Math 290 ELEMENTARY LINEAR ALGEBRA

REVIEW OF LECTURES – XI

September 27 (Wed), 2017

Instructor: Yasuyuki Kachi

Line #: 25751.

§11. Trace. Transpose.

Last time I declared to change narrative. I forgo my old ways and emulate the 'erudite' writing style à la research journal articles. So, from now on I define concepts for $n \times n$ matrices (and $m \times n$ matrices with $m \neq n$ if applicable) for all n (and m) at once, instead of isolating the 2×2 case (and the 3×3 case) to get a good feel first, whenever applicable/feasible. Today's first topic is 'trace'.

• From now on, a square (size) matrix means a matrix of size $\boxed{n} \times \boxed{n}$, a matrix whose number of rows and the number of columns are the same.

• **Definition (The trace of a matrix).** Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

be a square size matrix. We define the trace of A as the sum of the main diagonal entries $a_{11}, a_{22}, \cdots, a_{nn}$:

$$\boxed{\operatorname{Tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}.}$$

Repeat:

$$A = \begin{bmatrix} \boxed{a_{11}} & a_{12} & \cdots & a_{1n} \\ a_{21} & \boxed{a_{22}} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & \boxed{a_{nn}} \end{bmatrix} \quad \Longrightarrow \quad \operatorname{Tr} A = \boxed{a_{11}} + \boxed{a_{22}} + \cdots + \boxed{a_{nn}} .$$

Stated in other words:

$$\operatorname{Tr}\begin{bmatrix} \boxed{a_{11}} & a_{12} & \cdots & a_{1n} \\ a_{21} & \boxed{a_{22}} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & \boxed{a_{nn}} \end{bmatrix} = \boxed{a_{11}} + \boxed{a_{22}} + \cdots + \boxed{a_{nn}} \,.$$

This is deceptively innocuous. You are probably 'cynical' as to why we bother to isolate such a notion, and spend time on it. I said I prioritize 'efficiency', and that means that I don't necessarily stop and tell you the reason why I introduce this or that concept. Only this time, some nuts and bolts: Trace proves to be indispensable in one big area of mathematics, called *representation theory.* A grossly oversimplified description of representation theory:

> "*Search for possibilities whether and how the structures of various mathematical objects* (*models*) *in various sectors of math can be* '*simulated*' *by* $\boxed{\textbf{matrix}}$ *operations*."

For multiple good reasons, in mathematicians' mind, if the mathematical model that you are hammering away on (that initially has nothing to do with matrices) can be replicated using matrices, then that's the best possible scenario. And that indeed frequently happens. When that happens, there is a good possibility that model is mathematically rich and entrancing. That replica is called a representation (of the model). I know the word 'representation' is a generic word, but when used in math, it has this specific meaning. (Often some adjectives come with it, such as '*faithful* representation', '*irreducible* representation', '*induced* representation', *etc.*, *etc.* So it's really a technical term.)

Agree that this synopsis is a robust statement about the position of matrices in math. Matrices play a pivotal role in math. Such a dogma is fundamentally ingrained in every mathematician's head. Representation theory encompasses many areas of math. Representation theory itself is way beyond the scope of Math 290, but I just threw the above line for some doubters among you. So, for doubters: Trace is one central notion in representation theory, and thus important in math. Another name for 'trace' is '*character*'. So anyhow, please don't render premature judgment and be dismissive about it.

• So, what is the trace of

$$A \;=\; \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}?$$

Yes. It is $\;a_{11} + a_{22}$.

$$\boxed{\quad A \;=\; \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad \Longrightarrow \qquad \mathrm{Tr}\, A \;=\; a_{11} + a_{22}. \quad}$$

Stated in other words:

$$\boxed{\quad \mathrm{Tr} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \;=\; a_{11} + a_{22}. \quad}$$

• In the above I used the double suffixes. Last time we started using them as it is the only way to go about it when writing up a *general* $n \times n$ matrix. But for 2×2, we can certainly use a, b, c and d. So, a paraphrase:

Paraphrase for $\mathbf{2 \times 2}$ case.

$$\boxed{\quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad \Longrightarrow \qquad \mathrm{Tr}\, A \;=\; a + d. \quad}$$

Or the same to say:

$$\boxed{\quad \mathrm{Tr} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \;=\; a + d. \quad}$$

Example 1. (1) For $A \;=\; \begin{bmatrix} 2 & -1 \\ 8 & 3 \end{bmatrix}$, we have $\mathrm{Tr}\, A \;=\; 2 + 3 \;=\; 5$.

(2) $\quad \mathrm{Tr} \begin{bmatrix} 6 & 4 \\ 4 & -3 \end{bmatrix} \;=\; 6 + \Big(-3 \Big) \;=\; 3.$

- Next, what is the trace of

$$A \;=\; \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}?$$

Yes. It is $\quad a_{11} + a_{22} + a_{33}$:

$$\boxed{A \;=\; \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \qquad \Longrightarrow \qquad \mathrm{Tr}\, A \;=\; a_{11} + a_{22} + a_{33}.}$$

Stated in other words:

$$\boxed{\mathrm{Tr} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11} + a_{22} + a_{33}.}$$

- For 3×3 we can certainly do away with the double suffixes:

Paraphrase for $\mathbf{3 \times 3}$ case.

$$\boxed{A \;=\; \begin{bmatrix} a & b & c \\ p & q & r \\ x & y & z \end{bmatrix} \qquad \Longrightarrow \qquad \mathrm{Tr}\, A \;=\; a + q + z.}$$

Or the same to say:

$$\boxed{\mathrm{Tr} \begin{bmatrix} a & b & c \\ p & q & r \\ x & y & z \end{bmatrix} = a + q + z.}$$

Example 2. (1) For $A = \begin{bmatrix} 2 & 4 & -3 \\ 6 & 5 & 6 \\ -2 & 2 & 0 \end{bmatrix}$, we have

$$\operatorname{Tr} A = 2 + 5 + 0 = 7.$$

(2) $$\operatorname{Tr} \begin{bmatrix} -1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix} = \Big(-1 \Big) + \Big(-1 \Big) + \Big(-1 \Big) = -3.$$

Exercise 1. (1) Let $A = \begin{bmatrix} 4 & -1 \\ 5 & -3 \end{bmatrix}$. Find $\operatorname{Tr}\big(A\big)$.

(2) Find $\operatorname{Tr} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Exercise 2. (1) Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & 6 \\ 3 & 1 & 4 \end{bmatrix}$. Find $\operatorname{Tr}\big(A\big)$.

(2) Let $B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Find $\operatorname{Tr}\big(B\big)$.

(3) Find $\operatorname{Tr} \begin{bmatrix} a^2 & ab & ac & ad \\ ab & b^2 & bc & bd \\ ac & bc & c^2 & cd \\ ad & bd & cd & d^2 \end{bmatrix}$.

(4) Find $\operatorname{Tr} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{5} \end{bmatrix}$.

Formula 1. Let A and B be $n \times n$ matrices $\big(A$ and B are both square, and of the same size$\big)$.

(1) $$\mathrm{Tr}\left(A + B\right) \;=\; \mathrm{Tr}\left(A\right) + \mathrm{Tr}\left(B\right).$$

(2) Let s be a scalar. Then

$$\mathrm{Tr}\left(sA\right) \;=\; s\,\mathrm{Tr}\left(A\right).$$

(3) $$\mathrm{Tr}\left(AB\right) \;=\; \mathrm{Tr}\left(BA\right).$$

Proof of Formula 1. Write

$$A \;=\; \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad \text{and} \quad B \;=\; \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}.$$

As for part (1), for A and B as above, we have

$$A + B \;=\; \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \cdots & a_{nn} + b_{nn} \end{bmatrix}.$$

Accordingly,

$$\mathrm{Tr}\left(A + B\right) \;=\; \left(a_{11} + b_{11}\right) + \left(a_{22} + b_{22}\right) + \;\cdots\; + \left(a_{nn} + b_{nn}\right)$$

$$=\; \left(a_{11} + a_{22} + \;\cdots\; + a_{nn}\right) + \left(b_{11} + b_{22} + \;\cdots\; + b_{nn}\right).$$

This is $\mathrm{Tr}\,A + \mathrm{Tr}\,B$.

As for part (2), for A as above, we have

$$sA = \begin{bmatrix} sa_{11} & sa_{12} & \cdots & sa_{1n} \\ sa_{21} & sa_{22} & \cdots & sa_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ sa_{n1} & sa_{n2} & \cdots & sa_{nn} \end{bmatrix}.$$

Accordingly,

$$\begin{aligned} \operatorname{Tr}\Big(sA\Big) &= sa_{11} + sa_{22} + \cdots + sa_{nn} \\ &= s\Big(a_{11} + a_{22} + \cdots + a_{nn}\Big). \end{aligned}$$

This is $s \operatorname{Tr} A$.

As for part (3), let A and B as above. Then AB equals

$$\begin{bmatrix} a_{11}b_{11}+a_{12}b_{21}+\cdots+a_{1n}b_{n1} & * & \cdots & * \\ * & a_{21}b_{12}+a_{22}b_{22}+\cdots+a_{2n}b_{n2} & & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & & a_{n1}b_{1n}+a_{n2}b_{2n}+\cdots+a_{nn}b_{nn} \end{bmatrix}$$

(just filled out the main diagonal entries) and accordingly

$$\begin{aligned} \operatorname{Tr}\Big(AB\Big) = \; & \Big(\boxed{a_{11}\,b_{11} + a_{12}\,b_{21} + \cdots + a_{1n}\,b_{n1}}\Big) \\ & + \Big(\boxed{a_{21}\,b_{12} + a_{22}\,b_{22} + \cdots + a_{2n}\,b_{n2}}\Big) \\ & + \cdots \\ & + \Big(\boxed{a_{n1}b_{1n} + a_{n2}b_{2n} + \cdots + a_{nn}b_{nn}}\Big). \end{aligned}$$

Likewise, BA equals

$$\begin{bmatrix} b_{11}a_{11}+b_{12}a_{21}+\cdots+b_{1n}a_{n1} & * & \cdots & * \\ * & b_{21}a_{12}+b_{22}a_{22}+\cdots+b_{2n}a_{n2} & & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & & b_{n1}a_{1n}+b_{n2}a_{2n}+\cdots+b_{nn}a_{nn} \end{bmatrix}$$

and accordingly

$$\begin{aligned}
\mathrm{Tr}\Big(BA\Big) = \quad & \Big(b_{11}\,a_{11} + b_{12}\,a_{21} + \cdots + b_{1n}\,a_{n1}\Big) \\
+ & \Big(b_{21}\,a_{12} + b_{22}\,a_{22} + \cdots + b_{2n}\,a_{n2}\Big) \\
+ & \quad \cdots \\
+ & \Big(b_{n1}a_{1n} + b_{n2}a_{2n} + \cdots + b_{nn}a_{nn}\Big) \\
= \quad & \left(\boxed{\begin{matrix} a_{11}\,b_{11} \\ a_{12}\,b_{21} \\ \vdots \\ a_{1n}b_{n1} \end{matrix}} + \boxed{\begin{matrix} a_{21}\,b_{12} \\ a_{22}\,b_{22} \\ \vdots \\ a_{2n}b_{n2} \end{matrix}} + \cdots + \boxed{\begin{matrix} a_{n1}\,b_{1n} \\ a_{n2}\,b_{2n} \\ \vdots \\ a_{nn}b_{nn} \end{matrix}}\right) \\
= \quad & \Big(\boxed{a_{11}\,b_{11} + a_{12}\,b_{21} + \cdots + a_{1n}\,b_{n1}}\Big) \\
+ & \Big(\boxed{a_{21}\,b_{12} + a_{22}\,b_{22} + \cdots + a_{2n}\,b_{n2}}\Big) \\
+ & \quad \cdots \\
+ & \Big(\boxed{a_{n1}b_{1n} + a_{n2}b_{2n} + \cdots + a_{nn}b_{nn}}\Big).
\end{aligned}$$

This is exactly $\mathrm{Tr}\Big(AB\Big)$. $\square$

Exercise 3. Prove that $\underline{\underline{\text{no}}}$ $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ satisfy

$$AB - BA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$\star$ **[Pointers for Exercise 3]:** For $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$,

$$\begin{aligned}
&AB - BA \\
&= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} - \begin{bmatrix} a_{11}b_{11} + a_{21}b_{12} & a_{12}b_{11} + a_{22}b_{12} \\ a_{11}b_{21} + a_{21}b_{22} & a_{12}b_{21} + a_{22}b_{22} \end{bmatrix} \\
&= \begin{bmatrix} \Big(a_{11}b_{11} + a_{12}b_{21}\Big) - \Big(a_{11}b_{11} + a_{21}b_{12}\Big) & * \\ * & \Big(a_{21}b_{12} + a_{22}b_{22}\Big) - \Big(a_{12}b_{21} + a_{22}b_{22}\Big) \end{bmatrix} \\
&= \begin{bmatrix} a_{12}b_{21} - a_{21}b_{12} & * \\ * & a_{21}b_{12} - a_{12}b_{21} \end{bmatrix} \\
&= \begin{bmatrix} a_{12}b_{21} - a_{21}b_{12} & * \\ * & -\Big(a_{12}b_{21} - a_{21}b_{12}\Big) \end{bmatrix}.
\end{aligned}$$

So

$$AB - BA = \begin{bmatrix} a_{12}b_{21} - a_{21}b_{12} & * \\ * & -\Big(a_{12}b_{21} - a_{21}b_{12}\Big) \end{bmatrix}.$$

Now, suppose, hypothetically, $AB - BA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Then we have

$$a_{12}b_{21} - a_{21}b_{12} = 1, \qquad \text{and} \qquad -\Big(a_{12}b_{21} - a_{21}b_{12}\Big) = 1.$$

But is that possible?

$\star$ $\Big[$ **Alternative method** $\Big]$**:** Assume $AB - BA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, to derive a contradiction. For that matter, take the trace of the two sides, so

$$(*) \qquad \operatorname{Tr}\Big(AB - BA\Big) = \operatorname{Tr}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The right-hand side of $(*)$ equals 2, whereas the left-hand side of $(*)$ equals

$$\begin{aligned} \operatorname{Tr}\Big(AB - BA\Big) &= \operatorname{Tr}\Big(AB + (-BA)\Big) \\ &= \operatorname{Tr}\Big(AB\Big) + \operatorname{Tr}\Big(-BA\Big) \\ &= \operatorname{Tr}\Big(AB\Big) - \operatorname{Tr}\Big(BA\Big). \end{aligned}$$

Based on Formula 1, (3), this last quantity equals 0. Isn't this a contradiction?

- **Warning.** Though I said

$$\operatorname{Tr}\Big(AB\Big) = \operatorname{Tr}\Big(BA\Big)$$

is true, I did not say that $\operatorname{Tr}\Big(AB\Big) = \Big(\operatorname{Tr} A\Big)\Big(\operatorname{Tr} B\Big)$ is true.

So, I repeat:

$$\boxed{\underline{\underline{\textit{In general,}}} \quad \operatorname{Tr}\Big(AB\Big) \quad \underline{\underline{\textit{and}}} \quad \Big(\operatorname{Tr} A\Big)\Big(\operatorname{Tr} B\Big) \quad \underline{\underline{\textit{are not equal.}}}}$$

Example 3. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Then $AB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Clearly $\operatorname{Tr}\Big(AB\Big) = 2$, whereas $\operatorname{Tr} A = 2$, $\operatorname{Tr} B = 2$, and hence $\Big(\operatorname{Tr} A\Big)\Big(\operatorname{Tr} B\Big) = 4$. Thus in this case $\operatorname{Tr}\Big(AB\Big) \neq \Big(\operatorname{Tr} A\Big)\Big(\operatorname{Tr} B\Big)$.

Now we move on to the second topic of the day.

- **Definition (Transpose).** For an $m \times n$ matrix (need not be a square size)

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

the transpose A^T of A is defined as an $n \times m$ matrix

$$\boxed{A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}.}$$

Stated in other words, the transpose A^T of A is the matrix formed by interchanging rows and columns of A. Or stated in other words, A^T is the matrix whose (i, j) entry matches with the (j, i) entry of A.

- So, what is the transpose of

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

Yes. It is $\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$. The two entries off the main diagonal got interchanged.

In short:

$$\boxed{A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad \Longrightarrow \qquad A^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}.}$$

Stated in other words:

$$\boxed{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}.}$$

Paraphrase for 2×2 case (Do away with suffixes).

$$\boxed{A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Longrightarrow \quad A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}.}$$

Stated in other words:

$$\boxed{\begin{bmatrix} a & b \\ c & d \end{bmatrix}^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}.}$$

- Next, what is the transpose of

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}?$$

Yes. It is $\begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$.

The process to generate this matrix out of the original matrix A involves the following three interchanges:

(i) $$A = \begin{bmatrix} * & \boxed{a_{12}} & * \\ \boxed{a_{21}} & * & * \\ * & * & * \end{bmatrix} \quad \Longrightarrow \quad A^T = \begin{bmatrix} * & \boxed{a_{21}} & * \\ \boxed{a_{12}} & * & * \\ * & * & * \end{bmatrix},$$

(ii) $$A = \begin{bmatrix} * & * & \boxed{a_{13}} \\ * & * & * \\ \boxed{a_{31}} & * & * \end{bmatrix} \quad \Longrightarrow \quad A^T = \begin{bmatrix} * & * & \boxed{a_{31}} \\ * & * & * \\ \boxed{a_{13}} & * & * \end{bmatrix},$$

(iii) $$A = \begin{bmatrix} * & * & * \\ * & * & \boxed{a_{23}} \\ * & \boxed{a_{32}} & * \end{bmatrix} \quad \Longrightarrow \quad A^T = \begin{bmatrix} * & * & * \\ * & * & \boxed{a_{32}} \\ * & \boxed{a_{23}} & * \end{bmatrix}.$$

In short,

$$\boxed{A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \qquad \Longrightarrow \qquad A^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}.}$$

Stated in other words:

$$\boxed{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}.}$$

Paraphrase for 3×3 case (Do away with suffixes).

$$\boxed{A = \begin{bmatrix} a & b & c \\ p & q & r \\ x & y & z \end{bmatrix} \qquad \Longrightarrow \qquad A^T = \begin{bmatrix} a & p & x \\ b & q & y \\ c & r & z \end{bmatrix}.}$$

Stated in other words:

$$\boxed{\begin{bmatrix} a & b & c \\ p & q & r \\ x & y & z \end{bmatrix}^T = \begin{bmatrix} a & p & x \\ b & q & y \\ c & r & z \end{bmatrix}.}$$

- Observe that the rows of A become the columns of A^T:

$$\boxed{A = \begin{bmatrix} \boxed{a \quad b \quad c} \\ \boxed{p \quad q \quad r} \\ \boxed{x \quad y \quad z} \end{bmatrix} \qquad \Longrightarrow \qquad A^T = \left[\boxed{\begin{matrix} a \\ b \\ c \end{matrix}} \; \boxed{\begin{matrix} p \\ q \\ r \end{matrix}} \; \boxed{\begin{matrix} x \\ y \\ z \end{matrix}} \right].}$$

- Transpose operation is defined for non-square matrices.

- For example, what is the transpose of

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix} ?$$

Yes. It is $\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \\ a_{14} & a_{24} \end{bmatrix}$. Observe that the rows of A are the columns of A^T.

Paraphrase for 2×4 case.

$$\boxed{A = \begin{bmatrix} a & b & c & d \\ p & q & r & s \end{bmatrix} \quad \Longrightarrow \quad A^T = \begin{bmatrix} a & p \\ b & q \\ c & r \\ d & s \end{bmatrix}.}$$

Stated in other words:

$$\boxed{\begin{bmatrix} a & b & c & d \\ p & q & r & s \end{bmatrix}^T = \begin{bmatrix} a & p \\ b & q \\ c & r \\ d & s \end{bmatrix}.}$$

Example 4. (1) $A = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \quad \Longrightarrow \quad A^T = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}.$

(2) $A = \begin{bmatrix} 4 & 2 & 3 \\ 9 & 1 & 6 \\ -1 & 0 & -5 \end{bmatrix} \quad \Longrightarrow \quad A^T = \begin{bmatrix} 4 & 9 & -1 \\ 2 & 1 & 0 \\ 3 & 6 & -5 \end{bmatrix}.$

(3) $\begin{bmatrix} 2 & 4 \\ 1 & 5 \\ 0 & 3 \end{bmatrix}^T = \begin{bmatrix} 2 & 1 & 0 \\ 4 & 5 & 3 \end{bmatrix}.$

Exercise 4. Find the following:

(1) A^T, where $A = \begin{bmatrix} -4 & 3 \\ 5 & 1 \end{bmatrix}$.

(2) A^T, where $A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$.

(3) $\begin{bmatrix} 1 & -2 & 4 \\ 2 & 5 & 3 \\ 0 & 6 & 2 \end{bmatrix}^T$.

(4) A^T, where $A = \begin{bmatrix} 5 & -4 & -7 & 6 \\ 4 & 5 & -6 & -7 \\ 7 & 6 & 5 & 4 \\ -6 & 7 & -4 & 5 \end{bmatrix}$.

(5) $\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \end{bmatrix}^T$.

Formula 2. (1) For any matrix A, we have

$$\left(A^T\right)^T = A.$$

(2) For two matrices A and B such that $A + B$ is defined, we have

$$\left(A + B\right)^T = A^T + B^T.$$

(3) For a matrix A and a scalar t, we have

$$\left(tA\right)^T = t\left(A^T\right).$$

(4) For two matrices A and B such that AB is defined, we have

$$\left(AB\right)^T = B^T A^T.$$

(5) For a matrix A, we have

$$\left(A + A^T\right)^T = A + A^T \qquad \left(\text{when } A \text{ is square}\right), \quad \text{and}$$

$$\left(AA^T\right)^T = A\,A^T.$$

• We need to prove Formula 2. Parts (1–3) are trivial. Also, part (5) is an immediate consequence of parts (1–4). Hence it suffices to prove part (4).

Proof of Formula 2, (4).

Suppose A is of size (m, k), and B is of size (k, n), so that AB is defined. The (i, j)-th entry of $\left(AB\right)^T$ is the same as the (j, i)-th entry of AB, which is

$$(*) \qquad a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jk}b_{ki}.$$

Meanwhile, the i-th row of B^T is $\begin{bmatrix} b_{1i} & b_{2i} & \cdots & b_{ki} \end{bmatrix}$, and the j-th column of A^T is $\begin{bmatrix} a_{j1} \\ a_{j2} \\ \vdots \\ a_{jk} \end{bmatrix}$. Thus the (i, j)-th entry of $B^T A^T$ is

$$(\#) \qquad b_{1i}a_{j1} + b_{2i}a_{j2} + \cdots + b_{ki}b_{jk}.$$

Clearly $(*)$ and $(\#)$ are equal. $\square$

Warning. <u>In general $A^T A$ and AA^T are not equal,</u> as the next example shows:

Example 5. For $A = \begin{bmatrix} -3 & 4 \\ 2 & 1 \end{bmatrix}$, we have $A^T = \begin{bmatrix} -3 & 2 \\ 4 & 1 \end{bmatrix}$. Accordingly,

$$A^T A = \begin{bmatrix} -3 & 2 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} -3 & 4 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 13 & -10 \\ -10 & 17 \end{bmatrix},$$

whereas

$$AA^T = \begin{bmatrix} -3 & 4 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -3 & 2 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 25 & -2 \\ -2 & 5 \end{bmatrix}.$$

Thus in this case $A^T A \neq AA^T$.

Exercise 5. Let $A = \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix}$, and $B = \begin{bmatrix} -3 & -1 \\ 2 & 1 \end{bmatrix}$.

(1) Find $\Big(AB\Big)^T$. (2) Find $B^T A^T$.

(3) Verify $\Big(AB\Big)^T = B^T A^T$ for the above A and B.

- A couple more things before wrapping up.

- **Symmetric and skew-symmetric matrices.**

Definition (Symmetric matrix).

If a matrix A satisfies $\boxed{A^T = A}$ then we call such A a symmetric matrix.

Definition (Skew-symmetric matrix).

If a matrix A satisfies $\boxed{A^T = -A}$ then we call such A a skew-symmetric matrix.

- Both symmetric and skew-symmetric matrices have to be in square size.
- The main diagonal of a skew-symmetric matrix is entirely 0.
- $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is the general 2×2 symmetric matrix.
- $\begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$ is the general 2×2 skew-symmetric matrix.
- $\begin{bmatrix} a & b & c \\ b & p & q \\ c & q & r \end{bmatrix}$ is the general 3×3 symmetric matrix.
- $\begin{bmatrix} 0 & -r & q \\ r & 0 & -p \\ -q & p & 0 \end{bmatrix}$ is the general 3×3 skew-symmetric matrix.

Exercise 6. Write out the general 4×4 symmetric matrix. Also write out the general 4×4 skew-symmetric matrix.

Exercise 7. Decide whether each of the following matrices is symmetric. Decide whether each of the following matrices is skew-symmetric.

(1) $\begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$. (2) $\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$. (3) $\begin{bmatrix} 0 & 2 & -1 \\ -2 & 0 & -3 \\ 1 & 3 & 0 \end{bmatrix}$. (4) $\begin{bmatrix} 1 & -1 & 0 & 2 \\ -1 & 3 & 4 & -3 \\ 0 & 4 & 2 & 1 \\ 2 & -3 & 1 & 0 \end{bmatrix}$.

Fact 1. Let A be a square size matrix. Then

$$A + A^T, \qquad AA^T, \qquad \text{and} \qquad A^T A$$

are all symmetric.

Proof. This is immediate from Formula 2, part (5). □

Fact 2. Assume A and B are both symmetric matrices and of the same size. Then

(1) $A + B$ is symmetric.

(2) Let s be a scalar. Then sA is symmetric.

(3) $A^2 = AA$ is symmetric.

Proof. This is immediate from Formula 2, part (5). □

Fact 3. Assume A and B are both skew-symmetric matrices and of the same size. Then

(1) $A + B$ is skew-symmetric.

(2) Let s be a scalar. Then sA is skew-symmetric.

(3) $A^2 = AA$ is symmetric.

Exercise 8. Prove Fact 3 above.

Fact 4. Let A be a square size matrix. Then

(1) $\frac{1}{2}\left(A + A^T\right)$ is symmetric.

(2) $\frac{1}{2}\left(A - A^T\right)$ is skew-symmetric.

(3) A is always written as

$$A = B + C, \qquad B \text{ is symmetric, and } C \text{ is skew-symmetric,}$$

for suitable B and C.

Exercise 9. Prove Fact 4 above.

Exercise 10. Write $A = \begin{bmatrix} 2 & 5 & 3 \\ -3 & 6 & 0 \\ 4 & 1 & 1 \end{bmatrix}$ as the sum of a symmetric matrix B and a skew-symmetric matrix C.