postpone today. Cover Tuesday.

The transpose operation. One reason for considering this particular operation will be more clear by the beginning of next week, but since the text introduces it in section 2.1, we will as well.

Definition: Let $B_{m \times n} = [b_{ij}]$. Then the transpose of B, denoted by B^T is an $n \times m$ matrix defined by

$1 \le i \le m$
$1 \le j \le n$

$$\underbrace{entry_{ij}(B^T)} := \underbrace{entry_{ji}(B)} = b_{ji}.$$

$1 \le j \le m$
$1 \le i \le n$

The effect of this definition is to turn the columns of B into the rows of B^T:

$$entry_i(col_j(B)) = b_{ij}.$$
$$entry_i(row_j(B^T)) = entry_{ji}(B^T) = b_{ij}.$$

And to turn the rows of B into the columns of B^T:

$$entry_j(row_i(B)) = b_{ij}$$
$$entry_j(col_i(B^T)) = entry_{ji}(B^T) = b_{ij}.$$

$$col_j(B) = \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{mj} \end{bmatrix}$$

$$row_j(B^T) = [b^T_{j1} \; b^T_{j2} \cdots b^T_{jm}]$$
$$= [b_{1j} \; b_{2j} \cdots b_{mj}]$$

Exercise 1) explore these properties with the identity

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

$$[1, 2, 3]^T = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$ in some texts, makes sense

Algebra of transpose:

a $(A^T)^T = A$ ✓

b $(A + B)^T = A^T + B^T$ ✓

c for every scalar r $(rA)^T = r A^T$ ✓

$$entry_{ij}(AB)^T = entry_{ji} AB$$
$$= row_j(A) \cdot col_i(B)$$
$$= col_j(A^T) \cdot row_i(B^T)$$
$$= row_i(B^T) \cdot col_j(A^T)$$
$$= entry_{ij}(B^T A^T)$$

d (The only surprising property, so we should check it.) $\underline{(A B)^T = B^T A^T}$ *

check this

if A^{-1} exists then so does $(A^T)^{-1}$. & it $= (A^{-1})^T$

one consequence: if $AA^{-1} = I$, $A^{-1}A = I$

* take transpose: $(A^{-1})^T A^T = I$, $A^T (A^{-1})^T = I$

Wed Sept 19

- 2.2-2.3 Matrix inverses: the product of elementary matrices approach to matrix inverses

Announcements: Wed notes, after fill in the $\underline{(AB)^T = B^T A^T}$

quiz...

'til 12:57

Warm-up Exercise: Compute

$$\begin{bmatrix} 3 & 2 & -1 \end{bmatrix} \begin{bmatrix} a & b & c & d \\ A & B & C & D \\ e & f & g & h \end{bmatrix}. = 1 \times 4 \text{ matrix}$$

Can you express the result as a linear combination of the rows in the 3x4 matrix?

$$= \begin{bmatrix} 3a + 2A - e & 3b + 2B - f & 3c + 2C - g & 3d + 2D - h \end{bmatrix}$$

$$= 3\begin{bmatrix} a & b & c & d \end{bmatrix} + 2\begin{bmatrix} A & B & C & D \end{bmatrix} - \begin{bmatrix} e & f & g & h \end{bmatrix}$$

Exercise 1) Show that if A, B, C are invertible matrices, then

a) $$(A\,B)^{-1} = B^{-1}\,A^{-1}.$$

b) $$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$$

a). $(AB)(B^{-1}A^{-1})$

$= A\,(BB^{-1})\,A^{-1}$

$= (A\,I)A^{-1}$

$= AA^{-1} = I$

b) $(AB(C)(C^{-1})B^{-1}A^{-1}) \stackrel{?}{=} I$

$A(B\,\cancel{I}\,B^{-1})A^{-1}$

$A\,\cancel{I}\,A^{-1} = I$

(all the steps)

$(AB)C = A(BC)$

so $(AB)(B^{-1}A^{-1})$

$= A(\underbrace{B(B^{-1}A^{-1})})$

$= A((BB^{-1})A^{-1})$

$= A(I\,A^{-1})$

$= AA^{-1}$

$= I$

Theorem The product of $n \times n$ invertible matrices is invertible, and the inverse of the product is the product of their inverses in reverse order.

Our algorithm for finding the inverse of a matrix can be reformulated in terms of a product of so-called "elementary" matrices. This product idea will pay off elsewhere. To get started, let's notice an analog of the fact that a matrix times a vector is a linear combination of the matrix columns. That was in fact how we defined matrix times vector in week 2.

<u>Definition</u> (from 1.4) If A is an $m \times n$ matrix, with columns $\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \ldots \underline{\boldsymbol{a}}_n$ (in $\mathbb{R}^m$) and if $\underline{\boldsymbol{x}} \in \mathbb{R}^n$, then $A\,\underline{\boldsymbol{x}}$ is defined to be the linear combination of the columns, with weights given by the corresponding entries of $\underline{\boldsymbol{x}}$. In other words,

$$A\,\underline{\boldsymbol{x}} = \left[\underline{\boldsymbol{a}}_1 \;\; \underline{\boldsymbol{a}}_2 \;\; \ldots \underline{\boldsymbol{a}}_n\right]\underline{\boldsymbol{x}} \coloneqq x_1\,\underline{\boldsymbol{a}}_1 + x_2\,\underline{\boldsymbol{a}}_2 + \;\ldots x_n\,\underline{\boldsymbol{a}}_n \,.$$

<u>Theorem</u> If we multiply a *row vector* times an $n \times m$ matrix B we get a linear combination of the *rows* of B: <u>proof:</u> We want to check whether

$$\underline{\boldsymbol{x}}^T B = \left[\; x_1 \;\; x_2 \;\; \cdots \;\; x_n \;\right] \begin{bmatrix} \underline{\boldsymbol{b}}_1 \\ \underline{\boldsymbol{b}}_1 \\ \vdots \\ \underline{\boldsymbol{b}}_n \end{bmatrix} = x_1\,\underline{\boldsymbol{b}}_1 + x_2\,\underline{\boldsymbol{b}}_2 + \;\ldots x_n\,\underline{\boldsymbol{b}}_n \,.$$

where the rows of B are given by the row vectors $\underline{\boldsymbol{b}}_1, \underline{\boldsymbol{b}}_2, \ldots \underline{\boldsymbol{b}}_n$. This proposed identity is true if and only if its transpose is a true identity. But the transpose of the left side is

$$\left(\underline{\boldsymbol{x}}^T B\right)^T = B^T \left(\underline{x}^T\right)^T = B^T\,\underline{\boldsymbol{x}}$$

$$= \left[\underline{\boldsymbol{b}}_1^T \;\; \underline{\boldsymbol{b}}_2^T \;\; \ldots \underline{\boldsymbol{b}}_n^T\right]\underline{\boldsymbol{x}}$$

$$x_1\,\underline{\boldsymbol{b}}_1^T + x_2\,\underline{\boldsymbol{b}}_2^T + \;\ldots x_n\,\underline{\boldsymbol{b}}_n^T$$

which is the transpose of the right side of the proposed identity. So the identity is true.

Q.E.D.

notationally cleaner <u>proof</u> of thm; illustrated in warmup:

know $A\vec{x} = \left[\vec{a}_1 \;\; \vec{a}_2 \;\cdots\; \vec{a}_n\right]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n$

take transpose of both sides:

$$\vec{x}^T A^T = x_1\vec{a}_1^T + x_2\vec{a}_2^T + \cdots + x_n\vec{a}_n^T$$

$$\left[x_1 \;\; x_2 \;\cdots\; x_n\right]\begin{bmatrix} \vec{a}_1^T \\ \vec{a}_2^T \\ \vdots \\ \vec{a}_n^T \end{bmatrix} = x_1\vec{a}_1^T + x_2\vec{a}_2^T + \cdots + x_n\vec{a}_n^T$$

Exercise 2a Use the Theorem on the previous page and work row by row on so-called "elementary matrix" E_1 on the right of the product below, to show that $E_1\ A$ is the result of replacing $row_3(A)$ with $row_3(A) - 2\ row_1(A)$, and leaving the other rows unchanged:

look at what happens row by row, using previous page.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ -2a_{11}+a_{31} & -2a_{12}+a_{32} & -2a_{13}+a_{33} \end{bmatrix}$$

2b) The inverse of E_1 must undo the original elementary row operation, so must replace any $row_3(A)$ with $row_3(A) + 2\ row_1(A)$. So it must be true that

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}. \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}}_{E_1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \checkmark$$

Check!

2c) What 3×3 matrix E_2 can we multiply times A, in order to multiply $row_2(A)$ by 5 and leave the other rows unchanged. What is E_2^{-1} ?

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

E_2^{-1} : replace "5" with $1/5$

2d) What 3×3 matrix E_3 can we multiply time A, in order to swap $row_1(A)$ with $row_3(A)$? What is E_3^{-1} ?

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$E_3^{-1} = E_3$

Definition An *elementary matrix* E is one that is obtained by doing a single elementary row operation on the identity matrix.

Theorem Let $E_{m \times m}$ be an elementary matrix. Let $A_{m \times n}$. Then the product $E\,A$ is the result of doing the same elementary row operation to A that was used to construct E from the identity matrix.

Algorithm for finding A^{-1} re-interpreted: Suppose a sequence of elementary row operations reduces the $n \times n$ square matrix A to the identity I_n. Let the corresponding elementary matrices, in order, be given by

$$E_1, E_2, \ldots\ E_p.$$

Then we have

$$E_p\left(E_{p-1} \ldots E_2\left(E_1(A)\right)\ldots\right) = I_n$$

$$\underbrace{\left(E_p\,E_{p-1} \ldots E_2\,E_1\right)}_{\text{so, } = A^{-1}} A = I_n\,.$$

So,

$$A^{-1} = E_p\,E_{p-1} \ldots E_2\,E_1.$$

Notice that

$$E_p\,E_{p-1} \ldots E_2\,E_1 = E_p\,E_{p-1} \ldots E_2\,E_1\,I_n$$

so we have obtained A^{-1} by starting with the identity matrix, and doing the same elementary row operations to it that we did to A, in order to reduce A to the identity. I find this explanation of our original algorithm to be somewhat convoluted, but as I said, the matrix product decomposition idea is going to pay dividends elsewhere.

Also, notice that we have ended up "factoring" A into a product of elementary matrices:

$$A = \left(A^{-1}\right)^{-1} = \left(E_p\,E_{p-1} \ldots E_2\,E_1\right)^{-1} = E_1^{-1}\,E_2^{-1} \ldots E_{p-1}^{-1}\,E_p^{-1}\,.$$