Explanatory Notes for 6.390

Shaunticlair Ruiz (Current TA)

Fall 2022

Matrix Math: Debugging

Matrices are useful, but we have to be careful of some **problems** when doing math with them.

These problems come up just when **multiplying** matrices like normal. But, they are much more common when trying to do **calculus**.

In this class, **matrix calculus** is entirely about taking **derivatives** of vectors and matrices.

We will not show here how to **find** these derivatives, but the important rules you need to compute our derivatives are in the **appendix**.

There's a document explaining vector derivatives coming soon!

Instead, we'll focus on those problems mentioned above:

Concept 1

Important things to remember about **matrix derivatives**:

- Often, matrix derivative rules look **similar** to regular derivatives rules. **HOWEVER**, they are **not** exactly the same.
- If you're confused about a **derivative**, or you aren't getting the result you expect, check the dimensions (**shape**) of your objects.

The second point is especially important: many problems in this course come down to problems with the **shape** of your matrix.

Issue 1: Your answer is transposed

Your final result might be the **transpose** of what it should be.

- **Example:** You wanted a **column** vector, but you got a **row** vector.

$$\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \neq \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}$$

- You might just need to **transpose** that result (or an earlier step), but not always.

Issue 2: Multiplication with invalid dimensions

You could be doing **multiplication** with **mismatched** dimensions.

- Remember that the "inner" dimensions need to **match**: you need to multiply $(a \times b) * (b \times c)$.
- **Example:** You are trying to multiply a (2×2) matrix by a (1×2) matrix.

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} * \begin{bmatrix} 3 & 4 \end{bmatrix} = \text{INVALID!}$$

- Sometimes you can fix it by **transposing** one of the vectors. Be **careful** which one you transpose.

$$\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} * \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \text{Valid multiplication!}$$

- You can also switch the **order**, but you will get a different result.

$$\begin{bmatrix} 3 & 4 \end{bmatrix} * \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} = \text{Also valid, but different!}$$

Because these give **different** answers, **only one** is going to be correct for your question.

Clarification 2

A **valid** expression just means it has an **answer**: **invalid** means you can't even **calculate** an answer.

But valid **does not** means we have a **correct** answer.

An expression can be **valid** and **incorrect**!

Example: $1/0$ is an **invalid** expression, while $1/2$ is **valid**. But that **doesn't** mean $1/2$ is the **answer** to our question!

It's up to you to figure out which one is **correct**.

Issue 3: Addition with Invalid Dimensions

You could be **adding** two matrices with **incompatible** dimensions.

- Both dimensions must match, **OR** the non-matching dimension is 1:
  - $(a \times b)$, $(a \times b)$, $(a \times 1)$, $(1 \times b)$, (1×1) can all add together.
- **Example:** You try to add a (3×2) matrix to a (2×3) matrix.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} + \begin{bmatrix} 7 & 8 & 9 \\ 1 & 2 & 3 \end{bmatrix} = ???$$

- You might have multiplied to get the **wrong shape** earlier, or, again, need a transpose.

$$\begin{bmatrix}1 & 2\\3 & 4\\5 & 6\end{bmatrix} + \begin{bmatrix}7 & 8\\9 & 1\\2 & 3\end{bmatrix} = \text{Valid!}$$

Be careful to figure out which one to transpose!

$$\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix} + \begin{bmatrix}7 & 8 & 9\\1 & 2 & 3\end{bmatrix} = \text{Also valid, but which one is right?}$$

Issue 4: Your answer has the wrong shape

Your **final** result is the completely **wrong shape**.

- **Example:** You got (3×3) instead of (2×2).

$$\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix} \neq \begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{bmatrix}$$

- You might have done $(3\times 2) * (2\times 3)$ instead of $(2\times 3) * (3\times 2)$. For a specific example:

$$\begin{bmatrix}50 & 14\\122 & 32\end{bmatrix} \neq \begin{bmatrix}11 & 16 & 21\\19 & 26 & 33\\27 & 36 & 45\end{bmatrix}$$

  The mystery here is revealed by seeing our **multiplication**: we **transposed** our matrices!

$$\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix}\begin{bmatrix}1 & 7\\2 & 8\\3 & 9\end{bmatrix} \neq \begin{bmatrix}1 & 4\\2 & 5\\3 & 6\end{bmatrix}\begin{bmatrix}7 & 8 & 9\\1 & 2 & 3\end{bmatrix}$$

  But we also could have switched the **order** of our matrices.

$$\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix}\begin{bmatrix}1 & 7\\2 & 8\\3 & 9\end{bmatrix} \neq \begin{bmatrix}1 & 7\\2 & 8\\3 & 9\end{bmatrix}\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix}$$

Summary

In general, you can solve most of these situations (**shape** problems) by taking a **transpose**, or changing **order** of multiplication.

If this doesn't work, you may have a more **significant** problem.

Concept 3

If you are struggling to compute **matrix** math, you may be having a problem with the **shape** of your matrices.

Often, the solution is either

- take a **transpose** to make shapes compatible
- change the **order** of multiplication.

However, there are often **multiple** ways to get the right **shape**, with the different answers.

If you have an official equation, you can try referencing that.

Otherwise, a good way to figure out the correct form is the use only one **data point** and see which equation seems logical.

Reference Equations

You might find these equations useful:

A is a matrix.

Key Equation 4

Taking a transpose twice gives the original matrix.

$$(A^T)^T = A$$

B is a matrix, and k is a real number.

Key Equation 5

A transpose is **linear**: it preserves addition and scalar multiplication.

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

$$(kA)^T = kA^T$$

And finally:

Key Equation 6

Transposes and multiplication can be swapped like this:

$$(AB)^{T} = B^{T}A^{T}$$