Math 1b	TA: Padraic Bartlett

Recitation 5: Elementary Matrices

Week 5	*Caltech 2011*

1 Random Question

Consider the following two-player game, called **Angels and Devils**:

- Our game is played on a $n \times n$ chessboard, with one token (the "angel") placed at the center of the board.
- Our two players – the "angel" and the "devil" alternate taking turns.
- On the angel's turn, they can move their token to any adjacent square.
- On the devil's turn, they can pick any one square on the board not occupied by the angel and remove it from play; for the rest of the game, the angel can never enter that square.
- The angel wins if it can get to the perimeter of the board (as it can presumably then walk off the board.) The devil wins if it can trap the angel; i.e make it so the angel has no possible moves from its current position.

Assuming perfect play, who wins this game on a given $n \times n$ board? Does it depend on n?

2 HW comments

- Section average: 92%.
- People did remarkably well! In particular, there was a remarkably small amount of deviation from the in-class average of 92%; almost everyone displayed a nice command of the material on this set. Specifically, people did the following things much better than on sets past:
  - People explained why they did things! That was excellent.
  - People wrote far clearer proofs, and really paid attention to basic logical things, like how to prove "if and only if" statements. That was also excellent.
- Also, as an aside: always attach mathematica or wolfram alpha or what-have-you work if you use it. We need this to figure out what you're doing! Most of you are doing this, but some aren't, so I'll keep saying this until it's not an issue.

3 Elementary Matrices

Today's recitation is much shorter than previous recitations, because (1) I've been giving the Ma1b lectures, and so I'm really not sure what topics I missed (if I knew, then I wouldn't have missed them in lecture...) and (2) you have a midterm, and probably would prefer to talk about older material rather than newer material.

So, we did two things in this talk:

1. Discussed elementary matrices in depth: we recalled their definitions, proved that each elementary matrix corresponds to performing some given row operation on a matrix (and furthermore that for any such row operation, there is an elementary matrix that performs that operation), and showed how to decompose a matrix into elementary matrices.

2. Discussed various review topics, as determined by the class. As this was mostly off-the-cuff, I don't have notes for this; I think everything's contained in my previous recitations, but if you have any specific questions feel free to email me! I'll be around for all of Friday night, Saturday night, and also Sunday night (where there will be office hours at 10, should anyone still have questions and not be taking the midterm yet...)

In class on Wednesday, we defined **elementary matrices**. We restate their definition here, for convenience:

Definition. There are three different kinds of **elementary matrices**, corresponding to the three different types of row operations. We list them here:

$$E_{\text{multiply row k by }\lambda} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

If the λ is in the (i, i)-th spot, multiplying A on the left by this matrix multiplies A's i-th row by λ.

$$E_{\text{switch rows i and j}} = \begin{pmatrix} 1 & 0 & \dots & \boxed{0} & \dots & \boxed{0} & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & \vdots & \dots & 1 & \dots & 0 \\ 0 & 0 & \dots & 1 & \dots & \vdots & \dots & 0 \\ 0 & 0 & \ddots & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix}$$

The matrix above is the standard identity matrix with its i-th and j-th columns (highlighted) switched. Multiplying A on the left by this matrix switches A's i-th and j-th rows.

$$E_{\text{add } \lambda \cdot \text{row j to row i}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

If the λ is in the (i, j)-th spot, multiplying A on the left by this matrix adds λ times row j of A to row i of A.

In class, we claimed that these matrices "did what they said." In other words:

Theorem 1 *If we took any $n \times n$ matrix A, and multiply it on the left by some $n \times n$ elementary matrix E,*

- *if $E = E_{\text{multiply row k by } \lambda}$, then EA would be the matrix A with its k-th row multipled by λ.*
- *if $E = E_{\text{swap rows i and j}}$, then EA would be the matrix A with its i-th and j-th rows swapped, and*
- *if $E = E_{\text{add } \lambda \cdot \text{row j to row i}}$, then EA would be the matrix A with λ times its j-th row added to its i-th row.*

We prove this claim below: it's not necessarily very interesting, so feel free to skip this if you prefer.

Proof. To prove these claims, we simply perform matrix multiplication.

To start, take any $n \times n$ matrix A, row k and constant λ, and examine the product

$$E_{\text{multiply row k by } \lambda} \cdot A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 1 & \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & \dots & a_{3n} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & \dots & a_{4n} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & \dots & a_{5n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} & a_{n5} & \dots & a_{nn} \end{pmatrix}.$$

What do entries in the resulting matrix look like? Well, there are two cases:

- in the location (i, j), for any $i \neq k$ and any j, we know that the entry there is just the dot product of E's i-th row and A's j-th column: i.e.

$$\text{entry } (i,j) = (0, \ldots, 1, \ldots 0) \cdot (a_{1j}, \ldots a_{nj})^T = a_{ij},$$

  because the 1 in the i-th row of E is in the i-th place.

- in the location (k, j), for any j, we know that the entry there is just the dot product of E's k-th row and A's j-th column: i.e.

$$\text{entry } (k,j) = (0, \ldots, \lambda, \ldots 0) \cdot (a_{1j}, \ldots a_{nj})^T = \lambda_{kj},$$

  because the λ in the k-th row of E is in the k-th place.

By inspection, this matrix is A with its k-th row multipled by λ: so this elementary matrix works as claimed.

The proofs for the other two elementary matrices are similar. For the matrix $E_{\text{swap rows k and l}}$, we again examine the product EA:

$$E \cdot A = \begin{pmatrix} 1 & 0 & \ldots & \boxed{0} & \ldots & \boxed{0} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & \ldots & 1 & \ldots & 0 \\ 0 & 0 & \ldots & 1 & \ldots & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \boxed{0} & \ldots & \boxed{0} & \ldots & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & \ldots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & \ldots & a_{3n} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & & a_{4n} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & \ldots & a_{5n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} & a_{n5} & \ldots & a_{nn} \end{pmatrix}$$

Again, what do entries in the resulting matrix look like? In this situation, there are three cases:

- In the location (i, j), for any $i \neq k, l$ and any j, we know that the entry there is just the dot product of E's i-th row and A's j-th column: i.e.

$$\text{entry } (i,j) = (0, \ldots, 1, \ldots 0) \cdot (a_{1j}, \ldots a_{nj})^T = a_{ij},$$

  because the 1 in the i-th row of E is in the i-th place.

- In the location (k, j), for any j, we know that the entry there is just the dot product of E's k-th row and A's j-th column: i.e.

$$\text{entry } (k,j) = (0, \ldots, 1, \ldots 0) \cdot (a_{1j}, \ldots a_{nj})^T = a_{lj},$$

  because the 1 in the k-th row of E is in the l-th place.

- In the location (l, j), for any j, we know that the entry there is just the dot product of E's l-th row and A's j-th column: i.e.

$$\text{entry } (l,j) = (0, \ldots, 1, \ldots 0) \cdot (a_{1j}, \ldots a_{nj})^T = a_{kj},$$

  because the 1 in the l-th row of E is in the k-th place.

By inspection, this matrix is A with its k-th and l-th rows swapped, as claimed.

Finally, we turn to $E_{\text{add } \lambda\cdot\text{row k to row l}}$, and once more look at EA:

$$E_{\text{add } \lambda\cdot\text{row k to row l}} \cdot A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & \dots & a_{3n} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & \dots & a_{4n} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & \dots & a_{5n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & a_{n4} & a_{n5} & \dots & a_{nn} \end{pmatrix}.$$

Again, what do entries in the resulting matrix look like? In this situation, there are just two last cases:

- In the location (i, j), for any $i \neq l$ and any j, we know that the entry there is just the dot product of E's i-th row and A's j-th column: i.e.

  $$\text{entry } (i, j) = (0, \dots, 1, \dots 0) \cdot (a_{1j}, \dots a_{nj})^T = a_{ij},$$

  because the 1 in the i-th row of E is in the i-th place.

- In the location (l, j), for any j, we know that the entry there is just the dot product of E's k-th row and A's j-th column: i.e.

  $$\text{entry } (k, j) = (0, \dots, 0, \lambda, 0, \dots, 0, 1, 0, \dots 0) \cdot (a_{1j}, \dots a_{nj})^T = \lambda a_{kj} + a_{lj},$$

  because the λ in the l-th row of E is in the k-th place, and the 1 is in the l-th place.

By inspection, this matrix is A with λ times its k-th row added to its l-th row, as claimed.

So, they work! In other words, we can perform any row operation r on a matrix A by multiplying A on the left by some elementary matrix E_r. In class we made a rather remarkable claim based on this observation:

Theorem 2 *If A is a nonsingular matrix, we can find elementary matrices $E_1, \dots E_k$ such that $A = E_1 \cdot \ldots \cdot E_k$.*

Proof. Well, we know that if A is nonsingular, we can perform some series of row operations $r_n, \dots r_1$ to A such that $r_n \cdot \ldots r_1 \cdot A = I$, the identity matrix.

What does this mean? Well, suppose that we "undo" these row operations on I: i.e. suppose we start with I, and then perform the row operation r_n^{-1} that undoes the row operation r_n. (I.e. if r_n was multiplying the second row by 4, r_n^{-1} is dividing the second row by 4.) If we undo all of these in reverse order, we'd then have that

$$A = r_1^{-1} \cdot \ldots \cdot r_n^{-1} \cdot I.$$

So, simply take E_i to be the elementary matrix corresponding to r_i^{-1}! Then, we have that

$$A = E_1 \cdot \ldots \cdot E_n \cdot I,$$

as claimed.

The cool thing about the proof above is that it actually gives us a concrete algorithm that will decompose any nonsingular matrix into elementary matrices! The following is an example of this:

Example. Decompose the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 4 & 8 & 9 \end{pmatrix}$$

into elementary matrices.

Proof. So, we first find a sequence of row operations to transform A into the 3×3 identity matrix:

$$\begin{aligned}
\begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 4 & 8 & 9 \end{pmatrix} &=_{\text{(add -2r1 to r2)}} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 4 & 8 & 9 \end{pmatrix} \\
&\to_{\text{(add -4r1 to r3)}} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix} \\
&\to_{\text{(add -2r2 to r1)}} \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix} \\
&\to_{\text{(add r3 to r1)}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix} \\
&\to_{\text{(multiply r3 by -1/3)}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\end{aligned}$$

Now, according to our theorem, we can recreate A by undoing all of these row operations on I, starting from the last and working our way backwards: i.e.

$$\begin{aligned}
A &= \text{(add -2r1 to r2)}^{-1} \circ \text{(add -4r1 to r3)}^{-1} \circ \text{(add -2r2 to r1)}^{-1} \circ \text{(add r3 to r1)}^{-1} \circ \text{(mult. r3 by -1/3)}^{-1} \circ I \\
&= \text{(add 2r1 to r2)} \circ \text{(add 4r1 to r3)} \circ \text{(add 2r2 to r1)} \circ \text{(add (-1)r3 to r1)} \circ \text{(mult. r3 by -3)} \circ I \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix}.
\end{aligned}$$

So we're done!