- **Usefulness:**
  - It's very concise and commonly used in advanced physics and computation.
  - It translates directly into loops in programming code. To calculate `AB`, you need 3 loops (for `i`, `k`, and `j` as the accumulator).

---

2. "Aha!" Moment #1: Non-Square Matrix Multiplication

Einstein notation makes the "matching" rule for matrix multiplication very clear.

$$(AB)_{ik} = A_{ij}B_{jk}$$

This multiplication is only possible if the dimension being summed over (`j`) is the same size for both matrices.

- **Meaning:** The number of **columns** of the left matrix (`A`) must be equal to the number of **rows** of the right matrix (`B`).
- **Consequence:** We can multiply non-square matrices, as long as their "inner dimensions" match.
  - **Example:** A `(2 x 3)` matrix can be multiplied by a `(3 x 4)` matrix.
  - The result will have the dimensions of the "outer numbers": `(2 x 4)`.

- **Note:** Concepts like the inverse and determinant become complicated or undefined for non-square matrices.

---

3. "Aha!" Moment #2: The Dot Product as Matrix Multiplication

Let's look at the Dot Product again with Einstein notation:

- $\vec{u} \cdot \vec{v} = \sum_i u_i v_i$
- Einstein Notation: $u_i v_i$ (the index `i` is repeated).

Now, imagine we convert the column vector $\vec{u}$ into a row vector $\vec{u}^T$ (its transpose).

- $\vec{u}^T = [u_1, u_2, \dots, u_n]$ (a `1 x n` matrix)
- $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ (an `n x 1` column vector)

If we multiply them as matrices: $\vec{u}^T \vec{v}$

- This is a `(1 x n) * (n x 1)` matrix multiplication. The inner dimension (`n`) matches.
- The result will be a `(1 x 1)` matrix, which is a single number (a scalar).
- The calculation is: $u_1 v_1 + u_2 v_2 + \dots + u_n v_n$.
- **Conclusion:**

> The Dot Product $\vec{u} \cdot \vec{v}$ is computationally **IDENTICAL** to the matrix multiplication $\vec{u}^T \vec{v}$.

This shows a deep connection between the dot product (which we often think of geometrically) and matrix multiplication (which is a transformative action).

---

4. Geometric Proof of Dot Product Symmetry (Revisited)

This video provides a second visual proof (similar to 3Blue1Brown's) for why the dot product is symmetric ($\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$).

- **Idea:**
  1. Take two unit vectors, $\hat{u}$ and $\hat{e}_1$ (which is $\hat{\imath}$).
  2. The projection of $\hat{u}$ onto $\hat{e}_1$ is, by definition, the x-component of $\hat{u}$.
  3. Consider the projection of $\hat{e}_1$ onto $\hat{u}$.

4. Using a triangle symmetry argument, it can be shown that the lengths of these two projections are exactly the same.

- **Conclusion:**
  Since $\text{Projection}(\hat{u} \text{ onto } \hat{e}_1) = \text{Projection}(\hat{e}_1 \text{ onto } \hat{u})$, and we know the dot product is related to projection, this serves as a geometric justification for why $\hat{u} \cdot \hat{e}_1 = \hat{e}_1 \cdot \hat{u}$.
- **Final Message:**
  There is a beautiful relationship between the computational/algebraic world (matrix multiplication, Einstein Convention) and the geometric world (projections, symmetry). They are two different languages describing the same fundamental ideas.

---

5. Worked Practice Problems

In the previous lecture we saw the Einstein summation convention, in which we sum over any indices which are repeated. In traditional notation we might write, for example, $\sum_{j=1}^{3} A_{ij}v_j = A_{i1}v_1 + A_{i2}v_2 + A_{i3}v_3$. With the Einstein summation convention we can avoid the big sigma and write this as $A_{ij}v_j$. We know that we sum over j because it appears twice.

We saw that thinking about this type of notation helps us to multiply non-square matrices together. For example, consider the matrices

$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$,

and remember that in the A_{ij} notation the first index i represents the row number and the second index j represents the column number. For example, $A_{12} = 2$.

Let's define the matrix $C = AB$. Then in Einstein summation convention notation $C_{mn} = A_{mj}B_{jn}$.

Using the Einstein summation convention, calculate $C_{21} = A_{2j}B_{j1}$.

Let's use the following non-square matrices for our practice problems:

- **Matrix A (size 2 x 3):**

$$A = \begin{bmatrix} 5 & 1 & 0 \\ 2 & 3 & 4 \end{bmatrix}$$

- **Matrix B (size 3 x 2):**

$$B = \begin{bmatrix} 1 & 6 \\ 2 & 7 \\ 3 & 8 \end{bmatrix}$$

We want to find the matrix $C = A \cdot B$.

Problem 1: Calculating a Specific Element (C_{12})

Question:
Using the Einstein Summation Convention, calculate the value of $C_{12} = A_{1j}B_{j2}$.

Step-by-Step Solution:

1. **Translate the Einstein Notation:**
   - We see the notation $A_{1j}B_{j2}$.
   - The index `j` is repeated, so we know we will be summing over all possible values of `j`.
   - The values for `j` will go from 1 to 3 (since `A` has 3 columns and `B` has 3 rows).
   - The first index of `A` (the row) is fixed at `1`.
   - The second index of `B` (the column) is fixed at `2`.
   - **Meaning:** We will perform a Dot Product between **Row 1 of A** and **Column 2 of B**.
2. **Write the Summation Explicitly:**
   $C_{12} = A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{32}$
3. **Identify the Relevant Row and Column:**
   - **Row 1 of A:** $[A_{11}, A_{12}, A_{13}] = [5, 1, 0]$
   - **Column 2 of B:** $\begin{bmatrix} B_{12} \\ B_{22} \\ B_{32} \end{bmatrix} = \begin{bmatrix} 6 \\ 7 \\ 8 \end{bmatrix}$
4. **Perform the Calculation:**
   $C_{12} = (5 \cdot 6) + (1 \cdot 7) + (0 \cdot 8)$
   $C_{12} = 30 + 7 + 0$
   $C_{12} = 37$

Answer to Problem 1:
The value of C_{12} is **37**.

Problem 2: Calculating the Entire Result Matrix (`C = AB`)

Question:
Using the same matrices `A` and `B`, what is $C = A \cdot B$?

Step-by-Step Solution:

1. **Determine the Size of the Result Matrix:**
   - `A` is `(2 x 3)`.
   - `B` is `(3 x 2)`.
   - The "inner dimensions" (`3` and `3`) match.
   - The "outer dimensions" are `2` and `2`.
   - Therefore, matrix `C` will be size **2 x 2**. It will have the form:

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

2. **Calculate Each Element One by One:**
   - C_{11} **(Row 1 of A · Col 1 of B):**
     $[5, 1, 0] \cdot [1, 2, 3] = (5 \cdot 1) + (1 \cdot 2) + (0 \cdot 3) = 5 + 2 + 0 = \mathbf{7}$
   - C_{12} **(Row 1 of A · Col 2 of B):**
     We already calculated this in Problem 1!
     $[5, 1, 0] \cdot [6, 7, 8] = (5 \cdot 6) + (1 \cdot 7) + (0 \cdot 8) = 30 + 7 + 0 = \mathbf{37}$
   - C_{21} **(Row 2 of A · Col 1 of B):**
     $[2, 3, 4] \cdot [1, 2, 3] = (2 \cdot 1) + (3 \cdot 2) + (4 \cdot 3) = 2 + 6 + 12 = \mathbf{20}$
   - C_{22} **(Row 2 of A · Col 2 of B):**
     $[2, 3, 4] \cdot [6, 7, 8] = (2 \cdot 6) + (3 \cdot 7) + (4 \cdot 8) = 12 + 21 + 32 = \mathbf{65}$
3. **Assemble the Result Matrix `C`:**
   Now we place all the calculated numbers into our 2x2 matrix.

Answer to Problem 2:

$$C = \begin{bmatrix} 7 & 37 \\ 20 & 65 \end{bmatrix}$$

Tags: #mml-specialization #linear-algebra #einstein-summation #dot-product #matrix-multiplication