Math 210 **Fall 2020**
Rules of Matrix Arithmetic **Drew Armstrong**

Let A be an $\ell \times m$ matrix (i.e., ℓ rows and m columns) and let B be an $m' \times n$ matrix. If $m = m'$ then we define the $\ell \times n$ *product matrix* AB by requiring that

$$(AB)\mathbf{x} = A(B\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{R}^n.$$

If $m \neq m'$, i.e., if #(columns of A) $\neq$ #(rows of B), then the product matrix is **not defined**. We can compute the matrix AB with the following rules:

$$\begin{aligned}(i\text{th row of } AB) &= (i\text{th row of } A)B,\\ (j\text{th col of } AB) &= A(j\text{th col of } B),\\ (i,j \text{ entry of } AB) &= (i\text{th row of } A)(j\text{th col of } B).\end{aligned}$$

Note that the product of a row (on the left) times a column (on the right) is just the **dot product**. Furthermore, matrices of the same shape can be added componentwise and multiplied by scalars, just like vectors. Now let A, B, C be matrices and let s, t be scalars. Let O denote a zero matrix of any shape and let I denote a (square) identity matrix. Then the following rules hold (as long as the shapes match):

- $A + B = B + A$
- $A + (B + C) = (A + B) + C$
- $A + O = A$
- $s(A + B) = sA + sB$
- $(s + t)A = sA + tA$
- $s(AB) = (sA)B = A(sB)$
- $A(BC) = (AB)C$
- $A(B + C) = AB + AC$
- $(A + B)C = AC + BC$
- $AO = O$ and $OA = O$
- $AI = A$ and $IA = A$

Note that these rules include the rules of vector arithmetic as a special case because vectors are $n \times 1$ matrices and the dot product is a matrix product. Furthermore, the rules are easy to memorize because they all look obvious. The only difference is that matrix multiplication is not generally commutative:

$$AB \neq BA.$$

Next, if A has shape $m \times n$ then we define the $n \times m$ *transpose matrix* A^T as follows:

$$(i,j \text{ entry of } A^T) = (j,i \text{ entry of } A).$$

This operations satisfies the following additional rules:

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

- $(AB)^T = B^T A^T$

Maybe this last rule is a bit surprising? Let A be $\ell \times m$ and let B be $m \times n$, so that A^T is $m \times \ell$ and B^T is $n \times m$. Then the matrix $A^T B^T$ is **not defined** unless $\ell = n$. However, the matrix $B^T A^T$ is always defined and has the same shape as $(AB)^T$. So it makes sense. One important use of the matrix transpose is to express the dot product of vectors. If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are $n \times 1$ column vectors then $\mathbf{x}^T \mathbf{y}$ is a 1×1 scalar, which is just the dot product:

$$\mathbf{x}^T \mathbf{y} = \mathbf{x} \bullet \mathbf{y}.$$

Furthermore, since every 1×1 matrix is equal to its own transpose we have $\mathbf{x}^T \mathbf{y} = (\mathbf{x}^T \mathbf{y})^T = \mathbf{y}^T (\mathbf{x}^T)^T = \mathbf{y}^T \mathbf{x}$. [Remark: On the other hand, $\mathbf{x}\mathbf{y}^T$ and $\mathbf{y}\mathbf{x}^T$ are $n \times n$ matrices.]

Finally, we consider matrix inversion. If A and B are square matrices of the same size then we say that $A = B^{-1}$ and $B = A^{-1}$ (i.e., the matrices are *inverses* of each other) when

$$AB = I = BA.$$

[Subtle Remark: In fact we only need to check one of the identities $AB = I$ and $BA = I$, since each implies the other. But this fact is quite difficult to prove.] If $A\mathbf{x} = \mathbf{0}$ for some $\mathbf{x} \neq \mathbf{0}$ then the matrix A^{-1} does not exist. Otherwise, it does exist, and we may compute it using Gaussian elimination:

$$\left(\; A \mid I \; \right) \overset{\text{RREF}}{\rightsquigarrow} \left(\; I \mid A^{-1} \; \right).$$

Now let A and B be **any** square matrices of the same size (not necessarily inverses of each other) and suppose that A^{-1} and B^{-1} both exist. Then the following additional rules hold:

- $(A^{-1})^{-1} = A$
- $(sA)^{-1} = \frac{1}{s} A^{-1}$
- $(A^T)^{-1} = (A^{-1})^T$
- $(AB)^{-1} = B^{-1} A^{-1}$

Proof. The first rule is obvious. For the second we use the identity $(sA)(tB) = (st)(AB)$:

$$(sA)\left(\frac{1}{s} A^{-1}\right) = \left(s \cdot \frac{1}{s}\right)(AA^{-1}) = AA^{-1} = I.$$

For the third rule we use the identities $I^T = I$ and $(AB)^T = B^T A^T$:

$$\begin{aligned} AA^{-1} &= I \\ (AA^{-1})^T &= I^T \\ (A^{-1})^T A^T &= I. \end{aligned}$$

For the fourth rule we use the associativity of multiplication:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I.$$

□

These rules show us how inversion interacts with scalar multiplication, transposition and matrix multiplication. Warning: Inversion and addition do not play well together:

$$(A + B)^{-1} = \text{ nothing good.}$$

And that's it. I encourage you to memorize these rules.