Math 2270-004 Week 6 notes
We will not necessarily finish the material from a given day's notes on that day. We may also add or subtract some material as the week progresses, but these notes represent an in-depth outline of what we plan to cover. These notes cover material in 2.3, 3.1-3.3. In our original course syllabus I had planned to cover 2.4-2.5. These sections are considered optional by the Math Department course coordinator, and in the interests of having enough time for core material that comes later, we'll skip them, at least for now.

2.4 partioned matrices
2.5 matrix product decompositions, as a tool in solving linear systems.
3.1-3.3: determinants

Mon Feb 12
- 2.3 Matrix inverses, the elementary matrix approach
- overview of skipped section 2.5.

Announcements:
- HW posted - due next week
- Quiz this week, on T, W material.
- I'm about 70% finished grading exams (for elementary)

Warm-up Exercise: 'til 12:57

a) Compute the matrix product

$$\underset{E}{\begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \underset{A}{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}}$$

b) describe what happened to the rows of A

think row by row

$$EA = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 3a_{11}+a_{21} & 3a_{12}+a_{22} & 3a_{13}+a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

row_1, row_3 stayed the same

$3R_1 + R_2 \rightarrow R_2$

today: thinking about rref & inverse matrices via "Elementary matrices" that do elementary row ops to matrix A, when we multiply on the left.

We didn't get a chance to discuss the last three statements in the invertible matrix theorem in class, so let's do that before getting to the main part of today's discussion, about the elementary matrix approach to matrix inverses:

<u>The invertible matrix theorem</u> (page 114)

Let A be a square $n \times n$ matrix. Then the following statements are equivalent. That is, for a given A, the statements are either all true or all false.

a) A is an invertible matrix.
b) The reduced row echelon form of A is the $n \times n$ identity matrix.
c) A has n pivot positions
d) The equation $A\,\underline{\boldsymbol{x}} = \underline{\boldsymbol{0}}$ has only the trivial solution $\underline{\boldsymbol{x}} = \underline{\boldsymbol{0}}$.
e) The columns of A form a linearly independent set.
f) The linear transformation $T(\underline{\boldsymbol{x}}) := A\,\underline{\boldsymbol{x}}$ is one-one.
g) The equation $A\,\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}$ has at least one solution for each $\underline{\boldsymbol{b}} \in \mathbb{R}^n$.
h) The columns of A span $\mathbb{R}^n$.
i) The linear transformation $T(\underline{\boldsymbol{x}}) := A\,\underline{\boldsymbol{x}}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.

j) There is an $n \times n$ matrix C such that $C A = I$.

j, k : half an inverse $\Rightarrow$ all of an inverse.

a) $\Rightarrow$ j) Let $C = A^{-1}$

j) $\Rightarrow$ a) : j) $\Rightarrow$ d) $\Rightarrow$ a : if $A\vec{x} = \vec{0}$

$\left.\begin{array}{l}\underbrace{CA}_{I}\vec{x} = C\vec{0} = \vec{0}\end{array}\right\} \Rightarrow \vec{x} = \vec{0}$.

k) There is an $n \times n$ matrix D such that A D $= I$.

a) $\Rightarrow$ k) Let $D = A^{-1}$

k) $\Rightarrow$ g) $\Rightarrow$ a.

to solve $A\vec{x} = \vec{b}$
know $AD = I$
$AD\vec{b} = I\vec{b} = \vec{b}$. Let $\vec{x} = D\vec{b}$!

l) A^T is an invertible matrix.

a) $\Rightarrow$ l). If A^{-1} exists. Call it "B"

so $AB = I$, $BA = I$

take transpose : $(AB)^T = I^T = I$ $\qquad (BA)^T = I^T = I$

$B^T A^T = I$ $\qquad A^T B^T = I$

so $(A^T)^{-1} = B^T = (A^{-1})^T$

l) $\Rightarrow$ a) just use a) $\Rightarrow$ l) on A^T, i.e. if A^T has inverse, C,

then $\underset{\overset{\shortparallel}{A}}{(A^T)^T}$ has inverse C^T

2.3: elementary matrix approach to invertible matrices.

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

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

$$(AB)(B^{-1}A^{-1}) = \underbrace{A(\underbrace{BB^{-1}}_{I})}_{A}A^{-1} = AA^{-1} = I$$

$$(B^{-1}A^{-1})(AB)$$
$$= B^{-1}IB$$
$$= B^{-1}B$$
$$= I$$

$$AB\underbrace{CC^{-1}}_{I}B^{-1}A^{-1}$$
$$= ABB^{-1}A^{-1}$$
$$= AA^{-1}$$
$$= I$$

As the examples above show, it is true that

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, although we usually use the dot product way of computing the product entry by entry, instead:

Definition (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,

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

Theorem If we multiply a *row vector* times an $n \times m$ matrix B we get a linear combination of the *rows* of B instead. We could check this from scratch, but it's convenient to make use of transposes, which covert column facts into row facts.

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \underline{x}^T = [x_1, x_2 \cdots x_n]$$

proof: We want to check whether

$$* \qquad \underline{\boldsymbol{x}}^T B = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \underline{\boldsymbol{b}}_1 \\ \underline{\boldsymbol{b}}_2 \\ \vdots \\ \underline{\boldsymbol{b}}_n \end{bmatrix} = x_1\,\underline{\boldsymbol{b}}_1 + x_2\,\underline{\boldsymbol{b}}_2 + \ldots x_n\,\underline{\boldsymbol{b}}_n\,.$$

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

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

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

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

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

Q.E.D.

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:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ -2a_{11}+a_{31} & -2a_{12}+a_{32} & -2a_{13}+a_{33} \end{bmatrix} = \begin{bmatrix} 1\cdot row_1(A) \\ 1\cdot row_2(A) \\ -2\,row_1(A)+row_3(A) \end{bmatrix}$$

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

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

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} ?

$$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

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} ?

$$E_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ \text{———} & & \\ a_{11} & a_{12} & a_{13} \end{bmatrix}$$

$$E_3^{-1} = E_3$$

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

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

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

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 pays dividends elsewhere.

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

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

Overview of skipped section, 2.5, matrix product decompositions:

Even if A is not invertible, and even if it's not a square matrix, the elementary matrix process gives a way to factor A into a product of an invertible matrix times one that is easier to work with than A, such as a row echelon form or the reduced row echelon form. This can be helpful if one needs an alorithm to quickly and repeatedly solve $A\,\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}$ for a large number of right hand sides $\underline{\boldsymbol{b}}$, and there is no inverse matrix A^{-1}.

Here's how the method goes works. Let's write G for the good matrix that is, for example, the row echelon or reduced row echelon form of the $m \times n$ matrix A. Reduce as we did before for invertible matrices, to get

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

Write

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

for the product of elementary matrices, and note that it's the matrix one obtains by augmenting A with the $m \times m$ identity matrix and doing the same elementary row operations to it as one does to A, just as in the algorithm for constructing inverses to invertible matrices. Then

$$BA = G$$

so

$$A = B^{-1}\, G.$$

Then to solve $A\,\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}$ repeatedly, we want

$$B^{-1}\, G\;\, \underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}$$

which is equivalent to the system

$$G\,\underline{\boldsymbol{x}} = B\,\underline{\boldsymbol{b}}\,.$$

Since G is a good matrix - like reduced row echelon form, this system is much quicker to solve repeatedly. These ideas are related to the concept of "preconditioning a matrix" before solving linear systems, which you can read about at Wikipedia.